Contributions à l'estimation des paramètres du canal MIMO

Ali Mohydeen

UNIVERSITE BRETAGNE LOIRE MATHSTIC

UNIVERSITÉ DE NANTES

THESE DE DOCTORAT DE

L'UNIVERSITE DE NANTES

COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Électronique*

Par

Ali MOHYDEEN

Contributions to MIMO channel parameter estimation

Thèse présentée et soutenue à Nantes, le 12 Septembre 2019
Unité de recherche : IETR UMR CNRS 6164

Rapporteurs avant soutenance :

Salah BOURENNANE	Professeur, Ecole Centrale de Marseille
Yannis POUSSET	Professeur, Université de Poitiers

Composition du Jury :

Président :	Camel TANOUGAST	Professeur, Université de Lorraine
Examinateurs :	Salah BOURENNANE	Professeur, Ecole Centrale de Marseille
	Yannis POUSSET	Professeur, Université de Poitiers
Directeur de thèse :	Pascal CHARGE	Professeur, Université de Nantes
Co-directeurs de thèse :	Yide WANG	Professeur, Université de Nantes
	Oussama BAZZI	Professeur, Université Libanaise, Beyrouth

Acknowledgments

First of all, I would like to thank my Ph.D supervisor, Professor Pascal Chargé, for his guidance, support and encouragement. I am so grateful for all the valuable and fruitful discussions we have had throughout the Ph.D period. I am also grateful to my co-supervisors, Professor Yide Wang and Professor Oussama Bazzi, for their support and encouragement.

I would like to thank my committee members, Professor Salah Bourennane, Professor Yannis Pousset, and Professor Camel Tanougast, for their valuable and encouraging feedback.

I would also like to thank the administrative and technical staff of the IETR Laboratory for their administrative and technical support during the Ph.D period.

I am thankful to my colleagues in the lab for creating a friendly working environment and for many helpful discussions.

Finally, my special thanks go to my family, to my loving parents, for their absolute and unconditional support. I dedicate this thesis to them.

Contents

List of Figures

List of Abbreviations

4G	Fourth Generation
5G	Fifth Generation
AIC	Akaike Information Criterion
AOA	Angle of Arrival
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
CIR	Channel Impulse Response
CS	Compressive Sensing
CSI	Channel State Information
CSIR	Channel State Information at the Receiver
CSIT	Channel State Information at the Transmitter
DFT	Discrete Fourier Transform
EGC	Equal Gain Combining
EM	ElectroMagnetic
ESPRIT	Estimation of Signal Parameter via Rotational Invariance
GSCM	Geometry-based Stochastic Channel Model
ISI	Inter-Symbol Interference
LS	Least Squares
MDL	Minimum Description Length
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output

ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MMV	Multiple Measurement Vector
MMW	Millimeter Wave
MP	Matching Pursuit
MPC	Multipath Component
MRC	Maximum Ratio Combining
MUSIC	MUltiple SIgnal Classification
MVDR	Minimum Variance Distortionless Response
OFDM	Orthogonal Frequency Division Multiplexing
OMP	Orthogonal Matching Pursuit
PDP	Power Delay Profile
RIP	Restricted Isometry Property
RMS	Root Mean Square
RMSE	Root Mean Square Error
SC	Selection Combining
SIC	Successive Interference Cancellation
SIMO	Single-Input Multiple-Output
SISO	Single-Input Single-Output
SMV	Single Measurement Vector
SNR	Signal-to-Noise Ratio
SOMP	Simultaneous Orthogonal Matching Pursuit
SVD	Singular Value Decomposition
TOA	Time of Arrival
US	Uncorrelated Scattering
UWB	UltraWideBand
VCR	Virtual Channel Representation
WSS	Wide Sense Stationary

WSSUS	Wide Sense Stationary Uncorrelated Scattering
ZF	Zero Forcing

Notations

x	Scalar
$\mathbf{x}$	Vector
$\mathbf{X}$	Matrix
$[.]^T$	Transpose
$[.]^*$	Conjugate
$[.]^H$	Conjugate transpose
$[.]^{-1}$	Inverse
$[.]^\dagger$	Pseudo inverse
E[.]	Expectation
$\|.\|$	Absolute value
$\|\|.\|\|_p$	l_p norm
$\mathbb{C}$	Set of complex numbers
$diag\{x_1 \dots x_n\}$	$n \times n$ diagonal matrix with the elements $x_1 \dots x_n$ on its diagonal
$\mathbf{I}_n$	$n \times n$ identity matrix
$det[\mathbf{X}]$	Determinant of $\mathbf{X}$
$\odot$	Hadamard product
$\otimes$	Kronecker product

General Introduction

In the ancient world, light and flags were used as a way of wireless communication. In 1867, James Clerk Maxwell predicted the existence of electromagnetic (EM) waves, proposing an interrelation between electric and magnetic fields. In 1887, Heinrich Rudolf Hertz confirmed the existence of EM waves traveling at the speed of light by performing experiments in his laboratory. The waves he produced and received are now called radio waves. A breakthrough came with Guglielmo Marconi who developed the wireless telegraph in 1895. Since then, he succeeded in transmitting radio signals through space, increasing the distance of communication gradually. In 1901, he established the first wireless communication across the ocean, by transmitting radio signals across the Atlantic ocean. From then until today, different wireless technologies have been developed, including radio and television broadcasting, radar communications, satellite communications, wireless networking, mobile wireless communications, etc.

With the rise of big data era, along with the increasing demand of wireless data services, the major goal of researchers over the years has been to support high data rates to satisfy needs. A major obstacle to build a reliable high speed wireless communication system is the wireless propagation medium. In wireless communication, the signal propagating through the wireless channel is exposed to different types of fading, especially multipath fading due to multipath propagation [1]. This impacts the reliability of the communication link and limits the data rate.

Multiple-input multiple-output (MIMO) technology has become an active research topic during the last decade due to its capability for achieving the high transmission rates required by an increasing number of data-demanding applications. MIMO technology provides a plenty of benefits that allow dealing with the challenges posed by the impairments in the wireless channel, especially multipath fading. It provides several important performance gains such as array gain, diversity gain, and multiplexing gain.

The benefits of MIMO are achieved through the exploitation of the spatial dimension across multiple antennas at the transmitter and receiver, in addition to the time and frequency dimensions already exploited in the conventional single-input single-output (SISO) systems [2].

MIMO, along with orthogonal frequency division multiplexing (OFDM), are key technologies used in 4G (fourth generation) wireless networks. MIMO is a key technology for the next fifth generation (5G) wireless networks that employ massive antenna arrays and millimeter wave (MMW) frequencies.

Knowledge of the wireless propagation channel characteristics is crucial for the reliability of wireless communications, especially in MIMO communications in order to fully benefit from the advantages provided by using multiple antennas at the transmitter and receiver sides. This information about the channel characteristics is referred to as channel state information (CSI). CSI represents the information about the signal propagation from the transmitter to the receiver, it represents wireless channel effects such as power attenuation and time spreading of signals. CSI plays an important role in the system performance, channel state information at the receiver (CSIR) can be used for equalization purpose against intersymbol interference (ISI) caused by multipath propagation, and channel state information at the transmitter (CSIT) can be used for optimal transmission design. Hence a performant "ideal" MIMO communication system would require an exact knowledge of the MIMO channel or CSI. CSI estimation approaches can be classified into parametric and nonparametric. In the nonparametric approach, the channel matrix is estimated directly without referring to any underlying physical propagation parameters. On the other hand, the parametric approach relies on physical channel models to estimate channel parameters, such parameters are useful for understanding the wireless channel, and can be utilized to improve the communication system performance by adapting the transmission and reception designs according to them.

Clustering due to scatterers is an important property that characterizes several wireless channels, where according to different wireless channel investigations, channel multipath components are modeled as clusters of multirays. For instance, this phenomenon (clustering) characterizes wideband/ ultrawideband (UWB) and MMW communication channels. These new channel characteristics have to be considered in the future channel estimation techniques.

The work in this thesis focuses on clustered MIMO channel parameter estimation,

specifically, time domain parameters. This thesis is divided into five chapters. The first part of the first chapter represents the basic wireless channel propagation characteristics, especially, fading in wireless channels. Fading channel is classified into large scale fading and small scale fading, large scale fading characterizes the channel behavior over large distances and incorporates path loss and shadowing. Small scale fading characterizes the channel behavior over short time periods or travel distances, and is further classified into two categories based on multipath delay spread and doppler spread. Based on multipath delay spread, fading is classified into flat fading and frequency selective fading. Based on doppler spread, fading is classified into slow fading and fast fading. In the second part, we introduce diversity techniques in wirless communication systems, and we focus on spatial diversity, by showing the benefits brought by using multiple transmit or/and receive antennas.

In the second chapter, we provide an overview of different MIMO channel models. The channel models are classified into physical and non-physical models. Physical models are further classified into deterministic and stochastic models, where stochastic models are classified into geometry-based models, nongeometrical models, and propagation based analytical models. Then we introduce sparsity and clustering properties in wireless channels, in addition to the common support property in outdoor MIMO channels. Finally, we introduce a sparse clustered MIMO channel model with common support, on which our proposed estimation methods are based.

Chapter 3 deals with channel parameter estimation. Estimation approaches are classified into parametric and nonparametric. The nonparametric approach refers to MIMO channel matrix estimation (MIMO channel estimation) while the parametric approach refers to channel parameters estimation. We classify methods that can be used for channel parameter estimation into 3 categories: beamforming methods, subspace based methods and compressive sensing methods.

Chapter 4 illustrates two proposed channel mean path (cluster) delays estimation approaches. The first approach is based on the first order Taylor expansion around the mean delay parameter, where a compressive sensing based method is proposed to estimate the channel mean delays. And the second approach is based on higher order Taylor expansion around the mean delay parameter, where a method for estimating the channel mean delays is then proposed based on the subspace approach, and on the tracking of the effective dimension of the signal subspace. The proposed methods are validated through computer simulations and estimation performance is illustrated.

In chapter 5, we propose to estimate the standard deviation of delay spreading within each cluster. A stochastic channel modeling approach is proposed, where the statistical distribution of multiray delays within each cluster is exploited. A subspace based method that allows to estimate both mean delay and standard deviation of delay spreading is then derived. An efficient approach is proposed to estimate the standard deviation of delay spreading based on the derived subspace method, and using the mean delays estimated by the method proposed in chapter 4. The proposed method is validated through computer simulations.

Chapter 1

Basic wireless channel propagation characteristics and introduction to multiple antenna systems

1.1 Introduction

Establishing a reliable wireless communication system requires deep understanding of wireless channel propagation models and characteristics. Several factors are involved in the process of determining the channel behavior: signal/channel bandwidth, environment or propagation medium, noise, etc. Due to these factors, a signal transmitted through the environment exhibits fluctuation and attenuation in its level at the receiver side. The phenomenon of fluctuation of the attenuation of the signal level, is referred to as "fading". In wireless channels, a signal emitted from a transmit antenna arrives at the receive antenna through multiple paths (Figure 1.1) with different amplitudes, phase shifts, and delays due to the reflection, diffraction or scattering of electromagnetic waves in the environment. This results in a constructive and destructive interference of signals from the different paths, leading to the fluctuation of the received signal level. This phenomenon is the so-called "multipath fading", and it has a significant impact on the reliability and performance of wireless communication.
Mainly, two types of fading characterize a wireless channel: large scale fading and small scale fading [1, 3–7]. Large scale fading refers to the signal power attenuation and fluctuation due to path loss and shadowing. Path loss refers to power loss due to the propagation over large distances. Shadowing is when signals are interrupted or blocked by large objects such as mountains and buildings over the propagation path between the transmitter and the receiver. Shadowing results in relatively slow fluctuation in

the signal level, its effect depends on the dimension of objects in the environment with respect to the wavelength or the radio frequency of the electromagnetic waves. Three basic phenomena, characterize wireless signal propagation: reflection, diffraction, and scattering.

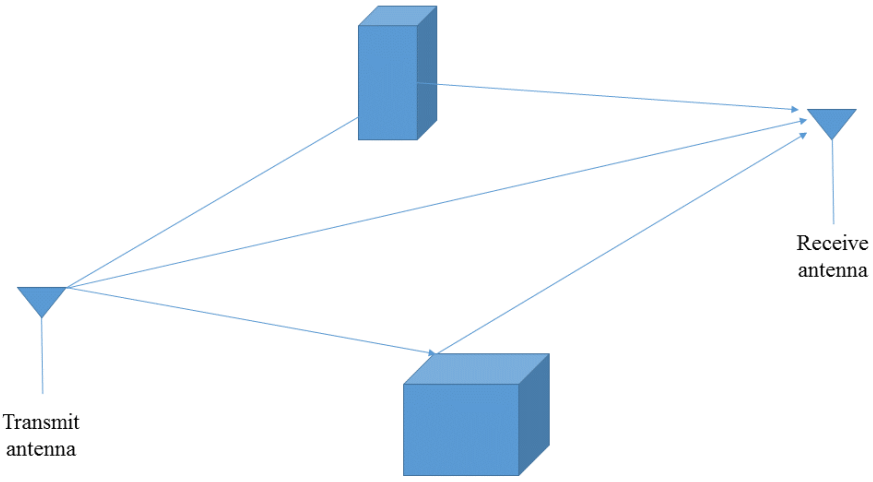

Figure 1.1: Multipath propagation

The free-space path loss is expressed as:

$$\frac{P_r}{P_t} = G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2 \tag{1.1}$$

where P_t and P_r are the transmitted and received signal powers respectively, G_t and G_r are the transmit and receive antenna gains respectively, λ is the signal bandwidth, and d is the distance between the transmit and receive antennas.

Small scale fading refers to the rapid fluctuation of the signal over short periods or short travel distances. The small scale fading scheme can be divided into two categories. The first one is related to the multipath delay spread, and the second is related to the Doppler spread. Within each category, depending on the relation between the different signal and channel parameters, different signals exhibit different types of fading. Figure 1.2 illustrates the different classifications of a wireless fading channel.

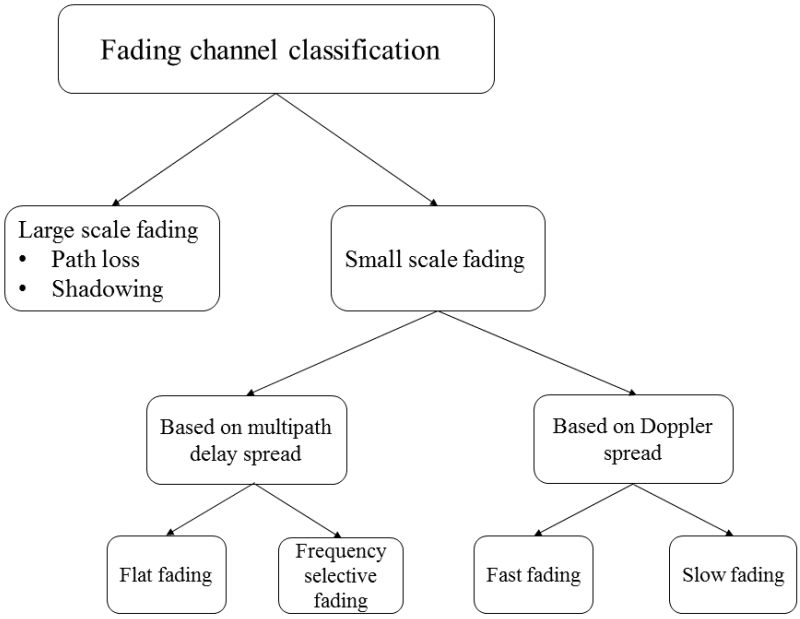

Figure 1.2: Fading channel classification

For small scale fading based on the multipath time delay spread, we have two types of fading, flat fading and frequency selective fading.

1.2 Flat fading

Consider a complex baseband signal modulated by a carrier frequency f_c and trasmitted through a wireless multipath channel that consists of L_0 propagation paths. Let $s(t)$ be the passband signal representation defined as:

$$s(t) = Re\{s_b(t)e^{j2\pi f_c t}\} \tag{1.2}$$

At the receiver, the corresponding received signal at a given instant t is given as

$$r(t) = Re\{\sum_{l=1}^{L_0} \gamma_l s_b(t-\tau_l) e^{j2\pi f_c(t-\tau_l)}\}$$
$$= Re\{(\sum_{l=1}^{L_0} \gamma_l s_b(t-\tau_l) e^{-j2\pi f_c \tau_l}) e^{j2\pi f_c t}\} \quad (1.3)$$

where $s_b(t)$ is the transmitted baseband signal, γ_l is the attenuation of path l and τ_l is the corresponding time delay.

The complex baseband received signal is given as:

$$r_b(t) = \sum_{l=1}^{L_0} \gamma_l s_b(t-\tau_l) e^{-j2\pi f_c \tau_l} \quad (1.4)$$

Assuming that we are dealing with narrowband signals, the narrowband assumption states that $s_b(t-\tau_l) \approx s_b(t)$ for all l. This is related to the limited time resolution due to signal's narrow bandwith in comparison with the channel coherence bandwidth. Limited time resolution means that signals from different paths cannot be distinguishable, as this time resolution is inversely proportional to the signal bandwidth. According to this, $r_b(t)$ is given as :

$$r_b(t) = s_b(t) \sum_{l=1}^{L_0} \gamma_l e^{-j2\pi f_c \tau_l} \quad (1.5)$$

Hence

$$h = \sum_{l=1}^{L_0} \gamma_l e^{-j2\pi f_c \tau_l} \quad (1.6)$$

is the complex channel coefficient. Depending on the values of the time delays τ_l, the different propagation paths can add up constructively or destructively causing fading. Constructive interference amplifies the signal amplitude at the receiver, while destructive interference attenuates the signal amplitude. An example of this phenomenon is illustrated in Figure 1.3, where it is assumed that at each time instant, the receive antenna (for example mobile user) is changing its position, where the power of the

received fading signal versus time is shown. For each time instant, the channel coefficient h is given as the constructive or destructive sum of a large number of multipath components.

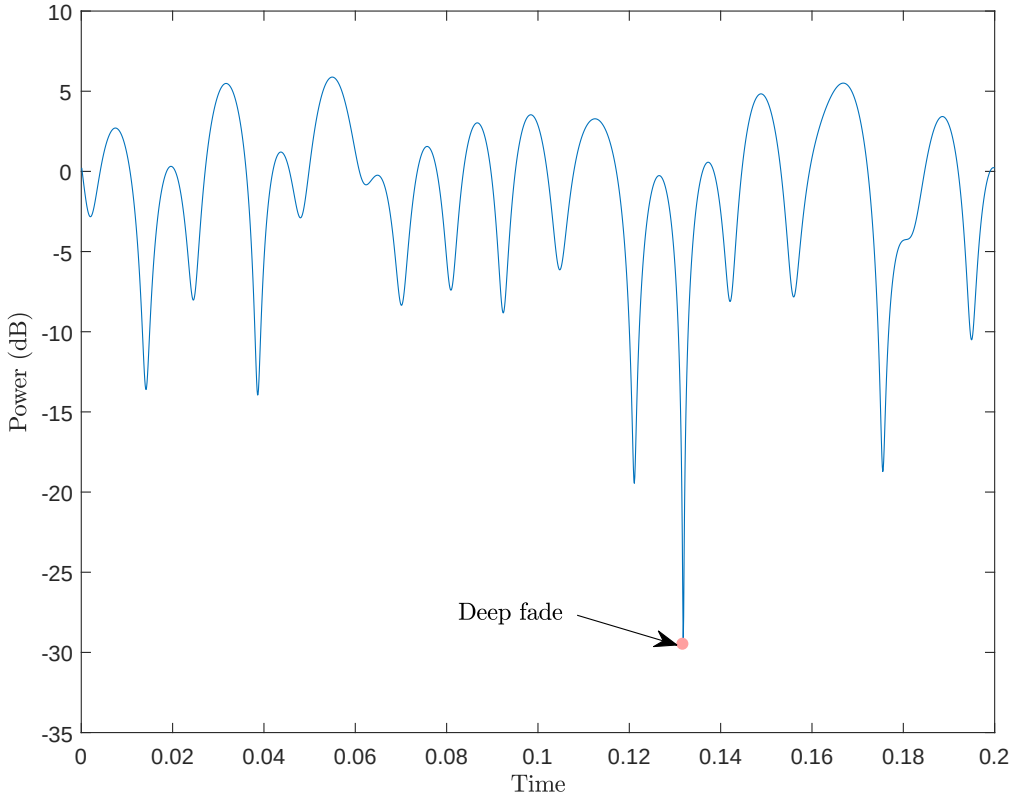

Figure 1.3: Received signal power versus time, multipath fading channel

As we can see, for the case of flat fading, at a given instant t, the channel is seen as a single coefficient (h), in other words, it is not described by its different propagation path delays, the channel impulse response is seen as a single impulse. Due to this, the channel over the given signal bandwidth looks flat in the frequency domain, that is, all the frequency components of the signal will experience the same level of fading.

Regarding the statistics of the fading coefficient h, let

$$\begin{aligned} h &= \sum_{l=1}^{L_0} \gamma_l e^{-j2\pi f_c \tau_l} \\ &= \sum_{l=1}^{L_0} \gamma_l cos(2\pi f_c \tau_l) - j \sum_{l=1}^{L_0} \gamma_l sin(2\pi f_c \tau_l) \\ &= X_r + jX_i \end{aligned} \tag{1.7}$$

where $X_r = \sum_{l=1}^{L_0} \gamma_l cos(2\pi f_c \tau_l)$ and $X_i = -\sum_{l=1}^{L_0} \gamma_l sin(2\pi f_c \tau_l)$.

Each of X_r and X_i is the sum of a large number of random components, hence according to the central limit theorem, X_r and X_i can be modeled as Gaussian random variables. Now let

$$h = X_r + jX_i = \gamma e^{j\phi} \tag{1.8}$$

where $\gamma = \sqrt{X_r^2 + X_i^2}$ is the magnitude and $\phi = \tan^{-1}(\frac{X_i}{X_r})$ is the phase.

Assuming that X_r and X_i are zero mean independent random variables with same variance σ_X^2. It follows that γ is Rayleigh distributed with the following probability density function

$$f_\gamma(\gamma) = \frac{\gamma}{\sigma_X^2} e^{\frac{-\gamma^2}{2\sigma_X^2}} \tag{1.9}$$

and ϕ is uniformly distributed over $[-\pi, \pi]$.

The Rayleigh probability density function is shown in Figure 1.4 for different σ_X^2.

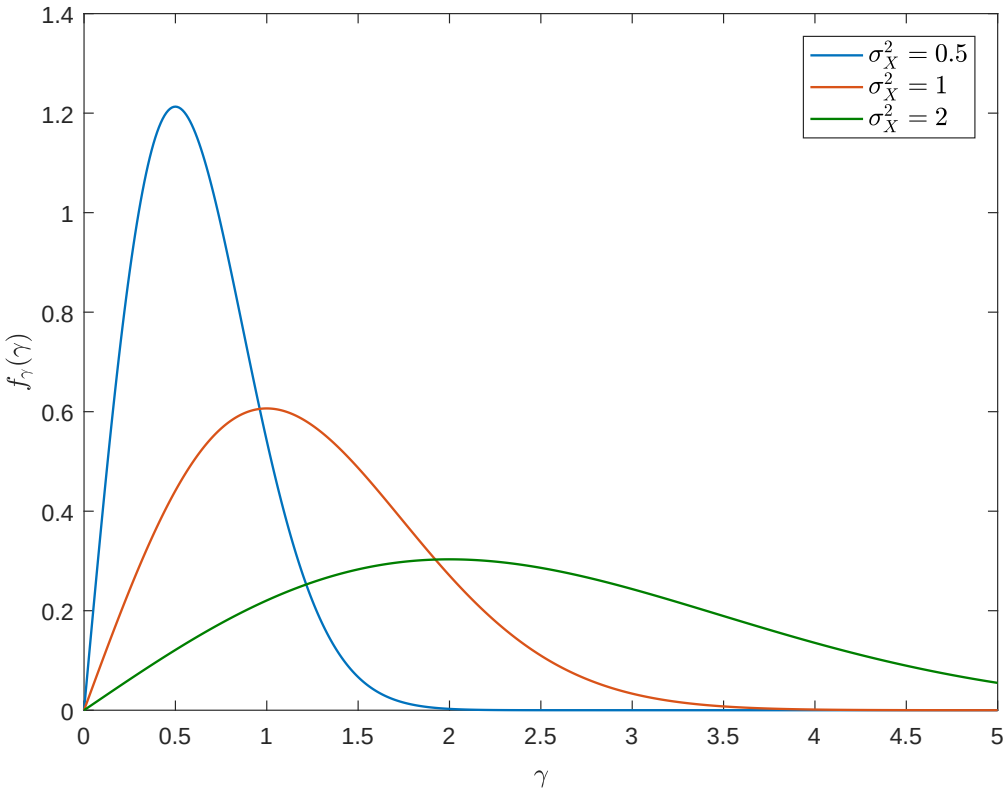

Figure 1.4: Rayleigh probability density function

The Rayleigh fading is reasonable when there is no line of sight path between the transmitter and receiver. When the line of sight path exists, Rician fading becomes a more reasonable model.

1.3 Frequency selective fading

Frequency selective fading refers to the case of wideband signals, where due to the relatively large bandwidth in comparison with channel coherence bandwidth, some path delays can now be resolvable.

The time varying response of a multipath channel is given as:

$$h(t,\tau) = \sum_{l=1}^{L} \alpha_l(t)\delta(\tau - \tau_l(t)) \tag{1.10}$$

If the channel is time invariant (static), the channel impulse response (CIR) is given as:

$$h(\tau) = \sum_{l=1}^{L} \alpha_l \delta(\tau - \tau_l) \tag{1.11}$$

where L is the number of resolvable path delays, τ_l is the time delay of path l and α_l is the corresponding complex gain coefficient (including phase parameter).

The number of resolvable multipaths depends on the available time resolution (or signal bandwidth). Consider Figure 1.5 that illustrates the signals scattering from different boundaries (ellipses).

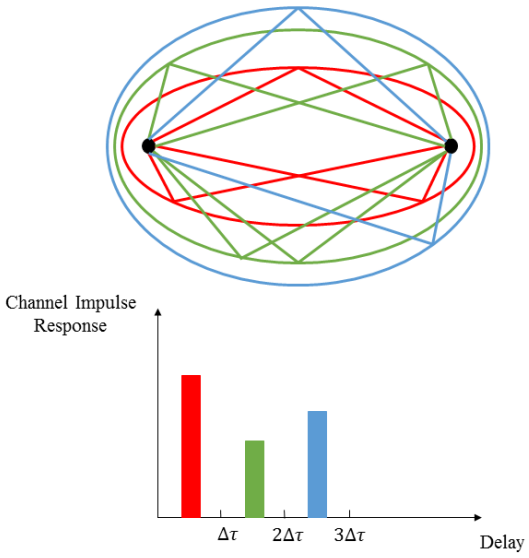

Figure 1.5: Resolvability of channel multipath components

Let $\Delta\tau$ be the available time resolvability, which is inversely proportional to the signal bandwidth. Due to the first boundary (ellipse), all multipath delays are arriving within 0 to $\Delta\tau$, according to the available time resolution, these are unresolvable, and hence seen as one path, similarly due to the second boundary (ellipse), multipath delays fall in the interval from $\Delta\tau$ to $2\Delta\tau$, then according to the available time resolution,

these are seen as one another path, and so on. Now if a sufficient number of unresolvable multipath components falls within the intervals $([0, \Delta\tau], [\Delta\tau, 2\Delta\tau], ...)$, then the envelope of each channel gain coefficient $|\alpha_l|$ within these intervals can be Rayleigh distributed (assuming no line of sight) with uniformly distributed phase.

1.3.1 Power delay profile

One of the important characteristics of frequency selective channels is the power delay profile (PDP). PDP is given as the measure of the received signal power as a function of the channel propagation delays, it is defined as:

$$P_{dp}(\tau) = R_{hh}(\tau) = E[|\alpha_l|^2]\delta(\tau - \tau_l) \tag{1.12}$$

where $E[|\alpha_l|^2]$ is the power associated to path delay τ_l, $R_{hh}(\tau)$ is the channel autocorrelation function.

The main parameters characterizing the PDP are the mean delay, the root mean square (RMS) delay spread, and the maximum excess delay (or maximum delay spread). The mean delay is given as:

$$\bar{\tau} = \frac{\sum_{l=1}^{L} \tau_l P_{dp}(\tau_l)}{\sum_{l=1}^{L} P_{dp}(\tau_l)} \tag{1.13}$$

The RMS delay spread is given as:

$$\sigma_\tau = \sqrt{\bar{\tau^2} - \bar{\tau}^2} \tag{1.14}$$

where $\bar{\tau^2} = \frac{\sum_{l=1}^{L} \tau_l^2 P_{dp}(\tau_l)}{\sum_{l=1}^{L} P_{dp}(\tau_l)}$, and the maximum excess delay is given as:

$$\tau_{max} = \tau_L - \tau_1 \tag{1.15}$$

If the channel coefficients at any two distinct delays τ_1 and τ_2 are uncorrelated, such that the channel autocorrelation function at these two delays is given as:

$$R_{hh}(\tau_1, \tau_2) = R_{hh}(\tau_1)\delta(\tau_1 - \tau_2) \tag{1.16}$$

then the channel is referred to as uncorrelated scattering (US) channel model.

1.3.2 Frequency domain characteristics

The channel can also be considered and characterized in the frequency domain, through the Fourier transform of the channel impulse response, as follows:

$$H(f) = \int_{-\infty}^{+\infty} h(\tau) e^{-j2\pi f\tau} d\tau \tag{1.17}$$

Similarly to what is done in the delay domain, the statistical properties of $H(f)$ can be obtained in the frequency domain, through the autocorrelation function. The correlation between frequency domain channel components at two distinct frequencies f_1 and f_2 is given as:

$$\begin{aligned} R_{HH}(f_1, f_2) &= E[H(f_1)H^*(f_2)] \\ &= E[\int_{-\infty}^{+\infty} h(\tau) e^{-j2\pi f_1\tau} d\tau \int_{-\infty}^{+\infty} h^*(\tau) e^{j2\pi f_2\tau} d\tau] \\ &= \int_{-\infty}^{+\infty} R_{hh}(\tau) e^{-j2\pi(f_1-f_2)\tau} d\tau \end{aligned} \tag{1.18}$$

Let $\Delta f = f_1 - f_2$, we have

$$R_{HH}(\Delta f) = \int_{-\infty}^{+\infty} R_{hh}(\tau) e^{-j2\pi\Delta f\tau} d\tau \tag{1.19}$$

which is the Fourier transform of the channel PDP.

$R_{HH}(\Delta f)$ gives information about the range of frequencies over which the fading pattern is highly correlated. This frequency range, over which the channel is considered flat, is the coherence bandwidth (B_c), that is, the frequency components of the signal over this range, will experience the same fading pattern.

B_c is a key parameter to characterize a wireless channel. Its value compared to the signal bandwith, provide information about the type of fading experienced. When the bandwidth of the transmitted signal (B_s) is less than B_c, the resulting phenomenon is flat fading, while when $B_s > B_c$, the fading is frequency selective.

The value of B_c is inversely proportional to the channel delay spread σ_τ, where generally, the relation is given as $B_c \approx 1/\sigma_\tau$. However there is no one universal definition. There are two common definitions that are based on the least correlation value assigned for the frequency correlation function $R_{HH}(\Delta f)$. If the coherence bandwidth is defined

as the frequency range over which $R_{HH}(\Delta f)/R_{HH}(0)$ is at least 0.9, then it is given as [3]:

$$B_c = \frac{1}{50\sigma_\tau} \tag{1.20}$$

However, if the coherence bandwidth is defined as the frequency range over which $R_{HH}(\Delta f)/R_{HH}(0)$ is at least 0.5, it is given as:

$$B_c = \frac{1}{5\sigma_\tau} \tag{1.21}$$

The main impairment introduced in the case of frequency selective fading is inter-symbol interference, that is, when the signal bandwidth is greater than the channel coherence B_c or in other words, when the symbol period ($T_s \approx 1/B_s$) is significantly less than the channel delay spread σ_τ, multiple copies of symbols (with different gains) arrive at different time delays causing different symbols to interfere with each other, resulting in signal distortion. This is not the case for flat fading, where the symbol period T_s is significantly greater than the channel delay spread σ_τ such that different multipath components arrive almost at the same time and the channel is modeled as a single complex coefficient.

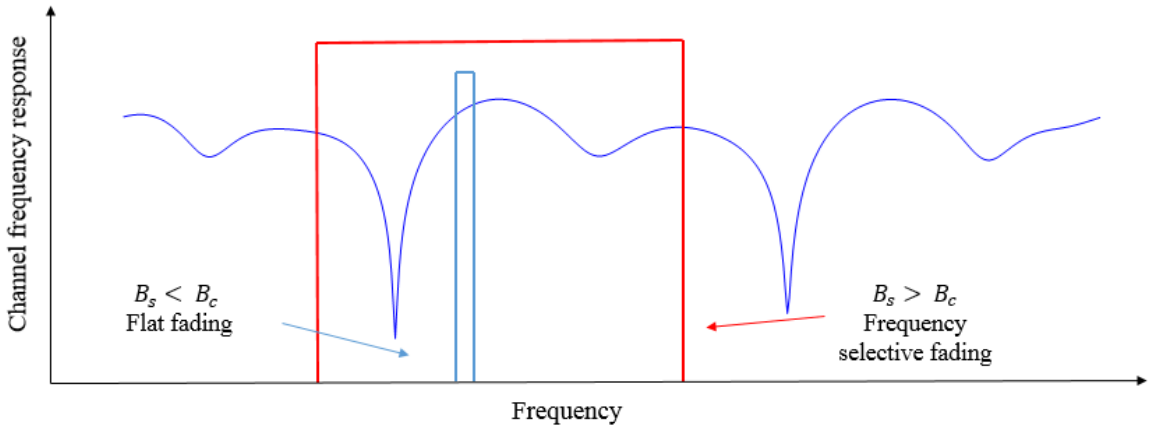

Figure 1.6: Flat fading vs Frequency selective fading

Figure 1.6 illustrates the flat fading and the frequency selective fading phenomena. As shown, for the case of flat fading ($B_s < B_c$), signal's frequency components undergo highly correlated fading pattern, while for frequency selective fading ($B_s > B_c$), signal's frequency components undergo different fading patterns, causing signal distortion.

1.4 Doppler effect

Doppler effect (or Doppler shift/spread) refers to the change in the frequency of the propagating electromagnetic wave due to the relative motion between the transmitter and the receiver, or due to the mobility of scattering objects in the environment.

Consider a mobile user (considered as transmitter) traveling through a base station (receiver) at a speed ϑ. Due to the doppler shift, the frequency received at the base station is given as:

$$f_r = f_c + f_d \tag{1.22}$$

where f_c is the carrier frequency and f_d the doppler frequency shift given as

$$f_d = \frac{\vartheta \cos\theta}{c} f_c \tag{1.23}$$

where c is the speed of light and θ is the angle between the direction of motion and the base station.

Consider again the flat fading case, if there is no motion between the transmitter and the receiver, the channel coefficient is given as in (1.6). Now, as the mobile station (transmitter) is moving toward the base station (receiver), after some time t has passed, the distance decreases by $(\vartheta \cos\theta)t$, the previous propagation delay (in the case where there was no motion) for a given path l will also decrease as:

$$\tau_l(t) = \tau_l - \frac{\vartheta \cos\theta_l}{c} t \tag{1.24}$$

Hence the time varying channel coefficient $h(t)$ is given as:

$$\begin{aligned} h(t) &= \sum_{l=1}^{L_0} \gamma_l e^{-j2\pi f_c \tau_l(t)} \\ &= \sum_{l=1}^{L_0} \gamma_l e^{-j2\pi f_c (\tau_l - \frac{\vartheta \cos\theta_l}{c} t)} \end{aligned} \tag{1.25}$$

In fact, in the above equation, channel amplitudes, delays, and doppler shifts are assumed to be approximately static (slow varying with time), however, the different channel parameters γ_l, τ_l and θ_l themselves can vary with time. To avoid confusion about time varying time delay, in (1.24), the time varying delay parameter will be

written as $\tau_l^{'}(t) = \tau_l(t) - \frac{\vartheta \cos\theta_l(t)}{c}t$, where $\tau_l(t)$ here (different from $\tau_l(t)$) in (1.24)) is the delay of the channel at time t (without doppler, due to the nature of the channel), $\tau_l^{'}(t)$ is the overall delay at time t when doppler is included (this means that $\tau_l(t)$ will be replaced by $\tau_l^{'}(t)$ in (1.25)), and the doppler frequency shift corresponding to path l is $f_{d_l}(t) = \frac{\vartheta \cos\theta_l(t)}{c} f_c$. Channel coefficient $h(t)$ can be generalized and written as:

$$h(t) = \sum_{l=1}^{L_0} \gamma_l(t) e^{-j\phi_l(t)} \tag{1.26}$$

where $\phi_l(t)$ is the phase parameter including the contribution of doppler in addition to the phase due to the propagation delay.

For the frequency selective case, the channel response is given as in (1.10), where the phase of $\alpha_l(t)$ contains also the contribution of doppler (each $\alpha_l(t)$ is of the form of $h(t)$ in(1.26)).

The doppler spread is sometimes called the channel fading rate or fading bandwidth. The key parameter related to doppler effect is the channel coherent time (T_c). T_c is the time interval over which the channel is almost invariant (static). T_c is inversely proportional to $f_{d_{max}}$ and it can have different definitions when it is related to it. A first approximative relationship between the coherence time and doppler spread is given as:

$$T_c \approx \frac{1}{f_{d_{max}}} \tag{1.27}$$

where $f_{d_{max}}$ is the maximum doppler shift given as $f_{d_{max}} = \frac{\vartheta}{c} f_c$.

Consider a frequency selective channel, the time varying channel complex coefficient is given as:

$$\alpha_l(t) = |\alpha_l(t)| e^{-j\phi_l(t)} \tag{1.28}$$

Assuming that the channel amplitudes, delays, and doppler shifts are approximately static, hence we have $|\alpha_l(t)| \approx |\alpha_l|$ and

$$\phi_l(t) = 2\pi f_c \tau_l - 2\pi f_{d_l} t \tag{1.29}$$

Now we want to know how the channel is changing with time, or for how long the channel will stay static (the duration over which the above assumptions still valid). In other words we want to know the channel coherence time T_c. This can be done through the autocorrelation function of the channel response for a given path l, given as:

$$\begin{aligned} R_{\alpha_l}(t, t+\Delta t) &= E[\alpha_l(t)\alpha_l^*(t+\Delta t)] \\ &= E[|\alpha_l|^2 e^{-j2\pi f_c \tau_l + j2\pi f_{d_l} t} e^{j2\pi f_c \tau_l - j2\pi f_{d_l}(t+\Delta t)}] \\ &= P_{\alpha_l} E[e^{-j2\pi f_{d_l} \Delta t}] \\ &= P_{\alpha_l} E[e^{-j2\pi f_{d_{max}} cos\theta_l \Delta t}] \end{aligned} \tag{1.30}$$

Assuming that θ_l is uniformly distributed over $[0, \pi]$, then

$$\begin{aligned} R_{\alpha_l}(t, t+\Delta t) &= P_{\alpha_l} \int_0^{\pi} \frac{1}{\pi} e^{-j2\pi f_{d_{max}} cos\theta_l \Delta t} d\theta_l \\ &= P_{\alpha_l} J_0(2\pi f_{d_{max}} \Delta t) = R_{\alpha_l}(\Delta t) \end{aligned} \tag{1.31}$$

where $J_0()$ is the zero-order Bessel function.

We can notice that the autocorrelation function depends on Δt and not on t. Hence the channel response is wide sense stationary (WSS). This WSS model combined with the US model, the overall model is called wide sense stationary uncorrelated scattering (WSSUS) model.

Now if T_c is defined as the time interval over which $R_{\alpha_l}(\Delta t)/R_{\alpha_l}(0)$ is greater than 0.5, then T_c is given as [3]:

$$T_c \approx \frac{9}{16\pi f_{d_{max}}} \tag{1.32}$$

In the frequency domain, the doppler power spectrum of the channel is given as:

$$\begin{aligned} S(f) &= \int_{-\infty}^{+\infty} \frac{R_{\alpha_l}(\Delta t)}{R_{\alpha_l}(0)} e^{-j2\pi f \Delta t} d\Delta t \\ &= \begin{cases} \frac{1}{\pi f_{d_{max}} \sqrt{1-(\frac{f}{f_{d_{max}}})^2}}, & \text{if } |f| \leqslant f_{d_{max}} \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{1.33}$$

hence the doppler spread is the frequency range over which $S(f)$ is non-zero. The above doppler spectrum is referred to as the Jakes' spectrum. Due to the doppler shift, the propagating signals exhibit what is called "spectral broadening".

Based on the value of T_c relative to the symbol period T_s, we have two types of fading: slow fading and fast fading. The fading phenomenon is referred to as slow fading when the transmitted symbol period is less than the channel coherence time, that is $T_s < T_c$ (or $B_s > f_{d_{max}}$), while when $T_s > T_c$ ($B_s < f_{d_{max}}$), it is referred to as fast fading.

The channel coherence time has an important impact on the wireless communication system design, since it affects the rate of channel estimation procedure that should be done at the receiver.

Figure 1.7 summarizes the different types of small scale fading.

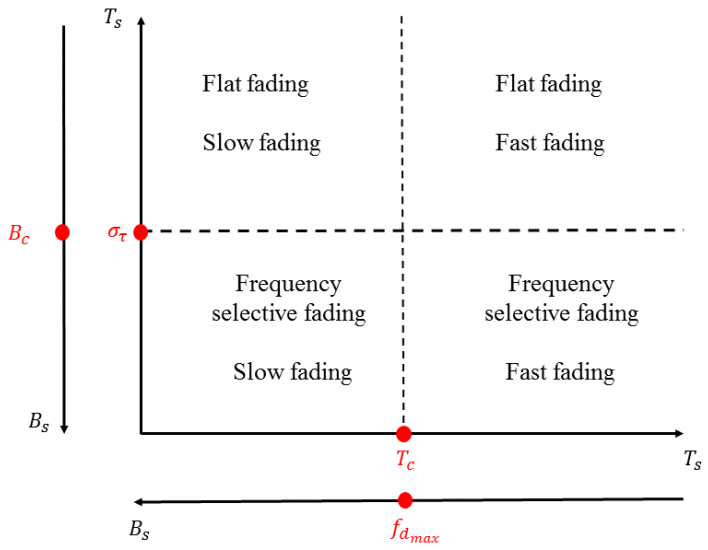

Figure 1.7: Classification of small scale fading.

Fading has a big impact on the amount of information that can be sent over a channel, the maximum data rate that a communication channel can support with small error probability is referred to as the channel capacity.

The capacity of an additive white Gaussian noise (AWGN) channel without fading is given as

$$C_{\text{AWGN}} = B\log_2(1 + \text{SNR}) \tag{1.34}$$

With the effect of fading, the channel capacity of a flat fading channel is given as

$$C_{\text{SISO}} = B\log_2(1 + |h|^2\text{SNR}) \tag{1.35}$$

1.5 Diversity and multiple antenna systems

Diversity is defined as the scheme or technique that can be used to improve the communication reliability against fading problem in wireless communications. Diversity techniques can be classified into three main categories: Time diversity, frequency diversity and spatial diversity (or antenna diversity).

In time diversity, multiple copies of the same signal are transmitted on different time slots seperated at least by the channel coherence time T_c, in order to assure that the signals will exhibit uncorrelated fadings. The disadvantage of time diversity is that different time slots are used to send the same data, hence resulting in lower data rate. In frequency diversity, multiple copies of the same signal are transmitted on different carrier frequencies seperated at least by the channel coherence bandwidth B_c to assure uncorrelated fadings. The disadvantage of frequency diversity is that different frequency bands are used to send the same data, hence requiring extra bandwidth and therefore decreasing spectral efficiency. Another form of diversity is the spatial diversity, which is also known as antenna diversity. Spatial diversity is a scheme that uses multiple antennas at the transmitter and/or receiver. The use of multiple antennas at the receiver is referred to as "receive diversity", and the use of multiple antennas at the transmitter is referred to as "transmit diversity". A system that consists of a single transmit antenna and multiple receive antennas is referred to as single-input multiple-output (SIMO), while a system consisting of multiple transmit antennas and single receive antenna is referred to as multiple-input single-output (MISO). A system consisting of multiple antennas at both the transmitter and receiver sides is referred to as MIMO.

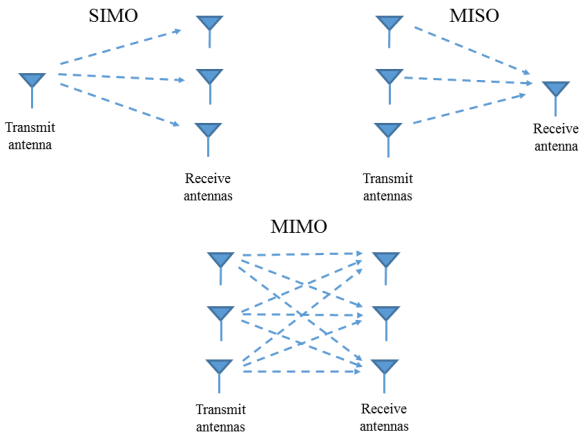

Figure 1.8: Illustration of SIMO, MISO and MIMO systems.

The improvement in the communication reliability against fading is measured by what is called "Diversity gain".

1.5.1 SIMO systems

In SIMO systems, multiple copies of the transmitted signal are received by the different receive antennas through different independent paths, where the received signals are processed and combined together in order to maximize the signal-to-noise ratio (SNR) at the receiver. Under the condition that the receive antennas should be separated at least by half of the signal carrier's wavelength ($\lambda/2$) to assure that the signals on different receive antennas will exhibit uncorrelated fadings. The increase in the average received SNR in SIMO systems is referred to as array gain. A SIMO system with M receive antennas can provide an array gain and a diversity gain of M.

Consider a flat fading channel, and a SIMO system with M receive antennas, the M-dimensional received vector is given as:

$$\mathbf{r} = \mathbf{h}s + \mathbf{z} \tag{1.36}$$

where s is the transmitted signal, $\mathbf{r} = [r_1 \dots r_M]^T$ such that r_m is received by the mth receive antenna, $\mathbf{h} = [h_1 \dots h_M]^T$ where h_m is the channel coefficient between the transmit antenna and the mth receive antenna and $\mathbf{z} = [z_1 \dots z_M]^T$ where z_m is an additive white Gaussian noise at antenna m. We need to find an optimal weighting vector that

maximizes the output SNR.

Consider the linear combination of the received signals as follows:

$$\tilde{r} = \mathbf{w}^H \mathbf{r} \tag{1.37}$$

where $\mathbf{w} = [w_1 \dots w_M]^T$ is a weighting vector. The operation in (1.37) is also known as beamforming where $\mathbf{w}$ can be termed as beamformer.

Hence at the output of the beamformer we have:

$$\tilde{r} = \mathbf{w}^H \mathbf{h} s + \mathbf{w}^H \mathbf{z} \tag{1.38}$$

The received signal component is $\mathbf{w}^H \mathbf{h} s$ and the noise component is $\mathbf{w}^H \mathbf{z}$. Hence the output signal power is given by $|\mathbf{w}^H \mathbf{h}|^2 \sigma_s^2$ where $\sigma_s^2 = E[|s|^2]$ is the transmitted signal power, and the output noise power is given as $E[|\mathbf{w}^H \mathbf{z}|^2] = ||\mathbf{w}||^2 \sigma_z^2$.

Hence the SNR at the output of the beamformer is given as

$$\text{SNR}_{\text{out}} = \frac{|\mathbf{w}^H \mathbf{h}|^2 \sigma_s^2}{||\mathbf{w}||^2 \sigma_z^2} \tag{1.39}$$

Now the goal is to find $\mathbf{w}$ that maximizes the output SNR. The optimal $\mathbf{w}$ that maximizes the ouput SNR is given as:

$$\mathbf{w}_{\text{out}} = \frac{\mathbf{h}}{||\mathbf{h}||} \tag{1.40}$$

The ouput SNR is then given as

$$\text{SNR}_{\text{out}} = \frac{||\mathbf{h}||^2 \sigma_s^2}{\sigma_z^2} \tag{1.41}$$

this is also known as spatial matched filter.

This type of receive diversity is also referred to as maximum ratio combining (MRC). Other receive diversity schemes are the selection combining (SC) and the equal gain combining (EGC). In the SC scheme, the antenna with the highest received signal power is only selected, ignoring the other received signals. In the EGC scheme, the signals at the different receive antennas are coherently combined or co-phased, where the magnitudes of the beamforming weights are set to unity. SC and EGC schemes require lower complexity than MRC, however MRC provides the best performance. In fact, in the SC

scheme, the channel state information is not required, simply the branch with highest SNR is selected. The EGC scheme requires the knowledge of the phase shifts caused by the fading channel. The MRC scheme requires full knowledge of the channel state information, in addition, it requires the estimation of the noise power at each receive antenna. When the noise power is not the same on the different receive antennas, the above defined weighting coefficients at each antenna are divided by the corresponding noise power. Generally, the weighting coefficients are approximated by considering the channel coefficients (as derived above) obtained through channel estimation [8].

In Figure 1.9, the bit error rate (BER) versus energy per bit to noise power spectral density ratio (E_b/N_0) is shown for BPSK modulation, using a SIMO system with MRC, for different number of receive antennas. As shown, the SIMO scheme provides better performance (less BER) compared to the SISO scheme, it is also shown that as the number of receive antennas increases, the BER decreases.

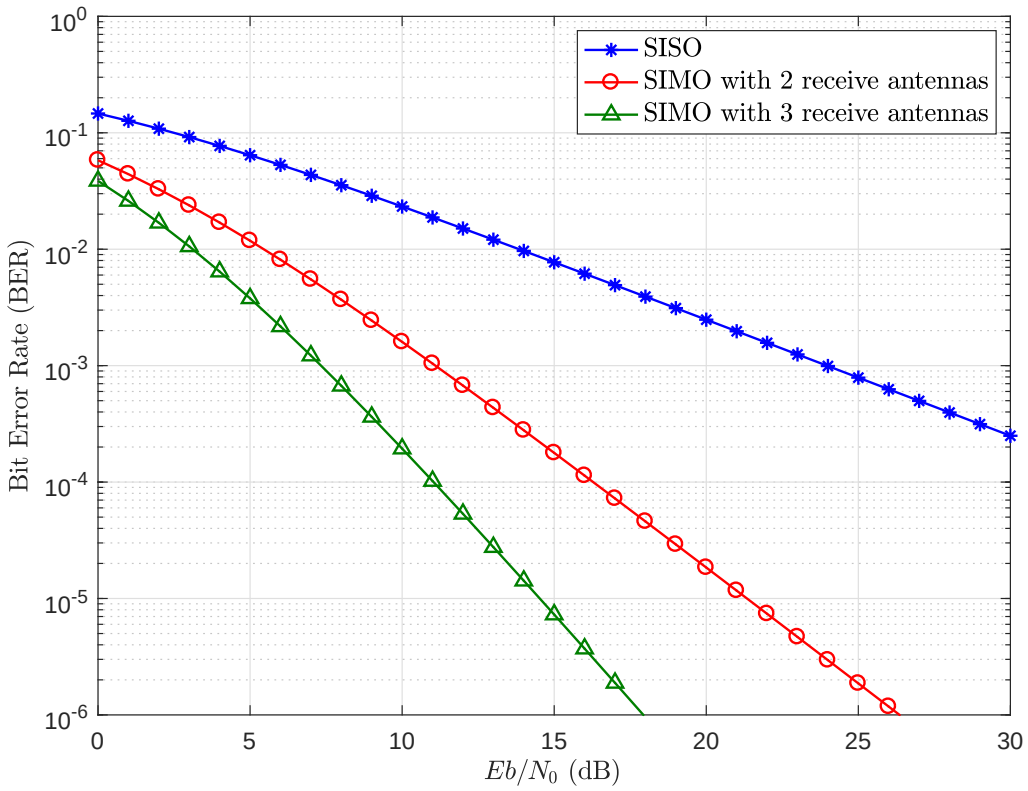

Figure 1.9: BER versus $E_b/N0$ for BPSK modulation, using SIMO system with MRC.

The capacity that can be achieved when using a SIMO system is given as:

$$C_{\text{SIMO}} = B\log_2(1 + \frac{||\mathbf{h}||^2\sigma_s^2}{\sigma_z^2}) \tag{1.42}$$

1.5.2 MISO systems

As mentioned before, a MISO system is composed of multiple transmit antennas and a single receive antenna. Diversity in MISO systems can be achieved through two schemes, depending on whether the CSI is available at the transmitter or not. First, consider the case where the CSI is estimated and available only at the receiver. In this case, MISO schemes exploit space and time dimensions, using the concept of space-time coding. A famous space-time coding scheme used is the Alamouti scheme where the data symbols and their conjugate are transmitted from the different transmit antennas at different times, without increasing the transmission power, having the same transmission power divided over the different transmit antennas. This transmission scheme provides diversity gain with no effect on the data rate.

Consider a MISO system having 2 transmit antennas, and assume Rayleigh flat fading channel, the data symbols to be transmitted are given as:

$$\mathbf{S} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix} \tag{1.43}$$

where in the first symbol period, data symbols s_1 and s_2 are transmitted from antenna 1 and antenna 2 respectively, then in the second symbol period, $-s_2^*$ and s_1^* are transmitted from antenna 1 and antenna 2 respectively.

At the receive antenna, at the first time instant, we have

$$r(1) = h_1 s_1 + h_2 s_2 + z(1) \tag{1.44}$$

where h_1 is the channel coefficient between the first transmit antenna and the receive antenna, h_2 is the channel coefficient between the second transmit antenna and the receive antenna, and $z(1)$ is the received noise sample at the first instant. At the second time instant

$$r(2) = -h_1 s_2^* + h_2 s_1^* + z(2) \tag{1.45}$$

Taking the conjugate of the second sample we have

$$r^*(2) = -h_1^* s_2 + h_2^* s_1 + z^*(2) \tag{1.46}$$

If we stack the first received sample and the conjugate of the second received sample as follows:

$$\begin{bmatrix} r(1) \\ r^*(2) \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} z(1) \\ z^*(2) \end{bmatrix} \tag{1.47}$$

Let $\tilde{\mathbf{H}}=\begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix}$, the columns of $\tilde{\mathbf{H}}$ are orthogonal where

$$\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H = \begin{bmatrix} |h_1|^2 + |h_2|^2 & 0 \\ 0 & |h_1|^2 + |h_2|^2 \end{bmatrix} \tag{1.48}$$

Let $\mathbf{w}_1 = [h_1 \quad h_2^*]^T$, then at the first time instant, the first symbol s_1 can be detected as:

$$\mathbf{w}_1^H \begin{bmatrix} r(1) \\ r^*(2) \end{bmatrix} = (|h_1|^2 + |h_2|^2)s_1 + h_1^* z(1) + h_2 z^*(2) \tag{1.49}$$

The output SNR is given as $\text{SNR}_{\text{out1}} = (|h_1|^2 + |h_2|^2)\frac{\sigma_{s_1}^2}{\sigma_z^2}$ where $\sigma_{s_1}^2$ is the power of the signal transmitted from the first antenna.

Let $\mathbf{w}_2 = [h_2 \quad -h_1^*]^T$, then at the second time instant, the second symbol s_2 can be detected as:

$$\mathbf{w}_2^H \begin{bmatrix} r(1) \\ r^*(2) \end{bmatrix} = (|h_1|^2 + |h_2|^2)s_2 + h_2^* z(1) - h_1 z^*(2) \tag{1.50}$$

The output SNR is given as $\text{SNR}_{\text{out}_2} = (|h_1|^2 + |h_2|^2)\frac{\sigma_{s_2}^2}{\sigma_z^2}$, where $\sigma_{s_2}^2$ is the power of the signal transmitted from the second antenna.

As the transmission power is divided equally among the transmit antennas such that $\sigma_{s_1}^2 = \sigma_{s_2}^2 = \frac{\sigma_s^2}{2}$. The output SNR for a given symbol is given as

$$\text{SNR}_{\text{out}} = \frac{||\mathbf{h}||^2}{2}\frac{\sigma_s^2}{\sigma_z^2} \tag{1.51}$$

As noticed, the Alamouti scheme can achieve a diversity gain of order 2, but results in a 3 dB loss (half of the signal power) in SNR compared to the MRC scheme with 2 receive antennas. Hence the above transmit Alamouti scheme provides diversity gain

without any array gain. This is the cost of achieving diversity in the absence of CSI at the transmitter.

In Figure 1.10, the BER versus E_b/N_0 is shown for BPSK modulation, using a MISO system employing the Alamouti code, for different number of transmit antennas.

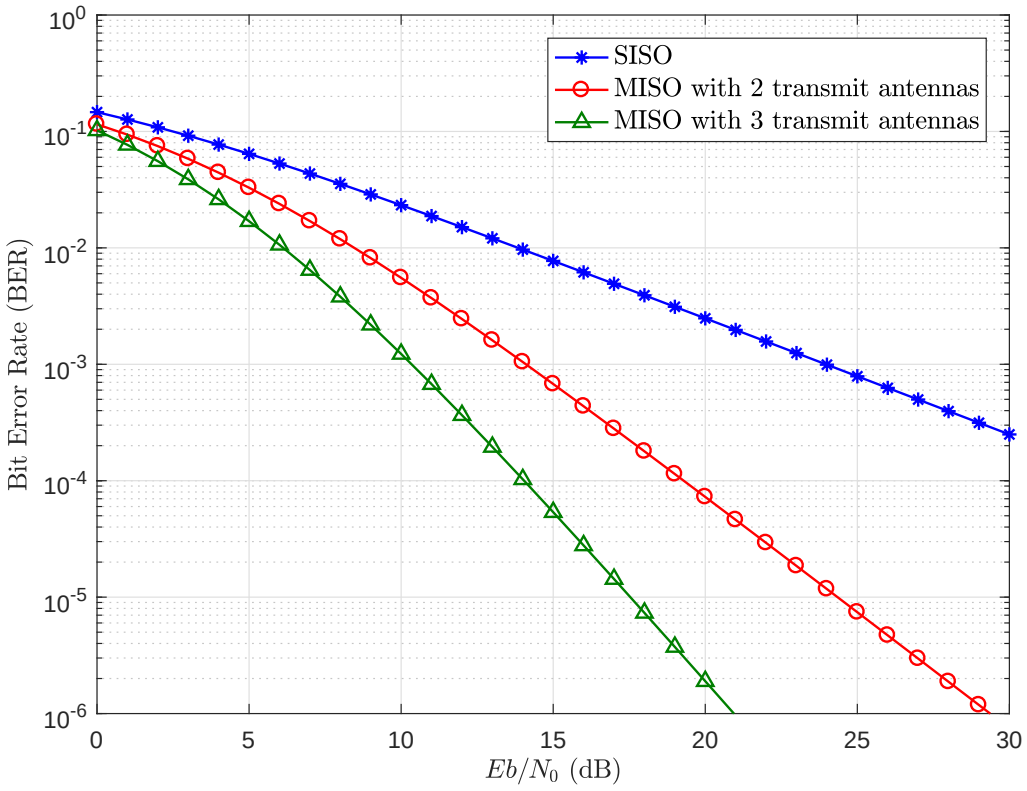

Figure 1.10: BER versus $E_b/N0$ for BPSK modulation

As shown, the MISO scheme allows to achieve less BER compared to the SISO scheme, where this BER decreases as the number of transmit antennas increases.

We can notice from Figures 1.9 and 1.10 that the SIMO scheme employing MRC provides better performance than the MISO scheme employing the Alamouti code. This is better illustrated in Figure 1.11. In fact, in addition to the diversity gain provided by the two schemes, the SIMO scheme (with MRC) can achieve array gain while the MISO-Alamouti scheme cannot.

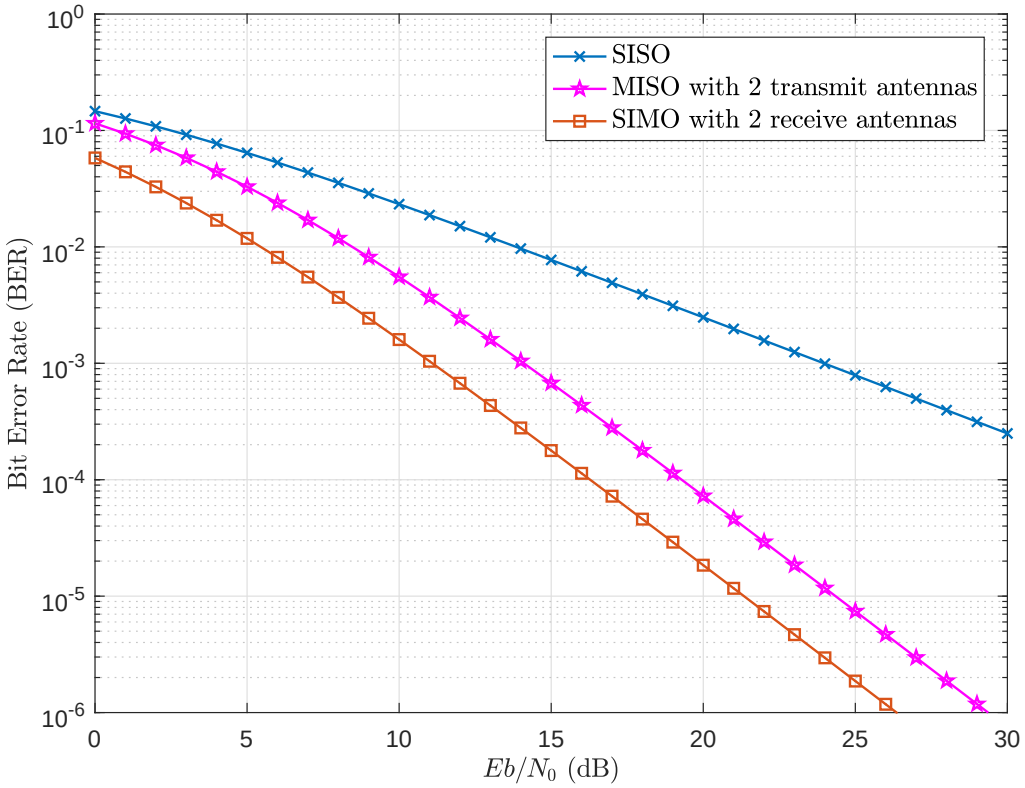

Figure 1.11: Comparison between MISO and SIMO performance

Although MISO scheme does not provide any array gain, that is, there is no increase in the input SNR ($\frac{\sigma_s^2}{\sigma_z^2}$), however due to the diversity gain, it results in an improvement of the output SNR in comparison with the SISO case.

For a MISO system with N transmit antennas, the MISO channel capacity is given as

$$C_{\text{MISO}} = B\log_2(1 + \frac{||\mathbf{h}||^2}{N}\frac{\sigma_s^2}{\sigma_z^2}) \tag{1.52}$$

Consider again a 2×1 MISO system, and a symbol s transmitted from the two transmit antennas, but this time, assume that the transmitter knows the CSI, such that the channel coefficients h_1 and h_2 are known to the transmitter. In this case, a kind of precoding can be applied such that at the transmitter side, the symbol s is multiplied by $\frac{h_1^*}{||\mathbf{h}||}$ and transmitted from the first antenna as $s_1 = \frac{h_1^*}{||\mathbf{h}||}s$, the symbol s is

also multiplied by $\frac{h_2^*}{||\mathbf{h}||}$ and transmitted as $s_2 = \frac{h_2^*}{||\mathbf{h}||}s$ from the second transmit antenna. At the receive antenna, we have:

$$r = (\frac{|h_1|^2}{||\mathbf{h}||} + \frac{|h_2|^2}{||\mathbf{h}||})s + z = ||\mathbf{h}||s + z \tag{1.53}$$

The output SNR is then given as

$$\text{SNR}_{\text{out}} = ||\mathbf{h}||^2 \frac{\sigma_s^2}{\sigma_n^2} \tag{1.54}$$

which is the same as MRC. This can be referred to as transmit beamforming.

1.5.3 MIMO systems

In MIMO systems, multiple antennas are used on both the transmitter and receiver sides, as seen before, the use of multiple transmit or receive antennas allows to provide diversity gain, thus to improve the communication reliability. In addition, the use of multiple transmit antennas (with CSI available at the transmitter) or multiple receive antennas provides array gain, that is, increasing the power of the transmitted signal, or SNR. Using multiple transmit and receive antennas is capable of providing an additional gain called the spatial multiplexing gain. The spatial multiplexing gain can be defined as the gain attained through transmitting different data streams from the different transmit antennas leading to an increase in the capacity of the system without the need for additional bandwidth or transmission power. However there is a tradeoff between the spatial multiplexing gain and the diversity gain, such that as diversity gain increases, multiplexing gain decreases and vice versa. With no spatial multiplexing, which means that the same data is transmitted over all the transmit antennas, a MIMO system with N transmit antennas and M receive antennas has the capability to achieve a diversity gain equal to $N \times M$. On the other hand, if different data are sent on the different transmit antennas, the MIMO system is able to achieve a maximum multiplexing gain equal to $\min(N, M)$ in the case of high SNR, without any diversity gain, where $\min(N, M)$ is defined as the number of degrees of freedom in the system.

Consider a flat fading MIMO system with N transmit antennas and M receive antennas, the MIMO system model is given as:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{z} \tag{1.55}$$

where $\mathbf{y} = [y_1 \dots y_M]^T$ is the received signal vector, $\mathbf{s} = [s_1 \dots s_N]^T$ is the transmitted signal vector, $\mathbf{z} = [z_1 \dots z_M]^T$ is the noise vector, and $\mathbf{H}$ is the MIMO channel matrix

given as:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & . & . & . & h_{1N} \\ h_{21} & h_{22} & . & . & . & h_{2N} \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ h_{M1} & h_{M2} & . & . & . & h_{MN} \end{bmatrix} \tag{1.56}$$

where h_{mn} is the channel coefficient between the mth receive antenna and nth transmit antenna.

Consider the singular value decomposition (SVD) of $\mathbf{H}$, then

$$\mathbf{H} = \mathbf{U}_H \mathbf{D}_H \mathbf{V}_H^H \tag{1.57}$$

where $\mathbf{D}_H$ is an $M \times N$ diagonal matrix, and the matrices $\mathbf{U}_H$ and $\mathbf{V}_H$ are $M \times M$ and $N \times N$ unitary matrices, respectively. If $\mathbf{H}$ is a full rank matrix then its rank is $\min(N, M)$, which is also the number of spatial degrees of freedom. Let N_d=$\min(N, M)$, then $\mathbf{H}$ has N_d non-zero singular values on its diagonal such that $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_{N_d}$.

Assuming the CSI is known at the transmitter side, the data symbols can be precoded as:

$$\tilde{\mathbf{s}} = \mathbf{V}_H \mathbf{s} \tag{1.58}$$

then $\tilde{\mathbf{s}}$ is transmitted. Consequently,

$$\bar{\mathbf{y}} = \mathbf{H}\tilde{\mathbf{s}} + \mathbf{z} \tag{1.59}$$

Now if we multiply at the receiver by $\mathbf{U}_H^H$, we have:

$$\begin{aligned} \tilde{\mathbf{y}} &= \mathbf{U}_H^H \bar{\mathbf{y}} + \mathbf{U}_H^H \mathbf{z} \\ &= \mathbf{U}_H^H \mathbf{U}_H \mathbf{D}_H \mathbf{V}_H^H \mathbf{V}_H \mathbf{s} + \tilde{\mathbf{z}} \\ &= \mathbf{D}_H \mathbf{s} + \tilde{\mathbf{z}} \end{aligned} \tag{1.60}$$

Hence we have:

$$\tilde{y}_m = \mu_m s_m \quad \text{for} \quad m = 1, \ldots N_d \tag{1.61}$$

when $M > N$ ($N_d = N$), then $\tilde{y}_m$ for $m = 1, \ldots N$ are the first N non-zero elements of $\mathbf{y}$, while when $M < N$ ($N_d = M$), only M symbols can be received, hence only M non-zero data symbols can be transmitted.

As shown in the above scheme, the MIMO system is equivalent to an N_d parallel SISO systems. It follows that the capacity of the MIMO system will be the sum of capacities of the different parallel SISO systems. The MIMO channel capacity is given as:

$$C_{\text{MIMO}} = \sum_{m=1}^{N_d} B\log_2(1 + \mu_m^2 \frac{\sigma_{s_m}^2}{\sigma_z^2}) \tag{1.62}$$

where the total transmission power σ_s^2 distributed over the different transmit symbols at the different transmit antennas is given as $\sigma_s^2 = \sum_{m=1}^{N_d} \sigma_{s_m}^2$.

The MRC scheme, which is illustrated before for the SIMO case can be employed in the MIMO case, where the received symbols are detected by multiplying the received vector $\mathbf{y}$ by the conjugate transpose of the channel matrix $\mathbf{H}$, as follows:

$$\hat{\mathbf{s}} = \mathbf{H}^H \mathbf{y} \tag{1.63}$$

If the channel matrix $\mathbf{H}$ is known at the transmitter, a precoding scheme (transmit beamforming) can be applied, by multiplying the data symbols to be transmitted by the conjugate of the channel matrix.

Other detection schemes are zero forcing (ZF) and minimum mean square error (MMSE) schemes. In the ZF scheme, the received signal vector is multiplied by the pseudo inverse of the channel matrix, as follows:

$$\hat{\mathbf{s}}_{\text{ZF}} = \mathbf{H}^\dagger \mathbf{y} \tag{1.64}$$

where $\mathbf{H}^\dagger = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H$. ZF receiver attempts to minimize inter-antenna interference, but results in noise enhancement. ZF can be used as a precoding scheme. The scheme involves multiplying the data symbols to be transmitted by $(\mathbf{H}^\dagger)^T$.

The MMSE detection provides better performance in comparison with the ZF scheme, with increased complexity. It attempts to simultaneously reduce both the interference and the noise. MMSE signal detection is given as:

$$\hat{\mathbf{s}}_{\text{MMSE}} = (\mathbf{H}^H \mathbf{H} + \sigma_z^2 \mathbf{I})^{-1} \mathbf{H}^H \mathbf{y} \tag{1.65}$$

When the channel matrix $\mathbf{H}$ is known at the transmitter, MMSE can be used as a precoding scheme. It involves multiplying the data symbols to be transmitted by $\mathbf{H}^*(\mathbf{H}^T\mathbf{H}^* + \sigma_z^2\mathbf{I})^{-1}$.

The precoding/detection schemes described above are linear schemes. There exists also a class of nonlinear schemes. Such schemes are more complex but can achieve better performance. Among several nonlinear detection schemes, the vertical bell lab layered space time (V-BLAST) scheme [9] is popular . V-BLAST relies on successive interference cancellation (SIC), where symbols are detected in an iterative manner.

In the above, for the sake of simplification, we considered a flat fading channel to illustrate the benefits achieved by using MIMO systems. Considering the frequency selective fading case, a time varying frequency selective channel matrix for an $N \times M$ MIMO system is given as:

$$\mathbf{H}(t,\tau) = \begin{bmatrix} h_{11}(t,\tau) & h_{12}(t,\tau) & . & . & . & h_{1N}(t,\tau) \\ h_{21}(t,\tau) & h_{22}(t,\tau) & . & . & . & h_{2N}(t,\tau) \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ h_{M1}(t,\tau) & h_{M2}(t,\tau) & . & . & . & h_{MN}(t,\tau) \end{bmatrix} \tag{1.66}$$

If the channel is time invariant, then $h_{nm}(t,\tau) = h_{nm}(\tau)$ for all n and m. Consider a frequency selective channel with L multipath components, the system model is then written as:

$$\mathbf{y}(t) = \sum_{l=1}^{L} \mathbf{H}(\tau_l)\mathbf{s}(t-\tau_l) + \mathbf{z}(t) \tag{1.67}$$

In the flat fading case, the narrowband assumption states that the different propogation paths cannot be resolved temporally, which is not the case for frequency selective fading. Based on that, flat fading or narrowband MIMO channel models are fully characterized by their spatial structure, while frequency selective or wideband channels are characterized by their spatial and temporal (multipath) structure.

1.6 Conclusion

This chapter provides a basic understanding of the fading characteristics in wireless channels. Different types of fading are presented, and the effects of fading on communication performance are addressed. Different diversity schemes used against fading are then introduced, focusing on spatial diversity, where a brief review of the basics of MIMO communication technology is provided.

Chapter 2

MIMO channel models

The design and performance of MIMO systems highly depend on the propagation environment [5, 10, 11]. A MIMO channel model that accurately describes the propagation channel is important to exploit the advantages provided by MIMO systems.

MIMO channel models can be categorized into physical and non-physical models. In non-physical models, the channel matrix coefficients are modeled statistically with respect to the correlation between them. On the other hand, physical models characterize the MIMO channel by means of physical parameters related to the signal propagation through the channel [12], such parameters can be the time of arrival (TOA), angle of departure (AOD), angle of arrival (AOA).

2.1 Physical models

Physical models can be divided into deterministic and stochastic models.

2.1.1 Deterministic models

In deterministic physical models, the channel is described in a deterministic fashion. Ray tracing [4, 13, 14] is a well known effective method used for deterministic channel modeling. It is used as an alternative to solving Maxwell's equations (approximate solution to Maxwell's equations) [6], that characterizes the propagation paths on the basis of geometrical optics. In ray tracing methods, prior information such as the positions of the transmitter, receiver and objects in the environment are known, channel propagation characteristics are obtained by tracking rays from the transmitter to the receiver, considering different types of physical propagation phenomena such as reflection, diffrac-

tion, etc., that occurs as a result of rays interaction with objects in the environment. Although deterministic physical models have the ability to provide accurate channel modeling, they are computationally intensive, which make statistical/stochastic physical models more convenient to consider, as they are more computationally efficient.

2.1.2 Stochastic models

In stochastic models, the physical propagation parameters are described in a stochastic or statistical manner. Stochastic models are divided into geometry-based stochastic models and nongeometric stochastic models.

2.1.2.1 Geometry-based stochastic channel models (GSCM)

In such models, the geometry of the channel is taken into account, where the physical propagation parameters are modeled by defining statistical distributions to scatterers in the environment, that is, the positions of the scatterers are modeled as random and ruled by some predetermined statistical distributions. It follows that the values of the physical parameters are influenced by these distributions.

Two main scattering configurations can be considered for geometry-based models: single-bounce scattering and multiple-bounce scattering [12, 15, 16].

A single-bounce scattering assumes that there exists a single scattering interface between the transmitter and the receiver, so that the transmitted signal bounces one time before it arrives at the receiver. The simple case of single-bounce scattering is when the scatterers are assumed to be uniformly distributed over space (spatially uniformly distributed) [12]. Another models assume that scatterers are distributed randomly around the transmitter or/and the receiver [12, 17].

Multi-bounce scattering is the case when the signal bounces several times before arriving at the receiver, this can be due to different scatterers distributed in the environment or/and local scatterers surrounding the transmitter or/and receiver.

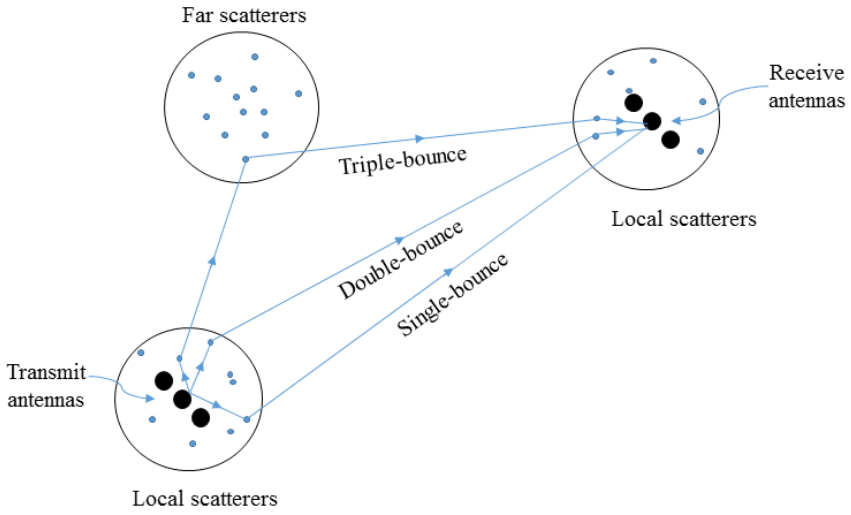

Figure 2.1: Illustration of GSCM (single-bounce and multiple-bounce scattering).

Figure 2.1 illustrates the phenomena of single-bounce and multiple-bounce scattering. As shown, we have three cases, for the single-bounce scattering case, the wave is only scattered by the local scatterers around the transmitter. For the double-bounce case, the wave is scattered by the local scatterers around the transmitter and receiver, and for the triple bounce case, in addition to wave scattering by local scatterers around the transmitter and receiver, the wave is also scattered by a cluster of scatterers faraway in the environment.

The geometry of scatterers in the environment has an impact on the AOA and TOA statistics, that is, the distributions of scatterer positions within clusters have an impact on AOA and TOA statistical distributions. [18, 19].

2.1.2.2 Nongeometrical stochastic channel models

In nongeometrical stochastic models, the physical propagation parameters are modeled in a statistical manner without referring to the geometry of scatterers in the environment. Two main nongeometrical models exist in the literature, the extended Saleh-Valenzuela model and Zwick model.

Extended Saleh-Valenzuela model

The Saleh-Valenzuela model is proposed for the indoor environments, it assumes that the multipath components (MPCs) arrive within clusters. The model was first proposed in SISO systems considering only the time of arrivals [20], but it was later developed to include the AOA and AOD [21–24], where the channel impulse response is given as :

$$h(t,\Theta_t,\Theta_r)=\sum_{l=1}^{L}\sum_{p=1}^{P_l}\beta_{lp}\delta(t-T_l-\tau_{lp})\delta(\theta^T-\theta_l^T-\theta_{lp}^T)\delta(\theta^R-\theta_l^R-\theta_{lp}^R) \tag{2.1}$$

where L is the number of clusters, P_l is the number of multirays within cluster l, T_l, θ_l^T and θ_l^R are the TOA, AOD and AOA of the lth cluster, respectively (T_l is defined as the TOA of the first ray in the cluster). τ_{lp},θ_{lp}^T and θ_{lp}^R are the TOA, AOD, AOA of the pth ray in the lth cluster, respectively, and β_{lp} is the corresponding complex gain.

The magnitude of β_{lp} ($|\beta_{lp}|$) is assumed to be Rayleigh distributed and its phase is uniform. The expected power of the pth ray in the lth cluster is given as

$$E[|\beta_{lp}|^2]=E[|\beta_{11}|^2]e^{-T_l/\Gamma}e^{-\tau_{lp}/\zeta} \tag{2.2}$$

where $E[|\beta_{11}|^2]$ is the expected power of the first ray in the first cluster, and Γ and ζ are the cluster and ray decay time constants.

Assuming that TOA, AOD and AOA statistics are independent, the cluster and ray arrival times follow a Poisson process, where the cluster arrival time is an exponentially distributed random variable conditioned on the TOA of the previous cluster, given as:

$$P(T_l/T_{l-1})=\Lambda e^{-\Lambda(T_l-T_{l-1})} \tag{2.3}$$

where Λ is the cluster arrival rate. Likewise, the ray arrival time within a cluster is an exponentially distributed random variable conditioned on the TOA of the previous ray, given as:

$$P(\tau_{lp}/\tau_{l,p-1}) = \eta e^{-\eta(\tau_{lp}-\tau_{l,p-1})} \tag{2.4}$$

where η is the ray arrival rate.

The distributions of the angles of arrival and departure θ_{lp}^T and θ_{lp}^R are assumed to be uniform over $[0, 2\pi)$ (note that θ_l^T and θ_l^R are meant to be the mean AOD and mean AOA respectively), and the rays within a cluster can be assumed to have a Laplacian distribution, such that

$$p(\theta_{lp}^T) = \frac{1}{\sqrt{2}\sigma_\theta^T} e^{-|\sqrt{2}\frac{\theta_{lp}^T}{\sigma_\theta^T}|} \tag{2.5}$$

where σ_θ^T is the standard deviation of the angular spread corresponding to AOD. Similarly

$$p(\theta_{lp}^R) = \frac{1}{\sqrt{2}\sigma_\theta^R} e^{-|\sqrt{2}\frac{\theta_{lp}^R}{\sigma_\theta^R}|} \tag{2.6}$$

where σ_θ^R is the standard deviation of the angular spread corresponding to AOA.

Zwick model

Zwick model treats the different MPCs separately, due to the reason that clustering and multipath fading do not occur in an indoor channel in case of sufficiently large sampling rate [12,25]. The MPCs are treated independently without amplitude fading, while the phase changes of the different MPCs are involved in the model through geometric basis, characterizing the motion of the transmitter, receiver and scatterers.

In physical models, the channel between the transmitter and receiver is modeled on the basis of electromagnetic waves propagation without referring to antenna configurations [12]. Alternatively, analytical models deal with MIMO channel matrix directly by providing it a mathematical representation. Analytical models depend on antenna configurations, and they can be classified based on correlation properties or through the decomposition of the channel matrix in terms of propagation parameters. Analytic models do not really rely on the physical propagation characteristics, however, when the MIMO channel matrix is expressed mathematically in terms of some propagation parameters, the model can be considered as a kind of propagation based analytical model. Hence we choose to classify these kinds of models as physical models. The other kind of analytical models are the correlation based models that totally ignore the

propagation parameters.

2.1.2.3 Propagation based analytical models

Finite scatterer model

In finite scatterer channel model [26], a finite number L of MPCs is considered, each of them is characterized by a complex amplitude α_l, AOD θ_l^T, AOA θ_l^R and delay τ_l.

Consider an $N \times M$ MIMO system, for the narrowband case, the delays can be neglected, and the channel matrix is expressed as:

$$\mathbf{H} = \sum_{l=1}^{L} \alpha_l \mathbf{a}(\theta_l^R)\mathbf{a}^T(\theta_l^T) = \mathbf{A}_R \mathbf{\Xi} \mathbf{A_T}^T \tag{2.7}$$

where $\mathbf{a}(\theta_l^T)$ and $\mathbf{a}(\theta_l^R)$ are the transmit and receive steering vectors and

$$\begin{aligned} \mathbf{A}_R &= [\mathbf{a}(\theta_1^R) \dots \mathbf{a}(\theta_L^R)] \\ \mathbf{A}_T &= [\mathbf{a}(\theta_1^T) \dots \mathbf{a}(\theta_L^T)] \\ \mathbf{\Xi} &= diag\{\alpha_1 \dots \alpha_L\} \end{aligned} \tag{2.8}$$

For wideband systems, the delay is included in the model, the tapped delay line representation of the channel matrix is then given as:

$$\mathbf{H}(\tau) = \sum_{\kappa_l=-\infty}^{+\infty} \mathbf{H}_{\kappa_l} \delta(\tau - \kappa_l T_s) \tag{2.9}$$

where $T_s = \frac{1}{B}$ with B the system bandwidth, and $\mathbf{H}_{\kappa_l}$ is given as

$$\begin{aligned} \mathbf{H}_{\kappa_l} &= \sum_{l=1}^{L} \alpha_l \text{sinc}(\tau_l - \kappa_l T_s)\mathbf{a}(\theta_l^R)\mathbf{a}^T(\theta_l^T) \\ &= \mathbf{A}_R(\mathbf{\Xi} \odot \mathbf{T}_{\kappa_l})\mathbf{A_T}^T \end{aligned} \tag{2.10}$$

where $\mathbf{T}_{\kappa_l} = diag\{\text{sinc}(\tau_1 - \kappa_l T_s), \dots, \text{sinc}(\tau_L - \kappa_l T_s)\}$.

The MPC parameters can be assigned some statistical distributions, and in some environments these distributions can be guessed from measurements [12].

Virtual channel representation

In the previous finite scatterer model, the channel matrix is expressed in terms of the actual physical AODs and AOAs. In virtual channel representation (VCR), the MIMO channel is modeled in the beam space, through expressing the channel matrix in terms of fixed spatial basis defined by fixed transmit and receive virtual angles associated to virtual scatterers in the environment, such that the channel matrix $\mathbf{H}$ is expressed as [27]:

$$\mathbf{H} = \mathbf{F}_M(\tilde{\mathbf{\Omega}} \odot \mathbf{G})\mathbf{F}_N^H \tag{2.11}$$

where $\mathbf{F}_N$ and $\mathbf{F}_M$ are the discrete Fourier transform (DFT) matrices (the fixed basis) that consists of the steering vectors corresponding to N transmit and M receive virtual angles (or scatterers), respectively, and $(\tilde{\mathbf{\Omega}} \odot \mathbf{G})$ is an $M \times N$ matrix representing the environment in between the virtual transmit and receive scatterers. The matrix $\tilde{\mathbf{\Omega}}$ is the element-wise square root of the coupling matrix $\mathbf{\Omega}$ whose elements represent the mean power coupled from the virtual transmit to the virtual receive angles (or scatterers), and $\mathbf{G}$ is an $M \times N$ independent and identically distributed (i.i.d.) zero mean complex Gaussian matrix. Hence the mnth element of $(\tilde{\mathbf{\Omega}} \odot \mathbf{G})$ corresponds to a pair of virtual transmit scatterer n and virtual receive scatterer m, if either the virtual scatterer n or the virtual scatterer m does not exist, the nmth element of $(\tilde{\mathbf{\Omega}} \odot \mathbf{G})$ will be zero. In fact, the matrix $(\tilde{\mathbf{\Omega}} \odot \mathbf{G})$ is a 2-D DFT of $\mathbf{H}$. The transmit and receive virtual angles defined by the transmit and receive DFT matrices are determined by the spatial resolution of the transmit and receive antenna arrays, which means that the accuracy of model increases as the number of antennas increases. It is important to mention that the virtual representation of the channel matrix in terms of DFT basis is only convenient for uniform linear antenna arrays. The VCR model is useful for the theoretical analysis of the capacity scaling in MIMO systems [12, 28].

Figure 2.2 illustrates both the physical channel model and the VCR model.

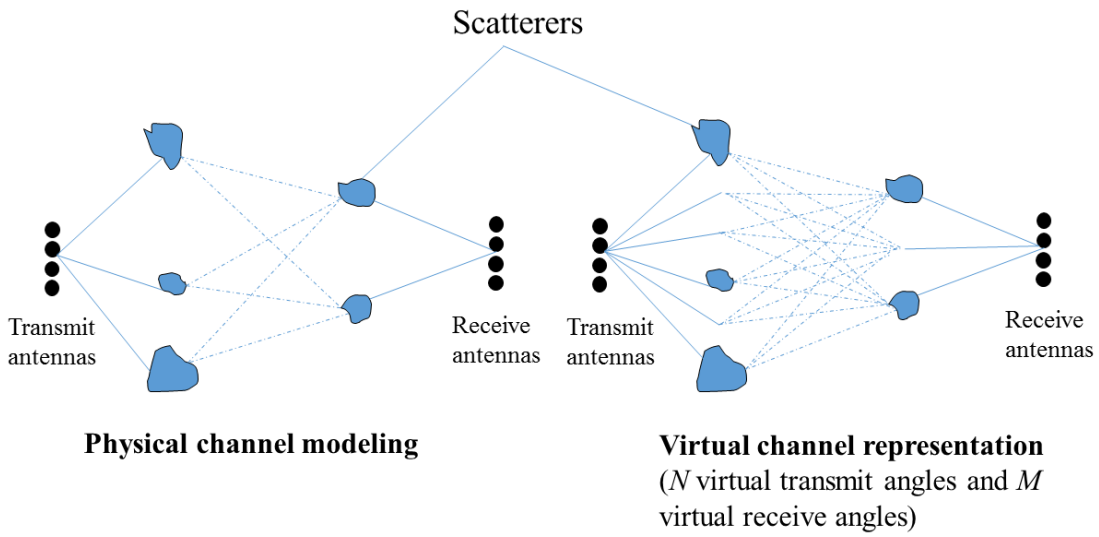

Figure 2.2: Physical channel modeling vs virtual channel representation.

Another propagation based analytical model is the Maximum entropy model, which is based on the principle of maximum entropy, that exploits the available prior information about the propagation environment (such as time delays, AODs, AOAs, bandwidth, etc.) to derive the model that expresses the available knowledge, that is, to find the probability distribution of the channel that best fits the available data [12, 29].

2.2 Non-physical models

We classify the analytical channel models that characterize the MIMO channel matrix statistically without referring to the physical propagation characteristics as non-physical channel models. Such models can be also referred to as correlation based models [12].

The i.i.d. model

The i.i.d model simply describes the MIMO channel matrix as random, with independent and identically distributed elements. It is the spatially white MIMO case that corresponds to a rich scattering environment. Referring to the VCR model, the

i.i.d model refers to the case where all the elements of $\tilde{\mathbf{\Omega}}$ are non-zero and identical. The i.i.d model is generally used for theoretical analysis of MIMO systems such as information-theoretic analysis [12].

Kronecker model

The Kronecker model [30–33] assumes that the correlation between antennas at the transmitter side is independent of the correlation between antennas at the receiver side, such that the channel covariance matrix can be written as:

$$\mathbf{R}_{\mathbf{H}} = \mathbf{R}_T \otimes \mathbf{R}_R \tag{2.12}$$

where $\mathbf{R}_T = E[\mathbf{H}^H\mathbf{H}]$ and $\mathbf{R}_R = E[\mathbf{H}\mathbf{H}^H]$ are the transmit and receive covariance matrices, respectively.

It follows that the channel matrix $\mathbf{H}$ can be written as:

$$\mathbf{H} = \mathbf{R}_R^{1/2}\mathbf{G}\mathbf{R}_T^{1/2} \tag{2.13}$$

where $\mathbf{G}$ is an $M \times N$ independent and identically distributed (i.i.d.) zero mean complex Gaussian matrix.

The Kronecker model results in poor estimation of channel capacity, due to the assumption of independent transmit and receive correlations, which is not accurate in the majority of physical channels [34]. Despite its popularity as a simple model, it has been shown to be only applicable in specific scenarios.

Weichselberger model

The Weichselberger model [35] removes the separability condition imposed by the Kronecker model, where it allows coupling between the transmit and receive eigenmodes. It replaces the transmit and receive DFT matrices in the VCR model with the eigenbasis of the transmit and receive spatial covariance matrices, $\mathbf{R}_T$ and $\mathbf{R}_R$. The eigendecompositions of $\mathbf{R}_T$ and $\mathbf{R}_R$ are given as:

$$\begin{aligned} \mathbf{R}_T &= \mathbf{U}_T\mathbf{E}_T\mathbf{U}_T^H \\ \mathbf{R}_R &= \mathbf{U}_R\mathbf{E}_R\mathbf{U}_R^H \end{aligned} \tag{2.14}$$

where the columns of $\mathbf{U}_T$ and $\mathbf{U}_R$ are the eigenvectors of $\mathbf{R}_T$ and $\mathbf{R}_R$, respectively, and $\mathbf{E}_T$ and $\mathbf{E}_R$ are diagonal matrices containing the corresponding eigenvalues. It follows that $\mathbf{H}$ is given as:

$$\mathbf{H} = \mathbf{U}_R(\tilde{\mathbf{\Omega}} \odot \mathbf{G})\mathbf{U}_T^T \tag{2.15}$$

where $\tilde{\mathbf{\Omega}}$ is the element-wise square root of the coupling matrix $\mathbf{\Omega}$ whose elements represent the mean power coupled from the transmit eigenvectors (or eigenmodes) to the receive eigenvectors and $\mathbf{G}$ is an $M \times N$ independent and identically distributed (i.i.d.) zero mean complex Gaussian matrix. The Weichselberger model is equivalent to the VCR model when the DFT matrix turns into the eigenvector matrix (as the number of antennas goes to infinity).

2.3 Sparsity

Recently, the issue of sparsity of the wireless channel has received a lot of attention. It is a fact that, in many scenarios, wireless channels can be characterized by few significant paths, this is mainly attributed to the limited number of dominant scatterers distributed in the environment, and especially for large bandwidths [36–41]. This sparsity property has been shown to be reasonable through physical investigations that have been reported for some physical channels [42].

2.4 Spatial correlation and the common support phenomenon

In MIMO outdoor communication scenarios, since the transmit and receive array dimensions are relatively small compared to the long transmission distance, the signals received by the different closely located antennas associated to the same scatterer (or same cluster of scatterers) may be spatially correlated, where CIRs of different transmit-receive antenna pairs may share a sparse "common support" [40, 43]. This spatial correlation is controlled by three parameters: signal bandwith, operating frequencies and antenna spacing [41]. When the difference between TOAs on different antennas is less than $\frac{1}{10B_s}$ (B_s is the signal bandwidth), the antennas will share the same path delay, as the individual path delays on these antennas cannot be resolved, while the associated channel gains are different [40, 41]. This is referred to as an exact common support, where different transmit-receive antenna pairs share exactly the same path delays [40, 44]. At least, depending on the size of the array, the exact common support phenomenon has been shown to be validated for certain groups of antennas within a

large antenna array [41].

2.5 Clustering in wireless channels

In practice, and according to various channel investigations, several wireless channels exhibit scattering (clustering) of the propagating signals, where the channel multipath components arrive in clusters [36, 45–49]. For example, this phenomenon (clustering) characterizes the wideband/ultrawideband communication channels [36, 45, 46, 50, 51], as well as the MMW communication channels [52–54], where the channels exhibit sparse clustering, that is, finite number of clusters are present in the channel. For instance, as stated in [36, 45, 46], the well known Saleh-Valenzuela (SV) model [20] developed for indoor channels, can be still adopted to model wideband outdoor communication channels, such that the channel MPCs are modeled as cluster of multirays. The number of clusters in a channel depends on both the propagation environment and the considered bandwidth [36, 50].

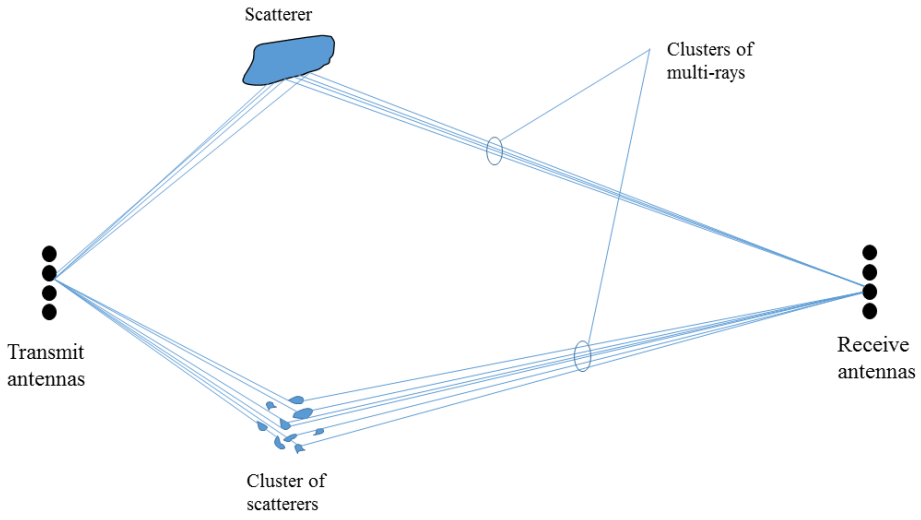

Figure 2.3: Clustering in wireless channels

2.6 Sparse clustered MIMO channel model with common support

For an $N \times M$ MIMO system, based on the clustering and sparsity properties of a wireless channel, the CIR $h^{(n,m)}(t)$ between the nth transmit antenna and mth receive antenna can be modeled as:

$$h^{(n,m)}(t) = \sum_{l=1}^{L} \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)} \delta(t - (t_l^{(n,m)} + \tau_{lp}^{(n,m)})), \quad 1 \leq n \leq N, 1 \leq m \leq M \tag{2.16}$$

where $\delta(.)$ is the Dirac function; L is the total number of propagation paths (clusters), which is considered finite, with P_l contributing rays for the lth path (cluster), $t_l^{(n,m)} + \tau_{lp}^{(n,m)}$ is the pth contributing ray delay in the lth cluster with $\tau_{lp}^{(n,m)}$ a small deviation from the cluster mean delay $t_l^{(n,m)}$ and $\alpha_{lp}^{(n,m)}$ is the corresponding complex gain.

In some existing works [40, 44], an exact common support condition is considered such that different transmit-receive antenna pairs share exactly the same path delays. However, as the bandwidth of the signal increases, the exact common support property starts losing its validity. For this reason, in the above channel model (2.16), multiray delays associated to a given cluster are not considered exactly the same. However, it can be assumed that the multirays associated to a given scatterer l share the same cluster mean delay at the different transmit-receive antenna pairs. This is a kind of common support hypothesis. Accordingly, the CIR $h^{(n,m)}(t)$ between the nth transmit antenna and mth receive antenna can be modeled as:

$$h^{(n,m)}(t) = \sum_{l=1}^{L} \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)} \delta(t - (t_l + \tau_{lp}^{(n,m)})), \quad 1 \leq n \leq N, 1 \leq m \leq M \tag{2.17}$$

where t_l is the mean delay associated to cluster l shared by all the transmit-receive antenna pairs.

2.7 Conclusion

In this chapter, an overview of MIMO channel models is provided. The channel models are divided into different categories based on different propagation scenarios under different conditions. Sparsity and clustering properties in wireless channels are then

introduced, followed by the common support property in outdoor MIMO channels. A sparse clustered MIMO channel model with common support, on which our proposed estimation methods in chapters 4 and 5 are based, is finally presented.

Chapter 3

Channel parameter estimation

A performant "ideal" MIMO communication system would require an exact knowledge of the MIMO channel or channel state information. MIMO channel parameters estimation is required for the equalization at the receiver side and precoding at the transmitter side. Hence, channel parameter estimation is a crucial part in MIMO communication. Channel estimation techniques are divided into training based, semi-blind and blind techniques. Training-based channel estimation methods are based on transmitting an a priori known training or pilot sequence to the receiver, and estimating the instantaneous channel coefficients based on the received signal and on the known transmitted sequence [43, 55–57]. The maximum likelihood (ML) and the MMSE are popular techniques used for channel estimation in the presence of training sequences. Blind channel estimation techniques do not consider training symbols, where the channel is estimated using only the statistical properties of the received signals and by exploiting the structure of the channel [56, 58]. Blind techniques have the capability to greatly improve the spectral efficiency of the system as no pilots are used for estimation. However they suffer from high complexity with low convergence. The more reliable semi-blind techniques refer to blind techniques which incorporate a short a priori known training sequences leading to a better performance and more robustness than purely blind techniques.

Channel estimation methods can be further divided into nonparametric and parametric categories.

3.1 Nonparametric approach (MIMO channel estimation)

In the nonparametric approach, the methods used are unconstrained methods, that is, they have no associated parameters to rely on. Hence, for MIMO channel estimation, the nonparametric approach involves estimating the channel matrix directly, ignoring its multipath structure, or without relying on any physical parameter. ML and MMSE are popular methods used for nonparametric channel estimation.

3.2 Parametric approach (MIMO channel parameter estimation)

A parametric channel estimation approach relies on a physical channel model that is defined by physical parameters related to the signal propagation through the channel. If wisely used, the parametric approach can achieve robust estimation and decrease the dimension of the estimation problem. Using the parametric approach, one can exploit the sparsity and common support properties, which is not possible in the non-parametric approach. Using a model with few number of parameters (sparsity) allows to obtain more robust estimation, and decrease the dimension of the estimation problem [59]. Moreover, by exploiting the exact common support property, less number of pilots is required for the estimation problem [40,41,44], hence achieving better spectral efficiency. In addition, a super resolution performance can be achieved through the parametric approach by using some super resolution estimation methods.

Hence for this approach, we consider the wideband (or frequency selective) channel model given in (1.11), where the channel is described in terms of different propagation path delays and their corresponding gains. For an $N \times M$ MIMO system, the CIR between the nth transmit antenna and the mth receive antenna is given as:

$$h^{(n,m)}(t) = \sum_{l=1}^{L} \alpha_l^{(n,m)} \delta(t - t_l^{(n,m)}), \quad 1 \leq n \leq N, 1 \leq m \leq M \tag{3.1}$$

where L is the total number of propagation paths, $t_l^{(n,m)}$ is the path delay between the nth transmit antenna and mth receive antenna, and $\alpha_l^{(n,m)}$ is the corresponding complex gain.

Assume that antennas at the transmitter side are transmitting pilot symbols on different carriers in order to identify the channels associated with different transmit-receive antenna pairs. Consider the known pulse shape $g(t)$ transmitted at the nth transmit antenna at a constant rate $1/T$ through the medium,

For the mth receive antenna, the received baseband signal is:

$$x_r^{(n,m)}(t) = \sum_{l=1}^{L} \alpha_l^{(n,m)} g(t - t_l^{(n,m)}) + z_r^{(n,m)}(t) \tag{3.2}$$

where $z_r^{(n,m)}(t)$ is an additive white Gaussian noise.

Assuming an exact common support [40, 41, 44], such that $t_l^{(n,m)} = t_l$. Hence the above equation is written as:

$$x_r^{(n,m)}(t) = \sum_{l=1}^{L} \alpha_l^{(n,m)} g(t - t_l) + z_r^{(n,m)}(t) \tag{3.3}$$

Applying the DFT, the Fourier coefficients of the received signal are given by:

$$X_r^{(n,m)}[k] = G[k] \sum_{l=1}^{L} \alpha_l^{(n,m)} e^{-j\frac{2\pi}{T}kt_l} + Z_r^{(n,m)}[k] \tag{3.4}$$

where $G[k]$ is the DFT of pulse $g(t)$ and $Z_r^{(n,m)}[k]$ is the DFT of $z_r^{(n,m)}(t)$. Writing (3.4) in the matrix form

$$\mathbf{x}_r^{(n,m)} = \mathbf{G}\mathbf{V}\boldsymbol{\alpha}^{(n,m)} + \mathbf{z}_r^{(n,m)} \tag{3.5}$$

$$\begin{aligned}
\mathbf{x}_r^{(n,m)} &= \left[X_r^{(n,m)}[0] \dots X_r^{(n,m)}[K-1]\right]^T \\
\mathbf{V} &= [\mathbf{v}_d(t_1) \dots \mathbf{v}_d(t_L)] \\
\mathbf{v}_d(t_l) &= [1 \quad e^{-j\frac{2\pi}{T}t_l} \dots e^{-j\frac{2\pi(K-1)}{T}t_l}]^T \\
\mathbf{G} &= diag\{G[0] \dots G[K-1]\} \\
\boldsymbol{\alpha}^{(n,m)} &= [\alpha_1^{(n,m)} \dots \alpha_L^{(n,m)}]^T \\
\mathbf{z}_r^{(n,m)} &= \left[Z_r^{(n,m)}[0] \dots Z_r^{(n,m)}[K-1]\right]^T
\end{aligned} \tag{3.6}$$

If we multiply by the inverse of $\mathbf{G}$ in (3.5), the model is written as:

$$\mathbf{y}_r^{(n,m)} = \mathbf{V}\boldsymbol{\alpha}^{(n,m)} + \mathbf{z}'^{(n,m)}_r \tag{3.7}$$

or

$$\mathbf{y}_r^{(n,m)} = \sum_{l=1}^{L} \alpha_l^{(n,m)} \mathbf{v}_d(t_l) + \mathbf{z}'^{(n,m)}_r \tag{3.8}$$

The model in (3.8) is common in the array signal processing context where several techniques have been proposed (or applied) to estimate the direction of arrival or time of arrival of signals impinging on an array of sensors [60].

3.2.1 Beamforming techniques

The main idea in beamforming is to scan through a dictionary of time delays and to form a linear combination of the output coefficients using a weighting vector $\mathbf{w}_{bf}$ as

$$q^{(n,m)} = \mathbf{w}_{bf}^H \mathbf{y}_r^{(n,m)} \tag{3.9}$$

and then to measure the output power.

At a certain time instant, for the $N \times M$ antennas, the output power is measured as

$$\begin{aligned} P(\mathbf{w}_{bf}) &= \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} |q^{(n,m)}|^2 \\ &= \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} \mathbf{w}_{bf}^H \mathbf{y}_r^{(n,m)} \mathbf{y}_r^{(n,m)H} \mathbf{w}_{bf} \\ &= \mathbf{w}_{bf}^H \hat{\mathbf{R}}_{\mathbf{y}_r} \mathbf{w}_{bf} \end{aligned} \tag{3.10}$$

where

$$\hat{\mathbf{R}}_{\mathbf{y}_r} = \frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} \mathbf{y}_r^{(n,m)} \mathbf{y}_r^{(n,m)H} \tag{3.11}$$

is the covariance matrix of the observation coefficient vectors.

3.2.1.1 Classical beamformer (Bartlett)

For the classical or conventional beamformer (also known as bartlett beamformer [61]), the weighting vector is chosen to be $\mathbf{v}_d(t)$ ($\mathbf{w}_B = \mathbf{v}_d(t)$). Then the normalized output of the beamformer is given by:

$$P_B(t) = \frac{\mathbf{v}_d(t)^H \hat{\mathbf{R}}_{\mathbf{y}_r} \mathbf{v}_d(t)}{\mathbf{v}_d(t)^H \mathbf{v}_d(t)} \tag{3.12}$$

The maximum output power refers to the true time delay where the time delays t_l correspond to the positions of the L peaks of $P_B(t)$.

The conventional Bartlett beamformer is a simple one, at the cost of limited resolution and interference of close-by delays.

3.2.1.2 Capon

The Minimum variance distortionless response (MVDR) also known as Capon's beamformer [62] was proposed to improve the resolution performance. It is based on minimizing the output power while keeping the signal of the desired delay undistorted:

$$\min_{w_{bf}} \; P(\mathbf{w}_{bf}) \qquad \text{s.t.} \quad \mathbf{w}_{bf}^H \mathbf{v}_d(t) = 1 \tag{3.13}$$

Then the weighting vector is given by:

$$\mathbf{w}_{CAP} = \frac{\hat{\mathbf{R}}_{\mathbf{y}_r}^{-1} \mathbf{v}_d(t)}{\mathbf{v}_d(t)^H \hat{\mathbf{R}}_{\mathbf{y}_r}^{-1} \mathbf{v}_d(t)} \tag{3.14}$$

Hence the power of the capon beamformer is given by

$$P_{CAP}(t) = \frac{1}{\mathbf{v}_d(t)^H \hat{\mathbf{R}}_{\mathbf{y}_r}^{-1} \mathbf{v}_d(t)} \tag{3.15}$$

With the MVDR beamformer, the resolution is greatly enhanced compared to the classical (Bartlett) beamformer.

3.2.2 Subspace based methods

A major breakthrough came out with the subspace based methods, especially the multiple signal classification (MUSIC) [63] and estimation of signal parameter via rotational invariance techniques (ESPRIT) [64] methods that provide a super resolution compared to the limited resolution in the beamforming techniques. The main principle in subspace methods is to divide the received data into signal and orthogonal (or noise) subspaces. This is done through the eigendecomposition of the observation data covariance matrix.

Second order statistical analysis

In the noise free case, the covariance matrix of the observation coefficient vectors can be written as

$$\mathbf{R}_{\mathbf{y}_r} = \mathbf{V}\mathbf{R}_{\boldsymbol{\alpha}}\mathbf{V}^H \tag{3.16}$$

where the channel gains covariance matrix $\mathbf{R}_{\boldsymbol{\alpha}}$ can be estimated as

$$\hat{\mathbf{R}}_{\boldsymbol{\alpha}} = \frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1}^{M}\boldsymbol{\alpha}^{(n,m)}\boldsymbol{\alpha}^{(n,m)H} \tag{3.17}$$

The eigendecomposition of $\mathbf{R}_{\mathbf{y}_r}$ is given as:

$$\mathbf{R}_{\mathbf{y}_r} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H \tag{3.18}$$

The matrix $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues in decreasing order on its diagonal ($\lambda_1 \geq \lambda_2 \cdots \geq \lambda_K$).

$\mathbf{R}_{\mathbf{y}_r}$ is of size $K \times K$, and $\mathbf{V}$ of size $K \times L$ where $K > L$. When $\mathbf{V}$ is full column rank and the channel gain coefficients are uncorrelated, the eigenvectors associated with the first L eigenvalues of $\boldsymbol{\Lambda}$ and column vectors of $\mathbf{V}$ will span the same subspace which is the signal subspace, and the eigenvectors associated with the last $K - L$ eigenvalues of $\boldsymbol{\Lambda}$ will span the orthogonal subspace. Then $\mathbf{U}$ is defined as

$$\mathbf{U} = [\mathbf{U}_s \quad \mathbf{U}_n] \tag{3.19}$$

such that $\mathbf{U}_s$ contains the L eigenvectors spanning the signal subspace and $\mathbf{U}_n$ contains the $K - L$ eigenvectors spanning the orthogonal subspace.

The L largest eigenvalues represent the contribution of signals arriving at L time delays and the remaining $K - L$ eigenvalues represent the noise power. In the absence

of noise, the smallest $K-L$ eigenvalues are zero and the rank of $\mathbf{R}_{\mathbf{y}_r}$ is L.

The eigendecomposition of $\mathbf{R}_{\mathbf{y}_r}$ can be also written as

$$\mathbf{R}_{\mathbf{y}_r} = \mathbf{U}_s \mathbf{\Lambda}_{ss} \mathbf{U}_s^H + \mathbf{U}_n \mathbf{\Lambda}_{nn} \mathbf{U}_n^H \tag{3.20}$$

where $\mathbf{\Lambda}_{ss}$ is an $L \times L$ diagonal matrix containing the L non-zero eigenvalues, and $\mathbf{\Lambda}_{nn}$ is an $(K-L) \times (K-L)$ zero matrix (noise-free case).

In the presence of white noise, the observation covariance matrix is written as

$$\begin{aligned} \mathbf{R}_{\mathbf{y}_r} &= \mathbf{V}\mathbf{R}_{\boldsymbol{\alpha}}\mathbf{V}^H + \sigma_z^2 \mathbf{I}_K \\ &= \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \\ &= \mathbf{U}_s(\mathbf{\Lambda}_{ss} + \sigma_z^2 \mathbf{I}_L)\mathbf{U}_s^H + \mathbf{U}_n(\sigma_z^2 \mathbf{I}_{K-L})\mathbf{U}_n^H \\ &= [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \mathbf{\Lambda}_{ss} + \sigma_z^2 \mathbf{I}_L & 0 \\ 0 & \sigma_z^2 \mathbf{I}_{K-L} \end{bmatrix} \begin{bmatrix} \mathbf{U}_s^H \\ \mathbf{U}_n^H \end{bmatrix} \end{aligned} \tag{3.21}$$

and has a full column rank.

Now the L largest eigenvalues of $\mathbf{R}_{\mathbf{y}_r}$ correspond to signal plus noise and the $K-L$ smallest eigenvalues correspond to noise only. By other words, the first L eigenvalues are greater than σ_z^2 and the last $K-L$ eigenvalues are equal to σ_z^2. The problem is then to define an algorithm to estimate the signal subspace, this can be done by comparing the eigenvalues based on a threshold defined by σ_z^2 or by using some information theoretic criteria.

3.2.2.1 MUSIC

MUSIC is one of the subspace based methods which is based on the eigendecomposition of the observation covariance matrix and relies on the orthogonality between the signal and noise subspaces.

The expression of the observation covariance matrix as shown before is given as:

$$\mathbf{R}_{\mathbf{y}_r} = \mathbf{V}\mathbf{R}_{\boldsymbol{\alpha}}\mathbf{V}^H + \sigma_z^2 \mathbf{I}_K = \mathbf{U}_s(\mathbf{\Lambda}_{ss} + \sigma_z^2 \mathbf{I}_L)\mathbf{U}_s^H + \mathbf{U}_n(\sigma_z^2 \mathbf{I}_{K-L})\mathbf{U}_n^H \tag{3.22}$$

As mentioned before, when matrix $\mathbf{R}_{\boldsymbol{\alpha}}$ is a full rank matrix (channel gains are uncorrelated), the columns of $\mathbf{V}$ and $\mathbf{U}_s$ will span the signal subspace, and as the signal and noise subspaces are orthogonal we have:

$$\mathbf{U}_n^H \mathbf{v}_d(t_l) = \mathbf{0} \quad for \quad l = 1 \dots L \tag{3.23}$$

When $\mathbf{V}$ has a full column rank such that the delays are distinct, and as $\mathbf{R}_{\boldsymbol{\alpha}}$ has a full rank, then we have L delays to estimate.

Given a dictionary of time delays, the algorithm used in MUSIC is to project $\mathbf{v}_d(t)$ for a given delay t chosen from the dictionary on the noise subspace in order to find the true delay, if the delay chosen from the dictionary corresponds to the true one, the projection will be zero. The MUSIC cost function is given by:

$$P_{MUSIC}(t) = \frac{\mathbf{v}_d(t)^H \mathbf{v}_d(t)}{\mathbf{v}_d(t)^H \hat{\mathbf{U}}_n \hat{\mathbf{U}}_n^H \mathbf{v}_d(t)} \tag{3.24}$$

where $\hat{\mathbf{U}}_n$ is the estimated noise subspace matrix.

Path delays then correspond to the L peaks of $P_{MUSIC}(t)$.

The steps involved in the MUSIC algorithm are summarized as follows:

MUSIC

1- Collect NM DFT-domain vectors.

2- Estimate the covariance matrix $\mathbf{R}_{\mathbf{y}_r}$.

3- Apply the eigenvalue decomposition on $\hat{\mathbf{R}}_{\mathbf{y}_r}$ and construct matrix $\hat{\mathbf{U}}_n$.

4- For each value of t, calculate the cost function $P_{MUSIC}(t)$

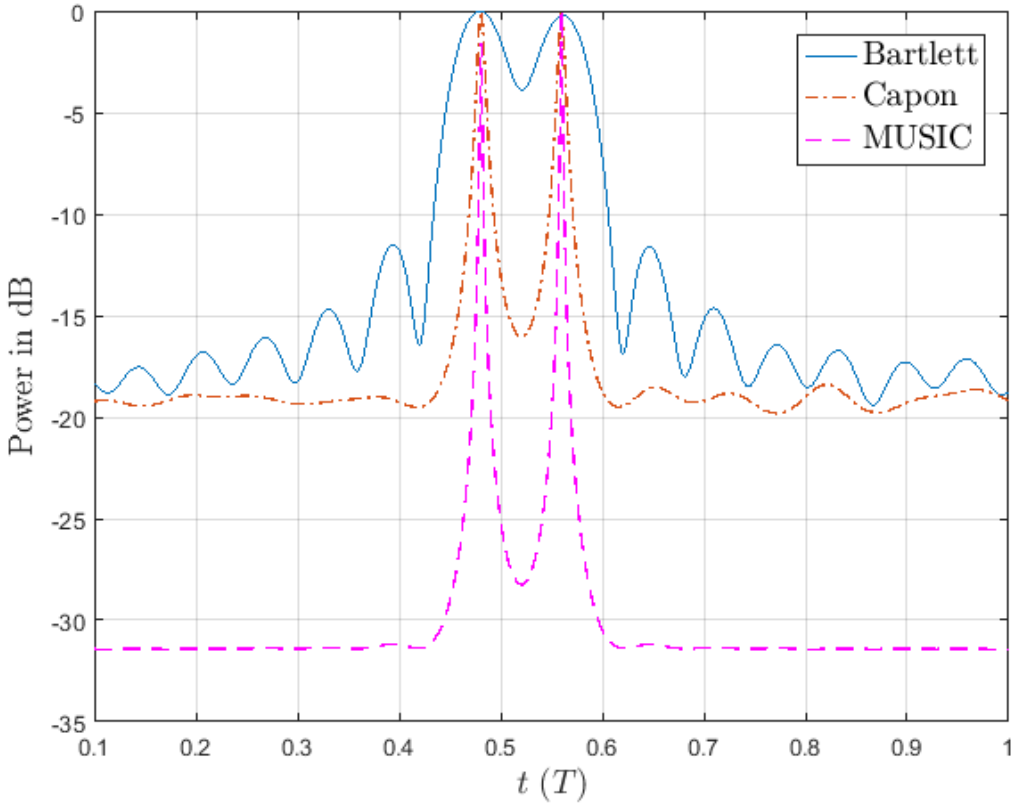

Figure 3.1: Power spectra of Bartlett, Capon and MUSIC, for a 12×12 MIMO system, $K = 32$ and two delays chosen as $\mathbf{t} = [0.48 \quad 0.56]T$ with SNR = 10 dB.

Figure 3.1 shows the spectra of the classical beamformer, capon's beamformer and MUSIC. It is noticed that the Capon beamformer provides higher resolution compared with the classical beamformer, and MUSIC outperforms both the classical and Capon beamformers in terms of resolution.

3.2.2.2 ESPRIT

Another subspace based method is ESPRIT. It relies on the Vandermonde structure of the matrix $\mathbf{V}$ and exploits its shift invariance property.

Let $\mathbf{V}_1$ be the first $K-1$ rows of $\mathbf{V}$ and $\mathbf{V}_2$ be the last $K-1$ rows of $\mathbf{V}$. Then $\mathbf{V}_1$ and $\mathbf{V}_2$ are related as:

$$\mathbf{V}_2 = \mathbf{V}_1\mathbf{\Phi} \tag{3.25}$$

where $\mathbf{\Phi} = diag\{e^{-j\frac{2\pi}{T}t_1}, \ldots, e^{-j\frac{2\pi}{T}t_L}\}$.

As matrices $\mathbf{U}_s$ and $\mathbf{V}$ span the same signal subspace, we have:

$$\mathbf{V} = \mathbf{U}_s\mathbf{J} \tag{3.26}$$

where $\mathbf{J}$ is a non singular matrix.

Let $\mathbf{U}_{s1}$ and $\mathbf{U}_{s2}$ be the first $K-1$ and the last $K-1$ rows of $\mathbf{U}_s$ respectively, we have:

$$\begin{aligned} \mathbf{V}_1 &= \mathbf{U}_{s1}\mathbf{J} \\ \mathbf{V}_2 &= \mathbf{U}_{s2}\mathbf{J} \end{aligned} \tag{3.27}$$

Using the rotational invariance property, we have:

$$\mathbf{U}_{s1}\mathbf{J}\mathbf{\Phi} = \mathbf{U}_{s2}\mathbf{J} \tag{3.28}$$

It follows that

$$\mathbf{U}_{s2} = \mathbf{U}_{s1}\mathbf{J}\mathbf{\Phi}\mathbf{J}^{-1} \tag{3.29}$$

Let $\mathbf{\Psi} = \mathbf{J}\mathbf{\Phi}\mathbf{J}^{-1}$, then $\mathbf{\Phi}$ is the matrix containing the eigenvalues of $\mathbf{\Psi}$, where $\mathbf{\Psi}$ can be estimated using the least squares method as:

$$\hat{\mathbf{\Psi}} = \mathbf{U}_{s1}^{\dagger}\mathbf{U}_{s2} \tag{3.30}$$

Let ψ_l for $l = 1 \ldots L$ be the eigenvalues of $\hat{\mathbf{\Psi}}$, the delays are then given by:

$$\hat{t}_l = -\frac{T}{2\pi} arg(\psi_l) \tag{3.31}$$

3.2.3 Compressive sensing

Compressive sensing (CS) is a breakthrough technology that has gained a lot of attention recently in the field of signal processing, as it provides the ability to recover signals with fewer measurements than what is required by the conventional Shannon-Nyquist theorem. In several applications, according to the Shannon-Nyquist theorem, a large

number of samples is needed to be acquired, then compression can be done prior to storage. However, CS allows to directly aquire fewer number of samples, through an efficient sampling/sensing scheme [65,66]. CS requires two conditions under which it can be applied, the first one is the sparsity of the signal to be recovered, and the second is the incoherence property, which is related to the sensing procedure. The sensing procedure involves linear projections of the sparse signal on the basis vectors that constitute the sensing or measurement matrix. Choosing a convenient measurement matrix can preserve the structure of the signal, the signal can then be recovered from the linear projections using several sparse recovery or reconstruction approaches.

Consider a sparse vector $\mathbf{x}_s$, and a sensing matrix $\mathbf{\Phi}_s$, the observed measurement vector $\mathbf{y}_o$ is given as:

$$\mathbf{y}_o = \mathbf{\Phi}_s \mathbf{x}_s \tag{3.32}$$

To achieve an efficient reconstruction, the sensing matrix $\mathbf{\Phi}_s$ should have low coherence, where the coherence of a matrix is defined as the largest absolute inner product between any two normalized columns of the matrix [67]. The lower the coherence of the sensing matrix, the better the reconstruction quality can be achieved. The incoherence property is related to the restricted isometry property (RIP) [68], which states that for a given k_s-sparse vector $\mathbf{x}_s$, there is a high probability to recover $\mathbf{x}_s$ from $\mathbf{y}_o$ if the sensing matrix $\mathbf{\Phi}_s$ obeys:

$$(1 - \delta_{k_s})||\mathbf{x}_s||_2^2 \leqslant ||\mathbf{\Phi_s x}_s||_2^2 \leqslant (1 + \delta_{k_s})||\mathbf{x}_s||_2^2 \tag{3.33}$$

where $0 < \delta_{k_s} < 1$, and $\mathbf{\Phi}_s$ obeys the RIP of order k_s if δ_{k_s} is not too close to 1 [66], which means that the measurement matrix $\mathbf{\Phi}_s$ approximately preserves the length of the k_s sparse vector.

In several cases, the signal is not sparse itself, but it can be sparsely represented in some basis. Consider a non-sparse signal $\mathbf{z}_s$ which is sparse in an orthogonal basis $\mathbf{\Psi}_s$, such that $\mathbf{z}_s = \mathbf{\Psi}_s \mathbf{x}_s$ where $\mathbf{x}_s$ is the sparse vector. The measurement vector $\mathbf{y}_o$ is then given as:

$$\mathbf{y}_o = \mathbf{\Phi}_s \mathbf{\Psi}_s \mathbf{x}_s \tag{3.34}$$

For signal recovery, it is required that matrix $\mathbf{\Phi}_s \mathbf{\Psi}_s$ has low coherence.

Consider the case where the signal is itself sparse (3.32), the sparse vector $\mathbf{x}_s$ can be recovered through the following optimization problem:

$$\min||\mathbf{x}_s||_0 \quad \text{s.t.} \quad \mathbf{y}_o = \mathbf{\Phi}_s \mathbf{x}_s \tag{3.35}$$

where the l_0 norm returns the number of non-zero elements in a vector. Hence in the above problem, the goal is to find the sparsest vector $\mathbf{x}_s$ subject to $\mathbf{y}_o = \mathbf{\Phi}_s \mathbf{x}_s$.

Although using l_0 norm allows perfect signal reconstruction with high probability, the above optimization problem is highly non-convex and thus NP-hard.

The methods used for signal recovery as a solution for the above non-convex problem are mainly classified into convex optimization (or relaxation) methods and greedy pursuit methods.

3.2.3.1 L1 minimization

The convex relaxation methods replace the l_0 norm with l_1 norm, the optimization problem turns out

$$\min||\mathbf{x}_s||_1 \quad \text{s.t.} \quad \mathbf{y}_o = \mathbf{\Phi}_s \mathbf{x}_s \tag{3.36}$$

The interest in using the l_1 norm is that it provides sparse solutions and results in an convex optimization problem, hence computationally tractable, where it can be handled as a linear programming problem.

In the presence of noise, the optimization problem becomes:

$$\min||\mathbf{x}_s||_1 \quad \text{s.t.} \quad ||\mathbf{y}_o - \mathbf{\Phi}_s \mathbf{x}_s||_2 \leqslant \epsilon \tag{3.37}$$

3.2.3.2 Greedy pursuit methods

Another class of methods includes greedy pursuit methods. Although convex optimization methods provide better performance when compared to the greedy methods with higher probability of signal recovery, the greedy methods are much faster [69].

Greedy methods work by selecting the elements of a dictionary (such as sensing matrix) that best approximate the measurement vector in an iterative manner. Methods like matching pursuit (MP) [70] and orthogonal matching pursuit (OMP) [71] are among the most popular greedy methods. In the MP method, the measurement vector

is set as a residual, and at each iteration, the signal is correlated with the column vectors given in the sensing matrix, the contribution of the vector with highest correlation is then removed from the residual.

Matching pursuit (MP)

Given (3.32), consider the problem of recovering $\mathbf{x}_s$ from $\mathbf{y}_o$, with $\mathbf{\Phi}_s$ the dictionary, the steps of the MP algorithm are summarized as follows:

Matching Pursuit

1- Initialize the residual vector as $\mathbf{r}_0 = \mathbf{y}_o$ and the iteration counter $i = 1$.

2- Assume that all the columns of $\mathbf{\Phi}_s$ are normalized, the algorithm selects the column that has the highest correlation with $\mathbf{y}_o$ by solving the following maximization problem

$$\zeta_i = \underset{n_s=1...N_s}{\operatorname{argmax}} | < \mathbf{r}_{i-1}, \boldsymbol{\phi}_{n_s} > | \tag{3.38}$$

where ζ_i is the index of the selected vector and N_s is the number of column vectors in $\mathbf{\Phi}_s$.

3- Let $c_i = | < \mathbf{r}_{i-1}, \boldsymbol{\phi}_{\zeta_i} > |$, the residual is updated as:

$$\mathbf{r}_i = \mathbf{r}_{i-1} - c_i \boldsymbol{\phi}_{\zeta_i} \tag{3.39}$$

where $\boldsymbol{\phi}_{\zeta_i}$ is the vector selected at the ith iteration.

4- Increment i and return to step 2, continue until the residual value $||\mathbf{r}_i||_2$ is below a predefined threshold or stop after K_s iterations if K_s (number of non-zero elements in $\mathbf{x}_s$) is known a priori.

Orthogonal matching pursuit (OMP)

The OMP method provides better performance than the MP method with higher complexity. In the OMP method, the basis vector is selected in the same manner as

in the MP method, however for each iteration, the measurement vector is projected on the subspace spanned by the current and previously selected vectors. The contribution of these vectors is then removed. In this manner, the previously selected vectors will not be selected again in the subsequent iterations. The steps of the OMP method are summarized as follows:

Orthogonal Matching Pursuit

1- Initialize the residual vector as $\mathbf{r}_0 = \mathbf{y}_o$ and the iteration counter $i = 1$.

2- Let φ_i be the set containing the indices of vectors selected until the ith iteration (where $\varphi_0 = \emptyset$) and ζ_i the index of vector selected at the ith iteration by solving the following maximization problem

$$\zeta_i = \underset{n_s=1...N_s}{\operatorname{argmax}} | < \mathbf{r}_{i-1}, \boldsymbol{\phi}_{ns} > | \tag{3.40}$$

3- Set $\varphi_i = \varphi_{i-1} \cup \zeta_i$

4- Find $\mathbf{c}_i$, the orthogonal projection of $\mathbf{y}_o$ on the subspace spanned by the vectors having their indices in φ_i as:

$$\mathbf{c}_i = \boldsymbol{\Phi}_{s_i}^{\dagger} \mathbf{y}_o \tag{3.41}$$

where $\boldsymbol{\Phi}_{s_i}$ is the matrix of vectors selected up to the ith iteration (having their indices in φ_i).

5- The approximation of the signal and the residual are then given as:

$$\begin{aligned} \mathbf{y}_{o_i} &= \boldsymbol{\Phi}_{s_i} \mathbf{c}_i \\ \mathbf{r}_i &= \mathbf{y}_o - \mathbf{y}_{o_i} \end{aligned} \tag{3.42}$$

6- Increment i and return to step 2, continue until the residual value $||\mathbf{r}_i||_2$ is below a predefined threshold or stop after K_s iterations if K_s (number of non-zero elements in $\mathbf{x}_s$) is known a priori.

OMP deals with the case of single measurement vector (SMV). For the case of multiple measurement vector (MMV), a method is proposed in [72] called simultaneous orthogonal matching pursuit (SOMP), given that the sparse input vectors share the

same location of non-zero elements. Consider again the model given in (3.5), the matrix of Fourier coefficient vectors at the receive antenna m due to the N transmit antennas is given as:

$$\mathbf{X}_r^{(m)} = \mathbf{G}\mathbf{V}\boldsymbol{\Upsilon}^{(m)} + \mathbf{Z}_r^{(m)} \tag{3.43}$$

where $\mathbf{X}_r^{(m)} = [\mathbf{x}_r^{(1,m)} \dots \mathbf{x}_r^{(N,m)}]$, $\boldsymbol{\Upsilon}^{(m)} = [\boldsymbol{\alpha}^{(1,m)} \dots \boldsymbol{\alpha}^{(N,m)}]$, and $\mathbf{Z}_r^{(m)} = [\mathbf{z}_r^{(1,m)} \dots \mathbf{z}_r^{(N,m)}]$.

Arranging the Fourier coefficient matrices of the different receive antennas as follows:

$$\mathbf{X}_r = [\mathbf{X}_r^{(1)} \dots \mathbf{X}_r^{(M)}] \tag{3.44}$$

we have

$$\mathbf{X}_r = \mathbf{G}\mathbf{V}\boldsymbol{\Upsilon} + \mathbf{Z}_r \tag{3.45}$$

where $\boldsymbol{\Upsilon} = [\boldsymbol{\Upsilon}^{(1)} \dots \boldsymbol{\Upsilon}^{(M)}]$ and $\mathbf{Z}_r = [\mathbf{Z}_r^{(1)} \dots \mathbf{Z}_r^{(M)}]$.

Multiplying by the inverse of $\mathbf{G}$, we have

$$\mathbf{Y}_r = \mathbf{V}\boldsymbol{\Upsilon} + \mathbf{Z}_r^{'} \tag{3.46}$$

where $\mathbf{Z}_r^{'} = \mathbf{G}^{-1}\mathbf{Z}_r$.

Given the model in (3.46), the matrix $\mathbf{Y}_r$ can be written as:

$$\mathbf{Y}_r = \mathbf{V}_{Q_l}\boldsymbol{\Upsilon}_{Q_l} + \mathbf{Z}_r \tag{3.47}$$

where $\mathbf{V}_{Q_l}$ is given as

$$\mathbf{V}_{Q_l} = [\mathbf{v}_d(t_{b_0}) \dots \mathbf{v}_d(t_{b_{Q_l}})] \tag{3.48}$$

and

$$t_{b_{q_l}} = \frac{q_l}{Q_l}T \quad \text{for} \quad q_l = 0 \dots Q_l \quad \text{s.t.} \quad Q_l >> L \tag{3.49}$$

The matrix $\boldsymbol{\Upsilon}_{Q_l}$ is a sparse matrix, such that its column vectors have the same indices of L non-zero gain coefficients associated to L column vectors in $\mathbf{V}_{Q_l}$, these L column vectors in $\mathbf{V}_{Q_l}$ are the column vectors of $\mathbf{V}$.

For sparse recovery, or for delays estimation, the steps of the SOMP method are given as follows:

Simultaneous Orthogonal Matching Pursuit

1- Initialize the residual matrix as $\mathbf{R}_0 = \mathbf{Y}_r$ and the iteration counter $i = 1$.

2- Let φ_i be the set containing the indices of vectors selected from $\mathbf{V}_{Q_l}$ until the ith iteration (where $\varphi_0 = \emptyset$) and ζ_i the index of vector selected at the ith iteration by solving the following maximization problem

$$\zeta_i = \underset{q_l=1\ldots Q_l}{\operatorname{argmax}} ||\mathbf{R}_{i-1}^H \mathbf{v}_d(t_{b_{q_l}})||_1 \tag{3.50}$$

3- Set $\varphi_i = \varphi_{i-1} \cup \zeta_i$

4- Find $\mathbf{C}_i$, the orthogonal projection of $\mathbf{Y}_r$ on the subspace spanned by the vectors having their indices in φ_i as:

$$\mathbf{C}_i = \mathbf{O}_i^\dagger \mathbf{Y}_r \tag{3.51}$$

where $\mathbf{O}_i$ is the matrix of vectors selected up to the ith iteration (having their indices in φ_i).

5- The approximation of the signal and the residual are then given as:

$$\begin{aligned} \mathbf{Y}_{r_i} &= \mathbf{O}_i \mathbf{C}_i \\ \mathbf{R}_i &= \mathbf{Y}_r - \mathbf{Y}_{r_i} \end{aligned} \tag{3.52}$$

6- Increment i and return to step 2, continue until the residual value $||\mathbf{R}_i||_F$ is below a predefined threshold or stop after L iterations if L is known a priori.

3.3 Conclusion

In this chapter, an overview of several channel parameter estimation techniques is provided. The estimation approaches are divided into parametric and nonparametric. The nonparametric approach estimates the channel matrix without referring to any physical propagation parameters. The parametric approach relies on physical channel models to estimate channel parameters. The chapter divides the channel parameter estimation methods into beamforming methods, subspace based methods and compressive sensing based methods, focusing on the parametric approach for channel estimation.

Chapter 4

Clustered MIMO channel delay estimation

In chapter 2, a clustered MIMO channel model is presented, where in addition to the sparsity property of a wireless channel, the channel multipath components are modeled as clusters of multirays. This is a typical channel configuration that future MIMO wireless communication systems will have to cope with. A cluster of multirays can be characterized by its mean delay and its delay spreading. The work in this chapter focuses on mean delays estimation. Two approaches are proposed, a subspace based approach and a compressive sensing based approach. The compressive sensing based approach is based on the first order Taylor expansion around the mean delay parameter where a modified SOMP method is proposed to estimate the channel mean delays. The subspace based approach is based on higher order Taylor expansion around the mean delay parameter, where a subspace based method is proposed to estimate the cluster mean delays. The latter approach uses the minimum description length (MDL) criterion [73, 74] to track the signal subspace in order to estimate its effective dimension that changes depending on the wireless environment and the SNR value.

4.1 Channel and system model

A sparse clustered MIMO channel model with common support is considered (model in 2.17). For an $N \times M$ MIMO system, each antenna at the emitter side is transmitting pilot symbols on a different carrier in order to identify the channels associated with different transmit-receive antenna pairs. Consider a known pulse shape $g(t)$ transmitted at the nth transmit antenna with a constant rate $1/T$ through a medium consisting of L clusters. For the mth receive antenna, the received baseband signal is:

$$x^{(n,m)}(t) = \sum_{l=1}^{L} \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)} g(t - (t_l + \tau_{lp}^{(n,m)})) + z^{(n,m)}(t) \tag{4.1}$$

with $z^{(n,m)}(t)$ an additive white Gaussian noise.

Applying the DFT, the Fourier coefficients of the received signal are given by:

$$X^{(n,m)}[k] = G[k] \sum_{l=1}^{L} \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)} e^{-j\frac{2\pi}{T}k(t_l+\tau_{lp}^{(n,m)})} + Z^{(n,m)}[k] \tag{4.2}$$

for $k = -K/2 + 1, \ldots, K/2$, where K is the considered number of Fourier coefficients, $G[k]$ is the DFT of pulse $g(t)$ and $Z^{(n,m)}[k]$ is the DFT of $z^{(n,m)}(t)$. Note that K is an even integer.

Let $v_k(t) = G[k]e^{-j\frac{2\pi}{T}kt}$; (4.2) can be rewritten as:

$$X^{(n,m)}[k] = \sum_{l=1}^{L} \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)} v_k(t_l + \tau_{lp}^{(n,m)}) + Z^{(n,m)}[k] \tag{4.3}$$

4.2 CS based channel delay estimation approach

Equation (4.3) can be written in the matrix form as:

$$\mathbf{x}^{(n,m)} = \mathbf{V}^{(n,m)} \bar{\boldsymbol{\alpha}}^{(n,m)} + \mathbf{z}^{(n,m)} \tag{4.4}$$

where

$$\begin{aligned}
\mathbf{x}^{(n,m)} &= \left[X^{(n,m)}[-k/2+1]\dots X^{(n,m)}[k/2]\right]^T \\
\mathbf{V}^{(n,m)} &= [\mathbf{V}_1^{(n,m)}\dots\mathbf{V}_L^{(n,m)}] \\
\mathbf{V}_l^{(n,m)} &= [\mathbf{v}(t_{l1}^{(n,m)})\dots\mathbf{v}(t_{lP}^{(n,m)})] \\
\mathbf{v}(t_{lp}^{(n,m)}) &= \left[v_{-K/2+1}(t_{lp}^{(n,m)})\dots v_{K/2}(t_{lp}^{(n,m)})\right]^T \\
\bar{\boldsymbol{\alpha}}^{(n,m)} &= [{\boldsymbol{\alpha}_1^{(n,m)}}^T\dots{\boldsymbol{\alpha}_L^{(n,m)}}^T]^T \\
\boldsymbol{\alpha}_l^{(n,m)} &= [\alpha_{l1}^{(n,m)}\dots\alpha_{lP}^{(n,m)}]^T \\
\mathbf{z}^{(n,m)} &= \left[Z^{(n,m)}[-k/2+1]\dots Z^{(n,m)}[k/2]\right]^T \\
t_{lp}^{(n,m)} &= t_l + \tau_{lp}^{(n,m)}
\end{aligned} \tag{4.5}$$

Arranging the Fourier coefficient vectors for the different receive antennas in matrices as $\mathbf{X} = [\mathbf{X}^{(1)}\dots\mathbf{X}^{(M)}]$ for $\mathbf{X}^{(m)} = [\mathbf{x}^{(1,m)}\dots\mathbf{x}^{(N,m)}]$, $\mathbf{X}$ can be written as

$$\mathbf{X} = \bar{\mathbf{V}}\bar{\mathbf{A}} + \mathbf{Z} \tag{4.6}$$

where $\bar{\mathbf{V}} = [\mathbf{V}^{(1)}\dots\mathbf{V}^{(M)}]$ for $\mathbf{V}^{(m)} = [\mathbf{V}^{(1,m)}\dots\mathbf{V}^{(N,m)}]$.

Given the model in (4.6), $\mathbf{X}$ can be written as:

$$\mathbf{X} = \mathbf{V}_{Q_i}\bar{\mathbf{A}}_{Q_i} + \mathbf{Z} \tag{4.7}$$

$$\mathbf{V}_{Q_i} = [\mathbf{v}(t_{b_0})\dots\mathbf{v}(t_{b_{Q_i}})] \tag{4.8}$$

and

$$t_{b_{q_i}} = \frac{q_i}{Q_i}T \quad \text{for} \quad q_i = 0\dots Q_i \quad \text{s.t.} \quad Q_i >> NMLP \tag{4.9}$$

where $P > P_l$ for all l.

The matrix $\bar{\mathbf{A}}_{Q_i}$ is a sparse matrix, since multiray delays associated to a scatterer are not the same on different transmit-receive antenna pairs, the column vectors of $\bar{\mathbf{A}}_{Q_i}$ do not share the same indices of the non-zero gain coefficients. However they share the same indices of the gain coefficients associated to the cluster mean delays, as the multirays associated to a given scatterer are assumed to share the same mean delay on the different transmit-receive antenna pairs.

Back to (4.3), as the deviations $\tau_{lp}^{(n,m)}$ are considered small, the first order Taylor expansion of (4.3) gives:

$$X^{(n,m)}[k] \approx \sum_{l=1}^{L}\sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)}\big(v_k(t_l) + \tau_{lp}^{(n,m)} d_k(t_l)\big) + Z^{(n,m)}[k] \tag{4.10}$$

where $d_k(t_l) = (-j\frac{2\pi}{T}k)v_k(t_l)$.

Equation (4.10) can be written as:

$$\tilde{X}^{(n,m)}[k] = \sum_{l=1}^{L} a_l^{(n,m)} v_k(t_l) + b_l^{(n,m)} d_k(t_l) + Z^{(n,m)}[k] \tag{4.11}$$

where $a_l^{(n,m)} = \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)}$ and $b_l^{(n,m)} = \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)}\tau_{lp}^{(n,m)}$.

Hence the vector $\tilde{\mathbf{x}}^{(n,m)}$ of the Fourier coefficients can be written as:

$$\tilde{\mathbf{x}}^{(n,m)} = \sum_{l=1}^{L} a_l^{(n,m)}\mathbf{v}(t_l) + b_l^{(n,m)}\mathbf{d}(t_l) + \mathbf{z}^{(n,m)} \tag{4.12}$$

where

$$\begin{aligned} \tilde{\mathbf{x}}^{(n,m)} &= \big[\tilde{X}^{(n,m)}[-K/2+1]\ldots\tilde{X}^{(n,m)}[K/2]\big] \\ \mathbf{v}(t_l) &= [v_{-K/2+1}(t_l)\ldots v_{K/2}(t_l)]^T \\ \mathbf{d}(t_l) &= [v_{-K/2+1}(t_l)\ldots d_{K/2}(t_l)]^T \end{aligned} \tag{4.13}$$

Due to the common support assumption on mean delays. The vector $\mathbf{v}(t_l)$ is contained in all the Fourier coefficient vectors in $\mathbf{X}$ (4.7). The concept used in the SOMP method can then be used to estimate the mean delays, with a modified selection step. The steps of the proposed mean delay estimation method are given as:

Algorithm of the proposed CS based method

1. Initialize the residual matrix as $\mathbf{R}_0 = \mathbf{X}$ and the iteration counter $i = 1$.

2. Let φ_i be the set containing the indices of vectors selected from $\mathbf{V}_{Q_i}$ until the ith iteration (where $\varphi_0 = \emptyset$) and ζ_i the index of vector selected at the ith iteration by solving the following maximization problem

$$\zeta_i = \underset{q_i=1\ldots Q_i}{\operatorname{argmax}} \left(||\mathbf{R}_{i-1}^H \mathbf{v}(t_{b_{q_i}})||_2 + ||\mathbf{R}_{i-1}^H \mathbf{d}(t_{b_{q_i}})||_2 \right) \tag{4.14}$$

3. Set $\varphi_i = \varphi_{i-1} \cup \zeta_i$.

4. Find $\mathbf{C}_i$, the orthogonal projection of $\mathbf{X}$ on the subspace spanned by the vectors having their indices in φ_i as:

$$\mathbf{C}_i = \mathbf{O}_i^{\dagger} \mathbf{X} \tag{4.15}$$

   where $\mathbf{O}_i$ is the matrix of vectors selected up to the ith iteration (having their indices in φ_i).

5. The approximation of the signal and the residual are then given as:

$$\begin{aligned} \mathbf{X}_i &= \mathbf{O}_i \mathbf{C}_i \\ \mathbf{R}_i &= \mathbf{X} - \mathbf{X}_i \end{aligned} \tag{4.16}$$

6. Increment i and return to step 2 if $i \leq L$, assuming that the number of mean delays L is known a priori.

The proposed estimation method differs from the classical SOMP method in the selection step, which is adapted according to the expression of the Fourier coefficient vectors given in (4.12) by taking into account the vector $\mathbf{v}(t_{b_{q_i}})$ and its derivative $\mathbf{d}(t_{b_{q_i}})$ for selection.

4.3 Subspace based channel delay estimation approach

As several subspace based methods are high resolution methods, we find it interesting to incorporate more details into the model rederived by means of Taylor expansion, seeking for better estimation performance. To achieve this, we consider the U^{th} order Taylor expansion of (4.3), given as:

$$X^{(n,m)}[k] = \sum_{l=1}^{L}\sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)}\big(v_k(t_l) + \sum_{u=1}^{U} \frac{(\tau_{lp}^{(n,m)})^u}{u!} v_k^{(u)}(t_l) + R_U(\tau_{lp}^{(n,m)})\big) + Z^{(n,m)}[k] \quad (4.17)$$

where $v_k^{(u)}(t_l)$ is the u^{th} order derivative of $v_k(t_l)$. In this last equation, $R_U(\tau_{lp}^{(n,m)})$ is the remaining term in the Taylor approximation, which is considered small, such that its impact in the approximation can be neglected.

Equation (4.17) can be written in the following form:

$$X^{(n,m)}[k] = G[k]\sum_{l=1}^{L}\sum_{u=0}^{U} a_{l,u}^{(n,m)} v_k^{(u)}(t_l) + Z^{(n,m)}[k] \quad (4.18)$$

where

$$a_{l,u}^{(n,m)} = \sum_{p=1}^{P_l} \alpha_{lp}^{(n,m)} \frac{(\tau_{lp}^{(n,m)})^u}{u!}. \quad (4.19)$$

Equation (4.18) is written in the following matrix form:

$$\mathbf{x}^{(n,m)} = \mathbf{W}\mathbf{a}^{(n,m)} + \mathbf{z}^{(n,m)} \quad (4.20)$$

where

$$\begin{aligned}
\mathbf{x}^{(n,m)} &= \left[X^{(n,m)}[-K/2+1],\ldots,X^{(n,m)}[K/2]\right]^T \\
\mathbf{W} &= [\mathbf{V}^{(0)}\mathbf{V}^{(1)}\ldots\mathbf{V}^{(U)}] \\
\mathbf{V}^{(u)} &= [\mathbf{v}^{(u)}(t_1),\ldots,\mathbf{v}^{(u)}(t_L)] \\
\mathbf{v}^{(u)}(t_l) &= [v^{(u)}_{-K/2+1}(t_l),\ldots,v^{(u)}_{K/2}(t_l)]^T \\
v^{(u)}_k(t_l) &= (-j\frac{2\pi}{T}k)^u v_k(t_l), \\
\mathbf{a}^{(n,m)} &= \left[\mathbf{a}_0^{(n,m)^T},\ldots,\mathbf{a}_U^{(n,m)^T}\right]^T \\
\mathbf{a}^{(n,m)}_u &= [a^{(n,m)}_{1,u},\ldots,a^{(n,m)}_{L,u}]^T \\
\mathbf{z}^{(n,m)} &= \left[Z^{(n,m)}[-K/2+1],\ldots,Z^{(n,m)}[K/2]\right]^T
\end{aligned} \tag{4.21}$$

with $\mathbf{x}^{(n,m)} \in \mathbb{C}^{K\times 1}$, $\mathbf{W} \in \mathbb{C}^{K\times(U+1)L}$, and $\mathbf{a}^{(n,m)} \in \mathbb{C}^{(U+1)L\times 1}$.

Matrix $\mathbf{W}$ is the same for all the transmit-receive antenna pairs, as the mean delays of each cluster are assumed to be the same for all the transmit-receive antenna pairs.

The matrix of Fourier coefficient vectors at antenna m due to the N transmit antennas is given by:

$$\mathbf{X}^{(m)} = [\mathbf{x}^{(1,m)}... \quad \mathbf{x}^{(N,m)}], \tag{4.22}$$

which can be written as:

$$\mathbf{X}^{(m)} = \mathbf{W}\mathbf{A}^{(m)} + \mathbf{Z}^{(m)}, \tag{4.23}$$

where $\mathbf{A}^{(m)} = [\mathbf{a}^{(1,m)}\ldots \quad \mathbf{a}^{(N,m)}]$ and $\mathbf{Z}^{(m)} = [\mathbf{z}^{(1,m)}\ldots\mathbf{z}^{(N,m)}]$

Arranging the Fourier coefficient matrices of different receive antennas as follows:

$$\mathbf{X} = [\mathbf{X}^{(1)}\mathbf{X}^{(2)}\ldots \quad \mathbf{X}^{(M)}] \tag{4.24}$$

$\mathbf{X}$ can be written as:

$$\mathbf{X} = \mathbf{W}\mathbf{A} + \mathbf{Z} \tag{4.25}$$

where $\mathbf{A} = [\mathbf{A}^{(1)}\mathbf{A}^{(2)}\ldots \quad \mathbf{A}^{(M)}]$ and $\mathbf{Z} = [\mathbf{Z}^{(1)}\ldots\mathbf{Z}^{(M)}]$.

4.3.1 Second order statistical analysis

The covariance matrix $\mathbf{R_X}$ of the observation Fourier coefficient vectors in $\mathbf{X}$ can be estimated as:

$$\hat{\mathbf{R}}_{\mathbf{X}} = \frac{1}{NM}\mathbf{X}\mathbf{X}^H. \tag{4.26}$$

For the noise-free case:

$$\mathbf{R_X} = \mathbf{W}\mathbf{R_a}\mathbf{W}^H, \tag{4.27}$$

where the covariance matrix $\mathbf{R_a}$ is given as:

$$\mathbf{R_a} = \frac{1}{NM}\sum_{n=1}^{N}\sum_{m=1}^{M}\mathbf{a}^{(n,m)}\mathbf{a}^{(n,m)^H} \tag{4.28}$$

Assuming that t_l for $l = 1, \ldots, L$ are different from one another, with $K > (U+1)L$ and $NM > (U+1)L$, it can be noted from the definition of matrix $\mathbf{W}$ that the dimension of its column space is equal to $(U+1)L$ with a full rank property.

It follows that the dimension of the signal subspace is equal to the rank of $\mathbf{R_a}$. In the following we will show that $\mathbf{R_a}$ is a full rank matrix of rank equal to $(U+1)L$, which implies that the dimension of the signal subspace is $(U+1)L$.

Proof Elements. Consider one cluster ($L = 1$) of mean delay t_1; for the sake of simplification, notations are abbreviated as $t_1 = t$, $P_1 = P$, $\alpha_{1,p}^{(n,m)} = \alpha_p$, $\tau_{lp}^{(n,m)} = \tau_p$, and $\mathbf{a}^{(n,m)}$ is a $(U+1)$-length random column vector which is replaced by a new vector $\mathbf{a}$ defined as:

$$\mathbf{a} = \big[\sum_{p=1}^{P}\alpha_p \quad \sum_{p=1}^{P}\alpha_p\frac{\tau_p^u}{u!}\cdots\sum_{p=1}^{P}\alpha_p\frac{\tau_p^U}{U!}\big]^T, \tag{4.29}$$

under the following assumptions:

- α_p for $p \in \{1, \ldots, P\}$ are independent complex random variables such that $\forall p, q \in \{1, \ldots, P\}$, $E[\alpha_p] = 0$; $E[\alpha_p\alpha_q^*] = \sigma_p^2$ if $p = q$, 0 otherwise; and $\sum_{p=1}^{P} E[|\alpha_p|^2] = 1$.
- τ_p for $p \in \{1, \ldots, P\}$ are independent real random variables having the same distribution.
- α_p and τ_p are independent.

Defining indices $u, v \in \{0, 1 \dots U\}$, the $(u+1, v+1)^{th}$ element of matrix $E[\mathbf{a}\mathbf{a}^H]$ is given as:

$$[E[\mathbf{a}\mathbf{a}^H]]_{(u+1,v+1)} = E\big[\sum_{p=1}^{P} \alpha_p \frac{\tau_p^u}{u!} \sum_{q=1}^{P} \alpha_q \frac{\tau_q^v}{v!}\big] = \frac{1}{u!v!} \sum_{p=1}^{P} \sum_{q=1}^{P} E[\alpha_p \alpha_q^* \tau_p^u \tau_q^v]. \tag{4.30}$$

It follows that:

$$[E[\mathbf{a}\mathbf{a}^H]]_{(u+1,v+1)} = \frac{1}{u!v!} \sum_{p=1}^{P} E[|\alpha_p|^2] E[\tau_p^{u+v}] = \frac{E[\tau_p^{u+v}]}{u!v!}. \tag{4.31}$$

For uniformly distributed τ_p between $[-\frac{s}{2}, \frac{s}{2}]$, we have:

$$[E[\mathbf{a}\mathbf{a}^H]]_{(u+1,v+1)} = \begin{cases} (\frac{s}{2})^{u+v} \frac{1}{u!v!} \frac{1}{u+v+1} & \text{if } u+v \text{ is even} \\ 0 & \text{if } u+v \text{ is odd} \end{cases}. \tag{4.32}$$

According to the elements of matrix $E[\mathbf{a}\mathbf{a}^H]$, it follows that its determinant is different from zero, $(det(E[\mathbf{a}\mathbf{a}^H]) \neq 0)$ and as it is a square matrix, then it is a full rank matrix with rank equal to $U + 1$. □

4.3.2 Channel delay estimation

The model in (4.25) suggests the possibility of using subspace based parameter estimation techniques. In the existing works with an exact common support [40,44], subspace methods such as MUSIC and ESPRIT are applied directly to estimate the delays. Although a non-strictly exact common support model is also taken into account in [40], the conventional ESPRIT and MUSIC methods are used with this model as well. In this thesis, in accordance with the delay spreading phenomenon, the channel model is reformulated, and a new subspace based method is proposed to cope with such a model. The principle of subspace based estimation methods is to find the parameters characterizing $\mathbf{W}$ by searching for the parameter vectors that best match the signal subspace, where the signal subspace can be estimated from the measurements, based on the eigendecomposition of the observation covariance matrix. The problem of estimating the correct dimension of signal or orthogonal subspaces is of great importance in such methods, especially in scattering situations where this problem becomes more complicated.

4.3.2.1 Asymptotic analysis of the signal subspace dimension

In the exact common support case with no delay spreading, and in the noise-free case, the covariance matrix $\mathbf{R_X}$ is of rank L. However, according to the U^{th} order Taylor-approximated model (4.18), due to the delay spreading, the rank of $\mathbf{R_X}$ is increased to $(U+1)L$, leading to a larger dimension of the signal subspace. More generally, considering the exact model (4.3) with additive noise, most of the energy of the signal will be concentrated in few eigenvalues of the covariance matrix depending on the standard deviation of the delay spreading, and on the level of noise. Due to these factors, different dimensions of the signal subspace can be obtained. It follows that determining the appropriate value of U is of great importance, and it should be chosen carefully depending on these factors in order to estimate the so-called effective dimension of the signal subspace. Consequently, based on the definition of matrix $\mathbf{W}$, an appropriate parametric expression of vectors can be chosen in the cluster mean delays estimation procedure.

4.3.2.2 Signal subspace tracking

From the generalized model in (4.18), the dimension of the signal subspace is $(U+1)L$, where U is the order of the Taylor expansion that should be estimated to define the effective dimension of the signal subspace. This can be done through analyzing the eigenvalues of the estimated covariance matrix. The MDL and the Akaike information criterion (AIC) [75] are the two most famous information theoretic criteria for estimating the number of sources impinging on an array of sensors [76]. Such methods test the eigenvalues of the estimated covariance matrix where the number of sources is estimated based on minimizing a derived criterion. In general, the MDL is preferred as it is considered to have better performance than AIC [77]. The MDL provides better performance in the presence of spatially and temporally white noise [78]. The AIC provides better estimation in low SNR and small observation sample conditions, however this estimator is not asymptotically consistent. The MDL provides better estimation in the case of large number of observation samples, and it is asymptotically consistent. Therefore, the MDL criterion is chosen in this thesis to estimate the effective dimension of the signal subspace. In fact, the MDL criterion will provide $(U+2)L$ as the dimension of the signal subspace. Despite the fact that the error term $R_U(\tau_{lp}^{(n,m)})$ in (4.17) is assumed to have no effect in defining the effective signal subspace, it will be detected by the MDL criterion as it is not equal to zero. Hence, during the signal subspace tracking process, the effective dimension of the signal subspace is estimated by subtracting L from the value provided by the MDL criterion. In fact, for $L > 1$, each of the L clusters is assumed to occupy the same degrees of freedom $(U+1)$ in the signal subspace, which should give rise to an overall subspace dimension equal to

$(U+1)L$, and hence $(U+2)L$ would be provided by the MDL criterion. However in practice, in the presence of noise, the MDL algorithm will sometimes provide an integer in the set $\{(U+1)L+1,(U+1)L+2,\dots(U+2)L\}$. The proposed solution for this case is to estimate U according to the following rule:

$$U=\left\lfloor \frac{MDLV}{L}-1 \right\rfloor, \tag{4.33}$$

where $MDLV$ is the value obtained by the MDL criterion and $\lfloor . \rfloor$ is the floor function.

4.3.2.3 Cluster mean delay estimation

The covariance matrix of $\mathbf{X}$ is estimated as in (4.26). As shown before, the delay spreading gives rise to an increase in the signal subspace dimension to $(U+1)L$, where the value of U changes according to the SNR as well as the standard deviation of the delay spreading. It follows that the eigenvectors associated with the $(U+1)L$ largest eigenvalues define the signal subspace. Hence, the eigendecomposition of the estimated covariance matrix provides the estimate of the noise subspace matrix $\mathbf{U}_n$. The estimated noise subspace matrix $\hat{\mathbf{U}}_n$ is composed of the $K-(U+1)L$ eigenvectors associated with the $K-(U+1)L$ smallest eigenvalues of $\hat{\mathbf{R}}_{\mathbf{X}}$.

For the case of no delay spreading, the conventional MUSIC can be applied to estimate the delays where the vector $\mathbf{v}^{(0)}(t)$ is projected on the noise subspace. In the case of delay spreading, as can be noticed from (4.18), vectors $\mathbf{v}^{(u)}(t)$ for $u=0\dots,U$ define the signal subspace. A new cost function is then proposed, such that vectors $\mathbf{v}^{(u)}(t)$ are jointly projected onto the estimated noise subspace of dimension $K-(U+1)L$. Hence, the following cost function can be defined:

$$P_{ss_1}(t)=\frac{1}{\sum_{u=0}^{U}\mathbf{v}^{(u)}(t)^H\hat{\mathbf{U}}_n\hat{\mathbf{U}}_n^H\mathbf{v}^{(u)}(t)} \tag{4.34}$$

where

$$\begin{aligned}\mathbf{v}^{(u)}(t)&=[v^{(u)}_{-K/2+1}(t),\dots,v^{(u)}_{K/2}(t)]^T,\\ v^{(u)}_k(t)&=(-j\frac{2\pi}{T}k)^u v_k(t).\end{aligned} \tag{4.35}$$

The derived cost function can be written in the following form:

$$P_{ss_1}(t)=\frac{1}{\mathbf{v}(t)^H(\sum_{u=0}^{U}\mathbf{D}_u^H\hat{\mathbf{U}}_n\hat{\mathbf{U}}_n^H\mathbf{D}_u)\mathbf{v}(t)} \tag{4.36}$$

where

$$\mathbf{v}(t) = \mathbf{v}^{(0)}(t) = [v_{-K/2+1}(t), \ldots, v_{K/2}(t)]^T$$
$$\mathbf{D}_u = diag\{(-j\frac{2\pi}{T}(-K/2+1))^u, \ldots, (-j\frac{2\pi}{T}K/2)^u\}. \quad (4.37)$$

The proposed cost function turns to be the conventional MUSIC when setting $U = 0$, which is the case for very small delay spreading (where it tends to be neglected). The proposed technique can be seen as an extension of MUSIC in the delay spreading case.

The algorithm used for the estimation is summarized as follows:

Algorithm of the proposed subspace based method

Number of delays L is assumed to be known.

1. Collect the NM DFT-domain vectors to build matrix $\mathbf{X}$ (4.25).
2. Estimate the covariance matrix $\mathbf{R_X}$ (4.26).
3. Apply the eigenvalue decomposition on $\hat{\mathbf{R}}_\mathbf{X}$.
4. Estimate the effective dimension of the signal subspace using the MDL criterion combined with the proposed rule (4.33).
5. Construct matrix $\hat{\mathbf{U}}_n$.
6. For each value of t, calculate the cost function $P_{ss_1}(t)$(4.36). Then, the cluster mean delays t_l are estimated by searching for the L peaks of $P_{ss_1}(t)$.

The proposed cost function shows no major difference from the conventional MUSIC in terms of the computational complexity as they share the same operations such as the FFT, $\mathbf{R_X}$ matrix construction, eigenvalue decomposition, and the spectral searching. The proposed cost function requires a slight increase in the computational complexity due to the addition of the terms corresponding to the delay spreading, which is the projection of $\mathbf{v}^{(u)}(t)$ for $u = 0 \ldots, U$ on the estimated noise subspace.

4.4 Simulation results

Simulations are carried out for a 24×24 MIMO system but similar conclusions can be obtained with other MIMO system configurations. The DFT-domain data are gener-

ated according to (4.3) with $K = 64$ Fourier coefficients. The delay deviations $\tau_{lp}^{(n,m)}$ of multirays within each cluster l for each (n, m) transmit-receive antenna pair are generated according to a uniform distribution and centered at t_l. The corresponding gains $\alpha_{lp}^{(n,m)}$ of the multirays are modeled as complex Gaussian random variables generated in a way that the effective power of each cluster is normalized, such that $\sum_{p=1}^{P_l} E[|\alpha_{lp}^{(n,m)}|^2] = 1, \forall\{l, n, m\}$. The added noise is modeled as complex white Gaussian noise. The estimation performance is assessed from $Q = 500$ independent simulations. The root mean square error (RMSE) of cluster mean delay estimation is calculated as $\sqrt{\frac{\sum_{i=1}^{Q}\sum_{l=1}^{L}(\hat{t}_l(i)-t_l)^2}{QT^2L}}$, where $\hat{t}_l(i)$ is the estimated mean delay for the ith experiment and t_l is the true mean delay.
Figure 4.1 shows the RMSE of mean delay estimation of the modified SOMP method and the classical SOMP method versus SNR.

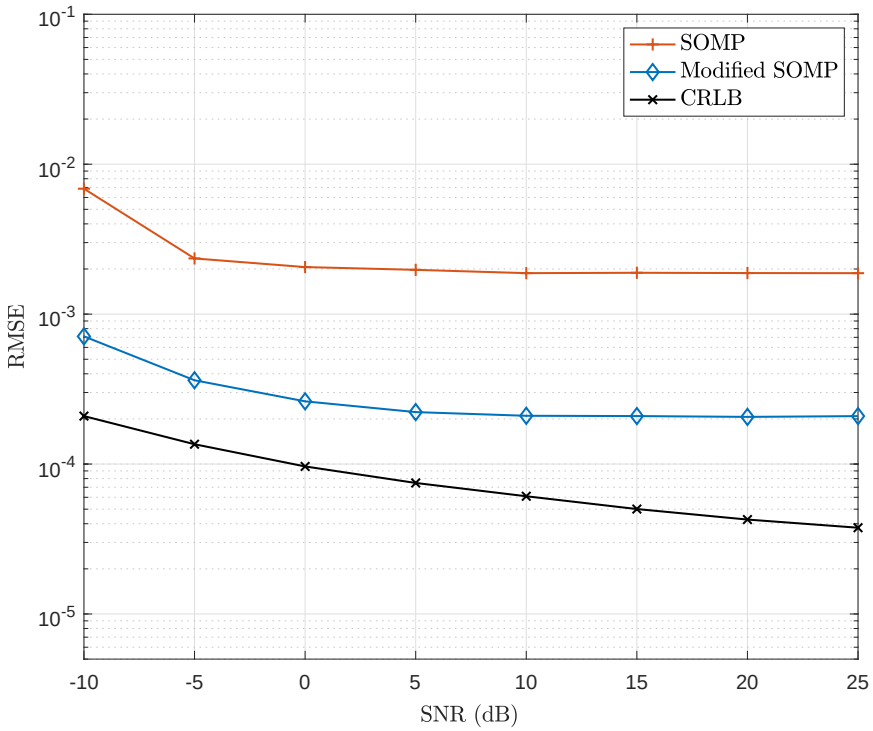

Figure 4.1: RMSE of mean delay estimation of the modified SOMP method and the classical SOMP method versus SNR; $P_l = 20$, $K = 64$, $L = 3$, vector of chosen delays is given as $\mathbf{t} = [0.32 \quad 0.45 \quad 0.61]T$, τ_{lp} are uniformly distributed with $0.008T$ chosen as the standard deviation of the delay spreading.

As shown in the figure, the proposed modified SOMP method provides better mean delay estimation in comparison with the classical SOMP method. By definition the CRLB is the lowest bound of the variance of any unbiased estimator. However, as we use the RMSE to evaluate the method, the CRLB in the figures represents the lower bound of the standard deviation of any unbiased estimator.

Figure 4.2 shows the RMSE of the mean delay estimation versus standard deviation of delay spreading for the modified SOMP method and the classical SOMP method.

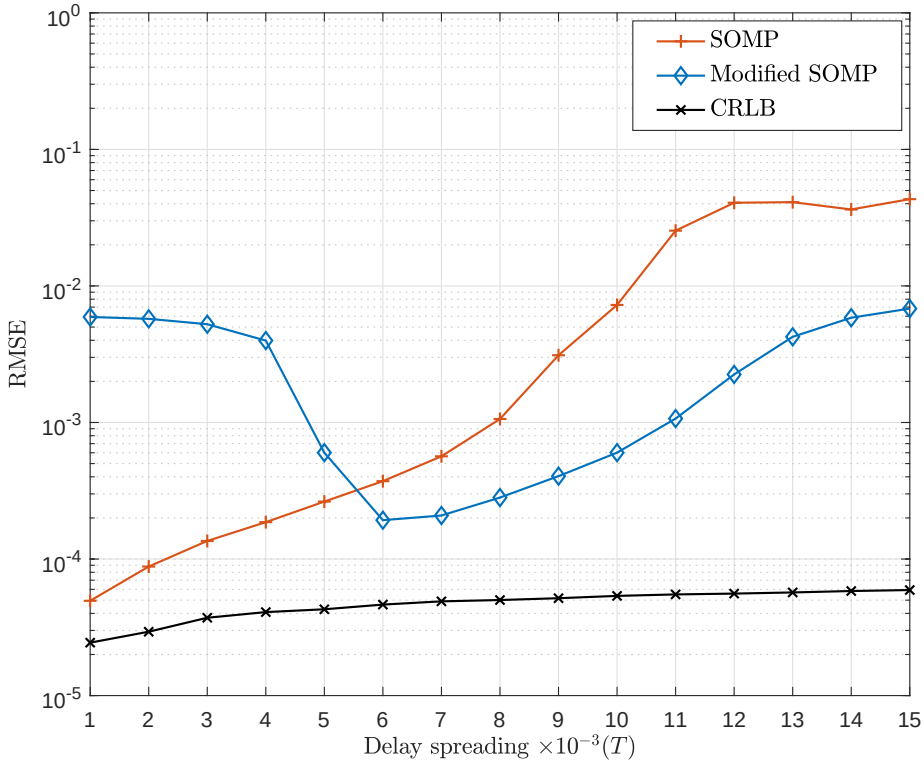

Figure 4.2: RMSE of mean delay estimation of the modified SOMP method and the classical SOMP method versus standard deviation of delay spreading; $P_l = 20$, $K = 64$, $L = 3$, vector of chosen delays is given as $\mathbf{t} = [0.32 \quad 0.45 \quad 0.61]T$, τ_{lp} are uniformly distributed, SNR=15 dB.

For relatively small delay spreading, the coefficient $b_l^{(n,m)}$ in (4.12) is relatively small. Hence it seems better to ignore the contribution of the vector $\mathbf{d}(t_l)$ where the classical

SOMP method can be used for estimation. However, as the standard deviation of delay spreading increases, the contribution of $\mathbf{d}(t_l)$ increases, where as shown in the above figure, the modified SOMP method provides better mean delay estimation in comparison with the classical SOMP method.

To illustrate the estimation performance of the subspace based approach, we consider first a unique delay group $L = 1$.

Figure 4.3 shows the RMSE of the mean delay estimation versus SNR of the proposed cost function (4.36) for different values of U and MUSIC.

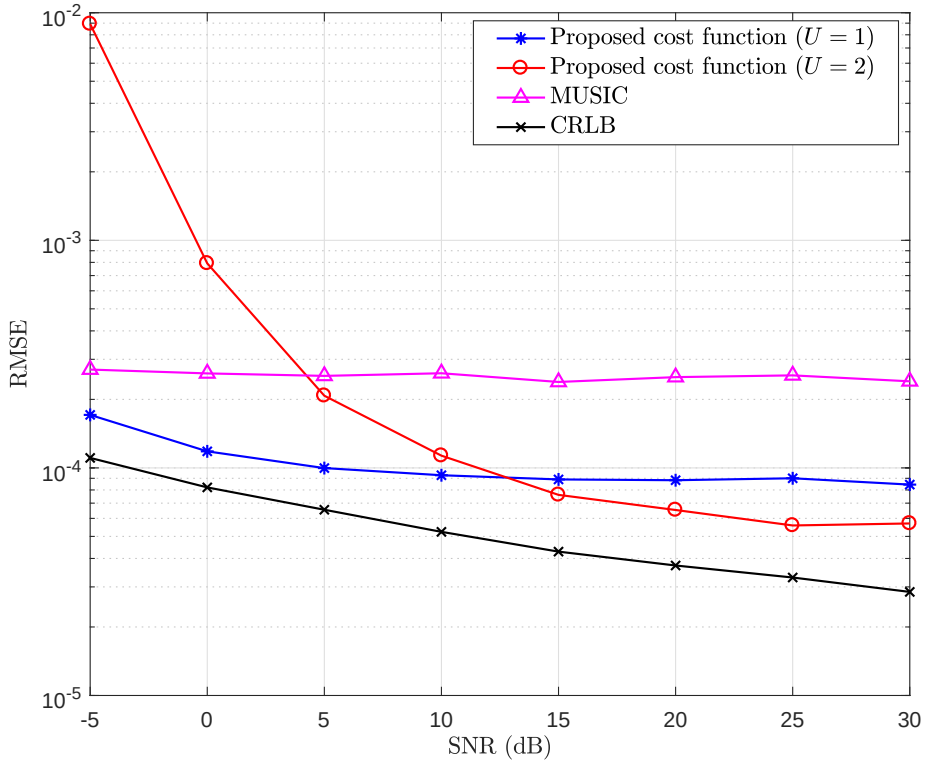

Figure 4.3: RMSE of mean delay estimation of the proposed cost function for $U = 1$, $U = 2$ and MUSIC versus SNR; $P_l = 20$, $K = 64$, τ_{lp} are uniformly distributed with $0.005T$ chosen as the standard deviation of the delay spreading.

Figure 4.3 shows that the proposed cost function (assuming $U = 1$) provides better performance of the cluster mean delay estimation than the conventional MUSIC.

According to the derivation of the model, in the noise-free case, it is better to choose higher values for U to benefit from a better approximation, where small details can be taken into account. On the other hand, with noise-contaminated data, the high order terms in Taylor expansion (4.17) may be too small with respect to the noise power. In fact, due to the property of Taylor expansion, it can be noted from (4.17) and (4.18), that:

$$|a_{l,0}^{(n,m)}| > |a_{l,1}^{(n,m)}| > \cdots > |a_{l,U}^{(n,m)}| \tag{4.38}$$

Hence, for $U = 2$ the elements in vector $a_{l,2}^{(n,m)}\mathbf{v}^{(2)}(t_l)$ will tend to be quite small. Thus, for a high level of noise, the norm of the vector $a_{l,2}^{(n,m)}\mathbf{v}^{(2)}(t_l)$ is very small compared to noise. Hence, this part of the signal is dominated by the noise, and it would be better not to consider this part of the signal to represent the signal subspace. For this reason, for low SNR, as can be observed in Figure 4.3, it would be better to consider $U = 1$ ($K - 2L$ as dimension of noise subspace). However as SNR increases, it would be worth considering vector $\mathbf{v}^{(2)}(t_l)$ as a part of the signal subspace, and then $U = 2$ ($K - 3L$ as dimension of noise subspace).

Figure 4.4 shows the comparison of RMSE of the cluster mean delay estimation versus standard deviation of delay spreading.

We can observe that for very small delay spreading with relatively low SNR (SNR = 5 dB), it would be better to consider that the dimension of the signal subspace is one ($U = 0$), for the same reason discussed before, which is the situation of the conventional MUSIC. However, as delay spreading increases, the proposed cost function with $U = 1$ outperforms MUSIC ($U = 0$). Again, as the delay spreading increases, the proposed cost function for $U = 3$ outperforms the proposed cost function for $U = 2$. The increase in the standard deviation of the delay spreading will give rise to an increase in the norm of $a_{l,u}^{(n,m)}\mathbf{v}^{(u)}(t_l)$. Then, this vector will have greater impact in formulating the signal subspace, where higher values of U can be considered.

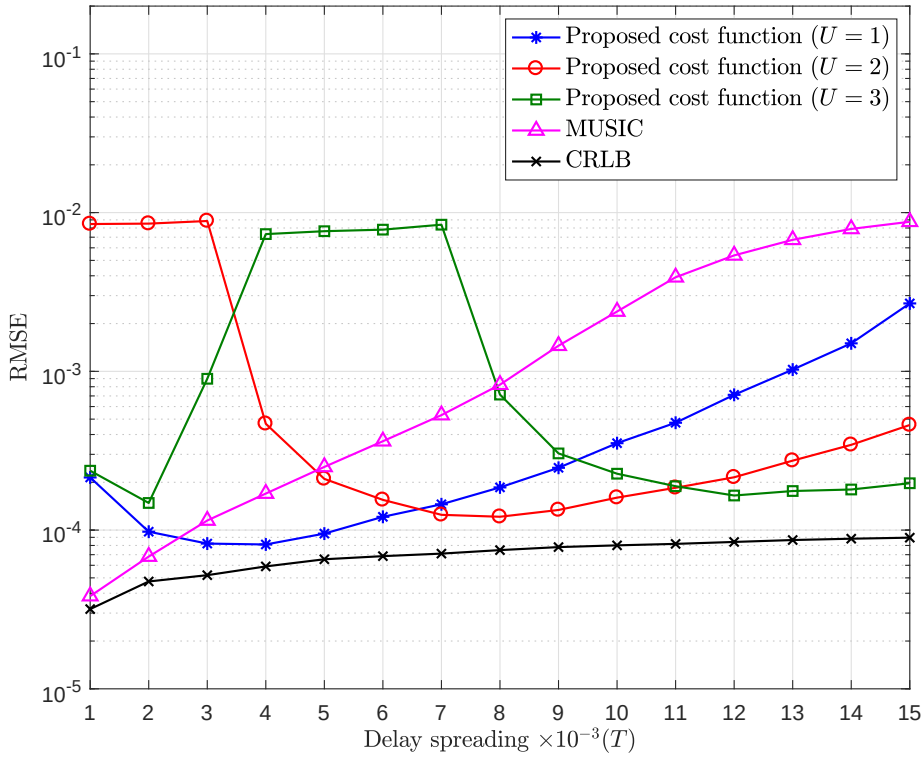

Figure 4.4: RMSE of mean delay estimation of the proposed cost function for $U = 1$, $U = 2$, $U = 3$, and MUSIC versus standard deviation of delay spreading; $P_l = 20$, $K = 64$, SNR = 5 dB.

We can also observe that as the delay spreading increases, the estimation is less accurate, which is probably due to the approximation error of Taylor expansion.

4.4.1 Selecting U according to the MDL criterion

To determine the most appropriate value of U, the MDL criterion is applied on the eigenvalues of $\mathbf{R_X}$. The calculated mean value of the MDL criterion (mean $MDLV$) versus the standard deviation of delay spreading is plotted in Figure 4.5 to show the performance of the MDL criterion.

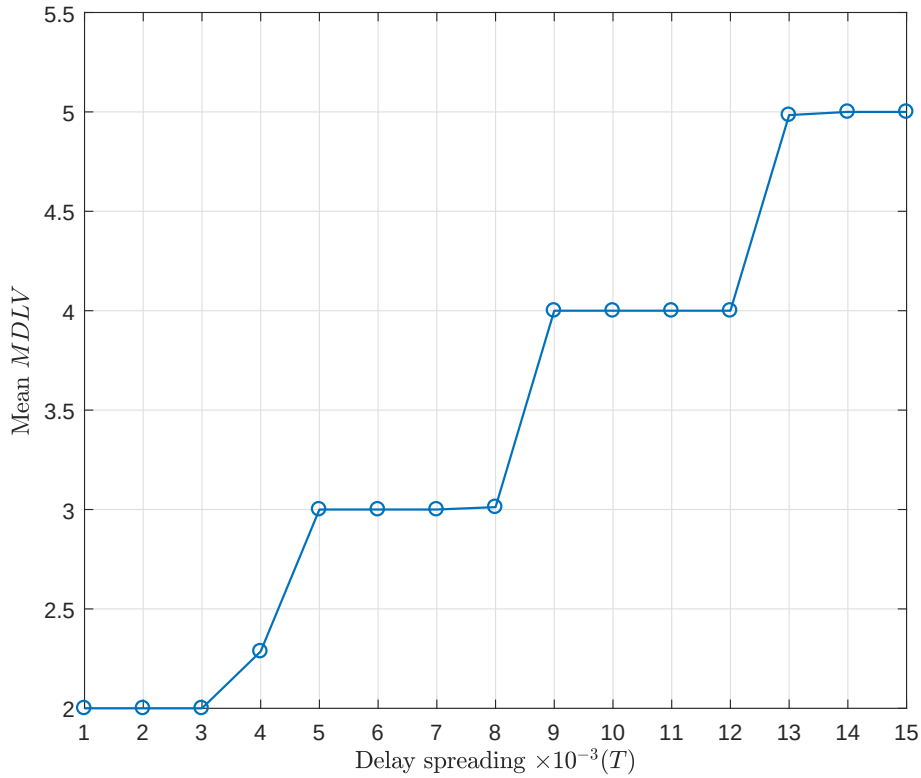

Figure 4.5: Mean value of the MDL criterion (mean $MDLV$) versus standard deviation of delay spreading; $P_l = 20$, $K = 64$, SNR $= 5$ dB, mean $MDLV$ values are obtained from 500 independent simulations each.

Observing Figures 4.4 and 4.5 simultaneously, four flat regions can be distinguished. When $MDLV = 2$, the best choice for the dimension of the signal subspace $(U + 1)L$ is 1 ($U = 0$). When $MDLV = 3, 4, 5$, the best dimensions of the signal subspace are 2 ($U = 1$), 3 ($U = 2$), and 4 ($U = 3$), respectively.

However, it can be noticed that the MDL criterion fails to estimate the effective dimension of the signal subspace correctly in some regions, it can be also noticed that flat regions are not always clearly distinguished. For example, when the delay spreading is $0.004T$, from the observed mean $MDLV$ values, it seems that the MDL criterion is providing the values 2 and 3 in different realizations. This shows that the MDL criterion, influenced by the random nature of scattering and the relatively high noise level, is not always a foolproof indicator.

In Figure 4.6, the RMSE of mean delay estimation of the proposed subspace tracking based method is shown. For each realization of received data, the decision about U is obtained from the rule in (4.33), which is used to estimate the effective dimension of the signal subspace. Then, the proposed cost function (4.36) is applied.
As shown in the figure, the proposed method allows for optimal selection of parameter U and provides the best cluster mean delay estimation. However some minor failures of the MDL criterion in systematically estimating the optimal effective dimension of the signal subspace can be noted.

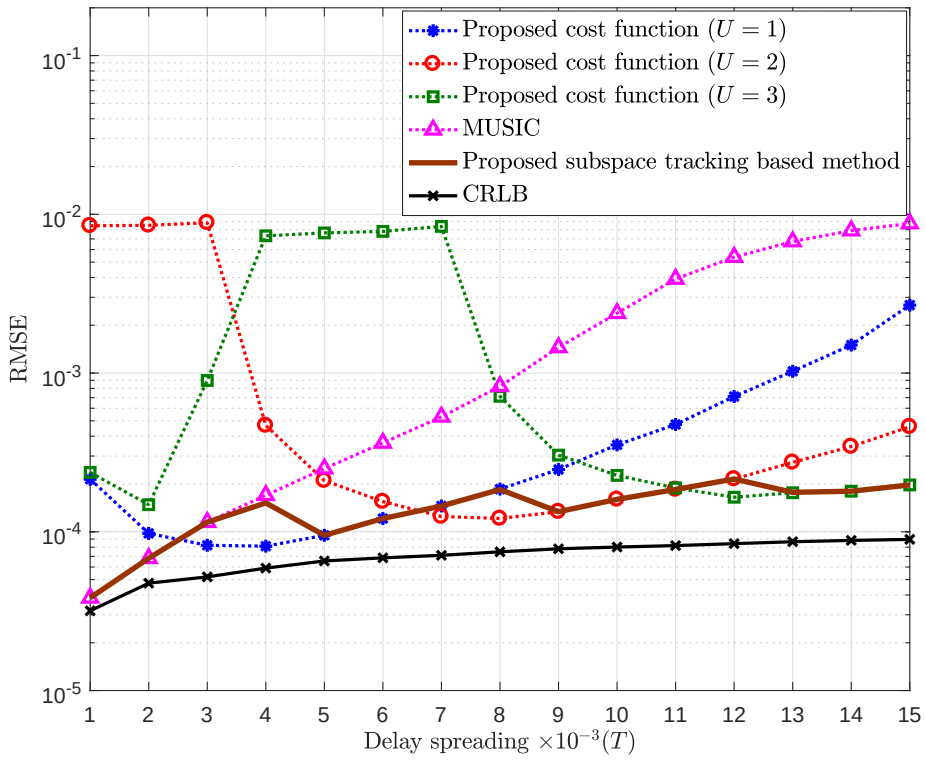

Figure 4.6: RMSE of mean delay estimation of the proposed cost function for $U = 1$, $U = 2$ and $U = 3$, the proposed subspace tracking based method and MUSIC versus standard deviation of delay spreading; $P_l = 20$, $K = 64$, SNR $= 5$ dB.

Figures 4.7 and 4.8 show the RMSE of the mean delay estimation and the mean $MDLV$ versus standard deviation of delay spreading, respectively for SNR = 15 dB.

The obtained results show that the MDL criterion provides better estimations of the different effective dimensions of the signal subspace for different standard deviations of delay spreading, leading to an improvement in the mean delay estimation performance of the proposed subspace tracking based method.

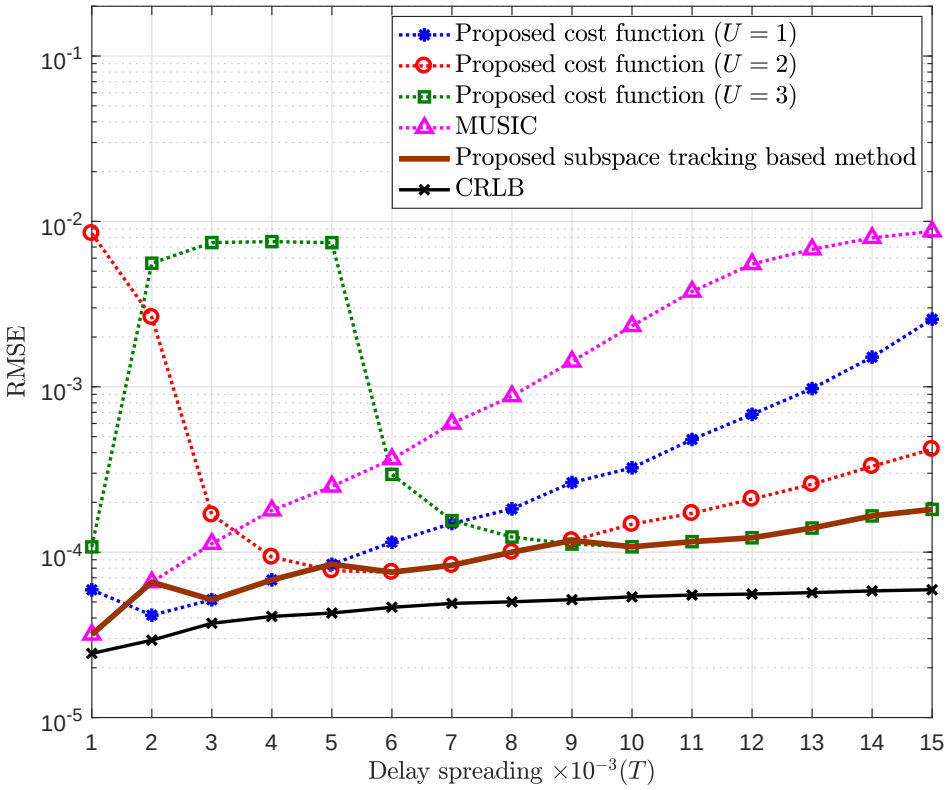

Figure 4.7: RMSE of mean delay estimation of the proposed cost function for $U = 1$, $U = 2$, and $U = 3$, the proposed subspace tracking based method and MUSIC versus standard deviation of delay spreading; $P_l = 20$, $K = 64$, SNR = 15 dB.

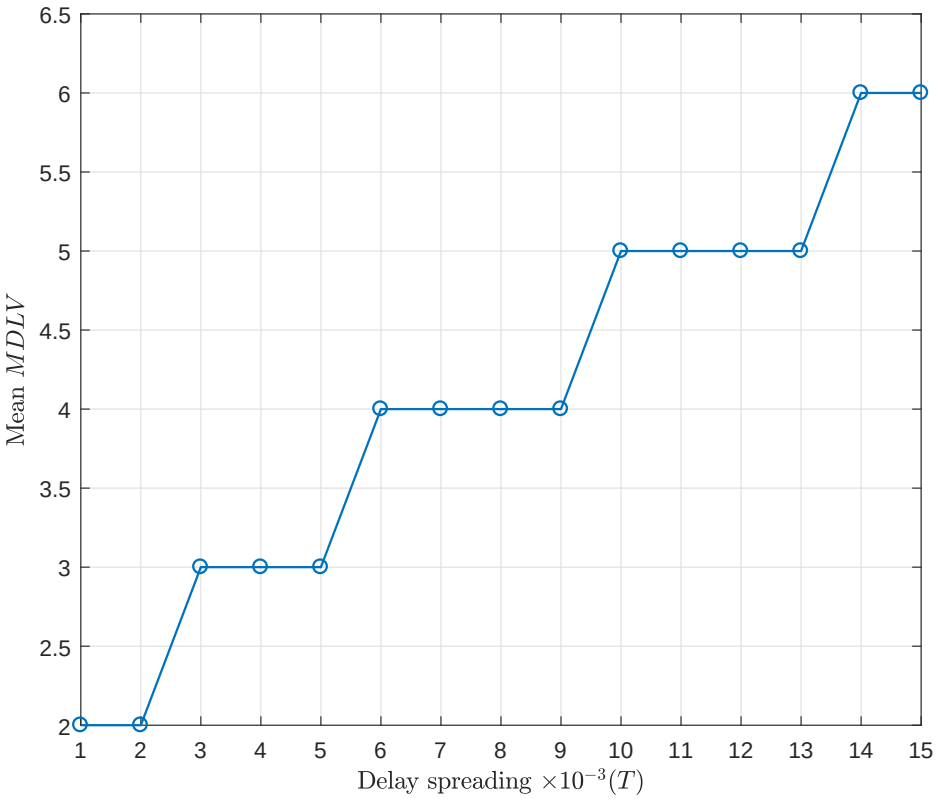

Figure 4.8: Mean $MDLV$ versus standard deviation of delay spreading; $P_l = 20$, $K = 64$, SNR = 15 dB

As mentioned before, the increase in the RMSE with respect to the delay spreading may be due to the approximation error in Taylor expansion. However, as shown in Figures 4.6 and 4.7, this increase is less significant for the proposed subspace tracking based method; this is the main advantage of tracking the effective dimension of the signal subspace that changes according to the value of the standard deviation of the delay spreading and the noise level.

Figure 4.9 shows the RMSE of mean delay estimation of the proposed subspace tracking based method and MUSIC versus standard deviation of delay spreading for different SNR values.

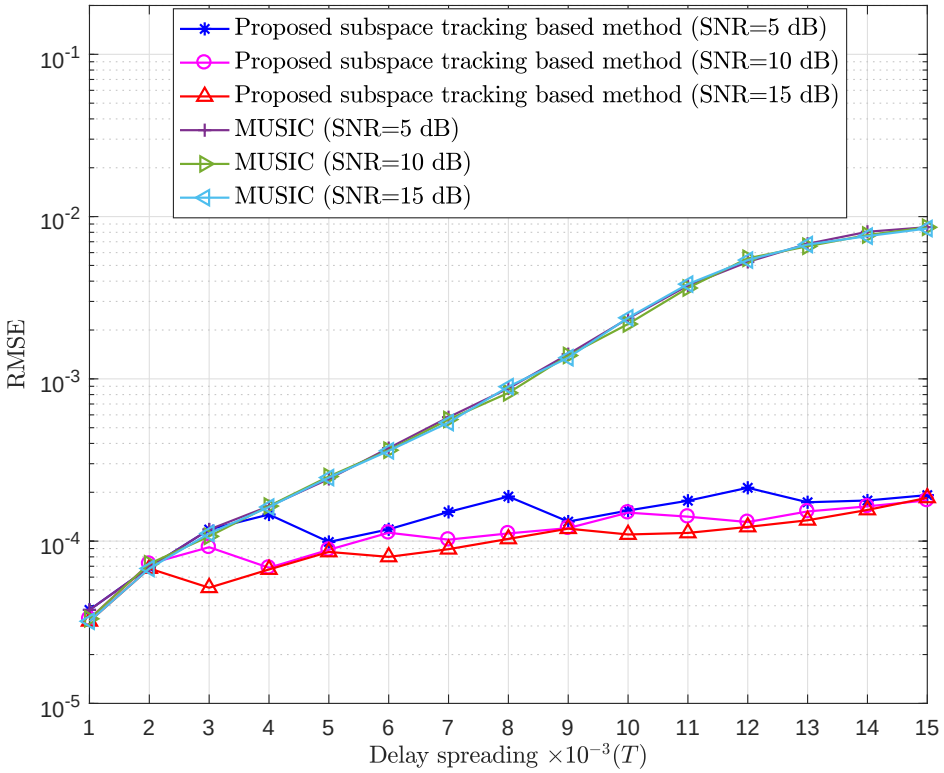

Figure 4.9: RMSE of mean delay estimation of the proposed subspace tracking based method and MUSIC versus standard deviation of delay spreading; $P_l = 20$, $K = 64$, SNR $= 5, 10, 15$ dB.

The obtained results show a moderate improvement in the estimation performance as SNR increases. In fact as SNR increases, better estimation of the covariance matrix is obtained, hence better estimation of signal or noise subspaces is attained.

As one cluster ($L = 1$) is considered before for the sake of simplification, Figures 4.10 and 4.11 show the RMSE of cluster mean delay estimation for $L = 2$ and $L = 3$ situations, respectively, at SNR $= 15$ dB.

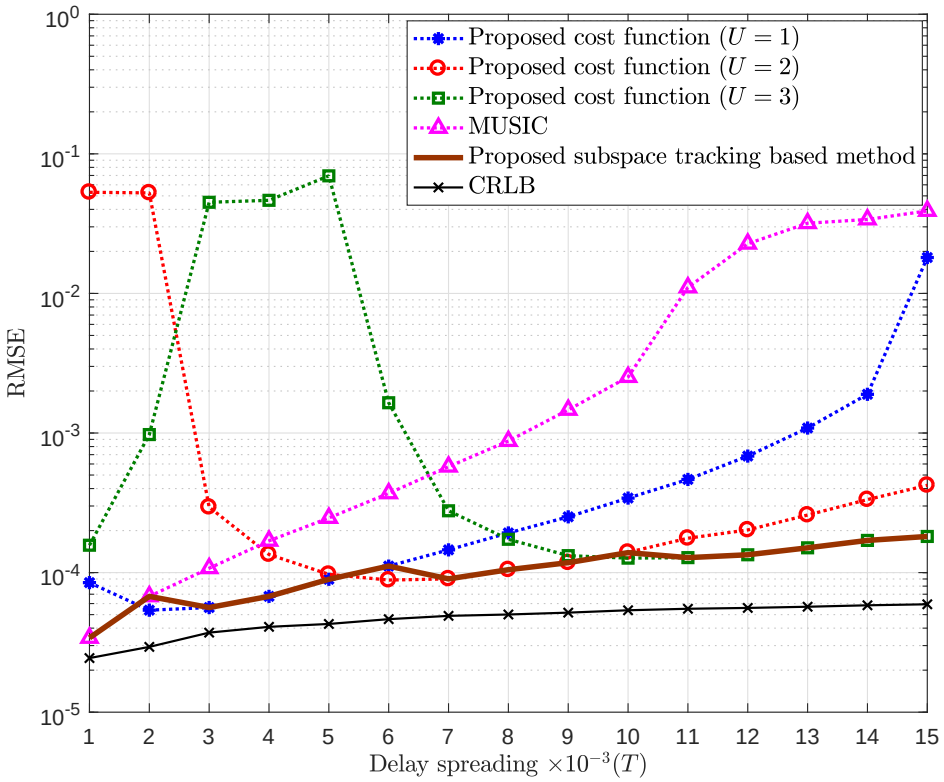

Figure 4.10: RMSE of mean delay estimation of the proposed cost function for $U = 1$, $U = 2$ and $U = 3$, the proposed subspace tracking based method and MUSIC versus standard deviation of delay spreading; $L = 2$, $\mathbf{t} = [0.37 \quad 0.51]T$, $P_l = 20$, $K = 64$, SNR = 15 dB.

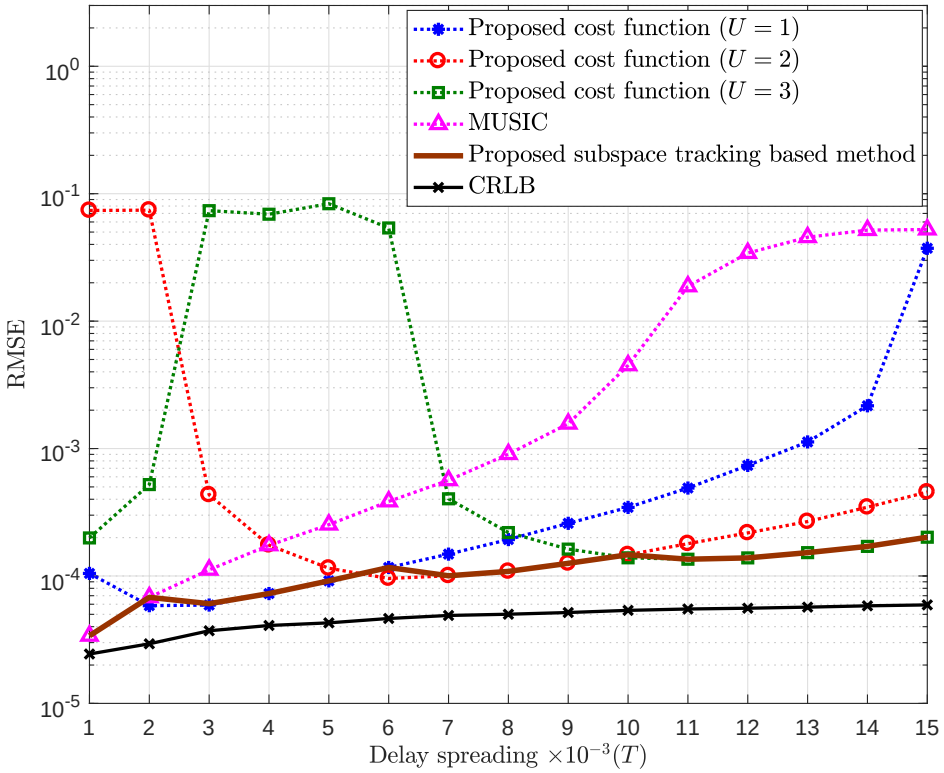

Figure 4.11: RMSE of mean delay estimation of the proposed cost function for $U = 1$, $U = 2$ and $U = 3$, the proposed subspace tracking based method and MUSIC versus standard deviation of delay spreading; $L = 3$, $\mathbf{t} = [0.37 \quad 0.51 \quad 0.67]T$, $P_l = 20$, $K = 64$, SNR = 15 dB.

Figure 4.12 shows the RMSE of mean delay estimation of the proposed modified SOMP method, the proposed subspace tracking based method, SOMP method and MUSIC versus standard deviation delay spreading. As shown in the figure, the modified SOMP method provides better performance than the conventional MUSIC and SOMP methods when the standard deviation of delay spreading is above a certain value. The proposed subspace tracking based method provides the best performance.

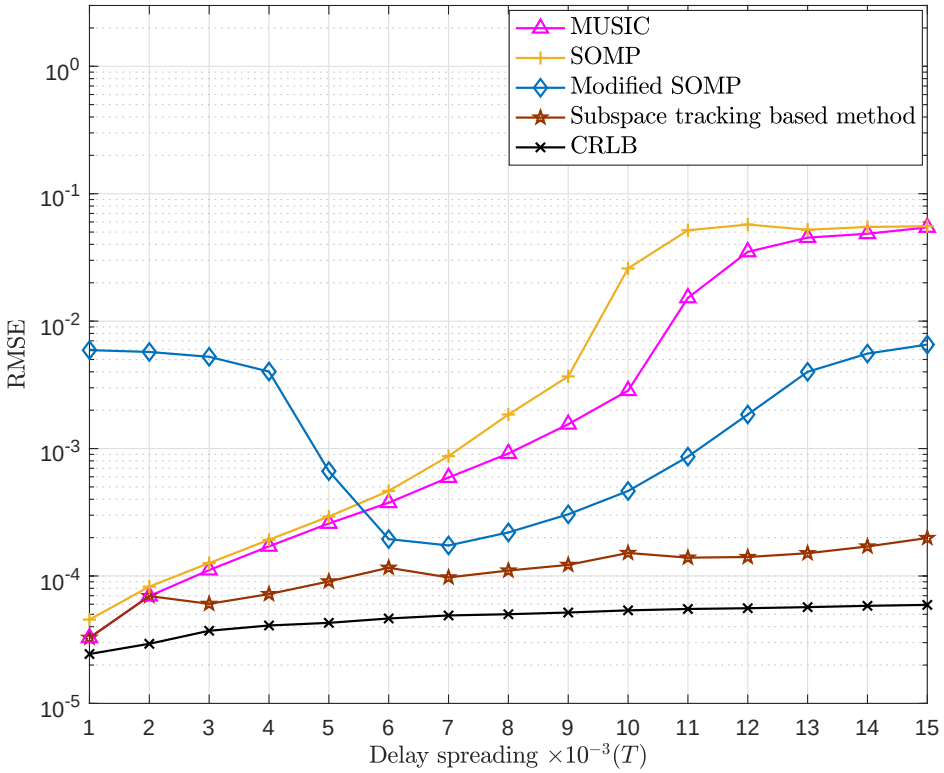

Figure 4.12: RMSE of mean delay estimation of the modified SOMP method, the proposed subspace tracking based method, SOMP method and MUSIC versus standard deviation of delay spreading; $L = 3$, $\mathbf{t} = [0.37 \quad 0.51 \quad 0.67]T$, $P_l = 20$, $K = 64$, SNR $=$ 15 dB.

4.5 Conclusion

In this chapter, two methods for channel mean delays estimation are proposed. A deterministic channel model is considered, and the DFT coefficients of the received signal are rederived by means of Taylor expansion around the mean delay parameter. Based on the first order Taylor expansion, a compressive sensing based method is proposed. Then based on higher order Taylor approximation, a subspace based method is developed based on the tracking of the effective dimension of the signal subspace, which depends on the channel features. The two proposed schemes are applied to estimate the channel mean delays. The proposed methods show better performance in compari-

son to the conventional methods. In comparison with the proposed compressive sensing based method, the proposed subspace based method allows estimating the cluster mean delays with more accuracy.

Chapter 5

Second order delay statistics estimation exploiting channel statistics - a stochastic approach

A cluster of multirays can be characterized by its mean delay and its delay spreading. In the previous chapter, we focus on estimating the mean delays of the different clusters based on a deterministic channel model. For the work to be complete, we propose in this chapter to estimate the standard deviation of the delay spreading of each cluster based on a stochastic model, exploiting time delays distribution of the clustered signals. Based on the stochastic model, a subspace based method is derived where both the mean delay and the standard deviation can be joinlty estimated but through a two-dimensional expensive search. Instead, the estimation procedure is divided into two steps. As a first step, the channel mean delays can be estimated using one of the methods proposed in the previous chapter based on the deterministic channel model. Then the associated standard deviations are estimated based on the stochastic model using the mean delays already estimated in the first step.

5.1 Stochastic model based approach

In this approach, the channel delay parameter is modeled in a stochastic manner, assuming a predefined statistical distribution for multiray delays. The Fourier coefficients for any transmit-receive antenna pair can be modeled with the following random function:

$$X[k] = \sum_{l=1}^{L} \int_{t \in \mathcal{T}} v_k(t) \alpha_l(t; \boldsymbol{\xi}_l) dt + Z[k] \tag{5.1}$$

where $\mathcal{T}$ is the interval in which the spreading for all clusters takes place; $\boldsymbol{\xi}_l = [t_l, \sigma_l]$ is the parameter vector characterizing the channel, such that t_l is the mean delay and σ_l is the standard deviation of the delay spreading of cluster l; $\alpha_l(t; \boldsymbol{\xi}_l)$ is the complex gain in the cluster, where for a fixed $\boldsymbol{\xi}_l$, $\alpha_l(t; \boldsymbol{\xi}_l)$ is a random process with respect to the delay variable t, and $Z[k]$ is the additive noise modeled as a Gaussian random variable with zero mean and variance σ_z^2. Since the gain coefficients are assumed to be identically distributed for all the transmit-receive antenna pairs with the same distribution of multiray delays, subscript (n, m) is omitted in the above equation.

The Fourier coefficients are concatenated to form the following random vector:

$$\mathbf{x} = \sum_{l=1}^{L} \int_{t \in \mathcal{T}} \mathbf{v}(t) \alpha_l(t; \boldsymbol{\xi}_l) dt + \mathbf{z} \tag{5.2}$$

with $\mathbf{x} = [X[-K/2+1], \ldots, X[K/2]]^T$, $\mathbf{v}(t) = [v_{-K/2+1}(t), \ldots, v_{K/2}(t)]^T$ and $\mathbf{z} = [Z[-K/2+1], \ldots, Z[K/2]]^T$.

The corresponding covariance matrix is given by

$$\mathbf{R}_{\mathbf{X}} = E[\mathbf{x}\mathbf{x}^H] =$$
$$\sum_{l,l'=1}^{L} \int_{\mathcal{T}} \int_{\mathcal{T}} E[\alpha_l(t; \boldsymbol{\xi}_l) \alpha_{l'}^*(t'; \boldsymbol{\xi}_{l'})] \mathbf{v}(t) \mathbf{v}(t')^H dt dt' + \sigma_z^2 \mathbf{I} \tag{5.3}$$

where $t, t' \in \mathcal{T}$.

Assuming that the different clustered signals are uncorrelated, and the multirays within each cluster are also uncorrelated. It comes that :

$$E[\alpha_l(t; \boldsymbol{\xi}_l) \alpha_{l'}^*(t'; \boldsymbol{\xi}_{l'})] = \delta_{ll'} \delta_{tt'} \sigma_{\alpha_l}^2 w_l(t; \boldsymbol{\xi}_l) \tag{5.4}$$

where δ_{pq} is the Kronecker delta.

The covariance matrix can then be written as follows:

$$\mathbf{R}_{\mathbf{X}} = \sum_{l=1}^{L} \mathbf{R}(t_l, \sigma_l) + \sigma_z^2 \mathbf{I} \tag{5.5}$$

where

$$\mathbf{R}(t_l, \sigma_l) = \sigma_{\alpha_l}^2 \int_{-\infty}^{+\infty} w_l(t; \boldsymbol{\xi}_l) \mathbf{v}(t) \mathbf{v}(t)^H dt \tag{5.6}$$

is the covariance matrix of the lth received clustered signal, $w_l(t; \boldsymbol{\xi}_l)$ is the normalized power delay function of the cluster and $\sigma_{\alpha_l}^2$ is its total mean power.

Assuming that multiray delays in each cluster are uniformly distributed, we have:

$$w_l(t; \boldsymbol{\xi}_l) = \frac{1}{2\sqrt{3}\sigma_l} Rect(t_l - \sqrt{3}\sigma_l, t_l + \sqrt{3}\sigma_l) \tag{5.7}$$

It turns that:

$$[\mathbf{R}(t_l, \sigma_l)]_{k+K/2, k'+K/2} = \frac{|G[k]|^2}{2\sqrt{3}\sigma_l} \int_{-\infty}^{+\infty} Rect(t_l - \sqrt{3}\sigma_l, t_l + \sqrt{3}\sigma_l) e^{-j\frac{2\pi}{T}(k-k')t} dt \tag{5.8}$$

where $k, k' = -K/2 + 1 \ldots K/2$. Let $\tilde{t} = t - t_l$, then

$$\begin{aligned} &[\mathbf{R}(t_l, \sigma_l)]_{k+K/2, k'+K/2} = \\ &|G[k]|^2 \frac{e^{-j\frac{2\pi}{T}(k-k')t_l}}{2\sqrt{3}\sigma_l} \int_{-\sqrt{3}\sigma_l}^{+\sqrt{3}\sigma_l} e^{-j\frac{2\pi}{T}(k-k')\tilde{t}} d\tilde{t} \\ &= |G[k]|^2 e^{-j\frac{2\pi}{T}(k-k')t_l} \mathrm{sinc}(\frac{2\pi}{T}(k-k')\sqrt{3}\sigma_l) \end{aligned} \tag{5.9}$$

The eigendecomposition of the observation covariance matrix can be given as

$$\mathbf{R}_{\mathbf{X}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H = \mathbf{U}_s\boldsymbol{\Lambda}_s\mathbf{U}_s^H + \mathbf{U}_n\boldsymbol{\Lambda}_n\mathbf{U}_n^H \tag{5.10}$$

where the columns of $\mathbf{U}_n$ are the $K-(U+1)L$ eigenvectors spanning the noise subspace, associated with the $K-(U+1)L$ smallest eigenvalues of $\mathbf{R}_{\mathbf{X}}$.

For the clustered signal l, the signal has most of its energy concentrated in the first few eigenvalues of the corresponding covariance matrix $\mathbf{R}(t_l, \sigma_l)$. The eigenvectors

associated to these eigenvalues are orthogonal to the noise subspace, as the remaining eigenvalues are considered so small, we have

$$\mathbf{R}(t_l, \sigma_l)\mathbf{U}_n \approx \mathbf{0} \tag{5.11}$$

$\mathbf{R}(t_l, \sigma_l)$ is given in (5.9) where it is expressed in terms of the mean delay of cluster l and the corresponding standard deviation.

Hence practically, the mean delay and the standard deviation can be jointly estimated by searching for the peaks of the following 2D cost function:

$$P_{ss_2}(t, \sigma) = \frac{1}{||\mathbf{R}(t, \sigma)\hat{\mathbf{U}}_n||_F^2} \tag{5.12}$$

where $\hat{\mathbf{U}}_n$ is the matrix of the estimated noise subspace eigenvectors, and $||.||_F$ is the Frobenius norm.

5.2 Joint deterministic-stochastic based approach

The above stochastic model based approach requires a two-dimensional search which is computationally expensive. Hence we propose a combination between the determinstic model and the stochastic model based approaches. Firstly, the mean delays are estimated through a one-dimensional search by the cost function given in (4.36) or by the modified SOMP method, then the value of each estimated mean delay is substituted in $\mathbf{R}(t, \sigma)$ where from (5.9),

$$[\mathbf{R}(t, \sigma)]_{k+K/2, k'+K/2} = |G[k]|^2 e^{-j\frac{2\pi}{T}(k-k')t}\text{sinc}(\frac{2\pi}{T}(k-k')\sqrt{3}\sigma) \tag{5.13}$$

the corresponding standard deviation can then be estimated by (5.12) through a one-dimensional search over σ.

The steps of the proposed approach are summarized as follows:

Algorithm of the proposed joint deterministic-stochastic based approach

1- Collect the NM DFT-domain vectors to build matrix $\mathbf{X}$.

2- Estimate the covariance matrix $\mathbf{R}_{\mathbf{X}}$ as in (4.26).

3- Estimate the effective dimension of the signal/noise subspace and construct matrix $\hat{\mathbf{U}}_n$.

4- Estimate the L mean delays using one of the proposed approaches in chapter 4.

5- For each estimated mean delay $\hat{t}_l$, substitute its value in $\mathbf{R}(t,\sigma)$ (5.13). Then the corresponding standard deviation σ_l is estimated as:

$$\hat{\sigma}_l = \underset{\sigma}{\text{argmax}} \frac{1}{||\mathbf{R}(\hat{t}_l,\sigma)\hat{\mathbf{U}}_n||_F^2} \tag{5.14}$$

5.3 Simulation results

In this section, some simulations are conducted to illustrate the behavior of the proposed method. Simulations are carried out for a 24×24 MIMO system. Three clusters ($L = 3$) with associated delays $t_1 = 0.37T$, $t_2 = 0.51T$ and $t_3 = 0.67T$ are considered. $P_l = 20$ cluster multirays are considered for $l = 1, \ldots, L$. $K = 64$ Fourier coefficients are taken, the delay deviations $\tau_{lp}^{(n,m)}$ are generated according to zero mean uniform distribution with same variance for the three clusters. The effective power of each cluster is normalized such that $\sum_{p=1}^{P_l} E[|\alpha_{lp}^{(n,m)}|^2] = 1$, $\forall\{l,n,m\}$. RMSE values are obtained from $Q = 500$ independent runs.

The RMSE of the standard deviation estimation is calculated as $\sqrt{\frac{\sum_{i=1}^{Q}\sum_{l=1}^{L}(\hat{\sigma}_l(i)-\sigma_l)^2}{QT^2L}}$ where $\hat{\sigma}_l(i)$ is the estimated standard deviation of the lth cluster for the ith experiment.

As the scope of this chapter is to estimate the standard deviation of the delay spreading, we show in the first figure the RMSE of the standard deviation estimation versus SNR, assuming that the mean delays are known.

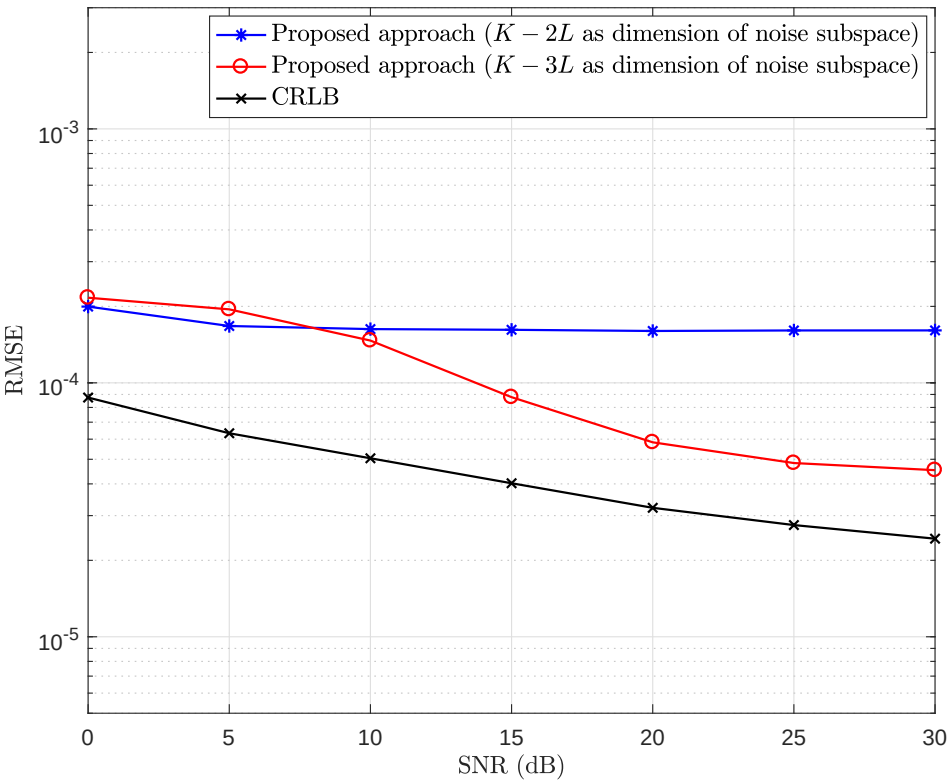

Figure 5.1: RMSE of standard deviation estimation versus SNR of the proposed approach with exact mean delay , $\sigma_l = 0.005T$ for all l.

In Figure 5.2, we show the RMSE of standard deviation estimation versus SNR provided by the proposed method (5.14) after substituting the mean delay already estimated using (4.36).

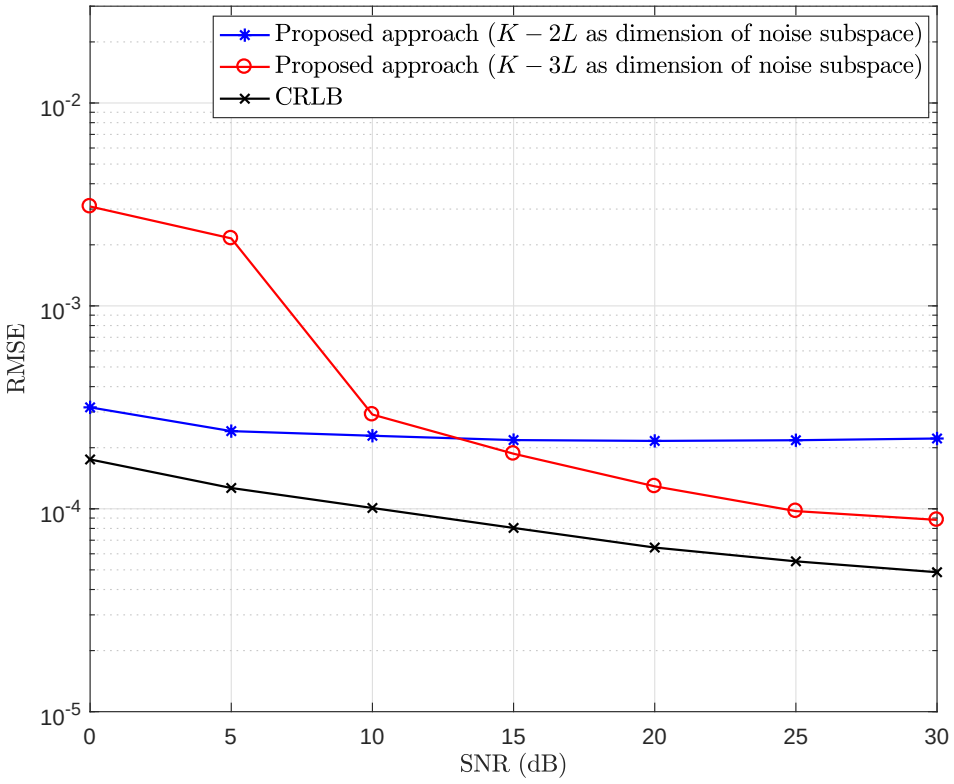

Figure 5.2: RMSE of standard deviation estimation versus SNR, $\sigma_l = 0.005T$ for all l.

As discussed in the previous chapter, two factors play the role in determining the effective dimension of the signal subspace, the standard deviation of the delay spreading and the noise level. As we can see in Figure 5.2, for relatively small SNR, it is better to use $2L$ as the dimension of the signal subspace ($K-2L$ as the dimension of the noise subspace). However as SNR increases, it is better to increase the considered dimension of the signal subspace to $3L$ ($K-3L$ as the dimension of the noise subspace). In fact for low SNR, small signal contributions are covered by noise, as a result, noise seems to occupy higher dimension in the measurement space, and therefore the effective dimension of the signal subspace decreases. On the other hand, as SNR increases, the effective dimension of the signal subspace increases, hence it is better to consider higher dimension for the signal subspace (or lower dimension for the noise subspace). This is illustrated in the above figures. By observing the two figures, we can notice the influence of the accuracy of the mean delay estimation on the standard deviation

estimation. This is expected as the mean delay estimated using (4.36) is used for the standard deviation estimation.

Figure 5.3 shows the RMSE of standard deviation estimation versus the standard deviation of delay spreading at SNR=15 dB.

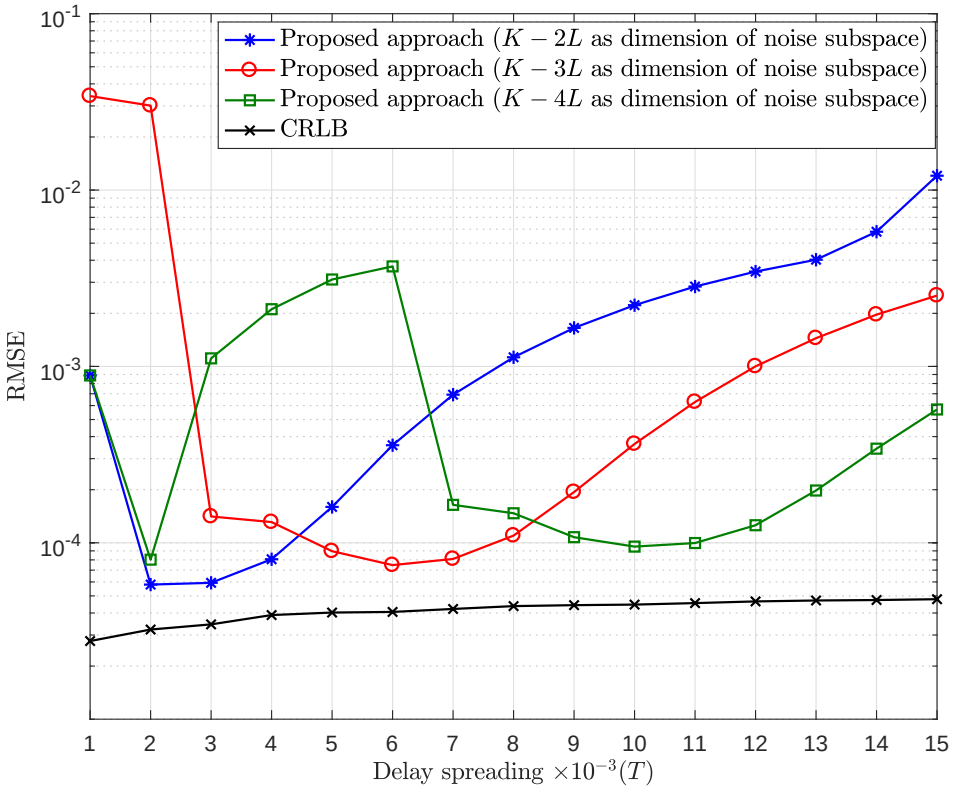

Figure 5.3: RMSE of standard deviation estimation versus standard deviation of delay spreading, SNR =15 dB.

As noticed, the effective dimension of the signal subspace increases (or the effective dimension of the noise subspace decreases) when the delay spreading increases. In fact, with the same principle discussed before, as the delay spreading increases, the signal contribution increases, subsequently, the effective dimension of the signal subspace increases, hence it is better to consider higher dimension for the signal subspace and lower dimension for the noise subspace.

5.4 Conclusion

A method for estimating the standard deviation of delay spreading parameter of clustered MIMO channel is proposed in this chapter. The channel delay parameter is treated as stochastic and the channel model is rederived and expressed in terms of the mean delay and the standard deviation of delay spreading parameters, exploiting the statistical distribution of multirays within clusters. A subspace based method that has the ability to estimate the mean delay and the standard deviation of delay spreading parameters jointly with two-dimensional search is then derived. A more interesting two-step approach is proposed where the mean delay parameter is estimated using one of the methods proposed in chapter 4, this estimated mean delay is then utilized to estimate the standard deviation of delay spreading parameter using the derived subspace based method.

General conclusion and perspectives

Conclusion

Understanding the characteristics of the wireless propagation channel is essential for the design of MIMO communication systems. Sparsity and clustering are two wireless channel properties that future wireless communication systems have to cope with. In order to permit communication over such channels, a deep knowledge of the channel characteristics and parameters is required. This thesis deals with sparsity and clustering properties of wireless channels. The channel is characterized in the time domain, where different schemes are proposed to estimate some time domain channel parameters.

The first chapter of the thesis provides an overview of the basic wireless propagation characteristics and introduces MIMO technology. The second chapter provides an overview of various MIMO channel models with different classifications based on different situations. The chapter then introduces the sparsity and clustering properties of several wireless channels, in addition to the common support property in MIMO outdoor communication scenarios. A sparse clustered MIMO channel model with common support hypothesis is then defined. The channel is considered sparse, such that it contains limited number of multipath components, where each multipath component is modeled as a cluster of multirays around a mean delay, and each cluster is parameterized by its mean delay and delay spreading. For this considered outdoor channel model, the multirays associated to the same scatterer are assumed to share the same delay parameters (mean and standard deviation) on the different transmit-receive antenna pairs. Chapter 3 deals with the MIMO channel parameter estimation problem. The estimation approaches are classified into parametric and nonparametric. The chapter focuses on the parametric approach for estimation as it allows exploiting some proper-

ties of the channel such as sparsity and common support.

Chapters 4 and 5 illustrate the methods proposed to estimate some channel parameters. The sparse clustered model defined in chapter 2 is considered. The work is focused on estimating the mean delay and the standard deviation of delay spreading parameters characterizing the clusters. Chapter 4 deals with mean delay estimation, where two approaches are proposed, a compressive sensing based approach and a subspace based approach, based on a deterministic channel model. The compressive sensing based approach is based on the first order Taylor expansion of the observation Fourier coefficients, where a modified SOMP method is proposed to estimate the cluster mean delays. The proposed modified SOMP method shows a better estimate of the mean delay parameter in comparison with the conventional SOMP method. The second scheme for estimation is based on the subspace approach. The expressions of the observation Fourier coefficients are rederived based on higher order Taylor expansion. Then, a subspace based method that exploits the rederived model is proposed to estimate the mean delays. The proposed approach is based on the tracking of the effective dimension of the signal subspace that changes depending on the standard deviation of delay spreading and SNR. The proposed subspace based approach outperforms both the conventional MUSIC method and the proposed compressive sensing based method (modified SOMP method) in terms of cluster mean delays estimation. In chapter 5, we focus on cluster delay spreads estimation. A stochastic modeling of the channel delay parameter is proposed, where the statistical distribution of multiray delays is exploited. A subspace based method is then derived where the mean delay and the delay spreading parameters can be estimated jointly through a two-dimensional search. More interesting, a two-step approach is proposed. The mean delays are estimated using one of the approaches proposed for mean delay estimation. The estimated mean delays are then exploited to estimate the corresponding standard deviation of delay spreading using the subspace approach.

Perspectives

The signal subspace tracking criterion proposed in chapter 4 does not seem to be efficient in some cases. As a future work, the problem of signal subspace tracking can be addressed where other approaches can be developed to attain better performance.

A time domain channel model is considered in this work. However, it will be inter-

esting to include some spatial domain parameters (angles), where different approaches can be proposed to estimate both time and angular domain parameters.

In addition, based on the estimated channel parameters, a study can be performed to analyze the effect of estimation error on the performance of the system. Moreover, a study can be carried out to analyze the capacity over the considered channel.

Finally, it would be interesting to conduct some experiments in real environments. On one side, the validity of the considered channel model can be investigated, on the other side, the effectiveness of the proposed estimation methods in real propagation environments can be examined.

Résumé en francais

Dans le monde antique, la lumière et les drapeaux étaient utilisés comme moyen de communication sans fil. En 1867, James Clerk Maxwell prédit l'existence d'ondes électromagnétiques (EM), proposant une interrelation entre les champs électriques et magnétiques. En 1887, Heinrich Rudolf Hertz a confirmé l'existence d'ondes électromagnétiques voyageant à la vitesse de la lumière en effectuant des expériences dans son laboratoire. Les ondes qu'il a produites et reçues sont maintenant appelées ondes radio. Guglielmo Marconi a fait une percée en mettant au point le télégraphe sans fil en 1895. Depuis lors, il a réussi à transmettre des signaux radio dans l'espace, augmentant progressivement la distance de communication. En 1901, il a établi la première communication sans fil à travers l'océan, en transmettant des signaux radio à travers l'océan Atlantique. Depuis lors jusqu'à aujourd'hui, différentes technologies sans fil ont été développées, notamment la radiodiffusion et la télédiffusion, les communications radar, les communications par satellite, les réseaux sans fil, les communications mobiles sans fil, etc.

Avec la montée de l'ère du *big data* et la demande croissante de services de données sans fil, l'objectif principal des chercheurs au fil des ans a été de soutenir des débits de données élevés pour répondre aux besoins. Un obstacle majeur à la construction d'un système de communication sans fil à haute vitesse fiable est le moyen de propagation sans fil. En communication sans fil, le signal se propageant par le canal sans fil est exposé à différents types d'évanouissements, en particulier les évanouissements dus à la propagation par trajets multiples. Cela affecte la fiabilité de la liaison de communication et limite le débit de données.

La technologie MIMO (Multiple-input multiple-output) est devenue un sujet de recherche actif au cours de la dernière décennie en raison de sa capacité d'atteindre les vitesses de transmission élevées requises par un nombre croissant d'applications exigeantes en données. La technologie MIMO offre de nombreux avantages qui permettent de relever les défis posés par les déficiences du canal sans fil, en particulier

l'évanouissement par trajets multiples. Il fournit plusieurs gains de performance importants tels que le gain d'antenne, le gain de diversité et le gain de multiplexage. Les avantages du MIMO sont obtenus grâce à l'exploitation de la dimension spatiale sur plusieurs antennes de l'émetteur et du récepteur, en plus des dimensions de temps et de fréquence déjà exploitées dans les systèmes conventionnels à entrée unique et sortie unique (SISO).

Le MIMO et le multiplexage par répartition en fréquence orthogonale (OFDM) sont des technologies clés utilisées dans les réseaux sans fil 4G (quatrième génération). MIMO est une technologie clé pour la prochaine cinquième génération (5G) de réseaux sans fil qui utilisent des réseaux d'antennes massifs et des fréquences millimétriques (MMW).

La connaissance des caractéristiques des canaux de propagation sans fil est cruciale pour la fiabilité des communications sans fil, en particulier dans les communications MIMO, afin de profiter pleinement des avantages offerts par l'utilisation de la technologie antennes multiples du côté de l'émetteur et du récepteur. Ces informations sur les caractéristiques du canal sont appelées informations sur l'état du canal (CSI). CSI représente l'information sur la propagation du signal de l'émetteur au récepteur, il représente les effets de canal sans fil tels que l'atténuation de puissance et l'étalement dans le temps des signaux. La connaissance des canaux de propagation sans fil est essentielle pour la conception des systèmes de communication MIMO. Les informations sur l'état des canaux au niveau du récepteur (CSIR) peuvent être utilisées à des fins d'égalisation contre les interférences intersymboles (ISI) causées par la propagation par trajets multiples, et les informations sur l'état des canaux au niveau de l'émetteur (CSIT) peuvent être utilisées pour concevoir une transmission optimale. Par conséquent, un système de communication MIMO "idéal" performant nécessiterait une connaissance exacte du canal MIMO ou CSI. Les approches d'estimation CSI peuvent être classées en deux catégories : paramétrique et non paramétrique. Dans l'approche non paramétrique, la matrice des canaux est estimée directement sans référence à aucun paramètre de propagation physique sous-jacent. D'autre part, l'approche paramétrique s'appuie sur des modèles de canaux physiques pour estimer les paramètres des canaux, ces paramètres sont utiles pour comprendre le canal sans fil et peuvent être utilisés pour améliorer les performances du système de communication en adaptant les conceptions de transmission et de réception en fonction de celles-ci. L'intérêt de l'approche paramétrique est qu'elle permet d'exploiter certaines propriétés du canal telles que la parcimonie et le support commun.

Le phénomène de regroupement dû aux diffuseurs est une propriété importante qui caractérise plusieurs canaux sans fil, où, selon les différentes recherches sur les canaux sans fil, les composantes de trajets multiples des canaux sont modélisées comme des grappes (“clusters”) de rayons multiples. Par exemple, ce phénomène (regroupement ou “clustering”) caractérise les canaux de communication à large bande/à bande ultralarge (UWB) et MMW. Par conséquent, la parcimonie et le regroupement sont deux propriétés des canaux sans fil auxquelles les futurs systèmes de communication sans fil devront faire face. Afin de permettre la communication sur de tels canaux, une connaissance approfondie des caractéristiques et des paramètres du canal est nécessaire, où les nouvelles caractéristiques du canal doivent être prises en compte dans les futures techniques d'estimation du canal. Les travaux de cette thèse se concentrent sur l'estimation de paramètres de canal MIMO en grappes, en particulier les paramètres du domaine temporel. Le canal est caractérisé dans le domaine temporel, où différents schémas sont proposés pour estimer certains paramètres de canal du domaine temporel.

La thèse est divisée en cinq chapitres. La première partie du premier chapitre représente les caractéristiques de base de la propagation d'un canal sans fil, en particulier l'évanouissement dans un canal sans fil. Le canal d'évanouissement est classé en évanouissement à grande échelle et évanouissement à petite échelle, l'évanouissement à grande échelle caractérise le comportement du canal sur de grandes distances et incorpore l'affaiblissement sur le trajet et l'ombrage. L'évanouissement à petite échelle caractérise le comportement du canal sur de courtes périodes de temps ou sur de courtes distances de déplacement et est classé en deux catégories en fonction de l'étalement du délai multi-trajet et de l'étalement Doppler. Sur la base de l'étalement du délai de propagation par trajets multiples, les évanouissements sont classés en évanouissements non sélectifs et en évanouissements sélectifs en fréquence. Basé sur l'étalement Doppler, l'évanouissement est classé en deux catégories : l'évanouissement lent et l'évanouissement rapide. Dans la deuxième partie, nous introduisons les techniques de diversité dans les systèmes de communication sans fil, et nous nous concentrons sur la diversité spatiale, en montrant les avantages apportés par l'utilisation d'antennes d'émission et/ou de réception multiples.

Dans le deuxième chapitre, nous donnons un aperçu des différents modèles de canaux MIMO. Les modèles de canaux sont classés en modèles physiques et non physiques. Les modèles physiques sont ensuite classés en modèles déterministes et stochastiques, où les modèles stochastiques sont classés en modèles géométriques, modèles non géométriques et modèles analytiques fondés sur la propagation. Ensuite, nous introduisons les pro-

priétés de parcimonie et de regroupement dans les canaux sans fil, en plus de la propriété de support commune dans les canaux MIMO extérieurs. Enfin, nous introduisons un modèle de canal MIMO en grappes parcimonieux avec un support commun, sur lequel sont basées les méthodes d'estimation que nous proposons.

Le chapitre 3 traite de l'estimation des paramètres de canal. Les méthodes d'estimation sont classées en deux catégories : paramétrique et non paramétrique. L'approche non paramétrique fait référence à l'estimation de la matrice de canaux MIMO tandis que l'approche paramétrique fait référence à l'estimat-ion des paramètres du canal. Nous classons les méthodes qui peuvent être utilisées pour l'estimation des paramètres de canal en 3 catégories : les méthodes de formation de faisceau, les méthodes basées sur le sous-espace et les méthodes de acquisition comprimée.

Les chapitres 4 et 5 illustrent les méthodes proposées pour estimer certains paramètres de canal. Le modèle en grappes parcimonieux défini au chapitre 2 est considéré. Le travail est axé sur l'estimation du retard moyen et de l'écart-type des paramètres d'étalement du retard caractérisant les grappes. Le chapitre 4 traite de l'estimation du retard moyen, où deux approches sont proposées, une approche fondée sur l'acquisition comprimée et une approche fondée sur le sous-espace, fondée sur un modèle de canal déterministe. L'approche fondée sur l'acquisition comprimée est basée sur l'expansion de Taylor du premier ordre autour du paramètre du retard moyen, où une méthode SOMP modifiée est proposée pour estimer les retards moyens des grappes. La méthode SOMP modifiée montre une meilleure estimation du retard moyen par rapport à la méthode SOMP conventionnelle. L'approche sous-espace est fondée sur l'expansion de Taylor d'ordre supérieur autour du paramètre du retard moyen, où une méthode fondée sur le sous-espace est proposée pour estimer les retards moyens de la grappe. L'approche proposée est fondée sur le suivi de la dimension effective du sous-espace du signal qui varie en fonction de l'écart-type de l'étalement du retard et du rapport signal/bruit. L'approche basée sur les sous-espaces proposée surpasse à la fois la méthode MUSIC conventionnelle et la méthode basée sur l'acquisition comprimée (méthode SOMP modifiée) en termes d'estimation des retards moyens des grappes. Dans le chapitre 5, nous nous concentrons sur l'estimation de l'écart-type de l'étalement du retard. Une modélisation stochastique du paramètre de retard du canal est proposée, où la distribution statistique des retards multi-rayons au sein de chaque grappe est exploitée. Une méthode basée sur le sous-espace est dérivée où le retard moyen et les paramètres d'étalement du retard peuvent être estimés conjointement par une recherche bidimensionnelle. Plus intéressant, une approche en deux étapes est proposée. Les retards moyens sont estimés à l'aide de l'une des approches proposées pour l'estimation du retard moyen. Les

retards moyens estimés sont ensuite exploités pour estimer l'écart-type de l'étalement des retards en utilisant l'approche de sous-espace.

Publications

Journal

A. Mohydeen, P. Chargé, Y. Wang, O. Bazzi, and Y. Ding, "Spatially correlated sparse MIMO channel path delay estimation in scattering environments based on signal subspace tracking," *Sensors (Basel, Switzerland)*,vol. 18, no. 5, 2018.

Conference

[1] A. Mohydeen, P. Chargé, Y. Wang, and O. Bazzi, "Estimation of clustered MIMO channel parameters exploiting channel statistics". *Fifth Sino-French Workshop on Information and Communication Technologies (SIFWICT 2019)*, June 2019, Nantes, France.

[2] A. Mohydeen, P. Chargé, Y. Wang, and O. Bazzi, "Spatially correlated sparse MIMO channel path delay estimation in scattering environments," in *2018 25th International Conference on Telecommunications (ICT)*, pp.138–142, IEEE, 2018.

Bibliography

[1] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*, vol. 2. prentice hall PTR New Jersey, 1996.

[2] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO wireless communications*. Cambridge university press, 2007.

[3] B. Sklar, "Rayleigh fading channels in mobile digital communication systems. I. characterization," *IEEE Communications magazine*, vol. 35, no. 9, pp. 136–146, 1997.

[4] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, "A survey of various propagation models for mobile communication," *IEEE Antennas and propagation Magazine*, vol. 45, no. 3, pp. 51–82, 2003.

[5] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[6] K. Pahlavan and A. H. Levesque, *Wireless information networks*, vol. 93. John Wiley & Sons, 2005.

[7] M. Dohler and Y. Li, *Cooperative communications: hardware, channel and PHY*. John Wiley & Sons, 2010.

[8] A. Sibille, C. Oestges, and A. Zanella, *MIMO: from theory to implementation*. Academic Press, 2010.

[9] P. W. Wolniansky, G. J. Foschini, G. Golden, R. A. Valenzuela, *et al.*, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *proc. ISSSE*, vol. 98, pp. 295–300, 1998.

[10] S. Primak and V. Kontorovich, *Wireless multi-antenna channels: modeling and simulation*, vol. 28. John Wiley & Sons, 2011.

[11] H. Ozcelik, N. Czink, and E. Bonek, "What makes a good MIMO channel model?," in *IEEE Vehicular Technology Conference*, vol. 61, p. 156, IEEE; 1999, 2005.

[12] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyösti, D. Laurenson, G. Matz, *et al.*, "Survey of channel and radio propagation models for wireless MIMO systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2007, no. 1, p. 019070, 2007.

[13] A. S. Glassner, *An introduction to ray tracing*. Elsevier, 1989.

[14] C. A. Balanis, *Advanced engineering electromagnetics*. John Wiley & Sons, 1999.

[15] Y. Ma and M. Pätzold, "Wideband two-ring MIMO channel models for mobile-to-mobile communications," in *Proc. 10th International Symposium on Wireless Personal Multimedia Communications, WPMC 2007*, pp. 380–384, 2007.

[16] X. Cheng, C.-X. Wang, H. Wang, X. Gao, X.-H. You, D. Yuan, B. Ai, Q. Huo, L.-Y. Song, and B.-L. Jiao, "Cooperative MIMO channel modeling and multi-link spatial correlation properties," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 2, pp. 388–396, 2012.

[17] A. F. Molisch, "A generic model for MIMO wireless propagation channels in macro- and microcells," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 61–71, 2004.

[18] J. Fuhl, A. F. Molisch, and E. Bonek, "Unified channel model for mobile radio systems with smart antennas," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 145, no. 1, pp. 32–41, 1998.

[19] A. Borhani and M. Patzold, "A unified disk scattering model and its angle-of-departure and time-of-arrival statistics," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 2, pp. 473–485, 2013.

[20] A. A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on selected areas in communications*, vol. 5, no. 2, pp. 128–137, 1987.

[21] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel," *IEEE Journal on Selected areas in communications*, vol. 18, no. 3, pp. 347–360, 2000.

[22] C.-C. Chong, C.-M. Tan, D. I. Laurenson, S. McLaughlin, M. A. Beach, and A. R. Nix, "A new statistical wideband spatio-temporal channel model for 5-GHz band WLAN systems," *IEEE Journal on selected areas in Communications*, vol. 21, no. 2, pp. 139–150, 2003.

[23] J. W. Wallace and M. A. Jensen, "Modeling the indoor MIMO wireless channel," *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 5, pp. 591–599, 2002.

[24] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-wave multipath clustering and channel modeling," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 3, pp. 1445–1455, 2014.

[25] T. Zwick, C. Fischer, and W. Wiesbeck, "A stochastic multipath channel model including path directions for indoor environments," *IEEE journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1178–1192, 2002.

[26] A. G. Burr, "Capacity bounds and estimates for the finite scatterers MIMO wireless channel," *IEEE Journal on Selected Areas in Communication*, pp. 812–818, 2003.

[27] A. M. Sayeed, "Deconstructing multiantenna fading channels," 2002.

[28] K. Liu, V. Raghavan, and A. M. Sayeed, "Capacity scaling and spectral efficiency in wide-band correlated MIMO channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2504–2526, 2003.

[29] M. Debbah and R. R. Muller, "MIMO channel modeling and the principle of maximum entropy," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1667–1690, 2005.

[30] C.-N. Chuah, J. M. Kahn, and D. Tse, "Capacity of multi-antenna array systems in indoor wireless environment," *GLOBECOM-NEW YORK-*, vol. 4, pp. 1894–1899, 1998.

[31] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Communications letters*, vol. 4, no. 11, pp. 337–339, 2000.

[32] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on communications*, vol. 48, no. 3, pp. 502–513, 2000.

[33] J.-P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on selected areas in Communications*, vol. 20, no. 6, pp. 1211–1226, 2002.

[34] V. Raghavan, J. H. Kotecha, and A. M. Sayeed, "Why does the kronecker model result in misleading capacity estimates?," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4843–4864, 2010.

[35] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 90–100, 2006.

[36] A. F. Molisch, "Ultrawideband propagation channels-theory, measurement, and modeling," *IEEE Trans. Vehicular Technology*, vol. 54, no. 5, pp. 1528–1545, 2005.

[37] A. Sayeed, "Sparse multipath wireless channels: Modeling and implications," in *Proc. ASAP*, 2006.

[38] W. U. Bajwa, A. Sayeed, and R. Nowak, "Sparse multipath channels: Modeling and estimation," in *Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009. IEEE 13th*, pp. 320–325, IEEE, 2009.

[39] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[40] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, "Estimation of sparse MIMO channels with common support," *IEEE Transactions on Communications*, vol. 60, no. 12, pp. 3705–3716, 2012.

[41] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 1, pp. 104–118, 2015.

[42] Y. Zhou, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," *IEEE Trans. Wireless Commun*, 2007.

[43] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM wireless communications with MATLAB*. John Wiley & Sons, 2010.

[44] Z. Gao, L. Dai, Z. Lu, C. Yuen, and Z. Wang, "Super-resolution sparse MIMO-OFDM channel estimation based on spatial and temporal correlations," *IEEE communications letters*, vol. 18, no. 7, pp. 1266–1269, 2014.

[45] M. Steinbauer, H. Aspuld, I. Coster, D. Hampicke, R. Heddergott, N. Lohse, and A. Molisch, "Mission report-modeling unification workshop," *European Cooperation in the Field of Scientific and Technical Research,, n. COST*, vol. 259, pp. 1–32, 1999.

[46] D. Shutin and G. Kubin, "Cluster analysis of wireless channel impulse responses with hidden markov models," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. iv–949, IEEE, 2004.

[47] N. Czink, E. Bonek, J. Ylitalo, and T. Zemen, "Measurement-based time-variant MIMO channel modelling using clusters," *Proceedings of the 29th General Assembly of the International Union of Radio Science (URSI'08)*, 2008.

[48] R. He, W. Chen, B. Ai, A. F. Molisch, W. Wang, Z. Zhong, J. Yu, and S. Sangodoyin, "On the clustering of radio channel impulse responses using sparsity-based methods," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 6, pp. 2465–2474, 2016.

[49] A. Mohydeen, P. Chargé, Y. Wang, O. Bazzi, and Y. Ding, "Spatially correlated sparse MIMO channel path delay estimation in scattering environments based on signal subspace tracking.," *Sensors (Basel, Switzerland)*, vol. 18, no. 5, 2018.

[50] W.-J. Chang and J.-H. Tarng, "Effects of bandwidth on observable multipath clustering in outdoor/indoor environments for broadband and ultrawideband wireless systems," *IEEE transactions on vehicular technology*, vol. 56, no. 4, pp. 1913–1923, 2007.

[51] X. Hong, C.-X. Wang, J. Thompson, B. Allen, W. Q. Malik, and X. Ge, "On space-frequency correlation of UWB MIMO channels," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 9, pp. 4201–4213, 2010.

[52] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2207–2225, 2016.

[53] J. Mo, P. Schniter, and R. W. Heath, "Channel estimation in broadband millimeter wave MIMO systems with few-bit adcs," *IEEE Transactions on Signal Processing*, vol. 66, no. 5, pp. 1141–1154, 2018.

[54] X. Li, J. Fang, H. Li, and P. Wang, "Millimeter wave channel estimation via exploiting joint sparse and low-rank structures," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1123–1133, 2018.

[55] N. Shariati, *Robust transmit signal design and channel estimation for multiantenna systems*. PhD thesis, KTH Royal Institute of Technology, 2014.

[56] E. De Carvalho and D. T. Slock, "Cramer-rao bounds for semi-blind, blind and training sequence based channel estimation," in *Signal Processing Advances in Wireless Communications, First IEEE Signal Processing Workshop on*, pp. 129–132, IEEE, 1997.

[57] G. Leus and A. van der Veen, "Channel estimation," *Smart Antennas—State of the Art, Hindawi*, pp. 293–319, 2005.

[58] D. Slock, A. Medles, and E. de Carvalho, *Blind and semiblind MIMO channel estimation*. Cambridge University Press, 2006.

[59] Y. Barbotin, *Parametric estimation of sparse channels: Theory and applications*. PhD thesis, Ph. D. dissertation, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, 2014.

[60] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE signal processing magazine*, 1996.

[61] M. Babtlett, "Smoothing periodograms from time-series with continuous spectra," *Nature*, vol. 161, no. 4096, p. 686, 1948.

[62] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[63] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE transactions on antennas and propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[64] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 37, no. 7, pp. 984–995, 1989.

[65] R. G. Baraniuk, "Compressive sensing," *IEEE signal processing magazine*, vol. 24, no. 4, 2007.

[66] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling [a sensing/sampling paradigm that goes against the common knowledge in data acquisition]," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[67] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE transactions on information theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[68] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes rendus mathematique*, vol. 346, no. 9-10, pp. 589–592, 2008.

[69] Y. Arjoune, N. Kaabouch, H. El Ghazi, and A. Tamtaoui, "Compressive sensing: Performance comparison of sparse recovery algorithms," in *2017 IEEE 7th annual computing and communication workshop and conference (CCWC)*, pp. 1–7, IEEE, 2017.

[70] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on signal processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[71] J. Tropp, A. C. Gilbert, *et al.*, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[72] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Simultaneous sparse approximation via greedy pursuit," in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 5, pp. v–721, IEEE, 2005.

[73] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[74] G. Schwarz *et al.*, "Estimating the dimension of a model," *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.

[75] H. Akaike, "A new look at the statistical model identification," in *Selected Papers of Hirotugu Akaike*, pp. 215–222, Springer, 1974.

[76] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[77] D. B. Williams, "Detection: Determining the number of sources," *Digital Signal Processing Handbook*, 1999.

[78] W. Cheng, Z. Zhang, H. Cao, Z. He, and G. Zhu, "A comparative study of information-based source number estimation methods and experimental validations on mechanical systems," *Sensors*, vol. 14, no. 5, pp. 7625–7646, 2014.

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE NANTES

Titre : Contributions à l'estimation des paramètres du canal MIMO

Mots clés : MIMO, estimation des paramètres de canal, méthodes de sous-espace, acquisition comprimée, statistiques de canal

Résumé : L'utilisation de la technologie MIMO dans les communications sans fil a augmenté de façon remarquable au cours des dernières années. En utilisant plusieurs antennes d'émission et de réception, cette technologie offre de nombreux avantages qui permettent de relever les défis posés par les déficiences du canal sans fil, en particulier les évanouissements par trajets multiples. Cela apporte plusieurs gains de performance importants tels que le gain d'antenne, le gain de diversité et le gain de multiplexage. Plusieurs canaux sans fil présentent des signaux de propagation parcimonieux et en "clusters". Ces propriétés caractérisent les canaux des différentes technologies de communication sans fil. Dans cette thèse, une classe de modèles de canaux MIMO en clusters est considérée, où les composantes de trajets multiples des canaux sont modélisées comme des clusters de rayons multiples autour de retards moyens, chaque cluster étant caractérisée par son retard moyen et son étalement de retards. Les travaux de recherche sont axés sur l'estimation de ces paramètres du domaine temporel. Pour l'estimation du paramètre du retard moyen, deux approches sont proposées, une approche fondée sur l'acquisition comprimée et une approche fondée sur le sous-espace, fondée sur un traitement déterministe du canal. Une approche basée sur le sous-espace est ensuite proposée pour estimer l'écart-type du paramètre d'étalement du retard, à partir d'une description stochastique du canal. Les résultats des estimations sont illustrés par des simulations informatiques.

Title : Contributions to MIMO channel parameter estimation

Keywords : MIMO, channel parameter estimation, subspace methods, Compressive sensing, channel statistics

Abstract : The use of MIMO technology in wireless communications has been increasing remarkably in the last years. By employing multiple transmit and receive antennas, the technology provides a plenty of benefits that allow dealing with the challenges posed by the impairments in the wireless channel, especially multipath fading. It provides several important performance gains such as antenna gain, diversity gain, and multiplexing gain. Several wireless channels exhibit sparsity and clustering of the propagating signals. Such properties characterize channels for different wireless communication technologies. In this thesis, a class of clustered MIMO channel models is considered, where channel multipath components are modeled as clusters of multirays around mean delays, with each cluster characterized by its mean delay and delay spreading. The research work is focused on estimating these time domain parameters. For the mean delay parameter estimation, two approaches are proposed, a compressive sensing based approach and a subspace based approach, based on a deterministic treatment of the channel. A subspace based approach is then proposed to estimate the standard deviation of delay spreading parameter, based on a stochastic description of the channel. Estimation results are illustrated through computer simulations.