ANGLE OF ARRIVAL ESTIMATION FOR CELLULAR NETWORKS IN MULTIPATH ENVIRONMENTS

by

Orçun TOPALOĞLU

B.S., Electronics and Communication Engineering, İstanbul Technical University, 2004

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Electrical and Electronics Engineering
Boğaziçi University

2007

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor, Prof. Dr. Emin Anarım for his guidance and motivation throughout my study. I would also like to thank my co-supervisor Assist. Prof. Dr. Kerem Harmancı for his continuous support and encouragement he provided in all stages of this thesis. Without their assistance and suggestions this thesis could not have been completed.

Finally, I would like to thank my family for their support during the period of study.

This work is supported by the State Planning Organization of Turkey under the Next Generation Satellite Networks Project, DPT 03K 120250.

ABSTRACT

ANGLE OF ARRIVAL ESTIMATION FOR CELLULAR NETWORKS IN MULTIPATH ENVIRONMENTS

Estimation of a mobile user's location within the wireless network has become a significant area of research within the last 10 years. Parallel to the increasing popularity of mobile phones, numerous location estimation techniques have been developed. These techniques aim to provide accurate position of mobile users and to open the way for the efficient usage of cellular services that can be provided with the help of location estimation.

Performance of location estimation schemes depends on the characteristics of the wireless channel. When the channel induces multipath effects, scattering objects can lead to errors in the estimated angle of arrival (AOA) of the mobile station.

This thesis studies the performance of different AOA estimation algorithms in the multipath propagation environment where uniform linear arrays (ULAs) are considered at the base stations. A nonisotropic scattering model is used and performances of different algorithms are analyzed with the results of several simulations.

ÖZET

ANGLE OF ARRIVAL ESTIMATION FOR CELLULAR NETWORKS IN MULTIPATH ENVIRONMENTS

Son on yılda, kablosuz haberleşme sistemindeki bir cep telefonunun yerini tespit etme problemi önemli bir araştırma alanı haline geldi. Cep telefonlarının daha sık kullanılmaya başlanmasına paralel olarak çok sayıda yer belirleme yöntemleri geliştirildi. Bu yöntemler, cep telefonu kullanıcılarının yerini en iyi şekilde tespit etmeyi ve yer belirleme yöntemleri sayesinde sunulacak servislerin etkili olarak kullanımını sağlamayı hedeflemekteler.

Yer belirleme yöntemlerinin performansı kablosuz kanalın karakteristik özelliklerine bağlıdır. Kanalda çokyolluluk etkileri mevcut ise, saçıcı nesneler cep telefonu için geliş açısı kestiriminde hatalara neden olabilirler.

Bu tez ile, çokyollu iletim ortamında ve baz istasyonlarında düzgün doğrusal anten kullanıldığı durumda, çeşitli geliş açısı kestirim yöntemlerinin performansları incelenmiştir. Yöne bağımlı bir saçılma modeli kullanılmıştır ve farkli yöntemlerin performansları simülasyon sonuçları ile birlikte analiz edilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS / ABBREVIATIONS

$\boldsymbol{a}$	Array transfer function vector
$\boldsymbol{A}$	Array transfer function matrix
c	Speed of light
d	Array separation in terms of wavelength
D	Distance between the MS and the BTS
$\boldsymbol{e}$	Noise vector
$\overline{e}$	Additive noise for a single waveform
e_l	Noise of the l-th array element
$\overline{E}$	Fourier transform of $\overline{e}$
f	Probability distribution function
$\overline{f}$	Asymptotic cost function
f_c	Carrier frequency
$f_{\max}$	Maximum Doppler frequency.
f_s	Spatial frequency
$\boldsymbol{h}$	Filter weight vector
$\overline{h}$	Impulse response of the kth array element
$\overline{H}$	Fourier transform of $\overline{h}$
$\boldsymbol{I}$	Identity matrix
I_0	Zero-order modified Bessel function of the first kind
m	Number of array elements
n	Number of sources
N	Number of samples
P	Number of clusters
P_{noise}	Noise power
P_{signal}	Signal power
$\boldsymbol{q}$	Eigenvector
$\boldsymbol{R_s}$	Correlation function matrix of incoming signals

$\boldsymbol{R}_y$	Correlation function matrix of received signals
$\hat{\boldsymbol{R}}_y$	Sample correlation function matrix
$\boldsymbol{R}_z$	Correlation function matrix of complex envelope LP equivalent of the signal
s	Baseband complex-valued signal
$\boldsymbol{s}$	Incoming signal vector
$\boldsymbol{S}$	Singular value matrix for SVD factorization
$\boldsymbol{u}$	Sinusoidal input vector
$\boldsymbol{U}$	Unitary matrix for SVD factorization
$\boldsymbol{V}$	Unitary matrix for SVD factorization
$\boldsymbol{V}_e$	Noise subspace matrix
w_c	Carrier angular frequency
w_s	Spatial angular frequency
x	Bandpass signal
X	Fourier transform of x
$\boldsymbol{y}$	Received signal vector
$\overline{y}$	Received single waveform
$\tilde{y}$	Demodulated signal
$\boldsymbol{y}_F$	FIR filter output vector
$\overline{Y}$	Fourier transform of $\overline{y}$
z	Roots of polynomial for Root-MUSIC
z_l	Complex envelope lowpass equivalent of the signal at the *l*-th array element
$\boldsymbol{\beta}$	Vector of unknown parameters for CRLB
δ	Dirac delta
ε	Mixture proportion of clusters
ϕ	Position of scatterer
ψ	Phase shift of scatterer
γ	Direction of motion for the scatterer
κ	Concentration parameter for the von Mises distribution
λ	Wavelength

λ_i	Eigenvalue
χ	Angle spread
σ	Size of the local scattering area around the MS
σ_s	Standard deviation of signal
σ_e	Standart deviation for noise
Ξ	Random reflection coefficient
$\boldsymbol{\rho}$	Vector of zero-mean complex Gaussian random variables
τ	Time delay
θ	Angle-of-arrival
$\overline{\theta}$	Maximizers/minimizers of the asymptotic cost function
$\hat{\theta}$	AOA estimates

AOA	Angle-of-Arrival
BER	Bit Error Ratio
BTS	Base Transceiver Station
CBF	Conventional Beamforming
CDMA2000	Code Division Multiple Access-2000
CRLB	Cramer-Rao Lower Bound
E112	Emergency 112
E911	Emergency 911
FCC	Federal Communications Commission
FIR	Finite Impulse Response
GPS	Global Positioning System
IMT2000	International Mobile Telecommunications-2000
ISI	Intersymbol Interference
LAN	Local Area Network
LOS	Line-of-Sight
MS	Mobile Station
MSE	Mean Squared Error
MUSIC	Multiple Signal Classification
MVDR	Minimum Variance Distortionless Response
PSAP	Public Safety Answering Point

PSTN	Public Switched Telephone Network
RF	Radio Frequency
RSS	Received Signal-Strength
SINR	Signal to Interference Plus Noise Ratio
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition
TDOA	Time-Difference-of-Arrival
TOA	Time-of-Arrival
ULA	Uniform Linear Array
W-CDMA	Wideband Code Division Multiple Access

1. INTRODUCTION

1.1. Motivation

Estimation of a mobile subscriber's location in a cellular environment has gained much attention over the past decade. This has been due to the increasing popularity of wireless networks and the proposed cellular services that can be provided with the help of location estimation.

The U.S. Federal Communications Commission (FCC) published some requirements regarding the location estimation of cellular phones in 1996. The FCC mandate stated that positions of all emergency 911 (E911) callers have to be determined within tens of meters by the wireless service providers. This mandate and the European recommendation for E112 have motivated the efforts to develop accurate algorithms for wireless location estimation. The main reason behind these emergency issues is that about one third of all emergency calls originate from cellular phones and also mobile users are more likely to be unaware of their exact location when compared with wireline PSTN users [1]. The order for the mobile service providers was to supply an estimate of the caller`s location 100m. within his actual location for at least 67 per cent of all wireless 911 calls. Moreover, an estimate within 300m. should have been possible for at least 95 per cent of all 911 calls.

There are some basic location estimation algorithms that gained attention in this research area. These radio location systems attempt to locate an MS by measuring radio signals between the MS and a set of base stations (BTSs). Such systems can be based on received signal-strength (RSS) [2], angle-of-arrival (AOA) [3], time-of-arrival (TOA) [4] or time-difference-of-arrival (TDOA) [5] and can be network-based or terminal-based.

The accuracy of these location estimation schemes depends on the propagation condition of the wireless channel. When the direct path from the MS to a BTS is blocked by buildings and other obstacles, the measurements include an error that is the result of multipath effects.

In this thesis, the focus is on the network-based AOA estimation algorithms. In next generation systems, AOA of a signal transmitted by an MS can be estimated at an antenna array which consists of multiple number of identical sensors placed at equal distances.

Multipath propagation due to scattering objects near and around the MS can lead to errors in the measured AOAs that cause the algorithms' performance to degrade. There are numerous researches in literature regarding the location estimation algorithms and their modification for the multipath environment [6, 7, 8].

The motivation of the thesis is to study the performance possibilities of using AOA estimation with ULAs in the outdoor multipath propagation environment where local scattering is present around the MS. Network-based approach is preferred and a nonisotropic scattering model is used to study the performances of different AOA estimation algorithms.

1.2. Outline of the Thesis

The thesis is organized as follows: Wireless positioning application areas and positioning categories are explained and some outdoor location estimation methods are introduced in Chapter 2.

In Chapter 3, wireless propagation mechanisms which include the basic fading models are described and multipath effects are analysed.

The historical development of antennas including smart antenna systems are explained and uniform linear array antenna model is studied in Chapter 4.

Chapter 5 introduces the von Mises distribution which is based on nonisotropic scattering model and some properties of the distribution function are analysed.

In Chapter 6, some AOA estimation methods are presented in detail. The von Mises pdf is used within these methods. Theoretical derivation of the Mean Squared Error (MSE) is made for these methods in order to use as a performance parameter in the simulations.

Cramer Rao Lower Bound for stochastic signals is given in Chapter 7. The derivation of CRLB for the von Mises model is also made in this chapter.

Chapter 8 presents simulation results and simulation analysis of different methods both with point source and with distributed source (using the von Mises pdf) assumption.

Chapter 9 concludes the thesis emphasizing the results of AOA estimation simulations with the von Mises pdf and also mentioning some suggestions for future work on this subject.

2. POSITIONING IN WIRELESS NETWORKS

2.1. Applications of Wireless Positioning

Although issues like emergency and security are the most efficient usage areas for location based services, there are many other useful future applications of wireless location technology ranging from navigation to billing [1]. Following are some application areas for wireless location estimation technology:

E911: Integration of wireless subscribers to the E911 network was planned by the FCC in 1996. The FCC order outlined the compliance in two phases: The first phase required the delivery of the caller`s number to public safety answering points (PSAPs) so that the dispatcher could dial back if the call was dropped. The second phase covered the display of the caller`s approximate location so that the nearest emergency personnel could be directed to the area.

Mobile advertising: Advertisements and sales news can be sent to the cellular phone users near the shops and malls with the help of location estimation.

Asset tracking: Efficient direction of people in sightseeing places, museums, malls, parking places can be carried on with cellular phones. In case of disasters like fire, flood and earthquakes locations of lost people can be determined easily. Moreover search and rescue teams can be directed to disaster hot spots in coordination.

Fleet management: Fleet management via cell phones may improve the coordination of police cars, fire brigades, taxi cabs and even the buses while trying to reach to emergency areas or bus stations. Drivers can be warned about accidents and traffic jams on highways so that they could choose alternative routes to reach to their destinations.

Location-based wireless access security: Employees working at high-tech firms and in large structures may only be allowed in certain zones. This control mechanism can be

established with the help of location-based tracking so that confidential information can be preserved efficiently.

Location-sensitive billing: Wireless positioning can also be used for special billing plans inside campuses, shopping malls and buses to attract new customers.

2.2. Positioning Categories

Wireless positioning systems can be studied in different classes. The most common classifications are according to the processing and the environment.

Mobile-based versus Network-based Positioning: If the MS does the processing and determines its location via received information from the BTSs or GPS satellites it is called mobile-based positioning. On the other hand, network-based positioning is based on the determination of the MSs position by the BTSs. Some wireless service providers chose to fulfill the FCC requirements via the network-based approach whereas others chose mobile-based solution which uses embedded GPS technology [1].

Network-based location technology does not require any changes in the existing cellular or wireless LANs. The base stations receive the signals from the cellular phones and location estimates from several BTSs are combined at a center to achieve a final result. The mobile stations do not require any modifications since they are not involved in the processing.

The companies who chose to use a network-based approach were able to meet the FCC deadline. However the ones who chose a phone-based solution that uses embedded GPS technology to identify the location had to ask for more time from the commission. The problem was not only the customers but also the call centers. The customers either did not want to spend money on a mobile phone with GPS technology or they did not want to learn to use a new one. GPS embedded phones are more expensive and larger in size with more battery consumption compared to the millions of classical cellular phones in the market. Another disadvantage is that the reception of GPS signals becomes very difficult and sometimes impossible in urban and indoor enviroments.

In some areas where the companies wanted to implement phone-based solution, digital service was not in good condition and the call centers were unable to receive the digital information from the cell phones. Since the states did not have the necessary funding or the authority to mandate these changes, the phone-based solution took a lot more time to be operational [9].

Indoor versus Outdoor Positioning: GPS technology and outdoor systems cannot provide accurate estimation inside buildings since much higher accuracy is needed when compared to outdoors. Outdoor systems can provide accurate results within tens of meters whereas indoor systems require an accuracy of a few meters.

Indoor location estimation systems use the existing wireless LANs. Scattering is a serious problem affecting the channel parameters in such places. Also shadow fading caused by walls and the path loss attenuate the signal strength even more. Another problem is that cellular phones operating indoors are subject to interference from other devices operating at the same frequency. For all these reasons, indoor location algorithms differ from outdoor solutions [10].

Outdoor positioning techniques differ according to the environment conditions (dense urban, urban, suburban) but they can provide long-range coverage with the help of existing wireless networks.

2.3. Outdoor Location Estimation Methods

There are several outdoor positioning methods based on different aspects of the signals in the propagation environment [1]. One of the methods is called the TOA (time of arrival) method. This method relies on the measurement of the time of arrival of the signal propagating between the mobile user and each base station receiver. Each TOA estimate is used to obtain a distance estimate for the MS. For a base station receiver this distance indicates a sphere around the BTS. This means, the MS may be located anywhere around the BTS at that distance. The combination of three different TOA estimates from three

BTSs results in an intersection point of three spheres. This point gives the location estimate for the MS as seen in Figure 2.1.

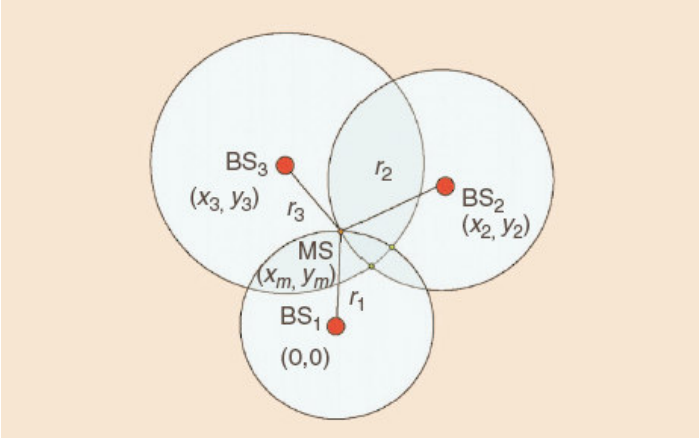


Figure 2.1. TOA Estimation [1]

Another method is the TDOA (time difference of arrival). The difference between the TOAs of the signal at BTS_i and BTS_1 is called the TDOA of BTS_i. The TDOA measurement indicates a hyperboloid of constant time differences with the two base stations at the foci of the hyperboloid. The measurements from several BTSs define multiple hyperboloids the intersection of which gives the MS location estimate.

Other than these two methods, received signal strength (RSS) can be used as an alternative. In this method, the mobile station collects RSS measurements of the signals transmitted by the surrounding base stations. These measurements are then used to determine the location of the MS with a one-to-one mapping between the RSS and the MS location values [2].

Another common positioning method is the AOA (angle of arrival). The angles of arrivals of the signals can be estimated with antenna arrays at the BTSs. The combination of two AOA estimates from two BTSs gives the location estimate of the MS as seen in Figure 2.2.

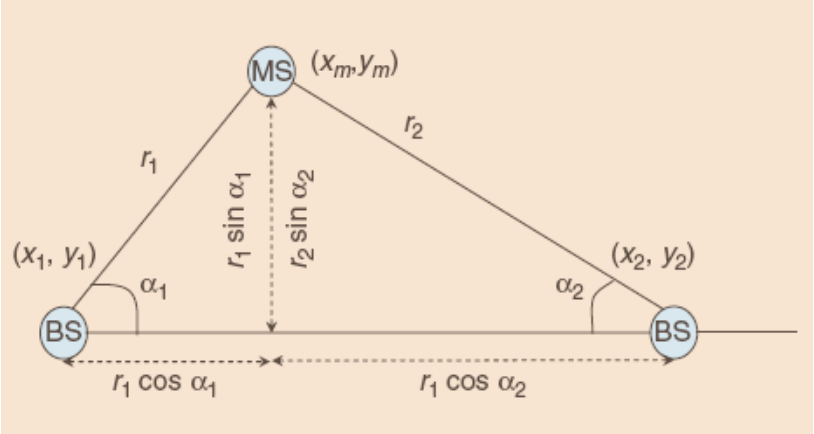


Figure 2.2. AOA Estimation [1]

The accuracy of all these methods relies on the properties of the propagation environment. Factors such as multipath effects, multiuser interference, low SNR and channel fading can lead to errors in the TOA/TDOA/AOA estimates which in turn can result in large errors in location estimates.

3. WIRELESS RADIO PROPAGATION MODELS

In a wireless propagation medium, signals propagating between the transmitter and receiver may directly reach the destination which means there is line-of-sight (LOS) between the transmitter and the receiver. On the other hand, signals may be scattered by nearby buildings, reflected from the ground or blocked by other obstacles. Generally, wireless propagation models are studied in two main groups: large-scale fading model and small-scale fading model [11].

Large-scale fading models are based on the analysis of the mean signal strength for large separation distances between the transmitter and the receiver. These models are used to obtain estimates for the radio coverage areas of the BTSs. In contrast, rapid changes in the signal strength within short periods of time or over small separation distances are considered as small-scale fading. The motion of a mobile station over very small distances can create rapid signal fluctuations whereas the motion over large distances can be observed as a decrease in the average received signal. Figure 3.1 shows the difference between the small-scale and large-scale fading with respect to signal fluctuations. Free space propagation model is one of the large-scale fading models used to estimate the average received signal strength when there is LOS between the transmitter and the receiver.

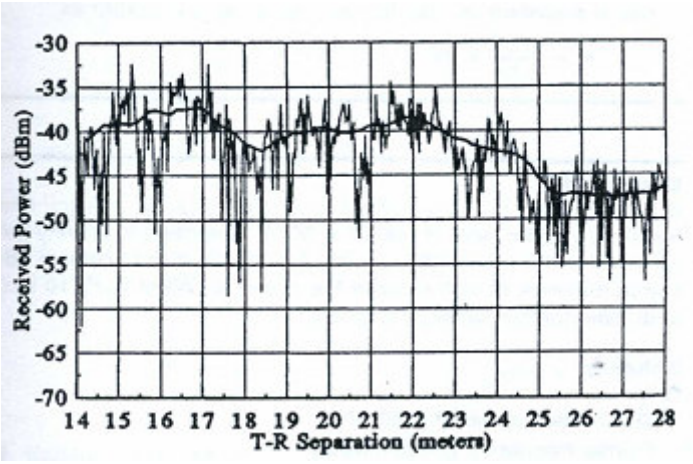


Figure 3.1. Small-scale and large-scale fading [11]

Small-scale fading occurs when two or more different versions of a signal arrive at the receiver with time delays. When these multipath waves are evaluated all together at the receiver, the final signal is observed to be distorted or faded.

3.1. Basic Propagation Mechanisms

There are several propagation mechanisms that can affect wireless signals [12]. These mechanisms cause the signal to have not only small-scale fading but also large-scale fading.

When a base station transmits a signal, it has to travel a distance to reach to a mobile station. Even if there is no atmosphere between the transmitter and the receiver, the power of the transmitted signal decreases with increasing distance. This loss of signal strength is called free space loss. Throughout the atmosphere, weather conditions like rain, snow, fog can also cause some loss in the signal strength. This path loss or attenuation is most significant for long distances.

When signals are blocked by walls, trees or large objects, they undergo extreme attenuation which is called blocking or shadowing. If the object is larger than the wavelength of the incoming signal, then some of the signal power is absorbed and the rest is reflected. In case of no LOS, reflection helps the transmission of signals. However power loss becomes greater as the number of reflections increases. Another mechanism affecting the propagating signals is refraction. Refraction means the bending of the LOS radio waves towards the earth. This effect is caused by the density difference between the higher and lower parts of the atmosphere. Since the atmosphere gets denser near the ground, velocity of the signals is decreased and they are bent towards the ground. Shadowing and reflection are examples of the particle behaviour of light. Shadowing, reflection and refraction mechanisms are shown in Figure 3.2.

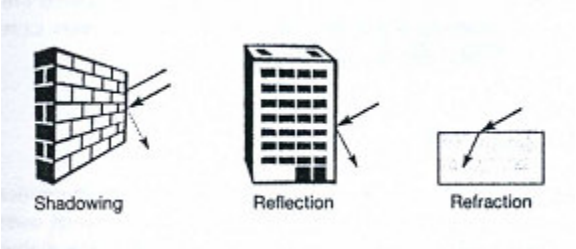


Figure 3.2. Shadowing, reflection and refraction [12]

If the transmitted signals hit obstacles smaller than the signal wavelength, they are broken into weaker signals and scattered. Scattering is likely to occur significantly in dense urban environments where there are many objects near the mobile station. Diffraction is another mechanism where the signals arrive at narrow obstacles like rooftop and are bent at an angle. The deflected signals then propagate in different directions. Scattering and diffraction both occur because of smaller obstacles than the signal wavelength and are examples of the wave behaviour of light. Scattering and diffraction are shown in Figure 3.3.

Figure 3.3. Scattering and Diffraction [12]

All these propagation mechanisms may be effective at a time and they all change with frequency and time. These mechanisms induce small-scale fading over short durations and large-scale fading over long durations of time. It is difficult to model the wireless channel when the case is small-scale fading causing rapid changes in signal parameters.

3.2. Factors Influencing Small-scale Fading

In the propagation medium, there are several factors that can induce small-scale fading.

3.2.1. Multipath Propagation

Propagation mechanisms mentioned above may cause changes in the incoming signal on its way to the receiver. The existence of scattering, reflection, refraction, diffraction and shadowing makes the signal arrive at the receiver from different paths with different amplitudes, phases or delays. When multiple versions of the signal are combined at the receiver, the resulting signal has small-scale fading caused by multipath signals. Multipath propagation often results in intersymbol interference [12]. Figure 3.4 shows intersymbol interference in the multipath environment.

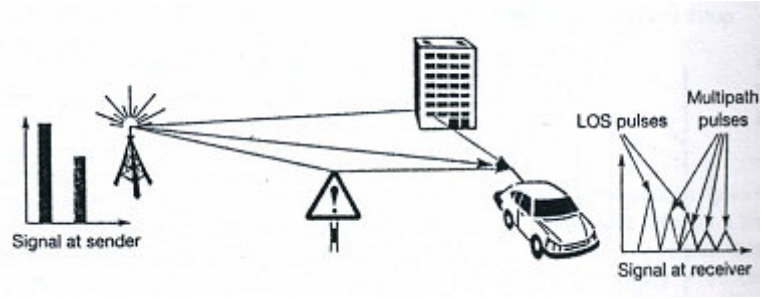


Figure 3.4. Intersymbol Interference [12]

3.2.2. Speed of the Mobile

When the mobile station is moving, the frequency seen by the base station is increased if the motion is towards the direction of arrival of the wave. Otherwise, if the motion is away from the direction of arrival of the wave, then the frequency is decreased. This effect called the Doppler shift is a result of the relative motion between the mobile station and the base station. Frequency of the received signal is changed according to this effect on each of the multipath components.

3.2.3. Speed of Surrounding Objects

Other than the speed of the mobile, Doppler shift may be caused by the speed of the objects in the propagation path. If the speed of objects is higher than the mobile station's speed, then Doppler shift can be considered to arise only from the objects' speed. If the

mobile is moving faster than the objects, then the speed of objects can be ignored. Doppler shift affects the staticness of the channel called the coherence time.

3.2.4. Transmission Bandwidth of the Signal

Small-scale fading is also influenced by the relationship between the transmission bandwidth of the signal and the bandwidth of the multipath channel. If the latter is smaller, then the received signal will have distortion but will not be affected much by fading. Otherwise, if the transmission bandwidth of the signal is smaller, then the signal will not be distorted but the amplitude will change rapidly leading to small-scale fading.

3.3. Types of Small-scale Fading

There are some definitions that are helpful in defining different types of small-scale fading. These are delay spread, coherence time, doppler spread and coherence bandwidth [11].

Delay spread is the result of multipath signals that are versions of the original signal with different arrival times at the receiver. This situation causes the original signal to be spread because of different delays of different signal versions. Delay spread is the time spread between the arrival times of the first and last multipath signals at the receiver. Coherence bandwidth is inversely proportional to delay spread and is defined as the frequency range over which the channel passes all spectral components with approximately equal gain and linear phase. Two frequency components outside this range are affected differently by the channel.

Doppler spread and coherence time are parameters which are useful for analysing the Doppler effect in a channel. The motion is effective in the channel as a shift in the spectral content and the spectral transmission range broadens as a result. Doppler spread is used to describe the amount of frequency shift due to motion. In other words, it is the change in frequency of the source seen by the receiver. Coherence time represents the time-domain version of the Doppler spread and is inversely proportional to it. It is defined as the time duration over which the channel response can be considered unchanging.

3.3.1. Fading due to Multipath Time Delay Spread

Time domain dispersion effects of the multipath channel lead to flat fading or frequency selective fading.

If the same degree of fading occurs for all of the frequency components of the transmitted signal, this type of fading is called flat fading. In flat fading channels, the bandwidth of the transmitted signal is narrow when compared to the coherence bandwidth of the channel. Thus, flat fading channels are also known as narrowband channels. Another characteristic of flat fading channels is that the rms delay spread is much smaller than the symbol period of the signal. The received signal is kept unchanged spectrally but in time domain the signal strength changes with time because of the multipath nature of the channel.

Frequency selective fading occurs if the coherence bandwidth of the channel is smaller than the bandwidth of the transmitted signal. In this case, the delay spread of the multipath channel is greater than the symbol period of the signal leading to intersymbol interference (ISI). Intersymbol interference is an effect induced by the overlapping versions of the incoming signal. Energy of each symbol is spread to the other symbols which results in distortion. In contrast to flat fading, different frequency components have different gains in a frequency selective channel. Wideband channel is another name for a frequency selective channel.

3.3.2. Fading due to Doppler Spread

Doppler spread effects of the multipath channel lead to fast fading or slow fading.

Fast fading is related with the velocity of the mobile or other objects in the channel. If the channel variations are faster compared to the signal variations and the coherence time of the channel is less than the symbol period, this type of distortion is called fast fading. High Doppler spread is a characteristic of fast fading scenario.

Another type of fading related to velocity and frequency dispersion is slow fading. Slow fading channels have a low Doppler spread and the channel variations are slower than the signal variations. Also the coherence time of the channel is greater than the symbol period of the signal. Generally, Doppler spread effects can be ignored at the receiver in a slow fading channel. Figure 3.5 shows types of small-scale fading.

These four types of small-scale fading can also be considered with respect to the baseband signal bandwidth and the transmitted symbol period. Relationships between the signal bandwidth, coherence bandwidth, Doppler spread and also between symbol period, coherence time and delay spread help us determine the final fading type. Figure 3.6 shows these relationships for types of fading.

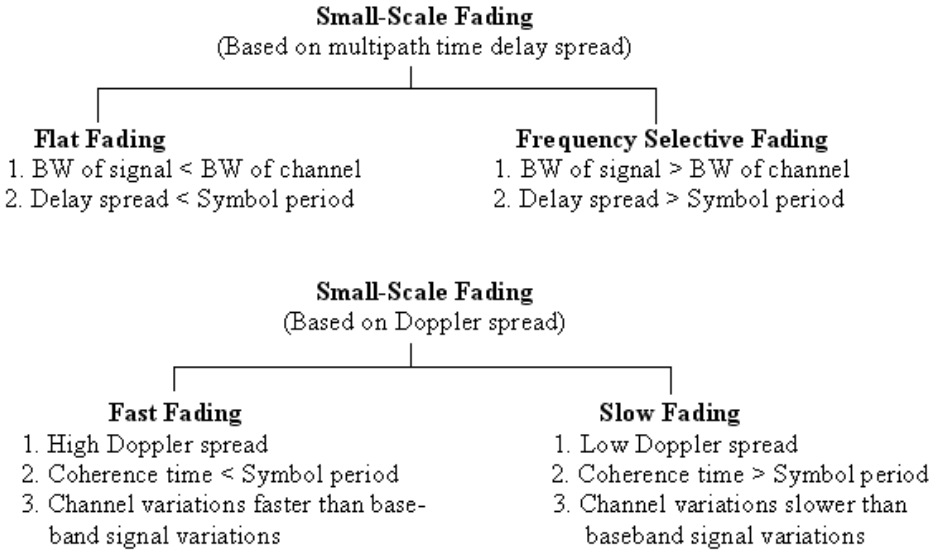


Figure 3.5. Types of small-scale fading [11]

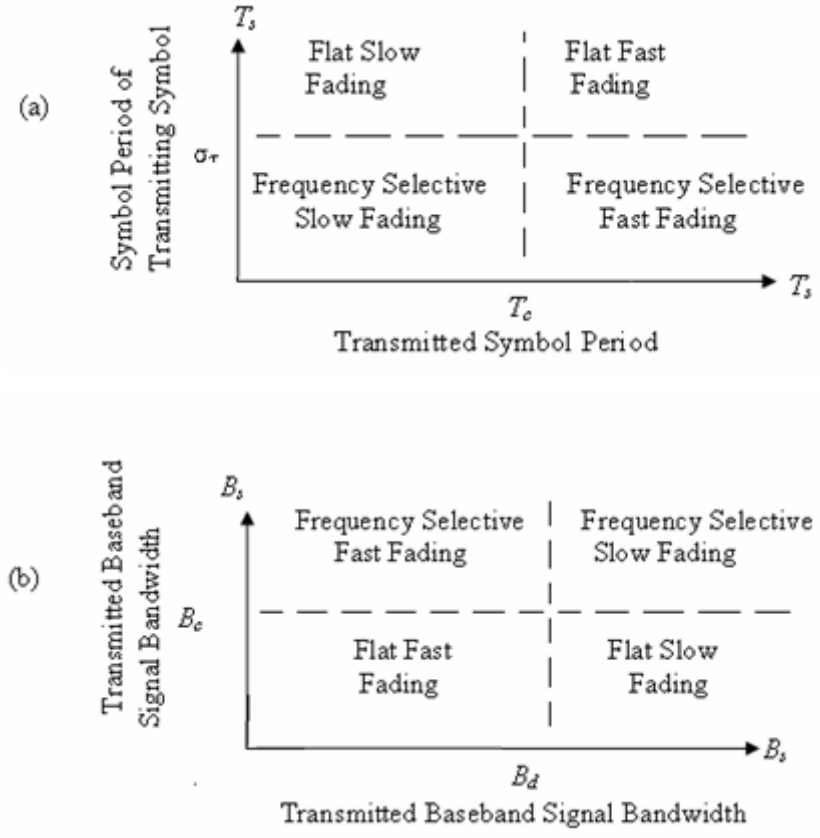


Figure 3.6. Types of fading as a function of: (a) symbol period; and (b) baseband signal bandwidth [11]

3.4. Spatial Fading and Angle Spread

Frequency and time distortion have been the most popular parameters regarding the wireless channel models. However, in recent years smart antennas and spatial channel models have received attention and new models have been formed which consider the effect of angular spreading (spatial fading). Besides doppler spread and delay spread, the third main parameter to cause distortion appears to be angle spread [13, 14]. Angular spread is the spread in arrival angle of the signals due to the scatterers around the mobile transmitter. Since the transmitted signals first hit the scatterers before reaching the base station, the resulting angle of arrival at the base station includes an angular spread caused by the location of the scatterer that reflects the signal. As a consequence, the source is viewed as a spatially distributed source with mean AOA and angular spread instead of a point source.

In order to make the comparison clear the three parameters can be identified as follows. The dispersion the transmitted signal experiences in the frequency domain is indicated with Doppler spread, or equivalently the temporal fading. The delay spread describes the temporal spreading observed at the receiver. In other words, it is equivalent to the frequency dependent fading. Finally, the angle spread describes the angular spreading observed at the receiver. This is equivalent to the spatial dependent fading experienced.

Angular spread depends on the mobile station's environment, on base station's height and also on the distance between the mobile station and the base station. Multipath environment, low base station and small distance between the MS and the BTS may result in higher values of angle spread. Depending on the parameters mentioned, angular spreads up to 10 degrees are commonly encountered in practice [15].

4. SMART ANTENNAS and THE ARRAY MODEL

In a telecommunications system, an antenna is the means of transferring radio frequency (RF) energy from the transmitter to the outside world and to the receiver from the outside world for communication purposes [16]. As wireless technologies and multimedia transfer evolves, antennas are becoming more significant for telecommunications because of the need to use the spectrum efficiently, to acquire better quality of service and to keep the installation costs at affordable levels.

Wireless systems like CDMA2000 and W-CDMA have been proposed under the generic name The International Mobile Telecommunications-2000 (IMT2000) as new generation wireless communication systems. These systems have to focus on extended capacity, superior coverage and better quality while providing low cost voice, data and multimedia services. In order to achieve these targets, new antennas have been developed to overcome the disadvantages of classical omnidirectional and directional antennas.

4.1. Smart Antennas

Smart antennas are composed of several antenna elements to form an array. Digital signal processing software running on these devices makes them smart antennas. Smart antennas can adjust the antenna pattern so that interference, noise and multipath effects can be reduced considerably. Smart antennas or antenna arrays are categorized on the basis of their operation principle as switched beam or adaptive array systems [16].

4.1.1. Switched Beam Antennas

Switched beam antennas are directional antennas with predefined fixed beams in particular directions. These antennas differ from sectorized antennas with the multiple antenna system that combines all the outputs of antenna elements to obtain spatial selectivity before choosing the appropriate beam. In other words, signal strength from each antenna element is detected and one of the predefined beams is chosen so that the received

signal arrives with most power. This switching operation provides better directivity when compared to a conventional antenna.

4.1.2. Adaptive Antenna Arrays

Since switched beam antennas can provide only limited improvement with their fixed beams, adaptive arrays have been proposed. Adaptive antennas have the capability of detecting interference and signal directions. Each antenna array element is associated with a weight that can be adapted according to signals or interference. If the array element senses interference in its look direction, then it changes its weight to minimize its gain. Otherwise, if the element detects signal, then the weight is adapted to maximize the gain in that direction. This way, the antenna pattern changes dynamically to minimize interference and to maximize received signal power. Thus, SNR (signal-to-noise ratio) improvement can be achieved. Adaptive antenna arrays are harder to implement but they are the most advanced smart antenna solution. Figure 4.1 shows the coverage pattern difference between switched beam antenna and adaptive array antenna.

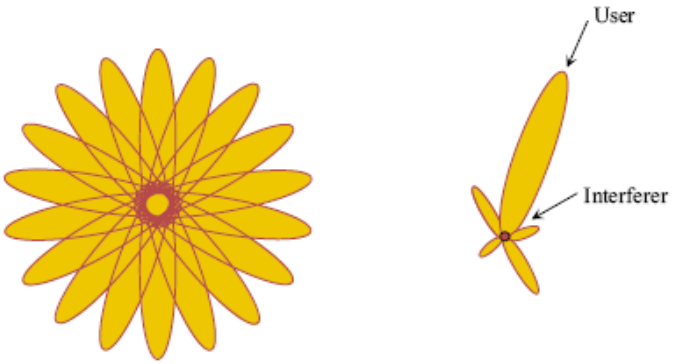


Figure 4.1. Switched beam antenna coverage pattern (on the left) and adaptive antenna array coverage (on the right) [16]

4.2. Uniform Linear Array

Array antennas are used to overcome some difficulties related to antenna characteristics such as directivity, beamwidth, shape of radiation pattern. Single antennas are generally unsuccessful in the attempt to obtain the required characteristics. On the other

hand, array antennas represent the core of smart antenna technology with their ability to produce low side-lobe patterns and desired radiation patterns [17, 18].

Antenna arrays have the advantage of providing acceptable error performance and maximizing the SINR for each user in the system. Multiple elements of the antenna are used to provide power gain and reduce the bit error ratio (BER). Spatial diversity of the antenna leads to improvement in the system coverage and capacity. Co-channel and intersymbol interference effects can also be reduced.

Following are the advantages of using array antennas:

- Number of active users for a given BER quality threshold can be increased.
- Improvement of the BER performance within a cell for a given number of users can be achieved.
- The SINR required at each antenna to achieve a target BER can be reduced leading to a reduction in the transmit power required by the mobile handset for the reverse link.
- The range of the base station receiver and thus the cell size can be increased.
- Reverse link power control can be managed within looser limits while maintaining acceptable BER performance.

Besides these advantages, the problems of cost and complexity may arise for antenna arrays.

- The hardware/software requirements increase with each array element.
- In order to provide effective performance array elements must be synchronized well.
- Array processing algorithms' computational complexity can be disadvantageous.
- Array size of the base station can be limited with the available space.
- Channel modeling errors, calibration errors, phase drift, and noise which is correlated between antennas may cause problems.

A ULA is a series of identical sensors linearly spaced at a distance *d* from each other as shown in Figure 4.2. Since ULA is a one dimensional antenna, it cannot distinguish between the signals arriving with the same angle from the front side and from the back side of the antenna array. This is usually called 'front-back ambiguity' [19]. Also, the ULA can only determine azimuth since it is able to observe the signals on one dimension. The azimuth, the elevation and the distance of the user can all be determined with a three dimensional array configuration. Figure 4.3 shows ULA ambiguity.

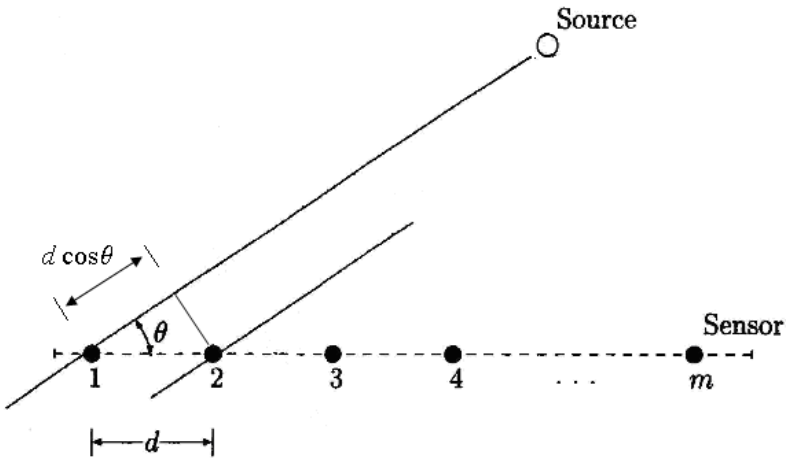


Figure 4.2. ULA (uniform linear array) antenna

In the analysis of an antenna array, the following assumptions are useful for simplicity [16]:

- All signals arriving at the antenna array are made up of finite number of plane waves.
- The transmitter, scatterers and other objects that cause multipath are in the far-field of the antenna array.
- The spacing between the sensors is small enough so that the amplitude of the signals received at different sensors is the same.
- Same radiation pattern and orientation is present for each sensor.
- No mutual coupling exists between antenna elements.

The property that the ULA has uniform antenna spacing leads to an important characteristic. When the signals arrive at the array elements, the time delay between any

two adjacent sensors appears to be the same. If the angle of arrival is assumed to be θ, then with the help of a simple trigonometric calculation, the additional distance the signal travels between adjacent sensors is $dcos\theta$. Selecting the first array element as the reference point and using the assumption of planar wave propagation, time delay between array elements can be expressed as

$$\tau_k = (k-1)\frac{d\cos\theta}{c} \qquad \theta \in [0,\pi] \qquad k = 1,...m \tag{4.1}$$

where θ represents the AOA of the incoming wave and c is the velocity of the wave. The AOA cannot be defined with an angle outside $[0,\pi]$ because the ULA cannot distinguish two signals arriving from symmetric angles with same delays at both sides of the array line.

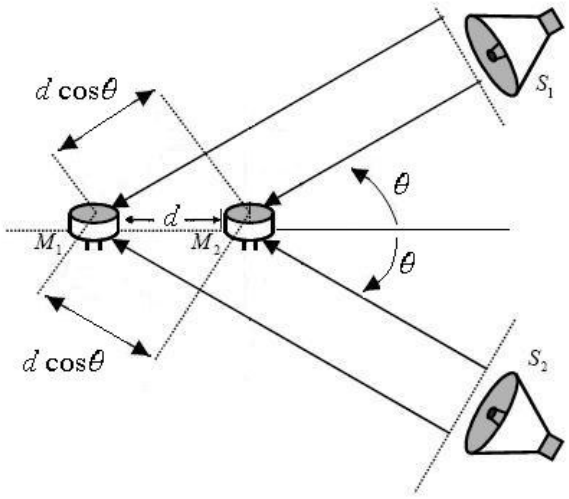


Figure 4.3. Front-back ambiguity [19]

Incident signals arriving at the antenna array are sampled in space and if the antenna elements are not close enough, spatial aliasing may occur [20]. This resembles the aliasing during analog to digital conversion when the sampling rate does not meet the Nyquist criterion. Spatial frequency defined by w_s is a measure of preventing spatial aliasing. Just like the Shannon sampling theorem, aliasing constraint is applicable to the ULAs. Since phase differences above π cause the array to sense incorrect time delays between received signals, spatial frequency should be limited to π.

$$|w_s| \leq \pi \tag{4.2}$$

$$w_s = 2\pi f_s = 2\pi f_c d\cos\theta / c = w_c(d\cos\theta / c) = 2\pi d\cos\theta / \lambda \tag{4.3}$$

Using the above equation, the inequality leads to

$$\frac{2\pi d\,|\cos\theta|}{\lambda} \leq \pi \tag{4.4}$$

$$2d\,|\cos\theta| \leq \lambda \tag{4.5}$$

$$d \leq \lambda / 2 \tag{4.6}$$

As a result, the spatial sampling period d should be smaller than half of the signal wavelength in order to avoid aliasing effects. Spatial sampling can be considered in a similar sense as temporal sampling: the closer sensors are, the more samples per unit distance are taken, analogous to a high sampling rate in time.

4.3. Array Model

Antenna arrays are preferred for location estimation purposes instead of single antennas because they provide spatial sampling with several elements which means more information can be obtained about the incoming waves. Array channel model can be explained by considering the case of a single channel at first sight.

Some assumptions will be made to simplify the channel model. First of all, the signal sources are assumed to be point emitters in the far field of the antenna array and they are assumed to be in the same plane with the antenna elements. Another assumption is that, the propagation medium is not dispersive (i.e., homogeneous) and thus the incoming waves are planar. These assumptions are made to eliminate the need to estimate all three parameters, azimuth, elevation and range for position determination in the general case. From this point of view, the only parameter that defines the source locations is the azimuth angle of arrival. Furthermore, the number of sources n is assumed to be known. Finally, the antenna array is

assumed to be calibrated; that is the array elements can be modeled as linear (time-invariant) systems and their transfer characteristics and locations are known.

Antenna arrays sample the incoming waves in both space and time. Supposing a single waveform incident on the *k*th (*k*=1,2,....,*m*) array element, the output of sensor *k* can be expressed as

$$\overline{y}_k(t) = \overline{h}_k(t) * x(t-\tau_k) + \overline{e}_k(t) \tag{4.7}$$

In the equation above, $x(t)$ is assumed to be the incoming wave measured on some reference point and τ_k indicates the travel time of the wave from the reference point to sensor *k*. Also $\overline{h}_k(t)$ is the impulse response of the *k*th element, and $\overline{e}_k(t)$ is additive noise. The reference point may be an element of the antenna array or some other point near the array that fits the planar wave assumption. The unknown parameters in the equation above are the incoming wave $x(t)$ and the time delay τ_k which also indicates the location of the source.

In order to obtain a more simplified model from the equation above, the signals will be assumed to be narrowband with the following steps [20]. Using the Fourier transform notation, the output signal can be expressed as

$$\overline{Y}_k(w) = \overline{H}_k(w) X(w) e^{-jw\tau_k} + \overline{E}_k(w) \tag{4.8}$$

where $\overline{Y}_k(w)$, $\overline{H}_k(w)$, $X(w)e^{-jw\tau_k}$ and $\overline{E}_k(w)$ are the Fourier transforms of $\overline{y}_k(t)$, $\overline{h}_k(t)$, $x(t-\tau_k)$ and $\overline{e}_k(t)$ respectively.

Let the received signal $x(t)$ be a bandpass signal with an even spectrum so that $x(t)$ is real valued. The demodulated form of $x(t)$ can be represented with a baseband complex-valued signal $s(t)$ whose spectrum may not be even because of an asymmetric

transmission channel. Figure 4.4 and Figure 4.5 show a bandpass signal spectrum and a baseband signal spectrum respectively.

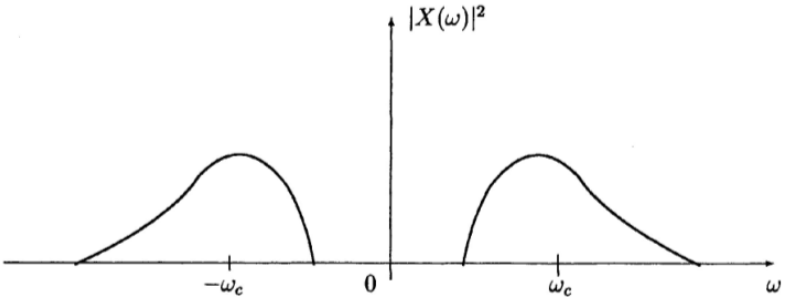


Figure 4.4. Spectrum of a bandpass signal [20]

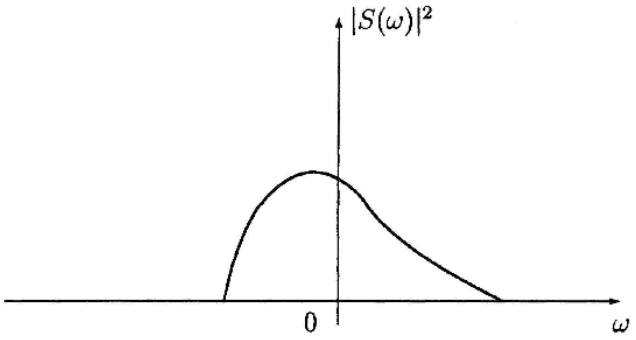


Figure 4.5. Spectrum of the baseband signal that forms into the bandpass signal above [20]

The right-shifted form of $s(t)$ in the frequency domain can be obtained by complex modulation process.

$$\int_{-\infty}^{\infty} s(t)e^{jw_c t}e^{-jwt}dw = \int_{-\infty}^{\infty} s(t)e^{-j(w-w_c)t}dw = S(w-w_c) \tag{4.9}$$

Since the received signal $x(t)$ is real-valued, it can be expressed as the sum of the right-shifted version of $S(w)$ and left-shifted version of folded complex-conjugated $S(w)$.

$$X(w) = S(w-w_c) + S^H(-(w+w_c)) \tag{4.10}$$

Inserting the above equation into (4.8) leads to

$$\overline{Y}_k(w) = \overline{H}_k(w)[S(w-w_c) + S^H(-w-w_c)]e^{-jw\tau_k} + \overline{E}_k(w) \tag{4.11}$$

Considering the demodulation process in the antenna array, the demodulated signal and its Fourier transform are as follows:

$$\tilde{y}_k(t) = \overline{y}_k(t)e^{-jw_c t} \tag{4.12}$$

$$\tilde{Y}_k(w) = \tilde{H}_k(w+w_c)[S(w) + S^H(-w-2w_c)]e^{-j(w+w_c)\tau_k} + \tilde{E}_k(w+w_c) \tag{4.13}$$

The demodulated signal $\tilde{y}_k(t)$ then goes through a lowpass filter of bandwidth matched to $S(w)$ and the components centered outside the cutoff frequency are left out. The resulting output signal is given as

$$Y_k(w) = H_k(w+w_c)S(w)e^{-j(w+w_c)\tau_k} + E_k(w+w_c) \tag{4.14}$$

In the equation above, it can be observed that when $\tilde{H}_k(w+w_c)$ and $\tilde{E}_k(w+w_c)$ are passed through the filter, the parts that stay within the passband Ω are shown as $H_k(w+w_c)$ and $E_k(w+w_c)$.

Another assumption will be made to finalize the output signal form. The received signals are assumed to be narrowband which means that as $|w|$ increases, $|S(w)|$ decreases rapidly. The output signal form will then include $H_k(w_c)$ but $E_k(w+w_c)$ will not be changed to $E_k(w_c)$ since noise may not be constant over the bandwidth. The output signal can be expressed as below in the frequency and time domain.

$$Y_k(w) = H_k(w_c)S(w)e^{-jw_c\tau_k} + E_k(w+w_c) \qquad w \in \Omega \tag{4.15}$$

$$y_k(t) = H_k(w_c)e^{-jw_c\tau_k}s(t) + e_k(t) \tag{4.16}$$

where $e_k(t)$ represents the inverse Fourier transform of $E_k(w+w_c)$ for ease of use.

The received signal is converted to digital form after demodulation process so that digital data can be used for AOA estimation.

The array output can be shown with a closed form including all the elements of the array as

$$\boldsymbol{y}(t) = \boldsymbol{a}(\theta)\boldsymbol{s}(t) + \boldsymbol{e}(t) \tag{4.17}$$

where $\boldsymbol{a}(\theta)$, the array transfer function is expressed as

$$\boldsymbol{a}(\theta) = [H_1(w_c)e^{-jw_c\tau_1} \quad \dots \quad H_m(w_c)e^{-jw_c\tau_m}]^T \tag{4.18}$$

Also the array's output vector and the additive noise vector can be written as

$$\boldsymbol{y}(t) = [y_1(t) \quad \dots \quad y_m(t)]^T \tag{4.19}$$

$$\boldsymbol{e}(t) = [e_1(t) \quad \dots \quad e_m(t)]^T \tag{4.20}$$

In some cases $\{H_k(w_c)\}_{k=1}^m$ can be considered independent of θ if omnidirectional antennas are used. Also the array elements may be considered identical which means similar $H_k(w_c)$ functions. From this point of view, if $H(w_c)$ is included in the expression of $s(t)$ (i.e., $s(t)$ represents $H(w_c)s(t)$) and the first array element is chosen the reference point, $\boldsymbol{a}(\theta)$ can be expressed as

$$\boldsymbol{a}(\theta) = [1 \quad e^{-jw_c\tau_2} \quad \dots \quad e^{-jw_c\tau_m}]^T \tag{4.21}$$

Using the delay expression in (47), $\boldsymbol{a}(\theta)$ can be written as

$$\boldsymbol{a}(\theta) = [1 \quad e^{-j2\pi f_c d\cos\theta/c} \quad \ldots \quad e^{-j2\pi f_c d(m-1)\cos\theta/c}]^T \tag{4.22}$$

Let the signal wavelength be defined by $\lambda = c/f_c$. Then the function becomes

$$\boldsymbol{a}(\theta) = [1 \quad e^{-j2\pi d\cos\theta/\lambda} \quad \ldots \quad e^{-j2\pi d(m-1)\cos\theta/\lambda}]^T \tag{4.23}$$

or

$$\boldsymbol{a}(\theta) = [1 \quad e^{-jw_s} \quad \ldots \quad e^{-j(m-1)w_s}]^T \tag{4.24}$$

where $w_s = 2\pi d\cos\theta/\lambda$

This model can be extended for the general case where *n* signals are incident on *m* array elements

$$\boldsymbol{y}(t) = \begin{bmatrix} a(\theta_1) & \ldots & a(\theta_n) \end{bmatrix} \begin{bmatrix} s_1(t) \\ \vdots \\ s_n(t) \end{bmatrix} + \boldsymbol{e}(t) \triangleq \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{e}(t) \tag{4.25}$$

where θ_k is the AOA of the *k*th source and $s_k(t)$ is the signal corresponding to the *k*th source.

5. NONISOTROPIC SCATTERING and THE VON MISES MODEL

Isotropic scattering model which indicates that the AOA of scatterers is distributed uniformly around the mobile station has been widely studied in location estimation models [21]. However recent researches have shown that non-isotropic scattering model is more suitable to model the multipath propagation environment since the scatterers are often distributed unevenly and the MS is not supposed to received signals form every direction all the time [21, 22]. Another reason is that for macrocells, the BTS antennas are usually mounted at a higher level than the surrounding scatterers and the received signal at the BTS mainly results from the local scatterers around the MS.

The multipath environment for the mobile station can be modeled as in Figure 5.1 where the user is at distance *D* from the base station ULA. The user is surrounded by scatterers which affect the signals propagated from the MS towards the ULA. As a result of the multipath propagation, the ULA cannot determine the location of the user exactly via the received signals.

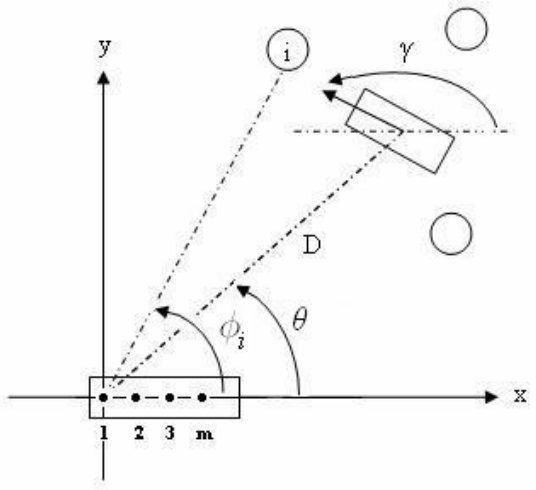


Figure 5.1. Mobile station surrounded by local scatterers and signals reaching the ULA

Richard von Mises proposed a continuous probability distribution function to examine why the atomic weights differ from exact integer values that are multiples of hydrogen in 1918. This pdf is used especially in circular statistics and statistical modeling to describe the density of a circular random variable with period 2π . It is also known as the circular normal distribution because of its similarity to the Gaussian distribution in the circular sense. The von Mises pdf can be used to model several distributions like uniform, cardioid and Gaussian [21]. This flexible pdf is also useful in modeling the distribution of nonisotropic scatterers around a mobile station which is significant to determine the position of the MS via AOA calculation. The angular nature of the pdf can be used to obtain a spatial model of the correlation function.

The von Mises probability distribution function for the angle θ is given by:

$$f(\phi) = \frac{\exp[\kappa \cos(\phi - \theta)]}{2\pi I_0(\kappa)} \qquad \phi \in [-\pi, \pi), \quad \theta \in [-\pi, \pi), \quad E[\phi] = \theta \tag{5.1}$$

where θ is the mean direction of AOA of scatter components seen by the BTS, ϕ indicates the scatterer position, $\kappa = D^2 / \sigma^2 \geq 0$ is a measure of the distribution of the scatterers around the mean direction and $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind defined by

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \exp(\kappa \cos \phi) d\phi \tag{5.2}$$

As seen from Figure 5.1, *D* is the distance between the mobile station and the ULA and σ is a measure of the size of the local scattering area around the MS. κ is also known as the “concentration parameter” and is inversely proportional with the angle spread.

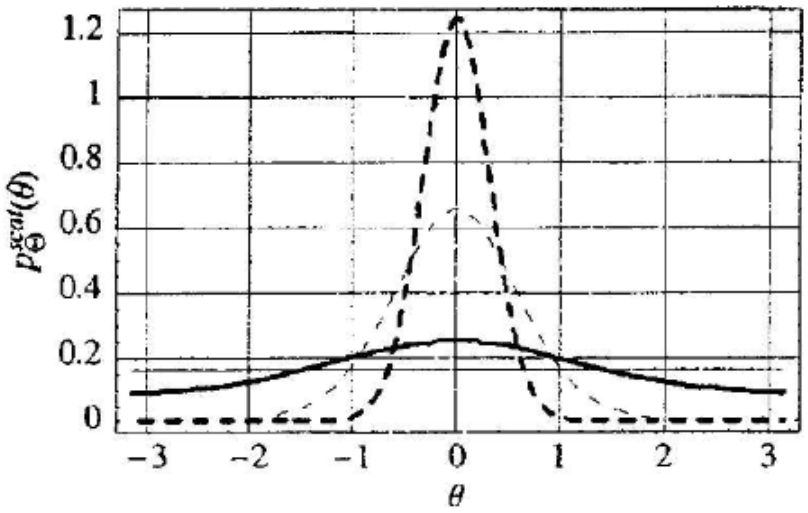


Figure 5.2. Von Mises pdf for the AOA of scatter components at the mobile station ($\theta = 0$, — $\kappa = 0$, — $\kappa = 0.5$, ···· $\kappa = 3$, ••• $\kappa = 10$) [21]

The effect of the κ parameter can be observed in Figure 5.2. For $\kappa = 0$ the scatterers have a uniform distribution (isotropic scattering).

$$I_0(0) = 1 \tag{5.3}$$

$$f(\phi) = 1/2\pi \tag{5.4}$$

As κ increases the scatterers get more concentrated around the mean direction θ. The von Mises pdf resembles the cardioid pdf for small κ.

$$I_0(\kappa) \approx 1 \tag{5.5}$$

$$\exp\left[\kappa\cos(\phi-\theta)\right] \approx 1+\kappa\cos(\phi-\theta) \tag{5.6}$$

$$f(\phi) \approx \frac{1+\kappa\cos(\phi-\theta)}{2\pi} \tag{5.7}$$

For large κ, the von Mises pdf approximates the Gaussian pdf with mean θ and standard deviation $1/\sqrt{\kappa}$ given as

$$f(\phi) = \frac{\exp\left[\frac{-(\phi-\theta)^2}{2\sigma^2}\right]}{\sigma\sqrt{2\pi}} \tag{5.8}$$

where $\sigma^2 = 1/\kappa$ [23].

The distribution is extremely nonisotropic for $\kappa = \infty$ which leads to

$$f(\phi) = \delta(\phi - \theta) \tag{5.9}$$

where $\delta(.)$ is the Dirac delta function.

When κ is not too small, angle spread can be expressed as $\chi = 2/\sqrt{\kappa}$ in radians [24]. For $\chi \leq 40^\circ$, the concentration parameter is $\kappa \geq 8.2$ and this condition makes the von Mises and Gaussian pdfs indistinguishable.

5.1. Correlation of the Complex Envelope of Multipath Components

Let us consider a user surrounded by *K* scatterers as in Figure 5.1. There is no line-of-sight between the user and the BTS in the model which is known as the model of Rayleigh fading [11]. The BTS consists of a uniform linear array antenna with *m* elements in the far field of the user. Also the antenna elements are omnidirectional unit-gain elements with element spacing *d*. The *i*-th scatterer at the angular position ϕ_i is moving in the direction γ with speed $\lambda f_{\max}$ at time instant *t*. If a unit-power unmodulated carrier is transmitted by the user, the complex envelope lowpass equivalent of the signal at the *l*-th element is given by [24]

$$z_l(t) = \lim_{K\to\infty} \frac{1}{\sqrt{K}} \sum_{i=1}^{K} \Xi_i \exp\{j\psi_i\} \exp\{-j2\pi f_{\max}[\cos(\phi_i - \gamma)]t + j2\pi(l-1)(d/\lambda)\cos(\phi_i)\} \tag{5.10}$$

where $1 \le l \le m$.

In the equation above, Ξ_i is the random reflection coefficient and ψ_i is the phase shift of the *i*-th scatterer. λ is the wavelength whereas $f_{\max}$ represents the maximum Doppler frequency.

The reflection coefficients $\{\Xi_i\}$ are independent positive random variables with finite variances, independent of the phase shift values $\{\psi_i\}$. Phase shift values are independent and identically distributed random variables which are distributed uniformly over $[-\pi, \pi)$.

The received signal is a unit-power signal, that is $E[|z_l(t)|^2] = 1$, which is satisfied by setting $\lim_{K\to\infty} \frac{1}{K} \sum_{i=1}^{K} E[\Xi_i^2] = 1$. $z_l(t)$ is a zero-mean complex Gaussian process since it consists of independent and identically distributed random variables with finite variances according to the central limit theorem. Assuming that $\{\Xi_i\}$ and $\{\psi_i\}$ are independent of $\{\phi_i\}$, the cross-correlation between $z_l(t)$ and $z_q(t)$, $1 \le l \le q \le m$, is given by

$$
\begin{aligned}
&E[z_l(t) z_q^{*}(t+\tau)] \\
&= E\left[\lim_{K\to\infty} \frac{1}{K} \left(\sum_{i=1}^{K} \Xi_i \exp\{j\psi_i\} \exp\{-j2\pi f_{\max}[\cos(\phi_i - \gamma)]t + j2\pi(l-1)(d/\lambda)\cos(\phi_i)\} \right) \right. \\
&\qquad \left. \left(\sum_{j=1}^{K} \Xi_j \exp\{-j\psi_j\} \exp\{j2\pi f_{\max}[\cos(\phi_j - \gamma)](t+\tau) - j2\pi(q-1)(d/\lambda)\cos(\phi_j)\} \right) \right]
\end{aligned}
$$

$$
\begin{aligned}
&= \lim_{K\to\infty} \frac{1}{K} \sum_{i=1}^{K} \sum_{j=1}^{K} E[\Xi_i \Xi_j] E[\exp\{j\psi_i - j2\pi f_{\max}[\cos(\phi_i - \gamma)]t \\
&+ j2\pi(l-1)(d/\lambda)\cos(\phi_i) - j\psi_j + j2\pi f_{\max}[\cos(\phi_j - \gamma)](t+\tau) \\
&- j2\pi(q-1)(d/\lambda)\cos(\phi_j)\}]
\end{aligned}
$$

$$= \lim_{K\to\infty} \frac{1}{K} \sum_{i=1}^{K} \sum_{j=1}^{K} E[\Xi_i^{\,2}]\ \delta(i-j)\, E[\exp\{ j\psi_i - j2\pi f_{\max}[\cos(\phi_i - \gamma)]t$$
$$+ j2\pi(l-1)(d/\lambda)\cos(\phi_i) - j\psi_j + j2\pi f_{\max}[\cos(\phi_j - \gamma)](t+\tau)$$
$$- j2\pi(q-1)(d/\lambda)\cos(\phi_j)\}]$$

$$= \lim_{K\to\infty} \frac{1}{K} \sum_{i=1}^{K} E[\Xi_i^{\,2}]\, E[\exp\{ j2\pi f_{\max}[\cos(\phi_i - \gamma)]\tau + j2\pi(l-q)(d/\lambda)\cos(\phi_i)\}] \tag{5.11}$$

Each obstacle around the mobile user scatters some part of the total transmitted power towards the BTS. The power received from the i-th scatterer can be expressed as $E[\Xi_i^{\,2}]/K$. If there are large number of scatterers, this power is found out to be equal to the infinitesimal power coming from the differential angle $d\phi$, centered at ϕ_i and with probability $f_\phi(\phi_i)d\phi$.

$$E[\Xi_i^{\,2}]/K = f_\phi(\phi_i)d\phi \tag{5.12}$$

Using this equation in (5.11), the cross-correlation becomes,

$$E[z_l(t)z_q^{\,*}(t+\tau)] = \int_{-\pi}^{\pi} \exp\{ j2\pi f_{\max}[\cos(\phi-\gamma)]\tau + j2\pi(l-q)(d/\lambda)\cos(\phi)\}\, f_\phi(\phi)d\phi \tag{5.13}$$

It is suitable to use the von Mises pdf in the above pdf expression since it is an angular distribution. After substitution of (5.1) into the equation and carrying out the integral with the help of integral tables, the spatial correlation function is found as

$$E[z_l(t)z_q^{\,*}(t+\tau)] = \frac{I_0\left(\sqrt{\kappa^2 - \eta^2 - \vartheta^2 - 2\eta\vartheta\cos(\gamma) + j2\kappa[\eta\cos(\gamma-\theta) + \vartheta\cos(\theta)]}\right)}{I_0(\kappa)} \tag{5.14}$$

where $\eta = 2\pi f_{\max}\tau$ and $\vartheta = 2\pi(l-q)(d/\lambda)$.

In order to simplify the case, the user is assumed to be stationary which means that $f_{\max} = 0$. The final equation then becomes

$$E[z_l(t)z_q^{*}(t)] = \boldsymbol{R}_{\boldsymbol{Z}}((q-l)(d/\lambda)) = \frac{I_0\left(\sqrt{\kappa^2 - \vartheta^2 - j2\vartheta\kappa\cos(\theta)}\right)}{I_0(\kappa)} \quad (5.15)$$

If the channel model is considered to be narrowband, the received signal at the *l*-th element of the antenna array can be written as

$$y_l(t) = \sqrt{P_{signal}}\, z_l(t) + \sqrt{P_{noise}}\, e_l(t) \quad (5.16)$$

where P_{signal} is the signal power, P_{noise} is the noise power and $e_l(t)$ is the independent zero-mean unit-variance and spatio-temporally white Gaussian noise of the *l*-th array element.

Generalizing in the vector formation, the cross-correlation of the received signals can be expressed as

$$\boldsymbol{R}_{\boldsymbol{y}} = P_{signal}\,\boldsymbol{R}_{\boldsymbol{z}} + P_{noise}\,\boldsymbol{I} \quad (5.17)$$

where $\boldsymbol{R}_{\boldsymbol{y}} = E[\boldsymbol{y}(t)\boldsymbol{y}^H(t)]$, $\boldsymbol{R}_{\boldsymbol{z}} = E[\boldsymbol{z}(t)\boldsymbol{z}^H(t)]$ and $E[\boldsymbol{e}(t)\boldsymbol{e}^H(t)] = \boldsymbol{I}$. With this notation, the matrix elements are given as $(\boldsymbol{R}_{\boldsymbol{y}})_{lq} = \boldsymbol{R}_{\boldsymbol{y}}((q-l)\Delta)$ and $(\boldsymbol{R}_{\boldsymbol{z}})_{lq} = \boldsymbol{R}_{\boldsymbol{z}}((q-l)\Delta)$.

5.2. The Case of Multiple Users

If multiple AOA estimations are required, then the angular distribution can be modeled as a mixture of von Mises distributions [25, 26]. Each von Mises distribution corresponds to a cluster of scatterers around one mobile user. The composite distribution can be written as

$$f(\phi) = \sum_{p=1}^{P} \varepsilon_p \frac{\exp\left[\kappa_p \cos(\phi - \theta_p)\right]}{2\pi I_0(\kappa_p)} \tag{5.18}$$

where P is the number of clusters and ε_p are unknown mixture proportions the sum of which is equal to 1.

$$\sum_{p=1}^{P} \varepsilon_p = 1 \tag{5.19}$$

If this distribution is used to generate the cross-correlation function for the multiple user case, the result is as

$$\boldsymbol{R}_z((q-l)(d/\lambda)) = \sum_{p=1}^{P} \varepsilon_p \frac{I_0\left(\sqrt{{\kappa_p}^2 - \vartheta^2 - j2\vartheta\kappa_p \cos(\theta_p)}\right)}{I_0(\kappa_p)} \tag{5.20}$$

where $\vartheta = 2\pi(l-q)(d/\lambda)$.

5.3. Received Signal Estimates for the von Mises Model

The signal model presented in section 5.1 can be used to obtain an estimate for the received signals at the base station antenna array.

The cross-correlation function of the received signals was shown to be

$$\boldsymbol{R}_y = P_{signal}\,\boldsymbol{R}_z + P_{noise}\,\boldsymbol{I} \tag{5.21}$$

The aim is first to obtain the model of received signals and then use it to generate an estimate of the received cross-correlation function.

$\boldsymbol{R}_z$ can be factorized using the LDL factorization so that a received signal model with the same statistical characteristic as the previously stated received signal model in (4.25) is obtained as

$$\boldsymbol{R}_y = P_{signal}\,\boldsymbol{R}_z + P_{noise}\,\boldsymbol{I} = P_{signal}\left(\boldsymbol{T}\boldsymbol{T}^H\right) + P_{noise}\,\boldsymbol{I} \tag{5.22}$$

$$\boldsymbol{y} = \sigma_s\,\boldsymbol{T}\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2 \tag{5.23}$$

where $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ are vectors consisting of zero-mean complex Gaussian random variables. Although $\sigma_s\,\boldsymbol{T}$ and $\boldsymbol{\rho}_1$ do not correspond exactly to $\boldsymbol{a}(\theta)$ and $\boldsymbol{s}(t)$, received signal statistics are the same as seen below.

$$E[(\sigma_s\,\boldsymbol{T}\boldsymbol{\rho}_1)(\sigma_s\,\boldsymbol{T}\boldsymbol{\rho}_1)^H] = \boldsymbol{T}\boldsymbol{T}^H = E[\boldsymbol{a}(\theta)\boldsymbol{a}(\theta)^H] \tag{5.24}$$

As a result, the received cross-correlation estimate can be written as

$$\hat{\boldsymbol{R}}_y = \frac{1}{N}\sum_{t=1}^{N}\boldsymbol{y}(t)\boldsymbol{y}^H(t) \tag{5.25}$$

The estimate above can be used in location estimation algorithms to obtain AOA estimates for the mobile users.

6. AOA ESTIMATION ALGORITHMS

6.1. Conventional Methods

Conventional estimation methods estimate the angle of arrival by computing the spatial spectrum and then determining the local maximas. These methods are based on spatial filtering or beamforming.

6.1.1. Conventional Beamforming (CBF)

Beamforming is one of the basic methods of AOA estimation and it provides the basis of following methods like "Capon's Minimum Variance Method" and MUSIC. The basic logic of AOA estimation is steering beams in all possible directions to find peaks in the output power. Peaks help to indicate the estimated AOA for the incoming signals [11].

In the classical beamforming structure the sensor outputs are multiplied by weights and then are summed up to form the output signal [27]. A spatial filter can be represented with a similar block diagram to that of a temporal filter as shown in Figure 6.1. A temporal filter can be used to enhance or attenuate some frequency components with suitable choice of the parameters. An FIR filter output for a sinusoidal input $\boldsymbol{u}(t)$ can be expressed as

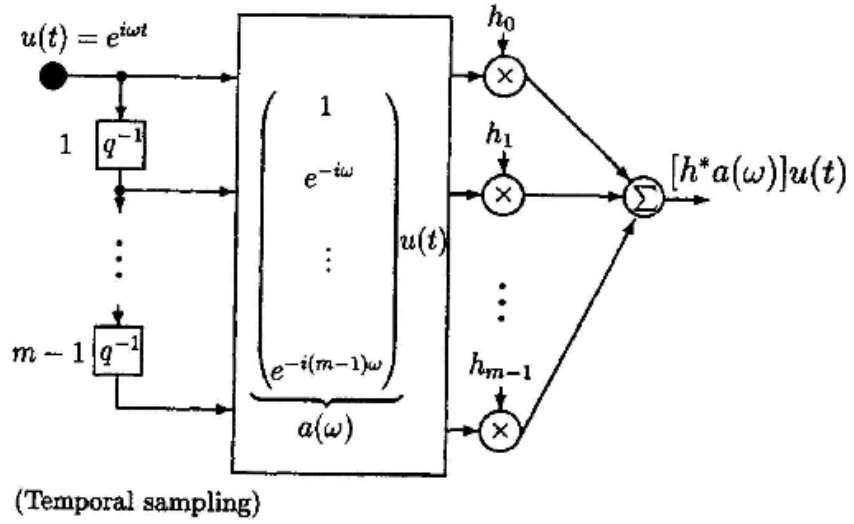
$$\boldsymbol{y}_F(t) = [\boldsymbol{h}^H \boldsymbol{a}(w)]\boldsymbol{u}(t) \tag{6.1}$$

where $\boldsymbol{a}(w)$ is given by

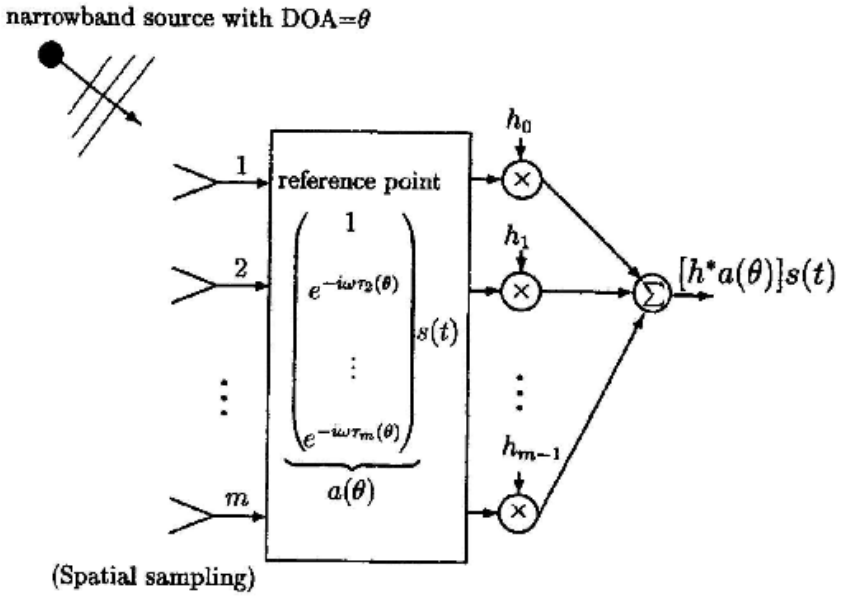
$$\boldsymbol{a}(w) = \begin{pmatrix} 1 & e^{-jw} & \dots & e^{-j(m-1)w} \end{pmatrix}^T \tag{6.2}$$

and $\boldsymbol{h}$ is

$$\boldsymbol{h} = [h_0 \quad \dots \quad h_{m-1}]^H \tag{6.3}$$



(a) Temporal filter



(b) Spatial filter

Figure 6.1. (a) Temporal filter and (b) spatial filter [20]

If $\boldsymbol{h}^H \boldsymbol{a}(w)$ is chosen large, the output power at frequency w can be increased and if it is chosen small, the power can be attenuated. For the spatial case, a spatial filter can be used to enhance or attenuate the signals with a specific AOA by selecting a large or small $\boldsymbol{h}^H \boldsymbol{a}(\theta)$ [20]. The output of a spatial filter can be expressed as

$$\boldsymbol{y}_F(t) = [\boldsymbol{h}^H \boldsymbol{a}(\theta)]\boldsymbol{s}(t) \tag{6.4}$$

where

$$\boldsymbol{a}(\theta) = \begin{pmatrix} 1 & e^{-jw\tau_2(\theta)} & \dots & e^{-jw\tau_m(\theta)} \end{pmatrix}^T \tag{6.5}$$

Then the output power is given by

$$E\left\{ | \boldsymbol{y}_F(t) |^2 \right\} = \boldsymbol{h}^H \boldsymbol{R}_y \boldsymbol{h} \tag{6.6}$$

The autocorrelation function of the received data can be given as below with respect to (4.25).

$$\boldsymbol{R}_y = E\left\{ \boldsymbol{y}(t) \boldsymbol{y}^H(t) \right\} \tag{6.7}$$

CBF estimates the mean power arriving from various directions by steering the array in those directions. $\boldsymbol{a}(\theta)$ denotes the response of the array to a unit power source in direction θ. This process is similar to steering the array mechanically in this direction and measuring the output power. However, the output power is not only contributed from the direction in which the array is steered. The directions where the sidelobes are pointing contribute to the power as well. Conventional beamforming is based on two conditions to be satisfied by the spatial filter:

(i). A distortionless response should be obtained from a specific AOA θ.

(ii). The filter should minimize the output power as much as possible so that signals with AOAs other than θ are attenuated.

The filters which satisfy the conditions above form a beam pointing the direction of the signal AOA; an example can be seen in Figure 6.2. The graph shows $\left|\boldsymbol{h}^H \boldsymbol{a}(\theta)\right|$ versus AOA for a signal arriving with $\theta = 25^\circ$.

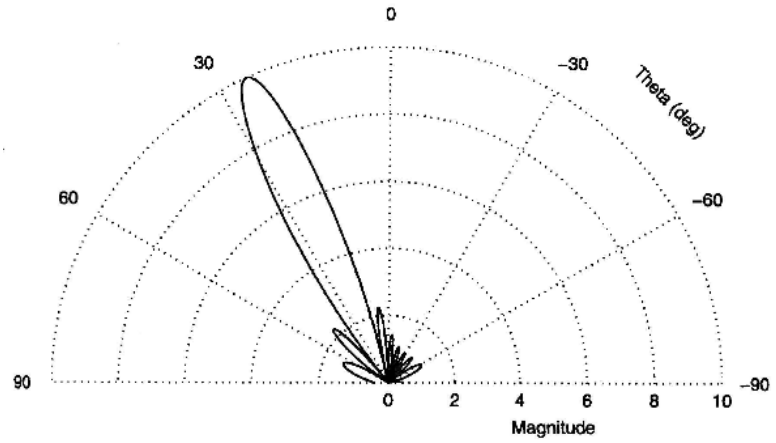


Figure 6.2. Beamforming ($\theta = 25^\circ$) [20]

Regarding the first condition, a distortionless response can be provided by setting $\boldsymbol{h}^H \boldsymbol{a}(\theta) = 1$.

For the second condition, it is assumed that the signal power incident on the array is equal at all angles. In other words, $\boldsymbol{y}(t)$ consists of spatially white signals with $\boldsymbol{R}_y = \boldsymbol{I}$. Then the output power that is to be minimized is given as

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \boldsymbol{h}^H \boldsymbol{h} \tag{6.8}$$

The two conditions explained above can now be redefined as:

$\min_h \boldsymbol{h}^H \boldsymbol{h}$ subject to $\boldsymbol{h}^H \boldsymbol{a}(\theta) = 1$.

This optimization problem can be solved using Lagrange multiplier.

$$\varepsilon = f + \lambda g \tag{6.9}$$

$$\varepsilon = E\left\{| \boldsymbol{y}_F(t) |^2\right\} + \lambda\left(\boldsymbol{h}^H \boldsymbol{a}(\theta) - 1\right) \tag{6.10}$$

$$\varepsilon = \boldsymbol{h}^H \boldsymbol{h} + \lambda\left(\boldsymbol{h}^H \boldsymbol{a}(\theta) - 1\right) \tag{6.11}$$

$$\nabla\varepsilon = \nabla f + \lambda \nabla g = 0 \tag{6.12}$$

$$\nabla\varepsilon = \frac{\partial \varepsilon}{\partial \boldsymbol{h}^H} = \boldsymbol{h} + \lambda \boldsymbol{a}(\theta) = 0 \tag{6.13}$$

Solving for $\boldsymbol{h}$,

$$\boldsymbol{h} = -\lambda \boldsymbol{a}(\theta) \tag{6.14}$$

Using the equation above in the first constraint

$$\boldsymbol{h}^H \boldsymbol{a}(\theta) = 1 \tag{6.15}$$

$$-\lambda^H \boldsymbol{a}^H(\theta)\boldsymbol{a}(\theta) = 1 \tag{6.16}$$

Assuming the array transfer function consists of identical sensors the following condition will be met:

$$\boldsymbol{a}^H(\theta)\boldsymbol{a}(\theta) = m \tag{6.17}$$

This equation can be used in (6.16) to obtain

$$-\lambda^H m = 1 \tag{6.18}$$

Then the parameter λ can be found as

$$\lambda = -1/m \tag{6.19}$$

Finally $\boldsymbol{h}$ can be obtained using this value of λ in (6.14)

$$\boldsymbol{h} = -\lambda \boldsymbol{a}(\theta) \tag{6.20}$$

$$\boldsymbol{h} = \boldsymbol{a}(\theta)/m \tag{6.21}$$

The output power has to be checked to look for peaks in specific directions. When $\boldsymbol{h}$ is inserted into the power equation

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \boldsymbol{h}^H \boldsymbol{R}_y \boldsymbol{h} \tag{6.22}$$

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \boldsymbol{a}^H(\theta) \boldsymbol{R}_y \boldsymbol{a}(\theta)/m^2 \tag{6.23}$$

The exact value of the correlation function used in this formula cannot be found by $\boldsymbol{R}_y = E\left\{\boldsymbol{y}(t)\boldsymbol{y}^H(t)\right\}$ since there are finite number of signals available and there is no information about the probability distribution function of signals. Instead, an estimate of the correlation function can be used by taking into account the number of samples.

$$\hat{\boldsymbol{R}}_y = \frac{1}{N}\sum_{t=1}^{N} \boldsymbol{y}(t)\boldsymbol{y}^H(t) \tag{6.24}$$

Rearranging the power equation with the correlation estimate and omitting $1/m^2$ for normalization,

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \boldsymbol{a}^H(\theta) \hat{\boldsymbol{R}}_y \boldsymbol{a}(\theta) \tag{6.25}$$

AOAs of n incident signals on the ULA can be estimated by finding the n highest peaks of the output power given above.

If a multipath model is considered, then using von Mises pdf can be another choice to obtain the correlation function to be used with conventional beamforming method. The correlation function derived from von Mises pdf was given in (5.15). Inserting it into the beamformer power equation results in a power equation for the distributed source model. *n* peaks of this new power statement will give the AOA estimates of *n* signals coming from mobile stations surrounded with scatterers in a multipath environment.

6.1.1.1. Mean Squared Error Calculation for CBF: MSE of CBF estimates of AOA can be calculated for large number of samples to analyse the performance of CBF [28]. AOA estimates obtained from CBF are asymptotically biased. The CBF asymptotic cost function is given as

$$\bar{f}(\theta) = \boldsymbol{a}^H(\theta)\boldsymbol{R}_y\boldsymbol{a}(\theta) \tag{6.26}$$

Let $\{\bar{\theta}\}_{k=1}^{n}$ be the maximizers of the asymptotic cost function above. When SNR is low, $\{\bar{\theta}_k\}$ values may not be exactly the same as true AOA values $\{\theta_k\}$. The CBF estimates $\{\hat{\theta}_k\}$ will fluctuate around $\{\bar{\theta}_k\}$ when the number of samples is large. The asymptotic MSE can be obtained by

$$MSE(\hat{\theta}_k) = \mathrm{var}(\hat{\theta}_k) + (\bar{\theta}_k - \theta_k)^2 \tag{6.27}$$

where $E[\hat{\theta}_k] = \bar{\theta}_k$.

Also,

$$\mathrm{var}(\hat{\theta}_k) = \frac{\Delta}{h^2(\bar{\theta}_k)} = \frac{\frac{1}{2N}\left\{(\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y^{-1}\boldsymbol{v}) - |\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{v}|^2\right\}}{\left[\mathrm{Re}[\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{u} + \boldsymbol{v}^H\boldsymbol{d}'(\bar{\theta}_k)]\right]^2} \tag{6.28}$$

where $\boldsymbol{u}^H = \boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y$ and $\boldsymbol{v} = \boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)$.

Derivation of the $\text{var}(\hat{\theta}_k)$ term above is given in Appendix A.

The steps to follow for the asymptotic MSE calculation are as below:

(i). Compute $\boldsymbol{R}_y$ for the scenario.

(ii). Maximize the asymptotic cost function to determine $\{\overline{\theta}_k\}$.

(iii). Determine $\text{var}(\hat{\theta}_k)$ using (6.28).

(iv). Calculate $MSE(\hat{\theta}_k)$ from (6.27) .

6.1.2. Minimum Variance Distortionless Response (MVDR) / Capon Method

Capon method is similar to conventional beamforming at first sight. The two constraints stated for conventional beamforming are still the main considerations. However, Capon method takes into account the actual array data fed into the array when calculating the output power. In other words, conventional beamformer scans all directions to suppress any signals arriving with an AOA other than θ whereas in the Capon method the only directions that are scanned are the ones which contain the incoming signals. For conventional beamformer, the incident signal is assumed to be white and steering is done uniformly regardless of the existence of signals [20].

MVDR method maximizes the output SNR and passes the look direction signal undistorted when there are no errors. Array weights are obtained by minimizing the mean output power subject to unity constraint in the look direction. In other words, MVDR is a spatial filter that passes the signal of a point source arriving from direction θ while maximally rejecting the signals coming from other directions. The design problem for the Capon method can be expressed as

$$\min_{h} \boldsymbol{h}^H \boldsymbol{R}_y \boldsymbol{h} \text{ subject to } \boldsymbol{h}^H \boldsymbol{a}(\theta) = 1$$

The solution of this optimization problem is similar to that of the conventional beamformer.

$$\varepsilon = f + \lambda g \tag{6.29}$$

$$\varepsilon = E\left\{| \boldsymbol{y}_F(t) |^2\right\} + \lambda\left(\boldsymbol{h}^H \boldsymbol{a}(\theta) - 1\right) \tag{6.30}$$

$$\varepsilon = \boldsymbol{h}^H \boldsymbol{R}_y \boldsymbol{h} + \lambda\left(\boldsymbol{h}^H \boldsymbol{a}(\theta) - 1\right) \tag{6.31}$$

$$\nabla\varepsilon = \nabla f + \lambda \nabla g = 0 \tag{6.32}$$

$$\nabla\varepsilon = \frac{\partial\varepsilon}{\partial \boldsymbol{h}^H} = \boldsymbol{R}_y \boldsymbol{h} + \lambda \boldsymbol{a}(\theta) = 0 \tag{6.33}$$

$\boldsymbol{h}$ can be found as,

$$\boldsymbol{h} = -\lambda \boldsymbol{R}_y{}^{-1} \boldsymbol{a}(\theta) \tag{6.34}$$

Using the equation above in the constraint of distortionless output

$$\boldsymbol{h}^H \boldsymbol{a}(\theta) = 1 \tag{6.35}$$

$$-\lambda^H \boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})^H \boldsymbol{a}(\theta) = 1 \tag{6.36}$$

Solving for λ,

$$\lambda = -1/\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta) \tag{6.37}$$

As a result $\boldsymbol{h}$ can be expressed as below using this value of λ in (6.34).

$$\boldsymbol{h} = -\lambda \boldsymbol{R}_y{}^{-1} \boldsymbol{a}(\theta) \tag{6.38}$$

$$\boldsymbol{h} = (\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta) / \boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta) \tag{6.39}$$

Substituting (6.39) into (6.34)

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \boldsymbol{h}^H \boldsymbol{R}_y \boldsymbol{h} \tag{6.40}$$

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \left[\frac{(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)}{\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)}\right]^H \boldsymbol{R}_y \left[\frac{(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)}{\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)}\right] \tag{6.41}$$

After some steps, the output power formula for Capon method can be found as below.

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \frac{\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{R}_y(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)}{\left[\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)\right]^2} \tag{6.42}$$

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \frac{\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{I}\boldsymbol{a}(\theta)}{\left[\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)\right]^2} \tag{6.43}$$

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \frac{\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)}{\left[\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)\right]^2} \tag{6.44}$$

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \frac{1}{\boldsymbol{a}^H(\theta)(\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\theta)} \tag{6.45}$$

Finally, replacing the correlation function with a proper estimate gives the Capon beamformer output power for AOA estimation.

$$E\left\{|\, \boldsymbol{y}_F(t)\,|^2\right\} = \frac{1}{\boldsymbol{a}^H(\theta)(\hat{\boldsymbol{R}}_y{}^{-1})\boldsymbol{a}(\theta)} \tag{6.46}$$

The assumption of nonisotropic scatterers can be used as in conventional beamformer to model the multipath environment with the von Mises pdf. The estimate replaced with the one derived with von Mises pdf results in a new power equation for the multipath model consisting of distributed sources.

6.1.2.1. Mean Squared Error Calculation for MVDR: AOA estimates obtained from MVDR are asymptotically biased. Large-sample MSE of MVDR estimates of AOA can be used to analyse the performance of MVDR [28]. The MVDR asymptotic cost function is given as

$$\bar{f}(\theta) = \boldsymbol{a}^H(\theta)\boldsymbol{R}_y^{-1}\boldsymbol{a}(\theta) \tag{6.47}$$

Let the minimizers of the asymptotic cost function above be $\{\bar{\theta}\}_{k=1}^{n}$. These values $\{\bar{\theta}_k\}$ may not be exactly the same as true AOA values $\{\theta_k\}$ under low SNR conditions. When there are large number of samples, the MVDR estimates $\{\hat{\theta}_k\}$ will fluctuate around $\{\bar{\theta}_k\}$. The large sample MSE can be obtained by

$$MSE(\hat{\theta}_k) = \text{var}(\hat{\theta}_k) + (\bar{\theta}_k - \theta_k)^2 \tag{6.48}$$

where $E[\hat{\theta}_k] = \bar{\theta}_k$.

Also,

$$\text{var}(\hat{\theta}_k) = \frac{\Delta}{h^2(\bar{\theta}_k)} = \frac{\frac{1}{2N}\left\{(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y\boldsymbol{v}) - |\,\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v}\,|^2\right\}}{\text{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}\boldsymbol{d}(\bar{\theta}_k) + \boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}\boldsymbol{d}'(\bar{\theta}_k)} \tag{6.49}$$

where $\boldsymbol{u}^H = \boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}$ and $\boldsymbol{v} = \boldsymbol{R}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)$.

Derivation of the $\text{var}(\hat{\theta}_k)$ term above is given in Appendix B.

The steps to follow for the asymptotic MSE calculation are as below:

(i). Compute $\boldsymbol{R}_y$ for the scenario.

(ii). Minimize the asymptotic cost function to determine $\{\bar{\theta}_k\}$.

(iii). Determine $\text{var}(\hat{\theta}_k)$ using (6.49).

(iv). Calculate $MSE(\hat{\theta}_k)$ from (6.48).

6.2. Subspace Based Methods

Subspace based methods are based on the property that the space spanned by the eigenvectors of the array correlation matrix may be divided in two subspaces called the signal subspace and the noise subspace and that the steering vectors corresponding to the directional sources are orthogonal to the noise subspace.

6.2.1. Multiple Signal Classification (MUSIC) Method

MUSIC algorithm was proposed in 1979 by Schmidt [20]. It is a high resolution algorithm compared to the conventional beamforming and MVDR methods. MUSIC estimates the noise subspace from the available samples and once the noise subspace has been estimated, a search for n directions is made by looking for steering vectors which are orthogonal to the noise subspace. MUSIC spectra is used to search for peaks indicating the directions.

If there are n signals incident on an m-element ULA, the multiple signal model for a uniform linear array is given as [27]

$$\boldsymbol{y}(t) \;=\; \boldsymbol{A}\boldsymbol{s}(t)+\boldsymbol{e}(t) \tag{6.50}$$

where

$$\boldsymbol{A}=\begin{pmatrix} a_1(\theta_1) & a_1(\theta_2) & \dots & a_1(\theta_n) \\ a_2(\theta_1) & a_2(\theta_2) & \dots & a_2(\theta_n) \\ \vdots & \vdots & \vdots & \vdots \\ a_m(\theta_1) & a_m(\theta_2) & \dots & a_m(\theta_n) \end{pmatrix}=\underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 \\ a_2(\theta_1) & a_2(\theta_2) & \dots & a_2(\theta_n) \\ \vdots & \vdots & \vdots & \vdots \\ a_m(\theta_1) & a_m(\theta_2) & \dots & a_m(\theta_n) \end{pmatrix}}_{mxn} \tag{6.51}$$

$$\boldsymbol{s}(t) = \underbrace{\begin{pmatrix} s_1(t_1) & s_1(t_2) & \dots & s_1(t_N) \\ s_2(t_1) & s_2(t_2) & \dots & s_2(t_N) \\ \vdots & \vdots & \vdots & \vdots \\ s_n(t_1) & s_n(t_2) & \dots & s_n(t_N) \end{pmatrix}}_{nxN} \tag{6.52}$$

$$\boldsymbol{e}(t) = \underbrace{\begin{pmatrix} e_1(t_1) & e_1(t_2) & \dots & e_1(t_N) \\ e_2(t_1) & e_2(t_2) & \dots & e_2(t_N) \\ \vdots & \vdots & \vdots & \vdots \\ e_m(t_1) & e_m(t_2) & \dots & e_m(t_N) \end{pmatrix}}_{mxN} \tag{6.53}$$

$$\boldsymbol{y}(t) = \underbrace{\begin{pmatrix} a_1(\theta_1) & a_1(\theta_2) & \dots & a_1(\theta_n) \\ a_2(\theta_1) & a_2(\theta_2) & \dots & a_2(\theta_n) \\ \vdots & \vdots & \vdots & \vdots \\ a_m(\theta_1) & a_m(\theta_2) & \dots & a_m(\theta_n) \end{pmatrix}}_{mxn} \underbrace{\begin{pmatrix} s_1(t_1) & s_1(t_2) & \dots & s_1(t_N) \\ s_2(t_1) & s_2(t_2) & \dots & s_2(t_N) \\ \vdots & \vdots & \vdots & \vdots \\ s_n(t_1) & s_n(t_2) & \dots & s_n(t_N) \end{pmatrix}}_{nxN}$$

$$+ \underbrace{\begin{pmatrix} e_1(t_1) & e_1(t_2) & \dots & e_1(t_N) \\ e_2(t_1) & e_2(t_2) & \dots & e_2(t_N) \\ \vdots & \vdots & \vdots & \vdots \\ e_m(t_1) & e_m(t_2) & \dots & e_m(t_N) \end{pmatrix}}_{mxN} \tag{6.54}$$

Then the input covariance matrix can be expressed as

$$\boldsymbol{R}_y = E[\boldsymbol{y}\boldsymbol{y}^H] = \boldsymbol{A}E[\boldsymbol{s}\boldsymbol{s}^H]\boldsymbol{A}^H + E[\boldsymbol{e}\boldsymbol{e}^H] \tag{6.55}$$

$$\boldsymbol{R}_y = E[\boldsymbol{y}\boldsymbol{y}^H] = \boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H + {\sigma_e}^2\boldsymbol{I} \tag{6.56}$$

where $\boldsymbol{R}_s$ is the signal correlation matrix.

MUSIC algorithm exploits the structure of the received signal as a composition of the signal subspace and the noise subspace. For this reason MUSIC is known as a

"subspace based" method. Eigenvalue decomposition is the key point in determining the AOAs of the incident signals on the array antenna.

The eigenvalues of input covariance matrix $\boldsymbol{R}_y$ can be found by

$$| \boldsymbol{R}_y - \lambda_i \boldsymbol{I} | = 0 \tag{6.57}$$

where $\{\lambda_1, \ldots, \lambda_m\}$ represent the m eigenvalues.

Inserting $\boldsymbol{R}_y$ from (6.56) and rearranging yields

$$\left| \boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H + {\sigma_e}^2 \boldsymbol{I} - \lambda_i \boldsymbol{I} \right| = \left| \boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H - (\lambda_i - {\sigma_e}^2)\boldsymbol{I} \right| = 0 \tag{6.58}$$

Comparing this equation with (6.57), it can be deduced that $v_i = \lambda_i - {\sigma_e}^2$ are the eigenvalues of $\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H$.

Assumptions in section 4.2 are valid for MUSIC since it uses the array model mentioned previously. In order to clarify the following steps of MUSIC algorithm, some more assumptions have to be made.

(i). The number of ULA elements (sensors) is greater than the number of sources. That is, $m > n$. Also the number of sources n is assumed to be known.

(ii). The steering vectors $\{\boldsymbol{a}(\theta_k)\}_{k=1}^{n}$ which comprise $\boldsymbol{A}$ are linearly independent. In other words, the matrix $\boldsymbol{A}$ has full column rank.

(iii). The noise $\boldsymbol{e}(t)$ is a zero-mean stationary complex Gaussian random vector which is both temporarily and spatially white. This statement can be expressed as

$$E\left[\boldsymbol{e}(k)\boldsymbol{e}^H(l)\right] = \sigma^2 \boldsymbol{I} \delta_{kl} \tag{6.59}$$

where δ_{kl} is the Kronecker delta which is equal to 1 for $k=l$ and 0 otherwise.

(iv). The signals $\{\boldsymbol{s}(t)\}_{t=1}^{N}$ are complex Gaussian random vectors that are independent and stationary with zero-mean. Which in turn leads to the result that received signals $\{\boldsymbol{y}(t)\}_{t=1}^{N}$ are also complex Gaussian random vectors, and they are independent and stationary with zero-mean.

Another consideration is, the matrix $\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^{\boldsymbol{H}}$ is positive semidefinite with rank n. This means that n eigenvalues of $\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^{\boldsymbol{H}}$ are nonzero and the remaining m-n are zero. The number of sources is assumed to be known for the scenario.

Since the eigenvalues of $\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^{\boldsymbol{H}}$ are expressed as

$$v_i = \lambda_i - {\sigma_e}^2 \tag{6.60}$$

zero eigenvalues of $\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^{\boldsymbol{H}}$ imply having m-n eigenvalues of $\boldsymbol{R}_y$ equal to noise variance ${\sigma_e}^2$. This can be written as

$$\lambda_{n+1},\ldots,\lambda_m = {\sigma_e}^2 \tag{6.61}$$

where the largest eigenvalue is supposed to be λ_1 and the smallest is λ_m.

Eigenvectors $\boldsymbol{q}_i$ of the correspoding eigenvalues λ_i satisfy the equation below.

$$(\boldsymbol{R}_y - \lambda_i \boldsymbol{I})\boldsymbol{q}_i = 0 \tag{6.62}$$

For the smallest eigenvalues the equation follows as

$$(\boldsymbol{R}_y - \sigma_e{}^2\boldsymbol{I})\boldsymbol{q}_i = \boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H\boldsymbol{q}_i + \sigma_e{}^2\boldsymbol{I} - \sigma_e{}^2\boldsymbol{I} = 0 \tag{6.63}$$

$$\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H\boldsymbol{q}_i = 0 \tag{6.64}$$

Full rank $\boldsymbol{A}$ matrix and nonsingular $\boldsymbol{R}_s$ matrix result in

$$\boldsymbol{A}^H\boldsymbol{q}_i = 0 \tag{6.65}$$

$$\left[\boldsymbol{a}^H(\theta_1)\boldsymbol{q}_i \quad \boldsymbol{a}^H(\theta_2)\boldsymbol{q}_i \quad \ldots \quad \boldsymbol{a}^H(\theta_n)\boldsymbol{q}_i\right]^T = \left[0 \quad 0 \quad \ldots \quad 0\right]^T \tag{6.66}$$

The key point in determining the AOAs via MUSIC resides in the orthogonality of the eigenvectors of *m-n* smallest eigenvalues of $\boldsymbol{R}_y$ to the *n* steering vectors of $\boldsymbol{A}$ matrix.

$$\left\{\boldsymbol{a}(\theta_1) \quad \boldsymbol{a}(\theta_2) \quad \ldots \quad \boldsymbol{a}(\theta_n)\right\} \perp \left\{\boldsymbol{q}_{n+1} \quad \boldsymbol{q}_{n+2} \quad \ldots \quad \boldsymbol{q}_m\right\} \tag{6.67}$$

The smallest eigenvalues equal to the noise variance represent the noise subspace whereas the *n* steering vectors of $\boldsymbol{A}$ are associated with the received signals making up the signal subspace. Thus the eigenvectors of covariance matrix $\boldsymbol{R}_y$ are either in the noise subspace or in the signal subspace.

In order to estimate the AOA of the incident signals, all the steering vectors of the array manifold have to be searched to find *n* vectors that are orthogonal to the noise subspace eigenvectors. The noise subspace eigenvectors can be written as

$$\boldsymbol{V}_e = \left[\boldsymbol{q}_{n+1} \quad \boldsymbol{q}_{n+2} \quad \cdots \quad \boldsymbol{q}_m\right] \tag{6.68}$$

Then if the array is steered through all possible angles, the AOAs of the multipath components can be determined by selecting the ones which satisfy the orthogonality below.

$$\boldsymbol{a}^H(\theta)\boldsymbol{V}_e\boldsymbol{V}_e{}^H\boldsymbol{a}(\theta)=0 \tag{6.69}$$

The AOAs of the incident signals which minimize the expression above will cause peaks when the above expression is inverted and put as denominator. This form gives the MUSIC spatial spectrum

$$P_{MUSIC}(\theta)=1/\boldsymbol{a}^H(\theta)\boldsymbol{V}_e\boldsymbol{V}_e{}^H\boldsymbol{a}(\theta) \tag{6.70}$$

or

$$P_{MUSIC}(\theta)=\boldsymbol{a}^H(\theta)\boldsymbol{a}(\theta)/\boldsymbol{a}^H(\theta)\boldsymbol{V}_e\boldsymbol{V}_e{}^H\boldsymbol{a}(\theta) \tag{6.71}$$

Selecting n highest peaks of the MUSIC spectrum means estimating AOAs of the n incident signals.

Finally, the steps to follow for the MUSIC algorithm can be stated as below:

(i). Estimate the covariance matrix of the received signals at the BTS.

$$\hat{\boldsymbol{R}}_y=\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{y}(t)\boldsymbol{y}^H(t) \tag{6.72}$$

(ii). Use eigenvalue decomposition to separate the signal and noise subspaces. One decomposition technique generally used for this purpose is the "singular value decomposition (SVD)" shown below.

$$\hat{\boldsymbol{R}}_y=\boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^H \tag{6.73}$$

$\boldsymbol{S}$ is a diagonal matrix of the same dimension as $\hat{\boldsymbol{R}}_y$, with nonnegative diagonal elements in decreasing order, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices. The diagonal elements of $\boldsymbol{S}$ are the singular values, and the columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right singular

vectors. Covariance matrix $\hat{\boldsymbol{R}}_y$ is a Hermitian matrix which is positive semi-definite, that is all its eigenvalues are real and non-negative. Then the singular values and singular vectors coincide with the eigenvalues and eigenvectors of $\hat{\boldsymbol{R}}_y$.

$$\boldsymbol{V}_e = \begin{bmatrix} \boldsymbol{q}_{n+1} & \boldsymbol{q}_{n+2} & \cdots & \boldsymbol{q}_m \end{bmatrix} \tag{6.74}$$

Noise eigenvectors above can be determined by extracting the right *m-n* singular vectors from $\boldsymbol{U}$.

(iii). Compute the MUSIC spectrum and find the n largest peaks of $P_{MUSIC}(\theta)$

$$P_{MUSIC}(\theta) = \boldsymbol{a}^H(\theta)\boldsymbol{a}(\theta) / \boldsymbol{a}^H(\theta)\boldsymbol{V}_e\boldsymbol{V}_e{}^H\boldsymbol{a}(\theta) \tag{6.75}$$

The difference of MUSIC method from conventional methods is that the peaks of MUSIC spatial spectrum correspond to the true AOAs under uncorrelated and identical noise conditions when the array input covariance matrix average is known. On the other hand, conventional methods like CBF and MVDR estimate the signal power incident from each arrival angle and help to determine the AOAs from these estimates.

The MUSIC algorithm is known to be an unbiased estimator and the estimation error variance approaches the Cramer-Rao lower bound (CRLB) when there are enough number of snapshots. AOA estimates from MUSIC asymptotically approach CRLB for the case of a single source. If SNR is large, multiple source estimates also approach CRLB.

6.2.2. Root-MUSIC Method

An improvement to MUSIC algorithm was proposed by Barabell to obtain better resolution. This algorithm called Root-MUSIC is based on polynomial rooting and is valid only for a uniformly spaced linear array antenna [27].

As previously stated in section 4.3, the steering vector for the array model is of the form

$$\boldsymbol{a}(\theta) = [1 \quad e^{-j2\pi d\cos\theta/\lambda} \quad \ldots \quad e^{-j2\pi d(m-1)\cos\theta/\lambda}]^T \tag{6.76}$$

where the p-th element of the steering vector can be expressed as

$$\boldsymbol{a}_p(\theta) = e^{-j2\pi dp\cos\theta/\lambda} \qquad p=0, ..., m\text{-}1 \tag{6.77}$$

The MUSIC spectrum is an all-pole function given as

$$\begin{aligned} P_{MUSIC}(\theta) &= 1/\boldsymbol{a}^H(\theta)\boldsymbol{V_e}\boldsymbol{V_e}^H\boldsymbol{a}(\theta) \\ &= 1/\boldsymbol{a}^H(\theta)\boldsymbol{C}^H\boldsymbol{a}(\theta) \end{aligned} \tag{6.78}$$

where $\boldsymbol{C} = V_e V_e^H$.

$$\begin{aligned} P_{MUSIC}^{-1}(\theta) &= \sum_{p=0}^{m-1}\sum_{k=0}^{m-1} e^{j2\pi dp\cos\theta/\lambda} C_{pk} e^{j2\pi dk\cos\theta/\lambda} \\ &= \sum_{p=1}^{m}\sum_{k=1}^{m} e^{j2\pi d(p-1)\cos\theta/\lambda} C_{(p-1)(k-1)} e^{j2\pi d(k-1)\cos\theta/\lambda} \end{aligned} \tag{6.79}$$

$C_{(p-1)(k-1)}$ is the element in the (p-1)th row and (k-1)th column of $\boldsymbol{C}$. Simplifying further,

$$P_{MUSIC}^{-1}(\theta) = \sum_{l=-m+1}^{m-1} C_l e^{j2\pi dl\cos\theta/\lambda} \tag{6.79}$$

where $C_l = \sum_{p-k=l} C_{(p-1)(k-1)}$ is the sum of the entries of $\boldsymbol{C}$ along l-th diagonal. The discussion continues with the usage of z-transform. The locations of AOAs determine the poles of a polynomial on the unit circle.

$$z = e^{j2\pi d\cos\theta/\lambda} \tag{6.81}$$

$$P_{MUSIC}^{-1}(\theta) = \sum_{p=1}^{m}\sum_{k=1}^{m} z^{(p-1)} C_{(p-1)(k-1)} z^{(k-1)} \tag{6.82}$$

$$P_{MUSIC}^{-1}(\theta) = \sum_{l=-m+1}^{m-1} C_l z^l \tag{6.83}$$

$$D(z) = \sum_{l=-m+1}^{m-1} C_l z^l \tag{6.84}$$

MUSIC algorithm imposes the constraint $|z| = 1$ and tries to find the local minima of the denominator of $P_{MUSIC}(\theta)$. Thus, each minima corresponds to a peak in the spectrum. If there is no noise, the poles of $D(z)$ are exactly on the unit circle. However, Root-MUSIC will try to locate all the roots of the polynomial without any constraint. The roots of $D(z)$ will be obtained in pairs, that is, if z is a root, so is z^{-1}. Possible directions of *n* sources can be found by filtering out the roots with magnitude greater than 1 and then choosing *n* roots that are closest to the unit circle. Angles of those roots will give the AOA estimates.

$$z_r = | z_r | e^{j \arg(z_r)} \tag{6.85}$$

$$\theta_r = \cos^{-1}\left(\frac{\lambda \arg(z_r)}{2\pi d}\right) \qquad r = 1, ...,n \tag{6.86}$$

Especially at low SNR conditions Root-MUSIC has better resolution than MUSIC algorithm.

6.2.2.1. Mean Squared Error Calculation for Root-MUSIC: MSE can be used to compare the performance of Root-MUSIC algorithm with the performance of the conventional algorithms. AOA estimations obtained from simulations can be used to calculate MSE as a

means of a measure of performance for Root-MUSIC. The expression below can be used for the MSE of an AOA estimate $\hat{\theta}_k$ where S is the number of simulations performed.

$$MSE(\hat{\theta}_k) = E\left[(\hat{\theta}_k - \theta)^2\right] \tag{6.87}$$

$$\widehat{MSE}(\hat{\theta}_k) = \frac{1}{S}\sum_{k=1}^{S}(\hat{\theta}_k - \theta)^2 \tag{6.88}$$

7. GENERAL CRLB AND DERIVATION OF CRLB FOR THE VON MISES MODEL

The Cramer-Rao Lower Bound (CRLB) gives the theoretically lowest value of the covariance of an unbiased estimator. Cramer-Rao Lower Bound (CRLB) for the von Mises model can be found under the assumption that the received signals are stationary stochastic processes and they have complex m-variate Gaussian distribution with zero-mean and correlation function $\boldsymbol{R}_y$.

7.1. General CRLB

The complex multivariate Gaussian pdf can be expressed as follows [29]. The only term dependant on the unknown parameters β is $\boldsymbol{R}_y$. The probability distribution function of a single observation $y(t_i)$ is

$$p_i(y;\beta) = \frac{1}{\pi^m \det[R_y(\beta)]} \exp\{-y^H [R_y^{-1}(\beta)]\, y\} \tag{7.1}$$

where β is the vector of unknown parameters.

If the snapshots are assumed to be independent and identically distributed, the probability distribution function of the complete data set $y(t_1),..., y(t_N)$ is given by

$$p\left[y(t_1),..., y(t_N);\beta\right] = \prod_{i=1}^{N} \frac{1}{\pi^m \det[\boldsymbol{R}_y(\beta)]} \exp\{-[y^H(t_i)][\boldsymbol{R}_y^{-1}(\beta)][\, y(t_i)]\} \tag{7.2}$$

After several calculations given in Appendix C, $\{i, j\}$ element of the Fisher information matrix can be written as

$$[FI(\boldsymbol{\beta})]_{ij} = N\,tr\left\{\boldsymbol{R}_{\boldsymbol{y}}^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \beta_i}\boldsymbol{R}_{\boldsymbol{y}}^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \beta_j}\right\} \tag{7.3}$$

where β_i represents the *i*-th element of the vector $\boldsymbol{\beta}$. Cramer Rao Lower Bound matrix is then calculated as below.

$$CRLB(\boldsymbol{\beta}) = [FI(\boldsymbol{\beta})]^{-1} \tag{7.4}$$

This is the large-sample "General CRLB" for stochastic signals that have complex *m*-variate Gaussian distribution with zero-mean and correlation function $\boldsymbol{R}_{\boldsymbol{y}}$.

7.2. Derivation of CRLB for the Von Mises Model

The cross-correlation function derived using the von Mises pdf can be inserted into the "General CRLB" expression.

In the case of von Mises model, the unknown parameters may be θ or κ, that is $\boldsymbol{\beta} = \{\theta, \kappa\}$. The derivative terms are then given by [24]

$$\boldsymbol{R}_{\boldsymbol{y}} = P_{signal}\,\boldsymbol{R}_{\boldsymbol{z}} + P_{noise}\,\boldsymbol{I} \tag{7.5}$$

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \theta} = \sigma_s^2\left(\frac{\partial \boldsymbol{R}_{\boldsymbol{z}}(\boldsymbol{\beta})}{\partial \theta}\right) \tag{7.6}$$

Referring to the von Mises cross-correlation function,

$$\boldsymbol{R}_{\boldsymbol{z}}(\boldsymbol{\beta}) = \frac{I_0\left[f(\boldsymbol{\beta})\right]}{I_0(\kappa)} \tag{7.7}$$

where $f(\boldsymbol{\beta}) = \sqrt{\kappa^2 - 4\pi^2(l-q)^2(d/\lambda)^2 - j4\pi\kappa(l-q)(d/\lambda)\cos(\theta)}$

Thus,

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \theta} = \sigma_s^2 \left\{ \frac{I_1\left[f(\boldsymbol{\beta})\right]\left[\partial f(\boldsymbol{\beta}) / \partial \theta\right]}{I_0(\kappa)} \right\} \tag{7.8}$$

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \theta} = \sigma_s^2 \left(\frac{I_1\left[\sqrt{\kappa^2 - 4\pi^2(l-q)^2(d/\lambda)^2 - j4\pi\kappa(l-q)(d/\lambda)\cos(\theta)}\right]}{I_0(\kappa)} \right) \left(\frac{j4\pi\kappa(l-q)(d/\lambda)\sin(\theta)}{2\sqrt{\kappa^2 - 4\pi^2(l-q)^2(d/\lambda)^2 - j4\pi\kappa(l-q)(d/\lambda)\cos(\theta)}} \right) \tag{7.9}$$

In the equations above $I_1(\cdot)$ is the first-order modified Bessel function of the first kind. Similar derivation can also be made for κ.

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \kappa} = \sigma_s^2 \left(\frac{\partial \boldsymbol{R}_{\boldsymbol{z}}(\boldsymbol{\beta})}{\partial \kappa} \right) \tag{7.10}$$

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \kappa} = \sigma_s^2 \left\{ \frac{I_0(\kappa)\left[\partial I_0\left[f(\boldsymbol{\beta})\right] / \partial \kappa\right] - I_0\left[f(\boldsymbol{\beta})\right]\left[\partial I_0(\kappa) / \partial \kappa\right]}{\left[I_0(\kappa)\right]^2} \right\} \tag{7.11}$$

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \kappa} = \sigma_s^2 \left\{ \frac{I_0(\kappa) I_1\left[f(\boldsymbol{\beta})\right](1/2)[f(\boldsymbol{\beta})]^{-1}[2\kappa - j2\vartheta\cos(\theta)] - I_0\left[f(\boldsymbol{\beta})\right] I_1(\kappa)}{\left[I_0(\kappa)\right]^2} \right\}$$

$$= \sigma_s^2 \left\{ [I_0(\kappa)]^{-1} I_1\left[f(\beta)\right][f(\beta)]^{-1}[\kappa - j\vartheta\cos(\theta)] - [I_0(\kappa)]^{-2} I_0\left[f(\beta)\right] I_1(\kappa) \right\} \tag{7.12}$$

where $\vartheta = 2\pi(l-q)(d/\lambda)$ and $f(\boldsymbol{\beta}) = \sqrt{\kappa^2 - \vartheta^2 - j2\kappa\vartheta\cos(\theta)}$

The Fisher information matrix can then be written as

$$[FI(\boldsymbol{\beta})] =$$

$$N\begin{bmatrix} tr\left\{\left(\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial\theta}\right)^2\right\} & tr\left\{\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial\theta}\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial\kappa}\right\} \\ tr\left\{\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial\theta}\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial\kappa}\right\} & tr\left\{\left(\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial\kappa}\right)^2\right\} \end{bmatrix} \tag{7.13}$$

The CRLBs for θ and κ are the first and second diagonal elements of the inverse Fisher information matrix respectively.

In case of multiple clusters and multiple unknown parameters, the Fisher information matrix for the sets of unknown parameters can also be derived similarly.

Assuming $\boldsymbol{\beta} = \{\kappa_p, \theta_p, \varepsilon_p\}$, p=1,..., P represents the set of three unknown parameters,

$$\begin{aligned} &\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \kappa_p} \\ &= \varepsilon_p \sigma_s^2 \left\{ \frac{I_0(\kappa_p) I_1\left[f(\boldsymbol{\beta})_p\right](1/2)[f(\boldsymbol{\beta})_p]^{-1}[2\kappa_p - j2\vartheta\cos(\theta_p)] - I_0\left[f(\boldsymbol{\beta})_p\right] I_1(\kappa_p)}{[I_0(\kappa_p)]^2} \right\} \\ &= \varepsilon_p \sigma_s^2 \left\{ [I_0(\kappa_p)]^{-1} I_1\left[f(\boldsymbol{\beta})_p\right][f(\beta)_p]^{-1}[\kappa_p - j\vartheta\cos(\theta_p)] - [I_0(\kappa_p)]^{-2} I_0\left[f(\boldsymbol{\beta})_p\right] I_1(\kappa_p) \right\} \end{aligned} \tag{7. 14}$$

where $\vartheta = 2\pi(l-q)(d/\lambda)$ and $f(\boldsymbol{\beta})_p = \sqrt{\kappa_p{}^2 - \vartheta^2 - j2\kappa_p\vartheta\cos(\theta_p)}$

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \theta_p} = \varepsilon_p \sigma_s^2 \left\{ \frac{I_1\left[f(\boldsymbol{\beta})_p\right]\left[\partial f(\boldsymbol{\beta})_p / \partial \theta_p\right]}{I_0(\kappa_p)} \right\}$$

$$= \varepsilon_p \sigma_s^2 \left(\frac{I_1\left[\sqrt{{\kappa_p}^2 - \vartheta^2 - j2\kappa_p \vartheta \cos(\theta_p)}\right]}{I_0(\kappa_p)} \right) \left(\frac{j2\kappa_p \vartheta \sin(\theta_p)}{2\sqrt{{\kappa_p}^2 - \vartheta^2 - j2\kappa_p \vartheta \cos(\theta_p)}} \right) \tag{7.15}$$

In order to include ε_p into the Fisher information matrix expressions, the condition about mixture proportions can be expressed as $\varepsilon_P = 1 - \sum_{p=1}^{P-1} \varepsilon_p$. The derivatives of the cross-correlation function with respect to ε_p, p=1,...,*P-1* can be given as

$$\frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \varepsilon_p} = \sigma_s^2 \left\{ \frac{I_0\left[f(\boldsymbol{\beta})_p\right]}{I_0(\kappa_p)} - \frac{I_0\left[f(\boldsymbol{\beta})_P\right]}{I_0(\kappa_P)} \right\} \tag{7.16}$$

Since there are *3P-1* unknown parameters, the Fisher information matrix is

$$[FI(\boldsymbol{\beta})]_{ij} = N\, tr\left\{ \boldsymbol{R}_{\boldsymbol{y}}^{-1}(\boldsymbol{\beta}) \frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_i} \boldsymbol{R}_{\boldsymbol{y}}^{-1}(\beta) \frac{\partial \boldsymbol{R}_{\boldsymbol{y}}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_j} \right\} \qquad i = 1, \ldots, 3P-1 \qquad j = 1, \ldots, 3P-1 \tag{7.17}$$

8. SIMULATIONS

8.1. Point Source Model Simulations

When the mobile users are assumed to be point sources, the performance of AOA estimation algorithms may differ from each other according to some model parameters. Simulations were made to compare the AOA estimation algorithms reviewed in Chapter 6. These simulations use the array model in section 4.3 and thus assume narrowband directional sources and white Gaussian noise. In case of multiple sources, these sources are also assumed to be uncorrelated.

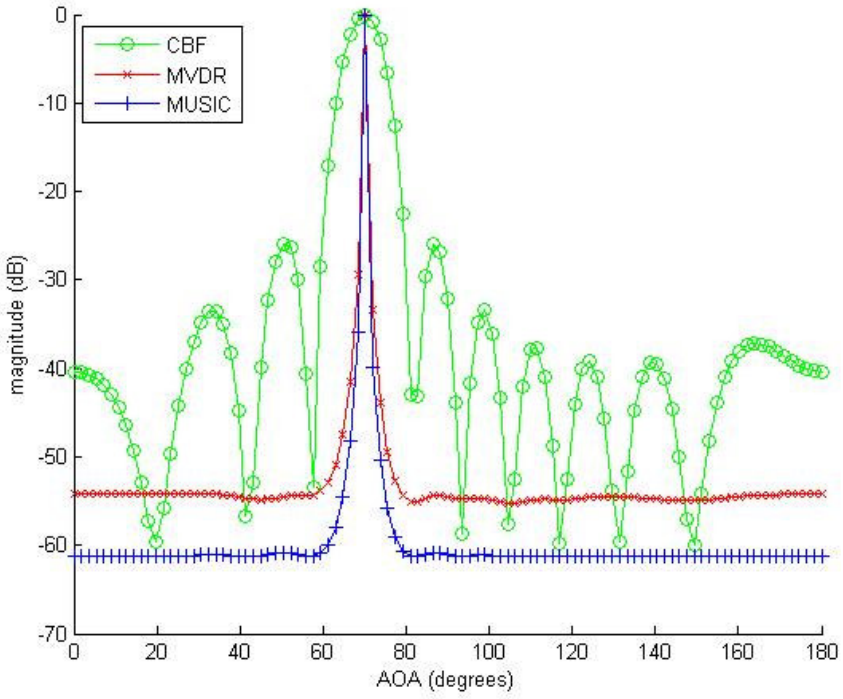


Figure 7.1. Estimated AOAs from different algorithms (m=10, $d = 0.5\lambda$, $\theta = 70^{\circ}$, SNR=20dB, $\sigma_e = 1$, N=1000)

The AOA spectrums of CBF, MVDR and MUSIC algorithms are compared to each other in Figure 7.1. A uniform linear array with 10 elements is considered. The mobile phone is situated at 70° with respect to the ULA and there is 0.5λ separation between the

array elements. 1000 samples of the received signal were used to obtain the results above. It is clear that the AOA estimated with MUSIC algorithm shows the best performance. The sharpest peak is observed to be the one from the MUSIC spectrum. The spectrum outside the AOA estimate is suppressed very well for MUSIC. For the MVDR spectrum, there are small fluctuations at the two sides of the AOA estimate but they are negligible. The worst performance appears to be of the CBF where smaller peaks can be observed at two sides of the AOA estimate. These extra peaks cause ambiguity if the number of sources is unknown since they may be counted as the other AOA estimates.

When there are multiple sources, another criterion that becomes important is resolution which is the ability of the location estimation algorithm to separate two sources near each other. The spectrum of the three methods is shown in Figure 7.2. First mobile phone is at 70 degrees while the second phone's AOA is 90 degrees. All methods provide acceptable results for these two sources 20 degrees apart from each other.

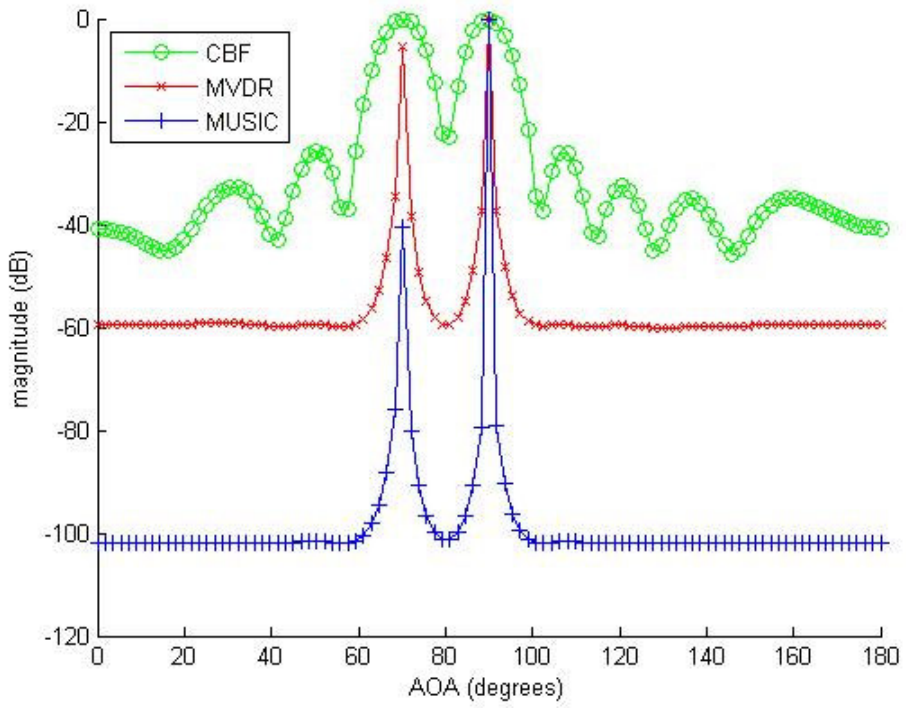


Figure 7.2. Estimated AOAs for two sources 20 degrees apart (m = 10, $d = 0.5\lambda$, $\theta_1 = 70^\circ$, $\theta_2 = 90^\circ$, SNR = 20dB, $\sigma_e = 1$, N = 1000)

If the sources are near each other and the resolution of one algorithm is not good enough, then two sources may appear as only one peak in the spectrum which leads to one missing source in the spectrum. In Figure 7.3, resolutions of the three algorithms are compared to each other. The peaks are magnified to obtain a better view of the resolution.

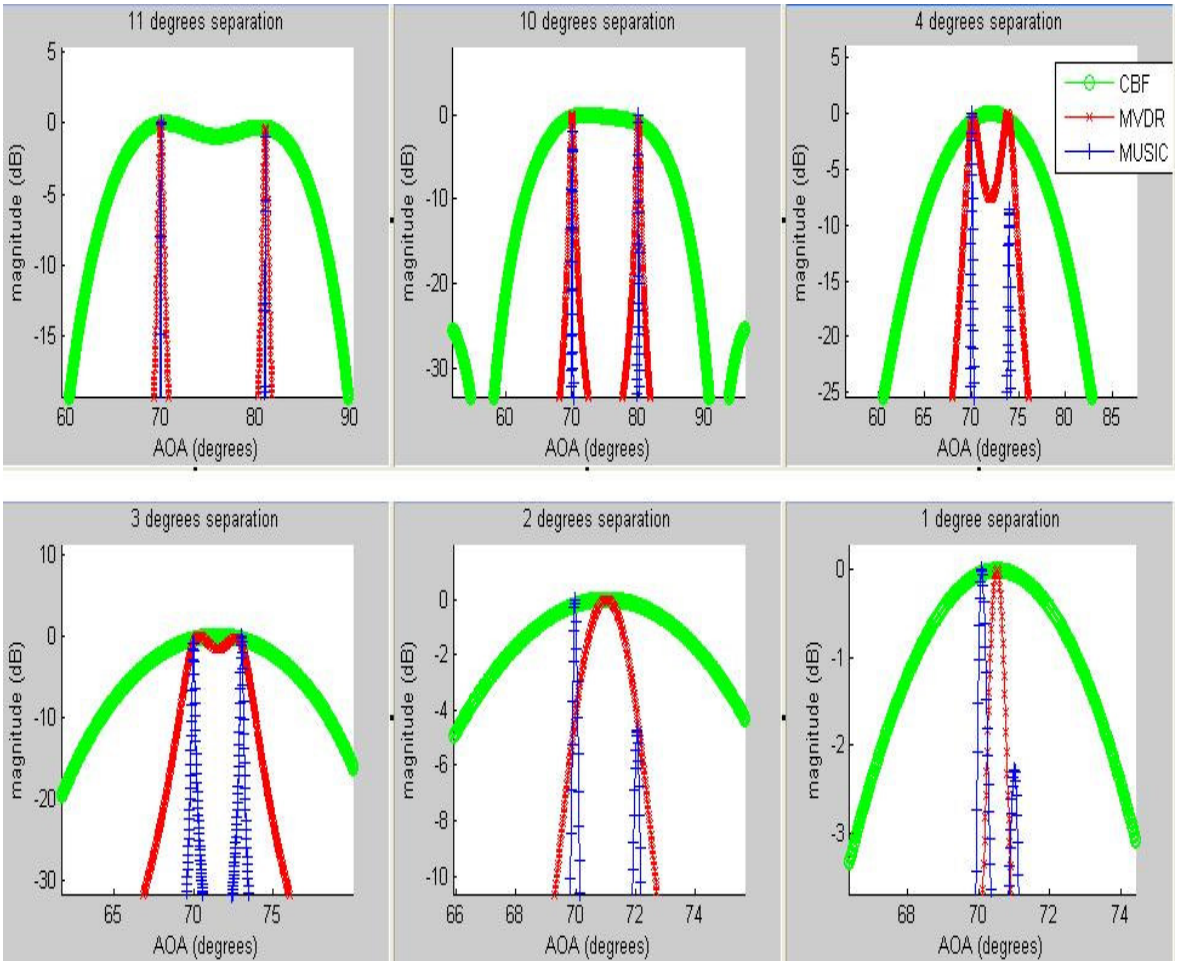


Figure 7.3. AOA resolutions for two sources (m = 10, $d = 0.5\lambda$, $\theta_1 = 70^\circ$, SNR = 20dB, $\sigma_e = 1$, N = 1000)

When two peaks unite to appear as one peak, the AOA estimate seen from that peak does not belong to either of the two sources. Instead it indicates the average of the two AOA estimates.

Simulations above were made with 10001 steering points; that is, the angular spectrum between 0-180 degrees was divided into 10000 points and numerical results were

collected from these points. The other parameters regarding the model are same as in the previous single source case.

For two sources 11 degrees apart, all three algorithms have two peaks for the two sources. When the separation becomes 10 degrees, CBF shows only one peak. Two distinct peaks can be observed with MVDR when there is 3 degrees separation. But these two peaks become one when the two sources are 2 degrees apart from each other. MUSIC can show two distinct peaks even for one degree separation between the sources. Hence, under these conditions MUSIC outperforms both MVDR and CBF when multiple sources are considered. In order to observe both AOA estimates, the minimum separation needed for CBF is 11 degrees, whereas this limit is 3 degrees for MVDR. MUSIC resolution is enough to verify the number of sources even for 1 degree separation.

8.1.1. Performance Graphs for Single Point Source

Some AOA estimation algorithms were presented and MSE's were derived for the algorithms in Chapter 6. In this section, these MSE's are used to analyse the performance of the algorithms with respect to changes in different parameters. Below are the performance graphs of the three algorithms and their comparison with the "General CRLB" given in section 7.1.

A single source with an AOA of 70 degrees is assumed to be present in the environment. The BTS is composed of a ULA with 10 elements. Results of 2000 simulations were used to obtain Root-MUSIC estimates. CBF and MVDR graphs were obtained from the theoretical MSE expressions derived in the previous chapter. The first graph in Figure 7.4 shows the difference in MSE when the number of samples N of received signals is increased. As the number of samples is increased CRLB becomes smaller and all algorithms tend to perform better. CBF, MVDR and Root-MUSIC graphs stay above the CRLB since CRLB is a lower bound for the variance of the AOA estimates. Among all algorithms, the least error is observed to be of the Root-MUSIC which is closest to the CRLB. The increase in N provides more samples to the algorithms and hence the AOAs are estimated with better accuracy.

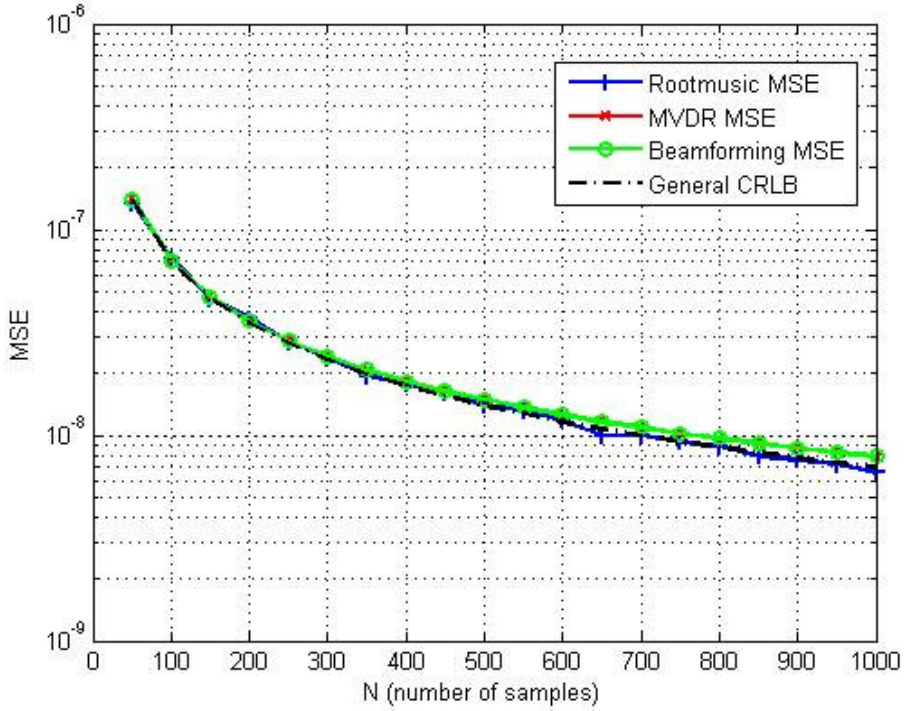


Figure 7.4. Performance graph with respect to N (m=10, $d = 0.5\lambda$, n=1, $\theta = 70^{\circ}$, SNR=20dB, $\sigma_e = 1$/// Simulations=2000)

Another parameter that can affect the AOA estimation performance is the number of antenna array elements. This has a similar reason as the number of samples to be effective on the results. When the number of elements in the ULA are increased, the number of received samples at the antenna array are increased. In Figure 7.5, it is clear that performances of all algorithms get better as the number of array elements increases. The smallest MSE is obtained by the Root-MUSIC as in the previous case.

The separation between the array elements may lead to "front-back ambiguity" as stated in section 4.2. This parameter can also have an effect on the AOA estimation accuracy. An increase in the separation between the elements results in a decrease in the antenna correlation. This in turn leads to better results for the AOA estimation process. In Figure 7.6, performances of all algorithms increase as the separation increases. In general antenna separation should be kept smaller than half wavelength in order to avoid aliasing effects.

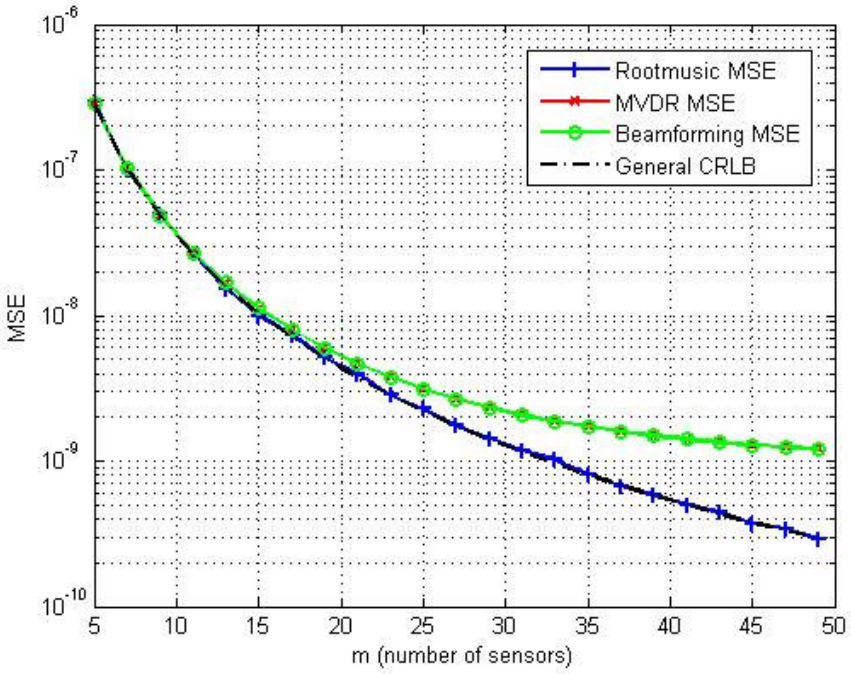


Figure 7.5. Performance graph with respect to m ($\mathrm{d} = 0.5\lambda$, n=1, $\theta = 70^{\circ}$, N=200, SNR=20dB, $\sigma_e = 1$ /// Simulations=2000)

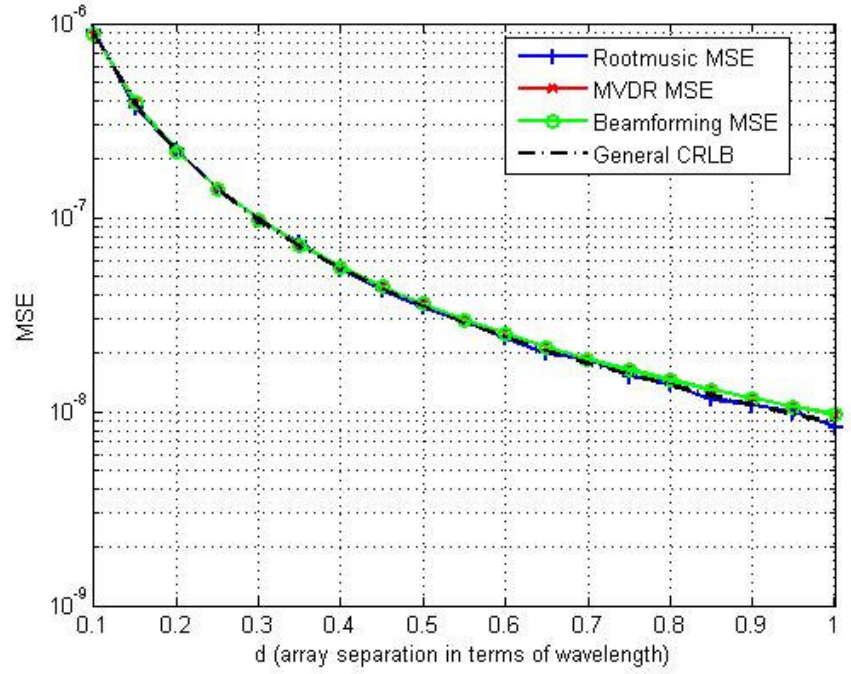


Figure 7.6. Performance graph with respect to d (m=10, n=1, $\theta = 70^{\circ}$, N=200, SNR=20dB, $\sigma_e = 1$ /// Simulations=2000)

For the point source model, signal to noise ratio is defined as

$$SNR = 10\log\left(\frac{\sigma_s^2}{\sigma_e^2}\right) \tag{8.1}$$

As expected, an increase in the SNR means noise effects will be reduced and MSE will decrease for AOA algorithms. Performance graph showing the comparison of three algorithms is seen in Figure 7.7.

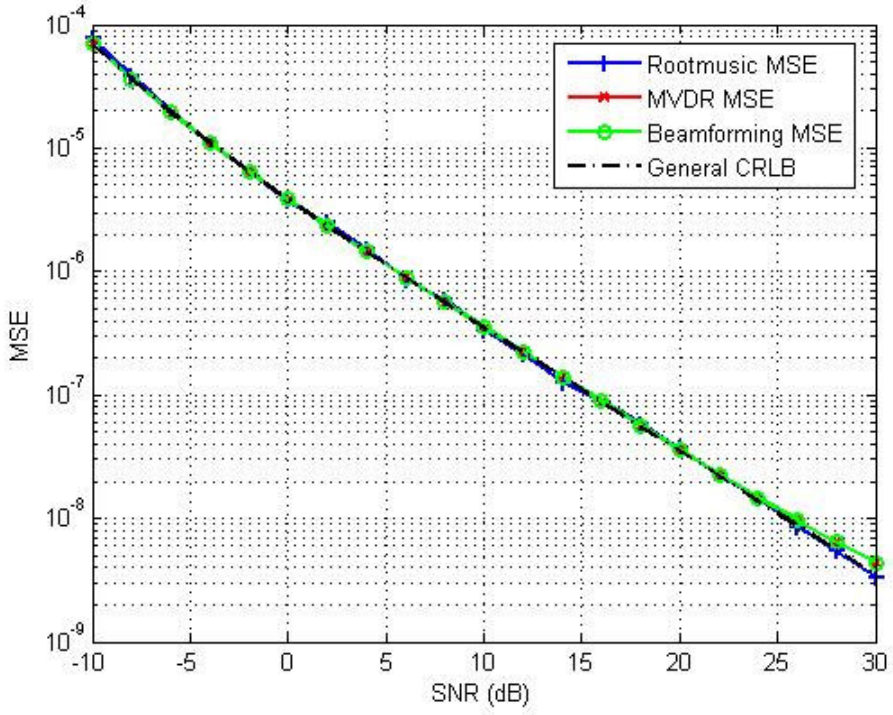


Figure 7.7. Performance graph with respect to SNR (m=10, $d = 0.5\lambda$, n=1, $\theta = 70^\circ$, N=200, $\sigma_e = 1$ /// Simulations=2000)

8.1.2. Performance Graphs for Multiple Point Sources

If the number of sources is increased, there will be a degrading in the estimation accuracy. The reason is that, complexity of estimation increases since the correlation function is made up of all the source signals. For instance, the case of three sources is

considered in the following simulations. Figure 7.8 shows the related graph. The first noticeable thing is the increase in the MSE for all simulations. For the case of multiple sources, graphs reflect the MSE of the source in the middle; that is the one with AOA of 70 degrees. As the number of samples increases, the MSE of Root-MUSIC and MVDR still stay close to CRLB but there is significant degradation for Conventional Beamforming. For *N* greater than 200, the MSE of CBF does not show much change.

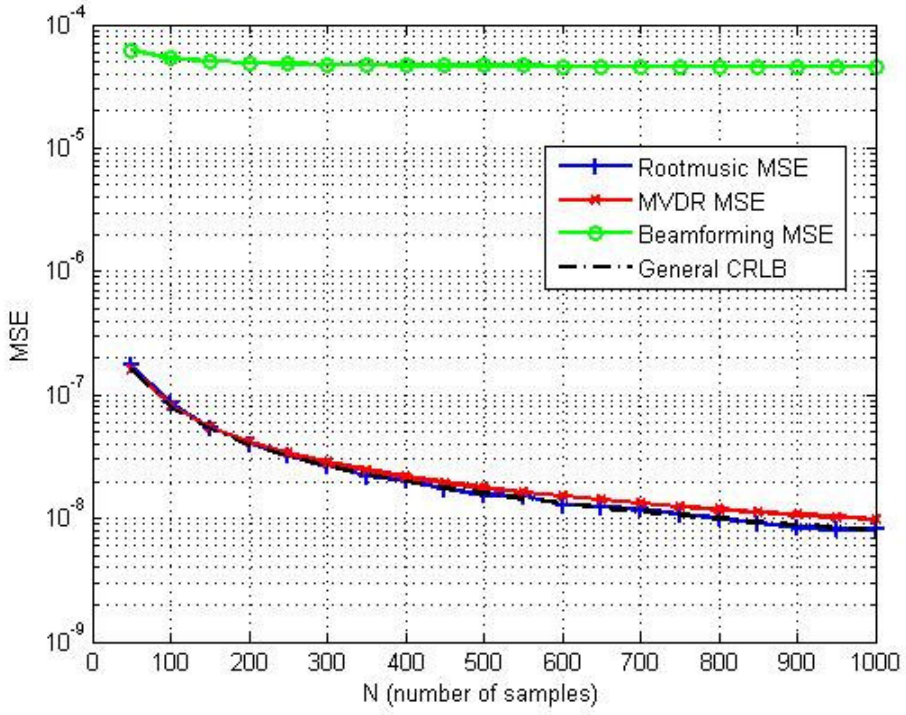


Figure 7.8. Performance graph with respect to *N* for three sources (m=10, $d = 0.5\lambda$, n=3, $\theta_1 = 40^\circ$, $\theta_2 = 70^\circ$, $\theta_3 = 90^\circ$, SNR=20dB, $\sigma_e = 1$/// Simulations=2000)

Although all methods performed similarly with increasing SNR for a single source, MSE in the multiple source case stays near 10^{-4} for CBF as seen in Figure 7.9. The worst performance appears to be of the CBF. Root-MUSIC and MVDR have good MSE values when compared to the CRLB.

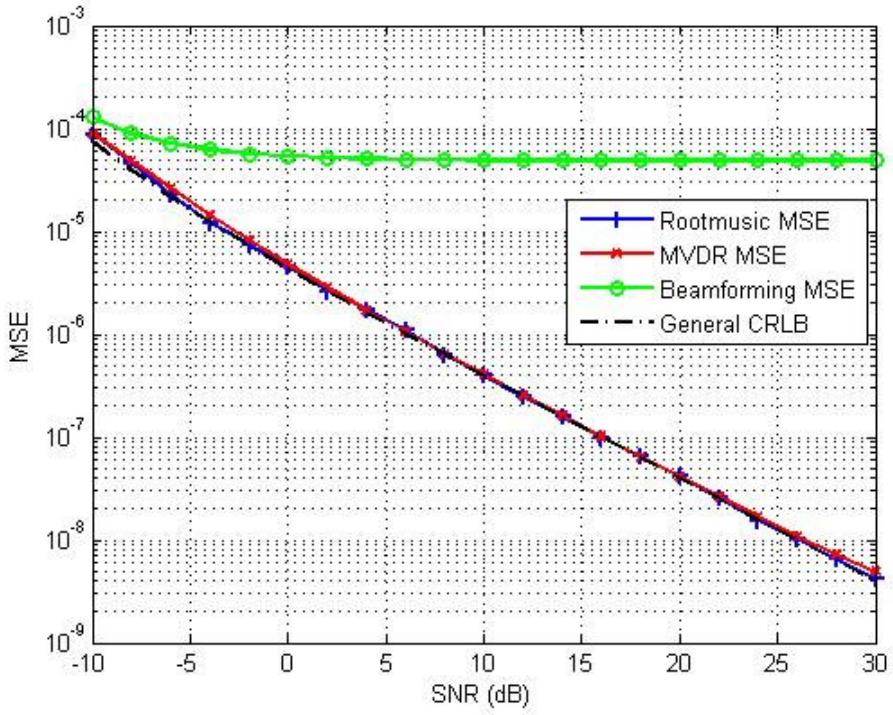


Figure 7.9. Performance graph with respect to SNR for three sources (m=10, $d = 0.5\lambda$, n=3, $\theta_1 = 40^\circ$, $\theta_2 = 70^\circ$, $\theta_3 = 90^\circ$, N=200, $\sigma_e = 1$/// Simulations=2000)

8.2. Distributed Source Model Simulations

In real life, especially in urban environment there is scarcely line-of-sight between a mobile station and a BTS antenna. Objects around the mobile may cause the signal to be reflected, refracted and scattered in different directions resulting in a multipath environment. Under these conditions, the point source model mentioned above may not provide reliable AOA estimates for the incoming signals. Von Mises model was presented in Chapter 5 for the situations where there are many local scatterers around the mobile station. Simulation results regarding the distributed source model will be interpreted in this section.

These simulations use the von Mises correlation model and thus assume narrowband directional sources and white Gaussian noise.

In the analysis of the von Mises model simulations, one of the most important parameters is κ. When κ is equal to or greater than 8.2, the pdf resembles Gaussian

model. Larger κ means that the scatterers are more concentrated at a certain direction between the mobile station and the ULA. The less the scatterers are distributed, the better the performance of the von Mises model is. Below are two AOA estimate graphs for two different values of κ. As seen in Figure 7.10, the peaks of all three AOA estimation methods get sharper as κ gets larger. In both cases, MUSIC performs better than the other two methods which is evident from the sharpness of its peak.

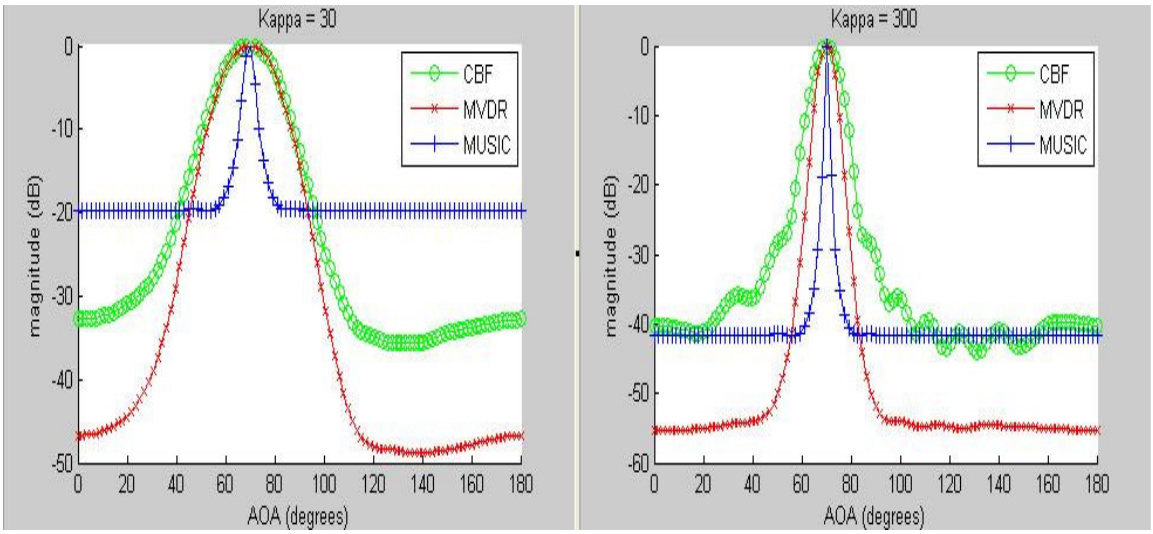


Figure 7.10. Estimated AOAs from different algorithms - Distributed source model (m=10, $d = 0.5\lambda$, $\theta = 70^{\circ}$, SNR=20dB, $\sigma_e = 1$, N=1000)

When there are two sources 20 degrees apart, CBF and MVDR resolutions are not good enough to separate the two sources from each other for $\kappa = 30$. Figure 7.11 shows the AOA estimation graphs for this case when $\kappa = 30$ and $\kappa = 300$. All methods provide good results for the two sources when $\kappa = 300$.

If compared to the point source model, these graphs reveal the fact that small values of κ do not provide as sharp peaks as the point source model for all three methods. MUSIC can be used to obtain the best results in both point source model and the distributed source model. For large values of κ, the AOA estimates from three methods are all acceptable results with small values of error.

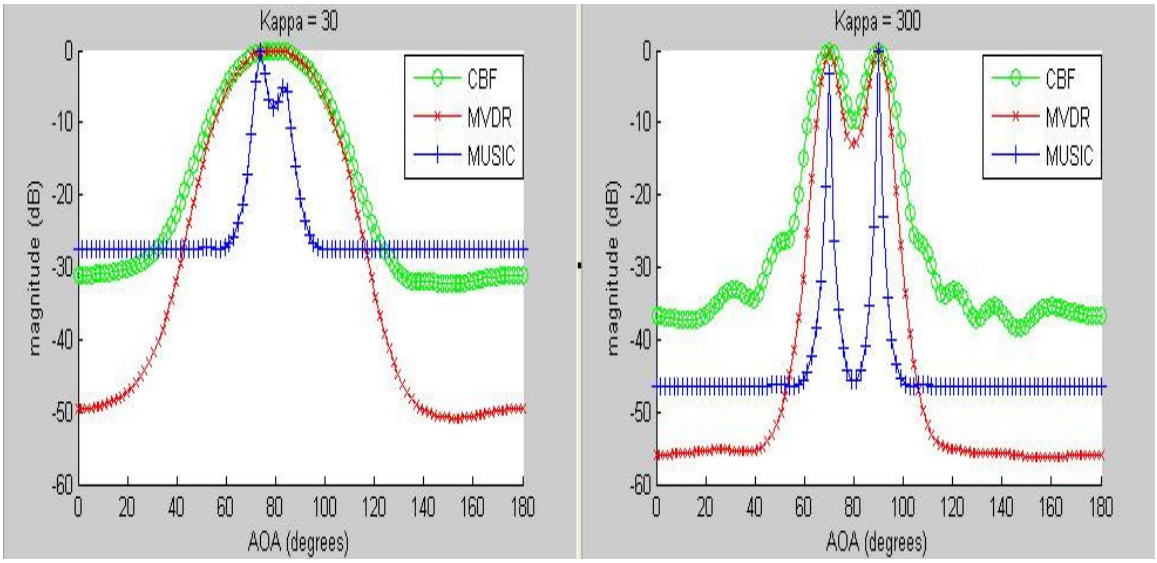


Figure 7.11. Estimated AOAs for two sources 20 degrees apart - Distributed source model

(m = 10, $d = 0.5\lambda$, $\theta_1 = 70^\circ$, $\theta_2 = 90^\circ$, SNR = 20dB, $\sigma_e = 1$, N = 1000)

In order to gain a deeper knowledge of the performance with the von Mises pdf, resolutions of the three algorithms are analysed with closer view of the peaks for different AOA separations of two sources.

The first figure below, Figure 7.12 displays the ability of the three algorithms to resolve the two sources from each other when $\kappa = 30$. For this value of κ, CBF and MVDR can generate two peaks when the AOA separation between the two sources is at least 23 degrees. Closer sources result in single peaks for these two methods. On the other hand, MUSIC can always show two peaks even when the sources are 1 degree apart. However with $\kappa = 30$, separation between the AOA estimates is not really observed to be 1 degree in the graph. It occurs to be around 10 degrees and the AOA estimates (66 and 76) have about 5 degrees of error. For larger AOA separations (e.g. 93-70 = 23 degrees), the estimates are closer (90-75 = 15 degrees) than they should be and for smaller separations (e.g. 71-70 = 1 degree) they are further (76-66=10 degrees) than real values.

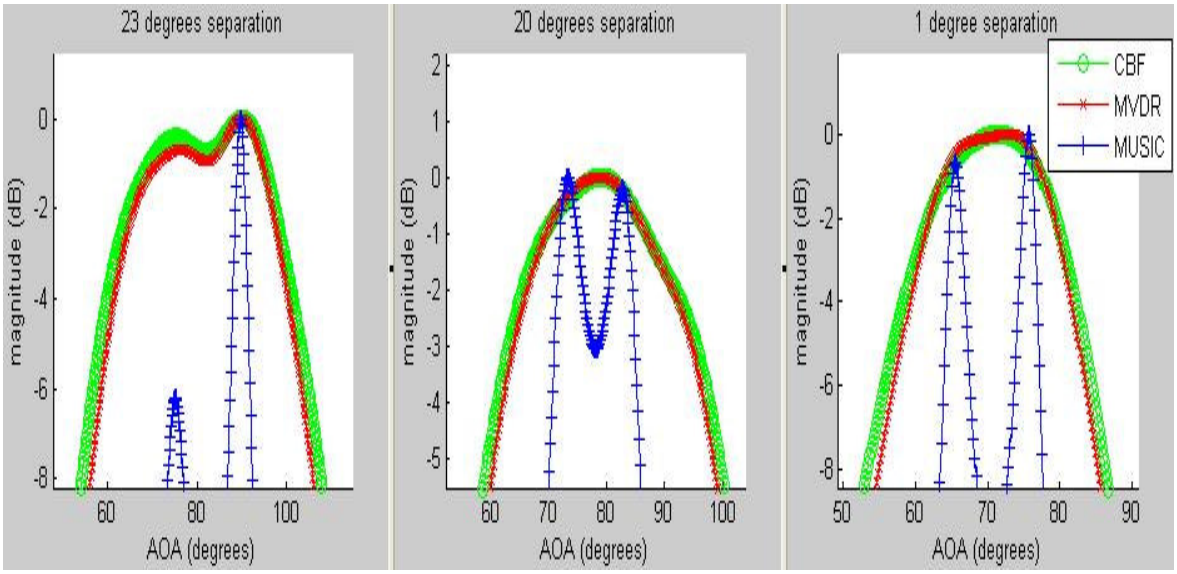


Figure 7.12. AOA resolutions for two sources (m = 10, $d = 0.5\lambda$, $\theta_1 = 70^\circ$, SNR = 20dB, $\sigma_e = 1$, N = 1000, $\kappa = 30$)

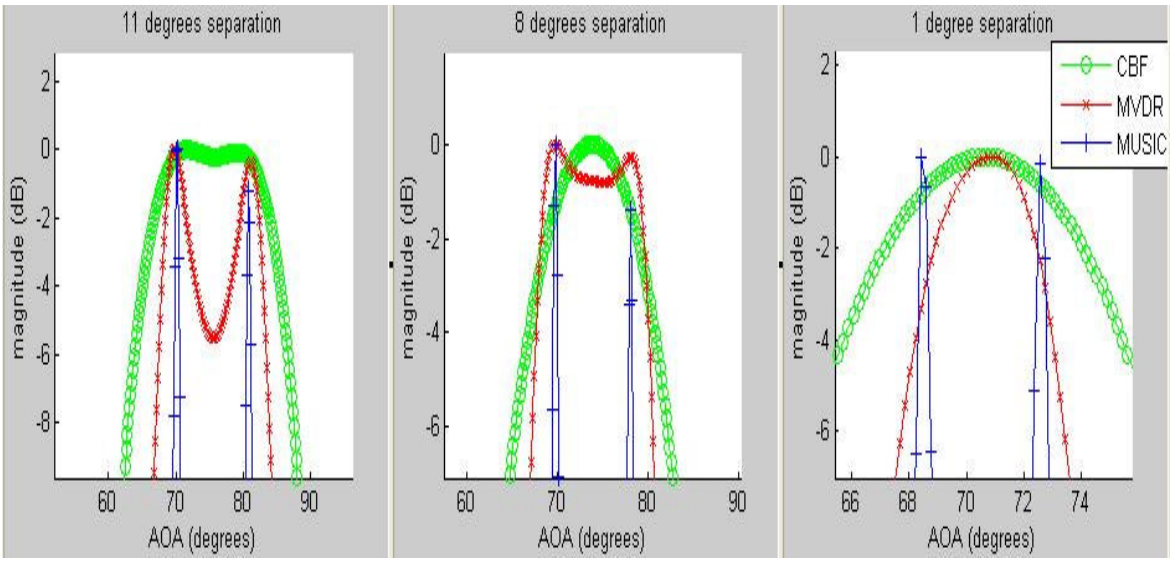


Figure 7.13. AOA resolutions for two sources (m = 10, $d = 0.5\lambda$, $\theta_1 = 70^\circ$, SNR = 20dB, $\sigma_e = 1$, N = 1000, $\kappa = 700$)

Figure 7.13 shows AOA resolutions for $\kappa = 700$. When the concentration parameter is much larger, CBF can separate two sources as close as 11 degrees. MVDR can also perform better and two sources 8 degrees apart can be identified with the AOA estimate peaks. MUSIC is still the one with best resolution and it provides sharper peaks with better resolution when compared to the results with smaller values of κ. The improvement with

larger κ shows itself also in the accuracy of the AOA estimates. With 11 and 8 degrees separation, both AOA estimates exhibit small errors and the AOA separation is reflected to the peaks. MUSIC performance still stays worse than the point source case for one degree separation and it reflects a gap of about four degrees between the AOA estimates in the graph. Each peak (at 68,5 and 72,5) varies about 1,5 degrees from the original AOAs (70 and 71).

8.2.1. Performance Graphs for Single Distributed Source

Distributed source simulations were made with the usage of the von Mises pdf in order to obtain some performance parameters for AOA estimation. Same AOA estimation methods were used also with this model to illustrate the effect of multipath environment in AOA estimation. Location estimation for cellular networks was studied with the distributed source model.

Assuming that signals from a single source arrive at the ULA, theoretical MSEs for CBF and MVDR were calculated. 2000 simulations were performed to obtain the MSE of Root-MUSIC. All three parameters were compared to the CRLB obtained with the correlation function from the von Mises model. MSE graphs were obtained by changing the parameters that effect the AOA estimation performance. κ, number of samples *N*, number of array elements *m*, array separation *d* and SNR were the varying parameters in the simulations.

Simulation results with respect to κ are shown in Figure 7.14. AOA estimation results for all algorithms provide similar MSEs for different values of κ. As κ gets larger, MSE gets better and the gap between the three algorithms and CRLB gets smaller which means better performance can be achieved for larger values of κ. Root-MUSIC MSE is the smallest for small values of κ but as it gets larger CBF performs slightly worse than Root-MUSIC and MVDR performance also improves. Since small values of κ indicate more distributed local scatterers around the mobile station, these results are expected consequences of the simulations.

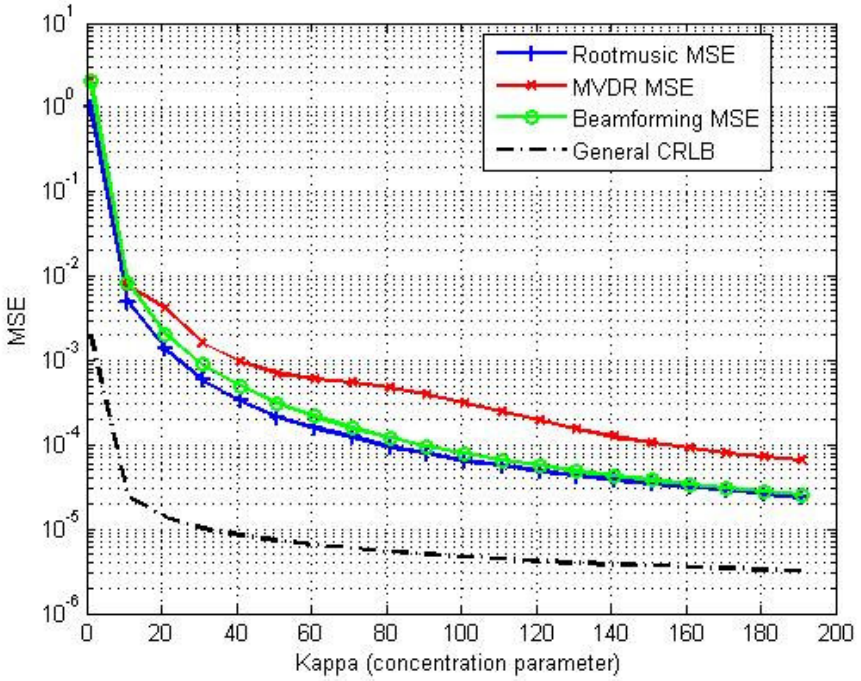


Figure 7.14. Performance graph with respect to κ – Distributed source model (m=10, $d = 0.5\lambda$, n=1, $\theta = 70^{\circ}$, N=200, SNR=20dB, $\sigma_e = 1$/// Simulations=2000)

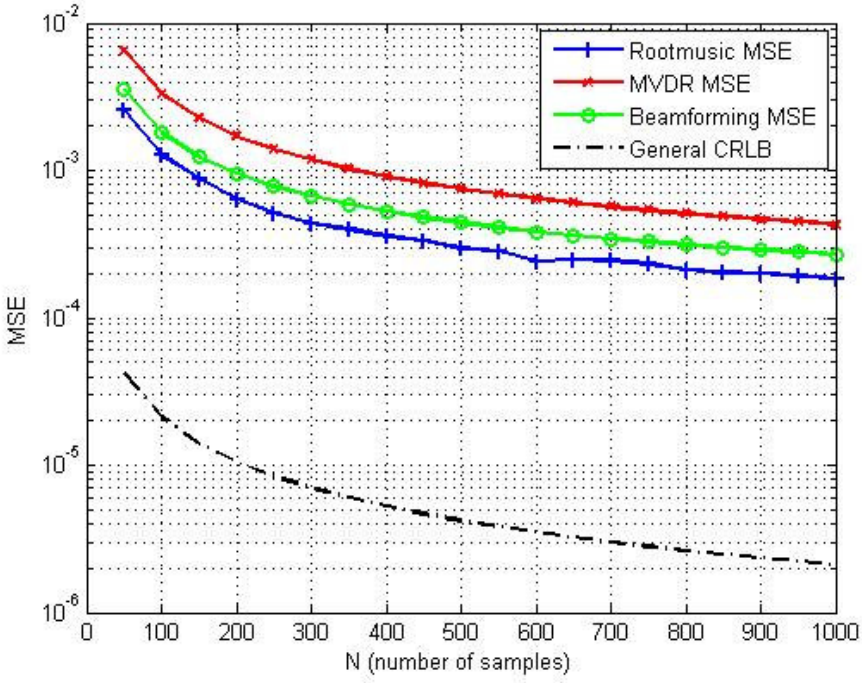


Figure 7.15. Performance graph with respect to N – Distributed source model (m=10, $\kappa = 30$, $d = 0.5\lambda$, n=1, $\theta = 70^{\circ}$, SNR=20dB, $\sigma_e = 1$/// Simulations=2000)

Figure 7.15 shows the MSE versus N graph for CBF, MVDR, Root-MUSIC algorithms and also CRLB for the multipath model. When compared to the point source model, MSEs for CBF, MVDR and Root-MUSIC are much larger than the CRLB values as the number of samples changes. This degradation is caused by the scatterers in the multipath environment which is outlined by a small concentration parameter. The increase in the number of samples produces better MSEs as in the point source model.

The following two graphs, Figure 7.16 and Figure 7.17 illustrate the change in AOA estimation performance due to multipath environment effects. First one is MSE versus number of array elements graph and the second one is MSE versus array separation graph. Noticeable difference is that MSEs of all algorithms get worse as the number of array elements and the separation between array elements increase.

The performance of AOA estimation algorithms degrade severely in multipath environments where multiple versions of the same signal arrive within the symbol duration. Multipath signal environment affects the structure of the received signal correlation functions.

Complex envelope lowpass equivalent of the received signal was used in section 5.1 to derive the correlation function and then to obtain the received signal form below.

$$\boldsymbol{R}_{\boldsymbol{y}} = P_{signal}\,\boldsymbol{R}_{\boldsymbol{z}} + P_{noise}\,\boldsymbol{I} = P_{signal}\left(\boldsymbol{T}\boldsymbol{T}^{H}\right) + P_{noise}\,\boldsymbol{I} \tag{8.2}$$

$$\boldsymbol{y} = \sigma_s\,\boldsymbol{T}\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2 \tag{8.3}$$

In the equation above $\sigma_s\,\boldsymbol{T}$ and $\boldsymbol{\rho}_1$ do not correspond exactly to $\boldsymbol{a}(\theta)$ and $\boldsymbol{s}(t)$ from the point source model. In this spatial model, $\boldsymbol{R}_z$ and thus $\boldsymbol{R}_y$ are made up of the sum of spatial multipath signals and the time sequence is lost in this form of spatial sum of signals. This spatial sum results in loss of phase information related to signals and also point source assumption becomes inappropriate due to angular spread caused by local scattering.

The signals are said to be fully coherent when the amplitude and phase of a signal at any sensor exactly determines its amplitude and phase at all other sensors. When the signals undergo scattering and angular spread, loss of spatial coherence can be observed for the signals across the ULA. Propagation of signals through randomly time-variant (refracting) medium or scattering of signals by a randomly time-varying surface may cause imperfect spatial coherence [30].

MUSIC is based on the assumption of received signals from uncorrelated point sources which means that $\boldsymbol{R}_s$ is full rank and MUSIC eigendecomposition gives satisfactory results. When multipath signals make up the correlation function, the rank of $\boldsymbol{A}\boldsymbol{R}_s\boldsymbol{A}^H$ is not equal to n and the rank condition is not satisfied [20]. Thus the assumption of point sources fails for MUSIC and AOA estimation results are degraded. The covariance matrix of a spatially distributed source is theoretically full rank and this leads to performance degradation when subspace based methods are used [31]. CBF and MVDR are affected in the same way because of the change in the correlation function formation. MVDR is affected more than CBF in some cases since MVDR assumes point source in its model and CBF is based only on measuring the output power from every direction. A perfectly coherent source's contribution to the spatial covariance matrix is a rank-one component whereas each source contributes a full-rank matrix in the case of angular spread [32].

In order to interpret the degradation of AOA estimates with increasing m and d, multipath environment effects have to be considered thoroughly. In point source simulations, the increase in these parameters provided more information for AOA estimation. However, signals passing through local scatterers and arriving at the ULA give rise to more received multipath signals as the number of array elements that receive these signals increases. When there are more ULA elements, more spatial signals are collected and complexity of the correlation function increases leading to less accuracy in AOA estimates. When the separation between the sensors increases, correlations between the sensors fall off and the signal at a sensor no longer exactly determines its value at all other sensors leading to loss of coherence. Loss of coherence in turn results in performance degradation [30].

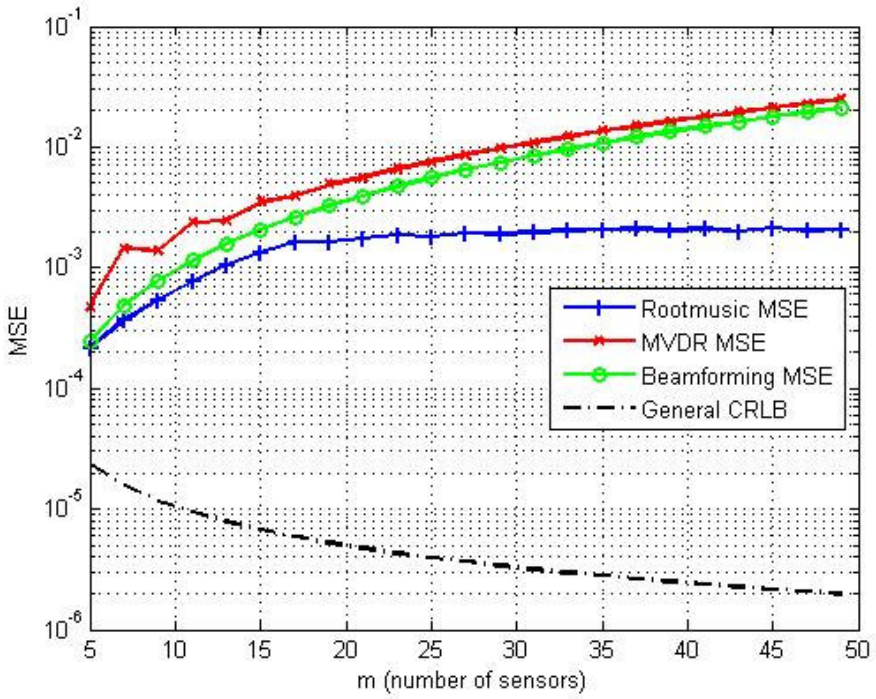


Figure 7.16. Performance graph with respect to m – Distributed source model ($\kappa = 30$, d $= 0.5\lambda$, n=1, $\theta = 70^{\circ}$, N=200, SNR=20dB, $\sigma_e = 1$/// Simulations=2000)

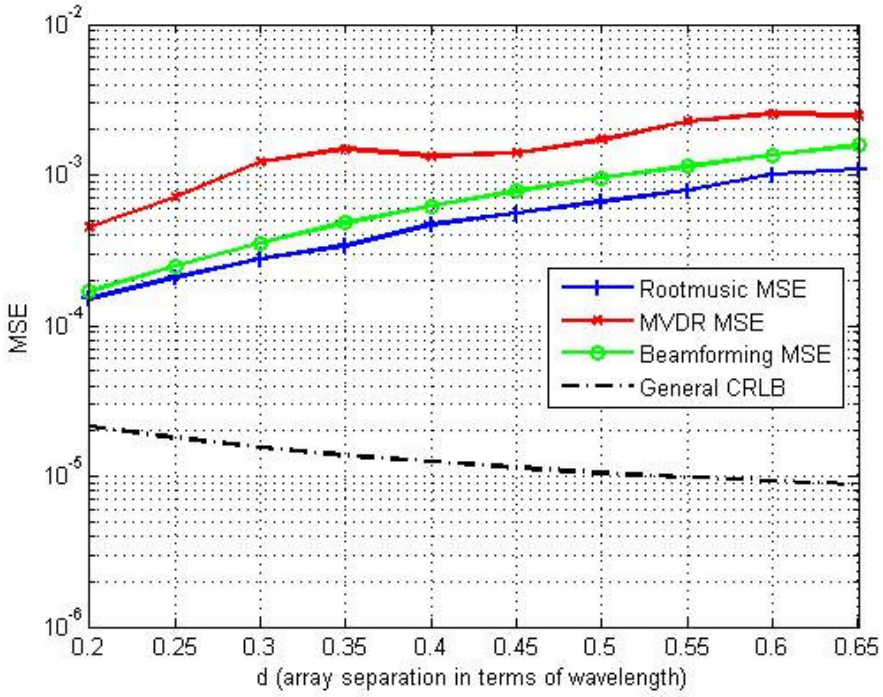


Figure 7.17. Performance graph with respect to d – Distributed source model (m=10, $\kappa = 30$, n=1, $\theta = 70^{\circ}$, N=200, SNR=20dB, $\sigma_e = 1$/// Simulations=2000)

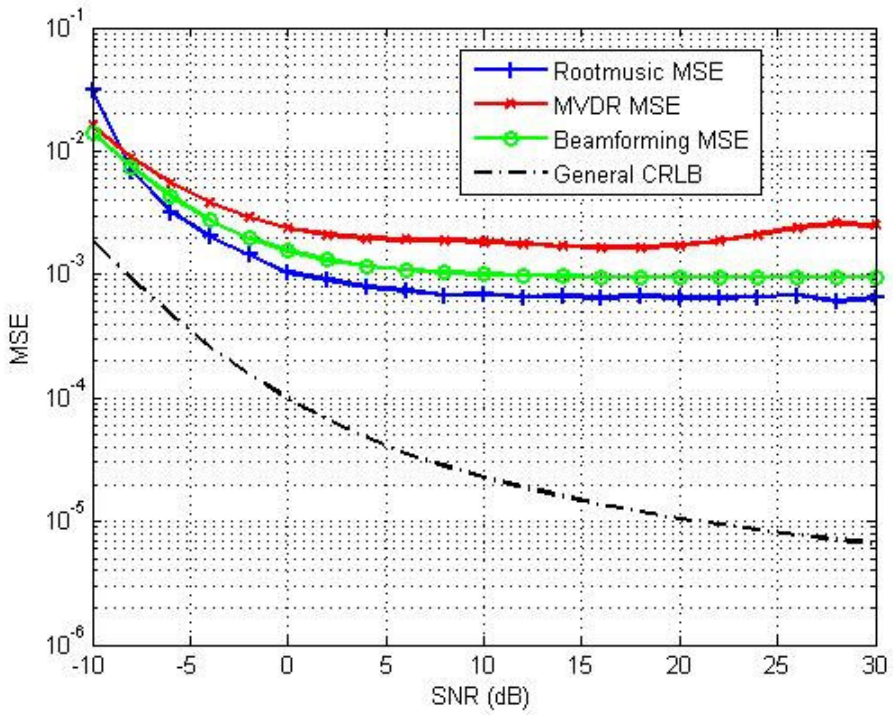


Figure 7.18. Performance graph with respect to SNR – Distributed source model (m=10, $\kappa = 30$, d $= 0.5\lambda$, n=1, $\theta = 70^{\circ}$, N=200, $\sigma_e = 1$ /// Simulations=2000)

MSEs of CBF, MVDR and Root-MUSIC are much higher than the CRLB for different values of the SNR with multipath environment simulations. For the distributed source model, signal to noise ratio is defined as

$$SNR = 10\log\left(\frac{\sigma_s^2}{\sigma_e^2}\right) \tag{8.4}$$

An increase in the SNR leads to a slight decrease in the MSEs but the MSE levels are kept unchanged around 10^{-3} for the three methods. Figure 7.18 shows the effect of SNR.

Changes in distributed source General CRLBs were also analysed by considering two parameters at a time. Below are General CRLB graphs of AOA and κ for a single source multipath environment.

It is obvious from Figure 7.19 that General CRLB of AOA estimate obtained for a single source gets lower as the number of samples and the number of antenna array elements increase.

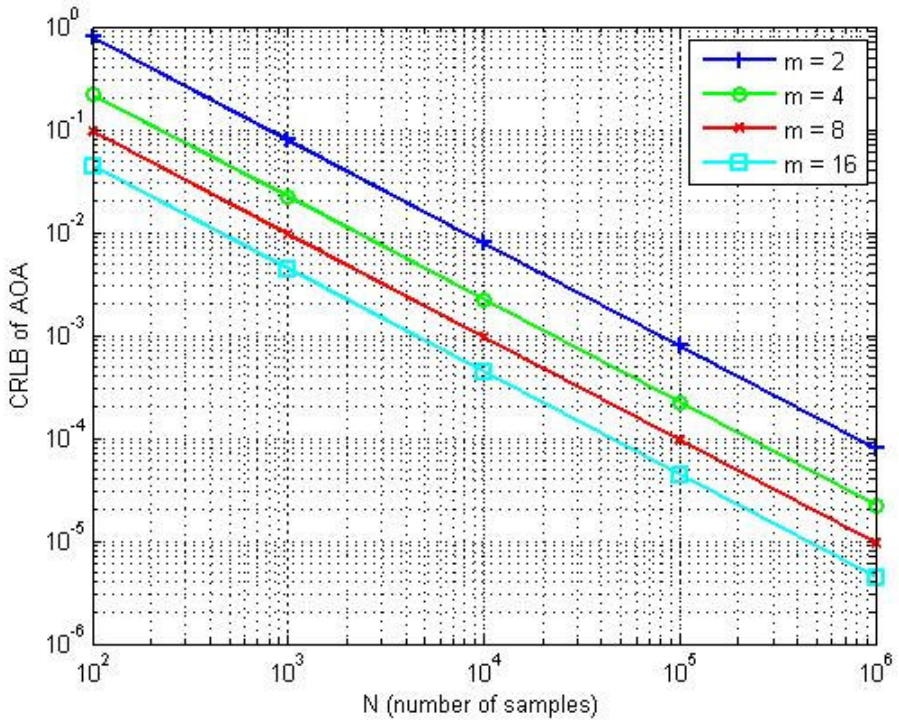


Figure 7.19. General CRLB graph for AOA changing according to N and m ($\kappa = 26.93$, $d = 0.5\lambda$, n=1, $\theta = 75^{\circ}$, SNR=20dB, $\sigma_e = 1$)

Another observation is that increasing m and SNR also lead to lower values of AOA CRLB for a single source. In the last graph for AOA CRLB, as SNR and array separation d increase, General CRLB values become smaller leading to a lower bound for AOA estimates. Figure 7.20 and Figure 7.21 show these results respectively.

Similar graphs can be drawn also for the CRLB values of the concentration parameter. A single source with AOA of 75 degrees was assumed and $\kappa = 26.93$ was chosen as the concentration parameter of the cluster. For increasing values of N, m, SNR and d, CRLB values of κ move down to lower values as shown in the following graphs. Figure 7.22, Figure 7.23 and Figure 7.24 show these performance graphs. CRLB changes with respect to two parameters can be seen at each graph. These graphs show consistency with the performance graphs obtained from the simulations.

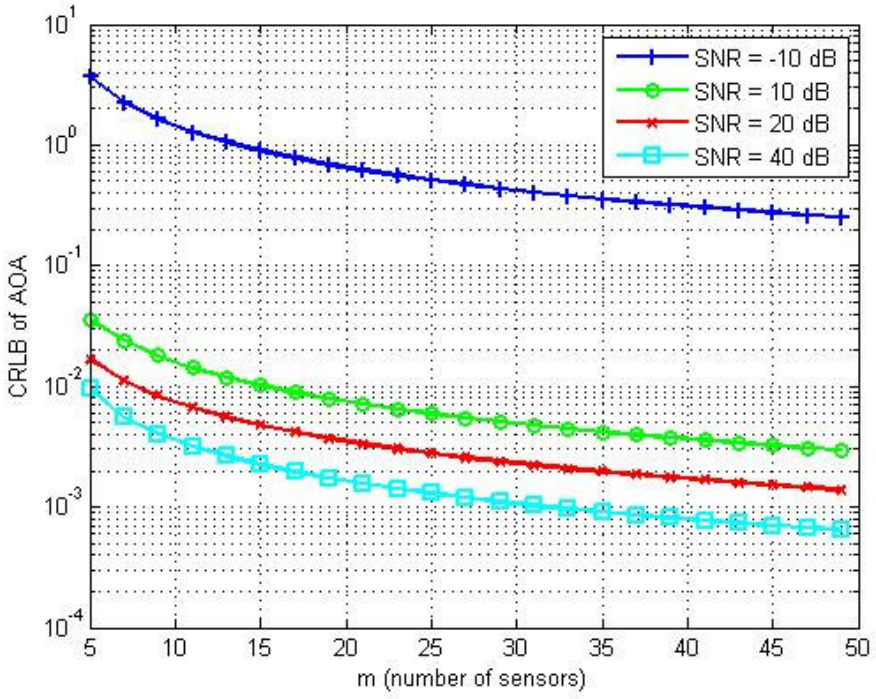


Figure 7.20. General CRLB graph for AOA changing according to m and SNR ($\kappa = 26.93$, $d = 0.5\lambda$, n=1, $\theta = 75^{\circ}$, N=1000, $\sigma_e = 1$)

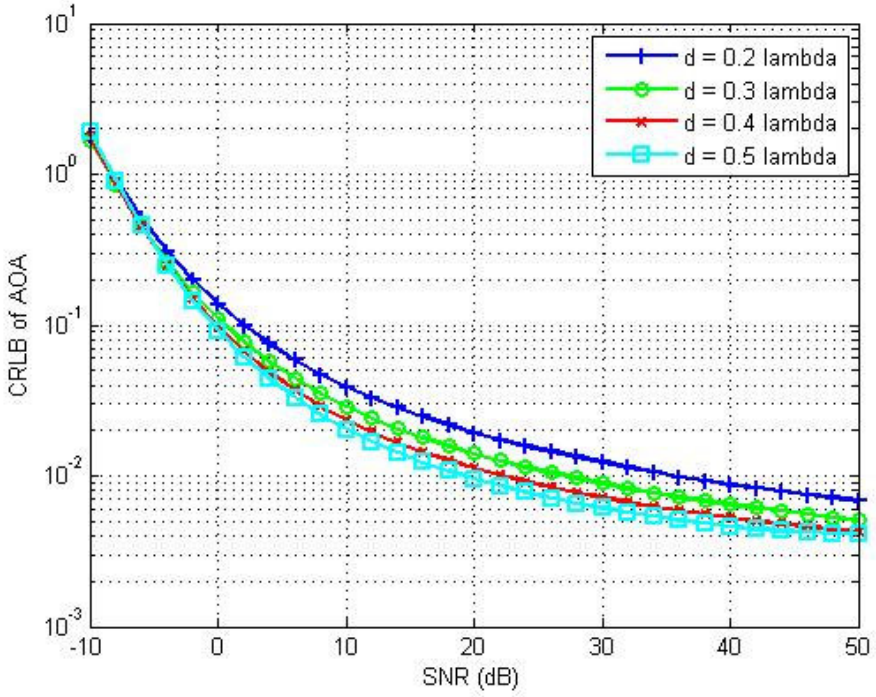


Figure 7.21. General CRLB graph for AOA changing according to SNR and d ($\kappa = 26.93$, m=8, n=1, $\theta = 75^{\circ}$, N=1000, $\sigma_e = 1$)

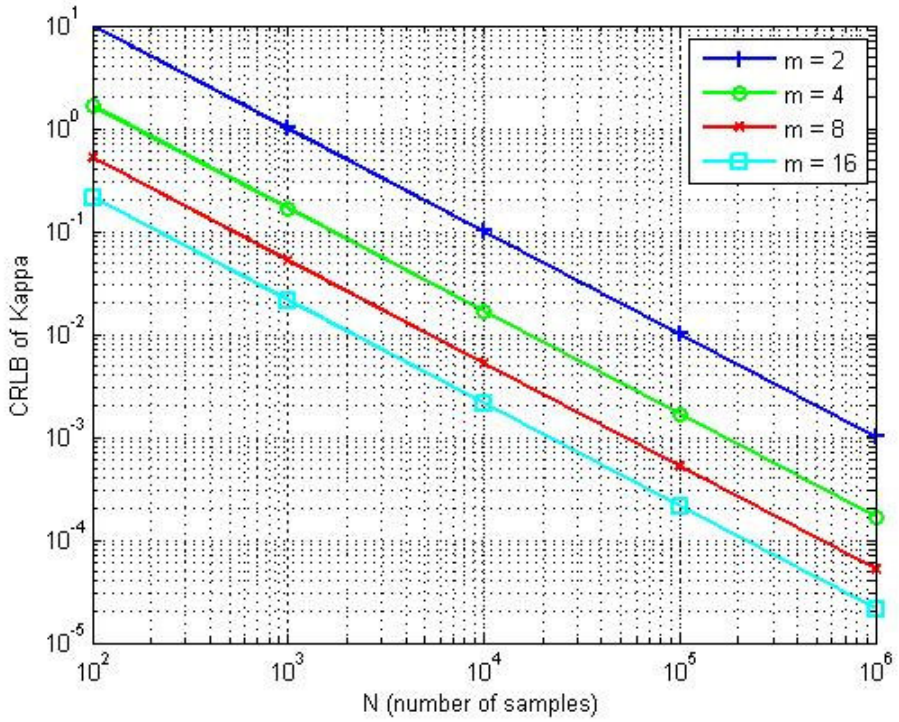


Figure 7.22. General CRLB graph for κ changing according to N and m ($\kappa = 26.93$, $d = 0.5\lambda$, n=1, $\theta = 75^{\circ}$, SNR=20dB, $\sigma_e = 1$)

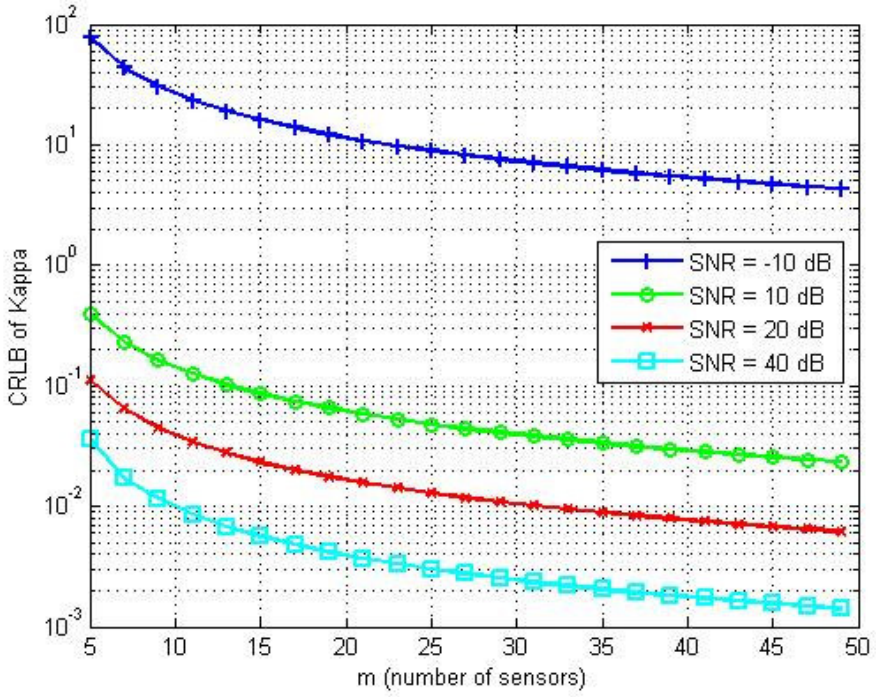


Figure 7.23. General CRLB graph for κ changing according to m and SNR ($\kappa = 26.93$, $d = 0.5\lambda$, n=1, $\theta = 75^{\circ}$, N=1000, $\sigma_e = 1$)

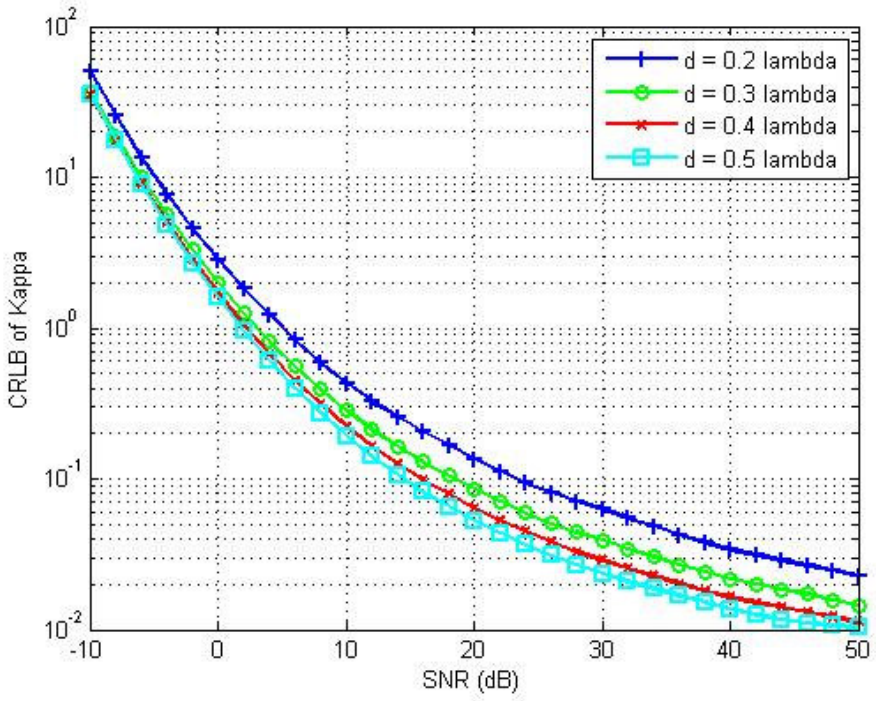


Figure 7.24. General CRLB graph for κ changing according to SNR and d ($\kappa = 26.93$, m=8, n=1, $\theta = 75^{\circ}$, N=1000, $\sigma_e = 1$)

8.2.2. Performance Graphs for Multiple Distributed Sources

In section 5.3 composite distribution is obtained by adding up the distributions from different sources in some proportion. Simulations were made to demonstrate the AOA estimation efficiency when there are multiple sources in a multipath environment with local scatterers. Three sources with different AOAs were aasumed to be existing in the environment and all three clusters were assumed to have the same concentration parameter. MSE was again chosen as the performance measure for AOA estimation by the three algorithms. For the case of multiple sources, graphs reflect the MSE of the source in the middle; that is the one with AOA of 70 degrees.

If compared to the single source case, there is slight degradation in the MSEs but it is around 10^{-4} for high values of the concentration parameter. As mentioned earlier, multiple sources can be identified with better accuracy for large values of κ and this is reflected by the smaller MSEs at the right hand side of Figure 7.25.

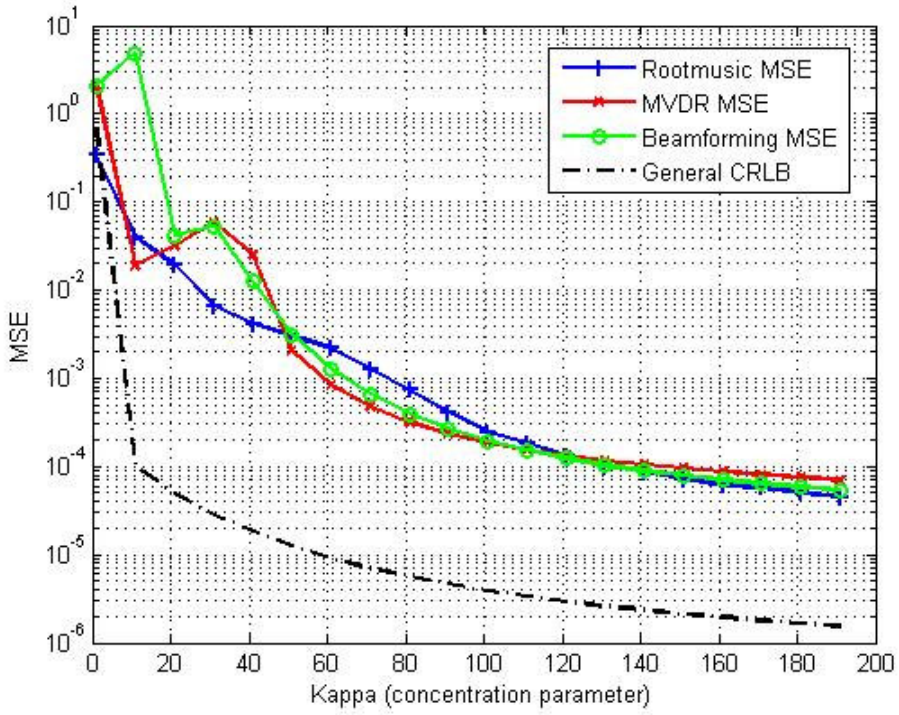


Figure 7.25. Performance graph with respect to κ for three sources – Distributed source model (m=10, $d = 0.5\lambda$, n=3, $\theta_1 = 40^\circ$, $\theta_2 = 70^\circ$, $\theta_3 = 90^\circ$, SNR=20dB, $\sigma_e = 1$, N = 200 /// Simulations=2000)

For increasing *N*, performance improvement is not as clear as for κ. CBF and MVDR have approximately same results for different values of *N*. Root-MUSIC estimates are better than these two at every point of the graph as seen in Figure 7.26. However the slope is not steep even for Root-MUSIC.

All sources covered with local scatterers have the same concentration parameter in the previous multiple-source graphs. If local scatterers are distributed with different concentration around each mobile station, the each cluster will have a different κ value. As mentioned in section 5.3, signals arriving at the ULA from different clusters can then be composed as the sum of distribution functions of each cluster.

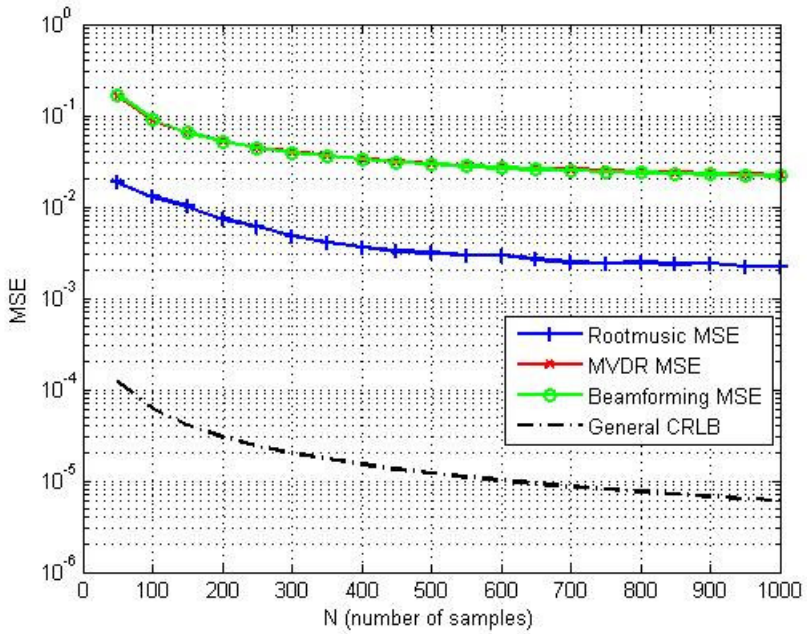


Figure 7.26. Performance graph with respect to *N* for three sources – Distributed source model (m=10, $d = 0.5\lambda$, n=3, $\theta_1 = 40^\circ$, $\theta_2 = 70^\circ$, $\theta_3 = 90^\circ$, SNR=20dB, $\sigma_e = 1$, $\kappa = 30$ /// Simulations=2000)

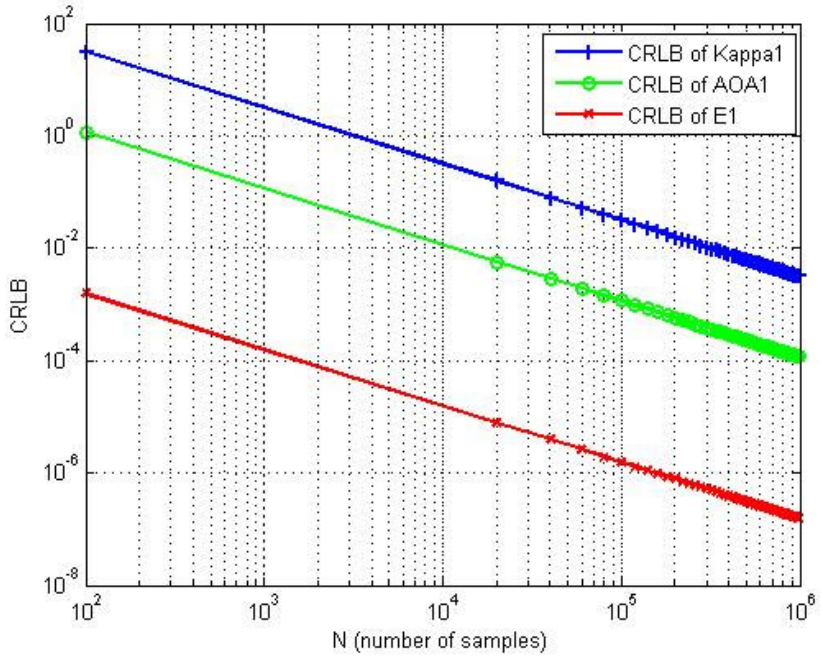


Figure 7.27. CRLB graph with respect to *N* for the first cluster – Distributed source model (m=8, $d = 0.5\lambda$, n=2, $\theta_1 = 60^\circ$, $\theta_2 = 90^\circ$, $\sigma_e = 1$, $\kappa_1 = 45$, $\kappa_2 = 30$, SNR = 20, E1 = 0.5, E2 = 0.5)

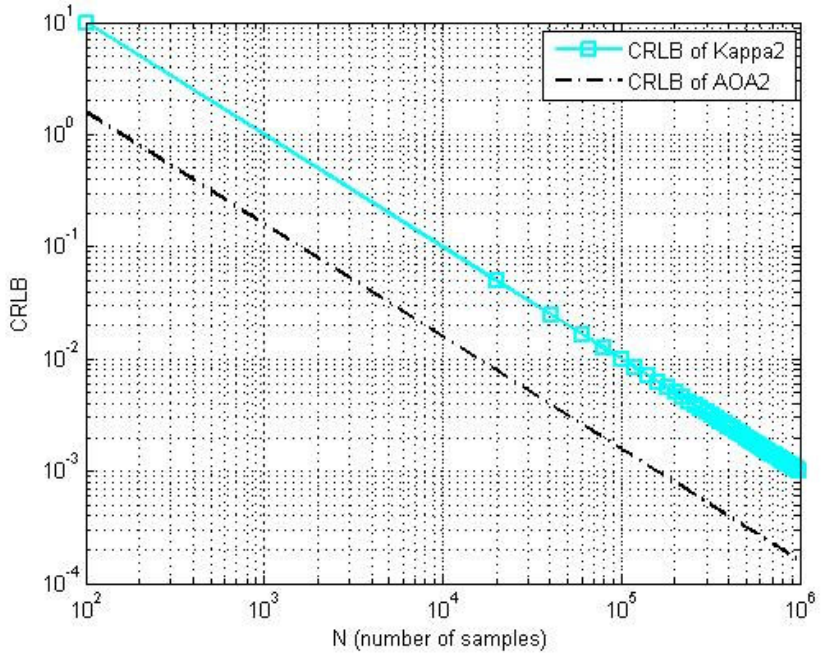


Figure 7.28. CRLB graph with respect to *N* for the second cluster – Distributed source model (m=8, $d = 0.5\lambda$, n=2, $\theta_1 = 60^\circ$, $\theta_2 = 90^\circ$, $\sigma_e = 1$, $\kappa_1 = 45$, $\kappa_2 = 30$, SNR = 20, E1 = 0.5, E2 = 0.5)

Two clusters with AOA values of 60 and 90 degrees were considered in the following case. Cluster concentration parameters were chosen as $\kappa_1 = 45$ and $\kappa_2 = 30$ respectively and the proportion of each cluster was set 0.5 in the composite function. A ULA with 8 elements was used to collect the incoming signals and CRLB values of κ_1, θ_1, E_1, κ_2 and θ_2 were obtained according to increasing number of samples. Figure 7.27 and Figure 7.28 show these CRLB graphs for the two clusters. As a result, it is shown that CRLBs for all 5 parameters have lower values for larger values of number of samples.

9. CONCLUSIONS

This study explored several AOA estimation methods intended to determine the location of the mobile station via uniform linear array antennas in a wireless network. Simulation results were obtained both for the point source model and for the distributed source model. Results from both models were compared with the help of previous researches in the literature.

Point source model assumed narrowband sources with line-of-sight between the MS and the BTS. Von Mises pdf was used to obtain a distributed source model with local scatterers around the MS. ULAs were assumed to be mounted on the base stations for both models.

Simulations included performance results for AOA estimates via four methods. These were conventional methods CBF and MVDR and subspace methods MUSIC and Root-MUSIC. Theoretical large-sample MSE formulas were derived for CBF and MVDR. MSE formula derivation for CBF is one of the contributions of this thesis. It was formulated with the help of a previous paper about the MSE of MVDR estimates. Simulation results were used to obtain the MSE for Root-MUSIC estimates. Asymptotic CRLB formulas for the point source and distributed source models were also derived. Simulation graphs included the comparison of MSEs and CRLBs for different variables.

Simulations for the single point source model revealed that CBF, MVDR, MUSIC and Root-MUSIC all provide good results with low MSEs. Increasing the number of samples, antenna elements, array separation and SNR leads to better performance for all methods in this model.

Resolution study showed that MUSIC can distinguish multiple point sources with better resolution than MVDR and CBF. The worst resolution capability was demonstrated by CBF. According to the results of simulations for multiple point sources, CBF performance for one source was affected most due to interference from other sources. A

degradation was observed for all simulations when compared to the single point source graphs.

The main contribution provided by this thesis is the analysis of CBF, MVDR, MUSIC and Root-MUSIC performances under multipath conditions with the von Mises model. Single distributed source simulations showed that higher values of concentration parameter, number of samples and SNR provide better results for all methods. MSE values were not as good as the point source MSEs but especially when the concentration parameter is high, good AOA estimates can be obtained. However, when the number of sensors and array separation increases, performance graphs indicate degradation of the AOA estimates for CBF, MVDR and Root-MUSIC. This degradation was found out to be caused by the phase and time sequence loss in the spatial correlation function which is made up of multipath signals arriving from numerous scatterers. This is referred to as "imperfect spatial coherence" in the literature.

When multiple distributed sources were assumed, performance degradation for one source became most obvious. This was caused by the nearby sources and by the complexity of the spatial correlation function. Resolutions of the three methods were also shown to be worse than that of the point source model. Root-MUSIC performance was found out to be the best also for the distributed source model.

To sum up, AOA estimation methods CBF, MVDR, MUSIC and Root-MUSIC have performance degradation under multipath conditions even when ULA is used at the base station with a composite scattering model. For future work, array manifold of the AOA estimation methods can be replaced with a more general model including the effects of multipath. Also performances of different AOA estimation methods can be studied with the von Mises model which includes several distributions and was found out to be a successful scattering model in the literature.

APPENDIX A: DERIVATION OF THE LARGE-SAMPLE MSE FOR CONVENTIONAL BEAMFORMING

In this section, an expression for the variance term $\text{var}(\hat{\theta}_k)$ will be obtained for (6.27) of the MSE. For the variance calculation, $\boldsymbol{y}(t)$ is assumed to be a Gaussian distributed random vector. Also $E[\boldsymbol{y}(t)\boldsymbol{y}^H(s)] = \boldsymbol{R}_{\boldsymbol{y}}\delta_{t,s}$ and $E[\boldsymbol{y}(t)\boldsymbol{y}^T(s)] = 0$. $\text{var}(\hat{\theta}_k)$ will be determined with the help of Taylor series expansion. For large N,

$$0 = f'(\hat{\theta}_k) \simeq f'(\overline{\theta}_k) + f''(\overline{\theta}_k)(\hat{\theta}_k - \overline{\theta}_k) \tag{A.1}$$

Since

$$f(\overline{\theta}) = \boldsymbol{a}^H(\overline{\theta})\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\overline{\theta}) \tag{A.2}$$

$$f'(\overline{\theta}_k) = \boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\overline{\theta}_k) + \boldsymbol{a}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{d}(\overline{\theta}_k) = 2\,\text{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\overline{\theta}_k)] \tag{A.3}$$

where $\boldsymbol{d}(\theta_k) = \dfrac{\partial \boldsymbol{a}(\theta_k)}{\partial \theta}$

$$f''(\overline{\theta}_k) = 2\,\text{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{d}(\overline{\theta}_k) + \boldsymbol{a}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{d}'(\overline{\theta}_k)] \tag{A.4}$$

Inserting the above result into (A.1),

$$0 = f'(\hat{\theta}_k) \simeq f'(\overline{\theta}_k) + f''(\overline{\theta}_k)(\hat{\theta}_k - \overline{\theta}_k) \simeq f'(\overline{\theta}_k) + 2h(\overline{\theta}_k)(\hat{\theta}_k - \overline{\theta}_k) \tag{A.5}$$

where

$$h(\theta_k) = \text{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{d}(\overline{\theta}_k) + \boldsymbol{a}^H(\overline{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{d}'(\overline{\theta}_k)] \tag{A.6}$$

$$(\hat{\theta}_k - \bar{\theta}_k) = \frac{-f'(\bar{\theta}_k)}{2h(\bar{\theta}_k)} \tag{A.7}$$

$$(\hat{\theta}_k - \bar{\theta}_k)^2 = \frac{[f'(\bar{\theta}_k)]^2}{4[h(\bar{\theta}_k)]^2} = \frac{4\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)]\}^2}{4[h(\bar{\theta}_k)]^2} = \frac{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)]\}^2}{[h(\bar{\theta}_k)]^2} \tag{A.8}$$

$$E(\hat{\theta}_k - \bar{\theta}_k)^2 = \frac{E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\}}{h^2(\bar{\theta}_k)} = \frac{\Delta}{h^2(\bar{\theta}_k)} = \mathrm{var}(\hat{\theta}_k) \tag{A.9}$$

If x is any complex-valued scalar, then

$$[\mathrm{Re}(x)]^2 = \frac{1}{2}\mathrm{Re}[|x|^2 + x^2] \tag{A.10}$$

Δ can be expressed as below in this form.

$$\Delta = E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\} = \frac{1}{2}\mathrm{Re}\left\{E\left[|\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)|^2\right] + E\left[(\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k))^2\right]\right\} \tag{A.11}$$

A useful formula for the expectation of the product of four Gaussian random variables is [33]

$$E[x_1x_2x_3x_4] = E[x_1x_2]E[x_3x_4] + E[x_1x_3]E[x_2x_4] + E[x_1x_4]E[x_2x_3] \tag{A.12}$$

$$E\left[|\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)|^2\right]=\frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}E[\underbrace{\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{y}(t)}_{x_1}\underbrace{\boldsymbol{y}^H(t)\boldsymbol{a}(\bar{\theta}_k)}_{x_2}\underbrace{\boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{y}(s)}_{x_3}\underbrace{\boldsymbol{y}^H(s)\boldsymbol{d}(\bar{\theta}_k)}_{x_4}]$$
$$=[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)][\boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{d}(\bar{\theta}_k)]$$
$$+\frac{1}{N}(\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{d}(\bar{\theta}_k))(\boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k))$$
$$=\left|\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)\right|^2+\frac{1}{N}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{d}(\bar{\theta}_k)][\boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)] \tag{A.13}$$

Similarly,

$$E\left[(\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k))^2\right]=\frac{1}{N^2}\sum_{t=1}^{N}\sum_{s=1}^{N}E[\underbrace{\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{y}(t)}_{x_1}\underbrace{\boldsymbol{y}^H(t)\boldsymbol{a}(\bar{\theta}_k)}_{x_2}\underbrace{\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{y}(s)}_{x_3}\underbrace{\boldsymbol{y}^H(s)\boldsymbol{a}(\bar{\theta}_k)}_{x_4}]$$
$$=\left[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)\right]^2+\frac{1}{N}\left[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)\right]^2 \tag{A.14}$$

As a result of (A.13) and (A.14),

$$\Delta=E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\}$$

$$=\frac{1}{2}\mathrm{Re}\left\{|\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)|^2+\left[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)\right]^2\right.$$
$$\left.+\frac{1}{N}\left[[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{d}(\bar{\theta}_k)][\boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)]+\left[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)\right]^2\right]\right\}$$

$$=\left\{\mathrm{Re}\left[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)\right]\right\}^2$$
$$+\frac{1}{2N}\mathrm{Re}\left[[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{d}(\bar{\theta}_k)][\boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)]+\left[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}\boldsymbol{a}(\bar{\theta}_k)\right]^2\right] \tag{A.15}$$

It is known that when the number of samples is large, $f'(\bar{\theta}_k)=0$. Thus,

$$\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)]=0 \tag{A.16}$$

Since $\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)]=0$, there is only imaginary part for $\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)$ and this results in

$$\mathrm{Re}\left[\left[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)\right]^2\right]=-|\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)|^2 \tag{A.17}$$

Inserting (A.16) and (A.17) into (A.15),

$$\Delta=\frac{1}{2N}\left\{[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{d}(\bar{\theta}_k)][\boldsymbol{a}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)]-\left|\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)\right|^2\right\} \tag{A.18}$$

Equation (A.18) can also be expressed in a simple form in terms of $\boldsymbol{u}$ and $\boldsymbol{v}$.

$$\Delta=\frac{1}{2N}\left\{(\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y^{-1}\boldsymbol{v})-|\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{v}|^2\right\} \tag{A.19}$$

where $\boldsymbol{u}^H=\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y$ and $\boldsymbol{v}=\boldsymbol{R}_y\boldsymbol{a}(\bar{\theta}_k)$.

Inserting the expression above for Δ into (A.9), the variance term is obtained for the asymptotic MSE.

$$\mathrm{var}(\hat{\theta}_k)=\frac{\Delta}{h^2(\bar{\theta}_k)}=\frac{\frac{1}{2N}\left\{(\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y^{-1}\boldsymbol{v})-|\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{v}|^2\right\}}{\left[\mathrm{Re}[\boldsymbol{u}^H\boldsymbol{R}_y^{-1}\boldsymbol{u}+\boldsymbol{v}^H\boldsymbol{d}'(\bar{\theta}_k)]\right]^2} \tag{A.20}$$

APPENDIX B: DERIVATION OF THE LARGE-SAMPLE MSE FOR MVDR

In this section, an expression for the variance term $\text{var}(\hat{\theta}_k)$ will be obtained for (6.48) of the MSE. There are some assumptions required for the variance calculation. $\boldsymbol{y}(t)$ is assumed to be a Gaussian distributed random vector. Also $E[\boldsymbol{y}(t)\boldsymbol{y}^H(s)] = \boldsymbol{R}_{\boldsymbol{y}}\delta_{t,s}$ and $E[\boldsymbol{y}(t)\boldsymbol{y}^T(s)] = 0$.

First of all, $\text{var}(\hat{\theta}_k)$ will be determined with the help of Taylor series expansion. For large *N*,

$$0 = f'(\hat{\theta}_k) \simeq f'(\overline{\theta}_k) + f''(\overline{\theta}_k)(\hat{\theta}_k - \overline{\theta}_k) \tag{B.1}$$

Since $f(\overline{\theta}) = \boldsymbol{a}^H(\overline{\theta})\hat{\boldsymbol{R}}_{\boldsymbol{y}}{}^{-1}\boldsymbol{a}(\overline{\theta})$,

$$f'(\overline{\theta}_k) = \boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}{}^{-1}\boldsymbol{a}(\overline{\theta}_k) + \boldsymbol{a}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}{}^{-1}\boldsymbol{d}(\overline{\theta}_k) = 2\,\text{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}{}^{-1}\boldsymbol{a}(\overline{\theta}_k)] \tag{B.2}$$

where $\boldsymbol{d}(\theta_k) = \dfrac{\partial \boldsymbol{a}(\theta_k)}{\partial \theta}$

$$f''(\overline{\theta}_k) = 2\,\text{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}{}^{-1}\boldsymbol{d}(\overline{\theta}_k) + \boldsymbol{a}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_{\boldsymbol{y}}{}^{-1}\boldsymbol{d}'(\overline{\theta}_k)] \tag{B.3}$$

Inserting the above result into (B.1),

$$0 = f'(\hat{\theta}_k) \simeq f'(\overline{\theta}_k) + f''(\overline{\theta}_k)(\hat{\theta}_k - \overline{\theta}_k) \simeq f'(\overline{\theta}_k) + 2h(\overline{\theta}_k)(\hat{\theta}_k - \overline{\theta}_k) \tag{B.4}$$

where

$$h(\theta_k) = \text{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}{}^{-1}\boldsymbol{d}(\overline{\theta}_k) + \boldsymbol{a}^H(\overline{\theta}_k)\boldsymbol{R}_{\boldsymbol{y}}{}^{-1}\boldsymbol{d}'(\overline{\theta}_k)] \tag{B.5}$$

$$(\hat{\theta}_k - \overline{\theta}_k) = \frac{-f'(\overline{\theta}_k)}{2h(\overline{\theta}_k)} \tag{B.6}$$

$$(\hat{\theta}_k - \overline{\theta}_k)^2 = \frac{[f'(\overline{\theta}_k)]^2}{4[h(\overline{\theta}_k)]^2} = \frac{4\{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_y{}^{-1}\boldsymbol{a}(\overline{\theta}_k)]\}^2}{4[h(\overline{\theta}_k)]^2} = \frac{\{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_y{}^{-1}\boldsymbol{a}(\overline{\theta}_k)]\}^2}{[h(\overline{\theta}_k)]^2} \tag{B.7}$$

$$E(\hat{\theta}_k - \overline{\theta}_k)^2 = \frac{E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_y{}^{-1}\boldsymbol{a}(\overline{\theta}_k)]\}^2\right\}}{h^2(\overline{\theta}_k)} = \frac{\Delta}{h^2(\overline{\theta}_k)} = \mathrm{var}(\hat{\theta}_k) \tag{B.8}$$

The second step is to derive an expression for Δ.

$$\begin{aligned}\hat{\boldsymbol{R}}_y{}^{-1} - \boldsymbol{R}_y{}^{-1} &= \hat{\boldsymbol{R}}_y{}^{-1}\boldsymbol{R}_y\boldsymbol{R}_y{}^{-1} - \hat{\boldsymbol{R}}_y{}^{-1}\hat{\boldsymbol{R}}_y\boldsymbol{R}_y{}^{-1} \\ &= \hat{\boldsymbol{R}}_y{}^{-1}(\boldsymbol{R}_y - \hat{\boldsymbol{R}}_y)\boldsymbol{R}_y{}^{-1} \\ &\simeq -\boldsymbol{R}_y{}^{-1}(\hat{\boldsymbol{R}}_y - \boldsymbol{R}_y)\boldsymbol{R}_y{}^{-1}\end{aligned} \tag{B.9}$$

It is known that when the number of samples is large, $f'(\overline{\theta}_k) = 0$. Thus,

$$\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\boldsymbol{R}_y{}^{-1}\boldsymbol{a}(\overline{\theta}_k)] = 0 \tag{B.10}$$

Using this equation,

$$\begin{aligned}\Delta &= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_y{}^{-1}\boldsymbol{a}(\overline{\theta}_k)]\}^2\right\} \\ &= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\hat{\boldsymbol{R}}_y{}^{-1}\boldsymbol{a}(\overline{\theta}_k)] - \underbrace{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)\boldsymbol{R}_y{}^{-1}\boldsymbol{a}(\overline{\theta}_k)]}_{0}\}^2\right\}. \\ &= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)(\hat{\boldsymbol{R}}_y{}^{-1} - \boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\overline{\theta}_k)]\}^2\right\} \\ &= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\overline{\theta}_k)(-\boldsymbol{R}_y{}^{-1}(\hat{\boldsymbol{R}}_y{}^{-1}-)\boldsymbol{R}_y{}^{-1})\boldsymbol{a}(\overline{\theta}_k)]\}^2\right\}\end{aligned} \tag{B.11}$$

Using (B.9) and (B.11),

$$\begin{aligned}
\Delta &= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\} \\
&= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\hat{\boldsymbol{R}}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)] - \underbrace{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)]}_{0}\}^2\right\} \\
&= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)(\hat{\boldsymbol{R}}_y^{-1} - \boldsymbol{R}_y^{-1})\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\} \\
&= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)(-\boldsymbol{R}_y^{-1}(\hat{\boldsymbol{R}}_y - \boldsymbol{R}_y)\boldsymbol{R}_y^{-1})\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\} \\
&= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}(\hat{\boldsymbol{R}}_y - \boldsymbol{R}_y)\boldsymbol{R}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\} \\
&= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}\hat{\boldsymbol{R}}_y\boldsymbol{R}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)] - \underbrace{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}\boldsymbol{R}_y\boldsymbol{R}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)]}_{0}\}^2\right\} \\
&= E\left\{\{\mathrm{Re}[\boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}\hat{\boldsymbol{R}}_y\boldsymbol{R}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)]\}^2\right\} \\
&= E\left\{\{\mathrm{Re}[\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v}]\}^2\right\}
\end{aligned} \tag{B.12}$$

where $\boldsymbol{u}^H = \boldsymbol{d}^H(\bar{\theta}_k)\boldsymbol{R}_y^{-1}$ and $\boldsymbol{v} = \boldsymbol{R}_y^{-1}\boldsymbol{a}(\bar{\theta}_k)$. The equation below can be used to obtain a more explicit expression. If x is any complex-valued scalar, then

$$[\mathrm{Re}(x)]^2 = \frac{1}{2}\mathrm{Re}[|x|^2 + x^2] \tag{B.13}$$

Δ can be expressed as below in this form.

$$\Delta = E\left\{\mathrm{Re}[\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v}]\right\}^2 = \frac{1}{2}\mathrm{Re}\left\{E\left[|\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v}|^2\right] + E\left[(\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v})^2\right]\right\} \tag{B.14}$$

A useful formula for the expectation of the product of four Gaussian random variables is [33]

$$E[x_1x_2x_3x_4] = E[x_1x_2]E[x_3x_4] + E[x_1x_3]E[x_2x_4] + E[x_1x_4]E[x_2x_3] \tag{B.15}$$

$$
\begin{aligned}
E\left[|\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v}|^2\right] &= \frac{1}{N^2}\sum_t^N\sum_s^N E[\underbrace{\boldsymbol{u}^H\boldsymbol{y}(t)}_{x_1}\underbrace{\boldsymbol{y}^H(t)\boldsymbol{v}}_{x_2}\underbrace{\boldsymbol{v}^H\boldsymbol{y}(s)}_{x_3}\underbrace{\boldsymbol{y}^H(s)\boldsymbol{u}}_{x_4}] \\
&= \frac{1}{N^2}\sum_t^N\sum_s^N\Big\{E[\boldsymbol{u}^H\boldsymbol{y}(t)\boldsymbol{y}^H(t)\boldsymbol{v}]E[\boldsymbol{v}^H\boldsymbol{y}(s)\boldsymbol{y}^H(s)\boldsymbol{u}] \\
&+ E[\boldsymbol{u}^H\boldsymbol{y}(t)\boldsymbol{v}^H\boldsymbol{y}(s)]E[\boldsymbol{y}^H(t)\boldsymbol{v}\,\boldsymbol{y}^H(s)\boldsymbol{u}] + E[\boldsymbol{u}^H\boldsymbol{y}(t)\boldsymbol{y}^H(s)\boldsymbol{u}]E[\boldsymbol{y}^H(t)\boldsymbol{v}\,\boldsymbol{v}^H\boldsymbol{y}(s)]\Big\} \\
&= \frac{1}{N^2}\sum_t^N\sum_s^N\Big[(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})(\boldsymbol{v}^H\boldsymbol{R}_y\boldsymbol{u}) + 0 + (\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y\boldsymbol{v})\delta_{t,s}\Big] \\
&= |\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v}|^2 + \frac{1}{N}(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y\boldsymbol{v})
\end{aligned}
\tag{B.16}
$$

$$
\begin{aligned}
E\left[(\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v})^2\right] &= \frac{1}{N^2}\sum_t^N\sum_s^N E[\underbrace{\boldsymbol{u}^H\boldsymbol{y}(t)}_{x_1}\underbrace{\boldsymbol{y}^H(t)\boldsymbol{v}}_{x_2}\underbrace{\boldsymbol{u}^H\boldsymbol{y}(s)}_{x_3}\underbrace{\boldsymbol{y}^H(s)\boldsymbol{v}}_{x_4}] \\
&= \frac{1}{N^2}\sum_t^N\sum_s^N\Big[(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v}) + 0 + (\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})\delta_{t,s}\Big] \\
&= \frac{1}{N^2}\sum_t^N\sum_s^N\Big[(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})^2 + (\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})^2\delta_{t,s}\Big] \\
&= (\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})^2 + \frac{1}{N}(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})^2
\end{aligned}
\tag{B.17}
$$

Inserting (B.16) and (B.17) into (B.14),

$$
\begin{aligned}
\Delta = E\left\{\mathrm{Re}[\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v}]\right\}^2 &= \frac{1}{2}\mathrm{Re}\left\{E\left[|\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v}|^2\right] + E\left[(\boldsymbol{u}^H\hat{\boldsymbol{R}}_y\boldsymbol{v})^2\right]\right\} \\
&= \frac{1}{2}\mathrm{Re}\left\{|\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v}|^2 + (\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})^2 + \frac{1}{N}\left[(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y\boldsymbol{v}) + (\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})^2\right]\right\} \\
&= \left[\mathrm{Re}(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})\right]^2 + \frac{1}{2N}(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{u})(\boldsymbol{v}^H\boldsymbol{R}_y\boldsymbol{v}) + \frac{1}{2N}\mathrm{Re}\left[(\boldsymbol{u}^H\boldsymbol{R}_y\boldsymbol{v})^2\right]
\end{aligned}
\tag{B.18}
$$

Equation (B.10) can also be expressed in terms of *u* and *v*. It can be used in (B.18) for simplification.

$$\mathrm{Re}[\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v}] = 0 \tag{B.19}$$

$$\Delta = \frac{1}{2N}\left\{(\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{u})(\boldsymbol{v}^H \boldsymbol{R}_y \boldsymbol{v}) + \mathrm{Re}\left[(\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v})^2\right]\right\} \tag{B.20}$$

Since $\mathrm{Re}[\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v}] = 0$, there is only imaginary part for $\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v}$ and this results in

$$\mathrm{Re}[(\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v})^2] = -|\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v}|^2 \tag{B.21}$$

$$\Delta = \frac{1}{2N}\left\{(\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{u})(\boldsymbol{v}^H \boldsymbol{R}_y \boldsymbol{v}) - |\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v}|^2\right\} \tag{B.22}$$

Inserting the expression above for Δ into (B.8), the variance term is obtained for the asymptotic MSE.

$$\mathrm{var}(\hat{\theta}_k) = \frac{\Delta}{h^2(\overline{\theta}_k)} = \frac{\frac{1}{2N}\left\{(\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{u})(\boldsymbol{v}^H \boldsymbol{R}_y \boldsymbol{v}) - |\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{v}|^2\right\}}{\left[\mathrm{Re}[\boldsymbol{u}^H \boldsymbol{R}_y \boldsymbol{u} + \boldsymbol{v}^H \boldsymbol{d}'(\overline{\theta}_k)]\right]^2} \tag{B.23}$$

APPENDIX C: DERIVATION OF THE LARGE-SAMPLE GENERAL CRLB FOR STOCHASTIC SIGNALS

Derivation of the CRLB expression in (7.4) is given in this section. The first step in the calculation of the CRLB is the Fisher information matrix.

$$FI(\theta) = N\,E\left[\left(\frac{\partial \ln p(\boldsymbol{y};\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)^2\right] \tag{C.1}$$

The natural logarithm expression can be expressed as below.

$$\frac{\partial \ln p(\boldsymbol{y};\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -\underbrace{\frac{\partial\{y^H[\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})]\boldsymbol{y}\}}{\partial \boldsymbol{\beta}}}_{\boldsymbol{A}} - \underbrace{\frac{\partial \ln\{\pi^m \det[\boldsymbol{R}_y(\boldsymbol{\beta})]\}}{\partial \boldsymbol{\beta}}}_{\boldsymbol{B}} \tag{C.2}$$

In order to clarify the calculation steps, the first term at the right hand side of the equation is named $\boldsymbol{A}$ and the second term is named $\boldsymbol{B}$.

Calculation of $\boldsymbol{A}$ is as follows:

$$\boldsymbol{A} = \frac{\partial\{\boldsymbol{y}^H[\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})]\boldsymbol{y}\}}{\partial \boldsymbol{\beta}} = \boldsymbol{y}^H\left[\frac{\partial \boldsymbol{R}_y^{-1}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\boldsymbol{y} \tag{C.3}$$

The term in the middle can be expressed differently so that the calculation becomes easier. The well-known equation below can be used to derive an expression for the derivative of the inverse function.

$$\boldsymbol{R}_y(\boldsymbol{\beta})\,\boldsymbol{R}_y^{-1}(\boldsymbol{\beta}) = \boldsymbol{I} \tag{C.4}$$

$$\frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})\,\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} = \frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}}\boldsymbol{R}_y^{-1}(\boldsymbol{\beta}) + \boldsymbol{R}_y(\boldsymbol{\beta})\frac{\partial[\boldsymbol{R}_y^{-1}(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} = 0 \tag{C.5}$$

$$\frac{\partial[\boldsymbol{R}_y{}^{-1}(\beta)]}{\partial\beta} = -\boldsymbol{R}_y{}^{-1}(\beta)\frac{\partial[\boldsymbol{R}_y(\beta)]}{\partial\beta}\boldsymbol{R}_y{}^{-1}(\beta) \tag{C.6}$$

Using this expression in (C.3), the calculation of $\boldsymbol{A}$ comes to an end.

$$\boldsymbol{A} = \frac{\partial\{\boldsymbol{y}^H[\boldsymbol{R}_y{}^{-1}(\beta)]\boldsymbol{y}\}}{\partial\beta} = -\boldsymbol{y}^H\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial\boldsymbol{R}_y(\beta)}{\partial\beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{y} \tag{C.7}$$

Definition of the determinant and trace operators are used to derive the expression for $\boldsymbol{B}$.

$$\begin{aligned}\boldsymbol{B} = \frac{\partial\ln\{\pi^m\det[\boldsymbol{R}_y(\beta)]\}}{\partial\beta} &= \frac{1}{\pi^m\det[\boldsymbol{R}_y(\beta)]}\frac{\partial\{\pi^m\det[\boldsymbol{R}_y(\beta)]\}}{\partial\beta} \\ &= \frac{1}{\det[\boldsymbol{R}_y(\beta)]}\frac{\partial\det[\boldsymbol{R}_y(\beta)]}{\partial\beta}\end{aligned} \tag{C.8}$$

Derivative of the determinant can be written in another form including the trace operator.

$$\begin{aligned}\frac{\partial\det[\boldsymbol{R}_y(\beta)]}{\partial\beta} &= \sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial\det[\boldsymbol{R}_y(\beta)]}{\partial[\boldsymbol{R}_y(\beta)]_{ij}}\frac{\partial[\boldsymbol{R}_y(\beta)]_{ij}}{\partial\beta} \\ &= tr\left(\frac{\partial\det[\boldsymbol{R}_y(\beta)]}{\partial[\boldsymbol{R}_y(\beta)]}\frac{\partial[\boldsymbol{R}_y{}^T(\beta)]}{\partial\beta}\right)\end{aligned} \tag{C.9}$$

where $\dfrac{\partial\det[\boldsymbol{R}_y(\beta)]}{\partial[\boldsymbol{R}_y(\beta)]}$ is an mXm matrix with $\{i,j\}$ element $\dfrac{\partial\det[\boldsymbol{R}_y(\beta)]}{\partial[\boldsymbol{R}_y(\beta)]_{ij}}$. In order to use the trace operator, the identity $tr(\boldsymbol{A}\boldsymbol{B}^T) = \sum_{i=1}^{m}\sum_{j=1}^{m}[\boldsymbol{A}]_{ij}[\boldsymbol{B}]_{ij}$ has been used above.

One of the definitions of the determinant can be used to get rid of the derivative. In the equation below $\boldsymbol{M}$ is the mXm cofactor matrix.

$$\det[\boldsymbol{R}_y(\boldsymbol{\beta})] = \sum_{i=1}^{m} [\boldsymbol{R}_y(\boldsymbol{\beta})]_{ij} [\boldsymbol{M}]_{ij} \qquad j = 1,...,m \tag{C.10}$$

Continuing with the derivative,

$$\frac{\partial \det[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial [\boldsymbol{R}_y(\boldsymbol{\beta})]_{ij}} = [\boldsymbol{M}]_{ij} \qquad or \qquad \frac{\partial \det[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial [\boldsymbol{R}_y(\boldsymbol{\beta})]} = [\boldsymbol{M}] \tag{C.11}$$

Using $\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta}) = \dfrac{\boldsymbol{M}^T}{\det[\boldsymbol{R}_y(\boldsymbol{\beta})]}$, the expression below can be written.

$$\frac{\partial \det[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial [\boldsymbol{R}_y(\boldsymbol{\beta})]} = [\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]^T \det[\boldsymbol{R}_y(\boldsymbol{\beta})] \tag{C.12}$$

Returning back to (C.9)

$$\begin{aligned} \frac{\partial \det[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} &= tr\left(\frac{\partial \det[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial [\boldsymbol{R}_y(\boldsymbol{\beta})]} \frac{\partial [\boldsymbol{R}_y{}^T(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} \right) \\ &= tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]^T \det[\boldsymbol{R}_y(\boldsymbol{\beta})] \frac{\partial [\boldsymbol{R}_y{}^T(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} \right) \\ &= tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})] \det[\boldsymbol{R}_y(\boldsymbol{\beta})] \frac{\partial [\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} \right) \end{aligned} \tag{C.13}$$

Finally, $\boldsymbol{B}$ can be written as

$$\begin{aligned} \boldsymbol{B} &= \frac{\partial \ln\{\pi^m \det[\boldsymbol{R}_y(\boldsymbol{\beta})]\}}{\partial \boldsymbol{\beta}} = \frac{1}{\det[\boldsymbol{R}_y(\boldsymbol{\beta})]} \frac{\partial \det[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} \\ &= \frac{1}{\det[\boldsymbol{R}_y(\boldsymbol{\beta})]} tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})] \det[\boldsymbol{R}_y(\boldsymbol{\beta})] \frac{\partial [\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} \right) \\ &= tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})] \frac{\partial [\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}} \right) \end{aligned} \tag{C.14}$$

Now that $\boldsymbol{A}$ and $\boldsymbol{B}$ have been expressed in simpler forms, calculation of the Fisher information matrix can proceed as follows:

$$\frac{\partial \ln p(\boldsymbol{y};\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta}) \left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\boldsymbol{y} - tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]\frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}}\right) \tag{C.15}$$

$$\begin{aligned}
FI(\boldsymbol{\beta}) &= NE\left[\left(\frac{\partial \ln p(\boldsymbol{y};\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)\left(\frac{\partial \ln p(\boldsymbol{y};\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)\right] \\
&= NE\left\{\left[\boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\boldsymbol{y} - tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]\frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}}\right)\right]\right. \\
&\qquad \left.\left[\boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\boldsymbol{y} - tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]\frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}}\right)\right]\right\} \\
&= NE\left\{\boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\boldsymbol{y}\boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\boldsymbol{y}\right\} \\
&\quad -2NE\left\{\boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] \boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\boldsymbol{y}\ tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]\frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}}\right)\right\} \\
&\quad + Ntr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]\frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}}\right)\right\}
\end{aligned} \tag{C.16}$$

A useful lemma will be used to express the first term above in a simpler way. If y is complex Gaussian with zero-mean and correlation function $\boldsymbol{R}_y$, then for $\boldsymbol{C}$ and $\boldsymbol{D}$ Hermitian matrices [29]

$$E[\boldsymbol{y}^H \boldsymbol{C}\boldsymbol{y}\boldsymbol{y}^H \boldsymbol{D}\boldsymbol{y}] = tr(\boldsymbol{C}\boldsymbol{R}_y)tr(\boldsymbol{D}\boldsymbol{R}_y) + tr(\boldsymbol{C}\boldsymbol{R}_y\boldsymbol{D}\boldsymbol{R}_y) \tag{C.17}$$

$$
\begin{aligned}
&E\left\{ \boldsymbol{y}^H \underbrace{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{y}}_{C}\, \boldsymbol{y}^H \underbrace{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{y}}_{D} \right\} \\
&= tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{R}_y(\beta)\right\} tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{R}_y(\beta)\right\} \\
&+ tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{R}_y(\beta)\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{R}_y(\beta)\right\} \\
&= tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\} tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\} + tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\}
\end{aligned}
\tag{C.18}
$$

For the second term of the Fisher information matrix,

$$
\begin{aligned}
&E\left\{\boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial R_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{y}\; tr\left([\boldsymbol{R}_y{}^{-1}(\beta)]\frac{\partial[\boldsymbol{R}_y(\beta)]}{\partial \beta}\right)\right\} \\
&= tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\} E\left\{\boldsymbol{y}^H \boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{y}\right\} \\
&= tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\} E\left\{tr\left[\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{y}\boldsymbol{y}^H\right]\right\} \\
&= tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\} tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\,E[\boldsymbol{y}\boldsymbol{y}^H]\right\} \\
&= tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\} tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\boldsymbol{R}_y{}^{-1}(\beta)\boldsymbol{R}_y(\beta)\right\} \\
&= tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\} tr\left\{\boldsymbol{R}_y{}^{-1}(\beta)\left[\frac{\partial \boldsymbol{R}_y(\beta)}{\partial \beta}\right]\right\}
\end{aligned}
\tag{C.19}
$$

Inserting theses expressions for the first and second term into the Fisher information matrix,

$$
\begin{aligned}
FI(\boldsymbol{\beta}) = & N tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\right\} tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\right\} \\
& + N tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\right\} \\
& - 2N tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\right\} tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\right\} \\
& + N tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right] tr\left([\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})]\frac{\partial[\boldsymbol{R}_y(\boldsymbol{\beta})]}{\partial \boldsymbol{\beta}}\right)\right\} \\
& = N tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\left[\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right]\right\}
\end{aligned}
\tag{C.20}
$$

For the N samples of the received signal, the Fisher information matrix is

$$
FI(\boldsymbol{\beta}) = N\, tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right\} \tag{C.21}
$$

Cramer Rao Lower Bound matrix is then,

$$
CRLB(\boldsymbol{\beta}) = [FI(\boldsymbol{\beta})]^{-1} \tag{C.22}
$$

The expression above is the large-sample "General CRLB" for stochastic signals that have complex m-variate Gaussian distribution with zero-mean and correlation function $\boldsymbol{R}_y$.

If there are several unknown parameters that belong to $\boldsymbol{\beta}$, then $\{i, j\}$ element of the Fisher information matrix can be written as

$$
[FI(\boldsymbol{\beta})]_{ij} = N\, tr\left\{\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_i}\boldsymbol{R}_y{}^{-1}(\boldsymbol{\beta})\frac{\partial \boldsymbol{R}_y(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_j}\right\} \tag{C.23}
$$

where β_i represents the i-th element of the vector β. CRLBs can be determined by the diagonal elements of the inverse Fisher information matrix.

REFERENCES

1. Sayed, A. H., A. Tarighat and N. Khajehnouri, "Network-Based Wireless Location", *IEEE Signal Processing Magazine*, Vol. 22, No. 4, pp. 24–40, July 2005.

2. Weiss, A. J., "On the Accuracy of a Cellular Location System Based on RSS Measurements", *IEEE Transactions on Vehicular Technology*, Vol. 52, No. 6, pp. 1508-1518, November 2003.

3. Klukas, R. and M. Fattouche, "Line-of-Sight Angle of Arrival Estimation in the Outdoor Multipath Environment", *IEEE Transactions on Vehicular Technology*, Vol. 47, No. 1, pp. 342-351, February 1998.

4. Al-Jazzar, S., J. Caffery Jr. and H.R. You, "A Scattering Model Based Approach to NLOS Mitigation in TOA Location Systems", *Proceedings of the 55th IEEE Vehicular Technology Conference*, Birmingham, 6-9 May 2002, Vol. 2, pp. 861-865.

5. Chan, Y. and K. Ho, "A Simple and Efficient Estimator for Hyperbolic Location", *IEEE Transactions on Signal Processing*, Vol. 42, No. 8, pp. 1905-1915, August 1994.

6. Hassanien, A., S. Shahbazpanahi and A. B. Gershman, "A Generalized Capon Estimator for Localization of Multiple Spread Sources", *IEEE Transactions on Signal Processing*, Vol. 52, No. 1, pp. 280-283, January 2004.

7. Besson, O. and P. Stoica, "Decoupled Estimation of DOA and Angular Spread for a Spatially Distributed Source", *IEEE Transactions on Signal Processing*, Vol. 48, No. 7, pp. 1872-1882, July 2000.

8. Astely, D. and B. Ottersten, "The Effects of Local Scattering on Direction of Arrival Estimation with MUSIC", *IEEE Transactions on Signal Processing*, Vol. 47, No. 12, pp. 3220-3234, December 1999.

9. *9-1-1 and Your Wireless Phone,* http://www.nena.org/pages/ContentList.asp?CTID=23.

10. Sun, G., J. Chen, W. Guo and K. J. R. Liu, "Signal Processing Techniques in Network-aided Positioning", *IEEE Signal Processing Magazine*, Vol. 22, No. 4, pp. 12–23, July 2005.

11. Rappaport, T. S., Wireless Communications: Principles and Practice, 2nd ed., Prentice-Hall, New Jersey, 2002.

12. Schiller, J. H., *Mobile Communications*, 2nd ed., Addison-Wesley, New York, 2003.

13. Fulghum, T. and K. Molnar, "The Jakes Fading Model Incorporating Angular Spread for a Disk of Scatterers", *Proceedings of the 48th IEEE Vehicular Technology Conference*, Ottawa, 18-21 May 1998, Vol. 1, pp. 489-493.

14. Buehrer, R. M., S. Arunachalam, K. H. Wu and A. Tonello, "Spatial Channel Model and Measurements for IMT-2000 Systems", *Proceedings of the 53rd IEEE Vehicular Technology Conference,* Rhodes, 6-9 May 2001, Vol. 1, pp. 342-346.

15. Shahbazpanahi, S., S. Valaee and A. B. Gershman, "A Covariance Fitting Approach to Parametric Localization of Multiple Incoherently Distributed Sources", *IEEE Transactions on Signal Processing*, Vol. 52, No. 3, pp. 592-600, March 2004.

16. Stevanović, I., A. Skrivervik and J. R. Mosig, "Smart Antenna Systems for Mobile Communications", Final Report, Laboratoire d'Electromagnétisme et d'Acoustique Ecole Polytechnique Fédérale de Lausanne, January 2003.

17. Nahi, P., C. Parini, L. Du, J. Bingham and L. Cuthbert, "Cell Shaping Using Pattern Synthesis for a Distributed Load Balancing Scheme in Cellular Networks", *Proceedings of the IEEE Wireless Communications and Networking Conference,* New Orleans, 16-20 March 2003, Vol. 1, pp. 93-97.

18. Thompson, J. S., P. Grant and B. Mulgrew, “Smart Antenna Arrays for CDMA Systems”, *IEEE Personal Communications Magazine,* Vol. 3, No. 5, pp. 16-25, October 1996.

19. McPheeters, C., J. Finnigan, J. Bass and E. Rodriguez, *Array Signal Processing: An Introduction,* http://cnx.org/content/m12561/latest/, 2005.

20. Stoica, P. and R. Moses, *Introduction to Spectral Analysis*, Prentice-Hall, New Jersey, 1997.

21. Abdi, A., H. A. Barger and M. Kaveh, “A Parametric Model for the Distribution of the Angle of Arrival and the Associated Correlation Function and Power Spectrum at the Mobile Station”, *IEEE Transactions on Vehicular Technology*, Vol. 51, No. 3, pp. 425-434, May 2002.

22. Abdi, A. and M. Kaveh, “A Versatile Spatio-Temporal Correlation Function for Mobile Fading Channels with Non-Isotropic Scattering”, Proceedings of the 10th IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, 14-16 August 2000, pp. 58–62.

23. Mardia, K. V. and P. E. Jupp, *Directional Statistics*, John Wiley, West Sussex, 2000.

24. A. Abdi and M. Kaveh, “Parametric Modeling and Estimation of the Spatial Characteristics of a Source with Local Scattering”, *IEEE International Conference on Acoustics, Speech, and Signal Processing,* Orlando, 13-17 May 2002, Vol. 3, pp. 2821-2824.

25. Ribeiro, C. B., E. Ollila and V. Koivunen, “Propagation Parameter Estimation in MIMO Systems Using Mixture of Angular Distributions Model”, *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Philadelphia, 18-23 March 2005, Vol. 4, pp. 885–888.

26. Ribeiro, C. B., E. Ollila and V. Koivunen, "Stochastic Maximum-Likelihood Method for MIMO Propagation Parameter Estimation", *IEEE Transactions on Signal Processing*, Vol. 55, No. 1, pp. 46-55, January 2007.

27. Liberti, J. C. and T. S. Rappaport, Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications, Prentice-Hall, New Jersey, 1999.

28. Händel, P., P. Stoica, T. Söderström, "Capon Method for DOA Estimation: Accuracy and Robustness Aspects", *IEEE Winter Workshop on Nonlinear Digital Signal Processing*, Tampere, 7-20 January 1993, pp. P_7.1-P_7.5.

29. Kay, S. M., Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice-Hall, New Jersey, 1993.

30. Paulraj, A. and T. Kailath, "Direction of Arrival Estimation by Eigenstructure Methods with Imperfect Spatial Coherence of Wave Fronts", *Journal of the Acoustical Society of America*, Vol. 83, No. 3, pp. 1034-1040, March 1988.

31. Besson, O. and P. Stoica, "Decoupled Estimation of DOA and Angular Spread for a Spatially Distributed Source", *IEEE Transactions on Signal Processing*, Vol. 48, No. 7, pp. 1872–1882, July 2000.

32. Ghogho, M., O. Besson and A. Swami, "Estimation of Directions of Arrival of Multiple Scattered Sources", *IEEE Transactions on Signal Processing*, Vol. 49, No. 11, pp. 3467–3480, November 2001.

33. Janssen, P. H. M. and P. Stoica, "On the Expectation of the Product of Four Matrix-Valued Gaussian Random Variables", *IEEE Transactions on Automatic Control*, Vol. 33, No. 9, pp. 867-870, September 1988.