IMPROVED MULTIPLE INPUT MULTIPLE OUTPUT BLIND EQUALIZATION ALGORITHMS FOR MEDICAL IMPLANT COMMUNICATION

LEE ZHI HOU, MEng.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

APRIL 2015

Abstract

Medical implant sensor that is used to monitor the human physiology signals is helpful to improve the quality of life and prevent severe result from the chronic diseases. In order to achieve this, the wireless implant communication link that delivers the monitored signal to a multiple antennas external device is an essential portion. However, the existing conventional narrow band Medical Implant Communications System (MICS) has low data rate because of the bandlimited channel is allocated. To improve the data rate in the radio frequency communication, ultra-wide band technology has been proposed. However, the ultra-wide band technology is relatively new and requires living human to be the test subject in order to validate the technology performance. In this condition, the test on the new technology can rise ethical challenge. As a solution, we improve the data rate in the conventional narrow band MICS. The improvement of data rate on the narrow band implies the information bandwidth is larger than the allocated channel bandwidth, and therefore the high frequency components of the information can loss. In this case, the signal suffers the intersymbol-interference (ISI). Instead of that, the multiple antennas external device can receive the signal from other transmitting implant sensor which has the same operating frequency. As a result, the signal is further hampered by co-channel interference (CCI). To recover the signal from the ISI and CCI, multiple-input multiple output (MIMO) blind equalization that has source separation ability can be exploited. Cross-Correlation Constant Modulus Algorithm (CC-CMA) is the conventional MIMO blind equalization algorithm that can suppress ISI and CCI

and able to perform source separation. However, CC-CMA has only been analyzed and simulated in the modulation of Phase Shift Keying (PSK). The performance of CC-CMA in multi-modulus modulation scheme such as 4-Pulse-amplitude modulation (PAM) and 16-Quadrature amplitude modulation (QAM), which has higher data rate than PSK, has not been analyzed. Therefore, our work is to analysis and optimize CC-CMA on the multi-modulus modulation scheme. From our analysis, we found that the cost function of CC-CMA is biased cost function. Instead of that, from our simulation, CC-CMA introduces an unexpected shrinking effect whereby the amplitudes of the equalizer outputs have been reduced, especially in multi-modulus modulation scheme. This shrinking effect is not severe in PSK because the decision of a PSK symbol is based on phase, but not amplitude. Unfortunately, this is severe in multi-modulus modulation scheme. To overcome this shrinking effect in multi-modulus modulation scheme, we propose Cross-Independent Constant Modulus Algorithm (CI-CMA). Based on the convergence analysis, we identify the new optimum dispersion value and mixing parameter in CI-CMA. From the simulation results, we confirm that CI-CMA is able to perform equalization and source separation in the multi-modulus modulation scheme. In order to improve the steady state performance of CI-CMA, we perform the steady state mean square error (MSE) analysis of CI-CMA using the energy preservation theorem that was developed by Mai and Sayed in 2001, and our result is more accurate than the previous work. From our analysis, only the reduction in adaptation step size can reduce the steady state MSE, but it is well known that the MSE is indeed a tradeoff with the speed of convergence. Therefore without sacrificing convergence speed, our last effort is to propose hybrid algorithms. The hybrid algorithms are done by combining a new adaptive constant modulus algorithm (ACMA), a decision directed algorithm and a cross-correlation function. From the simulation results, we found that the hybrid algorithms can show low steady state error and thereby improve the reliability of the communication link. The main achievement of this thesis is the discovery of new dispersion value through the convergence analysis.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr Lim Wee Gin for his guidance and encouragement. I have learned so much from him. Without his patience and understanding, this thesis would never be completed.

I am also grateful to my thesis examiners, Dr Amin Malek Mohammadi and Ir. Dr. Tiong Sieh Kiong. I also would like to thank to my 1st and 2nd year internal accessor, Dr Khalid Al Murrani for his useful feedback.

I am greatly indebted to the Department of Electrical and Electronic Engineering, the University of Nottingham Malaysia Campus for awarding me a PhD scholarship and a golden opportunity to become a research assistant.

Finally, I would like to thank my family and friends for their support and encouragement.

Contents

List of Abbreviations

ACMA	Adaptive Constant Modulus Algorithm
ADC	Analog to Digital Convertor
ASK	Amplitude Shift Keying
AWGN	Additive White Gaussian Noises
BG	Benveniste Goursat Algorithm
BPSK	Binary Phase Shift Keying
BSS	Blind Source Separation
BSS-CMA	Blind Source Separation Constant Modulus Algorithm
BSS-MMA	Blind Source Separation Multi Modulus Algorithm
CC	Cross Correlation
CC-CMA	Cross-Correlation Constant Modulus Algorithm
CCI	Co-Channel Interference
CC-SCMA	Cross-Correlation Simplified Constant Modulus Algorithm
CDMA	Code Division Multiplexing Access
CI	Cross Independent
CI-CMA	Crosee-Independent Constant Modulus Algorithm
CMA	Constant Modulus Algorithm
dB	Decibel
DD	Decision Directed
DNA	Deoxyribonucleic Acid
DPSK	Differential Phase Shift Keying
FA	Factor Analysis
FIR	Finite Impulse Responses
FSK	Frequency Shift Keying
HBC	Human Body Communication
HF	High Frequency
i.i.d.	Identical and Independent Distributed
ICA	Independent Component Analysis

IEEE	Institute of Electrical and Electronics Engineers
IF	Intermediate Frequency
IR	Impulse Radio
ISI	Inter-Symbol Interference
IT	Residual Interference
Kur	Kurtosis
L.H.S.	Left Hand Side
LAN	Local Area Network
LMS	Least Mean Square
MAP	Maximum A Posteriori
MATLAB	Matrix Laboratory
MB-OFDM	Multi Band Orthogonal Frequency-Division Multiple
MCIBG	Modified Cross Independent Benveniste-Goursat Algorithm
MCIRBA	Modified Cross Independent Reliability Based Algorithm
MCISAG	Modified Cross Independent Stop and Go Algorithm
MICS	Medical Implant Communications System
MIMO	Multiple Inputs Multiple Outputs
MIMO-CMA	Multiple Inputs Multiple Outputs Constant Modulus Algorithm
MSE	Mean Square Error
NB	Narrow Band
OFDM	Orthogonal Frequency-Division Multiple
O-MMA	Orthogonal Multi Modulus Algorithm
OOK	On/Off Keying
PAM	Pulse Amplitude Modulation
PAPR	Peak-to-Average Power Ratio
PCA	Principle Component Analysis
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
R.H.S.	Right Hand Side
RBA	Reliability Based Algorithm
RF	Radio Frequency
SAG	Stop-And-Go Algorithm Decision Directed Algorithm
SGA	Stochastic Gradient descent Algorithm
Sim	Simulation
SISO	Single-Input Single-Output
Th	Theoretical
UHF	Ultra High Frequency
UWB	Ultra Wide Band
VHF	Very High Frequency
WLAN	Wireless Local Area Network

List of Figures

List of Tables

Chapter 1

Introduction

Chronic diseases, also known as non-communicable diseases, are long-lasting illnesses that may lead to death and disability. These diseases are frequently preventable or controllable through early detection, medical treatment and proper life style. Examples of chronic diseases are heart disease, stroke, diabetes, cancer and etc. Unfortunately, World Health Organization has reported that chronic diseases caused 60 percent and 68 percent of all deaths in 2002 and 2012, respectively [3]. This indicates that 8 percent increment in 10 years and the percentage of death caused by chronic diseases probably increases to 76 percent in 2022. Too late in chronic diseases detection is the common reason that causes the death. Therefore, in order to reduce the percentage of death caused by chronic diseases and improve the quality of life, wireless medical sensor can be exploited to monitor the patient and take necessary action once abnormal condition is detected. More precisely, the wireless medical sensor can be used to collect various real time physiological signals such as heart rhythm, blood glucose level, body temperature, blood pressure and etc, without geographical limitation. Through telecommunication network, doctor can access the physiological signals to determine any possible harmful symptom of patient. If harmful symptom is detected by doctor, the patient can be informed to be admitted to hospital for further tests or treatment to avoid any dreadful event.

In general, the wireless medical sensor can be divided into wearable sensor and implant sensor, which are located on-body and in-body, respectively. Wireless wearable medical sensor, is placed on the body or patient's skin in non-invasive way and it is removable. The applications of this type sensor include blood pressure measurement, body temperature monitoring, heart rhythm monitoring [4], asthma monitoring [5] and etc. Instead of that, the implant sensor is an electronic sensor that is placed inside human body via surgery or swallowing. The applications of the implant sensor include blood glucose level monitoring [6], cardiovascular system monitoring [7], cancer detector, capsule endoscopy [8] and etc. Wireless communication is the essential element of the above sensors. Among the wearable and implant sensors, the wireless communication link of the implant sensor is more challenging than the wearable sensor because the implant sensor is located inside the human body. Therefore, we will focus on the improvement of the wireless communication link for the implant sensor.

1.1 Background

The wireless implant communication link, which connects an implant sensor to a multi-antennas external hand held device, is important in order to ensure the monitored data can be received by doctor in different geographic area. Due to the limited power in the implant sensor, the implant sensor is not able to directly send the monitored data to the cellular network or local area network (LAN). Therefore, the external device, which is required to be placed within 2 meter distance from the implant sensor, is functioned as a gateway because the external device receives the monitored data from the implant sensor and then re-transmits the data to the cellular or LAN. In general, the wireless implant communication is a two way communications link. Therefore, it can be divided into uplink and downlink, which are the communication link from implant to the external device and from the external device to implant, respectively. As a monitoring device, uplink is expected to have high data traffic than downlink because the implant sensor frequently sends the monitored signal to the external device from time to time. In contrary, downlink is expected to have low data

traffic because it is used to configure or reprogram the implant. Therefore, we will put our focus on the improvement of the uplink implant communication.

In the literature, the implant communication can be accomplished through Medical Implant Communications Systems (MICS) standard, Ultra-wideband (UWB) Communication and Human Body Communication. The MICS is a narrowband radio frequency (RF) communication that has been used in cardiac pacemaker and defibrillator to help heart disease patient since 1999 [9, 10, 11]. However, due to the limited channel bandwidth of 300 kHz, MICS has relatively low data rate. To overcome the data rate limitation, UWB which has the channel bandwidth of 500 MHz has been proposed. Instead of that, UWB is relatively immune to frequency selective channel fading and noise [8]. In spite of that, the application of UWB for implant sensor communication has encountered difficulty. Currently, the performance of UWB only has been tested on the simulations and models [11, 8]. The actual result on living human is still unknown. In order to conduct this test, living human is required to be the test subject. Therefore, this is unethical and probably illegal. As a result, the results of simulations and models of UWB are hardly to be validate. Instead of that, human body communication is a relatively new type of non-RF communication where human body is used as the transmission medium. This type of communication provides high security and high fidelity because the transmission medium is not shared by others. However, this type communication is unfriendly-user in setup and has a very short transmission range where it is limited to about 100 centimeter distance. Furthermore, the application of human body communication for the implant communication may rise ethical challenge because living human is required to be the test subject.

Due to the reason the application of UWB and human body communication for the implant communication may encounter the ethical challenge, we put our focus on the data rate improvement on the narrowband MICS.

1.2 Practical Challenges

Currently, the operating frequency range of implant communication has been defined from 402 to 405 MHz, which is also known as Medical Implant Communication Service band with the maximum channel bandwidth of 300 kHz since 1999, even 13 years earlier than the first publication of IEEE Body Area Network 802.15.6 [9, 10, 11]. The MICS, which is currently belonged to IEEE Body Area Network 802.15.6, also defines to the uplink and downlink communications. Therefore, the application of implant communication in MICS is legally defined by the document. In order to improve the data rate, the uplink communication of the implant communication can encounter some unique challenges and the challenges are described as following.

i Limited power resource of implant

The implant device requires a safe and reliable power source. In this case, battery can be a good choice since battery has been used in other in-body equipment such as pacemaker for many years and no severe issue has been reported. However, due to the reason that the battery together with implant is located inside human body, the battery is hardly to be recharged or replaced. Therefore, the battery has limited life time and thus it can be considered as expensive resource.

ii Limited computation power and memory of implant

Memory and computation operation consume power and cost. Therefore, in order to save power resource and memory cost, the computation operation should be designed as simple as possible and the memory size should be small enough to achieve the basic tasks. Therefore, in this computation and memory limitation condition, complicated digital signal processing is not encouraged to be performed on the implant.

iii Intersymbol interference (ISI) due to bandlimited channel

Currently, the maximum channel bandwidth of 300 kHz is allocated for implant communication and thus, the bandwidth of the information signal is traditionally restricted below

300 kHz to avoid any undesired distortion. However, the restriction on the information bandwidth also limits the data rate. In order to increase the data rate, we suggest that the information bandwidth to be increased more than the allocated channel bandwidth. In this case, to avoid the transmit signal does not exceed the allocated channel bandwidth, the information signal is passed though a bandlimited filter before the signal is transmitted. Obviously, high frequency components of the information signal are attenuated because the bandlimited filter has lower bandwidth compared to the bandwidth of the information signal. In this situation, due to the high frequency components of the information signal are lost, the receiving external device receives the information signal that is corrupted by intersymbol interference. The ISI can increase the error rate and cause the whole communication link becomes unreliable. Therefore, in order to achieve higher data rate and reduce the error rate in this new approach, the receiving external device is expected to equip with signal processing method that can cope the ISI.

iv Co-channel Interference (CCI)

In a special scenario such as two patients with the same operating frequency implant devices are standing side by side, the receiving node can receive a distorted signal which is the combined signal from the desired and undesired implant devices. This type signal distortion is called co-channel interference because two or more transmitting implant devices are operating in the same channel. CCI can strongly reduce the receiver performance. To overcome this problem, the receiving external device should have a signal processing method that can suppress CCI and perform source separation.

v Expensive communication link

In order to effectively setup the implant, the costs that include devices cost, deployment cost, surgery cost, medical cost and monitoring cost are expected to be paid. Instead of that, the limited battery life time has constraint the duration of the communication link to a finite time. Therefore, from the perspective of the above costs and power resource, the

implant communication link is expensive and thus the method that can produces high data throughput is required.

1.3 Practical Objectives and Proposed Solutions

Based on the above challenges, we can see that a new solution is demanded and the solution should able to achieve the following objectives.

- The solution should have light and moderate signal processing on the transmitting implant and the receiving external device, respectively.
- The solution should have the ability to suppress ISI and CCI simultaneously, and also able to perform source separation.
- The solution is able to improve data throughput.

In order to achieve the objectives, multiple inputs multiple outputs (MIMO) blind equalizer with source separation ability, is called MIMO blind equalizer, is proposed as the solution on the receiver side. Instead of that, pulse-amplitude modulation (PAM) and quadrature amplitude modulation (QAM) are suggested to be the modulation scheme. The reasons of the proposed solutions are justified as below.

- **MIMO blind equalizer does not require the transmitting implant to perform any complicated operation.**

  In order to established communication link, some common wireless technologies such as Orthogonal Frequency-Division Multiple (OFDM) and Code Division Multiplexing Access (CDMA), require the transmitter to perform Inverse Fast Fourier Transform operation and code spreading operation, respectively. Therefore, these operations can burden the transmitting implant. Instead of that, OFDM and CDMA are the wide band technologies. Thus, the performance of the technologies can degrade if they are applied in the narrow band.

Therefore, MIMO blind equalizer is suggested to be applied at the receiving external device because it does not require the above special operations to be performed on the transmitting implant. Due to the receiving external device is located outside human body, the battery replacement or recharge is relatively easy, and thus this allows MIMO blind equalizer to be applied.

- **PAM and QAM modulation schemes can improve the data throughput.**

  MICS communication has relatively low data throughput because of limited channel bandwidth (i.e. up to 300 kHz) and bandwidth inefficiency. The limited channel bandwidth issue has been explained. Now, we put focus on the bandwidth inefficiency issue. In the literature of MICS, the relatively simple communication modulation schemes such as on/off keying (OOK), amplitude shift keying (ASK) [12, 13], frequency shift keying (FSK) [14, 15, 16] and differential phase shift keying (DPSK) [10] can be found. Compared to multi-modulus PAM and QAM, the above modulations schemes are relatively bandwidth inefficiency and thereby can the data throughput is low.

1.4 Research Background

Multiple-input multiple-output (MIMO) technology such as spatial division multiple access (SDMA) has attracted strong interest in telecommunications field, because of the higher data throughput compared to single-input single-output (SISO) technology [17, 18, 19]. However, the signal recovery in the MIMO receiver is more difficult compared to the SISO receiver because two primary obstacles need to be overcome in order to retrieve the all the input signals. Firstly, the signals may suffer from ISI due to the bandlimited channel. The resulting channel is thus commonly known as a frequency selective channel where signals of different frequencies will suffer different levels of attenuation. Secondly, due to cochannel system sources, the received signals are overlapped versions of multiple source signals, i.e., a phenomenon called CCI. To overcome ISI and CCI simultaneously,

MIMO equalization which equips with open eye and source separation abilities is used.

Generally, there are two type equalization approaches such as trained equalization and blind equalization. Trained equalization requires the transmitter periodically sends training sequences to the receiver in order to open the channel eye. Least Mean Square (LMS) is an example of trained equalization algorithm [20]. Normally, trained equalization can rapidly suppress interferences. However, the periodical transmission of the training sequence will reduce the data throughput, and therefore blind equalization is developed. The blind equalizer exploits the statistical information of transmitted signals to recover signal instead of training signals. Thus, blind equalization is a good candidate when the communication link is expensive. In practice, blind equalizer depends on its algorithm to compute the optimum coefficient. Therefore, blind equalization algorithm is the key to determine the performance of a blind equalizer.

Many SISO blind equalization algorithms have been developed in the literature. Some blind equalization algorithms that are supported by theoretical analysis are Constant Modulus Algorithm (CMA) [21], Sato algorithm [22], Multimodulus algorithm [23] [24] and Shalvi Weistein Algorithm [25]. Among these algorithms, CMA is widely recognized as the most common algorithm due to its simplicity and its strong ability in open eye even in a severe channel. However, the SISO blind equalization algorithm was developed based on the SISO assumption, thus it may not optimum in MIMO case. To overcome the limitation, MIMO blind equalization algorithms are proposed. In the literature, many MIMO algorithms have been developed but not all algorithms are equipped with source separation ability. In order to identify this function of the algorithm, based on the presented results, we divide the MIMO algorithms into source separation algorithm, pure equalization algorithm, single task algorithm and two tasks algorithm, and the definitions are described as follows,

- **Source separation algorithm:** It is a blind algorithm that can perform source separation but it was designed and tested solely in CCI without ISI. This algorithm also known as blind source separation (BSS). [26, 27, 28, 29, 30, 31, 32, 33, 34]

- **Pure equalization algorithm:** It is a blind algorithm that can suppress interferences but cannot perform source separation. [35, 36]

- **Single task algorithm:** It is a blind algorithm that can perform solely one task in a time, either equalization or source separation but not the both. [37, 38]

- **Two tasks algorithm:** It is a blind algorithm that can perform two tasks of equalization and source separation simultaneously. [39, 40, 41, 2, 31, 42, 43, 44, 45, 46, 47]

Obviously, in order to achieve our objective of equalization and source separation, the two tasks algorithm are the candidate of MIMO equalizer. Instead of that, the detail description of the above algorithms can be found in the chapter of Literature Review.

The two task algorithm can be divided into orthogonal constraint cost function approach or non-constraint cost function approach. In the orthogonal constraint cost function approach, the equalizer is required to minimize a high order cost in order to open the channel eye, and perform matrix decomposition in order to ensure source separation [39, 40, 41]. However, this approach requires the channel order or channel length to be known. In practice, due to the channel is unknown, it is difficult to obtain the channel order. Furthermore, because matrix decomposition requires large computation efforts and large storage, the computation complexity of this method is high. Instead of that, noise has not been considered in the design. Therefore, we will pursue non-constraint cost function approach.

Cross-correlation Constant Modulus Algorithm (CC-CMA) are the non-constraint cost function approach that can mitigate both ISI and CCI and perform source separation. In this algorithm, the CMA which has the open eye ability is combined with a cross-correlation (CC) cost which penalizes correlations between the source signals and including their delayed versions thereby separating them. In fact, CC-CMA was independently proposed by Touzni et al [2] and Papadias et al [31]. CC-CMA has widely been accepted as a good MIMO blind equalization algorithm and the CC-CMA has been extended and enhanced by [48], [44], [49] and [50] based on the assumption

that the jointly goal of open eye and source separation always can be achieved perfectly.

1.5 Research Limitations and Objectives

Some research limitations of CC-CMA are addressed and the research objectives are stated as follows.

1.5.1 Robust to 4-PAM and 16-QAM

CC-CMA is a MIMO equalization algorithm that can be used to open the channel eye and perform source separation. Therefore, it is a potential candidate to be selected as the MIMO equalization algorithm for uplink implant communication. However, the theocratical analysis has only been done and tested for 2-Phase Shift Keying (PSK) or 4-PSK [2, 50, 43] whilst the theocratical analysis for higher order modulation schemes such as 4-PAM and 16-QAM are unknown. The 4-PAM and 16-QAM have different statistical properties with PSK, thus the theocratical analysis for PSK cannot be applied in the higher order modulation scheme. Therefore, a straight forward migration of CC-CMA to higher order modulation schemes are risky because the optimum parameters of CC-CMA for higher order modulation schemes are unclear. To overcome this issue, a new MIMO blind equalization algorithm that is supported by theocratical analysis and robust to 4-PAM and 16-QAM is required.

1.5.2 Superior Steady State Performance is required

Steady state is the condition that the equalizer reaches the stable condition and thereby the statistical properties of the equalizer output do not show any significant change. In other words, if the design is correct, the equalizer is able to produce a sufficiently low error rate in the steady state. Therefore, the performance of the equalizer in the steady state always has been used to imply the reliability of the communication link.

In order to obtain a reliable communication link, low steady state error value is always demanded. The steady state performance of MIMO blind equalization algorithm has been evaluated in the PSK cases [50], but no relevant information for 16-QAM case can be found. Therefore, the factors that can affect the steady state error for 16-QAM case are required to be found. Instead of that, a new algorithm that can improve the steady state performance is required.

1.6 Contributions

The main objective of the thesis is to develop MIMO blind equalization algorithms that can suppress ISI and CCI and automatically ensure all source sequences are retrieved without repetition. The contributions of the thesis are summarized as following.

1. **Identified the cost function of the CC-CMA is a biased cost function**

   We have identified the cost function of the CC-CMA, which was the widely accepted unbiased cost function, is a biased method and also we have established some mathematic proofs to prove the bias of the cost function of the CC-CMA. This finding has explained the shrinking effect and the contrary between the convergence analysis in [2] and its results. Furthermore, since the bias of the CC-CMA cost function has not been realized, some researches such as [51], [52] and [53] has claimed that the shrinking effect is unsolvable but only can be mitigated in a limited parameter range such as small number of source or small number of delay spread. Otherwise, the output can shrink to zero value and this implies symbol retrieval failed. This contribution can be found in Chapter 3.

2. **Proposed Cross Independent Constant Modulus Algorithm (CI-CMA)**

   In this thesis, we have proposed an unbiased cost function, CI-CMA to effectively solve the shrinking effect which was previously left unsolvable. Due to it is a new algorithm, we perform convergence analysis to confirm its blind source separation and open eye abilities. We emphasize that our works do not ignore the source statistic in the convergence analysis,

while the previous works such as [51], [52],[2] [53] have always ignored the source statistic for analysis simplicity. Specifically, the source is always assumed to be Binary Phase Shift Keying (BPSK) modulation by the previous works. With the presence of source statistic which we consider in our approach (therefore our result can be extended to any general modulation scheme), the bias that exists in the CC-CMA cost function becomes even clearer, even though the analysis become much more complicated. In our approach, a new dispersion constant value is determined in the CI-CMA to compensate the bias offset in the CC-CMA. Furthermore, the shrinking effect which is due to the bias cost function is simultaneously solved by the new dispersion constant. This contribution also can be found in Chapter 3.

3. **Perform Steady State Mean Square Error (MSE) analysis on CI-CMA**

   We have analytically predicted mean square error steady state (MSE) condition on the new unbiased algorithm, CI-CMA. The analytical MSE curve is closer to the practice MSE curve in comparison to the previous works [50]. It is worth to mention that MSE analytical curve is classically derived based on energy preservation theorem [54] that assumes an algorithm is unbiased. However, since the CC-CMA appears to be biased, the MSE analysis performed is not exactly accurate, because the criteria of energy preservation theorem would have then been violated. Furthermore, we have performed closed form mathematical manipulations on some key equations but the previous works have approximated some key equations without explicitly mathematic proofs given. Finally, with our analytical MSE equation has clearly determined the factors and how do these factors influence MSE value. This contribution can be found in Chapter 4.

4. **Proposed Hybrid Algorithms**

   Following our MSE analysis, we realize that the steady state MSE can only be reduced by minimizing the adaptation step size, which in doing so will slow the convergence of the algorithm. Therefore, three hybrid algorithms, which are Modified Cross Independent Benveniste-Goursat Algorithm (MCIBG), Modified Cross Independent Reliability Based Al-

gorithm (MCIRBA) and Modified Cross Independent Stop and Go Algorithm (MCISAG), are proposed to improve the steady state performance. The hybrid algorithms are the combination of a new adaptive constant modulus algorithm (ACMA), a decision-directed algorithm and a cross-correlation function. This contribution can be found in Chapter 5.

1.7 Structure of the thesis

The structure of the thesis is described below. Chapter 2 is the chapter of Literature Review that highlights different types of wireless medical sensor and applications, discusses on different types of implant communication system, provides the background of MIMO blind equalization, states the system model, defines a good MIMO equalization condition and reviews some related algorithms. Chapter 3 presents the problems of the CC-CMA, introduces the CI-CMA to overcome the problems and perform convergence analysis on the CI-CMA. Then, steady state MSE of the CI-CMA is analytically established in Chapter 4. Chapter 5 shows the new MIMO hybrid algorithms and the comparisons among the new algorithms. Finally, Chapter 6 describes the conclusion and future works.

Chapter 2

Literature Review

2.1 Introduction

This chapter begins with a general introduction of chronic diseases and wireless medical sensor and a discussion on different types of medical sensors from Section 2.2 to 2.5. The review will focus on the communication of implant sensor, and thus the discussion of current and potential implant communications can be found in Section 2.6. The chapter continues to highlight the interferences issue and MIMO channel equalization in Section 2.7. The typical system model and assumptions will be presented in Section 2.8. The theoretical background of MIMO equalizer can be found from Sec 2.9 to 2.11. Section 2.12 presents the performance measurements of MIMO equalizer. The chapter continues a literature review of blind equalization algorithms from Section 2.13 to 2.15.

2.2 Chronic diseases and wireless medical sensor

Chronic disease is a long-lasting illness that may cause death or disability if the disease has not been controlled well. The disease cannot be spread through virus or bacteria and the factors that can surely cause the disease still remain unknown. However, researches have shown that obesity, physical inactive, uncontrolled in smoking and drinking alcohol, insufficient nutrition, pollution and certain Deoxyribonucleic Acid (DNA) have strong correlation to chronic disease [55, 3, 56, 57]. Examples of chronic disease are heart disease, stroke, diabetes, cancer and etc. Early detection and frequent monitoring of chronic disease is often helpful to avoid severe result.

In order to achieve the detection and monitoring of chronic disease, the human body physiological signals, such as heart rhythm, blood glucose level, body temperature, blood pressure and etc, are always required to be observed for sufficient long period. In order to obtain the physiological signals, wired medical sensor has conventionally been used. However, the wired medical sensor is heavy and big equipment, and thereby it can restrict patient's movement. Therefore, battery-operated wireless medical sensor, which is relatively small and light, is developed to perform the similar task. The wireless medical sensor is not solely a sensor, but also has been integrated with processor, memory and radio frequency communication technology [58]. Hence, for the non-emergency case, the wireless medical sensor allows the patient to be home monitored and is helpful to reduce the face-to-face consultation times [59, 60, 61]. In this case, the resources such as patient's time and hospital space can be saved. In general, wireless medical sensor can be divided into wearable sensor and implant sensor, which are located on-body and in-body, respectively.

2.3 Wearable sensor

Wireless body surface sensor, also known as wireless wearable medical sensor, is placed on the body or patient's skin in non-invasive way and it is removable. Normally, the suspicious patient, who is suspected with certain disease and required sufficient long period physiological signals to

be confirmed, is advised to wear up this type sensor for several days or months. In contrast to the traditional wired medical sensor, wireless sensor does not restrict patient movement and patient is allowed to go home and work. In this case, since the patient does not need to be admitted in hospital immediately, hospital indirectly can save some resource as well. The applications of this type sensor include blood pressure measurement, body temperature monitoring, heart rhythm monitoring [4], asthma monitoring [5], sleep disorder monitoring [62], breathing monitoring [63], dementia brain disease detection [64, 65] and etc.

The wireless communication of wearable sensor is an essential element. In general, the wireless wearable sensor communication can be divided into narrowband communication, UWB communication and human body communication. The narrowband communication of wireless body surface has the bandwidth range from 300 kHz to 1 MHz and operates in various frequency bands that are within High Frequency (HF), Very High Frequency (VHF) and Ultra High Frequency (UHF). Except for the special case of 2.4GHz band (within UHF) with 10 MHz bandwidth, all the aforementioned operating frequency bands are licensed bands, and thereby the wireless communication link is legally protected from interferences by other wireless communications such as television signal or cellular signal. On the contrary, the 2.4GHz band is an unlicensed band and many wireless communications such as Wi-Fi or Wireless Local Area Network (WLAN) and Bluetooth are operating in this band. In this condition, the wireless body surface sensor that operates in this band is probably interfered by other wireless products. In contrast to narrowband communication, UWB communication is less susceptible to noise and interference. According to IEEE 802.15.6, the operating frequency range of UWB sensor is located within Microwave Frequency band which is 3.1 to 10 GHz with bandwidth nearly 500 MHz [10, 9]. Due to the bandwidth of UWB is higher than narrowband, the sensor that operates in UWB has higher data rate. Impulse radio are highly suggested to be the wireless technology for UWB wireless body surface sensor [66, 67, 68, 69, 70, 71]. Instead of that, human body communication performs data transfer by touching the wearable sensor and human body is used as the communication channel. This communication provides high

security benefit because the information has not been transfer to the air [72, 73, 74].

2.4 Implant sensor

Implant, also known as implant sensor or in-body medical sensor, is an electronic sensor that is placed inside human body via surgery or swallowing. Recent advanced in miniature technology that reduces the size and weight of traditional medical sensor has made implant becomes a realistic device because the device becomes small enough to be fitted into an organ [75]. Instead of that, with the advanced integrated circuit technology, implant is not just a solely sensor, but also has the ability to compute, memorize and perform wireless communication. Therefore, implant is expected to capture human physiology signals or in-body images and then send the signals or images to a multi-antennas external device, which is located outside the body.

2.5 Applications of implant sensor

Implant sensor has many potential medical applications. Some common application includes blood glucose level monitoring, cardiovascular system monitoring, cancer detector and capsule endoscopy are described as below.

2.5.1 Blood glucose level monitoring

Blood glucose level monitoring is critically important for diabetes patient to ensure the effectiveness of insulin dose and thereby avoid excessive blood sugar level which probability leads to complications. The complications include blindness, kidney damage, nerve damage and others. Traditionally, in order to get a blood glucose level reading, a blood sample is obtained by piercing on the finger and then the blood sample is analyzed chemically or electronically by a blood glucose meter. A severe diabetes patient, who requires insulin dose, may need to do the above blood glucose test 3 to 10 times a day. In this condition, the repetitive piercing process for the same area

over several years can damage the nearby tissues and blood vessels. Therefore, as an alternative solution, a implant to monitor the blood glucose level can be deployed into body. In this case, the implant can automatically update the blood glucose level reading to the external device without required any piercing. This reading is useful for the patient to makes decision on the amount of insulin dose, the types of meal and physical activities [76, 6].

2.5.2 Cardiovascular system monitoring

Ischemic and Arrhythmia are the two common heart diseases that may lead to dangerous events such as death, stroke and heart failure. Firstly, in Ischemic heart disease, the blood flow that supplies oxygen to the heart is partially or fully blocked by plaques, such as cholesterol and etc, and then the heart tissues can die because the heart tissues cannot obtain sufficient oxygen supply. Secondly, Arrhythmia is a heart disease that the heart occasionally beats too fast, too slow or irregular heart rhythm. In this unusual heart rhythm event, the blood pressure level is expected to be abnormal and thereby Arrhythmia has the potential to cause stroke and heart failure. For these heart diseases, doctor believes that occasional abnormal heart rhythm may be shown up before the dangerous events. Therefore, for prevention purpose, an implant for heart rhythm monitoring can be used to record and detect the occasional abnormal heart rhythm [7, 77, 78, 16]. Furthermore, a sufficient long record of heart rhythm that is generated by the implant allows doctor has more information to decide the optimum medical treatment.

2.5.3 Cancer detector

According to World Cancer Report 2014, cancer caused about 8.2 million deaths in 2012 [79]. The cancer death cases are expected to increase to 13.2 million in 2030 [80]. This has aroused the research on the prevention death from cancer. Cancer may not necessary cause death, some typical type cancers such as breast cancer, cervical cancer, oral cancer and colorectal cancer has high chance to be controlled or cured if the cancers are detected and treated in the early stage[81].

Therefore, cancer detection in the early stage is helpful to increase the survival rate. Past studies have shown that nitric oxide has an important role in the initiation and growth of cancer cell, and thus nitric oxide can be exploited to detect cancer cell [82]. For healthy people, nitric oxide is a signaling molecule that is important to deliver messages between cells, brain and immune system in order to help the immune system to reduce inflammation, kill bacteria, prevent tumor and etc [83]. On the other hand, for cancer patient, depend on the cancer types, many studies have indicated that unusual saturation level of nitric oxide can be observed in the patient's blood. For example, increased amount of nitric oxide has been observed in breast cancer patient [82]. Therefore, implant that monitors nitric oxide in the blood in long term can be used to identify the presence of cancer cell. This type implant can be applied for the people who has the frequent record on certain type cancer in the family history. Instead of cancer detection, this implant also provides a way for cancer study on alive human.

2.5.4 Capsule endoscopy

Capsule endoscopy can be used to capture the images along digestive tract in order to allow the doctor to diagnosis digestive tract tumor, ulcer or bleeding. In contrast to the above implants, capsule endoscopy is implant that stays inside human digestive tract for about 8 hours and then is expected to be flushed away naturally. More precisely, capsule endoscopy is a special pill which consists of camera and light, and it is used to capture the images of the entire digestive tract in real time. In order to capture the images in the digestive tract, patient is advised to swallow the pill and then the pill is moved by biological peristalsis. In this condition, the capsule endoscopy can capture the images of entire digestive tract, and the captured images are immediately sent to external device and then forwarded to doctor's computer through internet or WLAN. Therefore, doctor can diagnoses any tumor or disease in the digestive tract by observing the images. In contrast to the traditional endoscopy, without the restriction of wire, the wireless capsule endoscopy can observe the entire digestive tract [84, 85, 8, 86, 16].

2.6 Current and Potential Implant Communications

In order to connect implant sensor to external device, the wireless implant communication is an essential part. In general, the wireless implant communication is a two way communications link. Therefore, it can be divided into uplink and downlink, which are the communication link from implant to external device and from external device to implant, respectively. Currently, MICS standard is the only legal standard that can be used for implant two way communications. Instead of that, UWB and Human Body Communication for implant communication have been proposed by researchers. Therefore, MICS standard, UWB and Human Body Communication will be reviewed. Finally, a summary will be found the end of this section.

2.6.1 MICS standard

This narrowband RF communication standard has been established since 1999 for cardiac pacemaker and defibrillator to help heart disease patient. The RF operating frequency range is 402 MHz to 405 MHz with maximum channel bandwidth 300 kHz and maximum power of 2 microwatt, which roughly covers distance of 1 to 2 meter. Currently, this is the only approved communication standard for implant sensor on living human body [87, 88, 9, 10, 11, 89].

Compared to other communications, this type communication has two limitations. First, this communication has relatively low data rate because of limited channel bandwidth (i.e. up to 300 kHz) and bandwidth inefficiency. In comparison to 16-QAM modulation scheme, it is bandwidth inefficiency because relatively simple communication modulation schemes such as on/off keying (OOK), amplitude shift keying (ASK) [12, 13], frequency shift keying (FSK) [14, 15, 16] and differential phase shift keying (DPSK) [10] can be found. Second, due to low data rate, this type communication requires longer transmission time to transmit same amount of data, thus it can exhaust the battery faster than any other fast communications.

2.6.2 UWB communication

To overcome the low data rate issue in MISC, RF UWB communication for implant sensor communication has been proposed. The UWB has the operating frequency range from 3.1 to 10.6 GHz and the maximum channel bandwidth about 500 MHz. Instead of high data rate due to high bandwidth, UWB is relatively immune to frequency selective channel fading and noise [8].

However, the application of UWB for implant sensor communication has some difficulties. First, due to the reason that UWB is relatively new technology and such wide spectrum resource is an expensive resource, it may be difficult for all countries to allocate this band for this communication. Second, the UWB experiment may raise ethical and juridical issue because the UWB experiment requires living human to be a test subject [11]. To overcome this problem, model and simulation tools have been conducted. However, the models and tools are expensive and thereby limited number of model can be found. By 2013, only two research centers, such as Nagoya Institute of Technology in Japan [90] and Intervention Centre in Oslo University Hospital [91], are able to develop the models and tools. Instead of the high cost on models and tools, due to the prohibition of conducting test on living human, the above researches encounter the difficulty to validate the model results [11].

Currently, impulsive radio and MB-OFDM are the UWB technologies that have been proposed for implant sensor communication and the technologies are described as below.

2.6.2.1 MB-OFDM

MB-OFDM use OFDM technology to perform signal transmission in UWB. In this method, the available frequency bandwidth is divided into many orthogonal overlapping sub-bands and each of the sub-bands is carried by a sub-carrier. Because of the available bandwidth has been fully utilized, it has the highest bandwidth efficiency and highest overall data rate among all the mentioned technologies. Obviously, MB-OFDM is a multi-carrier technology because many sub-carriers can be found. Therefore, it encounters some multi-carrier issues and the issues are described as below.

- **Complex hardware and expensive cost are required.** In UWB MB-OFDM, intermediate frequency (IF) conversion is required to convert the carriers from baseband to the RF and vise versa. Therefore, the radio frequency hardware to perform the up and down conversions are required on transmitter and receiver, respectively. Instead of that, the high value in peak-to-average power ratio (PAPR), which is due to large number of carriers, can cause non-linear amplification and then destroys the orthogonality of OFDM signal. As a result, ISI can be introduced and thereby strongly degrade the performance. To overcome this issue, large linear range amplifier, which is expensive, is required [92, 93].

- **Sensitive to carrier frequencies offset.** In practice, a minor mismatch on the oscillators between the transmitter and the receiver causes carrier frequency offset in frequency domain, thereby the receiver cannot precisely sample at the sub-carrier frequencies. As a consequence, due to the receiver samples at the incorrect sub-carrier frequencies, the receiver suffers inter-carrier interference and then degrades the overall performance. To overcome this issue, some costly signal processing method is required [94, 95, 96, 97, 98].

2.6.2.2 Impulse Radio (IR)

To overcome the above issues, IR, which is a carrier-free technology, can be used. IR uses very short Gaussian pulses to perform data transmission for short distance communication. Ideally, the time length of the pulse is in a fraction of nanosecond and this implies that the information spectrum has been spread to wider spectrum. In this condition, the entire allocated UWB carries the same information, and thus this technology does not require the receiver to precisely demodulate the signal at certain carrier frequency. Therefore, in contrast to MB-OFDM, this technology is insensitive to carrier frequency offset. Furthermore, carrier-free technology of IR does not cause high PAPR value and thus high quality amplify is not required. Unlike MB-OFDM, direct baseband-to-RF or vise versa conversion can be used in IR. Therefore, the cost for IF conversion can be saved [8].

In spite of that, the IR technology has some weaknesses. First, high speed digital signal

processing at the receiver is required to detect the transmitting fraction-nanoseconds pulse length. Therefore, high sampling rate of analog to digital convertor (ADC) is required [99]. Second, in order to improve the overall performance, a correlation receiver with precise timing synchronization algorithm is required. However, the timing synchronization algorithm is still an opening challenge [100].

2.6.3 Human Body Communication

Instead of RF channel, data transfer can be done over an inductive link inside human body. In order to establish this link, an implant equipped with a small coil is placed inside body, and an external device with a large coil is located on the body surface. The inductive coupling between this two coils can form a below 30 MHz electromagnetic inductive loop that can be used to transfer data. [101, 102, 103, 88, 76]. This type communication provides higher data security than RF technology because the signal has not been broadcasted to the air. Moreover, due to the non-sharing human body communication channel, the signal does not suffer interference from other sources.

In spite of that, this method has some limitations. First, the setup method is not user-friendly for patient and doctor because the outside coil is required to be accurately positioned over the implant [104]. Second, the link only can provide a very short distance transmission range, roughly below 100 centimeter, whilst RF technology can provide 1 to 2 meter range [16]. Third, this communication probably can raise ethical and juridical challenge. The reason is the current standard does not define human body communication for implant sensor, but only defines human body communication for wearable sensor [10]. Therefore, this implant experiment or application on human body is prohibited. Model and simulation are allowed but the result is hard to be validated.

2.6.4 Summary

In summary, among the mentioned limitations, ethical and juridical challenge is considered as a serious challenge because this challenge cannot be overcome by high cost. Therefore, to avoid this

challenge, it is possible to improve the implant communication on the conventional MICS standard instead of UWB and Human Body Communication.

2.7 Interferences and MIMO channel equalization

In order to avoid signal distortion, the conventional narrow band MICS imposes that the information bandwidth is always lower than the allocated channel bandwidth of 300 kHz. However, this limited bandwidth can limit the data rate as well. As a consequence, a prolonged data transmission time is required and thereby it rapidly reduces the battery power. Furthermore, the limited data rate prohibits the capsule endoscopy to transmit high resolution images. Therefore, high data rate in the narrow band is demanded.

In order to achieve higher data rate, the information bandwidth should be increased in the narrow band channel. However, this increment can result on the case that the information bandwidth is larger than the allocation channel bandwidth. In this situation, due to the channel bandwidth imposed by the authority, the high frequency components of information bandwidth must be removed by a bandlimited filter, and thus the transmitting signal always has the bandwidth lower than the allocated channel bandwidth. In this scenario, the information is equivalently passed through a frequency selective channel. Because of the high frequency components of information are discarded, the current information symbol is inevitably combined with the subsequent symbols, therefore the receiver obtains the signal which has been corrupted by ISI. In order to retrieve the symbol from the ISI signal, equalizer is conventionally applied on the receiver. However, the above situation only implies that case of one transmitter and one receiver. In practice, one receiver can obtain the signals from multiple receivers. For example, a multi-antennas external device can receive signals from multiple transmitting implants in the same operating frequency. Therefore, the receiver unavoidably obtains the signals that has been corrupted by CCI due to multiple transmitters. In overall, due to multiple transmitters and frequency selective channel, the receiver obtains the combined signal that is corrupted by ISI and CCI simultaneously. The presence of ISI and

CCI at the receiver can strongly degrade the performance of the communication link. Therefore, these interferences must be overcome.

To cope with ISI and CCI, MIMO equalizer can be applied at the receiver. In general, MIMO equalization can be achieved through trained approach and blind approach. In the trained approach, the receiver requires the transmitters periodically sends out pilot sequences in order to open the channel eye. Least Mean Square (LMS) is an example of the trained approach. Normally, trained equalization can rapidly suppress interferences. However, the periodical transmission of the pilot sequence will reduce the data throughput, and thus blind approach is developed. Without any pilot sequence, the blind approach exploits the statistical information of transmitted signals to recover signal from the interferences. Therefore, blind approach is helpful to increase the data throughput especially for the expensive implant communication link.

2.8 System model and Assumptions

Multiple transmitted sources and multiple received antennas are considered here. In this condition, the received antenna receives the superposition of several source signals. Therefore the MIMO receiver has the burden to mitigate both ISI and CCI simultaneously. Such a scenario can be modeled as a *convolutive* MIMO system as shown in Fig. 2.1, rather than the simpler *instantaneous* MIMO system, where the channel impulse response is one tap coefficient rather than a vector coefficient. In other words, the difference in the two systems lie in the channel impulse response of the former having memory (and hence a vector), whilst the latter is memoryless (just a complex number). In general, the system model has M_{t} transmit sources and M_{r} received antennas. Furthermore, $M_{\mathrm{r}} \geq M_{\mathrm{t}}$ is assumed in this thesis. This assumption means we have more received antennas than sources, and therefore a unique inverse system possibly exists. Therefore zero forcing equalization is possible in this *overdetermined* MIMO system. The opposite, $M_{\mathrm{r}} < M_{\mathrm{t}}$ is called *underdetermined* MIMO system which is another complicated research topic and will not been pursued in this thesis.

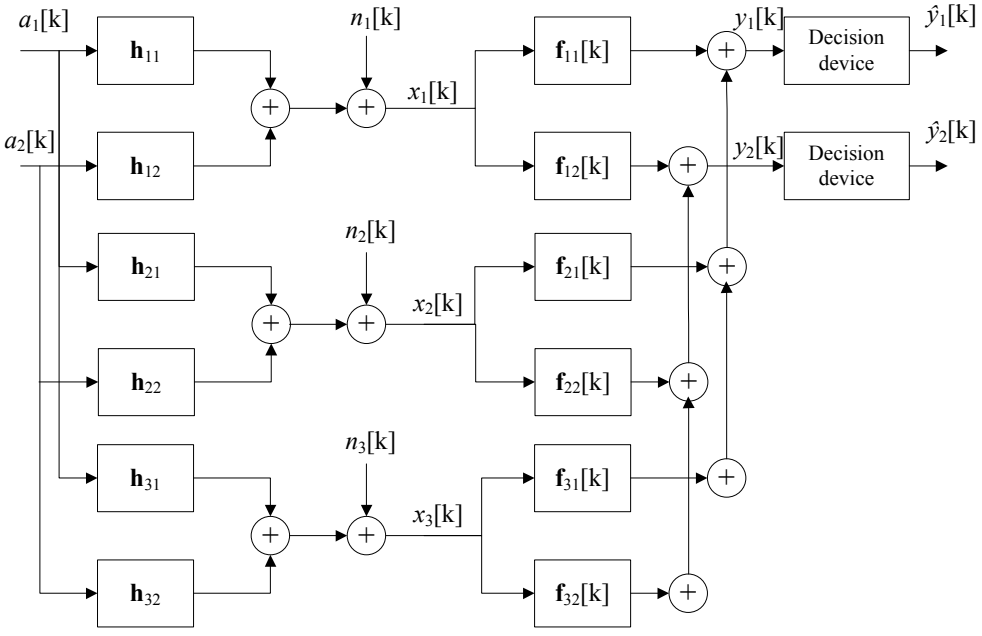

Figure 2.1: Baseband equivalent system for $M_t = 2$ and $M_r = 3$.

The *underdetermined* MIMO system is complicated because its inverse system is ambiguous from a mathematical perspective [105]. Finally, we highlight that the trivial case of $M_{\mathrm{r}} = 1$ and $M_{\mathrm{t}} = 1$, the system model is reduced to a convolutive SISO system model.

We further make the following assumptions, which are common:

A1 Assumption on sources:

The j-th source, $a_j[k]$, where $j \in (1, M_{\mathrm{t}})$, is an independently and identically distributed (i.i.d.) zero mean discrete time sequence. All sources, $a_j[k]$, $j \in (1, M_{\mathrm{t}})$ are uniformly selected from a PAM or QAM alphabet set, $\mathbf{A}$ and therefore $a_j[k]$ has zero mean, $\mathrm{E}\{a_j[k]\} = 0$, a finite power, $\sigma_{\mathrm{A}}^2 = \mathrm{E}\{|a_j[k]|^2\} > 0$ and a finite fourth order moment, $m_4 = \mathrm{E}\{|a_j[k]|^4\} > 0$ where $\mathrm{E}\{\cdot\}$ denotes statistical expectation. Since $a_j[k]$ is uniformly picked from a PAM or QAM alphabet set, $a_j[k]$ is circular and sub-Gaussian. Therefore, its normalized kurtosis, $\mathrm{Kur} \triangleq \frac{m_4}{\sigma_{\mathrm{A}}^4}$ must satisfy $\mathrm{Kur} < 3$ for real-valued case [2] and $\mathrm{Kur} < 2$ for complex-valued case [36]. In addition, $a_j[k]$ and $a_l[k+\delta]$, where $\delta \in \mathbb{Z}$ is any integer, $l \neq j$ and $1 \leq l \leq M_{\mathrm{t}}$, are mutually independent so that

$$\mathrm{E}\{a_j[k]a_l^*[k+\delta]\} = \mathrm{E}\{a_j[k]\}\mathrm{E}\{a_l^*[k+\delta]\} = 0 \tag{2.1}$$

where * denotes complex conjugate.

A2 Assumption on channels:

We model the channels from j-th input to i-th channel output as static finite impulse response (FIR) vectors and denote them as $\mathbf{h}_{ij} = [h_{ij}[0], h_{ij}[1], \cdots, h_{ij}[N_h - 1]]^T$, where $i \in (1, M_{\mathrm{r}})$, $j \in (1, M_{\mathrm{t}})$, and all channels have the same length, N_h. The properties and descriptions of channel are stated below. First, the channel is modeled as FIR vector because we focus on a stable bounded-input bounded-output channel which possibly allows a stable inverse system to be exist. Second, the channel is modeled as N_h length FIR vectors because we assume the channel has memory due to multipath propagation. This type of channel is called frequency selective channel which causes ISI. Third, the channels, which is FIR vectors, can

be rearranged and combined into a channel convolution matrix which is assumed to satisfy left invertibility condition. The left invertibility of the channel convolution matrix is the most important criteria to allow the existence of its inverse system. The conditions of the invertible channel convolution matrix are formulated below. The channel convolution matrix, $\mathbf{H}$ which has the dimension $M_{\mathrm{r}}L \times M_{\mathrm{t}}N_s$ where $N_s = L + N_h - 1$ is the convoluted length, is defined as below

$$\mathbf{H} = \begin{bmatrix} \ddot{\mathbf{H}}_{11} & \ddot{\mathbf{H}}_{12} & \dots & \ddot{\mathbf{H}}_{1M_{\mathrm{t}}} \\ \ddot{\mathbf{H}}_{21} & \cdots & \cdots & \ddot{\mathbf{H}}_{2M_{\mathrm{t}}} \\ \vdots & \cdots & \cdots & \vdots \\ \ddot{\mathbf{H}}_{M_{\mathrm{r}}1} & \cdots & \cdots & \ddot{\mathbf{H}}_{M_{\mathrm{r}}M_{\mathrm{t}}} \end{bmatrix}$$

where $\ddot{\mathbf{H}}_{ij}$ is a $L \times N_s$ toeplitz matrix which is formulated as

$$\ddot{\mathbf{H}}_{ij} = \begin{bmatrix} h_{ij}[0] & h_{ij}[1] & \cdots & h_{ij}[N_h-1] & 0 & 0 & \dots & 0 \\ 0 & h_{ij}[0] & h_{ij}[1] & \cdots & h_{ij}[N_h-1] & 0 & \dots & 0 \\ 0 & 0 & h_{ij}[0] & h_{ij}[1] & \cdots & h_{ij}[N_h-1] & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & \dots & 0 & h_{ij}[0] & \cdots & h_{ij}[N_h-1] \end{bmatrix}.$$

We assume the channel convolution matrix, $\mathbf{H}$ is full column rank matrix and therefore $\mathbf{H}$ is left invertible which also implies that its inverse system or zero forcing solution always exists.

A3 Assumption on sensor noise(s):

Some noises which are introduced by electronic sensors are modeled as additive white Gaussian noise(AWGN). In this condition, $n_i[k]$ for $i \in (1, M_{\mathrm{t}})$ is defined as the zero mean complex value additive white Gaussian noise with a constant variance, $\mathrm{E}\{|n_i[k]|^2\} = \sigma_{\mathrm{n}}^2$ at the i-th channel output. Furthermore, $n_i[k]$ is independent from all sources, $a_j[k]$.

Assumptions above are frequently encountered in the related literatures such as [106, 36, 107, 2, 51].

The observed signal at the i-th sensor, for $i \in (1, M_{\mathrm{r}})$, at time k is defined as:

$$x_i[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij} \otimes \mathbf{a}_j[k] + n_i[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij}^T \mathbf{a}_j[k] + n_i[k] \tag{2.2}$$

where $\mathbf{a}_j[k] = \big[a_j[k], \cdots, a_j[k - N_h + 1]\big]^T$, $\otimes$ and ^{T} denote discrete time convolution and vector transposition, respectively. The equalizer output corresponding to the j-th source, where $j \in$

$(1, M_{\mathrm{t}})$, at time k is denoted as:

$$y_j[k] = \sum_{i=1}^{M_{\mathrm{r}}} \mathbf{f}_{ij}^T[k]\mathbf{x}_i[k] \tag{2.3}$$

where $\mathbf{x}_i[k] = [x_i[k], \cdots, x_i[k-L+1]]^T$ is the input regressor vector for the equalizers of the i-th seonsor (there are a total of M_{t} equalizers associated with each of the i sensors), L is sufficiently large to eliminate the ISI of the channel, and $\mathbf{f}_{ij}[k] = \left[f_{ij}^{[0]}[k], \cdots, f_{ij}^{[L-1]}[k]\right]^T$ is a $L \times 1$ vector that represents the equalizer connecting the observed signal at i-th sensor to the j-th equalizer output at time k. There are a total of $M_{\mathrm{t}} \cdot M_{\mathrm{r}}$ equalizers at the receiver end (see Fig. 2.1). Furthermore, Eq. 2.3 can be written as

$$y_j[k] = \mathbf{F}_j^T[k]\mathbf{X}[k] \tag{2.4}$$

where $\mathbf{F}_j[k] = \left[\mathbf{f}_{1j}^T[k], \cdots, \mathbf{f}_{M_{\mathrm{r}}j}^T[k]\right]^T$ is a $M_{\mathrm{r}}L \times 1$ vector and $\mathbf{X}[k] = \left[\mathbf{x}_1^T[k], \cdots, \mathbf{x}_{M_{\mathrm{r}}}^T[k]\right]^T$ is a $M_{\mathrm{r}}L \times 1$ vector.

Then, $y_j[k]$ is fed into decision device and the decision device generates $\hat{y}_j[k]$. $\hat{y}_j[k]$ is chosen based on $y_j[k]$ and the alphabet set, $\mathbf{A}$. A symbol in $\mathbf{A}$ which has the smallest Euclidean distance to $y_j[k]$ is chosen as $\hat{y}_j[k]$. For example in 4-QAM, the relationship between $y_j[k]$ and $\hat{y}_j[k]$ is mathematically expressed as

$$\mathrm{Re}\{\hat{y}_j[k]\} = \begin{cases} -1 & \text{if } \mathrm{Re}\{y_j[k]\} \leq 0 \\ 1 & \text{otherwise} \end{cases} \tag{2.5}$$

and

$$\mathrm{Im}\{\hat{y}_j[k]\} = \begin{cases} -1 & \text{if } \mathrm{Im}\{y_j[k]\} \leq 0 \\ 1 & \text{otherwise} \end{cases} \tag{2.6}$$

Another example for 16-QAM, the mathematic relationship between $y_j[k]$ and $\hat{y}_j[k]$ is

$$\mathrm{Re}\{\hat{y}_j[k]\} = \begin{cases} -3 & \text{if } \mathrm{Re}\{y_j[k]\} \leq -2 \\ -1 & \text{if } -2 < \mathrm{Re}\{y_j[k]\} \leq 0 \\ 1 & \text{if } 0 < \mathrm{Re}\{y_j[k]\} \leq 2 \\ 3 & \text{otherwise} \end{cases} \tag{2.7}$$

and

$$\mathrm{Im}\{\hat{y}_j[k]\} = \begin{cases} -3 & \text{if } \mathrm{Im}\{y_j[k]\} \leq -2 \\ -1 & \text{if } -2 < \mathrm{Im}\{y_j[k]\} \leq 0 \\ 1 & \text{if } 0 < \mathrm{Im}\{y_j[k]\} \leq 2 \\ 3 & \text{otherwise} \end{cases} \tag{2.8}$$

2.9 Definition of a Good MIMO Blind Equalizer

We are going to mathematically describe a good MIMO blind equalizer that not only suppresses ISI and CCI but also retrieves all symbol sequences without repetition in the steady state [36]. First, we formulate a global system which combines channels and equalizers into a single system so that it directly connects input sources to outputs. The pattern of impulse responses in the global system is the key to determine a MIMO blind equalizer. In short, the suppression of ISI and CCI doesn't mean that it has achieved all its objectives, because the same source may be extracted twice (or more) in the receiver outputs while missing some, which in this case is bad. This activity is often known as source separation. In what follows we will show in the ideal case (i.e. equalization and source separation achieved), the global transfer function reduces to a special type of identity matrix, where the diagonal elements, which are polynomial functions, have only one non-zero coefficient. While this is achievable with trained systems (i.e. a known training

sequence is transmitted along the message signal), blind equalizers can only achieve such a matrix that is possibly permutated and/or phase rotated.

A good MIMO blind equalizer is one that successfully suppress all ISI, CCI and retrieving all sources once. In order to show this mathematically, we will pursue it in the Z-domain instead of time domain because time domain convolution is much more complex and tedious. In the end, we will get complete source separation and interferences suppression but subject to an ambiguous permutation and possible phase offset, as seen in the final Eq. 2.16.

To pursue the systems in Z-domain, let us define Unilateral Z-transform of some FIR systems, such as Unilateral Z-transform of channel $\mathbf{h}_{ij}$ as

$$h_{ij}(z) = \sum_{p=0}^{N_h-1} h_{ij}[p]z^{-p} \tag{2.9}$$

and Unilateral Z-transform of equalizer $\mathbf{f}_{ij}[k]$ as

$$f_{ij}(z) = \sum_{q=0}^{L-1} f_{ij}^{[q]}[k]z^{-q} \tag{2.10}$$

where $f_{ij}(z)$ is assumed has reached the steady state and therefore, it is time k independent. It is obvious that $h_{ij}(z)$ and $f_{ij}(z)$ are also polynomial functions which has variable z^{-1}. We further define $a_j(z)$ and $y_j(z)$ are Z-transforms of source $a_j[k]$ and output $y_j[k]$ respectively.

For simplicity, we assume noiseless condition (i.e. $\sigma_n^2 = 0$). Without loss of generality, we firstly assume $M_\mathrm{t} = 2$ transmit sources and $M_\mathrm{r} = 3$ receive sensors. This condition allows us to show the linear mathematical manipulations of the global system easily. After that, the overall system will be generalized into any MIMO system and will be described in matrix equation.

With the $M_\mathrm{t} = 2$ and $M_\mathrm{r} = 3$ assumption, the system model in Fig. 2.1 can be decomposed and split into equalizer 1 and equalizer 2. Fig. 2.2 illustrates the way to form the global transfer function $S_{(i),j}(z)$ from MIMO channel $h_{ij}(z)$ and FIR transfer function $f_{ij}(z)$. Fig. 2.2 (i) is the portion of the model 2.1 that illustrates the relationship between sources (i.e. $a_1(z)$ and $a_2(z)$) and equalizer output, 1 $y_1(z)$. The sources transmit signals through MIMO channel which is represented by MIMO channel transfer function $h_{ij}(z)$ for $i = 1, 2, 3$ and $j = 1, 2$ where $h_{ij}(z)$

represents the channel transfer function that connects source i to the received antenna j. After that, the MIMO equalizer 1, which is represented by FIR transfer functions $f_{11}(z)$, $f_{21}(z)$ and $f_{31}(z)$, receives the signal through MIMO channel. $f_{11}(z)$, $f_{21}(z)$ and $f_{31}(z)$ are the FIR transfer functions that connect received antenna 1, 2 and 3 that belongs to equalizer 1, respectively. The outputs of the FIR transfer functions are summed up to be equalizer output 1, $y_1(z)$. After the sources and the output are clearly identified, the MIMO transfer function and FIR transfer functions can be combined into global transfer functions. Fig. 2.2 (ii) illustrates the global transfer functions $S_{(1),1}(z)$ and $S_{(2),1}(z)$ that are the results of combining MIMO channel transfer functions and FIR transfer functions, while the sources and output are remained to be unchanged. The correspond equations of the global transfer functions will be shown later. Similarly, the above descriptions of Fig. 2.2 (i) and (ii) can be re-used to describe Fig. 2.2 (iii) and (iv) for equalizer 2, instead of equalizer 1.

From mathematically perspective, $y_1(z)$ in Fig. 2.2 (i) can be established as

$$\begin{aligned} y_1(z) &= \overbrace{\Big(f_{11}(z)h_{11}(z) + f_{21}(z)h_{21}(z) + f_{31}(z)h_{31}(z)\Big)}^{=S_{(1),1}(z)} a_1(z) \\ &\quad + \underbrace{\Big(f_{11}(z)h_{21}(z) + f_{21}(z)h_{22}(z) + f_{31}(z)h_{32}(z)\Big)}_{=S_{(2),1}(z)} a_2(z) \\ &= S_{(1),1}(z)a_1(z) + S_{(2),1}(z)a_2(z) \end{aligned} \tag{2.11}$$

where $S_{(i),j}(z)$ is the Z domain global system or is called global transfer function that directly connects input source i to output j and it is generally defined as

$$S_{(i),j}(z) = \sum_{p=1}^{M_r} f_{pj}(z)h_{pi}(z) = \sum_{r=0}^{N_s-1} S_{(i),j}^{[r]} z^{-r} \tag{2.12}$$

and due to the reason that $f_{pj}(z)$ and $h_{pi}(z)$ are polynomial functions, the final result of $S_{(i),j}(z)$ is also a polynomial function with coefficients $S_{(i),j}^{[r]}$ and the highest order N_s-1 where $N_s = L+N_h-1$ is the convolution length. After that, we re-apply the same manipulations to output 2, $y_2(z)$, we

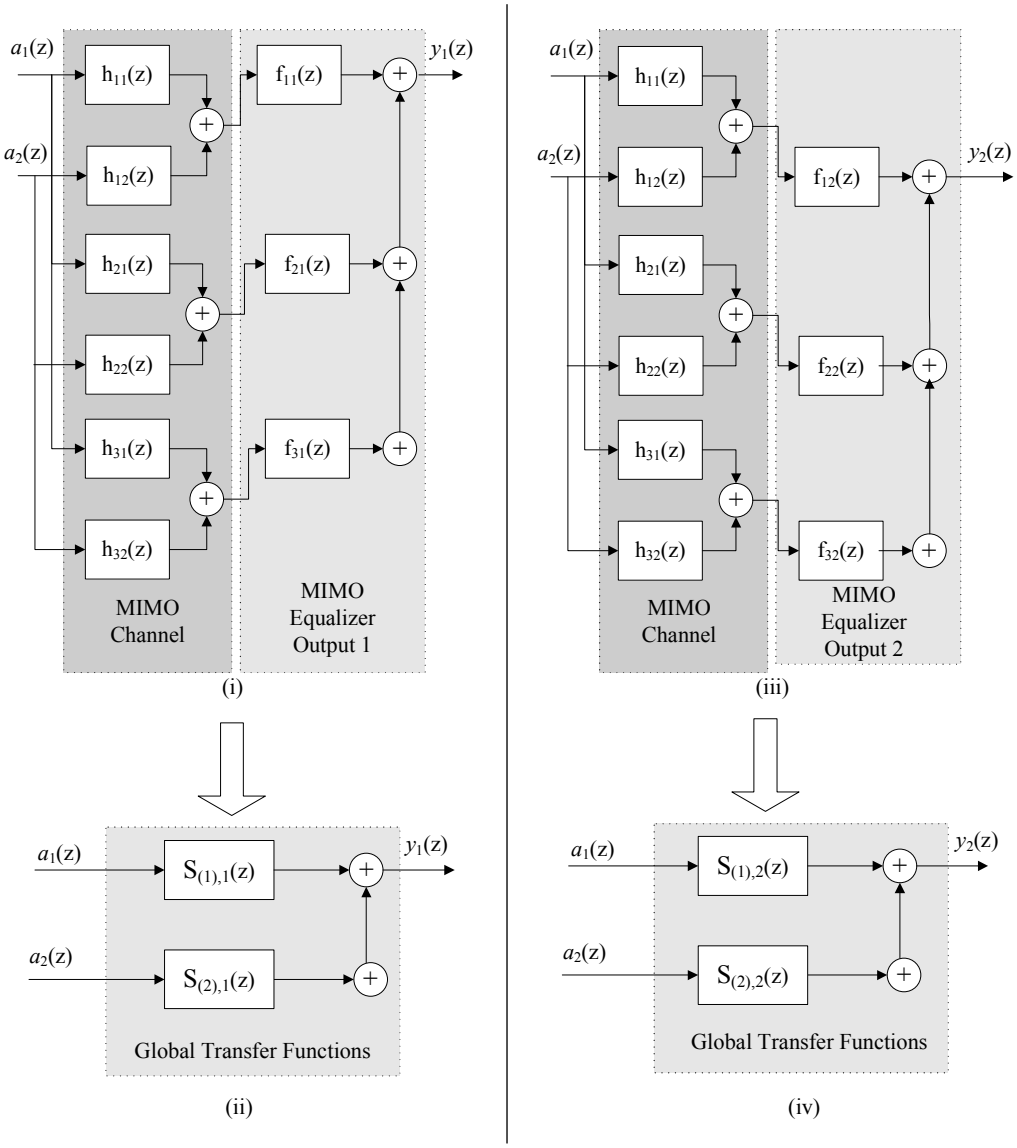

Figure 2.2: MIMO system

eventually obtain

$$\begin{aligned} y_2(z) &= \overbrace{(f_{12}(z)h_{11}(z) + f_{22}(z)h_{21}(z) + f_{32}(z)h_{31}(z)\ )}^{=S_{(1),2}(z)} a_1(z) \\ &\quad + \underbrace{(f_{12}(z)h_{21}(z) + f_{22}(z)h_{22}(z) + f_{32}(z)h_{32}(z)\ )}_{=S_{(2),2}(z)} a_2(z) \\ &= S_{(1),2}(z)a_1(z) + S_{(2),2}(z)a_2(z) \end{aligned} \tag{2.13}$$

. The role of global transfer functions $S_{(i),j}(z)$ are illustrated in Fig. 2.2 (ii) and (iv) for output 1 and 2 respectively. It is worth to mention that Equations 2.11 and 2.13 are linear functions and therefore they can be combined into a matrix equation in which the role of equalizer $f_{ij}(z)$ can be easily identified in the matrix equation.

In order to generalize the condition above to M_{t} transmit sources and M_{r} receive sensors, we can form matrix equation below,

$$\mathbf{y}(z) = \mathbf{F}^{\mathrm{T}}(z)\mathbf{H}(z)\mathbf{a}(z) = \mathbf{S}^{\mathbf{T}}(z)\mathbf{a}(z) \tag{2.14}$$

where $()^{\mathrm{T}}$ denotes matrix transpose, $\mathbf{F}(z)$ is a $M_r \times M_t$ polynomial matrix that describes equalizer transfer functions and is defined as

$$\mathbf{F}(z) = \begin{bmatrix} f_{11}(z) & f_{12}(z) & \cdots & f_{1M_t}(z) \\ f_{21}(z) & f_{22}(z) & \vdots & f_{2M_t}(z) \\ \vdots & \vdots & \vdots & \vdots \\ f_{M_r1}(z) & \cdots & \cdots & f_{M_rM_t}(z) \end{bmatrix},$$

$\mathbf{H}(z)$ is a $M_r \times M_t$ polynomial matrix that describes channel transfer functions and is defined as

$$\mathbf{H}(z) = \begin{bmatrix} h_{11}(z) & h_{12}(z) & \cdots & h_{1M_t}(z) \\ h_{21}(z) & h_{22}(z) & \vdots & h_{2M_t}(z) \\ \vdots & \vdots & \vdots & \vdots \\ h_{M_r1}(z) & \cdots & \cdots & h_{M_rM_t}(z) \end{bmatrix}.$$

Furthermore, $\mathbf{S}(z)$ is a $M_t \times M_t$ polynomial matrix. It describes global transfer functions which

combines channel and equalizer transfer functions and is defined as

$$\mathbf{S}(z) = \begin{bmatrix} S_{(1),1}(z) & S_{(1),2}(z) & \cdots & S_{(1),M_t}(z) \\ S_{(2),1}(z) & S_{(2),2}(z) & \vdots & S_{(2),M_t}(z) \\ \vdots & \vdots & \vdots & \vdots \\ S_{(M_t),1}(z) & \cdots & \cdots & S_{(M_t),M_t}(z) \end{bmatrix}.$$

Finally, $\mathbf{a}(z) = [a_1(z),\ a_2(z),\ \ldots\ ,\ a_{M_t}(z)]^T$ and $\mathbf{y}(z) = [y_1(z),\ y_2(z),\ \ldots\ ,\ y_{M_t}(z)]^T$ are $M_t \times 1$ polynomial matrices which represent Z-transform of input sources and outputs respectively.

Ideally, to recover all source sequences without any interferences, the equalizer matrix $\mathbf{F}(z)$ in Eq. 2.14 is chosen, so that the global matrix, $\mathbf{S}(z)$ should satisfy,

$$\mathbf{S}(z) = \mathbf{H}^{\mathrm{T}}(z)\mathbf{F}(z) = \mathrm{diag}(z^{-\mathrm{d}_1}, z^{-\mathrm{d}_2}, \ldots, z^{-\mathrm{d}_{\mathrm{M_t}}}) \tag{2.15}$$

where $\mathrm{diag}(z^{-\mathrm{d}_1}, z^{-\mathrm{d}_2}, \ldots, z^{-\mathrm{d}_{\mathrm{M_t}}})$ denotes a $M_t \times M_t$ diagonal matrix and $d_1,\ d_2, \ldots, d_{M_t} \in \{0, 1, \ldots, N_s - 1\}$. Eq. 2.15 implies that the output j may retrieves the d_j time delay version of symbol sequence from source j where $j = 1, \ldots, M_t$. However, the ideal solution in Eq. 2.15 cannot be achieved in MIMO blind equalization since the exact source sequences and the position j of source are unknown by the receiver. Therefore, output j does not always retrieve input source j and this weakness is called permutation ambiguity. Instead of that, due to the statistical circular symmetry condition on input sources, MIMO blind equalization cannot identify the phase of sources and this problem is called phase ambiguity. Hence, the best possible solution that is provided by MIMO blind equalization is

$$\mathbf{S}(z) = \mathbf{H}^{\mathrm{T}}(z)\mathbf{F}(z) = \mathrm{diag}(e^{j\theta_{d_1}} z^{-\mathrm{d}_1}, e^{j\theta_{d_2}} z^{-\mathrm{d}_2}, \ldots, e^{j\theta_{d_{M_t}}} z^{-\mathrm{d}_{\mathrm{M_t}}})\mathbf{P} \tag{2.16}$$

where $\mathbf{P}$ is a permutation matrix and $e^{\mathbf{j}\theta_{d_n}}$ denotes any possible phase offset with $-\pi \leq \theta_{d_n} < \pi$, $\mathbf{j} = \sqrt{-1}$ and $n = 1, \ldots, M_t$. Thus, a good MIMO blind equalizer should satisfy Eq. 2.16 which retrieves all source sequences and suppresses interferences but includes permutation ambiguity and phase ambiguity.

2.10 Time Domain Identification of Good MIMO Equalizer (with examples)

The Z domain matrix form in previous section may be difficult to visualize. Therefore we will provide the equivalent presentation in the time-domain vector form. We will present the absolute global impulse responses which are formed by the coefficient magnitudes of global transfer functions. Furthermore, the absolute global impulse responses are widely been used as the tool to identify a good equalizer in the steady state [36]. This section is structured as follows. First, we describe the global impulse response which is the time domain counterpart of global transfer function. Then, we present the absolute global impulse responses that are magnitude version of global impulse response and does not depend on phase.

The global impulse response, $\mathbf{S}_{(i),j}[k]$ that connects input source i with output j is the inverse Z transform of the global transfer function $S_{(i),j}(z)$ in which $\mathbf{S}_{(i),j}[k]$ is a $N_s \times 1$ vector and is defined as

$$\mathbf{S}_{(i),j}[k] = \left[S_{(i),j}^{[0]}, S_{(i),j}^{[1]}, \ldots, S_{(i),j}^{[N_s-1]} \right]^T. \tag{2.17}$$

It is obvious to see that the $r+1$-th element in $\mathbf{S}_{(i),j}[k]$ above is the coefficient $S_{(i),j}^{[r]}$ that belongs to variable z^{-r} in $S_{(i),j}(z)$ (see Eq. 2.12). Therefore, the global impulse response $\mathbf{S}_{(i),j}[k]$ is a straight forward transformation from the global transfer function $S_{(i),j}(z)$. Due to phase ambiguity issue in blind equalizer, the result of coefficient $S_{(i),j}^{[0]}$ is subject to phase ambiguity as well. For simplicity, we observe the magnitude of coefficient $S_{(i),j}^{[0]}$, $|S_{(i),j}^{[0]}|$ which does not depend on phase. In this condition, the absolute global impulse response, $|\mathbf{S}_{(i),j}[k]|$ where

$$|\mathbf{S}_{(i),j}[k]| = \left[|S_{(i),j}^{[0]}|, |S_{(i),j}^{[1]}|, \ldots, |S_{(i),j}^{[N_s-1]}| \right]^T \tag{2.18}$$

which only computes coefficient magnitude.

The success of a MIMO equalizer can be visually determined by observing the coefficients of $|\mathbf{S}_{(i),j}[k]|$ in the steady state. In particular, we look for source retrieval of the intended source and

source suppression for the neighbor sources. To demonstrate how we judge the success of a MIMO equalizer, we provide the following 3 examples which utilizes a simple $M_t = 2$ MIMO equalizer.

2.10.0.1 Example 1: Good MIMO Equalizer

A good MIMO equalizer would generate an impulse for $|\mathbf{S}_{(1),1}[k]|$ and $|\mathbf{S}_{(2),2}[k]|$, as well as an all-zero vector for $|\mathbf{S}_{(2),1}[k]|$ and $|\mathbf{S}_{(1),2}[k]|$ at steady state. For example this set of $|\mathbf{S}_{(i),j}[k|]$ vectors implies our objectives of equalization and source separation are met: $|\mathbf{S}_{(1),1}[k]| = [0, 0, 1, 0]$, $|\mathbf{S}_{(2),1}[k]| = [0, 0, 0, 0]$, $|\mathbf{S}_{(1),2}[k]| = [0, 0, 0, 0]$ and $|\mathbf{S}_{(2),2}[k]| = [0, 0, 0, 1]$.

Let us firstly consider $\mathbf{S}_{(1),1}[k]$ and $\mathbf{S}_{(2),1}[k]$. The single impulse in the 3rd coefficient in $\mathbf{S}_{(1),1}[k]$ implies perfect equalization of source 1 but delayed by 2 symbol periods, i.e. the retrieval of $|a_1[k-2]|$. It also means source 2 is completely suppressed at the output of equalizer 1.

In a similar fashion, the all-zero $\mathbf{S}_{(1),2}[k]$ and the single impulse $\mathbf{S}_{(2),2}[k]$ implies the successful retrieval of source 2 and suppression of source 1 at the output equalizer 2.

Therefore, $y_1[k]$ retrieves the symbol sequence from source 1 alone and $y_2[k]$ from source 2 alone, albeit being delayed, which does not affect the bit error rate performance. We consider this as a completely successful MIMO equalization.

2.10.0.2 Example 2: Good MIMO Equalizer with source permutation

We consider another good case where $|\mathbf{S}_{(1),1}[k]| = [0, 0, 0, 0]$, $|\mathbf{S}_{(2),1}[k]| = [0, 1, 0, 0]$, $|\mathbf{S}_{(1),2}[k]| = [0, 0, 1, 0]$ and $|\mathbf{S}_{(2),2}[k]| = [0, 0, 0, 0]$ in the steady state. Using the same explanation in Example 1, here $y_1[k]$ retrieves source 2 while $y_2[k]$ retrieves source 1. Therefore source permutation is present, but this can be easily resolved and therefore this setting is considered successful.

2.10.0.3 Example 3: MIMO Equalizer with source separation failed

We consider an unsuccessful case where $|\mathbf{S}_{(1),1}[k]| = [0, 1, 0, 0]$, $|\mathbf{S}_{(2),1}[k]| = [0, 0, 0, 0]$, $|\mathbf{S}_{(1),2}[k]| = [0, 0, 1, 0]$ and $|\mathbf{S}_{(2),2}[k]| = [0, 0, 0, 0]$ in the steady state.

Here, it can be seen that both equalizers retrieve the same source, i.e. source 1, thus missing source 2 entirely. This is an example of retrieving the same source repeatedly and missing at least one other source. Although the ISI has been successfully mitigated, we do not consider this as a successful MIMO equalizer.

2.11 The way of adaptation in MIMO equalizer

We briefly describe the way of a MIMO equalizer adapts in this section. We recall Eq. 2.16 and then we define a good MIMO equalizer to have the optimum transfer function $\mathbf{F}_{\text{opt}}(z)$. Therefore, $\mathbf{F}_{\text{opt}}(z)$ should satisfy

$$\mathbf{S}_{\text{opt}}(z) = \mathbf{H}^{\text{T}}(z)\mathbf{F}_{\text{opt}}(z) = \text{diag}(e^{j\theta_{d_1}} z^{-\text{d}_1}, e^{j\theta_{d_2}} z^{-\text{d}_2}, \ldots, e^{j\theta_{d_{M_t}}} z^{-\text{d}_{\text{M}_\text{t}}})\mathbf{P} \tag{2.19}$$

where $\mathbf{S}_{\text{opt}}(z)$ is the optimum global transfer function that indicates a good equalization has been achieved. Furthermore, we define the time domain counterpart of $\mathbf{F}_{\text{opt}}(z)$ as $\mathbf{F}_{\text{opt}}[k]$, which is the impulse responses of MIMO equalizer, and is defined as

$$\mathbf{F}_{\text{opt}}[k] = \left[\mathbf{F}^T_{1,\text{opt}}[k], \mathbf{F}^T_{2,\text{opt}}[k], \ldots, \mathbf{F}^T_{M_r,\text{opt}}[k]\right]^T \tag{2.20}$$

where $\mathbf{F}^T_{j,\text{opt}}[k]$ is the optimum equalizer for output j.

In order to achieve the above optimum condition in the steady state, it is obvious that the equalizer j, $\mathbf{F}_j[k]$ should be adapted by an algorithm towards its optimum $\mathbf{F}_{j,\text{opt}}[k]$, which is a static value. The above adaptation process can be done by an optimization algorithm which normally achieves optimum situation by minimizing a non-zero real function. In the literature, stochastic gradient descent algorithm (SGA) is widely been chosen as the optimization algorithm because of its simplicity in hardware implementation. Ideally, SGA searches for a minimum point of an user-defined non-negative function, which is called cost function. After SGA found a minimum point of the cost function, SGA will stay steadily on the minimum point. Therefore, researchers normally design a cost function that is a function of multivariate $\mathbf{F}_j[k]$ and locate the global minimum of the cost function on the multivariate optimum value $\mathbf{F}_{j,\text{opt}}[k]$.

In practice, $\mathbf{F}_j[k]$ is initialized on any value $\mathbf{F}_j[0]$ that corresponds to a non-minimum point in the cost function. After that, $\mathbf{F}_j[k]$ is started to be iteratively adapted by SGA towards $\mathbf{F}_{j,\text{opt}}[k]$. Eventually, when $\mathbf{F}_j[k] \simeq \mathbf{F}_{j,\text{opt}}[k]$ happened, the value $\mathbf{F}_j[k]$ is expected to be unchanged by SGA since the steady state is reached. This condition is called *global convergence* because the MIMO equalizer converges to the global minimum point and a good MIMO equalizer is achieved.

However, due to the reason that some cost functions have local minima, SGA possibly causes $\mathbf{F}_j[k]$ to stay steadily on the local minima instead of the global minimum if certain criterions have not been fulfilled. If so, this situation is known as *ill-convergence* or *local convergence* because the equalizer converges to a local minimum point instead of the global minimum point. This implies MIMO equalization is unsuccessful.

In the literature, many SISO and MIMO blind equalization algorithms are developed based on SGA . Generally, these algorithms can be categorized into robust algorithm and non-robust algorithm. Robust algorithm is a blind equalization algorithm that has *avoidable* local minima. This means equalizer possibly converges to the global minima if suitable parameters are chosen. On the other hand, non-robust algorithm is a blind equalization algorithm that has *unavoidable* local minima. Robust and non-robust algorithms will be reviewed in Section 2.13-2.15.

2.12 Performance Measurements

The absolute global impulse responses in the previous section are effective tools to identify a good MIMO equalizer that achieves ISI and CCI suppression. The performance of such a good MIMO equalizer can be measured using either the residual interference of the equalizer output j ($\text{IT}_j[k]$) or the mean square error (MSE) of output j ($\text{MSE}_j[k]$). They are commonly used performance measurements in the literature and will be adopted in this thesis. $\text{MSE}_j[k]$ measures the residual interference with the aid of decision device and it is suitable to be used in the presence of AWGN noise. On the other hand, $\text{IT}_j[k]$ evaluates the amount of residual interference based on global impulse responses without the presence of AWGN noise but it requires knowledge of the channel

which is often absent in real implementation (although in simulation environments we usually have accessed to the channel). In fact, a low value of $\mathrm{IT}_j[k]$ or $\mathrm{MSE}_j[k]$ is always desired as it reflects good equalizer quality. Instead of that, $\mathrm{IT}_j[k]$ is insensitive to the scale because it is normalized by the largest amplitude. In contrary, no normalized operation has been done on $\mathrm{MSE}_j[k]$, therefore $\mathrm{MSE}_j[k]$ can be used to trace the presence of scale. Here we emphasize that although $\mathrm{IT}_j[k]$ or $\mathrm{MSE}_j[k]$ may be low, source separation might not have been attained so the technique of absolute global impulse response (outlined in Section 2.4) is always needed to first determine source separation.

First, the amount of residual interference of the equalizer outputs for the j-th equalizer output at time instant k, $\mathrm{IT}_j[k]$ is defined as below [36]

$$\mathrm{IT}_j[k] = 10\log_{10}\left(\frac{\sum_{i=1}^{M_\mathrm{t}}\sum_{n=0}^{N_s-1}\left|s_{ij}^{[n]}[k]\right|^2 - \max_{i,n}\left|s_{ij}^{[n]}[k]\right|^2}{\max_{i,n}\left|s_{ij}^{[n]}[k]\right|^2}\right) \tag{2.21}$$

where $s_{ij}^{[n]}[k]$ is the $(n+1)$-th element of the global impulse response $\mathbf{S}_{(i),j}[k]$, $\mathbf{S}_{(i),j}[k]$ being the global impulse response that connects the i-th source, $a_i[k]$, to $y_j[k]$ and $\max_{i,n}\left|s_{ij}^{[n]}[k]\right|^2$ is a function that selects the largest value of $|s_{ij}^{[n]}[k]|^2$ in $0 \leq n \leq N_s - 1$ and $1 \leq i \leq M_\mathrm{t}$.

Next, MSE of the j-th equalizer output, $\mathrm{MSE}_j[k]$ is computed as following,

$$\mathrm{MSE}_j[k] = 0.99\mathrm{MSE}_j[k-1] + 0.01\left|\hat{y}_j[k] - y_j[k]\right|^2 \tag{2.22}$$

where $\hat{y}_j[k]$ denotes the output of decision device and $\mathrm{MSE}_j[0] = 1$ is initialized. Furthermore, MSE in decibel (dB) unit is defined as $\mathrm{MSE}_j[k]$ (dB) $= 10\log_{10}\mathrm{MSE}_j[k]$.

2.13 Review on SISO Blind Equalization Algorithms

Since 1960s, many blind equalization algorithms have been developed for SISO system to suppress ISI from a frequency selective channel. It is worth to mention that the development of MIMO blind equalization algorithm are normally inherent the concept or idea of SISO blind equalization algorithm. Therefore, it is important to review some famous SISO blind equalization algorithms such

as Decision Directed (DD) algorithm [108], Sato algorithm [22] and Gordard algorithm or Constant Modulus algorithm (CMA) [21] in this section. The main differences among these algorithms lie in their cost functions. In the following, we will briefly review how does the previous works categorize the above algorithms into robust or non-robust algorithms. After that, each algorithms and their uniqueness will be presented.

All blind equalization algorithms have undesirable local minima in its global response (i.e. the combination of channel and equalizer) cost surface if the FIR filter (i.e. Assumption A2) is used to model both the channel and the equalizer [109]. If the coefficients of the equalizer converge to these local minima, the ISI will not be suppressed successfully. Therefore, a blind algorithm that has the fewest distinct local minima is widely preferred since it reduces the risk of local convergence. It is worth to mention that Li et al [109, 110, 111] had analyzed some popular algorithms and found that local minima in CMA can be avoided if correct parameters are chosen at the initial stage. Based on this reason, we categorize CMA as a robust algorithm. Furthermore, the same analysis performed on the Sato algorithm and the DD algorithm show that they too have many local minima, so the local convergence problem is more likely encountered in implementation. Due to this reason, we categorize the Sato and the DD algorithms as non-robust algorithms, make the CMA (or more generally known as the Godard algorithm) the first even robust blind equalization algorithm that was developed.

Next, we explain the cost function of an algorithm, before reviewing its place in the stochastic gradient algorithm (SGA) which is widely used due to its simplicity.

Generally, blind equalization algorithms are derived based on a non-negative cost function which has a general form,

$$J_1[k] = \mathrm{E}\left\{|\varepsilon_1[k]|^2\right\} \tag{2.23}$$

where $J_1[k]$ is the cost function of equalizer output 1 and $\varepsilon_1[k]$ is called error function of equalizer output 1 which is different for each algorithms. However, $\mathrm{E}\left\{|\varepsilon_1[k]|^2\right\}$ in Eq. 2.23 is impossible to be directly determined in real implementation due to the E{.} operator. Therefore, the past

literature normally estimate E{.} using an instantaneous value. This estimation method does not introduce additional offset and therefore it is an unbiased estimation [112, 113, 114]. Therefore, the estimation of the cost function becomes

$$J_1[k] \triangleq |\varepsilon_1[k]|^2. \tag{2.24}$$

After that, stochastic gradient descent algorithm (SGA) is applied to find the minimum point of the cost function in Eq. 2.24. After incorporating with SGA, the tap coefficients, $\mathbf{F}_1[k]$ have a general update equation,

$$\begin{aligned} \mathbf{F}_1[k+1] &= \mathbf{F}_1[k] - \mu_1 \frac{\partial}{\partial \mathbf{F}_1[k]} J_1[k] \\ &= \mathbf{F}_1[k] - \mu_1 \varepsilon_1[k] \mathbf{X}^*[k] \end{aligned} \tag{2.25}$$

where $\mathbf{F}_1[k]$ and $\mathbf{X}[k]$ were defined in Section 2.8, μ_1 is user-defined step size and $\frac{\partial}{\partial \mathbf{F}_1[k]} J_1[k]$ is the gradient of the cost function $J_1[k]$ (Note: the algebraic manipulation of $\frac{\partial}{\partial \mathbf{F}_1[k]} J_1[k]$ can be found in some popular books such as [115, 116, 117] or calculus books).

First, we review the DD algorithm [108] which is the simplest algorithm. The error function of the DD algorithm is

$$\varepsilon_1[k] = y_1[k] - \hat{y}_1[k] \tag{2.26}$$

where $\hat{y}_1[k]$ is the decision output and is shown in Section 2.8. The DD algorithm is poor to open the channel eye because the decision output, $\hat{y}_1[k]$ is always unreliable in the closed eye condition. Therefore, the DD algorithm, which is a non-robust algorithm, is easy converge to local minima. In practice, the DD algorithm is used after the channel eye is opened because the decision output $\hat{y}_1[k]$ is only reliable in open eye conditions. Furthermore, in an open eye condition, the DD algorithm can provide the lowest MSE in comparison with the Sato algorithm and the CMA.

Sato algorithm [22] has widely been recognized as the first blind equalization algorithm because it is the first algorithm that can open the channel eye. However, the open eye ability of Sato algorithm is restricted to minimum phase channels. Due to the reason that Sato algorithm only exploits the second order statistic, it cannot distinguish between common magnitude response

channels such as $a + bz^{-1}$ or $b + az^{-1}$ [114]. Therefore, the algorithm is not able to open the non-minimum phase channel eye [118] because Sato algorithm mistakenly assumes the channel is a minimum phase channel. As a result, Sato algorithm equalizes the magnitude response of the non-minimum phase channel but it wrongly equalizes the phase response of the non-minimum phase channel. This implies that Sato algorithm will often converge to a spurious local minimum when the channel is the non-minimum phase channel. Hence, Sato algorithm is a non-robust algorithm since it cannot avoid local convergence for the non-minimum phase channel due to imperfect cost function regardless of parameters. The error function of Sato algorithm for implementation is given as

$$\varepsilon_1[k] = (|y_1[k]| - \mathrm{R_{Sato}}) \tag{2.27}$$

where $\mathrm{R_{Sato}} = \mathrm{E}\{|a_1[k]|^2\}/\mathrm{E}\{|a_1[k]|\}$.

To overcome the above problem, Godard has generalized Sato algorithm to take into account high order statistics and the generalized algorithm is called Godard algorithm [21]. Godard algorithm with fourth order statistic has been proven to successfully open the eye of even a non-minimum phase channel. At the same time, J. R. Treichler [119] independently proposed Constant Modulus Algorithm (CMA) which is exactly same as Godard algorithm with fourth order statistic. Therefore, CMA is also known as Godard algorithm. Subsequently, due to the widespread popularity of the CMA, its convergence behavior is thoroughly and exhaustively studied in [120], [109], [121], [122], [123] and [114]. These studies have identified the ill-convergence factors in CMA and proposed many solutions. Some important studies by [120] and [109] have found that proper filter taps initialization with a sufficiently small step size and at least one non-zero value in tap coefficients can avoid the occurrence of ill-convergence . This is because the previous works had identified the only undesired local minima of CMA is the origin, so initializing the otherwise all-zero equalizer coefficient vector with a single non-zero value is helpful to lead the equalizer "far away" from this specific undesirable local minimum at origin. As a result, CMA has widely been recognized as a good candidate for blind equalization since risks on CMA have been clearly

identified and effectively mitigated by the above studies. Generally, the cost function of the CMA is

$$J_1[k] = \mathrm{E}\left\{\left(|y_1[k]|^2 - \mathrm{R}_1\right)^2\right\} \tag{2.28}$$

where $\mathrm{R}_1 = \mathrm{E}\{|a_1[k]|^4\}/\mathrm{E}\{|a_1[k]|^2\}$ is the dispersion constant for equalizer output 1. However, the expectation value $\mathrm{E}\{.\}$ is unknown in practice, therefore the cost function of the CMA is widely estimated as

$$J_1[k] \triangleq \left(|y_1[k]|^2 - \mathrm{R}_1\right)^2 \tag{2.29}$$

where the instantaneous sample value is used to estimate its expectation value $\mathrm{E}\{.\}$. Finally, the error function of the CMA for implementation is given as

$$\varepsilon_1[k] = y_1[k]\left(|y_1[k]|^2 - \mathrm{R}_1\right). \tag{2.30}$$

Even though the global convergence of CMA was confirmed, CMA have some drawbacks in the quality. The drawbacks of CMA are low convergence rate and high steady state MSE for multi-modulus source symbol such as 16-QAM [54] [123]. To hasten the convergence and reduce the steady state MSE, hybrid algorithms (explained in the section below) which combine CMA and DD algorithm have been proposed.

2.14 Review on SISO Hybrid Algorithms

A hybrid algorithm that is employed by the equalizer is called that way because it opens the channel eye using an acquisition algorithm, for example CMA, and tracks the channel (after the channel eye is open) using a DD algorithm. The acquisition algorithm is capable of opening the channel eye provided sufficiently small adaptation step size but usually yields large steady state errors [123], [21], [22],[120] and normally a robust algorithm will be selected as acquisition algorithm. That is why hybrid algorithms couple the acquisition algorithm with a tracking algorithm (e.g. DD algorithm) to exploit the discrete nature of the input data to achieve much lower error rates. Hybrid

algorithms include Stop-And-Go Algorithm Decision Directed Algorithm (SAG) [124], Benveniste-Goursat Algorithm (BG) [125] and Reliability Based Algorithm (RBA)[126].

The error function, $\varepsilon_1[k]$ (in Eq. 2.24) is a function of $\epsilon_{\text{acq}}[k]$ and $\epsilon_{\text{tr}}[k]$ where $\epsilon_{\text{acq}}[k]$ and $\epsilon_{\text{tr}}[k]$ are defined as the blind equalization error function for the acquisition mode and the tracking mode at time instant k respectively. In the literature, CMA is selected as the choice of the acquisition algorithm since CMA is a good open eye blind algorithm, while DD algorithm is used as the tracking algorithm because DD algorithm yields a very low steady state MSE. Therefore, $\epsilon_{\text{acq}}[k] = y_1[k]\left(|y_1[k]|^2 - \text{R}_1\right)$ and $\epsilon_{\text{tr}}[k] = y_1[k] - \hat{y}_1[k]$ are established.

2.14.1 Stop-And-Go Algorithm Decision Directed Algorithm (SAG)

SAG is introduced by [124] with the real and imaginary part of equalization error function below:

$$\text{Re}\{\varepsilon_1[k]\} = \begin{cases} \gamma_{\text{SAG}}\text{Re}\{\epsilon_{\text{tr}}[k]\} & \text{if sgn}\left(\text{Re}\{\epsilon_{\text{tr}}[k]\}\right) = \text{sgn}\left(\text{Re}\{\epsilon_{\text{acq}}[k]\}\right) \\ 0 & \text{otherwise} \end{cases} \tag{2.31}$$

and

$$\text{Im}\{\varepsilon_1[k]\} = \begin{cases} \gamma_{\text{SAG}}\text{Im}\{\epsilon_{\text{tr}}[k]\} & \text{if sgn}\left(\text{Im}\{\epsilon_{\text{tr}}[k]\}\right) = \text{sgn}\left(\text{Im}\{\epsilon_{\text{acq}}[k]\}\right) \\ 0 & \text{otherwise} \end{cases} \tag{2.32}$$

where Re{.} and Im{.} denote real part and imaginary part respectively, sgn(.) is a signum function and γ_{SAG} is a positive constant. SAG relies on the same gradient descent direction of the acquisition algorithm and the tracking algorithm (i.e. $\text{sgn}\left(\epsilon_{\text{tr}}[k]\right) = \text{sgn}\left(\epsilon_{\text{acq}}[k]\right)$) to approach the minimum point. The equalization error is updated if the matched sign condition happened, otherwise, no updated on tap coefficient is required. This means SAG updates error function only when the error function is reliable. According to Lim et al [127], SAG is the slowest hybrid algorithm because the tap is not updated every iteration. However, SAG give the best quality in the steady state since it provides the lowest mean square error MSE.

2.14.2 Benveniste-Goursat Algorithm (BG)

Benveniste et al [125] suggested the first smooth transition hybrid method with the following equalization error function:

$$\varepsilon_1[k] = \beta_1 \left|\epsilon_{\mathrm{tr}}[k]\right| \epsilon_{\mathrm{acq}}[k] + \beta_2 \epsilon_{\mathrm{tr}}[k] \tag{2.33}$$

where β_1 and β_2 are positive constants. BG provides a smooth transition between the acquisition algorithm and the tracking algorithm. The term $\beta_1 \left|\epsilon_{\mathrm{tr}}[k]\right|$ determines the weight of the acquisition algorithm. From Eq. (2.33), BG yields larger error function since it is always occupied by two terms rather than one term, $\epsilon_{\mathrm{tr}}[k]$ in SAG. Therefore, it has higher MSE than SAG [127].

2.14.3 Reliability Based Algorithm (RBA)

The above algorithms decide the switching criteria between the acquisition algorithm and the tracking algorithm based on a single factor - the instantaneous value of $\epsilon_{\mathrm{tr}}[k]$. Based on the reason that the instantaneous value does not exactly reflect the channel eye at the onset of equalization, false switching to the tracking algorithm in the closed eye condition may occur and may reduces the convergence rate. Hence, Lim et al [126] suggested the switching criteria is not only depends on the instantaneous value but also take into account of the chances of correct output and based on maximum a posteriori (MAP) criterion. The chances of the correct output are known as the probability of correctly detecting a symbol for the output at time instant k, $P_{\mathrm{c}}[k]$. Generally, this method has higher computation cost but it outperforms the other methods. We denote this method as RBA and the equalization error function is below:

$$\varepsilon_1[k] = (1 - \alpha[k])\, \epsilon_{\mathrm{acq}}[k] + \alpha[k] \gamma_{\mathrm{RBA}} \epsilon_{\mathrm{tr}}[k] \tag{2.34}$$

where γ_{RBA} is a positive constant and $\alpha[k]$ is defined as below after some modification:

$$\alpha[k] = \begin{cases} 0 & \text{if } P_{\mathrm{c}}[k] < P_{\mathrm{th}} \text{ or } \sigma^2[k] > \sigma^2_{\mathrm{th}} \\ 2P_{\mathrm{c},j}[k] - 1 & \text{otherwise} \end{cases}$$

where P_{th} and σ^2_{th} are the positive constant and $P_{\text{c},j}$ is defined as below:

$$P_{\text{c}}[k] = \frac{\exp\left(\frac{-\left|y[k]-\hat{y}_1[k]\right|^2}{\sigma^2[k]}\right)}{\sum_{t=1}^{N_{\text{QAM}}} \exp\left(\frac{-\left|y[k]-A_t)\right|^2}{\sigma^2[k]}\right)} \tag{2.35}$$

where A_t is the tth constellation point in the N_{QAM} QAM alphabet set $\mathbf{A}$ and $\sigma^2[k]$ is the variance of $y[k]$. The estimation of $\sigma^2[k]$ in practice, $\hat{\sigma}^2[k]$ (with $\hat{\sigma}^2[0] = 10$) as below:

$$\hat{\sigma}^2[k] = 0.99\hat{\sigma}^2[k-1] + 0.01\left|y[k] - \hat{y}_1[k]\right|^2 \tag{2.36}$$

2.15 MIMO Blind equalization and source separation algorithms

The suppression of ISI and CCI in the conventional blind equalization is useful to open the channel eye, but it does not imply the receiver can retrieve all the transmitted sources. This is because the receiver has the chance to recover at least two similar symbol sequences and thus the receiver probably misses to recover some sources. In order to overcome this, blind equalization that is equipped with source separation ability is required. However, in the previous works, many types of blind equalization algorithms have been developed but not all algorithms are equipped with source separation ability. In order to identify this type algorithm, we divide the blind equalization algorithms into 4 types based on the presented simulation results.

- **Source separation algorithm:** It is a blind algorithm that can perform source separation but is designed and tested solely in CCI without ISI. This algorithm also known as blind source separation (BSS).
- **Pure equalization algorithm:** It is a blind algorithm that can suppress interferences but cannot perform source separation.
- **Single task algorithm:** It is a blind algorithm that can perform solely one task in a time, either equalization or source separation but not the both.

- **Two tasks algorithm:** It is a blind algorithm that can perform two tasks of equalization and source separation simultaneously.

Obviously, the two tasks algorithm is the algorithm that we are interested. Therefore, this type algorithm will be focused and reviewed later. Instead of that, source separation algorithm, pure equalization algorithm and single task algorithm will be briefly reviewed as follows.

2.15.1 Source separation algorithm

This type algorithm is designed based on the assumption of memoryless MIMO channels where $N_h = 1$, and thus only CCI is considered in the model. In order to recover a source signal from CCI, source separation algorithm is used. In fact, based on the implementation method, the algorithm can be divided into batch algorithm and adaptive algorithm. Batch algorithm updates the filter coefficients after the receiver obtains certain number of symbols. On the contrary, the adaptive algorithm updates the filter coefficients when a symbol is received. Therefore, the adaptive algorithm is a real time instantaneous algorithm and produces output immediately. However, the batch algorithm typically computes average on a batch of symbols, therefore the batch algorithm is less susceptible to noise. Examples of batch algorithm and adaptive algorithm are described as follows.

Principle Component Analysis (PCA), Factor Analysis (FA) [26] and Independent Component Analysis (ICA) [27, 28, 29] are the examples of batch source separation algorithm. PCA and FA are second order statistic algorithms that are used to perform uncorrelated pre-processing. However, These pre-processing algorithms cannot perform well in severe corrupted channel, therefore, high order statistics algorithm such as ICA is always appended after the pre-processing algorithm. In contrast to the pre-processing algorithms, ICA is an algorithm to ensure the receiver outputs are mutually independent. ICA can be achieved by maximizing a kurtosis cost function or a negative entropy cost function. Maximizing the kurtosis cost function can ensure the outputs becomes less statistical"Gaussainity" and thereby increases the independency among the outputs. On the

contrary, maximizing the negative entropy increases the statistical "uniformly" of the outputs and thus increases the independency among the outputs. These two cost functions achieve the independency through changing the output density functions. Due to the reason that the original information sources are i.i.d. , which is a special case of uniform distribution, the negative entropy cost function can perform better than kurtosis cost function. However, because of unknown source signals, ICA has the limitations of scale and permutation ambiguities [30].

Instead of that, Blind Source Separation Constant Modulus Algorithm (BSS-CMA) [31], Blind Source Separation Multi Modulus Algorithm (BSS-MMA) [32], Orthogonal Multi Modulus Algorithm (O-MMA) [33], Cross-Correlation Simplified Constant Modulus Algorithm (CC-SCMA) [34] are the adaptive source separation algorithms that can be found.

Blind Source Separation Constant Modulus Algorithm (BSS-CMA), which is directly adopted from SISO CMA and applied in memoryless MIMO channels, can show the ability of open eye in CCI distortion but it has some issues in permutation ambiguity, computation complexity, convergence rate and steady state performance [31]. Therefore, several solutions are proposed to overcome the issues. Blind Source Separation Multi Modulus Algorithm (BSS-MMA) is introduced to reduce the infinite permutation ambiguity to modulo 90 degree phase ambiguity [32]. By enhancing BSS-MMA, Orthogonal Multi Modulus Algorithm (O-MMA) has been proposed. Different with BSS-MMA and BSS-CMA, Orthogonal Multi Modulus Algorithm (O-MMA) has lower computation complexity and an orthogonal constrained on the filter to ensure the outputs are always mutually orthogonal, thereby the source separation performance is better than BSS-MMA [33]. Instead of that, Cross-Correlation Simplified Constant Modulus Algorithm (CC-SCMA) [34] is developed to improve the steady state performance of BSS-CMA. However, these algorithms has been developed under the assumption of invertible channels. The performance of the algorithms in non-invertible channels is unknown.

Nevertheless, due to the ISI has not been considered, the source separation algorithm is not a good choice for MIMO equalization.

2.15.2 Pure equalization algorithm

This is a classical algorithm that was design for convolutive MIMO channel where $N_h > 1$. In this case, the receiver obtains the signals that have been corrupted by ISI and CCI. Therefore, in order to suppress the interferences, MIMO-CMA where CMA in SISO system was adopted and applied in MIMO equalizer [35] [36]. The MIMO-CMA can open the channel eye of the convolutive MIMO channel but its source separation ability is not guaranteed because the same source sequence is possibly recovered more than one time at the equalizer outputs [36]. Therefore, this type algorithm is not a good candidate for MIMO equalization. To overcome this issue, single task algorithm and two tasks algorithm are developed.

2.15.3 Single task algorithm

This type of algorithm can only one task in a time. It can be only used for blind equalization or blind source separation, but not for both cases. To the best of the author's knowledge only two algorithms can be found. One is second order quadratic programming approach [37] and the other one is high order statistics approach [38]. Nevertheless, due to this type algorithm cannot perform equalization and source separation simultaneously, it is not a good candidate for MIMO equalization.

2.15.4 Two tasks algorithm

Two tasks algorithm is aimed to perform equalization and source separation at the same time. In general, the tasks can be achieved through orthogonal constraint cost function approach or non-constraint cost function approach. Non-constraint cost function approach performs the equalization and source separation tasks by simply minimizing a cost function, whilst orthogonal constraint cost function approach performs the tasks by minimizing (or maximizing) a cost function with an extra orthogonal constraint function. Instead of that, the CC-CMA is the example of the non-constraint cost function that can be found. Orthogonal constraint cost function approach and CC-CMA are

reviewed as follows.

2.15.4.1 Orthogonal constraint cost function approach

Signal interferences are formed by the superposition of many symbols and thus interferences are a linear weight combination of symbols. Therefore, if the linear weight matrix is known, the symbols can be retrieved. In order to achieve this, the estimation of inverse weight matrix is the main key of this approach. Therefore, the algorithm in this approach is to indirectly estimate the inverse cross-correlation channel matrix in order to recover the symbols. Due to the reason that the inverse matrix is the product of several orthogonal matrices, the algorithm estimates the orthogonal matrices one by one through cascade system (or through iterative computing) and each of the orthogonal matrices can produce one output. In the literature, there are many ways to select the form or"pattern" of the orthogonal matrices and this is formally called matrix decomposition. Given Rotation Decomposition [128, 129] and Eigen Value Decomposition [39, 40] and PCA [41] are the matrix decomposition methods that can be found. The main purpose of the orthogonal matrices is to ensure the outputs are uncorrelated to each and thereby can achieve source separation task in the memory channels. At the same time to perform the orthogonalization process, high order statistics algorithm such as kurtosis or CMA is used to open the channel eye. However, this approach requires the channel order or channel length to be known. In practice, due to the channel is unknown, it is difficult to obtain the channel order. Furthermore, because matrix decomposition requires large computation efforts and large storage, the computation complexity of this method is high. Instead of that, noise has not been considered in the design.

2.15.4.2 Cross Correlation Constant Modulus Algorithm (CC-CMA)

To overcome the above issues, a relatively simple algorithm that is non-sensitive to channel order and less susceptible to noise is motivated. Therefore, Cross Correlation Constant Modulus Algorithm (CC-CMA) is proposed [2, 31, 42, 43, 44, 45, 46, 47]. In this algorithm, the CMA which has the open eye ability is combined with a cross-correlation (CC) cost which penalizes correlations

between the source signals and including their delayed versions thereby separating them. The CC-CMA will converge towards the desired solution which is the goal of open eye with all the sources separated, with a properly chosen mixing parameter which weighs the constant modulus error and the correlation penalty appropriately, and under some mild assumptions [51]. This subsection is structured as following. Firstly, we describe the approach of the CC-CMA in a two outputs MIMO system and then mathematically state the cost functions of the CC-CMA. After that, we state the error functions of the CC-CMA for implementation purpose. At the end of this subsection, we point out the weaknesses of CC-CMA which motivate our research.

For simplicity, we review the cost functions of the CC-CMA using a two outputs system where there are two equalizer outputs. In order to reduce the computation complexity, different cost functions were suggested for different equalizer outputs [2]. Therefore, in the system, the equalizer output 1 is equipped with a CMA cost function alone to open the channel eye without source separation ability. On the other hand, the equalizer output 2 is equipped with a CC-CMA cost function which consists of a CMA cost function to open the channel eye and a CC cost function to perform source separation. Hence, the cost functions of the CC-CMA for the first and the second equalizer outputs which are $J_{1,\mathrm{CC-CMA}}[k]$ and $J_{2,\mathrm{CC-CMA}}[k]$ respectively are given as

$$J_{1,\mathrm{CC-CMA}}[k] = J_{1,\mathrm{CMA}}[k] \tag{2.37}$$

$$J_{2,\mathrm{CC-CMA}}[k] = J_{2,\mathrm{CMA}}[k] + 2k_2 J_{2,\mathrm{CC}}[k] \tag{2.38}$$

where $J_{1,\mathrm{CMA}}[k]$ and $J_{2,\mathrm{CMA}}[k]$ denote the CMA cost functions of the equalizer output 1 and 2 respectively, and $J_{2,\mathrm{CC}}[k]$ denotes the CC cost function of the equalizer output 2. These cost functions will be explained in the next paragraph. In addition, $k_2 \in \mathbb{R}^+$ is known as the mixing parameter for the second equalizer and it can be used to achieve the source separation feature when a suitable k_2 is chosen. The above CC-CMA approach (Eq. 2.37–2.38) is also known as *hierarchical* CC-CMA since the CC cost function is developed for the second equalizer output instead of the first equalizer output. The other approach is called the *symmetrical* CC-CMA, adopted by [31, 36], where the CC cost function is developed for every equalizer output. *Symmetrical* CC-CMA is less

attractive because it has higher computation complexity and also generates the maximum amount excess MSE for every equalizer output. Therefore, we will pursue *hierarchical* CC-CMA in this thesis.

The CMA cost functions (that was initially developed for SISO systems) are given as

$$J_{1,\mathrm{CMA}}[k] = \mathrm{E}\left\{\left(|y_1[k]|^2 - \mathrm{R}_1\right)^2\right\} \tag{2.39}$$

$$J_{2,\mathrm{CMA}}[k] = \mathrm{E}\left\{\left(|y_2[k]|^2 - \mathrm{R}_2\right)^2\right\} \tag{2.40}$$

for the first and the second equalizer outputs respectively, where $\mathrm{R}_1 = \mathrm{R}_2 = \frac{m_4}{\sigma_{\mathrm{A}}^2}$ is the dispersion constant which is adopted from the SISO CMA as well. To achieve blind source separation on the second equalizer output, CC cost function that penalizes the cross correlations between all source signals and including their delayed versions is defined as

$$J_{2,\mathrm{CC}}[k] = \sum_{l=1}^{M_\mathrm{t}} \sum_{n=1;n\neq l}^{M_\mathrm{t}} \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \left|\mathrm{E}\{y_n[k]y_l^*[k-\delta]\}\right|^2 \tag{2.41}$$

where $\delta_{\max}$ accounts for the maximum channel delay spread between all equalizer output and we recall that $M_\mathrm{t} = 2$ in this example.

For implementation purpose, the coefficients of the equalizer output 1 and 2, $\mathbf{F}_1[k]$ and $\mathbf{F}_2[k]$ respectively, are updated as below

$$\mathbf{F}_1[k+1] = \mathbf{F}_1[k] - \mu_1\varepsilon_1[k]\mathbf{X}^*[k] \tag{2.42}$$

$$\mathbf{F}_2[k+1] = \mathbf{F}_2[k] - \mu_2\varepsilon_2[k]\mathbf{X}^*[k] \tag{2.43}$$

where μ_1 and μ_2 is the adaptation step size for the equalizer output 1 and 2 respectively, $\varepsilon_1[k]$ is the error function of the first equalizer given as

$$\varepsilon_1[k] = (|y_1[k]|^2 - \mathrm{R}_1)y_1[k] \tag{2.44}$$

and $\varepsilon_2[k]$ is the error function of the second equalizer given as

$$\varepsilon_2[k] = (|y_2[k]|^2 - \mathrm{R}_2)y_2[k] + k_2 \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \hat{r}_{1,\delta}[k]y_1[k-\delta] \tag{2.45}$$

where $\hat{r}_{1,\delta}[k]$ is an estimate of $r_{1,\delta}[k] \equiv \mathrm{E}\{y_2[k]y_1^*[k-\delta]\}$, which may be recursively computed as

$$\hat{r}_{1,\delta}[k] = \lambda \hat{r}_{1,\delta}[k-1] + (1-\lambda)y_2[k]y_1^*[k-\delta]. \tag{2.46}$$

where $0 \leq \lambda < 1$ is the forgetting factor of the recursion.

In spite of that, CC-CMA has some limitations. The limitations are described as follows.

First, CC-CMA has the problem of shrinking effect where the equalizer outputs are inevitable scaled down or the simulation results in [2]). In contrary, the convergence analysis in [2] has shown that the scaled down phenomenon should not happen. Due to the reason that the interference residual performance measurement (i.e. Eq. 2.21) is normalized by the largest scale, this measurement is scale insensitive. Therefore, it is no surprise that some previous works (such as [48], [44] and [49]) did not realize the shrinking effect in CC-CMA. Furthermore, Luo et al independently performed the CC-CMA convergence analysis in [51], [52] and [53] and they shown that CC-CMA cannot converge to the the desired minimum point if mixing parameter, k_2 is large because the desired global minimum point disappears in this situation. This means CC-CMA cannot converge well as expected due to the shrinking effect and the equalizer output 2 possibly converges to zero. Even though some mitigation methods had been proposed in [51] and [52], the problem has not fully been solved due to the root cause of shrinking effect is still unknown.

Furthermore, the simulation results in [50], [48] and [44] have shown that the CC-CMA suffers high steady state MSE compared to conventional CMA.

Instead of that, the simulation tests [2, 51, 52, 53] that are supported by theoretical analysis have been done on single-amplitude constellation such as BPSK or QPSK only, but no similar study on multi-amplitude constellation can be found. Multi-amplitude constellation, such as 8PAM and 16QAM, can provide higher data rate than single-amplitude constellation because a symbol in multi-amplitude constellation can carries higher number of bits than single-amplitude constellation. However, the statical properties of multi-amplitude constellation, such as second and forth order moments, are different with single-amplitude constellation, thereby the analysis for single-amplitude constellation cannot directly been applied on multi-amplitude constellation. Therefore,

study of CC-CMA on multi-amplitude constellation is required.

Chapter 3

Cross-Independent Constant Modulus Algorithm

3.1 Introduction

The existing MIMO blind equalization algorithms by [31, 36, 2] introduced an undesirable and unexpected amplitude shrinking in the equalizer output. The shrinking intensifies when the number of sources increases, resulting in prolonged or unsuccessful retrieval of the sources even with just a few sources. Therefore, the main challenge of existing MIMO blind equalization algorithms is to retrieve all source sequences with unity gain. However, none of the existing solution is considered effective since the root cause of shrinking effect has not been exactly determined until our work. It is worth to mention that an automatic gain compensator is not an effective solution, even though it can mitigate the shrinking, because the shrinking is not introduced by the system structure but it is caused by improper computation within the algorithm. In this chapter, we will identify and prove the root cause of the shrinking is due to the inappropriate selection of the cost function. Therefore, we will propose a new source separation cost that is mixed to the MIMO CMA. The new cost not only reduces correlation level among the equalizer outputs, but also increases the

outputs' interdependency. We call the new mixed cost as Cross Independence Constant Modulus cost function.

The main advantage of this new cost is to solve the shrinking issue so that its source separation ability does not deteriorate with any number of delay spread and any number of users.

The chapter is organized as follows. We present the system model and state several assumptions in Section 3.2.1, recall the traditional mixed cost functions in Section 3.2.2 and formulate the shrinking issue in Section 3.2.3. We propose new algorithm in Section 3.3.1. From Section 3.3.2 to Section 3.4, we provide vigorous convergence analysis on the algorithm to confirm its open eye and source separation abilities. Section 3.5 presents the simulation results and discussions.

3.2 Background

3.2.1 System model and assumptions

We consider a system model that has M_{t} transmit sources and $M_{\mathrm{r}} \geq M_{\mathrm{t}}$ receive sensors. Assumptions in Section 2.8 are reapplied as following:

A1 Assumptions on sources : Same as assumption A1 in Section 2.8.

A2 Assumption on channels: Same as assumption A2 in Section 2.8.

A3 Assumption on sensor noise(s): For simplicity, we assume no sensor noise, therefore $\sigma_{\mathrm{n}}^2 = 0$.

Again, we recall that the observed signal at the i-th sensor, for $i \in (1, M_{\mathrm{r}})$, at time k is defined as:

$$x_i[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij} \otimes \mathbf{a}_j[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij}^T \mathbf{a}_j[k] \tag{3.1}$$

where $\mathbf{a}_j[k] = \big[a_j[k], \cdots, a_j[k - N_h + 1]\big]^T$. The equalizer output corresponding to the j-th source, where $j \in (1, M_{\mathrm{t}})$, at time k is denoted as:

$$y_j[k] = \sum_{i=1}^{M_{\mathrm{r}}} \mathbf{f}_{ij}^T[k] \mathbf{x}_i[k] \tag{3.2}$$

where $\mathbf{x}_i[k] = [x_i[k], \cdots, x_i[k - L + 1]]^T$ is the input regressor vector for the equalizers of the i-th seonsor, $L \in \mathbb{Z}^+$ denotes the equalizer length, and $\mathbf{f}_{ij}[k] = \Big[f_{ij}^{[0]}[k], \cdots, f_{ij}^{[L-1]}[k]\Big]^T$ is a $L \times 1$ vector that represents the equalizer connecting the observed signal at the i-th sensor to the j-th output at time k. There are a total of $(M_{\mathrm{t}} \times M_{\mathrm{r}})$ equalizers at the receiver end (see Fig. 3.1). Furthermore, (3.2) can be compactly written as

$$y_j[k] = \mathbf{F}_j^T[k] \mathbf{X}[k] \tag{3.3}$$

where $\mathbf{F}_j[k] = \big[\mathbf{f}_{1j}^T[k], \cdots, \mathbf{f}_{M_{\mathrm{r}}j}^T[k]\big]^T$ is a $M_{\mathrm{r}}L \times 1$ vector and the input regressor $\mathbf{X}[k] = \big[\mathbf{x}_1^T[k], \cdots, \mathbf{x}_{M_{\mathrm{r}}}^T[k]\big]^T$ has the same dimensions too.

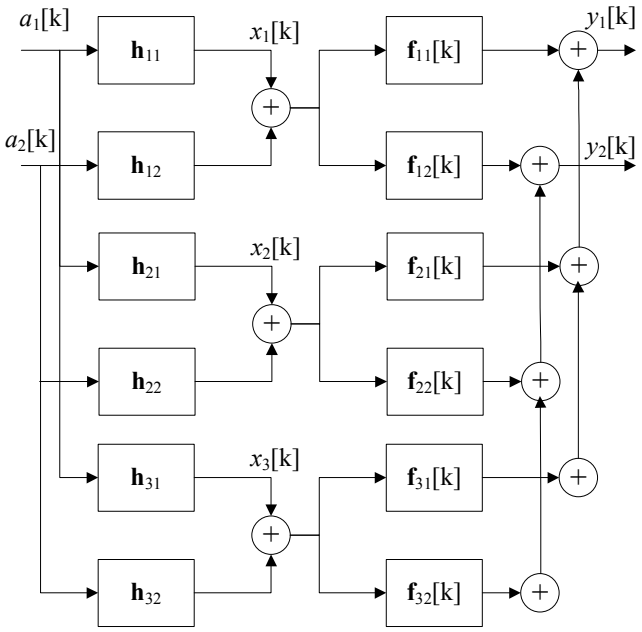

Figure 3.1: Baseband equivalent system for $M_t = 2$ and $M_r = 3$.

In the study of the convergence behavior of a blind algorithm later, the global impulse response that connects the j-th equalizer output with its respective sources is often preferred. Therefore we provide the following expression by substituting (3.1) into (3.3):

$$y_j[k] = \mathbf{F}_j^T[k]\mathbf{H}^T\overline{\mathbf{a}}[k] = \mathbf{S}_j^T[k]\overline{\mathbf{a}}[k] \tag{3.4}$$

where $\mathbf{S}_j[k] = \mathbf{H}\mathbf{F}_j[k]$ is the global impulse response with dimensions $M_{\mathrm{t}}N_s \times 1$, where $N_s = L + N_h - 1$, and $\overline{\mathbf{a}}[k] = \left[\overline{\mathbf{a}}_1^T[k], \overline{\mathbf{a}}_2^T[k], \ldots, \overline{\mathbf{a}}_{M_{\mathrm{t}}}^T[k]\right]^T$ denotes the $M_{\mathrm{t}}N_s \times 1$ vector where $\overline{\mathbf{a}}_j[k] = [a_j[k], a_j[k-1], \ldots, a_j[k-(N_s-1)]]^T$ is a $N_s \times 1$ source vector for $1 \leq j \leq M_{\mathrm{t}}$. In addition, $\mathbf{H}$, which was defined in Assumption A2 in Section 2.2, is a channel convolution matrix assumed to be full column rank [43] and has dimensions $M_{\mathrm{r}}L \times M_{\mathrm{t}}N_s$.

3.2.2 Classical mixed cost approach

A classical approach to perform blind signal separation (BSS) of a linear convolutive mixture of signals is to combine the costs of the CMA, $J_{j,\mathrm{CM}}(\mathbf{F})$, and a source separation cost function, $J_{j,\mathrm{SS}}(\mathbf{F})$. Therefore the mixed cost function responsible for the extraction of the j-th source is

$$J_{j,\mathrm{SS-CMA}}(\mathbf{F}) = J_{j,\mathrm{CMA}}(\mathbf{F}) + 2k_j J_{j,\mathrm{SS}}(\mathbf{F}) \tag{3.5}$$

where k_j is known as the mixing parameter that corresponds to the equalizer output j and $k_1 = 0$ was suggested to reduce the computation complexity [2]. The CMA opens the channel eye by minimizing the following cost [36]:

$$J_{j,\mathrm{CM}}(\mathbf{F}) = \mathrm{E}\left(|y_j|^2 - \mathrm{R_j}\right)^2 \tag{3.6}$$

where $\mathrm{R_j} = m_4/\sigma_{\mathrm{A}}^2$ is known as the dispersion constant [21], whereas the BSS algorithm minimizes this cost:

$$J_{j,\mathrm{SS}}(\mathbf{F}) = \sum_{n=1}^{j-1} \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \zeta_{n,\delta} \tag{3.7}$$

where $\zeta_{n,\delta}$ is a function of the j-th output and the n-th output which is delayed by $-\delta_{\mathrm{x}} \leq \delta \leq \delta_{\mathrm{x}}$ symbols relative to the j-th output, where $\delta_{\mathrm{x}} \in \mathbb{Z}^+ \geq N_s - 1$ is the maximum achievable delay between all equalized outputs. This function is responsible for blind source separation and may take the form of the following cross-correlation (CC) function [2, 31, 42, 43, 44, 45, 46],

$$\zeta_{n,\delta} \triangleq \zeta_{\mathrm{CC},n,\delta} = \left|\mathrm{E}\{y_j[k]y_n^*[k-\delta]\}\right|^2. \tag{3.8}$$

For practical implementation under stochastic gradient adaptation, the coefficients of the j-th equalizer are updated as follows:

$$\mathbf{F}_j[k+1] = \mathbf{F}_j[k] - \mu_j \frac{\partial J_{j,\mathrm{SS-CMA}}(\mathbf{F})}{\partial \mathbf{F}_j[k]} \tag{3.9}$$

$$= \mathbf{F}_j[k] - \mu_j \; e_{\mathrm{mix},j}[k] \; \mathbf{X}^*[k] \tag{3.10}$$

where μ_j is a small step size associated with the j-th equalizer and, with Eq. 3.5, the general

gradient, $\frac{\partial J_{j,\mathrm{SS-CMA}}(\mathbf{F})}{\partial \mathbf{F}_j[k]}$ is

$$\frac{\partial J_{j,\mathrm{SS-CMA}}(\mathbf{F})}{\partial \mathbf{F}_j[k]} \equiv \frac{\partial J_{j,\mathrm{CM}}(\mathbf{F})}{\partial \mathbf{F}_j[k]} + k_j \frac{\partial J_{j,\mathrm{SS}}(\mathbf{F})}{\partial \mathbf{F}_j[k]} \tag{3.11}$$

$$= (\varepsilon_{j,\mathrm{CM},j}[k] + k_j \varepsilon_{j,\mathrm{SS}}[k])\mathbf{X}^*[k] = e_{\mathrm{mix},j}[k]\mathbf{X}^*[k]. \tag{3.12}$$

where $\varepsilon_{j,\mathrm{CM},j}[k]$, $\varepsilon_{j,\mathrm{SS},j}[k]$ and $e_{\mathrm{mix},j}[k] = \varepsilon_{j,\mathrm{CM}}[k] + k_j\varepsilon_{j,\mathrm{SS}}[k]$ are known as the CMA error, the source separation error and the mixed error respectively for the j-th equalizer.

3.2.3 Motivation: shrinking of the equalizer output

We now highlight an implementation difficulty associated with the traditional mixed cost method, which arises due to the source separation cost, Eq. (3.7). Consider the CC function in Eq (3.8). It is a square-modulus function of an expectation function in the form of $|\mathrm{E}(\cdot)|^2$ whereas conventional costs are regular expectation functions, $\mathrm{E}(\cdot)$, which gives an unbiased estimate when instantaneous values are taken. With the modulus function, the expectation function within it cannot be simply omitted without incurring a bias in the estimator. A commonly adopted approach is to explicitly estimate $\frac{\partial J_{\mathrm{SS},j}(\mathbf{F})}{\partial \mathbf{F}_j[k]}$ based on instantaneous values or sample averaging [31]. In symbol-rate implementations, for example, a recursive estimator can be used to perform sample averaging as follows [46]

$$\hat{r}_{n,\delta}[k] = \lambda \hat{r}_{n,\delta}[k-1] + (1-\lambda) y_j[k] y_n^*[k-\delta] \tag{3.13}$$

and $\lambda \in (0, 1]$ is a forgetting factor. When $\lambda = 0$, the process becomes estimation by taking instantaneous values. The impact of using the above recursive estimator has not been studied before, and it is the purpose of this subsection.

Proposition: This estimator will introduce an undesirable bias that amounts to $(j-1)(2\delta_x + 1)\sigma_{\mathrm{A}}^4(1-\lambda)/(1+\lambda)$.

Proof: We determine the bias by evaluating the mean of the estimate as follows:

$$\mathrm{E}\{|\hat{r}_{n,\delta}[k]|^2\}$$

$$= \mathrm{E}\left\{\left|\lambda\hat{r}_{n,\delta}[k-1] + (1-\lambda)y_j[k]y_n^*[k-\delta]\right|^2\right\} \tag{3.14}$$

$$= \mathrm{E}\left\{\left|(1-\lambda)\Sigma_{d=0}^{\infty}\lambda^d y_j[k-d]y_n^*[k-\delta-d]\right|^2\right\} \tag{3.15}$$

$$= (1-\lambda)^2\sum_{d=0}^{\infty}\lambda^{2d}\mathrm{E}\left\{\left|a_j[k-d]\right|^2\right\}\mathrm{E}\left\{\left|a_n[k-\delta-d]\right|^2\right\} \tag{3.16}$$

$$= (\sigma_{\mathrm{A}}^2)^2(1-\lambda)^2\Sigma_{d=0}^{\infty}\lambda^{2d} \tag{3.17}$$

$$= \sigma_{\mathrm{A}}^4(1-\lambda)/(1+\lambda) > 0. \tag{3.18}$$

The important manipulations of Equations 3.14–3.18 are stated as below.

- From Eq. 3.14 to Eq. 3.15: Using Eq. 3.13, we evaluate all past CC estimates [1], i.e., $\lambda\hat{r}_{n,\delta}[k-1]$, $\lambda\hat{r}_{n,\delta}[k-2]$,..., $\lambda\hat{r}_{n,\delta}[k-\infty]$ and then recursively substituting them back until we get $\lambda\hat{r}_{n,\delta}[k-1]$ and therefore we get Eq. 3.15.

- From Eq. 3.15 to Eq. 3.16: The largest concern of this portion is how to mathematically describe the equalizer outputs ($y_j[k]$ and $y_n[k]$) in term of the source symbols ($a_j[k]$ and $a_n[k]$). In the asymptotic condition or the steady state (i.e. time $k \to \infty$), the source symbols are assumed to be perfectly recovered and separated without any interference in Eq. 3.4 (where this is a common assumption to verify a bias [113]), thus we have $y_j[k] \to a_j[k]$ and $y_n[k] \to a_n[k]$. In addition, $a_n[k]$ and $a_j[k]$ (note: $n \neq j$ is implicity stated in Eq.3.7) are mutually independent and discrete time stationary in Assumption A1 (Section 2.8), therefore $\mathrm{E}\left\{\left|a_j[k-d]\right|^2\left|a_n[k-\delta-d]\right|^2\right\}$ can be simplified as $\mathrm{E}\left\{\left|a_j[k-d]\right|^2\right\}\mathrm{E}\left\{\left|a_n[k-\delta-d]\right|^2\right.$. With the above information, we get Eq. 3.16 after some algebraic manipulations.

- From Eq. 3.16 to Eq. 3.17: We recall Assumption A1 in Section 2.8 that the source symbols has a common positive variance, σ_{A}^2. Therefore, we have $\sigma_{\mathrm{A}}^2 = \mathrm{E}\left\{\left|a_j[k-d]\right|^2\right\} = \mathrm{E}\left\{\left|a_n[k-\delta-d]\right|^2\right\}$ and we emphasize $(\sigma_{\mathrm{A}}^2)^2 = \sigma_{\mathrm{A}}^4$ to avoid confusion.

[1] For example, to evaluate $\hat{r}_{n,\delta}[k-d]$, we can replace k with $k-d$ in Eq. 3.13, therefore we get $\hat{r}_{n,\delta}[k-d] = \lambda\hat{r}_{n,\delta}[k-1-d] + (1-\lambda)y_j[k-d]y_n^*[k-\delta-d]$.

- From Eq. 3.17 to Eq. 3.18: We want to show the equality of $\Sigma_{d=0}^{\infty}\lambda^{2d} = \frac{1}{(1-\lambda)(1+\lambda)}$ in this portion. Due to the range of the forgetting factor, λ is given as $0 \leq \lambda < 1$ in Eq. 3.13, it is obvious that the range of its square, λ^2 is $0 \leq \lambda^2 < 1$. With $\lambda^2 < 1$, we can simplify the summation of the infinite geometric series, $\Sigma_{d=0}^{\infty}\lambda^{2d} = \Sigma_{d=0}^{\infty}(\lambda^2)^d = 1+\lambda^2+\lambda^4+\ldots$ as $\frac{1}{1-\lambda^2}$. Furthermore, we recall a common type of algebraic factoring which is $1-\lambda^2 = 1^2-\lambda^2 = (1+\lambda)(1-\lambda)$, therefore we can factorize the denominator of $\frac{1}{1-\lambda^2}$ and finally we obtain $\frac{1}{(1-\lambda)(1+\lambda)}$.

- **Comment on Eq. 3.18**: Since λ has the range $0 \leq \lambda < 1$, the value of Eq. 3.18 is always larger than zero, therefore the CC estimator, $\hat{r}_{n,\delta}[k]$ is an asymptotic bias estimator even in the perfect symbol recovery and the perfect source separation conditions. In fact, an unbiased estimator should give a value of zero in Eq. 3.18. Again, it is worth to mention that the previous works (such as [2, 31, 42, 43, 44, 45, 46]) had started their developments based on the assumption that the CC estimator, $\hat{r}_{n,\delta}[k]$ is an unbiased estimator. However, our works above have proven that the CC estimator, $\hat{r}_{n,\delta}[k]$ is a bias estimator.

The two outer sums in (3.7) introduce the factor $(j-1)(2\delta_x+1)$ to the total bias, which completes the proof.

This finite bias is the reason behind the equalizer output shrinking in magnitude, a behaviour that can be explained using our extrema (stationary point) analysis in Section 3.3.4. The shrinking effect is further accentuated with increased number of users and delay spread, hence restricting the usefulness of the algorithm to MIMO systems with few users and/or small delay spread, for example in polarized optical receivers where only two orthogonal polarizations (users) exist [46].

3.3 New Cross Independent Constant Modulus Algorithm (CI-CMA)

3.3.1 A new BSS cost: the cross independent function

Our intention is to propose a new BSS cost that is conventional in the sense that it conforms to $\mathrm{E}(\cdot)$. Therefore we propose the so-called "cross-independent" (CI) function

$$\zeta_{\mathrm{CI},n,\delta} = \mathrm{E}\left\{|y_j y_{n,\delta}^*|^2\right\} \tag{3.19}$$

where its derivative w.r.t. $\mathbf{F}_j[k]$ is

$$\frac{\partial}{\partial \mathbf{F}_j[k]}\zeta_{\mathrm{CI},n,\delta} = y_j[k]y_n^*[k-\delta]y_n[k-\delta]\mathbf{X}^*[k]. \tag{3.20}$$

The minimization of the CI cost leads to a stronger independence property (than decorrelation) among the equalizer outputs whose proof is documented in Appendix A.1 (at the end of this thesis). Therefore the complete algorithm for blind equalization and source separation is the combination of the MIMO CMA and this CI cost, called the CI-CMA, which has the following cost function,

$$J_{j,\mathrm{CI-CMA}}(\mathbf{F}) = J_{j,\mathrm{CMA}}(\mathbf{F}) + k_0 \sum_{n=1}^{j-1} \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \zeta_{\mathrm{CI},n,\delta} \tag{3.21}$$

and the corresponds error term is

$$e_{\mathrm{mix},j}[k] = \left(|y_j[k]|^2 - \hat{\mathrm{R}}_j\right) y_j[k] + \frac{k_0}{2} \sum_{n=1}^{j-1} \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \left(|y_n[k-\delta]|^2\right) y_j[k]. \tag{3.22}$$

where k_0 is the new mixing parameter of the CI cost and its value will be determined later. We would like to draw the attention of the reader to the dispersion constant of the CMA, $\hat{\mathrm{R}}_j$, which has been modified to ensure the original signal energy is preserved at its corresponding equalizer output. It will be proven in the sequel that the new value of $\hat{\mathrm{R}}_j$ solves the output-shrinking problem encountered by past approaches. Under the hierarchical criteria [2], we have $\hat{\mathrm{R}}_1 < \hat{\mathrm{R}}_2 < \cdots < \hat{\mathrm{R}}_{M_{\mathrm{t}}}$, and $\hat{\mathrm{R}}_1 = \frac{m_4}{\sigma_{\mathrm{A}}^2}$ because for $j = 1$ the CI-CMA reduces to the classical MIMO CMA without the BSS cost. For completeness, we state the value of $\hat{\mathrm{R}}_j$ under the symmetrical criteria [31] to be

$\hat{\mathrm{R}}_1 = \hat{\mathrm{R}}_2 = \cdots = \hat{\mathrm{R}}_{M_\mathrm{t}}$ but $\hat{\mathrm{R}}_j \neq \frac{m_4}{\sigma_\mathrm{A}^2}, \forall j$. The expression of the dispersion constant is left as the subject of the next subsection when the stationary points of the cost surface are derived.

In what follows, we first search for the global minimum points of the CI-CMA under gradient descent adaptation. We subsequently determine the conditions for which all spurious local minima of the CI-CMA would vanish, so that the global minima is always achieved. A real-valued PAM system is assumed here whereas the treatment of the complex modulation is found in Section 3.4.

3.3.2 Preliminary assumptions and useful notations

For simplicity but without losing generality, we confine our analysis to a $M_\mathrm{t} = 2$ MIMO system. Therefore for $j = 1$ the MIMO CMA extracts any source, but usually the one that enjoys a stronger SNR, while for $j = 2$ the equalizer must avoid the retrieval of the same source thanks to the CI cost.

For convenience, the global impulse response of the second equalizer is expressed as a concatenation of two vectors, i.e., $\mathbf{S}_2 = [\mathbf{S}_{(1),2}^T, \mathbf{S}_{(2),2}^T]^T$ which is $M_\mathrm{t}N_s = 2N_s$ long (see Eq (3.4)), where $\mathbf{S}_{(b),2} = [S_2^{[(b-1)N_s]}, S_2^{[(b-1)N_s+1]}, \cdots, S_2^{[bN_s-1]}]^T$ is the global filter vector that connects the b-th source to the second equalizer. The time dependency of $\mathbf{S}_2$ on k is dropped because the analysis is based on asymptotic conditions. The global impulse response of the first equalizer, which will be required later, are also expressed as vectors $\vec{\alpha} \triangleq \mathbf{S}_{(1),1}$ and $\vec{\beta} \triangleq \mathbf{S}_{(2),1}$ connecting the first and second sources, respectively, to the output of the first equalizer, i.e., $\mathbf{S}_1 \triangleq \left[\vec{\alpha}^T, \vec{\beta}^T\right]^T$. The use of new variables $\vec{\alpha}$ and $\vec{\beta}$ is to conveniently distinguish $\mathbf{S}_1$ from $\mathbf{S}_2$. Their respective i-th elements are $\alpha_i = S_1^{[i]}$ and $\beta_i = S_1^{[N_s+i]}$, where $i \in (0, N_s - 1)$.

We further assume that in analyzing the j-th equalizer that the previous $(j-1)$ equalizers up in the hierarchy have been perfectly equalized. In our $M_\mathrm{t} = 2$ case, we assume the first equalizer has perfectly recovered one of the sources (usually the stronger one) and thus $\mathrm{E}\{\alpha_i\}|_{i\neq d} = 0$ where the only non-zero impulse is at the $(d+1)$-th tap, where $d \in (0, N_s - 1)$. In fact $\mathrm{E}\{\alpha_d\} = 1$. We assume that it has also canceled all CCI so $\mathrm{E}\{\beta_i\} = 0, \forall i$. To be more practical, we further assume

the variance of the non-dominant taps are finite, i.e., $\Sigma_{i\neq d}\mathrm{E}\left\{|\alpha_i|^2\right\} > 0$ and $\Sigma_j \mathrm{E}\left\{|\beta_j|^2\right\} > 0$. Due to the unity-impulse in α_d, we have the relationship of $\mathrm{E}\|\vec{\alpha}\|^2 > \mathrm{E}\|\vec{\beta}\|^2$ where $\|\cdot\|$ is the Euclidean norm operator.

3.3.3 Extrema analysis

We now begin the stationary point analysis. For brevity, we denote $J_2[k] \triangleq J_{\mathrm{mix},2}(\mathbf{F})$. We evaluate the gradient of $J_2[k]$ w.r.t. $\mathbf{S}_2$ in the following equations from (3.25) to (3.33):

$$\frac{\partial J_2[k]}{\partial \mathbf{S}_2} = \left[\left(\frac{\partial J_2[k]}{\partial \mathbf{S}_{(1),2}}\right)^T, \left(\frac{\partial J_2[k]}{\partial \mathbf{S}_{(2),2}}\right)^T\right]^T \tag{3.25}$$

where

$$\frac{\partial J_2[k]}{\partial \mathbf{S}_{(1),2}} = \left[\frac{\partial J_2[k]}{\partial S_2^{[0]}}, \frac{\partial J_2[k]}{\partial S_2^{[1]}}, \ldots, \frac{\partial J_2[k]}{\partial S_2^{[N_s-1]}}\right]^T$$

and

$$\frac{\partial J_2[k]}{\partial \mathbf{S}_{(2),2}} = \left[\frac{\partial J_2[k]}{\partial S_2^{[N_s]}}, \frac{\partial J_2[k]}{\partial S_2^{[N_s+1]}}, \ldots, \frac{\partial J_2[k]}{\partial S_2^{[2N_s-1]}}\right]^T.$$

We would like to express $J_2[k]$ in terms of the global impulse response by using (3.4), so at first the CM cost is rewritten as

$$J_{\mathrm{CM},j} = \mathrm{E}\left\{\left(\left|\mathbf{S}_j^T[k]\bar{\mathbf{a}}[k]\right|^2 - \hat{\mathrm{R}}_j\right)^2\right\}. \tag{3.26}$$

Bearing in mind that the CI cost in Eq. (3.21) for $j = 2$ reduces to $\Sigma_\delta \zeta_{\mathrm{CI},1,\delta} = \Sigma_\delta \mathrm{E}\left\{|y_2 y_{1,\delta}^*|^2\right\} = \Sigma_\delta \mathrm{E}\left\{y_2^2 y_{1,\delta}^2\right\}$, the mixed-cost becomes

$$\begin{aligned} J_2[k] = \mathrm{E}\Big\{ &\left((\mathbf{S}_{(1),2}^T\bar{\mathbf{a}}_1[k] + \mathbf{S}_{(2),2}^T\bar{\mathbf{a}}_2[k])^2 - \hat{\mathrm{R}}_2\right)^2\Big\} \\ &+ k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \mathrm{E}\Big\{(\mathbf{S}_{(1),2}^T\bar{\mathbf{a}}_1[k] + \mathbf{S}_{(2),2}^T\bar{\mathbf{a}}_2[k])^2(\vec{\alpha}^T\bar{\mathbf{a}}_1[k-\delta] + \vec{\beta}^T\bar{\mathbf{a}}_2[k-\delta])^2\Big\}. \end{aligned} \tag{3.27}$$

After significant amount of mathematics, it can be shown that the gradient of the mixed cost is

$$\frac{\partial J_2[k]}{\partial \mathbf{S}_2} = 4\mathbf{\Lambda}\mathbf{S}_2 = 4(\mathbf{\Lambda}_{\mathrm{CM}} + \mathbf{\Lambda}_{\mathrm{CI}})\mathbf{S}_2 \tag{3.28}$$

where $\mathbf{\Lambda}$ is a $2N_s \times 2N_s$ diagonal matrix which can be split into $\mathbf{\Lambda}_{\mathrm{CM}}$ and $\mathbf{\Lambda}_{\mathrm{CI}}$, where they are derived from the CM and CI costs, respectively. To evaluate $\mathbf{\Lambda}_{\mathrm{CM}}$, we evaluate the gradient of $J_{\mathrm{CM},j}$ w.r.t. $S_2^{[d_0]}$ as follows

$$\frac{\partial J_{\mathrm{CM},j}}{\partial S_2^{[d_0]}} = 4S_2^{[d_0]}\Lambda_{\mathrm{CM}}^{[d_0]} \tag{3.29}$$

where

$$\Lambda_{\mathrm{CM}}^{[d_0]} = \left(S_2^{[d_0]}\right)^2 m_4 - 3\left(S_2^{[d_0]}\right)^2 \sigma_A^4 + 3\sigma_A^4\|\mathbf{S}_2\|^2 - \hat{\mathrm{R}}_2\sigma_A^2 \tag{3.30}$$

is the d_0-th, $d_0 \in (0, 2N_s - 1)$, diagonal element of $\mathbf{\Lambda}_{\mathrm{CM}}$. As for $\mathbf{\Lambda}_{\mathrm{CI}}$, the mathematics is more involved. The derivation is recorded in Appendix A.2 (in the end of the thesis) where the d_0-th diagonal element of $\mathbf{\Lambda}_{\mathrm{CI}}$ was shown to be

$$\Lambda_{\mathrm{CI}}^{[d_0]} = \frac{k_0}{2}\left(\left(m_4 - \sigma_A^4\right) \mathrm{E}\left\|\vec{\gamma}(d_0)\right\|^2 + \Omega_o \right) \tag{3.31}$$

where

$$\vec{\gamma}(d_0) = \begin{cases} \vec{\alpha} & \text{if } d_0 \in (0, N_s - 1) \\ \vec{\beta} & \text{if } d_0 \in (N_s, 2N_s - 1), \end{cases} \tag{3.32}$$

$\|\cdot\|$ is the Euclidean norm operator and

$$\Omega_o = \sigma_A^4 \left(2\delta_{\mathrm{x}} + 1\right)\left(\mathrm{E}\left\|\vec{\alpha}\right\|^2 + \mathrm{E}\left\|\vec{\beta}\right\|^2\right). \tag{3.33}$$

At this point, due to the complicated nature of equations (3.29)-(3.33) we categorize the stationary points (where Eq. 3.28 equals to zero) under the following four cases where they could be found:

Case 1: $S_2^{[d_0]} = 0$, for all $d_0 \in (0, 2N_s - 1)$. This *undesirable* case implies a zero-valued output at the second equalizer since the global impulse response does not have any non-zero taps. $\therefore \|\mathbf{S}_2\|^2 = 0$.

Case 2: $S_2^{[d_0]}|_{d_0 \neq b} = 0$, $\forall(d_0 \neq b)$ where $0 \leq b \leq N_s - 1$. This *undesirable* case implies the successful retrieval of the *first* source when the intended source is the second. Since only one tap

is non-zero, $\therefore \|\mathbf{S}_2\|^2 = \left(S_2^{[b]}\right)^2$, where $\left(S_2^{[b]}\right)^2$ is found by solving $4S_2^{[b]}(\Lambda_{\mathrm{CM}}^{[b]} + \Lambda_{\mathrm{CI}}^{[b]}) = 0$, i.e.,

$$(S_2^{[b]})^2\ m_4 - 3(S_2^{[b]})^2\ \sigma_A^4 + 3\sigma_A^4\|\mathbf{S}_2\|^2 - \hat{\mathrm{R}}_2\sigma_A^2 + \frac{k_0}{2}\Big(\left(m_4 - \sigma_A^4\right)\mathrm{E}\{\|\vec{\alpha}\|^2\} + \Omega_o\Big) = 0. \tag{3.34}$$

Since $\|\mathbf{S}_2\|^2 = (S^{[b]})^2$, so after some rearrangement, (3.34) becomes

$$(S_2^{[b]})^2 = \frac{\hat{\mathrm{R}}_2\sigma_A^2}{m_4} - \frac{k_0}{2m_4}\Big(\left(m_4 - \sigma_A^4\right)\mathrm{E}\left\{\|\vec{\alpha}\|^2\right\} + \Omega_o\Big). \tag{3.35}$$

Case 3: $S_2^{[d_0]}|_{d_0 \neq c} = 0$, $\forall(d_0 \neq c)$ where $N_s \leq c \leq 2N_s - 1$. This is the *desirable* case where the *second* source is successfully retrieved with all its ISI removed. Since only one tap is non-zero, $\therefore \|\mathbf{S}_2\|^2 = \left(S_2^{[c]}\right)^2$, where $\left(S_2^{[c]}\right)^2$ is found by solving $4S_2^{[c]}(\Lambda_{\mathrm{CM}}^{[c]} + \Lambda_{\mathrm{CI}}^{[c]}) = 0$ as in Case 2, so we get

$$(S_2^{[c]})^2 = \frac{\hat{\mathrm{R}}_2\sigma_A^2}{m_4} - \frac{k_0}{2m_4}\Big(\left(m_4 - \sigma_A^4\right)\mathrm{E}\{\|\vec{\beta}\|^2\} + \Omega_o\Big). \tag{3.36}$$

Case 4: $S_2^{[d_0]}|_{d_0 \neq e} = 0$, $\forall d_0$ and $e = \{e_1, e_2, \cdots, e_v\} \in (0, 2N_s - 1)$, where $1 < v \leq 2N_s$. This is another *undesirable* case which covers all other cases that are not covered above, where $v > 1$ non-zero elements exist in $\mathbf{S}_2$ at steady state. Since more than one taps are non-zero, the exact value of $\|\mathbf{S}_2\|^2$ is not known. The stationary points are given as

$$(S_2^{[e]})^2 = \frac{1}{m_4 - 3\sigma_A^4}\left(\hat{\mathrm{R}}_2\sigma_A^2 - 3\sigma_A^4\|\mathbf{S}_2\|^2 - \frac{k_0}{2}\left((m_4 - \sigma_A^4)\mathrm{E}\left\{\|\vec{\gamma}(e)\|^2\right\} + \Omega_o\right)\right). \tag{3.37}$$

3.3.4 Design of new dispersion constant, $\hat{\mathrm{R}}_j$

The main purpose of the dispersion constant is to ensure the energy of the transmit sequence is preserved at the equalizer output [21]. To do so, we examine Case 3, which corresponds to the desirable extraction of the correct source. Using (3.36) and assuming ideal conditions where $\mathrm{E}\{\|\vec{\beta}\|^2\} = 0$ and $\mathrm{E}\{\|\vec{\alpha}\|^2\} = 1$, which are similar conditions considered in [21] for the SISO case, (3.36) becomes

$$(S_2^{[c]})^2 = \frac{\hat{\mathrm{R}}_2\sigma_A^2 - \frac{1}{2}k_0\sigma_A^4\left(2\delta_{\mathrm{x}} + 1\right)}{m_4} \tag{3.38}$$

In order to avoid any scaling in the equalizer output, we choose $\hat{\mathrm{R}}_2$ so that $(S_2^{[c]})^2 = 1$. Therefore, the modified dispersion constant is given as

$$\hat{\mathrm{R}}_2 = \frac{m_4}{\sigma_A^2} + \frac{1}{2}\, k_0 \sigma_A^2 \left(2\delta_{\mathrm{x}} + 1\right). \tag{3.39}$$

This result is easily generalized for an arbitrary j-th equalizer as follows:

$$\hat{\mathrm{R}}_j = \frac{m_4}{\sigma_A^2} + \frac{j-1}{2}\, k_0 \sigma_A^2 \left(2\delta_{\mathrm{x}} + 1\right). \tag{3.40}$$

Note the first part of the R.H.S. of (3.40), i.e., m_4/σ_A^2 is the original dispersion constant [21], which is also used in previous works. If the second part of the R.H.S. of (3.40) is missing, we have $(S_2^{[c]})^2 < 1$ from (3.38). The value that is less than unity is the reason behind the shrinking of the equalizer output that has been observed in previous related works. In fact, it is the second part of the R.H.S. of (3.40) that is responsible for overcoming the adverse scaling effect of mixed-cost algorithms, introduced by Ω_o of (3.33). More importantly, in fulfilling this unity-gain condition, the desirable stationary points of Case 3 are "locked" to the location that is identical to that of the Godard criterion for the single user case, irrespective of k_0, j and δ_x, i.e., the mixing parameter, number of users and maximum delay. Historically, this feature is not available in the CC-CMA as the global minima are expected to shift towards the origin with increasing k_0, number of users and delay spread, as shown by the contour plots of Fig. 4 in [43]. In addition, the location of the desired minima is unaffected by additive noise, unlike Fig. 6 in [43].

3.3.5 Stability analysis

In order to determine whether or not the stationary point is stable, we perform the second derivative test which evaluates the Hessian matrix, Ψ which is generally defined as

$$\Psi = \frac{\partial}{\partial \mathbf{S}_2}\left(\frac{\partial J_2[k]}{\partial \mathbf{S}_2}\right)^T = \begin{bmatrix} \frac{\partial^2 J_2[k]}{\partial S_2^{[0]} \partial S_2^{[0]}} & \frac{\partial^2 J_2[k]}{\partial S_2^{[0]} \partial S_2^{[1]}} & \cdots & \cdots & \frac{\partial^2 J_2[k]}{\partial S_2^{[0]} \partial S_2^{[2N_s-1]}} \\ \frac{\partial^2 J_2[k]}{\partial S_2^{[1]} \partial S_2^{[0]}} & \frac{\partial^2 J_2[k]}{\partial S_2^{[1]} \partial S_2^{[1]}} & \cdots & \cdots & \frac{\partial^2 J_2[k]}{\partial S_2^{[1]} \partial S_2^{[2N_s-1]}} \\ \frac{\partial^2 J_2[k]}{\partial S_2^{[2]} \partial S_2^{[0]}} & \cdots & \frac{\partial^2 J_2[k]}{\partial S_2^{[3]} \partial S_2^{[3]}} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 J_2[k]}{\partial S_2^{[2N_s-1]} \partial S_2^{[0]}} & \cdots & \cdots & \cdots & \frac{\partial^2 J_2[k]}{\partial S_2^{[2N_s-1]} \partial S_2^{[2N_s-1]}} \end{bmatrix} \tag{3.41}$$

under the following four cases where they could be found:

Case 1: $S_2^{[d_0]} = 0$, for all $d_0 \in (0, 2N_s - 1)$. The Hessian of (3.41) can be shown to be

$$\Psi = \text{diag}(\underbrace{\psi_0, \cdots, \psi_0}_{N_s}, \underbrace{\psi_1, \cdots, \psi_1}_{N_s})$$

where $\text{diag}(\cdots)$ denotes diagonal matrix and

$$\psi_0 = \frac{\partial^2 J_2[k]}{\partial S_2^{[d_1]} \partial S_2^{[d_1]}} = -4\hat{\text{R}}_2 \sigma_A^2 + 2k_0\Big(\left(m_4 - \sigma_A^4\right) \text{E}\{\|\vec{\alpha}\|^2\} + \Omega_o \Big) \tag{3.42}$$

for $0 \leq d_1 \leq N_s - 1$, and

$$\psi_1 = \frac{\partial^2 J_2[k]}{\partial S_2^{[d_2]} \partial S_2^{[d_2]}} = -4\hat{\text{R}}_2 \sigma_A^2 + 2k_0\Big(\left(m_4 - \sigma_A^4\right) \text{E}\{\|\vec{\beta}\|^2\} + \Omega_o \Big) \tag{3.43}$$

for $N_s \leq d_2 \leq 2N_s - 1$.

Case 2: $S_2^{[d_0]}|_{d_0 \neq b} = 0$, $\forall (d_0 \neq b)$ where $0 \leq b \leq N_s - 1$. In this case, the Hessian has a special diagonal element at position b:

$$\Psi = \text{diag}(\psi_0, \cdots, \psi_0, \underbrace{\psi_b}_{\uparrow}, \psi_0, \cdots, \psi_0, \psi_1, \cdots, \psi_1)$$

where

$$\begin{aligned} \psi_b &= \frac{\partial^2 J_2[k]}{\partial S_2^{[b]} \partial S_2^{[b]}} \\ &= 4\Big\{ 3(S^{[b]})^2 m_4 - \hat{\text{R}}_2 \sigma_A^2 + \frac{k_0}{2}\Big(\left(m_4 - \sigma_A^4\right) \text{E}\left\{\|\vec{\alpha}\|^2\right\} + \Omega_o \Big) \Big\}, \end{aligned} \tag{3.44}$$

and

$$\begin{aligned}\psi_0 &= \frac{\partial^2 J_2[k]}{\partial S_2^{[d_1]} \partial S_2^{[d_1]}} \\ &= 4\Big\{3(S^{[b]})^2\sigma_A^4 - \hat{\mathrm{R}}_2\sigma_A^2 + \frac{k_0}{2}\Big(\left(m_4 - \sigma_A^4\right) \mathrm{E}\left\{\|\vec{\alpha}\|^2\right\} + \Omega_o\Big)\Big\} \end{aligned} \tag{3.45}$$

for $d_1 \neq b$, $0 \leq d_1 \leq N_s - 1$, and

$$\begin{aligned}\psi_1 &= \frac{\partial^2 J_2[k]}{\partial S_2^{[d_2]} \partial S_2^{[d_2]}} \\ &= 4\Big\{(3(S^{[b]})^2\sigma_A^4 - \hat{\mathrm{R}}_2\sigma_A^2 + \frac{k_0}{2}\Big(\left(m_4 - \sigma_A^4\right) \mathrm{E}\{\|\vec{\beta}\|^2\} + \Omega_o\Big)\Big\} \end{aligned} \tag{3.46}$$

for $N_s \leq d_2 \leq 2N_s - 1$.

Case 3: $S_2^{[d_0]}|_{d_0 \neq c} = 0$, $\forall (d_0 \neq c)$ where $N_s \leq c \leq 2N_s - 1$. The Hessian becomes

$$\Psi = \mathrm{diag}(\psi_0, \cdots, \psi_0, \psi_1, \cdots, \psi_1, \underbrace{\psi_c}_{\uparrow}, \psi_1, \cdots, \psi_1)$$

where

$$\psi_c = \frac{\partial^2 J_2[k]}{\partial S_2^{[c]} \partial S_2^{[c]}} = 16\hat{\mathrm{R}}_2\sigma_A^2 - 8k_0\left(\left(m_4 - \sigma_A^4\right) \mathrm{E}\{\|\vec{\beta}\|^2\} + \Omega_o\right), \tag{3.47}$$

and $\psi_0 = \frac{\partial^2 J_2[k]}{\partial S_2^{[d_1]} \partial S_2^{[d_1]}}$

$$\begin{aligned}\psi_0 &= \frac{\partial^2 J_2[k]}{\partial S_2^{[d_1]} \partial S_2^{[d_1]}} \\ &= 4\left(\xi - 1\right)\hat{\mathrm{R}}_2\sigma_A^2 - 2k_0\Big(\left(\xi - 1\right)\Omega_o + \left(m_4 - \sigma_A^4\right)\left(\xi \mathrm{E}\{\|\vec{\beta}\|^2\} - \mathrm{E}\left\{\|\vec{\alpha}\|^2\right\}\right)\Big) \end{aligned} \tag{3.48}$$

for $0 \leq d_1 \leq N_s - 1$, where ξ is a constant and is defined as$\xi = \frac{3\sigma_A^4}{m_4}$, and

$$\begin{aligned}\psi_1 &= \frac{\partial^2 J_2[k]}{\partial S_2^{[d_2]} \partial S_2^{[d_2]}} \\ &= 4\left(\xi - 1\right)\hat{\mathrm{R}}_2\sigma_A^2 - 2k_0\Big(\left(\xi - 1\right)\Omega_o + \left(m_4 - \sigma_A^4\right)\left(\xi - 1\right)\mathrm{E}\{\|\vec{\beta}\|^2\}\Big) \end{aligned} \tag{3.49}$$

for $d_2 \neq c$, $N_s \leq d_2 \leq 2N_s - 1$.

Case 4: $S_2^{[d_0]}|_{d_0 \neq e} = 0$, $\forall d_0$ and $e = \{e_1, e_2, \cdots, e_v\} \in (0, 2N_s - 1)$, where $1 < v \leq 2N_s$. The resulting Hessian is a sparse symmetric matrix, where non-zero elements are found not only on the

diagonal but also at the (i,j) and (j,i) positions of the matrix, where $i \neq j$ and $i,j \in e$. Unlike cases 1-3, the exact value of $\|\mathbf{S}_2\|^2$ is not known here. Using the same approach as in Cases 2 and 3, the stationary point is obtained as

$$(S_2^{[e]})^2 = \frac{1}{m_4 - 3\sigma_A^4}\left(\hat{\mathrm{R}}_2\sigma_A^2 - 3\sigma_A^4\|\mathbf{S}_2\|^2 - \frac{k_0}{2}\big((m_4 - \sigma_A^4)\mathrm{E}\left\{\|\vec{\gamma}(e)\|^2\right\} + \Omega_o\big)\right). \tag{3.50}$$

Therefore the Hessian becomes

$$\Psi = \mathrm{diag}(\cdots, \psi, \psi_{e_1}, \psi, \cdots, \psi, \psi_{e_2}, \psi, \cdots, \psi, \psi_{e_v}, \psi, \cdots) + \sum_{i=e_1}^{e_v}\sum_{\substack{j=e_1 \\ j\neq i}}^{e_v} \Upsilon(i,j) \tag{3.51}$$

where $\Upsilon(i,j)$ is a $2N_s \times 2N_s$ square matrix where all its elements are zero except for $\frac{\partial^2 J_2[k]}{\partial S_2^{[i]}\partial S_2^{[j]}}$ at the i-th row and j-th column. It can be shown that the non-diagonal elements are

$$\frac{\partial^2 J_2[k]}{\partial S_2^{[e_a]}\partial S_2^{[e_b]}} = \frac{\partial^2 J_2[k]}{\partial S_2^{[e_b]}\partial S_2^{[e_a]}} = 24\sigma_A^4 S_2^{[e_a]} S_2^{[e_b]} \tag{3.52}$$

where $S_2^{[e_a]}$ and $S_2^{[e_b]}$ are two arbitrary stationary points in $\{e\}$. Therefore the Hessian is non-diagonal and symmetrical. As for its diagonal elements,

$$\begin{aligned}
\psi_e &= \frac{\partial^2 J_2[k]}{\partial S_2^{[e]}\partial S_2^{[e]}} \\
&= 12(S_2^{[e]})^2(m_4 - \sigma_A^4) + 12\sigma_A^4\|\mathbf{S}_2\|^2 - 4\hat{\mathrm{R}}_2\sigma_A^2 + 2k_0\Big(\left(m_4 - \sigma_A^4\right)\mathrm{E}\left\{\|\vec{\gamma}(e)\|^2\right\} + \Omega_o\Big) \\
&= 8(S_2^{[e]})^2 m_4.
\end{aligned} \tag{3.53}$$

3.3.6 Design of mixing parameter, k_0

The objective of this subsection is to search for and determine a range of values for k_0 that admits the stationary points of Case 3 to be the only ones that are minimum points, while those of Cases 1,2 and 4 are non-minimum. If this is possible, the CI-CMA will become globally convergent to the desired minima where the desired source is retrieved accurately and without repetition. In this subsection, we will perform the positive definite test in order to achieve our objective.

Case 1: Referring to equations (3.42) and (3.43), the Hessian matrix is positive definite if and only if $\psi_0 > 0$ and $\psi_1 > 0$. However since $\mathrm{E}\{\|\vec{\alpha}\|^2\} > \mathrm{E}\{\|\vec{\beta}\|^2\}$, we have $\psi_0 > \psi_1 > 0$. A necessary

and sufficient condition for the stationary point to be unstable is for the opposite to be true, i.e., $\psi_1 \leq 0$. Therefore after substituting $\hat{\mathrm{R}}_2$ of (3.40) into (3.43) we arrive at the inequality (3.54) at below.

$$\frac{m_4}{(m_4 - \sigma_A^4)\,\mathrm{E}\{\|\vec{\beta}\|^2\} + \sigma_A^4\,(2\delta_{\mathrm{x}} + 1)\left(\mathrm{E}\{\|\vec{\alpha}\|^2\} + \mathrm{E}\{\|\vec{\beta}\|^2\} - 1\right)} \geq \frac{k_0}{2} \tag{3.54}$$

Case 2: Referring to equations (3.44), (3.45) and (3.46), it is easy to see that the Hessian matrix is positive definite if and only if $\psi_b > 0$, $\psi_0 > 0$ and $\psi_1 > 0$. Thanks to the assumptions of $m_4 \geq \sigma_A^4$ and $\mathrm{E}\{\|\vec{\alpha}\|^2\} > \mathrm{E}\{\|\vec{\beta}\|^2\}$, we have $\psi_b \geq \psi_0 > \psi_1 > 0$. and it is sufficient to show that if $\psi_1 > 0$ is true, the Hessian matrix is a positive definite matrix. Since this stationary point is undesirable, the Hessian should be non-positive definite, i.e., $\psi_1 \leq 0$. By substituting (3.40) and (3.35) into (3.46), we arrive at the inequality (3.55), where $\xi = \frac{3\sigma_A^4}{m_4}$.

$$\frac{3\sigma_A^4 - m_4}{(m_4 - \sigma_A^4)\left(\xi\mathrm{E}\{\|\vec{\alpha}\|^2\} - \mathrm{E}\{\|\vec{\beta}\|^2\}\right) + (\xi - 1)\sigma_A^4\,(2\delta_{\mathrm{x}} + 1)\left(\mathrm{E}\{\|\vec{\alpha}\|^2\} + \mathrm{E}\{\|\vec{\beta}\|^2\} - 1\right)} \leq \frac{k_0}{2} \tag{3.55}$$

Case 3: In this desirable case, we wish to find the range of k_0 for which the Hessian matrix is a positive definite matrix. It is necessary and sufficient to satisfy $\psi_c > 0$, $\psi_0 > 0$ and $\psi_1 > 0$. Due the the assumption that $\mathrm{E}\{\|\vec{\alpha}\|^2\} > \mathrm{E}\{\|\vec{\beta}\|^2\}$, we have $\psi_0 > \psi_1 > 0$. Therefore, we only determine $\psi_c > 0$ and $\psi_1 > 0$. The stationary point has been derived in (3.36). Substituting the expressions of $S_2^{[\mathrm{c}]}$ and $\hat{\mathrm{R}}_2$ from (3.36) and (3.39) into (3.47) and (3.49), we found that both inequalities $\psi_c > 0$ and $\psi_1 > 0$ are identical. The resulting inequality is shown in (3.56).

$$\frac{m_4}{(m_4 - \sigma_A^4)\,\mathrm{E}\{\|\vec{\beta}\|^2\} + \sigma_A^4\,(2\delta_{\mathrm{x}} + 1)\left(\mathrm{E}\{\|\vec{\alpha}\|^2\} + \mathrm{E}\{\|\vec{\beta}\|^2\} - 1\right)} > \frac{k_0}{2} \tag{3.56}$$

Case 4: Since this is an undesirable stationary point we wish to perform the following non-positive definite test (p. 128 [130]) on the Hessian, so that if the following test fails, it is conclusive that this stationary point is not a minimum point:

$$\left(\frac{\partial^2 J_2[k]}{\partial S_2^{[e_a]} \partial S_2^{[e_b]}}\right)^2 < \frac{\partial^2 J_2[k]}{\partial S_2^{[e_a]} \partial S_2^{[e_a]}} \cdot \frac{\partial^2 J_2[k]}{\partial S_2^{[e_b]} \partial S_2^{[e_b]}} \tag{3.57}$$

Using (3.52) and (3.53), the above test simplifies to $3\sigma_A^4 < m_4$, which cannot be true since the source is circular and sub-Gaussian (see Assumption A1). Therefore the test fails and it is concluded that the stationary point is unstable.

Finally, the inequalities (3.54), (3.55) and (3.56) are combined to yield the range of k_0 values for which the desired global convergence of the CI-CMA, i.e., blind equalization and source separation can be achieved:

$$k_0 \geq \frac{2m_4\left(3\sigma_A^4 - m_4\right)}{3\sigma_A^4\left(m_4 - \sigma_A^4\right)}. \tag{3.58}$$

Because the denominator is zero when $m_4 = \sigma_A^4$, there is no finite value of k_0 that can satisfy the inequality when constant modulus modulation schemes such as the 2-PAM are used.

3.4 Extension to complex modulations

In this section, the results and analysis of Section 3.3 are extended to complex modulations, such as QAM and PSK. The approach is fairly straight forward with much of its analysis identical to that of the real-valued case. We must only pay attention to the different statistical properties of the complex-valued signals, where some values and notations may be different. In particular, we emphasize that in complex modulations, $\mathrm{E}\{a_j^2[k]\}$ and $\mathrm{E}\{a_j^4[k]\}$ are zero-valued due to the circular symmetry of the i.i.d. sources, while $\sigma_A^2 = \mathrm{E}\{|a_j[k]|^2\}$ and $m_4 = \mathrm{E}\{|a_j[k]|^4\}$ are the variance and the fourth order moment of the source, respectively, which have finite positive values in practice.

It can be shown that the gradient of the mixed cost is

$$\frac{\partial J_2[k]}{\partial \mathbf{S}_2} = 2\mathbf{\Lambda}\mathbf{S}_2^* = 2(\mathbf{\Lambda}_{\mathrm{CM}} + \mathbf{\Lambda}_{\mathrm{CI}})\mathbf{S}_2^* \tag{3.59}$$

where $(\cdot)^*$ denotes complex conjugate, while $\mathbf{\Lambda}$, $\mathbf{\Lambda}_{\mathrm{CM}}$ and $\mathbf{\Lambda}_{\mathrm{CI}}$ are all $2N_s \times 2N_s$ diagonal matrices, and $\mathbf{\Lambda}_{\mathrm{CM}}$ can be derived by extending the result of Johnson [121] to complex modulations so that $\Lambda_{\mathrm{CM}}^{[d_0]}$, the d_0-th diagonal element of $\mathbf{\Lambda}_{\mathrm{CM}}$, is expressed as

$$\Lambda_{\mathrm{CM}}^{[d_0]} = \left|S_2^{[d_0]}\right|^2 m_4 - 2\left|S_2^{[d_0]}\right|^2 \sigma_A^4 + 2\sigma_A^4\|\mathbf{S}_2\|^2 - \hat{\mathrm{R}}_2\sigma_A^2 \tag{3.60}$$

where $d_0 \in (0, 2N_s - 1)$. The d_0-th diagonal element of $\mathbf{\Lambda}_{\mathrm{CI}}$ is identical to that of the real-valued case in (3.31)-(3.33).

The Hessian for complex modulations is defined as $\Psi = \frac{\partial}{\partial \mathbf{S}_2^*} \left(\frac{\partial J_2[k]}{\partial \mathbf{S}_2} \right)^T$, where $\frac{\partial}{\partial S_2^{*[i]}} \left(\frac{\partial J_2[k]}{\partial S_2^{[j]}} \right)$ refers to the element in the i-th row and j-th column of Ψ. The subsequent work is identical to that of the real case, so we simply state the expression of the dispersion constant and the range of k_0 values, derived from the positive definite test which evaluated the Hessian at stationary points of four cases admitting only the desired solution (Case 3) while discarding the others.

The dispersion constant, $\hat{\mathrm{R}}_j$, is selected in the same way as (3.40) to avoid undesirable scaling, with $m_4 = \mathrm{E}\{|a_j[k]|^4\}$. As for the mixing parameter, the range of k_0 values for which source separation and equalization can be *guaranteed* is

$$k_0 \geq \frac{m_4 \left(2\sigma_A^4 - m_4\right)}{\sigma_A^4 \left(m_4 - \sigma_A^4\right)} \tag{3.61}$$

where we have assumed $\mathrm{E}\{\|\vec{\alpha}\|^2\} = 1$ and $\mathrm{E}\{\|\vec{\beta}\|^2\} = 0$. Similar to the real-valued case, no finite value of k_0 can satisfy the inequality when constant modulus modulation schemes such as the m-ary PSK are used.

3.5 Simulations

The simulations are performed in order to identify the open eye and source separation abilities of the new algorithm CI-CMA for 2-PAM, 4-PAM, 4-QAM and 16-QAM. Furthermore, the shrinking of equalizer outputs will be identified in the simulations as well.

3.5.1 Simulation Setup

In order to simulate the system model in Section 3.2.1 and Fig. 3.1 with ($M_{\mathrm{t}} = 2$ and $M_{\mathrm{r}} = 3$) MIMO channel, simulations are performed using MATLAB(Matrix Laboratory) version 8.0 . Two random 2-PAM symbol sequences that have N_{seq} symbols in one sequence are generated as source 1, $a_1[k]$ and source 2, $a_2[k]$. The sources are convoluted with the memory channels $\mathbf{h}_{ij}$, and this process is equivalent to $\mathbf{h}_{ij} \otimes \mathbf{a}_j[k]$ where $j = 1, 2$. This is aimed to simulate the condition that the sources are passed through the memory channels and thereby the sources are interfered by ISI. At the i-th received sensor, the ISI signals from two sources $a_1[k]$ and $a_2[k]$ (but not a single source) have been observed, thus the observed signal at i-th sensor $x_i[k]$ is corrupted by CCI due to receiving two sources and ISI due to memory channels. This scenario can be simulated through computing Eq. 3.1 and the impulse responses of channels $\mathbf{h}_{ij}$ are stated in Table 3.1.

Next, to simulate the algorithm in the equalizer, we will test using the hierarchical CI-CMA, i.e., the first equalizer is allowed to recover any of the two sources (usually the stronger one) and the second equalizer should be able to avoid the extracted source by the first equalizer if BSS works. The filter weights $\mathbf{f}_{11}[0]$ and $\mathbf{f}_{22}[0]$ are initialized with a center spike [36] whilst $\mathbf{f}_{ij}[0]$, $\forall i \neq j$, are initialized with zero-vectors. After the initialization, the observed signals $x_i[k]$ are fed through the filters $\mathbf{F}_j[k]$, which is the concatenation vector by several filters $\mathbf{f}_{ij}[k]$, and generated the two outputs $y_j[k]$ where $j = 1, 2$ because two sources were transmitted. In this case, this process can be simulated by executing Eq. 3.3 and the filter values $\mathbf{F}_j[k]$ are iteratively updated for every time k through computing the filter update equation in Eq. 3.10 which has the error term of CI-CMA in Eq. 3.22. In order to ensure that different sources are not extracted by both equalizers by chance,

at mid way through the simulation we change the taps of the first equalizer to $\mathbf{f}_{11}[k_s] = \mathbf{0}$ and $\mathbf{f}_{21}[k_s] = g[0, \cdots, 0, 1, 0, \cdots, 0]$, $g > 1$, and k_s denotes the time of half-way point of the simulation. This new setting induces the filter of first equalizer ($\mathbf{F}_1[k]$) to recover the second source, $a_2[k]$ using a larger-than-unity center spike g, where we have used $g = 2$ in all our simulations. We emphasize that the filter of second equalizer ($\mathbf{F}_2[k]$) is unchanged in the above setting. While the first equalizer turns to extract the other source, we observe the way the second equalizer reacts to this switch in order to validate the source separation ability.

The above simulation setup is repeated for 4-PAM, 4-QAM and 16-QAM cases. Furthermore, for the cases of 4-QAM and 16-QAM, the complex number MIMO channel in Table 3.2 is used to replace Table 3.1.

Table 3.1: Impulse response of a two-input/three-output real channel [1]

$\mathbf{h_{11}}$	$[0.9844, 0.3365, 0.1002]^T$
$\mathbf{h_{12}}$	$[0.6303, 0.1575, 0.3791]^T$
$\mathbf{h_{21}}$	$[0.5232, 0.0396, 0.0121]^T$
$\mathbf{h_{22}}$	$[0.21, 0.2855, 0.6461]^T$
$\mathbf{h_{31}}$	$[0.9651, 0.1543, 0.5371]^T$
$\mathbf{h_{32}}$	$[0.2547, 0.2695, 0.7178]^T$

Table 3.2: Impulse response of a two-input/three-output complex channel [1]

$\mathbf{h_{11}}$	$[-0.6 + 0.4j, 1.2 - 0.2j]^T$
$\mathbf{h_{12}}$	$[-0.6 + 0.8j, 0.9 - 0.1j]^T$
$\mathbf{h_{21}}$	$[0.1 + 0.7j, -0.2 - 0.5j]^T$
$\mathbf{h_{22}}$	$[0.4 - 0.3j, -0.2 + 0.2j]^T$
$\mathbf{h_{31}}$	$[0.5 + 0.4j, -1 + 0.3j]^T$
$\mathbf{h_{32}}$	$[-0.1 + 0.8j, 0.4 + 0.1j]^T$

3.5.2 Simulation Parameters

In order to perform the simulations, the simulation parameters are required to be defined for the 2-PAM, 4-PAM, 4-QAM and 16-QAM.

First, 2-PAM has the constellation set $\mathbf{A} = \{-1, 1\}$ and this constellation set has defined statistical properties such as fourth order moment $m_4 = \mathrm{E}\left\{|a[k]|^4\right\} = 1$ and variance $\sigma_A^2 = 1$. In other words, the values of m_4 and σ_A^2 are depended on the constellation set $\mathbf{A} = \{-1, 1\}$ and cannot be arbitrary chosen. Next, the step sizes of the equalizers are assigned to be $\mu_1 = 0.001$ and $\mu_2 = 0.0001$ for equalizer 1 and 2, respectively. Small value in the step sizes are typically chosen, but it can cause the convergence becomes slow. On the contrary, large value in the step sizes can increase the convergence rate but it can cause the equalizer diverges and unstable. Therefore, to avoid unstable and slow convergence, the above values of step sizes are empirically found after several simulation attempts had been done. μ_2 is smaller than μ_1 because the equalizer 2 uses CI-CMA whilst equalizer 1 uses CMA. Due to CI-CMA can produce larger error than CMA, μ_2 is always chosen to be smaller than μ_1 to avoid divergence. The maximum delay is set at $\delta_{\mathrm{x}} = 9$ and the lengths of filter are $L = 7$. Ideally, δ_{x} and L that are always larger than the channel length are theoretically required. However, because we assume the channel is unknown, therefore, these value are arbitrary chosen. Large values in δ_{x} and L is helpful to reduce the error, but it increases the memory cost. Since $L = 7$ is chosen, the center tap strategy initializes $f_{11}^{[(L-1)/2]}[0] = f_{22}^{[(L-1)/2]}[0] = 1$ and all other taps to zero. Furthermore, based on the analysis in the previous section, it is clear that no suitable value of $\frac{k_0}{2}$ can satisfy Eq. 3.58 since its denominator is zero due to $m_4 = \sigma_A^4$. Thus we simply choose $\frac{k_0}{2} = 2$, a value that was suggested in [2].

Instead of that, the constellation sets $\mathbf{A}$ are $\{\pm 1, \pm 3\}$, $\{\pm 1 \pm j\}$ and $\{\pm 1 \pm j, \pm 3 \pm j, \pm 1 \pm 3j, \pm 3 \pm 3j\}$ for 4-PAM, 4-QAM and 16-QAM, respectively. The reasons to select parameters m_4, σ_A^2, μ_1, μ_2, δ_{x} and L of 4-PAM, 4-QAM and 16-QAM cases are same as the reasons of 2-PAM case, and the values of the parameters are stated in Table 3.3.

We discuss the way to choose the mixing parameter. For 4-PAM case, based on Eq. 3.58, we must choose $\frac{k_0}{2} > 1.1617$ to achieve the desired global convergence. Therefore, we choose $\frac{k_0}{2} = 1.41$ for 4-PAM case. Since 4-QAM is constant modulus and $m_4 = \sigma_A^4$, no value of k_0 satisfies the condition of Eq. 3.61 and we do not expect source separation to work properly, so we arbitrary choose $\frac{k_0}{2} = 1.41$. For 16-QAM, based on Eq. 3.61, we must choose $\frac{k_0}{2} > 1.4025$ to achieve the desired global convergence. Thus we choose $\frac{k_0}{2} = 1.41$.

Finally, k_s is the time of half-way point in the simulation and this value is required to be chosen empirically. The time k_s must be sufficient large in order to ensure the equalizer filters have reached the stability points and this stage is called the steady state. There are several ways to empirically identify the steady state. One way is to observe the error in the performance measurement. If the error is sufficiently small and does not have significant changes for a long time, we can conclude that the equalizer has reached the steady state. Another way is to observe the equalizer output. If the equalizer output converges to certain points after the sufficient long time and no significant changes can be observed, we also can conclude the equalizer has reached steady state. $k_s = 200,000$ is empirically chosen after several simulation attempts have been done. Furthermore, the number of symbols in a sequence, N_{seq} is chosen to be $2k_s$.

Table 3.3: Simulation Parameters [2]

Parameters	**2-PAM**	**4-PAM**	**4-QAM**	**16-QAM**	**Description**
m_4	1	41	4	132	Forth-order moment
σ_A^2	1	5	2	10	Variance
μ_1	0.001	5×10^{-6}	1.5×10^{-5}	4×10^{-6}	Step size of equalizer 1
μ_2	0.0001	1×10^{-6}	2×10^{-6}	2×10^{-7}	Step size of equalizer 2
δ_{x}	9	11	10	10	Channel delay spread
L	7	9	9	9	Length of filter $\mathbf{f}_{ij}[k]$
$\frac{k_0}{2}$	2	1.41	1.41	1.41	Mixing parameter by 2
k_s	200,000	600,000	600,000	600,000	the time of half-way point
N_{seq}	400,000	1,200,000	1,200,000	1,200,000	Number of symbols in a sequence

3.5.3 Performance measurements

In order to measure the performance of algorithm, we compute the amount of residual interference of equalizer output 1 and 2 at time instant k, $\mathrm{IT}_1[k]$ and $\mathrm{IT}_2[k]$ as defined in Section 2.12 using MATLAB. $\mathrm{IT}_j[k]$ where $i = 1, 2$ is aimed to measure the remaining interference and thus it is a measurement of error. Therefore, the low value of $\mathrm{IT}_j[k]$ are demanded.

However, the $\mathrm{IT}_j[k]$ only implies the amount of interference but cannot validate the source separation ability. In order to validate the source separation ability, we further require the information of the global impulse response, $s_{ij}^{[n]}[k]$ (Note: we use short notation $s_{ij}^{[n]}[k]$ here, where $s_{ij}^{[n]}[k] = S_{(i),j}^{[n]}$ and $S_{(i),j}^{[n]}$ was defined in Section 2.10), for $i = 1, 2$ and $j = 1, 2$ at the steady state. However, the value of $s_{ij}^{[n]}[k]$ probably can be a complex value or negative value. Therefore, to overcome the difficulty to plot complex number or negative number of $s_{ij}^{[n]}[k]$, we are mainly interested on the absolute $|s_{ij}^{[n]}[k]|$ and this represents the absolute impulse response that connects from source i to output j. Instead of that, the value of $|s_{ij}^{[n]}[k]|$ can fluctuate from time to time due to the non-zero error. To overcome this issue, we take the 1000 samples of $|s_{ij}^{[n]}[k]|$ and display the averaged of these samples. Therefore, the averaged global impulse responses at the steady state will be plotted in order to validate the source separation ability.

Moreover, the symbols that are produced by equalizer outputs 1 and 2 (i.e. $y_1[k]$ and $y_2[k]$) are plotted. In order to identify the open eye condition and shrinking issue, the positions and the concentrations of the symbols in the constellation are observed in the graph.

3.5.4 Simulation results for 2-PAM signal

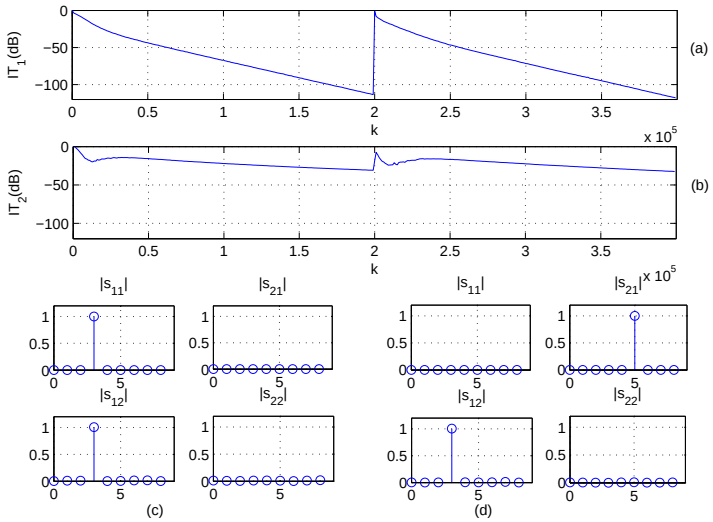

Figure 3.2: (a) Residual interference (IT) of equalizer 1. (b) IT of equalizer 2. (c) First half: the averaged global impulse responses taken from time $k = 199,000$ to time $k = 200,000$. Source separation is not successful. (d) Second half: the averaged global impulse responses taken from time $k = 399,000$ to time $k = 400,000$. Source is not changed in equalizer 2.

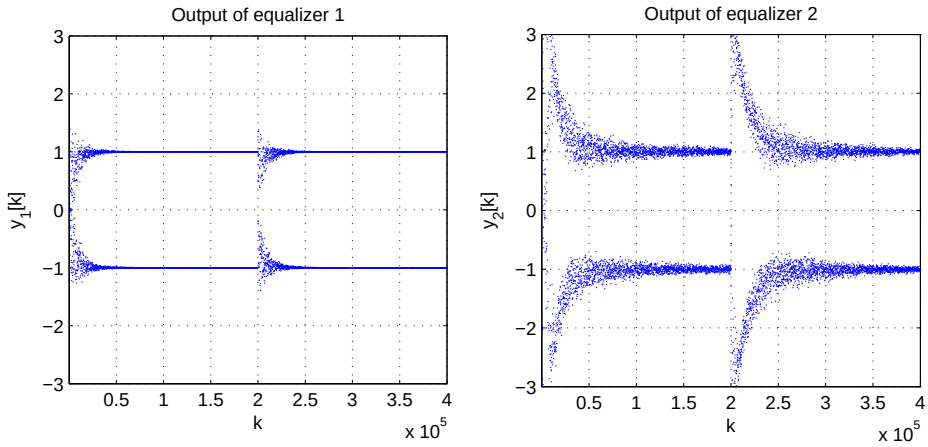

Figure 3.3: Output signals of equalizers 1 and 2.

3.5.4.1 Discussions of 2-PAM signal

Fig. 3.2 (a) and (b) illustrate the IT of equalizer 1 and 2 versus time k. In the first half (i.e., from $k = 1$ to $k = 2 \times 10^5$), the IT values are decreasing as the time k increases. This implies that CMA and CI-CMA are able to suppress the interferences of equalizer 1 and 2, respectively. Thus, the equalization task can be achieved.

Next, in order to validate the source separation ability in the first half, we can observe the averaged global impulse responses in Fig. 3.2 (c). Among the average global impulse responses $|s_{11}|$, $|s_{21}|$, $|s_{12}|$ and $|s_{22}|$, only $|s_{11}|$ and $|s_{12}|$ present the spikes. The spike in $|s_{11}|$ indicates that the equalizer 1 recovers source 1, whilst the spike in $|s_{12}|$ implies the equalizer 2 recovers source 1. Therefore, due to equalizers 1 and 2 recover the same source 1, we can ensure that the source separation ability is failed. This is expected in our analysis because no suitable value of $\frac{k_0}{2}$ can satisfy Eq. 3.58 due to $m_4 = \sigma_A^4$ in 2-PAM.

Instead of that, Fig. 3.3 illustrates the outputs of equalizer 1 and 2 versus time k. In the first half, at the beginning of time where k is near zero, the symbols of output 1 and 2 are scattered and this implies the outputs are unreliable. As the time k increasing, due to the equalizers have successfully suppressed the interferences, the outputs converge towards the points 1 and -1, and this implies the channel eye can be open. Furthermore, the converged points (1 and -1) are same as the constellation set of 2-PAM, thus we can validate that no shrinking issue has been found.

Again, in the first half of Fig. 3.3, by observing the signals at time $k = 0.5 \times 10^5$ to $k = 2 \times 10^5$ of equalizer 1 and time $k = 1 \times 10^5$ to $k = 2 \times 10^5$ of equalizer 2, the signals do not show any significant change within these time ranges. Therefore, these time ranges can be defined at the steady state time ranges of the equalizers. In the steady state, the global impulse responses within the steady state time ranges are nearly same and only present minor fluctuations. Therefore, we take the averaged global impulse responses from the time range $k = 199,000$ to time $k = 200,000$, which is within the steady state time ranges.

In the second half (i.e., time $k = 2 \times 10^5 + 1$ to time $k = 4 \times 10^5$), during the initial time (i.e.

near $k = 2 \times 10^5 + 1$), the IT values in Fig. 3.2 (a) and (b) are high and the signals in Figure 3.3 are scattered. This is because we intentionally change the taps of the first equalizer (as mentioned in the Simulation Setup) at time $k = 2 \times 10^5$. In this case, the equalizers require some time to suppress the interference. By observing Fig. 3.3, we can see that after the time $k = 2.5 \times 10^5$ of equalizer 1 and the time $k = 3 \times 10^5$ of equalizer 2, the signals of the equalizers do not show any significant change. Thus, we can conclude that the equalizers enter the steady state after the mentioned times. Therefore, we plot the averaged global impulse responses from the time range $k = 399,000$ to time $k = 400,000$ in Fig 3.2 (d). Among the four impulse responses, we can see that $|s_{21}|$ and $|s_{12}|$ have spikes. Therefore, in the second half, the equalizer 1 recovers source 2 and the equalizer 2 recovers source 1. Even though different sources are recovered, the source separation task is considered failed because the source separation task was failed in the first half test.

Next, we compare IT of equalizer 1 and 2 in Fig. 3.2 (a) and (b). We can observe that the IT of equalizer 2, which employs CI-CMA, suffers higher residual interference than IT of equalizer 1, which uses CMA. This limitation is due to the extra CI term of the CI-CMA.

3.5.5 Simulation results for 4-PAM signal

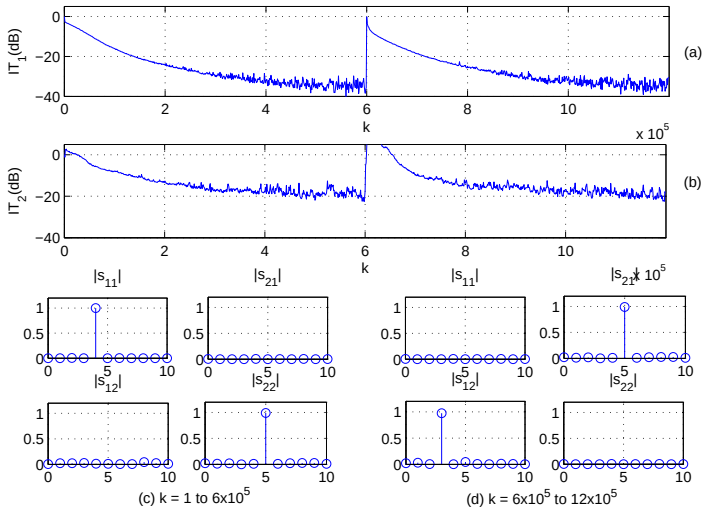

Figure 3.4: (a) Residual interference (IT) of equalizer 1. (b) IT of equalizer 2. (c) First half: the averaged global impulse responses taken from time $k = 599,000$ to time $k = 600,000$. The apparent success in source separation is not yet conclusive. (d) Second half: the averaged global impulse responses taken from time $k = 1,199,000$ to time $k = 1,200,000$. Source separation is proven successful.

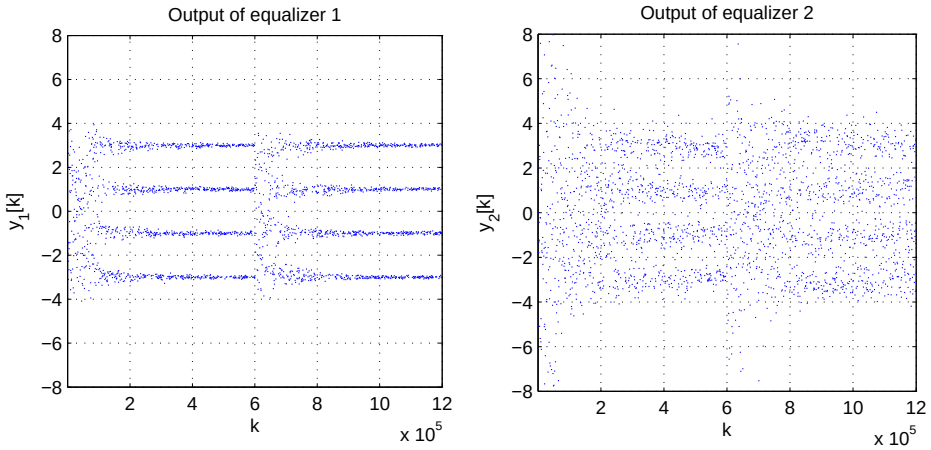

Figure 3.5: Output signals of equalizers 1 and 2.

3.5.5.1 Discussions of 4-PAM signal

Fig. 3.4 (a) and (b) show the IT of equalizer 1 and 2 versus time k. In the first half (i.e., from $k = 1$ to $k = 6 \times 10^5$), the IT values are decreasing as the time k increases. This implies that CMA and CI-CMA are able to suppress the interferences of equalizer 1 and 2, respectively. Thus, the equalization task can be achieved.

Next, in order to validate the source separation ability in the first half, we can observe the averaged global impulse responses in Fig. 3.4 (c). Among the average global impulse responses $|s_{11}|$, $|s_{21}|$, $|s_{12}|$ and $|s_{22}|$, only $|s_{11}|$ and $|s_{22}|$ present the spikes. The spike in $|s_{11}|$ indicates that the equalizer 1 recovers source 1, whilst the spike in $|s_{22}|$ implies the equalizer 2 recovers source 2. Therefore, due to equalizers 1 and 2 recover the different sources, we *temporarily* define that the source separation ability is fine, but need to be confirmed in the second half test later.

Instead of that, Fig. 3.5 illustrates the outputs of equalizer 1 and 2 versus time k. In the first half, at the beginning of time where k is near zero, the symbols of output 1 and 2 are scattered and this implies the outputs are unreliable. As the time k increasing, due to the equalizers have successfully suppressed the interferences, the outputs converge towards the points 3, 1, -1 and -3, and this implies the channel eye can be open. Furthermore, the converged points are same as the constellation set of 4-PAM, thus we can validate that no shrinking issue has been found.

In the first half of Fig. 3.5, by observing the signals at time $k = 2 \times 10^5$ to $k = 6 \times 10^5$ of equalizer 1, the signals do not show any significant change within the time range. Therefore, the time range can be defined at the steady state time range of the equalizer 1. However, this type of steady state identification is not suitable for equalizer 2 here. As we can see the equalizer output 2 in Fig. 3.5, the signals are relatively vague and fluctuating, and thus it is hard to find the steady state through such observation. As an alternative, we can identify the steady state through Fig.3.4 (b). We can see that the IT value of equalizer 2 does not show significant change from $k = 4 \times 10^5$ to $k = 6 \times 10^5$. Therefore, we define this time range as the steady state time range of equalizer 2 in the first half. As a result, we take the averaged global impulse responses from the time range

$k = 599,000$ to time $k = 600,000$, which is within the steady state time ranges above.

In the second half (i.e., time $k = 6 \times 10^5 + 1$ to time $k = 12 \times 10^5$), during the initial time (i.e. near $k = 6 \times 10^5 + 1$), the IT values in Fig. 3.4 (a) and (b) are high and the signals in Figure 3.5 are scattered. This is because we intentionally change the taps of the first equalizer (as mentioned in the Simulation Setup) at time $k = 6 \times 10^5$. In this case, the equalizers require some time to suppress the interference. By observing Fig. 3.5, we can see that after the time $k = 8 \times 10^5$ of equalizer 1, the signals of the equalizer 1 do not show any significant change. Thus, the steady state time range for Equalizer 1 in the second half is from $k = 8 \times 10^5$ to $k = 12 \times 10^5$. Through the observation on IT of equalizer 2 in Fig. 3.4 (b), we can see that the IT value does not show significant change after time $k = 10 \times 10^5$. Therefore, time range from $k = 10 \times 10^5$ to $k = 12 \times 10^5$ is defined as the steady state time range for equalizer 2 at the second half. Therefore, we plot the averaged global impulse responses from the time range $k = 1,199,000$ to time $k = 1,200,000$ in Fig 3.4 (d). Among the four impulse responses, we can see that $|s_{21}|$ and $|s_{12}|$ have spikes. Therefore, in the second half, the equalizer 1 recovers source 2 and the equalizer 2 recovers source 1. This indicates that different sources have been recovered. Due to the reason that the source separation task is achieved in the first and the second half, we can confirm that the source separation can be achieved in 4-PAM.

Next, we compare IT of equalizer 1 and 2 in Fig. 3.4 (a) and (b). Obviously, the IT of equalizer 2, which employs CI-CMA, suffers higher residual interference than IT of equalizer 1, which uses CMA. This is the limitation of the CI-CMA.

3.5.6 Simulation results for 4-QAM signal

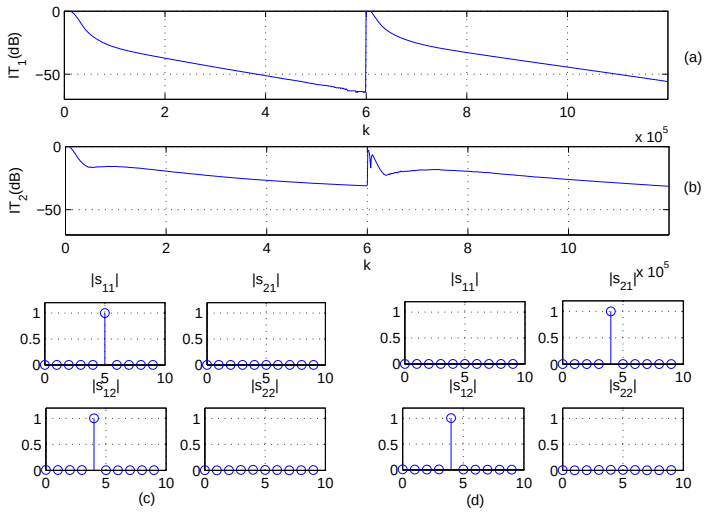

Figure 3.6: (a) Residual interference (IT) of equalizer 1. (b) IT of equalizer 2. (c) First half: the averaged global impulse responses taken from $k = 599,000$ to $k = 600,000$. Source separation is not successful. (d) Second half: the averaged global impulse responses taken from $k = 1,199,000$ to $k = 1,200,000$. Source is not changed in equalizer 2.

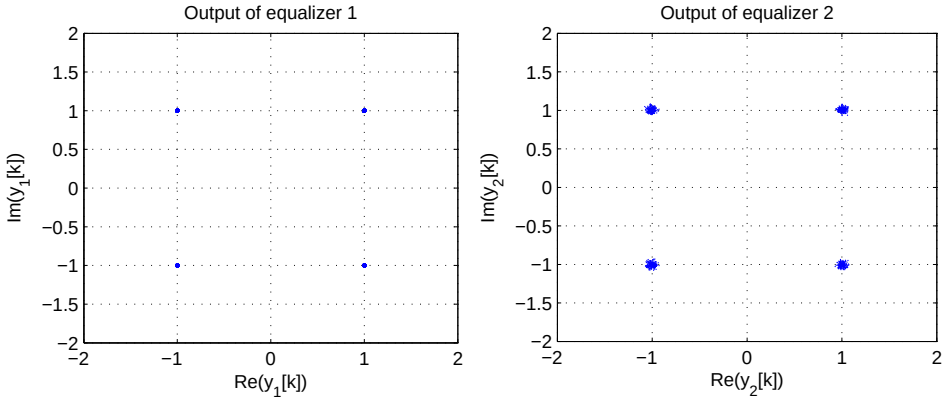

Figure 3.7: Output signals of equalizers 1 and 2 at steady state.

3.5.6.1 Discussions of 4-QAM signal

Fig. 3.6 (a) and (b) show the IT of equalizer 1 and 2 versus time k. In the first half (i.e., from $k = 1$ to $k = 6 \times 10^5$), the IT values are decreasing as the time k increases. This implies that CMA and CI-CMA are able to suppress the interferences of equalizer 1 and 2, respectively. Thus, the equalization task can be achieved.

Next, in order to validate the source separation ability in the first half, we can observe the averaged global impulse responses in Fig. 3.6 (c). Among the average global impulse responses $|s_{11}|$, $|s_{21}|$, $|s_{12}|$ and $|s_{22}|$, only $|s_{11}|$ and $|s_{12}|$ present the spikes. The spike in $|s_{11}|$ indicates that the equalizer 1 recovers source 1, whilst the spike in $|s_{12}|$ implies the equalizer 2 recovers source 1. Therefore, due to equalizers 1 and 2 recover the same source 1, the source separation ability is unsuccessful.

Moreover, Fig. 3.7 illustrates the outputs of equalizer 1 and 2, where imaginary part of outputs versus real part of outputs. Due to the symbols are complex value, the time axis is difficult to be plotted. In this case, unlike the 2-PAM and 4-PAM cases, we cannot observe the symbols change over time. However, we can see that the symbols closely match the 4-QAM constellation symbol set. Thus, the channel eye can be open and no shrinking issue can be observed.

In order to identify the time range of the steady state in the first half, we can observe IT of equalizer 1 and 2 in Fig. 3.6 (a) and (b). The IT value of -30 dB indicates that the amount of residual interference is small enough and less likely to influence the outputs. More precisely, -30 dB is equivalent to 0.001 residual interference power, whilst the 4-QAM signal has the power of 2 (i.e. $\sigma_A^2 = 2$). In this case, the signal power is 2000 times larger than the interference power, and thereby the interference is relatively small and can be ignored. Therefore, we can identify the steady state based on the IT value of -30 dB. In Fig. 3.6 (a), -30 dB corresponds to time $k = 1 \times 10^5$. Thus, the steady state time range for equalizer 1 in the first half starts from $k = 1 \times 10^5$ to $k = 6 \times 10^5$. Through the observation of Fig. 3.6 (b), we can see that the time range from $k = 4 \times 10^5$ to $k = 6 \times 10^5$ does not show any significant change. Therefore, this range is selected as the steady

state time range of equalizer 2 in the first half. As a result, we take the averaged global impulse responses from the time range $k = 599,000$ to time $k = 600,000$, which is within the steady state time ranges above.

Similarly, in the second half, we can identify the steady state time ranges for equalizer 1 and 2 as $k = 8 \times 10^5$ to $k = 12 \times 10^5$ and $k = 10 \times 10^5$ to $k = 12 \times 10^5$, respectively. Therefore, we plot the averaged global impulse responses from the time range $k = 1,199,000$ to time $k = 1,200,000$ in Fig 3.6 (d). Among the four impulse responses, we can see that $|s_{21}|$ and $|s_{12}|$ have spikes. Thus, the equalizer 1 recovers source 2 and the equalizer 2 recovers source 1. This indicates that different sources have been recovered. Due to the reason that the source separation task was failed in the first half, the source separation ability is unsuccessful even though it works for the second half.

3.5.7 Simulation results for 16-QAM signal

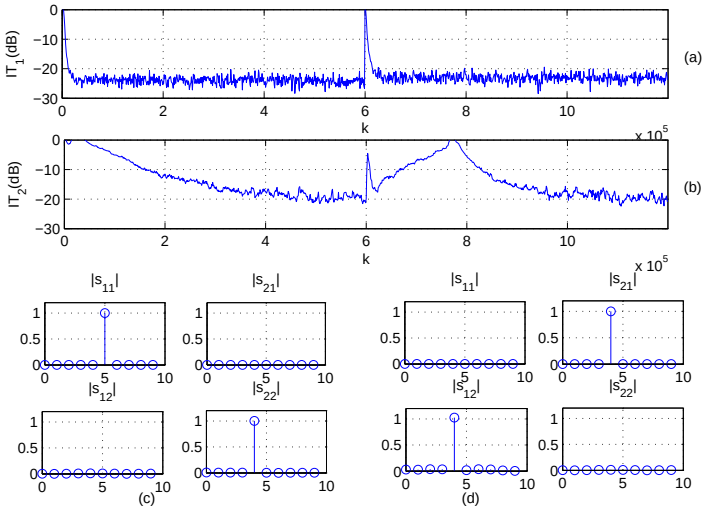

Figure 3.8: (a) Residual interference (IT) of equalizer 1. (b) IT of equalizer 2. (c) First half: the averaged global impulse responses taken from $k = 599,000$ to $k = 600,000$. The apparent success in source separation is not yet conclusive. (d) Second half: the averaged global impulse responses taken from $k = 1,199,000$ to $k = 1,200,000$. Source separation is proven successful.

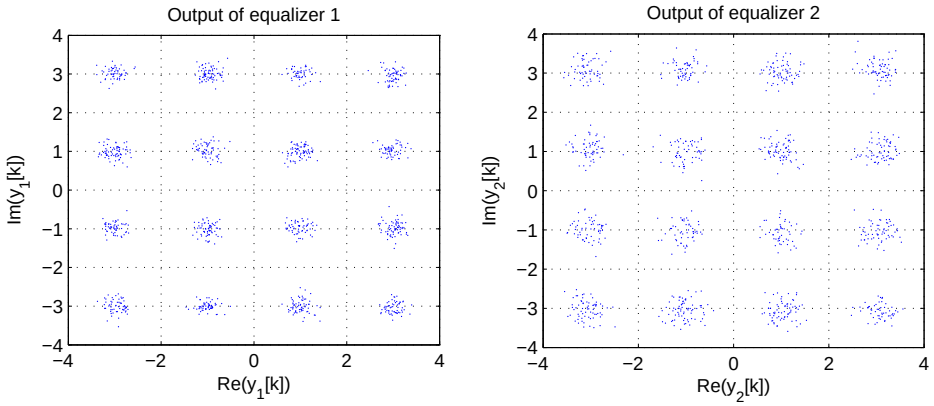

Figure 3.9: Output signals of equalizers 1 and 2 at steady state.

3.5.7.1 Discussions of 16-QAM signal

Fig. 3.8 (a) and (b) show the IT of equalizer 1 and 2 versus time k. In the first half (i.e., from $k = 1$ to $k = 6 \times 10^5$), the IT values are decreasing as the time k increases. This implies that CMA and CI-CMA are able to suppress the interferences of equalizer 1 and 2, respectively. Thus, the equalization task can be achieved.

Next, in order to validate the source separation ability in the first half, we can observe the averaged global impulse responses in Fig. 3.8 (c). Among the average global impulse responses $|s_{11}|$, $|s_{21}|$, $|s_{12}|$ and $|s_{22}|$, only $|s_{11}|$ and $|s_{22}|$ present the spikes. The spike in $|s_{11}|$ indicates that the equalizer 1 recovers source 1, whilst the spike in $|s_{22}|$ implies the equalizer 2 recovers source 1. Therefore, due to equalizers 1 and 2 recover the different sources, the source separation ability is *temporarily* successfully.

Moreover, Fig. 3.9 illustrates the outputs of equalizer 1 and 2, where imaginary part of outputs versus real part of outputs. Due to the symbols are complex value, the time axis is difficult to be plotted. In this case, unlike the 2-PAM and 4-PAM cases, we cannot observe the symbols change over time. However, we can see that the symbols closely match the 16-QAM constellation symbol set. Thus, the channel eye can be open and no shrinking issue can be observed.

In order to identify the time range of the steady state in the first half, we can observe IT of equalizer 1 and 2 in Fig. 3.8 (a) and (b). The time ranges from $k = 0.5 \times 10^5$ to $k = 6 \times 10^5$ and from $k = 4 \times 10^5$ to $k = 6 \times 10^5$ are chosen to be the steady state time ranges for output 1 and 2, respectively, because no significant change in IT values can be observed within these time ranges. As a result, we take the averaged global impulse responses from the time range $k = 599,000$ to time $k = 600,000$, which is within the steady state time ranges above.

Similarly, in the second half, we can identify the steady state time ranges for equalizer 1 and 2 as $k = 6.5 \times 10^5$ to $k = 12 \times 10^5$ and $k = 10 \times 10^5$ to $k = 12 \times 10^5$, respectively. Therefore, we plot the averaged global impulse responses from the time range $k = 1,199,000$ to time $k = 1,200,000$ in Fig 3.8 (d). Among the four impulse responses, we can see that $|s_{21}|$ and $|s_{12}|$ have spikes.

Thus, the equalizer 1 recovers source 2 and the equalizer 2 recovers source 1. This indicates that different sources have been recovered. Since the source separation ability can be found in the first and the second half, we conclude that the source separation ability can be achieved in 16-QAM.

3.5.8 Summary

We have presented the simulation results for CI-CMA on 2-PAM, 4-PAM, 4-QAM and 16-QAM. CI-CMA can open the channel eye and does not introduce any shrinking issue on the 4 cases. For the multi-modulus modulation schemes such as 4-PAM and 16-QAM, CI-CMA can perform equalization and source separation ability simultaneously. However, CI-CMA is not able to perform the source separation task in the single-modulus modulation schemes such as 2-PAM and 4-QAM.

Chapter 4

Steady State MSE Analysis of the CI-CMA

4.1 Introduction

We have developed a new MIMO algorithm, CI-CMA in the previous chapter. Indeed, we have shown its convergence analysis and source separation ability. It is also important to evaluate the steady steady MSE of the algorithm because this is often used as a performance measurement for equalization algorithm. Therefore, the chief purpose of this chapter is to perform an accurate steady state MSE analysis for the CI-CMA using the energy-preserving theorem [54] initially developed for the single-input/single-output (SISO) blind deconvolutive system. The steady state MSE analysis has been carried out for the CC-CMA in [50, 131]. In [50], Luo *et al* have made a good contribution in analyzing the steady state MSE of the CC-CMA in MIMO systems. Unfortunately, the analysis in [50, 131] have been done in the presence of the shrinking issue, which results in noticeable mismatches between the simulation and theoretical results for certain practical range of values. Therefore, this motivates us to analysis the steady state MSE of the CI-CMA since the shrinking issue has been resolved in the CI-CMA.

We wish to emphasize that the subsequent steady state MSE analysis is carried out with two major differences than the previous works of [50, 44]. Firstly, we use new dispersion constants, R_j (shown in the previous chapter) as compared to the conventional dispersion constant, $\mathrm{R} = \frac{m_4}{\sigma_{\mathrm{A}}^2}$. Secondly, our MSE analysis is more accurate than [50, 44]. This is because the previous works analyzed the CC-CMA which is a biased cost function based on energy preservation theorem. It is worth to mention that the energy preservation theorem was originally established for an unbiased cost function [54]. Therefore, we improve the previous works by analyzed the new CI-CMA which is an unbiased cost function based on the energy preservation theorem.

The chapter is organized as follows. Section 4.2 describes the system model and assumptions made. Section 4.3 briefly describes the generalized CI-CMA. Section 4.4 describes in detail how the steady state MSE of the CI-CMA is evaluated. Several terms which have previously been approximated as zero in [50] are re-derived and mathematically proven as finite values that contribute significantly to the steady state MSE. Section 4.5 shows that the MSE obtained through simulations closely matches the MSE obtained analytically.

4.2 System model and assumptions

We consider the system model that has M_{t} transmit sources and M_{r} receive sensors. We further make the following assumptions:

A1 Assumptions on sources : Same as assumption A1 in Section 2.8. In additional, the signal has finite positive values on sixth order moment $m_6 = \mathrm{E}\{|a_j[k]|^6\}$, fourth order moment $m_4 = \mathrm{E}\{|a_j[k]|^4\}$ and variance $\sigma_{\mathrm{A}}^2 = \mathrm{E}\{|a_j[k]|^2\}$.

A2 Assumption on channels: Same as assumption A2 in Section 2.2.

A3 Assumption on sensor noise(s): We assume no sensor noise, therefore $\sigma_{\mathrm{n}}^2 = 0$.

We explicitly recall MIMO equalizer structure here because it will be used to derive our new MSE expression. The observed signal at the i-th sensor, for $i \in (1, M_{\mathrm{r}})$, at time k is defined as:

$$x_i[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij} \otimes \mathbf{a}_j[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij}^T \mathbf{a}_j[k] \tag{4.1}$$

where $\mathbf{a}_j[k] = \big[a_j[k], \cdots, a_j[k - N_h + 1]\big]^T$, $\otimes$ and ^{T} denote discrete time convolution and vector transposition, respectively. The equalizer output corresponding to the j-th source, where $j \in (1, M_{\mathrm{t}})$, at time k is denoted as:

$$y_j[k] = \sum_{i=1}^{M_{\mathrm{r}}} \mathbf{f}_{ij}^T[k] \mathbf{x}_i[k] \tag{4.2}$$

where $\mathbf{x}_i[k] = [x_i[k], \cdots, x_i[k - L + 1]]^T$ is the input regressor vector for the equalizers of the i-th seonsor (there are a total of M_{t} equalizers associated with each of the i sensors), L is sufficiently large to eliminate the ISI of the channel, and $\mathbf{f}_{ij}[k] = \left[f_{ij}^{[0]}[k], \cdots, f_{ij}^{[L-1]}[k]\right]^T$ is a $L \times 1$ vector that represents the equalizer connecting the observed signal at i-th sensor to the j-th equalizer output at time k. There are a total of $M_{\mathrm{t}} \cdot M_{\mathrm{r}}$ equalizers at the receiver end (see Fig. 4.1). Furthermore, Eq. 4.2 can be written as

$$y_j[k] = \mathbf{F}_j^T[k] \mathbf{X}[k] \tag{4.3}$$

where $\mathbf{F}_j[k] = \big[\mathbf{f}_{1j}^T[k], \cdots, \mathbf{f}_{M_{\mathrm{r}}j}^T[k]\big]^T$ is a $M_{\mathrm{r}}L \times 1$ vector and $\mathbf{X}[k] = \big[\mathbf{x}_1^T[k], \cdots, \mathbf{x}_{M_{\mathrm{r}}}^T[k]\big]^T$ is a $M_{\mathrm{r}}L \times 1$ vector.

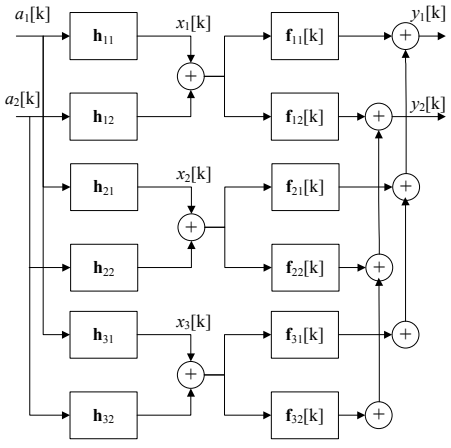

Figure 4.1: Baseband equivalent system for $M_{\mathrm{t}} = 2$ and $M_{\mathrm{r}} = 3$.

4.3 The CI-CMA

The CI-CMA which can mitigate the ISI and CCI simultaneously was proposed in the previous chapter. In this section, we recall the CI-CMA for M_t sources. For simplicity, we will pursue $\mathrm{M}_t = 2$. We first state the estimated cost functions and then show its SGA (stochastic gradient algorithm).

We recall that the cost function of the first equalizer output is the CMA cost function and the cost function of the jth equalizer output is the CI-CMA cost function. Therefore, the cost functions of the first and j-th equalizer outputs, $J_1[k]$ and $J_j[k]$ are given by

$$J_1[k] \quad = \quad \mathrm{E}\left\{\left(|y_1[k]|^2 - \mathrm{R}_1\right)^2\right\} \tag{4.4}$$

and

$$J_j[k] \quad = \quad \mathrm{E}\left\{\left(|y_j[k]|^2 - \mathrm{R}_j\right)^2\right\} + 2k_j \sum_{\substack{n=1 \\ n\neq m}}^{M_\mathrm{t}} \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \mathrm{E}\left\{\left|y_n[k]y_m^*[k-\delta]\right|^2\right\} \tag{4.5}$$

respectively, where $\mathrm{R}_1, \mathrm{R_j} \in \mathbb{R}^+$ are the dispersion constants where the values will be stated later, $k_j \in \mathbb{R}^+$ is known as the mixing parameter for the j-th equalizer output and $\delta_{\max} \in \mathbb{Z}^+$ accounts for the maximum channel delay spread between all equalizer outputs.

Now suppose that we wish to extract the j-th source only where $j \in (1, M_\mathrm{t})$. By the SGA, the update equation of the equalizer coefficients, $\mathbf{F}_j[k]$, is given as

$$\mathbf{F}_j[k+1] = \mathbf{F}_j[k] - \mu_j \left(e_{\mathrm{CMA},j}[k] + k_j e_{\mathrm{CI},j}[k]\right) \mathbf{X}^*[k] \tag{4.6}$$

where μ_j is the adaptation step size for the j-th equalizer output, $e_{\mathrm{CMA},j}[k]$ and $e_{\mathrm{CI},j}[k]$ are the respective error functions of the CMA and cross-independent (CI) terms given as

$$e_{\mathrm{CMA},j}[k] = (|y_j[k]|^2 - \mathrm{R}_j) y_j[k] \tag{4.7}$$

and

$$e_{\mathrm{CI},j}[k] = \sum_{\substack{l=1 \\ l\neq j}}^{M_\mathrm{t}} \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} y_j[k]\left|y_l[k-\delta]\right|^2, \tag{4.8}$$

where $y_l[k-\delta]$ that involves the past, current or future time l-th equalizer output sample can be estimated as below [113]

$$y_l[k-\delta] = \mathbf{F}_l^T[k]\mathbf{X}[k-\delta] \tag{4.9}$$

and

$$\mathrm{R}_j = \frac{m_4}{\sigma_{\mathrm{A}}^2} + \beta_j \sigma_{\mathrm{A}}^2 \tag{4.10}$$

is the dispersion constant of the j-th equalizer output with [1]

$$\beta_j = k_j(j-1)(M_{\mathrm{t}}-1)(2\delta_{\max}+1) \tag{4.11}$$

designed to ensure all the equalizer output energies are restored to the sources' energies. In additional, we highlight that the dispersion constant of the first ($j=1$) equalizer output, R_1 is $\frac{m_4}{\sigma_{\mathrm{A}}^2}$.

4.4 Steady state MSE analysis of CI-CMA

4.4.1 Recap on energy-preserving theorem and assumptions

In order to evaluate the steady state MSE of a blind adaptive equalization algorithm, Mai and Sayed [54] developed an elegant approach by exploiting the energy-preserving relation of the filter update process which we now recap. At steady state where we have assumed the equalizer weights have converged to their global minima, the output of the j-th equalizer will converge towards

$$y_j[k] \rightarrow a_j[k-d_j]\exp(\mathbf{j}\theta_j) - e_{\mathrm{a},j}[k] \tag{4.12}$$

where d_j is an integer that represents the delay for j-th source and θ_j is an arbitrary phase offset which for the purpose of this analysis is assumed $\theta_j = 0$[2], and $e_{\mathrm{a},j}[k]$ is due to the gradient noise

[1] $\mathrm{R_j}$ was derived in Section 3.3.4 and the mixing parameter k_j is defined as $k_j = \frac{k_0}{2}$ (where $\frac{k_0}{2}$ was determined in Section 3.3.6.) in this chapter.

[2] A phase lock loop (PLL) or the multi-modulus algorithm (MMA) may be used to correct the phase offset experienced by the CMA in the steady state, subject to an arbitrary $\pi/2$ offset [132].

of the stochastic update algorithm. In fact, the steady state MSE is defined as

$$\text{MSE} = \lim_{k\to\infty} \text{E}\left\{|e_{\text{a},j}[k]|^2\right\}. \tag{4.13}$$

According to the energy-preserving relation, $e_{\text{a},j}[k]$ is the *a-prioiri* estimation error and $e_{\text{p},j}[k]$ is the *a-posteriori* estimation error, defined as

$$e_{\text{a},j}[k] = \Delta\mathbf{F}_j^T[k]\mathbf{X}[k] \tag{4.14}$$

$$e_{\text{p},j}[k] = \Delta\mathbf{F}_j^T[k+1]\mathbf{X}[k] \tag{4.15}$$

where $\Delta\mathbf{F}_j[k] = \mathbf{F}_{\text{ZF},j} - \mathbf{F}_j[k]$ is the difference between the weight vectors of the optimal zero-forcing (ZF) equalizer and those of the j-th equalizer at time k. Note the ZF equalizer yields $\mathbf{F}_{\text{ZF},j}^T\mathbf{X}[k] = a_j[k-d_j]\exp(\mathbf{j}\theta_j)$.

Mai and Sayed [54] utilized $e_{\text{a},j}[k]$, $e_{\text{p},j}[k]$ and the update equation 4.6 to establish

$$e_{\text{p},j}[k] = e_{\text{a},j}[k] + \mu_j e_{\text{o},j}[k]\|\mathbf{X}[k]\|^2 \tag{4.16}$$

where $e_{\text{o},j}[k] = e_{\text{CMA},j}[k] + k_j e_{\text{CI},j}[k]$ accounts for the CI-CMA error function at time k of the j-th equalizer output. This leads the following energy-preserving relation

$$\text{E}\left\{\|\Delta\mathbf{F}_j[k+1]\|^2\right\} + \text{E}\left\{\frac{|e_{\text{a},j}[k]|^2}{\|\mathbf{X}[k]\|^2}\right\} = \text{E}\left\{\|\Delta\mathbf{F}_j[k]\|^2\right\} + \text{E}\left\{\frac{|e_{\text{p},j}[k]|^2}{\|\mathbf{X}[k]\|^2}\right\} \tag{4.17}$$

At this point, we state the following assumptions so that the MSE may be extracted from Eq. 4.17 after some mathematical manipulation. At steady state (i.e., $k \to \infty$),

A3 $\text{E}\left\{\|\Delta\mathbf{F}_j[k+1]\|^2\right\} = \text{E}\left\{\|\Delta\mathbf{F}_j[k]\|^2\right\}$, which is a common assumption, c.f. [54, 44, 50, 131];

A4 For a sufficient long equalizer (i.e., $L \gg N_h$) and a sufficient small step size (i.e., $\mu_j \approx 0$), $\|\mathbf{X}[k]\|^2$ is independent of $|e_{\text{o},j}[k]|^2$ and $y_j[k]$ (the details of this assumption can be found in [54]);

A5 The j-th source, $a_j[k]$, is independent with the estimation error at the l-th equalizer output, $e_{\text{a},l}[k]$, i.e. $\text{E}\left\{a_j[k-d_j]e_{\text{a},l}^*[k-\delta]\right\} = 0$ where $j \in (1, M_\text{r})$, $l \in (1, M_\text{r})$ and $d_j, \delta \in \mathbb{Z}$ is an arbitrary integer, c.f. [50, 131];

A6 For $L \gg N_h$ and $\mu_j \approx 0$, the estimation error at different equalizer outputs are assumed to be uncorrelated, i.e. $\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{a},l}[k-\delta]\right\} = 0$, c.f. [50] [131];

A7 When μ_j is small enough, the estimation of $\mathrm{E}\left\{\mu_j \left|e_{\mathrm{a},j}[k]\right|^{2p}\right\}$, where $p \in \mathbb{Z}^+$, is negligible and can be ignored, c.f. [54, 50].

Equipped with assumptions A3, A4 and Eq. 4.16, we can simplify Eq. 4.17 to

$$\underbrace{\mu_j \mathrm{E}\left\{\|\mathbf{X}[k]\|^2\right\}\mathrm{E}\left\{|e_{\mathrm{o},j}[k]|^2\right\}}_{\text{L.H.S.}} = -\underbrace{\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{o},j}[k] + e^*_{\mathrm{a},j}[k]e_{\mathrm{o},j}[k]\right\}}_{\text{R.H.S.}} \tag{4.18}$$

The subsequent derivations are aimed at extracting the MSE (c.f. Eq. 4.13) from Eq. 4.18. It is achieved by systematically expanding the L.H.S. and expressing it in terms of the source statistics, i.e., $\mathrm{E}\left\{|a_j[k]|^2\right\}$ and $\mathrm{E}\left\{|a_j[k]|^4\right\}$, etc. This manipulation of the L.H.S. is pursued in the next subsection (Section 4.4.2). After that we expand the R.H.S. and express it in terms of the MSE, i.e., $\mathrm{E}\left\{|e_{\mathrm{a},j}[k]|^2\right\}$, and this is done in Section 4.4.3. Finally we equate the L.H.S. and the R.H.S. to arrive at an expression of the steady state MSE.

4.4.2 Analysis on L.H.S. of Eq. 4.18

The L.H.S. of Eq. 4.18 can be expanded as follows into terms X, A, B, C, and D:

$$\begin{aligned}\mu_j \mathrm{E}\left\{\|\mathbf{X}[k]\|^2\right\}\mathrm{E}\left\{|e_{\mathrm{o},j}[k]|^2\right\} \;=\; & \mu_j \underbrace{\mathrm{E}\left\{\|\mathbf{X}[k]\|^2\right\}}_{\text{Term X}}\Bigg(\underbrace{\mathrm{E}\{|e_{\mathrm{CMA},j}[k]|^2\}}_{\text{Term A}} + \underbrace{k_j^2\mathrm{E}\{|e_{\mathrm{CI},j}[k]|^2\}}_{\text{Term B}} \\ & + \underbrace{k_j\mathrm{E}\{e^*_{\mathrm{CMA},j}[k]e_{\mathrm{CI},j}[k]\}}_{\text{Term C}} + \underbrace{k_j\mathrm{E}\{e_{\mathrm{CMA},j}[k]e^*_{\mathrm{CI},j}[k]\}}_{\text{Term D}}\Bigg).\end{aligned} \tag{4.19}$$

4.4.2.1 Term X

Utilizing Assumptions A1, A2 and Eq. 4.1, it can be shown that

$$\mathrm{E}\left\{\|\mathbf{X}[k]\|^2\right\} = L\sigma_{\mathrm{A}}^2 \sum_{i=1}^{M_{\mathrm{r}}}\sum_{b=1}^{M_{\mathrm{t}}} \mathbf{h}_{ib}^H \mathbf{h}_{ib} \tag{4.20}$$

where ^{H} denotes Hermitian operator and $\sum_{i=1}^{M_{\mathrm{r}}}\sum_{b=1}^{M_{\mathrm{t}}} \mathbf{h}_{ib}^H\mathbf{h}_{ib}$ denotes the total power gain of the MIMO channel.

4.4.2.2 Term A

Term A, which was derived in [54], is simply

$$\mathrm{E}\{|e_{\mathrm{CMA},j}[k]|^2\} = \mathrm{E}\left\{\mathrm{R}_j^2|a_j[k]|^2 - 2\mathrm{R}_j|a_j[k]|^4 + |a_j[k]|^6\right\}. \tag{4.21}$$

4.4.2.3 Term B

Before we proceed, we wish to highlight that the evaluation of Term B is most different from previous works [44, 50]. This is because we found that Term B is a function of m_4 which had not been found by the previous works.

By substituting Eq. 4.8 into Term B, we get the following expression

$$k_j^2\mathrm{E}\{|e_{\mathrm{CI},j}[k]|^2\} = k_j^2 \sum_{\substack{n=1\\n\neq j}}^{M_\mathrm{t}} \sum_{\substack{l=1\\l\neq j}}^{M_\mathrm{t}} \sum_{\delta_A=-\delta_{\max}}^{\delta_{\max}} \sum_{\delta_B=-\delta_{\max}}^{\delta_{\max}} g_j(n,l,\delta_A,\delta_B) \tag{4.22}$$

where

$$g_j(n,l,\delta_A,\delta_B) = \mathrm{E}\left\{|y_j[k]|^2\,|y_n[k-\delta_A]|^2\,|y_l[k-\delta_B]|^2\right\}. \tag{4.23}$$

To evaluate Eq. 4.23, we replace $y_j[k] \rightarrow a_j[k_j] - e_{\mathrm{a},j}[k]$, where $j \in (1, M_\mathrm{t})$ and we abbreviated $k_j = k - d_j$ to save space, and then utilizing Assumptions A1, A5, A6 and A7. Then, we get

$$g_j(n,l,\delta_A,\delta_B) \approx \mathrm{E}\left\{|a_j[k_j]|^2\right\}\mathrm{E}\left\{|a_n[k_n-\delta_A]|^2\,|a_l[k_l-\delta_B]|^2\right\}. \tag{4.24}$$

Now, we substitute Eq. 4.24 into Eq. 4.22, we then get the new Term B as

$$k_j^2\mathrm{E}\{|e_{\mathrm{CI},j}[k]|^2\} = k_j^2 \sum_{\substack{n=1\\n\neq j}}^{M_\mathrm{t}} \sum_{\substack{l=1\\l\neq j}}^{M_\mathrm{t}} \sum_{\delta_A=-\delta_{\max}}^{\delta_{\max}} \sum_{\delta_B=-\delta_{\max}}^{\delta_{\max}} \mathrm{E}\left\{|a_j[k_j]|^2\right\}\mathrm{E}\left\{|a_n[k_n-\delta_A]|^2\,|a_l[k_l-\delta_B]|^2\right\} \tag{4.25}$$

Regarding to Eq. 4.25, we can identify four scenarios that could occur:

Case 1 : $n = l$ and $\delta_A = \delta_B$ with $n \neq j$ and $l \neq j$;

Case 2 : $n = l$ and $\delta_A \neq \delta_B$ with $n \neq j$ and $l \neq j$;

Case 3 : $n \neq l$ and $\delta_A = \delta_B$ with $n \neq j$ and $l \neq j$;

Case 4 : $n \neq l$ and $\delta_A \neq \delta_B$ with $n \neq j$ and $l \neq j$.

We are going to evaluate the four cases. Among the four cases, Case 1 was not been evaluated in [50]. In order to obtain an accurate MSE analysis, we determine Case 1 now. In Case 1, we can evaluate $\mathrm{E}\{|a_j[k_j]|^2\}\mathrm{E}\{|a_n[k_n-\delta_A]|^2|a_l[k_l-\delta_B]|^2\}$ as $\sigma_{\mathrm{A}}^2 m_4$. In order to complete the evaluation of Case 1, we wish to know the number of occurrences of Case 1 in Eq. 4.25. Via the summations $\sum_{\substack{n=1 \\ n\neq j}}^{M_{\mathrm{t}}} \sum_{\substack{l=1 \\ l\neq j}}^{M_{\mathrm{t}}} \sum_{\delta_A=-\delta_{\max}}^{\delta_{\max}} \sum_{\delta_B=-\delta_{\max}}^{\delta_{\max}}$ in Eq. 4.25, we can determine the number of occurrence of Case 1 is $(M_{\mathrm{t}}-1)(2\delta_{\max}+1)$. Therefore, the final answer of Case 1 is

$$k_j^2(M_{\mathrm{t}}-1)(2\delta_{\max}+1)\sigma_{\mathrm{A}}^2 m_4 \tag{4.26}$$

Next, we evaluate Case 2, 3 and 4 together. In those cases, we notice $\mathrm{E}\{|a_j[k_j]|^2\}\mathrm{E}\{|a_n[k_n-\delta_A]|^2|a_l[k_l-\delta_B]|^2\} = \mathrm{E}\{|a_j[k_j]|^2\}\mathrm{E}\{|a_n[k_n-\delta_A]|^2\}\mathrm{E}\{|a_l[k_l-\delta_B]|^2\}$, and therefore we can simplify it as σ_{A}^6. Then, we wish to determine the number of occurrences of those cases. In fact, the number of occurrences of Case 2, 3 and 4 can be obtained by the number of occurrences of the four cases subtracts the number of occurrences of Case 1. From Eq. 4.25, the number of occurrences of the four cases is $(M_{\mathrm{t}}-1)^2(2\delta_{\max}+1)^2$, therefore, we can get the number of occurrences of Case 2, 3 and 4 as $(M_{\mathrm{t}}-1)^2(2\delta_{\max}+1)^2-(M_{\mathrm{t}}-1)(2\delta_{\max}+1)$. Eventually, we get the combined answer of Case 2, 3 and 4 which is

$$k_j^2\big((M_{\mathrm{t}}-1)^2(2\delta_{\max}+1)^2-(M_{\mathrm{t}}-1)(2\delta_{\max}+1)\big)\sigma_{\mathrm{A}}^6 \tag{4.27}$$

Finally, combining the answer of Case 1 (Eq. 4.26) and the answer of Case 2, 3 and 4 (Eq. 4.27), we obtain Term B as

$$\begin{aligned} k_j^2\mathrm{E}\{|e_{\mathrm{CI},j}[k]|^2\} &= k_j^2\Bigg((M_{\mathrm{t}}-1)(2\delta_{\max}+1)\sigma_{\mathrm{A}}^2 m_4 \\ &\quad +\Bigg((M_{\mathrm{t}}-1)^2(2\delta_{\max}+1)^2-(M_{\mathrm{t}}-1)(2\delta_{\max}+1)\Bigg)\sigma_{\mathrm{A}}^6\Bigg) \end{aligned} \tag{4.28}$$

4.4.2.4 Term C & Term D

Terms C and D have identical values as they are complex conjugate of each other. We first consider Term C in Eq. (4.19). Substituting Eq. 4.7 and Eq. 4.8 into Term C, we then get

$$k_j \mathrm{E}\{e^*_{\mathrm{CMA},j}[k] e_{\mathrm{CI},j}[k]\} \quad = \quad k_j \sum_{\substack{l=1 \\ l \neq j}}^{M_\mathrm{t}} \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \mathrm{E}\Big\{ \left|y_l[k-\delta]\right|^2 \left(\left|y_j[k]\right|^4 - \mathrm{R}_j \left|y_j[k]\right|^2 \right) \Big\} \tag{4.29}$$

To evaluate Eq. 4.29, we replace $y_j[k] \rightarrow a_j[k_j] - e_{\mathrm{a},j}[k]$, where $j \in (1, M_\mathrm{t})$ and we abbreviated $k_j = k - d_j$ to save space, then we apply Assumption A7 to ignore small error terms, i.e., $\mathrm{E}|e_{\mathrm{a},j}[k]|^2$ and $\mathrm{E}|e_{\mathrm{a},j}[k]|^4$. Thus Eq. 4.29 can be simplified as

$$k_j \mathrm{E}\{e^*_{\mathrm{CMA},j}[k] e_{\mathrm{CI},j}[k]\} \quad = \quad k_j (M_\mathrm{t} - 1)(2\delta_{\max} + 1)\sigma_\mathrm{A}^2 \Big(m_4 - \mathrm{R_j} \sigma_\mathrm{A}^2 \Big). \tag{4.30}$$

We notice that if $\mathrm{R_j}$ is replaced with $\mathrm{R} = \frac{m_4}{\sigma_\mathrm{A}^2}$, then Eq. 4.30 becomes zero. In this case, we have proven that the approximation of Term C to zero in [50] is valid. The distinction between our analysis and the previous works of [50, 44] is that we managed to mathematically show that Term C is actually zero and not merely by assuming it is. However, in this chapter, we consider a different dispersion constant, i.e., $\mathrm{R}_j = m_4/\sigma_\mathrm{A}^2 + \beta_j \sigma_\mathrm{A}^2$. Therefore we emphasize that Term C in Eq. 4.30 is non-zero for the CI-CMA except the first equalizer output (i.e. $j = 1$). Since Term D is conjugate counterpart of Term C, the final expression of Term D is same as Eq. 4.30.

4.4.2.5 Combined Expression for L.H.S. of Eq. 4.18

Finally, combining Term X (Eq. 4.20), Term A (Eq. 4.21), Term B (Eq. 4.28), Term C (Eq. 4.30), Term D (also Eq. 4.30), and replacing R_j with $m_4/\sigma_\mathrm{A}^2 + \beta_j \sigma_\mathrm{A}^2$, the L.H.S. of Eq. 4.19 can be simplified as Eq. 4.31 as below after some algebra manipulations.

$$\begin{aligned} &\underbrace{\mu \mathrm{E}\left\{ \|\mathbf{X}[k]\|^2 \right\} \mathrm{E}\left\{ |e_{\mathrm{o},j}[k]|^2 \right\}}_{\text{L.H.S. of Eq. 4.18}} \\ = \quad & \mu L \sigma_\mathrm{A}^2 \Bigg(\sum_{i=1}^{M_\mathrm{r}} \sum_{b=1}^{M_\mathrm{t}} \mathbf{h}_{ib}^H \mathbf{h}_{ib} \Bigg) \Bigg(m_6 - \frac{(m_4)^2}{\sigma_\mathrm{A}^2} + k_j \beta_j \sigma_\mathrm{A}^2 \left(m_4 - \sigma_\mathrm{A}^4 \right) \Bigg) \end{aligned} \tag{4.31}$$

4.4.3 Analysis on R.H.S. of Eq. 4.18

The purpose of expanding the R.H.S. of Eq. 4.18 is to find an expression of the MSE, $\mathrm{E}\{|e_{\mathrm{a},j}[k]|^2\}$, in a similar fashion as [54, 50, 44]. This can be done by expanding its first term to

$$\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{o},j}[k]\right\} = \mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{CMA},j}[k] + k_j e_{\mathrm{a},j}[k]e^*_{\mathrm{CI},j}[k]\right\} \tag{4.32}$$

since $e^*_{\mathrm{o},j}[k] = e^*_{\mathrm{CMA},j}[k] + k_j e^*_{\mathrm{CI},j}[k]$. The first term of Eq. 4.32, $\mathrm{E}\{e_{\mathrm{a},j}[k]e^*_{\mathrm{CMA},j}[k]\}$, has been derived in [54] before. We simply re-state the result but with a different dispersion constant:

$$\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{CMA},j}[k]\right\} = \mathrm{E}\{|e_{\mathrm{a},j}[k]|^2\}\left(\mathrm{R}_j - 2\sigma^2_{\mathrm{A}}\right). \tag{4.33}$$

Now we consider the second term in Eq. 4.32. By substituting 4.8 into the term, we get

$$k_j\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{CI},j}[k]\right\} \quad = \quad k_j\mathrm{E}\Bigg\{e_{\mathrm{a},j}[k]\sum_{\substack{l=1\\l\neq j}}^{M_{\mathrm{t}}}\sum_{\delta=-\delta_{\max}}^{\delta_{\max}}\left(y^*_j[k]|y_l[k-\delta]|^2\right)\Bigg\} \tag{4.34}$$

We replace $y_j[k] \to a_j[k_j] - e_{\mathrm{a},j}[k]$ where $k_j = k - d_j$ for $j \in (1, M_{\mathrm{t}})$ into Eq. 4.34, we then get

$$\begin{aligned} & k_j\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{CI},j}[k]\right\} \qquad (4.35)\\ = \quad & k_j\mathrm{E}\Bigg\{\left(a^*_j[k_j]e_{\mathrm{a},j}[k] - |e_{\mathrm{a},j}[k]|^2\right)\times \\ & \left(\sum_{\substack{l=1\\l\neq j}}^{M_{\mathrm{t}}}\sum_{\delta=-\delta_{\max}}^{\delta_{\max}}\left(|a_l[k_l-\delta]|^2 + |e_{\mathrm{a},l}[k]|^2 + a^*_l[k_l-\delta]e_{\mathrm{a},l}[k] + a_l[k_l-\delta]e^*_{\mathrm{a},l}[k]\right)\right)\Bigg\} \end{aligned}$$

We apply assumptions A1, A5 and A6 into Eq. 4.35, then Eq. 4.35 can be simplified as

$$\begin{aligned} & k_j\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{CI},j}[k]\right\} \qquad (4.36)\\ = \quad & k_j\mathrm{E}\big\{-|e_{\mathrm{a},j}[k]|^2\big\}\sum_{\substack{l=1\\l\neq j}}^{M_{\mathrm{t}}}\sum_{\delta=-\delta_{\max}}^{\delta_{\max}}\left(\mathrm{E}\big\{|a_l[k_l-\delta]|^2\big\} + \mathrm{E}\big\{|e_{\mathrm{a},l}[k]|^2\big\}\right) \end{aligned}$$

Now, we introduce an assumption from [50] that the error magnitudes are approximately same which is $\mathrm{E}\big\{|e_{\mathrm{a},l}[k]|^2\big\} \approx \mathrm{E}\big\{|e_{\mathrm{a},j}[k]|^2\big\}$. Equipped with the above assumption and the stationary assumption of the sources in A1, Eq. 4.36 can be further simplified as

$$-k_j(M_{\mathrm{t}}-1)(\delta_{\max}+1)\left(\sigma^2_{\mathrm{A}}\mathrm{E}\left\{|e_{\mathrm{a},j}[k]|^2\right\} + \left(\mathrm{E}\left\{|e_{\mathrm{a},j}[k]|^2\right\}\right)^2\right) \tag{4.37}$$

It is not impractical to assume that $(\mathrm{E}\{|e_{\mathrm{a},j}[k]|^2\})^2 \ll \mathrm{E}\{|a_j[k]|^2\}\mathrm{E}\{|e_{\mathrm{a},j}[k]|^2\}$ (which is the assumption from [50]), and therefore Eq. 4.37 can be further simplified to

$$k_j \mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{CI},j}[k]\right\} \approx -\beta_j \sigma_{\mathrm{A}}^2 \mathrm{E}\left\{|e_{\mathrm{a},j}[k]|^2\right\} \tag{4.38}$$

where $\beta_j = k_j(M_{\mathrm{t}} - 1)(2\delta_{\max} + 1)$ as before, c.f. Eq. 4.11.

After deriving both terms in Eq. 4.32, i.e., Eq. 4.33 and Eq. 4.38, we substitute them back in Eq. 4.32. By also replacing R_j with $m_4/\sigma_{\mathrm{A}}^2 + \beta_j\sigma_{\mathrm{A}}^2$, we then get

$$\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{o},j}[k]\right\} = -\mathrm{E}\left\{|e_{\mathrm{a},j}[k]|^2\right\}\left(2\sigma_{\mathrm{A}}^2 - \frac{m_4}{\sigma_{\mathrm{A}}^2}\right). \tag{4.39}$$

Now the conjugate counterpart of Eq. 4.32, i.e., $\mathrm{E}\left\{e^*_{\mathrm{a},j}[k]e_{\mathrm{o},j}[k]\right\}$, would have the same expression as Eq. 4.39. Since both terms are identical, the R.H.S. of Eq. 4.18 is thus

$$-\mathrm{E}\left\{e_{\mathrm{a},j}[k]e^*_{\mathrm{o},j}[k] + e^*_{\mathrm{a},j}[k]e_{\mathrm{o},j}[k]\right\} = 2\mathrm{E}\left\{|e_{\mathrm{a},j}[k]|^2\right\}\left(2\sigma_{\mathrm{A}}^2 - \frac{m_4}{\sigma_{\mathrm{A}}^2}\right). \tag{4.40}$$

Note that since Eq. 4.40 has an MSE term in it, our objective to expand the R.H.S. of Eq. 4.18 is achieved.

4.4.4 Expression of the Steady State MSE

Finally, with the L.H.S. of Eq. 4.18 expressed in terms of the source statistics, c.f. (Eq. 4.31), and the R.H.S. of Eq. 4.18 in terms of the MSE, c.f. (Eq. 4.40), we can proceed to derive an expression of the steady state MSE of the CI-CMA. By equating the L.H.S. and the R.H.S., the steady state MSE for the j-th equalizer output can be approximated as Eq. 4.41 which is found below. We have also intentionally placed the (simplified[3]) MSE expression derived by Luo *et al* [50] in Eq. 4.42 below our MSE expression for the purpose of comparison. For values of k_j which are very close to zero, both Eq. 4.41 and Eq. 4.42 become almost identical. By inspection of Eq. 4.31, it is clear that not only is the steady state MSE being influenced by the statistical properties of the sources, but it is also influenced by the total power gain of MIMO channels, the selection

[3]Simplification is achieved by making the assumption of stationarity of the sources.

of μ_j, filter length L, mixing parameter k_j, maximum delay spread $\delta_{\max}$, no. of transmit M_{t} and received antennas M_{r} (some of these terms are absorbed into β_j). In addition, note that if $M_{\mathrm{t}} = 1$ and $M_{\mathrm{r}} = 1$, then Eq. 4.41 reduces to the SISO case and agrees with the expression in [54].

$$\begin{aligned} \mathrm{MSE}_{\mathrm{CI-CMA}_j} &= \mathrm{E}\left\{|e_{\mathrm{a},j}[k]|^2\right\} \\ &= \mu_j L \sigma_{\mathrm{A}}^2 \sum_{i=1}^{M_{\mathrm{r}}} \sum_{b=1}^{M_{\mathrm{t}}} \mathbf{h}_{ib}^H \mathbf{h}_{ib} \left(\frac{m_6 - \frac{m_4^2}{\sigma_{\mathrm{A}}^2} + k_j \beta_j \sigma_{\mathrm{A}}^2 \left(m_4 - \sigma_{\mathrm{A}}^4\right)}{2\left(2\sigma_{\mathrm{A}}^2 - \frac{m_4}{\sigma_{\mathrm{A}}^2}\right)} \right) \end{aligned} \tag{4.41}$$

$$\mathrm{MSE}_{\mathrm{previous,j}} = \mu_j L \sigma_{\mathrm{A}}^2 \sum_{i=1}^{M_{\mathrm{r}}} \sum_{b=1}^{M_{\mathrm{t}}} \mathbf{h}_{ib}^H \mathbf{h}_{ib} \left(\frac{m_6 - \frac{m_4^2}{\sigma_{\mathrm{A}}^2} + k_j \beta_j \sigma_{\mathrm{A}}^6}{2\left(2\sigma_{\mathrm{A}}^2 - \frac{m_4}{\sigma_{\mathrm{A}}^2}\right)} \right) \tag{4.42}$$

4.5 Simulations

4.5.1 Simulation Setup

In order to simulate the system model in Section 4.2 and Fig. 4.1 with ($M_{\mathrm{t}} = 2$ and $M_{\mathrm{r}} = 3$) MIMO channel, simulations are performed using MATLAB. Two random 16-QAM symbol sequences that have $N_{seq} = 3 \times 10^6$ symbols in one sequence are generated as source 1, $a_1[k]$ where $j = 1$ and source 2, $a_2[k]$ where $j = 2$. The sources are convoluted with the memory channels $\mathbf{h}_{ij}$, and this process is equivalent to $\mathbf{h}_{ij} \otimes \mathbf{a}_j[k]$ where $j = 1, 2$. This is aimed to simulate the condition that the sources are passed through the memory channels and thereby the sources are interfered by ISI. At the i-th received sensor, the ISI signals from two sources $a_1[k]$ and $a_2[k]$ (but not a single source) have been observed, thus the observed signal at i-th sensor $x_i[k]$ is corrupted by CCI due to receiving two sources and ISI due to memory channels. This scenario can be simulated through computing Eq. 4.1 and the impulse responses of channels $\mathbf{h}_{ij}$ are stated in Table 4.1.

Next, to simulate the algorithm in the equalizer, we will test using the hierarchical CI-CMA, where the first equalizer uses CMA and the second equalizer uses CI-CMA. The filter weights

Table 4.1: Impulse response of two-input/three-output complex channel [1]

$\mathbf{h_{11}}$	$[-0.6+0.4\mathbf{j}, 1.2-0.2\mathbf{j}]^T$
$\mathbf{h_{12}}$	$[-0.6+0.8\mathbf{j}, 0.9-0.1\mathbf{j}]^T$
$\mathbf{h_{21}}$	$[0.1+0.7\mathbf{j}, -0.2-0.5\mathbf{j}]^T$
$\mathbf{h_{22}}$	$[0.4-0.3\mathbf{j}, -0.2+0.2\mathbf{j}]^T$
$\mathbf{h_{31}}$	$[0.5+0.4\mathbf{j}, -1+0.3\mathbf{j}]^T$
$\mathbf{h_{32}}$	$[-0.1+0.8\mathbf{j}, 0.4+0.1\mathbf{j}]^T$

$\mathbf{f}_{11}[0]$ and $\mathbf{f}_{22}[0]$ are initialized with a center spike [36] whilst $\mathbf{f}_{ij}[0]$, $\forall i \neq j$, are initialized with zero-vectors. After the initialization, the observed signals $x_i[k]$ are fed through the filters $\mathbf{F}_j[k]$, which is the concatenation vector by several filters $\mathbf{f}_{ij}[k]$, and generated the two outputs $y_j[k]$ where $j = 1, 2$ because two sources were transmitted. In this case, this process can be simulated by executing Eq. 4.3 and the filter values $\mathbf{F}_j[k]$ are iteratively updated for every time k through computing the filter update equation in Eq. 4.6. Due to the reason that the steady state MSE of the CMA was done in [54], we will not pursue the first equalizer output, which exploits the CMA. Therefore, we only focus on the second equalizer where $j = 2$. In this case, the second equalizer $\mathbf{F}_j[k]$ where $j = 2$ is updated by CI-CMA and thus the filter update equation in Eq. 4.6 has the CI-CMA errors that were defined in Equations 4.7 and 4.8. As a result, the equalizer output 2, $y_j[k]$ where $j = 2$ can be obtained.

Regarding to the performance measurement, the equalizer output 2, $y_j[k]$ where $j = 2$, is used to compute the decision-directed MSE as follows.

$$\mathrm{M}_{\mathrm{DD},j}[k] = 0.99\mathrm{M}_{\mathrm{DD},j}[k-1] + 0.01\left|y_j[k] - \hat{y}_j[k]\right|^2 \tag{4.43}$$

where a symbol within the 16-QAM alphabet set $\mathbf{A} = \{\pm 1 \pm j, \pm 3 \pm j, \pm 1 \pm 3j, \pm 3 \pm 3j\}$ which has the minimum Euclidean distance with $y_j[k]$ is selected to be $\hat{y}_j[k]$. Furthermore, the initial value of MSE is assigned to be $\mathrm{M}_{\mathrm{DD},j}[0] = 1$. Because of we are interested on the steady state of MSE but not the whole series of MSE, only the final value of MSE, $\mathrm{M}_{\mathrm{DD},j}[k]$ where time $k = N_{seq}$, is selected and plotted. In other word, only one MSE value is obtained in a simulation. Instead of

that, the theoretical values that are produced by CI-CMA (i.e. Eq. 4.41) and previous work (i.e. Eq. 4.42) are plotted

Regarding to the simulation parameters, since 16-QAM alphabet set $\mathbf{A} = \{\pm 1 \pm j, \pm 3 \pm j, \pm 1 \pm 3j, \pm 3 \pm 3j\}$ is applied, the statistics of this source such as sixth order moment, fourth order moment and variance have the value of $m_6 = \mathrm{E}\{|a_j[k]|^6\} = 1960$, $m_4 = \mathrm{E}\{|a_j[k]|^4\} = 132$ and $\sigma_{\mathrm{A}}^2 = \mathrm{E}\{|a_j[k]|^2\} = 10$, respectively. Instead of that, the step size of equalizer 1, the length of channel spread value and the length of filter are assigned to be $\mu_1 = 10^{-6}$, $\delta_{\max} = 10$ and $L = 9$.

We are interested to investigate the steady state MSE performance by the various values of mixing parameter, k_j and step size, μ_j in the equalizer 2, $j = 2$. Therefore, $\mu_j = 1 \times 10^{-7}$ and $k_j = 1.41$ are firstly selected and then the simulation is performed as described as in the first paragraph and eventually a result MSE value is obtained. The simulations are repeated for $k_j = 1.46$ and $k_j = 1.51$, while the other parameters are unchanged. The MSE results versus the correspond mixing parameter, k_j are plotted as Fig. 4.2.

Next, we set $k_j = 1.41$ and aim to test various values of step size μ_j in the equalizer 2. The step sizes values of 0.75×10^{-7}, 1×10^{-7}, 1.25×10^{-7} and 1.5×10^{-7} are selected. The simulations are performed and the correspond MSE values are plotted in Fig. 4.3.

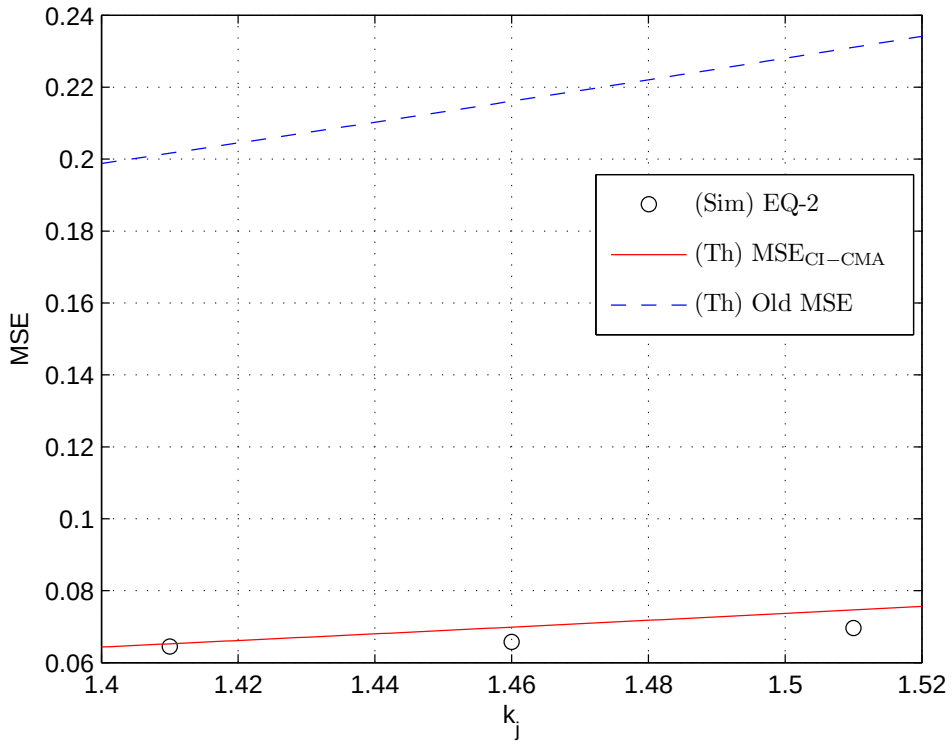

Figure 4.2: Simulation and theoretical curves for the steady state MSE as a function of mixing parameter, k_j with $\mu_j = 1 \times 10^{-7}$ from a 16-QAM constellation for equalizer 2, $j = 2$.

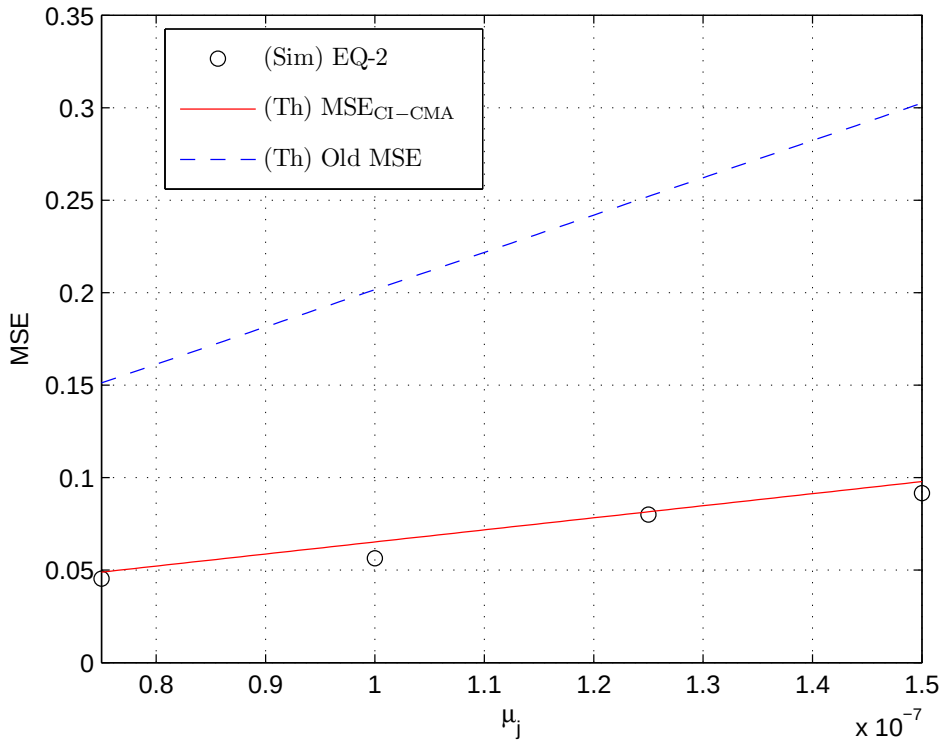

Figure 4.3: Simulation and theoretical curves for the steady state MSE as a function of step size, μ_j with mixing parameter, $k_j = 1.41$ from a 16-QAM constellation for equalizer 2, $j = 2$.

4.5.2 Discussion

Fig. 4.2 illustrates the steady state MSE versus mixing parameter k_j. The steady state simulation result from the second equalizer output is called "(Exp) EQ-2". Our CI-CMA theoretical MSE is called "(Th) $\text{MSE}_{\text{CI}-\text{CMA}}$" (i.e. Eq. 4.41) and the theoretical MSE by the previous work is called "(Th) Old MSE" (i.e. Eq. 4.42). As predicted, CI-CMA theoretical MSE follows closely to the simulation MSE of the equalizer 2. From the figure, we notice that the gap between CI-CMA theoretical result and the simulation result is not larger than 0.01 for the range from 1.4 to 1.52. The theoretical MSE by the previous work shows rather poor match with large gap between theory and simulation. This is not surprising as their MSE was derived with the inappropriate estimation of the CC term. In 16-QAM case, the value of the mixing parameter, k_j must be larger than 1.40

in order to ensure source separation ability. From the figure, we can see that the steady state MSE can be increased as k_j increasing. Therefore, we suggest that $k_j = 1.41$ should be chosen in order to ensure source separation ability and also remain the lowest steady state MSE value.

Next, Fig. 4.3 presents the steady state MSE versus step size of the equalizer 2, μ_j. The figure shows CI-CMA theoretical MSE follows closely to the simulation MSE of the equalizer 2 and a poorer match with the MSE of Eq. 4.42. Large value of step size is helpful to increase the convergence rate, but it has been shown in the figure that it increases the MSE value as well. This reveals the limitation of CI-CMA because the high convergence rate and low steady state MSE cannot be achieved at the same time.

Chapter 5

Hybrid Algorithms for MIMO Equalization

5.1 Introduction

High reliability communication link can be achieved through the reduction of the error rate at the receiver. However, in the presence of noises, ISI and CCI, the error reduction is a challenging task for the receiver. Therefore, we propose hybrid algorithms to improve the performance of MIMO blind equalization algorithm. The hybrid algorithms are constructed by some blind algorithms. In this chapter, we discuss the way of the hybrid algorithms have been constructed and then validate the hybrid algorithms through simulations.

The chapter is structured as follows, In Section 5.2, the system model is briefly presented. In Section 5.3, the general ways to hybrid the algorithms in MIMO equalization are shown. The performance comparison among the hybrid algorithms and the previous works are presented in Section 5.4 with discussions in Section 5.4.4.

5.2 System model and assumptions

We re-use the system model and assumptions in Section 2.8 and some parameters in the assumptions are defined as follows.

A1 All sources, $a_j[k]$, $j \in (1, M_{\mathrm{t}})$ are uniformly selected from 16 QAM alphabet set, $\mathbf{A}$ which is defined as

$$\mathbf{A} = \{\pm 1 \pm \mathbf{j}1, \mp 1 \pm \mathbf{j}1, \pm 1 \pm \mathbf{j}3, \mp 1 \pm \mathbf{j}3, \pm 3 \pm \mathbf{j}1, \mp 3 \pm \mathbf{j}1, \pm 3 \pm \mathbf{j}3, \mp 3 \pm \mathbf{j}3\} \tag{5.1}$$

and therefore $a_j[k]$ has a finite power, $\sigma_{\mathrm{A}}^2 = \mathrm{E}\{|a_j[k]|^2\} = 10$ and fourth order moment, $m_4 = \mathrm{E}\{|a_j[k]|^4\} = 132$.

A2 The impulse response of $\mathbf{h}_{ij}$ is shown in Table 5.1 [50] where $M_{\mathrm{t}} = 2$ and $M_{\mathrm{r}} = 3$.

A3 The zero mean complex value additive white Gaussian noise (AWGN), $n_i[k]$ has a constant variance, $\mathrm{E}\{|n_i[k]|^2\} = \sigma^2 = 0.001$ at the i-th channel output.

Again, we recall that the observed signal at the i-th sensor, for $i \in (1, M_{\mathrm{r}})$, at time k is defined as:

$$x_i[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij} \otimes \mathbf{a}_j[k] + n_i[k] = \sum_{j=1}^{M_{\mathrm{t}}} \mathbf{h}_{ij}^T \mathbf{a}_j[k] + n_i[k] \tag{5.2}$$

where $\mathbf{a}_j[k] = \big[a_j[k], \cdots, a_j[k - N_h + 1]\big]^T$. The equalizer output corresponding to the j-th source, where $j \in (1, M_{\mathrm{t}})$, at time k is denoted as:

$$y_j[k] = \sum_{i=1}^{M_{\mathrm{r}}} \mathbf{f}_{ij}^T[k] \mathbf{x}_i[k] \tag{5.3}$$

Table 5.1: Impulse response of two-input/three-output complex channel [1]

$\mathbf{h_{11}}$	$[-0.6 + 0.4\mathbf{j}, 1.2 - 0.2\mathbf{j}]^T$
$\mathbf{h_{12}}$	$[-0.6 + 0.8\mathbf{j}, 0.9 - 0.1\mathbf{j}]^T$
$\mathbf{h_{21}}$	$[0.1 + 0.7\mathbf{j}, -0.2 - 0.5\mathbf{j}]^T$
$\mathbf{h_{22}}$	$[0.4 - 0.3\mathbf{j}, -0.2 + 0.2\mathbf{j}]^T$
$\mathbf{h_{31}}$	$[0.5 + 0.4\mathbf{j}, -1 + 0.3\mathbf{j}]^T$
$\mathbf{h_{32}}$	$[-0.1 + 0.8\mathbf{j}, 0.4 + 0.1\mathbf{j}]^T$

where $\mathbf{x}_i[k] = [x_i[k], \cdots, x_i[k-L+1]]^T$ is the input regressor vector for the equalizers of the i-th seonsor, $L \in \mathbb{Z}^+$ denotes the equalizer length, and $\mathbf{f}_{ij}[k] = \left[f_{ij}^{[0]}[k], \cdots, f_{ij}^{[L-1]}[k]\right]^T$ is a $L \times 1$ vector that represents the equalizer connecting the observed signal at the i-th sensor to the j-th output at time k. There are a total of $(M_\mathrm{t} \times M_\mathrm{r})$ equalizers at the receiver end (see Fig. 5.1). Furthermore, (5.3) can be compactly written as

$$y_j[k] = \mathbf{F}_j^T[k]\mathbf{X}[k] \tag{5.4}$$

where $\mathbf{F}_j[k] = \left[\mathbf{f}_{1j}^T[k], \cdots, \mathbf{f}_{M_\mathrm{r}j}^T[k]\right]^T$ is a $M_\mathrm{r}L \times 1$ vector and the input regressor $\mathbf{X}[k] = \left[\mathbf{x}_1^T[k], \cdots, \mathbf{x}_{M_\mathrm{r}}^T[k]\right]^T$ has the same dimensions too.

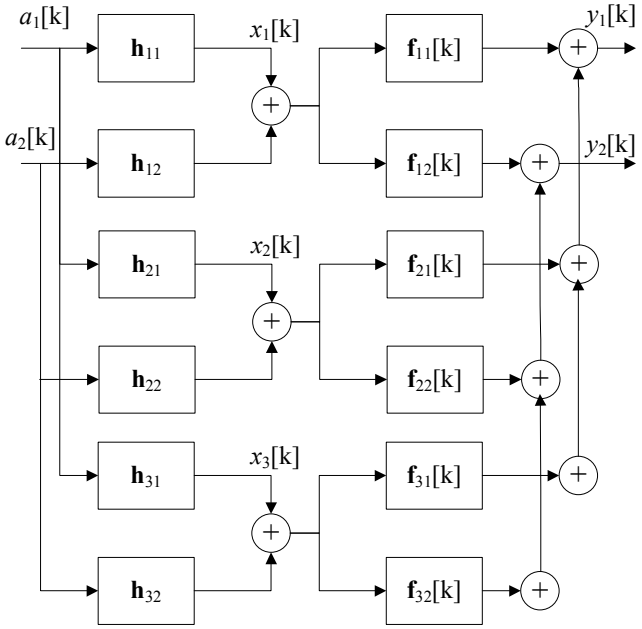

Figure 5.1: Baseband equivalent system for $M_t = 2$ and $M_r = 3$.

5.3 Blind Adaptive Hybrid Algorithms For MIMO Systems

A hybrid algorithm that is employed by the MIMO linear equalizer is called that way because it opens the channel eye using an acquisition algorithm, a new Adaptive Constant Modulus Algorithm (ACMA), tracks the channel (after the channel eye is open) using a decision directed (DD) algorithm, and perform source separation algorithm using a cross correlation function (CC). The acquisition and source separation algorithm can open the channel eye and ensure the outputs are mutually uncorrelated but usually yields large steady state errors. That is why hybrid algorithms couple the acquisition and source separation algorithms with a tracking algorithm (e.g. DD algorithm) to exploit the discrete nature of the input data to achieve much lower error rates. To the best of the author's knowledge, such algorithms have not been purposed for the MIMO system, even though a similar idea can be found in SISO system. We emphasize that the migration of some hybrid algorithms from the SISO to MIMO systems, although may seem straight forward initially, often results in divergence unless a very small step size (which is undesirable because it slows convergence) is used. To overcome the issue, we will propose a new acquisition algorithm. Instead of that, the past SISO approaches are not equipped with source separation algorithm, therefore, the past approaches cannot perform source separation. Our approach in this chapter is to propose the new hybrid algorithms that can open the channel eye, performance source separation and produce significant low steady state MSE.

In what follows, we will provide a general cost function that we employ in our hybrid algorithms. The update equations are also provided thereafter.

5.3.1 General cost function

In this section, we define the general cost function of hybrid algorithm and then present the stochastic gradient descent algorithm (SGA).

For simplicity, we consider two sources (i.e. $M_t = 2$) case. $J_1[k]$ and $J_2[k]$ are denoted as the cost functions of the equalizer output 1 and 2 respectively. Furthermore, the j-th cost function,

$J_j[k]$ where $j = 1, 2$ is defined as

$$J_j[k] \quad = \quad \frac{1}{2}\mathrm{E}\left\{|\epsilon_j[k]|^2\right\} \tag{5.5}$$

where ϵ_j is known as equalization error function that is a function of j-th equalizer output, $y_j[k]$.

In practice, we use SGA to minimize the cost function 5.5. As a result, the equalizer coefficients after dropping E{.} are updated as below:

$$\mathbf{F}_j[k+1] = \mathbf{F}_j[k] - \mu_j \epsilon_j[k]\mathbf{X}^*[k] \tag{5.6}$$

where μ_j is a suitable step size value for the j-th equalizer output and $\mathbf{X}^*[k] = \left[\mathbf{x}_1^T[k], \mathbf{x}_2^T[k], ..., \mathbf{x}_M^T[k]\right]^H$ is a $M_r L \times 1$ vector. Finally, $y_j[k]$ can be obtained as Eq. 5.4.

We emphasize that the main distinction among the hybrid algorithms is the error function, $\epsilon_j[k]$ in Eq. 5.6. Therefore, we will concentrate on $\epsilon_j[k]$ in the below section.

5.3.2 Acquisition, Source Separation and Tracking algorithms

The acquisition, source separation and tracking algorithms are the single task algorithms that will be combined as hybrid algorithms. Therefore, we describe the single task algorithms before the description of hybrid algorithms.

In order to open the channel eye, the acquisition algorithm must be employed. In this case, we propose a new acquisition algorithm, which is call adaptive constant modulus algorithm (ACMA) as the acquisition algorithm. ACMA is different with the conventional CMA because adaptive dispersion value is used in ACMA instead of fixed value. The error functions of acquisition (i.e. $e_{\mathrm{acq},j}[k]$) and ACMA (i.e. $e_{\mathrm{ACMA},j}[k]$) for j-th equalizer are defined as

$$e_{\mathrm{acq},j}[k] = e_{\mathrm{ACMA},j}[k] = (|y_j[k]|^2 - \mathrm{R_j[k]})y_j[k] \tag{5.7}$$

where the adaptive dispersion is defined as

$$\mathrm{R_j}[k+1] \quad = \quad \mathrm{R_j}[k] - 0.001\big(|y_j[k]|^2 - \sigma_{\mathrm{A}}^2\big) \tag{5.8}$$

with the initial value $R_j[0] = 1$. The main purpose of the adaptive dispersion constant is to maintain the power of the j-th equalizer output and avoid the output to be shrunk.

Next, in order to perform source separation, source separation algorithm is required to be applied. The cross-independent (CI) function is suggested to be the source separation algorithm. In this approach, the cross-correlation function is used to ensure the interdependency among the equalizer outputs. The error functions of the source separation algorithm (i.e. $e_{\mathrm{ss},j}[k]$) and CI function (i.e. $e_{\mathrm{CI},j}[k]$) for the j-th equalizer are defined as

$$e_{\mathrm{ss},j}[k] = e_{\mathrm{CI},j}[k] = k_j \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} |y_1[k-\delta]|^2 \, y_2[k] \tag{5.9}$$

where $\delta_{\max}$ is the length of channel spread and k_j is the mixing parameter of the j-th equalizer. Due to the reason that *hierarchical* approach will be used in the simulations, k_j for equalizer 1 (where $j = 1$) is always 0, whilst k_j for equalizer 2 (i.e. k_2, where $j = 2$) has the value of 1.41. In this case, the equalizer 1 does not have the source separation ability while equalizer 2 has the source separation ability.

The tracking algorithm is important to reduce the steady state MSE. Decision directed (DD) algorithm is selected as the tracking algorithm to achieve this. The error functions of tracking algorithm and DD algorithm for the j-th equalizer are defined as

$$e_{\mathrm{tr},j}[k] = e_{\mathrm{DD},j}[k] = \hat{y}_j[k] - y_j[k] \tag{5.10}$$

where $\hat{y}_j[k]$ for 16-QAM is defined as

$$\mathrm{Re}\{\hat{y}_j[k]\} = \begin{cases} -3 & \text{if } \mathrm{Re}\{y_j[k]\} \leq -2 \\ -1 & \text{if } -2 < \mathrm{Re}\{y_j[k]\} \leq 0 \\ 1 & \text{if } 0 < \mathrm{Re}\{y_j[k]\} \leq 2 \\ 3 & \text{otherwise} \end{cases} \tag{5.11}$$

and

$$\mathrm{Im}\{\hat{y}_j[k]\} = \begin{cases} -3 & \text{if } \mathrm{Im}\{y_j[k]\} \leq -2 \\ -1 & \text{if } -2 < \mathrm{Im}\{y_j[k]\} \leq 0 \\ 1 & \text{if } 0 < \mathrm{Im}\{y_j[k]\} \leq 2 \\ 3 & \text{otherwise} \end{cases} \tag{5.12}$$

where $\mathrm{Re}\{y_j[k]\}$ and $\mathrm{Im}\{y_j[k]\}$ are the real and imaginary part of $y_j[k]$.

5.3.3 MIMO Hybrid algorithms

5.3.3.1 Modified Cross Independent Benveniste-Goursat Algorithm (MCIBG)

We propose Modified Cross Independent Benveniste-Goursat Algorithm (CIBG) by modified the conventional Benveniste-Goursat Algorithm (BG) [125] from SISO system with new ACMA and CI. This is aimed to open the channel eye, perform source separation and control the output power. In this case, the error function of MCIBG for the j-th equalizer is defined as

$$\epsilon_j[k] = \beta_{\mathrm{BG},1}\left|e_{\mathrm{tr},j}[k]\right| \left(e_{\mathrm{acq},j}[k] + e_{\mathrm{ss},j}[k]\right) + \beta_{\mathrm{BG},2} e_{\mathrm{tr},j}[k] \tag{5.13}$$

where $\beta_{\mathrm{BG},1} = 1$ and $\beta_{\mathrm{BG},2} = 10$ are the weights of this hybrid algorithm.

5.3.3.2 Modified Cross Independent Stop and Go Algorithm (MCISAG)

We propose the second hybrid algorithm by merging the conventional Stop and Go algorithm from [124] with the acquisition algorithm and source separation algorithm. In this way, we improve the conventional Stop and Go algorithm, which can conventionally reduce the steady state MSE, with new features on the source separation and MIMO equalization abilities. We call the new hybrid algorithm as Modified Cross Independent Stop and Go Algorithm (MCISAG) and the error function of the j-th equalizer is separated into real and imaginary parts as

$$\epsilon_j[k] = \mathrm{Re}\{\epsilon_j[k]\} + \mathbf{j}\mathrm{Im}\{\epsilon_j[k]\} \tag{5.14}$$

where $\mathbf{j} = \sqrt{-1}$. The real and imaginary parts of $\epsilon_j[k]$ are defined as follows

$$\mathrm{Re}\{\epsilon_j[k]\} = \begin{cases} \mathrm{Re}\{e_{\mathrm{tr},j}[k]\} & \text{if } \mathrm{sgn}\left(\mathrm{Re}\{e_{\mathrm{tr},j}[k]\}\right) = \mathrm{sgn}\left(\mathrm{Re}\{e_{\mathrm{acq},j}[k] + e_{\mathrm{ss},j}[k]\}\right) \\ 0 & \text{otherwise} \end{cases} \tag{5.15}$$

and

$$\mathrm{Im}\{\epsilon_j[k]\} = \begin{cases} \mathrm{Im}\{e_{\mathrm{tr},j}[k]\} & \text{if } \mathrm{sgn}\left(\mathrm{Im}\{e_{\mathrm{tr},j}[k]\}\right) = \mathrm{sgn}\left(\mathrm{Im}\{e_{\mathrm{acq},j}[k] + e_{\mathrm{ss},j}[k]\}\right) \\ 0 & \text{otherwise} \end{cases} \tag{5.16}$$

where sgn(x) is a signum of x, for example, sgn(x) is 1 if $x > 0$, sgn(x) is 0 if $x = 0$ and sgn(x) is -1 if $x < 0$.

5.3.3.3 Modified Cross Independent Reliability Based Algorithm (MCIRBA)

Lastly, we propose the third hybrid algorithm by merging the SISO RBA from [126] with source separation and MIMO equalization abilities. We call this new algorithm as Modified Cross Independent Reliability Based Algorithm (CIRBA). The error function of CIRBA for the j-th equalizer is defined as

$$\epsilon_j[k] = (1 - \alpha_2[k])\left(e_{\mathrm{acq},j}[k][k] + e_{\mathrm{ss},j}[k]\right) + \alpha_2[k]\gamma_{\mathrm{RBA}} e_{\mathrm{tr},j}[k] \tag{5.17}$$

$\gamma_{\mathrm{RBA}} = 15$ is a weight and $\alpha_j[k]$ is defined as below:

$$\alpha_j[k] = 2P_{\mathrm{c},j}[k] - 1$$

where $P_{\mathrm{c},j}$ for 16-QAM is defined as below:

$$P_{\mathrm{c},j}[k] = \frac{\exp\left(\frac{-\left|\hat{y}_j[k] - y_j[k]\right|^2}{\sigma_j^2[k]}\right)}{\sum_{t=1}^{16} \exp\left(\frac{-\left|y_j[k] - A_t)\right|^2}{\sigma_j^2[k]}\right)} \tag{5.18}$$

where A_t is the tth constellation point in the 16-QAM alphabet set $\mathbf{A}$ and $\sigma_j^2[k]$ is the variance of $y_j[k]$. The estimation of $\sigma_j^2[k]$ in practice, $\hat{\sigma}_j^2[k]$ (with $\hat{\sigma}_j^2[0] = 10$) as below:

$$\hat{\sigma}_j^2[k] = 0.99\hat{\sigma}_j^2[k-1] + 0.01\left|\hat{y}_j[k] - y_j[k]\right|^2 \tag{5.19}$$

5.3.4 CI-CMA

CI-CMA was proposed in Chapter 3 and is re-stated here for the purpose of comparison. The error function of CI-CMA of j-th equalizer is stated as

$$\epsilon_j[k] = e_{\mathrm{CMA},j}[k] + k_j e_{\mathrm{CI},j}[k] \tag{5.20}$$

$$e_{\mathrm{CMA},j}[k] = (|y_j[k]|^2 - \mathrm{R_j})y_j[k] \tag{5.21}$$

where $\mathrm{R_j}$ is the dispersion constant for the j-th output is defined as

$$\mathrm{R}_j = \frac{m_4}{\sigma_{\mathrm{A}}^2} + k_j(2\delta_{\max} + 1)\sigma_{\mathrm{A}}^2 \tag{5.22}$$

$$e_{\mathrm{CI},j}[k] = \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} |y_1[k-\delta]|^2 \, y_2[k] \tag{5.23}$$

where k_j is the mixing parameter where $k_j = 0$ for equalizer 1 (i.e. $j = 1$) and $k_j = 1.41$ for equalizer 2 (i.e. $j = 2$) due to *hierarchical* approach.

5.3.5 Previous works

In order to compare the simulation results, the error functions of the previous work are stated here.

5.3.5.1 MIMO-CMA

MIMO-CMA is a classical MIMO blind equalization algorithm that can open the channel eye but cannot perform source separation [35] [36]. The error function of MIMO-CMA, is the conventional CMA, for the j-th equalizer is stated as follows

$$\epsilon_j[k] = e_{\mathrm{CMA},j}[k] = (|y_j[k]|^2 - \mathrm{R})y_j[k] \tag{5.24}$$

where R is the dispersion constant that is defined as

$$\mathrm{R} = \frac{m_4}{\sigma_{\mathrm{A}}^2} \tag{5.25}$$

5.3.5.2 CC-CMA

Cross Correlation Constant Modulus Algorithm (CC-CMA) is a conventional MIMO equalizer algorithm with source separation ability that is proposed [2, 31, 43]. The error function for the j-th equalizer is defined as

$$\epsilon_j[k] \quad = \quad (|y_2[k]|^2 - \mathrm{R})y_2[k] + k_j \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \hat{r}_{1,\delta}[k]y_1[k-\delta] \tag{5.26}$$

where R is the dispersion constant that is defined as

$$\mathrm{R} = \frac{m_4}{\sigma_{\mathrm{A}}^2} \tag{5.27}$$

and k_j is the mixing parameter where $k_j = 0$ for equalizer 1 (i.e. $j = 1$) and $k_j = 1.41$ for equalizer 2 (i.e. $(j = 2)$ due to *hierarchical* approach. Furthermore, $\hat{r}_{1,\delta}[k]$ is an estimate of $r_{1,\delta}[k] \equiv \mathrm{E}\{y_2[k]y_1^*[k-\delta]\}$, which may be recursively computed as

$$\hat{r}_{1,\delta}[k] = \lambda_{\mathrm{CC}}\hat{r}_{1,\delta}[k-1] + (1-\lambda_{\mathrm{CC}})y_2[k]y_1^*[k-\delta]. \tag{5.28}$$

where $\lambda_{\mathrm{CC}} = 0.9$ is the weight.

5.4 Simulations

5.4.1 Simulation Setup

In order to simulate the system model in Section 5.2 and Fig. 5.1 with ($M_{\mathrm{t}} = 2$ and $M_{\mathrm{r}} = 3$) MIMO channel, simulations are performed using MATLAB. Two random 16-QAM symbol sequences that have $N_{seq} = 3.8 \times 10^5$ symbols in one sequence are generated as source 1, $a_1[k]$ and source 2, $a_2[k]$. The sources are convoluted with the memory channels $\mathbf{h}_{ij}$, and this process is equivalent to $\mathbf{h}_{ij} \otimes \mathbf{a}_j[k]$ where $j = 1, 2$. This is aimed to simulate the condition that the source symbols are passed through the memory channels and thereby the symbols are interfered by ISI. At the i-th received sensor, the ISI signals from two sources $a_1[k]$ and $a_2[k]$ (but not a single source) have been observed, thus the observed signal at i-th sensor $x_i[k]$ is corrupted by CCI due to receiving

two sources and ISI due to memory channels. Instead of that, a random zero mean complex value additive white Gaussian noises (AWGN), $n_i[k]$ that have the power of $\mathrm{E}\{|n_i[k]|^2\} = \sigma^2 = 0.001$ are generated and summed with the the observed signal at i-th sensor $x_i[k]$. In this cases, the observed signal at i-th sensor $x_i[k]$ is not only corrupted by interferences, but also is distorted by noises. This scenario can be simulated through computing Eq. 5.2 and the impulse responses of channels $\mathbf{h}_{ij}$ are stated in Table 5.1.

Next, to simulate the equalization process, we will test using the hierarchical approach, i.e., the first equalizer is allowed to recover any of the two sources (usually the stronger one) and the second equalizer should be able to avoid the extracted source by the first equalizer if BSS works. In this case, equalizer 1 uses the algorithm without source separation ability, whilst equalizer 2 uses the algorithm that has source separation ability (except the case of MIMO-CMA algorithm).

The filter weights $\mathbf{f}_{11}[0]$ and $\mathbf{f}_{22}[0]$ are initialized with a center spike [36] whilst $\mathbf{f}_{ij}[0]$, $\forall i \neq j$, are initialized with zero-vectors. After the initialization, the observed signals $x_i[k]$ are fed through the filters $\mathbf{F}_j[k]$, which is the concatenation vector by several filters $\mathbf{f}_{ij}[k]$, and generated the two outputs $y_j[k]$ where $j = 1, 2$ because two sources were transmitted. This can be achieved executing the output equation in Eq. 5.4. At the same time, the filter vector $\mathbf{F}_j[k]$ are iteratively updated for every time k through computing the filter update equation in Eq. 5.6.

In order to simulate an algorithm, the general error function, $\epsilon_j[k]$ that is located inside the filter update equation (i.e. Eq. 5.6) is replaced by the correspond error function of the algorithm. Therefore, to perform simulation on MIMO-CMA, we uses the error function in Eq. 5.24 to update the filter update equation (i.e. Eq. 5.6). Similarly, the simulation is repeated for CC-CMA (i.e. Eq. 5.26), CI-CMA (i.e. Eq. 5.20), MCIBG (i.e. Eq. 5.13), MCIRBA (i.e. Eq. 5.17) and MCISAG (i.e. Eq. 5.14).

Regarding to the simulation parameters, the filter length $L = 3$ and channel spread length $\delta_{\max} = 4$ are assigned. The step sizes of equalizer 1 and 2 are empirically chosen. This means several simulations have been run to find the step sizes that can avoid divergence. The step size of

equalizer 1 for μ_1 and the step size of equalizer 2, μ_2 for various algorithms are listed in Table 5.2.

Table 5.2: Simulation Parameters [2]

Step sizes	MIMO-CMA	CC-CMA	CI-CMA	MCIBG	MCIRBA	MCISAG
μ_1 ($\times 10^{-5}$)	1	1	1	0.8	2	100
μ_2 ($\times 10^{-5}$)	0.4	0.4	0.4	0.4	1	100

5.4.2 Performance measurements

In order to measure the performance of the algorithms, MSE of the j-th equalizer output, $\text{MSE}_j[k]$ is computed as following,

$$\text{MSE}_j[k] = 0.99\text{MSE}_j[k-1] + 0.01\left|\hat{y}_j[k] - y_j[k]\right|^2 \tag{5.29}$$

where $\text{MSE}_j[0] = 1$ is initialized. Furthermore, MSE in decibel (dB) unit is defined as $\text{MSE}_j[k]$ (dB) $= 10\log_{10}\text{MSE}_j[k]$. We use the dB unit in the simulation to trace the small value of MSE. MATLAB is used to compute MSE performance.

However, the $\text{MSE}_j[k]$ only implies the amount of interference but cannot validate the source separation ability. In order to validate the source separation ability, we further require the information of the global impulse response, $s_{ij}^{[n]}[k]$ (Note: we use short notation $s_{ij}^{[n]}[k]$ here, where $s_{ij}^{[n]}[k] = S_{(i),j}^{[n]}$ and $S_{(i),j}^{[n]}$ was defined in Section 2.10), for $i = 1, 2$ and $j = 1, 2$ at the steady state. However, the value of $s_{ij}^{[n]}[k]$ probably can be a complex value or negative value. Therefore, to overcome the difficulty to plot complex number or negative number of $s_{ij}^{[n]}[k]$, we are mainly interested on the absolute $|s_{ij}^{[n]}[k]|$ and this represents the absolute impulse response that connects from source i to output j. Instead of that, the value of $|s_{ij}^{[n]}[k]|$ can fluctuate from time to time due to the non-zero error. To overcome this issue, we take the 1000 samples of $|s_{ij}^{[n]}[k]|$ and display the averaged of these samples. Therefore, the averaged global impulse responses at the steady state will be plotted in order to validate the source separation ability.

Moreover, the symbols that are produced by equalizer outputs 1 and 2 (i.e. $y_1[k]$ and $y_2[k]$) are plotted in order to identify the open eye condition and shrinking issue.

5.4.3 Simulation results

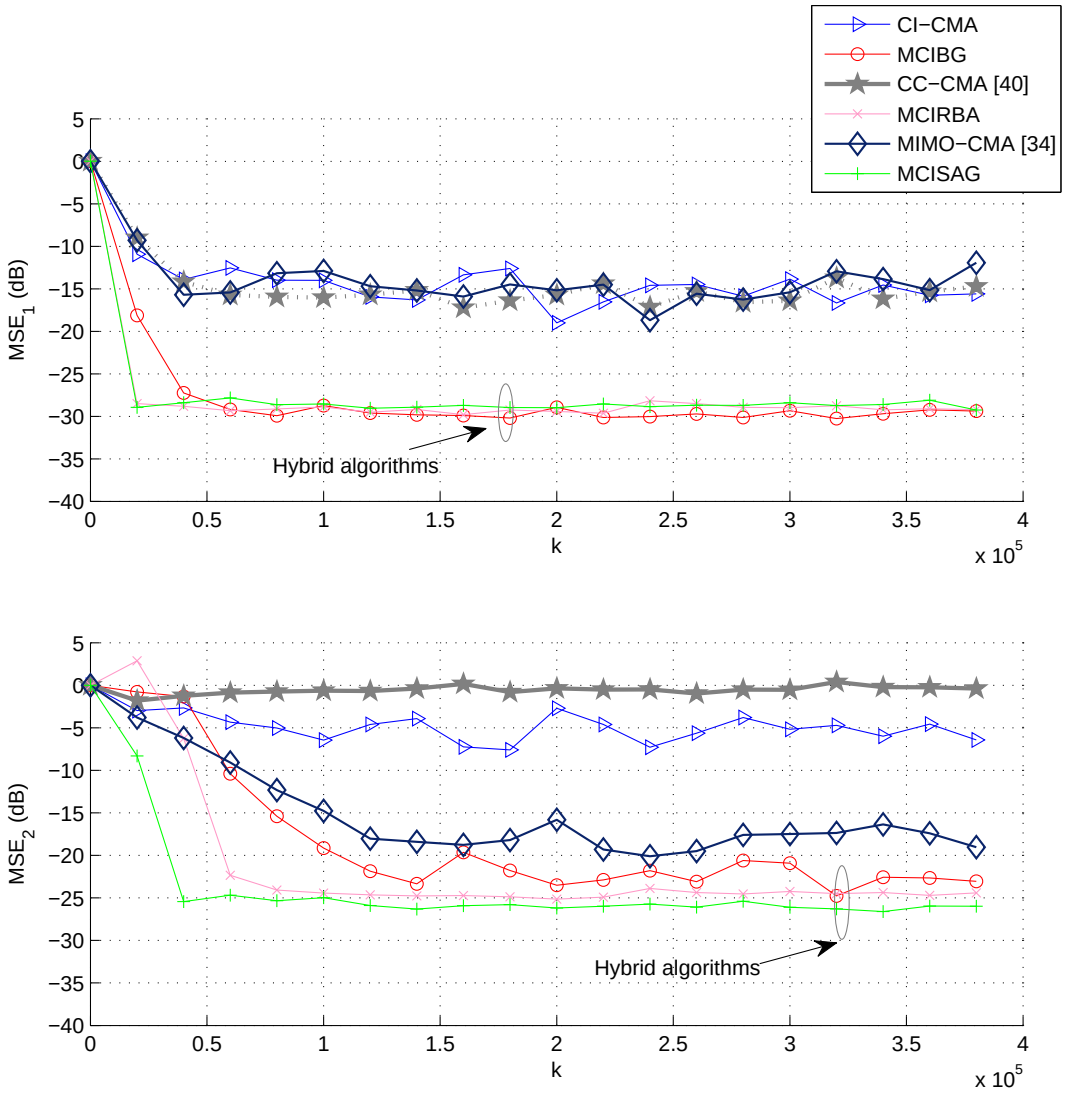

Figure 5.2: Performance comparison of different algorithms for equalizer 1 and 2. Hybrid algorithms provide superior performance than others.

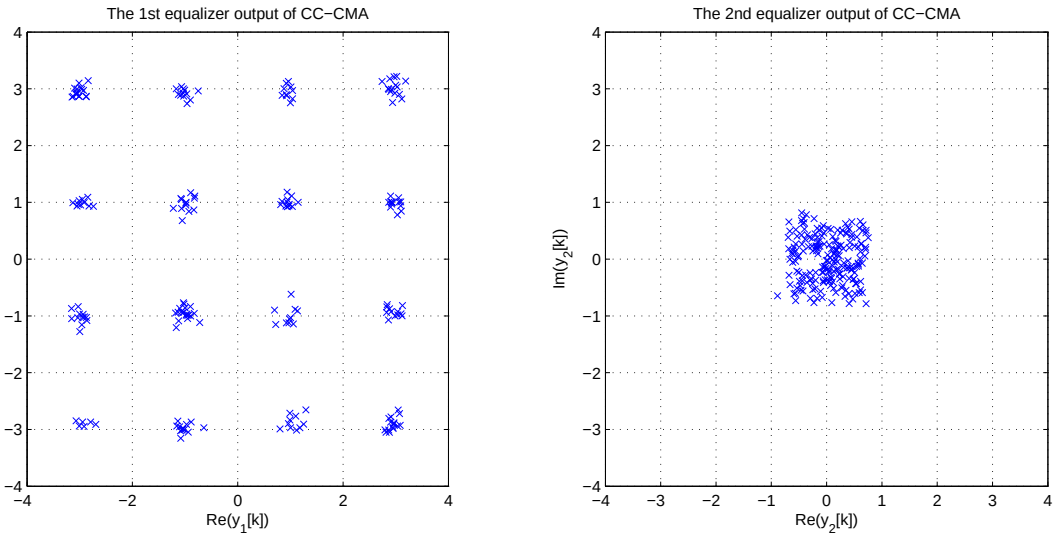

Figure 5.3: Equalizer outputs of CC-CMA at the steady state. The symbols closely match with the 16-QAM constellation set for equalizer 1 but the symbols of equalizer 2 have been shrunk.

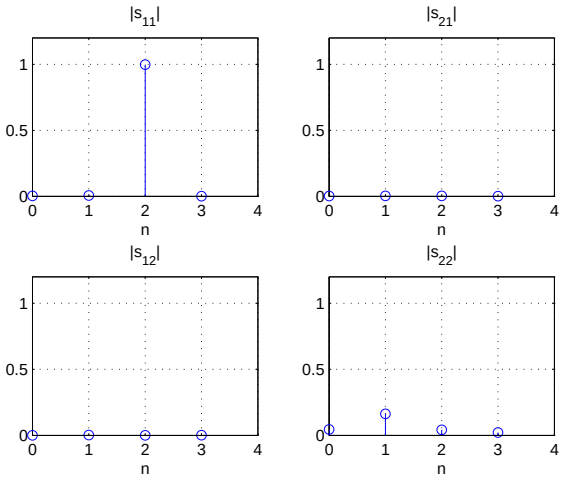

Figure 5.4: The averaged global impulse responses of CC-CMA taken in the steady state. Source separation is successful but a down-scaling of 0.2 has been introduced to the equalizer 2.

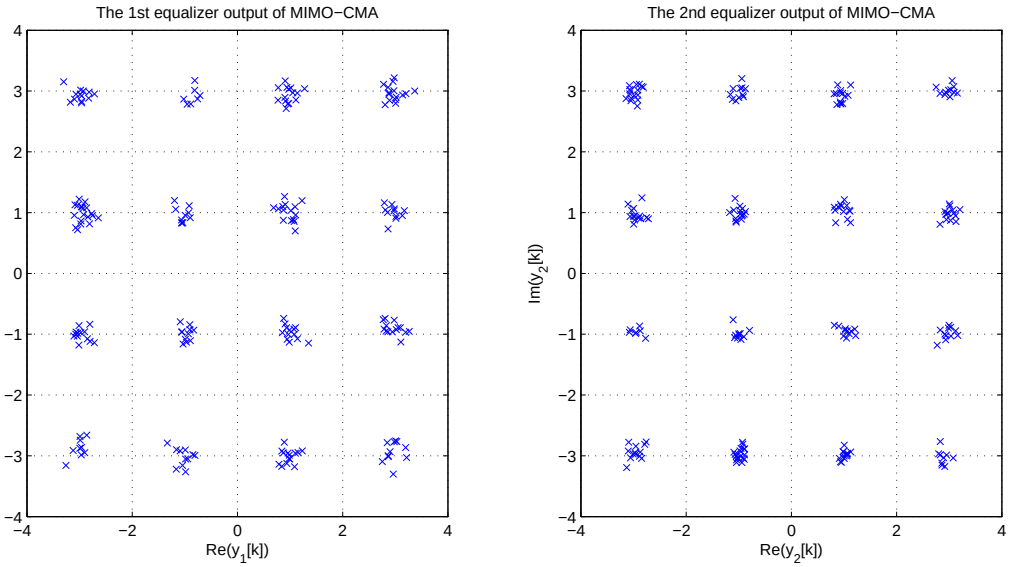

Figure 5.5: Equalizer outputs of MIMO-CMA at the steady state. The symbols closely match with the 16-QAM constellation set for equalizer 1 and 2

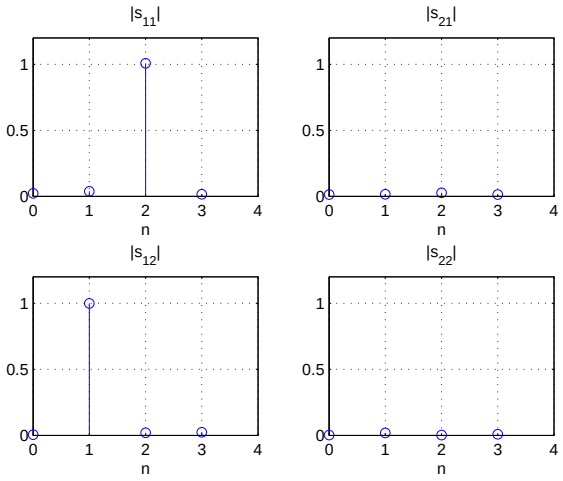

Figure 5.6: The averaged global impulse responses of MIMO-CMA taken in the steady state. Source separation is not successful.

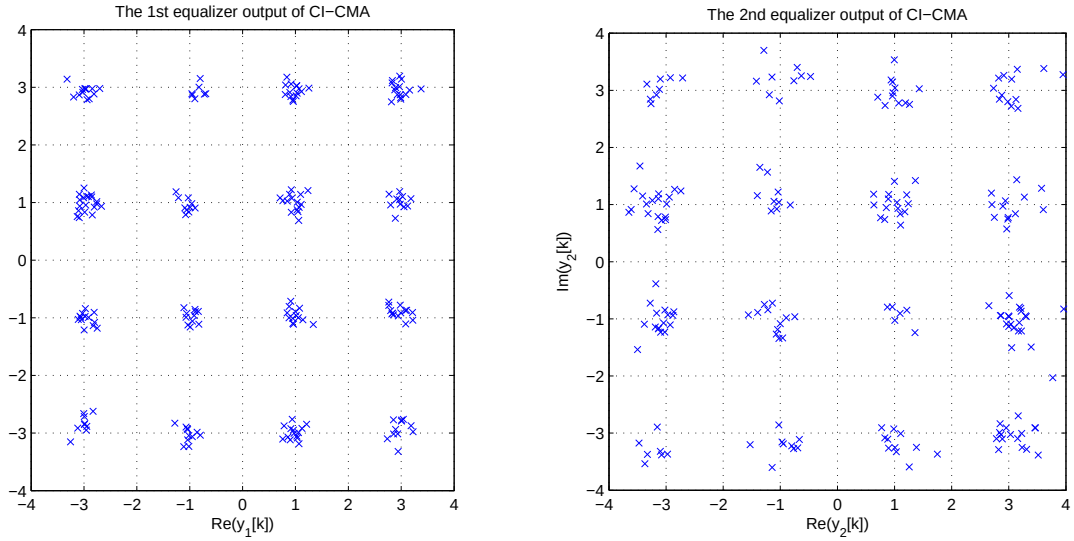

Figure 5.7: Equalizer outputs of CI-CMA at the steady state.

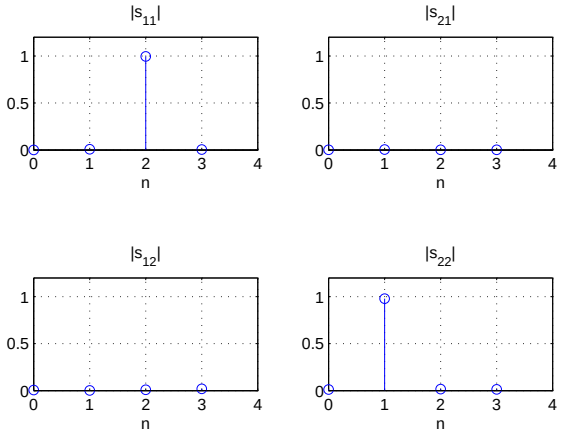

Figure 5.8: The averaged global impulse responses of CI-CMA taken in the steady state. Source separation is successful.

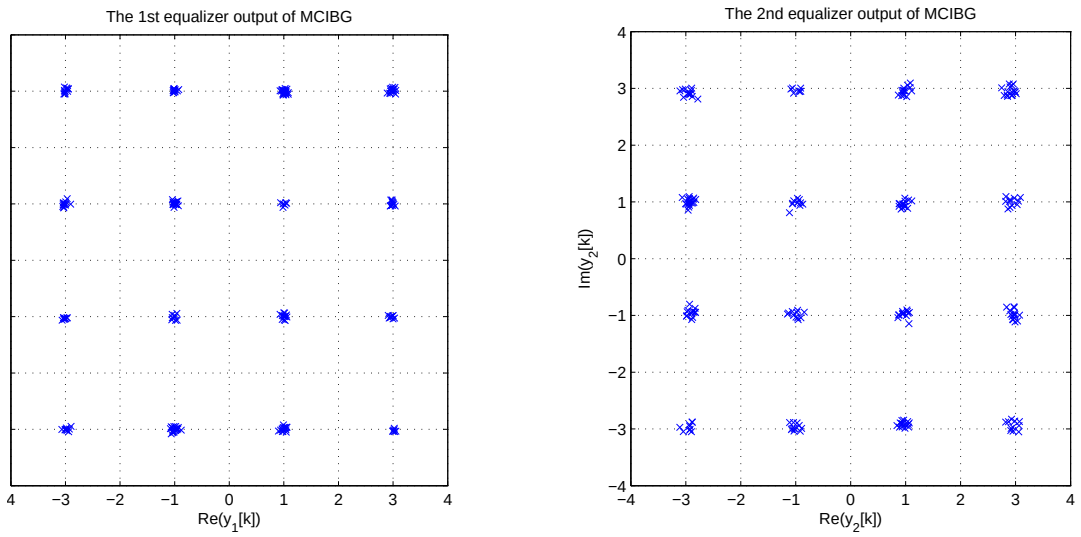

Figure 5.9: Equalizer outputs of MCIBG at the steady state.

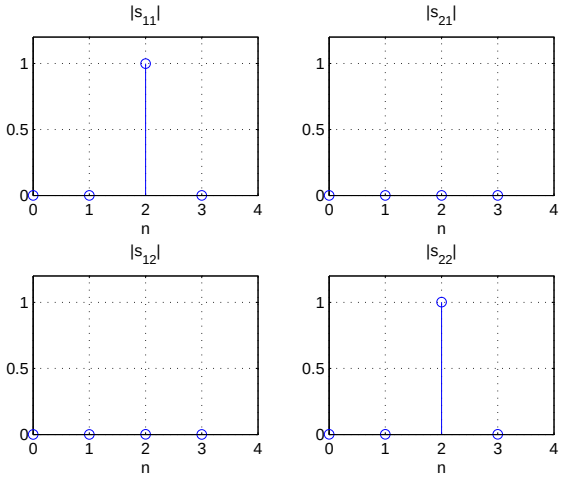

Figure 5.10: The averaged global impulse responses of MCIBG taken in the steady state. Source separation is successful.

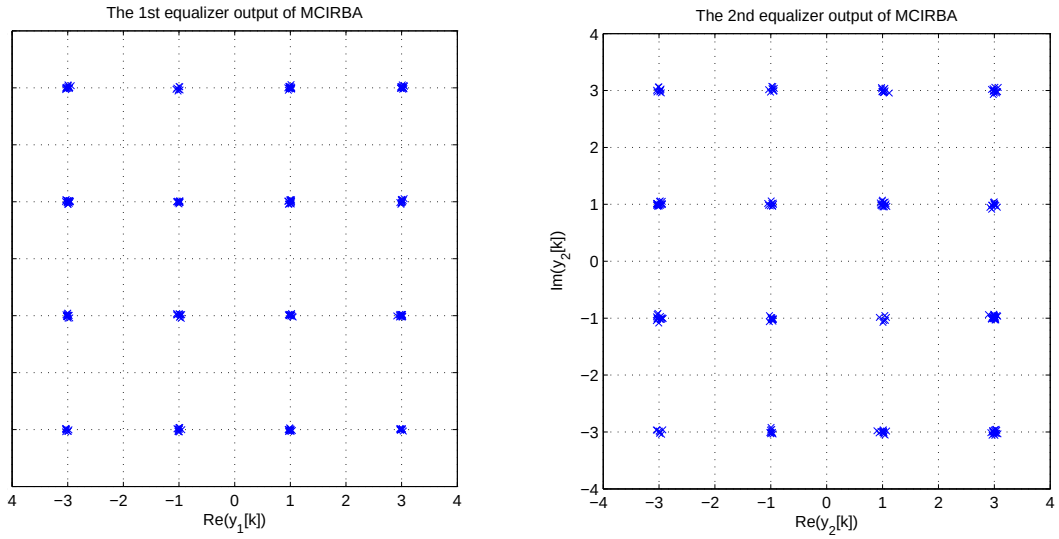

Figure 5.11: Equalizer outputs of MCIRBA at the steady state.

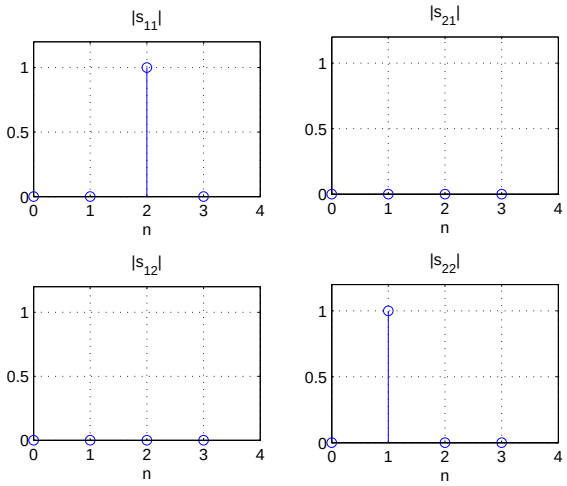

Figure 5.12: The averaged global impulse responses of MCIRBA taken in the steady state. Source separation is successful.

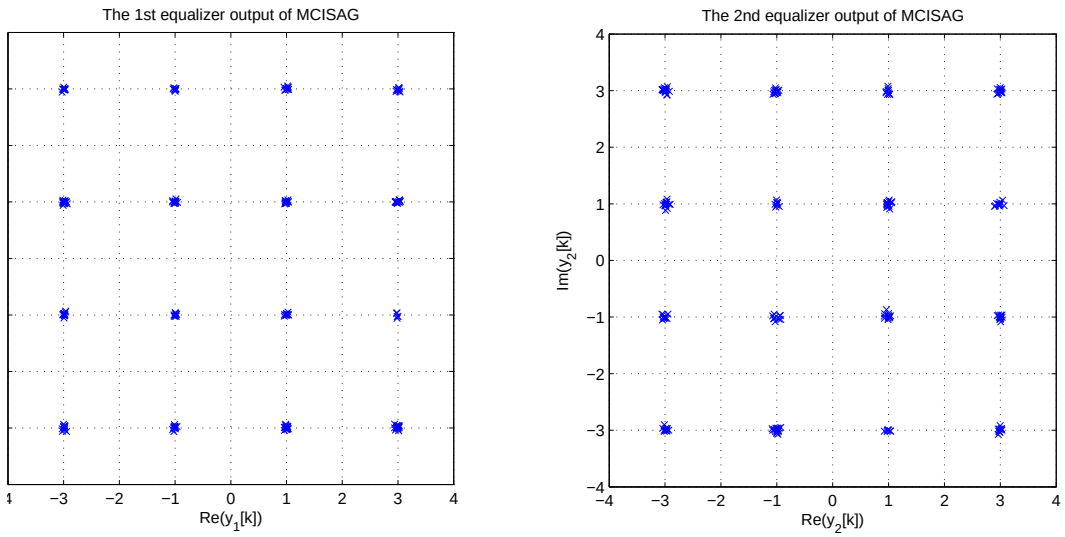

Figure 5.13: Equalizer outputs of MCISAG at the steady state.

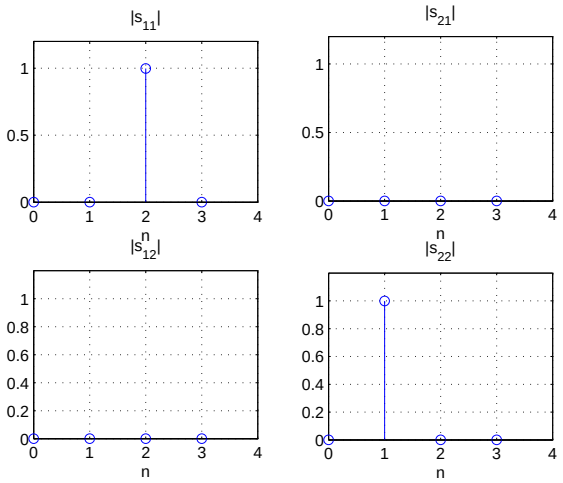

Figure 5.14: The averaged global impulse responses of MCISAG taken in the steady state. Source separation is successful.

5.4.4 Discussion

Fig. 5.2 illustrates the MSE performance of various algorithms versus time k for equalizer 1 and 2, MSE_1 and MSE_2, respectively. Before any description, we firstly need to identify the common steady state time range for equalizer 1 and 2. In the figure, we can see that after the time $k = 1.5 \times 10^5$, all of the the algorithms in equalizer 1 and 2 do not show any significant change in the MSE curves. Therefore, the common steady state time happens from the time $k = 1.5 \times 10^5$ to $k = 3.8 \times 10^5$. The equalizer outputs and the averaged global impulse responses in Figures 5.3 to 5.14 are taken in the steady state because the information of the system in the steady state is always accurate enough to represent the condition.

Obviously, in Fig. 5.2, we can see that the hybrid algorithms such as MCIBG, MCIRBA and MCISAG can provide lower steady state MSE compared to the non-hybrid algorithms for equalizer 1 and 2. Among the algorithms, MCISAG shows the best performance in term of convergence rate and steady state error in the equalizer 1 and 2. Large allowable step size without divergence in MCISAG causes MCISAG can reduce the error faster than any other algorithms. Instead of that, we can see that MCISAG can produce the sufficient low steady state MSE values which are below -25 dB for equalizer 1 and 2, and this implies MCISAG is more reliable than others. Now, we focus on the MSE of equalizer 1 (i.e. MSE_1) in Fig. 5.2, we can see that the non-hybrid algorithms, such as MIMO-CMA, CI-CMA and CC-CMA, show almost same MSE performance. This is because they use the same algorithms, which are CMA, due to hierarchical approach as explained. In contrast, the hybrid algorithms, which are equipped with tracking algorithm, show significant MSE reduction and thereby they produce superior performance.

Next, we put our focus on the MSE of equalizer 2 (i.e. MSE_2) in Fig. 5.2. We inspect the conventional algorithms, which are CC-CMA [2] and MIMO-CMA [36]. We can see the MSE of CC-CMA is high throughout all times, whilst the MSE of MIMO-CMA can be reduced and eventually reaches -20 dB. In order to find out the reason of CC-CMA shows poor performance, we can see the equalizer outputs of CC-CMA in Fig. 5.3 and the averaged global impulse response

in Fig. 5.8. In Fig. 5.3, we can see the outputs of equalizer 1 closely match with the 16-QAM constellation set, while the outputs of equalizer 2 has been shrunk and become vague. The shrunk equalizer output 2 explains the reason of poor MSE performance of CC-CMA. Instead of that, by observing Fig. 5.8, we can find that the global impulses of $|s_{11}|$ and $|s_{22}|$ have a spike. This implies that the equalizer 1 recovers source 1 and equalizer 2 recovers source 2, therefore the source separation task is fine. However, we further look into $|s_{22}|$ in the figure, we can see the highest magnitude of the impulse is 0.2 but not 1. Therefore, we can see that a down-scaling of 0.2 has been introduced by the algorithm and thereby causes the shrunk of equalizer output 2 and then produces poor MSE performance. This is because CC-CMA exploits the conventional dispersion constant, $\frac{m_4}{\sigma_{\mathrm{A}}^2}$ (i.e. Eq .5.27) and the detail will be explained in the next paragraph.

Now, we put our focus on MIMO-CMA [36] in Fig. 5.2. We can observe for the case of MIMO-CMA, MSE of equalizer 1 is higher than the MSE of equalizer 2. This is not surprise because the step size of MIMO-CMA in equalizer 2 is smaller than the step size in equalizer 1. Next, we concentrate on the equalizer outputs and global impulses of MIMO-CMA in Figures 5.5 and 5.6, respectively. In Fig. 5.5, we can see the output symbols closely match with the 16-QAM constellation set and no serious issue can be found. In Fig. 5.6, we can see $|s_{11}|$ and $|s_{12}|$ have a spike. This means equalizer 1 and 2 recover the same source 1. Thus, the source separation is not successful. This is expected because the previous work [36] has pointed out that MIMO-CMA does not have the source separation ability. Instead of that, we can observe the error function of MIMO-CMA (i.e. Eq. 5.24) and found that it solely has $e_{\mathrm{CMA},j}[k]$ without $e_{\mathrm{CI},j}[k]$. In other words, MIMO-CMA can only open the channel eye by the term $e_{\mathrm{CMA},j}[k]$, but cannot perform source separation due to the absence of $e_{\mathrm{CI},j}[k]$. Different with CC-CMA, MIMO-CMA does not introduce any down-scaling at equalizer output even though the same conventional dispersion constant $\frac{m_4}{\sigma_{\mathrm{A}}^2}$ (i.e. Eq. 5.25) has been exploited. This is because the conventional dispersion constant was originally designed and optimized for the term $e_{\mathrm{CMA},j}[k]$ alone [21]. Hence, MIMO-CMA, which contains only $e_{\mathrm{CMA},j}[k]$ without other extra term, does not introduce any down-scaling. On the other hand,

CC-CMA has an extra term of "$k_j \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \hat{r}_{1,\delta}[k] y_1[k-\delta]$" for source separation in Eq. 5.26, but the conventional dispersion constant, which was initially designed without considering the extra term, has been exploited in CC-CMA. Therefore, due to the inappropriate dispersion constant has been exploited, CC-CMA introduces down-scaling at equalizer output.

Now, we put our attention on CI-CMA. In Fig. 5.2, we can see CI-CMA shows relatively high MSE values in equalizer 1 and 2. This is expected because CI-CMA only has the ability of open eye and source separation, but it does not been equipped with the tracking algorithm. Next, we put the target on the equalizer outputs and global impulses of CI-CMA in Figures 5.7 and 5.8, respectively. In Fig. 5.7, the symbols of equalizer 1 and 2 match with the 16-QAM constellation set. We can further observe that the symbols of equalizer 2 is more scattering than equalizer 1. This is because, due to the hieratical approach, the equalizer 2 has an extra source separation algorithm and thus produce high error rate. In Fig. 5.8, we can see a spike can be found in $|s_{11}|$ and $|s_{22}|$. This implies equalizer 1 retrieves source 1 and equalizer 2 retrieves source 2. Thus, source separation is achieved as expected. Furthermore, different with CC-CMA, no down-scaling can be found in CI-CMA because of new dispersion value in Eq. 5.22 has been used.

In Fig. 5.2, especially MSE of equalizer 1, we can observe the hybrid algorithms have almost same MSE value after time $k = 1 \times 10^5$. This is because in this low error condition, the tracking algorithms that are within the hybrid algorithms are almost dominant the hybrid algorithms. Due to the reason that the hybrid algorithms use the same tracking algorithms (i.e. decision directed algorithms), the hybrid algorithms show the almost same results after time $k = 1 \times 10^5$. This case can also be applied on the MCISAG and MCIRBA in equalizer 2. Instead of that, in the MSE of equalizer 2, we can see MCIBG always shows higher MSE value than MCISAG and MCIRBA. This is because the coexistence of tracking, source separation and acquisition algorithm in the error function of MCIBG causes MCIBG is more sensitive to error compared to MCISAG and MCIRBA.

We put our focus on the the equalizer outputs and global impulses of hybrid algorithms. We check the equalizer outputs of hybrid algorithms in Figures 5.9, 5.11 and 5.13. The symbols of

equalizer 1 and 2 of the hybrid algorithms show closely match with the 16-QAM constellation set and no serious issue can be traced. In Fig. 5.9, we can see the symbols of equalizer 2 for MCIBG is slightly scatter than the outputs of other hybrid algorithms. This is because MCIBG is more sensitive to error. In Figures 5.10, 5.12 and 5.14, we can observe that, a spike can be found in $|s_{11}|$ and $|s_{22}|$. This implies equalizer 1 retrieves source 1 and equalizer 2 retrieves source 2. Thus, source separation is successfully.

For simplicity purpose, the hybrid algorithms currently have been tested for 16-QAM. However, higher contestation symbol set such as 64-QAM, has the statistical property of the fourth order moment is greater than variance (i.e. $m_4 > \sigma^2$), the extension of the hybrid algorithms to 64-QAM is straight forward because 16-QAM has the similar property.

In spite of that, the hybrid algorithms have some limitations. First, due to transmitted source symbols in blind approach, the hybrid algorithms is slower than any trained algorithm. This is because the hybrid algorithm require longer time to estimate the output without any correct reference. In this situation, the hybrid algorithm requires longer time to reach the steady state compared to trained algorithm and thus the response time of hybrid algorithm is larger than trained algorithm. Instead of that, due to the reason that the stationary channels have been assumed, the hybrid algorithms is not expected to show good performance in fast varying channel. Nevertheless, since the practice channels are time varying channels, the extension of the hybrid algorithms to fast varying channel are suggested as future works.

Chapter 6

Conclusions and Future works

6.1 Conclusions

In this thesis, several MIMO blind equalization algorithms were proposed. These algorithms include the new CI-CMA and its associated hybrid algorithms. The CI-CMA was mainly proposed to improve the existing CC-CMA while the hybrid algorithms were proposed to reduce the steady state error and increase the convergence rate making them suitable for real-life implementations for multi-modulus modulation scheme.

One of the most significant contributions of the thesis lies in the identification of the bias of the CC-CMA cost function where it has been thought to be unbiased since it was first proposed in 1997. Based on the awareness of the bias issue, some differing interpretations and discrepancies in the literature can now be explicitly explained. Our main finding is to show that the CC cost (of the CC-CMA) is a fourth order function of the source statistics, contradicting the past assumption or belief that CC term is a mere second order function. Although the CC-CMA has bias issue, which was neglected in the convergence analysis of previous authors, the impact of not realizing this bias is much more significant than one would imagine. It is because the assumption that the cost involves only 2nd order statistic has mis-communicated the actual convergence behavior of

the algorithm. In realizing this issue, we subsequently took a more vigorous approach to perform the convergence analysis that is based on the correct fourth order statistics. Our finding that the bias of the CC-CMA will always shrink the equalizer output (not enlarge) has led to the novel unbiased solution called the cross-independence constant modulus algorithm (CI-CMA). Satisfied with the performance of the CI-CMA (in terms of overcoming the shrinking effect, and achieving source separation and equalization simultaneously), which we analytically proved and then verified through numerous computer simulations, we further performed the steady state MSE analysis which closely matches the simulated MSE. In our analysis, we found that among all the factors, only the step size can be reduced in order to achieve lower steady state error. However, value reduction in the step size has significant negative impact on the convergence rate. Therefore by retaining as large a step size as possible for the CI-CMA (to retain quick convergence), a hybrid solution that is able to reduce the steady state error without sacrificing or compromising on the speed of convergence is very attractive. Consequently, we proposed several MIMO hybrid algorithms that are based on the new CI-CMA, which naturally allows the transition between the DD algorithm and itself to be performed smoothly in the absence of the shrinking effect. Relative to our CI-CMA, the conventional CC-CMA, which always shrinks the output, is hard to be merged into a hybrid solution.

6.2 Future Works

In this thesis, the analytical MSE prediction is restricted to the steady state stage only. In the future, the analytical MSE prediction should be done on transient stage as well. The MSE prediction on the transient stage is helpful to identify the factors that influence the speed of convergence. Furthermore, the analytical MSE prediction currently has been done on the absence of additive noise. In the future, the analytical MSE prediction in the presence of additive noise should be performed.

Currently, convergence analysis is hard to be done on hybrid algorithms because of the nonlinearity function from decision device. The convergence analysis of hybrid algorithms should be done in the future. One other area which we can look into is to extend the attractive equalization and source separation features of the CI-CMA to single-modulus modulation schemes such as the m-ary PSK and 4-QAM. Therefore, a new algorithm that can perform source separation feature on the single-modulus modulation scheme should be proposed as a future work.

Appendix A

Proof of Equations

A.1 Proof of independence of the CI cost (3.7) and (3.19)

A necessary and sufficient condition to prove expectation independence between y_j and $y_{n,\delta}$ is to show that the equality of the joint probability density function $p(y_j, y_{n,\delta}) = p(y_j)p(y_{n,\delta})$ is achieved when the CI cost, $\sum_n \sum_\delta \zeta_{n,\delta}$, has been minimized. We rewrite the CI subcost $\zeta_{n,\delta} = \mathrm{E}\{|y_j y^*_{n,\delta}|^2\} = \mathrm{E}\{|y_j|^2|y_{n,\delta}|^2\}$. The covariance of two random variables is always equals or larger than zero so we have

$$\begin{aligned} & C(|y_j|^2,|y_{n,\delta}|^2) \\ & \quad \triangleq \mathrm{E}\{(|y_j|^2 - \mathrm{E}(|y_j|^2))(|y_{n,\delta}|^2 - \mathrm{E}(|y_{n,\delta}|^2))\} \\ & \quad = \mathrm{E}\{|y_j|^2|y_{n,\delta}|^2\} - \mathrm{E}(|y_j|^2)\mathrm{E}(|y_{n,\delta}|^2) \geq 0 \end{aligned} \tag{A.63}$$

which leads to the inequality $\mathrm{E}\{|y_j|^2|y_{n,\delta}|^2\} \geq \mathrm{E}(|y_j|^2)\mathrm{E}(|y_{n,\delta}|^2)$. Expanding the terms we get

$$\mathrm{E}\{|y_j|^2|y_{n,\delta}|^2\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |y_j|^2|y_{n,\delta}|^2 p(y_j, y_{n,\delta}) dy_j dy_{n,\delta} \tag{A.64}$$

$$\mathrm{E}\{|y_j|^2\} = \int_{-\infty}^{\infty} |y_j|^2 p(y_j) dy_j \tag{A.65}$$

$$\mathrm{E}\{|y_{n,\delta}|^2\} = \int_{-\infty}^{\infty} |y_{n,\delta}|^2 p(y_{n,\delta}) dy_{n,\delta} \tag{A.66}$$

Therefore for the equality to occur which corresponds to the CI subcost being minimized, the independence between y_j and $y_{n,\delta}$ must have been achieved giving $p(y_j, y_{n,\delta}) = p(y_j)p(y_{n,\delta})$.

A.2 Stationary Point Analysis: Derivation of (3.31)

Recall that we are primarily interested in the second equalizer's CI cost only, where $j = 2$, thus $n = 1$ and $\Sigma_\delta \zeta_{\mathrm{CI},n,\delta} = \Sigma_\delta \mathrm{E}|y_j y^*_{n,\delta}|^2 = \Sigma_\delta \mathrm{E}\{y_2^2 y_{1,\delta}^2\}$. First we obtain the expression of $\zeta_{\mathrm{CI},1,\delta}$ in terms of the global impulse response, splitting it into four terms, before differentiating each term w.r.t. the d-th, $d \in (0, 2N_s - 1)$, element of $\mathbf{S_2}$, i.e., $\frac{\partial \zeta_{\mathrm{CI},1,\delta}}{\partial S_2^{[d]}}$, and finally evaluating the full expression of $k_0 \sum_{\delta=-\delta_\mathrm{x}}^{\delta_\mathrm{x}} \frac{\partial \zeta_{\mathrm{CI},1,\delta}}{\partial S_2^{[d]}}$. For mathematical traceability, we split the derivation of $\frac{\partial \zeta_{\mathrm{CI},1,\delta}}{\partial S_2^{[d]}}$ into two portions: $\frac{\partial \zeta_{\mathrm{CI},1,\delta}}{\partial S_2^{[d_1]}}$ where $0 \le d_1 \le N_s - 1$ and $\frac{\partial \zeta_{\mathrm{CI},1,\delta}}{\partial S_2^{[d_2]}}$ where $N_s \le d_2 \le 2N_s - 1$. We begin now by expressing $\zeta_{\mathrm{CI},n,\delta}$ in terms of (3.4):

$$
\begin{aligned}
\zeta_{\mathrm{CI},1,\delta} &= \mathrm{E}\left\{y_2^2 y_{1,\delta}^2\right\} \\
&= \mathrm{E}\Big\{ \left(\mathbf{S}_{(1),2}^T \bar{\mathbf{a}}_1[k] + \mathbf{S}_{(2),2}^T \bar{\mathbf{a}}_2[k]\right)^2 \left(\vec{\alpha}^T \bar{\mathbf{a}}_1[k-\delta] + \vec{\beta}^T \bar{\mathbf{a}}_2[k-\delta]\right)^2 \Big\} \\
&= \mathrm{E}\Big\{ \left(\left(\mathbf{S}_{(1),2}^T \bar{\mathbf{a}}_1[k]\right)^2 + \left(\mathbf{S}_{(2),2}^T \bar{\mathbf{a}}_2[k]\right)^2 \right) \left(\left(\vec{\alpha}^T \bar{\mathbf{a}}_1[k-\delta]\right)^2 + \left(\vec{\beta}^T \bar{\mathbf{a}}_2[k-\delta]\right)^2 \right) \Big\} \\
&= \underbrace{\mathrm{E}\Big\{ \left(\mathbf{S}_{(1),2}^T \bar{\mathbf{a}}_1[k]\right)^2 \left(\vec{\alpha}^T \bar{\mathbf{a}}_1[k-\delta]\right)^2 \Big\}}_{\mathrm{T}_1} + \underbrace{\mathrm{E}\Big\{ \left(\mathbf{S}_{(1),2}^T \bar{\mathbf{a}}_1[k]\right)^2 \left(\vec{\beta}^T \bar{\mathbf{a}}_2[k-\delta]\right)^2 \Big\}}_{\mathrm{T}_2} \\
&\quad + \underbrace{\mathrm{E}\Big\{ \left(\mathbf{S}_{(2),2}^T \bar{\mathbf{a}}_2[k]\right)^2 \left(\vec{\alpha}^T \bar{\mathbf{a}}_1[k-\delta]\right)^2 \Big\}}_{\mathrm{T}_3} + \underbrace{\mathrm{E}\Big\{ \left(\mathbf{S}_{(2),2}^T \bar{\mathbf{a}}_2[k]\right)^2 \left(\vec{\beta}^T \bar{\mathbf{a}}_2[k-\delta]\right)^2 \Big\}}_{\mathrm{T}_4}
\end{aligned}
\tag{A.72}
$$

Since $\vec{\alpha}^T\overline{\mathbf{a}}_1[k-\delta] = \sum_{i=0}^{N_s-1}\alpha_i a_1[k-\delta-i]$ and $(\mathbf{S}^T_{(1),2}\overline{\mathbf{a}}_1[k])^2 = \sum_{p=0}^{N_s-1}\sum_{q=0}^{N_s-1} S_2^{[p]}a_1[k-p]S_2^{[q]}a_1[k-q]$, we derive $\frac{\partial \mathrm{T}_1}{\partial S_2^{[d_1]}}$ as shown in (A.73) at below. In a similar fashion, we get the rest of the derivatives in (A.74)-(A.77).

$$
\begin{aligned}
\frac{\partial \mathrm{T}_1}{\partial S_2^{[d_1]}} &= \sum_{i=0}^{N_s-1}\mathrm{E}\{\alpha_i^2\}\frac{\partial}{\partial S_2^{[d_1]}}\Big(\sum_{p=0}^{N_s-1}\sum_{q=0}^{N_s-1}\mathrm{E}\{S_2^{[p]}S_2^{[q]}a_1[k-p]a_1[k-q]a_1^2[k-\delta-i]\}\Big) \\
&= \sum_{i=0}^{N_s-1}\mathrm{E}\{\alpha_i^2\}\frac{\partial}{\partial S_2^{[d_1]}}\Big(\mathrm{E}\Big\{\left(S_2^{[d_1]}\right)^2 a_1^2[k-d_1]a_1^2[k-\delta-i] + S_2^{[d_1]}a_1[k-d_1]\sum_{p=0;p\neq d_1}^{N_s-1}S_2^{[p]}a_1[k-p]a_1^2[k-\delta-i] \\
&\quad + S_2^{[d_1]}a_1[k-d_1]\sum_{q=0;q\neq d_1}^{N_s-1}S_2^{[q]}a_1[k-p]a_1^2[k-\delta-i] + \sum_{p=0;p\neq d_1}^{N_s-1}\sum_{q=0;q\neq d_1}^{N_s-1}S_2^{[p]}S_2^{[q]}a_1[k-p]a_1[k-q]a_1^2[k-\delta-i]\Big\}\Big) \\
&= \sum_{i=0}^{N_s-1}\mathrm{E}\{\alpha_i^2\}\Big(\mathrm{E}\Big\{2\left(S_2^{[d_1]}\right)a_1^2[k-d_1]a_1^2[k-\delta-i] + 2a_1[k-d_1]\sum_{p=0;p\neq d_1}^{N_s-1}S_2^{[p]}a_1[k-p]a_1^2[k-\delta-i]\Big\}\Big) \\
&= 2\sum_{i=0}^{N_s-1}\mathrm{E}\{\alpha_i^2\}\sum_{p=0}^{N_s-1}S_2^{[p]}\mathrm{E}\left\{a_1[k-d_1]a_1[k-p]a_1^2[k-\delta-i]\right\} \qquad \text{(A.73)} \\
\frac{\partial \mathrm{T}_2}{\partial S_2^{[d_1]}} &= 2\sum_{j=0}^{N_s-1}\mathrm{E}\{\beta_j^2\}\sum_{p=0}^{N_s-1}S_2^{[p]}\mathrm{E}\{a_1[k-d_1]a_1[k-p]a_2^2[k-\delta-j]\} \qquad \text{(A.74)} \\
\frac{\partial \mathrm{T}_3}{\partial S_2^{[d_2]}} &= 2\sum_{i=0}^{N_s-1}\mathrm{E}\{\alpha_i^2\}\sum_{p=0}^{N_s-1}S_2^{[p+N_s]}\mathrm{E}\{a_2[k-d_1]a_2[k-p]a_1^2[k-\delta-i]\} \qquad \text{(A.75)} \\
\frac{\partial \mathrm{T}_4}{\partial S_2^{[d_2]}} &= 2\sum_{j=0}^{N_s-1}\mathrm{E}\{\beta_j^2\}\sum_{p=0}^{N_s-1}S_2^{[p+N_s]}\mathrm{E}\{a_2[k-d_1]a_2[k-p]a_2^2[k-\delta-j]\} \qquad \text{(A.76)} \\
\text{Also,}\quad & \frac{\partial \mathrm{T}_3}{\partial S_2^{[d_1]}} = \frac{\partial \mathrm{T}_4}{\partial S_2^{[d_1]}} = \frac{\partial \mathrm{T}_1}{\partial S_2^{[d_2]}} = \frac{\partial \mathrm{T}_2}{\partial S_2^{[d_2]}} = 0. \qquad \text{(A.77)}
\end{aligned}
$$

We now turn our attention to the evaluation of the full sum of terms 1 and 2, starting with T_1:

$$\begin{aligned} k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \frac{\partial \mathrm{T}_1}{\partial S_2^{[d_1]}} &= 2k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} \sum_{p=0}^{N_s-1} \left(S_2^{[p]} \mathrm{E}\{ \underbrace{a_1[k-d_1]}_{\mathsf{P}} \underbrace{a_1[k-p]}_{\mathsf{Q}} \underbrace{a_1^2[k-\delta-i]}_{\mathsf{R}} \} \right) \\ &= \mathrm{E}\{\mathsf{PQR}\} + \mathrm{E}\{\mathsf{PQ}\}\mathrm{E}\{\mathsf{R}\} \end{aligned} \tag{A.78}$$

We utilize a tagging notation where the respective source signals are tagged with an alphabet P, Q or R for easier reference. The expectation function yields non-zero values only for the above two cases of (A.78), since the other two possible combinations, i.e., $\mathrm{E}\{\mathsf{P}\}\mathrm{E}\{\mathsf{QR}\}$ and $\mathrm{E}\{\mathsf{Q}\}\mathrm{E}\{\mathsf{PR}\}$, would have zero values at steady state because the sources are zero mean and i.i.d.. Consider now $\mathrm{E}\{\mathsf{PQR}\}$ which could be non-zero only when $p = \delta + i = d_1$. It can be simplified as below in a progressive manner:

$$\begin{aligned} \mathrm{E}\{\mathsf{PQR}\} &= 2S_2^{[d_1]} k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} \mathrm{E}\left\{a_1^2[k-d_1] a_1^2[k-\delta-i]\right\} \\ &= 2S_2^{[d_1]} k_0 \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} \mathrm{E}\left\{a_1^4[k-d_1]\right\} \\ &= 2S_2^{[d_1]} k_0 m_4 \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} \end{aligned} \tag{A.79}$$

The evaluation of $\mathrm{E}\{\mathsf{PQ}\}\mathrm{E}\{\mathsf{R}\}$ proceeds by examining the case when $p = d_1$ and $d_1 \neq \delta + i$, which gives

$$\begin{aligned} \mathrm{E}\{\mathsf{PQ}\}\mathrm{E}\{\mathsf{R}\} &= 2k_0 S_2^{[d_1]} \sigma_{\mathrm{A}}^2 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \sum_{\substack{i=0 \\ d_1 \neq \delta+i}}^{N_s-1} \mathrm{E}\{\alpha_i^2\} \mathrm{E}\left\{a_1^2[k-\delta-i]\right\} \\ &= 2k_0 S_2^{[d_1]} \sigma_{\mathrm{A}}^4 \left((2\delta_{\mathrm{x}}+1) \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} - \sum_{j=0}^{N_s-1} \mathrm{E}\{\alpha_j^2\} \right). \end{aligned} \tag{A.80}$$

As a result, combining (A.79) and (A.80) we get

$$k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \frac{\partial \mathrm{T}_1}{\partial S_2^{[d_1]}} = 2S_2^{[d_1]} k_0 \left(\left(m_4 - \sigma_{\mathrm{A}}^4\right) \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} + (2\delta_{\mathrm{x}}+1)\, \sigma_{\mathrm{A}}^4 \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} \right) \tag{A.81}$$

As for term 2, the full sum is expressed and tagged as follows:

$$
\begin{aligned}
k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \frac{\partial \mathrm{T}_2}{\partial S_2^{[d_1]}} &= 2k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \sum_{j=0}^{N_s-1} \mathrm{E}\{\beta_j^2\} \sum_{p=0}^{N_s-1} S_2^{[p]} \mathrm{E}\Big\{ \underbrace{a_1[k-d_1]}_{\mathsf{F}} \underbrace{a_1[k-p]}_{\mathsf{G}} \underbrace{a_2^2[k-\delta-j]}_{\mathsf{D}} \Big\} \\
&= \mathrm{E}\{\mathsf{FG}\}\mathrm{E}\{\mathsf{D}\} \\
&= 2k_0 \sum_{\delta=-\delta_{\mathrm{x}}}^{\delta_{\mathrm{x}}} \sum_{j=0}^{N_s-1} \mathrm{E}\{\beta_j^2\} S_2^{[d_1]} \Big(\mathrm{E}\big\{a_1^2[k-d_1]\big\} \mathrm{E}\big\{a_2^2[k-\delta-j]\big\} \Big) \\
&= 2k_0 S_2^{[d_1]} \sigma_{\mathrm{A}}^4 \left(2\delta_{\mathrm{x}}+1\right) \sum_{j=0}^{N_s-1} \mathrm{E}\{\beta_j^2\} \qquad \text{(A.82)}
\end{aligned}
$$

The only non-zero value is obtained by $\mathrm{E}\{\mathsf{FG}\}\mathrm{E}\{\mathsf{D}\}$, so we only need to consider the case when $p = d_1$, which gives the remaining two equations of (A.82). Thus, the complete derivation w.r.t. $S_2^{[d_1]}$ is

$$
k_0 \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \frac{\partial f_{\mathrm{CI},\delta}}{\partial S_2^{[d_1]}} = 4 S_2^{[d_1]} \frac{k_0}{2} \Bigg(\left(m_4 - \sigma_{\mathrm{A}}^4\right) \sum_{i=0}^{N_s-1} \mathrm{E}\{\alpha_i^2\} + \Omega_o \Bigg) \qquad \text{(A.83)}
$$

where $\Omega_o = \left(2\delta_{\max}+1\right) \sigma_{\mathrm{A}}^4 \big(\mathrm{E}\{\|\vec{\alpha}\|^2\} + \mathrm{E}\{\|\vec{\beta}\|^2\} \big)$. Then the entire process is repeated to obtain the expression w.r.t. $S_2^{[d_2]}$ where eventually the result is

$$
k_0 \sum_{\delta=-\delta_{\max}}^{\delta_{\max}} \frac{\partial f_{\mathrm{CI},\delta}}{\partial S_2^{[d_2]}} = 4 S_2^{[d_2]} \frac{k_0}{2} \Bigg(\left(m_4 - \sigma_{\mathrm{A}}^4\right) \sum_{j=0}^{N_s-1} \mathrm{E}\{\beta_j^2\} + \Omega_o \Bigg). \qquad \text{(A.84)}
$$

Combining (A.83) and (A.84) gives Eq (3.31), which completes our proof.

References

[1] S. Haykin, *Unsupervised Adaptive Filtering, Volumne 2: Blind Deconvolution.* John Wiley and Sons Inc., 2000.

[2] A. Touzni, I. Fijalkow, M. G. Larimore, and J. R. Treichler, "A globally convergent approach for blind MIMO adaptive deconvolution," *IEEE Trans. Signal Processing*, vol. 49, pp. 1166–1178, Jun 2001.

[3] World Health Organization, "Factsheets: Noncommunicable diseases." `http://www.who.int/mediacentre/factsheets/fs355/en/`, 2013. [Online accessed 19-Sep-2014].

[4] R. Dilmaghani, H. Bobarshad, and M. Ghavami, "Wireless Sensor Networks for Monitoring Physiological Signals of Multiple Patients," *IEEE Transactions on Biomedical Circuits and Systems*, pp. 347–356, 2011.

[5] D. Oletic, B. Arsenali, and V. Bilas, "Low-Power Wearable Respiratory Sound Sensing," *Open Access Journal Sensors*, pp. 6535–6566, 2014.

[6] D. C. Klonoff, "The Benefits of Implanted Glucose Sensors," *Journal Diabetes Science Technology*, vol. 1, no. 6, p. 797800, 2007.

[7] R. Sobot, "Implantable RF telemetry for cardiac monitoring in the murine heart: a tutorial review," *EURASIP Journal on Embedded Systems*, vol. 1, 2013.

[8] A. Khaleghi, R. Chavez-Santiago, and I. Balasingham, "Ultra-wideband pulse-based data communications for medical implants," *IET Communications Journal*, vol. 4, pp. 1889–1897, 2010.

[9] K. S. Kwak, S. Ullah, and N. Ullah, "An overview of IEEE 802.15.6 standard," *3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, vol. 2, pp. 1–6, Mar 2012.

[10] "IEEE Standard for Local and Metropolitan Area Networks - Part 15.6: Wireless Body Area Networks. IEEE STD 802.15.6-2012," *IEEE Computer Society*, vol. 2, Mar 2012.

[11] R. Chvez-Santiago and K. Sayrafian-Pour and A. Khaleghi et al, "Propagation Models for IEEE 802.15.6 Standardization of Implant Communication in Body Area Networks," *IEEE Communications Magazine*, pp. 80–87, Aug 2013.

[12] T. Akin, K. Najafi, and R. M. Bradley, "A wireless implantable multichannel digital neural recording system for a micromachined sieve electrode," *IEEE J. Solid-State Circuits*, vol. 33, no. 1, pp. 109–118, 1998.

[13] P. R. Troyk and M. A. K. Schwan, "Closed-loop class E transcutaneous power and data link for microimplants," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 6, p. 589599, 1998.

[14] M. Ghovanloo and K. Najafi, "A wide-band frequency-shift keying wireless link for inductively powered biomedical implants," *IEEE Trans. Circuits Syst. I*, vol. 51, no. 12, p. 23742383, 2004.

[15] P. Mohseni and K. Najafi, "Wireless multi-channel biopotential recording using an integrated FM telemetry circuit," *IEEE Trans. Neural. Syst. Rehabil. Eng*, vol. 13, p. 263271, 2005.

[16] R. Bashirullah, "Wireless Implants," *IEEE Microwave Magazine*, vol. 11, no. 7, pp. S14–S23, 2010.

[17] H. Tsuji and J. Zheng and K. Mostafa, "Experiments in radio location estimation using an airbone array," *IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 189–193, July 2008.

[18] H. Artés, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection," *IEEE Trans. Signal Processing, Special Issue on MIMO Communications Systems*, vol. 51, pp. 2808–2820, Nov 2003.

[19] M. Chouayakh, A. Knopp, and B. Lankl, "Low-effort near maximum likelihood MIMO detection with optimum hardware resource exploitation," *Electronics Letters*, vol. 43, pp. 1104–1106, Sep 2007.

[20] O. Macchi, *Adaptive processing. The LMS Approach with Applications in Transmission.* New York: Wiley, 1995.

[21] D. N. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication system," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov 1980.

[22] Y. Sato, "A method of self-recovering equalization for multi-level amplitude-modulations systems," *IEEE Trans. Commun.*, pp. 679–682, Jun 1975.

[23] J. Yang, J.-J. Werner, and G. A. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE Journal on selected areas in communications*, vol. 20, pp. 997–1015, Jun 2002.

[24] K. N. Oh and Y. O. Chin, "New blind equalization techniques based on constant modulus algorithm," *Proc. 1995 IEEE Global Telecommunications Conf., Singapore*, pp. 865–869, Nov 14-16 1995.

[25] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Information Theory*, vol. 36, Mar 1990.

[26] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. John Wiley and Sons, 1st ed., 2001.

[27] P. Comon, "Independent component analysis, a new concept?," *Elsevier Sigal Processing*, vol. 36, pp. 287–314, Apr 1994.

[28] C. B. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion," *IEEE Trans. Signal Processing*, vol. 48, pp. 3508–3519, Dec 2000.

[29] M. Phegade, P. Mukherji, and U. S. Sutar, "Hybrid ICA algorithm for ECG analysis," *International Conference on Hybrid Intelligent Systems*, vol. 35, pp. 478–483, Mar 2012.

[30] G. R. Naik and D. K. Kumar, "An Overview of Independent Component Analysis and Its Applications," *Informatica*, vol. 35, Mar 2011.

[31] C. B. Papadias and A. J. Paulraj, "A constant modulus algorithm for multiuser signal separation inpresence of delay spread using antenna arrays," *IEEE Signal Processing Lett.*, vol. 2, pp. 178–181, Jun 1997.

[32] P. Sansrimahachai, D. B. Ward, and A. G. Constantinides, "Blind source separation for BLAST," *Conf. Digital Signal Processing*, vol. 1, pp. 139–142, July 2002.

[33] A. Ikhlef and D. L. Guennec, "Blind recovery of MIMO QAM signals: a criterion with its convergence analysis," *Proceedings of the 14th European Signal Processing Conference (EUSIPCO'06)*, no. 1, 2006.

[34] A. Ikhlef and D. L. Guennec, "A simplified constant modulus algorithm for blind recovery of MIMO QAM and PSK signals: a criterion with convergence analysis," *EURASIP Journal on Wireless Communications and Networking 2007*, vol. 15, 2007.

[35] R. Gooch and J. Lundell, "The cm array: An adaptive beamformer for constant modulus signals," *Proc. IEEE Acoustics, Speech, and Signal Process. (ICASSP)*, pp. 2523–2526, 1986.

[36] Y. Li and K. J. R. Liu, "Adaptive blind source separation and equalization for multiple-input/multiple-output systems," *IEEE Trans. Information Theory*, vol. 44, pp. 2864–2876, Nov 1998.

[37] C. Meng, J. Tuqan, and Z. Ding, "A Quadratic Programming Approach to Blind Equalization and Signal Separation," *IEEE Trans. Signal Processing*, vol. 57, pp. 3196–3203, Dec 2009.

[38] H. D. Han and Z. Ding, "Steepest descent algorithm implementation for multichannel blind signal recovery," *IET Communications Journal*, vol. 6, pp. 2232–2244, Dec 2012.

[39] M. Castella, S. Rhioui, E. Moreau, and J. Pesquet, "Quadratic Higher-Order Criteria for Iterative Blind Separation of a MIMO Convolutive Mixture of Sources," *IEEE Trans. Signal Processing*, vol. 55, Jan 2007.

[40] M. Kawamoto, Y. Inouye, K. Kohno, and T. Maeda, "An Eigenvector Algorithm with Reference Signals Using a Deflation Approach for Blind Deconvolution," *7th International Conference, ICA 2007*, 2007.

[41] J. Via, I. Santamaria, and J. Perez, "Deterministic CCA-Based Algorithms for Blind Equalization of FIR-MIMO Channels," *IEEE Trans. Signal Processing*, vol. 55, pp. 3867–3878, July 2007.

[42] L. Castedo, C. J. Escudero, and A. Dapena, "A blind signal separation method for multiuser communications," *IEEE Trans. Signal Process.*, vol. 45, pp. 1343–1348, May 1997.

[43] Y. Luo, J. A. Chambers, and S. Lambotharan, "Global convergence and mixing parameter selection in the cross-correlation constant modulus algorithm for the multi-user environment," *IEE Proc. Visual Image Signal Proceessing*, vol. 148, pp. 9–20, Feb 2001.

[44] S. Li, T. S. Qiu, and S. F. Zhang, "Space-time blind equalisation in impulsive noise," *IET Signal Process.*, vol. 3, no. 6, pp. 445–458, 2009.

[45] S. Li, T. Qui, and D. Zha, "Adaptive blind equalization for MIMO systems under alpha-stable noise environment," *IEEE Commun. Letter*, vol. 13, pp. 609–611, Aug 2009.

[46] A. Vgenis, C. S. Petrou, C. B. Papadias, I. Roudas, and L. Raptis, "Nonsingular constant modulus equalizer for PDM-QPSK coherent optical receivers," *IEEE Photnics Technology Letters*, vol. 22, pp. 45–47, Jan 2010.

[47] S. Li, "Space-Time Blind Equalization under alpha-Stable Noise Environment," *Wireless Communications, Networking and Mobile Computing*, pp. 1–4, Sept 2009.

[48] M. Silva, M. Miranda, and A N Licciardi Jr, "A robust algorithm for space-time equalization," *Proc. ICASSP 04*, pp. 857–860, 2004.

[49] A. M. Nassar and E. W. E. Nahal, "Blind equalization technique for cross correlation constant modulus algorithm," *WSEAS Transaction on Signal Process.*, vol. 6, no. 2, pp. 23–32, 2010.

[50] Y. Luo and J. A. Chambers, "Steady state mean-square error analysis of the cross-correlation and constant modulus algorithm in a mimo convolutive system," *Proc. 2002 IEE Vision, Image and Signal Processing*, vol. 149, pp. 196–203, 2002.

[51] Y. Luo, S. Lambotharan, and J. A. Chambers, "Global convergence and mixing parameter selection in the cross correlation constant modulus algorithm for the multi-user environment," *Proc. IEE Vision, Image and Signal Processing*, vol. 148, pp. 9–20, 2001.

[52] Y. Luo and J. A. Chambers, "Bounds for the mixing parameter within the cc-cma algorithm applied in non ideal multiuser environments," *Proc. 2001 IEEE Acoustics, Speech, and Signal Process. (ICASSP)*, vol. 4, pp. 2173–2176, 2001.

[53] S. Lambotharan and J. A. Chambers, "On the surface characteristics of a mixed constant modulus and cross-correlation for the blind equalisation of a mimo channel," *Elsevier Signal Processing*, vol. 72, no. 2, pp. 209–216, 1999.

[54] J. Mai and A.H. Sayed, "A feedback approach to the steady-state performance of fractionally spaced blind adaptive equalizers," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 80–91, 2000.

[55] Centers for Disease Control and Prevention, "Chronic Diseases and Health Promotion." `http://www.cdc.gov/chronicdisease/overview/`, 2014. [Online accessed 19-Sep-2014].

[56] W. Chodzko-Zajko, D. Proctor, , M. F. Singh, C. Minson, C. Nigg, G. Salem, and J. Skinner, "American College of Sports Medicine Position Stand. Exercise and Physical Activity for Older Adults," *Med. Sci. Sport. Exerc*, vol. 41, pp. 1510–1530, 2009.

[57] L. Allet, R. H. Knols, K. Shirato, and E. D. de Bruin, "Wearable Systems for Monitoring Mobility-Related Activities in Chronic Disease: A Systematic Review," *Open Access Journal Sensors*, pp. 9026–9052, 2010.

[58] L. Wang, G. Yang, J. Huang, J. Zhang, L. Yu, and Z. Nie, "A wireless biomedical signal interface system-on-chip for body sensor networks," *IEEE Transaction on Biomedical Circuits and Systems*, vol. 4, pp. 112–117, 2010.

[59] S. Patel, H. Park, , P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 21, 2012.

[60] S. Gulley, E. Rasch, and L. Chan, "If we build it, who will come? Working-age adults with chronic health care needs and the medical home," *Medical Care*, vol. 49, no. 2, pp. 149–155, 2011.

[61] P. Bonato, "Wearable sensors and systems. From enabling technology to clinical applications," *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, pp. 25–36, 2010.

[62] N. Oliver and F. Flores-Mangas, "HealthGear: Automatic Sleep Apnea Detection and Monitoring with a Mobile Phone," *Journal Of Communications*, vol. 2, Mar 2007.

[63] P. Corbishley and Rodriguez-Villegas, "Breathing detection: towards a miniaturized, wearable, battery-operated monitoring system," *IEEE Transaction on Biomedical Engineering*, vol. 55, pp. 196–204, 2008.

[64] W. Haiying, Z. Huiru, and J. A. JC, "Monitoring and analysis of sleep pattern for people with early dementia," *IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*, pp. 405–410, 2010.

[65] J. Jimison, M. Pavel, J. Pavel, and J. McKanna, "Home monitoring of computer interactions for the early detection of dementia," *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4533–4536, 2004.

[66] J. Zhang, P. Orlik, Z. Sahinoglu, A. Molisch, and P. Kinney, "UWB systems for wireless sensor networks," *Proceedings of the IEEE*, vol. 97, pp. 313–331, 2009.

[67] M. Hernandez and R. Kohno, "UWB systems for body area networks in IEEE 802.15. 6," *IEEE International Conference on Ultra-Wideband (ICUWB)*, pp. 235–239, 2011.

[68] M. H. V. Niemel and J. Iinatti, "IEEE 802.15. 4a UWB Receivers in medical applications," *International Journal of Ultra Wideband Communications and Systems*, vol. 2, pp. 73–82, 2011.

[69] B. Z. W. Rhee, N. Xu and Z. Wang, "Low power, non invasive UWB systems for WBAN and biomedical applications," *International Conference on Information and Communication Technology Convergence (ICTC)*, vol. 2, pp. 35–40, 2010.

[70] M. H. H. Viittala, B. N. Nahar and J. Iinatti, "Medical applications adapting ultra wideband: a system study," *International Journal of Ultra Wideband Communications and Systems*, vol. 1, pp. 237–247, 2010.

[71] M. Hernandez and R. Kohno, "Ultra low power UWB transceiver design for body area networks," *2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, vol. 1, pp. 1–4, 2009.

[72] N. Kernen, M. Srestniemi, and J. Partala, "IEEE802.15.6 based Multi-Accelerometer WBAN System for Monitoring Parkinson's Disease," *35th Annual International Conference of the IEEE EMBS*, pp. 1656–1659, 2013.

[73] A. Darwish and A. E. Hassanien, "Wearable and Implantable Wireless Sensor Network Solutions for Healthcare Monitoring," *Sensors*, vol. 11, no. 6, p. 55615595., 2011.

[74] Masihpour Mehrnoush, *Cooperative communication in near field magnetic induction communication systems*. PhD Thesis, University of Technology, Sydney, 2012.

[75] O. Brand, "Microsensor integration into systems-on-chip," *IEEE Engineering in Medicine and Biology Magazine*, vol. 94, pp. 1160–1176, 2006.

[76] Vemund Svanes Bertelsen, *Communication System for a Medical Implant*. Msc Thesis, Universtiy of Oslo, 2004.

[77] Eliza Strickland, "Medtronic Wants to Implant Sensors in Everyone, IEEE Spectrum." `http://spectrum.ieee.org/tech-talk/biomedical/devices/medtronic-wants-to-implant-sensors-in-everyone`, 2014. [Online accessed 19-Sep-2014].

[78] F. M. Merchant, G. W. Dec, and J. P. Singh, "Implantable Sensors for Heart Failure," *Circulation: Arrhythmia and Electrophysiology, American Heart Association*, p. 657667, 2010.

[79] "World Cancer Report 2014," *World Health Organization*, 2014.

[80] A. C. Society, "Global cancer facts and figures." `http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-027766.pdf`, 2011. [Online accessed 19-Sep-2014].

[81] World Health Organization, "Factsheets: Cancer." http://www.who.int/mediacentre/factsheets/fs297/en/, 2014. [Online accessed 19-Sep-2014].

[82] K. Choudhari et al, "Nitric oxide and cancer: a review," *World Journal of Surgical Oncology*, 2013.

[83] Medical News Today, "Implantable sensor may monitor cancer and diabetes." http://www.medicalnewstoday.com/articles/268347.php, 2013. [Online accessed 19-Sep-2014].

[84] G. Ciuti, A. Menciassi, and P. Dario, "Capsule Endoscopy: From Current Achievements to Open Challenges," *IEEE Review in Biomedical Engineering*, vol. 4, pp. 59–72, 2012.

[85] Y. Gu, X. Xie, Z. Wang, and G. Li, "A new globularity capsule endoscopy system with multi-camera," *IEEE Biomedical Circuits and Systems Conference*, pp. 289–292, 2009.

[86] L. Dong and Y. Wu, "A Wireless Narrowband Imaging Chip for Capsule Endoscope," *IEEE Transactions On Biomedical Circuits Aand Systems*, vol. 4, no. 6, pp. 462–468, 2010.

[87] E. Jovanov and A. Milenkovic, "Body area networks for ubiquitous healthcare applications: opportunities and challenges," *Journal of Medical Systems*, vol. 35, pp. 1245–1254, 2011.

[88] Anders J. Johansson, *Wireless Communication with Medical Implants: Antennas and Propagation*. PhD Thesis, Lund University, Sweden, 2004.

[89] K. Y. Yazdandoost and R. Kohno, "Wireless Communications for Body Implanted Medical Device," *Microwave Conference*, 2007.

[90] J. Wang and Q. Wang, "Channel Modeling and BER Performance of an Implant UWB Body Area Link," *Proc. 2nd Intl. Symp. Applied Sciences in Biomedical and Commun. Technologies*, pp. 24–27, 2009.

[91] A. Khaleghi, R. Chvez-Santiago, and I. Balasingham, "An Improved Ultra Wideband Channel Model Including the Frequency-Dependent Attenuation for In-Body Communications," *Proc. 34th Annual Intl. Conf. IEEE Eng. in Medicine and Biology Society*, p. 16311634, 2012.

[92] G. Wunder, R. F. H. Fischer, H. Boche, S. Litsyn, and J.-S. No, "The PAPR Problem in OFDM Transmission: New Directions for a Long-Lasting Problem," *IEEE Signal Processing Magazine*, pp. 130–144, Nov 2013.

[93] J.-H. Wen, G.-R. Lee, C.-C. Kung, and C.-Y. Yang, "Coding Schemes Applied to Peak-to-Average Power Ratio (PAPR) Reduction in OFDM Systems," *International Wireless Communications and Mobile Computing Conference*, pp. 807–812, Aug 2008.

[94] Z. Zhang, K. Long, M. Zhao, and Y. Liu, "Joint Frame Synchronization and Frequency Offset Estimation in OFDM Systems," *IEEE Transaction On Broadcasting*, vol. 51, pp. 389–394, Sept 2005.

[95] M. H. Akbari, M. Ahmadian, and S. Salari, "Joint Frame and Frequency Synchronization for MIMO OFDM Systems with ZCZ Sequences," *3rd IEEE/IFIP International Conference in Central Asia on Internet*, pp. 1–5, 2007.

[96] Z. Ye, C. Duan, P. V. Orlik, J. Zhang, and A. A. Abouzeid, "A Synchronization Design for UWB-Based Wireless Multimedia Systems," *IEEE Transaction On Broadcasting*, pp. 211–225, 2010.

[97] Y. Mostofi, "Analysis of the effect of timing synchronization errors on pilot-aided ofdm systems," *IEEE Conference on Signals, Systems and Computers*, vol. 1, pp. 638–642, Nov 2003.

[98] T. Haritha, S. SriGown, and D. E. Rani, "Frequency offset estimation in coherent ofdm systems using different fading channels," *ARPN Journal of Engineering and Applied Science*, vol. 7, pp. 691–695, June 2013.

[99] G. Tatsis, C. Votis, V. Raptis, V. Christofilakis, and P. Kostarakis, "A/D Restrictions (Errors) in Ultra-Wideband Impulse Radios ," *Int. J. Communications Network and System Sciences*, vol. 3, pp. 425–429, 2010.

[100] R. Akbar and E. Radoi, "An overview of synchronization algorithms for IR-UWB systems," *International Conference on Computing, Networking and Communications (ICNC)*, pp. 573–577, 2012.

[101] H. Aubert, "RFID Technology for Human Implant Devices ," *Special Issue on Nanoscience and nanotechnologies: hopes and concerns*, vol. 12, no. 7, pp. 675–683, 2011.

[102] M. Sun, S. A. Hackworth, Z. Tang, G. Gilbert, S. Cardin, and R. J. Sclabassi, "How to pass information and deliver energy to a network of implantable devices within the human body," *Proc. IEEE Conf. on Engineering in Medicine and Biology Society (EMBS 2007)*, pp. 5286–5289, 2007.

[103] Z. Nie, T. Leng, W. Wang, F. Guan, and L. Wang, "Experimental Characterization of human body communication in shield chamber," *2012 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 759–762, 2012.

[104] H. Higgins, "Body Implant Communication - Is It A Reality?," *IET Seminar on Antennas and Propagation for Body-Centric Wireless Communications*, pp. 33–36, 2007.

[105] Y. Yu, *Blind Identification of Possibly Under-determined Convolutive MIMO Systems.* PhD Thesis, Drexel University, July 2007.

[106] C. Y. Chi, C. C. Feng, C. H. Chen, and C. Y. Chen, *Blind Equalization and System Identification: Batch Processing Algorithms, Performance and Applications.* Springer-Verlag London Limited, 1st ed., 2006.

[107] R. T. Causey, *Blind Multiuser Detection Based on Second-Order Statistics.* PhD Thesis, Georgia Institutue of Technology, 1999.

[108] R. W. Lucky, "Techniques for adaptive equalization of digital communciation systems," *Bell System Tehnical Journal*, vol. 45, Feb 1966.

[109] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 43, pp. 2120–2129, Sep 1995.

[110] Y. Li and Z. Ding, "Static and dynamic convergence behavior of adaptive blind equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 2736–2745, Nov 1996.

[111] Y. Li, K. J. R. Liu, and Z. Ding, "Length and cost dependent local minima of blind channel equalizers," *IEEE Trans. Signal Processing*, vol. 44, pp. 2726–2735, Nov 1996.

[112] S. Bellini, "Bussgang techniques for blind equalization," *GLOBECOM*, pp. 1634–1640, 1986.

[113] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory.* Prentice Hall, 1st ed., 1993.

[114] S. Haykin, *Blind Deconvolution.* Prentice Hall Information and System Sciences Series, 1994.

[115] S. Haykin, *Adaptive Filter Theory.* Prentice Hall Information and System Sciences Series, 3 ed., 1996.

[116] A. H. Sayed, *Adaptive Filters.* John Wiley and Sons, 2008.

[117] J. G. Proakis, *Digital Communications.* McGraw-Hill International Editions, 2 ed., 1989.

[118] R. A. Kennedy, B. Anderson, Z. Ding, and C. R. Johnson, Jr., "Local stable minima of the Sato recursive identification scheme," *Decision and Control, 1990., Proceedings of the 29th IEEE Conference on*, vol. 6, pp. 3194–3199, Dec 1990.

[119] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Signal Processing*, vol. 31, pp. 349–372, Apr 1983.

[120] Z. Ding, R. A. Kennedy, C. R. Johnson, Jr., and B. D. O. Anderson, "Ill-convegence of Godard blind equalizers in data communication systems," *IEEE Trans. Commun.*, vol. 39, pp. 1313–1326, Sep 1991.

[121] C. R. Johnson, Jr. *et al.*, "Blind equalization using the constant modulus criterion: A review," *Proc. IEEE*, vol. 86, pp. 1927–1950, Oct 1998.

[122] M. G. Larimore and J. R. Treichler, "Convergence behavior of the Constant Modulus Algorithm," *ICASSP (Boston)*, pp. 13–16, Apr 1983.

[123] I. Fijalkow, C. Manlove, and C. R. Johnson, Jr., "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Processing*, vol. 46, pp. 227–231, Jan 1998.

[124] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a "Stop-and-Go" decision-directed algorithm," *IEEE Trans. Commun.*, vol. COM-35, pp. 877–887, Sep 1987.

[125] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COM-32, pp. 871–883, Aug 1984.

[126] W. G. Lim, T. D. Abhayapala, and R. A. Kennedy, "Reliability based soft transition technique for dual-mode blind equalizers," *IEEE International Conference on Communications, Paris*, vol. 5, pp. 2631–2635, Jun 2004.

[127] W. G. Lim, "New soft transition dual mode type algorithms for blind equalization," *IEEE Wireless Communications and Networking Conference (WCNC 2007), Hong Kong*, pp. 504–508, Mar 11–15 2007.

[128] N. Delfosse and P. Loubaton, "Adaptive separation of independent sources: a deflation approach," *ICASSP*, vol. 4, 1994.

[129] S. Icart, P. Comon, and L. Rota, "Blind paraunitary equalization," *Signal Processing (Elsevier)*, pp. 283–290, 2009.

[130] R. Bronson, *Schaum's Theory and Problems of Matrix Operations.* McGraw Hill, 1st ed., 1989.

[131] M. Silva and M. Miranda, "Tracking analysis of some space-time blind equalization algorithms," *Proc. 2005 IEEE/SP 13th Workshop on Statistical Signal Processing*, pp. 167–172, 2005.

[132] W. G. Lim, "Automatic phase recovery properties of blind adaptive equalization algorithms," *IEEE 8th Malaysia International Conference on Communications, Penang*, May 14–17 2007.

List of Publications

- Zhi-Hou Lee and Wee Gin Lim,"Multi-user multimodulus algorithm in blind source separation and equalization for MIMO systems, *IEEE Malaysia International Conference on Communications (MICC)*, Kuala Lumpur, 2009, pp 234-237.

- Zhi-Hou Lee and Wee Gin Lim,"Hybrid methods in MIMO blind equalization and source separation, *IEEE Symposium on Industrial Electronic and Applications (ISIEA)*, 2010, pp 5-9.

- Zhi-Hou Lee, and David Wee-Gin Lim, "Energy normalized cross-correlation constant modulus algorithm in a MIMO convolutive system, *IEEE 10th International Conference on Signal Processing (ICSP)*, 24-28 Oct 2010, pp. 1445-1448.