Channel Estimation and Self-Interference Cancellation in Full-Duplex Communication Systems

Abbas Koohian Mohammadabadi

May 2019

A thesis submitted for the degree of Doctor of Philosophy
of The Australian National University

Research School of Electrical, Energy, and Materials Engineering
College of Engineering and Computer Science
The Australian National University

Declaration

The contents of this thesis are the results of original research and have not been submitted for grant of a higher degree to any other university or institution.

The work in this thesis has been published or has been submitted for publication as journal papers or conference proceedings. These papers are:

Journal articles:

J1. **A. Koohian**, H. Mehrpouyan, A. A. Nasir, S. Durrani, and S. D. Blostein, "Superimposed signaling inspired channel estimation in full-duplex systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 8, Jan. 2018. [Online]. Available: https://doi.org/10.1186/s13634-018-0529-9.

J2. **A. Koohian**, H. Mehrpouyan, A. A. Nasir, and S. Durrani, "Joint channel and phase noise estimation for mmWave full-duplex communication systems," EURASIP Journal on Advances in Signal Processing, vol. 2019, no. 1, p. 18, Mar 2019. [Online]. Available: https://doi.org/10.1186/s13634-019-0614-8

Conference paper:

C1. **A. Koohian**, H. Mehrpouyan, A. A. Nasir, S. Durrani, and S. D. Blostein, "Residual self-interference cancellation and data detection in full-duplex communication systems," in *Proc. IEEE ICC*, May 2017, pp.1–6.

Prepared for submission: The following manuscript has also been prepared based

on the material of this thesis for future submission.

A. Koohian and S. Durrani, "Self-interference suppression in full-duplex massive MIMO communication," *To be submitted*

The following publication is also the result from my PhD study but is not included in this thesis:

A. Koohian, H. Mehrpouyan, M. Ahmadian, and M. Azarbad, "Bandwidth efficient channel estimation for full-duplex communication systems," in *Proc. IEEE ICC*, Jun. 2015, pp. 4710–4714.

The research work presented in this thesis has been performed jointly with Assoc. Prof. Salman Durrani (The Australian National University, Australia), Dr. Hani Mehrpouyan (Boise State University, USA), Dr. Ali Nasir (King Fahd University of Petroleum and Minerals, Saudi Arabia), and Prof. Steven D. Blostein (Queen's University, Canada). The substantial majority of this work was my own.

Abbas Koohian
Research School of Electrical, Energy, and Materials Engineering
The Australian National University,
Canberra, ACT 2601
Australia.

Acknowledgements

The work presented in this thesis would not have been possible without the support of a number of individuals and organizations and they are gratefully acknowledged below:

- First and foremost, I would like to thank my wife and my dearest friend Nozhan for her understanding, patience, kind care, and encouragement. I would also like to thank my father Ata, and my mother Shahrbanoo, for always having my back and supporting me in every possible way. They have always been great sources of inspiration to me.

- I would like to express my sincere thanks to my supervisors, Assoc. Prof. Salman Durrani, and Dr. Hani Mehrpouyan and Dr. Ali Nasir. I am very much grateful of their guidance, support and encouragement throughout my PhD studies. They have taught me incredibly valuable lessons both in my research and personal life. I have had many in-depth research discussions with them, which made my whole PhD journey more enjoyable. They have been great role models for me, and I will never forget the pleasure of spending precious research time with them.

- I would also like to thank Dr. Taneli Riihonen (Tampere University of Technology, Finland). I had many meaningful research discussions with him and was taught a fair bit of knowledge in statistics, and signal processing by him. He has always been welcoming and easily approachable for research discussions.

- It is my great pleasure to study in the Communications research group at the Research School of Electrical, Energy, and Materials Engineering. I would

like to thank everyone for making this group a friendly and relaxing research environment. In particular, I would like to thank Assoc. Prof. Xiangyun (Sean) Zhou and Dr. Nan (Jonas) Yang for giving me the opportunity to deliver tutorials for digital communication and wireless communication courses.

- Thanks must go to the Australian government and the Australian National University for funding me through research training program (RTP). I am also grateful for the CECS top-up scholarship granted to me by the College of Engineering and Computer Science (CECS).

Abstract

Full-duplex (FD) wireless communications, along with millimeter wave (mmWave), and massive multiple-input multiple-output (MIMO) are key technologies for future communication networks, known as 5G networks. The main challenge in exploiting the full potential of FD communication systems lies in cancellation of strong self-interference (SI) signal. In particular, since SI cancellation requires accurate knowledge of both SI and communication channels, bandwidth efficient channel estimation techniques are of practical interest. Furthermore, SI cancellation encounters new challenges, once FD technology is combined with mmWave or massive MIMO technologies. This is because FD communication at mmWave frequencies needs to be able to deal with fast phase noise (PN) variation, and FD massive MIMO base station (BS) requires simultaneous cancellation of SI and multi-user interference (MUI).

The first half of this thesis investigates channel estimation techniques to simultaneously estimate both SI and communication channels for FD communication at microwave and mmWave frequencies. We first consider FD communication at microwave frequencies and inspired by superimposed signalling, we propose a novel bandwidth efficient channel estimation technique for estimating the SI and communication channels. To evaluate the performance of the proposed estimator, we derive the lower bound for the estimation error, and show that the proposed estimator reaches the performance of the bound. In contrast to microwave frequencies, at mmWave frequencies the challenge lies in jointly estimating the channels and tracking the fast varying PN process. We address this problem by proposing an Extended Kalman filter to jointly estimate the channels and track the PN process. We derive a lower bound for the estimation error of PN at mmWave, and numerically show that the mean square error performance of the proposed estimator approaches the lower bound.

The second half of this thesis focuses on the SI cancellation and data detection

problems. The ultimate goal of SI cancellation in FD communication is to allow reliable data detection. However, achieving perfect SI cancellation is not always feasible. This is because accurate channel estimates might not be available. In this regard, we investigate blind data detection problem, when only statistical properties of SI and communication channels are available. We propose a maximum aposterior probability (MAP) based blind detector, which allows for data detection without channel estimation and SI cancellation stages. This blind detection is achieved by using the statistical properties of the SI and communication channels instead of accurate channel estimation and SI cancellation. Finally, we rigorously study precoder design for a FD enabled massive MIMO BS. The main design challenge in here is to design precoders that can simultaneously cancel SI and MUI. We prove that in order to suppress both SI and MUI, the number of transmit antennas must be greater than or equal to the sum of the number of receive antennas and the number of uplink users. In addition, we rigorously show that the problem of simultaneous suppression of SI and MUI has a solution with probability 1. These results validate previous heuristic assumptions made in the literature.

List of Acronyms

ADC	analog to digital converter
BCRLB	Bayesian Cramér-Rao lower bound
BER	bit error rate
BPSK	binary phase-shift keying
BS	base station
CSI	channel state information
DC	digital cancellation
DL	downlink
EKF	extended Kalman filter
EM	expectation maximization
FIM	Fisher information matrix
FD	full duplex
FDD	frequency division duplexing
HD	half duplex
i.i.d.	independent and identically distributed
LHS	left hand side
LMS	least mean square
LoS	line of sight
LGMs	linear Gaussian models
LTE	long term evolution
MAP	maximum aposterior probability
ML	maximum likelihood
MSE	mean square error
MMSE	minimum mean square error
mmWave	millimeter wave
MIMO	multiple-input multiple-output
MUI	multi-user interference
M-PSK	M phase-shift keying
NLoS	non-line of sight
OFDM	orthogonal frequency division multiplexing
PN	phase noise
QAM	quadrature-amplitude modulation

RF	radio frequency
RHS	right hand side
SI	self-interference
SINR	signal to interference plus noise ratio
SIR	signal to interference ratio
SNR	signal to noise ratio
SISO	single-input single-output
TDD	time division duplexing
UL	uplink
w.p.	with probability

Notations and Symbols

j	$\sqrt{-1}$
a	Scalar
$\mathbf{a}$	Column vector
$\mathbf{A}$	Matrix
$[\mathbf{A}]_{m,n}$	Element in the mth row and nth column of $\mathbf{A}$
$(\cdot)^{\mathrm{T}}$	Vector and matrix transpose operator
$(\cdot)^{\dagger}$	Vector and matrix complex conjugate transpose operator
$\det(\mathbf{A})$	Determinant of a square matrix
$\mathrm{diag}(\mathbf{a})$	Square matrix with $\mathbf{a}$ as the main diagonal
$\mathrm{Tr}(\mathbf{A})$	Trace of a square matrix
$\|\cdot\|$	Euclidean norm of a vector
$\mathbf{0}_N$	All-zeros column vector of length N
$\mathbf{0}_{N\times N}$	All-zeros matrix of length $N \times N$
$\mathbf{1}_N$	All-ones column vector of length N
$\mathbf{1}_{N\times N}$	All-ones matrix of length $N \times N$
$\mathbf{I}_N$	Identity matrix of size $N \times N$
$\mathbb{E}[\cdot]$	Expectation operator
$\Re\{\cdot\}$	Real part
$\Im\{\cdot\}$	Imaginary part
$\odot$	Hadamard product
$\mathbb{R}^N$	N dimensional real vector space
$\mathbb{R}^{N\times M}$	$N \times M$ dimensional real matrix space
$\mathbb{C}^N$	N dimensional complex vector space
$\mathbb{C}^{N\times M}$	$N \times M$ dimensional complex matrix space
$f_Y(\cdot)$	The probability density function (PDF) of a random variable Y
$p_X(\cdot)$	The probability mass function (PMF) of a random variable X
$\mathcal{CN}(\mu, \sigma^2)$	A complex Gaussian distribution with mean μ and variance σ^2
z^*	Complex conjugate of complex number z
$\|z\|$	The absolute value of complex number z
$\angle z$	The phase of the complex number z

$\mathrm{g}(\cdot)$	Function of real or complex number
$e^{j\boldsymbol{\theta}}$	The multivariate complex exponential function
$\sum_{\sim i}$	Summation over all possible values except i
$\mathbb{I}\{\cdot\}$	indicator function

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The next generation of wireless communication networks, i.e., 5G networks, are promising faster, more reliable, and better quality communications [1–3]. 5G networks are expected to allow for many new wireless applications, which were deemed impossible with older generation networks [4–6]. It is expected that these networks will be significantly different from their predecessor networks and introduce a paradigm shift in wireless communication research [7–9]. This paradigm shift requires technologies to revolutionize the performance of the current networks in terms of data rates, latency, connectivity, and energy efficiency. This expected technological leap then allows for the realization of important applications such as internet of things (IoT), and virtual reality [10]. In this regard, many technologies have been proposed for 5G networks and among the most prominent ones are full duplex, millimeter wave (mmWave), and massive multiple-input multiple-out (MIMO) technologies [11–13].

In the remainder of this chapter we first review each of these technologies individually. We then identify the scope of this thesis, and subsequently, review the literature of the problems that will be discussed in this thesis. Finally, this section is concluded by discussing contributions and organization of this thesis.

1.1.1 Full Duplex Communication

Traditionally, communication systems are designed to communicate in half duplex (HD) mode. This means that separate time and frequency resources are allocated to transmission and reception [14]. This separation in time and frequency is achieved by using time division duplexing (TDD), and frequency division duplexing (FDD), respectively. However, the inefficiency of allocating different resources to transmission and reception has encouraged research in FD communication, where devices are allowed to transmit and receive over the same temporal and spectral resources [15, 16]. FD communication can potentially double the spectral efficiency of future wireless communication systems [17–19].

FD communication has been used in the context of wired communication, where an interference signal is generated due to the electromagnetic coupling between the receive and transmit wires [20]. This electromagnetic coupling creates a channel, which is known as the echo channel, and is normally weaker than the communication channel [20, 21]. In the context of wireless communication however, the interference is due to in-band transmission and reception. The transmitted signal intended for the destination node is also received by all the receiver antennas of the transmitter. Consequently, the interference channel is significantly stronger compared to wired communication. This interference channel is known as self-interference (SI) channel and it is a major obstacle in realising the full potential of wireless FD communication systems, i.e., doubling the spectral efficiency [22–25].
Fig. 1.1 shows the transceiver structure of a FD communication device as proposed in [26]. The transmitter chain of the structure contains similar functional blocks to a HD device, and includes a digital to analog converter (DAC), low pass filter (LPF), IQ mixer, variable again amplifier (VGA), and power amplifier blocks. However, in comparison to hardware architecture for HD devices, the receiving chain additionally includes analog and digital cancellation blocks to deal with the strong SI signal. The figure also illustrates a link between the transmit and receive chains. This link is necessary to reliably feed the self-interference signal back to the receive chain, and subsequently use this reliable signal to cancel the SI signal received through the receiving antenna, i.e., the wireless channel.

The received SI signal through the wireless channel has a significantly higher

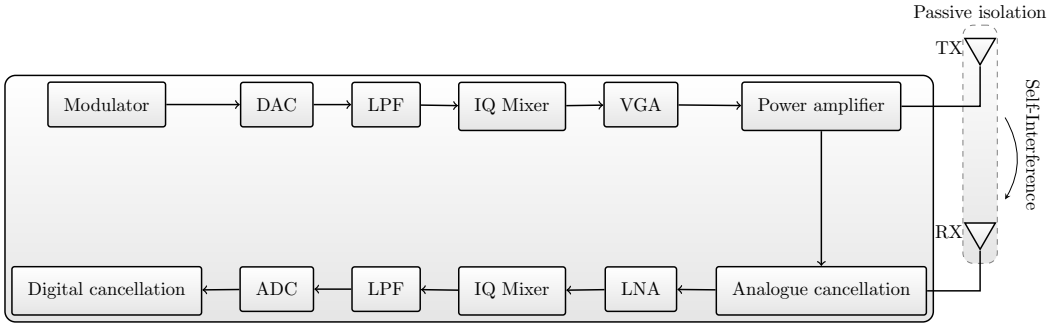


Figure 1.1: Transceiver structure of a FD device, where DAC: digital to analog converter, LPF: low pass filter, VGA: variable gain amplifier, LNA: low noise amplifier, and ADC: analog to digital converter.

power compared to the intended signal. This is because the SI signal travels a smaller distance in orders of centimetres to reach the receiving antenna, as opposed, to the intended signal which might be travelling hundreds of meters to reach the receiving antenna. If this strong SI signal is left untreated it will saturate the ADC and result in total loss of the intended signal. For example, if 8-bit ADC is used, the quantization error is approximately given by 48 dBm. On the other hand, the received SI power can be in order of 100 to 150 dB [15, 27]. This means that the received SI power will dominate the quantization error, and the accuracy of the ADC is predominantly determined by the received SI power.

Recently, there has been a lot of interest in designing SI cancellation techniques to deal with the strong SI signal [15,26–31]. Different techniques for SI cancellation can be divided into two main categories [15]: (i) passive suppression in which the SI signal is suppressed by suitably isolating the transmit and receive antennas [15,27, 32], and (ii) active cancellation which uses knowledge of the SI signal to cancel the interference in either the analog domain (i.e., before the signal passes through the analog-to-digital converter (ADC) [15,28,33] and/or the digital domain [26,29,30].

Passive suppression techniques normally require directional and polarized antennas to deal with the strong SI signal [28]. For example, the authors of [15] use polarized antennas to provide 40 dB attenuation. Furthermore, the authors of [31] propose a passive suppression technique to provide for up to 60 dB suppression of the SI signal. Analog suppression techniques use the knowledge of SI channel to suppress the SI signal via an analog circuitry. For example, the authors of [26, 29]

propose a direct conversion FD transceiver architecture to suppress the SI signal using an analog circuit. In the proposed FD architecture of [26, 29], the SI signal is fed back to the receiver chain via a wired link. This fed-back SI signal is then used to cancel the SI signal receive through the wireless channel before the ADC block. Furthermore, a transceiver architecture is proposed for FD communication in [30], where the wired link that feeds back the SI signal shares oscillator with the receiver chain. The proposed architecture is then shown to have a better robustness to oscillator impairments compared to the architecture proposed by [26, 29].

Depending upon the design, passive suppression and analog cancellation can provide about $40 - 80$ dB cancellation in total [31], which is not sufficient for reliable communication. Hence, in practice, the SI is cancelled in multiple stages, beginning with passive suppression and followed by cancellation in the analog and digital domains.

1.1.2 Millimeter Wave Communication

MmWave networks are expected to explore the large unprecedented frequency bands available at mmWave to provide fast and directional communication [34]. For example, the available unlicensed spectrum at 60 GHz is ten times larger than the spectrum used for 4G cellular communication, which can be used to provide faster communications compared to the current cellular and WiFi networks [35]. A number of successful experiments with mmWave communication has persuaded the regularity bodies such as U.S. Federal Communication Commission (FCC) to allow for the use of mmWave spectrum for 5G communications [36].

MmWave communication is one of the key technologies, which allows many important fantasy applications, such as IoT, drive-less cars, and smart cities, to become realities [34]. However, there are certain impairments associated with mmWave communication that can significantly affect quality of reliable communications at mmWave band frequencies [35–37]

Firstly, due to short wave length of mmWave frequencies, the signal power is largely attenuated before reaching receiving antennas [35, 36]. This is because diffraction and material penetration at mmWave frequencies cause more signal attenuation in contrast to microwave frequencies, i.e., signal blockage [37]. Con-

sequently, a communication system operating at mmWave frequencies requires to rely more on line-of-sight (LoS) communication, which is not always feasible due to physical constraints. A large body of research investigates the problem with short effective range for mmWave communication [38–40]. The principle idea behind the proposed techniques is to explore the small size of antennas at mmWave frequencies to pack more antennas into the antenna array. Once antenna arrays are constructed for mmWave communication, then beamforming can be used to increase the effective communication range at mmWave frequencies. However, two types of beamforming have received more attention compared to the other types due to their simplicity and practical feasibility: (i) analog beamforming [40], and (ii) hybrid beamforming [38, 39]. Analog beamforming does not require channel knowledge and it involves using phase shifters to steer the signal into the desired direction. Despite, the simplicity of analog beamforming techniques, they can only steer the transmission in a single direction. Hybrid beamforming allows for directional transmission in multiple directions, however, they require channel knowledge. Hybrid beamforming is achieved by using analog and digital beamforming techniques in combination. The advantage of hybrid beamforming techniques compared to fully digital beamforming techniques is that they allow for directional communication with less complexity [38].

Secondly, the oscillators designed to operate at mmWave frequencies suffer from a large phase noise (PN) [41, 42]. If the oscillator PN is not accurately tracked and eliminated, it will have detrimental effect on the communication performance at mmWave frequencies. This is because PN causes the constellation diagram to rotate [43]. This rotation in turn results in erroneous detection. For a HD single-input single-out (SISO) system the PN problem at mmWave bands has been thoroughly investigated in [44], where the authors derive the lower bound for the PN estimation error and show that the proposed estimation algorithm reaches the performance of the derived bound. A joint channel and PN estimation algorithm is proposed in [43] for a HD MIMO communication system. The performance of the estimator proposed in [43] for HD MIMO system is improved in [45] by using a soft-input estimator.

Moreover, since for hybrid and digital beamforming require the knowledge of the channels, the importance of PN estimation problem is compounded. This is

because without an accurate PN tracking mechanism, accurate channel knowledge cannot be obtained.

1.1.3 Massive MIMO

Massive MIMO is a physical layer technology, which proposes to deploy a large antenna array at base stations (BS) to serve smaller number of users in the downlink [46, 47]. Massive MIMO systems are expected to significantly improve the performance of current MIMO systems. Furthermore, the use of large antenna arrays at BS allows for effective way of dealing with the wireless channel impairments including fast fading effect [46].

Massive MIMO systems promote network architectures, where minimum processing is required by the users. This is achieved by using TDD and assuming channel reciprocity for downlink and uplink [47]. In this case, only the BS needs to know the channels, and since the uplink and downlink channels are the same, beamforming matrices can be designed at the BS to direct the communication in the downlink and remove the interference in the uplink. The advantage of this network architecture is that the complexity of the system only grows with the number of users and is independent of number of BS antennas.

Despite the potential benefits of massive MIMO systems, important drawbacks are associated with these networks. In particular, [46] shows how pilot contamination can act as a bottleneck and does not allow for exploiting the full potential of HD massive MIMO. Pilot contamination is a term used to describe the inaccuracy of the channel estimates, when a set of pilots are shared and transmitted by a number of base stations in a cellular network [48]. The phenomenon of pilot contamination emerges in massive MIMO communication because of the TDD based nature of these networks. In TDD based networks, the length of the pilots are limited by the coherence of the channel. Since the length of pilot sequences are limited, a large set of independent sequences cannot be generated, and hence, the pilots need to be reused in the network. However, reusing the pilots causes an interference for channel estimation, which in turns results in inaccurate channel estimates. A large body of research has investigated the pilot contamination problem for HD massive MIMO systems [48–52].

The second associated problem with massive MIMO systems, which is of practical importance, is the problem of low complexity channel estimation and data detection algorithms [53, 54]. The importance of low complexity algorithms for massive MIMO systems is two fold, not only they need to be low complexity to be feasible for large antenna systems, but also they need to be low complexity to limit the power consumption of the system [54]. For this reason, linear processing is more desirable for massive MIMO systems [47].

In the following section, we discuss the focus of this thesis. In particular, we show how this thesis considers solutions for the problems involving FD, mmWave, and massive MIMO technologies.

1.2 Research Challenges and Motivation

While this thesis primarily focuses on FD communication, it also investigates the FD communication at mmWave frequency bands, and FD massive MIMO BS. In the remainder of this section we present the problems that will be discussed in this thesis and their relevant literature review.

In this thesis, we focus on the SI cancellation and data detection after the passive suppression and analog cancellation, termed *residual SI*. Since, the effectiveness of any digital interference cancellation technique depends strongly on the quality of the available channel estimates for both the SI and desired communication channels [55–57], we first study the channel estimation problems for FD communication., and then study the problems of residual SI cancellation and data detection based on available channel estimates.

Section 1.1.1 has identified that the main obstacle in enabling FD communication in 5G networks is the SI cancellation. We focus on the SI cancellation and data detection after the passive suppression and analog cancellation, termed residual SI. Since, the effectiveness of any digital interference cancellation technique depends strongly on the quality of the available channel estimates for both the SI and desired communication channels [55–57], we first study the *channel estimation problems for FD communication.* In particular by focusing on FD communication

at microwave frequencies, we formulate the first problem considered in this thesis as follows:

- *Problem 1:* Bandwidth efficient channel estimation techniques for FD communication systems.

Section 1.1.2 has highlighted oscillator PN as an important issue associated with HD mmWave communication systems. However, this issue is compounded for FD communication systems. This is because for FD communication the channel estimation and PN tracking needs to be done in the presence of a strong SI signal. In addition, since PN is a hardware impairment the algorithms developed for PN tracking for HD devices cannot be used for FD devices, which have fundamentally different hardware architecture as shown in Fig. 1.1. Consequently, we formulate the second problem considered in this thesis as follows:

- *Problem 2:* Joint channel and PN estimation algorithms for FD communication at mmWave frequencies.

Once the channels are estimated then they are used to: (i) cancel the strong SI to below the receiver noise floor, and (ii) detect the desired communication symbols. For this reason, this thesis also investigates the problem of *SI cancellation and data detection* based on the availability of channel estimates. In particular, since obtaining accurate channel estimates is not always feasible, this thesis seeks to address the following problem:

- *Problem 3:* Data detection for FD communication using statistical properties of the channels.

Section 1.1.3 has discussed the problem associated with HD massive MIMO systems. While a large body of research studies the associated problems of HD massive MIMO, achieving FD communication with massive MIMO BS is still challenging. This is because even in the presence of accurate channel estimates, SI cancellation is a very challenging problem for massive MIMO BS. Furthermore, this challenge becomes even harder in the context of cellular communication, when both SI and MUI should be cancelled simultaneously. Hence, we formulate the last problem considered in this thesis as follows:

- *Problem 4:* Precoder design for joint SI and MUI suppression for FD massive MIMO BS.

We can put the considered problems in this thesis into two groups: (i) channel estimation problems, and (ii) SI cancellation and data detection problems. In the remainder of this section, we review the literature of the four problems presented here, and provide background and motivation for each problem.

1.2.1 Channel Estimation

Accurate channel estimates are fundamentally important for reliable wireless communication [58–62]. The accuracy of these estimates becomes even more important in FD communication [15, 55, 63]. This is because these estimates serve two purposes. On the one hand, the estimate of communication channel is needed to reliably detect the transmitted symbols, and on the other hand, the SI channel estimation is needed to cancel the residual SI signal at the baseband.

The channel estimation for FD communication at microwave frequency differs substantially from the channel estimation at millimeter wave (mmWave) communication [64,65]. This is because the mmWave channel suffers from the rapid variation of phase noise (PN) [66]. In this thesis we first consider the channel estimation at microwave frequencies, where the channel is not affected by the random variation of the PN. Given the importance and emergence of mmWave communication systems, we then study the joint channel and PN estimation problem at mmWave frequencies.

In what follows the literature of channel estimation techniques for FD communication at microwave and mmWave is reviewed.

1.2.1.1 Channel Estimation at Microwave Frequencies

Conventionally in FD communication, the baseband channels are estimated by using *data-aided channel estimation techniques*, where a portion of the data frame is allocated for known training sequences or pilot symbols [56, 67–69]. In this regard, a maximum-likelihood (ML) approach was proposed in [56] to jointly estimate the residual SI and communication channels by exploiting the known transmitted SI

symbols, and both the known pilot and unknown data symbols from the other intended transceiver. Another approach was proposed in [70] where a sub-space based algorithm was developed to jointly estimate the residual SI and communication channels, using known pilots.

Bandwidth efficient channel estimation techniques for FD communication can be inspired from the HD counterparts. A bandwidth efficient channel estimation technique in HD systems is *superimposed training* [71, 72]. The authors of [72] propose a superimposed training based channel estimation technique for HD communication. The proposed technique alleviates the need to explicit time slots allocation for channel estimation, by superimposing a periodic low power training sequence onto the data symbols. The downside of this approach is that some power is consumed in superimposed training which could have otherwise been allocated to the data transmission. This lowers the effective signal-to-noise ratio (SNR) for the data symbols, and affects the bit error rate (BER) at the receiver.

In contrast to data-aided and superimposed training based channel estimation techniques, *blind techniques* avoid the use of pilots altogether by exploiting statistical and other properties of the transmitted signal [73–78]. A blind channel estimation technique for HD communication systems is proposed in [73]. The authors consider a single-input multiple-output (SIMO) system, and through identifiability analysis show that blind estimation can only recover the channel coefficients up to a scaling factor, i.e., ambiguous channel estimation. This means that the channel phases cannot be fully recovered. To overcome the ambiguity problem, semi-blind approaches, where the HD channel is estimated using the combination of known pilots and unknown data symbols, are proposed in [79, 80]. Moreover, in the context of cooperative communication blind and semi-blind channel estimation techniques have been proposed in [74, 75], for a HD two way relay network, respectively.

The above literature review motivates the study of bandwidth efficient channel estimation techniques for FD communication. Furthermore, it indicates that this study should include an investigation of ambiguity problem, which is associated with the estimation techniques that do not use conventional data-aided piloting.

1.2.1.2 Channel Estimation at mmWave Frequencies

Recently, SI channel measurements have been carried out for FD communication at mmWave frequencies [81, 82]. The measurements indicate that, as opposed to the microwave frequency band, the SI channel at mmWave has a non-line-of-sight (NLoS) component, which cannot be cancelled using passive and active analog suppression techniques. This partial suppression of the SI signal results in a large residual SI signal at baseband, which is still significantly higher than the receiver noise floor [81]. Digitally cancelling the residual SI signal at the baseband, requires tracking the large and rapidly changing oscillator PN [66, 83]. The existing techniques for residual SI signal cancellation at baseband assume a very steady oscillator PN [70, 84–86], and hence, they cannot be used for FD mmWave communication.

The authors in [66] consider a FD communication system, where the digital SI cancellation is hampered by oscillator PN. However, the authors assume that the PN stays constant over a number of orthogonal frequency division multiplexing (OFDM) symbols. This assumption is in contrast to fast variation of PN at mmWave communication [43]. Furthermore, channel and PN estimation problems for FD communication is also investigated in [87], where the authors propose two separate maximum likelihood (ML) estimators for channel and PN estimation. The proposed ML estimators require two separate set of piloting. This makes them less bandwidth efficient compared to the joint channel and PN estimation techniques. In addition, the authors of [84] assume a slow varying PN process, which is a valid assumption for FD communication at microwave frequency. For a single-input single-output (SISO) FD communication system the PN estimation problem is considered in [86]. Beside the SISO system assumption, there are two other limiting factor associated with the proposed PN estimator in [86]. Firstly, the ML based estimator is complex and cannot be directly applied to more general system models, such as MIMO FD communication system, and secondly, the authors assumptions about the PN process is only valid when the PN variations are small.

The above mentioned works motivate the study of joint channel and PN estimation problem for FD communication system at mmWave frequency band, where the PN variations are significantly larger compared to the microwave frequencies.

1.2.2 SI Cancellation and Data Detection for FD Communication

Ultimate goal of any communication system is to reliably detect the transmitted symbols to allow robust end-to-end communication [88–90]. For this reason, once the channels are accurately estimated, for reliable data detection, the SI signal needs to be suppressed to below the receiver noise floor. However, the desired SI suppression cannot be achieved using a single cancellation technique. For instance, for small-cells in Long Term Evolution (LTE), the maximum transmit power is typically 23 dBm (200 mW) and the typical noise floor is -90 dBm [91]. Ideally, this requires a total of 113 dB SI cancellation for realizing the full potential of FD systems [91]. While passive and active RF cancellation can provide some cancellations, the residual SI can still be relatively strong in the baseband digital signal. For example, for the LTE small-cell example, it can be as high as 50 dB assuming state-of-the-art passive suppression and analog cancellation provide 60 dB of the total required SI cancellation of 113 dB. Thus, accurate digital SI cancellation is also required to bring the SI as close to the noise floor as possible.

However, achieving effective digital SI cancellation might not always be feasible. This can be due to the lack of availability of accurate channel estimates. In this case, it is important to investigate data detection techniques for FD communication, which only require statistical properties of the channels instead of accurate channel estimates. The detection techniques that work with empirical statistical knowledge of the channels do not need channel estimation and SI cancellation stages. Hence, they can save both bandwidth and processing power.

Moreover, even with the perfect channel knowledge, SI cancellation and data detection in cellular networks with FD enabled massive MIMO BS imposes new challenges that are not encountered in point to point FD communication systems. In particular, the problem of spatial suppression of the SI signal via transmit precoding is both interesting and challenging. This is because there are two major challenges associated with precoder design for FD enabled massive MIMO BS: (i) the challenge to overcome multi-user interference (MUI) in the downlink (DL), which is caused by the inherent multi-user communication [14], and (ii) the challenge to overcome SI in the uplink (UL), which is caused by the simultaneous

in-band transmission (in the DL) and reception (in the UL) [92, 93].

This thesis studies the SI cancellation and data detection for both point-to-point FD communication, and cellular network with FD massive MIMO BS. In what follows the SI cancellation and data detection literature for these systems is reviewed.

1.2.2.1 Blind Data Detection for Point-to-Point FD System

Conventionally, for data detection, the SI signal is cancelled using a replica of the SI signal and available SI channel estimates, which are obtained through pilot-based estimation techniques [15, 28, 29, 84].

For SI cancellation and data detection in a SISO FD communication system, the authors of [15, 16] propose an initial HD phase, in which the transmitting node is switched off, while pilots are sent from the receiving node for SI channel estimation. The SI channel is then estimated using a least square (LS) based estimator. The LS-based channel estimates are subsequently used to cancel the residual SI signal at the baseband and detect communication signals. However, not only does the initial HD phase for SI channel estimation waste the bandwidth, but it also introduces processing delays. The associated problem with the initial HD phase needed for SI cancellation for FD SISO system proposed by [15, 16], is resolved in [67]. The authors of [67] allow for the transmitting node to transmit during the SI channel estimation phase, however, they treat the intended received signal as noise, which significantly degrades the quality of channel estimates and subsequently the reliability of the detected symbols. The reliability of the detected symbols is improved in [94, 95], by obtaining the SI channel estimates through an adaptive least mean square (LMS) algorithm. Although, in contrast to [15, 16], the proposed SI cancellation and data detection technique does not need an initial HD phase, the proposed algorithm requires many iterations to converge.

SI cancellation and data detection for MIMO FD communication is considered in [70, 84]. The authors in [84] use extensive two-stage piloting to obtain the estimates of the SI channel for digital SI cancellation and data detection. The extensive two-stage piloting of [84] is omitted in [70]. This is done by exploiting the sparsity of the channel. Despite better bandwidth efficiency, the proposed

technique in [70] still requires significant processing to obtain the estimates of both SI and communication channel. This in turn adds undesirable processing delays.

Blind data detection techniques, where the data symbols are detected without obtaining channel estimates, have been investigated for HD communication systems [96–98]. In [96], an iterative sequential Monte Carlo based method is proposed for blind data detection in the presence of PN in HD communication. The authors have shown how blind detection cannot recover the phase of the transmitted data symbols, and have used some channel estimates to remove this ambiguity. In [97], the authors propose to use cyclic prefix to avoid the ambiguity of blind data detection for HD OFDM systems. Finally, the authors of [98] consider blind detection problem for a HD two way relay network, and investigate the ambiguity problem associated with it.

The above mentioned works clearly show that the channel estimation and SI cancellation stages require significant processing and bandwidth resources. This makes the problem of data detection using statistical properties of the channels, instead of actual channel estimates (blind data detection), of great practical importance and interest. Furthermore, any such study also needs to address the ambiguity problem associated with blind data detection.

1.2.2.2 SI Cancellation for Cellular Network with FD Massive MIMO Base Station

Point-to-point communication systems are considered in the reviewed works in Section 1.2.2.1. However, the SI cancellation for FD MIMO communication systems in the context of cooperative and cellular communications has been studied in [92, 99–102].

In the context of cooperative communication, the authors of [92] consider a MIMO FD relay and propose two different SI cancellation techniques based on antenna and beam selection techniques. However, the issue associated with both the techniques is the complexity associated with selecting a set of antennas or beams to suppress the SI signal. While the proposed techniques completely suppress the SI signal for MIMO system with low number of antennas, they might not be directly applicable to massive MIMO systems, where the BS is expected to

have hundreds of antennas. The authors of [99] take a non-linear approach in designing precoder matrix for MIMO FD relays to jointly perform beamforming and SI suppression. However, the proposed non-linear approach, which is based on non-convex optimization is computationally complex for massive MIMO BS with large number of antennas.

Recently, in the context of cellular communication, the problem of linear precoder design to suppress both SI and MUI for massive MIMO BS has been investigated [100–102]. In this context, the existing solutions in the literature [100–102], assume the number of transmit antennas is greater than the number of receive antennas and heuristically design the precoding matrices based on the transmission of zeros to all or a subset of the receive antennas. In [102], the authors consider a FD massive MIMO BS, which serves a number of single antenna HD users. A transmit precoder is designed by first adding a number of zeros to the vector of data symbols and making it as large as the number of transmit antennas. Then zero forcing (ZF) based precoder is designed by exploiting the added zeros, such that the receive antenna array receives all the zeros. This work is extended to HD users with MIMO users in [103], which uses the same design procedure for transmit precoder design. Furthermore, the authors of [101] consider an FD massive BS with MIMO FD users, and design precoder matrix to suppress both SI and MUI in two stages. First, assuming that a subset of receive antennas receives zeros, a precoder matrix is designed such that the transmit vector lies in the null space of the interfering channel. Second, a precoder matrix is designed such that the overall effect of applying the first and second precoder matrices removes MUI. However, these works lack a well-established mathematical foundation to allow validating their core central assumptions. In particular, the assumption of being able to send zeros to all or a subset of receive antennas cannot be validated using any existing framework.

The above works motivate the study of precoding design in FD massive BS to provide a rigorous mathematical foundation for the design of linear low complexity precoders.

In what follows, we provide an overview of the thesis and show how this thesis addresses the questions that were not investigated in the existing works.

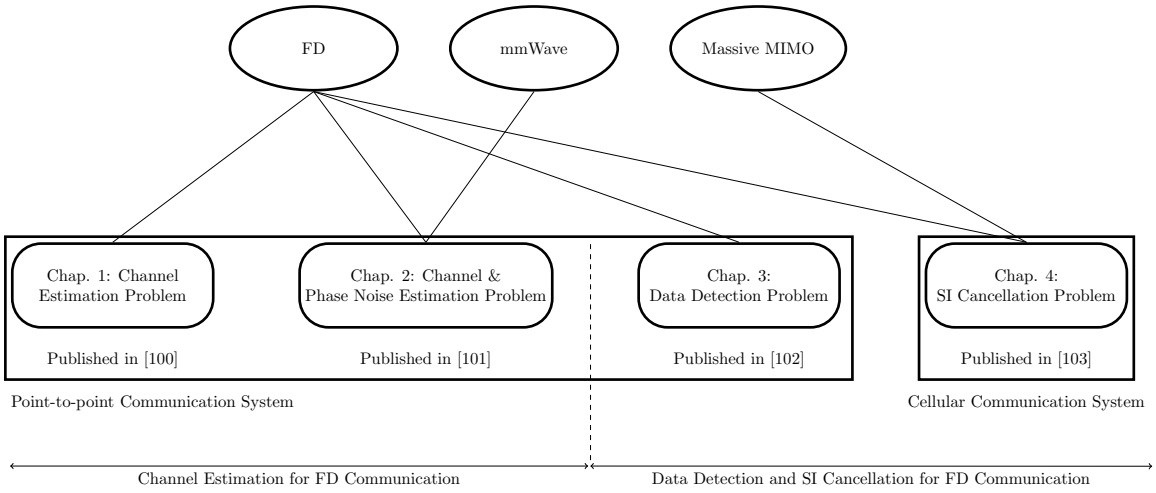


Figure 1.2: Overview of the thesis and its contributions

1.3 Overview and Contributions of Thesis

Fig. 1.2 presents an overview of the thesis. The content of this thesis can be divided into two halves. In the first of half of the thesis the channel estimation problems for FD communication systems are studied. The second half of the thesis is dedicated to the study of SI cancellation and data detection techniques for FD communication systems. In particular, this thesis seeks to provide solution to the existing problems in the literature. More specifically, as outlined in Section 1.1, the following problems are studied in this thesis:

- *Problem 1:* Bandwidth efficient channel estimation techniques for FD communication systems.

- *Problem 2:* Joint channel and PN estimation algorithms for FD communication at mmWave frequencies.

- *Problem 3:* Data detection for FD communication using statistical properties of the channels.

- *Problem 4:* Precoder design for joint SI and MUI suppression for FD massive MIMO BS.

1.3.1 Questions to be Answered

Solutions to the above mentioned problems will answer the following open questions:

Q1. Can both SI and communication channels be estimated without sending data-aided pilot sequences?

Q2. Can the inherent ambiguity of blind channel estimation techniques for FD communication be resolved?

Q3. What is the lower bound on the estimation error estimating both SI and communication channels in a FD communication system?

Q4. How can we jointly estimate the channel and a fast varying PN process for FD communication at mmWave frequencies?

Q5. What is the lowest achievable mean square error (MSE) for joint channel and PN estimation for mmWave FD communication?

Q6. Can the transmitted symbols be detected without SI and channel estimation stages, solely based on the available statistical properties of the SI and communication channels?

Q7. Can transmit precoders be designed to jointly cancel SI and MUI for FD-enabled massive MIMO by sending zeros to all or a subset of receive antennas?

1.3.2 Thesis Contributions and Organization

The first part of this thesis focuses mainly on channel estimation techniques for FD communication (Chapters 2 and 3), while in the second part of the thesis (Chapters 4 and 5), the focus is on SI cancellation and data detection for FD communication.

In this regard, the problem of bandwidth-efficient channel estimation for FD communication is considered in Chapter 2. Chapter 3 discusses the problem of joint channel and PN estimation for mmWave FD communication. In Chapter 4, we study data detection problem when channel estimates are not available, and

only the statistical properties of the SI and communication channels are known. Chapter 5 studies the precoder design for both SI and MUI suppression for FD-enabled massive MIMO BS. The chapter-wise summary of the contributions is given as follows:

Chapter 2 – Superimposed Signalling Inspired Channel Estimation in Full-Duplex Systems

In Chapter 2, we consider the problem of bandwidth efficient channel estimation in a SISO FD communication system. We propose a new technique for channel estimation and residual SI cancellation in FD systems. Our approach draws inspiration from (i) blind channel estimation techniques in that we examine the condition for identifiability of channel parameters in FD systems, and (ii) superimposed signalling in that we superimpose (i.e., add) a constant real number to each constellation point of the modulation constellation. However, our proposed technique is distinct from superimposed signalling. In superimposed signalling, the superimposed signal is typically a periodic training sequence that is added to the data signal after the data symbols are modulated. Hence, the additional power of the superimposed signal is only used for channel estimation. In our proposed technique, the superimposed signal is a constant (non-random) signal and the objective is to shift the modulation constellation away from the origin, which we exploit for estimating the SI and communications channels without ambiguity. In addition, the additional power of the superimposed signal is used for both modulating the data symbols and channel estimation, which does not reduce the effective SNR as in superimposed signalling. The novel contributions are as follows:

- We derive the condition for blind identifiability of channel parameters in a FD system (cf. Theorem 2.1) and show that symmetric modulation constellations with respect to the origin cannot be used for ambiguity-free channel estimation in a FD system. Based on Theorem 2.1, our proposed technique is able to resolve the inherent ambiguity of blind channel estimation in FD communication via shifting the modulation constellation away from origin.

- Using the proposed technique, we derive a computationally efficient expectation maximization (EM) estimator for simultaneous estimation of both SI and communication channels. We derive a lower bound for the channel estimation

error, which depends on the energy used for shifting the modulation constellations, and use it to find the minimum signal energy needed for accurate channel estimation in a given FD communication system.

- We use simulations to compare the performance of the proposed technique against that of the data-aided channel estimation method, under the condition that the pilots use the same extra power as the shift. Our results show that the proposed technique performs better than the data-aided channel estimation method both in terms of the MSE of channel estimation and BER. In addition, the proposed technique is robust to an increasing SI power.

The results in Chapter 2 have been presented in the following publication [104], which is listed again for ease of reference:

J1. **A. Koohian**, H. Mehrpouyan, A. A. Nasir, S. Durrani, and S. D. Blostein, "Superimposed signaling inspired channel estimation in full-duplex systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 8, Jan. 2018. [Online]. Available: https://doi.org/10.1186/s13634-018-0529-9.

Chapter 3 – Joint Channel and Phase Noise Estimation for mmWave Full-Duplex Communication Systems

In Chapter 3, we consider the problem of joint channel and PN estimation for a mmWave FD MIMO communication systems. We assume that oscillator PN varies from one communication symbol to another. This assumption best captures the fast variation of PN at mmWave frequencies [43]. The main contributions of this work are as follows:

- We construct a state vector for the joint estimation of the channel and PN, and propose an algorithm based on extended Kalman Filtering (EKF) technique to track the fast PN variation at mmWave band.

- We derive the lower bound on the estimation error of the proposed estimator, and numerically show that the proposed estimator reaches the performance of the lower bound. We also show the effectiveness of a digital SI cancellation, which uses the proposed estimation technique to estimate the SI channel.

- We present simulation results to show the MSE and BER performance of a mmWave FD MIMO system with different PN variances, and signal-to-interference-ratios (SIR). The results show that for a 2×2 FD system with 64− quadrature amplitude modulation (QAM), and PN variance of 10^{-4}, the residual SI power can be reduced to -25 dB and -40 dB, respectively, for SIR of 0 and 15 dB.

The results in Chapter 3 have been presented in the following publication [105], which is listed again for ease of reference:

J2. **A. Koohian**, H. Mehrpouyan, A. A. Nasir, and S. Durrani, "Joint channel and phase noise estimation for mmWave full-duplex communication systems," EURASIP Journal on Advances in Signal Processing, vol. 2019, no. 1, p. 18, Mar 2019. [Online]. Available: https://doi.org/10.1186/s13634-019-0614-8

Chapter 4 – Data Detection in Full-Duplex Communication Systems

In Chapter 4, we focus on the received signal after the passive and RF cancellation stages in a point-to-point FD communication system. Different from existing works, we propose a data detection technique based on superimposed signalling which does not require any channel estimates. We show that superimposed signalling can overcome the inherent ambiguity of blind data detection problem when channel estimates are not used. The main contributions of this work are:

- We formulate a maximum a posterior (MAP) detector, based on the posterior probability distribution (PDF) function of the data, to detect the data symbols in FD communication with no need for obtaining channel estimates (blind detection).

- We show that if the modulation constellation is symmetric around the origin, blind data detection in FD communication results in ambiguity when the channel estimates are not available. We demonstrate that one simple method to resolve this detection ambiguity is to use superimposed signalling, i.e., to shift the modulation constellation away from the origin and create an asymmetric modulation constellation.

- We compare the BER performance of the proposed detection method to that of the conventional channel estimation-based detection method, where the unknown channels are first estimated and then the data signal is detected, under the constraint of same average energy over a transmission block. The results show that the proposed method outperforms the conventional method. Since the proposed method does not require any channel estimates, it enhances bandwidth and power efficiency.

The results in Chapter 4 have been presented in the following publication [106], which are listed again for ease of reference:

C1. **A. Koohian**, H. Mehrpouyan, A. A. Nasir, S. Durrani, and S. D. Blostein, "Residual self-interference cancellation and data detection in full-duplex communication systems," in *Proc. IEEE ICC*, May 2017, pp. 1–6.

Chapter 5 – Self-Interference Suppression in Full-Duplex Massive MIMO Communication

In Chapter 5, we consider a single-cell multi-user scenario with a FD massive MIMO BS. In particular, we focus on the problem of SI suppression via spatial precoding. The main contributions of this work are:

- We develop a mathematical foundation for spatial suppression of SI and MUI.
- We prove that in order to suppress both SI and MUI, the number of transmit antennas must be greater than or equal to the sum of the number of receive antennas and the number of UL user equipments (UEs). In addition, we rigorously show that the problem of simultaneous suppression of SI and MUI has a solution with probability 1. This validates previous heuristic assumptions in the literature [100–102].
- We also study the precoder design for a special case, when the number of transmit antennas is equal to the sum of the number of receive antennas and the number of UL UEs.
- The simulation results demonstrate the validity of the proposed framework and the effectiveness of the SI and MUI suppression via transmit precoding.

The results in Chapter 5 have been prepared for submission in the following manuscript [107], which is listed again for ease of reference:

J3. **A. Koohian** and S. Durrani, "Self-interference suppression in full-duplex massive MIMO communication," *to be submitted*.

Finally, Chapter 6 provides a summary of the thesis results and makes suggestions for future research work.

Chapter 2

Superimposed Signaling Inspired Channel Estimation in Full-Duplex Systems

Residual SI cancellation in the digital baseband of an FD communication system requires an accurate knowledge of the SI channel. Furthermore, robust data detection requires a reliable estimate of the communication channel. Hence, in contrast to HD communication, where only the communication channel is estimated, both SI and communication channels need to be estimated for reliable end-to-end FD communication. Obtaining accurate estimates for both SI and communication channels via piloting wastes a significant portion of the sacred bandwidth. Blind channel estimation techniques estimate the channel with no piloting and are more bandwidth efficient compared to their piloting counterparts, but they have an inherent ambiguity problem.

This chapter of the thesis investigates the problem of bandwidth efficient channel estimation for FD communication. In particular, it mathematically studies the ambiguity problem associated with blind channel estimation techniques for FD communication, and shows how this ambiguity can be resolved. In addition, this chapter proposes a novel bandwidth efficient channel estimation techniques to estimate both the SI and communication channels with no ambiguity.

The remainder of this chapter is organised as follows. The system model is

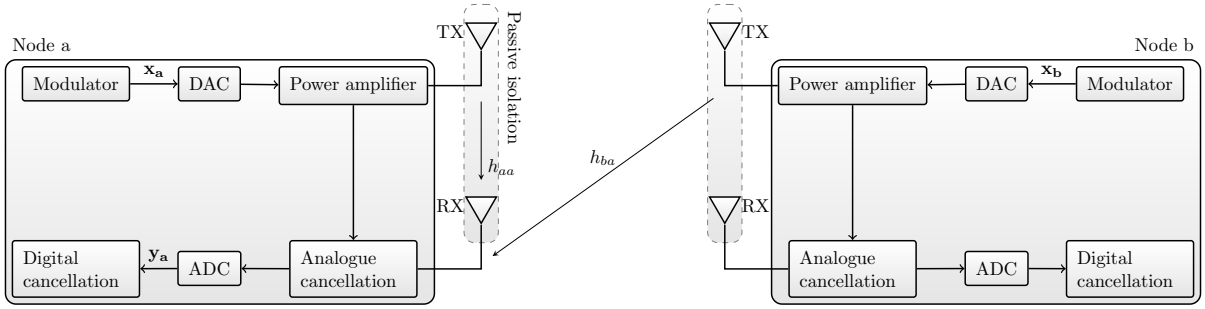


Figure 2.1: Illustration of full duplex communication between two transceivers, each with a single transmit and a single receive antenna. ADC = analog to digital converter, DAC = digital to analog converter. TX = transmit. RX = receive.

presented in Section 2.1. The bandwidth efficient channel estimation problem and the proposed technique are formulated in Section 2.2. The EM estimator and the lower bound on the channel estimator error are derived in Section 2.3. The performance of the proposed technique is assessed in Section 2.4. Finally, conclusions are presented in Section 2.5.

2.1 System Model

Consider the channel estimation problem for a SISO FD communication system between two nodes a and b, as illustrated in Fig. 2.1. Transceiver nodes a and b are assumed to have passive suppression and analog cancellation stages and we only consider the digital cancellation to remove the residual SI, i.e., the SI, which is still present after the passive suppression and analog cancellation. We consider the received signal available at the output of the analog to digital converter (ADC). The received signal at node a is given by[1]

$$\mathbf{y_a} = h_{aa}\mathbf{x_a} + h_{ba}\mathbf{x_b} + \mathbf{w_a}, \tag{2.1}$$

where $\mathbf{x_a} \triangleq [x_{a_1}, \cdots, x_{a_N}]^T$, $\mathbf{x_b} \triangleq [x_{b_1}, \cdots, x_{b_N}]^T$ are the $N \times 1$ vectors of trans-

[1]Note that the system model for FD communication in (2.1) is applicable to per subcarrier communication in orthogonal frequency division multiplexing (OFDM) FD communication [15, 23, 108].

mitted symbols from nodes a and b, respectively, $\mathbf{y_a} \triangleq [y_{a_1}, \cdots, y_{a_N}]^T$ is the $N \times 1$ vector of observations, $\mathbf{w_a}$ is the noise vector, which is modeled by N independent Gaussian random variables, i.e., $f_{\mathbf{W_a}}(\mathbf{w_a}) = \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_N)$, and h_{aa}, h_{ba} are the residual SI and communication channel gains, respectively. Furthermore, we model h_{aa} and h_{ba} as independent random variables that are constant over one frame of data and change independently from frame to frame [55]. Finally, N is the number of transmitted symbols in a given communication packet, i.e., the size of communication packet.

Remark 2.1 *Including all the hardware impairments and unknown parameters in mathematical modeling of parameter estimation problem in FD communication results in a highly non-linear system model, which may not have a tractable solution. The current approach is to separate the estimation of the linear and non-linear parameters [56, 84]. In this work, we focus on the estimation of linear parameters, while the estimation of non-linear parameters can be the topic of future works.*

Modulation Assumptions and Definitions: In this thesis, we assume that the transmitted symbols are all equiprobable and call the set $\mathcal{A} \triangleq \{x_1, x_2, ..., x_M\}$, which contains an alphabet of M constellation points, *a modulation set*. Let $\mathcal{K} \triangleq \{1, \cdots, M\}$ denote set of indices of the constellation points.

We define E as the average symbol energy of a given constellation, i.e.,

$$E \triangleq \mathbb{E}_{X_k}[|x_k|^2] = \frac{\sum_{k=1}^{M} |x_k|^2}{M}, \tag{2.2}$$

where $x_k \in \mathcal{A}$. Note that the average symbol energy can be related to the average bit energy as $E_b \triangleq E/\log_2(M)$.

2.2 Channel Estimation for FD Systems

In this section, we first formulate the blind channel estimation problem for the FD system considered in Section 2.1. Based on this formulation, we present a theorem which provides the necessary and sufficient condition for ambiguity-free channel estimation. Finally, we discuss the proposed technique to resolve the ambiguity problem.

2.2.1 Problem Formulation

Without loss of generality, we consider the problem of baseband channel estimation at node a only (similar results apply at node b). In formulating the problem, we make the following assumptions: (i) the transmitter is aware of its own signal, i.e., $\mathbf{x_a}$ is known at node a, which is a commonly adopted assumption in the literature [15, 23], (ii) the interference channel h_{aa}, and the communication channel h_{ba} are unknown deterministic parameters, (iii) the transmit symbol from node b is modelled using a discrete random distribution, and (iv) we observe N independent received symbols.

The blind channel estimation problem requires the knowledge of the joint probability density function (PDF) of all observations, which is derived from the conditional PDF of a single observation. Given the system model in (2.1), the conditional PDF of a single observation is given by

$$f_{Y_{a_i}}(y_{a_i}|x_{b_i}; h_{aa}, h_{ba}) = \frac{1}{\pi\sigma^2} \exp\left(\frac{-1}{\sigma^2}|y_{a_i} - h_{ba}x_{b_i} - h_{aa}x_{a_i}|^2\right), \tag{2.3}$$

where $i \in \mathcal{I} \triangleq \{1, \cdots, N\}$, y_{a_i} is the i^{th} received symbol, and x_{a_i} and x_{b_i} are the i^{th} transmitted symbols from nodes a and b, respectively.

The marginal PDF of a single observation is then found by multiplying (2.3) by the uniform distribution $p_{X_{b_i}}(x_{b_i}) = \frac{1}{M}\mathbb{I}_{\{\mathcal{A}\}}(x_{b_i})$, and summing the results over all the possible values of x_{b_i}, where, $\mathbb{I}_{\{\mathcal{A}\}}(x) = 1$ if $x \in \mathcal{A}$ and 0 otherwise. Therefore, we have

$$\begin{aligned} f_{Y_{a_i}}(y_{a_i}; h_{aa}, h_{ba}) &= \sum_{\forall x_{b_i}} f_{Y_{a_i}}(y_{a_i}|x_{b_i}; h_{aa}, h_{ba}) p_{X_{b_i}}(x_{b_i}) \\ &= \frac{1}{M\pi\sigma^2} \sum_{\forall x_{b_i}} \exp\left(\frac{-1}{\sigma^2}|y_{a_i} - h_{ba}x_{b_i} - h_{aa}x_{a_i}|^2\right) \mathbb{I}_{\{\mathcal{A}\}}(x_{b_i}) \\ &= \frac{1}{M\pi\sigma^2} \sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2}|y_{a_i} - h_{ba}x_k - h_{aa}x_{a_i}|^2\right), \end{aligned} \tag{2.4}$$

where the last step follows from the fact that $\mathbb{I}_{\{\mathcal{A}\}}(x_{b_i}) = 1$ if and only if $x_{b_i} = x_k$, where $x_k \in \mathcal{A}$. Finally, since the transmitted symbols are assumed independent,

and we observe N independent observations, the joint PDF of all the observations is given by

$$f_{\mathbf{Y_a}}(\mathbf{y_a}; h_{aa}, h_{ba}) = \prod_{i=1}^{N} f_{Y_{a_i}}(y_{a_i}; h_{aa}, h_{ba}) \\ = \left(\frac{1}{M\pi\sigma^2}\right)^N \prod_{i=1}^{N}\sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2}|y_{a_i} - h_{ba}x_k - h_{aa}x_{a_i}|^2\right). \quad (2.5)$$

where we substitute the value of $f_{Y_{a_i}}(y_{a_i}; h_{aa}, h_{ba})$ from (2.4).

Using (2.5), we can state the channel estimation problem as shown in the proposition below.

Proposition 2.1 *The blind maximum likelihood (ML) channel estimation problem in a SISO FD system is given by*

$$\arg\max_{h_{aa}, h_{ba}} f_{\mathbf{Y_a}}(\mathbf{y_a}; h_{aa}, h_{ba}), \quad (2.6)$$

where $f_{\mathbf{Y_a}}(\mathbf{y_a}; h_{aa}, h_{ba})$ *is given by* (2.5).

In the next subsection, we show that (2.6) does not have a unique solution if modulation sets which are symmetric around the origin are used.

2.2.2 Identifiability Analysis

In this subsection, we present the identifiability analysis for the blind channel estimation problem in (2.6), which allows us to determine when ambiguity-free channel estimation is possible. For ease of analysis, we first define $\boldsymbol{\theta} \triangleq [h_{aa}, h_{ba}]$ and rewrite (2.5) as

$$f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta}) = \left(\frac{1}{M\pi\sigma^2}\right)^N \prod_{i=1}^{N}\sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \boldsymbol{\theta}(1)x_{a_i} - \boldsymbol{\theta}(2)x_k|^2\right), \quad (2.7)$$

where $\boldsymbol{\theta}(1)$ and $\boldsymbol{\theta}(2)$ represent the first and second elements of $\boldsymbol{\theta}$.

We start the identifiability analysis by presenting the following definition and remark:

Definition 2.1 *[109, Definition 5.2] If* $\mathbf{Y}$ *is a random vector distributed according to* $f_{\mathbf{Y}}(\mathbf{y};\boldsymbol{\theta})$*, then* $\boldsymbol{\theta}$ *is said to be unidentifiable on the basis of* $\mathbf{y}$*, if* $\forall \mathbf{y}$ *there exists* $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ *for which* $f_{\mathbf{Y}}(\mathbf{y};\boldsymbol{\theta}) = f_{\mathbf{Y}}(\mathbf{y};\boldsymbol{\theta}')$*.*

Remark 2.2 *Definition 2.1 states that* $\boldsymbol{\theta}$ *and* $\boldsymbol{\theta}'$ *(*$\boldsymbol{\theta} \neq \boldsymbol{\theta}'$*) cannot be distinguished from a given set of observations if they both result in the same probability density function for the observations. This implies that if* $\boldsymbol{\theta}$ *is unidentifiable, then it is impossible for any estimator to uniquely determine the value of* $\boldsymbol{\theta}$*.*

In order to present the main result in this subsection, we first give the definitions of a symmetric modulation constellation [110] and a bijective function [111].

Definition 2.2 *We mathematically define modulation constellation as the graph of the function* $\mathrm{f}(x_k) = x_k$*, where* $x_k \in \mathcal{A}\ \forall k \in \mathcal{K}$*. Then a modulation constellation is symmetric with respect to the origin if and only if* $\mathrm{f}(-x_k) = -\mathrm{f}(x_k)\ \forall x_k \in \mathcal{A}$ *[110].*

Definition 2.3 *Let* $\mathcal{C}$ *and* $\mathcal{D}$ *be two sets. A function from* $\mathcal{C}$ *to* $\mathcal{D}$ *denoted* $\mathrm{t} : \mathcal{C} \rightarrow \mathcal{D}$ *is a bijective function if and only if it is both one-to-one and onto.*

The above definition states that a bijective function is a function between the elements of two sets, where each element of one set is paired with exactly one element of the other set and there are no unpaired elements. Note that a bijective function from a set to itself is also called a permutation [111].

In this work, we define and use the bijective function g: $\mathcal{K} \rightarrow \mathcal{K}$, i.e., g is a one-to-one and onto function on $\mathcal{K} \rightarrow \mathcal{K}$. Using this bijective function, we present the main result as below.

Theorem 2.1 *There exists a* $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ *for which the joint probability density* $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta})$ *given by* (2.7) *is equal to* $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}')\ \forall \mathbf{y}_a$*, if and only if there exists a bijective function* g*:* $\mathcal{K} \rightarrow \mathcal{K}$*, such that* $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$*, where* $c \neq 1$ *is a constant and* $|c| = 1$*, i.e., the modulation constellation is symmetric about the origin.*

Proof: We prove the result in Theorem 2.1 in three steps. First, we assume $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ for which $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}) = f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}')\ \forall \mathbf{y_a}$ exists, and show that it leads to a bijective function g satisfying $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$. Then, we assume that a bijective function g satisfying $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$ exists, and show that there exists a $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$

for which $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}) = f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}')\ \forall \mathbf{y_a}$. Finally, using Definition 2.2, we show that the condition $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$ is equivalent to the modulation constellation being symmetric with respect to the origin. The details are in Appendix A.1.

Remark 2.3 *From Theorem 2.1, we can see that since the modulation constellations, such as M-QAM, satisfy the definition of symmetric modulation constellations in Definition 2.2, the blind channel estimation problem in* (2.6) *does not have a unique solution and suffers from an ambiguity problem.*

2.2.3 Proposed Technique

In this subsection, we present our proposed technique to resolve the ambiguity problem in (2.6).

The rationale behind the proposed technique comes from the fact that Theorem 2.1 shows that symmetry of the modulation constellation with respect to the origin is the cause of the ambiguity. A simple way to achieve constellation asymmetry[2] is to add a constant s to each element of $\mathcal{A}$. The resultant asymmetric shifted modulation constellation is formally defined as follows:

Definition 2.4 *The asymmetric shifted modulation constellation, $\overline{\mathcal{A}}$, is defined as*

$$\overline{\mathcal{A}} \triangleq \{x_k + s, \quad \forall\, x_k \in \mathcal{A}, \quad s \in \mathbb{R}^+\}, \tag{2.8}$$

where $\mathbb{R}^+$ is the set of positive real numbers.

In the rest of the thesis, we also use $\overline{x}_k = x_k + s$ to denote the kth element of $\overline{\mathcal{A}}$.

For illustration, Fig. 2.2 shows the effect of the proposed technique on the 16-QAM constellation. We can see that the resulting modulation constellation is shifted along the horizontal axis, which increases the average energy per symbol of the modulation constellation. *This increase in the average energy per symbol can be justified as follows: in reality it is inevitable to use some extra energy to estimate*

[2]Note that it may be possible to achieve constellation asymmetry through other means, such as design of irregular modulation constellations. The optimum design of such modulation constellations is outside the scope of this work.

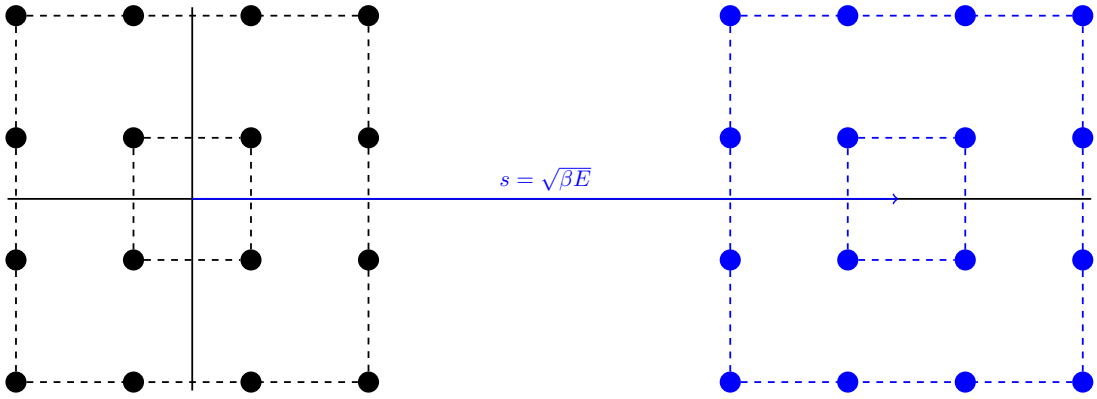


Figure 2.2: Effect of the proposed technique on the constellation of 16-QAM. The resulting constellation is shifted along the horizontal axis, i.e., it is asymmetric around the origin.

the unknown channels, whether it is done by pilots or by the proposed technique. In this regard, it is important to note that the smaller the energy used for shifting the modulation constellation, the closer the average energy of the proposed technique is to the ideal scenario where the channels are perfectly known at the receiver and no extra energy is needed for channel estimation.

Remark 2.4 *The addition of the DC component lowers power efficiency similar to the use of superimposed training [72][3]. However, the proposed shifted modulation technique has the offsetting advantages that i) bandwidth efficiency is not reduced, and ii) the DC offset can be used to reduce the peak-to-average power of the signal envelope during transmissions resulting in lowered cost/complexity power amplifiers. Moreover, the proposed scheme is well-suited to MQAM as investigated here. Since, this is a spectrally efficient modulation scheme used where power efficiency is not critical*

For the sake of numerically investigating the problem of smallest possible shift energy, we define β as the portion of the average energy per symbol that is allocated to the shift and use the real constant $s \triangleq \sqrt{\beta E}$, where $0 < \beta < 1$ to shift the symmetric modulation constellation, and E as the average symbol energy of a given constellation as defined by (2.2). In this case, the problem of smallest shift energy

[3]Note that the existing hardware implementations for superimposed training [112], can also be used here for shifting the modulation constellation.

corresponds to the problem of finding the minimum value of β. The minimum value of β is an indication of how much extra energy is needed compared to the perfect channel knowledge scenario.

In Section 2.3.1, we derive a lower bound on the estimation error, which allows us to numerically find the minimum value of β.

2.3 EM-Based Estimator

In this section, we derive an EM estimator to obtain channel estimates in a FD system with asymmetric shifted modulation constellation defined in Definition 2.4. We derive a lower bound on the estimation error of the estimator. Finally, we investigate the complexity of the proposed estimator.

For the sake of notational brevity, we first define

$$\boldsymbol{\phi} \triangleq [\Re(h_{aa}), \Im(h_{aa}), \Re(h_{ba}), \Im(h_{ba})]. \tag{2.9}$$

We can then reformulate the ML problem in (2.6) as follows

$$[\widehat{\Re(h_{aa})}, \widehat{\Im(h_{aa})}, \widehat{\Re(h_{ba})}, \widehat{\Im(h_{ba})}] \triangleq \arg\max_{\boldsymbol{\phi}} (\ln f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\phi})), \tag{2.10}$$

where $f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\phi})$ is given by (2.7) and $\ln f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\phi})$ is known as the log-likelihood function.

In formulating the channel estimation problem in (2.6) (and hence in (2.10)), we assumed unknown transmitted symbols. These unknown transmitted symbols can be treated as hidden data. A common approach to solving the maximization problem in (2.10) in the presence of hidden data is the EM algorithm [113], which is adopted in this work. The main steps of EM algorithm are

1. *Expectation step:* In the E-step, the expectation of the log-likelihood is taken over all the values of the hidden variable, conditioned on the vector of observations, and the nth estimate of $\boldsymbol{\phi}$ ($\boldsymbol{\phi}^{(n)}$). In (2.1), the hidden variable is $\overline{\mathbf{x}}_\mathbf{b}$ and consequently, we need to evaluate $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) \triangleq \mathbb{E}_{\overline{\mathbf{X}}_\mathbf{b}|\mathbf{y_a},\boldsymbol{\phi}^{(n)}}[\ln f_{\mathbf{Y_a}}(\mathbf{y_a}, \overline{\mathbf{x}}_\mathbf{b}|\boldsymbol{\phi})]$.

2. *Maximization step:* In the M-step, the function $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$ obtained from the

E-step is maximized with respect to $\boldsymbol{\phi}$.

3. *Iterations:* We iterate between the E- and M-steps until convergence is achieved.

The equations needed for the E- and M-steps are summarized in the propositions below.

Proposition 2.2 *The E-step during nth iteration of the algorithm is given by*

$$Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) = -N\ln(M\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)}|y_{a_i} - h_{ba}\overline{x}_k - h_{aa}\overline{x}_{a_i}|^2, \tag{2.11}$$

where $\boldsymbol{\phi}^{(n)} \triangleq [\widehat{h}_{aa}^{(n)}, \widehat{h}_{ba}^{(n)}]$ *are the estimates of the channels obtained from* $\boldsymbol{\phi}^{(n)}$ *during the nth iteration of the algorithm, and* $T_{k,i}^{(n)}$ *is defined as*

$$T_{k,i}^{(n)} \triangleq \frac{\exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \widehat{h}_{ba}^{(n)}\overline{x}_k - \widehat{h}_{aa}^{(n)}\overline{x}_{a_i}|^2\right)}{\sum_{\overline{k}=1}^{M}\exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \widehat{h}_{ba}^{(n)}\overline{x}_{\overline{k}} - \widehat{h}_{aa}^{(n)}\overline{x}_{a_i}|^2\right)}, \tag{2.12}$$

where $k \in \mathcal{K}$, $i \in \mathcal{I} \triangleq [1, 2, \cdots, N]$, $\overline{x}_k \in \overline{\mathcal{A}}$, $\overline{x}_{a_i} \in \overline{\mathcal{A}}$ *and* $\overline{x}_{\overline{k}} \in \overline{\mathcal{A}}$.

Proof: See Appendix A.2.

Proposition 2.3 *The M-step during the nth iteration of the algorithm is given by*

$$\boldsymbol{\phi}^{(n+1)} = \frac{1}{s_1 s_4 - s_2^2 - s_3^2}\begin{bmatrix} -s_2 v_3 - s_3 v_4 + s_4 v_1 \\ -s_2 v_4 + s_3 v_3 + s_4 v_2 \\ s_1 v_3 - s_2 v_1 + s_3 v_2 \\ s_1 v_4 - s_2 v_2 - s_3 v_1 \end{bmatrix}, \tag{2.13}$$

where

$$s_1 \triangleq \sum_{i=1}^{N}|\overline{x}_{a_i}|^2,\ s_2 \triangleq \sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)}\Re(\overline{x}_{a_i}\overline{x}_k^*), \tag{2.14a}$$

$$s_3 \triangleq \sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)}\Im(\overline{x}_{a_i}\overline{x}_k^*),\ s_4 \triangleq \sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)}|\overline{x}_k|^2, \tag{2.14b}$$

$$v_1 \triangleq \sum_{i=1}^{N} \Re(\overline{x}_{a_i}^* y_{a_i}), \quad v_2 \triangleq \sum_{i=1}^{N} \Im(\overline{x}_{a_i}^* y_{a_i}), \tag{2.14c}$$

$$v_3 \triangleq \sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)} \Re(y_{a_i} \overline{x}_k^*), \quad v_4 \triangleq \sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)} \Im(y_{a_i} \overline{x}_k^*). \tag{2.14d}$$

Proof: See Appendix A.2.

Remark 2.5 *It is well-known that the EM algorithm may be very sensitive to initialization [114]. Although different methods exist for EM initialization, generally they are not computationally efficient [114, 115]. For the given channel assumptions in Section 2.4, our empirical results showed that initializing the EM algorithm by* $\boldsymbol{\phi}^{(0)} \triangleq [0, 0, 0, 0]$ *resulted in the lowest estimation error. Hence, this initialization is used in this work.*

2.3.1 Lower Bound on the Estimation Error

In this section we derive a closed-form lower bound on the estimation error of the proposed estimator. The derived lower bound directly links the channel estimation error to the parameter β, defined in Section 2.2.3.

The EM algorithm, defined in Propositions 2.2 and 2.3, is a ML estimator for the parameter vector $\boldsymbol{\phi}$ in (2.9). Hence, we aim to derive the lower bound for the variance of the proposed ML estimator. The ML estimator is asymptotically efficient [116] and its MSE is lower bounded by the inverse of the Fisher information matrix (FIM) [116]. This result is known as Cramer Rao lower bound (CRLB) and is given by

$$\mathbb{E}_{\widehat{\Phi}_l}[|\widehat{\phi}_l - \phi_l|^2] \geq \left[\mathbf{I}^{-1}\left[f_{\mathbf{Y_a}}(\mathbf{y}_a; \boldsymbol{\phi})\right]\right]_{l,l}, \tag{2.15}$$

where ϕ_l is the lth element of the parameter vector $\boldsymbol{\phi}$, $\widehat{\phi}_l$ is an estimate of ϕ_l, for $l \in \{1, 2, 3, 4\}$, $[\cdot]_{l,l}$ is the lth diagonal element of a square matrix, and $\mathbf{I}^{-1}\left[f_{\mathbf{Y_a}}(\mathbf{y}_a; \boldsymbol{\phi})\right]$ is the inverse of FIM. Since the inverse of FIM in (2.15) cannot be found in closed-form [74, 117], we derive a lower bound on the MSE of the proposed estimator, which is in closed-from. The result is presented in the proposition below.

Proposition 2.4 *The variance of the proposed estimator is lower bounded by*

$$\mathbb{E}_{\widehat{\Phi}_l}[|\widehat{\phi}_l - \phi_l|^2] \geq \left(\frac{\sigma^2}{2NE}\right) \frac{1+\beta}{(1+2\beta)}, \tag{2.16}$$

where $l \in \{1, \cdots, 4\}$*,* N *is the number of observations,* E *is the average symbol energy of the modulation constellation before the shift, and* β *is the portion of* E *that is allocated to the shift.*

Proof: See Appendix A.3.

Remark 2.6 *The result in* (2.16) *links the closed-form lower bound of the estimation error to the average energy of the modulation constellation before the shift and the portion of this average energy allocated to the shift. This is important because in Section 2.4.1, we will use* (2.16) *to find the minimum shift energy needed for the proposed technique.*

2.3.2 Complexity Analysis

To evaluate the feasibility in implementing the proposed estimator, we investigate the computational complexity of the estimator in terms of required floating point multiplications and additions (flops) [118].

Table 2.1 shows the number of multiplications and additions needed for the EM estimator for h_{ba}. Although we only present the complexity analysis of h_{ba}, similar complexity is also observed for estimating h_{aa}. In each row of the table, the number of required additions and multiplications to implement a given equation is presented and are then summed to obtain overall complexity.

It is clear from Table 2.1 that the complexity of EM estimator per iteration is proportional to NM^2. This analysis shows that the EM algorithm is computationally very efficient since, for a given modulation constellation with size M, the computational complexity of the EM estimator only grows linearly with the number of observations, N.

Remark 2.7 *In data-aided approaches to channel estimation both* $\mathbf{x}_a$ *and* $\mathbf{x}_b$ *in* (2.1) *are assumed to be known. Consequently, linear channel estimation can be performed for linear Gaussian models (LGMs) as explained in [116]. It is a well-known fact*

Table 2.1: Complexity analysis of the EM estimator.

EM - Complexity per iteration		
(Eq. No.)	Additions	Multiplications
(2.12)	$3M+2$	$6M+6$
(2.11)	$NM(3M+4)$	$NM(6M+9)+5$
(2.13)	$4NM(3M+2)+3N+14$	$4NM(6M+8)+3N+33$
$\widehat{h}_{ba}$	$15NM^2+12NM+3N+3M+16$	$30NM^2+41NM+3N+6M+44$

that linear estimator complexity estimators for LGM is independent of the modulation size M, and only grows linearly with the number of observations [116]. The extra complexity of the proposed algorithm compared to linear estimators is expected. This is because as opposed to linear estimators, the proposed estimator requires no data-aided piloting, and hence, it allows for efficient use of the bandwidth for channel estimation.

2.4 Simulation Results

In this section, we present numerical and simulation results to investigate the performance of the proposed estimator with asymmetric shifted modulation constellation. We consider a FD communication system as illustrated in Fig. 2.1. The analysis in Section 2.3.1 shows an identical lower bound for the estimation error of both h_{aa} and h_{ba}. Hence, in this section, we only present the results for the communication channel h_{ba} since identical results are obtained for the SI channel h_{aa}.

For each simulation run, N data and interfering symbols are randomly generated assuming shifted 16-QAM modulation constellation is used ($M = 16$). The channels are constant for the transmission of N symbols, i.e., the quasi static assumption. We assume that there is no line-of-sight (LOS) communication link between the transmitter of node b and the receiver of node a. Hence, the communication channel h_{ba} can be modelled as a Rayleigh fading channel, i.e., $h_{ba} \sim \mathcal{CN}(0, \sigma^2_{h_{ba}})$. For the SI channel, experimental results have shown that before passive and active cancellation the SI channel has a strong LOS component and can be modelled as a Rician distribution with a large K factor (approximately 20-25 dB). After passive

suppression and analog cancellation, the strong LOS component is significantly reduced but still present and can be modelled as a Rician distribution with $K = 0$ dB [15]. Hence, we generate the SI channel as

$$h_{aa} = \sqrt{\frac{K}{K+1}}\sigma_{h_{aa}} e^{\mathrm{j}\zeta} + \sqrt{\frac{1}{K+1}}\mathcal{CN}(0, \sigma^2_{h_{aa}}), \tag{2.17}$$

where ζ is uniformly distributed angle of arrival of the LOS component of the SI channel [119].

For the simulations, the signal-to-interference-noise ratio (SINR) is given by [15]

$$\text{SINR} = \frac{1}{\frac{1}{\text{SIR}} + \frac{1}{\text{SNR}}}, \tag{2.18}$$

where the signal-to-interference ratio $\text{SIR} = \frac{\sigma^2_{h_{ba}}}{\sigma^2_{h_{aa}}}$ assuming both nodes use constellations with the same average energy, the desired-signal-to-noise ratio $\text{SNR} = \frac{\sigma^2_{h_{ba}} \log_2(M) E_b}{N_0}$, E_b is the average bit energy which is defined below (2.2) and N_0 is the noise power spectral density.

As discussed in Section 1.1, even with state-of-the-art passive suppression and analog cancellation, the SIR can still be around -50 dB [15, 30]. Hence, we adopt this value of the SIR in the simulations while assuming that the communication channel has average energy of unity, i.e., $\mathbb{E}[|h_{ba}|^2] = \sigma^2_{h_{ba}} = 1$. Furthermore, in order to investigate the performance of the proposed estimator over a range of SINR, we fix $N_0 = 1$ and run the simulations for different values of E_b/N_0 (in dB). The figures of merit used are the average mean square error (MSE) and the BER, which are obtained by averaging over 5000 Monte Carlo simulation runs.

2.4.1 Minimum Energy Needed for Channel Estimation

In this subsection, we are interested in finding the minimum value of β, for a given E_b/N_0 and N. As discussed in Section 2.2.3, we use $s \triangleq \sqrt{\beta E}$, where $0 < \beta < 1$, to shift the symmetric modulation constellation. Hence, a lower value of β is desirable since it means less energy is used to shift the modulation constellation.

In order to find a minimum value of β suitable for a practical range of E_b/N_0 and N, we use the average MSE lower bound in (2.16) to observe the behavior

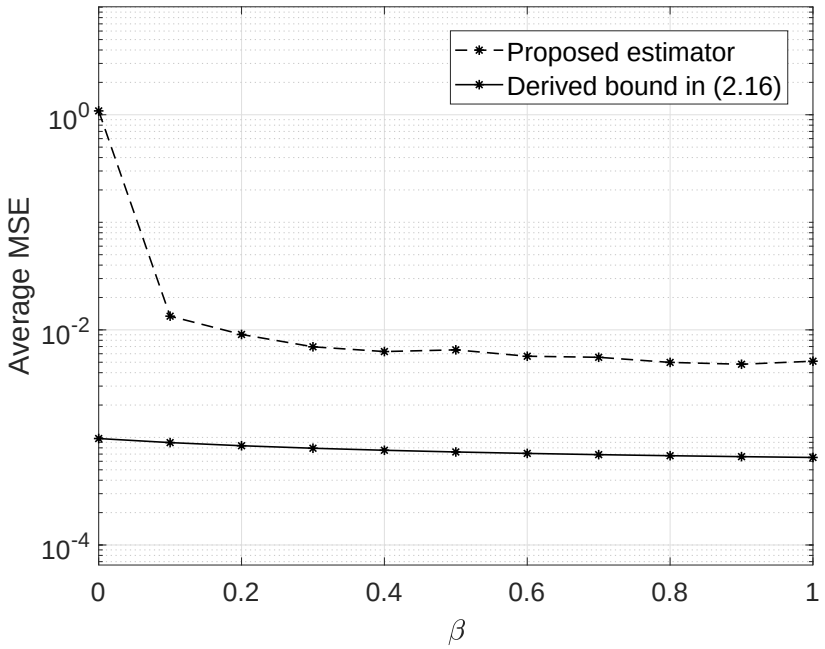


Figure 2.3: MSE performance of the proposed channel estimator for different values of β for $E_b/N_0 = 0$ dB, $N = 128$ and SIR $= -50$ dB.

of the proposed estimator as a function of β at low N and low E_b/N_0. This is motivated by the fact that the minimum value of β found for low N and low E_b/N_0 will ensure that the desired estimation error will also be achieved for high E_b/N_0 and/or when the number of observations N is large. This intuition is confirmed from (2.16), which indicates that higher values of β are needed at low E_b/N_0 to reach a given estimation error. Furthermore, since the lower bound on the estimation error also decreases with N, the minimum value of β found for smaller N can also serve for larger N. Since the experimental results of [15, 108] show that the FD communication channel is normally constant for more than $N > 128$ symbols, we propose to find the minimum β at $N = 128$ and $E_b/N_0 = 0$ dB.

Fig. 2.3 shows the MSE performance of the proposed technique versus β for $E_b/N_0 = 0$ dB, $N = 128$ and SIR $= -50$ dB. If the desired estimation error is taken to be within 10% of the lower bound error, then we can see from the figure that for $\beta < 0.2$, the MSE of the proposed estimator is within 10% of the lower bound. Consequently, the minimum value of β is 0.2. This minimum value of β in

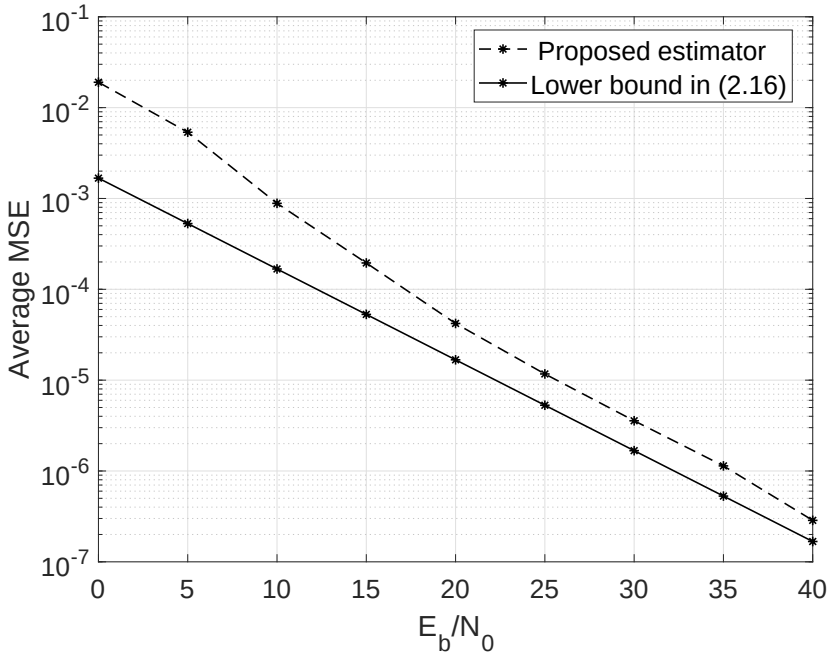


Figure 2.4: MSE performance of the proposed channel estimator vs. E_b/N_0 for $\beta = 0.2$, $N = 128$ and SIR $= -50$ dB.

turn determines the minimum required shift energy, i.e., the optimum shift energy. This is because the energy of the shift for a fixed constellation energy (E) is given by s^2, where s is given by $s = \sqrt{\beta E}$.

Fig. 2.4 shows the MSE performance of the proposed estimator with $\beta = 0.2$ (the selected minimum value of β) vs. E_b/N_0 (dB) for $N = 128$ and SIR $= -50$ dB. The lower bound in (2.16) is plotted as a reference. The figure shows that as E_b/N_0 increases, the gap between the performance of the proposed estimator and the lower bound decreases. The gap is less than 2 dB after $E_b/N_0 = 20$ dB.

In the following sections, we set $\beta = 0.2$ and $N = 128$ to study the performance of the FD communication system.

2.4.2 Comparison with Data-Aided Channel Estimation

In this section we compare the MSE and BER performance of the proposed estimator against a data-aided channel estimator for the case that the average energy

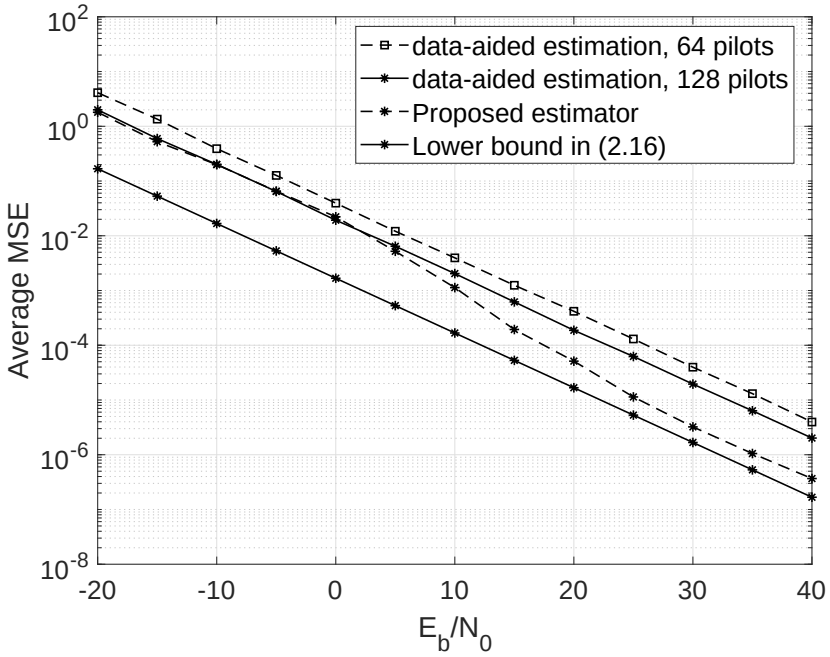


Figure 2.5: MSE performance of the proposed technique.

per transmitted frame is the same for both methods [4]. For the proposed technique, we assume that (i) all the transmitted symbols are data symbols, and (ii) shifting the modulation constellation increases the average energy by 20% compared to the ideal scenario when no channel estimation is needed (corresponds to $\beta = 0.2$). For the data-aided channel estimation, we assume that (i) 64 pilot symbols are used in a frame of 128 symbols and (ii) these pilots also require an extra 20% energy.

MSE performance: The average MSE reveals the accuracy of the channel estimation. Fig. 2.5 plots the average MSE vs. E_b/N_0 with $\beta = 0.2$, $N = 128$ and SIR $= -50$ dB. The lower bound from (2.16) is plotted as a reference. We also plot the MSE for data-aided channel estimation with (i) 64 pilot symbols in a frame of 128 symbols and (ii) 128 pilot symbols in a frame of 128 symbols. Fig. 2.5 shows that the proposed technique outperforms data-aided channel estimation when both methods use the same extra amount of energy for channel estimation. At high

[4]Note that the simulation results have been obtained by normalizing the shifted modulation so the extra power needed to shift the modulation does not push the power amplifiers into saturation and hence, power amplifiers do not experience any non-linearities.

E_b/N_0, the MSE performance of the proposed technique is within $3-4$ dB of the lower bound.

It has been shown in [26] that the effect of quantization error in FD communication system can be modeled as an additive Gaussian noise. This means the system model given by (2.1) implicitly includes the effect of quantization noise as well as the effect of thermal noise in the Gaussian noise term $\mathbf{w}_a$. Consequently, the effect of quantization noise on the performance of the proposed estimator can be studied by observing the MSE results of the proposed estimator in the low SNR region as shown by Fig. 2.5. The results of Fig. 2.5 show that at low SNR region a noticeable MSE gain is not obtained by using the proposed estimator instead of the data-aided technique that uses 128 pilots. However, the proposed technique is still more attractive compared to the data-aided technique because of the bandwidth efficiency.

BER performance: Fig. 2.6 shows the average BER vs. E_b/N_0 (dB) with $\beta = 0.2$, $N = 128$ and SIR $= -50$ dB. The BER performance with perfect channel knowledge is plotted as a reference. We also plot the BER for data-aided channel estimation with 64 pilot symbols in a frame of 128 symbols. Fig. 2.6 shows that the proposed technique outperforms the data-aided channel estimation in terms of the BER. This is to be expected since, as shown in Fig. 2.5, for the same extra amount of energy for channel estimation the proposed technique has much lower MSE compared to data-aided channel estimation. In addition, at high E_b/N_0, the BER performance of the proposed technique is within 1 dB of the ideal performance obtained with perfect channel knowledge.

As shown in this sub-section, In comparison to the data-aided algorithms, the proposed algorithm is more bandwidth efficient. Also comparison to Compared to the existing blind algorithms, which suffer from phase ambiguity [73], the proposed algorithm can estimate the channel with no phase ambiguity. These advantages have been obtained by an increase in complexity as explained in Remark 2.7. However, the superior performance of the proposed algorithm as shown in Figs. 2.5 and 2.6 produces attractive tradeoffs compared to the existing data-aided and blind algorithms.

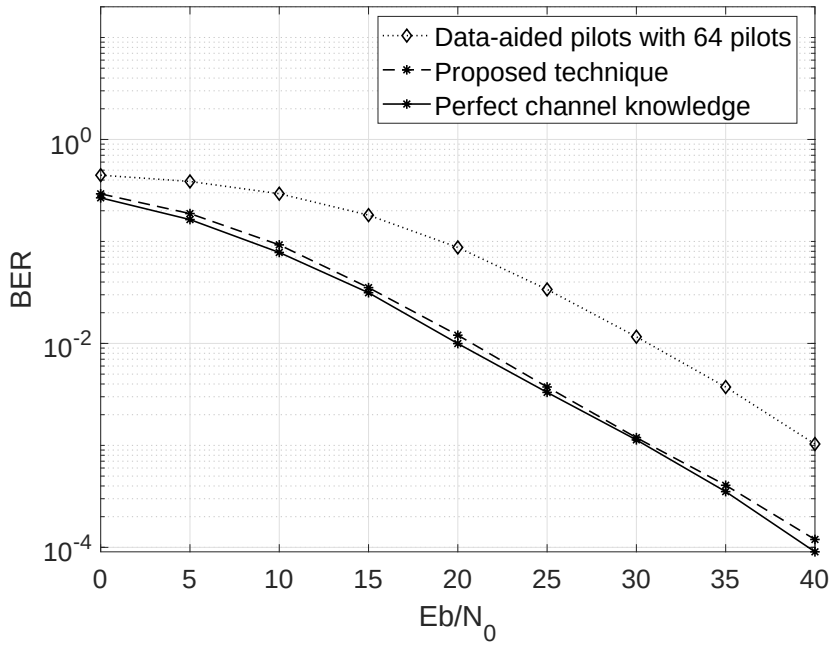


Figure 2.6: BER performance of the proposed technique.

2.4.3 Effect of Power of SI Signal

In the results so far, we have set the SIR to -50 dB. In this section, we assess the impact of the SI power level on the performance of the proposed technique.

Fig. 2.7 plots the BER versus the SIR (dB) for $E_b/N_0 = 0, 10, 20$ dB, with $\beta = 0.2$ and $N = 128$. *We can see that as the SI power increases, the BER performance of the proposed technique remains nearly constant. This is because in FD communication the self-interference signal is completely known to the receiver [23]. Consequently, for relatively small channel estimation error of the proposed estimator, the SI can be cancelled regardless of its power.* Fig. 2.7 illustrates the robustness of the proposed technique, i.e., even with weak passive suppression and analog cancellation requiring digital SI cancellation to handle a very large SIR (e.g., -100 dB), the BER is not significantly altered.

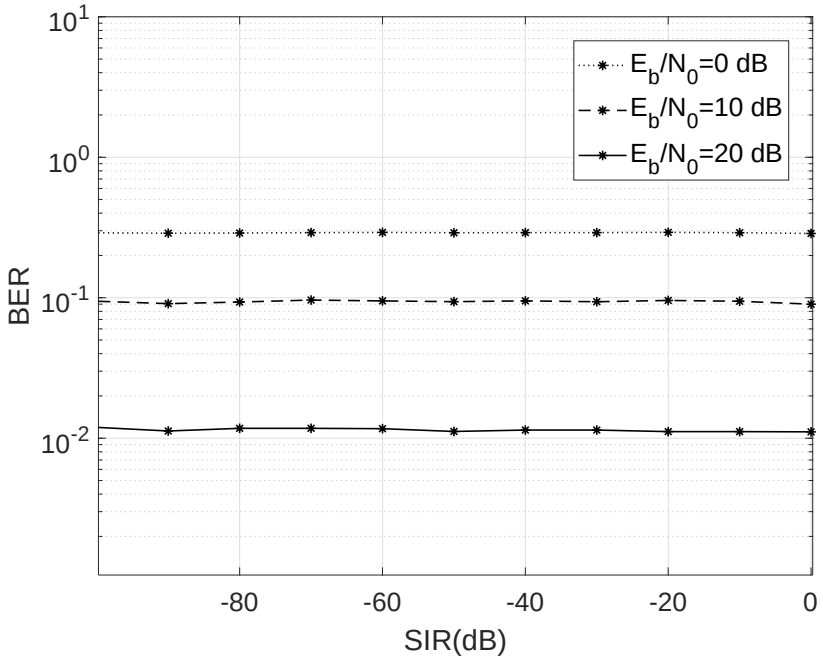


Figure 2.7: Effect of SI power level on the BER performance of the proposed technique.

2.5 Conclusions

In this chapter, we have proposed a new technique to estimate the SI and communication channels in a FD communication systems for residual SI cancellation. In the proposed technique, we add a real constant number to each constellation point of a modulation constellation to yield asymmetric shifted modulation constellations with respect to the origin. Using identifiability analysis, we showed mathematically that such a modulation constellation can be used for ambiguity-free channel estimation in FD communication systems. We proposed a computationally efficient EM-based estimator to estimate the SI and communication channels simultaneously using the proposed technique. We also derived a lower bound for the estimation error of the proposed estimator. The results showed that the proposed technique is robust to the level of SI power.

Chapter 3

Joint Channel and Phase Noise Estimation for mmWave Full-Duplex Communication Systems

The PN process associated with local oscillators varies significantly faster at mmWave frequency band as opposed to microwave frequency band. This fast variation of the PN process reduces the effectiveness of digital SI cancellation techniques. However, in contrast to HD communication, tracking the PN process for FD enabled systems is more challenging. This is due to the presence of strong SI signal, which corrupts the observations needed for tracking. This chapter of the thesis considers a MIMO mmWave FD communication system, and proposes an EKF based estimation algorithm to track the rapid variation of PN at mmWave frequencies in the presence of a strong SI signal.

The remainder of this chapter is organised as follows. In Section 3.1, we present the system model. Then in Section 3.2, we propose an EKF based algorithm to jointly estimate the channel and track the PN process. In this same section we also discuss the complexity of the proposed algorithm, and present Proposition 3.1, which derives a lower bound for the estimation error of PN at mmWave. In Section 3.3, we numerically show that the mean square error performance of the

proposed estimator approaches the lower bound. We also simulate the bit error rate performance of the proposed system and show the effectiveness of a digital canceller, which uses the proposed estimator to estimate the SI channel.

3.1 System Model

We consider the MIMO communication system between two mmWave FD nodes a and b, each with N_t transmit and N_r receive antennas as illustrated in Fig. 3.1. The considered communication system can be a model for backhaul communication for cellular systems [43]. In this work, we make the following assumptions:

1. *The same number of transmit and receive antennas for both nodes*: We assume both nodes in the considered FD communication system have the same number of transmit and receive antennas.

2. *Modeling of RF impairments*: RF impairments due to imperfect transmitter and receiver chain electronics have been shown to significantly degrade the performance of the analogue cancellation techniques [26, 29]. Since the focus of this work is residual SI cancellation, we only include PN in our model and assume that the other hardware impairments are dealt by a RF canceller. Such an assumption is also made in [70, 84, 120, 121].

3. *Assumptions on oscillators*: We make two assumptions about the oscillators. First, we assume that free running oscillators are used. The assumption of using free running oscillators for mmWave communications has also been made in [43, 122]. Second, we assume each transmit and receive antenna is equipped with an independent oscillator.

4. *Quasi-static flat fading channel assumption*: The SI measurement results of [81] show that even with omnidirectional dipole antennas, the delay spread of the channel does not exceed 800 ns. This delay is significantly smaller than the proposed symbol durations for 5G communication [123, 124], which are in order of μs. Hence, not only can the channel be assumed flat but it can also be assumed to remain constant over transmission of one block of data (quasi-static). Similarly, measurement results of the desired communication

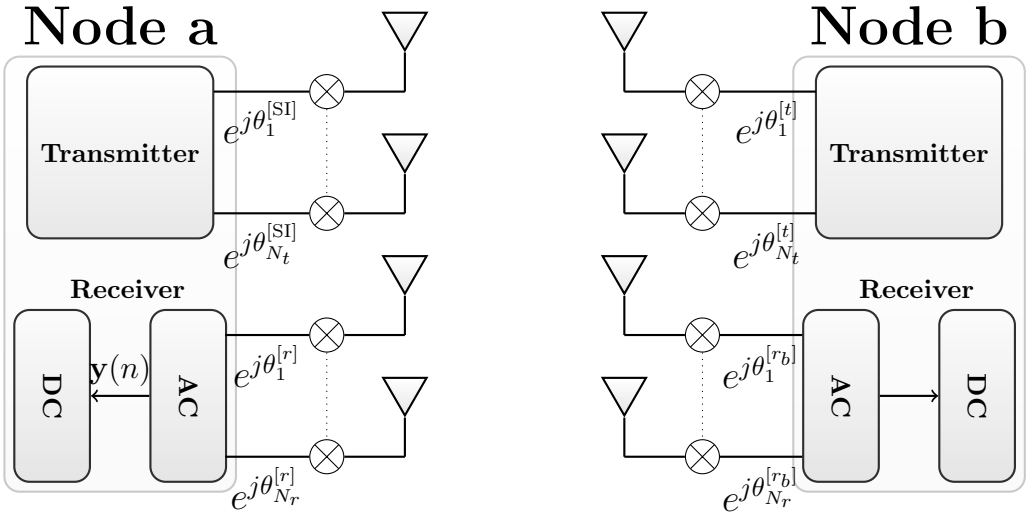


Figure 3.1: System model block diagram of FD communication, where AC stands for analogue SI cancellation, DC stands for digital SI cancellations, $e^{j\theta_l^{[r/r_b/t/SI]}}$ represents the PN at the lth antenna.

channel show that the channel delays are relatively small compared to the symbol durations [125].[1]

5. *Synchronized transmission and reception*: Although synchronizing transmission and reception of analogue desired communication signal with the reception of analogue SI signal is an important practical problem and requires its own detailed investigation, the synchronized FD communication assumption is widely used in the literature of channel and PN estimation for digital SI cancellation (DC) [55, 70, 84].

Before proceeding further, we tabulate the mathematical notations used in this chapter in Table 3.1, for the ease of reference.

[1]While some studies consider multiple non line of sight (NLoS) for mmWave channels [126–128], in this work we only consider a mmWave communication channel with a single-tap multipath channel. The same assumption has also been made in [43,122,129], where PN problem at mmWave communication is studied.

Table 3.1: Important symbols used in Chapter 3.

Symbol	Description
$\mathbf{y}(n)$	The $(N_r \times 1)$ vector of received symbols.
$\mathbf{x}(n)$	The $(N_t \times 1)$ vector of transmitted symbols.
$\mathbf{x}^{\mathrm{SI}}(n)$	The $(N_t \times 1)$ vector of self-interfering (SI) symbols.
$\mathbf{w}(n)$	The $(N_r \times 1)$ Gaussian noise vector.
$\theta_i^{[m]}(n)$	The time varying phase noise of ith oscillator and $m \in \{t = \text{transmit}, r = \text{receive}, \mathrm{SI}\}$.
$\mathbf{H}(n)$	The $(N_r \times N_t)$ communication channel.
$\mathbf{H}^{\mathrm{SI}}(n)$	The $(N_r \times N_t)$ self-interference (SI) channel.
$\overline{\mathbf{H}}(n)$	The $(N_r \times 2N_t)$ state transition matrix for joint PN and channel estimation.
$\boldsymbol{\beta}(n)$	The $(2N_t \times 1)$ state vector for joint channel and PN estimation.

3.1.1 Mathematical Representation of Received Vector

In this subsection, we present a mathematical model for the received vector of a FD MIMO communication system at mmWave frequencies. The received vector at node a and at time instant n, is $\mathbf{y}(n)$, and is given by

$$\mathbf{y}(n) = \mathbf{H}(n)\mathbf{x}(n) + \mathbf{H}^{\mathrm{SI}}(n)\mathbf{x}^{\mathrm{SI}}(n) + \mathbf{w}(n), \tag{3.1}$$

where $\mathbf{y}(n) \triangleq [y_1(n), \cdots, y_{N_r}(n)]^T$, and $y_i(n)$ is the received symbol at the ith antenna. For $i \in \{1, \cdots, N_r\}$ and $k \in \{1, \cdots, N_t\}$, the element in the ith row and kth column of $N_r \times N_t$ channel matrix $\mathbf{H}(n)$ is given by $h_{i,k}e^{j\left(\theta_i^{[r]}(n)+\theta_k^{[t]}(n)\right)}$, where, $h_{i,k}$ is the communication channel between the kth transmit antenna of node b to the ith receive antenna of node a, for $m \in \{r, t, \mathrm{SI}\}$, $\theta_j^{[m]}(n)$ is the oscillator PN at

the jth antenna and m determines the type of antenna such that $m = r$ indicates a receive antenna, $m = t$ means a transmit antenna, and $m = \text{SI}$ indicates an interfering antenna. Furthermore, PN variation of a free running oscillator follows a Wiener process [130], i.e.,

$$\theta_j^{[m]}(n) = \theta_j^{[m]}(n-1) + \delta(n), \tag{3.2}$$

where $\delta(n)$ is Gaussian noise with mean 0 and variance $\sigma^2_{[m]}$, i.e., $\delta(n) \sim \mathcal{N}(0, \sigma^2_{[m]})$.

Similarly, the element in the ith row and kth column of $N_r \times N_t$ SI channel matrix $\mathbf{H}^{\text{SI}}(n)$ is given by $h^{\text{SI}}_{i,k} e^{j\left(\theta_i^{[r]}(n) + \theta_k^{[\text{SI}]}(n)\right)}$, where $h^{\text{SI}}_{i,k}$ is the interference channel between the kth transmit antenna to the ith receive antenna of node a.

In addition, the kth elements of $N_t \times 1$ vectors $\mathbf{x}(n)$ and $\mathbf{x}^{\text{SI}}(n)$ are given by $x_k(n)$ and $x^{\text{SI}}_k(n)$, respectively, which are the transmitted symbols from the kth transmit antenna of nodes b and a, respectively.

Finally, $\mathbf{w}(n) \triangleq [w_1(n), \cdots, w_{N_r}(n)]^T$, where $w_i(n)$ is the complex Gaussian noise, i.e., $w_i(n) \sim \mathcal{CN}(0, \sigma^2)$.

3.1.2 Mathematical Representation for Joint Channel and PN Estimation

For received vector $\mathbf{y}(n)$ and noise vector $\mathbf{w}(n)$ in (3.1), a useful mathematical model for joint channel and PN estimation, is of the form [116, Ch. 13, pp.450, eq.13.66]

$$\mathbf{y}(n) = \overline{\mathbf{H}}(n) f\left(\boldsymbol{\beta}(n)\right) + \mathbf{w}(n), \tag{3.3}$$

where $\overline{\mathbf{H}}(n)$ is the state transition model matrix, f is a nonlinear function, and $\boldsymbol{\beta}(n)$ is the state vector to be estimated.

A fundamental step in the problem of joint channel and PN estimation is the construction of the state vector and the state transition matrix based on the system model given by (3.1). The state vector and the state transition matrix for the joint PN and channel estimation in the presence of SI signal are given by (3.4) and (3.5), respectively.

- ***The state vector***: We construct the state vector using the time evolution of phase given by (3.2) as follows:

$$\boldsymbol{\beta}(n) \triangleq [\overline{\boldsymbol{\beta}}_1(n), \cdots, \overline{\boldsymbol{\beta}}_{N_r}(n)]^T \tag{3.4}$$

- ***The state transition matrix***:

$$\overline{\mathbf{H}}(n) \triangleq \begin{bmatrix} \overline{\mathbf{h}}_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \overline{\mathbf{h}}_{N_r} \end{bmatrix} \tag{3.5}$$

where

$$\overline{\boldsymbol{\beta}}_i(n) \triangleq [\overline{\beta}_{i,1}, \cdots, \overline{\beta}_{i,2N_t}], \tag{3.6a}$$

$$\overline{\beta}_{i,\overline{k}}(n) \triangleq \begin{cases} \theta_i^{[r]}(n) + \theta_k^{[t]}(n), & \overline{k} \text{ is odd;} \\ \theta_i^{[r]}(n) + \theta_k^{[\mathrm{SI}]}(n), & \overline{k} \text{ is even} \end{cases}, \tag{3.6b}$$

$$\overline{\mathbf{h}}_i \triangleq [\overline{h}_{i,1}, \cdots, \overline{h}_{i,2N_t}] \odot [\overline{x}_1(n), \cdots, \overline{x}_{2Nt}(n)], \tag{3.6c}$$

$$\overline{h}_{i,\overline{k}} \triangleq \begin{cases} h_{i,k}, & \overline{k} \text{ is odd;} \\ h_{i,k}^{\mathrm{SI}}, & \overline{k} \text{ is even} \end{cases}, \tag{3.6d}$$

$$\overline{x}_{\overline{k}}(n) \triangleq \begin{cases} x_k(n), & \overline{k} \text{ is odd;} \\ x_k^{\mathrm{SI}}(n), & \overline{k} \text{ is even} \end{cases}, \tag{3.6e}$$

$$\overline{k} = \{1, \cdots, 2N_t\} \tag{3.6f}$$

$$k = \begin{cases} \overline{k}, & \overline{k} < N_t \text{ ;} \\ \overline{k} - N_t, & \overline{k} > N_t \end{cases}. \tag{3.6g}$$

The principle idea behind the design of the state vector $\boldsymbol{\beta}(n)$ and the state transition matrix $\overline{\mathbf{H}}(n)$ as given by (3.4) and (3.5), respectively, is the fact that the PN noise is the only random variable that varies from one symbol to another, and needs to be tracked. On the other hand, because of the quasi-static nature of the communication and SI channels, they remain constant over transmission of a single data packet. Therefore, these channels need to be estimated only once at the

beginning of data transmission. This initial channel estimation for the constant channels can be done using pilot transmission.

Furthermore, we note that at each receive antenna there are $2N_t$ parameters that need to be estimated, N_t parameters for the communication channel and N_t parameters for the SI channel. This explains the existence of index $\overline{k} \in \{1, \cdots, 2N_t\}$.

Finally, with the state vector $\boldsymbol{\beta}(n)$ and the state transition matrix $\overline{\mathbf{H}}(n)$ given by (3.4) and (3.5), the discrete-time received vector at time instant n and at the baseband of node a is given by

$$\mathbf{y}(n) = \overline{\mathbf{H}}(n)e^{j\boldsymbol{\beta}(n)} + \mathbf{w}(n). \tag{3.7}$$

3.2 Joint Channel and PN Estimation

In this section we use the state vector (3.4) and the state transition matrix (3.5), and present a joint channel and PN estimator based on the concept of extended Kalman filtering (EKF) [116]. The observation vector of EKF is given by $\mathbf{y}(n)$ in (3.7), which is a non-linear function of the states $\boldsymbol{\beta}(n)$. The EKF state equation is given by

$$\boldsymbol{\beta}(n) = \boldsymbol{\beta}(n-1) + \mathbf{u}(n), \tag{3.8}$$

where $\mathbf{u}(n)$ is Gaussian with mean zero and covariance $\mathbf{Q} \triangleq \mathbb{E}[\boldsymbol{\beta}(n)\boldsymbol{\beta}^T(n)]$, i.e., $\mathbf{u}(n) \sim \mathcal{N}(\mathbf{0}_{2N_rN_t}, \mathbf{Q})$. The $2N_tN_r \times 2N_tN_r$ covariance matrix $\mathbf{Q}$ is given by

$$\mathbf{Q} \triangleq \mathbb{E}\left[\boldsymbol{\beta}(n)\boldsymbol{\beta}^T(n)\right] = \begin{bmatrix} \mathbf{R}_{1,1} & \cdots & \overline{\mathbf{R}}_{1,N_r} \\ \vdots & \vdots & \vdots \\ \mathbf{R}_{N_r,1} & \cdots & \overline{\mathbf{R}}_{N_r,N_r} \end{bmatrix}, \tag{3.9}$$

where, for $m, n \in \{1, \cdots, N_r\}$, $\mathbf{R}_{m,n}$ is $2N_tN_t$ matrix given by (3.10), where σ_r^2, σ_t^2,

and σ^2_{SI} are PN variances due to receive, transmit, and SI antennas, respectively.

$$\mathbf{R}_{m,n} = \begin{cases} \sigma_r^2 \mathbf{1}_{2N_t \times 2N_t} + diag\left(\underbrace{\sigma_t^2, \cdots, \sigma_t^2}_{2N_t}\right), & m = n \text{ and is odd} \\ \sigma_r^2 \mathbf{1}_{2N_t \times 2N_t} + diag\left(\underbrace{\sigma^2_{\text{SI}}, \cdots, \sigma^2_{\text{SI}}}_{2N_t}\right), & m = n \text{ and is even} \\ diag\left(\underbrace{\sigma_t^2, \cdots, \sigma_t^2}_{2N_t}\right), & m \neq n \text{ and is odd} \\ diag\left(\underbrace{\sigma^2_{\text{SI}}, \cdots, \sigma^2_{\text{SI}}}_{2N_t}\right), & m \neq n \text{ and is even} \end{cases} \tag{3.10}$$

The EKF state update equations are given by [116]

$$\widehat{\boldsymbol{\beta}}(n|n) = \widehat{\boldsymbol{\beta}}(n|n-1) + \Re\left\{\mathbf{K}(n)\left(\mathbf{y}(n) - \overline{\mathbf{H}}(n)e^{j\widehat{\boldsymbol{\beta}}(n|n-1)}\right)\right\}, \tag{3.11}$$

$$\widehat{\boldsymbol{\beta}}(n|n-1) = \widehat{\boldsymbol{\beta}}(n-1|n-1), \tag{3.12}$$

$$\mathbf{K}(n) = \mathbf{M}(n|n-1)\mathbf{D}^{\dagger}(n) \times \left(\sigma^2 \mathbf{I}_{N_r} + \mathbf{D}(n)\mathbf{M}(n|n-1)\mathbf{D}^{\dagger}(n)\right)^{-1}, \tag{3.13}$$

$$\mathbf{M}(n|n-1) = \mathbf{M}(n-1|n-1) + \mathbf{Q}, \tag{3.14}$$

$$\mathbf{M}(n|n) = \Re\left\{\left(\mathbf{I}_{N_r} - \mathbf{K}(n)\mathbf{D}(n)\right)\mathbf{M}(n|n-1)\right\}, \tag{3.15}$$

where, for $k \in \{1, \cdots, N_t\}$,

$$\mathbf{D}(n) = \frac{\partial \overline{\mathbf{H}}(n)e^{j\boldsymbol{\beta}(n)}}{\partial \boldsymbol{\beta}^T(n)} = \begin{pmatrix} \mathbf{z}_1 & \mathbf{0}^T_{2Nt} & \mathbf{0}^T_{2Nt} \\ \mathbf{0}^T_{2Nt} & \ddots & \vdots \\ \mathbf{0}^T_{2Nt} & \mathbf{0}^T_{2Nt} & \mathbf{z}_{N_r} \end{pmatrix}, \tag{3.16}$$

$$\mathbf{z}_i = \begin{cases} h_{i,k}x_k(n)e^{j\widehat{\beta}_{i,k}(n|n-1)} & k \text{ is even} \\ h^{\text{SI}}_{i,k}x^{\text{SI}}_k(n)e^{j\widehat{\beta}_{i,k}(n|n-1)} & k \text{ is odd} \end{cases}, \tag{3.17}$$

and $\widehat{\beta}_{i,k}(n|n-1)$ is the $2(i-1)N_t + k$ element of vector $\widehat{\boldsymbol{\beta}}(n|n-1)$.

Remark 3.1 *We note that the state vector, as given by* (3.8)*, is a real vector. This is because the state vector only contains the phases, which are real numbers. The*

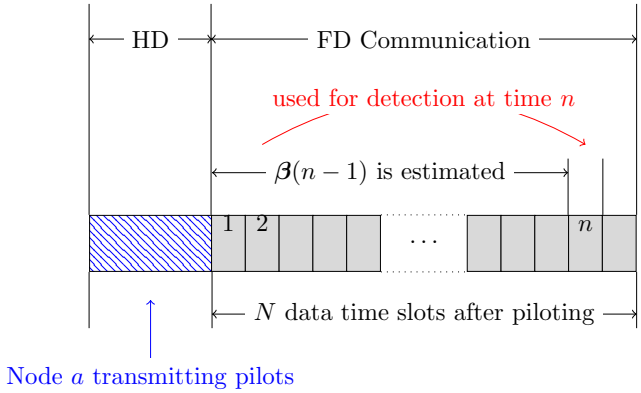


Figure 3.2: Time diagram of modified EKF.

complex channel coefficients are estimated using this estimated real vector and using the complex exponential function as given by (3.7)*. Since the states are all real, when updating the mean of the states in EKF we can safely discard the imaginary part of the updated mean as in* (3.11)*.*

3.2.1 Symbol Detection

The EKF equation (3.17) shows that $\mathbf{z}_i$ requires the knowledge of the constant channels $h_{i,k}$, $h_{i,k}^{\mathrm{SI}}$ and the transmitted symbols. Note that x_k^{SI}, the SI symbol is perfectly known at the receiver.

The knowledge of the constant channels can be obtained using pilot based estimation during the initial half-duplex (HD) phase of the communication. In addition, the transmitted symbols at time n are detected using the initial channel estimates and the estimates of the state vector $\boldsymbol{\beta}$ at time $n-1$. This is because at time n of the EKF algorithm $\boldsymbol{\beta}(n-1)$ has been successfully estimated. This procedure is shown in Fig. 3.2.

3.2.2 Lower Bound of Estimation Error

In this section we derive a lower bound on the estimation error of the estimator proposed in the previous subsection. We first note that mean square error (MSE)

Equation No.	Complexity
(3.11)	$\mathcal{O}(N_t N_r)$
(3.13)	$\mathcal{O}(N_t^2 N_r^3)$
(3.14)	$\mathcal{O}(N_t^2 N_r^2)$
(3.15)	$\mathcal{O}(N_t^3 N_r^3)$

Table 3.2: Complexity of each step of EKF algorithm.

for estimating the state vector $\boldsymbol{\beta}(n)$ is given by

$$\text{MSE} = trace\left(\mathbb{E}\left[\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(\boldsymbol{n})\right)\left(\boldsymbol{\beta}_n - \widehat{\boldsymbol{\beta}}_{\boldsymbol{n}}\right)^T\right]\right) \tag{3.18}$$

With above definition of the MSE vector, we present the following proposition.

Proposition 3.1 *MSE of the EKF is lower bounded by trace* $(\boldsymbol{Q})$*, i.e.,*

$$MSE \geq trace\left(\boldsymbol{Q}\right), \tag{3.19}$$

where $\boldsymbol{Q}$ *is the state covariance matrix given by* (3.9)*.*

Proof

See Appendix B.2.

Remark 3.2 *We note that* (3.19) *shows that the lower bound on the estimation error increases as the sum of diagonal elements of the covariance matrix of the states increases. Furthermore,* (3.9) *indicates that the diagonal elements of the state covariance matrix are the function of PN variance. Consequently, increasing the PN variance will result in worse estimation error. Since the residual SI cancellation is performed using the estimated SI channel, increasing the PN variance will result in worse SI cancellation performance. It is also worth to note that [130] shows that the PN variance is a monotonic increasing of function of carrier frequency. This means that the estimation error increases with increasing the carrier frequency and vice versa.*

3.2.2.1 Complexity Analysis of EKF

We present the following proposition on the complexity of the proposed EKF estimator

Proposition 3.2 *Table 3.2 shows the complexity of each step of EKF algorithm using $\mathcal{O}-$notation.*

Proof
For the complexity analysis of the proposed joint channel and PN estimation technique, we take the approach used by [104, 131] and count the number of multiplications and additions used in each step of EKF algorithm. The corresponding complexity calculations for this table can be found in Appendix B.1.

Remark 3.3 *According to Table* (3.2)*, the EKF has a polynomial complexity as a function of number of transmit N_t and receive N_r antennas. We can justify the increased complexity as follows. In [84], the authors propose an algorithm for channel estimation with linear complexity. However, the algorithm in [84] assumes a constant PN for a block of data. This could be an acceptable scenario in microwave communication but does not suit mmWave communication. Hence, the increased complexity of the proposed algorithm is justified because of fast variation of PN, i.e., PN variation over symbol time.*

3.3 Simulation Results

In this section we present simulation results for MIMO FD systems at 60 GHz frequency, which corresponds to mmWave frequency band [132]. For each simulation run we assume a communication packet is 40 symbol long, i.e., $N = 40$. This communication packet is transmitted after the training packet, which is $2N_t$ symbols long, and is used for estimating the constant channels for EKF initialization as described in Section 3.2.1. We then use $10,000$ simulation runs to obtain the desired simulation results.

Moreover, we use the assumptions presented in Section 3.1 to generate the random noise and PN. As summarized in [133], there are many mmWave channel models available for mmWave systems. In this work, similar to a large number

of existing works in [34, 43, 122, 134, 135], we adopt a general Rician model. Note that the proposed estimator is independent of the adopted model. A performance comparison of the different mmWave channel models is outside the scope of this work.

We generate the random SI and communication channel ($\mathbf{H}_{\text{SI/COM}}$) using Rician distribution as follows

$$\mathbf{H}_{\text{SI/COM}} = \sqrt{\frac{K}{K+1}}\mathbf{H}_{\text{LoS}} + \sqrt{\frac{1}{K+1}}\mathbf{H}_{\text{NLoS}}, \tag{3.20}$$

where, K is the Rician distribution K-factor, $\mathbf{H}_{\text{LoS}}$ is the LoS component of the channel and is generated assuming uniform distribution for angle of arrival, using the approach presented in [43], $\mathbf{H}_{\text{NLoS}}$ is the NLoS component of the channel and for both SI and communication channel is generated assuming Rayleigh fading. Furthermore, for both the SI and communication channel we set the K-factor to 2 dB.

We note that the SI and communication channels have different power intensities, i.e., $\mathbb{E}\left[\mathbf{H}_{\text{SI}}\mathbf{H}_{\text{SI}}^{\dagger}\right] \neq \mathbb{E}\left[\mathbf{H}_{\text{COM}}\mathbf{H}_{\text{COM}}^{\dagger}\right]$. Assuming that the LoS power of the residual SI (SI signal after the passive and analog cancellation) is the same as the LoS power of the communication signal, the signal to interference ratio (SIR) is given by

$$\text{SIR} = \frac{\sigma_{\text{COM}}^2}{\sigma_{\text{SI}}^2}, \tag{3.21}$$

where σ_{COM}^2 and σ_{SI}^2 are the variances of NLoS components of the communication and SI channels, respectively.

In addition, SNR is defined as

$$\text{SNR} \triangleq \frac{\mathbb{E}[E_s]}{\sigma^2}, \tag{3.22}$$

where E_s is the symbol energy, $\mathbb{E}[E_s] = 1$, and σ^2 is the noise variance.

Finally, we use the MSE for the state vector at time $N = 40$. This MSE is

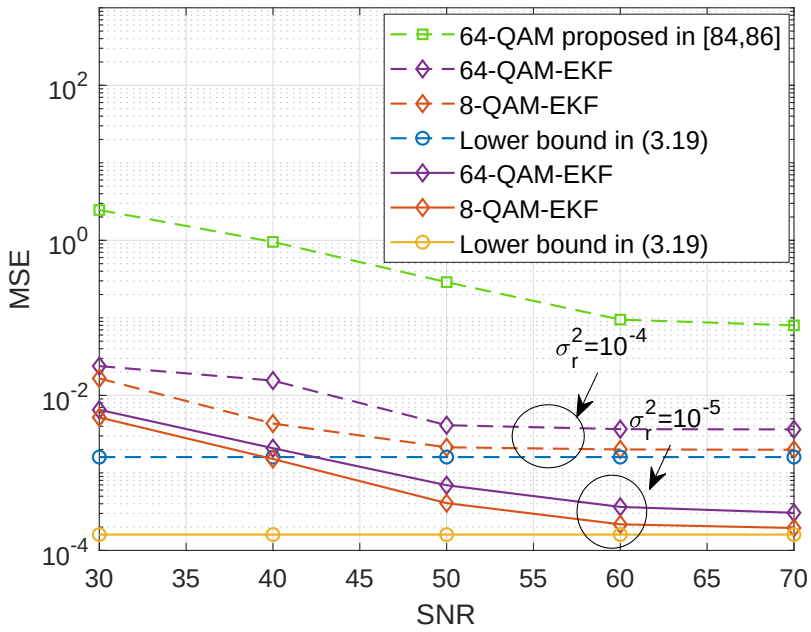


Figure 3.3: MSE performance for PN variances $\sigma_r^2 = \sigma_t^2 = \sigma_{\text{SI}}^2 = 10^{-4}, 10^{-5}$ and different QAM modulations for a 2×2 FD MIMO system with SIR$= 0$ dB.

given by rewriting (3.18) in terms of the Euclidean norm of a vector, i.e., $||\cdot||_2$,

$$\mathbb{E}\left[\left|\left|\boldsymbol{\beta}(N) - \widehat{\boldsymbol{\beta}}(N)\right|\right|_2\right]. \tag{3.23}$$

In what follows we first present the MSE results for different FD MIMO communication systems. We then investigate the residual SI power after digital cancellation and the bit error rate (BER) performance of these systems with the proposed PN estimation technique.

3.3.1 MSE Performance

In this section we investigate the MSE performance of the proposed PN estimation technique for a 2×2 FD MIMO system, and assume that SIR $= 0$ dB, i.e., the SI signal is as strong as the desired communication signal.

Fig. 3.3 shows the MSE performance of the proposed system against the derived theoretical bound in Section 3.2.2 for different quadrature amplitude mod-

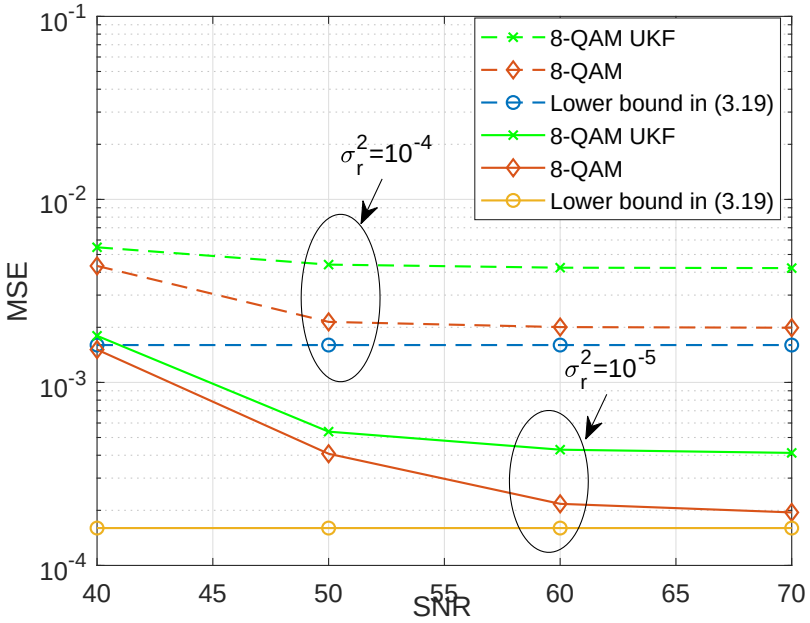


Figure 3.4: MSE performance of the UKF and proposed EKF for PN variances $\sigma_r^2 = \sigma_t^2 = \sigma_{\text{SI}}^2 = 10^{-4}, 10^{-5}$ and 8-QAM modulation for a 2×2 FD MIMO system with SIR$=0$ dB.

ulations (QAM) and different PN variances. Firstly, as discussed in Remark 3.2, with increasing PN variance the estimation performance degrades. Secondly, it can be observed from this figure that lower order modulations have better performance compared to the higher order modulations. This is because as shown in Section 3.2.1, the EKF algorithms requires to detect the transmitted symbols. Hence, the MSE of EKF is affected by the detection error. Finally, Fig. 3.3 shows that at high SNRs the MSE performance of the proposed estimator approaches the lower bound.

In Fig. 3.3, we also plot the MSE result of the state-of-the-art pilot-based phase noise estimator in [84,86] for microwave frequency. As expected, this estimator does not perform well compared to our proposed estimator. This is because it assumes that the PN variations are small, which is not applicable for the case for mmWave frequency. Note that we only show the MSE result of the estimator in [84, 86] for $64-$QAM modulation since the MSE performance is invariant with respect to the modulation order (the estimator uses pilots and does not require detection).

3.3.2 Comparison With Unscented Kalman Filter

We compare the performance of the proposed EKF estimator with Unscented Kalman Filter (UKF). UKF provides an alternative for linearizing the observations. The detailed implementation of the UKF is provided in Appendix B.3. Fig. 3.4 shows the performance of the EKF and UKF estimators for 8−QAM modulation, SIR= 0 dB and different PN variances. We can see that the MSE performance of the proposed EKF estimator is better than the UKF estimator. This is because: (i) UKF estimator works with the sigma points approximation of the mean of the state process, while EKF tracks the PN based on the true mean of the linear state vector, (ii) while the MSE performances of both EKF and UKF are degraded because of the detection error, this error affects UKF algorithm more than EKF. This is because the sigma points calculation are affected more by the error due to the symbol detection (Section 3.2.1), and (iii) UKF is inherently more suitable for the systems which experience high non-linearities, i.e., both the state and process models are non-linear and noise is non-linear too. In our case, only the process model in (3.7) is non-linear.

3.3.3 Residual SI Power

In this section we numerically investigate the remaining SI power after digital cancellation for a 2×2 MIMO FD system with 64−QAM modulation, assuming the PN variance for all the oscillators is 10^{-4}. This residual power is given by

$$P_{\mathrm{SI}} = \left|\left|\left(\mathbf{H}^{\mathrm{SI}}(n) - \overline{\mathbf{H}}^{\mathrm{SI}}(n)\right)\mathbf{x}^{\mathrm{SI}}(n)\right|\right|_2, \tag{3.24}$$

where $||\cdot||_2$ is the Euclidean norm of a vector, and $\overline{\mathbf{H}}^{\mathrm{SI}}(n)$ is an estimate of the SI channel using the proposed EKF estimator. Fig. 3.5 shows the residual SI power for different SIR values, where a SIR value of 0 dB indicates that passive and analogue cancellation stages have managed to reduce the SI power to the same level as the desired signal power.

The numerical result of Fig. 3.5 shows that the performance of digital canceller depends on the residual SI power after passive and analog cancellation stages. As the residual SI power after passive and analogue cancellation decreases, so does the

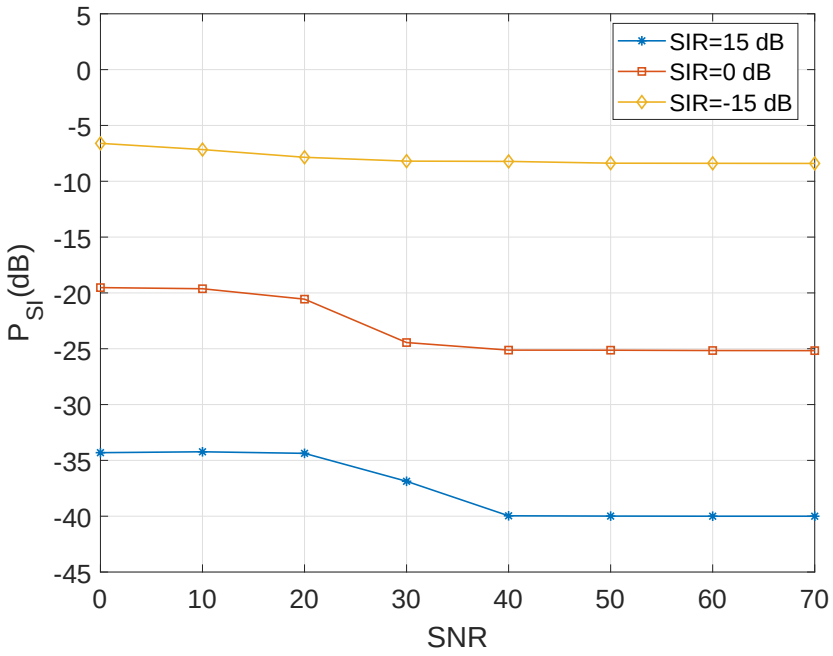


Figure 3.5: The residual SI power P_{SI} after digital cancellation.

residual SI power after the digital cancellation. The results show that the residual SI power can be reduced to -25 and -40 dB for SIR of 0 and 15 dB, respectively. This is important as it shows the effectiveness of digital SI cancellation after passive and analogue cancellation.

3.3.4 BER Performance

Finally, in this section we present the BER results of a 2×2 FD MIMO system with different QAM modulations, assuming that PN variance for all oscillators is 10^{-4}. Fig. 3.6 shows the BER performance of the system for different values of SIR. The results are consistent with the results of the residual SI power in Fig. 3.5, i.e., the higher the SIR the better the BER results. Furthermore, $8-$QAM system performs better than the $64-$QAM system, which is consistent with the results of Fig. 3.6.

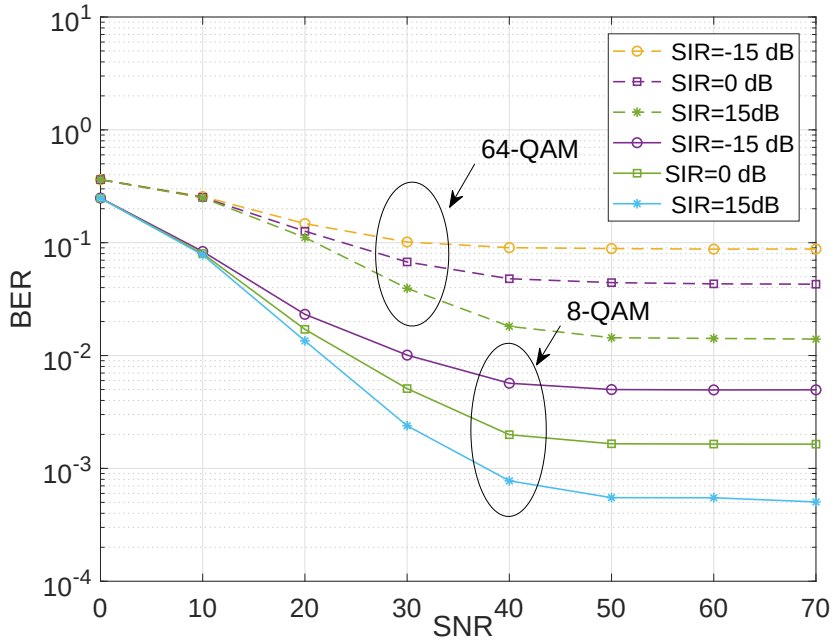


Figure 3.6: BER performance of the proposed system for a 2×2 MIMO FD system with different QAM modulations.

3.4 Conclusions

In this chapter we considered a MIMO FD system for mmWave communication and proposed a joint channel and PN estimation algorithm[2]. We also derived a lower bound on the estimation error and numerically showed that the MSE of the proposed estimator approaches the error bound. Furthermore, we investigated the residual SI power after the digital cancellation and showed that the digital canceller, which uses the estimated SI channel can reduce the SI power to -25 to -40 dB. These results indicate the effectiveness of digital cancellation after passive and analogue cancellation stages.

[2]Indeed, the main focus of this work is to correctly estimate the channel and PN for effective SI cancellation. In case of inter-node interference [136], the proposed estimator would need to be modified. However, in the special case, if the inter-node interference can be treated as Gaussian, then the system model given by (3.1) can capture the effect of the inter-node interference by including an additional Gaussian noise term due to inter-node interference.

Chapter 4

Data Detection in Full-Duplex Communication Systems

The corrupted received signal due to the presence of strong SI signal makes data detection in FD communication challenging. To remove the SI signal and recover the SI-free signal, the channels are estimated and used to subtract a copy of the known SI signal from the received signal. Consequently, not only the quality of the received signal depends on the accuracy of the channel estimates and SI cancellation, but also processing delay is added due to the extra processing stages. Channel estimation and SI cancellation can be avoided if the statistical properties of the channels are known. However, detecting the data symbols solely using the statistical properties of the channels instead of accurate channel estimates, introduces an ambiguity that needs to be resolved.

This chapter of the thesis investigates the problem of blind data detection, where detection is based on the statistical properties of the channels. In doing so it mathematically formulates the ambiguity problem associated with blind detection, and shows how it can be resolved. The remainder of this chapter is organised as follows. Section 4.1 presents the system model. Section 4.2 formulates the MAP detector for data detection in the absence of channel estimates and illustrates the ambiguity problem associated with the MAP detector. Section 4.3 proposes a superimposing technique to resolve the detection ambiguity problem. Section 4.4 presents and discusses the simulation results. Finally, Section 4.5 concludes the

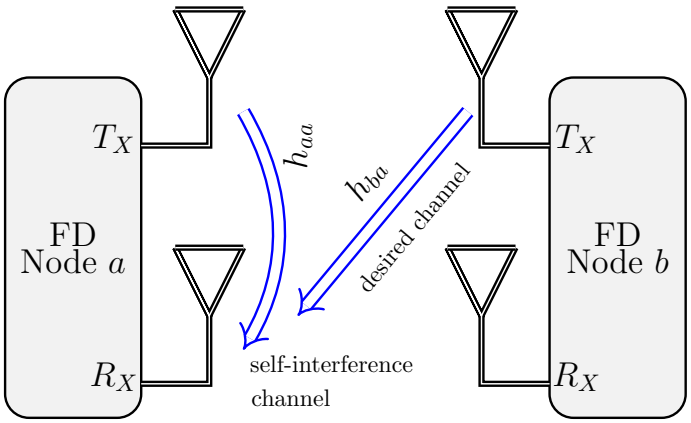


Figure 4.1: Full Duplex System with single transmit and receive antenna. The single antenna at each node is shown separately for the transmission and reception for ease of illustration.

chapter.

4.1 System Model

We consider the data detection problem for the single-input single-output (SISO) FD communication system, as shown in Fig. 4.1. Nodes a and b each have a pair of antennas, which is used for simultaneously transmit and receive on the same frequency band. Due to the inherent symmetry of the problem, we only investigate the data detection problem for node a, as identical results are expected for node b.

The received signal at node a is given by

$$\mathbf{y}_a = h_{aa}\mathbf{x}_a + h_{ba}\mathbf{x}_b + \mathbf{w}_a, \tag{4.1}$$

where, $\mathbf{y}_a \triangleq [y_{a_1}, \cdots, y_{a_N}]^T$ is the $N \times 1$ vector of received symbols, $\mathbf{x}_a \triangleq [x_{a_1}, \cdots, x_{a_N}]^T$ is the $N \times 1$ vector of self-interference symbols, $\mathbf{x}_b \triangleq [x_{b_1}, \cdots, x_{b_N}]^T$ is the $N \times 1$ vector of desired communication symbols, $\mathbf{w}_a \triangleq [w_{a_1}, \cdots, w_{a_N}]^T$ is the $N \times 1$ vector of independent identically distributed (i.i.d.) Gaussian noise with zero mean and variance σ^2, i.e., $w_{ai} \sim \mathcal{CN}(0, \sigma^2)$.

We make the following assumptions in this thesis:

- Since the digital channels are the channels observed after the passive and

RF cancellation stages, the direct line-of-sight (LoS) components of these channels have already been canceled and the residual components are due to the scatterers [15, 27]. Consequently, similar to [30, 55], we assume h_{aa} and h_{ba} are flat-fading and Rayleigh distributed with zero mean and variance one, i.e., $h_{aa}, h_{ba} \sim \mathcal{CN}(0, 1)$. Furthermore, we note that the assumption of same variance for both SI and communication channels corresponds to the worst case scenario, where the SI channel is as strong as the communication channel.

- The transmitted symbols are modulated using the modulation set $\mathcal{A} = \{A_1, A_2, ..., A_M\}$, with size M. Modulation set $\mathcal{A}$ contains all constellation points of any given standard modulation constellation, such as M-ary phase shift keying (MPSK) modulation, and the transmitter is likely to send each constellation point with equal probability.

4.2 Blind Data Detection in FD Communication

In this section, we first derive a blind MAP symbol detector for the FD communication system, which does not require channel estimates and SI cancellation. Then we show that this blind detector suffers from the detection ambiguity problem because of the symmetry of conventional modulation constellations around the origin.

4.2.1 MAP Detector

The main results in this section are presented in the following propositions.

Proposition 4.1 *The blind MAP symbol detector for the SISO FD communication system presented in Section 4.1 is given by*

$$\widetilde{x}_{b_i} = \max_{x_{b_i}} f(x_{b_i}|\mathbf{y}_a). \tag{4.2}$$

where the marginal probability distribution $f(x_{b_i}|\mathbf{y}_a)$ *is proportional to*

$$f(x_{b_i}|\mathbf{y}_a) \propto \underbrace{\sum_{j_N=1}^{M} \cdots \sum_{j_1=1}^{M}}_{\sim j_i} \frac{1}{\lambda} \exp\left(\frac{|\xi|^2}{\lambda\sigma^2}\right), \tag{4.3}$$

where M *is the size of modulation set* $\mathcal{A}$*,* N *is the length of the transmitted vector, i.e., number of transmitted symbols in a transmission block, and*

$$\lambda \triangleq \sum_{n=1,n\neq i}^{N} |A_{j_n}|^2 - \frac{1}{\gamma}\left|\sum_{n=1,n\neq i}^{N} x_{a_n}^* A_{j_n} + x_{a_i}^* x_{b_i}\right|^2 + |x_{b_i}|^2 + \sigma^2, \tag{4.4}$$

$$\xi \triangleq \sum_{n=1,n\neq i}^{N} y_{a_n} A_{j_n}^* + y_{a_i} x_{b_i}^* - \frac{1}{\gamma}\sum_{n=1}^{N} y_{a_n} x_{a_n}^* \left(\sum_{n=1,n\neq i}^{N} x_{a_n}^* A_{j_n} + x_{a_i}^* x_{b_i}\right)^*, \tag{4.5}$$

$$\gamma \triangleq \sum_{n=1}^{N} |x_{a_j}|^2 + \sigma^2. \tag{4.6}$$

Proof: See Appendix C.1.

Note that the proportionality in (4.3) does not depend on the residual self-interference symbol x_{b_i} and, hence, does not affect the decision in (4.2).

Remark 4.1 *The posterior PDF* $f(x_{b_i}|\mathbf{y}_a)$ *is independent of both the SI and communication channels, i.e.,* h_{aa} *and* h_{ba}*. Hence, the MAP detector as proposed by Proposition 4.1 is independent of the channel estimates. In other words, the symbols can be detected without requiring the interference or communication channel to be estimated. The MAP detector also directly detects the symbols without requiring a separate SI cancellation stage. Consequently, the proposed MAP detector is a blind detection technique that only uses the statistical properties of the channels, i.e., the Rayleigh fading assumption in Section 4.1.*

Proposition 4.2 *We call* $\mathcal{A}$ *a symmetric modulation set, if and only if for* $x_k \in \mathcal{A}$*, there exists* $-x_k \in \mathcal{A}$*,* $\forall k$*. The posterior PDF* $f(x_{b_i}|\mathbf{y}_a)$ *does not have a unique maximum if and only if* x_{b_i} *in* (4.3) *comes from a symmetric modulation set* $\mathcal{A}$*.*

Proof: See Appendix C.2.

Corollary 4.2.1 *Since conventional modulation constellations are symmetric around the origin, data detection in FD communication with no channel estimation will result in ambiguity.*

4.3 Superimposed Signaling for Resolving the Ambiguity of Blind Data Detection

In this section we present a superimposed signalling technique to tackle the inherent ambiguity problem in data detection with no available channel estimates.

4.3.1 Why Superimposed Signalling?

The rationale for using superimposed signalling is as follows. From Proposition 4.2, the data detection ambiguity in FD communication in the absence of channel estimates, arises because of the symmetry of the modulation constellation around the origin. Consequently, an obvious approach to resolve the data detection ambiguity is to alter the symmetry of the modulation constellation around the origin and create a suitable asymmetric modulation constellation.

One simple way to achieve an asymmetric modulation constellation around the origin is to add (superimpose) a constant known signal to the transmitted signal.[1] We call this approach superimposed signalling. For illustration, Fig. 4.2 shows the effect of superimposed signalling with constant P on the constellation of an $M = 4$-PSK modulation set. Once the M-PSK constellation is shifted, then the new constellation is asymmetric around the origin and can be used for ambiguity-free MAP detection with no need for channel estimation.

[1]The design of optimum asymmetric modulation constellations is outside the scope of this thesis and is the subject of future work [104].

4.3.2 Modified System Model

If both nodes a and b superimpose a common constant and known signal P to the transmitted symbols, then (4.1) can be written as:

$$\mathbf{y}_a = h_{aa}(\mathbf{x}_a + P) + h_{ba}(\mathbf{x}_b + P) + \mathbf{w}_a. \tag{4.7}$$

It is again clear from (4.7) that the effect of superimposed signalling with constant signal P is the same as shifting the modulation constellation by P along the horizontal axis.

4.3.3 Power Normalization

As illustrated above, superimposed signalling increases the average energy per symbol of the modulation constellation. Conventional (symmetric) modulations operate under an average transmit power constraint, which places limits on the average energy per symbol. A fundamental question regarding superimposed signaling is, therefore, how to choose a fair value of the extra power which is required to superimpose a known signal on the data symbols to shift the modulation constellation.

If the channels were perfectly known there would be no need to allocate power for channel estimation. However, in reality the channels are unknown and hence it is inevitable to expand extra power for channel estimation. The proposed superimposed signalling approach is similar in spirit to superimposed training in the literature, which has been extensively used as a bandwidth-efficient channel estimation technique in half-duplex (HD) communication systems [137, 138]. In superimposed training, the extra power in the superimposed pilots is used for channel estimation. In our case, we do not use the extra power for channel estimation. Rather, we use it only for achieving an asymmetric modulation constellation. Consequently, to ensure that the proposed method does not exceed the average transmit power constraint, we shift the modulation by $P \triangleq \sqrt{E_p}$, where E_p is the average energy used for channel estimation in conventional pilot based channel estimation systems.

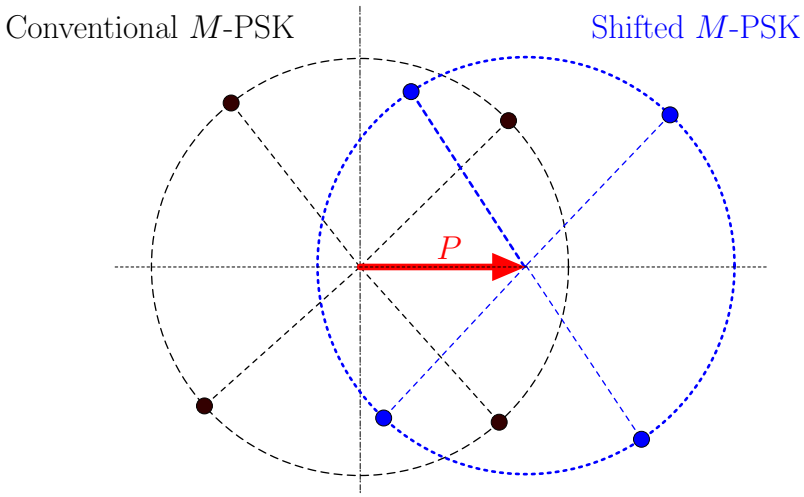


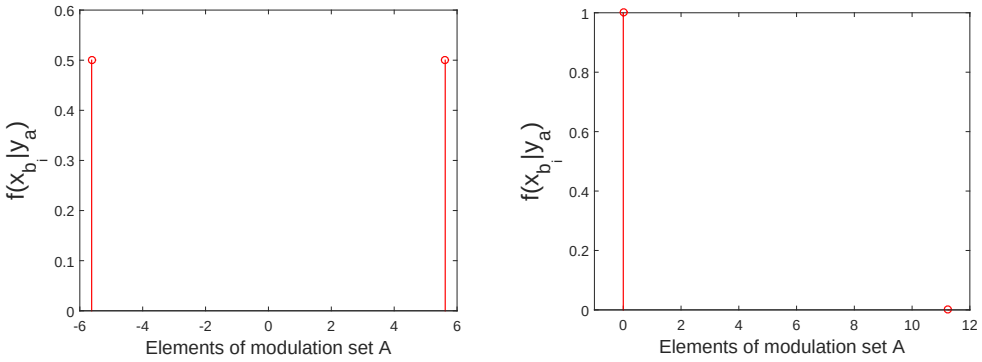
Figure 4.2: Effect of superimposed signalling on the modulation constellation of $M = 4$-PSK.

4.4 Simulation Results

In this section, we present the simulation results. First we demonstrate that detection without channel estimation, using symmetric modulation constellation can result in ambiguity. Then we show that this ambiguity is resolved once the modulation set is shifted to a asymmetric modulation set, i.e., a known signal is superimposed on the data signal. We find the minimum power required for superimposed signaling to resolve the ambiguity problem. Finally, we investigate the BER performance of the proposed detector. Throughout this section we make the following assumptions:

- *Channel and noise*: For each run of the simulation, the random channels h_{aa} and h_{ba} are generated according to a Rayleigh distribution and are assumed constant for blocks of N symbols, i.e. block fading. We assume independent block fading for simulation purposes which means channels are independent from block to block, i.e., quasi-static.

- *Modulation*: For the sake of simplicity, we only present the result for binary shift keying (BPSK) modulation. Consequently, the modulation set $\mathcal{A}$ has two elements.



(a) Symmetric BPSK modulation. (b) Asymmetric BPSK modulation.

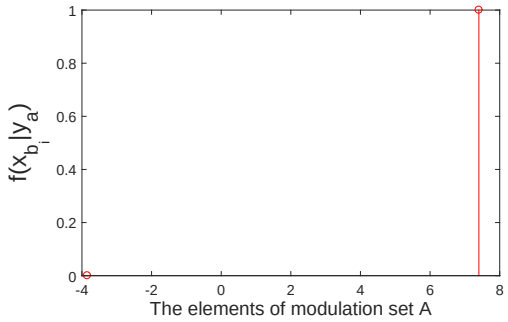
Figure 4.3: Posterior function $f(x_{b_i}|\mathbf{y}_a)$ at $\frac{E_b}{N_o} = 15$ dB.

- *Noise and shift powers*: We assume the average bit energy of the modulation is E_b and noise power is $N_0 = 1$.

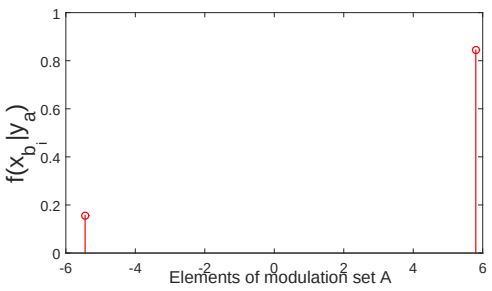
4.4.1 Symmetric Modulation Set

In this section we highlight the result of Proposition 4.2 through simulations.

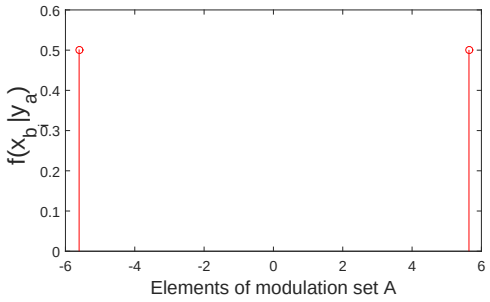
For symmetric BPSK modulation the posterior function $f(x_{b_i}|\mathbf{y}_a)$ takes two discrete values. Fig. 4.3(a) shows the posterior function at $\frac{E_b}{N_o} = 15$ dB when symmetric BPSK modulation is used. It is clear from Fig. 4.3(a) that when this modulation constellation is used the posterior function does not have a unique maximum and hence the MAP detector of (4.2) results in ambiguity. This ambiguity is seen as equal probability for the elements of modulation set $\mathcal{A}$ in Fig. 4.3(a). Fig. 4.3(b) shows the posterior function at $\frac{E_b}{N_o} = 15$ dB when the modulation constellation is shifted by $P \triangleq \sqrt{E_b}$. It is clear that in this case, the posterior function has only one maximum and consequently the MAP detector as proposed by Proposition 4.1 results in no ambiguity. This is because now the elements of modulation set $\mathcal{A}$ have different probabilities, hence, the detector can determine which element is more likely to be transmitted given the received data.



(a) $\beta = 0.1$



(b) $\beta = 0.001$



(c) $\beta = 0.00001$

Figure 4.4: Posterior function $f(x_{b_i}|\mathbf{y}_a)$ for different values of β

4.4.2 Minimum Required Energy for Superimposed Signalling

Although Fig. 4.3(b) shows that the ambiguity of the MAP detector is resolved by shifting the modulation constellation, this comes at the cost of increasing the transmit power by the shift power ($|P|^2 \triangleq E_b$). We are interested in the minimum required power for ambiguity-free MAP detector. Consequently, for $0 < \beta < 1$, we set the shift to $P \triangleq \sqrt{\beta E_b}$ and numerically investigate the minimum value for β.

Fig. 4.4 shows the posterior function $f(x_{b_i}|\mathbf{y}_a)$ for different values of β. Clearly, as β decreases, the difference between the maximum and minimum value of the posterior function increases, such that for $\beta = 0.00001$, the posterior function does not have a unique maximum. Fig. 4.4(b) shows that $\beta = 0.001$ is sufficient enough for ambiguity-free MAP detection. However, our simulation results show that for the FD system under consideration to have a stable detection performance for

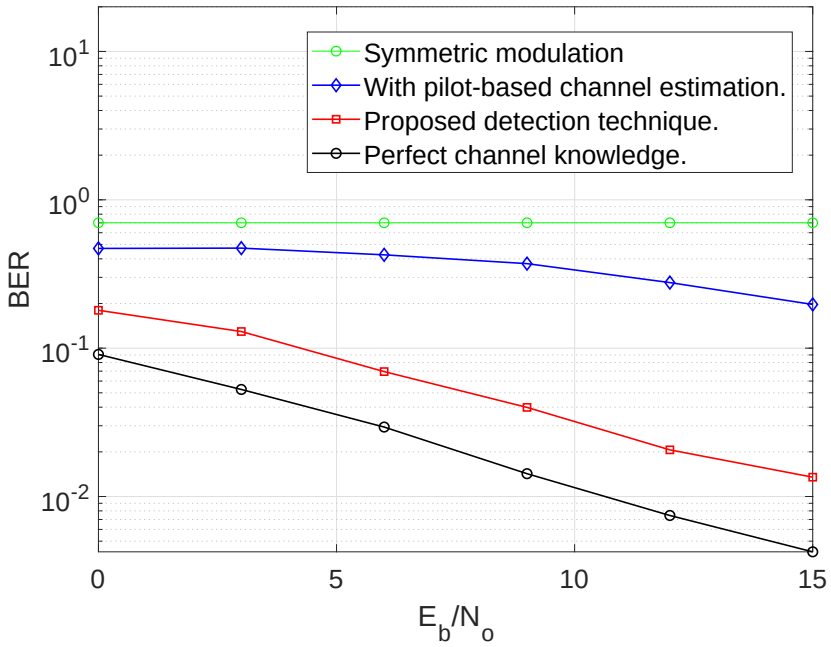


Figure 4.5: BER performance of FD communication system with different availability of channel estimates.

different channel realizations, the minimum value for β is 0.1.

4.4.3 BER Performance

In this section we investigate the BER performance of the proposed detector. For simplicity, we only present the results for BPSK modulation in the presence of a self-interference signal which is as strong as the desired signal. We also set the shift to $P \triangleq \sqrt{0.1E_b}$. The BPSK BER with perfect channel knowledge is plotted as a reference. The performance of the proposed detector is compared with a conventional channel estimation-based detection method, assuming the channel estimation uses the same extra power as the superimposed signal for channel estimation. In the channel estimation-based detection method, the channels are first estimated using the same extra energy as the superimposed signal and then these estimates are used for data detection.

Fig. 4.5 shows that when the modulation constellation is symmetric around

the origin and no channel estimates are available, then the detector fails to detect the symbols, i.e., all the possible outcomes are equally likely for the transmitted symbols (c.f. Fig. 4.3(a)). However, shifting the modulation set to an asymmetric modulation set resolves the ambiguity. In addition, the performance of the proposed detection method is better than the conventional pilot-based detection method.

4.5 Conclusion

In this thesis, we demonstrated that the detection of symbols in FD communication systems with no channel estimation results in ambiguity. We proposed a solution to this ambiguity problem using superimposed signaling, which involves shifted modulation constellations. We proposed a MAP detector to be used with the shifted modulation constellation in FD communication system for data detection without channel estimation. Our results showed that the proposed detection method has better BER performance, compared to conventional channel estimation-based detection method. The proposed method is bandwidth efficient and can be used in any system model where the self-interference signal is known, such as in two-way relay networks and multi-hop one way relay networks [76, 139].

Chapter 5

Self-Interference Suppression in Full-Duplex Massive MIMO Communication

FD enabled massive MIMO BS can be more bandwidth efficient compared to its HD counterpart, if the strong SI signal associated with FD communication is suppressed. This can be done by exploiting the extra degree of freedom provided by the antenna array in the spatial domain. One way of exploiting this degree of freedom is to spatially suppress the SI signal using transmit precoder. However, the challenge in designing transmit precoders is to simultaneously suppress both SI in uplink and MUI in downlink. Hence, this problem first requires a rigorous mathematical foundation to show that transmit precoders can be designed that can suppress both SI and MUI. If this problem has a solution then transmit precoders can be designed to suppress SI and MUI simultaneously.

This chapter of thesis first mathematically formulates the SI and MUI suppression problem, and discusses the existence of a solution for it. It then considers a massive MIMO BS with particular antenna array configuration, and proposes a precoder to suppress both SI and MUI. The remainder of this chapters is organised as follows. In Section 5.1 we present the system model. Then in Section 5.2 the problem of SI and MUI suppression for FD enabled massive MIMO is formulated and its solution is discussed. In this same section, we also propose a precoder matrix for FD

massive MIMO system when the number of transmit antennas is the same as the sum of the number of receiver antennas and the number of uplink users. Finally, Section 5.4 concludes the chapter.

5.1 System Model

We consider a single-cell multi-user scenario, as illustrated in Fig. 5.1, where a single FD massive MIMO BS with N_{T} transmit and N_{R} receive antennas, simultaneously serves K single-antenna HD users in the UL, and U single-antenna HD users in the DL[1]. For this system model, the received and transmitted vectors by the BS in the UL and the DL are given by

$$\mathbf{y}_{\mathrm{UL}} = \underbrace{\mathbf{P}_{\mathrm{R}}\mathbf{H}_{\mathrm{UL}}\mathbf{x}_{\mathrm{UL}}}_{\text{Desired signal}} + \underbrace{\mathbf{P}_{\mathrm{R}}\mathbf{H}_{\mathrm{SI}}\mathbf{P}_{\mathrm{T}}\mathbf{x}_{\mathrm{DL}}}_{\text{SI}} + \mathbf{n}_{\mathrm{DL}}, \tag{5.1a}$$

$$\mathbf{y}_{\mathrm{DL}} = \underbrace{\mathbf{H}_{\mathrm{DL}}\mathbf{P}_{\mathrm{T}}\mathbf{x}_{\mathrm{DL}}}_{\text{Desired signal}} + \mathbf{n}_{\mathrm{DL}}, \tag{5.1b}$$

respectively, where, $\mathbf{y}_{\mathrm{UL}}$ is the $K \times 1$ received vector by the BS in the UL, $\mathbf{y}_{\mathrm{DL}}$ is the $U \times 1$ received vector sent to user equipments (UEs) in the DL, $\mathbf{x}_{\mathrm{UL}}$ is $K \times 1$ vector of UL symbols, $\mathbf{x}_{\mathrm{DL}}$ is $U \times 1$ vector of DL symbols, $\mathbf{P}_{\mathrm{R}}$ is the $K \times N_{\mathrm{R}}$ receive beamforming matrix, $\mathbf{P}_{\mathrm{T}}$ is the $N_{\mathrm{T}} \times U$ transmit beamforming (precoder) matrix, $\mathbf{n}_{\mathrm{DL}}$ is $U \times 1$ Gaussian noise vector of DL UEs, i.e., $\mathbf{n}_{\mathrm{DL}} \sim \mathcal{CN}(\mathbf{0}_U, \sigma^2_{\mathrm{DL}}\mathbf{I}_U)$, and $\mathbf{n}_{\mathrm{UL}}$ is $K \times 1$ Gaussian noise vector at BS, i.e., $\mathbf{n}_{\mathrm{UL}} \sim \mathcal{CN}(\mathbf{0}_K, \sigma^2_{\mathrm{UL}}\mathbf{I}_U)$, where, $\mathbf{I}_U$ and $\mathbf{I}_K$ are $U \times U$ and $K \times K$ identity matrices, respectively, $\mathbf{H}_{\mathrm{UL}}$ is $N_{\mathrm{R}} \times K$ channel matrix between the N_{R} receive antennas of the BS and the K UL users, $\mathbf{H}_{\mathrm{SI}}$ is $N_{\mathrm{R}} \times N_{\mathrm{T}}$ SI channel matrix between the transmit and the receive antennas of the BS, $\mathbf{H}_{\mathrm{DL}}$ is the $U \times N_{\mathrm{T}}$ DL channel matrix. Furthermore, we make the following assumptions about the statistical properties of the channels. The UL, and DL channels are Rayleigh fading channels with their entries drawn from $\mathcal{CN}(0, 1)$ [100,101]. On the other hand, because of the strong line of sight component (LoS), the SI channel

[1]We note that considering two sets of different users for UL and DL allows for the problem to be formulated in its most general form. This system model has also been considered in [102], however, the authors of [100,140] consider a special case where there is only one set of users being served in DL or UL.

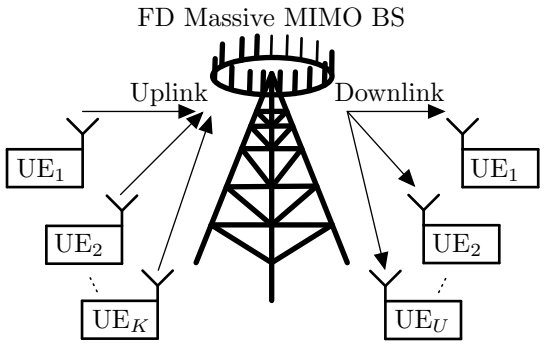


Figure 5.1: A FD massive MIMO BS serving U UL and K DL users.

is Rician fading with the entries drawn from $\mathcal{CN}(\mu, \sigma^2)$ [100, 101]. We ignore the large-scale fading as it does not impact the main analysis in this work. Finally, we assume there are no hardware imperfections at the FD massive MIMO BS [141]. In addition, perfect knowledge of channel state information is available at the BS. Note that the channel estimation can be carried out using the techniques such as [142] and it is outside the scope of this work.

5.2 Spatial Suppression of SI and MUI

5.2.1 Problem Formulation

For spatial suppression of the SI and MUI, the transmit precoding matrix should be designed such that the following two equations hold simultaneously

$$\mathbf{H}_{\mathrm{DL}}\mathbf{P}_{\mathrm{T}}\mathbf{x}_{\mathrm{DL}} = \mathbf{x}_{\mathrm{DL}}, \tag{5.2a}$$

$$\mathbf{H}_{\mathrm{SI}}\mathbf{P}_{\mathrm{T}}\mathbf{x}_{\mathrm{DL}} = \mathbf{0}. \tag{5.2b}$$

Note that (5.2a) allows for the DL communication with no MUI, and (5.2b) completely suppresses SI signal in the UL. Using (5.2), the spatial SI and MUI suppression can be re-formulated as the following system of equations:

$$\mathbf{H}\mathbf{u} = \mathbf{b}, \tag{5.3}$$

where, $\mathbf{H} \triangleq \begin{bmatrix} \mathbf{H}_{\mathrm{DL}} \\ \mathbf{H}_{\mathrm{SI}} \end{bmatrix}$ is a $N \times N_{\mathrm{T}}$ matrix, $\mathbf{u} \triangleq \mathbf{P}_{\mathrm{T}}\mathbf{x}_{\mathrm{DL}}$ is a $N_{\mathrm{T}} \times 1$ vector, $\mathbf{b} \triangleq \begin{bmatrix} \mathbf{x}_{\mathrm{DL}} \\ \mathbf{0}_{(N-U)\times 1} \end{bmatrix}$ is a $N \times 1$ vector, and N is a design parameter. By designing a particular value of N, we can transmit zeros to all or subset of receive antennas and/or UEs.

Existence of Solution: If (5.3) has a solution then the precoding matrix $\mathbf{P}_{\mathrm{T}}$ can be designed. However, because of the random channel matrices $\mathbf{H}_{\mathrm{DL}}$ and $\mathbf{H}_{\mathrm{SI}}$, $\mathbf{H}$ in (5.3) is also a random matrix, which cannot simply assumed to be full-rank. This means the existence of a solution for (5.3) is not always guaranteed. In addition, since the dimensions of $\mathbf{H}_{\mathrm{DL}}$ and $\mathbf{H}_{\mathrm{SI}}$ are determined by the number of transmit and receive antennas, for (5.3) to have a solution a condition is imposed on the size of transmit and receive antenna arrays as presented in the following proposition.

Proposition 5.1 *For a FD massive MIMO BS with N_T transmit antennas and N_R receive antennas, serving K single-antenna HD users in the UL and U single-antenna HD users in the DL,* (5.3) *has a solution only if $N_T \geq N_R + U$.*

Proof: We prove Proposition 5.1 using the method of forward and backwards reasoning. With forward reasoning, we show that if there is a solution to the system of equations given by (5.3) then $N_{\mathrm{T}} \geq N_{\mathrm{R}} + U$. With backwards reasoning, we show that if $N_{\mathrm{T}} \geq N_{\mathrm{R}} + U$, then (5.3) has a solution. The details are in Appendix D.1.

Remark 5.1 *The result of Proposition 5.1 explains why the number of transmit and receive antennas cannot be the same in a FD massive MIMO BS. In addition, it allows us to relate to and validate the assumptions made in the earlier works. For example, the authors in [102] assume that $N_T = 3N_R$ and $N_R \gg U$, and $\mathbf{H}$ in* (5.3) *has a right inverse. Then, they solve* (5.3) *for $\mathbf{b} = \begin{bmatrix} \mathbf{x}_d \\ \mathbf{0}_{N_R\times 1} \end{bmatrix}$. Similarly, the authors of [100] assume that $N_T \geq N_R + U$ and $\mathbf{H}$ in* (5.3) *has a right inverse and solve* (5.3) *for $\mathbf{b} = \begin{bmatrix} \mathbf{x}_d \\ \mathbf{0}_{U\times 1} \end{bmatrix}$. Both of these assumptions are special cases of* (5.3).

5.2.2 General Solution

Remark 5.1 illustrates the importance of finding a solution for the general problem in (5.3), i.e., solution of the system when the vector $\mathbf{b}$ in (5.3) has an arbitrary dimension of $N \times 1$. We present the solution in the theorem below.

Theorem 5.1 *For a FD massive MIMO BS, with the system model as given in Section 5.1, if $N_T \geq N_R + U$, then the matrix $\mathbf{H}$ given by* (5.3) *is full-rank with probability (w.p.)* 1*. Hence,* (5.3) *has a unique solution w.p.* 1*, given by*

$$\mathbf{u} = (\mathbf{H}\mathbf{H}^*)^{-1}\mathbf{H}^*\mathbf{b}, \tag{5.4}$$

where $(\mathbf{H}\mathbf{H}^*)^{-1}\mathbf{H}^*$ is the right inverse of $\mathbf{H}$, and $\mathbf{b}$ is a $N \times 1$ vector.

Proof: To prove Theorem 5.1, the main contribution is to rigorously show that $\mathbf{H}\mathbf{H}^*$ is full-rank (See Appendix D.2). Once it is established that $\mathbf{H}\mathbf{H}^*$ is full-rank, this implies that $\mathbf{H}$ is also full-rank [143]. Thus the right inverse exists and the unique solution is as given by (5.4).

Special Case: For the special case $N_\mathrm{T} = N_\mathrm{R} + U$, $\mathbf{H}$ becomes a square matrix and (5.4) simplifies to $\mathbf{u} = \mathbf{H}^{-1}\mathbf{b}$, where, $\mathbf{b} = \begin{bmatrix} \mathbf{x}_d \\ \mathbf{0}_{(N_\mathrm{R}+U)\times 1} \end{bmatrix}$.

Then it is easy to see that the precoder matrix that allows for both SI and MUI suppression is given by,

$$\mathbf{P}_\mathrm{T} = \mathbf{H}^{-1} \begin{bmatrix} \mathbf{I}_U \\ \mathbf{0}_{(N_\mathrm{R}+U)\times U} \end{bmatrix}, \tag{5.5}$$

i.e., the design parameter $N = N_\mathrm{R} + U$ in this case.

5.3 Simulation Results

In this section we present simulation results to validate and illustrate the performance of the proposed precoder design. Unless otherwise stated, we set the parameters as follows: $U = K = 5$, $N_\mathrm{T} = N_\mathrm{R} + U$, $\sigma^2_\mathrm{DL} = \sigma^2_\mathrm{UL}$, transmit power $E[\mathbf{x}_\mathrm{DL}\mathbf{x}^*_\mathrm{DL}] = N_\mathrm{T}$, where, $E[\cdot]$ is the expectation operator, signal-to-noise-ratio SNR $\triangleq \frac{E[\mathbf{x}^*_\mathrm{DL}\mathbf{x}_\mathrm{DL}]}{\sigma^2_\mathrm{UL}}$, and residual SI power is given by $E\left[\mathbf{H}_\mathrm{SI}\mathbf{P}_\mathrm{T}\mathbf{x}_\mathrm{DL}\left(\mathbf{H}_\mathrm{SI}\mathbf{P}_\mathrm{T}\mathbf{x}_\mathrm{DL}\right)^*\right]$.

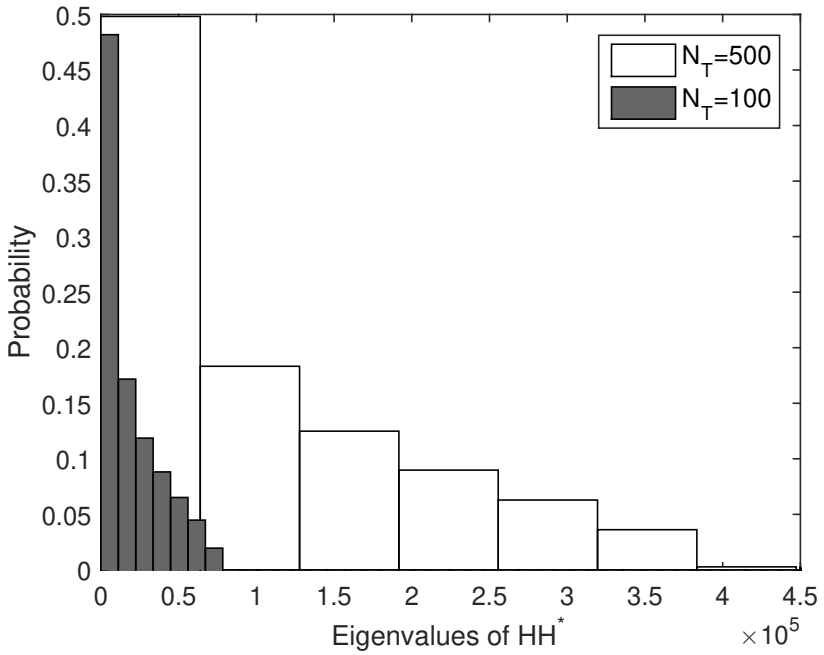


Figure 5.2: Empirical PDF of eigenvalues of $\mathbf{HH}^*$.

5.3.1 Validation

First we present results to validate Theorem 5.1. According to Theorem 5.1, $\mathbf{H}$ is full-rank w.p. 1. This means that the eigenvalues of $\mathbf{HH}^*$ must be positive [143]. Fig. 5.2 plots the empirical probability density function (PDF) of eigenvalues of $\mathbf{HH}^*$, where $\mathbf{H}$ is given by (5.3), for $N_{\mathrm{T}} = 100$ and 500. The results are generated using 10^5 Monte Carlo simulation runs. We can see that the PDFs are all zero for negative values of Eigenvalues. This agrees with Theorem 5.1.

5.3.2 Comparison with Existing Work

Fig. 5.3 plots the residual SI power versus number of transmit antenna elements, assuming SNR = 10 dB. We compare the performance of the following schemes: (i) no SI suppression, (ii) standard massive MIMO transmit beamforming, i.e., $\mathbf{P}_{\mathrm{T}} = \mathbf{H}_{\mathrm{DL}}^*$ [144], (iii) the precoder matrix based on the right inverse of $\mathbf{H}$ as used in [100–102], i.e., $\mathbf{P}_{\mathrm{T}} = (\mathbf{HH}^*)^{-1}\mathbf{H}^* \begin{bmatrix} \mathbf{I}_U \\ \mathbf{0}_{(N_{\mathrm{R}}+U)\times U} \end{bmatrix}$, (iv) the precoder matrix

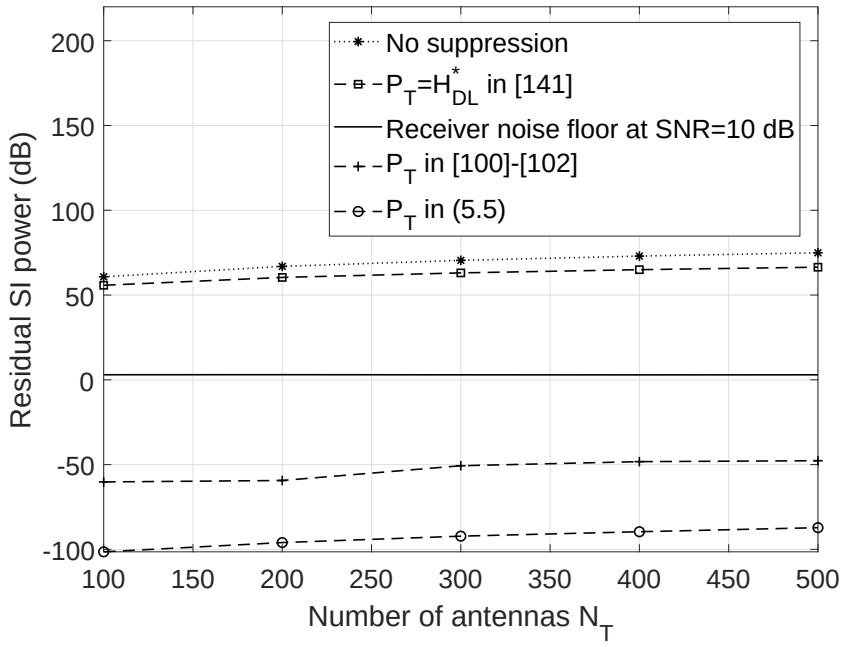


Figure 5.3: Residual SI power versus number of transmit antennas for a FD massive MIMO BS.

in (5.5). The receiver noise floor at the assumed SNR is also shown. We can see that the standard massive MIMO transmit beamforming, which ignores SI suppression, performs the worst. The schemes based on the right inverse and the proposed scheme can successfully reduce the residual SI power below the receiver noise floor. However, the precoder given by (5.5) has better performance. This is because for square matrix $\mathbf{H}$, (5.5) is the optimum solution to (5.3).

5.3.3 BER Performance

Fig. 5.4 plots the average bit error rate (BER) in the uplink and downlink versus SNR (dB) assuming 4-quadrature amplitude modulation (4−QAM). The results for UL with no SI and DL with no MUI are plotted as a benchmark. We can see that the BER results with the proposed precoder design in (5.5) overlap the benchmark results.

Fig. 5.4 shows that the BER in the UL is better than the DL. This is explained

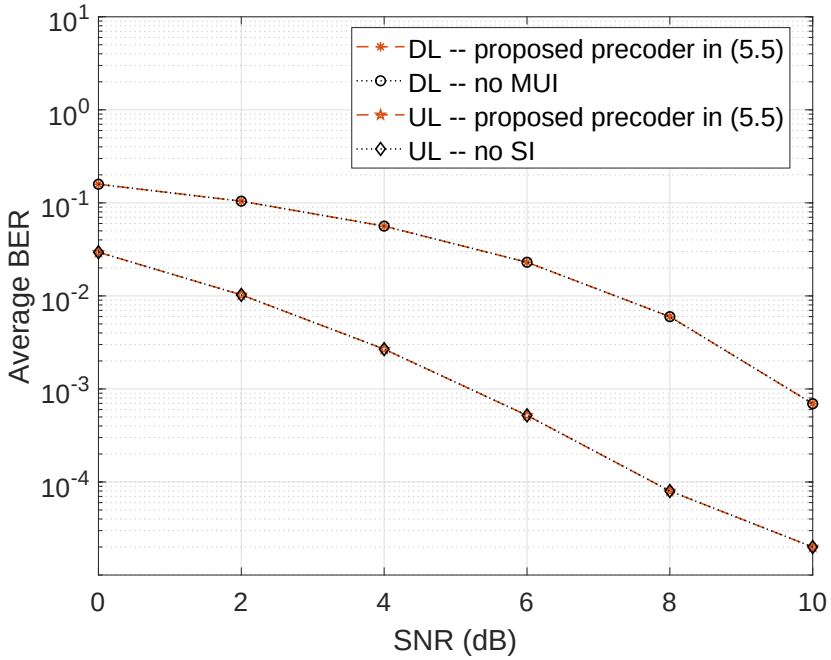


Figure 5.4: Average BER performance of a FD massive MIMO BS.

as follows. In the DL, following [46], we do not employ any receive filtering to keep the processing at the users at minimum. However, in the uplink, the receiver beamforming matrix $\mathbf{P}_{\mathrm{R}}$ in (5.1a) is set to $\frac{1}{N_{\mathrm{R}}}\mathbf{H}_{\mathrm{UL}}^{*}$. This allows us to take advantage of the spatial diversity in the UL, as stated in [46]. This explains the better performance in the UL.

Furthermore, although in this work we only consider one type of linear precoder design (zero-forcing (ZF) precoder) based on the well-cited work of [102], other linear precoders such a regularised inverse channel (RCI) precoders [145, 146] can also be considered. However, further performance improvement is not expected. This is because as shown in Fig. 5.4 the ZF based precoder allows for complete suppression of SI and MUI.

5.4 Conclusions

In this chapter, we developed a mathematical foundation which allows systematic design of precoding matrix at the BS to suppress SI and MUI in a single-cell,

multi-user, FD massive MIMO system. We rigorously derived the condition on the number of transmit and receive antennas. As a main finding, we proved that the simultaneous suppression of SI and MUI has a solution with probability 1. Our results validate previous heuristic assumptions made in the literature.

Chapter 6

Conclusions and Future Research Directions

In this chapter, we provide a summary of general conclusions that can be drawn from this thesis. In particular, we focus on the open research problems highlighted in Section 1.3.1. We also outline some future research directions arising from this work.

6.1 Conclusions

The open research questions were discussed in Section 1.3.1. The major contribution of the thesis is addressing those research questions.

Addressing Q1 in Section 2.3:

- We proposed a low complexity EM based estimator to estimate both SI and communication channels using shifted modulation. The proposed estimation techniques does not require data-aided piloting for channel estimation.

Addressing Q2 in Section 2.2.1:

- We thoroughly investigated the inherent ambiguity of blind channel estimation for estimating SI and communication channels in FD communication systems. We also presented Theorem 2.1 to design modulation constellations for ambiguity-free channel estimation in FD communication.

- Based on the result of Theorem 2.1, we designed a modulation constellation, which allows for estimating both SI and communication channels without sending data-aided pilots.

Addressing Q3 in Section 2.3.1:

- We derived a closed-form lower bound for the estimation error for estimating the SI and communication channels in a FD communication system.

- In Section 2.4, we showed that the MSE performance of the proposed EM-based estimator reaches the derived lower bound.

Addressing Q4 in Section 3.2:

- We first mathematically formulated the problem in Section 3.1.2. The rigorous problem formulation allowed us to proposed an EKF based estimator in Section 3.2 to jointly estimate the channel and track the fast varying PN process.

Addressing Q5 in Sections 3.1 and 3.3 :

- We derived the lower bound for the estimation error of the EKF-based estimator in Section 3.2.2. Proposition 3.1 showed that the error bound is reciprocal to the covariance matrix of PN random process.

- In Section 3.3 we showed that the MSE performance of the proposed EKF-based estimator reaches the performance of the bound at high SNR.

Addressing Q6 in Section 4.2:

- We answered this question through the results of Propositions 4.1 and 4.2. In Proposition 4.1, we first proposed a MAP-based detector to detect the desired symbols blindly, and then in Proposition 4.2, we showed how to remove the ambiguity of blind detection by designing asymmetric modulation constellations.

Addressing Q7 in Section 5.2:

- We presented Proposition 5.1 on the number of transmit and receive antennas for a FD-enabled massive MIMO BS.
- We presented Theorem 5.1 , which stated that for a FD-enabled massive MIMO BS with the number of transmit antennas greater than or equal to the sum of the number of receive antennas and the number of UL UEs, there exits a precoder matrix with probability 1 to suppress both SI and MUI.

In summary the major contributions of the thesis are as follows:

- A novel bandwidth efficient channel estimator for FD communication at microwave frequency and the derivation of the corresponding lower bound.
- An extended Kalman filter based approach to jointly estimate the channel and the PN process at mmWave frequency and the derivation of the theoretical lower bound for the estimation error of the PN.
- A MAP based blind detector to detect data using statistical properties of communication and SI channel without having the explicit channel state information.
- A criteria for simultaneous MUI and SI suppression for FD massive MIMO BS.

6.2 Future Research Directions

A number of future research directions arise from the work presented in this thesis.

- The channel estimation techniques considered in this thesis are applicable to frequency flat fading channels. A possible extension of these techniques is to consider frequency-selective channels.
- The analysis of inherent ambiguity problem for channel estimation in Chapter 2 was carried out for classical estimator, when the statistical property of the channels are unknown. An interesting extension to this work is to

consider Bayesian estimator such as minimum mean square error (MMSE) estimator and investigate the effect of different statistical properties on the ambiguity problem of Bayesian estimators.

- The joint channel and PN estimator in Chapter 3 can also be extended by applying the design principle outlined in this thesis to MIMO FD communication systems, where some of the transmit antennas share a single oscillator.
- The result of Proposition 5.1 imposes a condition on the number of transmit and receive antennas for a FD-enabled massive MIMO BS. An interesting research question that arises here is the application of antenna selection techniques to these systems.
- This thesis does not consider the impact of mutual coupling [147] or hardware impairments. This can be considered in future work.

Appendix A

This appendix contains proofs used in Chapter 2.

A.1 Proof of Theorem 2.1

The proof consists of three main steps.

Step 1: We show that if $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ exists such that $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}') = f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta})\ \forall \mathbf{y}_a$, then a bijective function $\mathrm{g} : \mathcal{K} \to \mathcal{K}$ exists, such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$, where $c \neq 1$ is a constant and $|c| = 1$. This is done as follows.

If the two joint probability densities $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}')$ and $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta})$ are equal $\forall \mathbf{y}_a$, then it easily follows that the marginal densities $f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta}')$ and $f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta})$ are also equal $\forall y_{a_i}$. From (2.4), $f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta}')$ and $f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta})$ are given by

$$f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta}) = \left(\frac{1}{M\pi\sigma^2}\right)\sum_{k=1}^{M}\exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \boldsymbol{\theta}(1)x_{a_i} - \boldsymbol{\theta}(2)x_k|^2\right), \tag{A.1a}$$

$$f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta}') = \left(\frac{1}{M\pi\sigma^2}\right)\sum_{k=1}^{M}\exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \boldsymbol{\theta}'(1)x_{a_i} - \boldsymbol{\theta}'(2)x_k|^2\right). \tag{A.1b}$$

If $f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta}')$ and $f_{Y_{a_i}}(y_{a_i};\boldsymbol{\theta})$ are equal $\forall y_{a_i}$, they should also be equal for $y_{a_i} = \boldsymbol{\theta}(1)x_{a_i} + \boldsymbol{\theta}(2)x_1$. In this case, we have

$$\sum_{k=1}^{M}\exp\left(\frac{-1}{\sigma^2}|\boldsymbol{\theta}(2)(x_1 - x_k)|^2\right) =$$

$$\sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2}\left|(\boldsymbol{\theta}(1)-\boldsymbol{\theta}'(1))x_{a_i}+\boldsymbol{\theta}(2)x_1-\boldsymbol{\theta}'(2)x_k\right|^2\right). \tag{A.2}$$

The left hand side (LHS) of (A.2) is independent of i, while the right hand side (RHS) of (A.2) depends on i through x_{a_i}. Consequently, for (A.2) to hold for $\forall y_{a_i}$, the coefficient of x_{a_i} should be zero, i.e., $\boldsymbol{\theta}'(1) = \boldsymbol{\theta}(1)$. Knowing that $\theta'(1) = \theta(1)$ and equating (A.1a) and (A.1b), we have

$$\sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2}|y_{a_i}-\boldsymbol{\theta}(1)x_{a_i}-\boldsymbol{\theta}(2)x_k|^2\right) = \sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2}|y_{a_i}-\boldsymbol{\theta}(1)x_{a_i}-\boldsymbol{\theta}'(2)x_k|^2\right). \tag{A.3}$$

By taking the first and second order derivatives of both sides of (A.3) with respect to y_{a_i}, it can be shown that $\forall k \in \mathcal{K}$, the points $y_{a_i} = \boldsymbol{\theta}(1)x_{a_i} + \boldsymbol{\theta}(2)x_k$ and $y_{a_i} = \boldsymbol{\theta}(1)x_{a_i} + \boldsymbol{\theta}'(2)x_k$ maximize the summations of the M exponential functions on the LHS and RHS of (A.3), respectively. Consequently, since (A.3) holds $\forall y_{a_i}$, the points that maximize the summation of M exponential on the LHS of (A.3) are the same as the points that maximize the summation of M exponentials on the RHS of (A.3). Hence, for a bijective function $\mathrm{g} : \mathcal{K} \to \mathcal{K}$

$$\boldsymbol{\theta}(1)x_{a_i} + \boldsymbol{\theta}(2)x_k = \boldsymbol{\theta}(1)x_{a_i} + \boldsymbol{\theta}'(2)x_{\mathrm{g}(k)} \tag{A.4}$$

It can easily be verified that if (A.4) holds $\forall y_{a_i}$, then,

$$\boldsymbol{\theta}(2)x_k = \boldsymbol{\theta}'(2)x_{\mathrm{g}(k)}, \tag{A.5}$$

or

$$\frac{\boldsymbol{\theta}'(2)}{\boldsymbol{\theta}(2)} = \frac{x_k}{x_{\mathrm{g}(k)}}. \tag{A.6}$$

The LHS of (A.6) does not depend on k, consequently, the RHS of (A.6) should also be independent of k and should be a constant. Hence, for bijective function g, $\frac{x_k}{x_{\mathrm{g}(k)}} = c$, where $c \neq 1$ is a constant. We note that if $c = 1$ then $\theta(2) = \theta'(2)$

and hence $\boldsymbol{\theta} = \boldsymbol{\theta}'$, which violates the assumption that $f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta}') = f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta})$ for $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$.

Let us now define permutation Π on the ordered set $\mathcal{A} = \{x_1, \cdots, x_M\}$ as

$$\Pi \triangleq \begin{pmatrix} x_1 & x_2 & \cdots & x_M \\ \Pi(x_1) = x_{\mathrm{g(1)}} & \Pi(x_2) = x_{\mathrm{g(2)}} & \cdots & \Pi(x_M) = x_{\mathrm{g(M)}} \end{pmatrix}. \tag{A.7}$$

The sequence $(x_k, \Pi(x_k), \Pi(\Pi(x_k)), \cdots, x_k)$ forms an orbit of the permutation Π [111]. If $x_k = c x_{\mathrm{g(k)}}\ \forall k \in \mathcal{K}$, then from the definition of the orbit it is clear that $x_k = c^m x_k\ \forall k \in \mathcal{K}$, where m is the length of the orbit sequence and $c \neq 1$ is a constant. Since, $x_k = c^m x_k\ \forall k \in \mathcal{K}$, we can conclude that $c^m = 1$ and $|c| = 1$.

Step 2: We show that if the bijective function g: $\mathcal{K} \to \mathcal{K}$ exists, such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$, then there exists a $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ for which $f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta}') = f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta})\ \forall \mathbf{y}_a$. This is done as follows.

Firstly, in Section 2.2.2 we showed that the joint PDF of all the observations is given by

$$f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta}) = \left(\frac{1}{M\pi\sigma^2}\right)^N \prod_{i=1}^{N} \sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2} |y_{a_i} - \boldsymbol{\theta}(1) x_{a_i} - \boldsymbol{\theta}(2) x_k|^2\right), \tag{A.8}$$

We can see that for a fixed i in (A.8), M different exponential functions corresponding to different values of $x_k \in \mathcal{A}$ are summed together. Since, g : $\mathcal{K} \to \mathcal{K}$ is a bijective function and consequently replacing x_k by $x_{\mathrm{g}(x)}$ only affects the order of the exponential functions, we can rewrite (A.8) as

$$f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta}) = \left(\frac{1}{M\pi\sigma^2}\right)^N \prod_{i=1}^{N} \sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2} |y_{a_i} - \boldsymbol{\theta}(1) x_{a_i} - \boldsymbol{\theta}(2) x_{\mathrm{g}(k)}|^2\right), \tag{A.9}$$

Secondly, for $\boldsymbol{\theta}' \triangleq [\boldsymbol{\theta}(1), \frac{\boldsymbol{\theta}(2)}{c}]$, $f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta}')$ is given by

$$f_{\mathbf{Y_a}}(\mathbf{y_a}; \boldsymbol{\theta}') = \left(\frac{1}{M\pi\sigma^2}\right)^N \prod_{i=1}^{N} \sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2} |y_{a_i} - \boldsymbol{\theta}(1) x_{a_i} - \frac{\boldsymbol{\theta}(2)}{c} x_k|^2\right), \tag{A.10}$$

and since $\frac{x_k}{c} = x_{\mathrm{g}(k)}\ \forall k \in \mathcal{K}$, (A.10) can be written as

$$f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}') = \left(\frac{1}{M\pi\sigma^2}\right)^N \prod_{i=1}^{N}\sum_{k=1}^{M} \exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \boldsymbol{\theta}(1)x_{a_i} - \boldsymbol{\theta}(2)x_{\mathrm{g}(k)}|^2\right), \quad \text{(A.11)}$$

Comparing (A.11) with (A.9) reveals that for $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$, $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}') = f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta})$ $\forall \mathbf{y}_a$. Consequently, if bijective function g: $\mathcal{K} \to \mathcal{K}$ exits, such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$, then there exists a $\boldsymbol{\theta}' \neq \boldsymbol{\theta}$ for which $f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta}') = f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\theta})\ \forall \mathbf{y}_a$.

Step 3: Thirdly, we show that the condition $\frac{x_k}{x_{\mathrm{g}(k)}} = c\ \forall k \in \mathcal{K}$ is equivalent to the modulation constellation being symmetric around the origin. To prove this equivalency, we need to consider the following two sub-cases:

(i) Firstly, we need to show that if a bijective function g : $\mathcal{K} \to \mathcal{K}$ exists such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c$, then the modulation constellation is symmetric with respect to origin. Equivalently, we can show that if a bijective function g : $\mathcal{K} \to \mathcal{K}$ does not exist such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c$, then the modulation constellation is not symmetric with respect to origin. To prove this equivalent statement, we use proof by contradiction. We assume g : $\mathcal{K} \to \mathcal{K}$ does not exist such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c$, but the modulation constellation is symmetric with respect to origin. If the modulation is symmetric with respect to the origin then it satisfies the condition of Definition 2.2 and hence,

$$\mathrm{f}(-x_k) = -\mathrm{f}(x_k). \quad \text{(A.12)}$$

However, since the function $\mathrm{f}(x_k)$ is defined on set $\mathcal{A}$, (A.12) holds if and only if both x_k and $-x_k$ are in the set $\mathcal{A}$. Consequently, set $\mathcal{A}$ can be represented by

$$\mathcal{A} = \{x_1, x_2, \cdots, x_{\frac{M}{2}}, -x_1, -x_2, \cdots, -x_{\frac{M}{2}}\}. \quad \text{(A.13)}$$

Now if the bijective function g is defined as $\mathrm{g}(k) = k + \frac{M}{2}$, then $\frac{x_k}{x_{\mathrm{g}(k)}} = -1$. However, this contradicts the assumption that bijective function g : $\mathcal{K} \to \mathcal{K}$ does not exist, such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c \forall k \in \mathcal{K}$. Hence, if g : $\mathcal{K} \to \mathcal{K}$ does not exist such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c$, then the modulation constellation cannot be symmetric

with respect to origin.

(ii) Secondly, we need to show that if the modulation is symmetric then a bijective function $\mathrm{g} : \mathcal{K} \to \mathcal{K}$ exists such that $\frac{x_k}{x_{\mathrm{g}(k)}} = c \; \forall k \in \mathcal{K}$. This easily follows from the proof of previous step, where we showed that if the modulation constellation is symmetric then $\mathcal{A}$ can be represented by (A.13). Consequently, a bijective function $\mathrm{g} : \mathcal{K} \to \mathcal{K}$ exists such that $\frac{x_k}{x_{\mathrm{g}(k)}} = -1 \; \forall k \in \mathcal{K}$, i.e., $\mathrm{g}(k) = k + \frac{M}{2} \; \forall k \in \mathcal{K}$.

Combining the proofs of the three steps, Theorem 2.1 is proved.

A.2 Proof of Propositions 2.2 and 2.3

Propositions 2.2 and 2.3 correspond to the E and M steps of the EM algorithm. We assume that both transmitters at nodes a and b use the asymmetric shifted modulation constellation $\overline{\mathcal{A}}$ defined in Definition 2.4, i.e., $\overline{x}_{a_i}, \overline{x}_{b_i} \in \overline{\mathcal{A}}$, and assume a uniform discrete distribution for the transmitted symbols.

A.2.1 Proof of E-Step

In the E-step of the algorithm function $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$ is given by

$$Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) = \mathbb{E}_{\overline{\mathbf{X}}_\mathbf{b}|\mathbf{y_a},\boldsymbol{\phi}^{(n)}}[\ln f_{\mathbf{Y_a}}(\mathbf{y_a}, \overline{\mathbf{x}}_\mathbf{b}|\boldsymbol{\phi})]. \tag{A.14}$$

To calculate (A.14), we require $\ln f_{\mathbf{Y_a}}(\mathbf{y}_a, \overline{\mathbf{x}}_\mathbf{b}|\boldsymbol{\phi})$. Hence, we start with the following joint PDF

$$\begin{aligned} f_{Y_{a_i}}(y_{a_i}, \overline{x}_{b_i}; \boldsymbol{\phi}) &= f_{Y_{a_i}}(y_{a_i}|\overline{X}_{b_i} = \overline{x}_{b_i}; \boldsymbol{\phi}) p_{\overline{X}_{b_i}}(\overline{x}_{b_i}) \\ &= \frac{1}{M\pi\sigma^2} \sum_{k=1}^{M} \delta_{\overline{x}_k, \overline{x}_{b_i}} \exp\left(\frac{-1}{\sigma^2}|y_{a_i} - h_{ba}\overline{x}_k - h_{aa}\overline{x}_{a_i}|^2\right), \end{aligned} \tag{A.15}$$

where, $\delta_{\overline{x}_k, \overline{x}_{b_i}}$ is the Kronecker delta function and $\delta_{\overline{x}_k, \overline{x}_{b_i}} = 1$ if $\overline{x}_{b_i} = \overline{x}_k$ and 0

otherwise [110]. Consequently, the log-likelihood of all the observations is given by

$$\begin{aligned}\ln(f_{\mathbf{Y_a}}(\mathbf{y_a},\mathbf{x_b};\boldsymbol{\phi})) &= \sum_{i=1}^{N}\ln(f_{Y_{a_i}}(y_{a_i},\overline{x}_{b_i};\boldsymbol{\phi}))\\ &= -N\ln(M\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{i=1}^{N}\sum_{k=1}^{M}\delta_{\overline{x}_k,\overline{x}_{b_i}}|y_{a_i} - h_{ba}\overline{x}_{b_i} - h_{aa}\overline{x}_{a_i}|^2. \end{aligned} \tag{A.16}$$

The expectation in (A.14) is conditioned on knowing $\boldsymbol{\phi}^{(n)}$ during the nth iteration of the algorithm, which is obtained from the M-step. Substituting (A.16) in (A.14), we have

$$\begin{aligned}Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) &= \mathbb{E}_{\overline{\mathbf{x}}_\mathbf{b}|\mathbf{y_a},\boldsymbol{\phi}^{(n)}}[\ln f_{\mathbf{Y_a}}(\mathbf{y_a},\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})] = -N\ln(M\pi\sigma^2)\\ &- \frac{1}{\sigma^2}\mathbb{E}_{\overline{\mathbf{x}}_\mathbf{b}|\mathbf{y_a},\boldsymbol{\phi}^{(n)}}\left[\sum_{i=1}^{N}\sum_{k=1}^{M}\delta_{\overline{x}_k,\overline{x}_{b_i}}|y_{a_i} - h_{ba}\overline{x}_{b_i} - h_{aa}\overline{x}_{a_i}|^2\right]. \end{aligned} \tag{A.17}$$

The assumption of independent transmitted symbols allows to rewrite (A.17) as follows

$$\begin{aligned}Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) &= -N\ln(M\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{i=1}^{N}\sum_{k=1}^{M}\mathbb{E}_{\overline{x}_{b_i}|\mathbf{y_a},\boldsymbol{\phi}^{(n)}}\left[\delta_{\overline{x}_k,\overline{x}_{b_i}}|y_{a_i} - h_{ba}\overline{x}_{b_i} - h_{aa}\overline{x}_{a_i}|^2\right],\\ &= -N\ln(M\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{i=1}^{N}\sum_{k=1}^{M}P(\overline{x}_{b_i} = \overline{x}_k|\mathbf{y}_a,\boldsymbol{\phi}^{(n)})|y_{a_i} - h_{ba}\overline{x}_k - h_{aa}\overline{x}_{a_i}|^2. \end{aligned} \tag{A.18}$$

We define

$$T_{k,i}^{(n)} \triangleq P(\overline{x}_{b_i} = \overline{x}_k|\mathbf{y}_a,\boldsymbol{\phi}^{(n)}). \tag{A.19}$$

Then it can easily be shown that

$$T_{k,i}^{(n)} = \frac{\exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \widehat{h}_{ba}^{(n)}\overline{x}_k - \widehat{h}_{aa}^{(n)}\overline{x}_{a_i}|^2\right)}{\sum_{\overline{k}=1}^{M}\exp\left(\frac{-1}{\sigma^2}|y_{a_i} - \widehat{h}_{ba}^{(n)}\overline{x}_{\overline{k}} - \widehat{h}_{aa}^{(n)}\overline{x}_{a_i}|^2\right)}. \tag{A.20}$$

Finally, substituting (A.20) into (A.18), $Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)})$ can be found as in (A.14). This concludes the proof of Proposition 2.2.

A.2.2 Proof of M-step

The maximization-step of the EM algorithm is given by

$$\begin{aligned}\boldsymbol{\phi}^{(n+1)} &= \arg\max_{\boldsymbol{\phi}} Q(\boldsymbol{\phi}|\boldsymbol{\phi}^{(n)}) \\ &= \arg\min_{\boldsymbol{\phi}} \sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)} |y_{a_i} - h_{ba}\overline{x}_k - h_{aa}\overline{x}_{a_i}|^2. \end{aligned} \tag{A.21}$$

We define the following function

$$\mathrm{r}(\boldsymbol{\phi}) \triangleq \sum_{i=1}^{N}\sum_{k=1}^{M} T_{k,i}^{(n)} |y_{a_i} - h_{ba}\overline{x}_k - h_{aa}\overline{x}_{a_i}|^2, \tag{A.22}$$

The minimum of function $\boldsymbol{\phi}$ (the maximum of the likelihood function), which corresponds to the solution of the $M-$ step of the EM algorithm during the nth iteration, happens at the critical point $\boldsymbol{\phi}^{(n+1)}$ for which the Jacobian is zero, i.e., $\mathbf{J} = 0$ [148]. To find this critical point the Jacobian matrix should be constructed and set equal to zero. This is done by taking the derivative of $\mathrm{r}(\boldsymbol{\phi})$ with respect to the four elements of vector $\boldsymbol{\phi}$, as defined by (2.9), to construct the Jacobian matrix and then set it equal to zero. Then, it can be easily shown that the critical point $\boldsymbol{\phi}^{(n+1)}$ is given by

$$\boldsymbol{\phi}^{(n+1)} = \mathbf{S}^{-1}\boldsymbol{v}, \tag{A.23}$$

where

$$\mathbf{S} \triangleq \begin{bmatrix} s_1 & 0 & s_2 & s_3 \\ 0 & s_1 & -s_3 & s_2 \\ s_2 & -s_3 & s_4 & 0 \\ s_3 & s_2 & 0 & s_4 \end{bmatrix}, \quad \boldsymbol{v} \triangleq \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}, \tag{A.24}$$

where the elements of $\mathbf{S}$ and $\boldsymbol{v}$ are given by (2.14b)-(2.14d).

However, to ensure that the critical point $\boldsymbol{\phi}^{(n+1)}$ is the minimum of function $\mathrm{r}(\boldsymbol{\phi})$, the Hessian matrix $\mathbf{H}$ should be positive semi-definite [148]. By taking the second derivatives of $\mathrm{r}(\boldsymbol{\phi})$ with respect to the four elements of vector $\boldsymbol{\phi}$, we can show that $\mathbf{H} = 2\mathbf{S}$. Then, according to Sylvester's criterion [148], $\mathbf{H}$ is positive semi-definite if and only if all the following are positive

$$s_1, \quad \det\left(\begin{bmatrix} s_1 & 0 \\ 0 & s_1 \end{bmatrix}\right), \quad \det\left(\begin{bmatrix} s_1 & 0 & s_2 \\ 0 & s_1 & -s_3 \\ s_2 & -s_3 & s_4 \end{bmatrix}\right), \quad \det(\mathbf{S}). \tag{A.25}$$

It can easily be shown that $\det(\mathbf{S}) = (s_1 s_4 - s_2^2 - s_3^2)^2$, and is always positive. According to (2.14b), s_1 is always positive, and it is clear that the second determinant is always positive. However, the positivity of the third determinant, i.e., $s_1 s_4 - s_2^2 - s_3^2$, directly depends on the initialization. This is evident from definitions in (2.14b)-(2.14d), which link s_1, s_2, s_3 and s_4 to the function $T_{k,i}^{(n)}$ and the derivation of function $T_{k,i}^{(n)}$ in (A.20), which is a function of $\widehat{h}_{aa}^{(n)}$ and $\widehat{h}_{ba}^{(n)}$, i.e., the estimates from the nth iteration. Our numerical investigation shows that for the initialization vector $\boldsymbol{\phi}^{(0)} \triangleq [0, 0, 0, 0]$, the Hessian matrix $\mathbf{H}$ is always positive and hence the critical point $\boldsymbol{\phi}^{(n+1)}$ is indeed the minimum of function $\mathrm{r}(\boldsymbol{\phi})$. Consequently, the EM algorithm with initialization vector $\boldsymbol{\phi}^{(0)} \triangleq [0, 0, 0, 0]$ converges to the maximum of the likelihood function.

This concludes the proof of Proposition 2.3.

A.3 Proof of Proposition 2.4

It is shown in [149] that for any random variables X and Y and any parameter θ, if the probability distribution of X is independent of the parameter θ, then $\mathbf{I}[f_Y(y;\theta)] < \mathbf{I}[f_Y(y|x;\theta)]$, where $\mathbf{I}[\cdot]$ is the FIM and $f(\cdot;\theta)$ is the probability density function parameterized by θ. Consequently, the performance of the proposed estimator is lower bounded by the inverse of $\mathbf{I}[f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})]$, i.e.,

$$\mathbb{E}_{\widehat{\Phi}_l}[|\widehat{\phi}_l - \phi_l|^2] \geq \left[\mathbf{I}^{-1}\left[f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})\right]\right]_{l,l} \quad \forall l \in \{1,2,3,4\}, \tag{A.26}$$

where ϕ_l is the lth element of the parameter vector $\boldsymbol{\phi}$, $\widehat{\phi_l}$ is an estimate of ϕ_l, and $[\cdot]_{l,l}$ is the lth diagonal element of matrix. This is because (i) the performance of the proposed estimator is lower bounded by $\mathbf{I}[f_{\mathbf{Y_a}}(\mathbf{y_a};\boldsymbol{\phi})]$ according to (2.15), and (ii) $p_{\overline{\mathbf{X}}_\mathbf{b}}(\overline{\mathbf{x}}_\mathbf{b})$ is independent of $\boldsymbol{\phi}$. Furthermore, since (A.26) holds $\forall \overline{\mathbf{x}}_\mathbf{a}, \overline{\mathbf{x}}_\mathbf{b}$, then the variance is also lower-bounded by

$$\mathbb{E}_{\widehat{\Phi}_l}[|\widehat{\phi_l} - \phi_l|^2] \geq \left[\mathbf{I}_{avg}^{-1}\left[f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})\right]\right]_{l,l} \quad \forall l \in \{1,2,3,4\}, \tag{A.27}$$

where $\mathbf{I}_{avg} = \mathbb{E}_{\overline{\mathbf{X}}_\mathbf{b},\overline{\mathbf{X}}_\mathbf{a}}\left[\mathbf{I}[f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})]\right]$. The value of $\mathbf{I}[f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})]$, needed to evaluate $\mathbf{I}_{avg}$, is presented in the Lemma below.

Lemma A.3.1 $\mathbf{I}[f(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})]$ *is a* 4×4 *matrix with its elements given by*

$$i_{1,1} = i_{2,2} = \frac{2}{\sigma^2}\sum_{i=1}^{N}|\overline{x}_{a_i}|^2, \qquad i_{3,3} = i_{4,4} = \frac{2}{\sigma^2}\sum_{i=1}^{N}|\overline{x}_{b_i}|^2, \tag{A.28a}$$

$$i_{2,4} = i_{4,2} = i_{1,3} = i_{3,1} = \frac{2}{\sigma^2}\sum_{i=1}^{N}\left(\Re\{\overline{x}_{a_i}\}\Re\{\overline{x}_{b_i}\} + \Im\{\overline{x}_{a_i}\}\Im\{\overline{x}_{b_i}\}\right), \tag{A.28b}$$

$$i_{2,3} = i_{3,2} = -i_{1,4} = -i_{4,1} = \frac{2}{\sigma^2}\sum_{i=1}^{N}\left(\Re\{\overline{x}_{b_i}\}\Im\{\overline{x}_{a_i}\} - \Re\{\overline{x}_{a_i}\}\Im\{\overline{x}_{b_i}\}\right). \tag{A.28c}$$

Proof: It can be easily seen from (2.5) that $f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};h_{aa},h_{ba})$ is given by

$$f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};h_{aa},h_{ba}) = \left(\frac{1}{\pi\sigma^2}\right)^N \exp\sum_{i=1}^{N}\left(\frac{-|y_{a_i} - h_{aa}\overline{x}_{a_i} - h_{ba}\overline{x}_{b_i}|^2}{\sigma^2}\right). \tag{A.29}$$

Then, for $l,l' \in \{1,2,3,4\}$, $\mathbf{I}[f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};h_{aa},h_{ba})]$ is [116]

$$\mathbf{I}_{l,l'} = -\mathbb{E}_{\mathbf{Y_a}}\left[\frac{\partial^2}{\partial_{\phi_m}\partial_{\phi_n}}\ln f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};h_{aa},h_{ba})\right], \tag{A.30}$$

where $m,n \in \{1,2,3,4\}$, ϕ_m and ϕ_n are the mth and nth elements of $\boldsymbol{\phi} = [\Re\{h_{aa}\},\Im\{h_{aa}\},\Re\{h_{ba}\},\Im\{h_{ba}\}]$. By evaluating (A.30), using the joint PDF given by (A.29), the non-zero elements of $\mathbf{I}[f_{\mathbf{Y_a}}(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};h_{aa},h_{ba})]$ can be found and are given by (A.28b)-(A.28c).

Using the value of $\mathbf{I}[f(\mathbf{y_a}|\overline{\mathbf{x}}_\mathbf{b};\boldsymbol{\phi})]$ given by the above lemma, we need to evaluate the expectations in order to find $\mathbf{I}_{avg}$. As discussed in Section 2.2.3, we assume

that both nodes a and b use a real constant $s \triangleq \sqrt{\beta E}$ to shift the modulation constellation. Since all the constellation points are equally likely to be transmitted, before shifting the modulation constellation we have $\mathbb{E}_{\Im\{\overline{X}_{a_i}\}}[\Im\{x_{a_i}\}] = \mathbb{E}_{\Im\{\overline{x}_{b_i}\}}[\Im\{x_{b_i}\}] = \mathbb{E}_{\Re\{\overline{X}_{a_i}\}}[\Re\{x_{a_i}\}] = \mathbb{E}_{\Re\{\overline{X}_{b_i}\}}[\Re\{x_{b_i}\}] = 0$. However, after the shift, $\mathbb{E}_{\Re\{\overline{X}_{a_i}\}}[\Re\{\overline{x}_{a_i}\}] = \mathbb{E}_{\Re\{\overline{X}_{b_i}\}}[\Re\{\overline{x}_{b_i}\}] = \sqrt{\beta E}$ and $\mathbb{E}_{\Im\{\overline{X}_{a_i}\}}[\Im\{\overline{x}_{a_i}\}] = \mathbb{E}_{\Im\{\overline{X}_{b_i}\}}[\Im\{\overline{x}_{b_i}\}] = 0$. Furthermore, $\mathbb{E}_{X_{a_i}}[|\overline{x}_{a_i}|^2] = \mathbb{E}_{X_{b_i}}[|\overline{x}_{b_i}|^2] = E + \beta E$ since the average energy of the constellation after the shift is increased by the shift energy ($|s|^2 = \beta E$). Consequently, the average FIM with respect to $\overline{\mathbf{x}}_\mathbf{a}$ and $\overline{\mathbf{x}}_\mathbf{b}$ is given by

$$\mathbf{I}_{avg} = \frac{2NE}{\sigma^2} \begin{pmatrix} 1+\beta & 0 & \beta & 0 \\ 0 & 1+\beta & 0 & \beta \\ \beta & 0 & 1+\beta & 0 \\ 0 & \beta & 0 & 1+\beta \end{pmatrix}, \tag{A.31}$$

and $\mathbf{I}_{avg}^{-1}$ is given by

$$\mathbf{I}_{avg}^{-1} = \frac{\sigma^2}{2NE} \begin{pmatrix} \frac{\beta+1}{(2\beta+1)} & 0 & -\frac{\beta}{(2\beta+1)} & 0 \\ 0 & \frac{\beta+1}{(2\beta+1)} & 0 & -\frac{\beta}{(2\beta+1)} \\ -\frac{\beta}{(2\beta+1)} & 0 & \frac{\beta+1}{(2\beta+1)} & 0 \\ 0 & -\frac{\beta}{(2\beta+1)} & 0 & \frac{\beta+1}{(2\beta+1)} \end{pmatrix}. \tag{A.32}$$

Using (A.27) and considering the diagonal elements of (A.32), we arrive at the result in (2.16).

Appendix B

This appendix contains proofs and derivation of Unscented Kalman Filter (UKF) used in Chapter 3.

B.1 Proof of Proposition 3.2

In this section we provide the complexity analysis of the EKF algorithm by counting the number of multiplications and additions. However, before we proceed it can easily be shown that every entry of product of a $K \times L$ matrix by a $L \times M$ matrix requires L multiplications and $L-1$ additions, and hence, the whole matrix requires KML multiplications and $KM(L-1)$ additions, where KM is the size of the resulting matrix. Furthermore, it is known that matrix inversion has the same complexity in terms of additions and multiplication as the matrix multiplication, up to a multiplicative constant γ [150]. We can now proceed with calculating the complexity of EKF algorithm in (B.1) to (B.4).

$$\widehat{\boldsymbol{\beta}}(n|n) = \underbrace{\widehat{\boldsymbol{\beta}}(n|n-1) \underbrace{\Re\left\{\mathbf{K}(n)\left(\mathbf{y}(n) - \underbrace{\overline{\mathbf{H}}(n)e^{j\widehat{\boldsymbol{\beta}}(n|n-1)}}_{N_r(2N_tN_r)+N_r(2N_tN_r-1)}\right)\right\}}_{N_r+N_r(2N_tN_r)+N_r(2N_tN_r-1)}}_{2N_tN_r^2+2N_tN_r(N_r-1)+N_r+N_r(2N_tN_r)+N_r(2N_tN_r-1)}, \tag{B.1}$$
$$2N_tN_r + 2N_tN_r^2 \quad +2N_tN_r(N_r-1) + N_r + N_r(2N_tN_r) + N_r(2N_tN_r-1)$$

$$\mathbf{K}(n) = \mathbf{M}(n|n-1)\mathbf{D}^{\dagger}(n) \underbrace{\left(\sigma^2\mathbf{I}_{N_r} + \mathbf{D}(n)\mathbf{M}(n|n-1)\mathbf{D}^{\dagger}(n)\right)^{-1}}, \tag{B.2}$$

$$\begin{aligned} &4N_t^2N_r^3 + 2N_tN_r^2(2N_tN_r-1) + 2N_tN_r^3 + 2N_tN_r^2(N_r-1) \\ &+ \gamma N_r^3 + \gamma N_r^2(N_r-1) + N_r^2 + 2N_tN_r^2 + N_r(2N_tN_r-1) \\ &+ 4N_t^2N_r^3 + 2N_tN_r^2(2N_tN_r-1) \end{aligned}$$

$$\mathbf{M}(n|n-1) = \underbrace{\mathbf{M}(n-1|n-1) + \mathbf{Q}}_{4N_t^2N_r^2}, \tag{B.3}$$

$$\mathbf{M}(n|n) = \Re\left\{ \underbrace{\mathbf{M}(n|n-1)\underbrace{\left(\mathbf{I}_{2N_tN_r} - \underbrace{\mathbf{K}(n)\mathbf{D}(n)}_{2N_tN_r^3+2N_tN_r^2(N_r-1)}\right)}_{4N_t^2N_r^2+2N_tN_r^3+2N_tN_r^2(N_r-1)}}_{8N_t^3N_r^3+4N_t^2N_r^2(2N_tN_r-1)+2N_tN_r^3+2N_tN_r^2(N_r-1)} \right\}, \tag{B.4}$$

B.2 Proof of Proposition 3.1

In this section we derive the lower bound of the estimation error. We start the proof by expanding $\mathbb{E}\left[\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)^T\right]$.

$$\begin{aligned} &\mathbb{E}\left[\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)^T\right] = \\ &\mathbb{E}\left[\boldsymbol{\beta}(n)\boldsymbol{\beta}^T(n)\right] + \mathbb{E}\left[\widehat{\boldsymbol{\beta}}(n)\widehat{\boldsymbol{\beta}}^T(n)\right] - \mathbb{E}\left[\boldsymbol{\beta}(n)\widehat{\boldsymbol{\beta}}^T(n)\right] - \mathbb{E}\left[\widehat{\boldsymbol{\beta}}(n)\boldsymbol{\beta}^T(n)\right] \end{aligned} \tag{B.5}$$

Next we show that the last two terms of (B.5) are zero. We do this by showing only $\mathbb{E}\left[\boldsymbol{\beta}(n)\widehat{\boldsymbol{\beta}}^T(n)\right] = 0$ as a similar approach can be used to show that $\mathbb{E}\left[\widehat{\boldsymbol{\beta}}(n)\boldsymbol{\beta}^T(n)\right] = 0$.

We first note that $\boldsymbol{\beta}(n)$ given by (3.8) is a Gaussian autoregressive model (AR)

with mean zero, i.e., $\mathbb{E}\left[\boldsymbol{\beta}(n)\right] = 0$. Hence,

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{\beta}(n)\widehat{\boldsymbol{\beta}}^{T}(n)\right] &= \int\int \boldsymbol{\beta}(n)\widehat{\boldsymbol{\beta}}(\boldsymbol{n})\; p(\boldsymbol{\beta}(n), \mathbf{y}(n))\; d\boldsymbol{\beta}(n)\; d\mathbf{y}(n) \\ &= \int \widehat{\boldsymbol{\beta}}(\boldsymbol{n}) \int \boldsymbol{\beta}(n)\; p(\boldsymbol{\beta}(n))\; d\boldsymbol{\beta}(n)\; p(\mathbf{y}(n)|\boldsymbol{\beta}(n))\; d\mathbf{y}(n) \\ &= \int \widehat{\boldsymbol{\beta}}(\boldsymbol{n})\mathbb{E}\left[\boldsymbol{\beta}(n)\right]\; p(\mathbf{y}(n)|\boldsymbol{\beta}(n))\; d\mathbf{y}(n) = 0. \end{aligned} \tag{B.6}$$

Consequently, we can rewrite (B.5) as follow

$$\mathbb{E}\left[\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)^{T}\right] = \mathbb{E}\left[\boldsymbol{\beta}(n)\boldsymbol{\beta}^{T}(n)\right] + \mathbb{E}\left[\widehat{\boldsymbol{\beta}}(n)\widehat{\boldsymbol{\beta}}^{T}(n)\right] \tag{B.7}$$

It is easy to show that $\mathbb{E}\left[\widehat{\boldsymbol{\beta}}(n)\widehat{\boldsymbol{\beta}}^{T}(n)\right]$ is a positive semi-definite matrix and hence

$$\mathbb{E}\left[\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)^{T}\right] \geq \mathbb{E}\left[\boldsymbol{\beta}(n)\boldsymbol{\beta}^{T}(n)\right] \tag{B.8}$$

Furthermore, the properties of $\mathrm{Tr}(\cdot)$ allows us to write

$$\mathrm{Tr}\left(\mathbb{E}\left[\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)\left(\boldsymbol{\beta}(n) - \widehat{\boldsymbol{\beta}}(n)\right)^{T}\right]\right) \geq \mathrm{Tr}\left(\mathbb{E}\left[\boldsymbol{\beta}(n)\boldsymbol{\beta}^{T}(n)\right]\right) \tag{B.9}$$

Finally, using (B.9) and the definitions of $\boldsymbol{Q}$ and MSE in (3.9) and (3.18), we can establish the proof of the proposition.

B.3 Derivation of Unscented Kalman Filter (UKF)

Unscented Kalman Filter (UKF) provides an alternative to EKF for non-linear state vector estimation. In UKF instead of linearizing the observation vector, the probability distributions of states and observations are approximated using *sigma points* [151]. UKF can solve a very general class of problems, where both state process and observations are nonlinear. However, the joint channel and PN estimation problem, as given by the observation vector (3.7) and the state vector (3.8), has a linear state process and additive noise. This allows for the use of non-augmented state vectors for UKF [152]. For the state vector $\boldsymbol{\beta}(n)$ in (3.8), the sigma points

$\mathcal{B}(i,n)$ are given by

$$\mathcal{B}(0,n|n-1) = \mathcal{B}(0,n-1), \tag{B.10a}$$

$$\mathcal{B}_(i,n|n-1) = \mathcal{B}(i,n-1) + \left(\sqrt{\{(L+\lambda)\,\mathbf{Q}\}}\right)_i, \quad i = 1,\cdots,L, \tag{B.10b}$$

$$\mathcal{B}(i,n|n-1) = \mathcal{B}(i,n-1) - \left(\sqrt{\{(L+\lambda)\,\mathbf{Q}\}}\right)_i, \quad i = L+1,\cdots,2L \tag{B.10c}$$

where, $\sqrt{\{\cdot\}}$ is the matrix square root, $(\cdot)_i$ is the ith column of the matrix, $L = 4N_tN_r$, $\lambda = \alpha^2 L - L$, where $\alpha = 10^{-3}$ [151] , and $\mathbf{Q}$ is the state covariance matrix given by (3.9). Subsequently, the mean of the sigma points, which is used as an approximate to the true mean of the probability distribution of states, is given by

$$\overline{\boldsymbol{\beta}}(n) = \sum_{i=0}^{2l} \mathcal{W}_i^m \mathcal{B}(i,n|n-1), \tag{B.11}$$

where,

$$\mathcal{W}_0^m = \frac{\lambda}{L+\lambda} \tag{B.12}$$

$$\mathcal{W}_i^m = \frac{1}{2(L+\lambda)}, \quad i = 1,\cdots,2L. \tag{B.13}$$

Similarly, the covariance of the state vector based on the sigma points approximation is given by

$$\overline{\mathbf{P}}_n = \sum_{i=0}^{2L} \mathcal{W}_i^m \left[\mathcal{B}_(i,n|n-1) - \overline{\boldsymbol{\beta}}(n)\right]\left[\mathcal{B}_(i,n|n-1) - \overline{\boldsymbol{\beta}}(n)\right]^* \tag{B.14}$$

Moreover, the sigma points for the observations, and the corresponding approximate mean of probability distribution of observations are given by

$$\mathcal{Y}(n|n-1) = \overline{\mathbf{H}}(n) e^{j\mathcal{B}(i,n|n-1)}, \tag{B.15a}$$

$$\overline{\mathbf{y}}(n) = \sum_{i=0}^{2l} \mathcal{W}_i^c \mathcal{Y}(n|n-1), \tag{B.15b}$$

where, $\mathcal{W}_0^c = \mathcal{W}_0^m + (1-\alpha^2+\beta)$, $\beta = 2$, and $\mathcal{W}_i^c = \mathcal{W}_i^m$ for $i = 1,\cdots,2L$. Once the

Algorithm 1 UKF for joint channel and PN estimation

1: **Initialize:**
   $\widehat{\boldsymbol{\beta}}(0) = \mathbf{0}_{2N_tN_r \times 1}$
   $\mathbf{P}_0 = \mathbf{Q}$
2: **for** n = 0 to N **do**
3: Calculate sigma points using (B.10a) to (B.10c).
4: Calculate mean of state $\overline{\boldsymbol{\beta}}(n)$ using (B.11).
5: Calculate the covariance matrix of the state vector using (B.14).
6: Calculate the sigma points for observations using (B.15a).
7: Calculate the mean of observation using (B.15b).
8: Update the mean $\widehat{\boldsymbol{\beta}}(n)$ using (B.16).
9: Update the variance $\widehat{\mathbf{P}}_n$ using (B.17).

state and the process models are approximated by the sigma points using (B.10a)-(B.10c), and (B.15a), respectively, the updated mean $\widehat{\boldsymbol{\beta}}(n)$ and variance $\widehat{\mathbf{P}}_n$ can be calculated as follow

$$\widehat{\boldsymbol{\beta}}(n) = \overline{\boldsymbol{\beta}}(n) + \mathcal{K}\left(\mathbf{y}(n) - \overline{\mathbf{y}}(n)\right), \tag{B.16}$$

$$\widehat{\mathbf{P}}_n = \overline{\mathbf{P}}_n - \mathcal{K}\mathbf{P}_{y,y}\mathcal{K}^T, \tag{B.17}$$

where,

$$\mathcal{K} = \mathbf{P}_{x,y}\mathbf{P}_{y,y}^{-1}, \tag{B.18}$$

$$\mathbf{P}_{x,y} = \sum_{i=0}^{2L} \mathcal{W}_i^c \left[\mathcal{B}_(i, n|n-1) - \overline{\boldsymbol{\beta}}(n)\right] \left[\mathcal{Y}(n|n-1) - \overline{\mathbf{y}}(n)\right]^*, \tag{B.19}$$

$$\mathbf{P}_{y,y} = \sum_{i=0}^{2L} \mathcal{W}_i^c \left[\mathcal{Y}(n|n-1) - \overline{\mathbf{y}}(n)\right] \left[\mathcal{Y}(n|n-1) - \overline{\mathbf{y}}(n)\right]^*. \tag{B.20}$$

Algorithm 1 summarizes the UKF joint channel and PN estimation algorithm.

Appendix C

This appendix contains proofs used in Chapter 4.

C.1 Proof of Proposition 4.1

We start the proof by deriving the conditional density function $f(\mathbf{y}_a|\mathbf{x}_b)$ as follows

$$\begin{aligned} f(\mathbf{y}_a|\mathbf{x}_b) &= \int_{h_{aa}} \int_{h_{ba}} f(\mathbf{y}_a|\mathbf{x}_b, h_{aa}, h_{ba}) f(h_{aa}, h_{ba}) \quad dh_{aa} dh_{ba}, \\ &= \int_{h_{aa}} \int_{h_{ba}} f(\mathbf{y}_a|\mathbf{x}_b, h_{aa}, h_{ba}) f(h_{aa}) f(h_{ba}) \quad dh_{aa} dh_{ba}. \end{aligned} \tag{C.1}$$

where in (C.1),

$$f(\mathbf{y}_a|\mathbf{x}_b, h_{aa}, h_{ba}) = \frac{1}{(\pi\sigma^2)^N} \prod_{i=1}^{N} \exp\left(-\frac{|y_{a_i} - h_{aa} x_{a_i} - h_{ba} x_{b_i}|^2}{\sigma^2} \right), \tag{C.2}$$

$$f(h_{aa}) = \frac{1}{\pi} \exp\left(-|h_{aa}|^2 \right), \tag{C.3}$$

$$f(h_{ba}) = \frac{1}{\pi} \exp\left(-|h_{ba}|^2 \right). \tag{C.4}$$

Rewriting (C.1), we arrive at

$$\begin{aligned} f(\mathbf{y}_a|\mathbf{x}_b) &= \frac{1}{(\pi\sigma^2)^N \pi^2} \int_{h_{ba}} \exp\left(-|h_{ba}|^2 \right) \\ &\times \int_{h_{aa}} \exp\left(-\sum_{i=1}^{N} \frac{|y_{a_i} - h_{aa} x_{a_i} - h_{ba} x_{b_i}|^2}{\sigma^2} \right) \exp\left(-|h_{aa}|^2 \right) \quad dh_{aa} dh_{ba}. \end{aligned} \tag{C.5}$$

Note that in performing the integration in (C.5), we can use the fact that the total probability of a complex Gaussian random variable is one.

Using the Bayes' rule

$$f(\mathbf{x}_b|\mathbf{y}_a) = \frac{f(\mathbf{y}_a|\mathbf{x}_b)f(\mathbf{x}_b)}{f(\mathbf{y}_a)}, \tag{C.6}$$

where $f(\mathbf{x}_b) = \left(\frac{1}{M}\right)^N$ since the transmitted symbols come from a equiprobable modulation set, i.e., $f(x_{b_i}) = \frac{1}{M}$ and $f(\mathbf{y}_a|\mathbf{x}_b)$ is given in (C.5). Substituting and simplifying, we can obtain the result in (4.3).

C.2 Proof of Proposition 4.2

To prove Proposition 4.2, we first define permutation $\Pi(\cdot)$ as a one-to-one and onto function on the index set of modulation set $\mathcal{A}$, i.e., $\mathcal{K} \triangleq \{1, 2, \cdots, M\}$. If $x_k \in \mathcal{A}$, then, $x_{\Pi(k)} \in \mathcal{A}'$, $\forall k \in \mathcal{K}$, where $\mathcal{A}'$ is one possible permutation of original modulation set $\mathcal{A}$. Without loss of generality, we further assume that both $\mathcal{A}$ and $\mathcal{A}'$ are ordered set and A_k and A'_k are the kth elements of $\mathcal{A}$ and $\mathcal{A}'$, respectively.

For simplicity of analysis, we show the proof for constant power M-PSK modulation sets in here. The extension to QAM modulation is straightforward and omitted here [104].

Lemma C.2.1 *For the ith transmitted symbol, the posterior function $f(x_{b_i}|\mathbf{y}_a)$ does not have a unique maximum if and only if for the modulation set $\mathcal{A}$ there exists a permuted set $\mathcal{A}'$ for which $\frac{x_k}{x_{\Pi(k)}} = -1$, $\forall k \in \mathcal{K}$.*

Proof: For the first part of the proof we assume that the permutation $\pi(\cdot)$ that satisfies the condition of the lemma exists and then for a permuted set $\mathcal{A}'$, for which $\frac{x_k}{x_{\Pi(k)}} = -1$, $\forall k \in \mathcal{K}$, we assume that $x'_{b_i} = x_{\Pi(k)} = A'_{\Pi(k)}$ and $x_{b_i} = A_k$ maximizes the posterior density. Then $f(x_{b_i}|\mathbf{y}_a)$ can be rewritten as

$$f(x_{b_i} = A_k|\mathbf{y}_a) \propto \underbrace{\sum_{j_N=1}^{M} \cdots \sum_{j_1=1}^{M}}_{\sim j_i} \frac{1}{\sum_{n=1, n\neq i}^{N} |A_{j_n}|^2 - \vartheta + |A_k|^2 + \sigma^2}$$

$$\times \exp\left(\frac{|\xi|^2}{\left(\sum_{n=1,n\neq i}^{N}|A_{j_n}|^2-\vartheta+|A_k|^2+\sigma^2\right)\sigma^2}\right), \tag{C.7}$$

where,

$$\vartheta \triangleq \frac{|A_k|^2}{\gamma}\left|\sum_{n=1,n\neq i}^{N} x_{a_n}^{*}\frac{A_{j_n}}{A_k}+x_{a_i}^{*}\right|^2. \tag{C.8}$$

It is easy to see from (C.7) that the posterior PDF of $x'_{b_i}=A'_{\Pi}(k)$ differs from the posterior PDF of x_{b_i} only in the term ϑ as the rest of the terms depend on the power of the modulation constellation, which is constant for M-PSK modulation set. We define

$$\vartheta' \triangleq \frac{|A'_{\Pi(k)}|^2}{\gamma}\left|\sum_{n=1,n\neq i}^{N} x_{a_n}^{*}\frac{A'_{\Pi(j_n)}}{A'_{\Pi(k)}}+x_{a_i}^{*}\right|^2. \tag{C.9}$$

To prove the lemma we need to show that

$$f(x_{b_i}=A_k|\mathbf{y}_a)=f(x'_{b_i}=A'_{\Pi(k)}|\mathbf{y}_a), \tag{C.10}$$

and hence no unique maximum. (C.10) holds true if and only if

$$\vartheta=\vartheta'. \tag{C.11}$$

Finally, (C.11) holds true if and only if

$$\frac{A_{j_n}}{A_k}=\frac{A'_{\Pi(j_n)}}{A'_{\Pi(k)}}. \tag{C.12}$$

Since we know that $\frac{x_k}{x'_{\Pi(k)}}=\frac{A_k}{A'_{\Pi(k)}}=-1\ \forall k$, consequently, (C.12) is valid, which in turn means (C.10) holds true and the posterior function does not have a unique maximum. For the second part of the proof, it is clear that when no permutation exists to satisfy the condition of the lemma then ϑ can never be equal to ϑ' and consequently, the lemma holds if and only if such a permutation exists.

It is easy to see that the condition of Lemma C.2.1 is met if and only if the

modulation constellation is symmetric around the origin. This is because with symmetric modulations around the origin if $x_k \in \mathcal{A}$ so is $-x_k \in \mathcal{A}$. Consequently, there always exists a permutation for which the condition of Lemma C.2.1 holds. Therefore, the posterior function does not have a unique maximum if and only if the modulation constellation is symmetric around the origin.

This concludes the proof of the proposition.

Appendix D

This appendix contains proofs used in Chapter 5.

D.1 Proof of Proposition 5.1

The proof uses tools from linear algebra, in particular, the rank-null theorem which relates the dimensions of a linear map's kernel and image with the dimension of its domain [143].

Forward reasoning: Assuming (5.3) has a solution, then (5.2b) holds. Since (5.2b) holds, then the following must hold [143],

$$\dim \mathrm{Im}\left(\mathbf{P}_{\mathrm{T}}\right) \leq \dim \ker\left(\mathbf{H}_{\mathrm{SI}}\right). \tag{D.1}$$

This is because if $\mathrm{Im}\left(\mathbf{P}_{\mathrm{T}}\right)$ is not a subspace of $\ker\left(\mathbf{H}_{\mathrm{SI}}\right)$, then there can be a $\mathbf{x}_{\mathrm{DL}} \in \mathrm{Im}\left(\mathbf{P}_{\mathrm{T}}\right)$ such that $\mathbf{P}_{\mathrm{T}}\mathbf{x}_{\mathrm{DL}} \notin \ker\left(\mathbf{H}_{\mathrm{SI}}\right)$, and hence $\mathbf{H}_{\mathrm{SI}}\mathbf{P}_{\mathrm{T}}\mathbf{x}_{\mathrm{DL}} \neq 0$, i.e., (5.2b) does not hold.

Next we use the rank-null theorem and apply it to matrices $\mathbf{P}_{\mathrm{T}}$ and $\mathbf{H}_{\mathrm{SI}}$, to further simplify (D.1). Applying the rank-null theorem to $\mathbf{P}_{\mathrm{T}}$, we have $\dim \mathrm{Im}\left(\mathbf{P}_{\mathrm{T}}\right) + \dim \ker\left(\mathbf{P}_{\mathrm{T}}\right) = U$. Now $\dim \ker\left(\mathbf{P}_{\mathrm{T}} = 0\right)$. This is because, if it is not zero, then the desired signal term in (5.1b) can be zero for some $\mathbf{x}_{\mathrm{DL}} \in \ker\left(\mathbf{P}_{\mathrm{T}}\right)$. Thus, $\dim \mathrm{Im}\left(\mathbf{P}_{\mathrm{T}}\right) = U$. Applying the rank-null theorem to $\mathbf{H}_{\mathrm{SI}}$, we have $\dim \mathrm{Im}\left(\mathbf{H}_{\mathrm{SI}}\right) + \dim \ker\left(\mathbf{H}_{\mathrm{SI}}\right) = N_{\mathrm{T}}$. Now $\dim \mathrm{Im}\left(\mathbf{H}_{\mathrm{SI}}\right) = N_{\mathrm{R}}$. This is because $\mathbf{H}_{\mathrm{SI}}$ is a linear transformation from transmit antenna array to the receive antenna array. Thus, $N_{\mathrm{R}} + \dim \ker\left(\mathbf{H}_{\mathrm{SI}}\right) = N_{\mathrm{T}}$ or $\dim \ker\left(\mathbf{H}_{\mathrm{SI}}\right) = N_{\mathrm{T}} - N_{\mathrm{R}}$. Substituting values in (D.1) we have $U \leq N_{\mathrm{T}} - N_{\mathrm{R}} \implies N_{\mathrm{T}} \geq N_{\mathrm{R}} + U$.

Proof by backwards reasoning: Assuming $N_{\mathrm{T}} \geq N_{\mathrm{R}} + U$, we can easily use the rank-null theorem to show that $\dim \ker(\mathbf{H}_{\mathrm{SI}}) \geq U$. This means the image space of $\mathbf{P}_{\mathrm{T}}$ is a subspace of the $\ker(\mathbf{H}_{\mathrm{SI}})$. This implies that (5.2b) holds, and subsequently, (5.3) has a solution.

D.2 Proof of Theorem 5.1

In this appendix we show that $\mathbf{H}\mathbf{H}^*$ is full-rank with probability 1. The key result used in the proof is the following [153]: For a general random matrix $\mathbf{G}$, $\mathbf{G}\mathbf{G}^*$ is full-rank w.p. 1, if $\forall \mathbf{g}_i$, every hyperplane has probability 0, where $\mathbf{g}_i$ is the ith column of the random matrix $\mathbf{G}$. We apply this to our system model in Section 5.1.

Step 1: From (5.3) and the system model channel assumptions, we can see that the columns of $\mathbf{H}$ are independent and identically distributed (i.i.d). We denote the ith column of $\mathbf{H}$, by $\mathbf{h}_i$, which is a complex vector. Since the real and imaginary parts of $\mathbf{h}_i$ are independent, the complex vector $\mathbf{h}_i$ has the same distribution as the real vector $\mathbf{h}_i^{\mathbb{R}}$ with independent elements, such that the first $2U$ elements of $\mathbf{h}_i^{\mathbb{R}}$ are drawn from $\mathcal{N}(0,1)$, and the remaining elements are drawn from $\mathcal{N}(\mu, \sigma^2)$.

Step 2: Based on the distribution of the elements of the real vector $\mathbb{R}^{2N}$, we can say that the joint probability distribution of $\mathbf{h}_i^{\mathbb{R}}$ is a Gaussian probability measure on $\mathbb{R}^{2N}$, where $\mathbb{R}$ is the set of real numbers. We denote this Gaussian probability measure by $\mathcal{N}_{2N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean and covariance matrix of the vector $\mathbf{h}_i^{\mathbb{R}}$, respectively. Consequently, $\mathbf{H}\mathbf{H}^*$ is full-rank if given the probability Gaussian measure $\mathcal{N}_{2N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the probability of every hyperplane is 0. The Gaussian measure $\mathcal{N}_{2N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is absolutely continuous (a.c.) with respect to the Lebesgue measure λ_{2N} on $\big(\mathbb{R}^{2N}, \mathscr{B}(\mathbb{R}^{2N})\big)$, if and only if, $\boldsymbol{\Sigma}$ is positive definite, where $\mathscr{B}(\mathbb{R}^{2N})$ defines Borel $\sigma-$algbera on $\mathbb{R}^{2N}$ [154]. From Step 1 and knowing that the elements of $\mathbf{h}_i^{\mathbb{R}}$ are independent, it can be concluded that $\det(\boldsymbol{\Sigma}) = \sigma^{4(N-U)}$. Hence, $\boldsymbol{\Sigma}$ is positive definite and the Gaussian measure given by $\mathcal{N}_{2N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is (a.c.) with respect to the Lebesgue measure λ_{2N}. This means that for every measurable set $\mathscr{A}$, if $\lambda_{2N}(\mathscr{A}) = 0$, then probability of $\mathscr{A}$ is also zero, i.e., $P(\mathscr{A}) = 0$. Consequently, to prove that the probability of every hyperplane is zero for $\mathbf{h}_i^{\mathbb{R}}$, we can show that the Lebesgue measure of every hyperplane is zero, i.e., $\lambda_{2N}(\text{every flat hyperplane}) = 0$.

Step 3: In order to show $\lambda_{2N}(\text{every flat hyperplane}) = 0$, we proceed as follows. Mathematically, hyperplane is defined as [155], $\mathscr{A} \triangleq \{\mathbf{h}_i^{\mathbb{R}} \in \mathbb{R}^{2N} | \mathbf{a}^T \mathbf{h}_i^{\mathbb{R}} = \overline{\mathbf{b}}\}$, where $\mathbf{a} \in \mathbb{R}^{2N}$ and $\overline{\mathbf{b}} \in \mathbb{R}$. We also define, $\overline{\mathscr{A}} \triangleq \{\mathbf{h}_i^{\mathbb{R}} \in \mathbb{R}^{2N} | \mathbf{a}^T \mathbf{h}_i^{\mathbb{R}} = 0\}$, which is a shifted version of $\mathscr{A}$. Since, a Lebesgue measure is invariant under shift [154], we can conclude that $\lambda_{2N}(\mathscr{A}) = \lambda_{2N}(\overline{\mathscr{A}})$. Furthermore, we note that $\mathscr{E} \triangleq \{\mathbf{h}_i^{\mathbb{R}} \in \mathbb{R}^{2N} | h_{i,2N}^{\mathbb{R}} = 0\}$, where $h_{i,2N}^{\mathbb{R}}$ is the $2N$th element of vector $\mathbf{h}_i^{\mathbb{R}}$, is a subspace of $\mathbb{R}^{2N}$, and there exists a linear transformation $\mathbf{T}$ such that $\overline{\mathscr{A}} = \mathbf{T}(\mathscr{E})$ [154]. Then the following relationship holds [154, Theorem 12.1], $\lambda_{2N}(\overline{\mathscr{A}}) = |\det(\mathbf{T})| \lambda_{2N}(\mathscr{E})$. Consequently, we can show $\lambda_{2N}(\mathscr{A}) = 0$, if we can show that $\lambda_{2N}(\mathscr{E}) = 0$.

Step 4: In order to show $\lambda_{2N}(\mathscr{E}) = 0$, we proceed as follows. According to the definition of Lebesgue measure given in [154], we can define the box, $\mathbf{B} \triangleq \prod_{j=1}^{2N} [a_j, b_j]$, where $\prod[\cdot]$ is the Cartesian product, $a_j = h_{i,j}^{\mathbb{R}} - 1/2$, $b_j = h_{i,j}^{\mathbb{R}} + 1/2$, $\forall j = \{1, \cdots, 2N-1\}$, $a_{2N} = b_{2N} = 0$, and $h_{i,j}^{\mathbb{R}}$ is the jth element of $\mathbf{h}_i^{\mathbb{R}} \in \mathbb{R}^{2N}$. It is clear that $\forall \mathbf{h}_i^{\mathbb{R}} \in \mathscr{E}$, there exists no element of $\mathbf{h}_i^{\mathbb{R}}$ that lies outside the box. Hence, $\mathbf{B}$ covers $\mathscr{E}$, and the following holds [154], $\lambda_{2N}(\mathscr{E}) = \lambda_{2N}(\mathbf{B}) = \prod_{j=1}^{2N}(b_j - a_j) = 0$, where $\prod(\cdot)$ is the arithmetic product. Consequently,

$$\lambda_{2N}(\mathscr{E}) = 0 \implies \lambda_{2N}(\mathscr{A}) = 0 \overset{\text{a.c.}}{\implies} P(\mathscr{A}) = 0. \tag{D.2}$$

Step 5: (D.2) shows that the probability of hyperplane is zero for a single column of $\mathbf{H}$. Since the columns of $\mathbf{H}$ are i.i.d. this holds for all the other columns as well. Consequently, we can conclude that $\mathbf{H}\mathbf{H}^*$ is full-rank.

Bibliography

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be ?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[2] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[3] D. Wang, D. Chen, B. Song, N. Guizani, X. Yu, and X. Du, "From IoT to 5G I-IoT: The next generation IoT-based intelligent algorithms and 5g technologies," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 114–120, Oct. 2018.

[4] S. A. A. Shah, E. Ahmed, M. Imran, and S. Zeadally, "5G for vehicular communications," *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 111–117, Jan. 2018.

[5] N. Kshetri, "5G in E-commerce activities," *IT Professional*, vol. 20, no. 4, pp. 73–77, Jul. 2018.

[6] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[7] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, May 2014.

[8] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5G backhaul challenges and emerging research directions: A survey," *IEEE Access*, vol. 4, pp. 1743–1766, 2016.

[9] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," vol. 30, no. 1, pp. 40–60, Jan. 2013.

[10] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. D. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[11] C. Wang, F. Haider, X. Gao, X. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[12] B. Bangerter, S. Talwar, R. Arefi, and K. Stewart, "Networks and devices for the 5G era," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 90–96, Feb. 2014.

[13] E. Dahlman, G. Mildh, S. Parkvall, J. Peisa, J. Sachs, Y. Selen, and J. Skold, "5G wireless access: requirements and realization," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 42–47, Dec. 2014.

[14] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.

[15] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[16] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. K. Ramakrishnan, C. W. Rice, and N. K. Shankaranarayanan, "Design and characterization of a full-duplex multiantenna system for WiFi networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1160–1177, Mar. 2014.

[17] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2017–2046, Fourthquarter 2015.

[18] G. Liu, F. R. Yu, H. Ji, V. C. M. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 500–524, Secondquarter 2015.

[19] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y. S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: pushing the limits," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 80–87, Sep. 2016.

[20] T. Gansler and G. Salomonsson, "Nonintrusive measurements of the telephone channel," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 158–167, Jan. 1999.

[21] M. Ho, J. M. Cioffi, and J. A. C. Bingham, "Discrete multitone echo cancelation," *IEEE Trans. Commun.*, vol. 44, no. 7, pp. 817–825, Jul. 1996.

[22] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 128–137, May 2015.

[23] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[24] Z. He, S. Shao, Y. Shen, C. Qing, and Y. Tang, "Performance analysis of RF self-interference cancellation in full-duplex wireless communications," *IEEE Wireless Commun. Lett.*, vol. 3, no. 4, pp. 405–408, Aug. 2014.

[25] S. Wang, Y. Liu, W. Zhang, and H. Zhang, "Achievable rates of full-duplex massive MIMO relay systems over rician fading channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 9825–9837, Nov. 2017.

[26] D. Korpi, L. Anttila, V. Syrjala, and M. Valkama, "Widely-linear digital self-interference cancellation in direct-conversion full duplex transceiver," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.

[27] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[28] B. Kaufman, J. Lilleberg, and B. Aazhang, "An analog baseband approach for designing full-duplex radios," in *Proc. Asilomar Conf. on Signals, Syst. and Computers*, Nov. 2013.

[29] D. Korpi, T. Riihonen, V. Syrjala, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.

[30] E. Ahmed and A. M. Eltawil, "All-digital self-interference cancellation technique for full-duplex systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3519–3532, Jul. 2015.

[31] A. Nadh, J. Samuel, A. Sharma, S. Aniruddhan, and R. K. Ganti, "A linearization technique for self-interference cancellation in full-duplex radios," *arXiv Technical Report*, 2017. [Online]. Available: https://arxiv.org/abs/1605.01345

[32] D. van den Broek, E. A. M. Klumperink, and B. Nauta, "An in-band full-duplex radio receiver with a passive vector modulator downmixer for self-interference cancellation," *IEEE J. Solid-State Circuits*, vol. 50, no. 12, pp. 3003–3014, Dec. 2015.

[33] B. Debaillie, D. van den Broek, C. Lavin, B. van Liempd, E. A. M. Klumperink, C. Palacios, J. Craninckx, B. Nauta, and A. Parssinen, "Analog/RF solutions enabling compact full-duplex radios," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1662–1673, Sep. 2014.

[34] T. S. Rappaport, G. R. MacCartney, S. Sun, H. Yan, and S. Deng, "Small-scale, local area, and transitional millimeter wave propagation for 5g communications," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6474–6490, Dec. 2017.

[35] T. K. Vu, M. Bennis, M. Debbah, M. Latva-aho, and C. S. Hong, "Ultra-reliable communication in 5G mmwave networks: A risk-sensitive approach," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 708–711, Apr. 2018.

[36] A. Ali, N. Gonzalez-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, Feb. 2018.

[37] A. W. Mbugua, W. Fan, Y. Ji, and G. F. Pedersen, "Millimeter wave multi-user performance evaluation based on measured channels with virtual antenna array channel sounder," *IEEE Access*, vol. 6, pp. 12 318–12 326, 2018.

[38] S. Han, C. I, Z. Xu, and S. Wang, "Reference signals design for hybrid analog and digital beamforming," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1191–1193, Jul. 2014.

[39] D. Zhu, B. Li, and P. Liang, "A novel hybrid beamforming algorithm with unified analog beamforming by subspace construction based on partial CSI for massive MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 594–607, Feb. 2017.

[40] S. Zhang, C. Guo, T. Wang, and W. Zhang, "ONOFF analog beamforming for massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4113–4123, May 2018.

[41] A. A. Nasir, S. Durrani, H. Mehrpouyan, S. D. Blostein, and R. A. Kennedy, "Timing and carrier synchronization in wireless communication systems: a survey and classification of research in the last 5 years," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, Aug. 2016. [Online]. Available: https://doi.org/10.1186/s13638-016-0670-9

[42] H. Mehrpouyan, M. R. Khanzadi, M. Matthaiou, A. M. Sayeed, R. Schober, and Y. Hua, "Improving bandwidth efficiency in E-band communication systems," *IEEE Commun. Mag.*, vol. 52, no. 3, pp. 121–128, Mar. 2014.

[43] H. Mehrpouyan, A. A. Nasir, S. D. Blostein, T. Eriksson, G. K. Karagiannidis, and T. Svensson, "Joint estimation of channel and oscillator phase

noise in MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.

[44] L. Zhao and W. Namgoong, "A novel phase-noise compensation scheme for communication receivers," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 532–542, Mar. 2006.

[45] A. A. Nasir, H. Mehrpouyan, R. Schober, and Y. Hua, "Phase noise in MIMO systems: Bayesian CramerRao bounds and soft-input estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2675–2692, May 2013.

[46] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[47] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[48] B. Wang, Y. Chang, and D. Yang, "On the SINR in massive MIMO networks with MMSE receivers," *IEEE Commun. Lett.*, vol. 18, no. 11, pp. 1979–1982, Nov 2014.

[49] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[50] A. Khansefid and H. Minn, "On channel estimation for massive MIMO with pilot contamination," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1660–1663, Sep. 2015.

[51] J. K. Tugnait, "Self-contamination for detection of pilot contamination attack in multiple antenna systems," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 525–528, Oct. 2015.

[52] X. Zhu, L. Dai, and Z. Wang, "Graph coloring based pilot allocation to mitigate pilot contamination for multi-cell massive MIMO systems," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1842–1845, Oct. 2015.

[53] J. Chen, "A low complexity data detection algorithm for uplink multiuser massive MIMO systems," vol. 35, no. 8, pp. 1701–1714, Aug. 2017.

[54] L. Fang, L. Xu, and D. D. Huang, "Low complexity iterative MMSE-PIC detection for medium-size massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 108–111, Feb. 2016.

[55] D. Kim, H. Ju, S. Park, and D. Hong, "Effects of channel estimation error on full-duplex two-way networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4666–4672, Nov. 2013.

[56] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator for self-interference cancelation in full-duplex systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5122–5132, Jul. 2016.

[57] X. Li, C. Tepedelenlioglu, and H. Senol, "Channel estimation for residual self-interference in full-duplex amplify-and-forward two-way relays," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 4970–4983, Aug. 2017.

[58] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[59] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.

[60] A. A. Nasir, D. T. Ngo, H. D. Tuan, and S. Durrani, "Iterative optimization for max-min sinr in dense small-cell multiuser miso swipt system," in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec 2015, pp. 1392–1396.

[61] A. A. Nasir, H. Mehrpouyan, S. Durrani, S. D. Blostein, R. A. Kennedy, and B. Ottersten, "Transceiver design for distributed STBC based AF cooperative networks in the presence of timing and frequency offsets," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3143–3158, Jun. 2013.

[62] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Mixture Kalman filtering for joint carrier recovery and channel estimation in time-selective Rayleigh fading channels," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 3496–3499.

[63] W. Li, J. Lilleberg, and K. Rikkinen, "On rate region analysis of half- and full-duplex OFDM communication links," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1688–1698, Sep. 2014.

[64] W. Ding, Y. Niu, H. Wu, Y. Li, and Z. Zhong, "Qos-aware full-duplex concurrent scheduling for millimeter wave wireless backhaul networks," *IEEE Access*, vol. 6, pp. 25 313–25 322, 2018.

[65] A. H. M. R. Islam, M. Bakaul, A. Nirmalathas, and G. E. Town, "Simplified generation, transport, and data recovery of millimeter-wave signal in a full-duplex bidirectional Fiber-Wireless system," *IEEE Photon. Technol. Lett.*, vol. 24, no. 16, pp. 1428–1430, Aug 2012.

[66] E. Ahmed and A. M. Eltawil, "On phase noise suppression in full-duplex systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1237–1251, Mar. 2015.

[67] S. Li and R. D. Murch, "Full-duplex wireless communication using transmitter output based echo cancellation," in *Proc. IEEE GLOBECOM*, Dec. 2011.

[68] M. Sohaib, H. Nawaz, K. Ozsoy, O. Gurbuz, and I. Tekin, "A low complexity full-duplex radio implementation with a single antenna," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2206–2218, Mar. 2018.

[69] F. Shu, J. Wang, J. Li, R. Chen, and W. Chen, "Pilot optimization, channel estimation, and optimal detection for full-duplex OFDM systems with IQ imbalances," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6993–7009, Aug. 2017.

[70] A. Masmoudi and T. Le-Ngoc, "Channel estimation and self-interference cancelation in full-duplex communication systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 321–334, Jan. 2017.

[71] F. Mazzenga, "Channel estimation and equalization for M-QAM transmission with a hidden pilot sequence," *IEEE Trans. Broadcast.*, vol. 46, no. 2, pp. 170–176, Jun. 2000.

[72] J. K. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Commun. Lett.*, vol. 7, no. 9, pp. 413–415, Sep. 2003.

[73] E. de Carvalho and D. T. M. Slock, "Blind and semi-blind FIR multichannel estimation: (global) identifiability conditions," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1053–1064, Apr. 2004.

[74] S. Abdallah and I. N. Psaromiligkos, "EM-based semi-blind channel estimation in amplify-and-forward two-way relay networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 527–530, Oct. 2013.

[75] ——, "Blind channel estimation for amplify-and-forward two-way relay networks employing M-PSK modulation," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3604–3615, Jul. 2012.

[76] S. Zhang, S.-C. Liew, and H. Wang, "Blind known interference cancellation," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1572–1582, Aug. 2013.

[77] Y. Zhu, D. Guo, and M. L. Honig, "A message-passing approach for joint channel estimation, interference mitigation, and decoding," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6008–6018, Dec. 2009.

[78] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Modified constant modulus algorithm for joint blind equalization and synchronization," in *Proc. Australian Communications Theory Workshop (AusCTW)*, Feb 2010, pp. 59–64.

[79] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.

[80] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.

[81] B. Lee, J. B. Lim, C. Lim, B. Kim, and J. y. Seol, "Reflected self-interference channel measurement for mmwave beamformed full-duplex system," in *Proc. IEEE Golbecom Workshops*, Dec. 2015, pp. 1–6.

[82] A. Demir, T. Haque, E. Bala, and P. Cabrol, "Exploring the possibility of full-duplex operations in mmwave 5G systems," in *Proc. WAMICON*, Apr. 2016, pp. 1–5.

[83] T. A. Thomas, M. Cudak, and T. Kovarik, "Blind phase noise mitigation for a 72 GHz millimeter wave system," in *Proc. IEEE ICC*, June 2015, pp. 1352–1357.

[84] A. Masmoudi and T. Le-Ngoc, "Channel estimation and self-interference cancellation in full-duplex communication systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 321–334, Jan. 2017.

[85] X. Xiong, X. Wang, T. Riihonen, and X. You, "Channel estimation for full-duplex relay systems with large-scale antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6925–6938, Oct. 2016.

[86] R. Li, A. Masmoudi, and T. Le-Ngoc, "Self-interference cancellation with nonlinearity and phase-noise suppression in full-duplex systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2118–2129, Mar. 2018.

[87] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator in MIMO full-duplex systems," in *Proc. IEEE VTC*, Sep. 2014.

[88] A. A. Nasir, H. Mehrpouyan, S. Durrani, S. D. Blostein, R. A. Kennedy, and B. Ottersten, "Optimal training sequences for joint timing synchronization and channel estimation in distributed communication networks," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 3002–3015, Jul 2013.

[89] ——, "Transceiver design for distributed STBC based af cooperative networks in the presence of timing and frequency offsets," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3143–3158, Jun. 2013.

[90] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Blind timing and carrier synchronisation in distributed multiple input multiple output communication systems," *IET Communications*, vol. 5, no. 7, pp. 1028–1037, May 2011.

[91] M. S. Sim, M. Chung, D. Kim, J. Chung, D. K. Kim, and C. B. Chae, "Nonlinear self-interference cancellation for full-duplex radios: From link-level and system-level performance perspectives," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 158–167, Jun. 2017.

[92] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[93] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3933–3938, Aug. 2009.

[94] N. Li, W. Zhu, and H. Han, "Digital interference cancellation in single channel, full duplex wireless communication," in *Proc. Wireless Communications, Networking and Mobile Computing (WiCOM)*, Sep. 2012, pp. 1–4.

[95] J. R. Krier and I. F. Akyildiz, "Active self-interference cancellation of passband signals using gradient descent," in *Proc. IEEE Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2013, pp. 1212–1216.

[96] E. Panayirci, H. A. Cirpan, M. Moeneclaey, and N. Noels, "Blind data detection in the presence of PLL phase noise by sequential monte carlo method," in *Proc. ICC*, vol. 7, Jun. 2006, pp. 3202–3206.

[97] A. A. Quadeer, T. Y. Al-Naffouri, and M. Shadaydeh, "Iterative blind data detection in constant modulus OFDM systems," in *Proc. European Signal Processing Conference*, Aug. 2008, pp. 1–5.

[98] Z. Jiang, X. Shen, Y. Ge, and H. Wang, "Approximate gibbs algorithm for blind data detection in two-way relay networks," *IET Communications*, vol. 11, no. 8, pp. 1230–1240, 2017.

[99] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2205–2217, Apr. 2015.

[100] A. Shojaeifard, K. Wong, M. D. Renzo, G. Zheng, K. A. Hamdi, and J. Tang, "Massive MIMO-enabled full-duplex cellular networks," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4734–4750, Nov. 2017.

[101] ——, "Self-interference in full-duplex multi-user MIMO channels," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 841–844, Apr. 2017.

[102] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and J. Lilleberg, "Full-duplex in large-scale wireless systems," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Nov. 2013, pp. 1623–1627.

[103] K. Min, S. Park, Y. Jang, T. Kim, and S. Choi, "Antenna ratio for sum-rate maximization in full-duplex large-array base station with half-duplex multiantenna users," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 168–10 173, Dec. 2016.

[104] A. Koohian, H. Mehrpouyan, A. A. Nasir, S. Durrani, and S. D. Blostein, "Superimposed signaling inspired channel estimation in full-duplex systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, p. 8, Jan. 2018. [Online]. Available: https://doi.org/10.1186/s13634-018-0529-9

[105] A. Koohian, H. Mehrpouyan, A. A. Nasir, and S. Durrani, "Joint channel and phase noise estimation for mmWave full-duplex communication systems," *EURASIP Journal on Advances in Signal Processing*, vol. 2019, no. 1, p. 18, Mar 2019. [Online]. Available: https://doi.org/10.1186/s13634-019-0614-8

[106] A. Koohian, H. Mehrpouyan, A. A. Nasir, S. Durrani, and S. D. Blostein, "Residual self-interference cancellation and data detection in full-duplex communication systems," in *Proc. IEEE ICC*, May 2017, pp. 1–6.

[107] A. Koohian and S. Durrani, "Self-interference suppression in full-duplex massive MIMO communication," *to be submitted.*

[108] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. Asilomar Conf. on Signals, Syst. and Computers*, Nov 2010.

[109] E. Lehmann and G. Casella, *Theory of Point Estimation*. Springer Verlag, 1998.

[110] P. J. Pahl and R. Damrath, *Mathematical Foundations of Computational Engineering: A Handbook*. Springer, 2001.

[111] N. Loehr, *Bijective Combinatorics*, 1st ed. Chapman & Hall/CRC, 2011.

[112] E. Romero-Aguirre, R. Parra, A. G. Orozco, and R. Carrasco-Alvarez, "Full-hardware architectures for data-dependent superimposed training channel estimation," in *Proc. Workshop on Signal Processing Systems (SiPS)*, Oct. 2011.

[113] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, pp. 1–38, 1977.

[114] V. Melnykov and I. Melnykov, "Initializing the EM algorithm in Gaussian mixture models with an unknown number of components," *Comput. Stat. Data Analysis*, vol. 56, no. 6, pp. 1381–1395, Jun. 2012.

[115] C. Biernacki, G. Celeux, and G. Govaert, "Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models," *Comput. Stat. Data Analysis*, vol. 41, no. 3-4, pp. 561–575, Jan. 2003.

[116] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[117] S. Abdallah and I. N. Psaromiligkos, "Exact Cramer Rao bounds for semi-blind channel estimation in Amplify-and-Forward Two-Way Relay Networks employing square QAM," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 6955–6967, Dec. 2014.

[118] H. Mehrpouyan and S. D. Blostein, "Bounds and algorithms for multiple frequency offset estimation in cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1300–1311, Apr. 2011.

[119] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. New York, NY, USA: Cambridge University Press, 2005.

[120] L. Samara, M. Mokhtar, O. Ozdemir, R. Hamila, and T. Khattab, "Residual self-interference analysis for full-duplex OFDM transceivers under phase noise and i/q imbalance," *IEEE Commun. Lett.*, vol. 21, no. 2, pp. 314–317, Feb. 2017.

[121] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Rate gain region and design tradeoffs for full-duplex wireless communications," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3556–3565, Jul. 2013.

[122] A. A. Nasir, H. Mehrpouyan, R. Schober, and Y. Hua, "Phase noise in MIMO systems: Bayesian Cramer Rao bounds and soft-input estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2675–2692, May 2013.

[123] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, and A. Szufarska, "A flexible 5G frame structure design for frequency-division duplex cases," *IEEE Commun. Mag.*, vol. 54, no. 3, pp. 53–59, Mar. 2016.

[124] S. Dutta, M. Mezzavilla, R. Ford, M. Zhang, S. Rangan, and M. Zorzi, "Frame structure design and analysis for millimeter wave cellular systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1508–1522, Mar. 2017.

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

[126] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, "Multi-beam NOMA for hybrid mmWave systems," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1705–1719, Feb 2019.

[127] L. Zhao, Z. Wei, D. W. K. Ng, J. Yuan, and M. C. Reed, "Multi-cell hybrid millimeter wave systems: Pilot contamination and interference mitigation," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5740–5755, Nov 2018.

[128] L. Zhao, G. Geraci, T. Yang, D. W. K. Ng, and J. Yuan, "A tone-based aoa estimation and multiuser precoding for millimeter wave massive MIMO," *IEEE Trans. Commun.*, vol. 65, no. 12, pp. 5209–5225, Dec 2017.

[129] O. H. Salim, A. A. Nasir, H. Mehrpouyan, W. Xiang, S. Durrani, and R. A. Kennedy, "Channel, phase noise, and frequency offset in OFDM systems: Joint estimation, data detection, and hybrid Cramer-Rao lower bound," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3311–3325, Sep. 2014.

[130] M. R. Khanzadi, R. Krishnan, D. Kuylenstierna, and T. Eriksson, "Oscillator phase noise and small-scale channel fading in higher frequency bands," in *Proc. IEEE Globecom Workshops*, Dec. 2014, pp. 410–415.

[131] A. A. Nasir, H. Mehrpouyan, S. D. Blostein, S. Durrani, and R. A. Kennedy, "Timing and carrier synchronization with channel estimation in multi-relay cooperative networks," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 793–811, Feb. 2012.

[132] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[133] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, Secondquarter 2018.

[134] A. G. Siamarou, "Digital transmission over millimeter-wave radio channels: A review [wireless corner]," *IEEE Antennas Propag. Mag.*, vol. 51, no. 6, pp. 196–203, Dec. 2009.

[135] Z. Zhang, K. Long, A. V. Vasilakos, and L. Hanzo, "Full-duplex wireless communications: Challenges, solutions, and future research directions," *Proceedings of the IEEE*, vol. 104, no. 7, pp. 1369–1409, Jul. 2016.

[136] W. Feng, Y. Wang, D. Lin, N. Ge, J. Lu, and S. Li, "When mmWave communications meet network densification: A scalable interference coordination perspective," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1459–1471, Jul. 2017.

[137] B. Farhang-Boroujeny, "Pilot-based channel identification: proposal for semi-blind identification of communication channels," *Electronics Letters*, vol. 31, no. 13, pp. 1044–1046, Jun. 1995.

[138] R. W. Heath and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 848–856, Mar. 1999.

[139] C. Qin, N. Santhapuri, S. Sen, and S. Nelakuditi, "Known interference cancellation: Resolving collisions due to repeated transmissions," in *Proc. IEEE WIMESH*, Jun. 2010, pp. 1–6.

[140] A. Shojaeifard, K. Wong, M. D. Renzo, K. A. Hamdi, and J. Tang, "Design and analysis of full-duplex massive MIMO cellular networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–6.

[141] E. Bjornson, L. Sanguinetti, and J. Hoydis, "Hardware distortion correlation has negligible impact on UL massive MIMO spectral efficiency," *IEEE Trans. Commun.*, 2018 (to appear).

[142] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170–3184, Apr. 2017.

[143] T. S. Shores, *Applied Linear Algebra and Matrix Analysis*, 2007.

[144] E. Bjornson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[145] G. Geraci, M. Egan, J. Yuan, A. Razi, and I. B. Collings, "Secrecy sum-rates for multi-user MIMO regularized channel inversion precoding," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3472–3482, Nov 2012.

[146] G. Geraci, R. Couillet, J. Yuan, M. Debbah, and I. B. Collings, "Secrecy sum-rates with regularized channel inversion precoding under imperfect CSI at the transmitter," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 2896–2900.

[147] S. Durrani and M. E. Bialkowski, "Effect of mutual coupling on the interference rejection capabilities of linear and circular arrays in CDMA systems," *IEEE Trans. Antennas Propag.*, vol. 52, no. 4, pp. 1130–1134, Apr. 2004.

[148] G. Strang, *Linear Algebra and its applications*. Harcourt Brace Jovanovich Publishers, 1988.

[149] R. Zamir, "A proof of the Fisher information inequality via a data processing argument," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1246–1250, May 1998.

[150] S. Arora and B. Barak, *Computational Complexity: A Modern Approach*, 1st ed. New York, NY, USA: Cambridge University Press, 2009.

[151] E. A. Wan and R. V. D. Merwe, "The unscented Kalman filter for nonlinear estimation," in *in Proc. IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*, Oct. 2000, pp. 153–158.

[152] Y. Wu, D. Hu, M. Wu, and X. Hu, "Unscented Kalman filtering for additive noise case: augmented versus nonaugmented," *IEEE Signal Process. Lett.*, vol. 12, no. 5, pp. 357–360, May 2005.

[153] M. L. Eaton and M. D. Perlman, "The non-singularity of generalized sample covariance matrices," *The Annals of Statistics*, vol. 1, no. 4, pp. 710–717, 1973.

[154] P. Billingsley, *Probability and Measure*, 2nd ed. John Wiley and Sons, 1986.

[155] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.