UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA MAGISTRALE IN
ICT FOR INTERNET AND MULTIMEDIA - INGEGNERIA PER LE COMUNICAZIONI MULTIMEDIALI E INTERNET

Statistics estimation of interference in mmWave channels

Relatore: Prof. Stefano Tomasin

Laureando: Raphael Hasler
Matricola: 1197753

ANNO ACCADEMICO 2021/2022

Data di laurea 11.04.2022

Contents

Acknowledgement

I would like to thank my supervisor, Professor Stefano Tomasin, who supported and guided me throughout the experimental study and the thesis.
I am grateful to Professor Antonia Tulino of Università degli Studi Federico II di Napoli, Professor Matilde Sánchez-Fernándes of Universidad Carlos III de Madrid and Professor Monica Barbara Nicoli, Professor Umberto Spagnolini and Professor Dario Tagliaferri of Politecnico di Milano for their precious contribution in the realization of this project. Each advice has been a source of inspiration for me, allowing me to conduct this work.

Abstract

MmWave communication systems are exposed to fundamental problems when high interference is present, due to the high number of devices, strong attenuation and transmission frequencies above 28GHz. The main questions are how this interference can be described statistically using the correlation matrix, how this interference correlation can be estimated and how the estimated interference correlation can be used to improve the capacity of a mmWave transmission. Furthermore, changes in correlation matrices should be detected, searching for methods which can be used in mmWave environments. In this work a simple mmWave environment was reproduced. A focus was set to the angles of arrival (AoA), which is a parameter that defines the interference. Traditional and novel methods for AoAs estimation, like MUSIC and methods based on a new gridless multidimensional AoAs estimation method, are used in the mmWave environment to estimate the interference correlation. They are compared with the least square estimation method in terms of mean square error (MSE) and in terms of capacity in a mmWave uplink transmission.The characteristics of the mmWave interference was analyzed and used for the search of a possible algorithm to detect changes in the interference correlation. This thesis revealed the importance of knowing the AoAs for the interference correlation estimation and the differences in terms of performance between the various estimation methods, showing an estimation improvement in some novel methods taking into account the MSE. Considering the capacity of a mmWave transmission, all methods, which use AoAs for the estimation of the interference correlation, accelerate the reaching of the maximal capacity significantly. Also a slight improve in terms of capacity of the novel gridless AoA estimation methods, with respect to the MUSIC algorithm, was detected. Moreover a method for change detection, called ODIT, was found and tested. It was asserted that it satisfies the main properties to be used as mmWave interference correlation change detection algorithm.

Keywords: *mmWaves*, *interference correlation*, *change detection*

Chapter 1

Introduction

This thesis deals with a fundamental problem in mmWave communication systems, when high interference is present, due to the high number of devices, strong attenuation and transmission frequencies above 28GHz. The consequences of this is that only a few echos of the transmitted signal are available at the receiver, typically around 3 or 4. This results in a reduced capacity due to attenuation, environment noise and interference noise coming from the other transmitters. To mitigate the effects of correlated interference, the correlation matrix of the interference is estimated. Traditional methods like least square (LS) estimate the correlation matrix step by step, using the interference vectors. This results in an accurate, but slow process and is not useful in mobile environments, in which fast changes in the environments occur. For this reason new methods, which could improve the estimation of the correlation matrix, have been found. A first focus was on searching methods to estimate the angles of arrival (AoA) of the interference devices signal echos. Traditional methods use solutions which make use of the estimated covariance matrix, like the multiple signal classification method (MUSIC). In this work this traditional method is compared with a new method, which consists in a gridless technique, based on the atomic norm minimization. They are compared when used for estimating the interference correlation matrix in terms of mean square error(MSE) and in terms of capacity, when the estimated interference correlation matrix is used in an uplink transmission channel. The second part of the thesis deals with the change in time of the real interference correlation matrix. The estimated correlation matrix could be, due to the mobility of the interferer, increase again in terms of MSE, with respect to the real correlation matrix. This brings a further challenge which was analyzed in this thesis. Different meth-

ods are analyzed which could be used to detect a change of the interference correlation matrix in time.

Chapter 2

Estimation of the interference correlation

MmWave transmissions deal with carrier frequencies above 28GHz which leads to a large bandwidth [3] but also to higher vulnerability due interference. The increased bandwidth comes at the expense of an increased attenuation [2]. Different techniques have been developed to overcome this problem, which focused by the majority on channel estimation techniques. The problem of strong attenuation leads to a particular structure of the mmWave channel, which is utilized in mmWave channel estimation techniques. This property is visible in the 2-Dimensional discrete Fourier transform (DFT) plane of the channel, where it appears as a structured parametric matrix with few parameters. The advantage of this property is that only a few paths of each channel between base station (BS) and user end (UE), must be considered, which are the paths that suffers less attenuation. Often this property is also called "sparsity" in the angular domain of the channel. Techniques like compressed sensing (CS) have made use of this property for channel estimation[1]. Other techniques use the particular algebraic structure of the mmWave channel, which is obtained exploiting the durability of the directions of the angles of arrival (AoA), the angles of transmission (AoT), the delays and the average power[1]. In this section the problem of attenuation is approached in mitigating noise by estimating the correlation matrix of the interference devices and using it for transmission. Being part of the noise sent of by other devices interference is summed up with the additive white Gaussian noise (AWGN) noise. Also the so called interference environment channel has a particular non parametric structure with few parameters and different researches [4] tried to use this for estimating

the correlation matrix. In this work the basic least square (LS) estimation was used as reference method to estimate the interference correlation matrix and then a sort of advantage was obtained using traditional and novel methods which estimate the AoAs of the interferer devices. This advantage was then used for the continuous correlation matrix estimation. The methods used for AoA estimation were the multiple signal classification (MUSIC)[5] procedure and a recent procedure developed for automotive radars, based on a gridless AoA estimation [6] [7]. Furthermore, the latter method was adapted and we tried to improve the performance using a clustering method and using the subspace of a previous done LS estimation. The performance was then evaluated in terms of mean square error (MSE) and in terms of capacity.

2.1 mmWave single user channel

We consider a single-carrier mmWave multiple input-multiple output (MIMO) channel. The receiver has N_R antennas and the transmitter N_T antennas. At both the transmitter and the receiver the antennas are organized in linear arrays. In an uplink transmission, at the base station (BS) the received signal is

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{N}, \tag{2.1}$$

where $\boldsymbol{Y} \in \mathbb{C}^{N_R \times 1}$ is the column vector of received signals, $\boldsymbol{X} \in \mathbb{C}^{N_T \times 1}$ represents the column vector of transmitted symbols, $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$ represents the channel matrix, and $\boldsymbol{N} \in \mathbb{C}^{N_R \times 1}$ is the noise column vector. In particular, $\boldsymbol{N}$ is a vector of additive white Gaussian noise (AWGN), with independent entries having zero mean and variance σ^2.
Entry (r, t) of channel matrix $\boldsymbol{H}$ is

$$[\boldsymbol{H}]_{r,t} = \sum_{i=0}^{N_C-1} \breve{g}_i e^{-j2\pi(r-1)\eta_i^{(\mathrm{Rx})}} e^{-j2\pi(t-1)\eta_i^{(\mathrm{Tx})}}, \tag{2.2}$$

where $r = 1, ..., N_R$, $t = 1, ..., N_T$, N_C is the number of reflected rays, $j = \sqrt{-1}$, $\breve{g}_i$ is a complex gain and $\eta_i^{(\mathrm{Tx})}$/ $\eta_i^{(\mathrm{Rx})}$ are the transmitter/receiver phase shifts, defined as

$$\begin{aligned}\eta_i^{(\mathrm{Tx})} &= \frac{D_{min}^{(\mathrm{Tx})}}{\lambda}\sin\alpha_i^{(\mathrm{Tx})},\\ \eta_i^{(\mathrm{Rx})} &= \frac{D_{min}^{(\mathrm{Rx})}}{\lambda}\sin\alpha_i^{(\mathrm{Rx})},\end{aligned} \tag{2.3}$$

where $D_{min}^{(\mathrm{Tx})}$ and $D_{min}^{(\mathrm{Rx})}$ are the distances between the elements in the linear antenna transmitter/receiver arrays and λ is the wavelength of the carrier signal. At mmWaves we have $\lambda \in [1-10]$mm, which corresponds to a frequency of $30-300$GHz.
Angles $\alpha_i^{(\mathrm{Tx})}$ and $\alpha_i^{(\mathrm{Rx})}$ represent the angle of departure (AoD) and the angle of Arrival (AoA) of ray i. Variable i is the index of the ray and it varies from $i=0$ to N_C-1. In mmWaves, the attenuation is high, due to the high operating frequencies and only a few rays participate in the formation of the useful signal (typically $N_C=[1-3]$).

2.1.1 Angular domain representation

The mmWave channel can be well represented in the angular domain by taking the two-dimensional discrete Fourier transform (DFT) of the channel matrix. Let $\Omega=[\Omega_1,\Omega_2]$ and let us define matrix $\boldsymbol{W(\Omega)}\in\mathbb{C}^{M_1\times M_2}$ with entries

$$\begin{aligned}[\boldsymbol{W}(\Omega)]_{\phi_1,\phi_2} = &\left[N_R\mathrm{sinc}_{N_R}\left(\frac{N_R(\phi_1-\Omega_1)}{M_1}\right)\mathrm{e}^{-j\pi(\phi_1-\Omega_1)\frac{N_R-1}{M_1}}\right]\times\\ &\left[N_T\mathrm{sinc}_{N_T}\left(\frac{N_T(\phi_1-\Omega_1)}{M_2}\right)\mathrm{e}^{-j\pi(\phi_2-\Omega_2)\frac{N_T-1}{M_2}}\right],\end{aligned} \tag{2.4}$$

where $\phi_1=0,1,...,M_1-1$, and $\phi_2=0,1,...,M_2-1$ and

$$\mathrm{sinc}_N(a)=\begin{cases}\frac{\sin(\pi a)}{N\sin(\frac{\pi a}{N})}, a\neq 0,\\ 1, a=0.\end{cases} \tag{2.5}$$

The 2-D DFT of sizes M_1 and M_2 of $\boldsymbol{H}$ is matrix $\boldsymbol{\Phi}\in\mathbb{C}^{M_1\times M_2}$, with entries

$$[\boldsymbol{\Phi}]_{\phi_1,\phi_2}=\sum_{r=1}^{N_R}\sum_{t=1}^{N_T}[\boldsymbol{H}]_{r,t}\mathrm{e}^{\frac{-j2\pi\phi_1(r-1)}{M_1}}\mathrm{e}^{\frac{-j2\pi\phi_2(t-1)}{M_2}}=\sum_{i=0}^{N_C-1}\check{g}_i[\boldsymbol{W}(\boldsymbol{\Omega^{(i)}})]_{\phi_1,\phi_2}. \tag{2.6}$$

2.2 Interference in mmWave channels

Considering (2.1), vector $\boldsymbol{N}$ includes not only AWGN but also interference coming from other transmitters (e.g, in nearby cells), thus

$$\boldsymbol{N} = \boldsymbol{G}\boldsymbol{J} + \boldsymbol{Z}, \tag{2.7}$$

where $\boldsymbol{G} \in \mathbb{C}^{N_R \times N_I}$ is the complex interference channel matrix that describes the narrowband channel between the receiver and the interfering transmit antennas, N_I is the number of interference signals, $\boldsymbol{J} \in \mathbb{C}^{N_I \times 1}$ is the interference column vector, and $\boldsymbol{Z} \in \mathbb{C}^{N_R \times 1}$ is the AWGN vector with independent entries having zero mean and variance σ^2. Vectors $\boldsymbol{J}$ and $\boldsymbol{Z}$ are independent from each other, since they have different non correlated origins. The entries of $\boldsymbol{J}$ have zero mean and are assumed to be independent zero-mean Gaussian variables. For this reason, $\boldsymbol{N}$ is still zero mean. The cross-correlation of $\boldsymbol{N}$ is

$$\boldsymbol{R} = \mathbb{E}(\boldsymbol{N}\boldsymbol{N}^{\mathrm{H}}) = \mathbb{E}[(\boldsymbol{G}\boldsymbol{J} + \boldsymbol{Z})(\boldsymbol{G}\boldsymbol{J} + \boldsymbol{Z})^{\mathrm{H}}] = \boldsymbol{G}\boldsymbol{G}^{\mathrm{H}} + \sigma^2\boldsymbol{I}, \tag{2.8}$$

where $\boldsymbol{I} \in \mathbb{C}^{N_R \times N_R}$ is the identity matrix. We have

$$\boldsymbol{G}\boldsymbol{J} + \boldsymbol{Z} \sim \boldsymbol{G'}\boldsymbol{J'}, \tag{2.9}$$

in which $\sim$ means that the statistical distribution between the two sides is the same and the column vector $\boldsymbol{J'} \in \mathbb{C}^{N_I \times 1}$ are Gaussian distributed with zero mean independent entries. From (2.9) and (2.8) we immediately have

$$\mathbb{E}[(\boldsymbol{G}\boldsymbol{J}+\boldsymbol{Z})(\boldsymbol{G}\boldsymbol{J}+\boldsymbol{Z})^{\mathrm{H}}] = \mathbb{E}[\boldsymbol{G'}\boldsymbol{J'}\boldsymbol{J'}^{\mathrm{H}}\boldsymbol{G'}^{\mathrm{H}}] = \boldsymbol{G'}\boldsymbol{G'}^{\mathrm{H}} = \boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}+\sigma^2\boldsymbol{I}. \tag{2.10}$$

Now using the singular value decomposition (SVD), $\boldsymbol{G}\boldsymbol{R}_J^{\frac{1}{2}}$ can be decomposed as

$$\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}, \tag{2.11}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices and $\boldsymbol{\Lambda}$ is a diagonal matrix. From (2.7) and (2.11), we obtain

$$\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}} + \sigma^2\boldsymbol{I} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}} + \sigma^2\boldsymbol{U}\boldsymbol{U}^{\mathrm{H}} = \boldsymbol{U}(\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}} + \sigma^2\boldsymbol{I})\boldsymbol{U}^{\mathrm{H}}. \tag{2.12}$$

Choosing $\boldsymbol{G'} = \boldsymbol{U}\sqrt{\boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}} + \sigma^2\boldsymbol{I}}$ and considering (2.9), it can be stated that the AWGN vector $\boldsymbol{Z}$ can be absorbed by the interference.

2.3 Interference correlation estimation

The least square (LS) estimate of correlation matrix is defined as

$$\hat{\boldsymbol{R}_{ls}} = \frac{1}{T_{\max}} \sum_{t=0}^{T_{\max}-1} \boldsymbol{N}(t)\boldsymbol{N}(t)^{\mathrm{H}}, \tag{2.13}$$

where $T_{\max}$ is the number of samples of $\boldsymbol{N}$ on which we estimate the matrix. For $T_{\max} \to \infty$ the estimated correlation matrix $\hat{\boldsymbol{R}_{ls}}$ converges to $\boldsymbol{R}$. Considering the case with N_{int} interferers with N_T antennas each. We have $N_I = N_{int}N_T$ and we can write

$$\boldsymbol{G} = [\boldsymbol{G}_1, ..., \boldsymbol{G}_{N_{int}}]. \tag{2.14}$$

Considering that $\boldsymbol{G}_\ell$ where $\ell = 1, ..., N_{int}$, has the same model as $\boldsymbol{H}$, from (2.2) we have:

$$\boldsymbol{G_\ell} = \sum_{i=0}^{N_g-1} v_{i,\ell}\boldsymbol{a}_{N_R}(\beta_{i^{(\mathrm{Rx})},\ell})\boldsymbol{a}_{N_I}^{\mathrm{T}}(\beta_{i^{(\mathrm{Tx})},\ell}), \tag{2.15}$$

where N_g is the number of received rays per interferer and

$$\boldsymbol{a}_N(\beta) = [1, e^{-2\pi j\beta}, ..., e^{-2\pi j\beta(N-1)}]^{\mathrm{T}}. \tag{2.16}$$

Thus, we have

$$\boldsymbol{GJ} = \sum_{\ell} \sum_{i=0}^{N_g-1} \boldsymbol{a}_{N_R}(\beta_{i,\ell}^{(\mathrm{Rx})}) \underbrace{[v_{i,\ell}\boldsymbol{a}_{N_I}^{\mathrm{T}}(\beta_{i,\ell}^{(\mathrm{Tx})})\boldsymbol{J}_\ell]}_{x_{i,\ell}}, \tag{2.17}$$

where

$$\boldsymbol{J} = [\boldsymbol{J}^{\mathrm{T}}, ..., \boldsymbol{J}_{N_{int}}^{\mathrm{T}}], \tag{2.18}$$

and $x_{i,\ell}$ is zero-mean complex Gaussian with variance

$$\begin{aligned} \mathbb{E}(x_{i,\ell}x_{i^*,\ell}) &= \mathbb{E}(|x_{i,\ell}|^2) \\ &= v_{i,\ell}v_{i^*,\ell}\boldsymbol{a}_{N_I}^{\mathrm{T}}(\beta_{i,\ell}^{(\mathrm{Tx})})\boldsymbol{J}_\ell\boldsymbol{a}_{N_I}^{*}(\beta_{i,\ell}^{(\mathrm{Tx})})\boldsymbol{J}_\ell^{*} \\ &= |v_{i,\ell}|^2\boldsymbol{a}_{N_I}^{\mathrm{T}}(\beta_{i,\ell}^{(\mathrm{Tx})})\boldsymbol{a}_{N_I}^{*}(\beta_{i,\ell}^{(\mathrm{Tx})}), \end{aligned} \tag{2.19}$$

The product of $\boldsymbol{a}_{N_I}^{*}(\beta_{i,\ell}^{(\mathrm{Tx})})\boldsymbol{J}_\ell^{*}$ results in a scalar value and for this reason the two vectors can be exchanged to $\boldsymbol{J}_\ell^{*}\boldsymbol{a}_{N_I}^{*}(\beta_{i,\ell}^{(\mathrm{Tx})})$. Then (2.19) can be written as

$$\mathbb{E}(|x_{i,\ell}|^2) = |v_{i,\ell}|^2\boldsymbol{a}_{N_I}^{\mathrm{T}}(\beta_{i,\ell}^{(\mathrm{Tx})})\boldsymbol{a}_{N_I}^{*}(\beta_{i,\ell}^{(\mathrm{Tx})}), \tag{2.20}$$

where the average is taken with respect to $\boldsymbol{J}_\ell$. The cross-correlation of $x_{i,\ell}$ and $x_{j,\ell}$ with $i \neq j$ is

$$\mathbb{E}(x_{i,\ell}x_{j,\ell}^*) = v_{i,\ell}v_{j,\ell}^*\boldsymbol{a}_{N_I}^*(\beta_{i,\ell}^{(\mathrm{Tx})})\boldsymbol{a}_{N_I}^{\mathrm{T}}(\beta_{j,\ell}^{(\mathrm{Tx})}). \tag{2.21}$$

Since $\boldsymbol{J}_\ell$ are independent and zero-mean, we can write

$$\boldsymbol{N}(t) = \boldsymbol{G}\boldsymbol{J} + \boldsymbol{Z} = \boldsymbol{A}\boldsymbol{x}(t), \tag{2.22}$$

with

$$\boldsymbol{A} = [\boldsymbol{a}_{N_R}(\beta_{1,1}^{(\mathrm{Rx})}), ..., \boldsymbol{a}_{N_R}(\beta_{N_g,1}^{(\mathrm{Rx})}), \boldsymbol{a}_{N_R}(\beta_{1,2}^{(\mathrm{Rx})}), ..., \boldsymbol{a}_{N_R}(\beta_{N_g,N_{int}}^{(\mathrm{Rx})})] \tag{2.23}$$

and $\boldsymbol{x}(t) = \sum_{\ell=1}^{N_{int}} \boldsymbol{x}_\ell(t)$ a correlated Gaussian vector. In the following we estimate the angles $\{\beta_{i,\ell}^{(Rx)}\}$ and then we estimate the correlation matrix of $\boldsymbol{x}(t)$ to obtain the full interference correlation matrix.

2.4 Sparse representation with single interference device

Consider now a single interfering device. The 2D-DFT of the interference channel has a structure depending on few parameters, considering only a few paths to the interferer. This property will be used for the estimation of the interference correlation matrix. In formulas, the $M \times M$ 2D-DFT is

$$\mathrm{DFT2}[\boldsymbol{G}] = \boldsymbol{F}\boldsymbol{G}\boldsymbol{F}^{\mathrm{H}} = \sum_{\ell=0}^{N_g-1} \breve{g}_\ell \boldsymbol{W}(\boldsymbol{\Omega}^\ell), \tag{2.24}$$

where N_g is the number of interference paths, and $\boldsymbol{F}$ is the 2-D $M \times M$ DFT matrix. Using the property $\boldsymbol{F}^{\mathrm{H}}\boldsymbol{F} = \boldsymbol{I}$, we have:

$$\mathrm{DFT2}[\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}] = \boldsymbol{F}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{F}^{\mathrm{H}} = \boldsymbol{F}\boldsymbol{G}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{F}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{F}^{\mathrm{H}}. \tag{2.25}$$

Considering the following relations:

$$\boldsymbol{F}\boldsymbol{G}\boldsymbol{F}^{\mathrm{H}} = \mathrm{DFT2}[\boldsymbol{G}], \tag{2.26}$$

$$\boldsymbol{F}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{F}^{\mathrm{H}} = \mathrm{DFT2}[\boldsymbol{G}^{\mathrm{H}}] = (\mathrm{DFT2}[\boldsymbol{G}])^{\mathrm{H}}, \tag{2.27}$$

entry (m, n) of matrix DFT2$(\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}})$ is

$$[\mathrm{DFT2}(\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}})]_{m,n} = \sum_{i=1}^{M} [\mathrm{DFT2}(\boldsymbol{G})]_{m,i} [\mathrm{DFT2}(\boldsymbol{G})]^*_{n,i}, \tag{2.28}$$

and using (2.24), we have

$$[\mathrm{DFT2}(\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}})]_{m,n} = \sum_{i=1}^{M} \sum_{\ell_1=0}^{N_{g-1}} \sum_{\ell_2=0}^{N_{g-1}} \check{g}_{\ell_1} \check{g}^*_{\ell_2} [\boldsymbol{W}(\boldsymbol{\Omega}^{(\ell_1)})]_{m,i} [\boldsymbol{W}(\boldsymbol{\Omega}^{(\ell_2)})]^*_{n,i}. \tag{2.29}$$

Thus

$$\begin{aligned} \mathrm{DFT2}(\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}) &= \sum_{\ell_1} \sum_{\ell_2} \check{g}_{\ell_1} \check{g}_{\ell_2} \boldsymbol{W}(\boldsymbol{\Omega})^{(\ell_1)} \boldsymbol{W}(\boldsymbol{\Omega})^{(\ell_2)} \\ &= \sum_{\ell} \check{g}'_{\ell} \boldsymbol{X}(\boldsymbol{\Omega}'^{(\ell)}), \end{aligned} \tag{2.30}$$

with $\boldsymbol{\Omega}'^{(\ell)} = [\boldsymbol{\Omega}'^{(\ell_1)}, \boldsymbol{\Omega}'^{(\ell_2)}]$
Example: Having only one reflection

$$\begin{aligned} [\mathrm{DFT2}(\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}})]_{m,n} &= \sum_{i=1}^{M} |\check{g}_1|^2 [\boldsymbol{W}(\boldsymbol{\Omega}^{(1)})]_{m,i} [\boldsymbol{W}(\boldsymbol{\Omega}^{(1)})]^*_{n,i} = \\ &\quad |\check{g}_1|^2 \sum_{i=1}^{M} [\boldsymbol{W}(\boldsymbol{\Omega}^{(1)})]_{m,i} [\boldsymbol{W}(\boldsymbol{\Omega}^{(1)})]^*_{n,i}. \end{aligned} \tag{2.31}$$

with

$$\mathrm{DFT2}(\boldsymbol{R}) = |\check{g}_1|^2 \boldsymbol{X}(\boldsymbol{\Omega}^{(1)}) \tag{2.32}$$

where

$$[\boldsymbol{X}(\boldsymbol{\Omega}^{(1)})]_{m,n} = \sum_{i=1}^{M} [\boldsymbol{W}(\boldsymbol{\Omega}^{(1)})]_{m,i} [\boldsymbol{W}(\boldsymbol{\Omega}^{(1)})]^*_{n,i}. \tag{2.33}$$

2.4.1 MUSIC angle of arrival estimation

The MUSIC-Algorithm [5] is an estimation method, invented by Ralph Otto Schmidt and it permits to estimate the AoA, evaluating a series of received vectors which are affected by interference. The algorithm calculates the correlation matrix between the received vectors. Successively the eigenvalues and the eigenvectors of the correlation matrix are calculated and two subspaces (noise and signal) are created. These subspaces are used to calculate a spectrum in which it is possible to determine the largest peaks and then to estimate the angles of arrival (AoA).

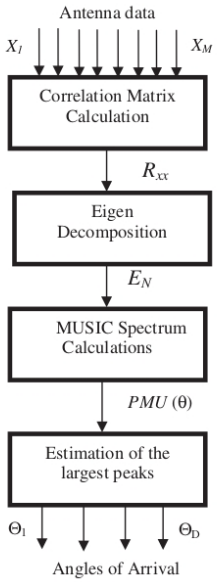

Figure 2.1: MUSIC AoA estimation steps [14]

2.4.2 Gridless angle of arrival estimation

This method is part of the gridless estimation techniques and it is based on the atomic norm minimization [6] [7]. This technique does not recover directly the AoA, but recovers the angular frequencies $\beta_{i,\ell}^{(\mathrm{Rx})}$, which are specified in Section 2.4.5. The main steps of the procedure can be summarized in the following points:

1. L0 atomic norm minimization of the received vectors.
2. Determination of the eigenvalues and eigenvectors

3. Eigendecomposition
4. Vandermode decomposition
5. Frequency pairing and acquisition of the angular frequencies

This new method should have advantages, compared to more traditional methods like MUSIC, especially in a rapidly changing environment where the number of homogeneous received vectors is limited. The estimation is done on a set of received vectors $\boldsymbol{N}$.

2.4.3 Gridless angle of arrival estimation - Short Method

The Gridless AoA - Short Method performs a LS estimation of the received vectors $\boldsymbol{N}$ using (2.13). The so estimated correlation matrix $\hat{\boldsymbol{R}}$ is cleaned of the noise

$$\hat{\boldsymbol{R}}_{NL} = \hat{\boldsymbol{R}} - \sigma^2 \boldsymbol{I}. \tag{2.34}$$

The obtained matrix is then transformed into a Hermitian matrix and the diagonal values are replaced with the mean power of the matrix $\hat{\boldsymbol{R}}$, i.e. ,

$$\mathrm{diag}(\hat{\boldsymbol{R}}_{NL}) = \mathrm{mean}(\mathrm{diag}(\hat{\boldsymbol{R}}_{NL})) \tag{2.35}$$

On the obtained matrix $\hat{\boldsymbol{R}}_{NL}$ an eigenvalue decomposition of the first N_g eigenvalues $\boldsymbol{\lambda}$, where $\boldsymbol{\lambda} \in \mathbb{C}^{N_g \times N_g}$, with the eigenvalues on the diagonal and the corresponding N_g largest eigenvectors $\boldsymbol{e}$, where $\boldsymbol{e} \in \mathbb{C}^{N_R \times N_g}$ is performed. Using the obtained vectors/values a subspace matrix $\boldsymbol{V}$ is calculated as

$$\boldsymbol{V} = \boldsymbol{e}\sqrt{\lambda}. \tag{2.36}$$

The so obtained matrix is then used in the previous described Gridless AoA estimation method on a fixed number of vectors N_g, to obtain the angular frequencies $\beta_i^{(\mathrm{Rx})}$ of the N_g rays. This method has the advantage to use in the fist step the LS estimation, which is not heavy in terms of calculation resources. Then in the second step the Gridless AoA estimation method is used on a fixed number of vectors N_g. In this way the AoAs are estimated continuously using new vectors $\boldsymbol{N}$ with a fixed amount of calculation power, due to the fixed vector size in Gridless AoA estimation.

2.4.4 Gridless angle of arrival estimation - Average Method

The Gridless AoA - Average Method performs an standard Gridless AoA estimation every t_{est} received vectors $\boldsymbol{N}(t)$ on t_{est} vectors. The estimated

angular frequencies $\beta_{i,\ell}^{(\mathrm{Rx})}$ are mapped in space. Using the new calculated $\beta_{i,\ell}^{(\mathrm{Rx})}$ and all the previous one, the N_g cluster centers are calculated and each $\beta_{i,\ell}^{(\mathrm{Rx})}$ assigned to a cluster. So the cluster centers are iteratively updated using new estimations of the angular frequencies $\beta_{i,\ell}^{(\mathrm{Rx})}$. Further, after the calculation of the cluster centers and the assignment to these, the clusters are cleaned removing far points from the clusters in clusters where more points than t/t_{est} are revealed. This because in optimal case every t_{est} noise vectors one point should be add to each cluster and if there are more points present, points from other clusters have been wrongly assigned. The obtained cluster centers are then used as $\beta_{i,\ell}^{(\mathrm{Rx})}$ in (2.16) for the calculation of $\hat{\boldsymbol{R}}$

2.4.5 Estimation of the correlation matrix with angle of arrivals (AoA)

It can be observed that the correlation matrix can be written as

$$\boldsymbol{R} = \boldsymbol{A}\mathbb{E}(\boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{H}}(t))\boldsymbol{A}^{\mathrm{H}} = \boldsymbol{A}\boldsymbol{R}_x\boldsymbol{A}^{\mathrm{H}} + \sigma^2\boldsymbol{I}. \tag{2.37}$$

$\boldsymbol{x}$ can be obtained from $\boldsymbol{N}$(t) as

$$\boldsymbol{A}^{\dagger}\boldsymbol{N}(t) = \boldsymbol{x}(t) + \boldsymbol{A}^{\dagger}\boldsymbol{Z}(t), \tag{2.38}$$

where $\boldsymbol{A}^{\dagger}$ is the pseudo-inverse of $\boldsymbol{A}$. Defining as

$$\boldsymbol{y} = \sum_t \boldsymbol{A}^{\dagger}\boldsymbol{N}\boldsymbol{N}^{\mathrm{H}}(\boldsymbol{A}^{\dagger})^{\mathrm{H}} = \boldsymbol{A}^{\dagger}\hat{\boldsymbol{R}}(\boldsymbol{A}^{\dagger})^{\mathrm{H}} = \boldsymbol{R}_x + \boldsymbol{A}^{\dagger}\boldsymbol{A}^{\dagger^{\mathrm{H}}}\sigma^2, \tag{2.39}$$

an estimate of $\boldsymbol{R}_y$ is

$$\hat{\boldsymbol{R}}_y = \frac{1}{T_{\max}} \sum_{t=0}^{T_{\max}-1} \boldsymbol{A}^{\dagger}\boldsymbol{N}(t)\boldsymbol{N}^{\mathrm{H}}(t)\boldsymbol{A}^{\dagger^{\mathrm{H}}}. \tag{2.40}$$

With $T_{max} \to \infty$ the estimated correlation matrix $\hat{\boldsymbol{R}}_y$ goes to $\boldsymbol{R}_x + \boldsymbol{A}^{\dagger}\boldsymbol{A}^{\dagger^{\mathrm{H}}}\sigma^2$. $\boldsymbol{R}_x$ can so be estimated subtracting the correlation matrix which belongs to the white Gaussian noise:

$$\hat{\boldsymbol{R}}_x = \hat{\boldsymbol{R}}_y - \boldsymbol{A}^{\dagger}\boldsymbol{A}^{\dagger^{\mathrm{H}}}\sigma^2. \tag{2.41}$$

Using $\hat{\boldsymbol{R}}_x$ and $\hat{\boldsymbol{R}}_y$ the estimated correlation matrix is obtained using

$$\hat{\boldsymbol{R}} = \boldsymbol{A}\hat{\boldsymbol{R}}_x\boldsymbol{A}^{\mathrm{H}} + \sigma^2\boldsymbol{I}. \tag{2.42}$$

2.5 Tested scenario and result

2.5.1 Interference environment

In the simulated scenario a Base Station (BS) is at the center of a quadratic radio cell. The cell of the receiver BS is surrounded with 8 neighborhood cells, in which interferers communicate with their own BS.

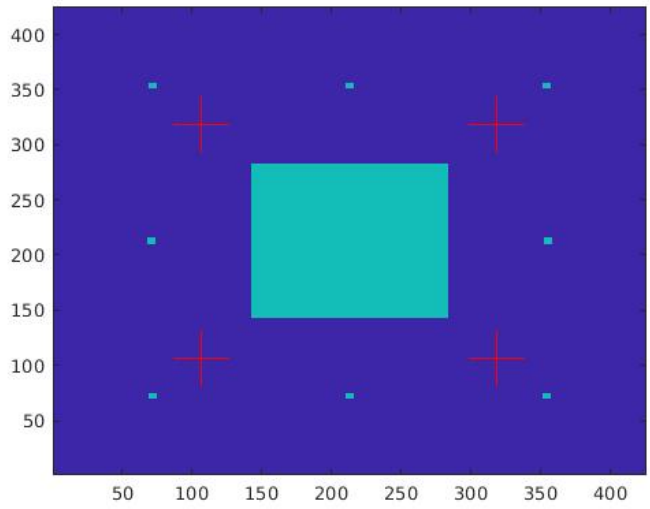

Figure 2.2: Interference environment

The quadratic cells have a diagonal size d_S of 200m. So the maximal distance off a Mobile Device (MD) to his BS is $d_{max} = 100m$. MDs must have a minimum distance of $d_0 = 3m$ from their BS. The reiceiver BS, in the center of the environment, suffers from interference, which is generated by the interferers, located in the neighborhood cells. Signals are generated by the MDs and sent to their BS. The interference, which is received at the receiver BS, arrives with i reflected rays, which in mmWave can be approximated to 3, due to the strong attenuation. The angles of transmission (AoT) at the MDs vary randomly between 0 and 2π and the angles of arrival (AoA) at the receiver BS varies randomly around the real positions of the MDs between $+\frac{\pi}{12}$ and $-\frac{\pi}{12}$. In the simulations the number of interferes $N_{int} = 4$ and all interferers have a distance $d = 149m$ from the BS.

Interference Channel between BS and a single MD

The MD communicate on a carrier frequency $f_c = 28\text{GHz}$, which corresponds to a wavelength of $\lambda = \frac{c}{f} = 10.7\text{mm}$. This is a typical wavelength used in mmWave communications. The amplitude of the echos are complex normally distributed and have the following exponential decreasing amplitude profile:

$$\exp\left(-\delta(i-1)\right). \tag{2.43}$$

where δ assumes the values $0, 1, 2$.

Interference Channel between BS and more MDs

The receiver BS works with N_R receiving antennas and the MDs with N_T transmission antennas. For the simulation these values were set as $N_R = 32$ and $N_T = 1$. To create the interference channel matrix $\boldsymbol{G}$ all interference channel matrices between the BS and the MDs $\boldsymbol{G}_x$, where the index x stays for the considered interferer, are merged together to form

$$\boldsymbol{G} = [\boldsymbol{G}_1, ..., \boldsymbol{G}_{N_{int}}], \tag{2.44}$$

of size $N_R \times N_R$.

Results

In the following scenarios three simulations has been done. In the diferent scenarios, in which the noise varies with respect to the mean power. The mean power of the interference channel $\boldsymbol{G}$ is calculated as

$$\sum_K \boldsymbol{G}\boldsymbol{G}^{\mathrm{H}} = \overline{\sigma_G^2}, \tag{2.45}$$

with K a sufficient high number (ex. 100). The noise power σ_Z^2 was then choosen as $\overline{\sigma_G^2}[dB] - 3dB$, $\overline{\sigma_G^2}$ and $\overline{\sigma_G^2}[dB] + 3dB$. Using the formulas in Section 2.4.5 In the following graphs the correlation matrix $\boldsymbol{R}$ was estimated the following different methods:

1. The LS estimation which is updated every new received $\boldsymbol{N}$ vector.
2. The reference estimation having real AoAs.
3. The gridless AoA-slow method with continuously estimating the AoAs using all previous received $\boldsymbol{N}$ vectors until a value $t_{est} = 10$.

4. The MUSIC method with continuously estimating the AoAs using new arrived $\boldsymbol{N}$ vectors

5. The gridless AoA-short method with continuously estimating the AoAs using new arrived $\boldsymbol{N}$ vectors

6. The gridless AoA-average method which performs an AoA estimation every $t_{est} = 10$ $\boldsymbol{N}$ vectors on $t_{est} = 10$ $\boldsymbol{N}$ vectors

Every graph is averaged by 100 distinct realizations.

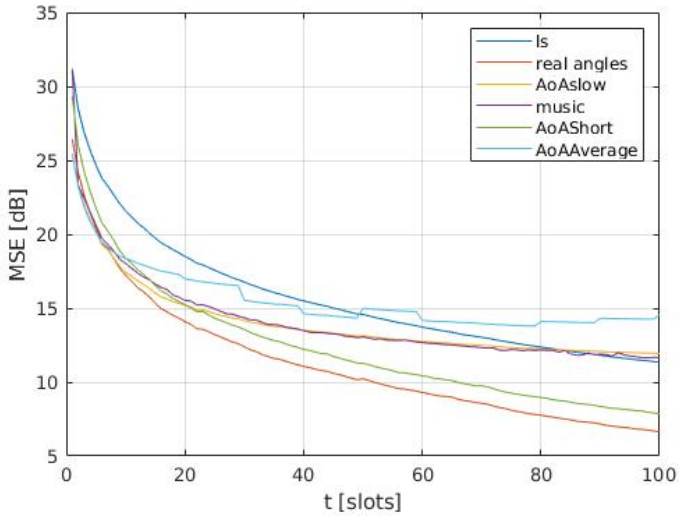

Figure 2.3: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] + 3dB$, $\delta = 0$

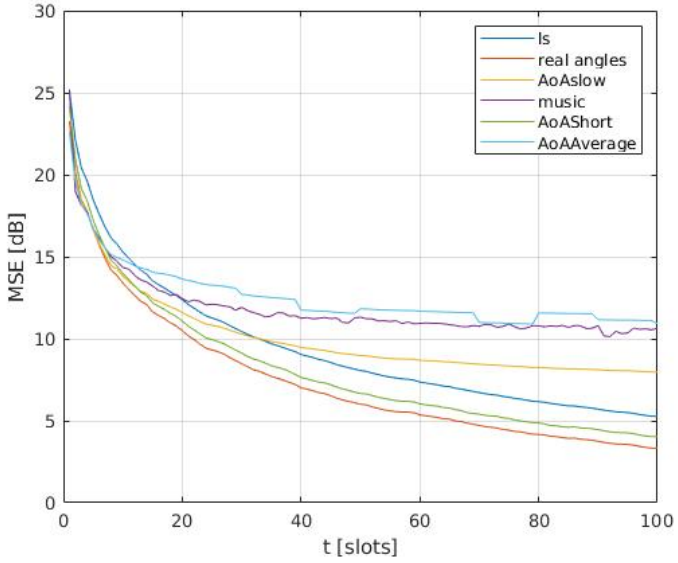

Figure 2.4: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] - 3dB$, $\delta = 0$

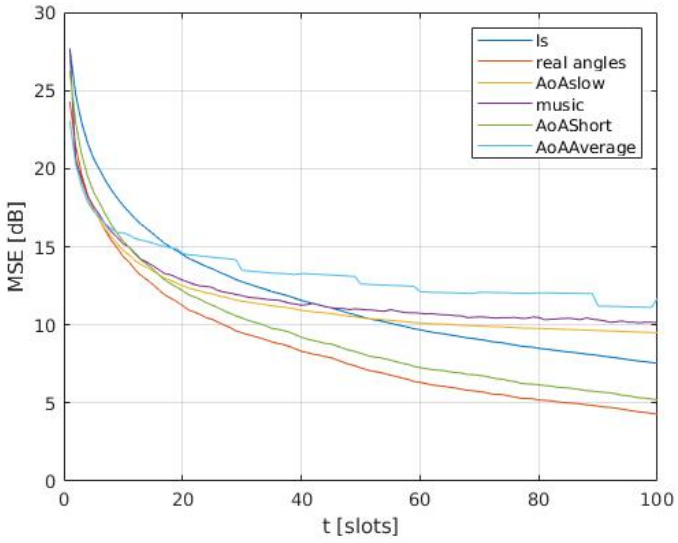

Figure 2.5: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB]$, $\delta = 0$

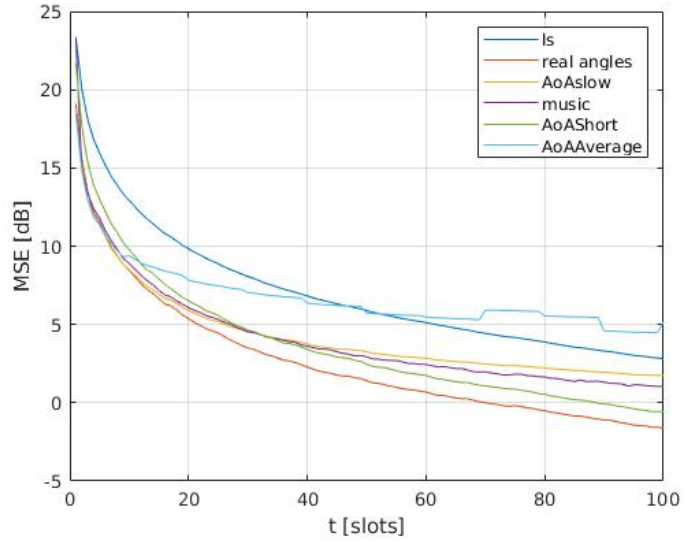

Figure 2.6: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] + 3dB$, $\delta = 1$

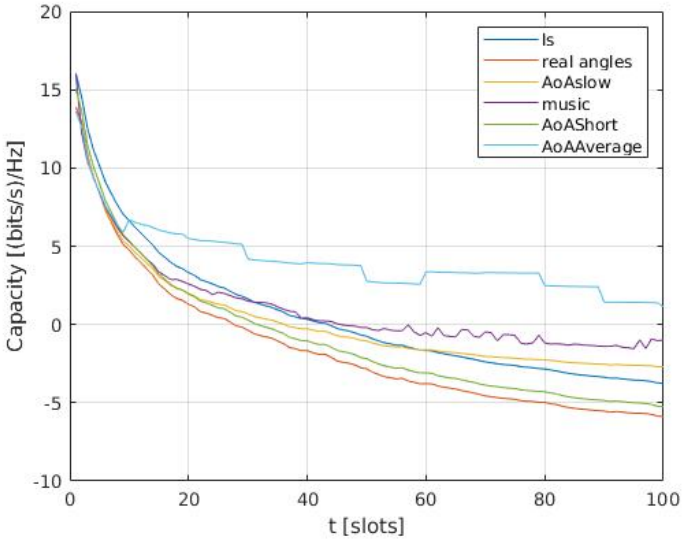

Figure 2.7: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] - 3dB$, $\delta = 1$

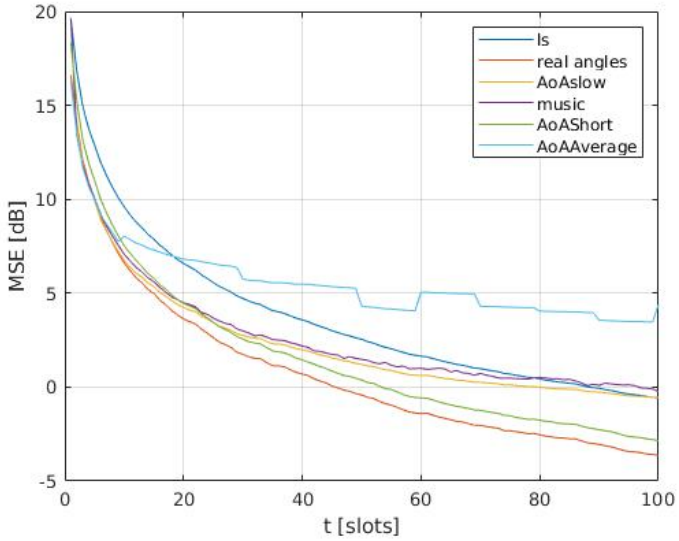

Figure 2.8: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB]$, $\delta = 1$

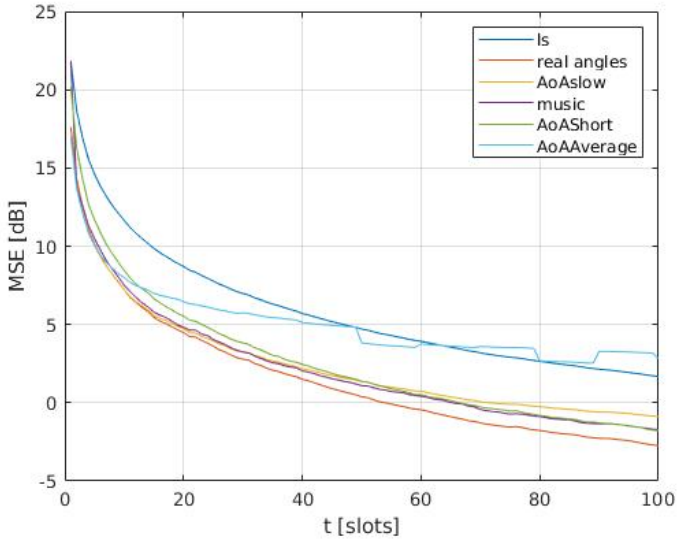

Figure 2.9: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] + 3dB$, $\delta = 2$

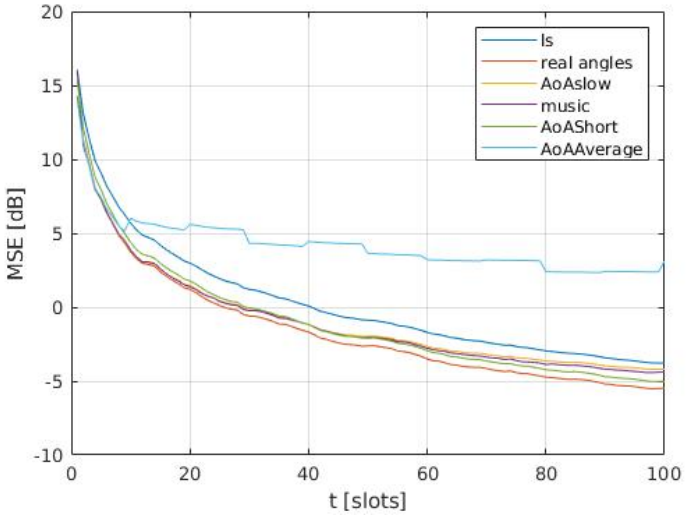

Figure 2.10: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] - 3dB$, $\delta = 2$

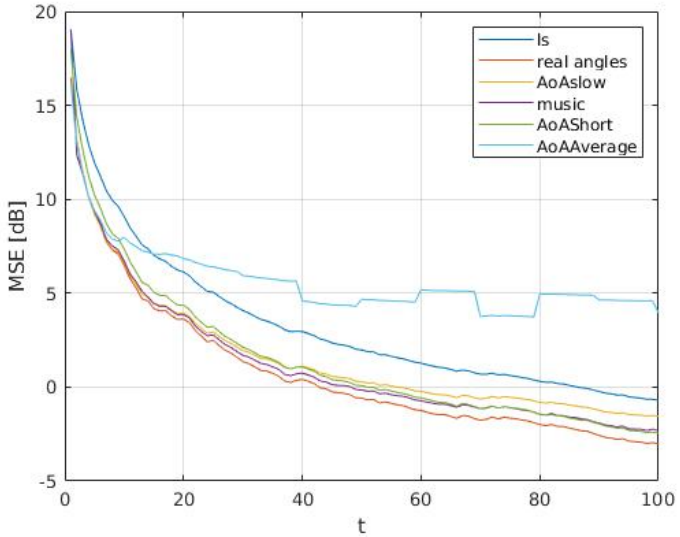

Figure 2.11: MSE[dB] vs t[slots] between $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ with $\sigma^2[dB] = \overline{\sigma_G^2}[dB]$, $\delta = 2$

From the graphs, it can be seen how the advantage increases when the

AoAs are known, especially when the noise power increases. Without noise no MSE reduction would be obtained with respect to the standard LS estimation method. In all graphs the AoA-short method performs best because it derives an advantage from a previous done LS estimation and its AoAs are continuously estimated with every new $\boldsymbol{N}$. Both the MUSIC and the standard AoA-slow methods perform very similar, which is also a very interesting result knowing that AoA-slow continues to estimate the AoAs only until $t = 10$. It seems that the MUSIC method, which performs the AoA estimation on the covariance matrix, reaches a certain limit, hence it does not improve the estimation. The AoA-average method is problematic in all graphs. A possible reason could be the limited number of estimations (with 100 vectors 11 estimations are done). Increasing the amplitude profile from $\delta = 0$ to $\delta = 2$ a general convergence of all methods towards the real angles curve can be observed (with exception of the average method).

2.6 Capacity evaluation

In this simplified scenario an uplink transmission between a BS with N_R receiver antennas and a MD with $N_T = 1$ transmitter antennas is considered. The received signal of the BS, $\boldsymbol{y}$ is determined by (2.1). A whitening filter is applied to the signal to transform the interference and noise to white noise. This is done decomposing the correlation matrix of the noise $\boldsymbol{R} = \mathbb{E}[\boldsymbol{N}\boldsymbol{N}^H]$ using SVD as

$$\boldsymbol{R} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}}. \tag{2.46}$$

The received vector with whitened noise can be obtained as

$$\boldsymbol{Y'} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{Y} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{H}X + \boldsymbol{N'} = \boldsymbol{g}X + \boldsymbol{N'}, \tag{2.47}$$

where $\boldsymbol{g} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{H}$. The obtained vector $\boldsymbol{N'}$ is white noise. The capacity of this upload transmission channel can be obtained using a Maximal Ratio Combiner (MRC) at the receiver. For this purpose $\hat{X} = \boldsymbol{g}^{\mathrm{H}}\boldsymbol{Y'}$ is defined and the SNR of the signal can be obtained decomposing $\hat{X}$ as

$$\hat{X} = \boldsymbol{g}^{\mathrm{H}}\boldsymbol{Y'} = \boldsymbol{g}^{\mathrm{H}}(\boldsymbol{g}X + \boldsymbol{N'}). \tag{2.48}$$

The signal part of $\hat{X}$ is composed out of $\boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}X$. The power of the signal part is obtained taking the power of two of the signal part, i.e.,

$$\mathbb{E}[\boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}X^2] = |\boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}|^2\mathbb{E}[|X|^2] = |\boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}|^2\sigma_x^2. \tag{2.49}$$

The noise part of $\hat{X}$ is composed out of $\boldsymbol{g}^{\mathrm{H}}\boldsymbol{N'}$. The power of the noise part is obtained taking the averaged power of the noise, i.e.,

$$\mathbb{E}[(\boldsymbol{g}^{\mathrm{H}}\boldsymbol{N'})(\boldsymbol{N'}^{\mathrm{H}}\boldsymbol{g})] = \boldsymbol{g}^{\mathrm{H}}\mathbb{E}[\boldsymbol{N'}\boldsymbol{N'}^{\mathrm{H}}]\boldsymbol{g} = \boldsymbol{g}^{\mathrm{H}}\boldsymbol{I}\sigma^2\boldsymbol{g} = \boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}\sigma^2, \tag{2.50}$$

where σ^2 is the power of the noise. The capacity can then be obtained as:

$$C = \log_2\left(1 + \frac{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}\sigma_x^2}{\sigma^2}\right). \tag{2.51}$$

Estimating the noise correlation matrix $\hat{\boldsymbol{R}}$ a reduced capacity can be obtained which depends on the estimation error. We decompose by SVD $\hat{\boldsymbol{R}}_{N'}$ as

$$\hat{\boldsymbol{R}}_{N'} = \hat{\boldsymbol{\Lambda}}^{-1}\hat{\boldsymbol{U}}^{\mathrm{H}}\hat{\boldsymbol{R}}\hat{\boldsymbol{U}}\hat{\boldsymbol{\Lambda}}^{-1}, \tag{2.52}$$

and define

$$\hat{\boldsymbol{g}} = \hat{\boldsymbol{\Lambda}}^{-1/2}\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{H}, \tag{2.53}$$

to obtain

$$\hat{\boldsymbol{Y'}} = \hat{\boldsymbol{\Lambda}}^{-1/2}\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{Y} = \hat{\boldsymbol{g}}X + \hat{\boldsymbol{N}}', \tag{2.54}$$

with $\hat{\boldsymbol{N}}'$ with correlation matrix

$$\boldsymbol{R}_{\hat{N}'} = \hat{\boldsymbol{\Lambda}}^{-1/2}\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{R}\hat{\boldsymbol{U}}\hat{\boldsymbol{\Lambda}}^{-1/2}. \tag{2.55}$$

The capacity is obtained as

$$\hat{C} = \log_2\left(1 + \frac{|\hat{\boldsymbol{g}}^{\mathrm{H}}\hat{\boldsymbol{g}}|^2\sigma_x^2}{\hat{\boldsymbol{g}}^{\mathrm{H}}\boldsymbol{R}_{\hat{N}'}\hat{\boldsymbol{g}}}\right). \tag{2.56}$$

The capacity, using the estimated correlation matrix is $\hat{C} < C$.

2.6.1 Results

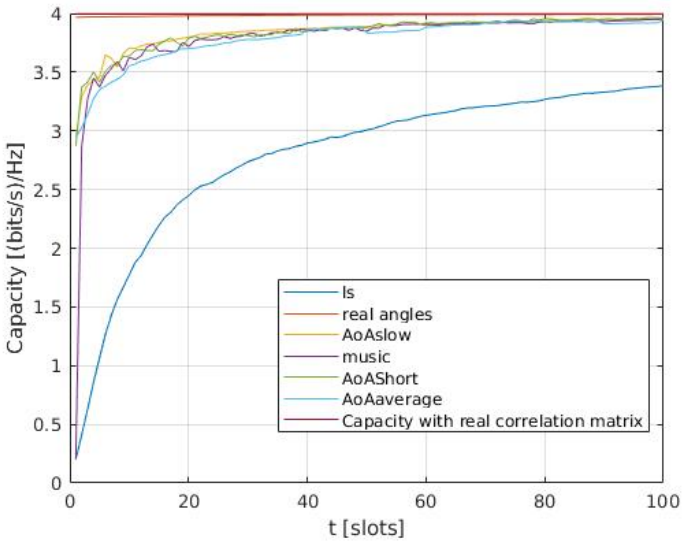

Figure 2.12: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] + 3dB$, $\delta = 0$.

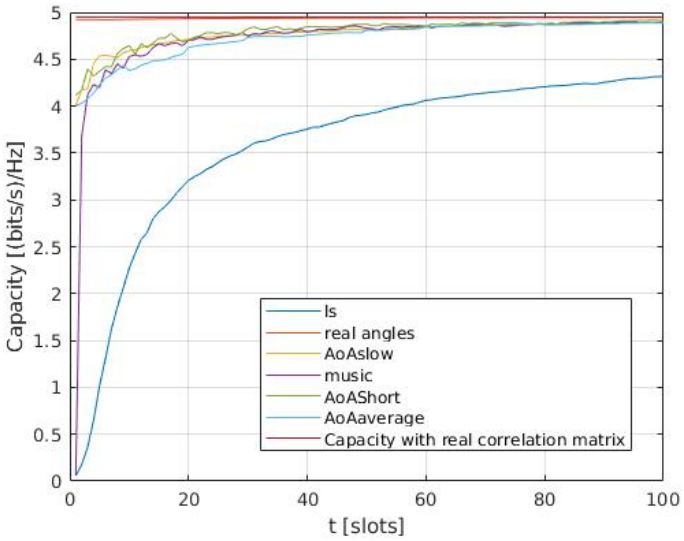

Figure 2.13: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] - 3dB$, $\delta = 0$.

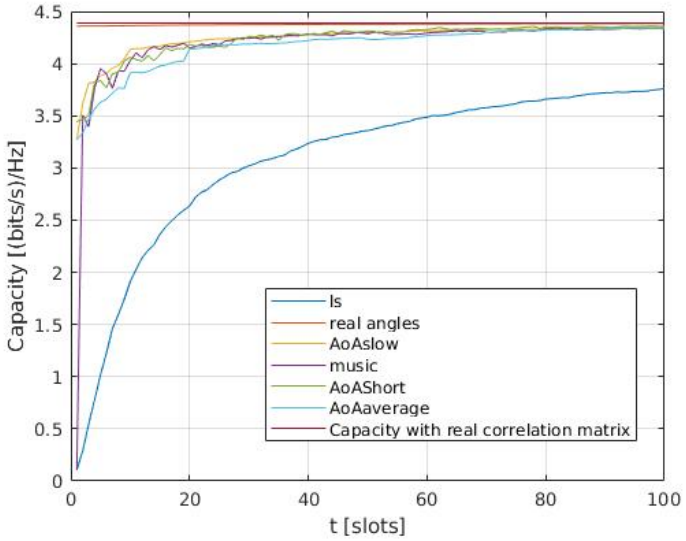

Figure 2.14: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB]$, $\delta = 0$.

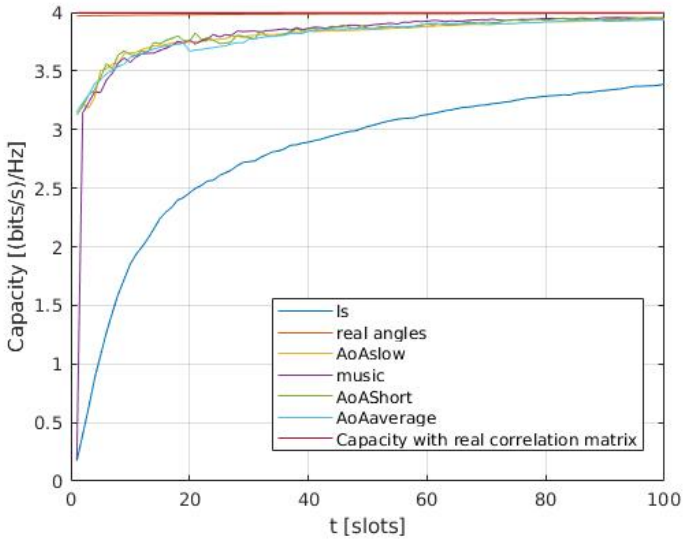

Figure 2.15: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] + 3dB$, $\delta = 1$.

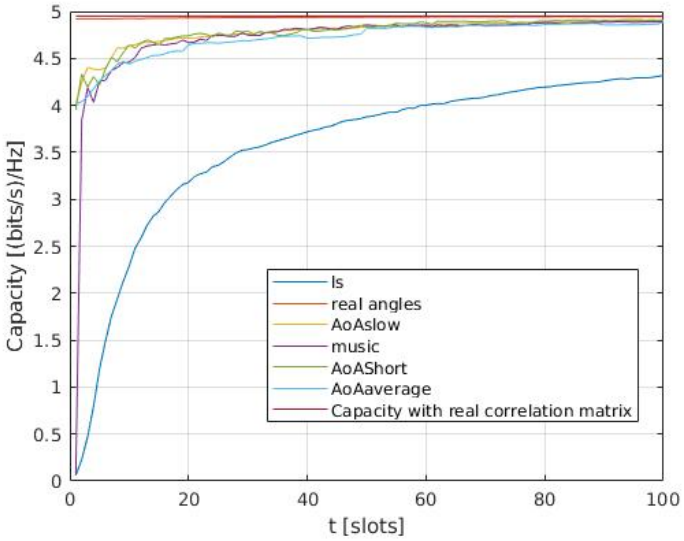

Figure 2.16: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] - 3dB$, $\delta = 1$.

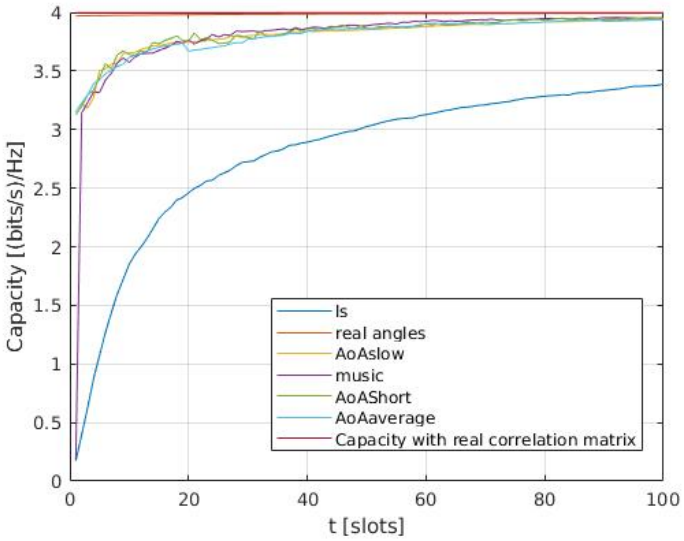

Figure 2.17: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB]$, $\delta = 1$.

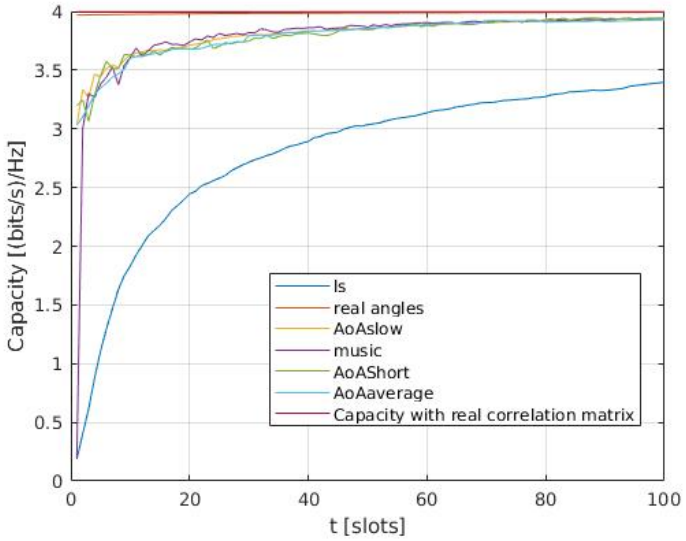

Figure 2.18: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] + 3dB$, $\delta = 2$.

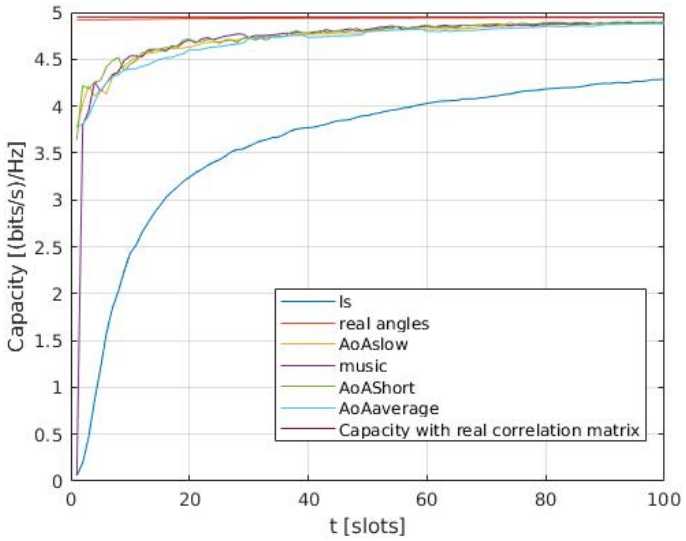

Figure 2.19: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB] - 3dB$, $\delta = 2$.

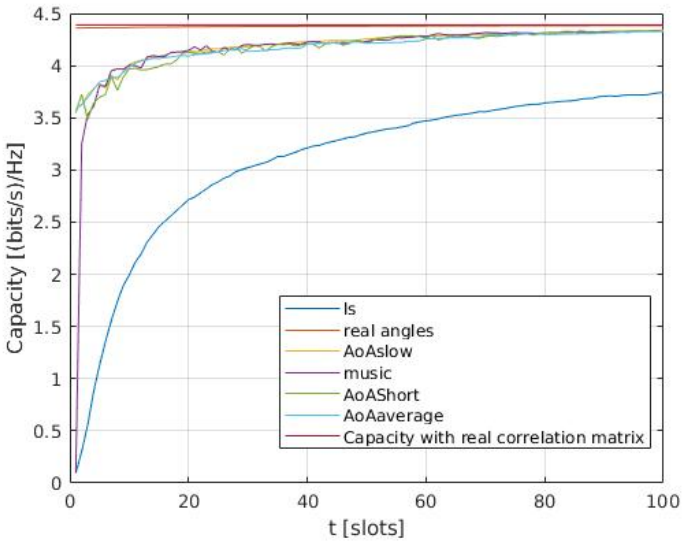

Figure 2.20: Capacity vs t with $\sigma^2[dB] = \overline{\sigma_G^2}[dB]$, $\delta = 2$.

The graphs above are more meaningful for our purpose. The first thing which can be seen is that with larger signal power, with respect to the noise power, the maximal capacity, which could be obtained, increases. The second important observation is that estimating the AoAs gives a clear advantage, in terms of capacity, with respect of estimation methods where these are unknown. So also if in some graphs after $t = 100$ the MSE is more than the MSE of LS, in terms of capacity all methods which estimate the AoAs perform better. The estimations which make use of the gridless AoA estimation method performs similar to the esimation which uses MUSIC. Nevertheless a small advantage in terms of capacity is visible. So with $t = 2$ and also sometimes with $t = 3$ there is an advantage using the Gridless estimation methods. Looking at the amplitude profile the difference between $\delta = 0$ and $\delta = 2$ is very limited. But in the first 10 vectors a slight sharper increase can be observed using $\delta = 0$.

2.7 Orthogonal matching pursuit

From (2.30) it can be observed that the orthogonal matching pursuit(OMP) algorithm on the estimate $\hat{\boldsymbol{R}}$. In fact, we first define $\hat{\boldsymbol{D}}^{(1)} = \text{DFT2}(\hat{\boldsymbol{R}})$, $\hat{\boldsymbol{\Omega}}'^{(1)} = [\hat{\boldsymbol{\Omega}}^{(1)}, \hat{\boldsymbol{\Omega}}^{(2)}]$ such that, for $\ell = 1$

$$(\check{g'}_1, \boldsymbol{\Omega}^{'(1)}) = \underset{\check{g'}_1, \boldsymbol{\Omega}^{'(1)}}{\operatorname{argmin}} ||\hat{\boldsymbol{D}}^{(1)} - \check{g'}_1 \boldsymbol{X}(\Omega^{'(1)})||^2 \tag{2.57}$$

Then we remove its contribution from defining, for $\ell = 1$,

$$\hat{\boldsymbol{D}}^{(\ell+1)} = \mathrm{DFT2}(\hat{\boldsymbol{D}}^{(\ell)} - \check{g'}_\ell \boldsymbol{X}(\Omega^{'(\ell)})) \tag{2.58}$$

The process is repeated for $\ell = 2, 3, 4, ...$ with (2.57) and (2.58) iterated at each step. A stopping criterion can be either a fixed maximum value for ℓ or the step $\ell^{'}$ when $||\hat{\boldsymbol{D}}^{(\ell+1)}|| < \epsilon$ with ϵ a suitable threshold.

Chapter 3

Change detection of the interference correlation

The previous described procedures for the estimation of the interference correlation matrix deals with a static environment, in which the correlation matrix does not change. The main focus was to find methods which monitor the correlation matrix in real time evaluating received vectors every time they are received by the BS. Procedures and algorithms which satisfy this purpose are also used in cybersecurity, health care and the automotive field. The correlation matrix is calculated out of vectors which are originated by a certain number of antennas. So an anomaly/change could only be detected if all parameters are monitored. The change detection must satisfy the requisite of a multivariate anomaly detection system. Another requisite is that the procedures should work sequentially having memory of the previous received information. One of the standard procedures, which satisfy this task is the Cumulative Sum (CUSUM) [7] [8]. In this thesis the standard algorithm was tested on a correlation matrix but it has problems for further use, because it must be aware of the statistics before and after the anomaly event [9]. This could be useful only in ideal static conditions with a fixed number of interferers in fixed positions and known behaviors. There are some CUSUM-like methods which deal with these problems [9] but not in a multivariate sense. Other papers dealt with this problem using kNN based solutions [12]. These methods work well considering the multivariate property and the unknown type of statistics of the data, but have problems with temporal anomalys. [10] proposed a method with a temporal window based kNN detection method in which the time is split in windows and tested when the statistics exceeds a certain threshold. Also these methods are not

fast in detecting changes. [11] proposed a method based on kNN, which resolved the problem of real time detection, called Online Discrepancy Test (ODIT). In the paper the CUSUM procedure was tested on a correlation matrix and the ODIT procedure analyzed. Both methods could be, as a further challenge beyond this thesis, tested on the correlation matrices of the previous section.

3.1 Notation

The interference has a statistic which changes in time, in particular with the correlation matrix. The change of the correlation matrix happens because of the On/Off behavior of the UDs, which follow the scheduling of their cells. Let $\boldsymbol{R}(s)$ be the correlation matrix at time s (slot duration). Let us suppose that L correlation matrices occur, $\boldsymbol{R}^{(\ell)}$, with $\ell = 1, ..., L$.
The change detection procedure supposes that $\boldsymbol{R}(s) = \boldsymbol{R}^{(\ell)}$ and tries to understand, in slot $\tau > s$, if the matrix changes ($\boldsymbol{R}(\tau) \neq \boldsymbol{R}^{(\ell)}$). Furthermore, the algorithm wants to reveal the new correlation matrix for a $\ell \neq \ell'$ so that $\boldsymbol{R}(\tau) = \boldsymbol{R}^{(\ell')}$ with the estimation time for the change τ. To understand if a change has happen, the noise vector $\boldsymbol{N}(s)$ is analyzed if it is a Gaussian vector with mean value $\mathbf{0}$ and correlation matrix $\boldsymbol{R}^{(\ell)}$ or not. In the first part only scenarios with $L = 2$ are considered, which means that only two possible correlation matrices are considered.

3.2 CUSUM

The multivariate cumulative sum (CUSUM) is a procedure to detect changes in the statistics on a sequence of observed vectors, $\boldsymbol{N}(s)$ in our case. Of these vectors a certain number m are part of a distribution function F_a and a certain number of a distribution function F_b.

$$\begin{cases} \boldsymbol{N}(1), ..., \boldsymbol{N}(m) \in & F_a \\ \boldsymbol{N}(m+1), ..., N(\infty) \in & F_b \end{cases} \tag{3.1}$$

A sum value S_t, is defined, which is updated at each received noise vector $\boldsymbol{N}$ recursively, having $S_0 = 0$ and

$$S_t = \max\left(S_{t-1} + \log\frac{f_a(\boldsymbol{N}(n))}{f_b(\boldsymbol{N}(n))}, 0\right), \tag{3.2}$$

where f_a and f_b are the probability density functions which correspond to the distribution functions F_a and F_b. The stopping vector, which determines a change in the distribution is

$$T_{stop} = \inf(t : S_t \geq h_c), \tag{3.3}$$

where h_c is a preset level which, if it is exceeded, reveals a change of the correlation matrix. Note that in order to apply this method we need the statistics F_a and F_b, i.e., we need F_b known in advance $\boldsymbol{R}^{(1)}$ and $\boldsymbol{R}^{(2)}$.

3.2.1 Tested scenarios and results

1. Convergence test:
   In the first scenario the CUSUM change detector is not used. A fictitious change is introduced every $n_{\text{change}} = 50$ vectors and and the following matrices are used to compose the covariance matrices:

$$\begin{cases} \boldsymbol{\Sigma}_1 = \text{diag}(1, ..., 1) \in \mathbb{C}^{64\times 64} \\ \boldsymbol{\Sigma}_2 = \text{diag}(2, ..., 2) \in \mathbb{C}^{64\times 64} \\ \boldsymbol{D} = \text{Discrete Fourier Transform Matrix} \in \mathbb{C}^{64\times 64} \end{cases} \tag{3.4}$$

   The received vectors $\boldsymbol{N}$ have a Gaussian distribution with mean value 0 and covariance matrices

$$\begin{aligned} \boldsymbol{R}_1 &= \boldsymbol{D}\boldsymbol{\Sigma}_1\boldsymbol{D}^{\mathrm{T}}, \\ \boldsymbol{R}_2 &= \boldsymbol{D}\boldsymbol{\Sigma}_2\boldsymbol{D}^{\mathrm{T}}. \end{aligned} \tag{3.5}$$

   Furthermore the level h_c is set to 200. This experiment is only a preliminary test to show the convergence between the estimated covariance matrices $\hat{\boldsymbol{R}}$ and the real covariance matrices $\boldsymbol{R}$. To measure the convergence the mean square error (MSE) is used

$$\text{MSE} = \frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k}([\boldsymbol{R}]_{i,j} - [\hat{\boldsymbol{R}}]_{i,j})^2, \tag{3.6}$$

   where k is set equal to 64 considering the size of the covariance matrices.

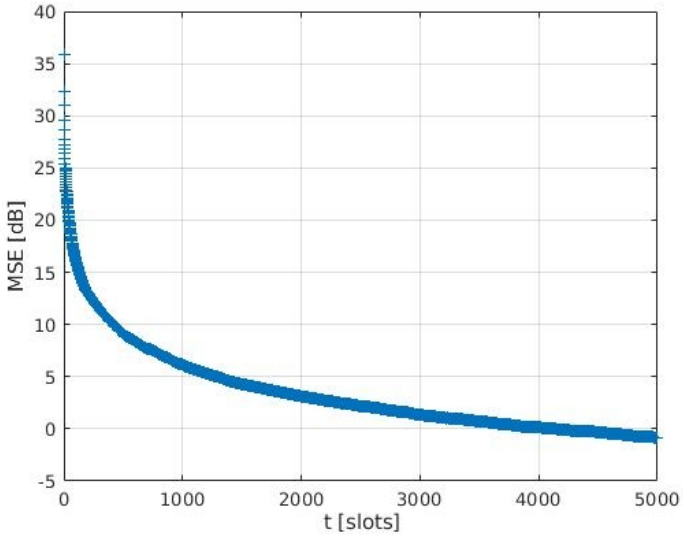

Figure 3.1: MSE vs T of between $\boldsymbol{R}_1$ and $\hat{\boldsymbol{R}}_1$

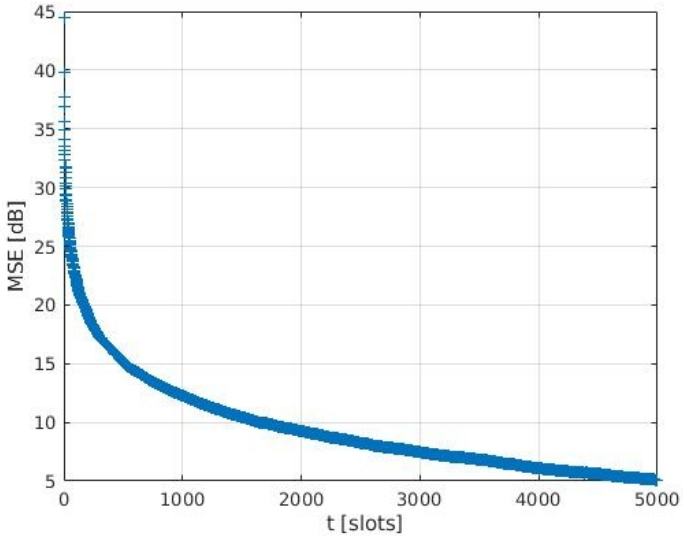

Figure 3.2: MSE vs T of between $\boldsymbol{R}_2$ and $\hat{\boldsymbol{R}}_2$

In Fig.3.1 and Fig.3.2 the convergence of the covariance matrices can

be seen when 10000 vectors are received. Due to the perfect conditions all the vectors $\boldsymbol{N}$ participate in the estimation of the correct covariance matrix. Both figures show also that with an increasing number of associated vectors the MSE decreases.

2. Perfect change detector:
   In the second scenario the CUSUM change detector is introduced. The covariance matrices $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ and the parameters h_c are maintained as in scenario 1 and n_{change} is set to 500. In this scenario the change detector is aware of the real correlation matrices $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$.

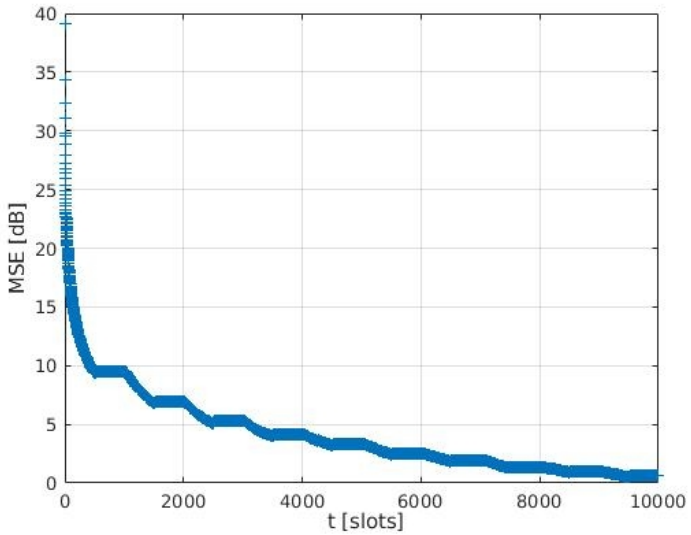

Figure 3.3: MSE vs T of $\hat{\boldsymbol{R}}_1$ and $\hat{\boldsymbol{R}}_2$ compared with $\boldsymbol{R}_1$

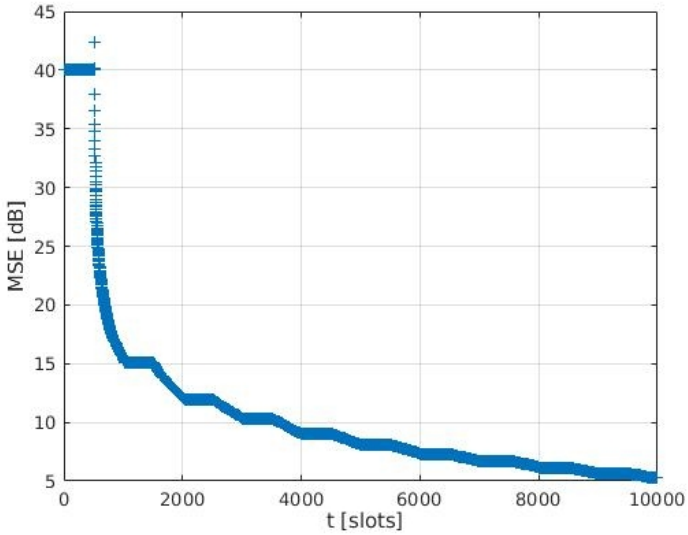

Figure 3.4: MSE vs T of $\hat{\boldsymbol{R}}_1$ and $\hat{\boldsymbol{R}}_2$ compared with $\boldsymbol{R}_1$

Fig.3.3 and Fig.3.4 show that three different behaviors can be observed:

(a) The MSE decreases when the new realizations of the received vectors $\boldsymbol{N}$ are associated to the correct $\boldsymbol{R}$ and participate so in the convergence of this matrix towards the real one.

(b) The MSE remains stable if the received vectors $\boldsymbol{N}$ are not used to estimate the selected correlation matrix.

(c) The MSE increases if the change detection algorithm does not reveal the change or it is revealed with a certain delay and the vectors $\boldsymbol{N}$ are used to estimate the wrong correlation matrix.

The change detector, in this setting, is able to detect very precisely the change. This can be seen in the figures as the MSE never increase significantly.

3. In the third scenario the CUSUM change detector is still used but the detector is not aware of the real correlation matrices $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$. The change detector has to use the estimated matrices $\hat{\boldsymbol{R}}_1$ and $\hat{\boldsymbol{R}}_2$ to detect a change in the distribution. In this scenario all parameters are the same as in scenario 1 whit exception of $n_{change} = 2000$.

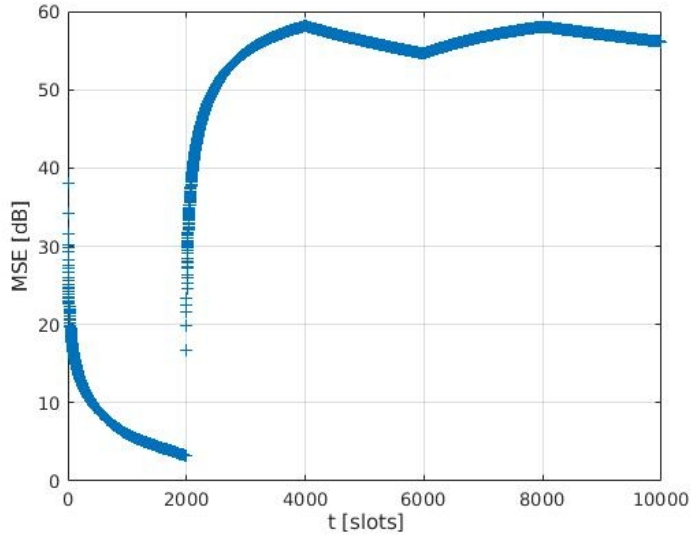

Figure 3.5: MSE vs T of $\hat{\boldsymbol{R}}_1$ and $\hat{\boldsymbol{R}}_2$ compared with $\boldsymbol{R}_1$

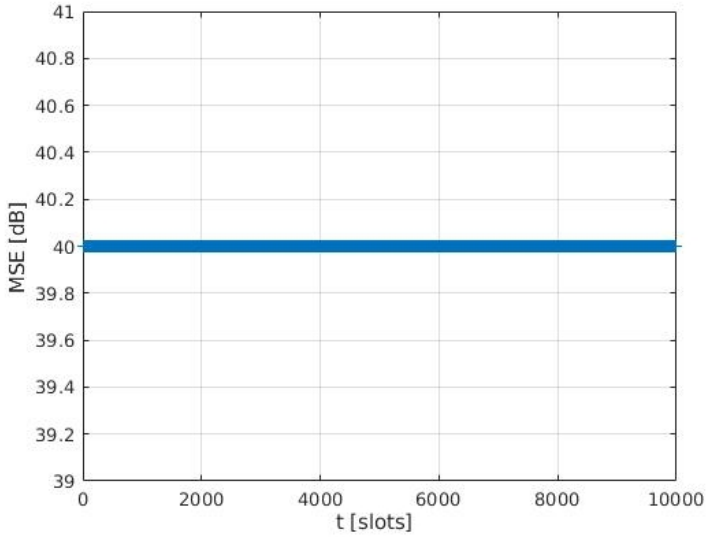

Figure 3.6: MSE vs T of $\hat{\boldsymbol{R}}_1$ and $\hat{\boldsymbol{R}}_2$ compared with $\boldsymbol{R}_2$

In figure 3.5 the first covariance matrix converges but at the change

point the change is not detected. This is due to the initial conditions. Even with a good estimation of the first covariance matrix $\hat{\boldsymbol{R}}_1$ (MSE $\approx 4dB$) after 2000 noise vectors $\boldsymbol{N}$, the second estimation matrix $\hat{\boldsymbol{R}}_2$ is not known until that point. The second initial estimated correlation matrix $\hat{\boldsymbol{R}}_2$ can be set very similar or very different with respect to the first estimation matrix. But both settings are problematic:

- If the two estimated matrices $\hat{\boldsymbol{R}}_1$ and $\hat{\boldsymbol{R}}_2$ are too similar a change detection oscillation is observable with a lot of false alarms and consequently an update of the wrong covariance matrix.
- If the two estimated matrices $\hat{\boldsymbol{R}}_1$ and $\hat{\boldsymbol{R}}_2$ are too different the changes are not detected (3.5 and 3.6). The MSE of the first estimated matrix $\hat{\boldsymbol{R}}_1$ continues to decrease and successively to increase. The second estimated covariance matrix $\hat{\boldsymbol{R}}_2$ is never estimated.

3.3 Online discrepancy test (ODIT)

A possible algorithm, which takes vectors as input for the change detection procedure, is the ODIT procedure proposed by [8].
A training set of noise vectors $\boldsymbol{\mathcal{N}}_C$ of size C is partitioned into two sets $\boldsymbol{\mathcal{N}}_{C_1}$ and $\boldsymbol{\mathcal{N}}_{C_2}$ with $C = C_1 + C_2$.
This training set consists in a nominal data-set which is used to learn the statistical properties of the system in the starting condition. This means that ODIT trains only on the initial conditions and compares then sequentially the observed noise vectors $\boldsymbol{N}$ with the training data, without having a training on the anomalous data.
The Euclidean distances between the points $\boldsymbol{N}(m) \in \boldsymbol{\mathcal{N}}_{C_1}$ and its k nearest neighbors in $\boldsymbol{\mathcal{N}}_{C_2}$ are calculated. We define the total k-nearest neighbor(kNN) distance of $\boldsymbol{N}(m)$ as

$$L(m) = \sum_{n=k-s+1}^{k} g_n^{\gamma}(\boldsymbol{N}(m)), \tag{3.7}$$

where $(g_n(\boldsymbol{N}(m)))$ is the Euclidean distance between point $\boldsymbol{N}(m) \in \boldsymbol{\mathcal{N}}_{C_1}$ and its nth nearest neighbors in $\boldsymbol{\mathcal{N}}_{C_2}$, $s \in \{1, ..., k\}$ is a fixed number, $\gamma > 0$ a weight introduced for flexibility and $\alpha \in (0, 1)$ is a significance level. The training phase is finished by choosing the $(1-\alpha)$th percentile of total k-NN distances $\{L(m)\}$.

The algorithm selects the Kth smallest distance $L(K)$ with $K = \lfloor C_1(1-\alpha) \rfloor$ as a baseline statistic for measuring the deviation of new observations from the nominal data-set in the test phase.

In the test phase, ODIT computes the total kNN distance $L(t)$ for each observation $\boldsymbol{N}(t)$ with respect to the nominal points in $\boldsymbol{\mathcal{N}}_{C_2}$. This provides the anomaly evidence

$$D(t) = N_I(\log L(t) - \log L(K)), \tag{3.8}$$

with N_I the dimension of the data, which in our case is the number of interference antennas. At this point, $D(t)$ shows a positive or negative evidence for statistical change/anomaly. A positive $D(t)$ for example shows that the observations lies outside the estimated most compact set of the nominal training set, which proves a positive evidence for a change.
The algorithm updates the detection statistic $\Delta(t)$ over time. This continues until the first time $\Delta(t)$ exceeds the predefined threshold h. This follows the following update and decision rule:

$$\begin{aligned} \Delta(t) = \max\{\Delta(t-1) + D(t), 0\}, \quad \Delta(0) = 0, \\ T_{stop} = \min\{t : \Delta(t) \geq h\}. \end{aligned} \tag{3.9}$$

The size of C_1 and C_2 should be so that C_2 is larger than C_1, because C_2 determines the accuracy of (3.8) and C_1 the significance of the accuracy parameter α.
ODIT is part of the non parametric techniques for multivariate change detection. This means that compared to CUSUM, which is a parametric technique, no specific probability distributions are assumed.

3.3.1 Tested scenarios and results

The algorithm was tested using multivariate vectors $\boldsymbol{N} \in \mathbb{R}^{1 \times N_I}$ choosing for $N_I = 64$. The vectors have a correlation matrix $\boldsymbol{R}_1 = (\text{eye}(64) + 1)$ and a mean value of 0. At a certain received vector $N(t)$ with $t = 50$ the statistics of the vector changed and they were produced with a correlation matrix $\boldsymbol{R}_2 = (\text{eye}(64) + 1)500$. The training phase was done using $C = 100$ $\boldsymbol{N}$ vectors, dividing the set in $C_1 = 0.2C$ and $C_2 = C - C1$. The training set so happened without having vectors from the anomalous data set. α was chosen as 0.05 and $s = 1$, following the suggestions in [8]. In the following

graph a cumulative distribution function (CDF) was done testing in 400 rounds when the change was detected.

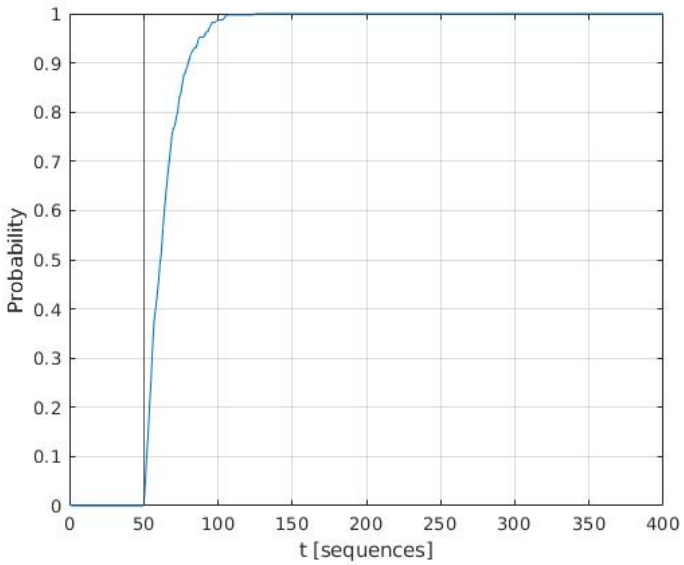

Figure 3.7: Cumulative Distribution Function with change point t=50

In figure 3.7 the false alarm probability was 0% and after around 21 $\boldsymbol{N}$ vectors 80% of the changes were detected. The correlation matrices $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ were very different and so the change of the correlation matrix was detected fast.

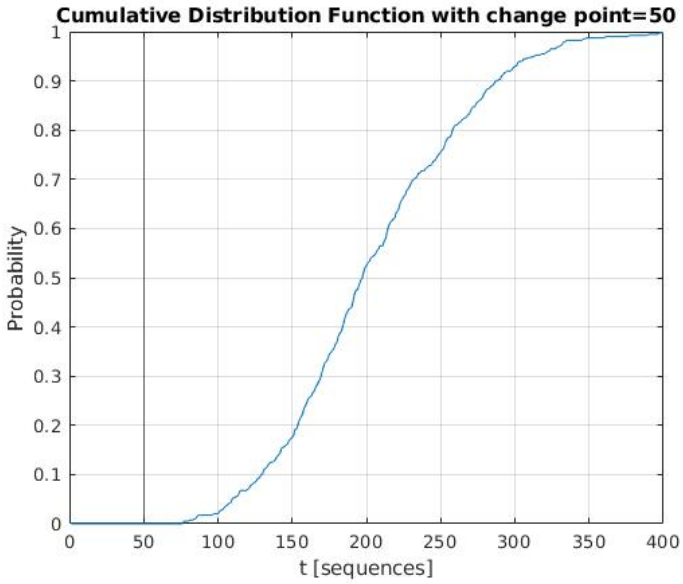

Figure 3.8: Cumulative Distribution Function with change point t=50

Contrary in figure 3.8 when $\boldsymbol{R}_2 = (\boldsymbol{R}_1 + 1)10$ the result is worse and around 209 vectors are needed to capture 80% of the changes.

3.3.2 Conclusion

ODIT could work well on correlation matrices determined in the previous chapter. Only some modifications shoud be done which allows to the algorithm to deal with complex vectors instead only with real vectors. But the fact that is is capable to do real time change detection on multivariate vectors without knowing the statistics of the anomalous data is perfect for the mmWave environment, in which the interference correlation matrix $\boldsymbol{R}$ changes fast and without a known behavior, because of the ON/OFF behavior of the UE and their mobility.

Chapter 4

Conclusion

About the interference correlation matrix estimation, the obtained results show a clear advantage of estimating the AoAs with respect to methods which are not aware of the AoAs. Also a little advantage in terms of capacity of the new Gridless AoA estimation methods can be seen. Nevertheless further analysis should be done improving the scenarios and trying to simulate a more realistic environment with more interferers and in particular interferers with active/inactive behaviours and mobility in the environment (fast and slow). Maybe the advantage of gridless AoAs estimation could be emphasized as it is developed for automotive and radar solutions. About the change detection, a further step could be then the simulation of the ODIT algorithm to detect behavioral changes of the interferers and to react against them, maintaining the capacity stable. ODIT satisfies all requirements for this, but due to time reasons and technical problems (implementation of ODIT with complex vectors) it could not be implemented.

Bibliography

[1] A. Brighente, M. Cerutti, M. Nicoli, S. Tomasin and U. Spagnolini,"Estimation of Wideband Dynamic mmWave and THz Channels for 5G Systems and Beyond," *IEEE Journal on Selected Areas in Communications*, vol. 38, n. 9, pp. 2026-2040, Sept. 2020.

[2] G. Soatti, A. Murtada, M. Nicoli, J. Gambini and U. Spagnolini,"Low-Rank Channel and Interference Estimation in mm- Wave Massive Antenna Arrays," *2018 26th European Signal Processing Conference (EUSIPCO)*, Sept. 2018.

[3] J. Deng, O. Tirkkonen, R. Freij-Hollanti, T. Chen, N. Nikaein,"Resource Allocation and Interference Management for Opportunistic Relaying in Integrated mmWave/sub-6 GHz 5G Networks," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 94-101, June 2017.

[4] M. Gallo, S. Tomasin,"Stima della statistica di interferenza per sistemi di telecomunicazione mm-wave," *Università di Padova - Dipartimento di Ingegneria dell'Informazione*, Nov. 2020.

[5] R. Schmidt,"Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276-280, March 1986.

[6] M. Sanchez-Fernandez,V. Jamali,J. Llorca and A. Tulino,"Gridless Multidimensional Angle-of-Arrival Estimation for Arbitrary 3D Antenna Arrays," *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4748-4764, July 2021.

[7] A. Vega Delgado, M. Sanchez-Fernandez, L. Venturino and A. Tulino,"Super-Resolution in Automotive Pulse Radars," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 91-926, June 2021.

[8] M. Mozaffari, Y. Yilmaz,"Online Multivariate Anomaly Detection and Localization for High-dimensional Settings," *arXiv*, July 2020.

[9] E. Page,"Continuous Inspection Schemes," *Biometrika*, vol. 41, n. 1/2, pp. 100-115, 1954

[10] G. Moustakides,"Optimal Stopping Times for Detecting Changes in Distributions," *Ann. Statist.*, vol. 14, n. 4, Dec. 1986.

[11] Y. Mei,"Efficient scalable schemes for monitoring a large number of data streams," *Biometrika.*, vol. 97, n. 2, pp. 419-433, June 2010.

[12] H. Chen,"Sequential change-point detection based on nearest neighbors," *Ann. Statist.*, vol. 47, n. 3, pp. 1381-1407, June 2019.

[13] A. Hero,"Geometric entropy minimization (GEM) for anomaly detection and localization,"*Advances in Neural Information Processing Systems 19 (NIPS 2006)*, 2006

[14] M. Mohanna, M. L. Rabeh, E. M. Zieur and S. Hekala,"Optimization of MUSIC algorithm for angle of arrival estimation in wireless communications,"*NRIAG Journal of Astronomy and Geophysics*, pp.116-124, 2013