TECHNICAL UNIVERSITY OF CRETE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING
TELECOMMUNICATIONS DIVISION

Subspace Tracking for Nested Arrays

by

Emmanouil Fountoulakis

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DIPLOMA OF

ELECTRONIC AND COMPUTER ENGINEERING

July 2016

THESIS COMMITTEE

Associate Professor George N. Karystinos, *Thesis Supervisor*
Associate Professor Aggelos Bletsas
Professor Vassilios Digalakis

Abstract

In radar, sonar, and mobile communications, the estimation of the directions from which multiple signals arrive at a point is called the direction of arrival (DoA) estimation problem and, over the past decades, has been performed often through uniform linear arrays (ULAs) in conjunction with high-resolution subspace-based algorithms. Such techniques, however, have limited capability of the number of directions they can estimate; if the ULA consists of N antenna elements, then high-resolution subspace-based algorithms can estimate the directions of up to $N-1$ signals. To increase this number, a novel structure that consists of N antenna elements and enables the estimation of $O\left(N^2\right)$ signal directions has been developed recently. It lies on specific nonuniform-array structures which are called nested arrays.

In this thesis, we first overviewed the structure and properties of nested arrays and evaluated their performance through computer simulations. We observed that the nested arrays with conventional subspace-based signal-processing algorithms can offer high performance but require high complexity, which makes them impractical for real-time applications. Then, we developed novel subspace tracking techniques for nested arrays that have lower complexity, are proven to converge to the optimal, subspace-based estimator, and are capable of tracking changes in the directions of the arriving signals (for example, when a source is moving with respect to the array receiver).

Acknowledgements

I would like to express my special thanks to my advisor Professor George Karystinos. Without his guidance and cooperation, this thesis would have never been accomplished.

Also, I would like to thank my family for their encouragement, support and love.

Finally, I would like to thank my friends for all the great and memorable moments we had together.

To my family.

Table of Contents

List of Figures

Chapter 1

Introduction

Direction-of-arrival (DoA) estimation is a fundamental problem in signal processing with important applications in radar, sonar, source localization, wireless and mobile communications. Existing DoA estimation techniques may be broadly categorized into spectral estimation methods, likelihood maximization methods, and high resolution subspace-based algorithms. Subspace-based methods enjoy great popularity, mostly due to their favourable trade-off between target-angle resolution ability and computational simplicity in implementation.

With conventional uniform linear arrays (ULAs), we can estimate the angles of arrival of $\mathcal{O}(N)$ source signals by N antennas at the receiver, using high resolution subspace-based algorithms. Algorithms such as the celebrated MUltiple SIgnal Classification (MUSIC) algorithm rely on the r (number of targets) principal components of the recorded snapshots. The r principal components are obtained either by means of singular-value decomposition (SVD) of the sensor-array data matrix or eigen-value decomposition (EVD) of the estimated received-signal autocorrelation matrix.

However, it was recently shown that we can estimate the angles of arrival of $\mathcal{O}(N^2)$ source signals by N antennas at the receiver using nested arrays. Nested arrays [2] constitute a novel structure which is capable to provide a dramatic increase of the degrees of freedom (DoF). Hence, we resolve significantly more sources than the actual number of physical antennas. That is, nested arrays are obtained by combining two or more ULAs with increasing intersensor spacing in order to exploit $\mathcal{O}(N^2)$ DoF from only $\mathcal{O}(N)$ physical antennas. The implementation of high resolution subspace-based algorithms for nested arrays offers a very good performance; on the other hand, the computation of the principal components by SVD requires a high complexity.

In this work, we investigate the implementation of subspace tracking techniques for nested arrays. We develop an algorithm that can estimate the signal subspace avoiding the SVD method. Our approach is based on the definition of a minimization problem and the use of Stochastic Approximation Theory [3]. We prove that the solution of the minimization problem is the subspace of the signal. Hence, high resolution subspace-based algorithms can be implemented avoiding the use of the SVD method. That is, by combining nested arrays and Stochastic Approximation Theory, we develop an efficient subspace-tracking algorithm with low complexity.

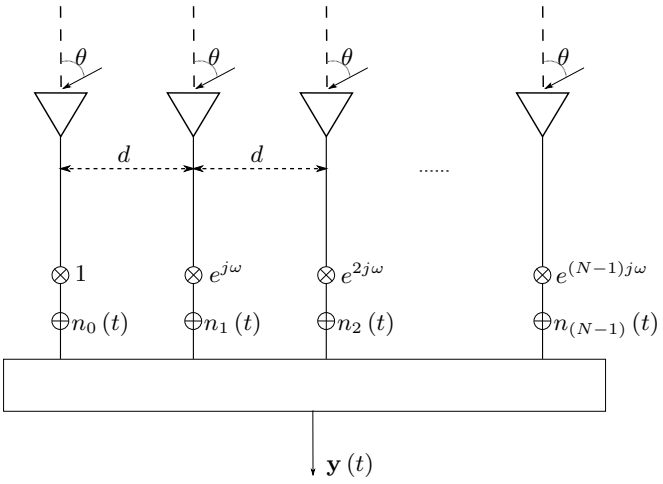

Figure 2.1: Uniform linear array structure. Case about one impinging signal.

Chapter 2

Uniform Linear Arrays

2.1 ULA Structure and Signal Model

In the uniform linear arrays (ULAs), the sensors (antennas) are placed linearly with equal distance. Consider a ULA consisting of N antennas for impinging signal of *one* source. The delay along the array can be defined as a phase shift due to the baseband assumption. Let $s(t)$ a signal, that is transmitted by a source from angle θ. The baseband received signal vector can be written as

$$\mathbf{y}(t) = \mathbf{a}(\theta)\, s(t) + \mathbf{n}(t) \tag{2.1}$$

where $t = 1, \ldots, T$ is the time index, $\mathbf{y}(t) \in \mathbb{C}^N$ is the received vector, $\mathbf{n}(t) \in \mathbb{C}^N$ is the white complex Gaussian noise vector with variance 1, and $\mathbf{a}(\theta) \in \mathbb{C}^N$ is the steering vector whose elements are the phase shifts experienced by the received signal at the elements of the array. If the first element of the arrays is considered as a reference point, the steering vector can be written as

$$\mathbf{a}(\theta) = \left[1, \phi, \ldots, \phi^{N-1}\right]^T$$

where

$$\phi = e^{j\omega}, \omega = j2\pi \frac{d}{\lambda} \sin(\theta)$$

with d, the distance between the antennas. In Figure 2.1, the structure of ULA for one impinging signal from angle θ and the output $\mathbf{y}(t)$ are shown.

We can generalize the signal model considering K signal sources. The signal model for the general case can be written as

$$\mathbf{y}(t) = \sum_{i=1}^{K} \mathbf{a}(\theta_i) s_i(t) + \mathbf{n}(t). \tag{2.2}$$

Relation (2.2) can also be written as

$$\mathbf{y}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{2.3}$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_k)], \in \mathbb{C}^{N\times K}$ is the steering matrix and $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^T \in \mathbb{C}^K$, is the source signal. The structure of the $\mathbf{A}$ steering matrix is shown in (2.4).

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ e^{j\omega_1} & e^{j\omega_2} & \ldots & e^{j\omega_K} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(N-1)\omega_1} & e^{j(N-1)\omega_2} & \ldots & e^{j(N-1)\omega_K} \end{bmatrix}. \tag{2.4}$$

The steering matrix $\mathbf{A}$ is a Vandermonde matrix.

Also, it should be noted that the distance d is defined as $d \leq \frac{\lambda}{2}$, where λ is the wavelength of received signals.

Proof : We know that

$$-\pi \leq \omega \leq \pi \tag{2.5}$$

where

$$\omega = 2\pi \frac{d}{\lambda} \sin(\theta) \tag{2.6}$$

and

$$-1 \leq \sin(\theta) \leq 1, \forall\theta \in [0, 2\pi]. \tag{2.7}$$

From (2.5) and (2.6),

$$-\pi \leq 2\pi \frac{d}{\lambda} \sin(\theta) \leq \pi \Rightarrow -1 \leq 2\frac{d}{\lambda} \sin(\theta) \leq 1 \Rightarrow -\frac{1}{2}\frac{\lambda}{d} \leq \sin(\theta) \leq \frac{1}{2}\frac{\lambda}{d}. \tag{2.8}$$

From (2.7) and (2.8), the distance between the antennas must be defined as $d \leq \frac{\lambda}{2}$.

Chapter 3

High Resolution Subspace-Based Algorithms

3.1 Background Knowledge, Basic Definitions, and Theorems

3.1.1 Vector Spaces

Vector Space Definition

The set $\mathcal{V}$ is called a ***vector space over*** $\mathcal{F}$ when the vector addition and scalar multiplication operations satisfy the following properties [4]

(**A1**) $\mathbf{x}+\mathbf{y} \in \mathcal{V}, \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$. This is called the ***closure property for vector addition***.

(**A2**) $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$, $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$.

(**A3**) $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}, \forall \mathbf{x}, \mathbf{y} \in \mathbf{V}$.

(**A4**) $\exists\ \mathbf{0} \in \mathcal{V}$ such that $\mathbf{x}+\mathbf{0} = \mathbf{x}, \forall \mathbf{x} \in \mathbf{V}$.

(**A5**) $\forall \mathbf{x} \in \mathcal{V}$, $\exists$ a vector $(-\mathbf{x}) \in \mathcal{V}$ such that $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$.

(**M1**) $\alpha\mathbf{x} \in \mathcal{V}, \forall \alpha \in \mathcal{F}$ and $\forall \mathbf{x} \in \mathcal{V}$. This is the ***closure property for scalar multiplication***.

(**M2**) $(\alpha\beta)\,\mathbf{x} = \alpha\,(\beta\mathbf{x})$ $\forall \alpha, \beta \in \mathcal{F}$ and $\forall \mathbf{x} \in \mathcal{V}$.

(**M3**) $\alpha\,(\mathbf{x}+\mathbf{y}) = \alpha\mathbf{x}+\alpha\mathbf{y}$, $\forall \alpha \in \mathcal{F}$ and $\forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$.

(**M4**) $(\alpha+\beta)\,\mathbf{x} = \alpha\mathbf{x}+\beta\mathbf{x}, \forall \alpha, \beta \in \mathcal{F}$ and $\forall \mathbf{x} \in \mathcal{V}$.

(**M5**) $1\mathbf{x} = \mathbf{x}, \forall \mathbf{x} \in \mathcal{V}$.

Subspaces

Les $\mathcal{S} \subseteq \mathcal{F}$. If $\mathcal{S}$ is also a vector space over $\mathcal{F}$ using the addition and scalar multiplication operations, then $\mathcal{S}$ is said to be a subspace of $\mathcal{V}$. It is not necessary to check all 10 of the defining conditions in order to determine if a subset is also a subspace.

A nonempty $\mathcal{S} \subseteq \mathcal{V}$ is a subspace of $\mathcal{V}$ if and only if

(A1) $\mathbf{x} + \mathbf{y} \in \mathcal{S} \Rightarrow \forall \mathbf{x}, \mathbf{y} \in \mathcal{S}$ and

(M1) $\mathbf{x} \in \mathbf{S} \Rightarrow \alpha\mathbf{x} \in \mathcal{S}, \forall \alpha \in \mathcal{F}$.

Range spaces

The ***range of matrix*** $\mathbf{A} \in \Re^{m\times n}$ is defined to be the subspace $\mathcal{R}(\mathbf{A})$ of $\Re^m$ that is generated by the range of $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$. That is,

$$\mathcal{R}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \Re^n\} \subseteq \Re^m. \tag{3.1}$$

Similarly, the $\mathcal{R}(\mathbf{A})$ is the subspace of $\Re^n$ defined by

$$\mathcal{R}\left(\mathbf{A}^T\right) = \left\{\mathbf{A}^T\mathbf{x} \mid \mathbf{x} \in \Re^m\right\} \subseteq \Re^n. \tag{3.2}$$

Null spaces

- For an $m \times n$ matrix $\mathbf{A}$, the set $\mathcal{N}(\mathbf{A}) = \{\mathbf{x}_{n\times 1} \mid \mathbf{A}\mathbf{x} = \mathbf{0}\} \subseteq \Re^n$ is called the ***nullspace*** of $\mathbf{A}$. In other words, $\mathcal{N}(\mathbf{A})$ is simply the set of all solutions to the homogeneous system $\mathbf{A}\mathbf{x} = \mathbf{0}$.
- The set $\mathcal{N}\left(\mathbf{A}^T\right) = \left\{\mathbf{y}_{m\times 1} \mid \mathbf{A}^T\mathbf{y}^T\right\} \subseteq \Re^m$ is called the ***left-hand nullspace*** of $\mathbf{A}$ because $\mathcal{N}\left(\mathbf{A}^T\right)$ is the set of all solutions to the left-hand homogeneous system $\mathbf{y}^T\mathbf{A} = \mathbf{0}$.

3.1.2 Singular-Value Decomposition (SVD)

For each $A \in \Re^{m\times r}$ of rank r, where $r = \min(m, n)$, there are orthogonal matrices $\mathbf{U}_{m\times m}$, $\mathbf{V}_{n\times n}$ and a diagonal matrix $\mathbf{\Sigma}_{m\times m} = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$ such that [4]

$$\mathbf{A} = \mathbf{U}\begin{pmatrix} \mathbf{\Sigma} & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)-r} & \mathbf{0}_{(m-r)\times(n-r)} \end{pmatrix}\mathbf{V}^T, \quad \text{with} \quad \sigma_1 \geq \sigma_1 \geq \ldots \geq \sigma_r \geq 1. \tag{3.3}$$

The σ_i's are called the nonzero ***singular values*** of $\mathbf{A}$. When $r \leq p = \min(m, n)$, $\mathbf{A}$ is said to have $p - r$ additional zero singular values. The columns in $\mathbf{U}$ and $\mathbf{V}$ are called left-hand and right-hand ***singular values*** for $\mathbf{A}$, respectively.
Also there are some properties that be shown below

i. $\mathbf{U}_{:,1:r}$ is an orthonormal basis for $\mathcal{R}(\mathbf{A})$.

ii. $\mathbf{U}_{:,r+1:m}$ is an orthonormal basis for $\mathcal{N}\left(\mathbf{A}^T\right)$.

iii. $\mathbf{V}_{:,1:r}$ is an orthonormal basis for $\mathcal{R}(\mathbf{A})$.

iv. $\mathbf{V}_{:,r+1:n}$ is an orthonormal basis for $\mathcal{N}\left(\mathbf{A}^T\right)$.

Proof.

i. $\mathcal{R}(\mathbf{A}) = \mathcal{R}\left(\mathbf{U}_{:,1:r}\Sigma\mathbf{V}_{:,1:r}^T\right) \subseteq \mathcal{R}(\mathbf{U}_{:,1:r}) = \mathcal{R}(\mathbf{A}\mathbf{V}_{:,1:r}\Sigma^{-1}) \subseteq \mathcal{R}(\mathbf{A}) \Rightarrow \mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{U}_{:,1:r})$.

ii. If $\mathbf{x} \in \mathcal{R}(\mathbf{U}_{:,r+1:m})$, then $\mathbf{x} = \mathbf{U}_{:,r+1:m}\mathbf{y} \Rightarrow \mathbf{A}^T\mathbf{x} = \mathbf{V}_{:,1:r}\Sigma\mathbf{U}_{:,1:r}^T\mathbf{U}_{:,r+1:m}\mathbf{y} = \mathbf{0} \Rightarrow \mathbf{x} \in \mathcal{N}\left(\mathbf{A}^T\right)$.

Hence,

$$\mathcal{R}(\mathbf{U}_{:,r+1:m}) \subseteq \mathcal{N}\left(\mathbf{A}^T\right), \tag{3.4}$$

$$dim\left(\mathcal{R}(\mathbf{U}_{:,r+1:m})\right) = m - r, \tag{3.5}$$

$$dim\left(\mathcal{N}(\mathbf{A})\right) = m - r. \tag{3.6}$$

Therefore, from (3.4)–(3.6), we obtain $\mathcal{R}(\mathbf{U}_{:,r+1:m}) = \mathcal{N}\left(\mathbf{A}^T\right)$.
$(iii), (iv)$: Apply (i) and (ii) to $\mathbf{A}^T$. □

3.1.3 Properties of Eigenvalues and Eigenvectors

$\forall \mathbf{A} \in \Re^{n\times n}$, scalar λ and vectors $\mathbf{x}_{n\times 1} \neq \mathbf{0}$ satisfying $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ are called ***eigen values*** and ***eigen vector*** of $\mathbf{A}$, respectively, and any such pair, $(\lambda, \mathbf{x})$, is called an ***eigenpair*** for $\mathbf{A}$. The set of the *distinct* ***eigenvalues***, denoted by $\sigma(\mathbf{A})$, is called the ***spectrum*** of $\mathbf{A}$ [4].

- $\lambda \in \sigma(\mathbf{A}) \Leftrightarrow \mathbf{A} - \lambda\mathbf{I}$ is signular $\Leftrightarrow det(\mathbf{A} - \lambda\mathbf{I}) = 0$.
- $\{\mathbf{x} \neq \mathbf{0} \mid \mathbf{x} \in N(\mathbf{A} - \lambda\mathbf{I})\}$ is the set of all eigenvectors associated with λ, where $N\{\mathbf{A} - \lambda\mathbf{I}\}$ is called an ***eigenspace*** for $\mathbf{A}$.
- Nonzero row vectors $\mathbf{y}^*$ such that $\mathbf{y}^*(\mathbf{A} - \lambda\mathbf{I}) = 0$ are called ***left-hand eigenvectors*** for $\mathbf{A}$.

3.1.4 Eigen-Value Decomposition (EVD)

$\forall \mathbf{A} \in \Re^{n\times n}$, $\exists$ non singular $\mathbf{Q}$ with $\|\mathbf{Q}_{:,i}\| = 1$, $i = 1, \ldots, n$, and diagonal matrix $\mathbf{\Lambda}$ such that

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} \tag{3.7}$$

where $\mathbf{Q} = [\mathbf{q_1}, \ldots, \mathbf{q_n}]$, and $\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$.

These are the ***eigenvectors*** and the ***eigenvalues*** of matrix $\mathbf{A}$, respectively.

Also it should be noted that it is proved the below property.

If $\mathbf{A} \in \mathbb{C}^{n \times n}$ and $\mathbf{A}^H = \mathbf{A}$ (Hermitian matrix), then

$$\lambda_i \in \Re, i = 1, \ldots, n,$$

and

$$\mathbf{Q}\mathbf{Q}^H = \mathbf{Q}^H\mathbf{Q} = \mathbf{I}_n \left(\mathbf{Q} \text{ is a Unitary matrix}\right).$$

Hence,

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H, \text{ where } \mathbf{U} = \mathbf{Q}.$$

3.2 MUltiple SIgnal Classification (MUSIC) Algorithm [Schmidt '79]

Let the signal model that we defined in Chapter 2, relation (2.3). Requirements of MUSIC $n < m$, and $\mathbf{n}$ white additional noise $\left(E\left\{\mathbf{n}\mathbf{n}^H\right\} = \sigma^2\mathbf{I}, E\left\{\mathbf{n}\right\} = \mathbf{0}\right)$. The autocorrelation matrices of signal $\mathbf{s}$, and output $\mathbf{y}$ are defined as

$$\mathbf{R}_s = E\left\{\mathbf{s}\mathbf{s}^H\right\}, \text{ where } E\left\{\mathbf{s}\right\} = 0,$$

$$\mathbf{R}_y = E\left\{\mathbf{y}\mathbf{y}^H\right\}, \text{ where } E\left\{\mathbf{y}\right\} = 0,$$

respectively. Also, it is proved that $\mathbf{A}\mathbf{R}_s\mathbf{A}^H$ is a positive definite matrix, and $rank\left(\mathbf{A}\mathbf{R}_s\mathbf{A}^H\right) = r$, where $r = \min\left(n, m\right)$. Therefore,

$$\mathbf{A}\mathbf{R}_s\mathbf{A}^H = \mathbf{Q}\mathbf{\Lambda}_0\mathbf{Q}^H = \mathbf{Q}_s\mathbf{\Lambda}_0\mathbf{Q}_s^H. \tag{3.8}$$

Where, $\mathbf{Q} = \left[\underbrace{\mathbf{q}_1, \ldots, \mathbf{q}_r}_{\mathbf{Q}_s (m \times r)} \; \underbrace{\mathbf{q}_{r+1} \ldots \mathbf{q}_m}_{\mathbf{Q}_n \times (m-r)}\right]$ is a unitary matrix, and

$$\mathbf{\Lambda}_0 = \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}, \mathbf{\Lambda}_s = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_r \end{bmatrix}, \; \lambda_1 \geq \ldots \geq \lambda_r > 0.$$

The autocorrelation matrix of output $\mathbf{y}$ is estimated as

$$\mathbf{R}_y = E\left\{\mathbf{y}\mathbf{y}^H\right\} = E\left\{(\mathbf{A}\mathbf{s} + \mathbf{n})(\mathbf{A}\mathbf{s} + \mathbf{n})^H\right\} = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma^2\mathbf{I} \tag{3.9}$$

$$= \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H + \sigma^2\mathbf{Q}\mathbf{Q}^H = \mathbf{Q}\underbrace{\left(\mathbf{\Lambda}_0 + \sigma^2\mathbf{I}\right)}_{\mathbf{\Lambda}}\mathbf{Q}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H. \tag{3.10}$$

Where, $\left(\mathbf{Q}, \begin{bmatrix} \lambda_1 + \sigma^2 & & & & & \\ & \ddots & & & & \\ & & \lambda_r + \sigma^2 & & & \\ & & & \sigma^2 & & \\ & & & & \ddots & \\ & & & & & \sigma^2 \end{bmatrix}\right) = \text{evd}\left(\mathbf{R}_y\right)$.

$\mathcal{R}(\mathbf{Q}_s)$ and $\mathcal{R}(\mathbf{Q}_n)$ are the ***signal subspace*** and the ***noise subspace***, respectively.

Assume distinct $\theta_1, \theta_2, \ldots, \theta_n$ and $m < n$.

Property:

$\theta \in \{\theta_1, \ldots, \theta n\}$ if and only if $\mathbf{a}(\theta) \in \mathcal{R}(\mathbf{A})$, if and only if $\mathbf{Q}_n^H \mathbf{a}(\theta) = 0$.

Proof. Since $\theta_i \neq \theta_j$, $r = rank(\mathbf{A}) = \min(m, n) = n$.
If $\theta = \theta_i$, then $\mathbf{a}(\theta) = \mathbf{A}\mathbf{e}_i \in \mathcal{R}(\mathbf{A})$. If $\mathbf{a}(\theta) \in \mathcal{R}(\mathbf{A})$, then

$$\mathbf{a}(\theta) = \mathbf{A}\mathbf{x} \Rightarrow \mathbf{A}\mathbf{x} - \mathbf{a}(\theta) = \mathbf{0} \Rightarrow \underbrace{[\mathbf{A} \;\; \mathbf{a}(\theta)]}_{m \times (n+1)} \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} = 0 \Rightarrow \mathcal{R}(\mathbf{A}) = \{\mathbf{0}\}$$

$$\Rightarrow rank([\mathbf{A} \;\; \mathbf{a}(\theta)]) < n+1 \Rightarrow \theta = \theta_i, \text{ for some } i = 1, \ldots, n. \tag{3.11}$$

Hence, $\theta \in \{\theta_1, \ldots \theta_n\} \Leftrightarrow \mathbf{a}(\theta) \in \mathcal{R}(\mathbf{A}) \Leftrightarrow \mathbf{a}(\theta) \in \mathcal{N}\left(\mathbf{Q}_n^H\right)$, equivalently,

$$\mathbf{Q}_n^H \mathbf{a}(\theta) = 0. \tag{3.12}$$

□

MUSIC algorithm

If $\hat{\mathbf{R}}_y = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H$, where $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N]$, and the vector $\mathbf{y}_i$, is the output for different samples, and $i = 1, \ldots, N$, then

- $\left(\left[\hat{\mathbf{Q}}_s, \hat{\mathbf{Q}}_n\right], \mathbf{\Sigma}, \mathbf{V}^H\right) = \text{svd}\left(\hat{\mathbf{R}}_y\right)$,
- $\hat{\theta}_1, \ldots, \hat{\theta}_n = \arg\max_\theta \left\{\hat{\mathbf{P}}_{\text{MUSIC}}(\theta) = \frac{1}{\|\mathbf{Q}_n^H \mathbf{a}(\theta)\|^2} = \frac{1}{\mathbf{a}^H(\theta)\mathbf{Q}_n\mathbf{Q}_n^H\mathbf{a}(\theta)}\right\}$,
- $\hat{\theta}_1, \ldots, \hat{\theta}_n = \arg\max_\theta \left\{\hat{\mathbf{P}}_{\text{MUSIC}}(\theta) = \|\mathbf{Q}_s^H \mathbf{a}(\theta)\|^2 = \mathbf{a}^H(\theta)\,\mathbf{Q}_s\mathbf{Q}_s^H\mathbf{a}(\theta)\right\}$.

The $\hat{\mathbf{P}}_{\text{MUSIC}}(\theta)$ is named as ***MUSIC spectrum***, and the vector $\mathbf{a}(\theta)$ is named as the ***scanner vector*** or steering vector, that scan over all possible directions in order to compute the MUSIC spectrum. The MUSIC algorithm can give a performance optical result, where the estimated angles present a local maximum in MUSIC spectrum graph, knowing the number of incoming sources.

3.3 Root-MUSIC Algorithm

The ***Root-MUSIC*** algorithm, as the MUSIC, use the autocorrelation matrix of received signal, and implement SVD to estimate the noise subspace, estimating algebraically the incoming angles. In MUSIC, primary motivation for computing the null space (noise subspace) was the fact that $\mathbf{Q}_n^H \mathbf{a}(\theta) = 0$ (3.12). Also, it is straightforward to check that the 'null spectum'

$$P(a) = \|\mathbf{Q}_n^H \mathbf{A}\|^2 = P_{-m+1}a^{-m+1} + \ldots P_0 + \cdots + P_{m-1}a^{m-1}. \tag{3.13}$$

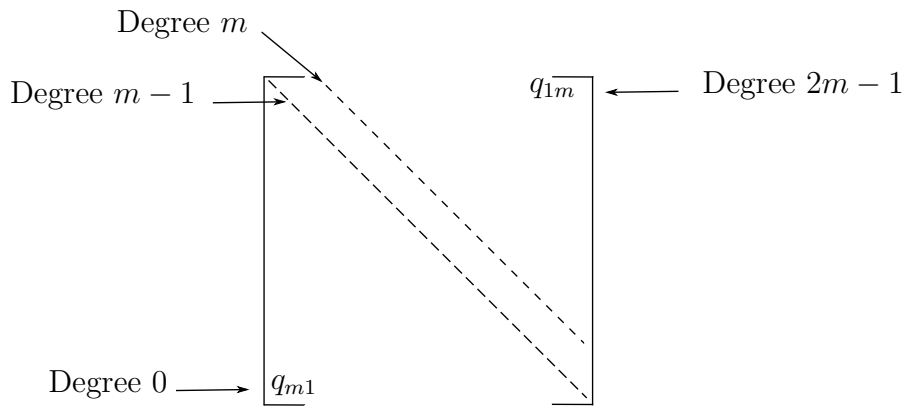

Figure 3.1: Calculation of the coefficients of $P(a)$ based on $\mathbf{Q}_n\mathbf{Q}_n^H$.

where, $a = j2\pi\frac{d}{\lambda}\sin(\theta)$, and the $P_{-m+1}, \ldots, P_{m-1}$ are the coefficients of the polynomial in (3.13) that is occurred using the norm property, has a conjugate symmetry property, i.e. $P_{-i} = P_i^*$. Let

$$r_i = \rho_i a^i \tag{3.14}$$

denote the roots of this polynomial which lie on inside or are the unit circle. Due to the conjugate symmetry of the polynomial $P(a)$, its roots come in pairs where one root is the conjugate reciprocal of the other. We select only one of each of these pairs. The coefficients of the polynomial can be calculated as the sum of the diagonal terms of the matrix $\mathbf{Q}_n\mathbf{Q}_n^H$ [5], according to the diagram in Figure 3.1.

For each root, the incoming agles is found by solving the following equation

$$\theta_k = \sin^{-1}\left[\frac{\lambda}{2\pi d}\arg(r_k)\right], k = 1, 2, \ldots, r. \tag{3.15}$$

Hence, to estimate the incoming angles using root-MUSIC algorithm, it calculates $2m-2$ roots of the polynomial $P(a)$, where m represents the number of antennas, then it finds the roots, which are on or are closest to the unit circle. The roots of the root-MUSIC in the z-plane are sketched in Figure 3.2, for 2 incoming signals, and 4 antennas at the receiver.

The root-MUSIC algorithm eliminates the need for performing a computationally intensive search, replacing it with a root-solving problem. The implementation of the root-MUSIC algorithm is simpler than that of the MUSIC algorithm which requires careful determination of the search step size, and the interpolation of the array manifold. Therefore it easy to be understood that the root-MUSIC algorithm has lower complexity than MUSIC.

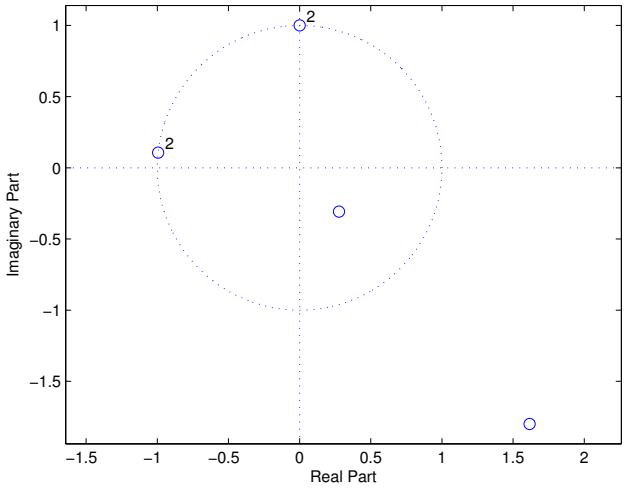

Figure 3.2: Root-MUSIC roots in z-plane

3.4 ESPRIT algorithm

ESPRIT properties

Let $\mathbf{A} \in \mathbb{C}^{m \times n}$, and $\mathbf{B} \in \mathbb{C}^{r \times n}$.

Property 1:

If $rank\left(\mathbf{A}\right) = n$ and $rank\left(\mathbf{B}\right) = n$, then $rank\left(\mathbf{A}\mathbf{B}^H\right) = n$.

Proof.

$$rank\left(\mathbf{A}\mathbf{B}^H\right) = rank\left(\mathbf{B}^H\right) - dim\left(\mathcal{N}\left(\mathbf{A}\right) \cap \mathcal{R}\left(\mathbf{B}^H\right)\right) \tag{3.16}$$

$$= rank\left(\mathbf{B}\right) - dim\left(\{\mathbf{O}\} \cap \mathcal{R}\left(\mathbf{B}\right)\right)$$

$$= rank\left(\mathbf{B}\right) - 0 = n. \tag{3.17}$$

□

Property 2:

If $rank\left(\mathbf{A}\right) = n$, then $rank\left(\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}\right) = n$.

Proof.

$$rank\left(\mathbf{A}_{m \times n}\right) = n \Rightarrow n \leq m. \tag{3.18}$$

$$rank\left(\underbrace{\begin{bmatrix} \mathbf{A}_{m \times n} \\ \mathbf{B}_{r \times n} \end{bmatrix}}_{(m+r) \times n}\right) \leq n. \tag{3.19}$$

$\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_n]$, $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_n]$.
If $rank\left(\begin{bmatrix} \mathbf{A}_{m\times n} \\ \mathbf{B}_{r\times n} \end{bmatrix}\right) < n$, then

$$\mathcal{N}\left(\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}\right) \neq \{\mathbf{0}\} \Rightarrow \exists\ \lambda \neq \mathbf{0} \text{s.t.} \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \lambda = 0 \Rightarrow \begin{bmatrix} \mathbf{A}\lambda \\ \mathbf{B}\lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tag{3.20}$$

$$\Rightarrow \mathbf{A}\lambda = \mathbf{0} \Rightarrow \mathcal{N}(\mathbf{A}) \neq \{\mathbf{0}\} \Rightarrow rank(\mathbf{A}) < n. \text{ (FALSE)} \tag{3.21}$$

□

Consider: $\begin{bmatrix} \mathbf{X_1} \\ \mathbf{X_2} \end{bmatrix} = \begin{bmatrix} \mathbf{AB}^H \\ \mathbf{ADB}^H \end{bmatrix}$
where

- $\mathbf{X}_1\,(m \times r)$, $\mathbf{X}_2\,(m \times r)$ are known,
- $\mathbf{A}\,(m \times n)$, $\mathbf{B}\,(r \times n)$ are unknown with $rank(\mathbf{A}) = rank(\mathbf{B}) = n$,
- $\mathbf{D}\,(n \times n)$ is diagonal and unknown matrix with $rank(D) = n$.

Objective:
Find $\mathbf{D}$ from $\mathbf{X}_1$, $\mathbf{X}_2$.
Solution:

$$\left.\begin{array}{l} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{AD} \end{bmatrix} \mathbf{B}^H \\ rank(\mathbf{A}_{m\times m}) = n \Rightarrow rank\left(\begin{bmatrix} \mathbf{A} \\ \mathbf{AD} \end{bmatrix}\right) = n \end{array}\right\} \Rightarrow rank\left(\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}\right) = n.$$

Compact SVD : $\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} = \mathbf{U}_{2m\times n}\mathbf{\Sigma}_{n\times n}\mathbf{V}^H_{r\times n}$.

$$\Rightarrow \begin{bmatrix} \mathbf{A} \\ \mathbf{AD} \end{bmatrix} \mathbf{B}^H = \mathbf{U\Sigma V}^H \Rightarrow \begin{bmatrix} \mathbf{A} \\ \mathbf{AD} \end{bmatrix} \underbrace{\mathbf{B}^H\mathbf{B}}_{n\times n, rank=n} = \mathbf{U\Sigma V}^H\mathbf{B}$$

$$\Rightarrow \begin{bmatrix} \mathbf{A} \\ \mathbf{AD} \end{bmatrix} = \mathbf{U} \underbrace{\mathbf{\Sigma V}^H\mathbf{B}\left(\mathbf{B}^H\mathbf{B}\right)^{-1}}_{\mathbf{P}_{n\times n}} = \mathbf{UP}.$$

If $rank(\mathbf{P}) < n$, then $rank(\mathbf{UP}) < n = rank\left(\begin{bmatrix} \mathbf{A} \\ \mathbf{AD} \end{bmatrix}\right)$ (FALSE). Hence,
$rank(\mathbf{P}_{n\times n}) = n \Rightarrow \mathbf{P}^{-1}$ exists.

Then,

$$\mathbf{U} = \begin{bmatrix} \mathbf{A} \\ \mathbf{AD} \end{bmatrix} \mathbf{P}^{-1} = \begin{bmatrix} \mathbf{AP}^{-1} \\ \mathbf{ADP}^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix}, \underbrace{\mathbf{U}_1}_{(m\times n)} = \mathbf{AP}^{-1}, \underbrace{\mathbf{U}_2}_{(m\times n)} = \mathbf{ADP}^{-1}.$$

$$\mathbf{R}_1 = \mathbf{U}_1^H\mathbf{U}_1 = \left(\mathbf{P}^{-1}\right)^H \mathbf{A}^H\mathbf{AP}^{-1} \Rightarrow \left(\mathbf{P}^{-1}\right)^H \mathbf{A}^H\mathbf{A} = \mathbf{R}_1\mathbf{P}. \tag{3.22}$$

$$\mathbf{R}_2 = \mathbf{U}_1^H\mathbf{U}_2 = \left(\mathbf{P}^{-1}\right)^H \mathbf{A}^H\mathbf{ADP}^{-1} = \mathbf{R}_1\mathbf{PDP}^{-1}. \tag{3.23}$$

Therefore,

$$\begin{aligned} rank\left(\mathbf{R}_1\right) = rank\left(\mathbf{U}_1^H\mathbf{U}_1\right) = rank\left(\mathbf{U}_1\right) = rank\left(\mathbf{AP}^{-1}\right) = \\ = rank\left(\mathbf{A}\right) = n \Rightarrow \mathbf{R}_1^{-1} \text{ exists.} \end{aligned} \tag{3.24}$$

From (3.22)–(3.24), $\mathbf{R}_1^{-1}\mathbf{R}_2 = \mathbf{PDP}^{-1} \leftarrow$ EVD.

ESPRIT analysis

Let the signal model that we have defined in Chapter 2, the output $\mathbf{y} = \mathbf{As}+\mathbf{n}$, equation (2.3), and the autocorrelation matrix $\mathbf{R}_y = \mathbf{AR}_s\mathbf{A}^H + \sigma^2\mathbf{I} \Rightarrow \mathbf{R}_y - \sigma^2 = \mathbf{AR}_s\mathbf{A}^H$.
Let $\mathbf{X}_1 = [\mathbf{X}]_{1:m-1,:}$, $\mathbf{X}_2 = [\mathbf{X}]_{2:m,:}$, $\mathbf{A}_1 = [\mathbf{A}]_{1:m-1.:}$, $\mathbf{A}_2 = [\mathbf{A}_2]_{2:m,:}$.

Note that: $\mathbf{A}_2 = \mathbf{A}_1\mathbf{D}$, $\mathbf{D} = \begin{bmatrix} e^{j\omega_1} & & \\ & \ddots & \\ & & e^{j\omega_n} \end{bmatrix}_{n\times n}$.

Then,

$$\mathbf{X}_1 = \left[\mathbf{AR}_x\mathbf{A}^H\right]_{1:m-1,:} = [\mathbf{A}]_{1:m-1,:}\,\mathbf{R}_s\mathbf{A}^H = \mathbf{A}_1\mathbf{R}_s\mathbf{A}^H \tag{3.25}$$

and

$$\mathbf{X}_2 = \left[\mathbf{AR}_s\mathbf{A}^H\right]_{2:m,:} = [\mathbf{A}]_{2:m,:}\,\mathbf{R}_s\mathbf{A}^H = \mathbf{A}_2\mathbf{R}_x\mathbf{A}^H = \mathbf{A}_1\mathbf{DR}_s\mathbf{A}^H. \tag{3.26}$$

Set $\mathbf{B} = \mathbf{AR}_s$. Then, $\mathbf{X}_1 = \mathbf{A}_1\mathbf{B}^H$ and $\mathbf{X}_2 = \mathbf{A}_1\mathbf{DB}^H$ with $rank\left(\mathbf{A}_1\right) = n$, $rank\left(\mathbf{D}\right) = n$, $rank\left(\mathbf{B}\right) = rank\left(\mathbf{AR}_s\right) = rank\left(\mathbf{A}\right) = n$.

ESPRIT algorithm

- $[\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}] = \text{compact\_svd}\left(\begin{bmatrix} [\mathbf{R}_y - \sigma^2\mathbf{I}]_{1:m-1,:} \\ [\mathbf{R}_y - \sigma^2\mathbf{I}]_{2:m,:} \end{bmatrix}\right)$,
- $\mathbf{U}_{2(m-1)\times n} = \left[\frac{\mathbf{U_1}}{\mathbf{U}_2}\right]$,
- $\mathbf{R}_1 = \mathbf{U}_1^H\mathbf{U}_1$, $\mathbf{R}_2 = \mathbf{U}_1^H\mathbf{U}_2$,
- $[\mathbf{P}, \mathbf{D}] = \text{evd}\left(\mathbf{R}_1^{-1}\mathbf{R}_2\right)$, from $\mathbf{D}$ we get the $\omega_1, \ldots, \omega_r$,

- $\theta_k = \sin^{-1}\left(\frac{\arg(\omega_k)}{\pi}\right)$, for $k = 1, \ldots, r$.

3.5 Simulations

In this section we provide results for the simulations about high resolution subspace-based algorithms, using ULA.

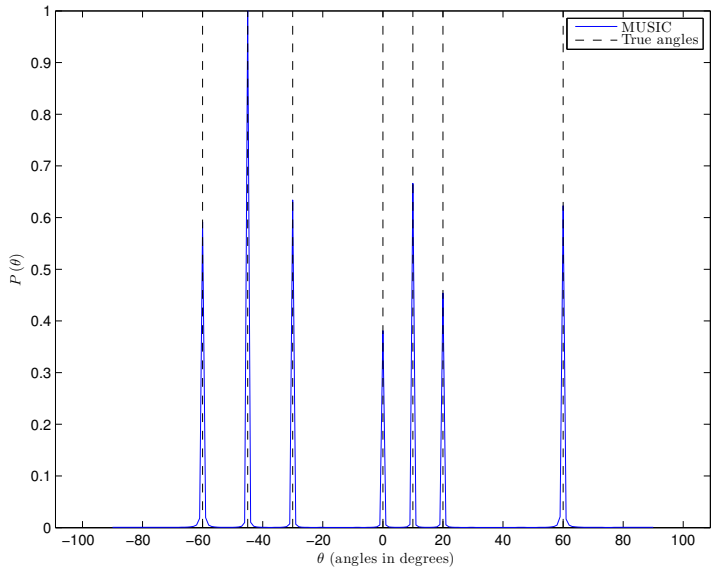

Figure 3.3: Linear arrays, 20 elements, 7 sources, 200 snapshots.

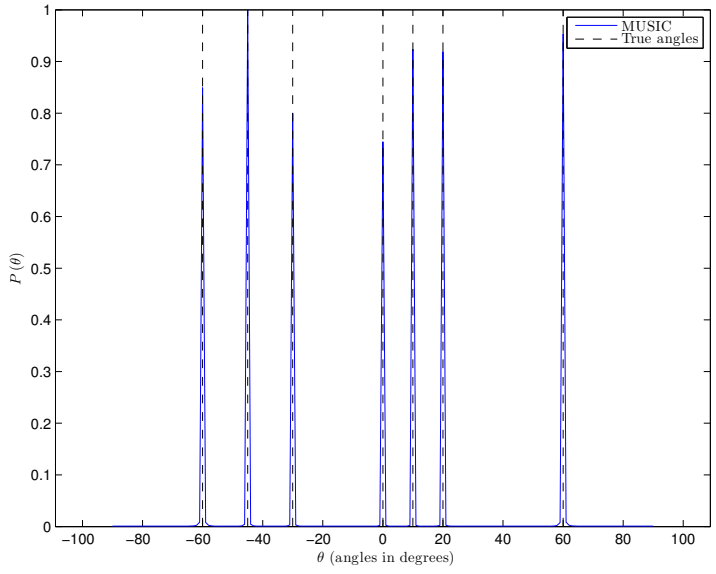

Figure 3.4: Linear arrays, 40 elements, 7 sources, 200 snapshots.

Chapter 4

Subspace Tracking for ULA

4.1 Stochastic approximation theorem

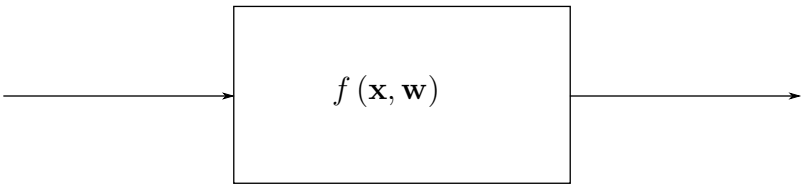

Figure 4.1: Input-Output.

In Figure 4.1, $\mathbf{x} \in \Re^m$ is the input, $\mathbf{y} \in \Re^l$ is the output and $\mathbf{w} \in \Re^d$. Find the parameter $\mathbf{w}$, say $\mathbf{w}^o$, that makes $E\{f(\mathbf{x}, \mathbf{w})\} = 0$.

Solution[Robins-Monro '51]: If $z_1, z_2, \ldots$ is an iid sequence of inputs, $E\{f(z, \ell)\} = 0$ has a unique solution, then

$$\ell_{n+1} = \ell_n - c_n f(z_n, \ell_n), n = 1, 2, \ldots \xrightarrow[m.s.]{\omega.p.1} \ell^o \tag{4.1}$$

where ℓ_0 is arbitrary and $\{c_n\}$ is a monotonically decreasing sequence of positive numbers s.t.

$$\sum_{n=1}^{\infty} c_n = \infty \text{ and } \sum_{n=1}^{\infty} c_n^2 < \infty. \tag{4.2}$$

4.2 Subspace Tracking

Let $\mathbf{x} \in \mathbb{C}^n$ be a complex valued random vector process with the correlation matrix $\mathbf{C} = E\left[\mathbf{x}\mathbf{x}^H\right]$. As [6],we consider the following scalar function

$$\begin{aligned}
J(\mathbf{W}) =& E\|\mathbf{x} - \mathbf{W}\mathbf{W}^H\mathbf{x}\|^2 = E\left\{\mathrm{tr}\left(\mathbf{x} - \mathbf{W}\mathbf{W}^H\mathbf{x}\right)\left(\mathbf{x} - \mathbf{W}\mathbf{W}^H\mathbf{x}\right)^H\right\} \\
=& E\left\{\mathrm{tr}\left(\mathbf{x}\mathbf{x}^H\right) - \mathrm{tr}\left(\mathbf{x}\mathbf{x}^H\mathbf{W}\mathbf{W}^H\right) - \mathrm{tr}\left(\mathbf{W}\mathbf{W}^H\mathbf{x}\mathbf{x}^H\right) + \mathrm{tr}\left(\mathbf{W}\mathbf{W}^H\mathbf{x}\mathbf{x}^H\mathbf{W}\mathbf{W}^H\right)\right\} \\
=& \mathrm{tr}\left(E\left\{\mathbf{x}\mathbf{x}^H\right\}\right) - \mathrm{tr}\left(E\left\{\mathbf{x}\mathbf{x}^H\right\}\mathbf{W}\mathbf{W}^H\right) - \mathrm{tr}\left(E\left\{\mathbf{x}\mathbf{x}^H\right\}\mathbf{W}\mathbf{W}^H\right) \\
& + \mathrm{tr}\left(\mathbf{W}\mathbf{W}^H E\left\{\mathbf{x}\mathbf{x}^H\right\}\mathbf{W}\mathbf{W}^H\right) \\
=& \mathrm{tr}(\mathbf{C}) - 2\mathrm{tr}\left(\mathbf{W}^H\mathbf{C}\mathbf{W}\right) + \mathrm{tr}\left(\mathbf{W}^H\mathbf{C}\mathbf{W}\mathbf{W}^H\mathbf{W}\right)
\end{aligned} \tag{4.3}$$

with a matrix argument $\mathbf{W} \in \mathbb{C}^{n\times r}$ $(r < n)$. Without loss of generality , we assume $\mathbf{W}$ to have full rank r. Otherwise, of the rank of $\mathbf{W}$ is $\tilde{r} < r$, $\mathbf{W}$ in (4.3) can always be replaced by a full rank $n \times \tilde{r}$ matrix satisfying $\tilde{\mathbf{W}}\tilde{\mathbf{W}}^H = \mathbf{W}\mathbf{W}^H$.

We want to know

- Is there a global minimum of $J(\mathbf{W})$?
- What is the relation between this minimum and the signal subspace of $\mathbf{C}$?
- Are there any other local minima of $J(\mathbf{W})$?

These questions are answered by the following theorems.

Theorem 1. $\mathbf{W}$ is a stationary point of $J(\mathbf{W})$ if and only if $\mathbf{W} = \mathbf{Q}_r\mathbf{U}$ where $\mathbf{Q}_r$ contains any r distinct eigenvectors of $\mathbf{C}$ and $\mathbf{U} \in \mathbb{C}^{r\times r}$ is an arbitrary unitary matrix.

Proof. First we calculate the gradient of $J(\mathbf{W})$ for real parameters. Let $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$ and $\nabla = [\nabla_1, \ldots, \nabla_r]$. ∇_i is the gradient operator with respect to $\mathbf{w}_r$. After some calculations, we get

$$\begin{aligned}\frac{1}{2}\nabla_i J &= \left[-2\mathbf{C} + \mathbf{C}\mathbf{W}\mathbf{W}^T + \mathbf{W}\mathbf{W}^H\mathbf{C}\right]\mathbf{w}_i,\\ \frac{1}{2}\nabla J &= \left[-2\mathbf{C} + \mathbf{C}\mathbf{W}\mathbf{W}^T + \mathbf{W}\mathbf{W}^T\mathbf{C}\right]\mathbf{W}.\end{aligned} \tag{4.4}$$

For complex valued data, we define $\nabla_{R,i}$ and $\nabla_{I,i}$ to be the gradient operator with respect to the real and imaginary part of $\mathbf{w}_i$. The complex gradient operator is defined as $\nabla_i = \frac{1}{2}\left[\nabla_{R,i} + j\nabla_{I,i}\right]$. After some calculations, we obtain from 4.4

$$\nabla J = \left[-2\mathbf{C} + \mathbf{C}\mathbf{W}\mathbf{W}^H + \mathbf{W}\mathbf{W}^H\mathbf{C}\right]\mathbf{W}. \tag{4.5}$$

Let $\mathbf{W} = \mathbf{Q}_r\mathbf{U}$, $\mathbf{Q}_r$ contains any r distinct eigenvectors of $\mathbf{C}$, and $\mathbf{U}$ is a orthogonal arbitrary matrix.

$$\begin{aligned}\nabla J &= \left[-2\mathbf{C} + \mathbf{C}\mathbf{W}\mathbf{W}^H + \mathbf{W}\mathbf{W}^H\mathbf{C}\right]\mathbf{W}\\ &= -2\mathbf{C}\mathbf{Q}_r\mathbf{U} + \mathbf{C}\mathbf{Q}_r\underbrace{\mathbf{U}\mathbf{U}^H}_{\mathbf{I}}\ \underbrace{\mathbf{Q}_r^H\mathbf{Q}_r}_{\mathbf{I}}\mathbf{U} + \mathbf{Q}_r\underbrace{\mathbf{U}\mathbf{U}^H}_{\mathbf{I}}\mathbf{Q}_r\mathbf{U}\\ &= -\mathbf{C}\mathbf{Q}_r\mathbf{U} + \mathbf{Q}\mathbf{Q}_r^H\mathbf{C}\mathbf{Q}_r\mathbf{U},\ \ \left(\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H = \mathrm{evd}\left(\mathbf{C}\right)\right)\\ &= -\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H\mathbf{Q}_r\mathbf{U} + \mathbf{Q}_r\mathbf{Q}_r^H\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H\mathbf{Q}_r\mathbf{U}\\ &= -\mathbf{Q}\mathbf{\Lambda}\begin{bmatrix}\mathbf{I}_{r\times r}\\ \mathbf{O}\end{bmatrix}\mathbf{U} + \mathbf{Q}_r\begin{bmatrix}\mathbf{I}_{r\times r} & \mathbf{O}\end{bmatrix}\mathbf{\Lambda}\begin{bmatrix}\mathbf{\Lambda}_{r\times r}\\ \mathbf{O}\end{bmatrix}\mathbf{U}\\ &= -\mathbf{Q}\begin{bmatrix}\mathbf{\Lambda}_s\\ \mathbf{0}\end{bmatrix}\mathbf{U} + \mathbf{Q}_r\begin{bmatrix}\mathbf{I}_{r\times r} & \mathbf{O}\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda}_s\\ \mathbf{0}\end{bmatrix}\mathbf{U}\\ &= -\mathbf{Q}\begin{bmatrix}\mathbf{\Lambda}_s\\ \mathbf{0}\end{bmatrix}\mathbf{U} + \mathbf{Q}_r\mathbf{\Lambda}_s\mathbf{U} = -\mathbf{Q}_s\mathbf{\Lambda}_s\mathbf{U} + \mathbf{Q}_s\mathbf{\Lambda}_s\mathbf{U} = 0.\end{aligned}$$

Therefore, $\mathbf{W} = \mathbf{Q}_r\mathbf{U}$ is solution of (4.5). □

Theorem 2. All stationary points of $J(\mathbf{W})$ are saddle points except when $\mathbf{U}_r$ contains the r dominant eigenvectors of $\mathbf{C}$.

Proof. See proof of Theorem 2 in [6]. □

Since (4.3) describes an unconstrained cost function to be minimized, and its gradient has solution the subspace signal of $\mathbf{C}$, we can apply the stochastic approximation theorem which we refer in Section 4.1.

Therefore, the subspace update can be written as

$$\mathbf{W}_n = \mathbf{W}_{n-1} - \mu\left[-2\mathbf{C}_n + \mathbf{C}_n\mathbf{W}_{n-1}\mathbf{W}_{n-1}^H + \mathbf{W}_{n-1}\mathbf{W}_{n-1}^H\mathbf{C}_n\right]\mathbf{W}_{n-1} \tag{4.6}$$

where $\mu > 0$ is a step size to be suitably chosen and $\mathbf{C}_n$ is an estimation for the correlation matrix $\mathbf{C}$ at the nth samples that is defined as $\mathbf{C}_n = \mathbf{x}_n\mathbf{x}_n^H$, as used in the least-mean-square (LMS) for adaptive filtering. The resulting subspace update is given by

$$\mathbf{y}_n = \mathbf{W}_{n-1}^H\mathbf{x}_n, \tag{4.7}$$

$$\mathbf{W}_n = \mathbf{W}_{n-1} + \mu\left[2\mathbf{x}_n\mathbf{y}_n^H - \mathbf{x}_n\mathbf{y}_n^H \times \mathbf{W}_{n-1}^H\mathbf{W}_{n-1} - \mathbf{W}_{n-1}\mathbf{y}_n\mathbf{y}_n^H\right]. \tag{4.8}$$

We observe that this algorithm has a computational complexity $\mathcal{O}(Nr)$.

Also, we can write the above algorithm, approximating $\mathbf{W}_{n-1}^H\mathbf{W}_{n-1} = \mathbf{I}$, as follows

$$\mathbf{W}_n = \mathbf{W}_{n-1} + \mu\left[\mathbf{x}_n - \mathbf{W}_{n-1}\mathbf{y}_n\right]\mathbf{y}_n^H. \tag{4.9}$$

Replacing the expectation in (4.3) with the exponentially weighted sum yields

$$\begin{aligned} J(\mathbf{W}_n) &= \sum_{i=1}^{n} b^{n-i}\|\mathbf{x}_i - \mathbf{W}_n\mathbf{W}_n^H\mathbf{x}_i\|^2 \\ &= \mathrm{tr}\left[\mathbf{C}_n\right] - 2\mathrm{tr}\left[\mathbf{W}_n^H\mathbf{C}_n\mathbf{W}_n\right] + \mathrm{tr}\left[\mathbf{W}_n^H\mathbf{C}_n\mathbf{W}_n\mathbf{W}_n^H\mathbf{W}_n\right]. \end{aligned} \tag{4.10}$$

The use of the forgetting factor $0 < b \leq 1$ is intended to ensure that data in the distant past are downweighted in order to afford the tracking capability when the system operates in a nonstationary environment. Relation (4.10) can be obtained from (4.3) if we use $\mathbf{W}_n$ instead of $\mathbf{W}$ and

$$\mathbf{C}_n = \sum_{i=1}^{n} b^{n-i}\mathbf{x}_i\mathbf{x}_i^H = b\mathbf{C}_{n-1} + \mathbf{x}_n\mathbf{x}_n^H \tag{4.11}$$

instead of $\mathbf{C}$.

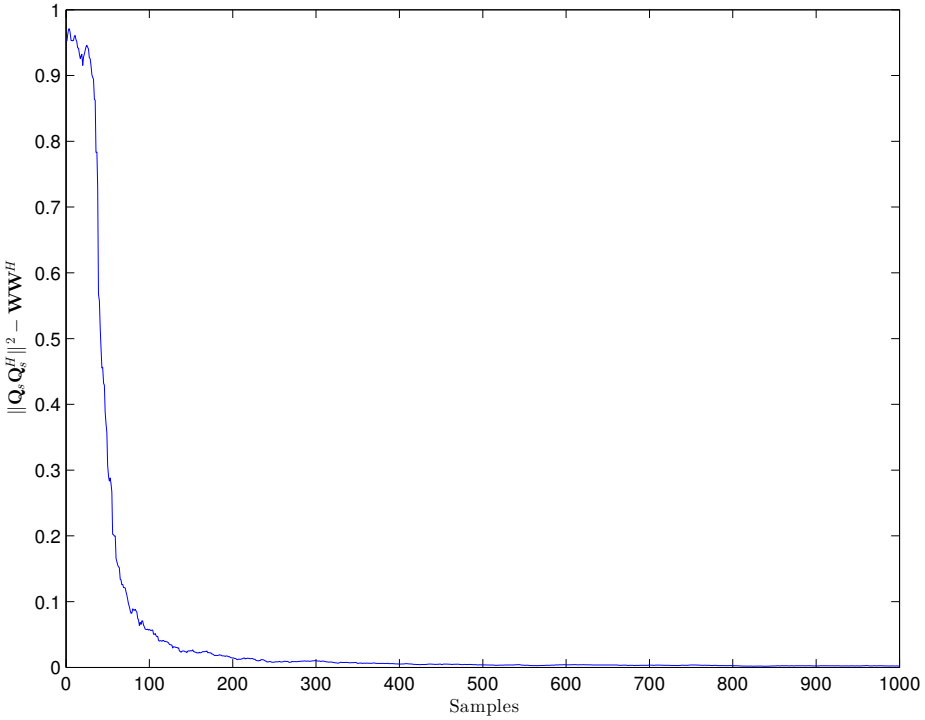

Figure 4.2: Stochastic Approximation (decreasing μ), 10 elements, 6 sources.

4.3 Simulations

In this section, we provide some results from various simulations for Subspace Tracking in ULA. Initially, we show how $\mathbf{W}\mathbf{W}^H$ converges to $\mathbf{Q}_s\mathbf{Q}_s^H$, where $\mathbf{Q}_s$ is the subspace of the received signal. The Stochastic Approximation Theorem requires decreasing μ. However, stable μ is used for real applications.In Figure 4.2, the decreasing μ case is showed. We observe that by using the Subspace Tracking algorithm (iteration (4.6)) we can track the subspace of the received signal with only a small number of samples. In Figure 4.3 we see the resulting Mean Square Error (MSE) by applying Subspace Tracking and through the use of decreasing μ, namely the Stochastic Approximation Theorem. We consider 6 incoming signals impinging on the array from direction of $\{-45^o, -30^o, 0^o, 10^o, 20^o, 45^o\}$. We observe that if we use stable μ, the algorithm is less efficient but this is true only for stationary environments. We choose a suitable stable μ intending to track the signal subspace for non-stationary environments. Firstly, we evaluate the performance of the implementation of iteration (4.6) using the forgetting factor b. The performance can be shown in Figure 4.6, for different values of b in a stationary environment. The case of $b = 0$ implies Stochastic Approximation. As expected, the closer b is to 1 the better the performance that can be achieved because we have better statistics estimation of the received signal. Also, the case about a non stationary environment is interesting to be studied. In Figure 4.7 the position of the target changes in the $400-th$ sample, from 20^o to 25^o. We can observe that the performance of LMS for $b = 0.97$ is better than both the performance of Stochastic Approximation ($b = 0$) and LMS for $b = 1$.

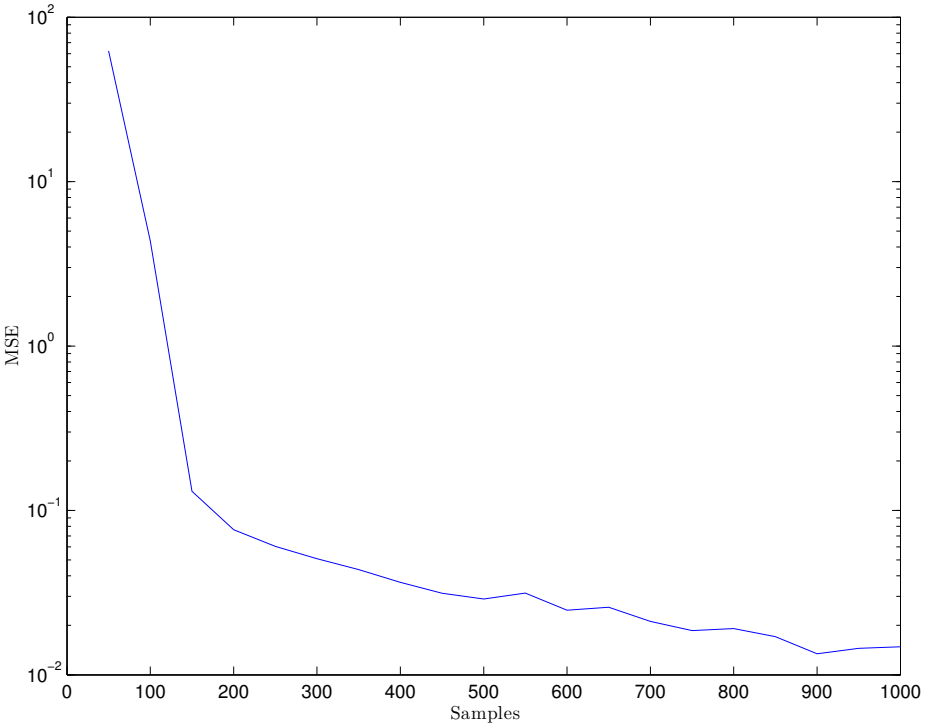

Figure 4.3: Stochastic Approximation (decreasing μ), 10 elements, 6 sources.

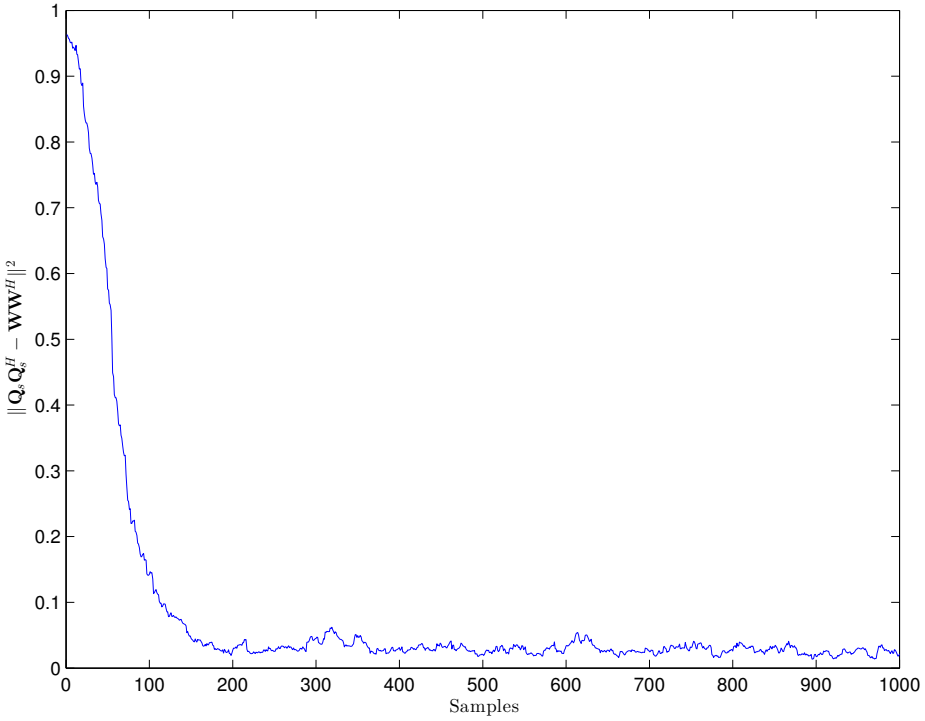

Figure 4.4: Subspace Tracking (stable μ), 10 elements, 6 sources.

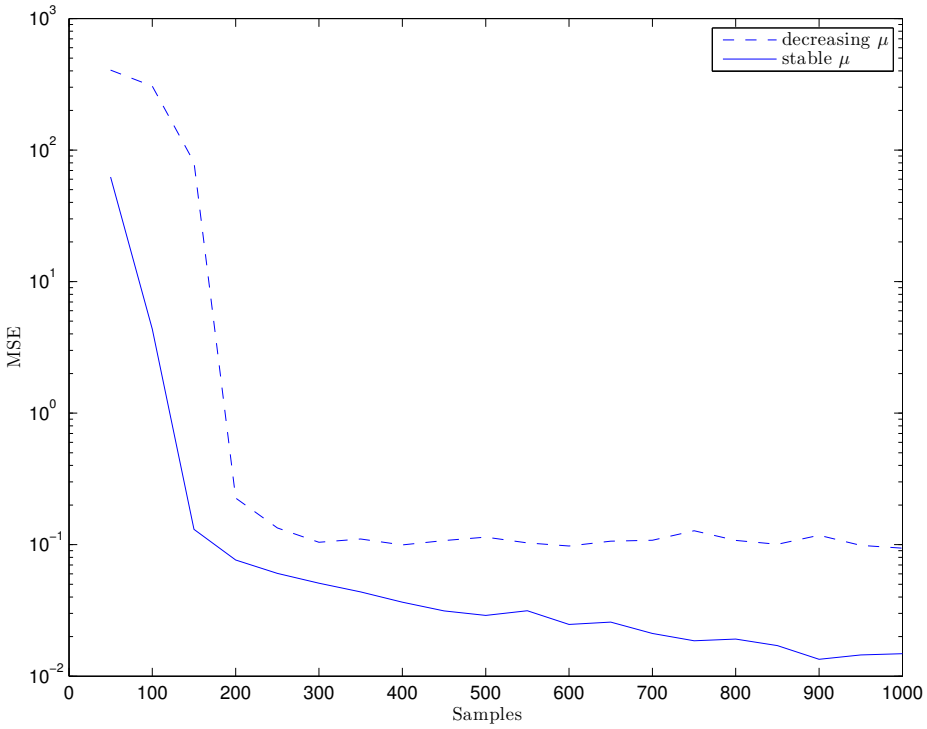

Figure 4.5: Linear Arrays (Stochastic Approximation), stable μ vs decreasing μ, 10 elements, 6 sources.

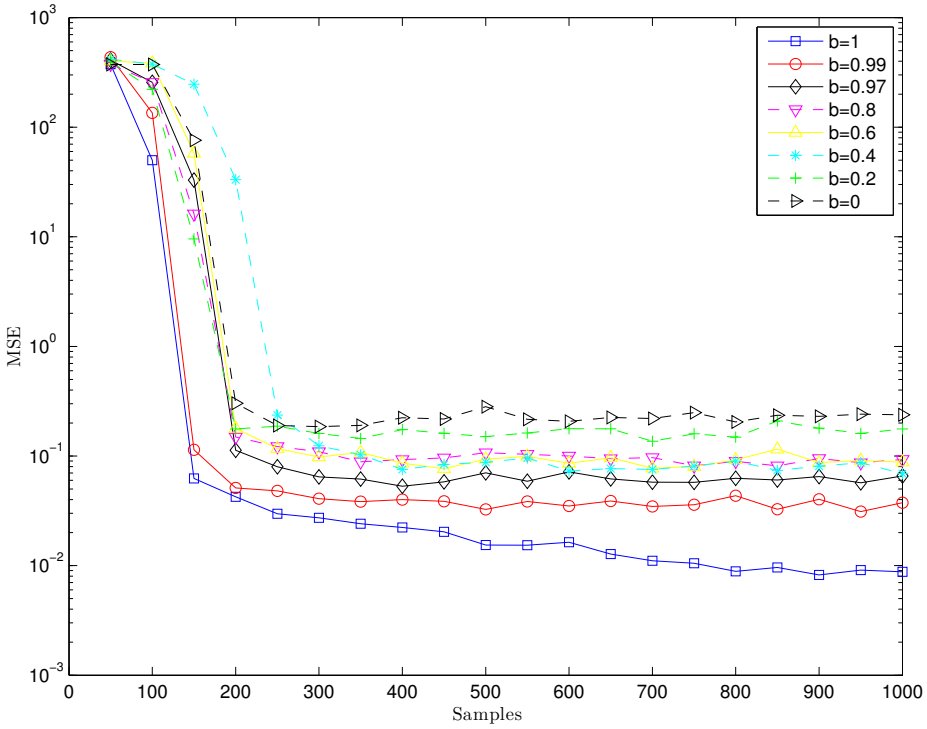

Figure 4.6: Performance of LMS, for different values of b μ, 10 elements, 6 sources, stationary environment.

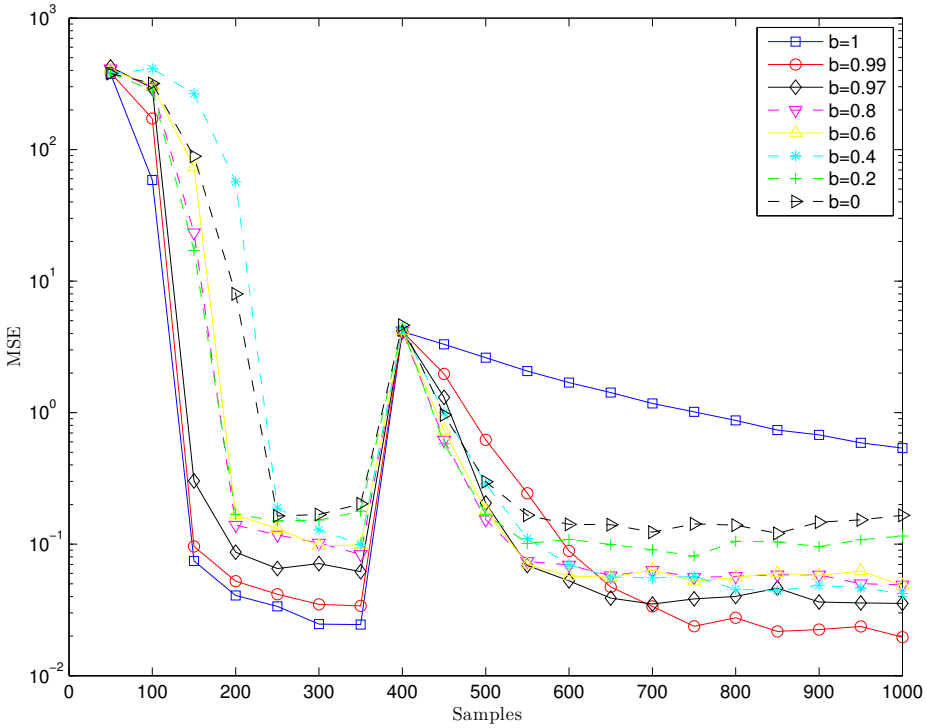

Figure 4.7: Performance of LMS, for different of b μ, 10 elements, 6 sources, non-stationary environment.

Chapter 5

Nested Arrays

5.1 Definitions and Signal Model Based on the Difference Co-array

5.1.1 Signal Model

Consider a N element non-uniform linear antenna array. Let $\mathbf{a}(\theta) \in \mathbb{C}^{N\times 1}$ be the corresponding to the θ direction. Let assume D narrowband sources impinging on this array from directions $\{\theta_i, i = 1, 2, \ldots, D\}$ with powers $\{\sigma^2, i = 1, 2, \ldots, D\}$, respectively. Hence the received signal vector is

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{5.1}$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_D)]$ denotes the steering matrix and $\mathbf{s}(t)$ denotes the source signal vector. We consider that $\mathbf{n}(t)$ is a white Gaussian noise, uncorrelated with the sources. Also we also assume that the sources to be temporally uncorrelated. Therefore, the autocorrelation matrix of $\mathbf{s}(t)$ is diagonal. Then,

$$\begin{aligned}\mathbf{R}_x &= E\left\{\mathbf{x}\mathbf{x}^H\right\} = E\left\{(\mathbf{A}\mathbf{s}+\mathbf{n})(\mathbf{A}\mathbf{s}+\mathbf{n})^H\right\} \\ &= E\left\{\mathbf{A}\mathbf{s}\mathbf{s}^H\mathbf{A}^H + \mathbf{A}\mathbf{s}\mathbf{n}^H + \mathbf{n}\mathbf{s}^H\mathbf{A}^H + \mathbf{n}\mathbf{n}^H\right\} \\ &= \mathbf{A}E\left\{\mathbf{s}\mathbf{s}^H\right\}\mathbf{A}^H + \mathbf{A}\mathbf{s}E\left\{\mathbf{n}^H\right\} + E\{\mathbf{n}\}\mathbf{s}^H\mathbf{A}^H + E\left\{\mathbf{n}\mathbf{n}^H\right\} = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma^2\mathbf{I}.\end{aligned} \tag{5.2}$$

Vectorizing the autocorrelation matrix of $\mathbf{x}$ according to the [7], we have

$$\begin{aligned}\mathbf{z} = \mathrm{vec}(\mathbf{R}_{xx}) &= \mathrm{vec}\left[\sum_{i=1}^{D}\sigma_i^2\left(\mathbf{a}(\theta_i)\mathbf{a}^H(\theta_i)\right)\right] + \sigma_n^2\mathbf{1}_n \\ &= (\mathbf{A}^*\odot\mathbf{A})\mathbf{p} + \sigma_n^2\mathbf{1}_n.\end{aligned} \tag{5.3}$$

Where $\mathbf{p} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_D^2]^T$ and $\mathbf{1}_n = \left[\mathbf{e}_1^T, \mathbf{e}_2^T, \ldots, \mathbf{e}_N^T\right]^T$ with $\mathbf{e}_i$ being a column vector of all zeros except ith position. Also $\odot$ denotes the Khatri-Rao (KR) product and can be defined as follows Let $\mathbf{A} \in \mathbb{C}^{n\times k}$ and $\mathbf{B} \in \mathbb{C}^{m\times k}$, where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k]$, and $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k]$, then

$$\mathbf{A}\odot\mathbf{B} = [\mathbf{a}_1\otimes\mathbf{b}_1, \mathbf{a}_2\otimes\mathbf{b}_2, \ldots, \mathbf{a}_k\otimes\mathbf{b}_k] \tag{5.4}$$

where $\otimes$ denotes the Kronecker product and be defined as follows

$$\mathbf{a} \otimes \mathbf{b} = \begin{bmatrix} \mathbf{a}_1\mathbf{b} \\ \mathbf{a}_2\mathbf{b} \\ \vdots \\ \mathbf{a}_n\mathbf{b} \end{bmatrix} = \text{vec}\left(\mathbf{b}\mathbf{a}^T\right). \tag{5.5}$$

5.1.2 Difference Co-Array perspective

- Consider an array of N sensors, with d_i denoting the position of the ith sensor. Define the difference co-array as

$$(d_1-d_1, d_1-d_2, \ldots, d_1-d_N, d_2-d_1, \ldots, d_N-d_N). \tag{5.6}$$

- Form the set D_u from the distinct elements of the co-array.
- The cardinality of D_u is $(N^2-2)/2+N$.

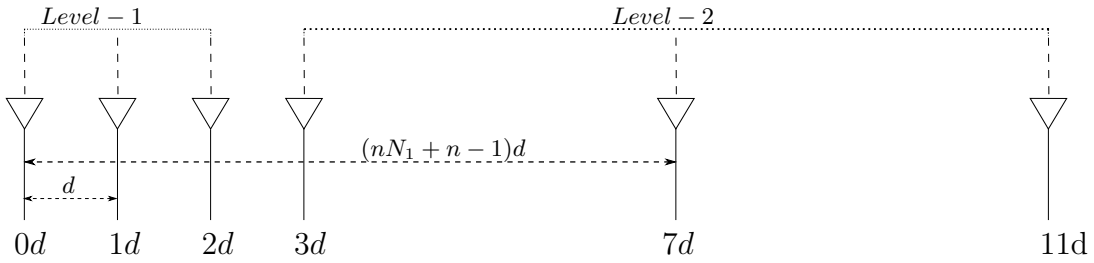

Figure 5.1: Nested arrays with 3 sensors in each level.

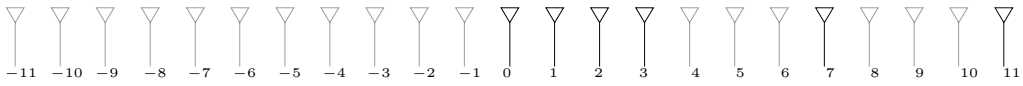

Figure 5.2: element positions in difference co-array.

5.2 Two Level Nested Array

A two-level nested array is a concatenation of a two ULAs (which we refer in chapter 2): inner and outer where the inner ULA has N_1 elements with spacing d_1 and the outer ULA has N_2 elements with spacing d_2 such that $d_2 = (N_1 + 1)\, d_1$. The locations of the sensors are given by the union of the follows sets

$$S_{inner} = \{nd_1,\ n = 1, 2, \ldots N_1\}\,.$$

$$S_{outer} = \{(nN_1 + n - 1)\, d_1,\ n = 1, 2, \ldots N_2\}\,.$$

The Figure 5.1 illustrates the nested arrays for 3 sensors in each level. Also the Figure 5.2 shows the difference co-array of this nested array. There are an important observation. It is a filled ULA with $2N_2\,(N_1 + 1) - 1$, elements whose positions are given by the set S_{sa} defined as

$$S_{sa} = \{nd_1, n = -M, \ldots, M,\ M = N_2\,(N_1 + 1) - 1\}\,. \tag{5.7}$$

The case above, is an example for a specific number of elements. We must define a equation that can give the optimal distribution of sensors in the two levels in order to find N_1, N_2 that maximize the degrees of freedom. Therefore, we must maximize the follows

$$N_2\,(N_1 + 1)\,. \tag{5.8}$$

Relation (5.8) is maximized when

$$N_1 = N_2 = \frac{N}{2},\ \text{for } N \text{ even} \tag{5.9}$$

and

$$N_1 = \frac{N+1}{2}, N_2 = \frac{N-1}{2},\ \text{for } N \text{ odd.} \tag{5.10}$$

Proof. Let $x = N_2$ and $y = (N_1 + 1)$, with constraint $N_1 + N_2 = N$.
Objective: $\max\limits_{x+y=N+1}(xy) = \max\limits_{x+y=N+1}(x(-x+N)) = -x^2 + xN + x = f(x)$.
Then we get the

$$f'(x) = 0 \Rightarrow -2x + N + 1 = 0 \Rightarrow x = \frac{N+1}{2} = N2.$$

Therefore,

$$N_1 = N - N_2 = N - \frac{N+1}{2} = \frac{N-1}{2}$$

$$\begin{aligned} &\text{for } x = \frac{N}{2}, \frac{N}{2}\left(-\frac{N}{2} + N + 1\right) = \frac{N}{2}\left(\frac{N}{2} + 1\right), \\ &\text{for } x = \frac{N}{2} + 1, \left(\frac{N}{2} + 1\right)\left(-\frac{N}{2} - 1 + N + 1\right) = \left(\frac{N}{2} + 1\right)\frac{N}{2}. \end{aligned} \tag{5.11}$$

□

Therefore,

N	optimal N_1, N_2	DOF
even	$N_1 = N_2 = \frac{1}{2}N$	$\frac{N^2-2}{2} + N$
odd	$N_1 = \frac{N-1}{2}, N_2 = \frac{N+1}{2}$	$\frac{N^2-2}{2} + N$

5.3 Spatial Smoothing

In this section, we apply spatial smoothing to exploit the increased DoFs offered by the co-array. Note that, we consider a two-level nested array with N sensors, and $\frac{N}{2}$ sensors in each level.

We remove the repeated rows from $\mathbf{A}^* \odot \mathbf{A}$ and also sort them so that the ith row corresponds to the sensor location $\left(-\frac{N^2}{4} - \frac{N}{2} + i\right) d$ in the difference co-array of the two-level nested array, giving a new vector

$$\mathbf{z_1} = \mathbf{A}_1\mathbf{p} + \sigma^2\bar{\mathbf{e}}_i \tag{5.12}$$

where $\bar{\mathbf{e}}_i \in \Re^{\left(\left(N^2-2\right)/2+N\right)\times 1}$ is a vector of all zeros except a 1 at the $(N^2/4 + N/2)$th position. The difference co-array of this two-level nested array has sensors located for $(-N^2/4 - N/2 + 1)\, d$ to $(N^2/4 + N/2 - 1)\, d$. We divide this co-array into $L = N^2/4 + N/2$ overlapping subarrays, each with $N^2/4 + N/2$ elements, where the ith subarray has sensors located at

$$\left\{(-i + 1 + n)\, d, n = 0, 1, \ldots, \frac{N^2}{4} + \frac{N}{2} - 1\right\}. \tag{5.13}$$

The *ith* subarray corresponds to the $(N^2/4 + N/2 - i + 1)$th to $((N^2 - 2)/2 + N - 1)$th rows

of $\mathbf{z}_1$, denoted as

$$\mathbf{z_{1}}_i = \mathbf{A_{1}}_i\mathbf{p} + \sigma^2\bar{\mathbf{e}}_i. \tag{5.14}$$

We can check that

$$\mathbf{z_{1}}_i = \mathbf{A_{11}}\mathbf{\Phi}^{i-1}\mathbf{p} + \sigma_n^2\bar{\mathbf{e}} \tag{5.15}$$

where

$$\mathbf{\Phi} = \begin{bmatrix} e^{-j\omega_1} & & & \\ & e^{-j\omega_2} & & \\ & & \ddots & \\ & & & e^{-j(L-1)\omega_D} \end{bmatrix}. \tag{5.16}$$

If $E\{\mathbf{Z}\} = [\mathbf{z_{11}}, \mathbf{z_{12}}, \ldots, \mathbf{z_{1}}_L]$ and $\mathbf{R} = \frac{1}{\sqrt{L}}\left(\mathbf{A_{11}}\mathbf{P}\mathbf{A}_{\mathbf{11}}^H + \sigma_n^2\mathbf{I}\right)$, b constant,

$\mathbf{A_{11}} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ e^{j\omega_1} & e^{j\omega_2} & \ldots & e^{j\omega_D} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(L-1)\omega_1} & e^{j(L-1)\omega_2} & \ldots & e^{j(L-1)\omega_D} \end{bmatrix}$ and $L = \frac{N^2}{4} + \frac{N}{2}$, then $E\{\mathbf{Z}\} = b\mathbf{R}$.

Proof.

$$\mathbf{z_{11}} = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ e^{j\omega_1} & e^{j\omega_2} & \ldots & e^{j\omega_D} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(L-1)\omega_1} & e^{j(L-1)\omega_2} & \ldots & e^{j(L-1)\omega_D} \end{bmatrix} \mathbf{\Phi}^0 \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_d^2 \end{bmatrix} + \sigma_n^2\bar{\mathbf{e}}_1 = \begin{bmatrix} \sum_{i=1}^{D}\sigma_i^2 \\ \sum_{i=1}^{D}e^{j\omega_i}\sigma_i^2 \\ \vdots \\ \sum_{i=1}^{D}e^{j(L-1)\omega_i}\sigma_i^2 \end{bmatrix} + \sigma_n^2\bar{\mathbf{e}}_1. \tag{5.17}$$

$$\mathbf{z_{12}} = \begin{bmatrix} \sum_{i=1}^{D}e^{-j\omega_i}\sigma_i^2 \\ \sum_{i=1}^{D}\sigma_i^2 \\ \vdots \\ \sum_{i=1}^{D}e^{j(L-2)\omega_i}\sigma_i^2 \end{bmatrix} + \sigma_n^2\bar{\mathbf{e}}_2. \tag{5.18}$$

Hence,

$$\mathbf{z}_{1i} = \begin{bmatrix} \sum_{k=1}^{D} e^{j(1-i)\omega_k}\sigma_k^2 \\ \sum_{k=1}^{D} e^{j(2-i)\omega_k}\sigma_k^2 \\ \vdots \\ \sum_{k=1}^{D} e^{j(L-i)\omega_k}\sigma_k^2 \end{bmatrix} + \sigma_n^2\bar{\mathbf{e}}_i,\ \forall i \in \{1,2,\ldots,L\}. \tag{5.19}$$

Therefore,

$$E\{\mathbf{Z}\} = \begin{bmatrix} \sum_{i=1}^{D}\sigma_i^2 & \sum_{i=1}^{D}\sigma_i^2 e^{-j\omega_i} & \ldots & \sum_{i=1}^{D}\sigma_i^2 e^{-j(L-1)\omega_i} \\ \vdots & \sum_{i=1}^{D}\sigma_i^2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{D}\sigma_i^2 e^{j(L-1)\omega_i} & \sum_{i=1}^{D}\sigma_i^2 e^{j(L-2)\omega_i} & \ldots & \sum_{i=1}^{D}\sigma_i^2 \end{bmatrix} + \sigma_n^2\mathbf{I}. \tag{5.20}$$

Also,

$$\mathbf{A}_{11}\mathbf{\Lambda}\mathbf{A}_{\mathbf{11}}^H = \begin{bmatrix} \sum_{i=1}^{D}\sigma_i^2 & \sum_{i=1}^{D}\sigma_i^2 e^{-j\omega_i} & \ldots & \sum_{i=1}^{D}\sigma_i^2 e^{-j(L-1)\omega_i} \\ \vdots & \sum_{i=1}^{D}\sigma_i^2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{D}\sigma_i^2 e^{j(L-1)\omega_i} & \sum_{i=1}^{D}\sigma_i^2 e^{j(L-2)\omega_i} & \ldots & \sum_{i=1}^{D}\sigma_i^2 \end{bmatrix}. \tag{5.21}$$

and

$$\mathbf{R} = \frac{1}{\sqrt{L}}\left(\mathbf{A}_{\mathbf{11}}\mathbf{P}\mathbf{A}_{\mathbf{11}}^H + \sigma_n^2\mathbf{I}\right). \tag{5.22}$$

From (5.20) and (5.22), we have

$$E\left\{\mathbf{Z}\right\} = b\mathbf{R}. \tag{5.23}$$

□

Therefore, we can apply high resolution subspace-based algorithms as MUSIC, root-MUSIC, and ESPRIT in $E\left\{\mathbf{Z}\right\}$ in order to estimate the angles of the impinging signals. Also it should be noted that the $E\left\{\mathbf{Z}\right\}$ is a positive definite matrix.

The matrix $E\left\{\mathbf{Z}\right\} \in \mathbb{C}^{L\times L}$ has $rank = L$ $(full - rank)$, therefore we can apply SVD method in $E\left\{\mathbf{Z}\right\}$ in order to estimate the subspace of the signal. It also should be noted that $L = \frac{N^2}{4} + \frac{N}{2} \geq N$, where N is the number of the physical elements. Therefore, using Nested-

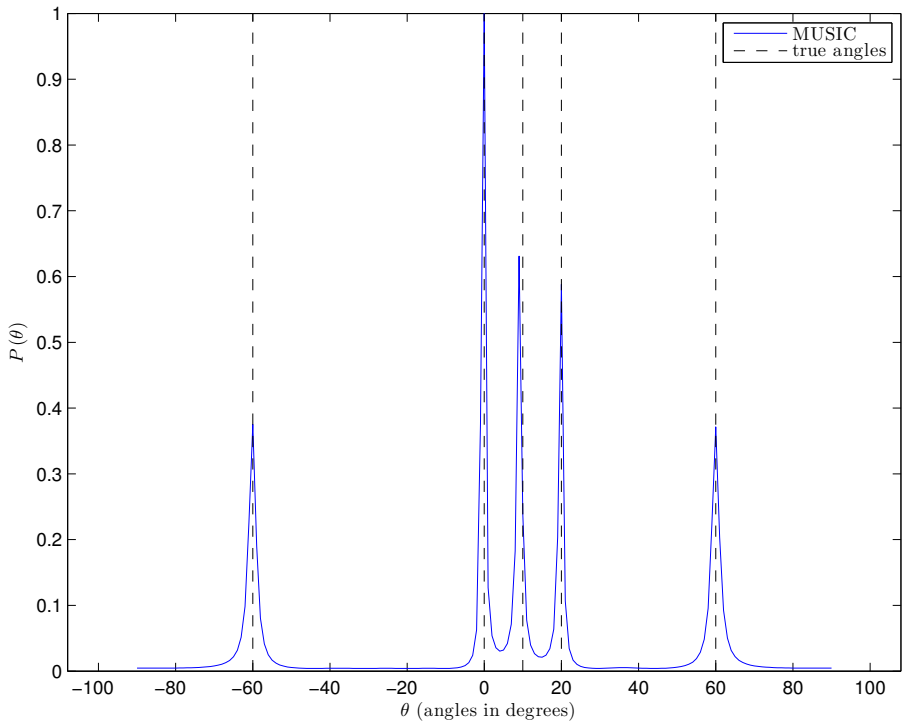

Figure 5.3: Nested arrays, 6 elements, 5 sources, 50 snapshots.

Arrays we can estimate more targets than the number of physical antennas as be used to linear arrays.

5.4 Simulations

In this section, we provide numerical examples to illustrate the performance of nested arrays. In Figure 5.3 we observe that by using a nested array we need only a small number of samples for the estimatation of the incoming angles. In Figures 5.4 and 5.5, the number of the source signals is greater than the number of elements. We observe that the nested array can estimate 7 incoming signals, but for 9 source signals it needs a larger number of snapshots, as shown in Figure 5.3. Also, an interesting example is shown in Figure 5.2. We observe that using a nested array with 10 elements at the receiver allows us to estimate 24 incoming angles by having collected 16000 samples.

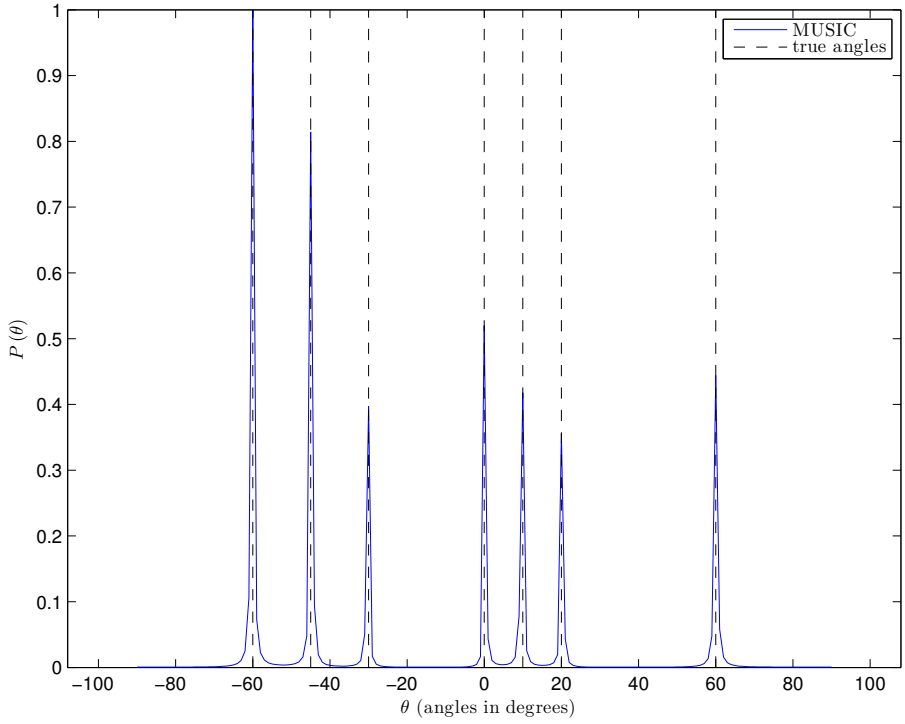

Figure 5.4: Nested arrays, 6 elements, 7 sources, 500 snapshots.

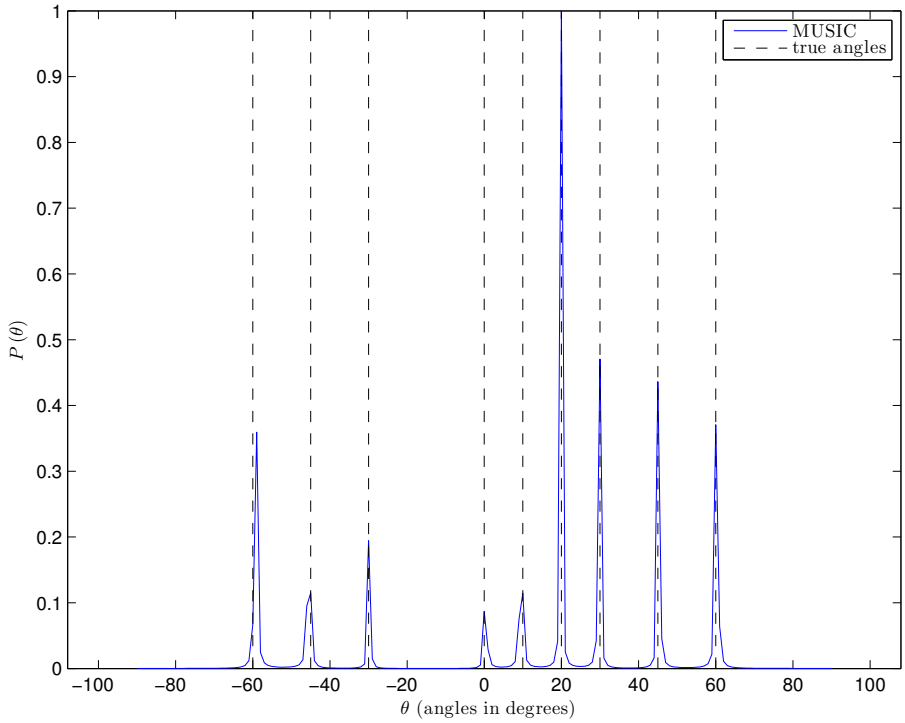

Figure 5.5: Nested arrays, 6 elements, 9 sources, 500 snapshots.

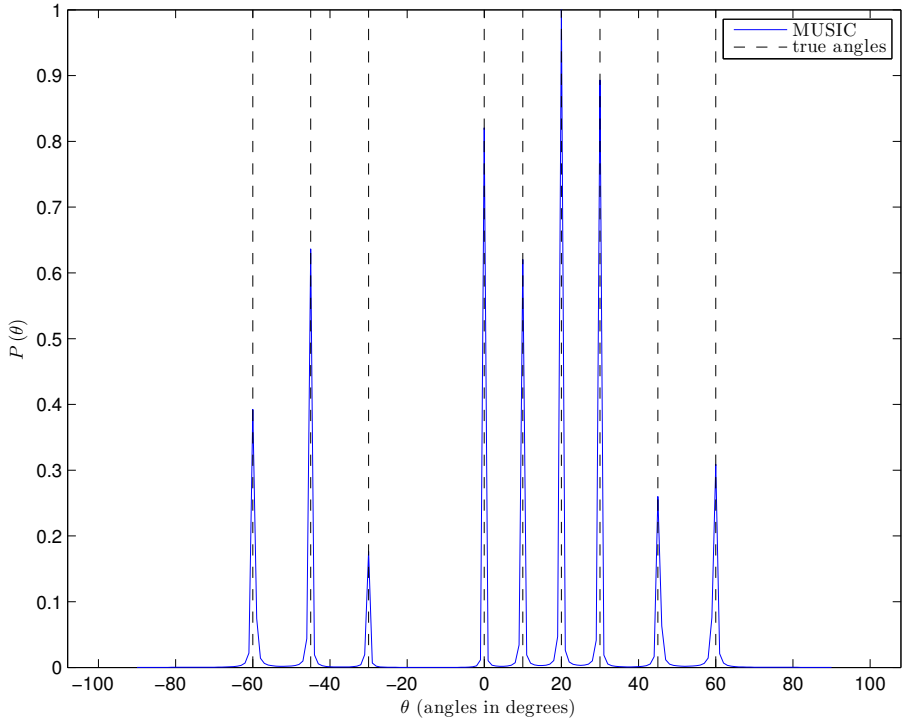

Figure 5.6: Nested arrays, 6 elements, 9 sources, 1000 snapshots.

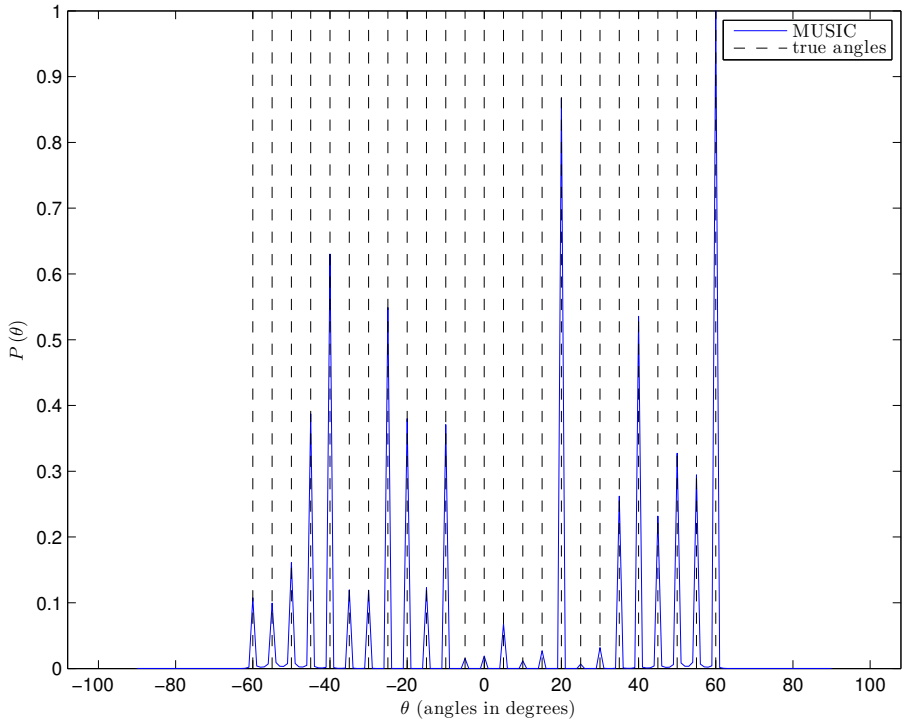

Figure 5.7: Nested arrays, 10 elements, 25 sources, 16000 snapshots.

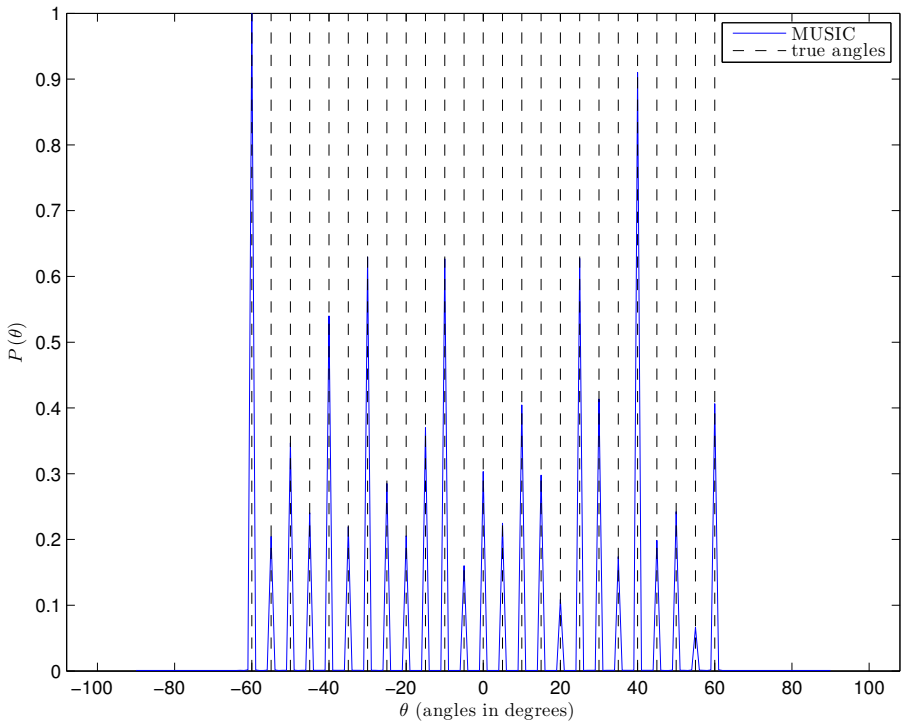

Figure 5.8: Nested arrays, 20 elements, 25 sources, 600 snapshots.

Chapter 6

Subspace Tracking for Nested Arrays

6.1 Definintion of The Problem

Let $\mathbf{Z} \in \mathbb{C}^{L\times L}$, non positive definite (pd) or positive semi definite (psd) matrix, that is an estimation of the correlation matrix $E\{\mathbf{Z}\}$ (equation 5.23) at the time instant t. Our goal is to modify the matrix $\mathbf{Z}$ suitably in order to be a positive definite matrix. We define

$$\lambda = \max_{i=1,2,\ldots,L} \left(\sum_{j\neq i} |z_{ij}| - z_{ii} \right) \tag{6.1}$$

where z_{ij}, $i,j \in \{1,\ldots,L\}$ are the entries of the matrix $\mathbf{Z}$, s.t. $\mathbf{Z}' = \mathbf{Z} + \lambda\mathbf{I}$ to be Hermitian diagonally dominant matrix.

Gershogorin Circle Theorem

Let $\mathbf{A}$, complex $n \times n$ matrix with entries a_{ij}. $\forall i \in \{1,2,\ldots,n\}$, let $\mathbf{R}_i = \sum\limits_{j\neq i} |a_{ij}|$ be the sum of the absolute values of the non-diagonal entries in the i-th row. Let λ eigenvalue of $\mathbf{A}$ and $\mathbf{x} = (\mathbf{x}_i)$ corresponding eigenvector. $\forall i \in \{1,2,\ldots,n\}$, s.t. $|x_i| = \max_j |x_i| \Rightarrow |x_j| > 0$, therefore $\mathbf{Ax} = \lambda\mathbf{x}$, and

$$\begin{aligned}
&\sum_{j\neq i} a_{ij}x_j = \lambda x_i - a_{ii}x_i \\
&\Rightarrow \lambda - a_{ii}x_i = \frac{\sum\limits_{j\neq i} a_{ii}x_i}{x_i} \\
&\Rightarrow |\lambda - a_{ii}| = \frac{|\sum\limits_{j\neq i} a_{ij}x_i|}{|x_i|} \leq \sum_{j\neq i} \frac{|a_{ij}x_i|}{|x_i|} \leq \sum_{j\neq i} |a_{ij}| = \mathbf{R}_i.
\end{aligned}$$

Where the last inequality is valid because $\frac{|x_j|}{|x_i|} \leq 1$, for $i \neq j$. Therefore

$$\lambda - a_{ii} \leq \sum_{j\neq i} |a_{ij}|$$

or

$$\lambda - a_{ii} \geq -\sum_{j\neq i} |a_{ij}|.$$

Hence,

$$\lambda \in \left[a_{ii} - \sum_{j \neq i} |a_{ij}|, a_{ii} + \sum_{j \neq i} |a_{ij}| \right]. \tag{6.2}$$

The $\mathbf{Z}'$ is a hermitian diagonally dominant matrix, therefore $z_{ii} \geq \sum_{j \neq i} |z_{ij}|, \forall i \in \{1, 2, \ldots, L\}$, hence from (6.2), $\lambda_i > 0$. Therefore $\mathbf{Z}'$ is a positive definite matrix. The positive definite matrices has the follows properties

Let positive definite matrix $\mathbf{A}$, $\in \mathbb{C}^{n\times}$, then

1. $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ for every non zero $\mathbf{x} \in \Re^{n\times 1}$.
2. All eigenvalues of $\mathbf{A}$ are positive.
3. $\mathbf{A} = \mathbf{T}\mathbf{T}^H$ for some nonsingular $\mathbf{T}$.

Following the $3th$ property of positive definite matrices, we consider that exists $\mathbf{T}$, such that $\mathbf{Z}' = \mathbf{T}\mathbf{T}^H$ and $E\{\mathbf{Z}'\} = E\left\{\mathbf{T}\mathbf{T}^H\right\}$.

We consider the following scalar function

$$\begin{aligned}
&E\left\{\|\mathbf{T} - \mathbf{W}\mathbf{W}^H\mathbf{T}\|_F^2\right\} \\
&= E\left\{\mathrm{tr}\left[\left(\mathbf{T} - \mathbf{W}\mathbf{W}^H\mathbf{T}\right)\left(\mathbf{T} - \mathbf{W}\mathbf{W}^H\mathbf{T}\right)^H\right]\right\} \\
&= E\left\{\mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\right) - \mathrm{tr}\left(\mathbf{T}\mathbf{T}^H\mathbf{W}\mathbf{W}^H\right) - \mathrm{tr}\left(\mathbf{W}\mathbf{W}^H\mathbf{T}\mathbf{T}^H\right) + \mathrm{tr}\left(\mathbf{W}\mathbf{W}^H\mathbf{T}\mathbf{T}^H\mathbf{W}\mathbf{W}^H\right)\right\} \\
&= \mathrm{tr}\left(E\left\{\mathbf{Z}'\right\}\right) - 2\mathrm{tr}\left(E\left\{\mathbf{Z}'\right\}\mathbf{W}\mathbf{W}^H\right) + \mathrm{tr}\left(\mathbf{W}^H E\left\{\mathbf{Z}'\right\}\mathbf{W}\mathbf{W}^H\mathbf{W}\right).
\end{aligned} \tag{6.3}$$

According to the (4.5), the gradient of $J(\mathbf{W})$ respect to $\mathbf{W}$ is

$$\bigtriangledown J(\mathbf{W}) = \left(-2E\{\mathbf{Z}'\} + E\{\mathbf{Z}'\}\mathbf{W}\mathbf{W}^H + \mathbf{W}\mathbf{W}^H E\{\mathbf{Z}'\}\right)\mathbf{W}. \tag{6.4}$$

Let $\mathbf{Q}_s$ subspace of $E\{\mathbf{Z}\}$ and $E\{\mathbf{Z}'\}$[1], $\mathbf{U} \in \mathbb{C}^{r\times r}$ orthogonal arbitrary matrix, and

$$\mathbf{W} = \mathbf{Q}_s\mathbf{U}. \tag{6.5}$$

[1] if $E\{\mathbf{Z}\} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H$ then $E\{\mathbf{Z}'\} = \mathbf{Q}\left(\mathbf{\Lambda} + E\{\lambda\}\mathbf{I}\right)\mathbf{Q}^H$.

Proof. Let,

$$\begin{aligned}
E\left\{\mathbf{Z}'\right\} &= \mathbf{Q}\left(\mathbf{\Lambda} + E\{\lambda\}\mathbf{I}\right)\mathbf{Q}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathbf{H}} + \mathbf{Q}E\{\lambda\}\mathbf{I}\mathbf{Q}^H \\
&= \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H + E\{\lambda\}\mathbf{Q}\mathbf{Q}^H = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H + E\{\lambda\}\mathbf{\Lambda}. \text{ (true)}
\end{aligned}$$

Therefore, $E\left\{\mathbf{Z}'\right\} = \mathbf{Q}\left(\mathbf{\Lambda} + E\{\lambda\}\mathbf{I}\right)\mathbf{Q}^H$. □

From (6.4) and (6.5),

$$
\begin{aligned}
&-2E\left\{\mathbf{Z}^{'}\right\}\mathbf{Q}_s\mathbf{U}+E\left\{\mathbf{Z}^{'}\right\}\mathbf{Q}_s\mathbf{U}\mathbf{U}^H\mathbf{Q}_s^H\mathbf{Q}_s\mathbf{U}+\mathbf{Q}_s\mathbf{U}\mathbf{U}^HE\{\mathbf{Z}^{'}\}\mathbf{Q}_s\mathbf{U}\\
&=-\mathbf{Q}\left(\mathbf{\Lambda}+\lambda\mathbf{I}\right)\begin{bmatrix}\mathbf{I}_{r\times r}\\ \mathbf{0}_{(L-r)\times r}\end{bmatrix}\mathbf{U}+\mathbf{Q}_s\begin{bmatrix}\mathbf{I}_{r\times r} & \mathbf{0}_{\mathbf{r}\times(L-r)}\end{bmatrix}\left(\mathbf{\Lambda}+\lambda\mathbf{I}\right)\times\begin{bmatrix}\mathbf{I}_{r\times r}\\ \mathbf{0}_{(L-r)\times r}\end{bmatrix}\mathbf{U}\\
&=-\mathbf{Q}\begin{bmatrix}\mathbf{\Lambda}_s+\lambda\mathbf{I}_{r\times r}\\ \mathbf{0}_{(L-r)\times r}\end{bmatrix}\mathbf{U}+\mathbf{Q}_s\begin{bmatrix}\mathbf{I}_{r\times r} & \mathbf{0}_{r\times(L-r)}\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda}_s+\lambda\mathbf{I}_{r\times r}\\ \mathbf{0}_{(L-r)\times r}\end{bmatrix}\mathbf{U}\\
&=-\mathbf{Q}_s\left(\mathbf{\Lambda}_s+\lambda\mathbf{I}_{r\times r}\right)\mathbf{U}+\mathbf{Q}_s\left(\mathbf{\Lambda}_s+\lambda\mathbf{I}_{r\times r}\right)\mathbf{U}=\mathbf{0}_{L\times r}.
\end{aligned}
$$

Therefore $\mathbf{Q}_s\mathbf{U}$ is a solution of (6.4) and a stationary point of the scalar function (6.3). Hence, according to the Theorem 2 in Section 4.2, the scalar function (6.3) is minimized for $\mathbf{W}=\mathbf{Q}_s\mathbf{U}$.

Since (6.3) describes an unconstrained cost function to be minimized and its gradient has solution the signal subspace of $E\{\mathbf{Z}\}$, we can apply the stochastic approximation theorem which we refer in section 4.1. Therefore, the subspace update can be written as

$$\mathbf{W}_n=\mathbf{W}_{n-1}-\mu\left[-2\mathbf{Z}_n^{'}+\mathbf{Z}_n^{'}\mathbf{W}_{n-1}\mathbf{W}_{n-1}^H+\mathbf{W}_{n-1}\mathbf{W}_{n-1}^H\mathbf{Z}_n^{'}\right]\mathbf{W}_{n-1} \tag{6.6}$$

where $\mu>0$ is a step size to be suitably chosen and $\mathbf{Z}_n^{'}$ is an estimate of the correlation matrix $E\left\{\mathbf{Z}^{'}\right\}$ at the nth sample. Also, we can write the above algorithm, approximating $\mathbf{W}_{n-1}^H\mathbf{W}_{n-1}=\mathbf{I}$, as follows

$$\mathbf{W}_n=\mathbf{W}_{n-1}+\mu\left[\mathbf{Z}_n^{'}\mathbf{W}_{n-1}^H-\mathbf{W}_{n-1}\mathbf{W}_{n-1}^H\mathbf{Z}_n^{'}\mathbf{W}_{n-1}\right]. \tag{6.7}$$

We observe that the complexity of above algorithm is $\mathcal{O}\left(L^2r\right)$. Comparing with the complexity of SVD method which is $\mathcal{O}\left(L^3\right)$, we can apply the above algorithm achieving high performance with low complexity. Also as the Chapter 4 we can replace the expectation in (6.3) with

$$J\left(\mathbf{W}_n\right)=\sum_{i=1}^{n}b^{n-i}\|\mathbf{T}_i-\mathbf{W}_n\mathbf{W}_n^H\mathbf{T}_i\|^2. \tag{6.8}$$

Relation (6.8) can be obtained from (6.3), if we use $\mathbf{W}_n$ instead of $\mathbf{W}$ and

$$
\begin{aligned}
\mathbf{Z}_n^{\prime}&=\sum_{i=1}^{n}b^{n-i}\mathbf{T}_i\mathbf{T}_i^H=b\mathbf{Z}^{\prime}{}_{n-1}+\mathbf{T}_n\mathbf{T}_n^H\\
&=b\mathbf{Z}_{n-1}^{\prime}+\mathbf{Z}_n^{\prime}
\end{aligned} \tag{6.9}
$$

instead of $E\{\mathbf{Z}^{\prime}\}$.

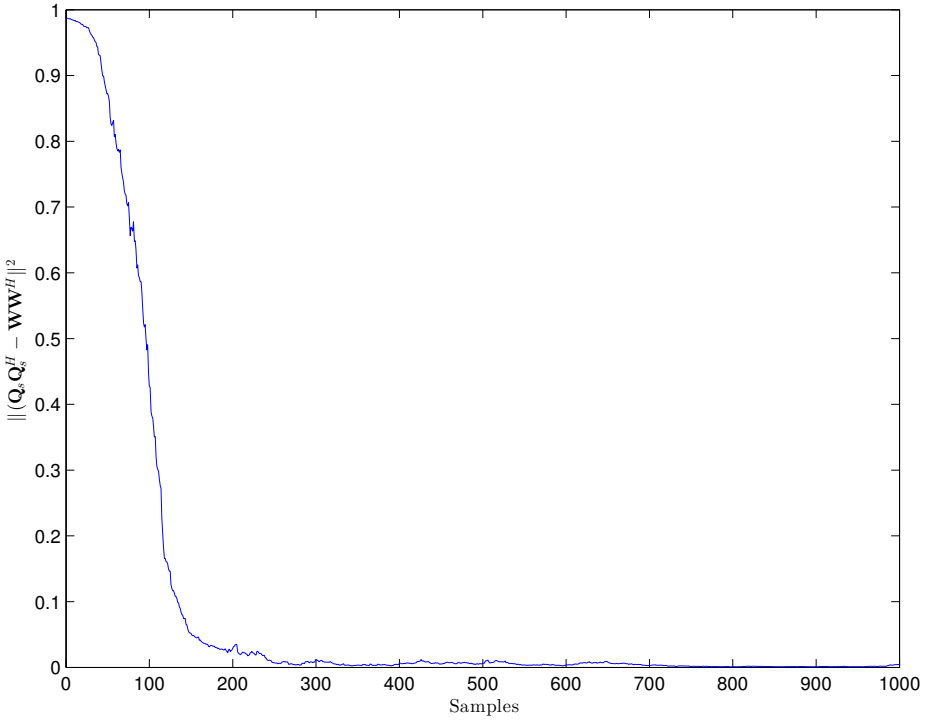

Figure 6.1: Stochastic Approximation (decreasing μ), 10 elements, 6 sources.

6.2 Simulations

In this section, we provide results from simulations about Subspace Tracking in nested arrays. Initially, it is important to show an example of how $\mathbf{W}\mathbf{W}^H$ converges to $\mathbf{Q}_s\mathbf{Q}_s^H$. In Figure 6.1 we show a case about 7 incoming signals at the receiver from direction $\{-45^o, -30^o, 0^o, 10^o, 20^o, 45^o\}$, where $\mathbf{Q}_s$ is the subspace of the received signal, and $\mathbf{W}$ is the estimated subspace using the iteration (6.6) for decreasing b.

Also, the case about stable μ is shown in Figure 6.2. As expected, we observe that using decreasing μ we can achieve better performance, as in the case of ULA. Subsequently, we evaluate the performance of iteration (6.6) using the forgetting factor b, where $0 \leq b \leq 1$ Figure 6.4. For $b = 1$, we have the Stochastic Approximation theorem. However, we are also interested in the case about non stationary environments. In Figure 6.5 results are provided for a change of target position from 20^o to 25^o for different values of b. From our simulations, we see that the algorithm provides better results for $b = 0.97$. We compare our algorithm with SVD and provide results from simulations in Figure 6.6. According to our simulations, we observed that the SVD method is efficient only for $b = 1$ and $b = 0.99$, therefore we decide to use these values for the rest of the simulations. Also, we observe that using Subspace Tracking with a forgetting factor we can achieve similar performance with the SVD method. We see that the SVD method (for $b = 0.99$) can re-estimate the angle of arrival of a source signal that moves faster than Subspace Tracking (ST), but after having collected more samples the performance of ST becomes more efficient compared to the the first. Also, we compare Subspace Tracking for nested arrays with Subspace Tracking for linear arrays, Figure 6.9.

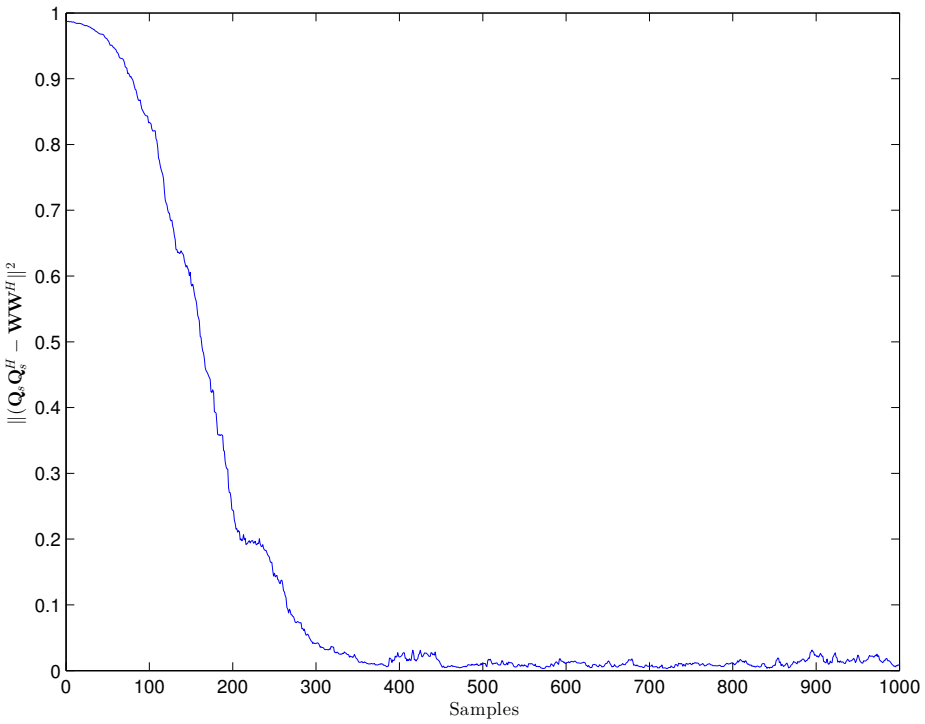

Figure 6.2: Stochastic Approximation (stable μ), 10 elements, 6 sources.

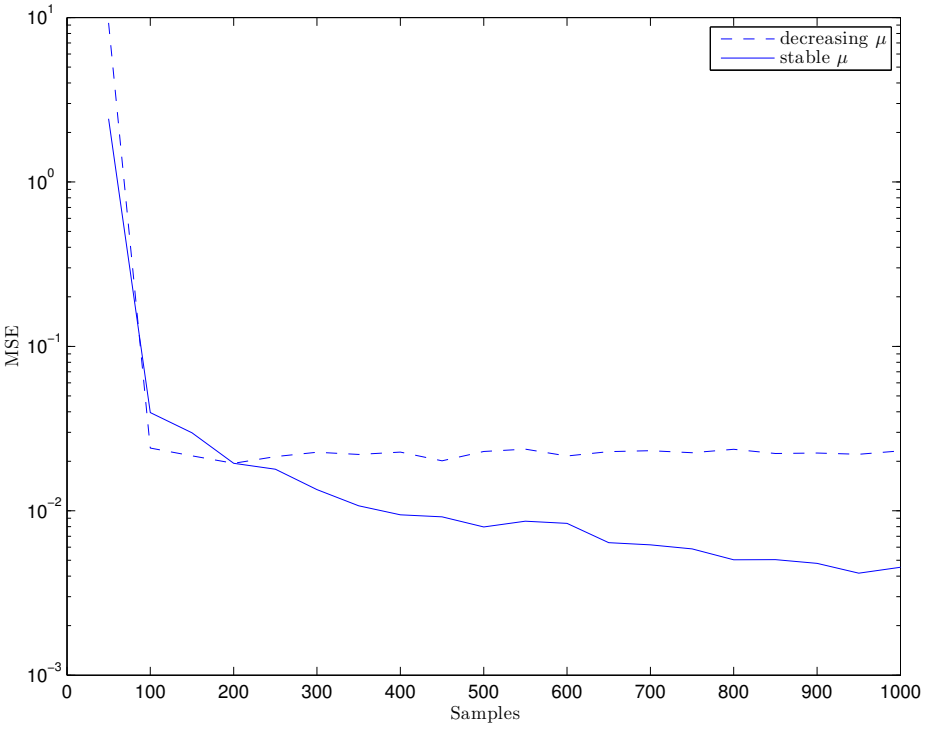

Figure 6.3: Nested Arrays (Stochastic Approximation), stable μ vs decreasing μ, 10 elements, 6 sources.

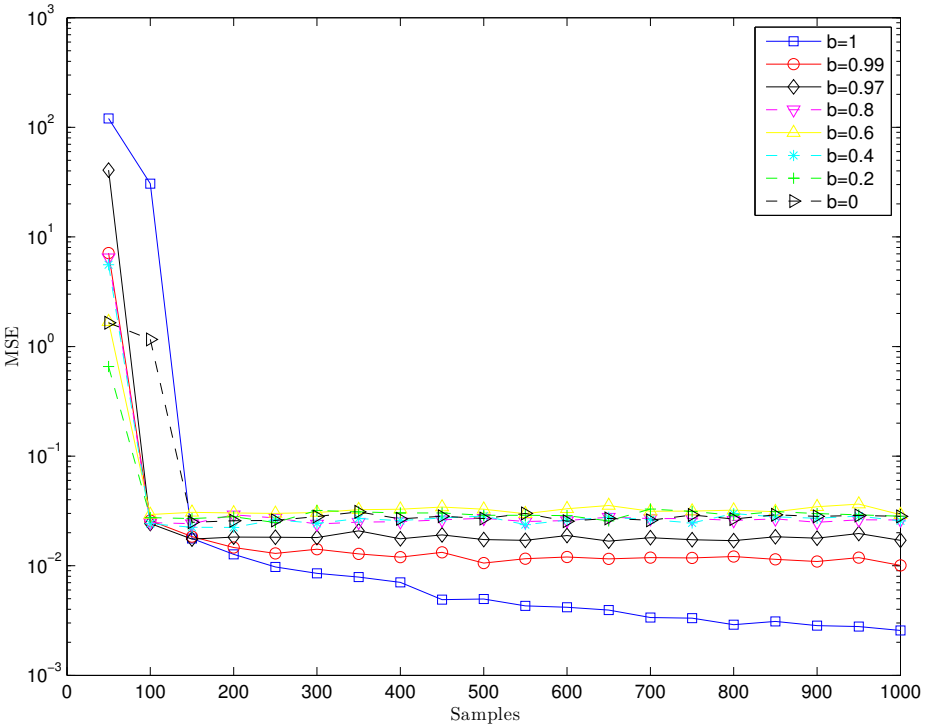

Figure 6.4: LMS for different values of b, 10 elements, 6 sources.

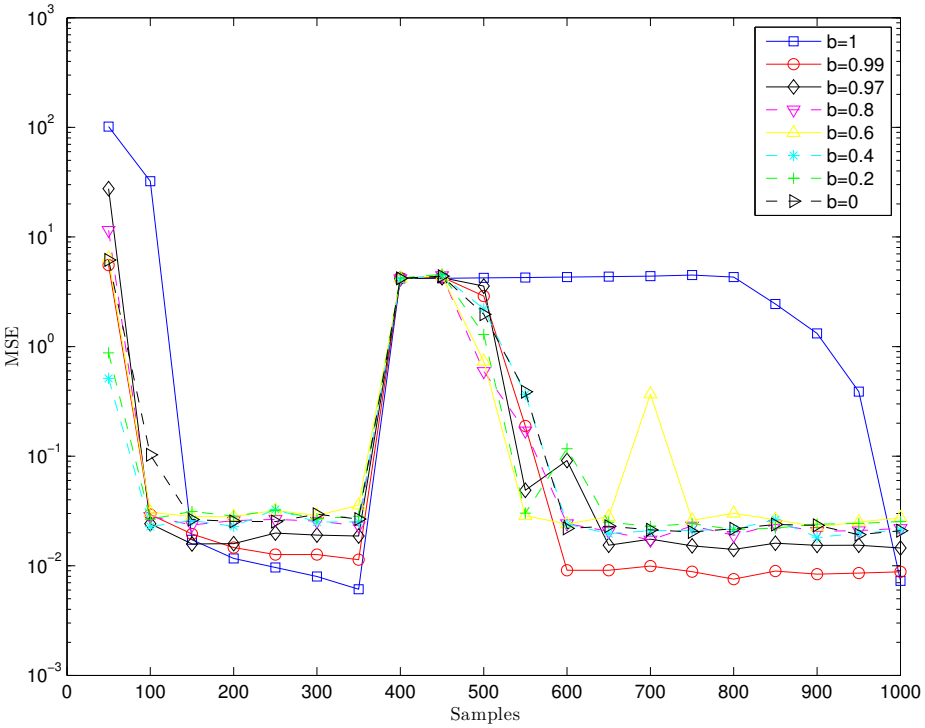

Figure 6.5: LMS for different values of b, 10 elements, 6 sources, non stationary environment.

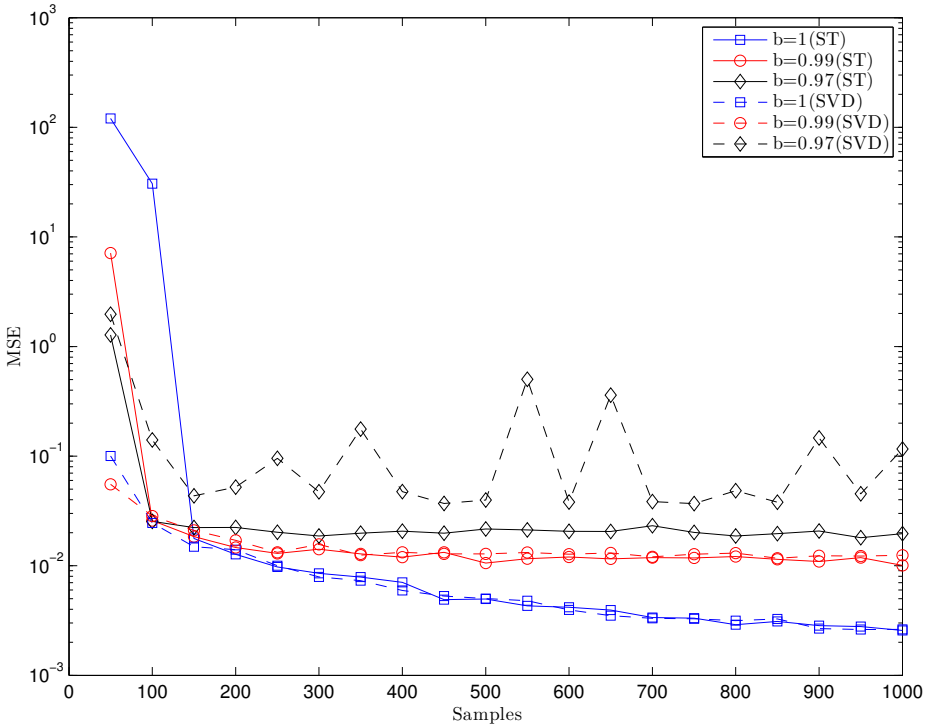

Figure 6.6: SVD vs Subspace Tracking, 10 elements, 6 sources, stationary environment.

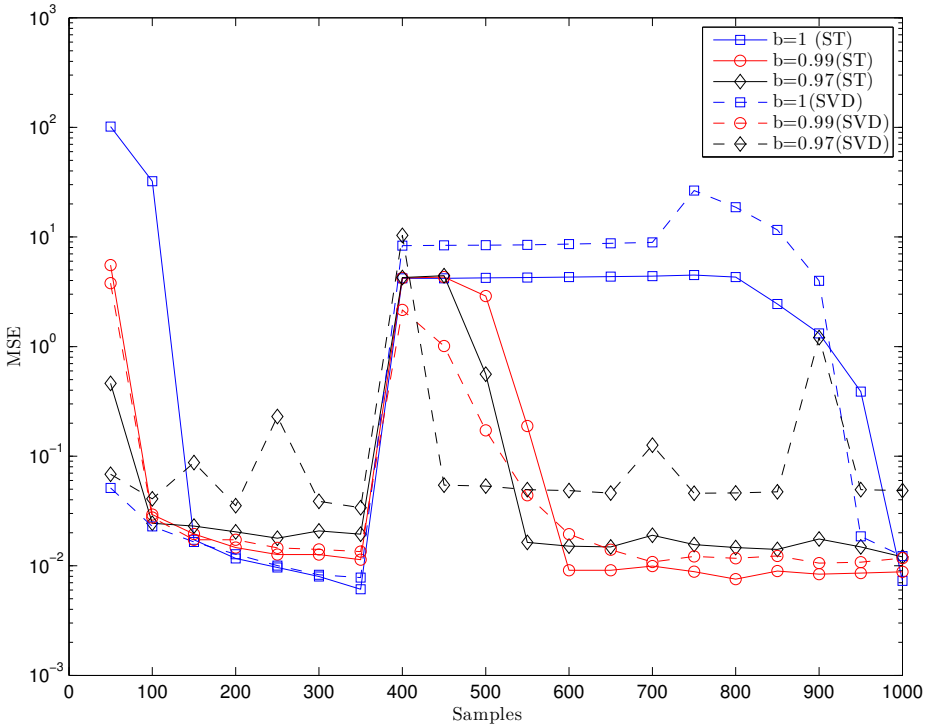

Figure 6.7: SVD vs Subspace Tracking, 10 elements, 6 sources, non stationary environment.

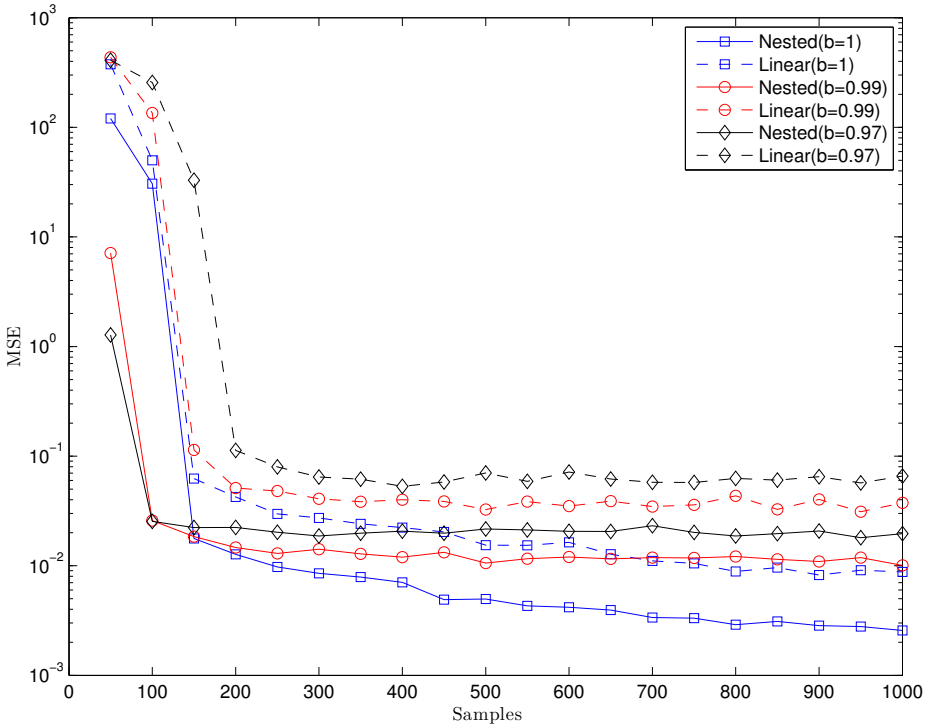

Figure 6.8: Subspace Tracking, Nested Vs Linear, stationary environment.

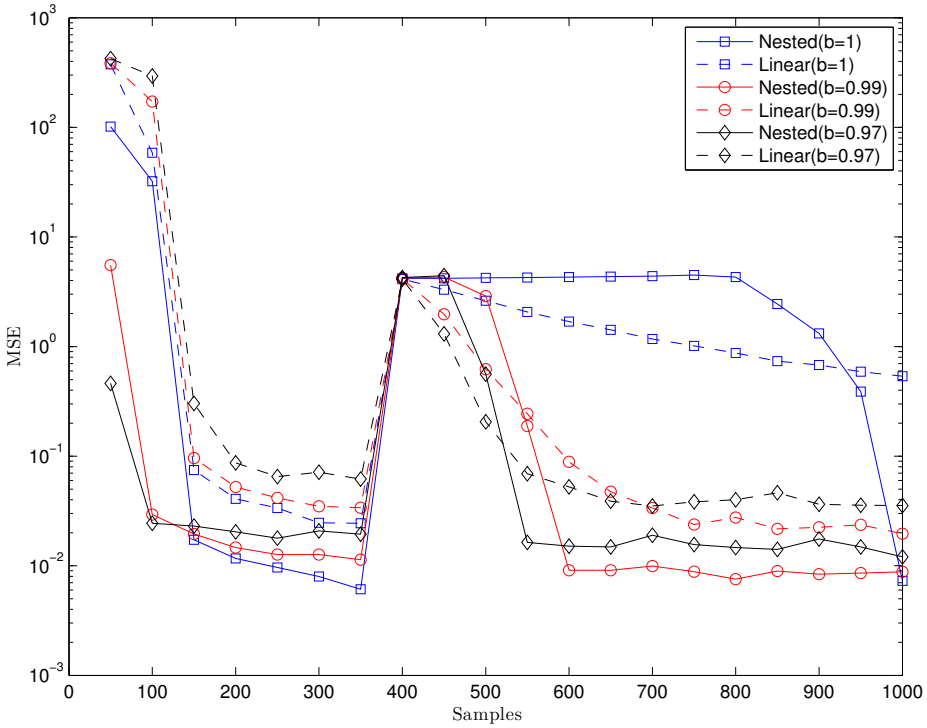

Figure 6.9: Subspace Tracking, Nested vs Linear, non stationary environment.

Bibliography

[1] Bhaskar D. Rao, and K. V. S. Hari. "Performance analysis of root-MUSIC," in *Proc. IEEE ICASSP 1989*, Glasgow, Scotland, Dec. 1989, vol. 37, no. 12, pp. 1939–1949.

[2] P. Pal and P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, 2010.

[3] H. Robbins and S. Monro,"A stochastic approximation method," *The Annals of Mathematical Statistics*, pp. 400–407, 1951.

[4] Meyer, Carl D, "Matrix analysis and applied linear algebra," vol. 2, 2000.

[5] Charbit, Maurice, "Digital signal and image processing using Matlab," vol. 666, 2010.

[6] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 95–107, 1995.

[7] Ma, Wing-Kin and Hsieh, Tsung-Han and Chi, Chong-Yung, "DOA estimation of quasi-stationary signals via Khatri-Rao product," in *Proc. IEEE ICASSP 2009*, Taipei International Convention Center Taipei, Taiwan, Apr. 2009, vol. 58, no. 4, pp. 2168–2180.