Least Mean Squares Channel Estimation for Downlink Non-Orthogonal Multiple Access

Lehlohonolo Sekokotoana

Johannesburg, 2020

A dissertation submitted to the Faculty of Engineering and the Built Environment, University of Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering.

DECLARATION

I declare that this dissertation is my own unaided work. It is being submitted to the degree of Master of Science in Engineering to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other University.

.....

(Signature of Candidate)

……. day of …… year …….

ABSTRACT

Wireless communications as a field of interest has evolved immensely with a seemingly ever-growing impact on modern life. Increasing service demands and innovations posed a need for the fifth generation (5G) wireless mobile network. Non-Orthogonal Multiple Access (NOMA) has been considered as an essential principle to realize some of the 5G requirements. NOMA is a multiple access scheme for accommodation of more users than the available resources. For reliable communication in wireless mobile systems, there is a need for channel estimation (CE). This dissertation considers CE for NOMA-based wireless communication systems. A simplified, two-user single-output-single-input (SISO)-NOMA scenario is considered in order to focus attention on the development of efficient CE approaches. These CE schemes can then be extended to a NOMA model that accommodates a large number of users and multiple antennas.

Four pilot aided CE approaches are presented in this dissertation. The first two approaches consider the constant step-size least mean squares (CSS-LMS) algorithm, the self-adaptive variable step-size LMS (VSS-LMS) algorithm and the self-adaptive VSS-LMS with adaptive speed (VSS-LMS-AS) algorithm. The last two approaches employ the CSS-LMS algorithm. The first approach entails a scenario whereby all the pilots are known to both users. In using the CSS-LMS algorithm, closed form expressions are derived for the optimal step-size and the corresponding error variance. The rest of the approaches assume partial knowledge of the training information. The second approach is the averaging sum-based CE scheme, while the third approach is a case-wise LMS-based CE and detection technique. A successive interference cancellation (SIC)-inspired iterative CE and detection approach is proposed as the fourth scheme.

For all four schemes, the simulation results were in accordance with the theoretical analysis. For both the all-known-pilots scheme and the averaging sum scheme, the CSS-LMS algorithm is the best algorithm in terms of CE speed. This is in comparison to the VSS-LMS and the VSS-LMS-AS. Power allocation was found to affect the CE performance in terms of both CE speed and accuracy for when using the averaging sum-based CE scheme. For the last two approaches detection error was shown to negatively affect the CE performance. As a result, motivating for detection schemes with a sufficiently low probability of error.

DEDICATION

To my family, friends, and my close circle.

ACKNOWLEDGEMENTS

I would like to extend my thanksgiving and appreciation to my supervisor, Prof. Fambirai Takawira. For the opportunity to work in his guidance and nurturing me from inexperience. For his assistance and encouragement in moments I doubted my learning capacity, I am sincerely grateful.

I would also like to express my gratitude to Prof. Olutayo Oyerinde, my co-supervisor. For his dedication in fostering my understanding and his timely counsel. I am genuinely grateful for his eye-opening advice and how he kept me motivated in all our interactions. His mentorship fuelled my strides even when I could not realize the bigger picture.

For financial support, I would like to thank the CeTAS program, the NRF sponsorship through Prof. Oyerinde and the Wits Postgraduate Merit Award.

Appreciation to my family, friends and acquaintances for the moral support and their efforts to ensure my wellbeing.

Table of Contents

Table of Figures

Abbreviations

MA	**M**ultiple **A**ccess
OMA	**O**rthogonal **MA**
NOMA	**N**on-**O**rthogonal **MA**
SISO-NOMA	**S**ingle-**i**nput-**s**ingle-**o**utput-**NOMA**
SIC	**S**uccessive **I**nterference **C**ancellation
BS	**B**ase **S**tation
DL	**D**own-**L**ink
SINR	**S**ignal-to-**I**nterference-and-**N**oise-**R**atio
SNR	**S**ignal-to-**N**oise-**R**atio
AWGN	**A**dditive **W**hite **G**aussian **N**oise
CSI	**C**hannel **S**tate **I**nformation
CIR	**C**hannel **I**mpulse **R**esponse
CE	**C**hannel **E**stimation
MSE	**M**ean **S**quare **E**rror
LMS	**L**east **M**ean **S**quares
ML	**M**aximum **L**ikelihood
LMMSE	**L**inear **M**inimum **M**ean **S**quare **E**rror
CCCP	**C**onstraint **C**oncave **C**onvex **P**rocedure
CSS-LMS	**C**onstant **S**tep-**S**ize-**LMS**
VSS-LMS	**V**ariable **S**tep-**S**ize-**LMS**
VSS-LMS-AS	**VSS-LMS-** with **A**daptive **S**peed
DDCE	**D**ecision **D**irected **C**hannel **E**stimation

CHAPTER 1

INTRODUCTION

1.1 Overview

The telecommunications field has seen notable growth in the recent years, owing to increasing diversity of service demands. Within telecommunications, the field of wireless mobile communication systems has seen a rapid and revolutionary change. This change is mainly influenced by the continual societal advancement and the requirement of state-of-the-art technologies to meet a wide and diverse market of service demands. Based on the requirements which can be referred to in the work presented in [1] and [2], the next generation of mobile networks has promises of possible improvements in smart cities, the internet of things (IoT) and virtual reality [3]. Meeting the fifth-generation mobile networks requirements proves pivotal in supporting the growing service demands.

The growth of wireless mobile communication systems from first generation to the fourth generation (1G – 4G) has been significantly influenced by the evolution of the multiple access schemes [4]. For each generation of wireless mobile communication, there is a wide domain of different areas of interest in terms of research. These research areas are usually centred around the fundamental concepts of wireless mobile communication systems. These concepts include modulation, coding, diversity, channel estimation and equalization, and other signal processing concepts for wireless communication systems. Articles [5] and [6] nominate potential research directions for wireless mobile communication systems. Amongst these, this work is focused on channel estimation for the fifth (5G) generation of mobile networks.

Multiple access (MA) technology design in the research space has documented contributions in advancing the state of wireless mobile networks. Frequency Division MA (FDMA) was the underlying MA technology for 1G. Predominantly, 2G is based on Time Division MA (TDMA), with FDMA also employed. Code Division MA (CDMA) was introduced for 3G while Orthogonal Frequency MA (OFDMA) was introduced for 4G [4], [7]. Spectral efficiency was one of the major issues being addressed in the introduction of each MA scheme. With the massive increase of traffic expected in the next generation of mobile networks, one of the major challenges is the efficient use of the spectrum. It is therefore essential that an effective MA scheme is employed in order to address this challenge.

1.2 Non-orthogonal Multiple Access (NOMA)

Faced with the challenges established for the next generation of wireless mobile networks, NOMA has been proposed. As a radio access technology (RAT), NOMA has been proposed as candidate to realizing some of the major concerns of the 5G of mobile networks [4], [7 – 19]. In various accounts, the underlying theme of NOMA entails accommodation of more users than the available resources. These include the frequency-domain, time-domain and code-domain resources. Essentially, the underlying concept of NOMA is to superimpose the signals of multiple users on the same resource block at the transmitter. A non-linear detection approach is then implemented on the receiver side to separate the superimposed signal of different users [4]. In contrast to the orthogonal multiple access (OMA) based systems, significant improvements introduced by NOMA have been reported. Such improvements include spectral efficiency, system throughput, user fairness, low latency, etc. [4], [7 – 8].

Various categories of NOMA based systems have been established and can loosely be grouped into power domain NOMA (PD-NOMA) and code domain NOMA. With PD-NOMA, multiple user signals are superimposed with different transmit powers and transmitted simultaneously on the same resource block. Successive interference cancellation (SIC) is usually employed at the receiver to differentiate the signal of each user by leveraging the different power levels. Code domain NOMA on the other hand has similarities to code division multiple access (CDMA). The key distinguishing factor of code domain NOMA from CDMA is the use of low cross-correlation non-orthogonal sequences or low-density sequences. In this dissertation, focus is given to PD-NOMA.

Numerous research efforts have been invested with respect to PD-NOMA and a notable knowledge base for potential advancement in wireless communications has been produced. In such accounts, both the uplink and the downlink perspectives are explored. Each perspective is presented in the subsequent chapter, chapter two, which touches on the literature review. Substantial improvements specific to PD-NOMA over OMA have been documented in literature. One of the most celebrated benefits of its use being spectral efficiency. Other benefits highlighted in literature include low signalling and low transmission latency. The improvement of the trade-off between system capacity and user fairness is also highlighted provided there are significant disparities between the user channel conditions.

1.3 Introduction to Channel Estimation

In the overview section of the current chapter, the importance of studying the fundamental concepts of wireless communication systems is pointed out. Channel estimation and equalization are detailed in the afore-mentioned list. Communication in wireless communication systems is achieved by propagation of radio waves from a transmitter to receiver through space. The radio channel, harbours effects that are imposed on the transmitted signal during transmission. For successful recovery of the transmission, these undesirable effects should be estimated and cancelled out [20 – 21]. The undesirable effects encompass signal reflections, diffractions, and signal scattering. Channel estimation is a process of estimating the potential modifications the channel may have effected on the transmitted signal. A certain channel model is adopted, and an approximation of the channel is provisioned for detection (usually at the receiver). As such, it is pivotal for the radio channel considered for transmission to be modelled accordingly. That is, for the channel model to closely reflect the actual channel of interest.

Channel estimation approaches are essentially categorized into pilot aided, decision directed, blind and semi-blind. Section 2.4 of the second chapter provides a detailed account of channel estimation in literature. Each category is attributed with advantages that may motivate for their employment and disadvantages that may otherwise dissuade their use in a given system. Estimation algorithms are employed to realize channel estimation schemes, and these have been vastly documented in literature. For the most part, channel estimation approaches available in documented material address OMA systems. Some of these are covered in Section 2.4 of the second chapter. As of recent, numerous works on channel estimation for NOMA based systems have been recorded. Most of these addresses cover channel estimation for the uplink perspective, with some covering the downlink perspective. Overall, channel estimation for PD-NOMA is a largely unexploited potential research direction.

1.4 Research Background and Motivation

Channel estimation for NOMA based systems is largely unexploited in literature and is considered a potential research direction in the work presented in [8]. Governed by the scope of channel estimation for wireless mobile systems, the work presented in [22] has been studied. In their work presented in [22], Gerzaguet et. al propose three low complexity least mean squares (LMS)-based channel estimation approaches for a Rayleigh flat fading channel. Their

first proposed LMS approach assumes *a priori* knowledge of the channel statistics, employing a conventional LMS estimator to provide the channel estimate. The other two approaches employ self-adaptive stochastic gradient LMS algorithms which can be distinguished by an added adaptive speed. For all three algorithms, a pilot aided scenario is assumed. As drafted, the work presented in [22] only presents the performance of their practical estimator in an OMA application and not for NOMA systems.

A novel training based linear channel estimator for a two-user downlink NOMA system is proposed in the work documented in [23]. In this account, a linear estimator is designed with the intention to maximize the effective signal-to-interference-and-noise-ratio (SINR) of the stronger user, while assuring a bounded effective SINR for the weaker user. A constraint concave convex procedure (CCCP)-based algorithm is employed in manifesting the solution. The CCCP-based approach had significant performance improvements when compared to the maximum likelihood (ML) and linear minimum mean square error (LMMSE) counterparts. By the nature of the work presented in [23], however, it is apparent that the highlight was the joint channel estimation and power allocation. The accuracy of the channel estimator employed was not the focus, given a simple linear estimator was considered. As such, the performance of the CCCP algorithm was predicated on the average effective SINR, which is not directly concerned with the channel estimation performance.

Given the prevailing conditions in the literature, the work in this dissertation is majorly focused on the development of a channel estimation approach for a SISO-NOMA based system, with the main object of interest being the speed and accuracy of channel estimation. As such, this dissertation seeks to undertake the following:

- Extension of the pilot-aided stochastic LMS channel estimation algorithm to a two-user SISO-NOMA system.
- Extension of two self-adaptive stochastic gradient LMS algorithms to a two-user SISO-NOMA system.
- Development of a channel estimation approach with speed and accuracy as the performance considerations.

The above-mentioned LMS algorithms are integrated in the channel estimation approaches as defined below:

- All known pilots channel estimation approach (purely training-based).

- Averaging sum-based channel estimation approach with partial pilot knowledge
- Iterative channel estimation and pilot detection approach assuming partial pilot knowledge.
- Case-wise channel estimation and pilot detection approach assuming partial pilot knowledge.

While the first two channel estimation schemes use all three LMS algorithm variants, the last two are set to use only the constant-step-size stochastic LMS algorithm. The performance of the channel estimation approaches is then evaluated using the MSE criterion.

With this setting, the work in this dissertation seeks to answer the following research question:

> *What Channel estimation approach can be used in the design of a downlink SISO-NOMA system to reduce the estimation error?*

1.5 Research Objectives and Contributions

1.5.1 Research Objectives

In an effort to address the research question highlighted in the previous subsection, the following were the main objectives developed:

A. To develop an all-known-pilots based channel estimation approach for a DL SISO-NOMA system, employing the following algorithms:
    - The conventional constant step-size stochastic gradient LMS algorithm
    - The self-adaptive stochastic gradient LMS algorithm
    - The self-adaptive stochastic gradient LMS algorithm with adaptive speed
B. To develop an averaging sum-based channel estimation approach for a DL SISO-NOMA system with partial pilots, employing the following algorithms:
    - The conventional constant step-size stochastic gradient LMS algorithm
    - The self-adaptive stochastic gradient LMS algorithm
    - The self-adaptive stochastic gradient LMS algorithm with adaptive speed
C. To develop an iterative channel estimation and pilot detection scheme for a DL SISO-NOMA system assuming partial pilot knowledge. The following algorithm is employed:
    - The conventional constant step-size stochastic gradient LMS algorithm

D. To develop a case-wise channel estimation and pilot detection scheme for a DL SISO-NOMA system with partial pilot knowledge. The following algorithm is employed:
    - The conventional constant step-size stochastic gradient LMS algorithm

1.5.2 Research Contributions

The main contributions of the work presented in this dissertation are as follows:

- Translation of the three stochastic gradient LMS algorithm variants to a DL SISO-NOMA system.
- Adoption of an averaging sum formulation and employing three stochastic LMS algorithms to solve an estimation problem for a DL SISO-NOMA system.
- Introduction of a SIC-inspired iterative channel estimation and pilot detection approach that employs a stochastic LMS algorithm for a DL SISO-NOMA system.
- Introduction of a case-wise channel estimation and pilot detection approach for a DL SISO-NOMA system.

1.6 Author Publication Material

Part of this dissertation has been submitted and presented for a publication, while other excerpts are in preparation. The publications the work in this dissertation is related to are as follows:

[1] L. E. Sekokotoana, F. Takawira and O. O. Oyerinde, "Least Mean Squares Channel Estimation for Downlink Non-Orthogonal Multiple Access," in Proceedings of *2019 IEEE AFRICON*, Accra, Ghana, pp.1-5, September 2019.

[2] L. E. Sekokotoana, F. Takawira, O. O. Oyerinde, "Successive Interference Cancellation Inspired Channel Estimation for Downlink Non-Orthogonal Multiple Access ", *submitted to the Transactions on Emerging Telecommunications Technologies*.

[3] L. E. Sekokotoana, F. Takawira, O. O. Oyerinde, "All-Known-Pilots Channel Estimation for Downlink Non-Orthogonal Multiple Access" *in preparation.*

1.7 Dissertation Organization

The organization of this dissertation is as follows.

Chapter 2 details the relevant survey carried out in the creation of this work. Initiated by an introduction, the survey then covers general information with regards to non-orthogonal

multiple access. The survey is thereafter narrowed down to power domain NOMA, the concept of which is presented in both the uplink and downlink perspective. Advancing the chapter, the wireless communication channel is briefed. Then a review of channel estimation in wireless mobile networks is presented. Before a summary of the chapter is provided, literature concerning channel estimation for NOMA systems is also reviewed.

In Chapter 3, the overall system model employed for channel estimation in this dissertation is provided. Both the data and the channel models are described in detail as a foundational measure to usher in the actual work in the subsequent chapters. Furthermore, the all-known-pilots LMS-based channel estimation scheme is presented. Which essentially is the application of the three stochastic LMS variants to the NOMA model. Firstly, covering the constant step-size algorithm, the corresponding closed-form expression for the error variance is presented. The derivations of the self-adaptive step-size algorithms are also effected and presented. Results are presented and evaluated, then a summary is provided.

Chapter 4 introduces the averaging sum problem formulation and presents the three LMS algorithm variants assuming partial knowledge of pilots. Employing the same algorithms as in Chapter 3, the error variance is also presented for the constant step-size algorithm. The three LMS algorithms are then evaluated through computer simulations, results of which are presented within this chapter. The chapter is then summarized.

Two channel estimation approaches are presented in Chapter 5. These are the iterative and the case-wise channel estimation and pilot detection approaches. For both algorithms, the conventional constant step-size LMS algorithm is assumed and presented. Closed-form expressions for the error variance are presented. Computer simulations are carried out in order to assess the performances of both approaches. A conclusive summary of the chapter is then presented.

Chapter 6 presents an overall conclusion of the dissertation and provides potential research directions given the current work.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter comprises the knowledge and information that has been documented thus far with regards to channel estimation for wireless communications. An investigation is carried out covering both channel estimation in Orthogonal Multiple Access (OMA) and Non-Orthogonal Multiple Access (NOMA). Firstly, a survey on NOMA as the radio access technology of interest is presented. Then the power domain NOMA is surveyed, whereby the downlink and uplink perspectives are defined. The brief survey on NOMA is then concluded by a brief account of the current challenges attributed to NOMA as a research area. In order to advance the survey on channel estimation in wireless communication, the wireless communication channel is studied and presented. Upon studying the wireless communication channel and its attributed influence on transmission, the importance of channel estimation is established. The survey on channel estimation methods for wireless communication systems is then presented. Subsequently, channel estimation literature for NOMA systems is addressed both the uplink and downlink perspective. The literature survey chapter is then concluded by a brief summary.

2.2 Non-Orthogonal Multiple Access (NOMA) Overview

To account for the growing service demands in wireless communications, requirements were established for the 5th generation (5G) technology. NOMA as a radio access technology has been deemed candidate to uphold some of the key performance indicators (KPIs) proposed for 5G solutions [8]. These include spectral efficiency, low latency and massive connectivity [1], [8 – 10]. The basic proposition in NOMA is to accommodate more users than the available orthogonal resources. This is contrary to OMA systems, where the orthogonal time-domain, frequency-domain or the code-domain is assigned to each of the multiple users [8]. Essentially, the NOMA concept is achieved by accommodating multiple users within one resource-block [4 – 5], [8], [11 – 17]. The approach with which users are assigned to resource blocks non-orthogonally gives rise to the different categories of NOMA.

Two classifications for NOMA have been identified, these are the code-domain NOMA and the power-domain NOMA. Low-density spreading (LDS), sparse code multiple access

(SCMA) are some of the code-domain NOMA techniques. Other NOMA schemes include pattern division multiple access (PDMA) and bit division multiplexing (BID) [8], [12]. NOMA in general has been documented to show significant performance increase as opposed to OMA. Typical performance improvements are in terms of spectral efficiency, throughput, user fairness, sum rate, etc. [4], [8], [11], [18]. With the advantages and disadvantages attributed to the NOMA schemes when compared, this work is focused on the Power-domain NOMA.

2.2.1 PD-NOMA

Power domain NOMA, herein referred to as NOMA, has been a subject of interest to numerous research efforts. Substantial improvements with regards to some of the performance metrics have been shown with the introduction of NOMA relative to OMA. Some of which include spectral efficiency, low latency support and an improvement in the trade-off between user fairness and system capacity [14 – 16].

Basically, NOMA is a MA scheme that allows for multiple users to transmit modulated signals within the same resource, by allocating each user a unique power level with regards to the relevant channel gain. In particular, more power is allocated to a user with poorer channel conditions relative to the user with better channel conditions [8], [18]. At the receiver end, the signals are differentiated by employment of successive interference cancellation (SIC) for multi-user detection and decoding [8]. The distinct user signal power allocations ensure the feasibility of SIC successfully detecting each signal separately.

The subsequent sub-sections present a two-users scenario for both the uplink and downlink perspectives, as a measure to showcase the basic idea of NOMA.

2.2.1.1 Uplink NOMA Transmission Scenario

Assuming a single-input-single-output (SISO) system, an uplink NOMA scenario entails multiple users transmitting to a base station (BS) at different power levels. Each independent user may transmit messages at any time using the entire allotted bandwidth through their respective channel. On reception of the signals at the BS, the sum of the signals and their attached interferences (channel and additive noise) are decoded employing SIC. Figure 2.1 illustrates a two-user uplink NOMA transmission scenario [14].

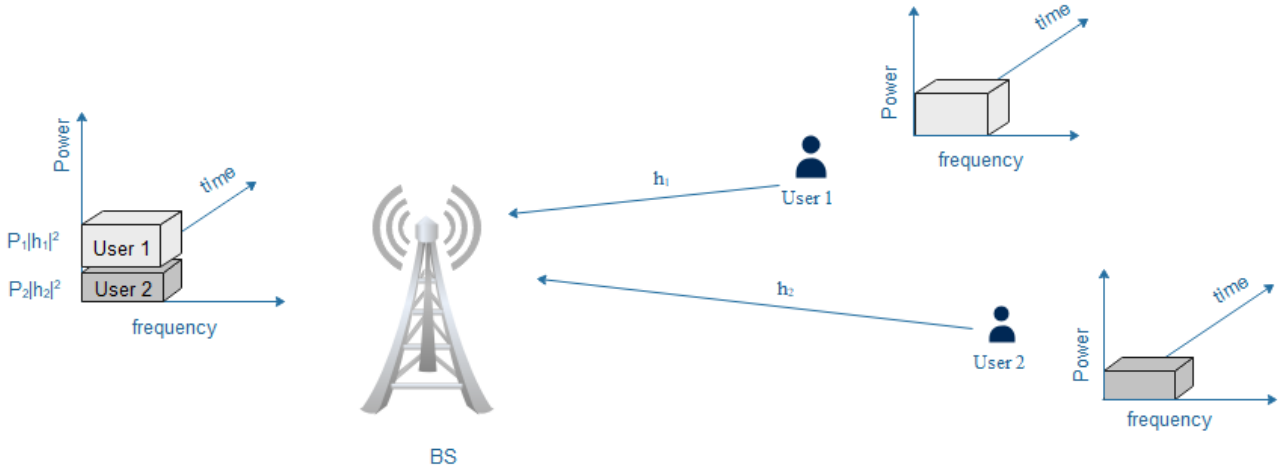


Figure 2.1 Uplink NOMA scenario for two users.

As depicted Figure 2.1, user 1 is assumed to transmit at a higher power than user 2. The signal received at the BS is modelled as shown below:

$$y = \sqrt{p_1}s_1h_1 + \sqrt{p_2}s_2h_2 + w, \tag{2.2.1}$$

where:

y is the received signal,

p_i is the transmit power of the i-th user, sum of which is the total transmit power,

s_i is the signal transmitted by the i-th user,

h_i is the channel coefficient of the i-th user,

and w is the additive white Gaussian noise.

On signal reception at the BS, SIC is performed on the superimposed message signal shown in equation (2.2.1). Firstly, the receiver decodes the stronger signal s_1 treating the rest of the signal as interference. Because s_1 is the dominant component in the superimposed signal y, it is more reliable to decode it prior to the weaker components. After successfully decoding s_1, its component is cancelled from the superimposed signal. The second phase then entails decoding the second signal, s_2. Generally, there are numerous metrics for performance analysis in multiple access (MA) techniques. Equations (2.2.2) gives the throughout analysis for the two-user uplink NOMA scenario:

$$R_i = \log_2\left(1 + \frac{p_i|h_i|^2}{\sum_{j=i+1}^{2} p_j|h_j|^2 + N_0}\right), \quad \forall i = 1,2, \tag{2.2.2}$$

where:

R_i is the throughput of the i-*th* user,

N_0 is the power density of noise component w.

2.2.1.2 Downlink NOMA Transmission Scenario

To present the basic idea of NOMA in downlink transmissions, a system composed of a BS and two users is considered. Information is transmitted from the BS to each user through their respective channels, with a single transmitter and receiver antenna considered. Channel gains between the users and the BS are denoted h_1 and h_2 for users 1 and 2 respectively, both of which are presumed to be known. It is further assumed that $|h_1|^2 \geq |h_2|^2$, thus user 1 is considered to be the strong user while user 2 is the weak user. Power allocation is performed with respects to the channel conditions between the users and the BS. User 1, with better channel conditions relative to user 2, is allocated the least power ratio. The BS transmits signal s_i for each user i where $E\{|s_i|^2\} = 1$ with transmission power p_i. Given total power P, the signal transmitted from the BS to both users is given as:

$$x = \sqrt{p_1}s_1 + \sqrt{p_2}s_2, \tag{2.2.3}$$

where:

x is overall superposition coded signal of signal s_1 and s_2,

s_i is the signal intended for user i,

p_i is the transmission power for user i'*s* signal,

$p_i = a_i P$ for power ratio a_i,

and $\sum_1^2 a_i = 1$ with $a_2 \geq a_1$ and $0 \leq a_i \leq 1$.

At user i, the received signal is expressed as:

$$y_i = h_i x + w_i, \tag{2.2.4}$$

where:

y_i is the overall signal received at user i,

h_i is the complex channel coefficient between user i and the base-station,

w_i is the receiver Gaussian noise and the inter-cell interference at user i's receiver with power density $N_{0,i}$.

Figure 2.2 illustrates the downlink NOMA system for a two-user scenario:

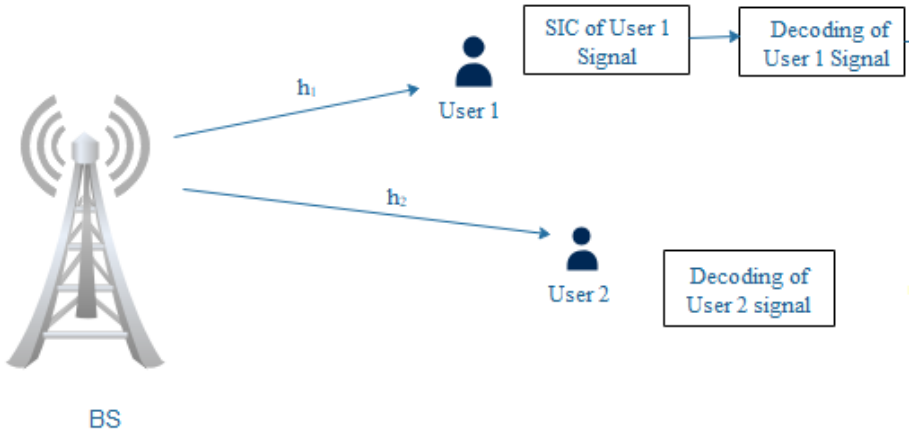


Figure 2.2 Downlink NOMA scenario for two users.

At the receiver of each user, SIC is employed for multi-user detection (MUD). The decoding order in downlink NOMA involves decoding of the strongest signal components prior to the weaker signal components within the superimposed signal. This is determined by ensuring that the decoding process follows the order of increasing channel gain normalized by noise and inter-cell interference power $h_i/N_{0,i}$ [8], [13 – 14]. With this setting, it is apparent that the user with the higher channel gain can successfully decode the signal before detecting its own signal. In the two-user scenario considered, user 2 decodes its signal treating the rest of the signal as interference. SIC is performed at user 1, whereby user 2's signal is decoded and subsequently cancelled out. The throughput expression for each user assuming perfect signal decoding with no error propagation is given in equation (2.2.5):

$$R_i = \log_2\left(1 + \frac{p_i|h_i|^2}{\sum_{j=1}^{i-1} p_j|h_i|^2 + N_{0,i}}\right), \quad \forall i = 1,2, \tag{2.2.5}$$

where:

$N_{0,i}$ is the power density of noise component w_i.

Various accounts in literature have presented the advantages of NOMA over the OMA techniques. The work presented in [14] highlights the potential for the improvement of the trade-off between user-fairness and system efficiency. This is for a scenario whereby there are significant differences in channel conditions amongst users. Robust performance in high mobility scenarios and vehicular communications is mentioned as a benefit in the work presented in [4]. The work presented in [4], [15] and [19] point out the affinity of NOMA to MIMO systems, suggesting high performance gains in NOMA when applied to MIMO systems.

2.2.1.3 NOMA Challenges and Research Opportunities

While NOMA has proven to be a viable means in the realization of the requirements for 5G technology, there are several issues that pose a challenge with the practical implementation. This passage serves as a measure to brief the reader on the challenges and issues that would potentially assert significant improvements in NOMA when investigated.

Optimal or near-optimal power allocation is identified as a viable research investment in [8] and [15]. In practice, the performance of NOMA is limited by intra-cell and inter-cell interferences [15]. It is pivotal therefore, to account for the problem of effective and efficient power allocation together with such limiting factors, which also include the total power constraint. An effective means of user-pairing or user-clustering coupled with limiting the number of users in a cluster will first be handled. Then the issue of dynamic power allocation can be done within a given cluster and/or among clusters to achieve different performance objectives. This solution can be considered alternatively to a brute-force search over the possible user pairs with dynamic power allocation mentioned in [24]. This is because this is a computationally expensive means of achieving the best throughput performance in NOMA.

The extension of NOMA to multiple input multiple output (MIMO) systems in order to derive the benefit of spectral efficiency is also suggested in the work presented in [8] and [25]. NOMA for cognitive radio networks is motivated in [8], and [18], with consideration of the possible benefits that can be derived for using both systems cooperatively. The work presented in [18]

also touches on the potential improvements that would be realized by applying the wireless power transfer principle (simultaneous wireless information and power transfer) to NOMA. Transmission security considerations are also a potential research direction as proposed by the authors of [18].

The need for informed system design is highlighted in [8] and [12]. To achieve this, a thorough theoretical analysis of the achievable rate and overloading bounds is considered to be of importance for NOMA systems. Code-domain NOMA systems relying on sophisticated spreading sequences can derive great benefit from the analysis of its attributed capacity bounds. Knowledge in this regard, as proposed both in [8] and [12], can contribute to an adequate system design in relation to specific application requirements. The overloading factor the system can support is also a matter of importance according to the work presented in [8] and [12]. This can be viewed as a performance metric when considering the interference cancellation capability and tolerable complexity at the receiver.

The receiver design is recognized as one of the challenges that can motivate for potential research directions [8], [12]. In this regard, a number of factors have to be taken into consideration. These factors include receiver complexity, receiver design complexity, SIC propagation errors and channel estimation errors, [8], [12], [14 – 15], [17], [24]. For optimal performance of NOMA based systems, it is essential that the receivers that are put in perspective have low complexity [8], [12]. Furthermore, for SIC based receivers, erroneous detection of signals which come earlier in the decoding order may result in error propagation. This in turn can lead to performance degradation for the succeeding users (in terms of decoding order) [8], [12], [15], [17]. The impact of these errors on the performance of NOMA based system should therefore be accounted for.

Most NOMA contributions rely on the assumption of perfect channel state information (CSI) in the analyses of various aspects of communication systems [8], [25]. Practically, channel estimation errors need to be factored in due to the infeasibility of a perfect CSI acquisition. Imperfect CSI information is therefore of importance when evaluating the achievable performance for NOMA systems. To advance this idea, the application of robust channel estimation techniques is desired in concurrence with the interference implications due to a large number of user support in NOMA. Also, channel estimation is essential for reliable retrieval of transmitted data, hence it is pivotal for the success of NOMA based systems.

Other challenges include system scalability for support of different traffic loading and radio environments in NOMA systems. Cross-layer resource allocation is covered as a challenge in the work presented in [18]. Research in [24] details other implementation issues that encompass signalling and processing overhead, limited number of user pairs and residual timing offset among NOMA users.

2.3 Wireless Communication Channel

Information in wireless mobile communications systems is transmitted by microwave radio signals through a wireless channel [20], [26 – 30]. In its nature, the transmission path between the transmitter and the receiver is variable. It changes from a simple line of sight (LOS) setting, to a path that is extremely obstructed by buildings, foliage and other significant objects that may eclipse the radiation [27], [28], [30]. In the design of a successfully operational wireless communication system, thorough understanding of the nature of the wireless channel of interest is inherent. A wireless channel is essentially characterized by the fluctuations of the channel strength over time and over frequency. Propagation models are fundamental to channel modelling and can be roughly categorized into large- and small-scale fading.

Large-scale fading characterizes the average signal power attenuation or path loss attributable to motion over large areas and it is typically frequency independent. Small-scale fading on the other hand is attributed with constructive and destructive interference of multiple signals travelling in different paths between the transmitter and receiver. These may result in significant alterations in the frequency, phase and amplitude of the transmitted signal in slight spatial changes. Figure 2.3 below illustrates a taxonomical summary of the categorization of channel fading.

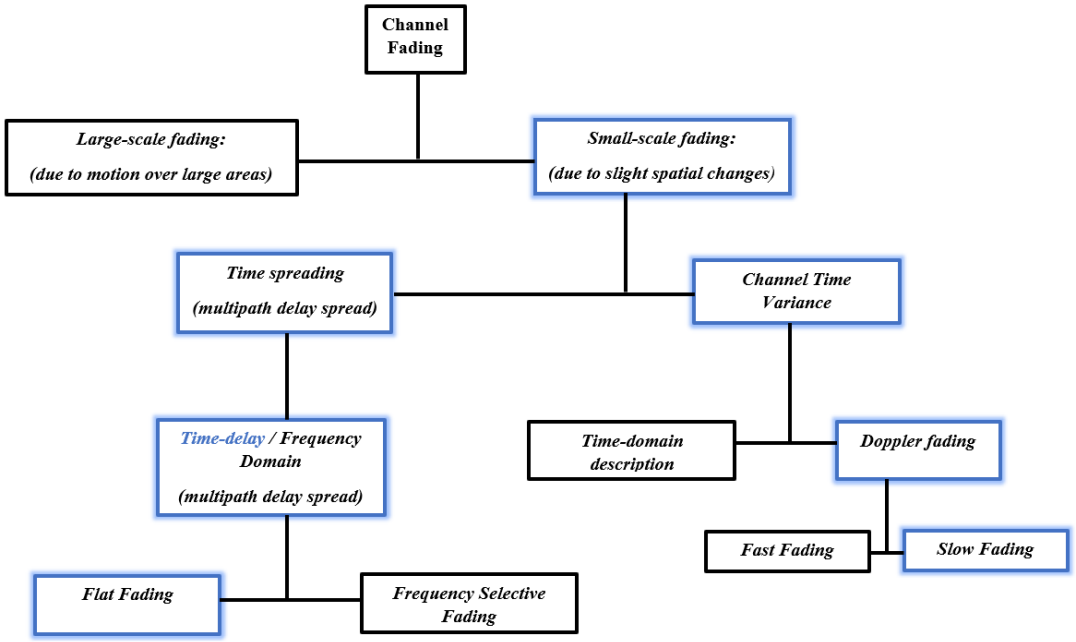


Figure 2.3 Types of fading in Mobile Wireless Communications (adapted from [31]).

The highlighted hierarchies in Figure 2.3 are of interest in this work. Typically, a channel may be modelled as a Rayleigh fading channel, provided there is a lack of a LOS signal component and there is a large number of reflective paths. Contrarily, if there is a dominant LOS propagation path, a Rician probability density function (pdf) is assumed [20], [28 – 29]. A Doppler spread is another property that is mainly employed in a scenario where there is relative motion between the transmitter and the receiver. In a general sense, the wireless channel is modelled as a linear time-varying multi-path channel with additive white Gaussian noise (AWGN). A channel model should closely resemble the actual physical channel for optimal performance to be achieved. The characterization of the channel model therefore serves as a fundamental building block for providing accurate channel state information (CSI).

2.4 Channel Estimation (CE) Methods

2.4.1 Overview

This subsection provides a brief review of CE, among other fundamental concepts, in literature. The wireless communication channel, as described in the previous section, harbours undesirable effects that are imposed on a transmitted signal. These effects result from typical

and arbitrary reflections, diffractions and scatterings of the transmitted signals arriving at the receiver along multiple paths. The interactions of the transmitted signal with the channel may cause multipath signals with different attenuations, distortions, delays and phase shifts [20]. In addition, the influence of additive thermal, man-made or atmospheric noise also impose significant effects on the signal. All these effects of the channel should be accounted for.

It is therefore inherent for the receivers to estimate the effect of the channel on the transmitted signal. Provided the receiver is able to identify how the channel modified the transmitted signal, the signal can be recovered with reasonable accuracy [21]. As such, provisioning of accurate channel state information (CSI) has been deemed a major challenge for optimal performance in wireless communication systems. This is especially required to ensure successful reconstruction of transmitted signals at the receiving end.

Channel estimation can be defined as a means of providing CSI for coherent signal detection. This mainly entails the procedures undertaken to estimate the channel impulse response (CIR) of the system. Defined as the instantaneous state of the dispersive channel encountered [20], CIR is typically modelled as a set of delays and channel coefficients. In this regard, estimating CIR means estimating the attributes that characterize the channel. In literature, CE approaches have been documented with different procedures employed. The disparities and similarities motivate for classifications of channel estimation approaches. Generally, CE approaches can be classified into blind, semi-blind and blind methods. The subsequent sub-sections provide brief surveys on each category. Figure 2.4 presents a taxonomy of CE techniques subject to study in literature:

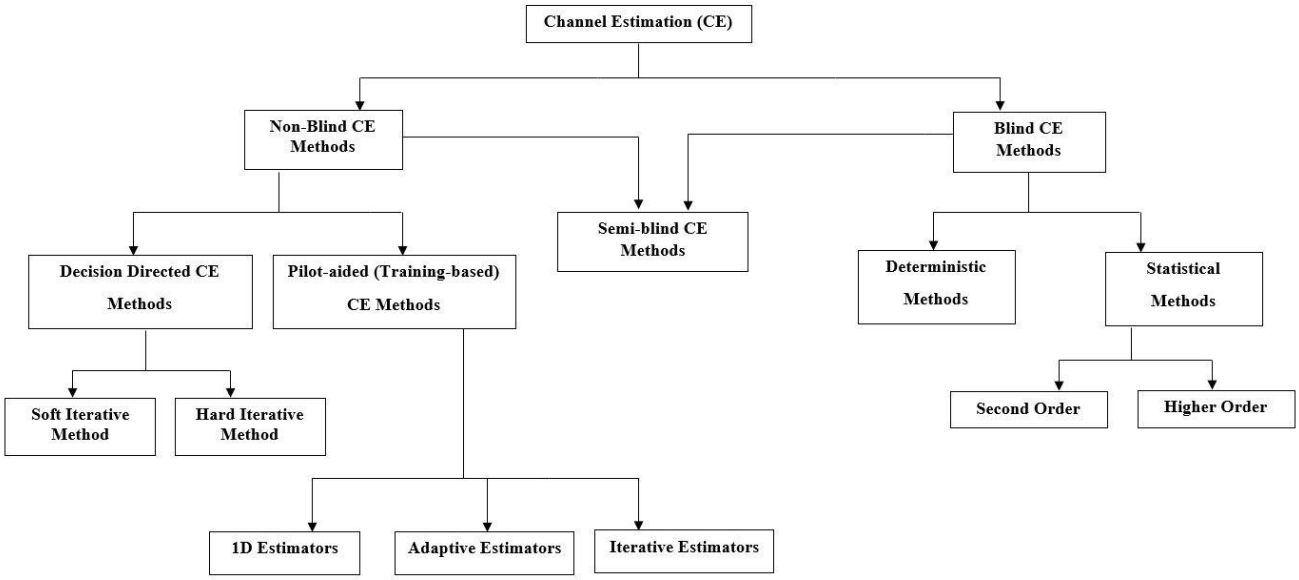


Figure 2.4 General classification of channel estimation techniques.

The following sub-sections expand on the Training-based CE methods as portrayed in Figure 2.4.

2.4.2 Pilot-aided

Training-based CE methods are CE techniques that employ the use of pilot symbols (data) known at the receiver to assist in estimating the channel. In this category, the issuance of pilot symbols may be at the beginning of the transmission or periodically during the transmission. The main advantage of this type of CE scheme is the improved level of estimation accuracy provided the number of pilots employed is high in quantities. However, the limiting factor to this accuracy is set by the spectral efficiency considerations. The main reason for this is that, the amount of pilot symbols employed introduce signal overhead [20], [29].

Various accounts of pilot-aided CE techniques have been documented in literature. The work presented in [22] addresses CE for a flat fading Rayleigh channel with a Jake's Doppler Spectrum. In the paper, Gerzaguet et. al present a Least Mean Square (LMS)-based CE approach with low complexity based on the minimum of asymptotic variance (MAV) criterion. Closed-form expressions for the optimal step-size and the corresponding mean squared error (MSE) of the LMS-based approach are also presented. The paper further introduces the self-adaptive, stochastic gradient approach CE approach that does not require *a priori* knowledge. The proposed approaches are compared to the Bayesian Cramer Rao Bound (BCRB), the auto-regressive Kalman filter (AR-KF) based on the correlation matching (CM) criterion and the AR-KF based on the MAV criterion.

The MAV-based algorithms show improved performance in varying signal-to-noise-ratio (SNR) conditions compared to the AR-KF method based on the CM criterion. These MAV-based algorithms include the LMS-based approach tuned with an optimal step-size. The self-adaptive LMS variants exhibit similar performance to the MAV-based algorithms, with improvements with the addition of adaptive speed. Overall, the BCRB on-line method exhibits the best performance.

M. Morelli and U. Mengali compare two pilot-aided CE methods for orthogonal frequency division multiplexed (OFDM) systems in [32]. The two schemes of interest use the maximum likelihood estimator (MLE) and the Bayesian minimum mean square error estimator (MMSEE). The MLE approach exhibits simplicity relative to the MMSEE counterpart. As reasoned by the authors, this is due to the fact that the MLE approach does not require knowledge of the channel operating conditions. Save for low SNR values and intermediate to high SNR conditions (where both schemes have similar performance), the MMSEE is reported to have better performance. The underlying requirement for the techniques to have similar performance lies with the sufficiently large pilot sequence length in comparison to the length of the CIR. However, for all the cases, the MMSEE is attributed with a higher complexity between the two.

The work presented in [33] proposes a pilot aided CE method for downlink wideband code division multiple access (WCDMA) systems employing an adaptive filter. This is for different fading channel speeds and frequency offsets. Using a fully adaptive smoother-type complex transversal FIR filters, the authors propose that the adaptation of the conventional LMS algorithm to the channel fading characteristics is sufficient. Channel estimation for multiple antenna systems has also been documented in literature. Some accounts of which can be found in the work presented in [34 – 37].

A two-dimensional (2D) pilot aided channel estimation approach for coherent OFDM systems with transmitter diversity is proposed in [34]. Deriving a minimum mean-square error (MMSE) channel estimator, the work documents the accuracy and efficiency of the estimator in tracking the variations of the channels in fast-fading scenarios. In comparison to the decision-directed counterpart, the 2D pilot aided CE approach shows peaked performance given the estimation process is performed in the frequency domain.

The work presented in [35] advances issues concerned with CE in multiple-input-multiple-output (MIMO) OFDM wireless communication systems. Specifically, the performance of

MIMO CE employing training sequences is studied. The estimators of interest in this study are the least squares (LS) estimator, MMSE estimator, and a new scaled LS (SLS) estimator. For each channel estimator, the performance when employing random training matrices is benchmarked to the performance attributed when optimal training matrices are employed. Performance is shown to be greatly improved when optimal training matrices are used in comparison with the scenario whereby random training sequences are employed.

Another LS based pilot aided CE technique is also presented in [36], for MIMO-OFDM systems. In this work, the authors also cover optimal pilot generation. Their work however encompasses both block-type and comb-type pilot structures. With both pilot designs assessed in varying Doppler scenarios, the authors of [36] highlight the significant degradation in performance when employing block-type pilots. This degradation is seen when comparing CE with block-type pilot design in low and intermediate mobile speeds. Relative to the block-type pilot design, the comb-type pilot design is shown to exhibit better performance in time variant scenarios.

More work on CE for MIMO-OFDM is documented in [37]. Hussein Hijazi et. al propose a CE algorithm for MIMO-OFDM systems in fast time-varying environments. The authors adopt the basis expansion model (BEM) as a measure to approximate the fast-varying channel. The coefficient dynamics of the model are used as a foundation for applying the auto-regressive (AR) model. Based on the Jakes process, the AR model introduces the feasibility of estimating and tracking BEM coefficients using a Kalman filter. The iterative algorithm proposed in the paper is a joint CE and data detection algorithm that employs the use of pilots for estimation and enhancement of the BEM coefficient estimates. Tallied with the theoretical analysis, the simulation of the algorithm is shown to have a considerably good performance for high Doppler scenarios.

A whole new paradigm of CE approaches is introduced in [38] and [39], employing deep learning for CE. These two accounts of CE using deep learning are loosely categorized as pilot-aided given there is usage of pilots. However, they are not necessarily pure pilot-aided CE approaches. In [38], deep learning for CE and signal detection is proposed for OFDM systems. For the deep learning approach, CSI is estimated implicitly and is used to recover transmitted symbols explicitly. The CE essentially entails employment of a deep neural network (DNN) which is trained offline using the simulation generated data of the channel statistics. The deep learning approach is shown to be more robust relative to the conventional CE methods in the

case whereby fewer training pilots are employed, among other conditions. It is also pointed out that the deep learning-based CE method is more advantageous compared to the MMSE estimator in cases whereby wireless channels are affected by serious distortions and interference.

A deep learning-based channel estimator for time-varying wireless channels is developed in [39]. Specifically designed for channels that are usually hard to estimate and track when modelled with finite parameters, the proposed estimator employs DNN based channel estimators to track channel variations. Firstly, the DNNs are trained offline, then the DNN estimator is initially trained with preamble training symbols. The estimators are then updated using pilots through incremental learning. The proposed scheme is shown to have better performance in comparison to the LS estimator for CE and tracking. Generally, CE and tracking can potentially be improved with the advancement of deep learning applications.

2.4.2.1 Pilot-aided Channel Estimation Algorithms

The following subsections explore simple versions of the conventional pilot aided CE algorithms as outlined in [21] and [22]. The motive behind the subsequent presentation being to document the general idea behind pilot aided CE sufficient for enabling application in any system. Orthogonal multiple access (OMA) systems are assumed for all the algorithms. Firstly, the least-squares (LS) algorithm is presented, followed by the maximum likelihood (ML) and the minimum mean square error (MMSE) algorithm for narrowband wireless communication systems. The least mean square algorithm is presented for a Rayleigh flat fading channel in the last part of this subsection.

2.4.2.2 System Model

To demonstrate the basic principle behind pilot-aided CE, we consider a simple single antenna OMA system as with the work presented in [21]. Signal bits s_n are transmitted from the sending end through a channel modelled by coefficients h_n. For this scenario, only coefficient estimation will be addressed, and noise samples are considered to be uncorrelated. At the receiver, the observation y_n in the presence of additive white Gaussian noise (AWGN) is generally expressed as:

$$y_n = \sum_{l=0}^{L-1} h_n(l) s_{n-l} + w_n, \tag{2.4.1}$$

where

` y_n is the received signal (observation) at instance n,

w_n corresponds to the AWGN samples for each received symbol at instance n, with variance σ_w^2,

$h_n(l)$ is the channel coefficient constituted by the influence of the channel conditions,

L represents the number of channel taps (multipaths),

and n denotes the symbol time dependency.

In the observation given by equation (2.4.1), the channel is modelled as a multipath channel. It should be noted that for a single path channel model (one channel tap), $L = 1$. The observation is therefore simplified to the expression given in (2.4.2):

$$y_n = h_n s_n + w_n. \tag{2.4.2}$$

In this survey, a Rayleigh slow fading channel is considered which assumes Jake's Doppler spectrum. The Jake's Doppler spectrum for the channel gain in (2.4.2) is expressed as [30], [22] and [40]:

$$\Gamma_h(f) = \begin{cases} \dfrac{\sigma_h^2}{\pi f_d \sqrt{1 - \left(\dfrac{f}{f_d}\right)^2}} & if\ |f| < f_d \\ 0 & if\ |f| > f_d \end{cases}, \tag{2.4.3}$$

where:

$f = f_d \cos\varphi$ is the doppler frequency of the wave signal arriving at angle φ,

f_d is the maximum doppler frequency,

σ_h^2 is the variance of the channel coefficients.

Assuming the mobile receiver is surrounded by uniformly distributed scattering objects, and that different channel coefficients are uncorrelated, each channel coefficient can be modelled as a complex Gaussian random variable. The autocorrelation function of the channel coefficients in (2.4.2) is expressed as [21]:

$$R_h[m] = E\{h_n h_{n+m}^*\}$$

$$= \sigma_h^2 J_0(2\pi f_d m T_s), \quad (2.4.4)$$

where:

mT_s is the autocorrelation delay,

σ_h^2 is the mean square value of the channel coefficient,

f_d is the doppler spread, which is proportional to the vehicular speed,

$J_0()$ is the zeroth Bessel function of the first kind.

This single-tap channel model is adopted for the least mean squares (LMS) approach presented in subsection 2.4.2.6. The rest of the approaches presented in this section assume the model given in equation (2.4.1). For convenience, equation (2.4.1) can be expressed in vector form. Assuming *N* total received samples and defining column vectors as follows:

$$\boldsymbol{y} = [y_0\ y_1 \dots\ y_{N-1}\ ]^T,$$

$$\boldsymbol{h} = [h(0), h(1), \dots, h(L-1)]^T,$$

$$\boldsymbol{w} = [w_0\ w_1 \dots\ w_{N-1}\ ]^T.$$

The expression in (2.4.2) is expressed in vector form is as shown below:

$$\boldsymbol{y} = \boldsymbol{B}\boldsymbol{h} + \boldsymbol{w}, \quad (2.4.5)$$

where:

$\boldsymbol{B}$ is a $N \times L$ matrix with rows corresponding to different shifts of the transmitted sequence of the symbols s_n.

The data model defined above is used throughout the survey, with further additions made for special cases in the subsequent subsections. Complimenting the data model, the channel coefficients are modelled as uncorrelated, zero-mean complex Gaussian random variables. Both the data model and the channel model constitute the system model.

Channel estimation performance can be evaluated using analyses of symbol error rate (SER), bit error rate (BER) and/or the mean square error (MSE) among other metrics. The MSE can be defined as the expected value of the error in channel estimation. To showcase this, the estimation error can be defined as shown below:

$$\boldsymbol{e} = \boldsymbol{h} - \widehat{\boldsymbol{h}}, \quad (2.4.6)$$

where $\boldsymbol{e} = [e_0, e_1, \ldots, e_{L-1}]$ is the estimation error vector, and $\widehat{\boldsymbol{h}} = [\hat{h}_0, \hat{h}_1, \ldots, \hat{h}_{L-1}]$ is the channel coefficient estimate.

Channel estimation performance can be quantified using the mean square error (MSE), otherwise referred to as the error covariance. In this scenario, the MSE matrix is defined as:

$$\boldsymbol{MSE} = E\{\boldsymbol{e} * \boldsymbol{e}^H\}, \tag{2.4.7}$$

where $E\{\cdot\}$ is denotes the expected value.

2.4.2.3 Least-squares (LS) Estimation

This sub-section assumes the system model described in the previous sub-section. In various applications of pilot aided estimation, LS estimation is usually employed [20 – 21], [29]. The channel coefficients are estimated as shown below, using the LS estimator [21]:

$$\widehat{\boldsymbol{h}}_{LS} = arg\min_{\boldsymbol{h}}(\boldsymbol{y} - \boldsymbol{B}\boldsymbol{h})^H(\boldsymbol{y} - \boldsymbol{B}\boldsymbol{h}),$$

where $\boldsymbol{y}$ is formed by $K = N - (L - 1)$ samples, N is the length of the training sequence (number of pilot symbols), L is the number of multipath channels. In single tap scenario, $K = N$.

Differentiating the estimator $\widehat{\boldsymbol{h}}_{LS}$ with respect to the channel coefficient and equating the result to zero, the LS channel estimate is given as:

$$\widehat{\boldsymbol{h}}_{LS} = (\boldsymbol{B}^H\boldsymbol{B})^{-1}\boldsymbol{B}^H\boldsymbol{y}. \tag{2.4.8}$$

On establishing the LS estimator, the estimation error variance analysis can be performed. By firstly substituting the expression of $\boldsymbol{y}$ from (2.4.5) into (2.4.8), we get:

$$\widehat{\boldsymbol{h}}_{LS} = \boldsymbol{h} + (\boldsymbol{B}^H\boldsymbol{B})^{-1}\boldsymbol{B}^H\boldsymbol{w}. \tag{2.4.9}$$

From (2.4.9), the expression of the MSE matrix can therefore be expressed as $(\boldsymbol{B}^H\boldsymbol{B})^{-1}\sigma_w^2$.

Provided the autocorrelation properties of the pilot symbols are ideal, the overall estimation variance error is minimized when $\boldsymbol{B}^H\boldsymbol{B} = (N - (L - 1))\boldsymbol{I}$, the resultant MSE matrix is therefore given as:

$$MSE_{LS} = \left(\frac{\sigma_w^2}{N - (L - 1)}\right)\boldsymbol{I}. \tag{2.4.10}$$

With LS estimation, the estimation error variance depends on the noise power, the length of the training (pilots) sequence, the number of channel taps and the correlation properties of the

training sequences. However, the case presented above assumes a scenario of uncorrelated channel coefficients.

2.4.2.4 Maximum Likelihood (ML) Estimation

This section provides an ML-based pilot aided CE approach. The system model described in the previous sub-section is also assumed. Maximum likelihood estimation maximizes the likelihood of received data samples given the channel coefficients, the solution of which is expressed:

$$\widehat{\boldsymbol{h}}_{ML} = arg \max_{\boldsymbol{h}} p(\boldsymbol{y}|\boldsymbol{h}),$$

where $p(\boldsymbol{y}|\boldsymbol{h})$ denotes the conditional probability density function of $\boldsymbol{y}$ given $\boldsymbol{h}$.

The solution to this problem is obtained by maximizing a quantity related to the log likelihood that is dependent on noise. For complex Gaussian noise, the solution is expressed as:

$$\widehat{\boldsymbol{h}}_{ML} = arg \max_{\boldsymbol{h}}[-(\boldsymbol{y} - \boldsymbol{B}\boldsymbol{h})^H \boldsymbol{R}_w^{-1}(\boldsymbol{y} - \boldsymbol{B}\boldsymbol{h})], \tag{2.4.11}$$

where $\boldsymbol{R}_w$ is the noise covariance matrix.

Differentiating (2.4.11) with respect to $\boldsymbol{h}$ and equating to zero, the estimate becomes:

$$\widehat{\boldsymbol{h}}_{ML} = (\boldsymbol{B}^H \boldsymbol{R}_w^{-1} \boldsymbol{B})^{-1}(\boldsymbol{B}^H \boldsymbol{R}_w^{-1})\boldsymbol{y}. \tag{2.4.12}$$

The corresponding MSE matrix is expressed as shown below:

$$MSE_{ML} = (\boldsymbol{B}^H \boldsymbol{R}_w^{-1} \boldsymbol{B})^{-1}. \tag{2.4.13}$$

2.4.2.5 Minimum Mean Square Error (MMSE) Estimation

In this sub-section, MMSE estimation in a pilot-based estimation scenario is presented as adapted from the work presented in [21]. Even though it is dependent on the moments of the channel coefficients, MMSE estimation is independent of the distribution of the channel coefficients that are to be estimated. The MMSE channel estimator is defined as shown below:

$$\begin{aligned}\widehat{\boldsymbol{h}}_{MMSE} &= \arg\min_{\boldsymbol{h}}\{(\boldsymbol{e})^H(\boldsymbol{e})\} \\ &= \arg\min_{\boldsymbol{h}}\left\{\left(\boldsymbol{h} - \widehat{\boldsymbol{h}}\right)^H\left(\boldsymbol{h} - \widehat{\boldsymbol{h}}\right)\right\}.\end{aligned} \tag{2.4.14}$$

Considering $\boldsymbol{h}$ to have a zero mean and covariance $\boldsymbol{R}_h$, equation (2.4.14) can be simplified to

$$\widehat{\boldsymbol{h}}_{MMSE} = \boldsymbol{R}_h \boldsymbol{B}^H (\boldsymbol{B}\boldsymbol{R}_h \boldsymbol{B}^H + \boldsymbol{R}_w)^{-1} \boldsymbol{y}. \quad (2.4.15)$$

Assuming that the channel coefficients are Gaussian, the MSE matrix of MMSE estimation is that of the LS estimation scaled by some factor $a = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_w^2/(N-J+1)}$, therefore the MSE matrix is given as:

$$MSE_{MMSE} = a \cdot MSE_{LS}. \quad (2.4.16)$$

$$MSE_{MMSE} = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_w^2/(N-J+1)} \cdot MSE_{LS}. \quad (2.4.17)$$

2.4.2.6 Least Mean Squares (LMS) Estimation

This sub-section explores a stochastic LMS CE approach presented in [22]. Contrary to the previously presented pilot-aided estimation approaches, estimation is performed at symbol level. That is, estimation is considered for each symbol in the training sequence, as opposed to the case of an entire sequence. As such, y_n is adopted as the observation, the estimation error is defined as $e_n = h_n - \hat{h}_n$ and the estimation error variance is given as $\sigma_e^2 = E\{|e_n|^2\}$

For simplicity it is assumed that the symbols are perfectly decided, and each symbol is given as $s_n = 1$. The following is adopted as the cost function to be used in performing a gradient descent:

$$J(\hat{h}_n) = E\left\{\left|y_n - \hat{h}_n\right|^2\right\}, \quad (2.4.18)$$

where:

$\hat{h}_n$ is the channel estimate,

$J(\hat{h}_n)$ is the cost function.

The update equation is of the form:

$$\hat{h}_n = \hat{h}_n - \mu \frac{\partial J(\hat{h}_n)}{\partial \hat{h}_n}. \quad (2.4.19)$$

In differentiating the cost function given in (2.4.18) and substituting the result into the equation given in (2.4.19), the update rule becomes:

$$\hat{h}_n = \hat{h}_{n-1} + \mu(y_n - \hat{h}_{n-1}), \quad (2.4.20)$$

where μ is the learning rate.

The authors of the work presented in [22] advance their work by providing closed-form expressions of the optimal step-size and that of the corresponding asymptotic MSE. These optimal expressions are wrought from minimizing the error variance expression approximated as given below:

$$\sigma_e^2 = \frac{1}{2}\left[\frac{2\pi f_d T}{\mu}\right]^2 \sigma_h^2 + \sigma_N^2 \frac{\mu}{2}, \tag{2.4.21}$$

where σ_e^2 is the asymptotic MSE.

The optimal step-size and asymptotic MSE, found by minimizing the error variance in (2.4.21), are respectively expressed as shown below:

$$\mu_{Opt} = 2(\pi f_d T)^{2/3}\left(\frac{\sigma_h^2}{\sigma_w^2}\right)^{\frac{1}{3}}. \tag{2.4.22}$$

$$\sigma_{e,(Opt)}^2 = \frac{3}{2}(\sigma_h^2)^{\frac{1}{3}}(\pi f_d T \sigma_w^2)^{\frac{2}{3}}. \tag{2.4.23}$$

where:

μ_{Opt} is the optimal step-size,

$\sigma_{e,(Opt)}^2$ is the optimal asymptotic MSE.

2.4.3 Decision-Directed Channel Estimation (DDCE) Methods

Decision-directed CE can be identified by the employment of pilot symbols as well as the re-modulated detected message symbols. Essentially, this approach provides more reliable channel estimates relative to purely pilot aided schemes [20]. Fundamentally, DDCE techniques initially use pilot symbols to provide the initial channel estimate, which in turn is used as the channel estimate for detection. The currently received pilot symbols and re-modulated symbols are then used in collectively to obtain a *posteriori* channel state function (CTF).

On reception of the subsequent symbols, the CTF is used as the *a priori* channel estimate for demodulation. As opposed to the purely pilot based CE techniques, DDCE techniques have less pilot symbol overhead. Two categories of DDCE methods are recognized in literature, these are the hard iterative DDCE methods and the soft iterative DDCE methods. The rest of the subsection documents accounts of DDCE-based CE approaches in literature.

An investigation on DDCE is conducted for multiple transmit antenna (MIMO) systems in [41]. In MIMO systems, optimal CE is limited by the high computation complexity involved in the inversion of data-dependent matrices. To address this limitation, the authors of [41] propose an iterative DDCE method that is set to avoid matrix inversion. This is achieved by the introduction of a channel estimator that is basically a data-independent filter, applied to the weighted sum of minimum norm LS estimate coupled with the estimate from the previous iteration. This proposed method reduces to an optimal estimator in single antenna applications, this is provided the weighted parameter is chosen optimally. The performance of the proposed method is assessed through simulation, employing a space-time code system with joint CE and decoding. With a reasonable number of iterations, the results show that the proposed iterative method achieves near optimal performance.

Another account of a DDCE method can be referred to in the work presented in [42]. Gao J. and Liu H. present a decision directed maximum *a posteriori* (MAP) CE scheme for MIMO systems in time-varying Rayleigh fading channels. Employing the use of a zero-forcing (ZF) receiver, spatially multiplexed symbols are detected within a given symbol interval. The detected symbols (symbol estimates) are then used in the decision directed MAP channel estimator to provide *a priori* channel coefficient estimates, for detection in subsequent symbol intervals.

In comparison with other schemes, the authors suggest that the proposed system operates with less pilot symbols for initial estimation. Furthermore, the proposed system is considered to have lower complexity and low pilot symbol overhead relative to other existing schemes. In assessing the proposed scheme, the authors of [42] show that it is more appropriate for channels with low to medium Doppler shifts. The results also highlight that the proposed scheme works fairly well in high SNR values without reaching an error floor implicated in schemes based on the qausi-static channel models.

In the work presented in [43], an enhanced DDCE scheme that assumes a projection approximation subspace tracking (PAST)-aided, fractionally-spaced CIR (FS-CIR) estimator is proposed. The performance of the proposed PAST-aided FS-CIR DDCE scheme is compared to the achievable performance of the Sample-Space CIR (SS-CIR) DDCE scheme. The motivation for the development of the proposed scheme was to provide a more realistic channel model and associated channel estimator assuming an OFDM-based system. Both methods were analysed in time-variant Rayleigh fading FS-CIR conditions.

With simulated results, the proposed PAST-aided FS-CIR DDCE scheme is shown to have a better achievable system performance as opposed to the SS-CIR DDCE scheme. This is in terms of the bit-error rate (BER) analysis, whereby the former outperforms the latter for the entire SNR values considered. In terms of the MSE analysis, the DDCE technique employing the SS-CIR estimator is attributed with a noise-floor at high SNR values, which is not the case with the DDCE technique based on the FS-CIR estimator.

In an effort to provide a more robust CE scheme, the authors of the work presented in [44] propose a fast data projection method (FDPM)-based CIR estimator of a DDCE scheme for OFDM systems. Assuming the context of the FS-CIR based channel model, the authors apply the FDPM CIR estimator based DDCE scheme and benchmark it with the PAST-aided FS-CIR DDCE scheme presented in [43]. Performance for both algorithms is analysed for fast and slow fading scenarios. The FDPM-based DDCE scheme is shown to outperform PAST-aided DDCE scheme in terms of the MSE for the entire range of SNR values considered. The performance gain of the FDPM-based FS-CIR estimator over the PAST-aided FS-CIR estimator is also documented in terms of the BER analysis.

Decision directed CE employing an LTE pilot structure in a MIMO-OFDM downlink system is presented in [45]. To improve the performance of the estimator in high velocity cases, the space-alternating generalized expectation-maximization (SAGE) algorithm is employed for low density of pilot symbols. Essentially, the SAGE algorithm is used to recompense the performance loss incurred due to decreased pilot symbol density in an iterative receiver. Ultimately, as the authors suggest, the SAGE algorithm reduces pilot overhead without performance degradation in the CE process.

Through computer simulation, the performance of the SAGE algorithm used with the LS estimator for the initial estimate is compared to that of a LS estimator. With the results, the SAGE algorithm is shown to compensate for the CE performance degradation caused by the long estimation latency in high user velocities. Furthermore, the reduction of pilot overhead is documented with the employment of the SAGE channel estimator.

2.4.4. Blind Channel Estimation

The major setback with the employment of pilot aided CE techniques is the overhead introduced by employment of training sequences. Pilot overhead is more significant in time varying channels as opposed to time-invariant channels due to the high number of pilots that

suffice for estimation. Motivated to improve the bandwidth utilization, blind CE approaches have since been investigated. Blind CE techniques fundamentally estimate the channel by employing intrinsic information within the received signals, the structural properties of the transmitted information, and the statistical information of the channel [20], [46].

In one blind approach, the sent signal maybe assumed to be random with recognised statistics. While in the other approach, the signal is assumed to be random without knowing the statistical description [47]. For the second approach, the observation is considered to be deterministic [20]. This gives rise to two categories of blind estimation methods: the former definition being the statistical blind CE while the latter refers to the deterministic blind CE. Blind identification in its various applications has been covered and documented in literature. The next paragraphs will delve into the literature in the two classes of blind channel estimation.

Firstly, statistical blind CE is explored, commencing with the account given in the work presented in [48]. A statistical blind CE approach for un-coded OFDM open-loop systems is proposed in the work presented in [48]. The proposed blind CE approach basically entails a decision algorithm, initially making a primary symbol estimate of the transmitted signal in each subcarrier based on a constraint linear MMSE (LMMSE) criterion. These primary data estimates are then employed for channel CE using the MMSE estimator. Subsequently, the channel estimates are used for standard CE.

In terms of performance, the blind CE technique presented in [48] is benchmarked with three approaches: the optimal coherent and the optimal differential detection approach, a decision-directed Kalman-filtering based technique and two pilot aided OFDM schemes. Evaluating the performance of the candidate techniques through simulations, the following were the interpretations of the results: through a BER analysis, the proposed blind method is shown to outperform the optimum differential detection and the Kalman-based channel tracking approaches. In comparison to the pilot aided systems, the comb-type and the block-type variants, the BER performance favours the proposed blind CE method for SNR values beyond 10dB. Additionally, the proposed blind CE method shows improved performance for varying channel conditions and it introduces better spectral efficiency relative to both pilot-aided variants. However, the blind CE technique has a significantly high complexity.

Another account of blind CE technique based on statistical information is presented in [49], where a blind CE algorithm for Space-Time Block Coded (STBC) systems is proposed. In this algorithm, the channel matrix is estimated by minimizing a Kurtosis-based cost function after

zero-forcing. The proposed method is suggested to be applicable to a general class of linear STBCs and advantageous to other approaches given it does not require transmitter modification.

In addition, the proposed method is declared to have low complexity, operating without initialization or employment of pilot sequences. The proposed algorithm outperforms the joint approximate diagonalization of eigen matrices (JADE) algorithm at its best and matches its performance otherwise, in the case of spatial multiplexing. In the case of identifiable orthogonal STBCs, the proposed approach yields a similar performance with the closed-form second-order statistics approach.

The work presented in [50] considers quadrature amplitude modulated (QAM) data links and proposes Second-Order Statistics (SOS) driven LMS blind channel equalization. By defining a novel cost function, a novel LMS adaptation rule is proposed with application to fractionally spaced blind channel equalization schemes. To complement the defined cost function, selected classical blind cost functions are also incorporated. The proposed adaptive scheme is shown to mitigate the divergence attributed to certain blind equalization algorithms, more significantly in large symbol constellation order and frequency selective channels. Computer simulation results show that the proposed LMS scheme has significant reductions of convergence failures in comparison to the classical LMS adaptation. Furthermore, the proposed CR based cooperative LMS rule is deemed candidate for joint usage with other blind equalization cost functions.

The works presented in [51 – 49] give accounts of the other class of blind CE, the deterministic blind CE approaches. Doubly-selective channels are considered in [51] and a deterministic subspace based blind CE algorithm is proposed. The proposed algorithm operates on time and frequency shifted versions of the received signal for a two receive antenna systems. This is contrary to the case where the received sequence is shifted solely in time or frequency, in which case multiple receive antennas are required.

Due to its deterministic nature, the approach needs only short data record to operate reliably. Benchmarked with the stochastic subspace based blind CE algorithm, computer simulations are presented with the normalized root mean square error (NRMSE) as the metric. Within the defined conditions for the proposed algorithm's operation, the results depict improved performance for the proposed algorithm relative to the stochastic subspace based blind CE algorithm.

A blind CE method for a block transmission system using fractional sampling at the receiver is introduced in [52]. The method capitalizes on the discrete-time, band-limited signal with a smooth waveform for estimating the CIR using an interpolation formula. The smooth waveform results from the received signal being sampled at a sufficiently high rate. For the method to work, knowledge of the input signal statistics or the received signal autocorrelation is not required. Furthermore, the proposed blind CE method is considered to be robust against overestimation, given that it is free of channel order estimation.

Through computer simulations, the proposed interpolation method is gauged against the residue polynomial-, fractionally sampled subspace-, the baud rate subspace- and the coding blind CE methods. The interpretation of the results show that the proposed method exhibits better performance relative to the four benchmark methods in the case of high SNR values. With an added advantage of not having to compute eigenvectors of matrices, the major setback with the proposed system is the need for more sampled points at the receiver.

Considering FIR filter channel models, the work presented in [53] proposes a blind deterministic channel and source estimation approach for sensor networks. In this work, a system identification technique is adapted for distributed CE, introducing two stable distributed algorithms. A Single-Input-Multiple-Output (SIMO) system is considered in a noise-free scenario as well as noisy scenario for application of the distributed algorithms. Both algorithms are compared with centralized algorithms through computer simulations. For less storage space and computation, the distributed algorithms yield results comparable to the centralized algorithms. In noisy environments, two stable distributed algorithms are provided with effective solutions for the blind CE problem.

2.4.5 Semi-Blind Channel Estimation

Semi-blind CE techniques can be thought of as hybrid techniques combining definitive principles of pilot aided and blind CE techniques. Basically, semi-blind CE methods employ both the known pilot information and the intrinsic information in the unknown received signals for CE purposes. Semi-blind CE methods belong to a class of algorithms that employ superimposed periodic training data to acquire channel coefficient estimates based on first or second order statistics of the channel [20], [54]. In essence, they are a mediation of the trade-offs that arise from either usage of blind or pilot aided CE methods. This subsection provides brief accounts of semi-blind CE in literature.

An adaptive semi-blind CE approach for single-cell massive MIMO systems is developed in [55]. Motivated by the complexity assumed by traditional CE schemes for massive MIMO, the authors of the work presented in [55] introduce an adaptive channel estimator. Primarily, the focus of the proposed estimator is to reduce the complexity usually exhibited by the CE schemes based on eigenvalue decomposition (EVD) or singular value decomposition (SVD). The estimator is based on the fast-single compensation approximated power iteration (FSCAPI) subspace tracking algorithm. The FSCAPI-based algorithm is attributed with fast convergence, good orthogonality and low computational complexity.

In comparison with the EVD-based estimator, the FSCAPI-based semi-blind channel estimator is shown to have improved performance in terms of the estimation error and channel capacity. Gauged with the SVD-based estimator, the adaptive FSCAPI-based estimator exhibits comparable performance. However, the complexity of the proposed estimator is significantly lower than that of the SVD-based channel estimator.

Assuming an OFDM-based amplify-and-forward (AF) two-way relay network (TWRN), a CE problem is considered in [56]. A more realistic assumption of frequency-selective environments is made by supposing non-reciprocal channels. Based on this channel assumption, a semi-blind CE approach is established for OFDM-based AF TWRNs. This semi-blind channel estimator is based on Gaussian maximum likelihood (GML) estimation, which treats data symbols as Gaussian-distributed nuisance parameters. An iterative quasi-Newton approach is used to acquire GML estimates. Additionally, a superimposed training strategy is employed at the relay to assist in the estimation of individual channels.

In advancing their idea, the authors of the work in [56] also propose conditions for optimality of training pilots. The performance of the proposed approach is then evaluated through simulations, comparing it to that of the derived semi-blind and pilot aided Cramer-Rao bound (CRB). The proposed estimator is shown to have significant improvements in accuracy over the conventional pilot-based estimator. On the other hand, the performance of the proposed estimator approaches that of the semi-blind CRB with the increase in SNR. Also, it is established that the proposed technique poses favourable performance gains at reasonable computational costs.

The work presented in [57] proposes a semi-blind channel estimator for massive MIMO systems based on an iterative space-alternating generalized expectation maximization (SAGE) approach. Initializing the estimation process with a pilot based MMSE estimator, the approach

is updated iteratively using the SAGE algorithm. The SAGE algorithm uses the data symbols received instead of pilots for updating the channel estimates, this in turn significantly reduces the pilot overhead. The proposed approach is centred around the premise that the base-station (BS) has knowledge of the large-scale fading coefficients of its cell and is unaware of the interfering cells. Essentially, the proposed method has two phases: the initialization phase and the implementation of the SAGE algorithm.

The initialization phase entails provisioning of the channel estimates through a pilot aided MMSE estimator. Phase two is the implementation of the SAGE algorithm that iteratively improves the initial estimate by employing both the training data and the soft information of the transmitted data. In terms of performance, the proposed system is largely affected by the number of data symbols used in the second phase. Benchmarked with existing pilot-assisted and data-aided channel estimators in the same premise, the proposed technique shows significant improvement at the cost of increased complexity. In high SNR cases and in pilot contamination regions, the proposed approach achieves convergence in almost one iteration. Without knowledge of large-scale fading coefficients of the interfering cells and without the need to employ extra pilots, the estimation accuracy is improved when employing the proposed technique.

Two semi-blind CE approaches for time-division duplex (TDD) massive MIMO systems are investigated in [58]. Firstly, an expectation-maximization (EM) algorithm that uses Gaussian distributions for unknown data symbols is derived. The semi-blind estimation approach based on the EM algorithm is contrasted with pilot-aided ML estimator and another ML estimator that assumes knowledge of the data symbols at the receiver. The performance criteria used for benchmarking include the MSE and the symbol error rate (SER). Secondly, an EM algorithm is derived based on *a priori* knowledge of the channel. The second approach is then evaluated against two CRB techniques. In the deterministic CRB, data symbols are assumed to be unknown deterministic values. The stochastic CRB on the other hand works on the assumption that data symbols are drawn from a Gaussian distribution.

Compared to the pilot aided ML estimator, the semi-blind CE scheme is shown to be more effective for both the uplink and the downlink scenarios in massive MIMO systems. The EM algorithm-based approach is shown to have similar performance with the genie aided (where data symbols are known) ML estimator when the number of BS antennas increases.

Furthermore, the boundless increase in the number of BS antennas expose a behaviour of CRBs that motivate for semi-blind CE for TDD massive MIMO applications.

2.5 Channel Estimation for Non-Orthogonal Multiple Access Systems

Literature has documented the importance of CE for OMA systems, NOMA based systems are of no exception. In literature, CE for NOMA based systems is largely unexploited. However, this excerpt provides a brief survey of the current literature pertaining to CE for NOMA systems. Channel estimation in the uplink NOMA scenario is presented and subsequently the accounts for the downlink scenario are addressed.

2.5.1 Channel Estimation for Uplink NOMA

Channel estimation for uplink NOMA can be referred to in the work presented in [55]. In the paper, the authors compare the performance of NOMA and interleave division multiple access (IDMA) employing a real estimator that exploits superimposed pilots. The cyclically repeated sequence superimposed pilots are assigned a given fraction of the total transmit power. These superimposed pilots are used in providing the initial channel estimate. Subsequently, a trellis-based channel estimator is introduced and incorporated in an iterative CE approach. A two-user system with a BS is assumed.

Channel estimation performance is considered to be enhanced with every iteration, as the superimposed pilots are reused. The influence of the trellis-based CE with superimposed pilots is assessed for both MA schemes. For performance evaluation of CE in each MA scheme, the behaviours of the BER and the MSE are investigated. Through simulations, the authors of the work presented in [59] highlight the robust CE performance for an IDMA system when the user power is balanced. However, this performance has a trade-off of high complexity. For NOMA, CE is effective when the user power allocation is imbalanced. Also, with NOMA, the trade-off between performance and complexity can be easily exploited to meet certain performance criteria.

Channel estimation for a code domain NOMA system variant, sparse code multiple access (SCMA), is considered in [60]. In this work, the authors modify a sparse Bayesian learning based CE algorithm and apply it to the SCMA uplink system. The modification is targeted to achieve better convergence rates, which are of importance in a situation of a large number of

active users. The proposed CE approach employs an iterative scheme for approximate Bayesian CE. The performance metric of interest in evaluation is the MSE and the user detection error rate. Through computer simulations, performance improvements are documented with the introduction of the proposed CE approach, when compared to its benchmark.

Seeking to improve the performance of the work presented in [61], the authors of the paper in [62] propose a disjoint user detection and CE approach for SCMA uplink systems. The main motivation for their proposition was the reduction of pilot overhead. Assuming a dispersive channel, the authors also propose a SCMA layer to physical resources mapping pattern with regards to the sparsity of SCMA codebooks. This measure is geared towards enhancing CE performance with limited pilot symbols.

With simulation results, the performance of the proposed channel estimator is presented in the MSE criteria. The proposed CE scheme is shown to reduce the pilot overhead by thirty percent as opposed to the joint CE and detection method proposed in [61]. Furthermore, the mapping pattern can significantly reduce the number of colliding users in each fading block, transitively improving the MSE performance. The performance of user detection is also presented in the paper [62], using the user detection error rate as a metric to benchmark it with the work presented in [61].

A novel joint CE and multi-user detection (CE) framework for uplink grant-free NOMA is proposed in the work presented in [63]. Considering the frame-wise joint sparsity of the pilot and data phases for a given frame, the authors formulate the multiple measurement vector-compressive sensing (MMV-CS) framework. The MMV-CS framework is then converted to a block-sparse single measurement vector-CS (BS-SMV-CS) model. A block sparsity adaptive subspace pursuit (BSASP) algorithm is then developed to derive benefit from the preferred block sparsity. For the BSASP algorithm to function, *a priori* information on the user sparsity level is not required. In terms of performance, the joint active user detection (AUD), CE and MUD framework is evaluated using the successful activity detection rate (SADR), normalized mean square error (NMSE) and the SER.

The proposed system is benchmarked with the conventional oracle LS approach, which assumes knowledge of actual active user support set at the BS. Another benchmark considered is the joint estimation and activity detection presented in [64]. In terms of the NMSE, the proposed approach is shown to outperform the technique presented in [64]. Compared to the oracle LS, the proposed algorithm is shown to approach the performance of the oracle LS

approach in high SNR values. Similar trends can be seen in detection performance, where the proposed technique shows better performance relative to the one in [64]. However, the oracle LS technique has better performance in this regard.

Another account of CE for an uplink SCMA system is given in [65]. In their paper, the authors proposed a CE method for frequency selective channel, which is set to use minimal training sequences. The pilot structure is designed based on the autocorrelation property of the Zadoff-Chu (ZC) sequences. In this scheme, users are overlapped in the same resource and pilot sequences are allocated to individual users. With the proposed CE technique, the channel information of each user is separated (in the autocorrelation domain) and the channel frequency response is estimated as per user basis.

In simulation, the proposed algorithm is evaluated against the unstructured channel estimator (UCE) [66] and the sparse pilot channel estimator (SPCE) [67]. Compared to these estimators that use orthogonality of the pilot sequence, the proposed method uses a windowing-based separation of channel information, which relatively minimizes the pilot overheads. In terms of MSE performance, the proposed system is shown to deliver similar performance to its counterparts with less pilot blocks employed for training.

2.5.2 Channel Estimation for Downlink (DL) NOMA

This subsection explores literature on CE for downlink NOMA. Even though the literature on DL NOMA is yet to be densely populated, the available accounts are presented.

In the papers [68 – 76], the importance of CE in the design of DL radio access is acknowledged. Even though the mentioned accounts are not necessarily concerned with CE as a principle, they incorporate the effects of estimation errors in their respective research interests. The research areas in these papers include system-level throughput and performance of NOMA for MIMO based systems and other environments. Other areas of focus include resource scheduling and optimal power allocation in NOMA based systems. For all the fore-mentioned accounts, imperfect channel state information (CSI) is assumed, and CE errors are included in respective analyses. This accommodation of CE errors in respective propositions for the DL NOMA technology can be viewed as a more practical approach. By virtue of the direction assumed in this dissertation, the work advanced in [68 – 76] will not be presented in detail.

Even more notable accounts of CE for DL NOMA can be referred to in the work presented in [23], [77 – 78]. The work presented in [23] is an account of joint CE and power allocation for a two-user single-input-single-output (SISO) NOMA DL system. In their work, the authors present a linear estimator that maximizes the average effective signal-to-interference-and-noise-ratio (SINR) of the strong user, while guaranteeing a bounded average effective SINR of the weak user. They further proposed a constrained concave convex procedure (CCCP)-based iterative algorithm as a solution for the CE and power allocation problem. The proposed solution is then compared to its LMMSE-based and ML-based estimator counterparts. For benchmarking, the performance metric used is the maximum average effective SINR. An outline of the work in [23] is given in the following passage.

Tang et. al consider a SISO NOMA downlink system with a base station (BS) and two users – $UE_i, i \in M = \{1,2\}$. The channel response from BS to user UE_i, denoted h_i is assumed to be a zero-mean circular symmetric complex Gaussian random variable with variance:

$$\sigma_{h_i}^2 = E\{|h_i|^2\}.$$

i.e. $h_i \backsim \mathcal{CN}(0, \sigma_{h_i}^2)$.

UE_1 is assumed to be the strong user and UE_2 is the weak user, thus the following relationship is established $\sigma_{h_1}^2 \geq \sigma_{h_2}^2$.

The signal sent out by the BS is given as:

$$x = \sqrt{\alpha_1 P} s_1 + \sqrt{\alpha_2 P} s_2, \tag{2.5.1}$$

where:

s_i is the signal intended for the i-th user satisfying $E\{|s_i|^2\} = 1$,

P is the total transmit power of the BS,

α_i is the power allocation ratio of the i-th user, $0 \leq \alpha_i \leq 1$ and $\sum_{i \in M} \alpha_i = 1$.

Subjected to modifications due to the channel response h_i relative to user i, and with addition of additive white Gaussian noise, the signal received at the i-th user is given by:

$$y_i = h_i \sum_{m=1}^{2} \sqrt{\alpha_m P}\, s_m + w_i, \tag{2.5.2}$$

where $w_i \backsim \mathcal{N}(0, \sigma_{w,i}^2)$ is the additive white Gaussian noise (AGWN) at the i-th user.

Pilot Symbols Transmission

Each user should estimate the channel response prior to signal detection during the CE phase. Initially, CE entails employment of pilot signal $\boldsymbol{x}_p$ of length N sent to both users. At each user i, the received training signal is given as:

$$\boldsymbol{z}_{i,p} = h_i \boldsymbol{x}_p + \boldsymbol{w}_{i,p}, \tag{2.5.3}$$

where $\boldsymbol{w}_{i,p} \backsim \mathcal{N}(0, \sigma_{w,i,p}^2 \boldsymbol{I})$ is the additive white Gaussian noise (AGWN) vector at the i-th user.

Each user employs CE methods based on $\boldsymbol{z}_{i,p}$ to estimate h_i

Proposed Channel Estimation Scheme for NOMA [23]

This sub-section showcases the CE process by first establishing the decoding of the received signal as per user and formulation of the linear channel estimator. The linear channel estimator is derived such that the SINR of the strong user is maximized in assurance of the minimum average effective SINR of the weak user. This is achieved using the constraint convex concave procedure (CCCP) based algorithm.

The following section entails the problem formulation to the CCCP algorithm. Consider a linear estimator as given below:

$$\hat{h}_i = \boldsymbol{v}_i \boldsymbol{z}_{i,p}, \tag{2.5.4}$$

where $\boldsymbol{v}_i$ is the unknown vector to be designed.

User Signal Reception and Decoding

When each user receives a signal, successive interference cancellation (SIC) plays a major role as a preferable method of signal decoding. In terms of signal decoding order, the signal that has the most interference (i.e. the signal of the weaker user) is decoded first. Thus, on reception of the signal, strong user UE_1 performs SIC before detecting its own signal. For SIC decoding to be realized, an estimation of the channel $\hat{h}_1$ has to be made, which is carried out in the subsequent section. SIC decoding is modelled as expressed below:

$$y_1^{(2)} = \hat{h}_1 \sqrt{\alpha_1 P} s_1 + (h_1 - \hat{h}_1) \sum_{m=1}^{2} \sqrt{\alpha_m P}\, s_m + \hat{h}_1 \sqrt{\alpha_2 P} s_2 + w_1, \tag{2.5.5}$$

where the second term and the third term are the CE error interferences and other-user interferences respectively.

In performing SIC, UE_1 first decodes signal s_2 and successively removes the component containing s_2 from the equation (2.5.5) above, resulting in:

$$y_1^{(1)} = \hat{h}_1\sqrt{\alpha_1 P}s_1 + \left(h_1 - \hat{h}_1\right)\sum_{m=1}^{2}\sqrt{\alpha_m P}\, s_m + w_1, \tag{2.5.6}$$

with which the signal for user UE_1 is decoded.

At weak user UE_2, the signal is detected using channel estimate $\hat{h}_2$ treating the other signal parameters or components as noise. The used signal model is as follows:

$$y_2^{(2)} = \hat{h}_2\sqrt{\alpha_2 P}s_2 + \left(h_2 - \hat{h}_2\right)\sum_{m=1}^{2}\sqrt{\alpha_m P}\, s_m + \hat{h}_2\sqrt{\alpha_1 P}s_1 + w_2. \tag{2.5.7}$$

Average Effective Signal-to-Interference- and-Noise-Ratio:

The average effective SINR for the strong user to detect the signal of the weak user as derived from (2.5.5) is given as:

$$\gamma_1^{(2)} = \frac{(1-\alpha_1)\boldsymbol{v}_1^H\boldsymbol{R}_1\boldsymbol{v}_1}{(1+\alpha_1)\boldsymbol{v}_1^H\boldsymbol{R}_1\boldsymbol{v}_1 - \sigma_{h_1}^2\left(\boldsymbol{v}_1^H\boldsymbol{x}_p + \boldsymbol{x}_p^H\boldsymbol{v}_1\right) + \eta_1}, \tag{2.5.8}$$

where $\boldsymbol{R}_1 = \sigma_{h_1}^2\boldsymbol{x}_p\boldsymbol{x}_p^H + \sigma^2\boldsymbol{I}$ is the covariance matrix of the received pilot signal and $\eta_1 = \sigma_{h_1}^2 + \sigma_{w,1}^2/P$.

Based on (2.5.6), the average effective SINR of the strong user decoding its own signal is obtained as:

$$\gamma_1^{(1)} = \frac{\alpha_1\boldsymbol{v}_1^H\boldsymbol{R}_1\boldsymbol{v}_1}{\boldsymbol{v}_1^H\boldsymbol{R}_1\boldsymbol{v}_1 - \sigma_{h_1}^2\left(\boldsymbol{v}_1^H\boldsymbol{x}_p + \boldsymbol{x}_p^H\boldsymbol{v}_1\right) + \eta_1}. \tag{2.5.9}$$

Unlike the strong user, which has to perform SIC to detect its own signal, the weak user detects its own signal treating the strong user's signal as interference. The average effective SINR for the weak user detecting its own signal is obtained as:

$$\gamma_2^{(2)} = \frac{(1-\alpha_1)\boldsymbol{v}_2^H\boldsymbol{R}_2\boldsymbol{v}_2}{(1+\alpha_1)\boldsymbol{v}_2^H\boldsymbol{R}_2\boldsymbol{v}_2 - \sigma_{h_2}^2\left(\boldsymbol{v}_2^H\boldsymbol{x}_p + \boldsymbol{x}_p^H\boldsymbol{v}_2\right) + \eta_2}. \tag{2.5.10}$$

where $\boldsymbol{R}_2 = \sigma_{h_2}^2 \boldsymbol{x}_p \boldsymbol{x}_p^H + \sigma_{w,2}^2 \boldsymbol{I}$ is the covariance matrix of the pilot symbol received at the weak user and $\eta_2 = \sigma_{h_2}^2 + \sigma_{w,2}^2/P$.

The proposed scheme is aimed at maximizing the SINR of the strong user while assuring that the minimum average effective SINR of the weak user is not less than a predefined threshold, γ_0. Below is the optimization problem:

$$\max_{\{0\le\alpha_1\le 1\},\{\boldsymbol{v}_i\}} \gamma_1^{(1)} \ \ s.t. \min_{i\in\{1,2\}} \left\{\gamma_i^{(2)}\right\} \ge \gamma_0. \tag{2.5.11}$$

CCCP-Based Iterative Algorithm

The CCCP-Based Iterative Algorithm is proposed as a solution to (2.5.11). Equation (2.5.11) can be recast as:

$$\max_{\{0\le\alpha_1\le 1\},\{\boldsymbol{v}_i\}} \gamma_1^{(1)} \ \ s.t. \ \gamma_i^{(2)} \ge \gamma_0 \ \ \forall i \in \{1,2\}. \tag{2.5.12}$$

Letting $\mu_1 = {}^1/_{\alpha_1}$, and introducing the slack variable τ, problem (2.5.12) can be equivalently rewritten and simplified to the following:

$$\min_{\{\mu_1\ge 1\},\{\boldsymbol{v}_i\},\tau} -\ln\tau + \ln\mu_1 \tag{2.5.13a}$$

$$s.t. \ \boldsymbol{v}_1^H \boldsymbol{R}_1 \boldsymbol{v}_1 - \sigma_{h_1}^2 \left(\boldsymbol{v}_1^H \boldsymbol{x}_p + \boldsymbol{x}_p^H \boldsymbol{v}_1\right) + \eta_1 - \frac{\boldsymbol{v}_1^H \boldsymbol{R}_1 \boldsymbol{v}_1}{\tau} \le 0 \tag{2.5.13b}$$

$$\boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i + \left(1 + \frac{1}{\gamma_0}\right)\frac{\boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i}{\mu_1} - \sigma_{h_i}^2 \left(\boldsymbol{v}_i^H \boldsymbol{x}_p + \boldsymbol{x}_p^H \boldsymbol{v}_i\right)$$

$$+\eta_i - \frac{1}{\gamma_0} \boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i \le 0. \tag{2.5.13c}$$

As $-\ln\mu_1$, $\frac{\boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i}{\mu_1}$ and $\boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i$ where $\mu_1 > 0, \boldsymbol{R}_i \ge 0$ are convex, the problem is a difference of convex (dc) programming and can therefore be solved using the CCCP algorithm. By Letting

$$\xi(\boldsymbol{v}_1, \tau) = \frac{\boldsymbol{v}_1^H \boldsymbol{R}_1 \boldsymbol{v}_1}{\tau},$$

$$\varsigma_i(\boldsymbol{v}_i) = \boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i, \quad i \in \{1,2\},$$

$$\rho(\mu_1) = -\ln\mu_1.$$

The first-order Taylor expansions of the fore-listed equations around the point $(\tilde{\boldsymbol{v}}_i, \tilde{\tau}, \tilde{\mu}_1)$ are computed as:

$$\xi(\boldsymbol{v}_1, \tau, \tilde{\boldsymbol{v}}_1, \tilde{\tau},) = \frac{2Re\{\tilde{\boldsymbol{v}}_1^H \boldsymbol{R}_1 \boldsymbol{v}_1\}}{\tilde{\tau}} - \frac{\tilde{\boldsymbol{v}}_1^H \boldsymbol{R}_1 \tilde{\boldsymbol{v}}_1}{\tilde{\tau}^2}\tau, \tag{2.5.14}$$

$$\varsigma_i(\boldsymbol{v}_i, \tilde{\boldsymbol{v}}_1) = 2Re\{\tilde{\boldsymbol{v}}_i^H \boldsymbol{R}_i \boldsymbol{v}_i\} - \tilde{\boldsymbol{v}}_i^H \boldsymbol{R}_i \tilde{\boldsymbol{v}}_i, \quad i \in \{1,2\}, \tag{2.5.15}$$

$$\rho(\mu_1, \tilde{\mu}_1) = -\ln \tilde{\mu}_1 - \frac{(\mu_1 - \tilde{\mu}_1)}{\tilde{\mu}_1}. \tag{2.5.16}$$

For the $(l+1)$-th iteration of the proposed CCCP-based iterative algorithm, the following optimization problem is solved:

$$\begin{aligned} \min_{\{\mu_1 \geq 1\}, \{v_i\}, \tau} \; & -\ln \tau - \rho\left(\mu_1, \tilde{\mu}_1{}^{(l)}\right) \\ s.t. \; & \boldsymbol{v}_1^H \boldsymbol{R}_1 \boldsymbol{v}_1 - \delta_1^2\left(\boldsymbol{v}_1^H \boldsymbol{x}_p + \boldsymbol{x}_{1,p}^H \boldsymbol{v}_1\right) + \eta_1 - \xi\left(\boldsymbol{v}_1, \tau, \tilde{\boldsymbol{v}}_1{}^{(l)}, \tilde{\tau}^{(l)},\right) \leq 0 \\ & \boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i + \left(1 + \frac{1}{\gamma_0}\right)\frac{\boldsymbol{v}_i^H \boldsymbol{R}_i \boldsymbol{v}_i}{\mu_1} - \delta_1^2\left(\boldsymbol{v}_1^H \boldsymbol{x}_p + \boldsymbol{x}_{1,p}^H \boldsymbol{v}_1\right) \\ & + \eta_i - \frac{1}{\gamma_0}\,\varsigma_i\left(\boldsymbol{v}_i, \tilde{\boldsymbol{v}}_1{}^{(l)}\right) \leq 0 \quad \forall i \in \{1,2\}. \end{aligned} \tag{2.5.17}$$

where the point $\left(\tilde{\boldsymbol{v}}_i{}^{(l)}, \tilde{\boldsymbol{\tau}}^{(l)}, \tilde{\boldsymbol{\mu}}_1{}^{(l)}\right)$, denotes the solution to problem (2.5.17) at the *l*th iteration. The CCCP-Based iterative algorithm for CE is summarized in Algorithm 2.2, with initial point assumed as $\left(\tilde{\boldsymbol{v}}_i{}^{(0)}, \tilde{\boldsymbol{\tau}}^{(0)}, \tilde{\boldsymbol{\mu}}_1{}^{(0)}\right)$:

Algorithm 2.2 CCCP-Based Iterative Algorithm
1: Initialization: $l = 0$, $\tilde{v}_i{}^{(0)}$, $\tilde{\tau}^{(0)}$, $\tilde{\mu}_1{}^{(0)}, i \in \boldsymbol{M}$;
2: Repeat:
Solve problem (2.5.17) to obtain $\tilde{v}_i{}^{(l+1)}$, $\tilde{\tau}^{(l+1)}$, $\tilde{\mu}_1{}^{(l+1)}, i \in \boldsymbol{M}$;
$l := l + 1$;
3. Until: Convergence.

Maximum Average Effective SINR Under ML and LMMSE Channel Estimation

ML and LMMSE channel estimators are of the linear form, as $\boldsymbol{v}_i^{(k)}\ for\ k \in \{ML, LMMSE\}$ and they are given by:

$$\boldsymbol{v}_i^{(k)} = \begin{cases} \dfrac{\boldsymbol{x}_p}{\|\boldsymbol{x}_p\|^2}, & k = ML \\ \dfrac{\sigma_{h_i}^2 \boldsymbol{x}_p}{\sigma_{w,i}^2 + \sigma_{h_i}^2 \|\boldsymbol{x}_p\|^2}, & k = LMMSE \end{cases}. \tag{2.5.18}$$

The maximum average effective SINR of the strong users under ML and LMMSE CE is given threshold γ_0 is expressed as [23]:

$$\gamma_1^{(k)} = \frac{\left(\epsilon_2^{(k)} - \gamma_0 b_2^{(k)}\right)\epsilon_1^{(k)}}{(1 + \gamma_0) b_1^{(k)} \epsilon_2^{(k)}}, \tag{2.5.19}$$

where:

$$\epsilon_i^{(k)} = \begin{cases} \sigma_{h_i}^2 + \dfrac{\sigma_{w,i}^2}{\|\boldsymbol{x}_p\|^2}, & k = ML \\ \dfrac{\sigma_{h_i}^4 \|\boldsymbol{x}_p\|^2}{\sigma_{w,i}^2 + \sigma_{h_i}^2 \|\boldsymbol{x}_p\|^2}, & k = LMMSE \end{cases},$$

$$b_i^{(k)} = \begin{cases} \dfrac{\sigma_{w,i}^2}{\|\boldsymbol{x}_p\|^2} + \dfrac{\sigma_{w,i}^2}{P}, & k = ML \\ \sigma_{h_i}^2 + \dfrac{\sigma_{w,i}^2}{P} + \dfrac{\sigma_{h_i}^4 \|\boldsymbol{x}_p\|^2}{\sigma_{w,i}^2 + \sigma_{h_i}^2 \|\boldsymbol{x}_p\|^2}, & k = LMMSE \end{cases}.$$

From simulation results, the authors of [23] show significant improvements introduced by the proposed scheme as opposed to LMMSE and ML, in terms of the average effective SINR of the strong user in different transmission power levels. Efficient allocation of transmission power is also effected with the proposed method as opposed to the evaluated traditional methods, revealing the necessity of joint CE and power allocation.

A novel CE and transmission strategy for millimetre wave (mmWave) NOMA communication system with hybrid architecture is presented in [77]. This work basically encompasses CE, hybrid precoding, user scheduling and power allocation for the considered system. An iterative

index detection-based CE algorithm (IDCEA) is proposed for mmWave NOMA systems, in order to obtain both the direction of arrivals (DOAs) and the channel coefficient of each channel path. Following this, an enhanced hybrid precoding scheme is developed to reduce the inter-beam interference using the recently obtained DOA information. A sum rate maximization problem is then formulated subject to joint user scheduling and power allocation strategy. This problem is identified to be non-convex. In order to solve it, the original optimization is dissociated into two sub-problems: one concerned with user scheduling and the other concerned with power allocation.

User scheduling is addressed by the development of the matching theory-based user scheduling algorithm (MTBUSA) in order to maximize the achievable sum rate. The realization of the user scheduling solution is aided by the usage of DOAs of different users to design a heuristic initialization user scheduling algorithm. For dynamic power allocation, an iterative optimization algorithm is designed, governed by the transmit power constraint. Simulation results are presented to evaluate the proposed approach, using the mean square error (MSE) criteria to evaluate the CE performance.

The proposed estimation approach is benchmarked with the maximum likelihood (ML) theoretical bound and Cramer-Rao bound (CRB) presented in [79], and the initial estimation and the angle rotation-based estimation presented in [80]. The proposed CE approach is shown to outperform the initial estimation and angle rotation-based algorithms for the entire SNR range considered. On the other hand, the CRB and ML estimators pose better performance in comparison with the proposed technique. With increase in BS antennas, each estimation method is shown to improve in estimation accuracy.

The MSE performance of the proposed system is also compared with the compressive-sensing (CS) estimation method proposed in [81], the spatial basis expansion model method (SBEM) [82] and the eigen-decomposition based method [83]. The proposed method is shown to have better performance in contrast with the CS-based and SBEM CE methods. However, the eigen-decomposition based method out-performs the proposed method. The setback with the eigen-decomposition based method was highlighted to be the impracticality of obtaining the accurate channel covariance matrix. With the presented advantages of the proposed CE technique over some conventional CE methods, the proposed CE can be deemed preferable. That and the fact that it is used in conjunction with the MTBUSA and power allocation schemes that show improved performance relative to the conventional mmWave beam-space MIMO system.

An enhanced NOMA system using adaptive coding and modulation is considered in the work presented in [78]. For this NOMA system, an effective CE algorithm based on long-short term memory neural network (LSTM-NN) is proposed. The added advantage of the LSTM-NN based CE algorithm is the dynamic adaptability to the behaviour of the fluctuating channel conditions. The Pascal's triangle is then used as the basis to develop a novel power allocation algorithm, which basically allocates power to different users based on the channel conditions attributed to the respective user. The authors of [78] also introduce an adaptive coding scheme based on Bose-Chaudhuri-Hocquenghem (BCH), and an adaptive modulation scheme based on constellation rotation and cyclic Q-delay. Holistically, the proposed system is evaluated in terms of the bit error rate (BER), outage probability and sum rates under varied conditions.

In terms of performance evaluation, employing the LSTM-NN based CE approach for the conventional DL NOMA introduced a 10% average reduction in the overall outage probability, and a 50% percent maximum reduction in the BER in probability. The average sum rate increase was reported to be 37%. In addition to the performance improvements attributed to using the LSTM-NN based CE approach, performance increase was reported with the employment of the proposed adaptive modulation and coding schemes.

2.6 Chapter Summary

In this chapter, a brief survey of NOMA has been presented, with interest directed towards the power domain variant. Assuming a general standpoint, the wireless channel is also studied showing the importance of CE in wireless communication systems. Furthermore, the importance of channel modelling in wireless communication systems is highlighted. Various accounts of CE techniques in literature are then studied for different underlying systems.

In surveying the different CE techniques, four categories were identified: pilot-aided methods, decision-directed methods, blind methods and semi-blind methods. Each category was deemed more suitable than the other depending on the communication system of application. Pilot-aided CE methods are regarded to pose low complexity and are more suitable for slow fading channels. Blind CE methods are were mainly motivated by reduction of pilot overhead. The semi-blind and decision-directed CE mainly strike a balance between the complexity and spectral efficiency trade-off. A survey on the new paradigm of CE based on deep learning principles is also effected.

The chapter is then advanced by a comprehensive survey on CE techniques documented for NOMA applications. With both the uplink and downlink scenarios explored, the apparent shortage of literature in CE for downlink NOMA systems is recognized.

CHAPTER 3

SYSTEM MODEL DESCRIPTION AND THE ALL-KNOWN-PILOTS CHANNEL ESTIMATION APPROACH

3.1 Introduction

This chapter provides detailed information about the system model adopted for the proposed channel estimation (CE) approaches in this dissertation. A two-user downlink (DL) power domain single-input-single-output non-orthogonal multiple access system (SISO-NOMA) is addressed. NOMA systems were essentially developed for accommodation of a large number of users within a resource block as measure for spectral efficiency. However, in this dissertation, a simplified two-user scenario is assumed. The reason for this approach is to shift the focus to developing efficient CE approaches to fill the dent that currently exist for CE in DL NOMA systems. By using this scenario, the proposed CE approaches can therefore be extended to models that entail multiplicity in terms of user numbers and antennas.

Three least mean squares (LMS) algorithms adopted from the work presented in [22] are extended to the NOMA model in this chapter. Unlike the work in [22] which assumes an OMA system, the algorithms are adapted to a NOMA system. For the CE process, a pilot scenario whereby both pilots within the superimposed signal are known to both users is considered. The constant step-size LMS (CSS-LMS) algorithm is presented in Section 3.3. For the CSS-LMS algorithm, an analytical expression is provided for the optimal step-size and the error variance.

Section 3.4 presents the variable step-size LMS (VSS-LMS) algorithm. For the third algorithm, adaptive speed is added to the second algorithm, giving rise to the variable step-size LMS algorithm with adaptive speed (VSS-LMS-AS). The VSS-LMS-AS algorithm is presented in Section 3.5. The performance of the three algorithms is assessed through simulations and the results are presented with their complementary evaluations in Section 3.6. Finally, chapter three is summarized in Section 3.7.

Overall, the all-known-pilots CE scheme posed pivotal performance with the employment of the CSS-LMS algorithm. The VSS-LMS-AS algorithm on the other hand has superior performance over the VSS-LMS algorithm in terms of CE speed. Insignificant CE performance

differences were observed when evaluating the CE performance at each user. Simulation of the CSS-LMS algorithm was shown to confirm the theoretical analysis.

3.2 System Model Description

In this system, we adopt a downlink system model that can be referred to in the work presented in [23], which is documented in Section 2.5.2. However, a few modifications are introduced. Assuming the base-station (BS) and two-users scenario as given in Section 2.5.2 from the work documented in [23], the pilot signal model scenario is fashioned to resemble the information data model. Furthermore, for estimation, the pilot symbols are considered on a sample basis instead as a sequence of a given length *N*. And finally, the work presented herein focuses on CE, instead of CE and power allocation as with the work in [23]. Thus, the model for the superimposed signal of pilot symbols transmitted from the BS are expressed as:

$$x_n = \sqrt{\alpha_1 P} s_{1,n} + \sqrt{\alpha_2 P} s_{2,n}, \tag{3.2.1}$$

where $s_{1,n}$ and $s_{2,n}$ are pilot symbols intended for user 1 and user 2 respectively, n denotes the symbol index and $E\{|x_n|^2\} = P$.

The observation of the pilot signal at the receiver of the i-th user is given by:

$$y_{i,n} = h_i \sum_{m=1}^{2} \sqrt{\alpha_m P}\, s_{m,n} + w_{i,n}, \tag{3.2.2}$$

where $w_{i,n} \backsim \mathcal{N}(0, \sigma_{w_i}^2)$ is the additive white Gaussian noise (AGWN) at the *i*th user.

The channel coefficient h_i is modelled as a Rayleigh flat fading channel with zero mean and variance $\sigma_{h_i}^2$. Each channel assumes Jake's Doppler spectrum, of which the one defined in subsection 2.4.2.2 can be modified to suit the NOMA model as follows:

$$\Gamma_{h_i}(f) = \begin{cases} \dfrac{\sigma_{h_i}^2}{\pi f_d \sqrt{1 - \left(\dfrac{f}{f_d}\right)^2}} & if\ |f| < f_d \\ 0 & if\ |f| \geq f_d \end{cases}. \tag{3.2.3}$$

where:

$f = f_d \cos\varphi$ is the frequency of arriving signal,

f_d is the maximum doppler frequency and φ is the phase of arrival uniformly distributed in $[0\ 2\pi)$,

and $\sigma_{h_i}^2$ is the variance of the channel coefficients.

In the subsequent chapters that cover the three CE approaches, the estimation error will be considered. This is specifically for performance evaluation purposes. The estimation error is generally given by the following expression:

$$e_{i,n} = h_{i,n} - \hat{h}_{i,n}, \tag{3.2.4}$$

where *i* represents the user index and *n* is the symbol index.

The corresponding error variance is given by the following expression:

$$\sigma_{e_i}^2 = E\left\{\left|e_{i,n}\right|^2\right\}. \tag{3.2.5}$$

3.3 Constant Step-Size Least Mean Square (CSS-LMS) Algorithm (Adapted from [22])

This sub-section presents the CSS-LMS algorithm for DL SISO-NOMA CE assuming an all-known-pilots scenario. With the CSS-LMS algorithm, an estimate $\hat{h}_i$ is provided for a Rayleigh flat fading channel. For simplification of the estimation problem, the pilots intended for each user are considered as $s_{1,n} = s_{2,n} = 1$. The superimposed pilot signal sent from the base station (BS) to both users given in (3.2.1) is therefore expressed as:

$$x = \sqrt{\alpha_1 P} + \sqrt{\alpha_2 P}, \tag{3.3.1}$$

where $E\{|x|^2\} = P$

Considering the pilot scenario sets the stage for the construction of the cost function employed in finding the LMS estimator. Now given the model of the pilot signal, the adopted cost function is expressed as:

$$V_1(\hat{h}_i) = E\left\{\left|y_{i,n} - \hat{h}_i x\right|^2\right\}. \tag{3.3.2}$$

The LMS estimator is based on the stochastic gradient adaptation, thus the update rule is formed with the usage of the gradient descent of the cost function in terms of the estimate. In general terms, the update rule of the LMS algorithm is of the form:

$$\hat{h}_{i,n} = \hat{h}_{i,n-1} - \mu \frac{\partial V_1(\hat{h}_i)}{\partial \hat{h}_i}. \tag{3.3.3}$$

The result of differentiating the cost function given in (3.3.2) in terms of the channel estimate is given as:

$$\frac{\partial V_1(\hat{h}_i)}{\partial \hat{h}_i} = -2x(y_{i,n} - \hat{h}_{i,n-1}x). \tag{3.3.4}$$

Substituting the gradient in equation (3.3.4) into (3.3.3), the LMS algorithm is therefore defined by the following update rule:

$$\hat{h}_{i,n} = \hat{h}_{i,n-1} - \mu x(y_{i,n} - \hat{h}_{i,n-1}x), \tag{3.3.5}$$

where μ is the constant step-size such that $\mu < 1/SNR$,

i and n are the user and symbol indices respectively.

The CSS-LMS algorithm is, in essence, described by the expression given in (3.3.5). The asymptotic mean square error (MSE) analysis is based on this setting.

3.3.1 Error Variance of the CSS-LMS Algorithm

The CSS-LMS asymptotic MSE (error variance) analysis is presented in this segment. From the update rule given by (3.3.5), a first-order time-invariant filter can be obtained. The expression of which is found in the z-domain as:

$$L(z) = \frac{\hat{h}_i(z)}{y_i(z)}. \tag{3.3.6}$$

Taking into account the property $E\{|x|^2\} = P$, (3.3.5) is given in the z-domain, as:

$$\hat{h}_i(z) = \hat{h}_i(z)z^{-1} + \mu x(y_i(z) - \hat{h}_i(z)z^{-1}x),$$

$$\hat{h}_i(z)(1 - z^{-1} + \mu P z^{-1}) = \mu x y_i(z).$$

Thus, the low-pass filter is expressed as:

$$L(z) = \frac{\mu x}{1 - (1 - \mu P)z^{-1}}. \tag{3.3.7}$$

Given the observation expressed in (3.2.2), the error signal described in (3.2.4) and the low-pass filter in (3.3.6), the CE error is therefore expressed in the z-domain as:

$$\begin{aligned} e_i(z) &= h_i(z) - L(z)\big(h_i(z)x + w_i(z)\big) \\ &= (1 - L(z)x)h_i(z) - L(z)w_i(z). \end{aligned} \tag{3.3.8}$$

Now, using the definition of the asymptotic MSE given in (3.2.5) and the CE error given in (3.3.8), the asymptotic is derived as an expression given below:

$$\sigma_{e_i}^2 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - L\left(\boldsymbol{e}^{2j\pi fT}\right)x\right|^2 \Gamma_f(f)\, df + \sigma_{w_i}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df, \tag{3.3.9}$$

where T is the sampling duration.

The explicit terms of the overall error variance expressions can be evaluated independently. As such, new variables are defined for each term in the error variance. These are given below:

$$MSE1 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - L\left(\boldsymbol{e}^{2j\pi fT}\right)x\right|^2 \Gamma_f(f)\, df. \tag{3.3.10}$$

$$MSE2 = \sigma_{w_i}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df. \tag{3.3.11}$$

Firstly, the focus is with the derivation of the expression of $MSE1$. To realize this, an approximation of this term is made by assuming low frequencies relative to the sampling rate $f << 1/T$. In addition, the cut-off frequency of the filter is chosen to be higher than the maximum Doppler frequency and sufficiently lower than the sampling rate such that, $f_d T < f_c T << 1$ [84]. With the proposed design of the low-pass filter, it should be noted that the cut-off frequency is selected as $f_c T = \frac{\mu P}{2\pi(1-\mu P)}$. With this setting, z^{-1} is then approximated as the first two terms of its Taylor series. That is, $z^{-1} = \boldsymbol{e}^{2j\pi fT} \approx 1 - j2\pi fT$. With this approximation, we then have:

$$\begin{aligned} L\left(\boldsymbol{e}^{2j\pi fT}\right)x &= \frac{\mu P}{1 - (1 - j2\pi fT)(1 - \mu P)} \\ &= \frac{\mu P}{1 - (1 - \mu P) + (1 - \mu P) j2\pi fT} \\ &= \frac{\mu P}{(1 - \mu P)\left[\frac{\mu P}{(1 - \mu P)} + j2\pi fT\right]}. \end{aligned} \tag{3.3.12}$$

Substituting the cut-off frequency in the expression given in equation (3.3.12) results in:

$$L\left(\boldsymbol{e}^{2j\pi fT}\right)x = \frac{2\pi f_c T}{2\pi f_c T + j2\pi fT}. \tag{3.3.13}$$

So, we can then have,

$$1 - xL\left(\boldsymbol{e}^{2j\pi fT}\right) = \frac{2\pi f_c T - 2\pi f_c T + j2\pi fT}{2\pi f_c T + j2\pi fT}$$

$$= \frac{j2\pi fT}{2\pi f_c T + j2\pi fT}. \tag{3.3.14}$$

And,

$$\left|1 - xL\left(e^{2j\pi fT}\right)\right|^2 = \left|\frac{j2\pi fT}{2\pi f_c T + j2\pi fT}\right|^2$$

$$= \left(\frac{\sqrt[2]{(2\pi fT)^2}}{\sqrt[2]{(2\pi f_c T)^2 + (2\pi fT)^2}}\right)^2$$

$$= \frac{(2\pi fT)^2}{(2\pi f_c T)^2 + (2\pi fT)^2}. \tag{3.3.15}$$

Assuming that the cut-off frequency is adjusted such that $2\pi f_d T << \mu << 1$ and $f_d T < f_c T < < 1$, equation (3.3.15) can be approximated:

$$\left|1 - xL\left(e^{2j\pi fT}\right)\right|^2 = \left(\frac{2\pi fT}{2\pi f_c T}\right)^2. \tag{3.3.16}$$

Using the power spectral density of the channel coefficients in (3.2.3) and (3.3.16), $MSE1$ can therefore be approximated as shown below:

$$MSE1 \approx \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left(\frac{2\pi fT}{2\pi f_c T}\right)^2 \frac{\sigma_{h_i}^2}{\pi f_d \sqrt{1 - \left(\frac{f}{f_d}\right)^2}}\, df$$

$$\approx \int_{-f_d}^{f_d} \left(\frac{2\pi fT}{2\pi f_c T}\right)^2 \frac{\sigma_{h_i}^2}{\pi f_d \sqrt{1 - \left(\frac{f}{f_d}\right)^2}}\, df$$

$$\approx 2\int_{0}^{f_d} \left(\frac{2\pi fT}{2\pi f_c T}\right)^2 \frac{\sigma_{h_i}^2}{\pi\sqrt{f_d^2 - f^2}}\, df. \tag{3.3.17}$$

Thereafter, a variable change can be implemented using $f = f_d \cos\varphi$, as given in the channel model. This essentially introduces the following limits:

$\varphi = \cos^{-1}\left(\frac{0}{f_d}\right) = \frac{\pi}{2}$, for the lower limit and $\varphi = \cos^{-1}\left(\frac{f_d}{f_d}\right) = 0$ for the upper limit.

The approximation of the first term therefore becomes:

$$MSE1 \approx \frac{2\sigma_{h_i}^2(2\pi f_d T)^2}{\pi(2\pi f_c T)^2}\int_{\frac{\pi}{2}}^{0}\frac{\cos^2\varphi}{\sqrt{f_d^2(1-\cos^{-2}\varphi)}} - f_d \sin\varphi \, d\varphi$$

$$\approx \frac{2\sigma_{h_i}^2(2\pi f_d T)^2}{\pi(2\pi f_c T)^2}\int_{0}^{\frac{\pi}{2}}\cos^2\varphi \, d\varphi$$

$$\approx \frac{2\sigma_{h_i}^2(2\pi f_d T)^2}{\pi(2\pi f_c T)^2}\left[\frac{\varphi}{2} + \frac{\sin 2\varphi}{2}\right]_0^{\frac{\pi}{2}}$$

$$\approx \frac{2\sigma_{h_i}^2(2\pi f_d T)^2}{\pi(2\pi f_c T)^2}\frac{\pi}{4}.$$

By substituting the expression for $2\pi f_c T$, the solution for $MSE1$ then becomes:

$$MSE1 = \sigma_{h_i}^2\frac{(2\pi f_d T)^2}{2(\mu P)^2}(1-\mu P)^2. \qquad (3.3.18)$$

In order to derive $MSE2$, the focus is firstly directed to simplifying the expression $\left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2$ as:

$$\left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 = \left|\frac{\mu x}{1-e^{-2j\pi fT}(1-\mu P)}\right|^2$$

$$= \mu^2 P\left|\frac{e^{2j\pi fT}}{e^{2j\pi fT}-(1-\mu P)}\right|^2$$

$$= \mu^2 P\left|\frac{(\cos 2\pi fT + j\sin 2\pi fT)}{\cos 2\pi fT - (1-\mu P) + j\sin 2\pi fT}\right|^2$$

$$= \left(\frac{\mu^2 P}{(\cos 2\pi fT)^2 - 2(1-\mu P)\cos 2\pi fT + (1-\mu P)^2 + (\sin 2\pi fT)^2}\right)$$

$$= \left(\frac{\mu^2 P}{1+(1-\mu P)^2 - 2(1-\mu P)\cos 2\pi fT}\right)$$

$$= \left(\frac{\mu^2 P}{(1+(1-\mu P)^2)\left[1 - \frac{2(1-\mu P)}{1+(1-\mu P)^2}\cos 2\pi f T\right]} \right)$$

$$= \left(\frac{B}{1 + C\cos 2\pi f T}\right). \tag{3.3.19}$$

where,

$B = \frac{\mu^2 P}{(1+(1-\mu P)^2)}$ and $C = \frac{-2(1-\mu P)}{1+(1-\mu P)^2}$.

The simplified expression in equation (3.3.19) can therefore be substituted into the integral of $MSE1$:

$$MSE2 = \sigma_{w_i}^2 TB \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left(\frac{1}{1 + C\cos 2\pi f T}\right) df$$

$$= 2\sigma_{w_i}^2 TB \int_{0}^{\frac{1}{2T}} \left(\frac{1}{1 + C\cos 2\pi f T}\right) df. \tag{3.3.20}$$

A change of variables is introduced by letting $v = 2\pi f T$, then $dv = 2\pi T df$ and the limits $\left[0 \ \frac{1}{2T}\right]$ are translated to $\left[2\pi(0)T \ \ 2\pi\left(\frac{1}{2T}\right)T\right]$, which ultimately becomes $[0 \ \pi]$. Using the solution given in the work presented in [85] (3.613), $\int_0^{\pi} \left(\frac{1}{1+C\cos v}\right) dv$ can similarly be computed as:

$$\frac{1}{2\pi T}\int_0^{\pi} \left(\frac{1}{1 + C\cos v}\right) dv = \frac{1}{2\pi T}\frac{\pi}{\sqrt{1-C^2}}.$$

by assuming the limit $|C| < 1$. To verify this, it can be observed that:

$$1 + (1-\mu P)^2 - 2(1-\mu P) = (\mu P)^2 > 0.$$

As such, $1 + (1-\mu P)^2 > 2(1-\mu P)$ thus C is a fraction and the limit $|C| < 1$ persists.

With this setting, the solution to the equation in (3.3.20) is therefore given as

$$MSE2 = 2\sigma_{w_i}^2 TB \frac{1}{2\pi T}\frac{\pi}{\sqrt{1-C^2}}, \tag{3.3.21}$$

Substituting the expressions of B and C, the following is undertaken to simplify and ultimately find the final expression:

$$MSE2 = 2\sigma_{w_i}^2 T \frac{\mu^2 P}{1+(1-\mu P)^2} \frac{1}{2\pi T} \frac{\pi}{\sqrt{1-\left(\frac{-2(1-\mu P)}{1+(1-\mu P)^2}\right)^2}}$$

$$= 2\sigma_{w_i}^2 T \frac{\mu^2 P}{1+(1-\mu P)^2} \frac{1}{2\pi T} \frac{\pi}{\sqrt{\frac{1+2(1-\mu P)^2+(1-\mu P)^4-4(1-\mu P)^2}{(1+(1-\mu P)^2)^2}}}$$

$$= 2\sigma_{w_i}^2 T \frac{\mu^2 P}{1+(1-\mu P)^2} \frac{1}{2\pi T} \frac{\pi}{\sqrt{\frac{1-2(1-\mu P)^2+(1-\mu P)^4}{(1+(1-\mu P)^2)^2}}}$$

$$= 2\sigma_{w_i}^2 T \frac{\mu^2 P}{1+(1-\mu P)^2} \frac{1}{2\pi T} \frac{\pi(1+(1-\mu P)^2)}{1-(1-\mu P)^2}$$

$$= \sigma_{w_i}^2 \frac{\mu^2 P}{1-1+2\mu P-(\mu P)^2}$$

$$= \sigma_{w_i}^2 \frac{\mu^2 P}{\mu P(2-\mu P)}. \quad (3.3.22)$$

Therefore, the final expression of $MSE2$ is given as:

$$MSE2 = \sigma_{w_i}^2 \frac{\mu}{(2-\mu P)}. \quad (3.3.23)$$

Using the expressions given in (3.3.18) and (3.3.23), the error variance in (3.3.9) therefore becomes:

$$\sigma_{e_i}^2 = \sigma_{h_i}^2 \frac{(2\pi f_d T)^2}{2(\mu x^2)^2}(1-\mu P)^2 + \sigma_{w_i}^2 \frac{\mu}{(2-\mu P)}. \quad (3.3.24)$$

To find the optimal step-size, and transitively the optimal asymptotic MSE, the error variance in equation (3.3.24) is minimized as follows:

$$\min_{\mu} \sigma_{e_i}^2. \quad (3.3.25)$$

To achieve this, the learning rate μ is carefully selected such that $\mu P \ll 1$. The asymptotic MSE expression, $\sigma_{e_i}^2$, can therefore be simplified using the following approximations:

$$(1-\mu P)^2 \approx 1 \text{ and } 2-\mu P \approx 2.$$

Given these approximations, the error variance given in (3.3.24) becomes:

$$\sigma_{e_i}^{2\prime} = \sigma_{h_i}^2 \frac{(2\pi f_d T)^2}{2(\mu P)^2} + \sigma_{w_i}^2 \frac{\mu}{2}. \tag{3.3.26}$$

To advance, $\sigma_{e_i}^{2\prime}$ in equation (3.3.26) is minimized as follows:

$$\frac{d\sigma_{e_i}^{2\prime}}{d\mu} = \sigma_{h_i}^2 \frac{-2(2\pi f_d T)^2}{2(P)^2\mu^3} + \frac{1}{2}\sigma_{w_i}^2 = 0,$$

$$\sigma_{h_i}^2 \frac{(2\pi f_d T)^2}{(P)^2\mu^3} = \frac{1}{2}\sigma_{w_i}^2,$$

$$\mu^3 = 2\sigma_{h_i}^2 \frac{(2\pi f_d T)^2}{(P)^2\sigma_{w,i}^2},$$

$$\mu_{OPT} = 2\frac{\left(\sigma_{h_i}^2\right)^{1/3}}{\left(\sigma_{w_i}^2\right)^{1/3}} \frac{(\pi f_d T)^{2/3}}{(P)^{2/3}}. \tag{3.3.27}$$

To get the optimal Asymptotic MSE, the optimal step-size in (3.3.27) is substituted into $\sigma_{e_i}^{2\prime}$ in (3.3.26) to get:

$$\sigma_{e_i}^{2\prime}(opt) = \sigma_{h_i}^2 \frac{(2\pi f_d T)^2}{2\left(2\frac{\left(\sigma_{h_i}^2\right)^{\frac{1}{3}}}{\left(\sigma_{w_i}^2\right)^{\frac{1}{3}}} \frac{(\pi f_d T)^{\frac{2}{3}}}{(P)^{\frac{2}{3}}} P\right)^2} + \sigma_{w_i}^2 \frac{2\frac{\left(\sigma_{h_i}^2\right)^{1/3}}{\left(\sigma_{w_i}^2\right)^{1/3}} \frac{(\pi f_d T)^{2/3}}{(P)^{2/3}}}{2}$$

$$= \left(\sigma_{h_i}^2\right)^{\frac{1}{3}} \frac{4(\pi f_d T)^{\frac{2}{3}}\left(\sigma_{w_i}^2\right)^{\frac{2}{3}}}{8(P)^{\frac{2}{3}}} + \left(\sigma_{h_i}^2\right)^{\frac{1}{3}} \frac{2(\pi f_d T)^{\frac{2}{3}}\left(\sigma_{w_i}^2\right)^{\frac{2}{3}}}{2(P)^{\frac{2}{3}}}$$

$$= \frac{1}{2}\left(\sigma_{h_i}^2\right)^{\frac{1}{3}} \frac{(\pi f_d T)^{\frac{2}{3}}\left(\sigma_{w_i}^2\right)^{\frac{2}{3}}}{(P)^{\frac{2}{3}}} + \left(\sigma_{h_i}^2\right)^{\frac{1}{3}} \frac{(\pi f_d T)^{\frac{2}{3}}\left(\sigma_{w_i}^2\right)^{\frac{2}{3}}}{(P)^{\frac{2}{3}}}. \tag{3.3.28}$$

Therefore, the optimal MSE is given as:

$$\sigma_{e_i}^{2'}(opt) = \frac{3}{2}\left(\sigma_{h_i}^2\right)^{\frac{1}{3}} \frac{(\pi f_d T)^{\frac{2}{3}}\left(\sigma_{w_i}^2\right)^{\frac{2}{3}}}{(P)^{\frac{2}{3}}}. \tag{3.3.29}$$

3.4 Variable Step-Size Least Means Square (VSS-LMS) Algorithm (Adapted from [22])

This section presents the self-adaptive LMS variant algorithm with a variable step-size. Similar assumptions are made as with the CSS-LMS algorithm, and thus the pilot model given in equation (3.3.1) is also considered. Another distinguishing attribute of the VSS-LMS algorithm with respect to the CSS-LMS algorithm is the introduction of a new gradient. This gradient contributes to the gain of the LMS update rule being self-adaptive. The following is the adopted cost function:

$$V_2(\mu) = \lim_{n\to\infty} E\left\{\left|y_{i,n} - \hat{h}_i x\right|^2\right\}. \tag{3.4.1}$$

The update rule for the cost function is derived in the same approach as the one adopted in finding the one given in (3.3.5). A stochastic gradient descent is performed on the cost function to give the update rule of the variable step-size as:

$$\mu_{i,n} = \mu_{i,n-1} - \epsilon \nabla_\mu V_2(\mu, n)|_{n-1},$$

where $\mu_{i,n}$ is the variable step-size, *i* is the user index and *n* is the symbol index.

An approximation is made by using the instantaneous value of the cost function given as $V_2'(\mu) = \left|y_{i,n} - \hat{h}_i x\right|^2$. With this setting, the update rule for the variable step-size is therefore expressed:

$$\mu_{i,n} = \mu_{i,n-1} - \epsilon \nabla_\mu \left(\left|y_{i,n} - \hat{h}_{i,n-1} x\right|^2\right). \tag{3.4.2}$$

The gradient of which can be derived as:

$$\nabla_\mu V_2'(\mu) = 2x\left(y_{i,n} - \hat{h}_{i,n-1} x\right)(-1)\frac{\partial \hat{h}_{i,n-1}^{\,*}}{\partial \mu},$$

where $(\cdot)^*$ is the complex conjugate of $\frac{\partial \hat{h}_{i,n}}{\partial \mu}$, which is given by the update rule in (3.3.5).

For the sake of presentation, we take $G_{i,n} = \frac{\partial \hat{h}_{i,n}}{\partial \mu}$. The update rule of the adaptive gain can then be expressed as shown below:

$$\mu_{i,n} = \mu_{i,n-1} + 2x\epsilon\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1}x\right)G^*_{i,n-1}\right], \tag{3.4.3}$$

where $\mathbb{R}(\cdot)$ denotes the real part of a complex number.

In differentiating (3.3.5) in terms of the step-size, the update rule for $G_{i,n}$ is given as

$$G_{i,n} = \left(1 - \mu_{i,n-1}x^2\right)G_{i,n-1} + x\left(y_{i,n} - \hat{h}_{i,n-1}x\right). \tag{3.4.4}$$

With the given analysis, the VSS-LMS algorithm is essentially defined by (3.3.5) with an adaptive gain, (3.4.3) and (3.4.4). The summary of the self-adaptive VSS-LMS algorithm is given below:

$$\begin{aligned}
\hat{h}_{i,n} &= \hat{h}_{i,n-1} + \mu_{i,n-1}x\left(y_{i,n} - \hat{h}_{i,n-1}x\right), \\
G_{i,n} &= \left(1 - \mu_{i,n-1}x^2\right)G_{i,n-1} + x\left(y_{i,n} - \hat{h}_{i,n-1}x\right), \\
\mu_{i,n} &= \mu_{i,n-1} + 2x\epsilon\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1}x\right)G^*_{i,n-1}\right].
\end{aligned} \tag{3.4.5}$$

3.5 Variable Step-Size LMS Algorithm with Added Speed (VSS-LMS-AS) (Adapted from [22])

The VSS-LMS-AS algorithm presented in this section is built upon the VSS-LMS algorithm described in the preceding section. The VSS-LMS exhibits a trade-off between the speed and accuracy of the adaptive step-size $\mu_{i,n}$ because of the choice of ϵ. To improve the speed while ensuring a low error in the adaptive algorithm, an adaptive speed is added to the algorithm. To document this, a similar cost-function as with VSS-LMS is considered, which is given by:

$$V_3(\epsilon) = \lim_{n\to\infty} E\left\{\left|y_{i,n} - \hat{h}_i x_p\right|^2\right\}.$$

As with the CSS-LMS algorithm, similar pilot assumptions are made. The adaptive speed is introduced by considering additional gradient components defined as shown below:

$$\begin{aligned}
N_{i,n} &= \frac{\partial \hat{h}_{i,n}}{\partial \epsilon}, \\
M_{i,n} &= \frac{\partial G_{i,n}}{\partial \epsilon}, \\
\Psi_{i,n} &= \frac{\partial \mu_{i,n}}{\partial \epsilon},
\end{aligned} \tag{3.5.1}$$

Furthermore, the stochastic gradient descent is performed using the constant step ϵ which can be referred to in (3.4.3). An exponential forgetting factor $\rho < 1$ is used in conjunction with the multiplicative procedure given out in [22],[86]. The following update rule arises:

$$\epsilon_{i,n} = \epsilon_{i,n-1}\left(\rho - \lambda x \mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1}x\right)N^*_{i,n-1}\right]\right), \tag{3.5.2}$$

where λ is a constant step-size.

The factor $\epsilon_{i,n}$ is usually bounded $\epsilon_{i,min} \leq \epsilon_{i,n} \leq \epsilon_{i,max}$, with the value of $\epsilon_{i,max}$ designed such that the initial convergence is fast. On the other hand, $\epsilon_{i,min}$ is chosen to assume a sufficiently small value such that the convergence error can be reduced. The VSS-LMS-AS algorithm is therefore expressed as follows:

Channel Estimation

$$\hat{h}_{i,n} = \hat{h}_{i,n-1} + \mu_{i,n-1}x\left(y_{i,n} - \hat{h}_{i,n-1}x\right). \tag{3.5.3}$$

Adaptive Gain

$$\begin{aligned} G_{i,n} &= \left(1 - \mu_{i,n-1}x^2\right)G_{i,n-1} - x\left(y_{i,n} - \hat{h}_{i,n-1}x\right), \\ \mu_{i,n} &= \mu_{i,n-1} + x\epsilon_{i,n-1}\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1}x\right)G^*_{i,n-1}\right]. \end{aligned} \tag{3.5.4}$$

Adaptive speed

$$N_{i,n} = N_{i,n-1} + x\Psi_{i,n-1}\left(y_{i,n} - \hat{h}_{i,n-1}x\right) - x^2\mu_{i,n-1}N_{i,n-1},$$

$$\begin{aligned} \Psi_{i,n} = \Psi_{i,n-1} &+ x\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1}x_p\right)G^*_{i,n-1}\right] \\ &+ x\epsilon_{i,n-1}\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1}x\right)M^*_{i,n-1}\right] - x^2\epsilon_{i,n-1}\mathbb{R}\left[G^*_{i,n-1}N_{i,n-1}\right], \end{aligned}$$

$$M_{i,n} = \left(1 - x^2\mu_{i,n-1}\right)M_{i,n-1} - x^2\Psi_{i,n-1}G_{i,n-1} - x^2N_{i,n-1},$$

$$\epsilon_{i,n} = \left[\epsilon_{i,n-1}\left(\rho - \lambda x \mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1}x\right)N^*_{i,n-1}\right]\right)\right]^{\epsilon_{max}}_{\epsilon_{min}}, \tag{3.5.5}$$

where $[\epsilon]^{\epsilon_{max}}_{\epsilon_{min}}$ details that $\epsilon_{min} \leq \epsilon \leq \epsilon_{max}$.

3.6 Results and Evaluation

This subsection presents the performance evaluation of the all-known-pilots LMS-based CE technique presented in this chapter. To assess the three underlying LMS algorithms employed in this estimation approach, computer simulations were carried out. In obtaining the results, assumptions were made on some of the parameters. In accordance with the adopted system model, the power ratios of user 1 and user 2 were set as $\alpha_1 = 0.39$ and $\alpha_2 = 0.61$ respectively. The channel coefficients between each user and the base-station were generated using Jake's channel model. The corresponding variances were computed within the simulation. The normalized Doppler is given as $f_d T = 0.009$. The algorithm implementations to generate the following results are outlined by Algorithm 1 – Algorithm 8, in Appendix A.

3.6.1 Performance Evaluation of the CSS-LMS Algorithm

Firstly, the results for the all-known-pilots CE approach when employing the CSS-LMS algorithm are presented. Figure 3.6.1 and Figure 3.6.2 characterize the evolution of the analytical and the simulation MSEs for a range of step-sizes at user 1 and user 2 respectively. Comparatively, the performance at both users poses a significant resemblance. For both users, the analytical and simulated MSEs exhibit a similar curve. The differences between the analytical and the simulation are more obvious since the log scale was used in the depiction of both Figure 3.6.1 and Figure 3.6.2. However, the differences are very small since they are in the order of 10^{-4}. As such, the simulation results agree with the theoretical analysis.

The development of simulated MSE and the theoretical MSE for a range of SNR values is depicted in Figure 3.6.3 and Figure 3.6.4 for user 1 and user 2 respectively. Firstly, a step-size of $\mu = 0.21$ is employed for both users. This is the step-size at which the analytical and simulation MSE are identical in Figure 3.6.1 and Figure 3.6.2. In either account, the simulated results corroborate the theoretical analysis conducted in Section 3.3.

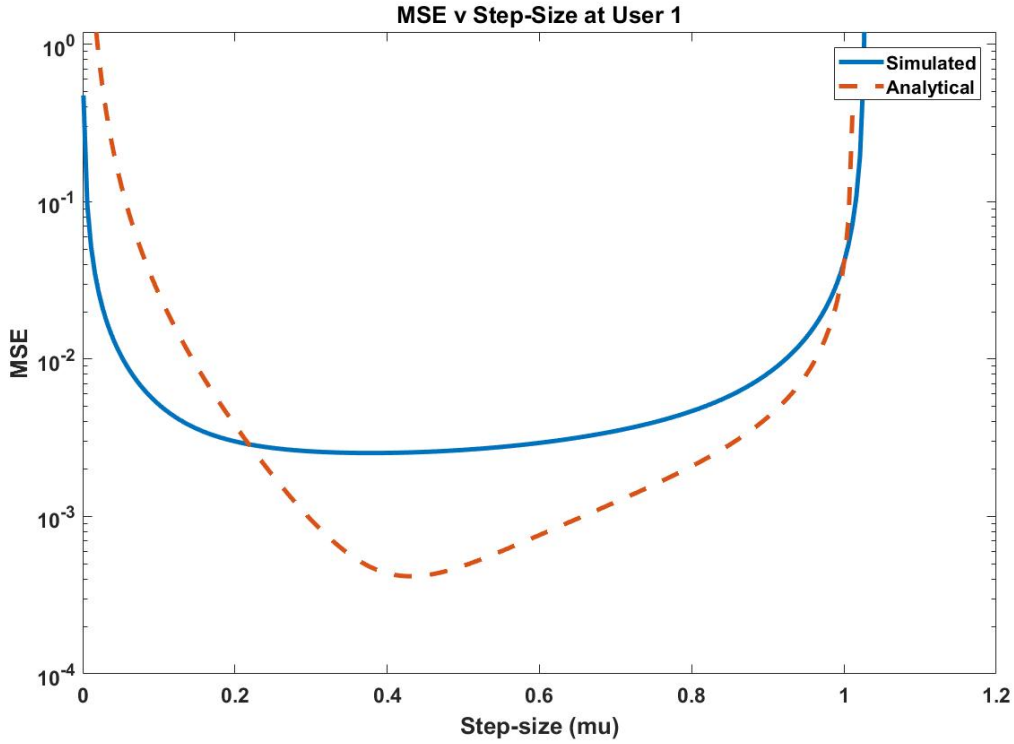


Figure 3.6. 1 User 1: Comparison of the evolution of the analytical MSE and the simulated MSE for different step-size.

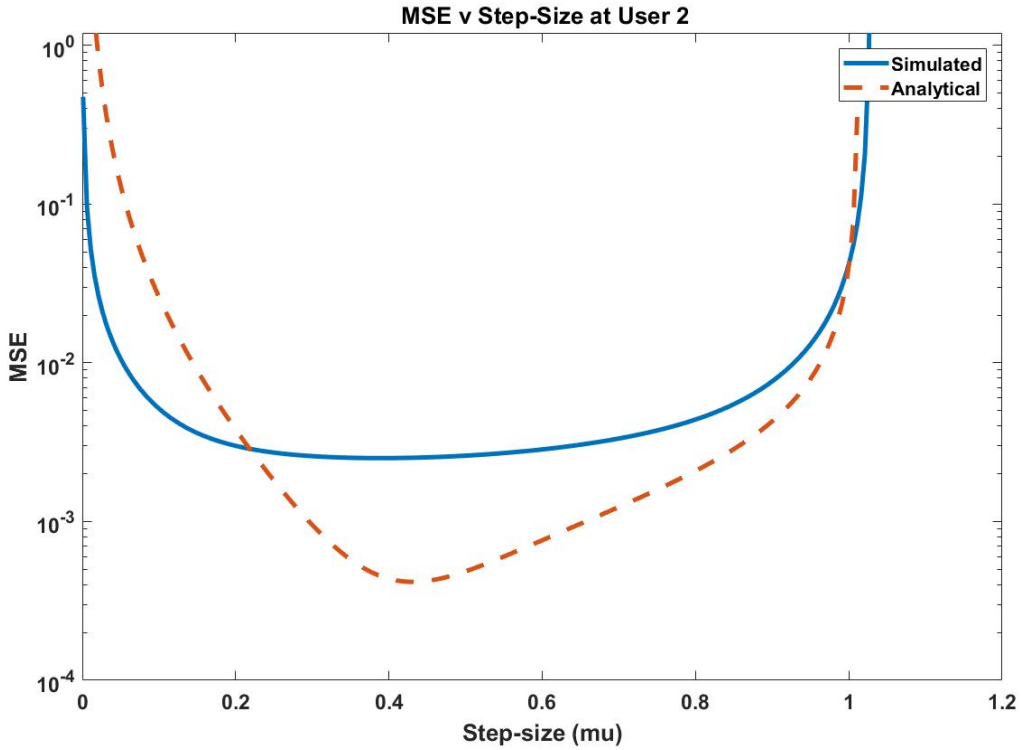


Figure 3.6. 2 User 2: Comparison of the evolution of the analytical MSE and the simulated MSE for different step-size.

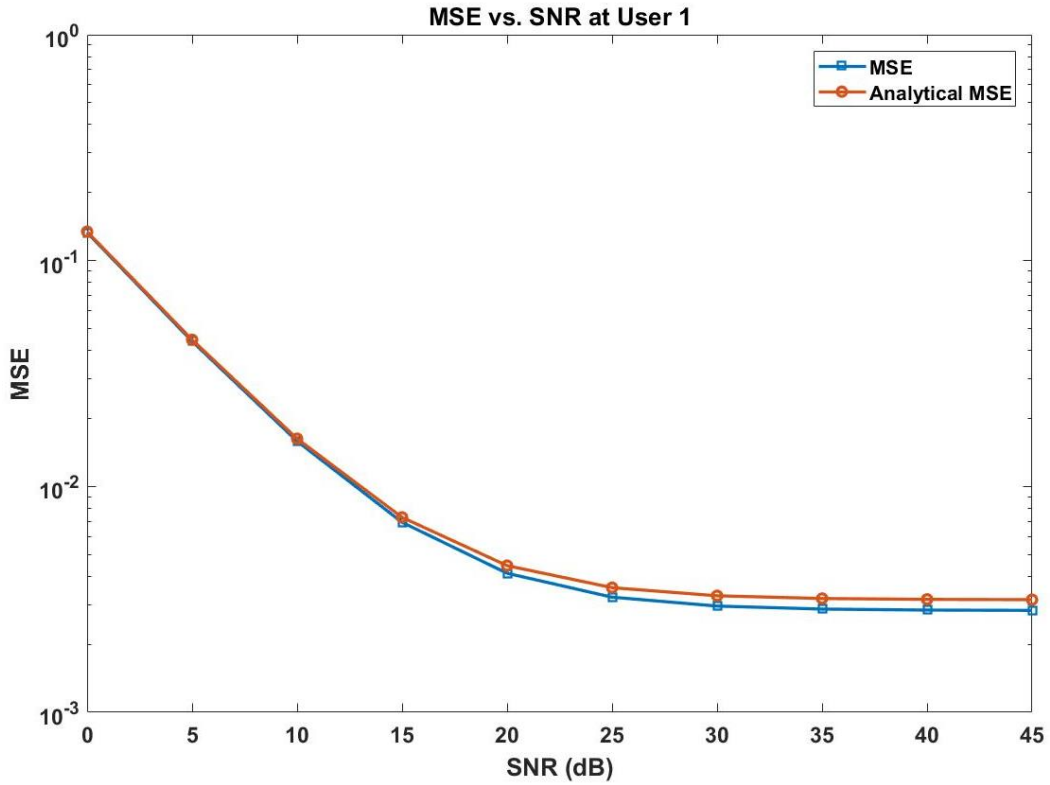
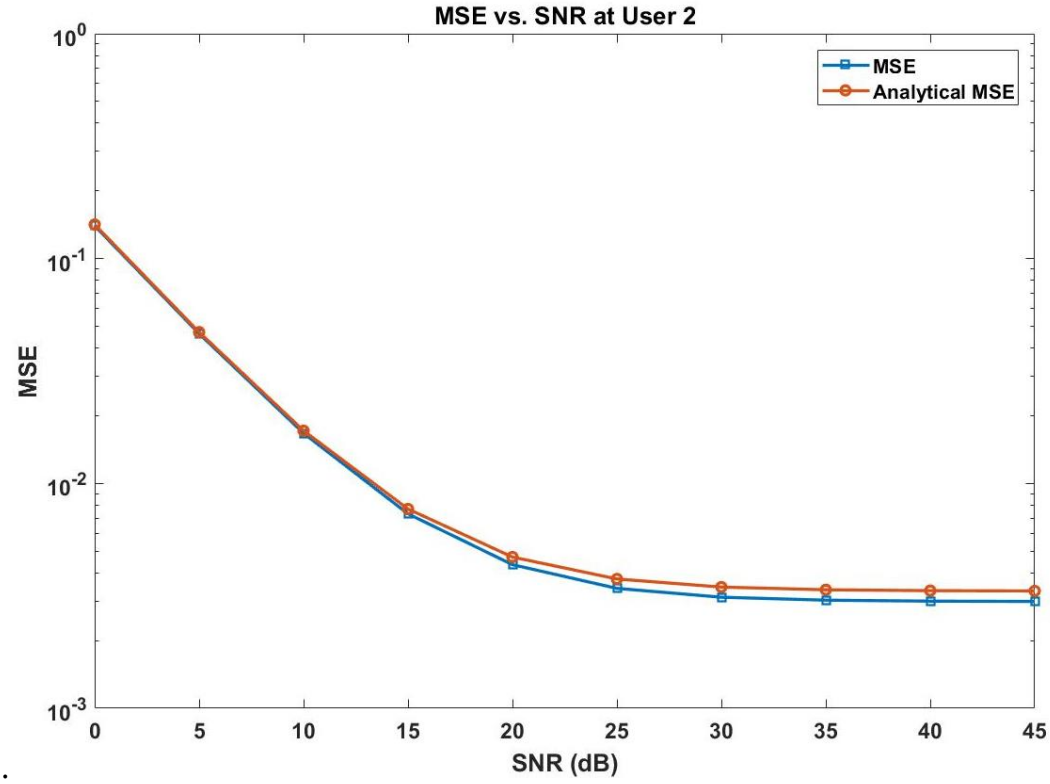


Figure 3.6. 3 User 1: Comparison between the analytical MSE and the simulated MSE using the CSS-LMS algorithm's MSE vs SNR plot.



.

Figure 3.6. 4 User 2: Comparison between the analytical MSE and the simulated MSE using the CSS-LMS algorithm's MSE vs SNR plot.

3.6.2 Comparison of the Three LMS Algorithms

Convergence characteristics for the CSS-LMS, VSS-LMS and the VSS-LMS-AS algorithms are presented and compared. For the CSS-LMS, convergence is assessed for the optimal learning rate $\mu = 0.21$. On the other hand, the VSS-LMS employs the constant step-size of $\epsilon = 10^{-5}$ while the VSS-LMS-AS also has a constant learning rate of $\lambda = 10^{-5}$. The selected constant step-sizes give the best achievable performance for both the VSS-LMS and the VSS-LMS-AS in the scenarios considered. For the VSS-LMS-AS, the forgetting factor is taken as $\rho = 0.98$. The SNR is fixed at 30dB for the convergence characteristics at both users and the results are averaged over 2000 Monte-Carlo runs.

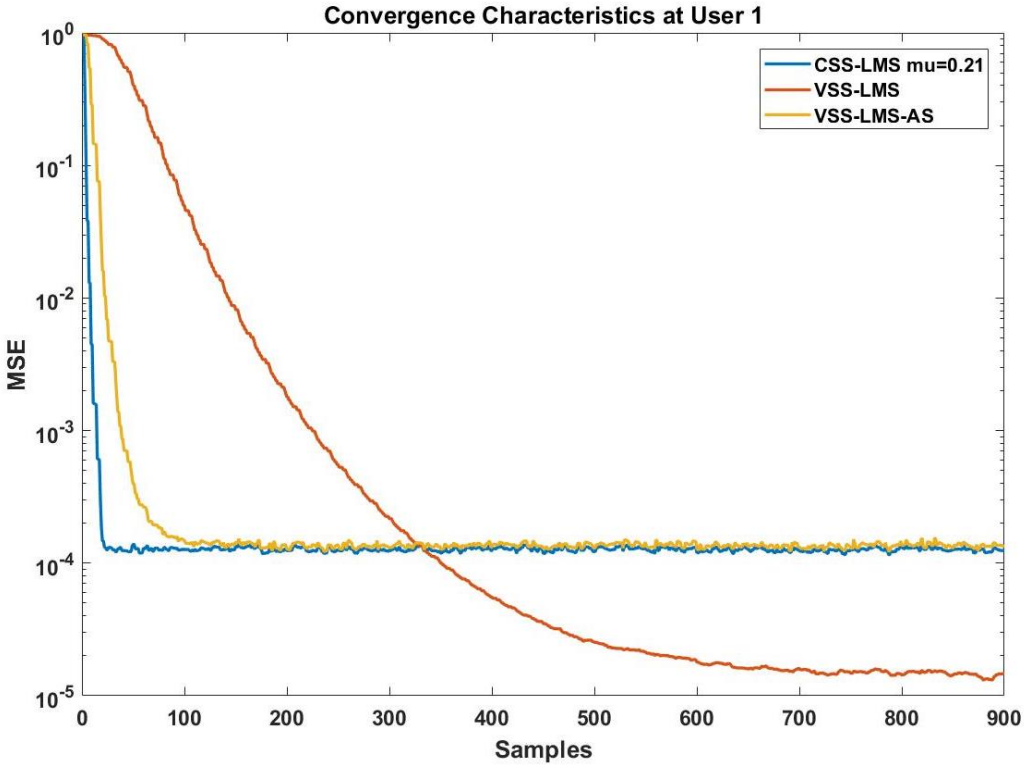


Figure 3.6. 5 User 1: MSE convergence characteristics for the three LMS algorithms.

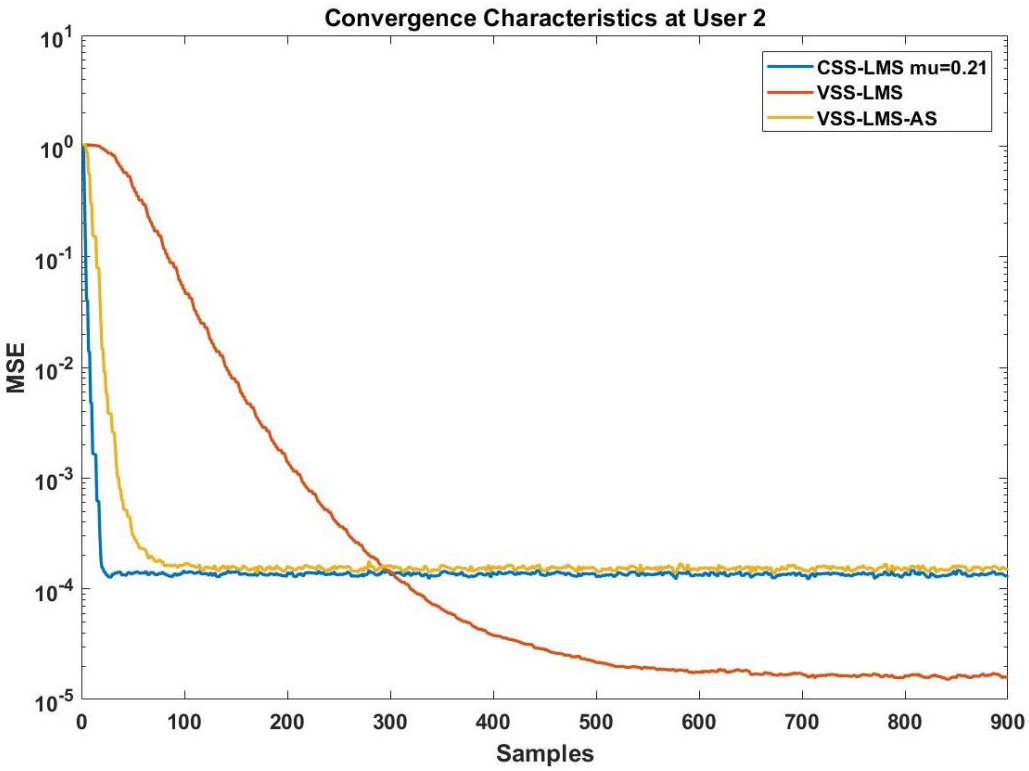


Figure 3.6. 6 User 2: MSE convergence characteristics for the three LMS algorithms.

Figure 3.6.5 presents the convergence characteristics of the all-known-pilots approach using the CSS-LMS, VSS-LMS and the VSS-LMS-AS algorithms at user 1's receiver. In terms of accuracy, the performances of the three algorithms are comparable. The VSS-LMS is slightly better than the other two algorithms in terms of accuracy. However, the difference of the VSS-LMS algorithm in accuracy to the other two is very small and hence is not significant. The CSS-LMS shows the best performance in terms of channel estimation speed. The additional derivatives that introduce adaptive speed to VSS-LMS algorithm enable the superior performance in terms of speed of the VSS-LMS-AS algorithm between the two.

The disparity between the channel estimation performance at user 1 and at user 2 is virtually insignificant. The dissimilarity is mainly influenced by the traits of the channel between each user and the base-station as well as the additive white Gaussian noise.

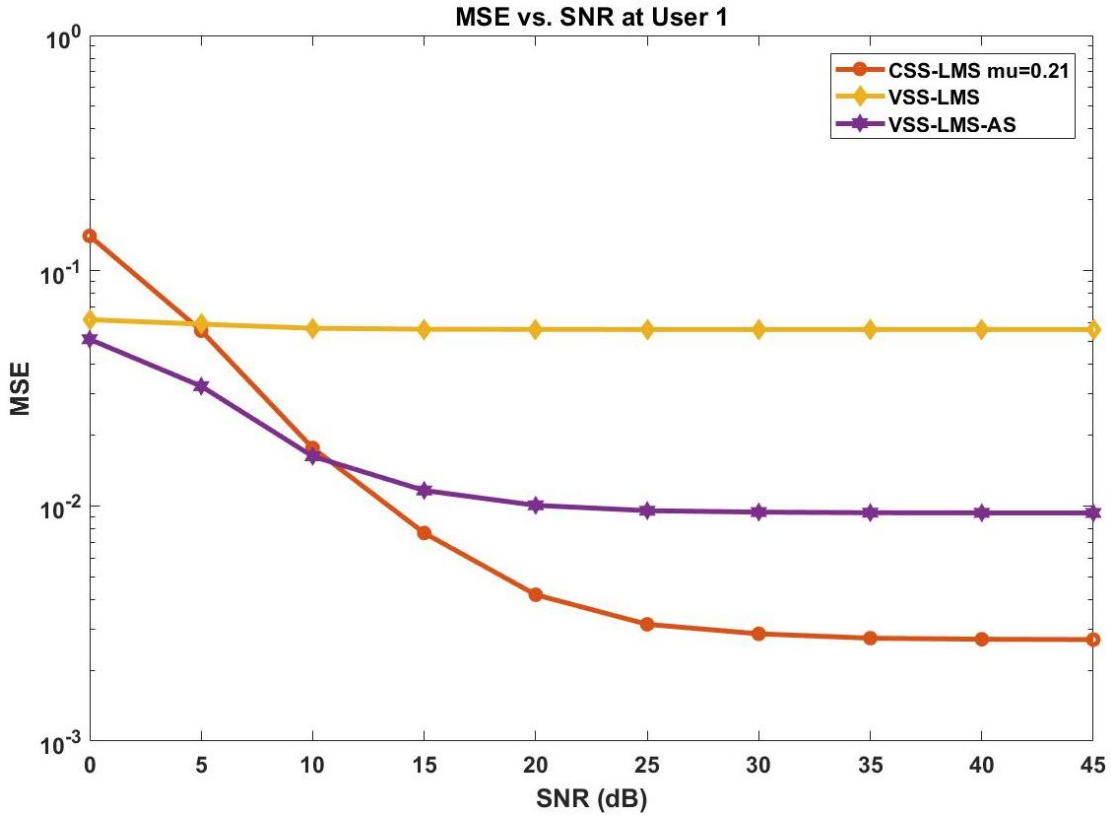


Figure 3.6. 7 User 1: MSE vs SNR curve for the three LMS algorithms.

Figure 3.6.7 and Figure 3.6.8 depict the MSE vs SNR curve for user 1 and user 2 respectively. Comparing the three algorithms, the CSS-LMS algorithm exhibits the best MSE performance relative to the self-adaptive LMS algorithm variants. The VSS-LMS-AS algorithm has the best MSE performance relative to the VSS-LMS, for the entire range of SNR values considered. The superior performance attributed to the VSS-LMS-AS algorithm as opposed to the VSS-LMS algorithm is a consequence of the adaptive speed derivatives included in the algorithm. Employment of the optimal step-size for the CSS-LMS algorithm is the reason for the best performance when compared to the self-adaptive algorithms.

Essentially, these results confirm the results presented in Figure 3.6.5 and Figure 3.6.6. Given the addition of new derivatives, the self-adaptive algorithms exhibit a slower CE speed in comparison with the CSS-LMS algorithm. At the user level, channel estimation performance between user 1 and user 2 is comparable with very small disparities.

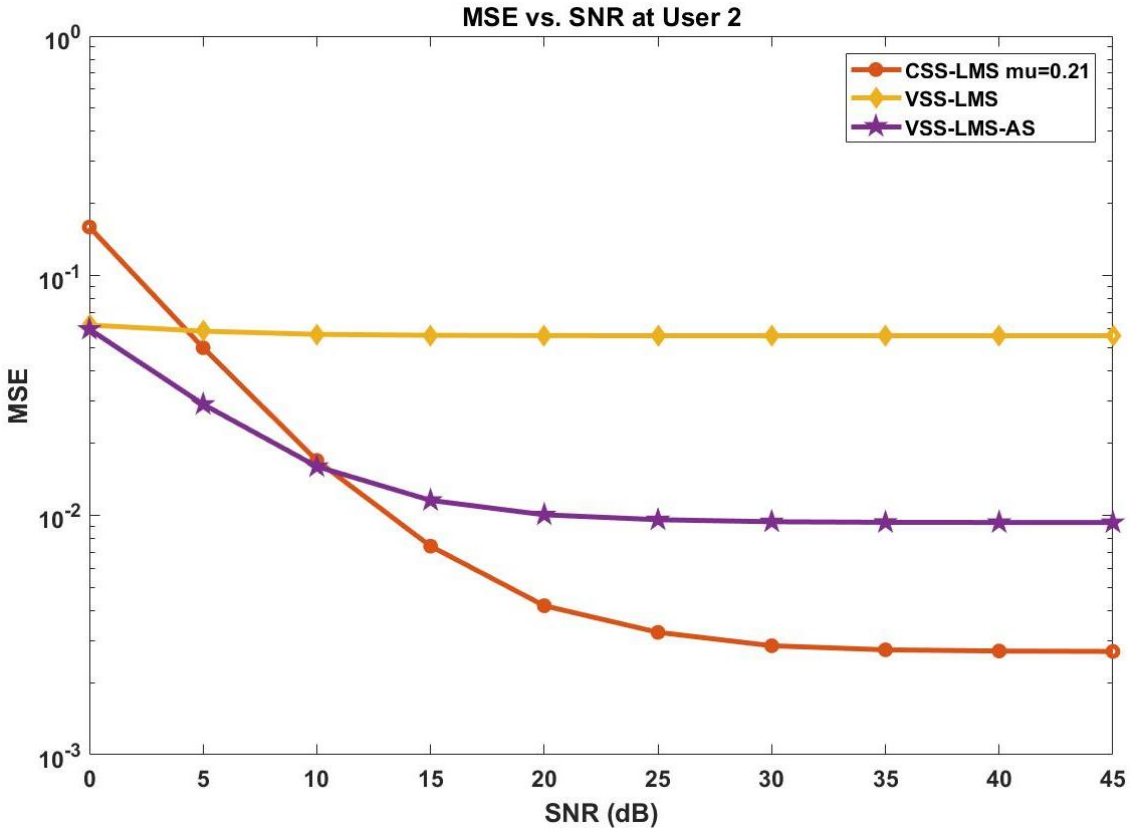


Figure 3.6. 8 User 2: MSE vs SNR curve for the three LMS algorithms.

In both Figure 3.6.7 and Figure 3.6.8, the VSS-LMS algorithm seemingly has a straight curve because of the small change it exhibits compared to the other two algorithms. Figure 3.6.9 and Figure 3.6.10 shows the explicit change of the VSS-LMS algorithm at user 1 and user 2 respectively. As shown in Figure 3.6.9 and Figure 3.6.10, the MSE of the VSS-LMS algorithm varies between around 0.07 and 0.05, which is a very small range when compared to the other two algorithms.

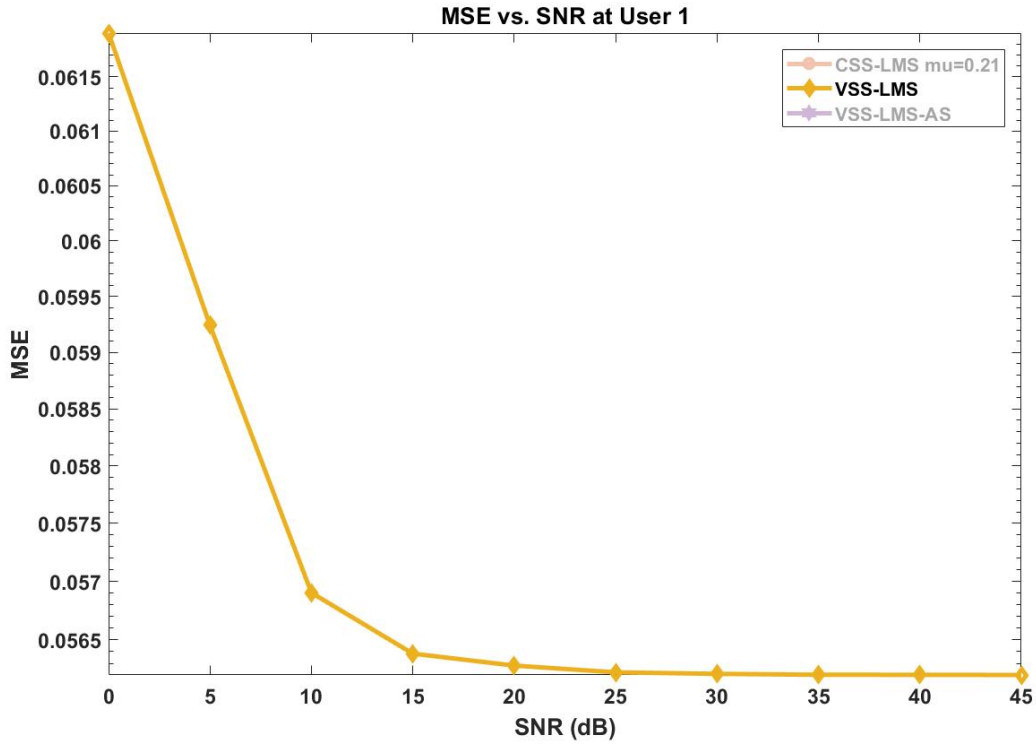


Figure 3.6. 9 User 1: MSE vs. SNR Curve for the VSS-LMS Algorithm.

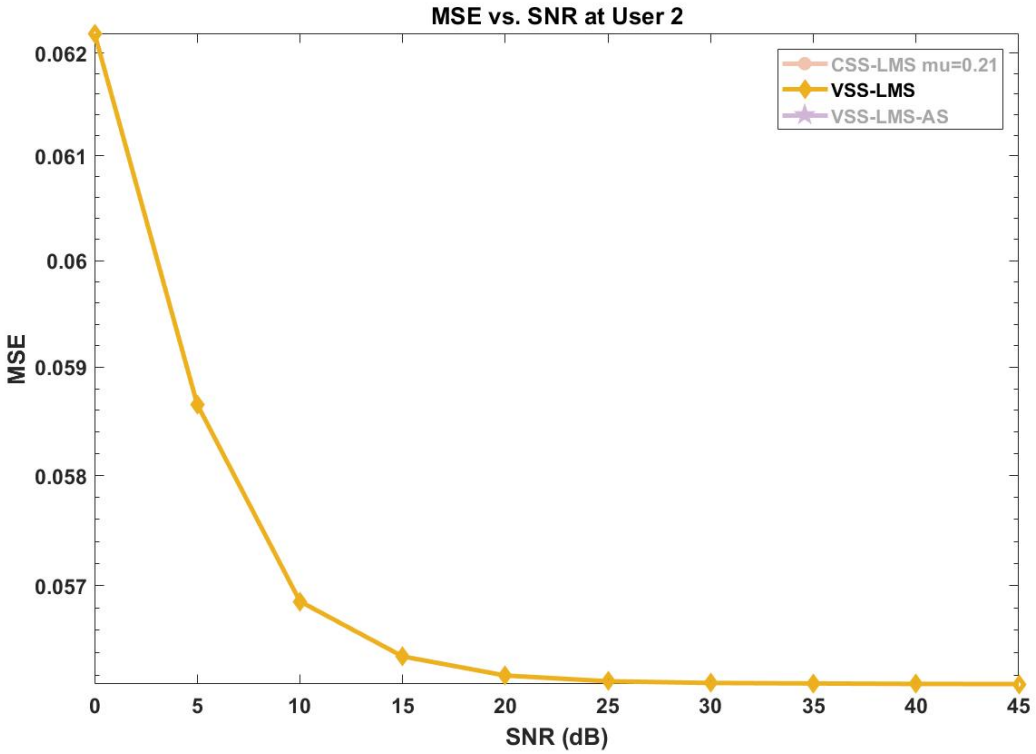


Figure 3.6.10 User 2: MSE vs. SNR Curve for the VSS-LMS Algorithm.

3.7 Chapter Summary

This chapter presented the system model considered throughout the dissertation. The description of the system model, comprised of both the pilot model and the channel model, was covered in the first section of the chapter.

The all-known-pilots based CE approach was then presented, employing three LMS algorithms. Adopting the work presented in [22], innovation was introduced with the translation and application of the three LMS variants to a SISO-NOMA model. Firstly, the CSS-LMS was presented with the corresponding closed form expressions of the error variance. Secondly, the self-adaptive VSS-LMS was presented with detail of the update rule and adaptive gain. Thirdly, the VSS-LMS-AS was also presented for the all-known-pilots based CE scheme.

Through computer simulations, the all-known-pilots scheme was assessed when using the three LMS variants. For the CSS-LMS algorithm, the CE performance of the simulation was compared to the analysis. In observing the simulation and the analytical MSE for a range of step-sizes, the two were identical for the step-size of 0.21 for both users. For the rest of the step-sizes, the analytical and the simulation MSEs showed similar curves with very small disparities. The analysis was therefore confirmed through simulation for both users. Similarly, the analysis was confirmed for the MSE vs. SNR curve. At a user level, CE performance was found to be virtually similar. The main reason for this is the fact that the same superimposed pilot is sent to both users.

Comparing the three LMS variants, the CSS-LMS algorithm out-performed the VSS-LMS and the VSS-LMS-AS algorithms in terms of channel estimation speed. The VSS-LMS-AS algorithm was the best in terms of the channel estimation speed, amongst the two self-adaptive algorithms. In terms of CE accuracy, the three algorithms showed comparable performance.

CHAPTER 4

LEAST MEAN SQUARES BASED AVERAGING SUM CHANNEL ESTIMATION APPROACH

4.1 Introduction

Chapter three presented a channel estimation approach aided by full knowledge of the training data employed. The main rationale for the approach is in essence the low complexity and accuracy attributed to it. However, the trade-off for the low complexity is the pilot overhead. As a measure to combat pilot overhead, a scheme that employs partial use of pilots or a completely blind scheme can be used. Hence, an LMS-based averaging sum CE technique for a downlink SISO-NOMA system that assumes partial pilot knowledge is proposed in this chapter.

Relative to the approach proposed in the previous chapter, only part of the superimposed pilot signal is known at the receiver. Similar to the previous approach, this approach exploits the three LMS variants as core, the CSS-LMS, VSS-LMS and the VSS-LMS-AS algorithms. In evaluating the three LMS variants through computer simulations, the following deductions on their respective performances were made: Using the MSE criterion, the three LMS algorithms exhibit similar CE accuracy at either user. At either user, the CSS-LMS algorithm exhibits the best performance in terms of speed, with a selection of an optimal step-size.

For the scenarios considered, the VSS-LMS-AS algorithm has better CE performance in terms of speed when compared to the VSS-LMS counterpart. At each user, the user with a higher power allocation ratio exhibits a superior performance for both speed and accuracy. This is when gauged against the user with a weaker power allocation. This observation signifies the importance and effect of power allocation on CE performance. The simulations of the CSS-LMS algorithm adhere to the theoretical analysis carried out for the asymptotic MSE.

The rest of the chapter adheres to the following organization: Section 4.2 constitutes the averaging sum formulation. Thereby introducing an evolved system model from the one presented in chapter three. The CSS-LMS algorithm is detailed in Section 4.3. Section 4.4 touches on the VSS-LMS based averaging sum CE approach, detailing the introduction of a new adaptive gain and highlighting the channel estimator. The VSS-LMS-AS is then presented

in Section 4.5. Subsequently, simulation results are presented in Section 4.6 alongside their complementary evaluation. Section 4.7 provides the conclusion to chapter four.

4.2 System Model

Assuming the system model as presented in the third chapter, this subsection presents the averaging sum operation on the observation. The resultant of the process is in turn used in the construction of the LMS cost function for the three algorithms. Relative to the previous all-known-pilots based CE approach, only the pilot sequence intended for the receiving user is known. Save for the signal intended for the receiving user, the rest of the signal is unknown.

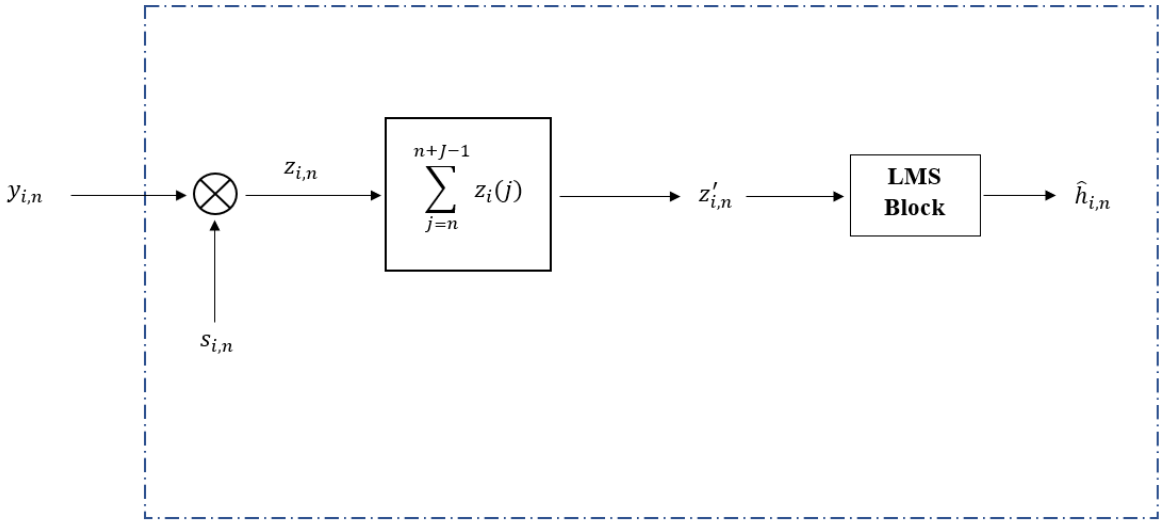


Figure 4.2. 1 Averaging Sum Operation.

Figure 4.2.1 demonstrates the averaging sum operation performed on the observation received at the *i-th* user. In essence, the averaging sum operation is performed in order to minimize the significance of the unknown portion of the overall received signal. The observation given in (3.2.2) can be re-written as shown in equation (4.2.1) for the receiving user *i*:

$$y_{i,n} = h_{i,n}\sqrt{\alpha_i P}s_{i,n} + h_{i,n}\sqrt{\alpha_k P}s_{k,n} + w_{i,n}, \qquad (4.2.1)$$

where *i* and *k* are indices of the receiving user and the "other" user respectively and *n* denotes the symbol index and the channel coefficient $h_{i,n}$ is constant for 1000 symbols.

At the *i-th* user's receiver, only the pilot s_i intended for user *i* is known within the observation given in (4.2.1). A product of the received signal and the known pilot is performed to define a new variable $z_{i,n} = y_{i,p,n} * s_{i,n}$, which is expressed as:

$$z_{i,n} = h_{i,n}\sqrt{\alpha_i P}s_{i,n}^2 + h_{i,n}\sqrt{\alpha_k P}s_{k,n}s_{i,n} + w_{i,n}s_{i,n}, \tag{4.2.2}$$

The elements of z_i are then summed up within a given frame of length *J*, and a new variable is defined as $z'_{i,n} = \sum_{j=n}^{n+J-1} z_i(j)$. This is expressed as:

$$z'_{i,n} = Jh_{i,n}\sqrt{\alpha_i P} + h_{i,n}\sqrt{\alpha_k P}\sum_{j=n}^{n+J-1} s_k(j)s_i(j) + \sum_{j=n}^{n+J-1} s_i(j)w_{i,n}, \tag{4.2.3}$$

where *n* denotes the symbol index, $s_i^2 = 1$, *k* is the index of the other user.

To reiterate, the averaging sum operation is performed in order to minimize the significance of the unknown pilot in the overall observation. This can be seen in the second term of (4.2.3), which is relatively smaller depending on the cross-correlation properties of s_k and s_i. In comparison, the unknown signal s_k is more significant in equation (4.2.1) than in (4.2.3). The three LMS algorithms are therefore built upon $z'_{i,n}$, employing it in the selection of the cost function for which the stochastic gradient is to be performed. For convenience, the following variables are introduced:

$$\theta_{s_k s_i} = \sum_{j=n}^{n+J-1} s_k(j)s_i(j) \text{ and } \theta_{s_i} = \sum_{j=n}^{n+J-1} s_i(j).$$

The introduction of variable θ_{s_i} is predicated on the assumption that the noise components are uncorrelated. However, it is important to note that the noise component $\sum_{j=n}^{n+J-1} s_i(j)w_{i,n}$ is a sum of correlated random variables (r.v.). To show this, a new variable $w'_{i,n}$ is defined at any index n such that:

$$w'_{i,n} = \sum_{j=n}^{n+J-1} s_i(j)w_{i,n}. \tag{4.2.4}$$

At index $n+1$, the following is true

$$w'_{i,n+1} = w'_{i,n} + s_i(j)w_{i,n+1} - s_i(j)w_{i,n-1}. \tag{4.2.5}$$

Strictly speaking, noise components $w'_{i,n}$ are correlated random variables. However, for deriving the error variance in Section 4.3.1 the noise components $w'_{i,n}$ are treated as

uncorrelated. As such, Figure 4.2.2, Figure 4.2.3 and Figure 4.2.4 compare the statistical properties of the noise components in both cases. That is, when correlation is assumed and when it is ignored. For different SNR instances, the power spectral density (PSD) and autocorrelation of the noise with and without the assumption of correlation is compared. In each of the subsequent figures, the results attributed to the case where the noise correlation is considered are labelled "Avg Sum Formulation". Observing the PSD and the autocorrelation of both accounts at each SNR, the disparities are less significant. Consequently, the theoretical analysis as conducted without accounting for the correlation is justified.

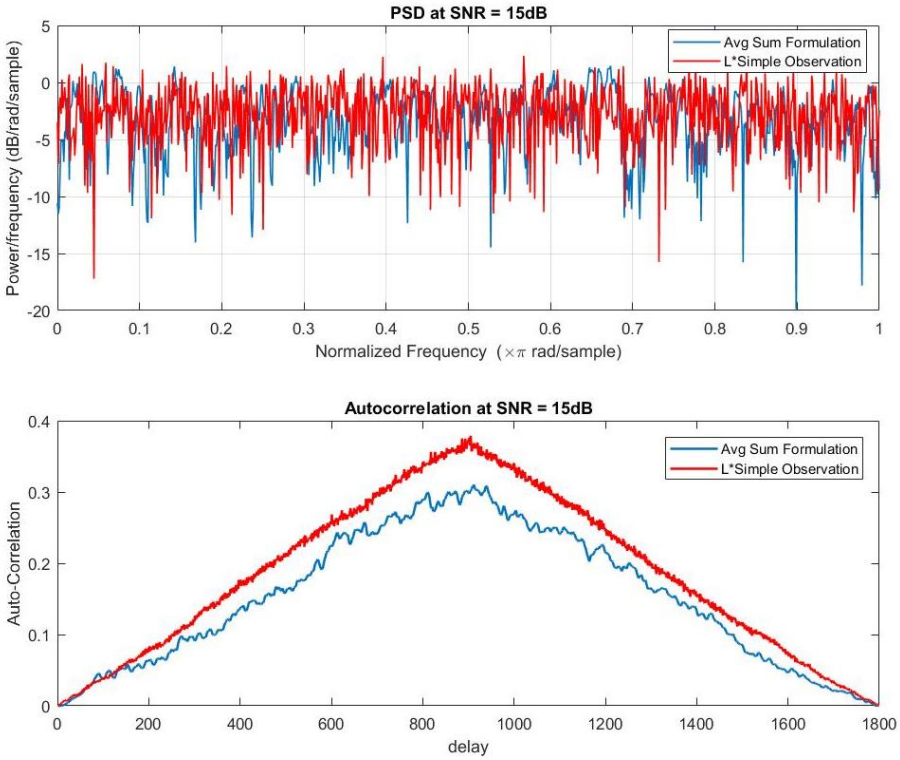


Figure 4.2. 2 Comparison of the statistical properties between the noise components using the Averaging sum formulation and when no correlation is assumed. SNR = 15dB.

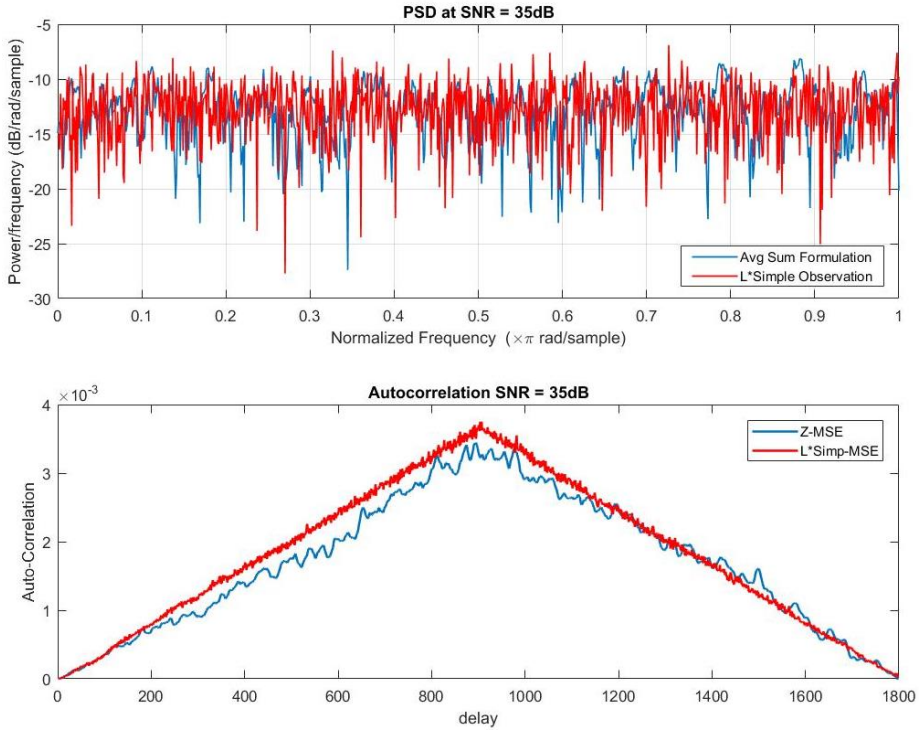


Figure 4.2. 3 Comparison of the statistical properties between the noise components using the Averaging sum formulation and when no correlation is assumed. SNR = 35dB.

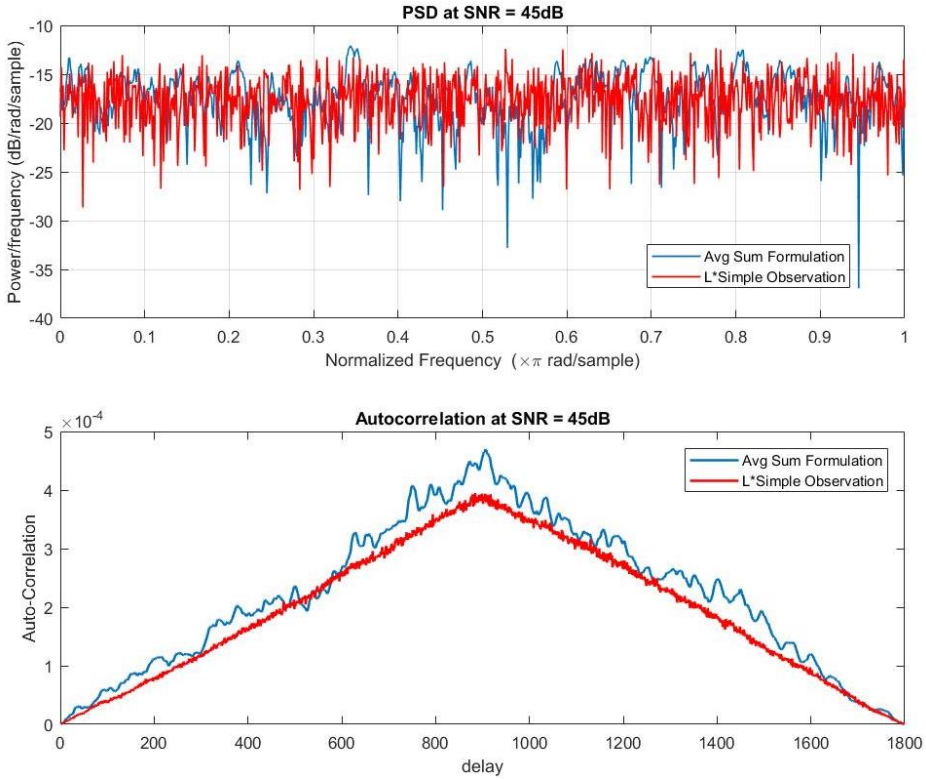


Figure 4.2. 4 Comparison of the statistical properties between the noise components using the Averaging sum formulation and when no correlation is assumed. SNR = 45dB.

4.3 CSS-LMS Algorithm (Adapted from [22])

Channel estimation in this segment is based on the CSS-LMS algorithm, assuming partial knowledge of the superimposed pilot signal. Essentially, this algorithm is the one presented in Section 3.3. The major difference is that, instead of a composite signal x of known pilots, only the known pilot factor $J\sqrt{\alpha_i P}$ is employed in the construction of the cost function. Taking the formulation segment in Section 4.2 as the foundation, the following is selected as the cost function:

$$V_1(\hat{h}_{i,n}) = E\left\{\left|z'_{i,n} - \hat{h}_{i,n} J\sqrt{\alpha_i P}\right|^2\right\}. \tag{4.3.1}$$

The stochastic gradient similar to the one shown in Section 3.3 is performed on the cost function in (4.3.1), the resultant update rule is given as:

$$\hat{h}_{i,n} = \hat{h}_{i,n-1} + \mu J\sqrt{\alpha_i P}\left(z'_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right), \tag{4.3.2}$$

where μ is the learning rate.

The CSS-LMS Algorithm using the averaging sum formulation is defined by the update equation in (4.3.2).

4.3.1 Error Variance of the Averaging Sum CSS-LMS Algorithm

Similar to the procedure followed in deriving the asymptotic MSE of the CSS-LMS algorithm in Section 3.3.1, the analytical analysis of the error variance is presented. However, instead of a composite signal x of known pilots, only the known pilot factor $J\sqrt{\alpha_i P}$ is employed. For the updated rule in (4.3.2), the low-pass filter is designed in the z-domain as follow:

$$L(z) = \frac{\mu J\sqrt{\alpha_i P}}{1 - z^{-1}(1 - \mu J^2 \alpha_i P)}, \tag{4.3.3}$$

with cut-off frequency selected to be $f_c T = \frac{\mu J^2 \alpha_i P}{2\pi(1-\mu J^2 \alpha_i P)}$.

Using the error expression in (3.2.4), the formulated equation (4.2.3) and the low-pass filter in (4.3.3), the channel estimation error is in turn expressed in the z-domain as follows:

$$e_i(z) = \left(1 - L(z) J\sqrt{\alpha_i P}\right) h_i(z) - L(z) h_i(z)\sqrt{\alpha_k P}\theta_{S_k S_i} - L(z) W_i(z)\theta_{S_i}. \tag{4.3.4}$$

From the CE error given in equation (4.3.4), the Asymptotic MSE in the frequency domain can be defined as:

$$\sigma_{e_i}^2 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - J\sqrt{\alpha_i P} L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_i}(f)\, df + \alpha_k P \theta_{s_k s_i}^2 \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_h(f)\, df$$

$$+\ \theta_{s_i}^2 \sigma_{w_i}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df, \qquad (4.3.5)$$

such that the MSE variance can be separated and evaluated disjointly as:

$$aMSE1 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - J\sqrt{\alpha_i P} L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_i}(f)\, df,$$

$$aMSE2 = \alpha_k P \theta_{s_k s_i}^2 \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_i}(f)\, df,$$

$$aMSE3 = \theta_{s_i}^2 \sigma_{w_i}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df. \qquad (4.3.6)$$

The approximation of $aMSE1$ in (4.3.6) is strikingly similar to the one made for (3.3.10) under Section 3.3.1. Given only one pilot is known in this analysis, the difference is the pilot factor and the low-pass filter design. In this approach, the known pilot factor is given by $J\sqrt{\alpha_i P}$ as opposed to x in (3.3.10). Furthermore, $aMSE3$ in this approach has an identical integral to the equation in (3.3.11). The low-pass filter is defined differently, even though it has the same form. A new term is introduced in this analysis, $aMSE2$, which embodies the unknown pilot term. As such, employing the procedure adopted in finding the expression for (3.3.10) in Section 3.3.1, the following is the resultant expression for $aMSE1$:

$$aMSE1 = \sigma_{h_i}^2 \frac{(2\pi f_d T)^2}{2(\mu J^2 \alpha_i P)^2} (1 - \mu J^2 \alpha_i P)^2. \qquad (4.3.7)$$

Similarly, using the corresponding approach in finding the integral of (3.3.11) in Section 3.3, the result of the integral can be expressed as: $\frac{\mu}{T(2-\mu J^2 \alpha_i P)}$. The resulting expression for $aMSE3$ is therefore given by:

$$aMSE3 = \theta_{s_i}^2 \sigma_{w_i}^2 \frac{\mu}{(2 - \mu J^2 \alpha_i P)}. \qquad (4.3.8)$$

For $aMSE2$, similar assumptions of low Doppler frequency scenarios are made as with the analysis given for *MSE1* in Section 3.3 in the error variance segment. These assumptions are originally adopted as with the work presented in [84]. The following derivation is effected:

$$aMSE2 = \alpha_k P\theta^2_{s_k s_i} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_i}(f)\, df$$

$$= \alpha_k P\theta^2_{s_k s_i} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|\frac{\mu J\sqrt{\alpha_i P}}{1-(1-j2\pi fT)(1-\mu L^2\alpha_i P)}\right|^2 \frac{\sigma^2_{h_i}}{\pi\sqrt{f_d^2 - f^2}}\, df$$

$$= \alpha_k P\theta^2_{s_k s_i} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|\frac{\mu J\sqrt{\alpha_i P}}{(1-\mu J^2\alpha_i P)\left[\frac{1-(1-\mu J^2\alpha_i P)}{(1-\mu J^2\alpha_i P)} + j2\pi fT\right]}\right|^2 \frac{\sigma^2_{h_i}}{\pi\sqrt{f_d^2 - f^2}}\, df.$$

Taking $K = \frac{\mu J\sqrt{\alpha_i P}}{1-\mu J^2\alpha_i P}$,

$$aMSE2 = \alpha_k P\theta^2_{s_k s_i} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|\frac{K}{2\pi f_c T + j2\pi fT}\right|^2 \frac{\sigma^2_{h_i}}{\pi\sqrt{f_d^2 - f^2}}\, df$$

$$= \alpha_k P\theta^2_{s_k s_i} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \frac{K^2}{(2\pi f_c T)^2 + (2\pi fT)^2} \frac{\sigma^2_{h_i}}{\pi\sqrt{f_d^2 - f^2}}\, df. \qquad (4.3.9)$$

With the assumptions considered, the following approximation persists $(2\pi f_c T)^2 + (2\pi fT)^2 \approx (2\pi f_c T)^2$. With this setting, $aMSE2$ becomes:

$$aMSE2 = 2\alpha_k P\theta^2_{s_k s_i} \int_0^{f_d} \frac{K^2}{(2\pi f_c T)^2} \frac{\sigma^2_{h_i}}{\pi\sqrt{f_d^2 - f^2}}\, df. \qquad (4.3.10)$$

Following the approach of finding MSE1 in Section 3.3, a change of variables is introduced with $f = f_d \cos\varphi$, which results in:

$$aMSE2 = 2\alpha_k P\theta^2_{s_k s_i}\sigma^2_{h_i} \frac{K^2}{(2\pi f_c T)^2} \int_{\frac{\pi}{2}}^{0} \frac{1}{\pi\sqrt{f_d^2(1-\cos^{-2}\varphi)}} - f_d \sin\varphi\, d\varphi$$

$$= 2\alpha_k P\theta^2_{s_k s_i}\sigma^2_{h_i} \frac{K^2}{\pi(2\pi f_c T)^2} \int_0^{\frac{\pi}{2}} 1 d\varphi$$

$$= 2\alpha_k P\theta^2_{s_k s_i}\sigma^2_{h_i} \frac{K^2}{\pi(2\pi f_c T)^2} \frac{\pi}{2}. \qquad (4.3.11)$$

Substituting for K and $2\pi f_c T$, results in:

$$aMSE2 = \alpha_k P\theta_{s_k s_i}^2 \sigma_{h_i}^2 \frac{\left(\mu J\sqrt{\alpha_i P}\right)^2}{(1-\mu J^2\alpha_i P)^2} \frac{(1-\mu J^2\alpha_i P)^2}{(\mu J^2\alpha_i P)^2}$$

$$= \sigma_{h_i}^2 \frac{\alpha_k P\theta_{s_k s_i}^2}{J^2\alpha_i P}. \tag{4.3.12}$$

The realization the sum of equations (4.3.7), (4.3.8) and (4.3.12) constitute to the overall asymptotic MSE, which is given by the following expression:

$$\sigma_{e_i}^2 = \sigma_{h_i}^2 \frac{(2\pi f_d T)^2}{2(\mu J^2\alpha_i P)^2}(1-\mu J^2\alpha_i P)^2 + \sigma_{h_i}^2 \frac{\alpha_k P\theta_{s_k s_i}^2}{J^2\alpha_i P} + \theta_{s_i}^2\sigma_{w_i}^2 \frac{\mu}{(2-\mu J^2\alpha_i P)}. \tag{4.3.13}$$

The optimal learning rate and the corresponding error variance follow the same derivations as previously mentioned in Section 3.3. The learning rates selected for (4.3.13) are sufficiently small such that the following approximations hold:

$(1-\mu J^2\alpha_i P)^2 \approx 1$ and $2-\mu J^2\alpha_i P \approx 2$.

The optimal learning rate is given as:

$$\mu_{OPT} = 2\frac{\left(\sigma_{h_i}^2\right)^{1/3}}{\left(\theta_{s_i}^2\sigma_{w_i}^2\right)^{1/3}} \frac{(\pi f_d T)^{2/3}}{(J^2\alpha_i P)^{2/3}}. \tag{4.3.14}$$

The corresponding optimal asymptotic MSE is expressed as:

$$\sigma_{e_{i,(OPT)}}^{2'} = \left(\sigma_{h_i}^2\right)^{\frac{1}{3}} \left(\frac{\pi f_d T}{J^2\alpha_i P}\right)^{\frac{2}{3}} \left(\theta_{s_i}^2\sigma_{w_i}^2\right)^{\frac{2}{3}} + \sigma_{h_i}^2 \frac{\alpha_k P\theta_{s_k s_i}^2}{J^2\alpha_i P}. \tag{4.3.15}$$

4.4 VSS-LMS Algorithm (Adapted from [22])

This section re-presents the VSS-Algorithm earlier derived in chapter three, Section 3.4, for the averaging sum-based CE scheme. With similar assumptions, the following is adopted as the cost function:

$$V_2\left(\hat{h}_{i,n}\right) = \lim_{n\to\infty} E\left\{\left|z'_{i,n} - \hat{h}_{i,n}J\sqrt{\alpha_i P}\right|^2\right\}. \tag{4.4.1}$$

Approximating the cost function by employing its instantaneous value, $V_2'(\mu) = \left|z'_{i,n} - \hat{h}_{i,n}J\sqrt{\alpha_i P}\right|^2$, the stochastic gradient descent is performed in terms of the learning rate. The VSS-LMS algorithm for the averaging sum formulation is therefore summarized as follows:

$$\hat{h}_{i,n} = \hat{h}_{i,n-1} + \mu_{i,n-1}J\sqrt{\alpha_i P}\left(z'_{i,n-1} - \hat{h}_{i,n-1}J\sqrt{\alpha_i P}\right),$$

$$G_{i,n} = \left(1 - \mu_{i,n-1} J^2 \alpha_i P\right) G_{i,n-1} + J\sqrt{\alpha_i P}\left(z'_{i,n-1} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right),$$

$$\mu_{i,n} = \mu_{i,n-1} + 2J\sqrt{\alpha_i P}\epsilon \mathbb{R}\left[\left(z'_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right) G^*_{i,n-1}\right]. \quad (4.4.2)$$

4.5 VSS-LMS-AS Algorithm (Adapted from [22])

The VSS-LMS-AS Algorithm using the averaging sum formulation is also built upon the VSS-LMS algorithm, which in this case is presented in the previous section. Adopted for CE in a DL SISO-NOMA system with partial pilot knowledge, the algorithm is summarized as follows:

Channel Estimation

$$\hat{h}_{i,n} = \hat{h}_{i,n-1} + \mu J\sqrt{\alpha_i P}\left(y_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right). \quad (4.5.1)$$

Gain Estimation

$$G_{i,n} = \left(1 - \mu_{i,n-1} J^2 \alpha_i P\right) G_{i,n-1} - J\sqrt{\alpha_i P}\left(y_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right),$$

$$\mu_{i,n} = \mu_{i,n-1} + J\sqrt{\alpha_i P}\epsilon_{i,n-1} \mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right) G^*_{i,n-1}\right]. \quad (4.5.2)$$

Adaptive speed

$$N_{i,n} = N_{i,n-1} + J\sqrt{\alpha_i P}\Psi_{i,n-1}\left(y_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right) - J^2 \alpha_i P \mu_{i,n-1} N_{i,n-1},$$

$$\Psi_{i,n} = \Psi_{i,n-1} + J\sqrt{\alpha_i P}\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right) G^*_{i,n-1}\right]$$

$$+ J\sqrt{\alpha_i P}\epsilon_{i,n-1}\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right) M^*_{i,n-1}\right]$$

$$- J^2 \alpha_i P \epsilon_{i,n-1}\mathbb{R}\left[G^*_{i,n-1} N_{i,n-1}\right],$$

$$M_{i,n} = \left(1 - J^2 \alpha_i P \mu_{i,n-1}\right) M_{i,n-1} - J^2 \alpha_i P \Psi_{i,n-1} G_{i,n-1} - J^2 \alpha_i P N_{i,n-1},$$

$$\epsilon_{i,n} = \left[\epsilon_{i,n-1}\left(\rho - \lambda J\sqrt{\alpha_i P}\mathbb{R}\left[\left(y_{i,n} - \hat{h}_{i,n-1} J\sqrt{\alpha_i P}\right) N^*_{i,n-1}\right]\right)\right]_{\epsilon_{min}}^{\epsilon_{max}}, \quad (4.5.3)$$

where $[\epsilon]_{\epsilon_{min}}^{\epsilon_{max}}$ details that $\epsilon_{min} \leq \epsilon \leq \epsilon_{max}$.

4.6 Results and Evaluation

In this section, the performance evaluation of the LMS-based averaging sum-based CE technique through computer simulations is presented. The theoretical analysis undertaken for the CSS-LMS algorithm is compared to the simulation of the algorithm. Furthermore, the three core algorithms are evaluated and compared in terms their respective CE performances. In order to achieve this, the power allocation ratios of user 1 and user 2 were set as $\alpha_1 = 0.39$ and $\alpha_2 = 0.61$ respectively. The channel coefficients between each user and the base-station were generated using Jake's channel model. The normalized Doppler is given as $f_d T = 2.4 \times 10^{-3}$. The algorithm implementation example for the convergence results is presented as Algorithm 9 in Appendix A. Other results can be generated in a similar manner as with the all-known pilots, using respective equations.

4.6.1 Performance Evaluation of the CSS-LMS Algorithm

Figure 4.6.1 and Figure 4.6.2 depict the MSEs for varying learning rates at user 1 and user 2 respectively. For both users, the analytical and the simulated MSEs exhibit a similar curve with the change in the step-sizes. The obvious disparities between the two MSEs arise from the approximations involved in the theoretical analysis. Furthermore, in the theoretical analysis, correlation of the noise components is ignored while it is accounted for in the simulated MSE. The aforementioned disparities between the analytical and the simulated are, however, very minute. They are virtually insignificant and are only visible due to the log scale used on the y-axis. The simulation therefore confirms the theoretical analysis.

In contrast, the CE performance at each user shows the significance of power allocation and how it affects CE performance. At user 2, convergence and divergence are realized at earlier step-sizes when compared to user 1. Furthermore, the improvement in CE performance at user 2 is apparent in terms of accuracy. This is depicted by the low simulated and the analytical MSEs at user 2 as opposed to the one at user 1.

The simulation MSE and the analytical MSE intersect for step-sizes $\mu = 1.4 \times 10^{-3}$ and $\mu = 9.7 \times 10^{-4}$ for user 1 and user 2 respectively. The aforementioned step-sizes are employed for evaluating the simulation and analytical MSEs in a range of SNR values at each user. Figure 4.6.3 and Figure 4.6.4 depict the performance evaluation of the CSS-LMS algorithm in relation to the theoretical analysis for user 1 and user 2 respectively. With the illustrated results, the theoretical performance is confirmed by the simulation results. This is more obvious when there is less noise relative to the signal power. User 2 is shown to have better performance

compared to user 1 in terms of CE accuracy. These performance disparities between the users can be attributed to the respective power allocation ratio. The user with the best accuracy was allocated a higher power ratio.

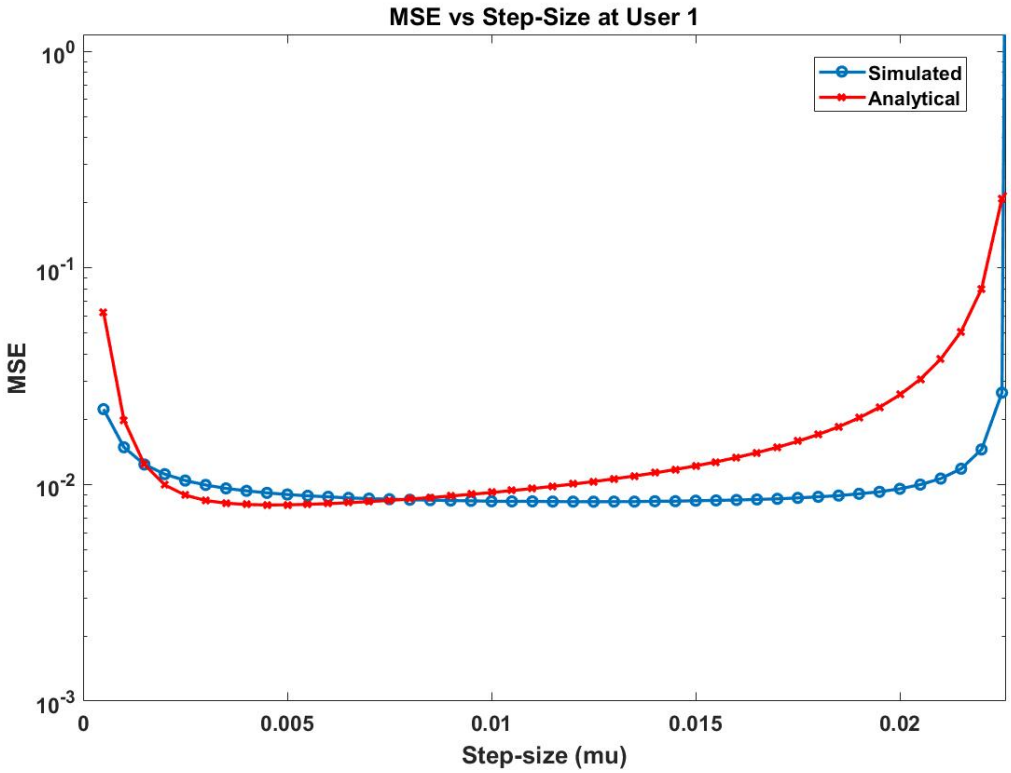


Figure 4.6. 1 User 1: Comparison of the evolution of the analytical MSE and the simulated MSE for different step-sizes.

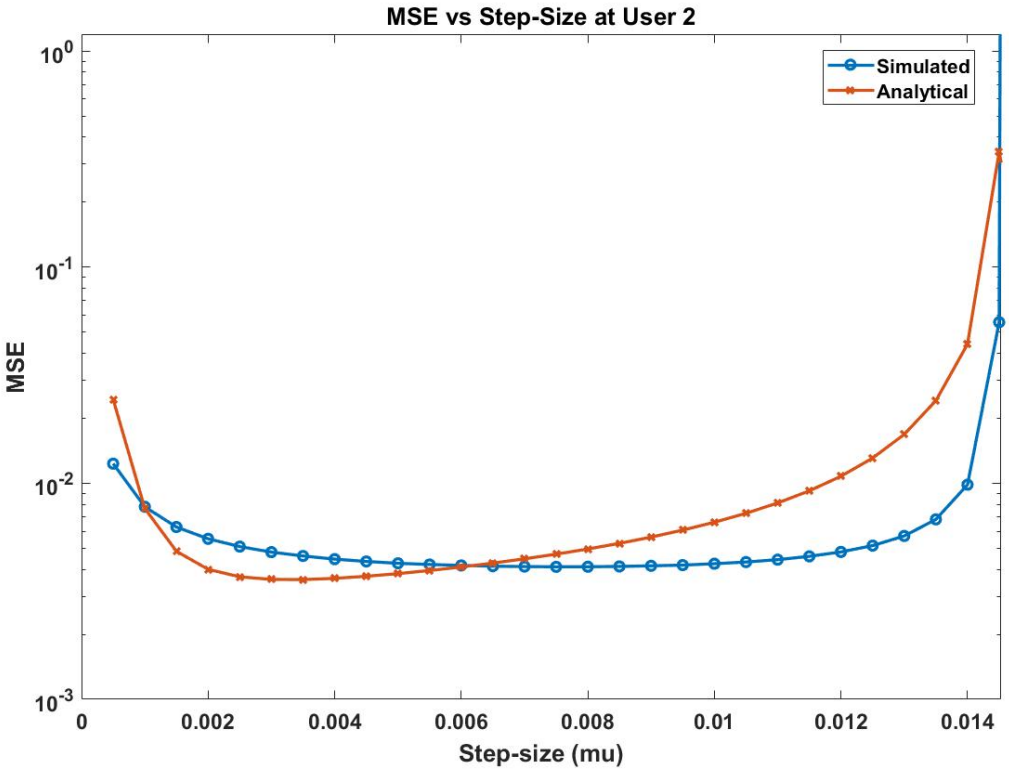


Figure 4.6. 2 User 2: Comparison of the evolution of the analytical MSE and the simulated MSE for different step-sizes.

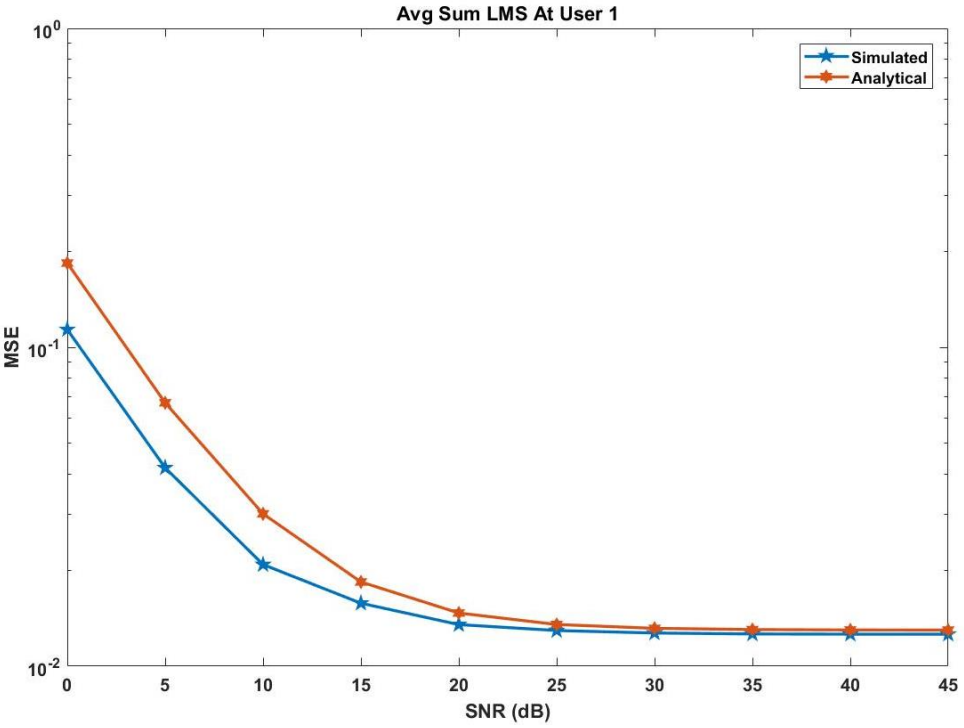


Figure 4.6. 3 User 1: Comparison between the analytical MSE and the simulated MSE using the CSS-LMS algorithm's MSE vs SNR plot.

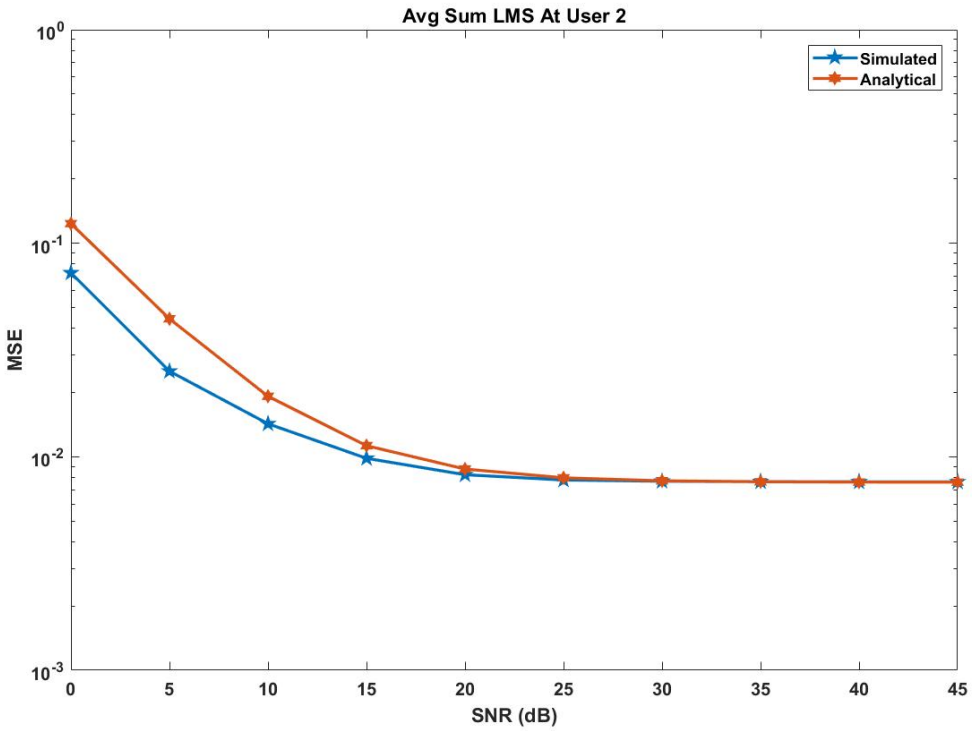


Figure 4.6. 4 User 2: Comparison between the analytical MSE and the simulated MSE using the CSS-LMS algorithm's MSE vs SNR plot.

4.6.2 Comparison of the Three LMS Algorithms

The averaging sum formulation-based CE scheme is assessed by comparing its performance when using the CSS-LMS, VSS-LMS and the VSS-LMS-AS algorithms. For the three LMS variants, optimal constant learning rates are employed. Concerning the CSS-LMS algorithm, step-sizes 7×10^{-3} and 10^{-2} are employed for user 1 and user 2 respectively. The constant step-sizes assumed for the VSS-LMS and the VSS-LMS-AS algorithms are $\epsilon = 10^{-9}$ and $\lambda = 10^{-12}$ respectively. In addition, the forgetting factor employed in the VSS-LMS-AS algorithm is given as $\rho = 0.98$. For all the three algorithms, the convergence characteristics were generated with a fixed SNR value of 30dB for both users.

Figure 4.6.5 and Figure 4.6.6 illustrate the convergence characteristics of the three LMS variants for user 1 and user 2 respectively. At both users, all three algorithms pose comparable performance in terms of CE accuracy. However, in terms of CE speed, the CSS-LMS algorithm has superior performance compared to the VSS-LMS and the VSS-LMS-AS algorithms. This is due to the self-adaption factor in the VSS-LMS and the VSS-LMS-AS algorithms, which introduces complexity. On account of the introduction of adaptive speed in the VSS-LMS-AS algorithm, it exhibits better performance as opposed to the VSS-LMS algorithm.

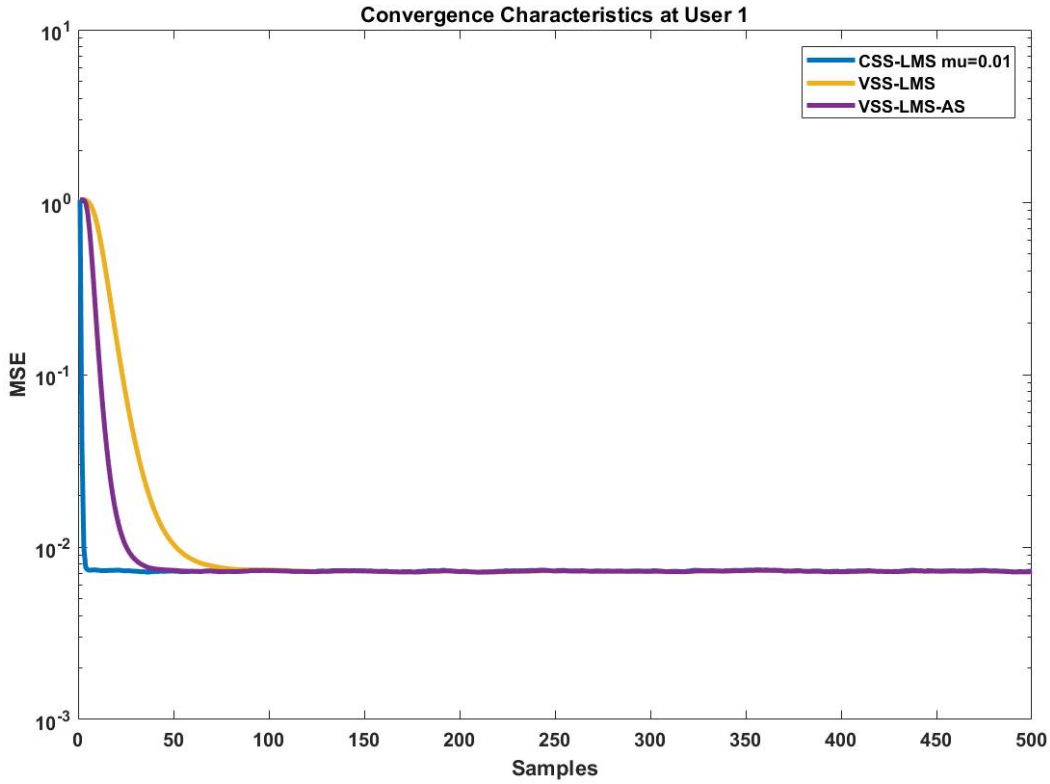


Figure 4.6. 5 User 1: MSE convergence characteristics for the three LMS algorithms.

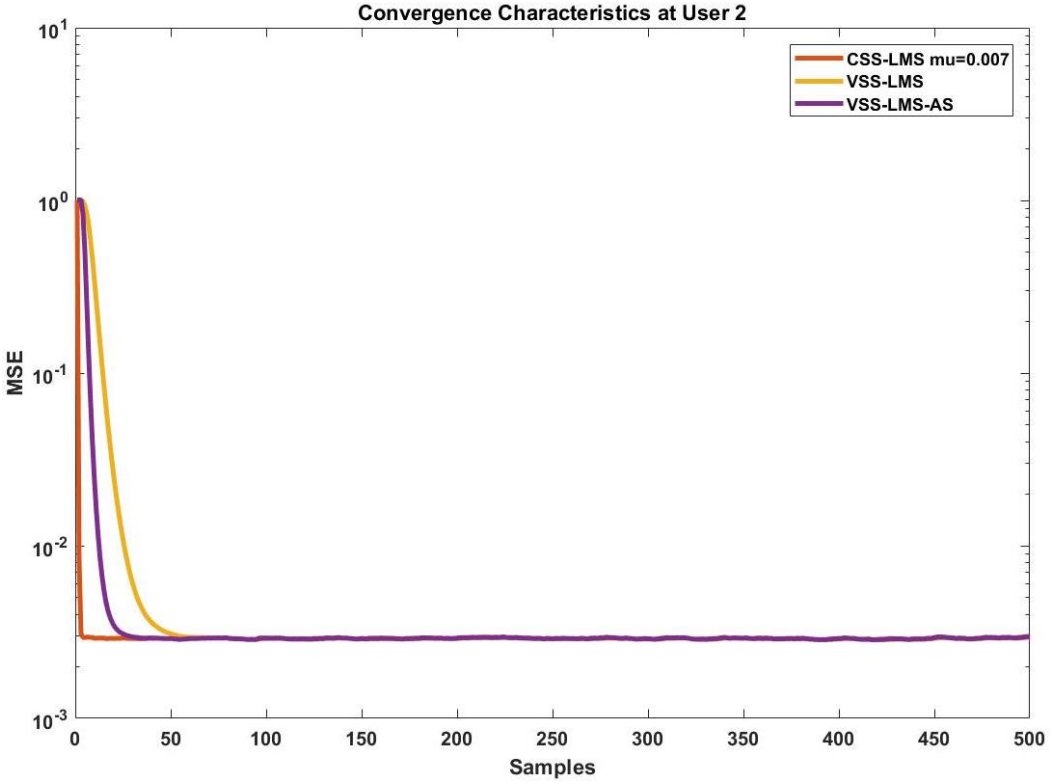


Figure 4.6. 6 User 2: MSE convergence characteristics for the three LMS algorithms

In terms of comparing the two users, user 2 is shown to exhibit the best performance in both speed and accuracy. This superior performance is attributed to the higher power ratio allocated to the afore-mentioned user. Figure 4.6.7 and Figure 4.6.8 depict the MSE versus SNR plot for the CSS-LMS, VSS-LMS and the VSS-LMS-AS algorithms at user 1 and user 2 respectively. For both users, the CSS-LMS algorithm exhibits the best MSE performance, more obviously for higher SNR conditions. These results tally with the convergence characteristics depicted in Figure 4.6.5 and Figure 4.6.6. Furthermore, as with the convergence characteristics, the VSS-LMS-AS algorithm out-performs the VSS-LMS algorithm for both users.

On the user level, CE performance is better at user 2 as opposed to user 1. This, as mentioned, is due to the power allocation ratios attributed to the respective users. The higher power allocation ratio attributed to user 2 is the main reason for the superior MSE performance. In this regard, the MSE vs. SNR results affirm the convergence characteristics.

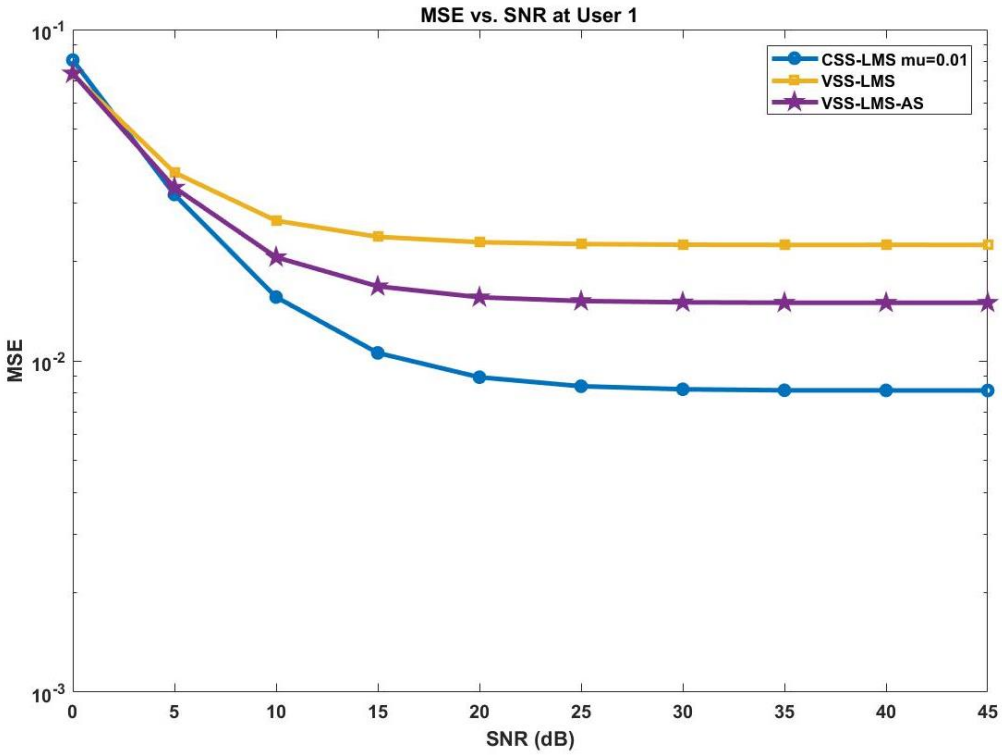


Figure 4.6. 7 User 1: MSE vs SNR curve for the three LMS algorithms.

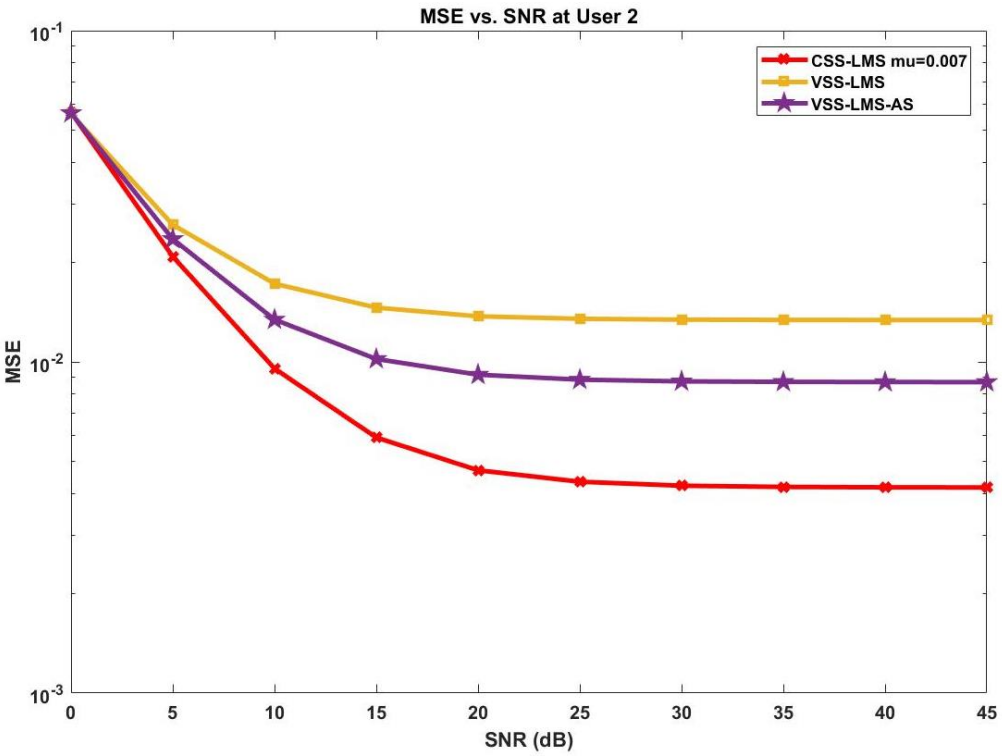


Figure 4.6. 8 User 2: MSE vs SNR curve for the three LMS algorithms.

4.7 Chapter Summary

This chapter proposed the LMS-based averaging sum CE approach for a DL SISO-NOMA system. This scheme was motivated by minimizing the pilot overhead attributed to the all-known-pilots approach. As such partial employment of pilot data for CE is considered in the LMS-based averaging sum CE scheme. Essentially, the design of this approach involves performing an averaging sum operation on the observation (received signal). The operation is done as a measure to minimize the significance of the unknown pilot symbols in the overall observation. The CE cost function is therefore constructed with accordance to the newly derived signal.

Within the chapter, the approach is first presented when the core algorithm adopted is the CSS-LMS. For the CSS-LMS algorithm, update rule is derived, and the corresponding low-pass filter is designed. Furthermore, an analytical expression for the error variance of the CSS-LMS approach is also derived. Expressions for the VSS-LMS and the VSS-LMS-AS algorithms are also introduced. These expressions in effect include the estimator update rules and the adaptive gains for both self-adaptive algorithms. Expressions for the adaptive speed in the case of the VSS-LMS-AS algorithms are also presented.

Computer simulations are employed to assess the averaging sum-based CE approach when each LMS algorithm is used. The highlight of the performance evaluation for the three algorithms aligns with the user power allocation. As deduced from the results of the three algorithms, a user with a higher power allocation ratio is attributed with superior CE performance relative to the one with a lower power allocation. This is apparent with the elevated CE performance at user 2 relative to user 1 in terms of speed of convergence and accuracy.

When evaluating the CSS-LMS algorithm, the simulation MSE performance corroborated the theoretical analysis carried out within the chapter. Comparing the three algorithms, the CSS-LMS algorithm demonstrates a leading performance relative to the self-adaptive LMS variants in terms of CE speed. This was the case for optimal step-size values. For the setting considered, the VSS-LMS-AS algorithm out-performs the VSS-LMS algorithm in terms of CE speed. Overall, channel estimation accuracy is indistinguishable between the three algorithms at either user.

CHAPTER 5

ITERATIVE AND CASE-WISE CHANNEL ESTIMATION AND DETECTION APPROACHES FOR SISO-NOMA

5.1 Introduction

Two least mean squares (LMS)-based CE approaches are advanced in this chapter. These are the iterative technique and the case-wise technique. Both approaches are of the pilot-aided CE group, with partial pilot knowledge assumed at each receiver. Applied to the two-user downlink (DL) SISO-NOMA system, the case-wise approach is designed for the stronger user (allocated a lower power). The iterative approach is transitively designed for the weaker user (allocated a higher power). The conventional LMS algorithm is employed as the core algorithm for CE.

Firstly, the case-wise approach considers two cases of the values that the unknown pilots may assume, finding the channel estimate for each case. The channel estimates are then employed in decoding the known pilots using simple detection. The results are then compared to the expected pilot and then the correct pilot is passed to the second LMS estimator for CE. Secondly, the iterative approach embeds the CSS-LMS algorithm for CE, initially making a rough estimate with the known pilot. A successive interference cancellation (SIC)-influenced detection approach is used to detect the unknown pilot. The initial channel estimate is refined in the subsequent iterations using the newly detected pilot and channel estimates.

Performance evaluations were carried out through computer simulations. Firstly, the case-wise CE and detection approach was evaluated for different probabilities of error. Secondly, the case-wise approach was evaluated against the theoretical analysis. For the iterative CE and detection approach, CE was also evaluated for different probabilities of error. Consequently, for a given probability of error, the iterative CE and detection scheme was evaluated against the theoretical analysis. For both approaches, detection error propagation was identified to have an adverse effect the CE performance. This in turn motivates for the implementation of a detection scheme with a low probability of error. For the probability of error of 0.001, the case-wise CE and detection approach was in agreement with the theoretical analysis. The iterative CE and detection approach were in agreement with the theoretical analysis for high SNR conditions.

The rest of the chapter is organized as follows: Section 5.2 presents the case-wise CE approach and presents the simulated results. Section 5.3 introduces the SIC-inspired iterative CE approach and the attributed results. Section 5.4 concludes chapter five.

5.2 Case-Wise Channel Estimation Approach

5.2.1 Description of the Case-wise CE Approach

In this section, the case-wise channel estimation approach is presented. It should be noted that the system model description assumed is given in chapter three, Section 3.2. This approach is considered for the stronger user, user 1. User 1's signal is allocated a lower power ratio in the superimposed pilot signal sent to both users. To reiterate, the following is the received signal at the receiver:

$$y_{1,n} = h_1\sqrt{\alpha_1 P}s_{1,n} + h_1\sqrt{\alpha_2 P}s_{2,n} + w_{1,n}. \tag{5.2.1}$$

At user 1, only pilot signal $s_{1,n}$ is known to the receiver, the rest of the signal is unknown. A simple SIC approach is not feasible, given that the known component of $y_{1,n}$ represents the weaker signal component in the superimposed observation. Also, the channel coefficients are unknown. Constructing a cost function for the LMS algorithm using this weak signal component would pose a high error. Figure 5.2.1 depicts the block diagram of the case-wise CE approach, considering a scenario whereby one symbol is sent for each observation $y_{1,n}$:

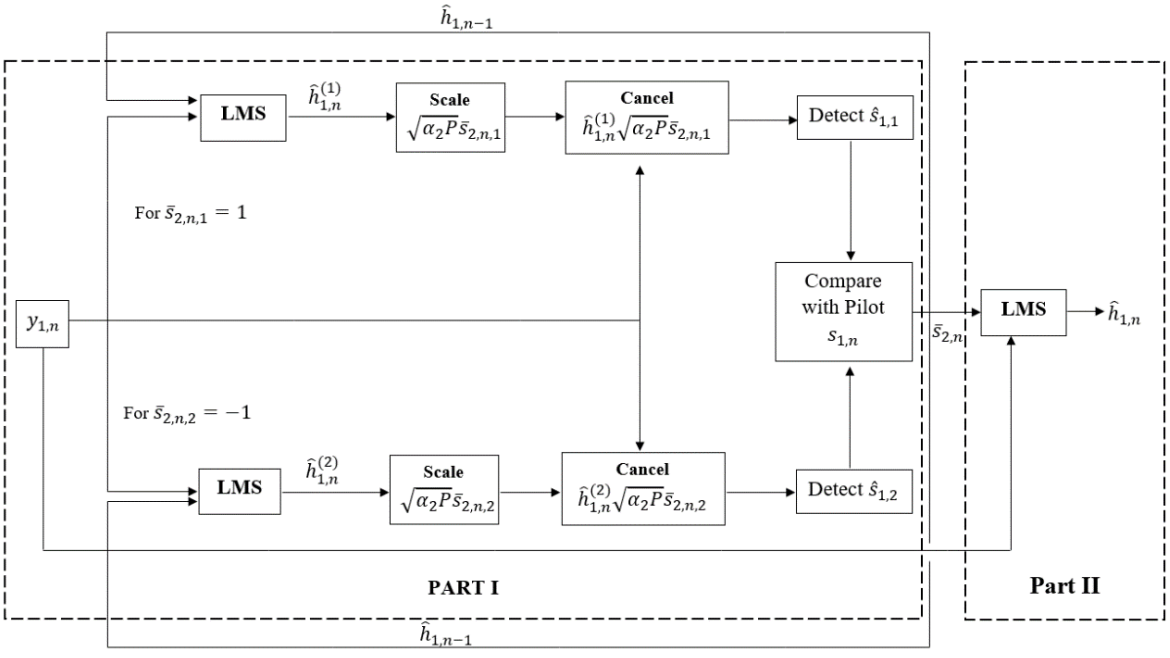


Figure 5.2. 1 LMS pilot aided channel estimation approach at User 1.

In Figure 5.2.1, the area labelled Part I is concerned with identifying the correct assumption made on the unknown signal, $s_{2,n}$. Estimation of the channel coefficients between user 1 and the BS is achieved by the area labelled Part II. On reception of signal $y_{1,n}$, the unknown pilot is assumed to be $+1$ and -1 on each branch. The assumptions are denoted $\bar{s}_{2,n,1}$ and $\bar{s}_{2,n,2}$ for user 1 and user 2 respectively. The received signal is used as the observation and input to the LMS block on each branch.

For each branch, the channel estimate is scaled and the component that contains the unknown pilot $s_{2,n}$ is cancelled out. The resultant signal $g_{1,k}$ is passed on to the detection block, where an estimate $\hat{s}_{1,k}$ is made (where *k* denotes the branch taken). After the detection block, each pilot estimate $\hat{s}_{1,k}$ is passed on to the comparison. The operations undertaken in the comparison block are described in Table 5.2.1.

Table 5.2.1 Comparison Block: LMS pilot aided channel estimation approach at User 1

Input: $\hat{s}_{1,1}, \hat{s}_{1,2}$ **Output:** $\bar{s}_{2,n}$ and correct estimate $\hat{h}_{1,n}$.

	Comparison of $\bar{s}_{1,1}$	**Comparison of $\bar{s}_{1,2}$**	**Decision**	**Correct Estimate**
Condition 1	$\hat{s}_{1,1} = s_{1,n}$	$\hat{s}_{1,2} \neq s_{1,n}$	$\bar{s}_{2,n,1}$ is the correct assumption	$\hat{h}_{1,n}$
Condition 2	$\hat{s}_{1,1} \neq s_{1,n}$	$\hat{s}_{1,2} = s_{1,n}$	$\bar{s}_{2,n,2}$ is the correct assumption	$-\hat{h}_{1,n}$
Condition 3	$\hat{s}_{1,1} \neq s_{1,n}$	$\hat{s}_{1,2} \neq s_{1,n}$	**if** $\lvert g_{1,1} - s_{1,n}\rvert > \lvert g_{1,2} - s_{1,n}\rvert$ then $\bar{s}_{2,n,2}$ is the correct assumption	$-\hat{h}_{1,n}$
Condition 4	$\hat{s}_{1,1} = s_{1,n}$	$\hat{s}_{1,2} = s_{1,n}$	**if** $\lvert g_{1,1} - s_{1,n}\rvert < \lvert g_{1,2} - s_{1,n}\rvert$ then $\bar{s}_{2,n,1}$ is the correct assumption	$\hat{h}_{1,n}$

The pilot estimates for each branch are compared to the known pilot $s_{1,n}$ as shown in Table 5.2.1. Various scenarios of the detected pilots in relation to the estimate of each branch are considered. The pilot estimate that is closest to the known pilot therefore implies the correct assumption in terms of pilot $s_{2,n}$. The correct assumption is denoted $\bar{s}_{2,n}$, and is output to the comparison block. Another output of the comparison block is the correct channel estimate,

$\hat{h}_{1,n}$, which is fed back for the next symbol at $n+1$. The $\bar{s}_{2,n}$ pilot is then used in the construction of the cost function for the second LMS block. Table 5.2.2 is outline of the operations involved in each block for every symbol received. The analysis is given for each branch:

Table 5.2.2: Block-level Operations for Each Branch

$\bar{s}_{2,n,1} = 1$ **Branch**	$\bar{s}_{2,n,2} = -1$ **Branch**
LMS Block *Input:* $y_{1,n}$ *with* $\bar{s}_{2,n,1} = 1$ *Output:* $\hat{h}_{1,n}^{(1)}$	**LMS Block** *Input:* $y_{1,n}$ *with* $\bar{s}_{2,n,2} = -1$ *Output:* $\hat{h}_{1,n}^{(2)}$
Scale Block *Input:* $\hat{h}_{1,n}^{(1)}$ *Output:* $\hat{h}_{1,n}^{(1)}\sqrt{\alpha_2 P}\bar{s}_{2,n,1}$	**Scale Block** *Input:* $\hat{h}_{1,n}^{(2)}$ *Output:* $\hat{h}_{1,n}^{(2)}\sqrt{\alpha_2 P}\bar{s}_{2,n,1}$
Cancellation Block *Input:* $\hat{h}_{1,n}^{(1)}\sqrt{\alpha_2 P}\bar{s}_{2,n,1}$, $y_{1,n}$ *Process:* $g_{1,1} = y_{1,n} - \hat{h}_{1,n}^{(1)}\sqrt{\alpha_2 P}\bar{s}_{2,n,1}$ $= h_1\sqrt{\alpha_1 P}s_{1,n} + h_1\sqrt{\alpha_2 P}s_{2,n} + w_{1,n} - \hat{h}_{1,n}^{(1)}\sqrt{\alpha_2 P}\bar{s}_{2,n,1}$ $\therefore$ $g_{1,1} = h_1\sqrt{\alpha_1 P}s_{1,n} + w'_{1,1}$ *where* $w'_{1,1} = \Delta_{h_1}\sqrt{\alpha_2 P}\Delta_{s_{2,p,1}} + w_{1,n}$, $\Delta_{h_1} = h_1 - \hat{h}_{1,n}^{(1)}$ and $\Delta_{s_{2,n,1}} = s_{2,n} - \bar{s}_{2,n,1}$ *Output:* $g_{1,1}$	**Cancellation Block** *Input:* $\hat{h}_{1,n}^{(2)}\sqrt{\alpha_2 P}\bar{s}_{2,n,2}$, $y_{1,n}$ *Process:* $g_{1,2} = y_{1,p} - \hat{h}_{1,n}^{(2)}\sqrt{\alpha_2 P}\bar{s}_{2,n,2}$ $= h_1\sqrt{\alpha_1 P}s_{1,n} + h_1\sqrt{\alpha_2 P}s_{2,n} + w_{1,n} - \hat{h}_{1,n}^{(2)}\sqrt{\alpha_2 P}\bar{s}_{2,n,2}$ $\therefore$ $g_{1,2} = h_1\sqrt{\alpha_1 P}s_{1,n} + w'_{1,2}$ *where* $w'_{1,2} = \Delta_{h_1}\sqrt{\alpha_2 P}\Delta_{s_{2,n,2}} + w_{1,n}$, $\Delta_{h_1} = h_1 - \hat{h}_{1,n}^{(2)}$ and $\Delta_{s_{2,n,2}} = s_{2,n} - \bar{s}_{2,n,2}$ *Output:* $g_{1,2}$

Detection Block: *Input:* $g_{1,1}$ *Process:* Simple detection *Output:* $\hat{s}_{1,1}$	**Detection Block:** *Input:* $g_{1,2}$ *Process:* Simple detection *Output:* $\hat{s}_{1,2}$

5.2.1.1 Channel Estimation Block: CSS-LMS Algorithm in Part I (Adapted from [22])

This subsection provides a "white-box" perspective of the LMS block in Part I of Figure 5.2.1. The input to the LMS block is the observation given in (5.2.1). Given the corresponding assumption made on branch k, the following is adopted as the cost-function:

$$V_1^{(k)}\left(\hat{h}_{1,n}^{(k)}\right) = E\left\{\left|y_{1,n} - \hat{h}_{1,n}^{(k)}\sqrt{\alpha_2 P}\bar{s}_{2,n,k}\right|^2\right\}. \tag{5.2.2}$$

Essentially, the LMS algorithm employed for the case-wise CE approach follows the same procedure outlined in chapter three, Section 3.3. The same procedure was used in chapter four, in Section 4.3. In adopting this algorithm, the following is the resulting update equation:

$$\hat{h}_{1,n}^{(k)} = \hat{h}_{1,n-1}^{(k)} + \mu\sqrt{\alpha_2 P}\bar{s}_{2,n,k}\left(y_{1,n} - \hat{h}_{1,n-1}^{(k)}\sqrt{\alpha_2 P}\bar{s}_{2,n,k}\right). \tag{5.2.3}$$

The update rule given by (5.2.3) serves as the channel estimator for the LMS block portrayed in Part I of Figure 5.2.1 and thus, summarizes the CE block for each branch.

5.2.2 Channel Estimation Block: CSS-LMS Algorithm in Part II (Adapted from [22])

This subsection provides a "white-box" perspective of the LMS blocks in Part II of Figure 5.2.1. The input to the LMS block is the observation given in (5.2.1) as well as the correct assumption $\bar{s}_{2,n}$. With the received signal $y_{1,n}$ and $\bar{s}_{2,n}$, the cost function is constructed as follows:

$$V_1\left(\hat{h}_{1,n}\right) = E\left\{\left|y_{1,n} - \hat{h}_{1,n}\bar{x}_n\right|^2\right\}. \tag{5.2.4}$$

where $\bar{x}_n = \sqrt{\alpha_2 P}s_{1,n} + \sqrt{\alpha_2 P}\bar{s}_{2,n}$ and $E\{|\bar{x}_n|^2\} = P$.

Following the procedure outlined in chapter three, Section 3.3. In adopting this algorithm, the following is the resulting update equation:

$$\hat{h}_{1,n} = \hat{h}_{1,n-1} + \mu\bar{x}_n\left(y_{1,n} - \hat{h}_{1,n-1}\bar{x}_n\right). \tag{5.2.5}$$

The update rule given by (5.2.5) serves as the channel estimator for the LMS block portrayed in Part II of Figure 5.2.1 and thus, summarizes the second CE block.

5.2.2.1 Error Variance Analysis

This sub-section presents the error variance analysis for the advanced CSS-LMS algorithm. Referring to Section 3.3.1 in chapter three, the same method is applied in this subsection. Thus, without having to reiterate each step, the following is the low-pass filter design for the update rule given in (5.2.6) in the z-domain:

$$L(z) = \frac{\mu \bar{x}_n}{1 - z^{-1}(1 - \mu P)}, \tag{5.2.6}$$

where $L(z)$ is the low pass filter with cut-off frequency chosen as:

$$f_c T = \frac{\mu P}{2\pi(1 - \mu P)}.$$

It should be noted that for any symbol n, $\bar{x}_n$ is fully known. Therefore, it is treated as a constant but kept as $\bar{x}_n$ for presentation. Following the referred to approach, the channel estimation error can be expressed in the z-domain as:

$$\begin{aligned} e_1(z) &= h_1(z) - L(z)\big(\, h_1(z)\bar{x}_n + W_1(z)\big) \\ &= (1 - L(z)\bar{x}_n)h_1(z) - L(z)W_1(z). \end{aligned} \tag{5.2.7}$$

From the error given in (5.2.7), the error variance is expressed in the frequency domain as follows:

$$\sigma_{e_1}^2 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - L\left(e^{2j\pi fT}\right)\bar{x}_n\right|^2 \Gamma_f(f)\, df + \sigma_{w_1}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(e^{2j\pi fT}\right)\right|^2 df, \tag{5.2.8}$$

such that the variance can be separated and evaluated disjointly as:

$$MSE1_1 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - L\left(e^{2j\pi fT}\right)\bar{x}_n\right|^2 \Gamma_{h_1}(f)\, df,$$

$$MSE2_1 = \sigma_{w_1}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df. \tag{5.2.9}$$

To advance the expressions of each term, it should be recalled that the form of the equation in (3.3.10) in Section 3.3.1 of chapter three has a similar form as $MSE1_1$ in (5.2.9). Similarly, the

same can be concluded for equation (3.3.11) in chapter three in contrast with $MSE2_1$ in (5.2.9). In this light, both terms of the asymptotic variance can undergo the same derivations as their counterparts. As such, the following are the final expressions of the terms:

$$MSE1_1 = \sigma_{h_1}^2 \frac{(2\pi f_d T)^2}{2(\mu P)^2}(1-\mu P)^2, \tag{5.2.10}$$

$$MSE2_1 = \sigma_{w_1}^2 \frac{\mu}{(2-\mu P)}. \tag{5.2.11}$$

Equations (5.2.10) and (5.2.11) are summed up together to give the final expression for the error variance which is expressed as:

$$\sigma_{e_1}^2 = \sigma_{h_1}^2 \frac{(2\pi f_d T)^2}{2(\mu P)^2}(1-\mu P)^2 \quad + \sigma_{w_1}^2 \frac{\mu}{(2-\mu P)}. \tag{5.2.12}$$

5.2.3 Results and Evaluation

This subsection presents the performance evaluation of the proposed CSS-LMS based case-wise CE approach. The MSE is employed as the metric, with accuracy and speed as the performance goals. For the simulation setup, power allocations are set as $\alpha_1 = 0.39$ and $\alpha_2 = 0.61$ for user 1 and user 2 respectively. The generation of the channel coefficients between each respective user and the base-station comply with Jake's Doppler spectrum. The corresponding normalized Doppler frequency set as $f_d T = 5.5 \times 10^{-5}$.

5.2.3.1 Probability of Error Propagation on CE Performance

In this subsection, CE performance is assessed for different probabilities of error. The probability of error considered is concerned with the detection block on each branch in part one of Figure 5.2.1. In addition to the set-up described in the previous subsection, the learning rates considered are taken as $\mu = 0.005, 0.05, 0.1$ and 0.26. Furthermore, the SNR is set to 30dB. The algorithm implementation example for the convergence results is presented in Appendix A, in Algorithm 10. Other results can be generated in a similar manner as with the all-known pilots, using respective equations.

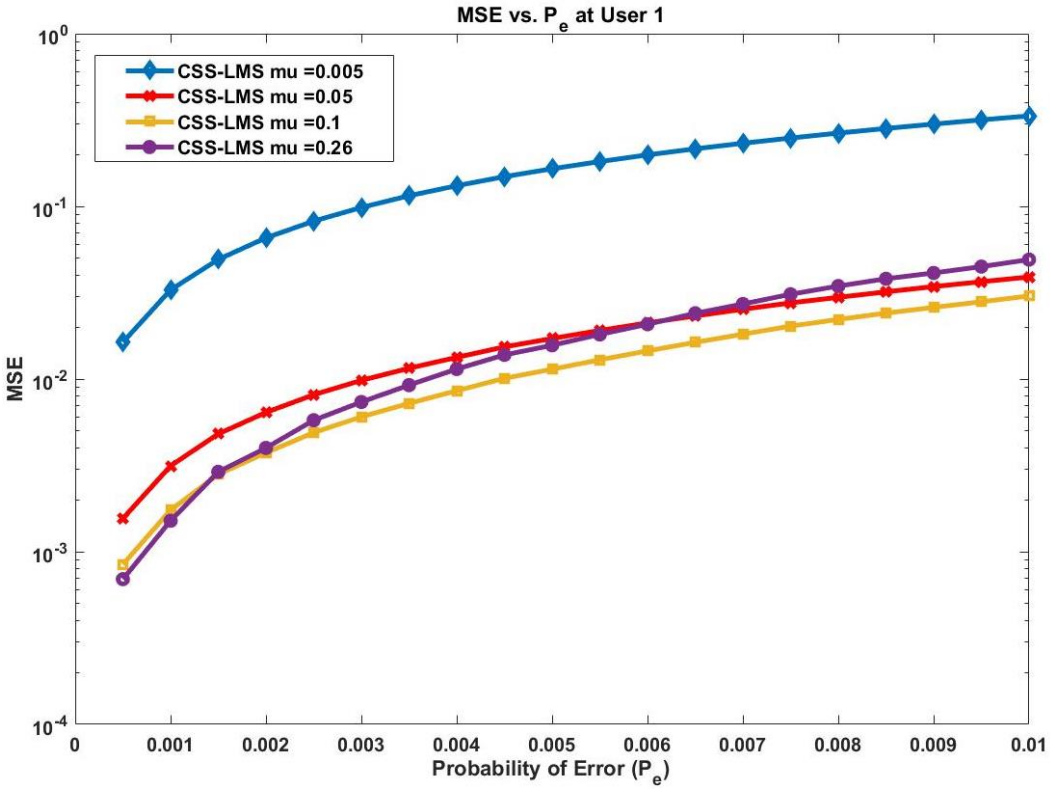


Figure 5.2. 2 User 1: MSE vs. Probability of Error Curve for different step-sizes.

Figure 5.2.2 depicts the change in the MSE for a range of probabilities of error, employing different step-sizes. The performance in terms of the MSE increases with increase of the values of the step-sizes. This is with the exception of the step-size of 0.26, which essentially shows deviation in performance when compared with other step-size performances. The main reason for this trend is that, sufficiently higher learning rates produce better performance when compared to lower step-size values. On the other hand, the learning rate that is beyond the maximum allowed step-size results in the curve concerned with the step-size of 0.26.

For each step-size, the MSE performance degrades with the increase in probability of error. The higher the value of the probability of error, the more bits are expected to be in error in the detected pilot. Due to detected pilots in error, a wrong branch in Figure 5.2.1 is assumed, and this error propagates to the second LMS CE block. As such, the CE performance is affected.

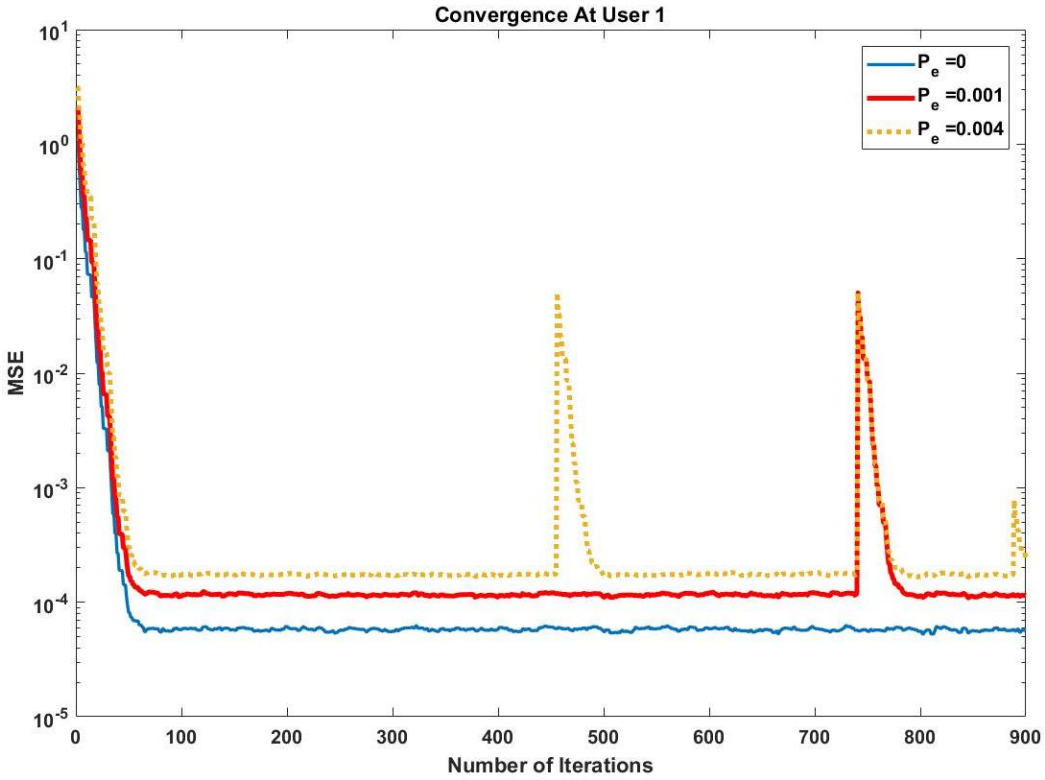


Figure 5.2. 3 User 1: Convergence characteristics for different Pe's.

Figure 5.2.3 shows the MSE convergence characteristics for a perfect detection case and considering the following probabilities of error: $P_e = 0.001$ and $P_e = 0.004$. When using the probability of error of 0.001, superior CE performance is exhibited relative to when using 0.004. This result concurs with the results given in Figure 5.2.2 in terms of performance for different probability of error. On the other hand, assuming perfect detection means there is no error propagation to affect the CE performance. As such, when the probability of error is 0, the best performance is achieved.

The spikes that appear for the convergence curves for $P_e = 0.001$ and $P_e = 0.004$ are results of the propagation of the detection error from Part I of Figure 5.2.1. In equation (5.2.5), the superimposed signal $\bar{x}_n$ has component $\bar{s}_{2,n}$ that is in error and thus resulting in the spikes shown in Figure 5.2.3. The frequency of the spikes increases with the increase as P_e increases.

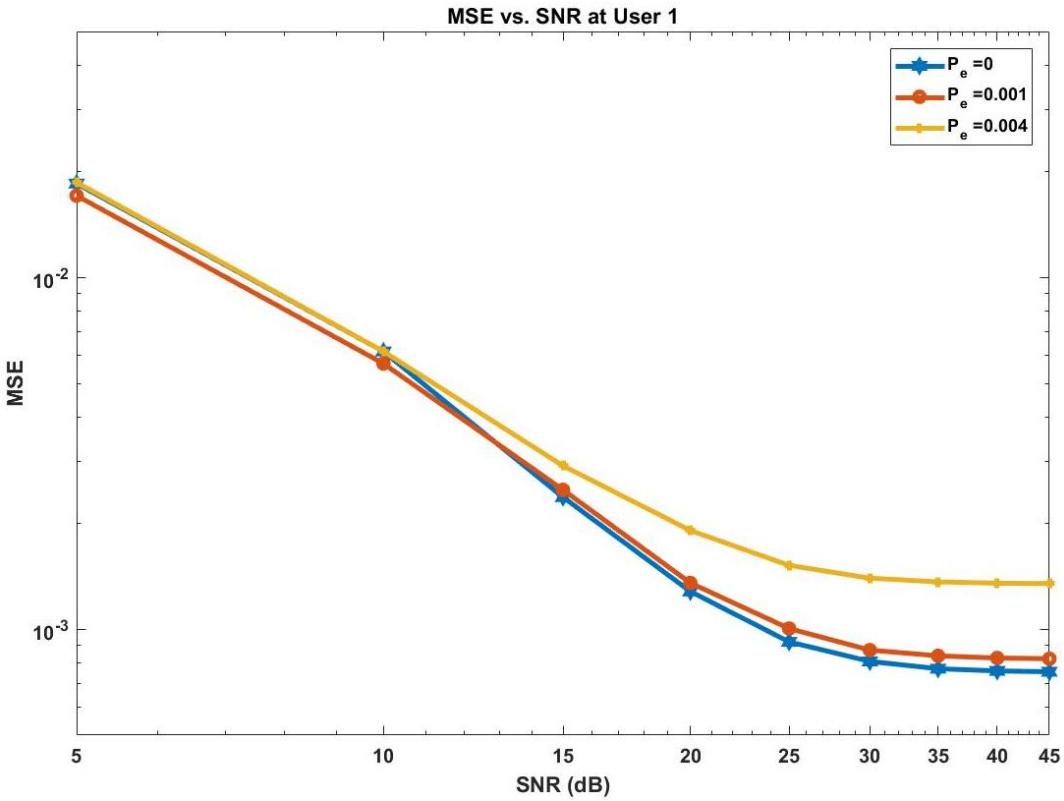


Figure 5.2. 4 User 1: MSE vs. SNR curve for different Pe's.

Similar observations in terms of CE performance when different $P_e{}'$s are use can be made in Figure 5.2.4, which shows the change in MSE for different signal-to-noise-ratios. From Figure 5.2.2, Figure 5.2.3 and Figure 5.2.4, the probability of error of 0.001 is selected as the reasonable probability of error for evaluating the CE performance in the rest of this section. This probability of error is selected because it is easily achievable when a practical detector is considered. Furthermore, the CE performance achieved when using $P_e = 0.001$ is comparable to when perfect pilot detection is considered.

5.2.3.2 Channel Estimation Analytical Performance Evaluation

This subsection evaluates the CE performance of the CSS-LMS algorithm in relation to the analysis carried out in subsection 5.2.2.1. In this light, the probability of error in detection before the comparison block in Figure 5.2.1 is taken as 0.001.

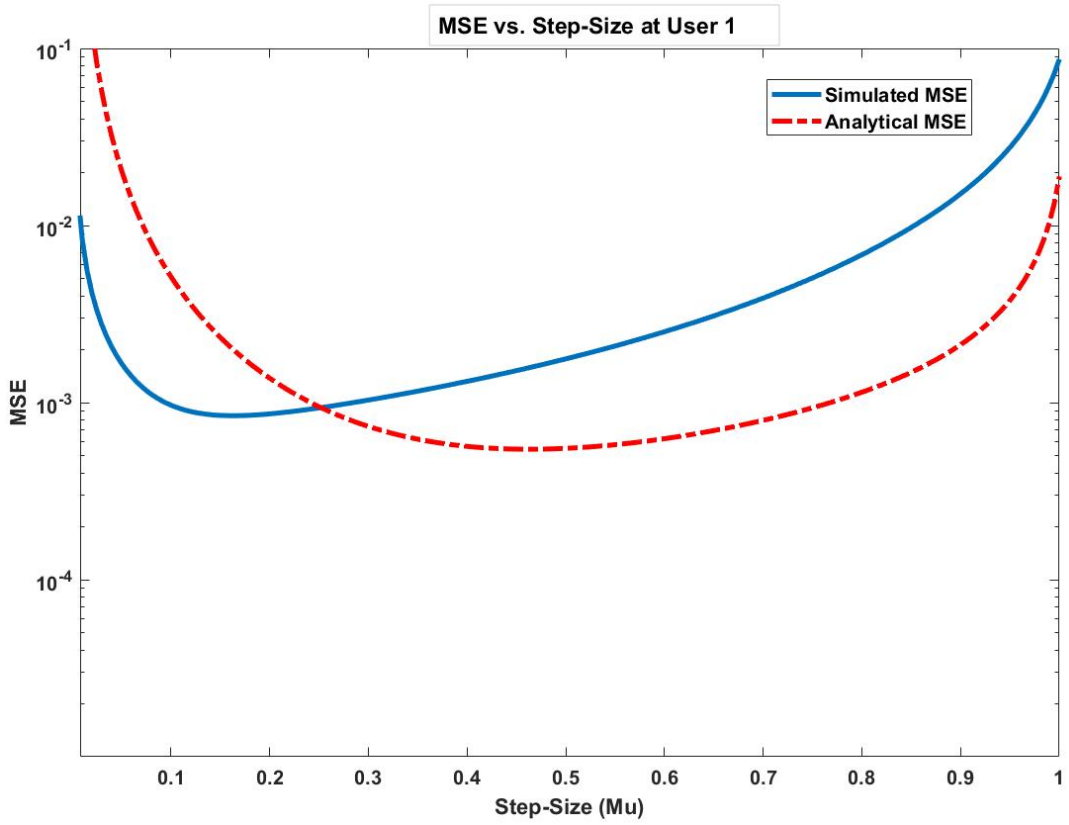


Figure 5.2. 5 User 1: MSE vs. Step-size Curve for the CSS-LMS Algorithm at Pe = 0.0001.

In Figure 5.2.5, the change in the MSE for different step-sizes is depicted for both the analytical and simulation MSE. The convergence range of step-sizes for both the analytical and the simulated MSEs are similar. Furthermore, the MSE performance achieved is virtually similar in terms of CE accuracy. The differences between the simulation and the analytical MSEs are insignificant, given they are in very low orders only visible through a log scale. The theoretical analysis is therefore confirmed by simulation.

For the step-size value of $\mu = 0.252$, the analytical and the simulation MSE have identical performance. The analytical and simulation performances of the CSS-LMS algorithm are therefore benchmarked using the learning rate of 0.25 in the MSE vs. SNR plot in Figure 5.2.6. Figure 5.2.6 compares the analytical MSE and the simulation MSE for a range of SNR values. For the entire range of SNR values considered, the analytical and simulations MSE have very little disparities. As such, the simulation confirms the analysis carried out in subsection 5.2.2.1.

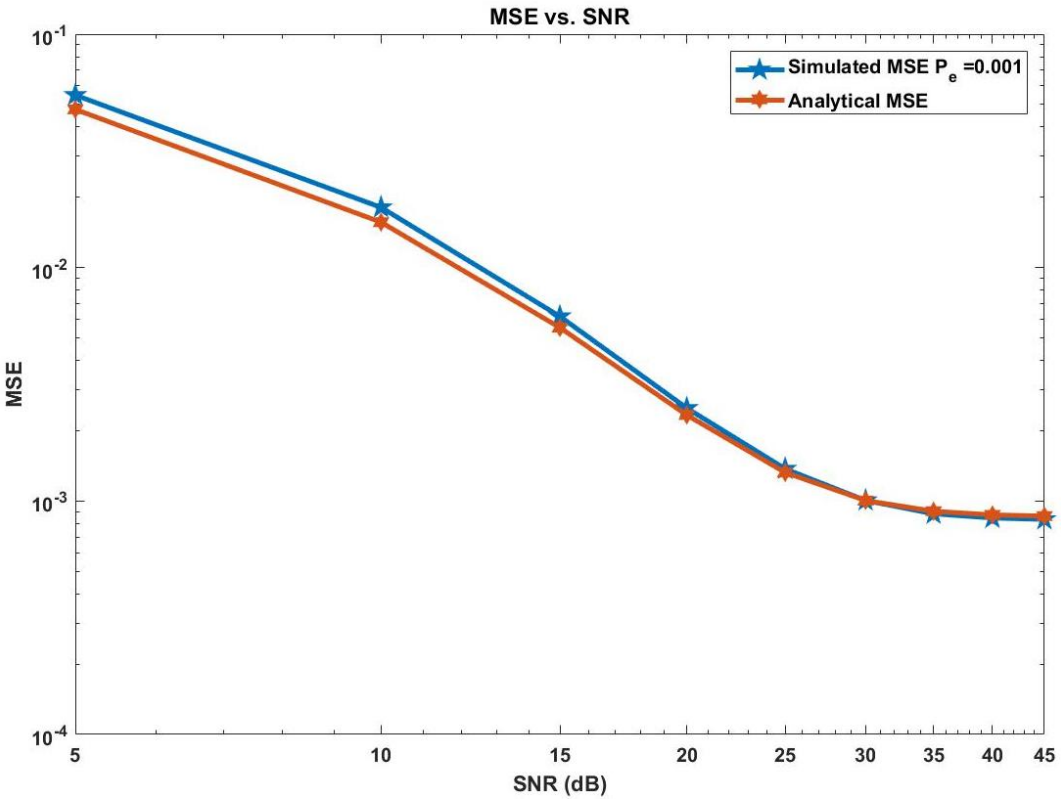


Figure 5.2. 6 User 1: MSE vs SNR Curve for the CSS-LMS Algorithm at Pe = 0.001.

5.3 LMS-based Iterative Channel Estimation Approach

5.3.1 Description of the Iterative CE Approach

In this subsection, the data model (adopted from chapter three) is defined for the LMS-based iterative CE technique. On defining the adopted model, the iterative scheme is advanced for user 2, considered to be the weaker user (with a stronger signal component in the observation). At user 2, the received pilot signal is modelled as shown below:

$$y_{2,n} = h_2\sqrt{\alpha_1 P}s_{1,n} + h_2\sqrt{\alpha_2 P}s_{2,n} + w_{2,n}. \quad (5.3.1)$$

From the observation $y_{2,n}$ received at user 2, only the pilot symbol $s_{2,n}$ is known. Now given the channel response is unknown, simply decoding $s_{2,n}$ and treating the rest of the signal as noise is not a viable approach. In order to estimate the channel, it is therefore essential to recover the known as well as the unknown pilots. The method proposed and illustrated in Figure 5.3.1 is an iterative successive cancellation approach with which the known pilot is used in providing the corresponding channel estimate. Respective channel estimates are refined in each iteration. Simultaneously, on each iteration, the unknown pilot is estimated. The block diagram below showcases the overall process of the iterative approach undertaken at *user 2.*

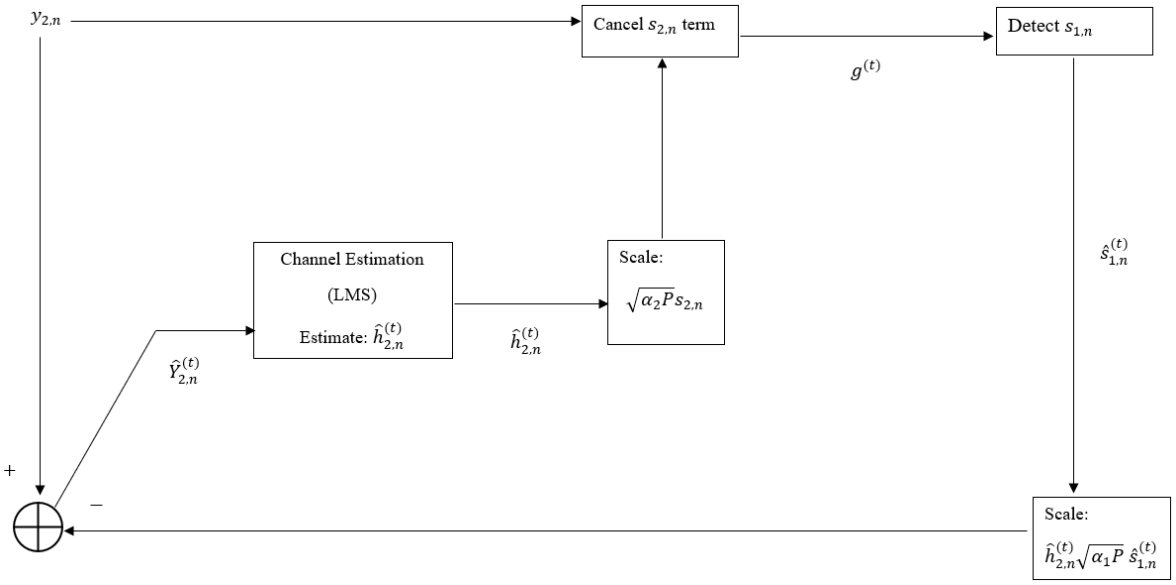


Figure 5.3. 1 Iterative pilot aided channel estimation approach at User 2.

This section constitutes an outline of the process depicted in Figure 5.3.1. The initial input to the channel estimation block is the signal received at the user.

$$Y_{2,n}^{(1)} = y_{2,n}. \tag{5.3.2}$$

In the following, each iteration is thoroughly examined, covering the inputs and outputs at each block. This is done for enough iterations to allow for generalization of the input to the LMS at iteration $t = i$.

Iteration 1 (t=1):

LMS Input: $Y_{2,n}^{(1)}$.

LMS Output: $\hat{h}_{2,n}^{(1)} = h_2 + \Delta_h^{(1)}$.

$s_{1,p}$ Detection Input: $g_2^{(1)} = h_2\sqrt{\alpha_1 P}s_{1,n} + w_2^{(1)\prime}$,

$$\text{where: } w_2^{(1)\prime} = h_2\sqrt{\alpha_2 P}s_{2,n} - \hat{h}_2^{(1)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n},$$

$$w_2^{(1)\prime} = \Delta_h^{(1)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n}.$$

Final Output to be fed back:

$$Y_{2,n}^{(2)} = y_{2,n} - \hat{h}_{2,n}^{(1)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(1)},$$

$$Y_{2,n}^{(2)} = h_2\sqrt{\alpha_1 P}s_{1,n} + h_2\sqrt{\alpha_2 P}s_{2,n} + w_{2,n} - \hat{h}_{2,n}^{(1)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(1)},$$

$$Y_{2,n}^{(2)} = h_2\sqrt{\alpha_2 P}s_{2,n} + h_2\sqrt{\alpha_1 P}\Delta_{s_1}^{(1)} - \Delta_h^{(1)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(1)} + w_{2,n}.$$

Iteration 2 (t=2):

LMS Input: $Y_{2,n}^{(2)}$.

LMS Output: $\hat{h}_{2,n}^{(2)} = h_2 + \Delta_h^{(2)}$.

$\boldsymbol{s_{1,p}}$ **Detection Input**: $g_2^{(2)} = h_2\sqrt{\alpha_1 P}s_{1,n} + w_2^{(2)\prime}$,

where: $w_2^{(2)\prime} = h_2\sqrt{\alpha_2 P}s_{2,n} - \hat{h}_{2,n}^{(2)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n}$,

$$w_2^{(2)\prime} = \Delta_h^{(2)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n}.$$

Final Output to be fed back:

$$Y_{2,n}^{(3)} = y_{2,n} - \hat{h}_{2,n}^{(2)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(2)},$$

$$Y_{2,n}^{(3)} = h_2\sqrt{\alpha_1 P}s_{1,n} + h_2\sqrt{\alpha_2 P}s_{2,n} + w_{2,n} - \hat{h}_{2,n}^{(2)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(2)},$$

$$Y_{2,n}^{(3)} = h_2\sqrt{\alpha_2 P}s_{2,n} + h_2\sqrt{\alpha_1 P}\Delta_{s_1}^{(2)} - \Delta_h^{(2)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(2)} + w_{2,n}.$$

Iteration 3 (t=3):

LMS Input: $Y_{2,n}^{(3)}$.

LMS Output: $\hat{h}_{2,n}^{(3)} = h_2 + \Delta_h^{(3)}$.

$\boldsymbol{s_{1,p}}$ **Detection Input**: $g_2^{(3)} = h_2\sqrt{\alpha_1 P}s_{1,n} + w_2^{(3)\prime}$,

where: $w_2^{(3)\prime} = h_2\sqrt{\alpha_2 P}s_{2,n} - \hat{h}_{2,n}^{(3)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n}$,

$$w_2^{(3)\prime} = \Delta_h^{(3)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n}.$$

Final Output to be fed back:

$$Y_{2,n}^{(4)} = y_{2,n} - \hat{h}_{2,n}^{(3)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(3)},$$

$$Y_{2,n}^{(4)} = h_2\sqrt{\alpha_1 P}s_{1,n} + h_2\sqrt{\alpha_2 P}s_{2,n} + w_{2,n} - \hat{h}_{2,n}^{(3)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(3)},$$

$$Y_{2,n}^{(4)} = h_2\sqrt{\alpha_2 P}s_{2,n} + h_2\sqrt{\alpha_1 P}\Delta_{s_1}^{(3)} - \Delta_h^{(3)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(3)} + w_{2,n}.$$

Iteration i (t=i):

LMS Input: $Y_{2,n}^{(i)} = h_2\sqrt{\alpha_2 P}s_{2,n} + h_2\sqrt{\alpha_1 P}\Delta_{s_1}^{(i-1)} - \Delta_h^{(i-1)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(i-1)} + w_{2,n}.$ (5.3.3)

LMS Output: $\hat{h}_{2,n}^{(i)} = h_2 + \Delta_h^{(i)}.$

$\boldsymbol{s_{1,p}}$ **Detection Input**: $g_2^{(i)} = h_2\sqrt{\alpha_1 P}s_{1,n} + w_2^{(i)\prime},$

where: $w_2^{(i)\prime} = h_2\sqrt{\alpha_2 P}s_{2,p} - \hat{h}_{2,n}^{(i)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n},$

$$w_2^{(i)\prime} = \Delta_h^{(i)}\sqrt{\alpha_2 P}s_{2,n} + w_{2,n}.$$

Final Output to be fed back:

$$Y_{2,n}^{(i+1)} = y_{2,n} - \hat{h}_{2,n}^{(i)}\sqrt{\alpha_1 P}\hat{s}_{1,n},$$

$$Y_{2,n}^{(i+1)} = h_2\sqrt{\alpha_1 P}s_{1,n} + h_2\sqrt{\alpha_2 P}s_{2,n} + w_{2,n} - \hat{h}_{2,n}^{(i)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(i)},$$

$$Y_{2,n}^{(i+1)} = h_2\sqrt{\alpha_2 P}s_{2,n} + h_2\sqrt{\alpha_1 P}\Delta_{s_1}^{(i)} - \Delta_h^{(i)}\sqrt{\alpha_1 P}\hat{s}_{1,n}^{(i)} + w_{2,n}.$$

From the equations outlined in each iteration, the uniformity of the inputs and outputs of each block become apparent from the second iteration to the fourth. As a result, these can be generalized for a later iteration *i*. However, the first iteration is treated differently, owing to the fact that it has some disparities from other iterations. The realization of these block-level equations allows for an initial analysis and a generalized analysis on the LMS block, which is the CE block. Channel estimation using the conventional least mean squares algorithm (coined as the CSS-LMS) is explored in the ensuing chapter sub-section.

5.3.2 Channel Estimation Block: CSS-LMS Algorithm (Adapted from [22])

In this sub-section, a "white-box" perspective of the LMS block in Figure 5.3.1 is presented. Essentially, the LMS algorithm employed for CE is equivalent to the conventional one employed in Section 3.3 and Section 4.3 in chapters three and four respectively. In a typical iteration, the input to the LMS block is an aggregate of the observation and the feedback after detection of the unknown pilot has been carried out. On the initial iteration, given there is no feedback, only the observation is input to the LMS block. The invariability of the form of the input to the LMS block from the second iteration onwards lead to a generalization, as mentioned in the previous section. From this perspective, the LMS algorithm will be considered for the observation given by (5.3.1), as well as the generalized feedback-observation sum given in (5.3.3).

5.3.2.1 First Iteration

This section expounds the CSS-LMS during the first iteration. The following is the selected cost function:

$$V\left(\hat{h}_{2,n}^{(1)}\right) = E\left\{\left|y_{2,n} - \hat{h}_{2,n}^{(1)}\sqrt{\alpha_2 P}s_{2,n}\right|^2\right\}. \tag{5.3.4}$$

A simplified approach is taken whereby the known pilot is taken as $s_{2,n} = 1$. On performing the stochastic gradient descent on the chosen cost function, the estimator of the channel is found to be:

$$\hat{h}_{2,n}^{(1)} = \hat{h}_{2,n-1}^{(1)} + \mu\sqrt{\alpha_2 P}\left(y_{2,n} - \hat{h}_{2,n-1}^{(1)}\sqrt{\alpha_2 P}\right). \tag{5.3.5}$$

The update equation in (5.3.5) summarizes the CSS-LMS algorithm for the first iteration. An error variance analysis can therefore be performed with the use of the update equation.

Error Variance Analysis

An asymptotic MSE analysis is presented in this subsection. Borrowing from the procedure followed in Section 3.3.1, additional components that are dissimilar from those presented in Section 3.3.1 are derived in detail. Given the update rule in (5.3.5), a low-pass filter is designed. The corresponding expression in the z-domain is given as follows:

$$L(z) = \frac{\mu\sqrt{\alpha_2 P}}{1 - z^{-1}(1 - \mu\alpha_2 P)}, \tag{5.3.6}$$

where $L(z)$ is the low pass filter with cut-off frequency chosen as:

$$f_c T = \frac{\mu\alpha_2 P}{2\pi(1-\mu\alpha_2 P)}.$$

Using the approach referred to in Section 3.3.1, the estimation error is expressed as:

$$e_2^{(1)}(z) = h_2(z) - L(z)(\, h_2(z)(\sqrt{\alpha_1 P}S_1(z) + \sqrt{\alpha_2 P}) + W_2(z)\,)$$

$$= \big(1 - L(z)(S_1(z)\sqrt{\alpha_1 P} + \sqrt{\alpha_2 P})\big)h_2(z) - L(z)W_2(z). \qquad (5.3.7)$$

With the expression given in (5.3.7), the corresponding asymptotic MSE in the frequency domain can be deduced as:

$$\sigma^2_{e_2^{(1)}} = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - \sqrt{\alpha_2 P}L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_2}(f)\, df + \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\sqrt{\alpha_1 P}\right|^2 S_1(f)\Gamma_{h_2}(f) df$$

$$+ \sigma^2_{w_2} T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df, \qquad (5.3.8)$$

such that the MSE variance can be split term by term as expressed hereunder:

$$iMSE1 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - \sqrt{\alpha_2 P}L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_2}(f)\, df,$$

$$iMSE2 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\sqrt{\alpha_1 P}\right|^2 S_1(f)\Gamma_{h_2}(f)\, df,$$

$$iMSE3 = \sigma^2_{w_2} T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df. \qquad (5.3.9)$$

As it appears, $iMSE1$ and $iMSE3$ in (5.3.9) are of the same form as (3.3.10) and (3.3.11) respectively, these can be checked in chapter three, Section 3.3.1. Consequently, following the same procedure undertaken in deriving the final expressions for (3.3.10) and (3.3.11), the following expressions are derived:

$$iMSE1 = \sigma^2_{h_2} \frac{(2\pi f_d T)^2}{2(\mu\alpha_2 P)^2}(1-\mu\alpha_2 P)^2, \qquad (5.3.10)$$

$$iMSE3 = \sigma^2_{w_2} \frac{\mu}{(2-\mu\alpha_2 P)}. \qquad (5.3.11)$$

A different approach is taken in evaluating $iMSE2$ in (5.3.9). However, similar assumptions as with the ones considered in Section 3.3.1 of chapter three are made. These assumptions include, among others, the low Doppler scenario considerations. The mathematical derivation is expressed as:

$$iMSE2 = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left| L\left(\boldsymbol{e}^{2j\pi fT}\right)\sqrt{\alpha_1 P}\right|^2 S_1(f)\Gamma_{h_2}(f)\, df$$

$$= \alpha_1 PT \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left| L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_2}(f)\, df$$

$$= \alpha_1 PT \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left| \frac{\mu\sqrt{\alpha_2 P}}{1-(1-j2\pi fT)(1-\mu\alpha_2 P)}\right|^2 \frac{\sigma_{h_2}^2}{\pi\sqrt{f_d^2 - f^2}}\, df$$

$$= \alpha_1 PT \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left| \frac{\mu\sqrt{\alpha_2 P}}{(1-\mu\alpha_2 P)\left[\frac{1-(1-\mu\alpha_2 P)}{(1-\mu\alpha_2 P)} + j2\pi fT\right]}\right|^2 \frac{\sigma_{h_2}^2}{\pi\sqrt{f_d^2 - f^2}}\, df. \quad (5.3.12)$$

To simplify the expression, the following variable is introduced $K = \frac{\mu\sqrt{\alpha_2 P}}{(1-\mu\alpha_2 P)}$,

$$iMSE2 = \alpha_1 PT \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left| \frac{K}{2\pi f_c T + j2\pi fT}\right|^2 \frac{\sigma_{h_2}^2}{\pi\sqrt{f_d^2 - f^2}}\, df$$

$$= \alpha_1 PT \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \frac{K^2}{(2\pi f_c T)^2 + (2\pi fT)^2} \frac{\sigma_{h_2}^2}{\pi\sqrt{f_d^2 - f^2}}\, df. \quad (5.3.13)$$

From the assumptions made in the referred section, Section 3.3.1, the following approximation persists: $(2\pi f_c T)^2 + (2\pi fT)^2 \approx (2\pi f_c T)^2$. As a result,

$$iMSE2 = \alpha_1 PT \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \frac{K^2}{(2\pi f_c T)^2} \frac{\sigma_{h_2}^2}{\pi\sqrt{f_d^2 - f^2}}\, df.$$

A change of variables, $f = f_d \cos\varphi$, is introduced to get:

$$iMSE2 = \alpha_1 PT\sigma_{h_2}^2 \frac{K^2}{(2\pi f_c T)^2} \int_{\frac{\pi}{2}}^{0} \frac{1}{\pi\sqrt{f_d^2(1-\cos^{-2}\varphi)}} - f_d \sin\varphi\, d\varphi$$

$$= \alpha_1 PT\sigma_{h_2}^2 \frac{K^2}{\pi(2\pi f_c T)^2}\int_0^{\frac{\pi}{2}} 1 d\varphi$$

$$= \alpha_1 PT\sigma_{h_2}^2 \frac{K^2}{\pi(2\pi f_c T)^2}\frac{\pi}{2}. \quad (5.3.14)$$

Given the above result, K and $2\pi f_c T$ is substituted back into the expression to give:

$$iMSE2 = \alpha_1 PT\sigma_{h_2}^2 \frac{\left(\mu\sqrt{\alpha_2 P}\right)^2}{(1-\mu\alpha_2 P)^2}\frac{(1-\mu\alpha_2 P)^2}{(\mu\alpha_2 P)^2}$$

$$= \sigma_{h_2}^2 \frac{\alpha_1 PT}{\alpha_2 P}$$

$$= \sigma_{h_2}^2 \frac{\alpha_1 T}{\alpha_2}. \quad (5.3.15)$$

The sum of the expressions in (5.3.10), (5.3.11) and (5.3.15) gives the overall error variance, which is expressed as:

$$\sigma_{e_2^{(1)}}^2 = \sigma_{h_2}^2 \frac{(2\pi f_d T)^2}{2(\mu\alpha_2 P)^2}(1-\mu\alpha_2 P)^2 - \sigma_{h_2}^2 \frac{\alpha_1 T}{\alpha_2} - \sigma_{w_2}^2\left(\frac{\mu}{2-\mu\alpha_2 P}\right). \quad (5.3.16)$$

5.3.2.2 CSS-LMS i-th Iteration

This subsection extends subsection 5.3.2.1 for iterations after the initial iteration. As a consequence, a similar setting in terms of the assumptions and considerations is presumed. Post the initial iteration, the chosen cost function for iteration i is expressed as:

$$V\left(\hat{h}_{2,n}^{(i)}\right) = E\left\{\left|Y_{2,n}^{(i)} - \hat{h}_{2,n}^{(i)}\sqrt{\alpha_2 P}\right|^2\right\}. \quad (5.3.17)$$

The attributed update equation is given as:

$$\hat{h}_{2,n}^{(i)} = \hat{h}_{2,n-1}^{(i)} + \mu\sqrt{\alpha_2 P}\left(Y_{2,n}^{(i)} - \hat{h}_{2,n-1}^{(i)}\sqrt{\alpha_2 P}\right). \quad (5.3.18)$$

The channel coefficients are therefore estimated by the update rule given in (5.3.18).

Error Variance Analysis

As with the previous subsections on asymptotic MSE analysis, the low-pass filter is considered. For the *i*-th iteration, the same simple first-order time-invariant low-pass filter as in (5.3.6) is

designed. The channel estimation error on the other hand, due to estimation error components introduced in subsequent iterations becomes different. Following is the CE error expressed in the z-domain:

$$e_2^{(i)}(z) = h_2(z) - \hat{h}_2^{(i)}(z),$$

$$\begin{aligned} e_2^{(i)}(z) &= h_2(z) - L(z)[\, h_2(z)(\sqrt{\alpha_2 P} + \sqrt{\alpha_1 P}\Delta_{s_1}^{(i-1)}(z)) + \Delta_h^{(i-1)}\sqrt{\alpha_1 P}\hat{s}_1^{(i-1)} \\ &\quad + W_{2,p}(z)] \\ &= h_2(z)\left[1 - L(z)\sqrt{\alpha_2 P} - L(z)\Delta_{s_1}^{(i-1)}(z)\sqrt{\alpha_1 P}\right) \\ &\quad -L(z)\Delta_h^{(i-1)}\sqrt{\alpha_1 P}\hat{s}_1^{(i-1)} - L(z)W_2(z). \end{aligned} \tag{5.3.19}$$

The error expression can be rearranged as shown in equation (5.3.20), giving four expressions which can be operated on separately in deriving the asymptotic variance expression:

$$\begin{aligned} e_2^{(i)}(z) = \; & h_2(z)\left[1 - L(z)\sqrt{\alpha_2 P}\right] \\ & - h_2(z)L(z)\Delta_{s_1}^{(i-1)}(z)\sqrt{\alpha_1 P}) \\ & + L(z)\Delta_h^{(i-1)}(z)\sqrt{\alpha_1 P}\hat{s}_1^{(i-1)} \\ & -L(z)W_2(z). \end{aligned} \tag{5.3.20}$$

From the above equation, $h_2(z)\Delta_{s_1}^{(i-1)}(z)\sqrt{\alpha_1 P}$ in the second term is reasonably assumed to be a Gaussian random variable with variance $\sigma^2_{\Delta_{s_1}^{(i-1)}}$. With this assumption, the asymptotic MSE is expressed in the frequency domain as shown below:

$$\begin{aligned} \sigma^2_{e_2^{(i)}} = & \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - \sqrt{\alpha_2 P}L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_2}(f)\, df \\ & + \sigma^2_{\Delta_{s_1}^{(i-1)}} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df \\ & + \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\sqrt{\alpha_1 P}\hat{s}_1^{(i-1)}\right|^2 \Delta_h^{(i-1)}(f) df \end{aligned}$$

$$+ \sigma_{w_2}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df. \tag{5.3.21}$$

The overall asymptotic MSE expression can be decomposed into the following expressions:

$$MSE1_i = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|1 - \sqrt{\alpha_2 P} L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 \Gamma_{h_2}(f)\, df,$$

$$MSE2_i = \sigma^2_{\Delta_{s_1}^{(i-1)}} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df,$$

$$MSE3_i = \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\sqrt{\alpha_1 P}\hat{s}_1^{(i-1)}\right|^2 \Delta_h^{(i-1)}(f) df,$$

$$MSE4_i = \sigma_{w_2}^2 T \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\right|^2 df. \tag{5.3.22}$$

The main idea behind disjoining the error variance terms as in (5.3.22) is to simplify the evaluation of the overall expression. Given the four terms, the first term, $MSE1_i$, can be evaluated in a similar manner to (3.3.10) in Section 3.3.1 of chapter three. On the other hand, $MSE2_i$ and $MSE4_i$ exhibit the same form as static part of the error variance in Section 3.3.1, given by (3.3.11). Taking into account the afore-mentioned, the resulting expressions for $MSE1_i$, $MSE2_i$, and $MSE3_i$ are as follows:

$$MSE1_i = \sigma_{h_2}^2 \frac{(2\pi f_d T)^2}{2(\mu\alpha_2 P)^2}(1 - \mu\alpha_2 P)^2, \tag{5.3.23}$$

$$MSE2_i = \sigma^2_{\Delta_{s_1}^{(i-1)}} \frac{\mu}{(2 - \mu\alpha_2 P)}, \tag{5.3.24}$$

$$MSE4_i = \sigma_{w_2}^2 \frac{\mu}{(2 - \mu\alpha_2 P)}. \tag{5.3.25}$$

The following subsection outlines the derivation of the expressions for $MSE3_i$. For $MSE3_i$, the variance of the error term carried from the previous iteration, $\Delta_h^{(i-1)}$, is considered to be flat and is given as $\sigma^2_{e_2^{(i-1)}}$. As such $MSE3_i$ can be expressed as:

$$MSE3_i = \sigma^2_{e_2^{(i-1)}} \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left|L\left(\boldsymbol{e}^{2j\pi fT}\right)\sqrt{\alpha_1 P}\hat{s}_1^{(i-1)}\right|^2 df$$

$$= \sigma^2_{e_2^{(i-1)}} \alpha_1 P \int_{\frac{-1}{2T}}^{\frac{1}{2T}} \left| L\left(\boldsymbol{e}^{2j\pi fT}\right) \right|^2 df. \tag{5.3.26}$$

In this form, $MSE3_i$ can easily be evaluated in the same way $MSE2_i$ and $MSE4_i$ are evaluated. Thus, the final expression of the term is given as:

$$MSE3_i = \sigma^2_{e_2^{(i-1)}} \alpha_1 P \frac{\mu}{(2 - \mu\alpha_2 P)}. \tag{5.3.27}$$

Therefore, in light of (5.3.23), (5.3.24), (5.3.25) and (5.3.27), the overall error variance expression is given as follows:

$$\sigma^2_{e_2^{(i)}} = \sigma^2_{h_2} \frac{(2\pi f_d T)^2}{2(\mu\alpha_2 P)^2} (1 - \mu\alpha_2 P)^2 + \sigma^2_{\Delta_{s_1}^{(i-1)}} \frac{\mu}{(2 - \mu\alpha_2 P)} + \sigma^2_{e_2^{(i-1)}} \alpha_1 P \left(\frac{\mu}{2 - \mu\alpha_2 P}\right)$$

$$+ \sigma^2_{w_2} \left(\frac{\mu}{2 - \mu\alpha_2 P}\right), \tag{5.3.28}$$

which is simplified:

$$\sigma^2_{e,i} = \sigma^2_{h_2} \frac{(2\pi f_d T)^2}{2(\mu\alpha_2 P)^2} (1 - \mu\alpha_2 P)^2 + \left(\frac{\mu}{2 - \mu\alpha_2 P}\right) \left[\sigma^2_{\Delta_{s_1}^{(i-1)}} + \sigma^2_{e,i-1} \alpha_1 P + \sigma^2_{w_2}\right]. \tag{5.3.29}$$

Expression of Variance $\sigma^2_{\Delta_{s_1}^{(i-1)}}$

This subsection provides the variance of the term $h_2(z)\Delta_{s_1}^{(i-1)}(z)\sqrt{\alpha_1 P}$ in $MSE2_i$, which is assumed to be Gaussian. The expression is instituted by providing the power spectral density (PSD) for $\Delta_{s_1}^{(i-1)}(z)$. To advance this, the following is recalled:

$$\Delta_{s_1}^{(i-1)} = s_{1,p} - \hat{s}_{1,p}^{(i-1)}.$$

Therefore, $\Delta_{s_1}^{(i-1)}$ is given by the following:

$$\Delta_{s_1}^{(i-1)} = \begin{cases} 0, & \left(1 - P_e^{(i-1)}\right) \\ +2, & \frac{1}{2} P_e^{(i-1)} \\ -2, & \frac{1}{2} P_e^{(i-1)} \end{cases}, \tag{5.3.30}$$

where $P_e^{(i-1)}$ is the probability of error in detecting s_1.

The autocorrelation of which is expressed as:

$$R_{\Delta_{s_1}^{(i-1)}}(l) = \begin{cases} 4P_e^{(i-1)}, & l = 0 \\ 0, & l \neq 0 \end{cases}.$$

By definition, the corresponding PSD is expressed as follows:

$$PSD\left(\Delta_{s_1}^{(i-1)}\right) = \frac{1}{T}(Tsinc(2\pi fT))^2 \sum_{-\infty}^{\infty} R_{\Delta_{s_1}^{(i-1)}}(l) e^{-j2\pi fT}$$

$$= Tsinc^2(2\pi fT) \sum_{-\infty}^{\infty} \left(R_{\Delta_{s_1}^{(i-1)}}(0) + R_{\Delta_{s_1}^{(i-1)}}(l) \right) e^{-j2\pi fT}$$

$$= Tsinc^2(2\pi fT) \sum_{-\infty}^{\infty} \left(4P_e^{(i-1)} + 0 \right) e^{-j2\pi fT}$$

$$= 4P_e^{(i-1)} T, \qquad (5.3.31)$$

where T is the symbol duration.

With this setting, the following can be obtained as the overall variance of term of interest:

$$\sigma^2_{\Delta_{s_1}^{(i-1)}} = 4P_e^{(i-1)} T \, E\left\{ \left| h_2 \sqrt{\alpha_1 P} \right|^2 \right\}$$

$$= 4\alpha_1 P \sigma^2_{h_2} P_e^{(i-1)} T. \qquad (5.3.32)$$

5.3.3 Results and Evaluation

In this subsection, the performance evaluation of the CSS-LMS-based iterative CE and detection scheme is examined. Since the focus of this research is CE, detection is treated as a black-box. That is, no practical detection algorithm is proposed. Instead, the iterative scheme is evaluated for different scenarios of probability of error. As the set-up for the computer simulations, the power allocations are set as $\alpha_1 = 0.39$ and $\alpha_2 = 0.61$ for user 1 and user 2 respectively. The channel coefficients between each user and the base-station are generated using Jake's channel model, with the corresponding variances computed within the simulation. For this, the normalized Doppler frequency is set as $f_d T = 1.2 \times 10^{-4}$. The algorithm implementation example for the convergence results is presented as in Algorithm 11, in

Appendix A. Other results can be generated in a similar manner as with the all-known pilots, using respective equations.

5.3.3.1 Probability of Error Propagation CE Performance

In this subsection, the effects of detection error on CE performance are presented. Firstly, the convergence characteristics of the CSS-LMS algorithm employed in the iterative scheme are presented for different probabilities of error. This is depicted in Figure 5.3.2, where the convergence of the MSE is shown for $P_e = 0, 0.001, 0.006$ and 0.01.

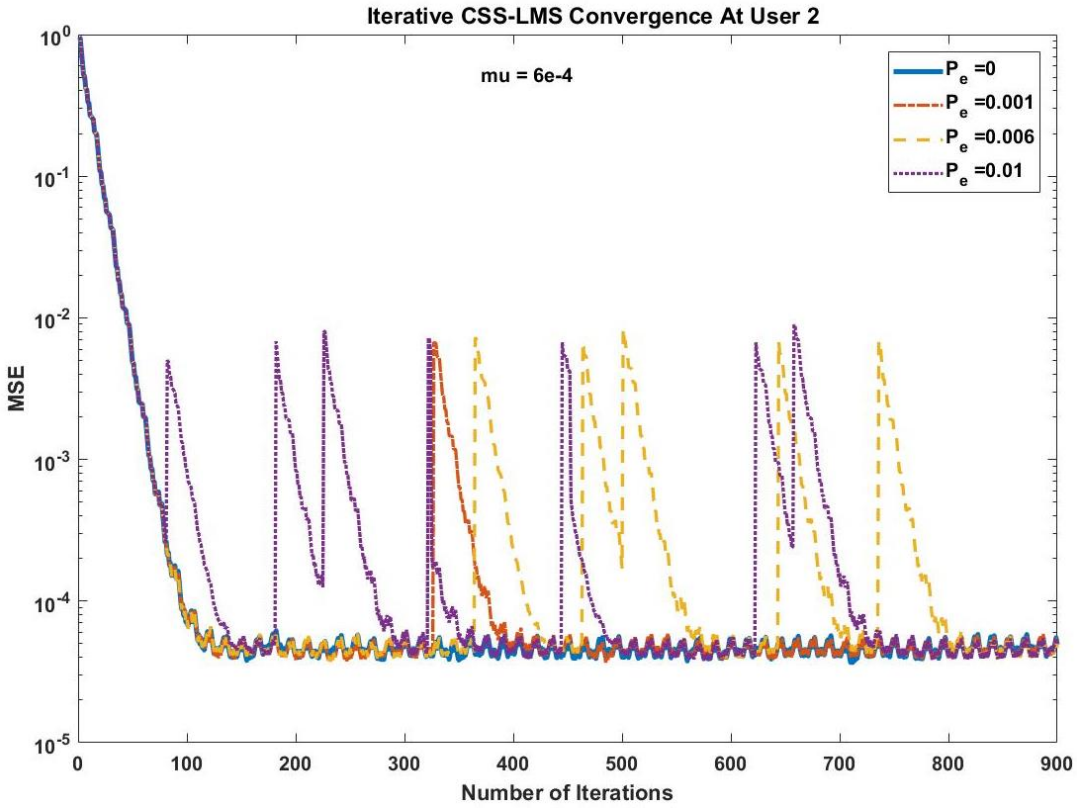


Figure 5.3. 2 User 2: Convergence Characteristics of the CSS-LMS algorithm for different probabilities of error (Pe's).

For $P_e = 0$, assuming perfect detection, convergence is observed in Figure 5.3.2 as a blue curve without spikes, mostly veiled by other curves. In assuming the error in detection $P_e = 0.001$, a similar trend as observed when assuming perfect detection. However, a spike in the MSE is observed at the stage of convergence. The disruption of the MSE at convergence is a result of the bit errors associated with the probability of error. From Figure 5.3.1, the spikes arise from the probability of error from the detection block. These spikes increase in number

with the increase in the P_e. As the P_e increases, the combined CE and detection error also increases, and this affects the CE accuracy. The effect propagation of the error on channel estimation is similar to the one observed in Section 5.2.3.

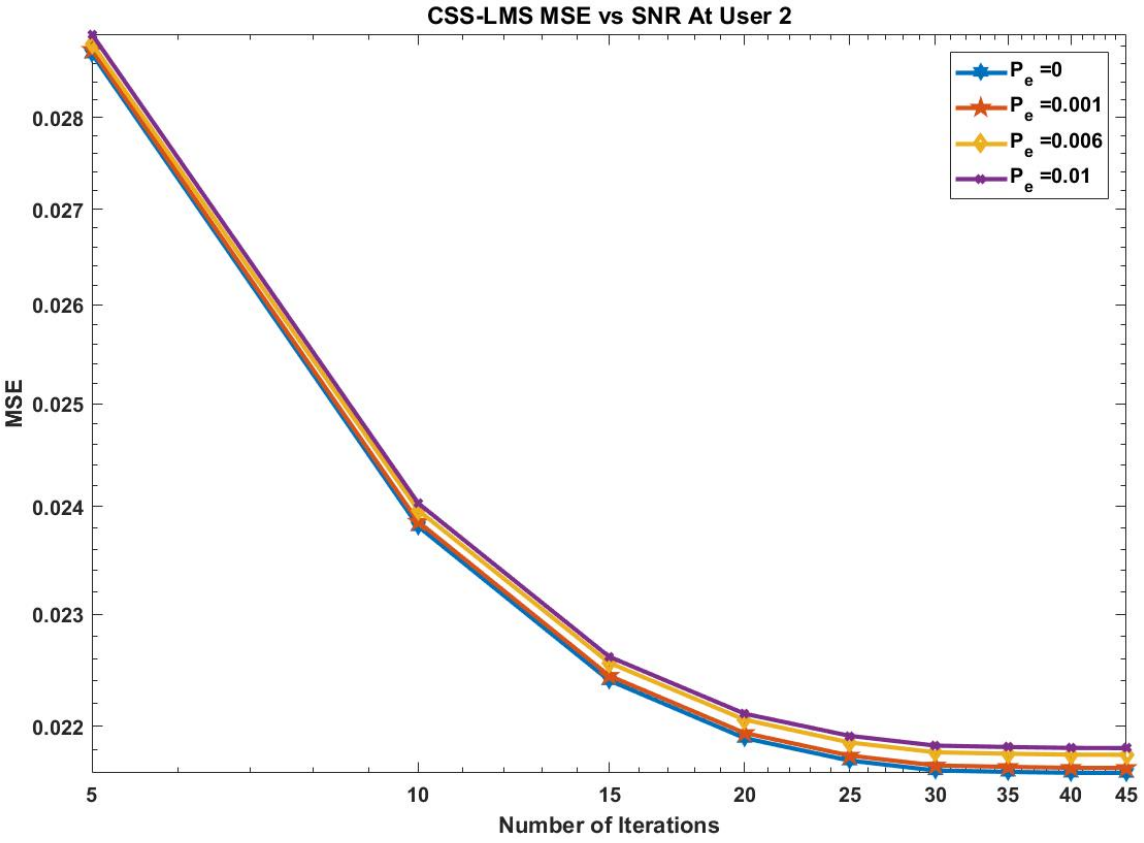


Figure 5.3. 3 User 2: The Evolution of the MSE at different probabilities of error, in varying SNR conditions.

In Figure 5.3.3, the evolution of the MSE for a range of SNR values is depicted. As with the convergence characteristics, the results were generated assuming perfect detection and for different probabilities of error: $P_e = 0, 0.001, 0.006$ and 0.01. The best CE performance is apparent when perfect detection is assumed. Performance degradation is witnessed with the consideration of probability of error in detection. With consideration of higher detection error values translates to poorer performance.

The signal detection error propagates to the subsequent iterations which are essentially meant for refining the channel estimate. As a result, the overall error is "magnified" as the error is compounded. It is of importance therefore, that an optimized signal detection algorithm is

employed when a practical detection system is considered. This is to guarantee a sufficiently low probability of error that does not compromise the channel estimation performance.

5.3.3.2 Channel Estimation Analytical Performance Evaluation

In this subsection, the simulation is compared to the analysis carried out in subsection 5.3.2.2. As the set-up, the probability of error considered for the signal detection block is $P_e = 0, 0.001$. This is the probability of error that has comparable CE performance when compared to the scenario of perfect detection. With this consideration, the theoretical analysis of the CSS-LMS algorithm for the Iterative approach is compared to the simulation. That is, the analytical asymptotic MSE of the CSS-LMS is compared to the simulated counterpart and this is illustrated in Figure 5.3.4.

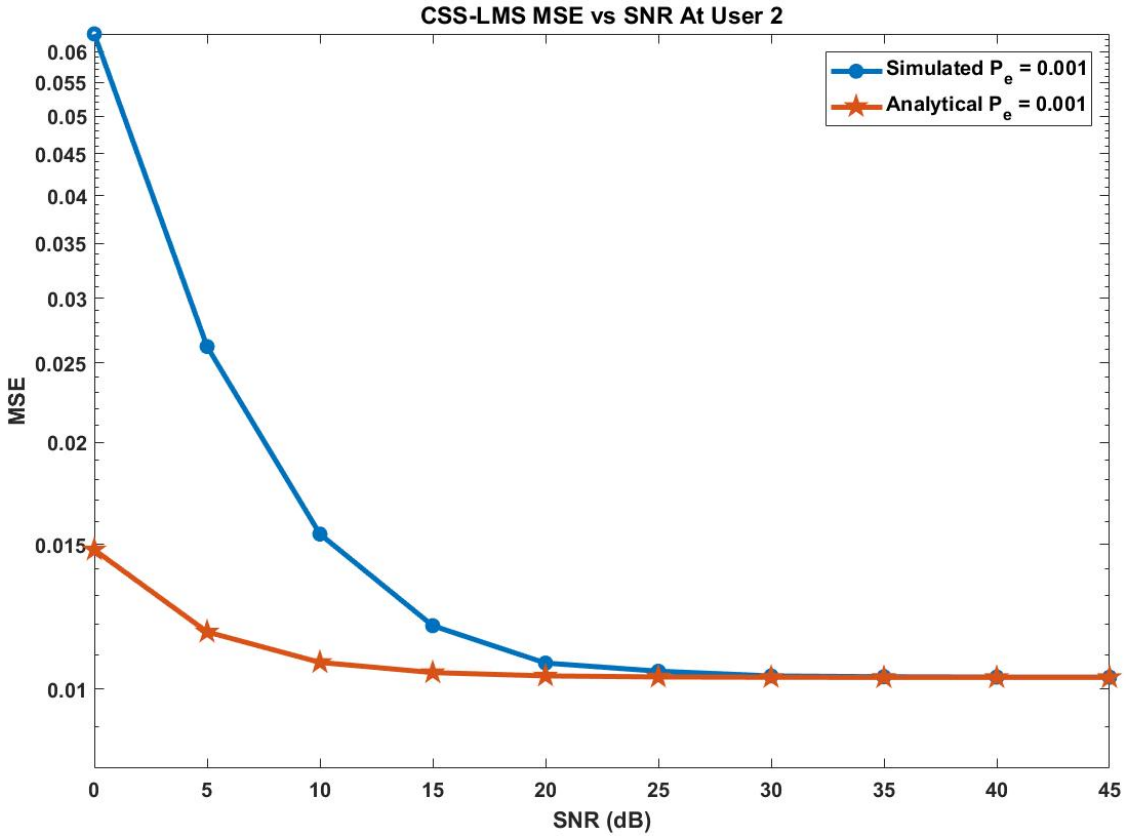


Figure 5.3. 4 User 2: Comparison between the analytical error variance and the simulated MSE in varying SNR Conditions.

In Figure 5.3.4, disparity between the theoretical error variance and the simulated is obvious for low SNR scenarios. This behaviour is mainly caused by the fact that the carried CE error, related to the third term in (5.3.29), is much more significant in low SNR scenarios. The change

of the MSE as the SNR values get higher however increasingly maintains the theoretical analysis carried out to derive the error variance. Therefore, the simulation of the CSS-LMS algorithm confirms the analytical derivations carried out in subsection 5.2.2.1 for high SNR scenarios.

5.4 Complexity Analysis of the CE Approaches Compared to the CCCP-Based CE Algorithm [23]

In this section, the computational complexity of the presented algorithms as well as the CCCP-based CE approach in [23] for the number of operations N. It should be noted that, while the presented LMS-based algorithms only focus on CE, the work in [23] also includes power allocation in the optimization problem. For each algorithm, the worst-case scenario is considered in the respective analyses. The CCCP-based algorithm in [23] has the complexity of $O(L * N^3)$. This is on account of the L iterations that are required for convergence. Furthermore, the vector-matrix operations of the optimization problem solution are of the cubic exponent.

The all-known-pilots and the case-wise CE approaches have the complexity of $O(N^2)$, which is essentially derived from the system model, since a simple LMS algorithm is employed. The worst-case scenario of the averaging sum CE approach introduces N multiple operations by virtue of the averaging-sum operation, and is given by $O(N^3)$. The SIC-inspired CE approach has an added loop for refining both the pilot and channel estimates, which at worst has J runs. As such, the complexity of the SIC-inspired CE approach is given by $O(J * N^2)$.

Table 5.4.1 summarizes the complexities of the CE approaches highlighted in the paragraph above. The CCCP-based CE approach of [23] has the highest computational complexity. This is since their problem formulation covers both CE and power allocation.

	Channel Estimation Approach	**Time Complexity**
1.	CCCP-Based Channel Estimation Approach in [23].	$O(L * N^3)$
2.	All-Known Pilots Channel Estimation approach using the LMS Algorithm.	$O(N^2)$

3.	Averaging Sum Based Channel Estimation Approach using the LMS Algorithm	$O(N^3)$
4.	Case-wise Channel Estimation Approach using the LMS Algorithm	$O(N^2)$
5.	SIC-inspired Iterative Channel Estimation Approach using the LMS Algorithm	$O(J * N^2)$

5.5 Chapter Summary

This chapter presented two CE techniques for a two-user downlink SISO-NOMA system. The first approach is the SIC-inspired iterative CE and detection scheme intended for the stronger user. For the weaker user, the case-wise CE and detection approach is presented.

Firstly, the case-wise CE scheme is presented, with the core algorithm employed as the CSS-LMS algorithm. The corresponding error variance expressions are then derived and presented. Upon presenting the expressions of the channel estimator, the case-wise CE approach was then evaluated. Through simulation results, it was established that the pilot detection error propagates and inherently affects the CE performance. This in turn motivates for a reasonable probability of error in implementing a practical detector. Channel estimation was therefore evaluated for a probability of error of 0.001 in comparison to the theoretical analysis. The simulation MSE confirms the analytical MSE for the entire range SNRs considered.

In progression, the CSS-LMS based iterative CE and detection approach is introduced in this chapter. A formulation is provided, which basically entails a high-level description of how the approach is set to work. Consecutively, the CSS-LMS algorithm is presented for the approach and the error variance expressions are established. Through computer simulations, the SIC-inspired iterative CE and detection scheme was evaluated. Through evaluations, the probability of error in detection was shown to degrade the CE performance. For various probabilities of error, the probability of error of 0.001 had comparable CE performance to a scenario of perfect detection. As such, the acceptable probability of error in detection was selected as 0.001. In turn, this motivates the employment of an optimized signal detection scheme to ensure a sufficiently low bit error rate. The iterative CE and detection approach was shown to confirm the theoretical analysis for high SNR conditions.

The computational complexity analysis comparing the CE estimation approaches presented herein and the approach presented in [23] was then conducted. From this exercise, the CCCP-based approach proposed in [23] was shown to have the highest worst-case scenario complexity. This can be attributed to the fact that the approach in [23] addresses both CE and power allocation in their letter. Amongst the CE approaches presented in this dissertation, the averaging-sum based channel estimation approach has the highest complexity. The SIC-inspired iterative CE approach has the second highest complexity. The all-known pilots CE approach and the case-wise approaches have the same worst-case scenario complexity, which is the lowest. With the provided information, trade-offs of each approach can be deduced considering complexity and performance in terms of CE accuracy and spectral efficiency.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

As a measure to address the requirements of the fifth generation of wireless mobile networks, non-orthogonal multiple access (NOMA) has been proposed. In designing multiple access schemes, channel estimation (CE) has proven to be a pivotal principle for reliable recovery of information. As such, the work in this dissertation focused on CE for downlink NOMA based systems. In this regard, the main objective of this work is the development of a CE scheme for downlink NOMA that is efficient in both speed and accuracy.

Through a literature survey, CE principles were explored and detailed in chapter two. For the most part, the work documented in literature entails CE for orthogonal multiple access (OMA)-based systems. Channel estimation for NOMA systems is mostly prominent for uplink applications. However, an account of CE for the downlink is presented in [23]. In this work, a simple linear estimator is developed with attention focused on optimizing power allocation amongst two users. With this direction, the work in [23] was therefore not concerned with establishing a CE scheme that guarantees accuracy or speed.

Given the current situation in literature, this dissertation proposed four pilot aided CE schemes essentially based on the least mean squares (LMS) algorithm. The first presented approach is the all-known-pilots CE technique, employing three LMS variants. These are the constant step-size LMS (CSS-LMS), the variable step-size LMS (VSS-LMS) and the VSS-LMS with added speed (VSS-LMS-AS). The three LMS variants are adapted from an OMA application presented in [22]. In principle, the all-known-pilots scheme essentially adopts the principle of using sufficient training data to ensure the reliability of CE.

Computer simulations were carried out in order to evaluate the performance of the all-known-pilots technique, using the three LMS variants. Through evaluations, the all-known-pilots CE scheme posed superior performance with employment of an optimal step-size value. The VSS-LMS-AS outperformed the VSS-LMS algorithm for all scenarios considered. The fore-mentioned conclusions were made with CE speed in consideration. In terms of CE accuracy however, the three algorithms exhibit comparable performance. The error variance theoretical analysis carried out for the CSS-LMS algorithm was confirmed through simulation.

The averaging sum formulation-based CE scheme is presented, also assuming the three LMS algorithm variants used for the all-known-pilots CE approach. As opposed to the all-known-pilots CE approach, the averaging sum approach essentially considered partial knowledge of the training data. Fundamentally, this approach entails performing an averaging sum on the received signal as a measure of minimizing the significance of the unknown pilot symbols in the overall observation. For assessment, computer simulations were employed.

Firstly, the highlight of the performance evaluations was the disparities between the CE performance at each user. These differences were identified to have resulted from different power allocations. The user with a higher power allocation demonstrated elevated performance in comparison to the user with a lower power allocation. Furthermore, the CE performance disparities manifested when employing all the three LMS variants in the averaging sum approach.

Evaluation of the theoretical analysis was carried out for the CSS-LMS algorithm. In this regard, the simulation MSE and the analytical MSE aligned for the SNR conditions considered and for different step-sizes. As such, the averaging sum-based scheme using the CSS-LMS confirmed the theoretical analysis. In comparing the application of the LMS variants, the CSS-LMS algorithm exhibits the best performance in terms of CE speed. In terms of channel estimation accuracy, the three algorithms show virtually similar CE performance when employed for the averaging sum CE approach. The VSS-LMS-AS algorithm on the other hand out-performed the VSS-LMS algorithm in terms of CE speed.

The third CE approach developed in this dissertation was the case-wise CE and detection approach for the weaker user. Essentially taking advantage of power allocation in NOMA, the SIC inspired iterative CE and detection approach was presented for the stronger user. For both of the approaches, partial knowledge of the pilot symbols was considered. Performance for both CE schemes was conducted in relation to the probability of error, error propagation and theoretical analysis.

For the case-wise CE scheme, it was established that the CE performance is affected by the probability of error in pilot detection. This is due to the fact that the error propagates to the CE block. As such, this motivates for a detection scheme with a sufficiently low probability of error. The case-wise scheme was then evaluated against the theoretical analysis for a probability of error of 0.001. With the considered probability of error, the simulation MSE corroborated the analysis.

The SIC-inspired iterative CE and detection approach was then presented along with the corresponding performance evaluations. For different probabilities of error, the CE performance was generally assessed. It was established that the CE accuracy degraded with increase in the probability of error in pilot detection. The acceptable probability of error was selected as 0.001. With this probability of error, the CE performance achieved was comparable to the one achieved when perfect detection was assumed. Furthermore, the simulation MSE and analytical MSE were consistent for high SNR scenarios.

6.2. Future Directions

The primary focus of the research was to introduce a competent channel estimation scheme for a downlink NOMA system with the main performance objectives being speed and accuracy. Four channel estimation approaches were proposed, with their attributed advantages and shortcomings. Given the nature of the work presented in this dissertation, the following are the suggested future works:

- In this research, a simplified downlink scenario comprised of two users and a base-station is considered. At heart, the notion of NOMA entails accommodation of multiple users in a given resource block. A viable research direction would therefore entail addressing channel estimation when a large number of users are considered.
- Combining the case-wise channel estimation approach with the SIC-inspired iterative channel estimation scheme counterpart employed for the stronger user is a potential research direction.
- Another potential research direction can be predicated on the fact that power allocation affects the channel estimation performance. This is especially apparent for the averaging sum approach. Sub-optimal power allocation could be addressed in conjunction with optimal channel estimation ensuring user fairness with regards to channel estimation performance.
- In terms of channel condition types, further research directions may encompass channel estimation approaches that will address other fading scenarios which were not covered in this research.
- In this research, the channel estimation approaches proposed are essentially pilot aided schemes. Introduction of blind, semi-blind or decision directed channel estimation schemes for downlink NOMA is a viable research direction.

- In advent of the current wave of deep-learning for channel estimation in the next generation of wireless mobile networks, practical deep-learning schemes are potential research directions. Potential performance improvements in channel estimation can be realized with the employment of deep-learning. This is in relation to traditional channel estimation approaches.

References

[1] E. Dahlman et al., "5G wireless access: requirements and realization," in IEEE Communications Magazine, vol. 52, no. 12, pp. 42-47, December 2014. doi: 10.1109/MCOM.2014.6979985

[2] S. Mattisson, "Overview of 5G requirements and future wireless networks," *ESSCIRC 2017 - 43rd IEEE European Solid State Circuits Conference*, Leuven, 2017, pp. 1-6, doi: 10.1109/ESSCIRC.2017.8094511.

[3] A. Gledhill, *5 reasons why 5G is the future*. [online] Information Age. Available at: https://www.information-age.com/5g-future-123472007/ [Accessed 22 Sept. 2020]

[4] Y. Wang, B. Ren, S. Sun, S. Kang and X. Yue, "Analysis of non-orthogonal multiple access for 5G," in *China Communications*, vol. 13, no. 2, pp. 52-66, 2016, doi: 10.1109/CC.2016.7405722.

[5] PHD Projects. *PhD research topics in Wireless communication*. [online] Available at: http://phdprojects.org/phd-research-topics-wireless-communication/ [Accessed 3 Sep. 2019].

[6] Engpaper.com. *Wireless communications and mobile technology research topics 2019*. [online] Available at: https://www.engpaper.com/wireless-communications-and-mobile-technology-research-topics-2019.htm [Accessed 3 Sep. 2019].

[7] Y. Tao, L. Liu, S. Liu and Z. Zhang, "A survey: Several technologies of non-orthogonal transmission for 5G," in China Communications, vol. 12, no. 10, pp. 1-15, Oct. 2015, doi: 10.1109/CC.2015.7315054.

[8] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen and L. Hanzo, "A Survey of Non-Orthogonal Multiple Access for 5G," in *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2294-2323, thirdquarter 2018, doi: 10.1109/COMST.2018.2835558.

[9] Y. Yifei and Z. Longming, "Application scenarios and enabling technologies of 5G," in China Communications, vol. 11, no. 11, pp. 69-79, Nov. 2014, doi:10.1109/CC.2014.7004525.

[10] Nakamura, Takehiro, Anass Benjebbour, Yoshihisa Kishiyama, Satoshi Suyama, and Tetsuro Imai. "5G radio access: Requirements, concept and experimental trials." IEICE Transactions on Communications 98, no. 8, pp. 1397-1406, Sept. 2018, doi:10.1587/transcom.E98.B.1397.

[11] Mehak Basharat, Waleed Ejaz, Muhammad Naeem, Asad Massod Khattak, Alagan Anpalagan, “A survey and taxonomy on nonorthogonal multiple-access schemes for 5G Networks”, Trans Emerging Tel Tech, 2017, doi: 10.1002/ett.3202

[12] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," in *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74-81, September 2015, doi: 10.1109/MCOM.2015.7263349.

[13] A. Benjebbour, K. Saito, A. Li, Y. Kishiyama and T. Nakamura, "Non-orthogonal multiple access (NOMA): Concept, performance evaluation and experimental trials," *2015 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Marrakech, 2015, pp. 1-6, doi: 10.1109/WINCOM.2015.7381343.

[14] Higuchi, Kenichi and Anass Benjebbour. "Non-orthogonal Multiple Access (NOMA) with Successive Interference Cancellation for Future Radio Access." IEICE Transactions 98-B, pp. 403-414, March 2015, doi: 10.1587/transcom.E98.B.403.

[15] H. Tabassum, M. S. Ali, E. Hossain, M. J. Hossain and D. I. Kim, "Uplink Vs. Downlink NOMA in Cellular Networks: Challenges and Research Directions," 2017 IEEE 85th Vehicular Technology Conference (VTC Spring), Sydney, NSW, 2017, pp. 1-7, doi: 10.1109/VTCSpring.2017.8108691.

[16] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li and K. Higuchi, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," 2013 IEEE 77th Vehicular Technology Conference (VTC Spring), Dresden, 2013, pp. 1-5, doi: 10.1109/VTCSpring.2013.6692652.

[17] Tabassum, Hina, Md Shipon Ali, Ekram Hossain, Md Hossain, and Dong In Kim. "Non-orthogonal multiple access (NOMA) in cellular uplink and downlink: Challenges and enabling techniques." arXiv preprint arXiv:1608.05783 (2016).

[18] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan and V. K. Bhargava, "A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends," in *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181-2195, Oct. 2017, doi: 10.1109/JSAC.2017.2725519.

[19] Y. Cai, Z. Qin, F. Cui, G. Y. Li and J. A. McCann, "Modulation and Multiple Access for 5G Networks," in IEEE Communications Surveys & Tutorials, vol. 20, no. 1, pp. 629-646, Firstquarter 2018. doi: 10.1109/COMST.2017.2766698.

[20] Olutayo O. Oyerinde & Stanley H. Mneney, "Review of Channel Estimation for Wireless Communication Systems," IETE Technical Review, vol. 29, no. 4, pp. 282-298, 2012, doi: 10.4103/0256-4602.101308.

[21] Arslan, Hüseyin, and Gregory E. Bottomley. "Channel estimation in narrowband wireless communication systems," Wireless Communications and Mobile Computing, vol. 1, no. 2, pp. 201-219, 2001, doi:10.1002/wcm.14.

[22] R. Gerzaguet, L. Ros, J. Brossier, S. Ghandour-Haidar and F. Belveze, "Self-adaptive stochastic rayleigh flat fading channel estimation," 2013 18th International Conference on Digital Signal Processing (DSP), Fira, 2013, pp. 1-6, doi: 10.1109/ICDSP.2013.6622690.

[23] Y. Tan, J. Zhou and J. Qin, "Novel Channel Estimation for Non-orthogonal Multiple Access Systems," in IEEE Signal Processing Letters, vol. 23, no. 12, pp. 1781-1785, Dec. 2016. doi: 10.1109/LSP.2016.2617897.

[24] S. M. R. Islam, N. Avazov, O. A. Dobre and K. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," in IEEE Communications Surveys & Tutorials, vol. 19, no. 2, pp. 721-742, Secondquarter 2017, doi: 10.1109/COMST.2016.2621116.

[25] Z. Ding *et al*., "Application of Non-Orthogonal Multiple Access in LTE and 5G Networks," in *IEEE Communications Magazine*, vol. 55, no. 2, pp. 185-191, February 2017, doi: 10.1109/MCOM.2017.1500657CM.

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

[27] Barsocchi, Paolo. "Channel models for terrestrial wireless communications: a survey." *CNR-ISTI technical report* 83, 2006.

[28] Arsal, Ali. *A study on wireless channel models: Simulation of fading, shadowing and further applications*. MS thesis. Izmir Institute of Technology, 2008.

[29] O. E. Ijiga, "Channel estimation techniques for filter bank multicarrier based transceivers for next generation of wireless networks, " Diss, University of the Witwatersrand, Johannesburg, 2017.

[30] Jakes, William C., and Donald C. Cox. "Microwave mobile communications. " Wiley-IEEE Press, 1994.

[31] T. K. Sarkar, Zhong Ji, Kyungjung Kim, A. Medouri and M. Salazar-Palma, "A survey of various propagation models for mobile communication," in IEEE Antennas and Propagation Magazine, vol. 45, no. 3, pp. 51-82, June 2003. doi: 10.1109/MAP.2003.1232163.

[32] M Morelli and U. Mengali., "A comparison of pilot-aided channel estimation methods for OFDM systems," in IEEE Transactions on Signal Processing, vol. 49, no. 12, pp. 3065-3073, Dec. 2001. doi: 10.1109/78.969514.

[33] P. Komulainen and V. Haikola, "Adaptive filtering for fading channel estimation in WCDMA downlink," 11th IEEE International Symposium on Personal Indoor and Mobile Radio Communications. PIMRC 2000. Proceedings (Cat. No.00TH8525), London, UK, 2000, pp. 549-553 vol.1. doi: 10.1109/PIMRC.2000.881484.

[34] Won Gi Jeon, Kyung Hyun Paik and Yong Soo Cho, "Two-dimensional MMSE channel estimation for OFDM systems with transmitter diversity," IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No.01CH37211), Atlantic City, NJ, USA, 2001, pp. 1682-1685 vol.3. doi: 10.1109/VTC.2001.956486.

[35] M. Biguesh and A. B. Gershman, "MIMO channel estimation: optimal training and tradeoffs between estimation techniques," 2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577), Paris, France, 2004, pp. 2658-2662 Vol.5. doi: 10.1109/ICC.2004.1313013.

[36] C. Yin, J. Li, X. Hou and G. Yue, "Pilot Aided LS Channel Estimation in MIMO-OFDM Systems," 2006 8th international Conference on Signal Processing, Beijing, 2006, pp. doi: 10.1109/ICOSP.2006.345832.

[37] H. Hijazi, E. P. Simon, M. Liénard and L. Ros, "Channel estimation for MIMO-OFDM systems in Fast Time-Varying Environments," 2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP), Limassol, 2010, pp. 1-6. doi: 10.1109/ISCCSP.2010.5463392.

[38] H. Ye, G. Y. Li and B. Juang, "Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems," in IEEE Wireless Communications Letters, vol. 7, no. 1, pp. 114-117, Feb. 2018. doi: 10.1109/LWC.2017.2757490.

[39] X. Ma, H. Ye and Y. Li, "Learning Assisted Estimation for Time- Varying Channels," 2018 15th International Symposium on Wireless Communication Systems (ISWCS), Lisbon, 2018, pp. 1-5. doi: 10.1109/ISWCS.2018.8491068.

[40] R. H. Clarke, "A statistical theory of mobile-radio reception," in The Bell System Technical Journal, vol. 47, no. 6, pp. 957-1000, July-Aug. 1968, doi: 10.1002/j.1538-7305.1968.tb00069.x.

[41] Xinmin Deng, A. M. Haimovich and J. Garcia-Frias, "Decision directed iterative channel estimation for MIMO systems," IEEE International Conference on

Communications, 2003. ICC '03., Anchorage, AK, 2003, pp. 2326-2329 vol.4. doi: 10.1109/ICC.2003.1204300.

[42] Jie Gao and Huaping Liu, "Decision-directed estimation of MIMO time-varying Rayleigh fading channels," in IEEE Transactions on Wireless Communications, vol. 4, no. 4, pp. 1412-1417, July 2005. doi: 10.1109/TWC.2005.852131.

[43] J. Akhtman and L. Hanzo, "Decision Directed Channel Estimation Aided OFDM Employing Sample-Spaced and Fractionally-Spaced CIR Estimators," in IEEE Transactions on Wireless Communications, vol. 6, no. 4, pp. 1171-1175, April 2007. doi: 10.1109/TWC.2007.348308.

[44] O. O. Oyerinde and S. H. Mneney, "Decision Directed Channel Estimation for OFDM Systems Employing Fast Data Projection Method Algorithm," 2009 IEEE International Conference on Communications, Dresden, 2009, pp. 1-5. doi: 10.1109/ICC.2009.5198884.

[45] J. Ketonen, M. Juntti and J. Ylioinas, "Decision directed channel estimation for improving performance in LTE-A," 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, 2010, pp. 1503-1507. doi: 10.1109/ACSSC.2010.5757787.

[46] Ganesh, R. S., and J. Jaya Kumari. "A survey on channel estimation techniques in mimo-ofdm mobile communication systems." International Journal of Scientific & Engineering Research 4.5 (2013): 1850-1855.

[47] Lang Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," in Proceedings of the IEEE, vol. 86, no. 10, pp. 1951-1968, Oct. 1998. doi: 10.1109/5.720247.

[48] S. Alireza Banani and R. G. Vaughan, "OFDM With Iterative Blind Channel Estimation," in IEEE Transactions on Vehicular Technology, vol. 59, no. 9, pp. 4298-4308, Nov. 2010. doi: 10.1109/TVT.2010.2080295.

[49] V. Choqueuse, A. Mansour, G. Burel, L. Collin and K. Yao, "Blind Channel Estimation for STBC Systems Using Higher-Order Statistics," in IEEE Transactions on Wireless Communications, vol. 10, no. 2, pp. 495-505, February 2011. doi: 10.1109/TWC.2010.112310.091576.

[50] G. Scarano, A. Petroni, M. Biagi and R. Cusani, "Second-Order Statistics Driven LMS Blind Fractionally Spaced Channel Equalization," in IEEE Signal Processing Letters, vol. 24, no. 2, pp. 161-165, Feb. 2017. doi: 10.1109/LSP.2016.2635034.

[51] G. Leus and M. Moonen, "Deterministic subspace based blind channel estimation for doubly-selective channels," 2003 4th IEEE Workshop on Signal Processing Advances in

Wireless Communications - SPAWC 2003 (IEEE Cat. No.03EX689), Rome, Italy, 2003, pp. 210-214. doi: 10.1109/SPAWC.2003.1318952.

[52] H. Murakami, "Deterministic Blind Channel Estimation for a Block Transmission System Using Fractional Sampling and Interpolation," in IEEE Transactions on Signal Processing, vol. 55, no. 10, pp. 4969-4978, Oct. 2007. doi: 10.1109/TSP.2007.896289.

[53] C. Yu, L. Xie and Y. C. Soh, "Blind Channel and Source Estimation in Networked Systems," in IEEE Transactions on Signal Processing, vol. 62, no. 17, pp. 4611-4626, Sept.1, 2014. doi: 10.1109/TSP.2014.2338837.

[54] Rashmi, N., and Mrinal Sarvagya. "A Survey on Channel Estimation for OFDM-IDMA Receivers Systems," in International Journal of Emerging Technology and Advanced Engineering, vol. 3, no. 1, January 2013.

[55] F. Xu, Y. Xiao and D. Wang, "Adaptive semi-blind channel estimation for massive MIMO systems," 2014 12th International Conference on Signal Processing (ICSP), Hangzhou, 2014, pp. 1698-1702. doi: 10.1109/ICOSP.2014.7015284.

[56] S. Abdallah and I. N. Psaromiligkos, "Semi-Blind Channel Estimation with Superimposed Training for OFDM-Based AF Two-Way Relaying," in IEEE Transactions on Wireless Communications, vol. 13, no. 5, pp. 2468-2467, May 2014. doi: 10.1109/TWC.2014.031714.130348.

[57] K. Mawatwal, D. Sen and R. Roy, "A Semi-Blind Channel Estimation Algorithm for Massive MIMO Systems," in IEEE Wireless Communications Letters, vol. 6, no. 1, pp. 70-73, Feb. 2017. doi: 10.1109/LWC.2016.2631535.

[58] E. Nayebi and B. D. Rao, "Semi-blind Channel Estimation for Multiuser Massive MIMO Systems," in IEEE Transactions on Signal Processing, vol. 66, no. 2, pp. 540-553, 15 Jan.15, 2018. doi: 10.1109/TSP.2017.2771725.

[59] Y. Chen, J. Schaepperle and T. Wild, "Comparing IDMA and NOMA with superimposed pilots based channel estimation in uplink," 2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Hong Kong, 2015, pp. 89-94. doi: 10.1109/PIMRC.2015.7343274.

[60] K. Struminsky, S. Kruglik, D. Vetrov and I. Oseledets, "A new approach for sparse Bayesian channel estimation in SCMA uplink systems," 2016 8th International Conference on Wireless Communications & Signal Processing (WCSP), Yangzhou, 2016, pp. 1-5. doi:10.1109/WCSP.2016.7752678.

[61] Y. Wang, S. Zhou, L. Xiao, X. Zhang and J. Lian, "Sparse Bayesian learning based user detection and channel estimation for SCMA uplink systems," 2015 International

Conference on Wireless Communications & Signal Processing (WCSP), Nanjing, 2015, pp. 1-5, doi: 10.1109/WCSP.2015.7341144.

[62] Y. Wang, X. Zhang, S. Zhou, J. Lian and L. Xiao, "User detection and channel estimation for SCMA uplink system in dispersive channel," 2016 IEEE International Conference on Communication Systems (ICCS), Shenzhen, 2016, pp. 1-5. doi: 10.1109/ICCS.2016.7833605.

[63] Y. Du et al., "Joint Channel Estimation and Multiuser Detection for Uplink Grant-Free NOMA," in IEEE Wireless Communications Letters, vol. 7, no. 4, pp. 682-685, Aug. 2018. doi: 10.1109/LWC.2018.2810278.

[64] G. Hannak, M. Mayer, A. Jung, G. Matz, N. Goertz, "Joint channel estimation and activity detection for multiuser communication systems", Proc. IEEE ICC, pp. 2086-2091, Jun. 2015.

[65] J. Heo, I. Jung, T. Kim, H. Kim and D. Hong, "Channel Estimation for Uplink SCMA Systems with Reduced Training Blocks," 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), Porto, 2018, pp. 1-5. doi: 10.1109/VTCSpring.2018.8417510.

[66] L. Sanguinetti, M. Morelli, U. Mengali, "Channel estimation and tracking for mc-cdma signals", Transactions on Emerging Telecommunications Technologies, vol. 15, no. 3, pp. 249-258, 2004, doi:10.1002/ett.970.

[67] E. Heo, N. Kim and H. Park, "Sparse Structure-Based Channel Estimation for Uplink SCMA System," in *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8037-8046, Sept. 2017, doi: 10.1109/TVT.2017.2679020.

[68] N. Nonaka, A. Benjebbour and K. Higuchi, "System-level throughput of NOMA using intra-beam superposition coding and SIC in MIMO downlink when channel estimation error exists," 2014 IEEE International Conference on Communication Systems, Macau, 2014, pp. 202-206. doi: 10.1109/ICCS.2014.7024794.

[69] A. Benjebbour, A. Li, Y. Kishiyama, H. Jiang and T. Nakamura, "System-level performance of downlink NOMA combined with SU-MIMO for future LTE enhancements," 2014 IEEE Globecom Workshops (GC Wkshps), Austin, TX, 2014, pp. 706-710, doi: 10.1109/GLOCOMW.2014.7063515.

[70] Y. Saito, A. Benjebbour, Y. Kishiyama and T. Nakamura, "System-Level Performance of Downlink Non-Orthogonal Multiple Access (NOMA) under Various Environments," 2015 IEEE 81st Vehicular Technology Conference (VTC Spring), Glasgow, 2015, pp. 1-5, doi: 10.1109/VTCSpring.2015.7146120.

[71] Z. Yang, Z. Ding, P. Fan and G. K. Karagiannidis, "On the Performance of Non-orthogonal Multiple Access Systems with Partial Channel Information," in IEEE Transactions on Communications, vol. 64, no. 2, pp. 654-667, Feb. 2016. doi: 10.1109/TCOMM.2015.2511078.

[72] J. Men, J. Ge and C. Zhang, "Performance Analysis for Downlink Relaying Aided Non-Orthogonal Multiple Access Networks with Imperfect CSI Over Nakagami-m Fading," in IEEE Access, vol. 5, pp. 998-1004, 2017. doi: 10.1109/ACCESS.2016.2631482.

[73] Z. Wei, D. W. K. Ng, J. Yuan and H. Wang, "Optimal Resource Allocation for Power-Efficient MC-NOMA With Imperfect Channel State Information," in IEEE Transactions on Communications, vol. 65, no. 9, pp. 3944-3961, Sept. 2017. doi: 10.1109/TCOMM.2017.2709301.

[74] F. Fang, H. Zhang, J. Cheng and V. C. M. Leung, "Energy-efficient resource scheduling for NOMA systems with imperfect channel state information," 2017 IEEE International Conference on Communications (ICC), Paris, 2017, pp. 1-5. doi: 10.1109/ICC.2017.7996360.

[75] K. Senel and S. Tekinay, "Optimal Power Allocation in NOMA Systems with Imperfect Channel Estimation," GLOBECOM 2017 - 2017 IEEE Global Communications Conference, Singapore, 2017, pp. 1-7. doi: 10.1109/GLOCOM.2017.8254919.

[76] F. Fang, H. Zhang, J. Cheng, S. Roy and V. C. M. Leung, "Joint User Scheduling and Power Allocation Optimization for Energy-Efficient NOMA Systems With Imperfect CSI," in IEEE Journal on Selected Areas in Communications, vol. 35, no. 12, pp. 2874-2885, Dec. 2017, doi: 10.1109/JSAC.2017.2777672.

[77] D. Fan, F. Gao, G. Wang, Z. Zhong and A. Nallanathan, "Channel Estimation and Transmission Strategy for Hybrid mmWave NOMA Systems," in IEEE Journal of Selected Topics in Signal Processing, vol. 13, no. 3, pp. 584-596, June 2019. doi: 10.1109/JSTSP.2019.2908652.

[78] AbdelMoniem, Mai, et al. "Enhanced NOMA System Using Adaptive Coding and Modulation Based on LSTM Neural Network Channel Estimation." Applied Sciences 9, no. 15:3022, (2019), doi: 10.3390/app9153022.

[79] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 5, pp. 720-741, May 1989, doi: 10.1109/29.17564.

[80] R. Cao, B. Liu, F. Gao and X. Zhang, "A Low-Complex One-Snapshot DOA Estimation Algorithm with Massive ULA," in IEEE Communications Letters, vol. 21, no. 5, pp. 1071-1074, May 2017, doi: 10.1109/LCOMM.2017.2652442.

[81] X. Rao and V. K. N. Lau, "Distributed Compressive CSIT Estimation and Feedback for FDD Multi-User Massive MIMO Systems," in IEEE Transactions on Signal Processing, vol. 62, no. 12, pp. 3261-3271, June15, 2014, doi: 10.1109/TSP.2014.2324991.

[82] H. Xie, F. Gao, S. Zhang and S. Jin, "A Unified Transmission Strategy for TDD/FDD Massive MIMO Systems With Spatial Basis Expansion Model," in IEEE Transactions on Vehicular Technology, vol. 66, no. 4, pp. 3170-3184, April 2017, doi: 10.1109/TVT.2016.2594706.

[83] A. Adhikary, J. Nam, J. Ahn and G. Caire, "Joint Spatial Division and Multiplexing—The Large-Scale Array Regime," in IEEE Transactions on Information Theory, vol. 59, no. 10, pp. 6441-6463, Oct. 2013, doi: 10.1109/TIT.2013.2269476.

[84] Soukayna Ghandour - Haidar, Laurent Ros, Jean-Marc Brossier," On the Use of First-order Autoregressive Modeling for Rayleigh Flat Fading Channel Estimation with Kalman Filter. Signal Processing, Elsevier", vol. 92, no. 2, pp.601-606, February 2012, doi: 10.1016/j.sigpro.2011.08.014.

[85] Gradshteyn, I. S., and I. M. Ryzhik. "Table of Integrals, Series and Products, ser. Alan Jeffrey Editor." 1994.

[86] Wee-Peng Ang and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms," in IEEE Transactions on Signal Processing, vol. 49, no. 4, pp. 805-810, April 2001, doi: 10.1109/78.912925.

Appendix

APPENDIX A: Results Generating Algorithms for each CE Approach

Algorithm 1: The CSS-LMS Algorithm using the all-known-pilots approach: MSE vs. $\boldsymbol{\mu}$

Initialization: $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{i,0} = 0, \forall i \in [1,2], h_{i,k}, \forall i \in [1,2]\ and\ \forall k \in [1 \dots K], \mu_v, \forall v \in [1, \dots, V],\ SNR = 30dB.$
for each μ_v in $\boldsymbol{\mu}$
  for each channel gain $h_{i,k}$
    Generate $y_{i,n}$ for current h_i
    for all the number of samples
      Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
      Solve (3.3.3) to obtain $\hat{h}_{i,n}$
      $n \coloneqq n + 1$;
    end for
    $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
  end for
  $\boldsymbol{amse}(v) \coloneqq \frac{1}{N} * sum(\boldsymbol{mse})$
end for
Output: $\boldsymbol{amse}$

Algorithm 2: The CSS-LMS Algorithm using the all-known-pilots approach: Convergence Characteristics

Initialization: $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{i,0} = 0, \forall i \in [1,2], h_{i,k}\ \forall i \in [1,2]$ and $\forall k \in [1 \dots K], SNR = 30\ dB$
for each channel gain $h_{i,k}$
    Generate $y_{i,n}$ for current h_i and SNR
    for all the number of samples n
      Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
      Solve (3.3.3) to obtain $\hat{h}_{i,n}$
      $n \coloneqq n + 1$;
    end for
    $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{mse}$;
end for
Average $\boldsymbol{mse}$ by the number until convergence runs
Output: Average MSE

Algorithm 3: The CSS-LMS Algorithm using the all-known-pilots approach: MSE vs. SNR

1: **Initialization:** $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{i,0} = 0, \forall i \in [1,2], h_{i,k}, \forall i \in [1,2]\ and\ \forall k \in [1 \dots K], SNR_v, \forall v \in [1, \dots, V]$
2: **for** each SNR_v in $\boldsymbol{SNR}$
3: **for** each channel gain $h_{i,k}$
4: Generate $y_{i,n}$ for current h_i and SNR_v
5: **for** all the number of samples
6: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
7: Solve (3.3.3) to obtain $\hat{h}_{i,n}$
8: $n \coloneqq n + 1$;
9: **end for**
10: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{mse}$;
11: end for
12: $\boldsymbol{amse}(v) \coloneqq \frac{\mathbf{1}}{\boldsymbol{N}} * sum(\boldsymbol{mse})$
13: **end for**
14: **Output:** $\boldsymbol{amse}$

Algorithm 4: The VSS-LMS Algorithm using the all-known-pilots approach: Convergence Characteristics

1: **Initialization:** $\alpha_1,\ \alpha_2, \hat{h}_{i,0} = 0, \forall i \in [1,2], h_{i,k}\ \forall i \in [1,2]$ and $\forall k \in [1 \dots K], SNR = 30\ dB, \mu_{i,n}, G_{i,n}, \forall n \in [1, \dots, N]$
2: **for** each channel gain $h_{i,k}$
3: Generate $y_{i,n}$ for current h_i and SNR
4: **for** all the number of samples n
5: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
6: Solve (3.4.5) to obtain $\hat{h}_{i,n}$, $G_{i,n}$ and $\mu_{i,n}$
7: $n \coloneqq n + 1$;
8: **end for**
9: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
10: **end for**
11: **Average** $\boldsymbol{mse}$ by the number until convergence runs
12: **Output: Average MSE**

Algorithm 5: The VSS-LMS Algorithm using the all-known-pilots approach: MSE vs. SNR

1: **Initialization:** $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{i,0} = 0, h_{i,k}, \forall i \in [1,2]$ and $\forall k \in [1 \dots K], SNR_v, \forall v \in [1, \dots, V], \mu_{i,n}\ \forall n \in [1, \dots, N], G_{i,n}$
2: **for** each SNR_v in $\boldsymbol{SNR}$
3: **for** each channel gain $h_{i,k}$
4: Generate $y_{i,n}$ for current h_i and SNR_v
5: **for** all the number of samples n
6: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
7: Solve (3.4.5) to obtain $\hat{h}_{i,n}$, $G_{i,n}$ and $\mu_{i,n}$
8: $n \coloneqq n + 1$;

9: **end for**
10: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
11: **end for**
12: $\boldsymbol{amse}(v) \coloneqq \frac{\mathbf{1}}{\boldsymbol{N}} * sum(\boldsymbol{mse})$
13: **end for**
14: **Output:** $\boldsymbol{amse}$

Algorithm 6: The VSS-LMS-AS Algorithm using the all-known-pilots approach: Convergence Characteristics

1: **Initialization:** $\alpha_1,\ \alpha_2, \hat{h}_{i,0} = 0, \forall i \in [1,2], h_{i,k}\ \forall i \in [1,2]$ and $\forall k \in [1 \dots K], SNR = 30\ dB, \mu_{i,n}, G_{i,n}, N_{i,n}, M_{i,n}, \Psi_{i,n}, \epsilon_{i,n}, \forall n \in [1, \dots, N]$
2: **for** each channel gain $h_{i,k}$
3: Generate $y_{i,n}$ for current h_i and SNR
4: **for** all the number of samples n
5: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
6: Solve (3.5.3), (3.5.4) and (3.5.5) to obtain $\hat{h}_{i,n}, G_{i,n}, \mu_{i,n}, N_{i,n}, M_{i,n}, \Psi_{i,n}$ and $\epsilon_{i,n}$
7: $n \coloneqq n + 1$;
8: **end for**
9: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
10: **end for**
11: **Average** $\boldsymbol{mse}$ by the number until convergence runs
12: **Output: Average MSE**

Algorithm 7: The VSS-LMS-AS Algorithm using the all-known-pilots approach: MSE vs. SNR

1: **Initialization:** $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{i,0} = 0, h_{i,k}, \forall i \in [1,2]$ and $\forall k \in [1 \dots K], SNR_v, \forall v \in [1, \dots, V], \mu_{i,n}, G_{i,n}, N_{i,n}, M_{i,n}, \Psi_{i,n}, \epsilon_{i,n}, \forall n \in [1, \dots, N]$
2: **for** each SNR_v in $\boldsymbol{SNR}$
3: **for** each channel gain $h_{i,k}$
4: Generate $y_{i,n}$ for current h_i and SNR_v
5: **for** all the number of samples n
6: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
7: Solve (3.5.3), (3.5.4) and (3.5.5) to obtain $\hat{h}_{i,n}, G_{i,n}, \mu_{i,n}, N_{i,n}, M_{i,n}, \Psi_{i,n}$ and $\epsilon_{i,n}$
8: $n \coloneqq n + 1$;
9: **end for**
10: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
11: **end for**
12: $\boldsymbol{amse}(v) \coloneqq \frac{\mathbf{1}}{\boldsymbol{N}} * sum(\boldsymbol{mse})$
13: **end for**
14: **Output:** $\boldsymbol{amse}$

Algorithm 8: The CSS-LMS Algorithm using the all-known-pilots approach: Convergence Characteristics

1: **Initialization:** $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{i,0} = 0, \forall i \in [1,2], h_{i,k}\ \forall i \in [1,2]$ and $\forall k \in [1 \dots K], SNR = 30\ dB$
2: **for** each channel gain $h_{i,k}$
3: Generate $y_{i,n}$ for current h_i and SNR
4: **for** all the number of samples n
5: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
6: Solve (3.3.3) to obtain $\hat{h}_{i,n}$
7: $n \coloneqq n + 1$;
8: **end for**
9: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
10: **end for**
11: **Average** $\boldsymbol{mse}$ by the number until convergence runs
12: **Output: Average MSE**

Algorithm 9: The CSS-LMS Algorithm using the averaging-sum approach: Convergence Characteristics

1: **Initialization:** $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{i,0} = 0, \forall i \in [1,2], h_{i,k}\ \forall i \in [1,2]$ and $\forall k \in [1 \dots K], SNR = 30\ dB, J = 15$
2: **for** each channel gain $h_{i,k}$
3: Generate $y_{i,n}$ for current h_i and SNR
4: Compute $z'_{i,n}$ using (4.2.3)
5: **for** all the number of samples n
6: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
7: Solve (4.3.2) to obtain $\hat{h}_{i,n}$
8: $n \coloneqq n + 1$;
9: **end for**
10: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
11: **end for**
12: **Average** $\boldsymbol{mse}$ by the number until convergence runs
13: **Output: Average MSE**

Algorithm 10: The CSS-LMS Algorithm using the case-wise CE approach: Convergence Characteristics

1: **Initialization:** $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{2,0} = 0,\ h_{2,k}$ and $\forall k \in [1 \dots K], SNR = 30\ dB, \boldsymbol{p_e} = [0, 10^{-3}, 4 * 10^{-3}]$
2: for each probability of error $\boldsymbol{p_e}(in)$ in $\boldsymbol{p_e}$
3: Generate $\bar{s}_1$ with the current probability of error.
4: **for** each channel gain $h_{2,k}$
5: Generate $y_{2,n}$ for current $h_{2,k}$ and SNR
6: **for** all the number of samples n
7: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
9: Solve (5.2.5) to obtain $\hat{h}_{1,n}$
10: $n \coloneqq n + 1$;
20: **end for**
21: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
22: **end for**
23: **Average** $\boldsymbol{mse}$ by the number until convergence runs, store as $\boldsymbol{Avmse}$
24: $\boldsymbol{PEmse}(:, \boldsymbol{in}) \coloneqq \boldsymbol{Avmse}$
25: **end for**
26: **Output:** $\boldsymbol{PEmse}$

Algorithm 11: The CSS-LMS Algorithm using the SIC-inspired iterative approach: Convergence Characteristics

1: **Initialization:** $\alpha_1 = 0.39,\ \alpha_1 = 0.61, \hat{h}_{2,0} = 0,\ \hat{h}_2^{(j)} = 0, \forall j \in [1 \dots J]\ h_{2,k}$ and $\forall k \in [1 \dots K], SNR = 30\ dB, p_e = [0, 10^{-3}, 6 * 10^{-3}, 10^{-2}]$
2: **for** each probability of error $\boldsymbol{p_e}(\boldsymbol{in})$ in $\boldsymbol{p_e}$
3: Generate $\bar{s}_1$ with the current probability of error.
4: **for** each channel gain $h_{2,k}$
5: Generate $y_{2,n}$ for current h_2 and SNR
6: **for** all the number of samples n
7: $Y_{2,n}^{(1)} \coloneqq y_{2,n}$;
8: **for** each refinement iteration *j*
9: Compute the instantaneous estimation error using (3.2.4) store in $\boldsymbol{imse}(n)$
10: **if** $j = 1$
11 Solve (5.3.5) to obtain $\hat{h}_2^{(1)}$
12: **else**
13: Solve (5.3.18) to obtain $\hat{h}_2^{(j)}$
14: **end**
15: $Y_2^{(j+1)} \coloneqq y_{2,n} - \hat{h}_2^{(j)} \sqrt{\alpha_1 P} \bar{s}_{1,n}$
16: $j \coloneqq j + 1$;
17: **end for**
18: $\hat{h}_{2,n} \coloneqq \hat{h}_2^{(J)}$;
19: $n \coloneqq n + 1$;
20: **end for**
21: $\boldsymbol{mse} \coloneqq \boldsymbol{mse} + \boldsymbol{imse}$;
22: **end for**
23: **Average** $\boldsymbol{mse}$ by the number until convergence runs, store as $\boldsymbol{Avmse}$

24: $\boldsymbol{PEmse}(:, in) \coloneqq \boldsymbol{Avmse}$
25: **end for**
26: **Output:** $\boldsymbol{PEmse}$