Synchronization in Cooperative Communication Systems

Ali Arshad Nasir

September 2012

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

Australian National University

Research School of 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 a higher degree to any other university or institution.

Much of 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. 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.

J2. **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.*, 2013. (accepted 20 Feb, 2013)

J3. **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," submitted to *IEEE Trans. Commun.*, July. 2012. (manuscript no. TCOM-TPS-12-0541, revised 9 Feb, 2013)

J4. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Blind Timing and Carrier Synchronization in Distributed MIMO Communication Systems," *IET Commun.*, vol. 5, no. 7, pp. 1028-1037, May 2011.

Conference papers

C1. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Estimation of Synchronization Parameters in AF Cooperative Networks", in *Proc. IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June 10-15, 2012.

C2. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Blind Timing and Carrier Synchronization in Decode and Forward Cooperative Systems", in *Proc. IEEE International Conference on Communications (ICC)*, Kyoto, Japan, June 5-9, 2011.

The following publications are also the result from my Ph.D. study but not included in this thesis:

Journal articles

J5. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Particle Filters for Joint Timing and Carrier Estimation: Improved Resampling Guidelines and Weighted Bayesian Cramer-Rao Bounds," in *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1407-1419, May 2012.

J6. 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. on Signal Process.*, vol. 60, no. 9, pp. 4790-4807, Sep. 2012.

J7. **A. A. Nasir**, H. Mehrpouyan, R. Schober, and Y. Hua, "Phase Noise in MIMO Systems: Bayesian Cramr-Rao Bounds and Soft-Input Estimation", *IEEE Trans. on Signal Process.*, 2013. (accepted 17 Jan, 2013)

Conference papers

C3. H. Mehrpouyan, **A. A. Nasir**, T. Eriksson, S. D. Blostein, G. K. Karagiannidis, and T. Svensson "Time-Varying Phase Noise and Channel Estimation in MIMO Systems", in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Cesme, Turkey, Jun. 17-20, 2012.

C4. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Achieving Cooperative Diversity with Multiple Frequency Offset Estimation", in Proc. *IEEE International Conference on Signal Processing and Communication Systems (ICSPCS)*, Honolulu, USA, Dec. 12-14, 2011.

C5. **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)*, Prague, Czech Republic, May 22-27, 2011.

C6. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Particle Filter for Joint Blind Carrier Frequency Offset Estimation and Data Detection", in *Proc. IEEE International Conference on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, Australia, Dec. 13-15, 2010.

C7. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Performance of Coarse and Fine Timing Synchronization in OFDM Receivers", in *Proc. IEEE International Conference on Future Computer and Communication (ICFCC)*, vol. 2, Wuhan, China, 21-24 May, 2010, pp. 412-416.

C8. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Blind Fractionally Spaced Equalization and Timing Synchronization in Wireless Fading Channels", in *Proc. IEEE International Conference on Future Computer and Communication (ICFCC)*, vol. 3, Wuhan, China, 21-24 May, 2010, pp. 15-19.

C9. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Modified Constant Modulus Algorithm for joint blind equalization and synchronization", in *Proc. IEEE Australian Communication Theory Workshop (AUSCTW)*, Canberra, Australia, pp. 59-64 Feb. 2010.

The research work presented in this thesis has been performed jointly with Dr. Salman Durrani (The Australian National University, Australia), Prof. Rodney A. Kennedy (The Australian National University, Australia), Dr. Hani Mehrpouyan (California State University, Bakersfield), Prof. Steven D. Blostein (Queen's University, Canada) and Prof. Björn Ottersten (University of Luxembourg, Luxembourg). The substantial majority of this work was my own.

Ali Arshad Nasir
Research School of Engineering,
The Australian National University,
Canberra, ACT 0200
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 the foremost, I would like to thank Almighty Allah for bestowing His blessings upon me and giving me the strength to carry out and complete this work.

- I would like to express my sincere thanks to my supervisors and friends Drs Salman Durrani and Hani Mehrpouyan and Profs Rodney A. Kennedy and Steven D. Blostein, for their guidance, support and encouragement throughout my PhD studies. They are always willing to have in-depth discussion whenever I approach them for help. Apart from my research studies, they have also given me great support in obtaining teaching and supervision experiences.

- I would like to thank Dr. Hani Mehrpouyan, Prof. Tommy Svensson and Prof. Thomas Eriksson from Chalmers University of Technology for kindly welcoming me to visit their research groups for six months, i.e., Apr. 2011 to Oct. 2011. The research in Chalmers University encouraged me to explore and publish in the research area of phase noise estimation in multiple-input-multiple-output (MIMO) systems.

- It is my great pleasure to study in the Applied Signal Processing (ASP) group at the Research School of Engineering. I would like to thank everyone for making ASP group a friendly and relaxing research environment. Special thanks go to Ms. Lesley Goldburg, our student administrator, for her assistance.

- Thanks must go to the Australian National University for providing me with the PhD scholarship, stipend, and a 2011 ANU Vice Chancellors Travel Grant

for supporting my conference attendances, living expenses, and university visit, respectively.

- I thank my parents for providing me the opportunity of studying abroad and their great financial support. My every little success would not have been possible without their prayers. I would like to give my special thanks to my wife for her constant support, encouragement, understanding, and prayers. Finally, I would like to thank my uncle and aunt in Canberra, who supported me throughout my PhD and provided me with a homely environment overseas.

Abstract

Cooperative communication is an attractive solution to combat fading in wireless communication systems. Achieving synchronization is a fundamental requirement in such systems. In cooperative networks, multiple single antenna relay terminals receive and cooperatively transmit the source information to the destination. The multiple distributed nodes, each with its own local oscillator, give rise to multiple timing offsets (MTOs) and multiple carrier frequency offsets (MCFOs). Particularly, the received signal at the destination is the superposition of the relays' transmitted signals that are attenuated differently, are no longer aligned with each other in time, and experience phase rotations at different rates due to different channels, MTOs, and MCFOs, respectively. The loss of synchronization due to the presence of MTOs and MCFOs sets up the recovery of the source signal at the destination to be a very challenging task. This thesis seeks to develop estimation and compensation algorithms that can achieve synchronization and enable cooperative communication for both decode-and-forward (DF) and amplify-and-forward (AF) relaying networks in the presence of multiple impairments, i.e., unknown channel gains, MTOs, and MCFOs.

In the first part of the thesis, a training-based transmission scheme is considered, in which training symbols are transmitted first in order to assist the joint estimation of multiple impairments at the destination node in DF and AF cooperative relaying networks. New transceiver structure at the relays and novel receiver design at the destination are proposed which allow for the decoding of the received signal in the presence of unknown channel gains, MTOs, and MCFOs. Different estimation algorithms, e.g., least squares (LS), expectation conditional maximization (ECM), space-alternating generalized expectation-maximization (SAGE), and differential evolution (DE), are proposed and analyzed for joint estimation of multiple impairments. In order to compare the estimation accuracy of the proposed estimators, Cramér-Rao lower bounds (CRLBs) for the multi-parameter estimation are derived. Next, in order to detect the signal from multiple relays in the presence of multiple impairments, novel optimal and sub-optimal minimum mean-square er-

ror (MMSE) compensation and maximum likelihood (ML) decoding algorithm are proposed for the destination receiver. It has been evidenced by numerical simulations that application of the proposed estimation and compensation methods in conjunction with space-time block codes achieve full diversity gain in the presence of channel and synchronization impairments. Considering training-based transmission scheme, this thesis also addresses the design of optimal training sequences for efficient and joint estimation of MTOs and multiple channel parameters.

In the second part of the thesis, the problem of joint estimation and compensation of multiple impairments in non-data-aided (NDA) DF cooperative systems is addressed. The use of blind source separation is proposed at the destination to convert the difficult problem of jointly estimating the multiple synchronization parameters in the relaying phase into more tractable sub-problems of estimating many individual timing offsets and carrier frequency offsets for the independent relays. Next, a criteria for best relay selection is proposed at the destination. Applying the relay selection algorithm, simulation results demonstrate promising bit-error rate (BER) performance and realise the achievable maximum diversity order at the destination.

List of Acronyms

AF	amplify-and-forward
AP	alternating projection
AWGN	additive white Gaussian noise
BCRLB	Bayesian Cramér-Rao lower bound
BER	bit error rate
BIM	Bayesian information matrix
BPSK	binary phase-shift keying
BSS	blind source separation
CRC	cyclic redundancy check
CSI	channel state information
CRLBs	Cramér-Rao lower bounds
DA	data-aided
DF	decode-and-forward
DE	differential evolution
DSTBC	distributed space-time block code
DTP	data transmission period
ECM	expectation conditional maximization
ECRLB	conditional Cramér-Rao lower bound
EM	expectation-maximization
EVD	eigen-value decomposition
FIM	Fisher's information matrix
i.i.d.	independent and identically distributed
JADE	joint approximate diagonalization of eigen-matrices
LLF	log-likelihood function
L-MMSE	low-complexity minimum mean-square error
LS	least squares
MAP	maximum a posteriori
MCFOs	multiple carrier frequency offsets
MIMO	multiple-input multiple-output

MISO	multiple-input single-output
ML	maximum likelihood
MMSE	minimum mean-square error
MSE	mean-square error
MTOs	multiple timing offsets
NDA	non-data-aided
NLS	non-linear least squares
PDF	probability density function
PSK	phase shift keying
QAM	quadrature amplitude modulation
QPSK	quadrature phase-shift keying
RRC	root-raised cosine
TDMA	time-division multiple-access
TSs	training signals
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
SAGE	space-alternating generalized expectation-maximization
SDMA	space division multiple access
SISO	single-input-single-output
SNR	signal-to-noise-ratio
STBC	space-time block code
TP	training period
ZF	zero-forcing

Notations

x	scalar variable
$\mathbf{x}$	vector variable
$\mathbf{X}$	matrix variable
$[\mathbf{X}]_{x,y}$	element in row x and column y of $\mathbf{X}$
$[\mathbf{X}]_{a:b,x:y}$	submatrix formed by the rows, a to b, and columns, x to y, of the matrix $\mathbf{X}$
$\mathbf{I}_X$	$X \times X$ identity matrix
$\mathbf{0}_{X\times X}$	$X \times X$ matrix of all zeros
$\mathbf{1}_{X\times X}$	$X \times X$ matrix of all ones
$\hat{x}$	estimated value of x
$(\cdot)^*$	conjugate operator
$(\cdot)^T$	transpose operator
$(\cdot)^T$	hermitian (conjugate transpose) operator
$(\cdot)'$	first derivative operator
$(\cdot)''$	second derivative operator
$(\cdot)^{[\mathrm{TP}]}$	signal during training period
$(\cdot)^{[\mathrm{DTP}]}$	signal during data transmission period
$(\cdot)^{[\mathrm{sr}]}$	source to relay parameter
$(\cdot)^{[\mathrm{rd}]}$	relay to destination parameter
$\odot$	Schur element-wise product (Hadamard product)
$\otimes$	Kronecker product
$\lvert x\rvert$	absolute value of a scalar x
$\lvert\mathbf{x}\rvert$	element-wise absolute value of $\mathbf{x}$
$\lVert\mathbf{x}\rVert$	ℓ_2 norm of $\mathbf{x}$
$\mathbb{E}\{\cdot\}$	mathematical expectation operator
$\Re\{\cdot\}$	real part of a complex quantity
$\Im\{\cdot\}$	imaginary part of a complex quantity
$\mathrm{diag}(\mathbf{X})$	a vector containing diagonal elements of the matrix $\mathbf{X}$
$\mathrm{diag}(\mathbf{x})$	a diagonal matrix with diagonal elements given by vector $\mathbf{x}$

$\text{diag}([\mathbf{X_1}, \mathbf{X_2}, \ldots])$	a block diagonal matrix, where $[\mathbf{X_1}, \mathbf{X_2}, \ldots]$ are the diagonal matrices
$\text{Tr}\{\mathbf{X}\}$	trace of the matrix $\mathbf{X}$
$\text{mod}(a/b)$	remainder of division of a by b
$\lfloor \cdot \rfloor$	floor function
$\mathcal{N}(\mu, \sigma^2)$	real Gaussian distribution with mean μ and variance σ^2
$\mathcal{CN}(\mu, \sigma^2)$	complex Gaussian distribution with mean μ and variance σ^2
$\mathcal{U}(-x, x)$	uniform distribution between $-x$ and x
$\Delta_{\mathbf{x}}^{\mathbf{x}} f(\cdot)$	$\triangleq \frac{\partial}{\partial \mathbf{x}} \left[\frac{\partial}{\partial \mathbf{x}} f(\cdot)\right]^T$ is the Hessian operator

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and Background

Wireless communication systems have been developing and improving in performance for nearly two decades. During that time, major theoretical advances such as the use of multiple antennas at the transmitter and receiver have been quickly incorporated into communication standards, underlying the value of theoretical research. However, future communication systems are diverging in their paradigm from the classical model. Communications systems were once thought of as well specified in terms of a transmitter with a fixed and known location transmitting over a known channel to a single receiver of fixed and known location. Now it is proposed that communication takes place almost opportunistically with a variable number of intermediate relays coordinating to get the information from the source node to the destination node [4–6].

A key research challenge in the design of future wireless networks is achieving accurate synchronization among cooperating nodes or users. Timing and carrier synchronization, also referred to as physical layer synchronization, is required before communication between any two users can begin [7]. Recent research has focused mainly on communication strategies for cooperative, multiuser and multiple antenna communications in wireless networks, assuming perfect synchronization. However the design of synchronization techniques to enable such wireless networks to be widely deployed in the communications mainstream has been largely unaddressed.

Synchronization is a common phenomenon in nature, e.g., the synchronized flashing of fireflies or the synchronous firing of neurons in the human brain [8]. In wireless communications, in general, a receiver node does not have prior knowledge of the physical wireless channel or the propagation delay associated with the

transmitted signal. Moreover, to keep the cost of the devices as low as possible, mass produced communication receivers use low cost oscillators which inherently have some drift [9]. Thus, synchronization is a fundamental requirement for any receiver to work properly. In this context:

1. *Timing synchronization* is the process by which a receiver node determines the correct instants of time at which to sample the incoming signal and

2. *Carrier synchronization* is the process by which a receiver adapts the frequency and phase of its local carrier oscillator with those of the received signal.

For instance, requiring two watches to be time synchronized means that they should both display the same time. However, requiring two watches to be carrier synchronized means that they should tick at the same speed, irrespective of what time they show [10]. Current cellular systems require carrier accuracy of 50 parts per billion and timing accuracy within the interval of 1 μs ($\pm$500 ns) [11]. The requirement in future wireless networks is towards tighter accuracies, e.g., timing accuracy of 200 ns to enable location-based services [10]. Both timing and carrier synchronization are necessary for successful communication between two nodes.

In conventional multiple-input-multiple-output (MIMO) systems, the antenna elements are co-located on a single device, which results in a single timing and carrier offset. However, in cooperative networks, multiple distributed nodes, each with its own local oscillator, gives rise to multiple timing offsets (MTOs) and multiple carrier frequency offsets (MCFOs), which need to be jointly estimated and compensated in order to decode the received signal at the destination [1, 2, 12]. On the other hand, it is well established that given perfect synchronization, the same spatial cooperative diversity as that of MIMO systems is achievable in cooperative networks without the need for multiple antennas at each node [13–17]. This reiterates the importance of achieving accurate synchronization in cooperative communication systems.

In the context of estimation of synchronization parameters, it is important to assess the achievable estimation accuracy for any proposed estimator. The *Cramer-Rao lower bound (CRLB)* is a lower bound on the variance of any unbiased estimator [18]. The CRLB is used as the performance benchmark for any unbiased estimator of a deterministic parameter. For a random parameter, Van Trees presented an analogous bound, the posterior CRLB (PCRLB) [19], which is also referred to as the Bayesian CRLB [20]. Such bounds are very important tools since they give the best performance that can be achieved by an ideal esti-

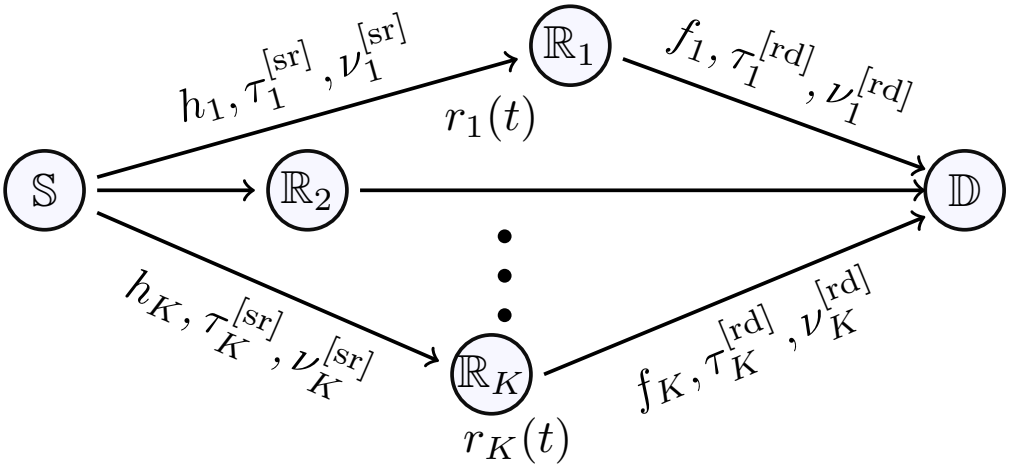


Figure 1.1: The system model for the cooperative network.

mator. Any practical estimator can then be compared with the bounds to assess its performance.

The typical communication scenario for a cooperative communication system is illustrated in Fig. 1.1. It comprises a half-duplex space division multiple access (SDMA) SISO cooperative communication system with one source node, $\mathbb{S}$, K relays, $\mathbb{R}_1, \cdots, \mathbb{R}_K$, and a single destination node, $\mathbb{D}$, as shown in Fig. 1.1. Unless otherwise specified, each node is equipped with a single omnidirectional antenna. The channel gains from the source to kth relay and the kth relay to destination are denoted by h_k and f_k, respectively. The index $k = 1, \ldots, K$ is used for the K relays. In Fig. 3.1, τ_k and ν_k are used to denote the timing and frequency offsets, respectively, where superscripts $(\cdot)^{[sr]}$ and $(\cdot)^{[rd]}$ denote offsets from $\mathbb{S}$ to $\mathbb{R}_k$ and $\mathbb{R}_k$ to $\mathbb{D}$, respectively.

The following assumptions have been made throughout the thesis:

- Quasi-static and frequency flat-fading channels are considered, which is motivated by prior research in this field in [1,2,12,21–25]. This allows the performance of the algorithms to be tested under worst case channel conditions, i.e., Rayleigh fading channel. The assumption of frequency flat channels can be broadened to frequency selective channels by employing orthogonal frequency division multiplexing (OFDM).

- Over a frame, the timing and frequency offsets, as shown in Fig. 1.1, are modeled as deterministic but unknown parameters.

- It is assumed that a coarse synchronizer is first applied and signals from the different nodes are within the same symbol period, i.e., the difference

between the timing offsets of any two nodes is given by $|\tau_k - \tau_{\bar{k}}| < 1$, for $k, \bar{k} = 1, \ldots, K$ and $k \neq \bar{k}$. Most communication systems use such coarse synchronizers before applying a fine synchronization algorithm to estimate the timing offsets within a symbol period [7, 25, 26].

During the broadcasting phase, the source node broadcasts its data to all the relays. During the relaying phase, each relay process the source signal and retransmit it to the destination. The received signal at the destination is a function of multiple impairments, i.e., multiple channel gains, $f_1, \ldots, f_K$, MTOs, $\tau_1^{[rd]}, \ldots, \tau_K^{[rd]}$, and MCFOs, $\nu_1^{[rd]}, \ldots, \nu_K^{[rd]}$. *Given the received signal is affected by multiple impairments at the destination, the challenging problem of this thesis is to recover the source signal from the received signal.*

In the following subsections, we provide detailed literature survey and highlight the research problems that are addressed in this thesis.

1.1.1 Synchronization in Data-Aided Cooperative Systems

In data-aided cooperative systems, a training-based transmission scheme is used, in which the training symbols are transmitted first in order to assist joint estimation of the channel parameters, MTOs, and MCFOs at the destination node in DF and AF cooperative relaying networks. Consequently, the transmission of signals from the source to relays to destination consists of a training period (TP) as well as a data transmission period (DTP). During the relaying phase, when relays communicate with the destination, all the relays transmit distinct training signals followed by the source data to the destination. Different training signals from the multiple relays are used to estimate multiple channels, MTOs, and MCFOs. Next, the estimated values of the multiple impairments are used to compensate the effect of the multiple channels, MTOs, and MCFOs and decode the received signal.

For single timing and carrier frequency offset estimation in *single-input-single-output* (SISO) systems, various algorithms have been proposed in the literature [27–32]. For multiple timing and carrier frequency offset estimation, most of the existing work in the literature focuses on estimating either MCFOs while assuming perfect timing synchronization [2, 21–24, 33, 34] or MTOs while assuming perfect carrier synchronization [1, 12, 25, 35].[1] *Cramér-Rao lower bounds* (CRLBs) and different techniques for estimating MCFOs in MIMO systems are addressed in [21–24]. However, the algorithms in [21–24] are not applicable to the case of *amplify-and-forward* (AF)-relaying cooperative networks and are based on the assumption of

[1]The algorithms for tracking multiple time varying phase noise processes are recently studied in [36–38].

perfect timing synchronization. Novel CRLBs and *maximum likelihood* (ML) estimators to obtain MTOs for detect-and-forward, *decode-and-forward* (DF), and AF systems are derived in [35], [25], and [1]- [12], respectively. Similarly, in [2, 33, 39] the estimation of multiple channel gains and MCFOs in DF and AF cooperative systems are analyzed. However, the analyses in [1,2,12,21–25,33,35,39] are focussed on estimating one set of system parameters while assuming that the remaining parameters are perfectly estimated and compensated. Admittedly, such assumptions do not hold in actual cooperative communication systems, where the channel gains, MCFOs, and MTOs need to be jointly estimated at the destination terminal.

Recently, a limited number of papers have investigated joint estimation of impairments. In [40], a new joint ML estimator for determining MCFOs, MTOs, and channel gains in DF cooperative networks is devised. Nevertheless, the ML estimator in [40] requires exhaustive search and is computationally very complex. In order to reduce the complexity of ML estimation, new iterative estimation schemes are proposed in [41] . However, the outcomes in [40, 41] are limited to DF cooperative networks, provide no specific initialization guidelines for the proposed estimators, and do not propose any means of decoding the received signals from multiple relays at the destination. Novel joint channel estimation and time-frequency synchronization for uplink *orthogonal frequency-division multiple access* (OFDMA) systems are proposed in [42–45] that exploit the cyclic prefix. However, depending on the number of sub-carriers used, the frequency acquisition range of the algorithms in [42–45] is very limited.

Estimating MCFOs, MTOs, and unknown channel gains is the first step to achieving synchronization. In fact, in order to achieve synchronous and coherent communication in multi-relay cooperative networks, the estimated MCFOs, MTOs, and unknown channel gains need to be applied to successfully equalize the received signal at the destination node. Recent literature has addressed either MCFO compensation [46–48] or MTO compensation [1, 25]. In [46], a new class of distributed linear convolutional space-time codes are proposed to compensate the effect of MCFOs. However, the application of the codes in [46] reduces overall cooperative network throughput since their effective rate is less than one. In [47, 48], various novel algorithms for compensating and detecting the received signal at the destination in the presence of MCFOs are proposed. Nevertheless, the algorithms in [47,48] can only detect the received signal at the destination over a narrow range of CFO values, e.g., $\left[-10^{-6}, 10^{-6}\right]$ in [48]. More importantly, the analyses in [46–48] focus on compensating one set of parameters and cannot equalize the received signal in the presence of both MCFOs and MTOs. Under the

assumption of perfect frequency synchronization, a novel resynchronization filter is proposed in [1] and [25] to compensate the effect of MTOs in DF and AF cooperative systems, respectively. However, as anticipated and confirmed in this thesis, in the presence of MCFOs, the resynchronization filter in [1] and [25] cannot compensate the effect of MTOs and results in erroneous decoding at the destination. In [26], a novel algorithm for compensating MCFOs and MTOs in *time-division multiple-access* (TDMA) distributed MIMO systems is proposed that exploits spatial diversity at the multi-antenna receiver. However, the method in [26] cannot be applied to the case of multi-relay *single-input-single-output* (SISO) cooperative networks, where multiple single antenna relays communicate with a single antenna destination. To the best of the authors' knowledge, an estimation and equalization scheme for non-OFDM based cooperative networks in the presence of MCFOs, MTOs, and unknown channels has not been proposed in the existing literature.

1.1.2 Synchronization in DSTBC based Data-Aided AF Cooperative Systems

Different relaying protocols have different cooperation strategies [17]. The *repetition cooperative strategy* allows relays to transmit sequentially in different time slots, which leads to inefficient bandwidth utilization [49]. In order to improve bandwidth efficiency, this approach can be modified to allow for the relays to transmit simultaneously in multi-relay cooperative networks. However, depending on how the relays' signals superimpose at the destination, achieving full diversity gain is not guaranteed [3, 50]. The *distributed space-time block code (DSTBC) cooperative strategy* allows relays to transmit in the same time slot using the DSTBC structure [25,51,52]. In association with DSTBC cooperative strategy, amplify-and-forward (AF) is an attractive relaying protocol due to its implementation simplicity, i.e., every relay node linearly processes the received signal before forwarding an appropriately scaled version to the destination. Given its potential, the combination of AF relaying and DSTBC has attracted recent research interest as an efficient means of providing cooperative diversity [53–55]. However, the performance benefits of DSTBC schemes depend on accurate synchronization amongst cooperating nodes, which increases the overhead at the relays.

The distributed nature of cooperative communication systems makes synchronization for AF relaying with DSTBC a non-trivial task. Unlike conventional multiple-input multiple-output systems, where co-located antenna elements may result in a single timing and a single carrier frequency offset, asynchronism amongst the multiple distributed relays in cooperative networks gives rise to *multiple timing*

offsets (MTOs) and *multiple carrier frequency offsets (MCFOs)* at the destination [3]. It has been shown that time asynchronism can destroy the elaborate structure of DSTBCs, reducing the achievable diversity gain [1, 56]. Similarly, carrier frequency asynchronism can lead to symbol-to-symbol channel time variation and to a violation of the essential requirement for successful DSTBC decoding and achieving spatial diversity [57].

Most of the literature on AF relaying with DSTBC has focused on analyzing the end-to-end system performance in terms of signal-to-noise ratio (SNR), outage probability, and symbol error probability while assuming perfect synchronization [53,54,58,59]. Others have focused on the design of delay tolerant DSTBC [60, 61] and carrier frequency offset tolerant space-frequency codes [62] for decode-and-forward (DF) systems. However, the algorithms in [60–62] do not consider the joint estimation and compensation of MTOs and MCFOs and cannot be applied to AF relaying due to the codes that rely on the assumption that the received signals at the relays are free of any timing and frequency offset. More importantly, [60] and [62] provide no means of estimating MTOs and MCFOs, even though they require these parameters to be obtained at the destination for successful decoding.

Many algorithms have been proposed for estimation and compensation of channel gains [55, 63], MTOs [64–66], or MCFOs [33, 67] in DSTBC-orthogonal frequency division multiplexing (OFDM) AF relaying systems. However, the algorithms in [33, 55, 63–67] cannot obtain both timing and carrier frequency offsets *jointly* at the destination. Even though joint synchronization schemes for obtaining and compensating multiple channel gains, MTOs, and MCFOs for DF relaying DSTBC-OFDM based cooperative systems are available in the literature [68–70], these solutions cannot be applied to AF relaying systems due to the fundamental differences in the relaying protocols and transceiver structures at the relays and destination. In addition, the algorithms in [55, 63–70] exploit the cyclic prefix and the frequency domain structure of the signal, which is specific to OFDM systems and depending on the number of sub-carriers used, the carrier frequency offset acquisition range of the algorithms in [55, 63–70] is very limited.

For general AF relaying with DSTBC, channel estimation and equalization [39, 53, 71], joint channel and MCFO [2, 46, 47], or joint channel and MTO estimation and compensation [1] are addressed in the literature. However, these schemes do not jointly estimate and compensate the multiple channel gains, MTOs and MCFOs, i.e., they estimate and compensate either MCFOs while assuming perfect timing synchronization [2, 33, 46, 47] or vice versa [1]. As shown in Section 3.5.2 in this thesis, such an approach does not allow for successful decoding of the received

signal at the destination in the presence of both MTOs and MCFOs.

1.1.3 Training Sequence Design for Data-Aided Cooperative Systems

Training sequences (TSs) are widely used in communication systems for channel estimation and synchronization [7]. Since the choice of TS significantly affects estimation and system performance [72], TS design is an important topic in the field of communications. It is also known that in distributed and cooperative communication systems, the receiver needs to detect and decode the signals from multiple nodes that are affected by multiple channels, multiple timing offsets (MTOs), and multiple carrier frequency offsets (MCFOs) [3, 26]. However, the main focus in the literature has been on the design of TSs for channel estimation [39, 73–88] or MCFO and channel estimation [89–93] while relatively little attention has been paid to the topic of TS design for the joint MTO and multiple channel parameter estimation [12].

To date, different approaches have been adopted to design TSs that improve the estimation performance. In [81,82,89], the authors consider specific estimation algorithms and attempt to find the TSs that minimize the corresponding variance of the estimation error. However, the resulting TSs are only optimal for the estimation method under consideration and may not have a broader application. In [73, 93], and [12], new TSs that minimize the Cramér-Rao lower bound (CRLB), which is the lower bound on the variance of the estimation error of any unbiased estimator, are derived. However, since the CRLBs in [73,93], and [12] are functions of specific channel realizations, the resulting TSs cannot be guaranteed to be optimal for all instances of random channels. In this regard, it is desirable to derive optimal TSs that are independent of the channel realizations.

Even though estimators for obtaining the MTOs and channel parameters in multi-user distributed networks, decode-and-forward (DF) cooperative systems, and amplify and forward (AF) cooperative systems are proposed in [26], [25], and [1], respectively, the design of optimal TSs is left as an open research area [1,25]. For the case of *single* timing offset estimation in point-to-point MIMO systems, [94] proposes optimal TSs based on the design in [84]. However, in [72], it is shown that the proposed TSs in [94] are not optimal in general and instead, Walsh sequences are proposed as a more appropriate alternative. More importantly, the solutions in [94] and [72] are not optimal and are only applicable to point-to-point MIMO systems, since they assume that the received signal is affected by a single timing offset. Thus, as shown in this thesis, they are not sufficient for the design

of optimal TSs for joint channel and MTO estimation in multi-user and cooperative networks. Finally, even though, [12] presents some guidelines for the design of optimal TSs for MTO and multiple channel estimation in cooperative multi-relay networks, they are mainly obtained using numerical investigations, and are not based on any analytical framework.

1.1.4 Synchronization in Non-Data-Aided Cooperative Systems

The use of training sequences can give rise to significant overhead with data rate reduction and may become unrealistic or impractical in certain scenarios. For instance, no training signal may be available to receivers in military communication scenarios and defence applications. In a multicast or broadcast system, it is highly undesirable for the transmitter to engage in a training session for a single user by temporarily suspending its normal transmission to a number of other users. Consequently, there is a strong and practical need for blind equalization without training [95].

Considering joint synchronization and channel estimation in distributed cooperative communication systems, most of the existing literature focuses on training based estimation and compensation methods. Only a few authors have proposed blind solutions. A blind method of multiple carrier frequency offset estimation is proposed in [96] based on formulating a virtual MIMO system using polyphase components of the oversampled signal. In [97, 98], the approach is shown to be able to separate two relays in a distributed multiple-input-single-output (MISO) system. However for more than two relays, the oversampling requirements become prohibitive and the approach becomes impractical. To the best of our knowledge, the joint solution for blind timing and carrier synchronization and channel estimation in DF cooperative communication systems does not exist in the existing literature.

1.2 Overview and Contribution of Thesis

This thesis seeks to develop estimation and compensation algorithms that can enable cooperative communications for both DF and AF relaying networks in the presence of multiple *impairments*, i.e., unknown channel gains, MTOs, and MCFOs. Both data-aided and non-data-aided systems are investigated for the joint estimation and compensation of multiple impairments. Moreover, the design of optimal

training sequences for efficient and joint estimation of MTOs and multiple channels in data-aided systems is addressed. *To the best of the our knowledge, the design of (i) a joint estimation and compensation scheme for the data-aided amplify-and-forward (AF) relaying and (ii) non-data-aided decode-and-forward (DF) relaying in the presence of multiple unknown channels, MTOs, and MCFOs, proposed in this thesis has not been considered before in the literature.*

1.2.1 Questions to be Answered

The following open questions are addressed in the thesis:

Q1. How can the received signal be decoded at the destination in DA cooperative communication systems in the presence of multiple impairments, i.e., unknown channel gains, MTOs, and MCFOs? What is the transceiver design at the relays and receiver design at the destination for the data-aided DF and AF cooperative communication systems in the presence of multiple impairments?

Q2. How can we jointly estimate unknown channel gains, MTOs, and MCFOs at the destination in DA DF and AF cooperative communication systems?

Q3. How can we compensate the effect of unknown channel gains, MTOs, and MCFOs and decode the received signal at the destination in DA DF and AF cooperative communication systems?

Q4. What is the lowest achievable estimation accuracy (CRLBs) for the joint estimation of multiple impairments at the destination in DA DF and AF cooperative communication systems?

Q5. How does the AF transceiver design change in the presence of DSTBCs? Does the application of DSTBCs help to achieve full order diversity gain in the presence of channel and synchronization impairments?

Q6. What are the optimal training sequences and training sequence design guidelines for the efficient and joint estimation of MTOs and multiple channel parameters?

Q7. How one can blindly (no training signals) decode the received signal at the destination in NDA cooperative communication systems in the presence of unknown channel gains, MTOs, and MCFOs?

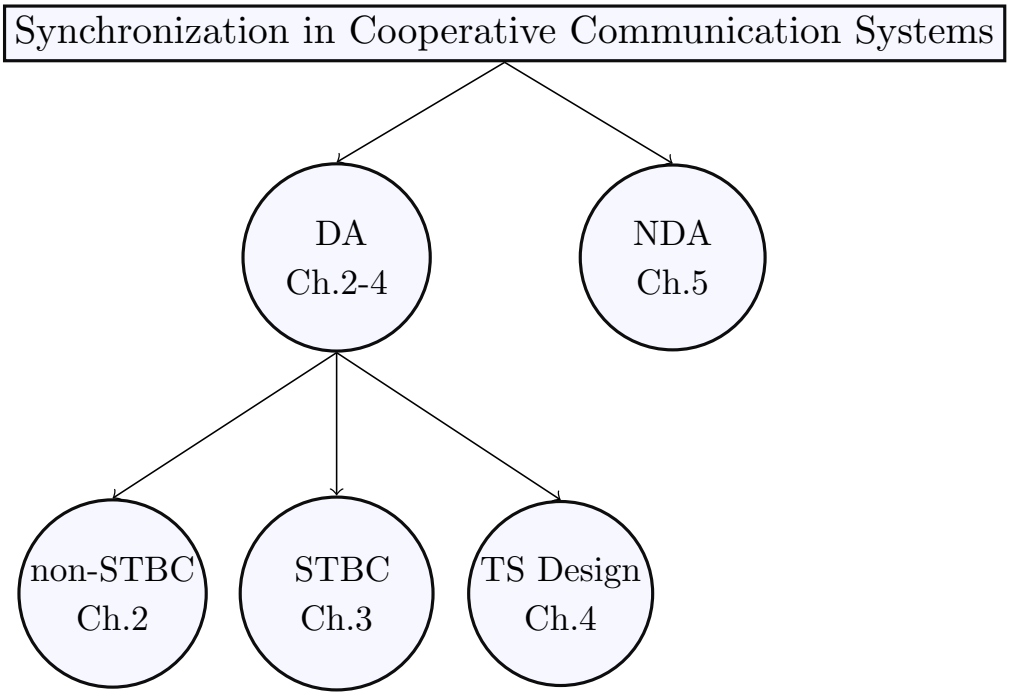


Figure 1.2: Thesis flowchart.

1.2.2 Thesis Contributions and Organization

Fig. 1.2 describes the flowchart of the thesis. The first part of the thesis (Chapter 2-4) discusses the synchronization in DA cooperative communication systems, whereas the second part of the thesis (Chapter 5) addresses the synchronization in NDA cooperative communication systems. The first part of the thesis can be further categorized to non-STBC based cooperative systems (Chapter 2), STBC based systems (Chapter 3), and TS design topics (Chapter 4). The chapter-wise summary of the contributions is given as follows:

Chapter 2 - Data-Aided Synchronization

In Chapter 2, the joint estimation of MCFOs, MTOs, and frequency flat-fading channel gains in both DF and AF multi-relay cooperative networks is addressed. A novel transceiver structure at the relays for achieving synchronization in AF-relaying networks is proposed and the signal model at the relays and destination in the presence of unknown channel gains, timing offsets, and frequency offsets is derived in detail. Next, the estimation problem is parameterized and the closed-form CRLBs for the multiple parameter estimation problem for both DF and AF cooperative networks are derived. A least squares (LS) estimator for joint estimation of MCFOs, MTOs, and channel gains in AF cooperative networks is proposed. In order to reduce the computational complexity associated with the estimation of synchronization parameters and channel gains, new iterative estimators based on

expectation conditional maximization (ECM), and space-alternating generalized expectation-maximization (SAGE) are derived for both DF and AF cooperative networks. Finally, a new ML decoder for detecting the source signal at the destination in the presence of MCFOs and MTOs for both DF and AF cooperative systems is derived. Simulation results show that through the application of the proposed estimators and ML decoder, the performance gains promised by cooperative networks are reachable even in non-ideal settings. The new contributions in Chapter 2 are:

- A novel training and data transmission method for AF cooperative networks in the presence of MCFOs, MTOs, and multiple *unknown* channel gains is proposed.
- New closed-form CRLBs for the multiple parameter estimation problem for both AF and DF cooperative networks are derived and used to assess the performance of the proposed estimators.
- An LS estimator for the estimation of multiple system parameters in AF cooperative networks is derived. Simulation results show that the performance of the proposed LS algorithm and the ML estimator in [40] are close to the CRLB at mid-to-high SNR values. In order to achieve significantly reduced computational complexity, ECM and SAGE algorithms for estimation of multiple system parameters for both AF and DF cooperative systems are derived. Finally, the computational complexity of the proposed estimators is analyzed.
- A new ML detector is derived that allows for the signal received from multiple relays to be successfully decoded in the presence of MCFOs and MTOs.
- Numerical and simulation results are presented, where it is shown that the mean-square error (MSE) performance of ML, and the proposed LS estimators for wide range of signal-to-noise-ratio (SNR) values are close to the derived CRLBs. Moreover, the MSE performance of the proposed ECM and SAGE estimators are evaluated. Finally, the effect of initialization on the performance of the proposed estimators is numerically investigated and specific guidelines that ensure the convergence of these estimators are outlined.

The block level descriptions for the proposed DF and AF asynchronous cooperative networks are shown in Figs. 1.3, and 1.4, respectively. The DF relay in Fig. 1.3 uses the source training signal to estimate source to relay channel, timing

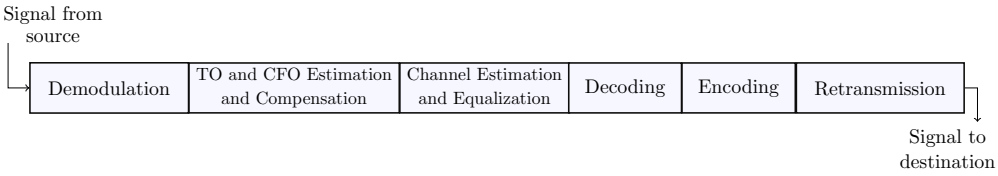


Figure 1.3: Relay block diagram for the DF cooperative systems

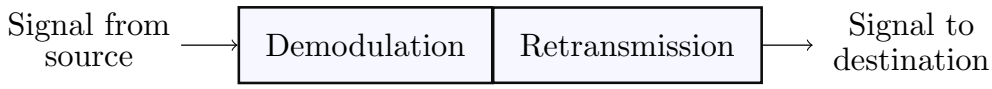


Figure 1.4: Relay block diagram for the AF cooperative systems

and carrier frequency offset. The estimated channel and offsets are used to equalize the channel and compensate single timing and carrier frequency offset from the data. Next, source data signal is decoded, encoded and retransmitted to the destination. Unlike the DF relay, the AF relay in Fig. 1.4 does not estimate the source to relay frequency offset and the channel gain and does not perform matched filtering, frequency offset correction, and equalization. The source data signal is not decoded and directly retransmitted to the destination.

It can be observed from Figs. 1.3 and 1.4 that among the two relaying protocols, the DF relay has the most computational complexity. On the other hand, the BER performance of the DF relaying cooperative network is better than the AF relaying network because of the amplified noise forwarded from the relays to destination in the AF relay [2].

The results in Chapter 2 have been presented in the following publications [3,99], which are listed again for ease of reference:

J1. **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.

C1. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Estimation of Synchronization Parameters in AF Cooperative Networks", in *Proc. IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June 10-15, 2012.

Chapter 3 - Data-Aided Synchronization in STBC based AF Systems

In Chapter 3, a complete synchronization, i.e., joint estimation and compensation of multiple channel gains, MTOs, and MCFOs, in DSTBC-AF relaying cooperative networks is proposed. The processing at each relay's transceiver consists of

estimating and then compensating the source to relay timing and frequency offsets before forwarding the linearly processed signals to the destination using DSTBC. At the destination, the channel gains, MTOs, and MCFOs are jointly estimated using transmitted training sequences and the proposed least squares (LS) or differential evolution (DE) based estimators. The DE algorithm, which is an iterative method, significantly reduces computational complexity at the destination. Next, a novel minimum mean-square error (MMSE) receiver for compensating for the effect of these impairments at the destination is derived. Note that Chapter 2 addresses the problem of timing and carrier synchronization in AF and DF cooperative networks. However, the transceiver designs and estimation algorithms proposed in Chapter 2 cannot be applied to DSTBC-AF cooperative networks due to the particular processing required at the relays and destination to enable transmission and decoding of DSTBC in AF cooperative networks. Further, the main focus of Chapter 2 is on the design of estimation algorithms for joint estimation of multiple impairments. The design of computationally efficient compensation algorithms is addressed in Chapter 3. In summary, the main contributions of Chapter 3 are as follows:

- A new transceiver structure at the relays and a receiver design at the destination for achieving synchronization in DSTBC-AF relaying cooperative network is proposed.

- New Cramér-Rao lower bounds (CRLBs) for the joint estimation of multiple system parameters at the destination are derived. Aside from benchmarking the performance of the proposed estimators, these bounds are applied to determine the statistics of the estimation errors for timing offset, frequency offset, and channels, which is in turn used in the design of the MMSE receivers at the destination.

- An LS estimator for the joint estimation of channels, MTOs, MCFOs is formulated. Next, in order to significantly reduce the computational complexity associated with the joint estimation problem, a DE based algorithm is applied. A parameterization of the proposed DE estimator is provided that achieves fast convergence. Simulation results show that the mean square error (MSE) performances of both the LS and the DE estimators are close to the CRLB at moderate-to-high SNRs.

- An MMSE receiver for compensating the effect of impairments and detecting the signal from the relays at the destination is derived. In order to reduce overhead, a low complexity version denoted by *L-MMSE* is also proposed that

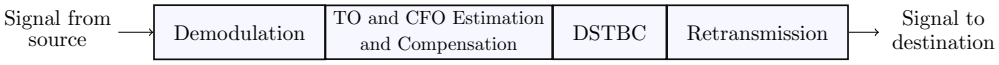


Figure 1.5: Relay block diagram for the DSTBC based AF cooperative systems

can detect the received signals using only the overall channel gains, MTOs, and MCFOs estimates obtained at the destination.

- Extensive simulations are carried out to investigate the performance of the overall proposed transceiver design in DSTBC-AF cooperative networks for different numbers of relays and system design parameters. It is shown that the proposed DSTBC-AF relaying system achieves full spatial diversity gain for 2 and 4 relay networks and outperforms existing cooperative strategies in the presence of MTOs, MCFOs, and unknown channels.

The block level descriptions for the proposed DSTBC based AF relaying in asynchronous cooperative network is shown in Fig. 1.5. The DSTBC based AF relay though does not perform channel estimation and equalization as the simple AF relay in Fig. 1.4 , but has to first estimate and compensate the source to relay timing and frequency offsets before applying DSTBC. Finally, similar to the simple AF relay, the DSTBC based AF relay directly retransmits the source data to the destination without decoding it. Note that the DSTBC based DF relay is not included in the above figure because it is similar to the simple DF relay with an additional DSTBC block after encoding.

Comparing the two AF relays in Figs. 1.4 and 1.5, though a less complex AF transceiver design is used in the simple AF relay in Fig. 1.4, it does not allow for the application of DSTBCs and as a result, sacrifices spatial diversity to achieve lower complexity. On the other hand, the DSTBC based AF transceiver design in Fig. 1.5 achieves full diversity gain while requiring additional synchronization overhead at the relays to allow for the application of DSTBCs.

The results in Chapter 3 have been presented in the following publication [100], which is listed again for ease of reference:

J2. **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.*, 2013. (accepted 20 Feb, 2013)

Chapter 4 - Training Sequence Design

Chapter 2 and Chapter 3 focus on the design of estimation and compensation algorithms for the joint estimation of multiple impairments in DA cooperative

communication systems. However, considering DA or training-based transmission, the design of optimal training sequences for the efficient estimation of multiple impairments is also critical.

In Chapter 4, we address the problem of designing optimal TSs for the joint estimation of MTOs and channels in distributed multi-user or cooperative communication networks. The effect of CFO on the received signal is not considered, since the topic of TS design for MCFO estimation has been extensively addressed in the literature, e.g., see [101] and references therein. The Hybrid Cramér-Rao lower bound (HCRLB) for joint estimation of MTOs and channel parameters is derived. By minimizing the HCRLB with respect to the TSs, three criteria for the design of optimal TSs is proposed. The use of the HCRLB results in new guidelines with broader applications since, unlike the CRLB, the HCRLB is not dependant on specific channel realizations. Moreover, in order to show that the proposed TSs improve MTO and channel estimation accuracy, the conditional Cramér-Rao lower bound (ECRLB), which is a tighter bound on the mean-square error (MSE) of estimators, and a new maximum a posteriori (MAP) estimator are derived. The main contributions of Chapter 4 can be summarized as follows:

- New HCRLBs, ECRLBs, and a MAP estimator for the joint estimation of multiple channels and MTOs are derived. The derivation of the ECRLB is motivated by the fact that it serves as a tight lower bound on the estimation accuracy of the MAP estimator.

- By minimizing the derived HCRLB for the joint estimation of the multiple channels and MTOs, three criteria for the design of optimal TSs are formulated. It is shown that the proposed guidelines not only lower the HCRLB, but also lower the ECRLB and the MSE of the MAP estimator.

- Based on the proposed TS design criteria, two TSs are proposed and it is demonstrated that these TSs result in a lower MSE for the joint estimation of the MTO and channel parameters when compared to TSs that violate the proposed guidelines. Moreover, simulations show that application of the proposed TSs in a multi-relay cooperative network significantly enhances the network's BER performance.

The results in Chapter 4 have been presented in the following publication [102], which is listed below for ease of reference:

J3. **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," submitted to *IEEE Trans. Commun.*, July. 2012. (manuscript no. TCOM-TPS-12-0541, revised 9 Feb, 2013)

Chapter 5 - Non-Data Aided Synchronization

Chapters 2-4 address synchronization for DA cooperative systems and assume the transmission of training sequences to assist the joint estimation of multiple channels, MTOs, and MCFOs. Chapter 5 focuses on blind synchronization and channel estimation in DF cooperative communication systems, where no training signals are used. For the broadcasting phase, the synchronization problem is identical to the single input single output (SISO) and we choose existing methods, available in the literature, for the blind synchronization and channel estimation. The important open research problem of joint synchronization and channel estimation lies in the relaying phase and is addressed in Chapter 5. The main contributions of Chapter 5 can be summarized as follows:

- A novel transceiver structure at the relays and a complete receiver design at the destination for achieving synchronization in blind DF relaying cooperative network is proposed.
- We propose to exploit blind source separation at the destination to convert the difficult problem of jointly estimating multiple synchronization parameters in the relaying phase into more tractable sub-problems of estimating many individual timing and carrier offsets for the independent relay. This allows the timing and carrier offsets to be estimated using the same algorithms as in the broadcasting phase. In order to implement blind source separation, we deploy multiple antennas at the destination, which are greater than or equal to the number of relays.
- We propose a novel relay selection algorithm at the destination. Simulation results demonstrate the excellent end-to-end BER performance of the proposed blind scheme in block Rayleigh flat-fading channels and show that it is able to achieve the maximum diversity order at the destination.

The results in Chapter 5 have been presented in the following publications [103, 104] which are listed below for ease of reference:

J4. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Blind Timing and Carrier Synchronization in Distributed MIMO Communication Systems," in *IET Commun.*, vol. 5, no. 7, pp. 1028-1037, May 2011.

C2. **A. A. Nasir**, S. Durrani and R. A. Kennedy, "Blind Timing and Carrier Synchronization in Decode and Forward Cooperative Systems", in *Proc. IEEE International Conference on Communications (ICC)*, Kyoto, Japan, June 5-9, 2011.

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

Chapter 2

Data-Aided Synchronization

This chapter seeks to develop estimation and detection algorithms that enable cooperative communications for both *decode-and-forward* (DF) and *amplify-and-forward* (AF) relaying networks in the presence of MCFOs, MTOs, and unknown channel gains. Chapter 2 is organized as follows.

In Section 2.1, a novel training and data transmission method for AF cooperative networks in the presence of MCFOs, MTOs, and multiple *unknown* channel gains is proposed. In Section 2.2, new closed-form CRLBs for the multiple parameter estimation problem for both AF and DF cooperative networks are derived and used to assess the performance of the proposed estimators. In Section 2.3, an LS estimator for the estimation of multiple system parameters in AF cooperative networks is derived. Simulation results show that the performance of the proposed LS algorithm and the ML estimator in [40] are close to the CRLB at mid-to-high SNR values. In order to achieve significantly reduced computational complexity, ECM and SAGE algorithms for estimation of multiple system parameters for both AF and DF cooperative systems are derived. Finally, the computational complexity of the proposed estimators is analyzed. In Section 2.4, a new ML detector is derived that allows for the signal received from multiple relays to be successfully decoded in the presence of MCFOs and MTOs. In Section 2.5, numerical and simulation results are presented, where it is shown that the *mean-square error* (MSE) performance of ML, and the proposed LS estimators for wide range of *signal-to-noise-ratio* (SNR) values are close to the derived CRLBs. Moreover, the MSE performance of the proposed ECM and SAGE estimators are evaluated. Finally, the effect of initialization on the performance of the proposed estimators is numerically investigated and specific guidelines that ensure the convergence of these estimators are outlined. Finally, Section 2.6 concludes the chapter.

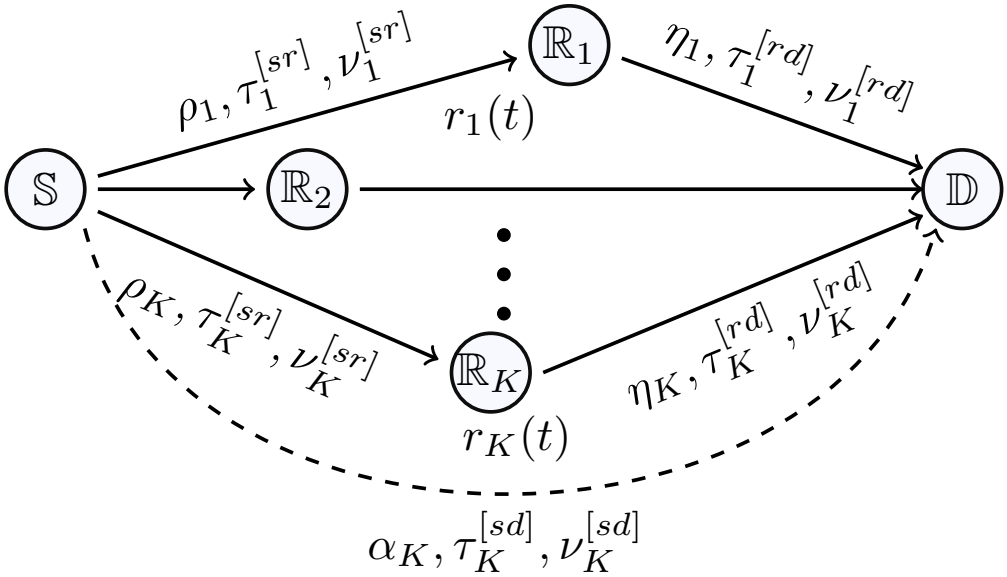


Figure 2.1: The system model for the cooperative network.

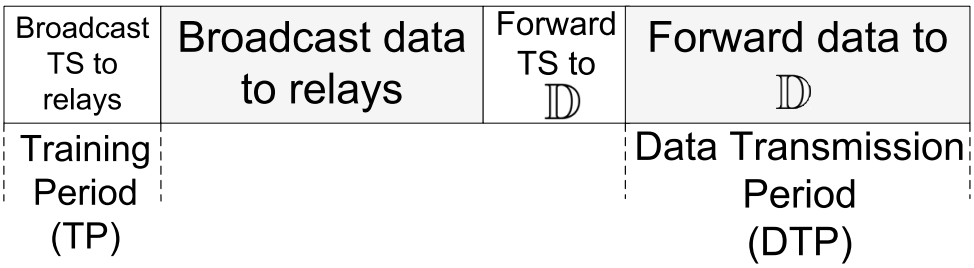


Figure 2.2: Scheduling diagram for the training and data transmission period.

2.1 System Model

We consider a half-duplex *space division multiple access* (SDMA) SISO cooperative communication system with one source node, $\mathbb{S}$, K relays, $\mathbb{R}_1, \cdots, \mathbb{R}_K$, and a single destination node, $\mathbb{D}$, as shown in Fig. 2.1. The channel gains from $\mathbb{S}$ to $\mathbb{R}_k$, $\mathbb{R}_k$ to $\mathbb{D}$, and $\mathbb{S}$-$\mathbb{R}_k$-$\mathbb{D}$ are denoted by ρ_k, η_k, and α_k, respectively, for $k = \{1, \cdots, K\}$. In Fig. 2.1, τ_k and ν_k are used to denote the timing offsets and CFOs, where superscripts $(\cdot)^{[sr]}$, $(\cdot)^{[rd]}$ and $(\cdot)^{[sd]}$ denote the offsets from $\mathbb{S}$ to $\mathbb{R}_k$, $\mathbb{R}_k$ to $\mathbb{D}$, and $\mathbb{S}$ to $\mathbb{D}$, respectively. Throughout this chapter, the following set of system design assumptions is considered:

A1. As shown in Fig. 2.2, each transmission frame from the source to relays and the relays to destination is comprised of two periods: a *training period (TP)* followed by a *data transmission period (DTP)*. Without loss of generality, it is

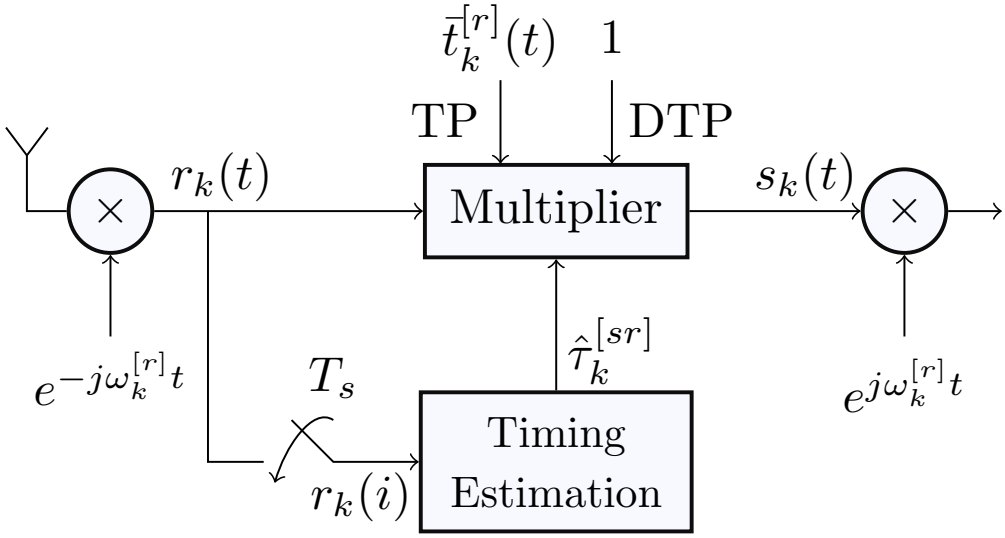


Figure 2.3: Block Diagram for the AF k^{th} Relay Transceiver.

assumed that during the TP, unit-amplitude phase shift keying (PSK) *training signals (TSs)* are transmitted from source to the kth relay and from the kth relay to the destination, $\forall\ k$.

A2. Quasi-static and frequency flat fading channels are considered, i.e., the channel gains do not change over the length of a frame but change from frame to frame according to a complex Gaussian distribution, $\mathcal{CN}(0, \sigma^2)$. The use of such channels is motivated by prior research in this field [1, 2, 12, 21–25].

A3. Over a frame, the timing and frequency offsets, as shown in Fig. 3.1, are modeled as deterministic but *unknown* parameters.

2.1.1 AF-Relaying Cooperative Network

The block diagrams for the AF transceiver at $\mathbb{R}_k$ and AF receiver at $\mathbb{D}$ are depicted in Figs. 2.3 and 2.4, respectively. The proposed training and data transmission methods for the TP and DTP are outlined in the following two subsections.

2.1.1.1 Training Period

The received signal at $\mathbb{R}_k$ is down converted by oscillator frequency, $\omega_k^{[r]}$, and then over sampled by the factor Q. The sampled received signal at the input of the

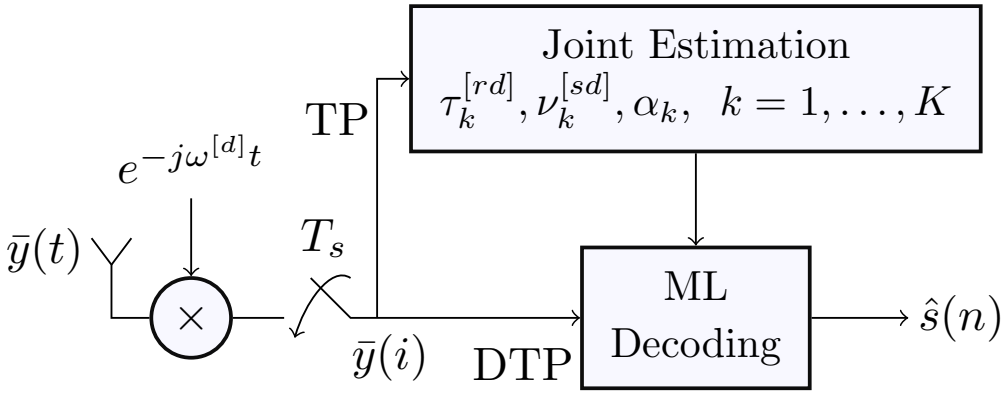


Figure 2.4: Block Diagram for the AF Destination Receiver.

timing estimation block, $r_k(i)$ is given by[1]

$$r_k(i) = \rho_k \sum_{n=0}^{L-1} t^{[s]}(n) g\left(iT_s - nT - \tau_k^{[sr]}T\right) e^{j2\pi i \nu_k^{[sr]}/Q} + u_k(i), \tag{2.1}$$

where

- ρ_k denotes the *unknown* channel gain from $\mathbb{S}$ to $\mathbb{R}_k$ that is assumed to be constant over a frame and distributed as $\mathcal{CN}(0, \sigma_\rho^2)$ from frame to frame,
- $t^{[s]}(n)$ is the source training signal (TS) of length L,
- $g(t)$ is the transmitter pulse shaping function, $\tau_k^{[sr]}$ is the normalized[2] fractional unknown timing offset of the sampler at $\mathbb{R}_k$, T_s is the sampling time period such that $T_s = T/Q$,
- $\nu_k^{[sr]}$ is the carrier frequency offset, normalized by the symbol duration T, between $\mathbb{S}$ and $\mathbb{R}_k$, and
- $u_k(i)$ is the zero-mean complex baseband *additive white Gaussian noise* (AWGN) at $\mathbb{R}_k$ with variance $\sigma_{u_k}^2$, i.e., $u_k(i) \sim \mathcal{CN}(0, \sigma_{u_k}^2)$.

Without loss of generality, it is assumed that the noise at all the relays has the same variance, i.e., $\sigma_u^2 = \sigma_{u_1}^2 = \cdots = \sigma_{u_K}^2$.

In order to ensure synchronous transmission and successful cooperation for AF networks, a timing detector at the k^{th} relay estimates the corresponding timing offset, $\hat{\tau}_k^{[sr]}$, using schemes available for point-to-point SISO systems as outlined

[1]For clarity, we reserve the index $n = \{0, \cdots, L-1\}$ for T-spaced samples and index $i = \{0, \cdots, QL-1\}$ for T_s-spaced samples.

[2]Throughout this section both the timing and carrier frequency offsets are normalized by T.

in [7]. The timing offset estimate $\hat{\tau}_k^{[sr]}$ is used as an input to the multiplier to ensure that the k^{th} relay's unit amplitude training signal, $\bar{t}_k^{[r]}(t)$,[3] is multiplied by the received signal $r_k(t)$ at the appropriate time. The training signal $\bar{t}_k^{[r]}(t)$ used for AF relaying here is given by $\bar{t}_k^{[r]}(t) = e^{-j\phi_k(n)}$ for $(n-1)T < t < nT$, where $\phi_k(n)$ in $[-\pi, \pi]$ and denotes the phase of the n^{th} symbol of the k^{th} relay's training signal, where $\phi_k(n) \neq \phi_{\bar{k}}(n)$, for $k \neq \bar{k}$. The output of the analog multiplier, $s_k(t)$, as shown in Fig. 2.3, is given by

$$s_k(t) = t_k^{[r]}(t)\rho_k e^{j2\pi\tilde{\nu}_k^{[sr]}t} \sum_{n=0}^{L-1} t^{[s]}(n) g\left(t - nT - \epsilon_k^{[sr]}T\right) + \bar{t}_k^{[r]}(t)u_k(t), \tag{2.2}$$

where $\epsilon_k^{[sr]} = \tau_k^{[sr]} - \hat{\tau}_k^{[sr]}$ is the timing estimation error between $\mathbb{S}$ and $\mathbb{R}_k$ and $\tilde{\nu}_k^{[sr]} = \nu_k^{[sr]}/T$ is the analog frequency offset between $\mathbb{S}$ and $\mathbb{R}_k$. The received signal at $\mathbb{D}$ for AF relaying, $\bar{y}(t)$, is affected by the timing offset from the k^{th} relay to the destination, $\tau_k^{[rd]}$, for $k = 1 \cdots, K$. Thus, the sampled received signal, $\bar{y}(i)$, is given by

$$\bar{y}(i) = \sum_{k=1}^{K} \zeta_k \eta_k s_k(iT_s - \tau_k^{[rd]}T) e^{j2\pi i\nu_k^{[rd]}/Q} + w(i)$$

$$\begin{aligned}\bar{y}(i) = &\sum_{k=1}^{K} \zeta_k \eta_k \rho_k \bar{t}_k^{[r]}(iT_s - \tau_k^{[rd]}T) e^{j2\pi i\nu_k^{[sd]}/Q} \sum_{n=0}^{L-1} t^{[s]}(n) g\left(iT_s - nT - \tau_k^{[sd]}T\right) \\ &+ \sum_{k=1}^{K} \zeta_k \eta_k u_k(i) \bar{t}_k^{[r]}(iT_s - \tau_k^{[rd]}T) e^{j2\pi i\nu_k^{[rd]}/Q} + w(i),\end{aligned} \tag{2.3}$$

where

- η_k denotes the complex *unknown* channel gain from $\mathbb{R}_k$ to $\mathbb{D}$ that is assumed to be distributed as $\mathcal{CN}(0, \sigma_\eta^2)$ from frame to frame,
- $\zeta_k = 1/\sqrt{\sigma_\rho^2 + \sigma_u^2}$ satisfies the k^{th} relay's power constraint,
- $\nu_k^{[sd]} \triangleq \nu_k^{[sr]} + \nu_k^{[rd]}$ is the sum of CFOs from $\mathbb{S}$–$\mathbb{R}_k$–$\mathbb{D}$, $\nu_k^{[rd]}$ is the normalized CFO from $\mathbb{R}_k$ to $\mathbb{D}$[4],
- $\tau_k^{[sd]} \triangleq \tau_k^{[rd]} + \epsilon_k^{[sr]}$, and

[3]Note that the bar operator, $\bar{\cdot}$ over the symbols, is used for the symbols in AF cooperative relaying systems in order to distinguish the symbols for AF and DF cooperative systems.

[4]The constant phase offset $e^{j2\pi\nu_k^{[sr]}\tau_k^{[rd]}}$ due to the timing offset $\tau_k^{[rd]}$ is incorporated in the baseband channel η_k.

- $w(i) \sim \mathcal{CN}(0, \sigma_w^2)$ is the AWGN at $\mathbb{D}$.

Note that $u_k(i)$ has been used in place of $u_k(iT_s - \tau_k^{[rd]}T)$ since $u_k(t)$ denotes the AWGN and its statistics are not affected by the change in the sampling point. Eq. (2.3) can be written in vector form as

$$\bar{\mathbf{y}} = \underbrace{\left[(\bar{\mathbf{\Lambda}}_1\bar{\mathbf{G}}_1\mathbf{t}^{[s]}) \odot \bar{\mathbf{t}}_1^{[r]}(\tau_k^{[rd]}), \cdots, (\bar{\mathbf{\Lambda}}_K\bar{\mathbf{G}}_K\mathbf{t}^{[s]}) \odot \bar{\mathbf{t}}_K^{[r]}(\tau_k^{[rd]})\right]}_{\triangleq\, \bar{\mathbf{\Omega}}} \boldsymbol{\alpha} + \underbrace{[\mathbf{\Lambda}_1\mathbf{v}_1, \cdots, \mathbf{\Lambda}_K\mathbf{v}_K]}_{\triangleq\, \bar{\mathbf{\Psi}}}\boldsymbol{\beta} + \mathbf{w}, \quad (2.4)$$

where

- $\bar{\mathbf{y}} \triangleq [\bar{y}(0), \bar{y}(1), \cdots, \bar{y}(QL-1)]^T$ is the received signal vector at $\mathbb{D}$,
- $\boldsymbol{\alpha} \triangleq [\alpha_1, \cdots, \alpha_K]^T$ is the overall channel gains from source to relay to destination, $\alpha_k \triangleq \zeta_k\eta_k\rho_k$,
- $\boldsymbol{\beta} \triangleq [\beta_1, \cdots, \beta_K]^T$, $\beta_k \triangleq \zeta_k\eta_k$,
- $\bar{\mathbf{t}}_k^{[r]}(\tau_k^{[rd]}) \triangleq \left[\bar{t}_k^{[r]}(\tau_k^{[rd]}T), \bar{t}_k^{[r]}(Ts - \tau_k^{[rd]}T) \cdots, \bar{t}_k^{[r]}((QL-1)T_s - \tau_k^{[rd]}T)\right]^T$ is the k^{th} relay's training signal,
- $\bar{\mathbf{\Lambda}}_k \triangleq \text{diag}\left([e^{j2\pi\nu_k^{[sd]}(0)/Q}, \cdots, e^{j2\pi\nu_k^{[sd]}(QL-1)/Q}]\right)$ is a $QL \times QL$ matrix depending on the frequency offset, $\nu_k^{[sd]}$,
- $\mathbf{\Lambda}_k \triangleq \text{diag}\left([e^{j2\pi\nu_k^{[rd]}(0)/Q}, \cdots, e^{j2\pi\nu_k^{[rd]}(QL-1)/Q}]\right)$ is a $QL \times QL$ matrix depending on the frequency offset, $\nu_k^{[sd]}$,
- $\left[\bar{\mathbf{G}}_k\right]_{m,\ell} \triangleq g\left(mT_s - \ell T - \tau_k^{[sd]}T\right)$ is a $QL \times L$ matrix depending on the timing offset, $\tau_k^{[sd]}$,
- $\mathbf{t}^{[s]} \triangleq \left[t^{[s]}(0), \cdots, t^{[s]}(L-1)\right]^T$ is the source training signal vector,
- $\mathbf{w} \triangleq [w(0), \cdots, w(QL-1)]^T$ is the AWGN vector at $\mathbb{D}$,
- $\mathbf{v}_k \triangleq [v_k(0), \cdots, v_k(QL-1)]^T$ is the forwarded AWGN from the relays to destination and $v_k(i) \triangleq u_k(i)\bar{t}_k^{[r]}\left(iT_s - \tau_k^{[rd]}T\right)$.

Note that $v_k(i)$ has the same statistical properties as $u_k(i)$, for $i = \{0, \cdots, QL-1\}$, due to the assumption of the unit-amplitude training signals.

Remark 2.1 *Unlike [1], the proposed processing structure at the relays in Fig. 2.3 is not based on the assumption of perfect timing and frequency offset estimation and matched-filtering at the relays [105]. In fact, [105] shows that the signal model in [1] does not resemble that of AF relaying. More importantly, in a follow up paper, the authors of [1] confirm that further research is needed to enable perfect timing synchronization in practical AF-relaying cooperative networks [106]. In order to address this issue, we have proposed a novel relaying structure that does not require perfect timing offset estimation and matched-filtering at the relays. In addition, in our proposed model, the relays do not perform frequency offset and channel estimation during the TP.*

2.1.1.2 Data Transmission Period

Modulated data symbol vector $\mathbf{s} \triangleq [s(0), s(1), \cdots, s(L-1)]^T$ is transmitted from $\mathbb{S}$ to the relays. As shown in Fig. 2.3, after performing timing correction using timing offset estimates obtained in the TP, $\mathbb{R}_k$ forwards the received signal to $\mathbb{D}$. The received signal at $\mathbb{D}$ in the DTP can be written as[5]

$$\bar{\mathbf{y}} = \sum_{k=1}^{K} \alpha_k \bar{\mathbf{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{s} + \sum_{k=1}^{K} \beta_k \mathbf{\Lambda}_k \mathbf{u}_k + \mathbf{w}, \tag{2.5}$$

where $\mathbf{u}_k \triangleq [u_k(0), \cdots, u_k(QL-1)]^T$. Fig. 2.4 summarizes the proposed transceiver structure at $\mathbb{D}$ for AF relaying.

2.1.2 DF-Relaying Cooperative Network

The block diagram for the DF transceiver at $\mathbb{R}_k$ is shown in Fig. 2.5. The receiver structure at $\mathbb{D}$ is similar to that of AF relaying in Fig. 2.4 with the exception that the received signal is represented by $y(t)$ instead of $\bar{y}(t)$.

2.1.2.1 Training Period

Similar to AF, the sampled received signal at $\mathbb{R}_k$, $r_k(i)$, is given by (2.1). However, in the case of DF relaying the received signal at $\mathbb{R}_k$ needs to be decoded. Therefore, during the TP, $\hat{\rho}_k$, $\hat{\tau}_k^{[sr]}$, and $\hat{\nu}_k^{[sr]}$ are jointly estimated and compensated using conventional schemes for SISO point-to-point systems [7]. Subsequently, $\mathbb{R}_k$ forwards its known distinct TS, $\mathbf{t}_k^{[r]} \triangleq \left[t_k^{[r]}(0), \cdots, t_k^{[r]}(L-1)\right]^T$ to $\mathbb{D}$. The received signal at

[5]The matrix $\bar{\mathbf{G}}_k$ takes into account the timing offset estimation error from $\mathbb{S}$ to $\mathbb{R}_k$ and the timing offset from $\mathbb{R}_k$ to $\mathbb{D}$.

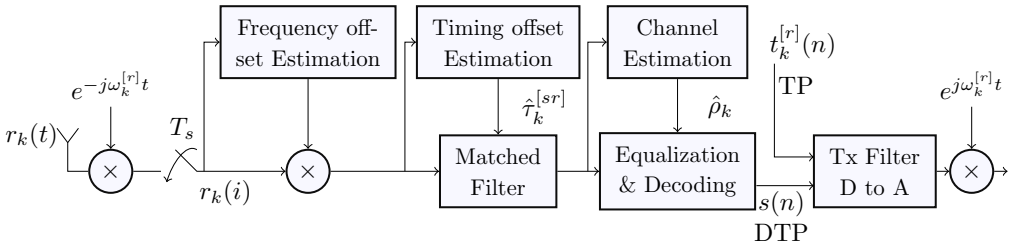


Figure 2.5: Block diagram for the DF transceiver at the k^{th} relay.

$\mathbb{D}$, $\mathbf{y} \triangleq [y(0), \cdots, y(QL-1)]^T$, is given by

$$\mathbf{y} = \underbrace{\left[\mathbf{\Lambda}_1 \mathbf{G}_1 \mathbf{t}_1^{[r]}, \cdots, \mathbf{\Lambda}_K \mathbf{G}_K \mathbf{t}_K^{[r]}\right]}_{\triangleq \mathbf{\Omega}} \boldsymbol{\eta} + \mathbf{w}, \tag{2.6}$$

where $\boldsymbol{\eta} \triangleq [\eta_1, \cdots, \eta_K]^T$, $\tau_k^{[rd]}$, is the normalized fractional unknown timing offset between $\mathbb{R}_k$ and $\mathbb{D}$, $\mathbf{\Lambda}_k$ is defined below (2.4), and $[\mathbf{G}_k]_{m,\ell} \triangleq g\left(mT_s - \ell T - \tau_k^{[rd]} T\right)$ is a $QL \times L$ matrix. Note that the notation, $\mathbf{t}_k^{[r]}$ is used instead of $\bar{\mathbf{t}}_k^{[r]}$ to distinguish between the TSs for the DF and AF relaying.

2.1.2.2 Data Transmission Period

For DF relaying, it is assumed that *cyclic redundancy checks* (CRC) are applied at the relays to ensure that the relays only forward correctly decoded signals to $\mathbb{D}$. Subsequently, the received signal at $\mathbb{D}$ in the DTP is given by

$$\mathbf{y} = \sum_{k=1}^{K} \eta_k \mathbf{\Lambda}_k \mathbf{G}_k \mathbf{s} + \mathbf{w}. \tag{2.7}$$

The receiver structure at $\mathbb{D}$ is given by Fig. 2.4 with the exception that $\tau_k^{[rd]}$, $\nu_k^{[rd]}$, and η_k are replaced by $\tau_k^{[sd]}$, $\nu_k^{[sd]}$, and α_k, for $k = \{1, \cdots, K\}$, respectively.

2.2 Cramer-Rao Lower Bound

In this section, new exact closed-form CRLBs for the joint estimation of multiple channel gains, MCFOs, and MTOs for AF-relaying cooperative networks are derived. For the case of DF relaying, we extend the results in [40] and present closed-form CRLBs for this multiple parameter estimation problem.

2.2.1 CRLB for AF Relaying

Throughout this section it is assumed that the forwarded AWGN from the relays, $\mathbb{R}_k$ and $\mathbb{R}_m$, $\mathbf{v}_k$ and $\mathbf{v}_m$, respectively, $\forall k \neq m$, and the AWGN at $\mathbb{D}$, $\mathbf{w}$, are mutually independent. Accordingly, the received signal vector at $\mathbb{D}$, $\bar{\mathbf{y}}$ in (2.4), is a circularly symmetric complex Gaussian random variable, i.e., $\bar{\mathbf{y}} \sim \mathcal{CN}(\boldsymbol{\mu}_{\bar{\mathbf{y}}}, \boldsymbol{\Sigma}_{\bar{\mathbf{y}}})$, with mean $\boldsymbol{\mu}_{\bar{\mathbf{y}}} = \bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}$ and covariance matrix $\boldsymbol{\Sigma}_{\bar{\mathbf{y}}} = \left(\sigma_u^2 \sum_{k=1}^{K} |\beta_k|^2 + \sigma_w^2\right) \mathbf{I}_{LQ}$, where $\mathbf{I}_{LQ}$ is the identity matrix of size $LQ \times LQ$ ($\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}$ is derived in Appendix A.1). As discussed in [2] and [39], in the case of AF relaying, only the overall channel gains, $\boldsymbol{\alpha}$, need to be estimated. As a result, the parameter vector of interest for AF relaying, $\bar{\boldsymbol{\theta}}$, is given by

$$\bar{\boldsymbol{\theta}} \triangleq \left[\Re\{\boldsymbol{\alpha}\}^T, \Im\{\boldsymbol{\alpha}\}^T, \bar{\boldsymbol{\nu}}^T, \bar{\boldsymbol{\tau}}^T\right]^T, \tag{2.8}$$

where $\bar{\boldsymbol{\nu}} \triangleq [\bar{\nu}_1, \cdots, \bar{\nu}_K]^T$ and $\bar{\boldsymbol{\tau}} \triangleq [\bar{\tau}_1, \cdots, \bar{\tau}_K]^T$ and for notational simplicity, $\bar{\nu}_k$ and $\bar{\tau}_k$ are used to denote $\nu_k^{[sd]}$ and $\tau_k^{[rd]}$ for AF relaying, respectively.

Theorem 2.1 *Based on the proposed training method, the Fisher's information matrix (FIM) for the estimation of $\bar{\boldsymbol{\theta}}$ is given by*

$$\mathbf{F}^{AF} = \frac{2}{\sigma_n^2} \underbrace{\begin{bmatrix} \Re\{\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}}\} & -\Im\{\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}}\} & -\Im\{\bar{\boldsymbol{\Omega}}^H\mathbf{D}\bar{\boldsymbol{\Omega}}\bar{\mathbf{H}}\} & \Re\{\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Gamma}}\bar{\mathbf{H}}\} \\ \Im\{\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}}\} & \Re\{\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}}\} & \Re\{\bar{\boldsymbol{\Omega}}^H\mathbf{D}\bar{\boldsymbol{\Omega}}\bar{\mathbf{H}}\} & \Im\{\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Gamma}}\bar{\mathbf{H}}\} \\ \Im\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Omega}}^H\mathbf{D}\bar{\boldsymbol{\Omega}}\} & \Re\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Omega}}^H\mathbf{D}\bar{\boldsymbol{\Omega}}\} & \underbrace{\Re\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Omega}}^H\mathbf{D}^2\bar{\boldsymbol{\Omega}}\bar{\mathbf{H}}\}}_{\triangleq\bar{\boldsymbol{\Phi}}_{11}} & \underbrace{\Im\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Omega}}^H\mathbf{D}\bar{\boldsymbol{\Gamma}}\bar{\mathbf{H}}\}}_{\triangleq\bar{\boldsymbol{\Phi}}_{12}} \\ \Re\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Gamma}}^H\bar{\boldsymbol{\Omega}}\} & -\Im\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Gamma}}^H\bar{\boldsymbol{\Omega}}\} & \underbrace{-\Im\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Gamma}}^H\mathbf{D}\bar{\boldsymbol{\Omega}}\bar{\mathbf{H}}\}}_{\triangleq\bar{\boldsymbol{\Phi}}_{21}} & \underbrace{\Re\{\bar{\mathbf{H}}^H\bar{\boldsymbol{\Gamma}}^H\bar{\boldsymbol{\Gamma}}\bar{\mathbf{H}}\}}_{\triangleq\bar{\boldsymbol{\Phi}}_{22}} \end{bmatrix}}_{\triangleq\bar{\boldsymbol{\mathcal{F}}}}, \tag{2.9}$$

where $\sigma_n^2 \triangleq \sigma_u^2 \sum_{k=1}^{K} |\beta_k|^2 + \sigma_w^2$, and $\bar{\boldsymbol{\Gamma}} \triangleq [(\bar{\boldsymbol{\Lambda}}_1\bar{\mathbf{R}}_1\mathbf{t}^{[s]}) \odot \bar{\mathbf{t}}_1^{[r]}(\tau_k^{[rd]}), \cdots, (\bar{\boldsymbol{\Lambda}}_K\bar{\mathbf{R}}_K\mathbf{t}^{[s]}) \odot \bar{\mathbf{t}}_K^{[r]}(\tau_k^{[rd]})]$, $\mathbf{D} \triangleq 2\pi/Q \times \mathrm{diag}\big([0, 1, \cdots, LQ-1]\big)$, $\bar{\mathbf{H}} \triangleq \mathrm{diag}(\alpha_1, \cdots, \alpha_K)$, and $\bar{\mathbf{R}}_k \triangleq \partial\bar{\mathbf{G}}_k/\partial\bar{\tau}_k$.

Proof: See Appendix A.1.

Let us rewrite (2.9) as

$$\mathbf{F}^{\mathrm{AF}} \triangleq \frac{2}{\sigma_n^2} \begin{bmatrix} \bar{\mathbf{F}}_{11} & \bar{\mathbf{Z}} \\ \bar{\mathbf{Z}}^T & \bar{\mathbf{F}}_{22,} \end{bmatrix}, \tag{2.10}$$

where $\bar{\mathbf{F}}_{11}$ and $\bar{\mathbf{F}}_{22}$ are the upper left and lower right $2K \times 2K$ sub matrices of $\bar{\boldsymbol{\mathcal{F}}}$, respectively, and $\bar{\mathbf{Z}}$ is the upper right $2K \times 2K$ sub matrix of $\bar{\boldsymbol{\mathcal{F}}}$. Using par-

titioned matrix inverse [2, 107], the closed-form CRLB for the estimation of $\bar{\boldsymbol{\theta}}$ can be determined as

$$\begin{aligned}\mathbf{CRLB}(\bar{\boldsymbol{\theta}}) =& \frac{\sigma_n^2}{2}\left(\begin{bmatrix} \bar{\mathbf{F}}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\bar{\mathbf{F}}_{11}^{-1}\bar{\mathbf{Z}} \\ \mathbf{I} \end{bmatrix} \underbrace{\left[\bar{\mathbf{F}}_{22} - \bar{\mathbf{Z}}^T\bar{\mathbf{F}}_{11}^{-1}\bar{\mathbf{Z}}\right]^{-1}}_{\triangleq\, \bar{\boldsymbol{\Upsilon}}} \begin{bmatrix} -\bar{\mathbf{Z}}^T\bar{\mathbf{F}}_{11}^{-1} & \mathbf{I} \end{bmatrix}\right) \\ =& \frac{\sigma_n^2}{2}\begin{bmatrix} \bar{\mathbf{F}}_{11}^{-1} + \bar{\mathbf{F}}_{11}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Upsilon}}\bar{\mathbf{Z}}^T\bar{\mathbf{F}}_{11}^{-1} & -\bar{\mathbf{F}}_{11}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Upsilon}} \\ -\bar{\boldsymbol{\Upsilon}}\bar{\mathbf{Z}}^T\bar{\mathbf{F}}_{11}^{-1} & \bar{\boldsymbol{\Upsilon}} \end{bmatrix}. \end{aligned} \tag{2.11}$$

Similarly, $\bar{\mathbf{F}}_{11}^{-1}$ and $\bar{\mathbf{F}}_{22}^{-1}$ are found as

$$\bar{\mathbf{F}}_{11}^{-1} = \begin{bmatrix} \Re\{(\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}})^{-1}\} & -\Im\{(\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}})^{-1}\} \\ \Im\{(\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}})^{-1}\} & \Re\{(\bar{\boldsymbol{\Omega}}^H\bar{\boldsymbol{\Omega}})^{-1}\}, \end{bmatrix}, \tag{2.12}$$

$$\bar{\mathbf{F}}_{22}^{-1} = \begin{bmatrix} (\bar{\boldsymbol{\Phi}}_{11} - \bar{\boldsymbol{\Phi}}_{12}\bar{\boldsymbol{\Phi}}_{22}^{-1}\bar{\boldsymbol{\Phi}}_{21})^{-1} & -\bar{\boldsymbol{\Phi}}_{11}^{-1}\bar{\boldsymbol{\Phi}}_{12}(\bar{\boldsymbol{\Phi}}_{22} - \bar{\boldsymbol{\Phi}}_{21}\bar{\boldsymbol{\Phi}}_{11}^{-1}\bar{\boldsymbol{\Phi}}_{12})^{-1} \\ -\bar{\boldsymbol{\Phi}}_{22}^{-1}\bar{\boldsymbol{\Phi}}_{21}(\bar{\boldsymbol{\Phi}}_{11} - \bar{\boldsymbol{\Phi}}_{12}\bar{\boldsymbol{\Phi}}_{22}^{-1}\bar{\boldsymbol{\Phi}}_{21})^{-1} & (\bar{\boldsymbol{\Phi}}_{22} - \bar{\boldsymbol{\Phi}}_{21}\bar{\boldsymbol{\Phi}}_{11}^{-1}\bar{\boldsymbol{\Phi}}_{12})^{-1} \end{bmatrix}. \tag{2.13}$$

The CRLB for the estimation of MCFOs and MTOs is given by diagonal elements of the matrix, $\bar{\boldsymbol{\Upsilon}}$ in (2.11), and can be written in closed-form as

$$\begin{aligned}\mathbf{CRLB}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\tau}}) &= \frac{\sigma_n^2}{2}\mathrm{diag}\Big((\bar{\mathbf{F}}_{22} - \bar{\mathbf{Z}}^T\bar{\mathbf{F}}_{11}^{-1}\bar{\mathbf{Z}})^{-1}\Big) \\ &= \frac{\sigma_n^2}{2}\mathrm{diag}\Big(\bar{\mathbf{F}}_{22}^{-1} + \bar{\mathbf{F}}_{22}^{-1}\bar{\mathbf{Z}}^T(\bar{\mathbf{F}}_{11} - \bar{\mathbf{Z}}\bar{\mathbf{F}}_{22}^{-1}\bar{\mathbf{Z}}^T)^{-1}\bar{\mathbf{Z}}\bar{\mathbf{F}}_{22}^{-1}\Big), \end{aligned} \tag{2.14}$$

where the second equality in (2.14) follows from the matrix inverse identity $(\mathbf{A} - \mathbf{C}^T\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{C}^T(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{C}^T)^{-1}\mathbf{C}\mathbf{A}^{-1}$ for invertible matrices $\mathbf{A}$ and $\mathbf{D}$ in [107]. In this case, $\bar{\mathbf{F}}_{22}$, $\bar{\mathbf{F}}_{11}$, and $\bar{\mathbf{Z}}$ in (2.14) correspond to $\mathbf{A}$, $\mathbf{D}$, and $\mathbf{C}$, respectively. Similarly, the CRLB for the estimation of the combined real and imaginary parts of $\boldsymbol{\alpha}$ is derived as

$$\mathbf{CRLB}(\Re\{\boldsymbol{\alpha}\}, \Im\{\boldsymbol{\alpha}\}) = \frac{\sigma_n^2}{2}\mathrm{diag}\Big(\bar{\mathbf{F}}_{11}^{-1} + \bar{\mathbf{F}}_{11}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Upsilon}}\bar{\mathbf{Z}}^T\bar{\mathbf{F}}_{11}^{-1}\Big). \tag{2.15a}$$

$$\mathbf{CRLB}(\boldsymbol{\alpha}) = \frac{\sigma_n^2}{2}\mathrm{diag}\Big(\mathbf{B}\bar{\mathbf{F}}_{11}^{-1}\mathbf{B}^H + \mathbf{B}\bar{\mathbf{F}}_{11}^{-1}\bar{\mathbf{Z}}\bar{\boldsymbol{\Upsilon}}\bar{\mathbf{Z}}^T\bar{\mathbf{F}}_{11}^{-1}\mathbf{B}^H\Big) \tag{2.15b}$$

where $\mathbf{B} \equiv [\mathbf{I}_K \; j\mathbf{I}_K]$ is used to obtain the CRLB of $\boldsymbol{\alpha}$ according to [18].

2.2.2 CRLB for DF Relaying

Similar to the case of AF relaying, the received signal vector in (2.6), $\mathbf{y} \sim \mathcal{CN}(\boldsymbol{\mu}_{\mathbf{y}}, \boldsymbol{\Sigma}_{\mathbf{y}})$, with mean $\boldsymbol{\mu}_{\mathbf{y}} = \boldsymbol{\Omega}\boldsymbol{\eta}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}} = \sigma_w^2 \mathbf{I}_{LQ}$. The parameter vector of interest for DF relaying, $\boldsymbol{\theta}$, is given by

$$\boldsymbol{\theta} \triangleq \Big[\Re\{\boldsymbol{\eta}\}^T, \Im\{\boldsymbol{\eta}\}^T, \boldsymbol{\nu}^T, \boldsymbol{\tau}^T\Big]^T \tag{2.16}$$

where $\boldsymbol{\nu} \triangleq [\nu_1, \cdots, \nu_K]^T$ and $\boldsymbol{\tau} \triangleq [\tau_1, \cdots, \tau_K]^T$ and for notational simplicity, ν_k and τ_k are used to denote $\nu_k^{[rd]}$ and $\tau_k^{[rd]}$, respectively. The FIM for the estimation of $\boldsymbol{\theta}$ is given by [40]

$$\mathbf{F}^{\mathrm{DF}} = \frac{2}{\sigma_w^2} \underbrace{\begin{bmatrix} \Re\{\boldsymbol{\Omega}^H\boldsymbol{\Omega}\} & -\Im\{\boldsymbol{\Omega}^H\boldsymbol{\Omega}\} & -\Im\{\boldsymbol{\Omega}^H\mathbf{D}\boldsymbol{\Omega}\mathbf{H}\} & \Re\{\boldsymbol{\Omega}^H\boldsymbol{\Gamma}\mathbf{H}\} \\ \Im\{\boldsymbol{\Omega}^H\boldsymbol{\Omega}\} & \Re\{\boldsymbol{\Omega}^H\boldsymbol{\Omega}\} & \Re\{\boldsymbol{\Omega}^H\mathbf{D}\boldsymbol{\Omega}\mathbf{H}\} & \Im\{\boldsymbol{\Omega}^H\boldsymbol{\Gamma}\mathbf{H}\} \\ \Im\{\mathbf{H}^H\boldsymbol{\Omega}^H\mathbf{D}\boldsymbol{\Omega}\} & \Re\{\mathbf{H}^H\boldsymbol{\Omega}^H\mathbf{D}\boldsymbol{\Omega}\} & \underbrace{\Re\{\mathbf{H}^H\boldsymbol{\Omega}^H\mathbf{D}^2\boldsymbol{\Omega}\mathbf{H}\}}_{\triangleq\boldsymbol{\Phi}_{11}} & \underbrace{\Im\{\mathbf{H}^H\boldsymbol{\Omega}^H\mathbf{D}\boldsymbol{\Gamma}\mathbf{H}\}}_{\triangleq\boldsymbol{\Phi}_{12}} \\ \Re\{\mathbf{H}^H\boldsymbol{\Gamma}^H\boldsymbol{\Omega}\} & -\Im\{\mathbf{H}^H\boldsymbol{\Gamma}^H\boldsymbol{\Omega}\} & \underbrace{-\Im\{\mathbf{H}^H\boldsymbol{\Gamma}^H\mathbf{D}\boldsymbol{\Omega}\mathbf{H}\}}_{\triangleq\boldsymbol{\Phi}_{21}} & \underbrace{\Re\{\mathbf{H}^H\boldsymbol{\Gamma}^H\boldsymbol{\Gamma}\mathbf{H}\}}_{\triangleq\boldsymbol{\Phi}_{22}} \end{bmatrix}}_{\triangleq\boldsymbol{\mathcal{F}}}, \tag{2.17}$$

where $\boldsymbol{\Gamma} \triangleq [\boldsymbol{\Lambda}_1\mathbf{R}_1\mathbf{t}_1^{[r]}, \cdots, \boldsymbol{\Lambda}_K\mathbf{R}_K\mathbf{t}_K^{[r]}]$, $\mathbf{H} \triangleq \mathrm{diag}\{\eta_1, \cdots, \eta_K\}$, and $\mathbf{R}_k \triangleq \partial\mathbf{G}_k/\partial\tau_k$. Let us rewrite (2.17) as

$$\mathbf{F}^{\mathrm{DF}} \triangleq \frac{2}{\sigma_w^2} \begin{bmatrix} \mathbf{F}_{11} & \mathbf{Z} \\ \mathbf{Z}^T & \mathbf{F}_{22} \end{bmatrix},$$

where $\mathbf{F}_{11}$ and $\mathbf{F}_{22}$ denote the upper left and lower right $2K \times 2K$ sub matrices of $\boldsymbol{\mathcal{F}}$, respectively, and $\mathbf{Z}$ is the upper right $2K \times 2K$ sub matrix of $\boldsymbol{\mathcal{F}}$. Using similar steps as that of AF relaying, the closed-form CRLB for the estimation of $\boldsymbol{\theta}$ can be determined as

$$\mathbf{CRLB}(\boldsymbol{\theta}) = \frac{\sigma_w^2}{2} \begin{bmatrix} \mathbf{F}_{11}^{-1} + \mathbf{F}_{11}^{-1}\mathbf{Z}\boldsymbol{\Upsilon}\mathbf{Z}^T\mathbf{F}_{11}^{-1} & -\mathbf{F}_{11}^{-1}\mathbf{Z}\boldsymbol{\Upsilon} \\ -\boldsymbol{\Upsilon}\mathbf{Z}^T\mathbf{F}_{11}^{-1} & \boldsymbol{\Upsilon} \end{bmatrix} \tag{2.18}$$

where $\boldsymbol{\Upsilon} \triangleq \big[\mathbf{F}_{22} - \mathbf{Z}^T\mathbf{F}_{11}^{-1}\mathbf{Z}\big]^{-1}$. Moreover, $\mathbf{F}_{11}^{-1}$ and $\mathbf{F}_{22}^{-1}$ are given by replacing $\bar{\boldsymbol{\Omega}}$ and $\bar{\boldsymbol{\Phi}}_{l,m}$ with $\boldsymbol{\Omega}$ and $\boldsymbol{\Phi}_{l,m}$, $\forall l, m$ in (2.12) and (2.13), respectively. Using (2.18) the closed-form CRLBs for the estimation of $\boldsymbol{\nu}$, $\boldsymbol{\tau}$, and $\boldsymbol{\eta}$ can be determined as

$$\mathbf{CRLB}(\boldsymbol{\nu}, \boldsymbol{\tau}) = \frac{\sigma_w^2}{2}\mathrm{diag}\Big(\mathbf{F}_{22}^{-1} + \mathbf{F}_{22}^{-1}\mathbf{Z}^T(\mathbf{F}_{11} - \mathbf{Z}\mathbf{F}_{22}^{-1}\mathbf{Z}^T)^{-1}\mathbf{Z}\mathbf{F}_{22}^{-1}\Big), \tag{2.19a}$$

$$\mathbf{CRLB}(\boldsymbol{\eta}) = \frac{\sigma_w^2}{2}\Big(\mathbf{B}\mathbf{F}_{11}^{-1}\mathbf{B}^H + \mathbf{B}\mathbf{F}_{11}^{-1}\mathbf{Z}\boldsymbol{\Upsilon}\mathbf{Z}^T\mathbf{F}_{11}^{-1}\mathbf{B}^H\Big). \tag{2.19b}$$

The following remarks are in order:

Remark 2.2 *Eqs. (2.9) and(2.11) for AF relaying, and Eqs. (2.17) and (2.18) for DF-relaying cooperative networks demonstrate that for both choices of protocols, the FIM and the CRLB for the estimation of MCFOs, MTOs, and channel gains are not block diagonal. Thus, there exists coupling between the estimation errors of MCFOs, MTOs, and channel gains. This shows the importance of jointly estimating MCFOs, MTOs, and channel gains in multi-relay cooperative networks. More importantly, this result indicates that the previously proposed methods that assume perfect frequency or timing synchronization while estimating MCFOs and MTOs in [1] and [2], respectively, cannot be applied to estimate MCFOs, MTOs, and multiple channel gains in distributed AF and DF cooperative networks.*

Remark 2.3 *The CRLBs for the estimation of $\bar{\boldsymbol{\nu}}$, $\bar{\boldsymbol{\tau}}$ and $\boldsymbol{\alpha}$ for AF relaying in (2.14) and (2.15b) depend on the source-relay-destination carrier frequency offsets, $\{\bar{\nu}_k\}_{k=1}^K$, timing offsets, $\{\bar{\tau}_k\}_{k=1}^K$, and channel gains, $\{\alpha_k\}_{k=1}^K$, through the matrices $\bar{\mathbf{F}}_{11}$ and $\bar{\mathbf{Z}}$ in (2.10). As anticipated, the CRLBs in (2.14) and (2.15b) are also dependent on the TS length and the choice of the TSs broadcasted from both $\mathbb{S}$ and the relays. Based on the CRLBs in (2.19a) and (2.19b), similar dependencies can be also deduced for the case of DF relaying. Therefore, we can conclude that the choice of TS and its length are important design parameters that significantly impact the performance of MCFO, MTO, and channel estimators in distributed multi-relay cooperative networks. Optimal TS design for the estimation of MCFOs is addressed in [21] and for the estimation of MTOs is addressed in Chapter 4 of the thesis.*

2.3 Joint Parameter Estimation

In this section in order to reduce the computational complexity associated with the estimation of MTOs, MCFOs and multiple channel gains for DF-and AF-relaying networks, two iterative estimators are proposed based on the ECM and SAGE algorithms and their computational complexity is analyzed.

2.3.1 Estimation Algorithms for DF Cooperative Networks

We first outline the ML estimator due to its desirable asymptotic properties [18]. The ML estimator presented here is also required to initialize the proposed ECM

and SAGE algorithms.

2.3.1.1 ML Estimator for DF Relaying

Based on the signal model in (2.6), the ML estimates of $\boldsymbol{\nu}$, $\boldsymbol{\tau}$, and $\boldsymbol{\eta}$ are given by [40]

$$\hat{\boldsymbol{\nu}}, \hat{\boldsymbol{\tau}} = \arg\max_{\boldsymbol{\nu},\boldsymbol{\tau}} \mathbf{y}^H \boldsymbol{\Omega}(\boldsymbol{\Omega}^H\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^H\mathbf{y}, \qquad \hat{\boldsymbol{\eta}} = (\boldsymbol{\Omega}^H\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^H\mathbf{y}. \tag{2.20}$$

Alternating projection (AP) is used to reduce the dimensionality of the maximization problem in (2.20) into a series of one-dimensional searches [108]. Even though AP is not guaranteed to converge to the true estimates, the results in [108, Sec. IV-A], demonstrate that AP always converges to a local maximum and through proper initialization, AP results in global convergence. Following the ML-based approach in [45], $\hat{\nu}_k$ and $\hat{\tau}_k$ are initialized to 0, $\forall$, k in our simulations. In addition, the numerical simulations in Figs. 2.6 indicate that using the above initialization, AP converges to the true timing and frequency offset estimates in 2 cycles only.

2.3.1.2 ECM Estimator for DF Relaying

The entries in the vector $\boldsymbol{\theta}$ in (2.16) can be rearranged into the new parameter vector $\boldsymbol{\lambda} \triangleq \left[\boldsymbol{\lambda}_1^T, \cdots, \boldsymbol{\lambda}_K^T\right]^T$, where $\boldsymbol{\lambda}_k \triangleq [\nu_k, \tau_k, \eta_k]^T$, for $k = \{1, \cdots, K\}$. In the *expectation-maximization* (EM) terminology, the received training signal $\mathbf{y}$ in (2.6) represents the *incomplete data set* [109]. Following [110], we define the hidden or *complete data set* as $\mathbf{z} \triangleq \left[\mathbf{z}_1^T, \cdots, \mathbf{z}_K^T\right]^T$, with

$$\mathbf{z}_k = \eta_k \boldsymbol{\Lambda}_k \mathbf{G}_k \mathbf{t}_k^{[r]} + \mathbf{w}_k, \tag{2.21}$$

where $\mathbf{w}_k \sim \mathcal{CN}(0, \gamma_k \sigma_w^2 \mathbf{I}_{LQ})$ with $\sum_{k=1}^K \gamma_k = 1$ and $\mathbf{z}_k$ is the hidden variable. In (2.21), $\mathbf{w}_k$ is obtained by decomposing the total noise vector $\mathbf{w}$ into K components such that $\sum_{k=1}^K \mathbf{w}_k = \mathbf{w}$ where γ_k may be chosen such that $\gamma_k = 1/K$, $\forall\ k$ [111]. Based on (2.21), the relationship between the complete and incomplete data sets is given by $\sum_{k=1}^K \mathbf{z}_k = \mathbf{y}$. In order to indicate the iterative processing let us define $\hat{\boldsymbol{\lambda}}^{[m]} \triangleq \left[\left(\hat{\boldsymbol{\lambda}}_1^{[m]}\right)^T, \cdots, \left(\hat{\boldsymbol{\lambda}}_K^{[m]}\right)^T\right]^T$ as the estimated value of $\boldsymbol{\lambda}$ at the m^{th} iteration, where $\hat{\boldsymbol{\lambda}}_k^{[m]} \triangleq \left[\hat{\nu}_k^{[m]}, \hat{\tau}_k^{[m]}, \hat{\eta}_k^{[m]}\right]^T$. Note that rough initial estimates, $\hat{\nu}_k^{[0]}, \hat{\tau}_k^{[0]}, \hat{\eta}_k^{[0]}$, may be obtained using alternating projection via (2.20) using a coarse step size. The *E-step* and *M-step* of the proposed ECM algorithms at the m^{th} iteration are derived next.

E-step Using the received signal $\mathbf{y}$ and the current estimates, $\hat{\boldsymbol{\lambda}}^{[m]}$, we compute the expectation of the *log-likelihood function* (LLF) of the complete data space given the parameter $\boldsymbol{\lambda}$, $\mathbf{N}\left(\boldsymbol{\lambda}|\hat{\boldsymbol{\lambda}}^{[m]}\right)$. That is

$$\mathbf{N}\left(\boldsymbol{\lambda}|\hat{\boldsymbol{\lambda}}^{[m]}\right) \triangleq \mathbb{E}\left\{\log f(\mathbf{z}|\boldsymbol{\lambda}) \;\middle|\; \mathbf{y}, \hat{\boldsymbol{\lambda}}^{[m]}\right\}, \tag{2.22}$$

where the probability density function of $\mathbf{z}$ given $\boldsymbol{\lambda}$ is determined as

$$f\Big(\mathbf{z}|\boldsymbol{\lambda}\Big) = \prod_{k=1}^{K} f\Big(\mathbf{z}_k|\boldsymbol{\lambda}_k\Big) = \prod_{k=1}^{K} \frac{1}{(\pi\gamma_k\sigma_w^2)^{LQ}} \exp\left\{\frac{\|\mathbf{z}_k - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\|^2}{\gamma_k\sigma_w^2}\right\}. \tag{2.23}$$

Substituting (2.23) into (2.22), we obtain

$$\begin{aligned}\mathbf{N}(\boldsymbol{\lambda}|\hat{\boldsymbol{\lambda}}^{[m]}) &= E_1 - \mathbb{E}\left\{\sum_{k=1}^{K}\frac{1}{\gamma_k\sigma_w^2}\left\|\mathbf{z}_k - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\right\|^2\Big|\mathbf{y}, \hat{\boldsymbol{\lambda}}^{[m]}\right\} \\ &= E_1 - \sum_{k=1}^{K}\frac{1}{\gamma_k\sigma_w^2}\left\|\hat{\mathbf{z}}_k^{[m]} - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\right\|^2, \end{aligned}\tag{2.24}$$

where $E_1 = -LQ\sum_{k=1}^{K}\log(\pi\gamma_k\sigma_w^2)$ is a constant that is independent of $\boldsymbol{\lambda}$. Since $\sum_{k=1}^{K}\mathbf{z}_k = \mathbf{y}$, it can be concluded that $\mathbf{z}_k$ and $\mathbf{y}$ are jointly Gaussian distributed, $\forall\ k$. Therefore, $\hat{\mathbf{z}}_k$ in (2.24) is given by

$$\hat{\mathbf{z}}_k^{[m]} \triangleq \mathbb{E}\left\{\mathbf{z}_k|\mathbf{y}, \hat{\boldsymbol{\lambda}}^{[m]}\right\} = \hat{\eta}_k^{[m]}\hat{\boldsymbol{\Lambda}}_k^{[m]}\hat{\mathbf{G}}_k^{[m]}\mathbf{t}_k^{[r]} + \gamma_k\left(\mathbf{y} - \sum_{k=1}^{K}\hat{\eta}_k^{[m]}\hat{\boldsymbol{\Lambda}}_k^{[m]}\hat{\mathbf{G}}_k^{[m]}\mathbf{t}_k^{[r]}\right), \tag{2.25}$$

where $\hat{\boldsymbol{\Lambda}}_k^{[m]} \triangleq \boldsymbol{\Lambda}_k\big|_{\nu_k=\hat{\nu}_k^{[m]}}$ and $\hat{\mathbf{G}}_k^{[m]} \triangleq \mathbf{G}_k\big|_{\tau_k=\hat{\tau}_k^{[m]}}$.

M-step The estimate of $\boldsymbol{\lambda}$ at the $(m+1)^{\text{th}}$ iteration, $\hat{\boldsymbol{\lambda}}^{[m+1]}$, is determined as

$$\hat{\boldsymbol{\lambda}}^{[m+1]} = \arg\max_{\boldsymbol{\lambda}}\mathbf{N}\left(\boldsymbol{\lambda}|\hat{\boldsymbol{\lambda}}^{[m]}\right) = \arg\min_{\boldsymbol{\lambda}}\sum_{k=1}^{K}\left\|\hat{\mathbf{z}}_k^{[m]} - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\right\|^2. \tag{2.26}$$

From (2.26) it can be observed that the process of updating $\boldsymbol{\lambda}$ can be decoupled into the processes of updating each of the $\boldsymbol{\lambda}_k$, for $k = \{1, \cdots, K\}$. As a result, the update-equation for calculating $\hat{\boldsymbol{\lambda}}_k^{[m+1]}$ can be determined as

$$\hat{\boldsymbol{\lambda}}_k^{[m+1]} = \arg\min_{\boldsymbol{\lambda}_k}\left\|\hat{\mathbf{z}}_k^{[m]} - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\right\|^2. \tag{2.27}$$

In order to further reduce the complexity associated with the *M-step* of the EM algorithm, the ECM scheme [112] is applied in this section, where the cost function in (2.27) is minimized with respect to to one of the parameters of interest while keeping the remaining parameters at their most recently updated values [111,112]. In the first step using the ECM approach, $\hat{\nu}_k^{[m+1]}$ can be determined as

$$\begin{aligned}\hat{\nu}_k^{[m+1]} &= \arg\min_{\nu_k} \left\| \hat{\mathbf{z}}_k^{[m]} - \eta_k \mathbf{\Lambda}_k \mathbf{G}_k \mathbf{t}_k^{[r]} \right\|^2 \Big|_{\substack{\tau_k=\hat{\tau}_k^{[m]}, \\ \eta_k=\hat{\eta}_k^{[m]}}} \\ &= \arg\max_{\nu_k} \sum_{i=0}^{LQ-1} \Re\left\{ \left(\hat{z}_k^{[m]}(i)\right)^* \hat{\eta}_k^{[m]} e^{j2\pi i\nu_k/Q} b_i\left(\hat{\tau}_k^{[m]}\right)\right\}, \end{aligned} \tag{2.28}$$

where $\hat{z}_k^{[m]}(i)$ is the i^{th} element of $\hat{\mathbf{z}}_k^{[m]}$ for $i = \{0, 1, \cdots, LQ-1\}$ and $b_i\left(\hat{\tau}_k^{[m]}\right)$ can be found using

$$b_i(\tau_k) = \sum_{\ell=-Lg}^{Lg} t_k^{[r]}(\ell + \lfloor i/Q \rfloor)\; g(\mathrm{mod}(i/Q)T_s - \ell T - \tau_k T), \tag{2.29}$$

where L_g is the selected pulse shaping filter lag in the TP. In order to handle the nonlinearity of (2.28), we can approximate the term $e^{j2\pi i\nu_k/Q}$ using Taylor series expansion to the second-order term as

$$\begin{aligned} e^{j2\pi i\nu_k/Q} \approx e^{j2\pi i\hat{\nu}_k^{[m]}/Q} &+ (\nu_k - \hat{\nu}_k^{[m]})(j2\pi i/Q)e^{j2\pi i\hat{\nu}_k^{[m]}/Q} \\ &+ \frac{1}{2}(\nu_k - \hat{\nu}_k^{[m]})^2 (j2\pi i/Q)^2 e^{j2\pi i\hat{\nu}_k^{[m]}/Q}. \end{aligned} \tag{2.30}$$

Using (2.30), (2.28) can be rewritten as

$$\begin{aligned}\nu_k^{[m+1]} = \arg\max_{\nu_k} \mathbf{\Xi} \triangleq &\sum_{i=0}^{LQ-1} \Re\left\{ \left(\hat{z}_k^{[m]}(i)\right)^* \hat{\eta}_k^{[m]} e^{j2\pi i\hat{\nu}_k^{[m]}/Q} b_i\left(\hat{\tau}_k^{[m]}\right)\right\} \\ &- (\nu_k - \hat{\nu}_k^{[m]}) \sum_{i=0}^{LQ-1} \left(\frac{2\pi i}{Q}\right) \Im\left\{ \left(\hat{z}_k^{[m]}(i)\right)^* \hat{\eta}_k^{[m]} e^{j2\pi i\hat{\nu}_k^{[m]}/Q} b_i\left(\hat{\tau}_k^{[m]}\right)\right\} \\ &- \frac{1}{2}(\nu_k - \hat{\nu}_k^{[m]})^2 \sum_{i=0}^{LQ-1} \left(\frac{2\pi i}{Q}\right)^2 \Re\left\{ \left(\hat{z}_k^{[m]}(i)\right)^* \hat{\eta}_k^{[m]} e^{j2\pi i\hat{\nu}_k^{[m]}/Q} b_i\left(\hat{\tau}_k^{[m]}\right)\right\}. \end{aligned} \tag{2.31}$$

Differentiating $\boldsymbol{\Xi}$ in (2.31) with respect to ν_k and equating the result to zero, the estimate of $\hat{\nu}_k$ at the $(m+1)^{\text{th}}$ iteration is obtained as

$$\hat{\nu}_k^{[m+1]} = \hat{\nu}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} \left(\frac{2\pi i}{Q}\right) \Im\left\{ \left(\hat{z}_k^{[m]}(i)\right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m]}/Q} b_i\left(\hat{\tau}_k^{[m]}\right) \right\}}{\sum_{i=0}^{LQ-1} \left(\frac{2\pi i}{Q}\right)^2 \Re\left\{ \left(\hat{z}_k^{[m]}(i)\right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m]}/Q} b_i\left(\hat{\tau}_k^{[m]}\right) \right\}}. \tag{2.32}$$

In the second step, by setting ν_k to its latest updated value, $\hat{\nu}_k^{[m+1]}$, the updated value of τ_k at the $(m+1)^{\text{th}}$ iteration, $\hat{\tau}_k^{[m+1]}$, can be determined as

$$\begin{aligned}
\hat{\tau}_k^{[m+1]} &= \arg\min_{\tau_k} \left\| \hat{\mathbf{z}}_k^{[m]} - \eta_k \boldsymbol{\Lambda}_k \mathbf{G}_k \mathbf{t}_k^{[r]} \right\|^2 \Big|_{\substack{\nu_k = \hat{\nu}_k^{[m+1]} \\ \eta_k = \hat{\eta}_k^{[m]}}} \\
&= \arg\max_{\tau_k} \sum_{i=0}^{LQ-1} \Re\left\{ \left(\hat{z}_k^{[m]}(i)\right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i(\tau_k) \right\}.
\end{aligned} \tag{2.33}$$

Taylor series expansion is again applied to linearize the maximization in (2.33), where the Taylor series expansion of $b_i(\tau_k)$, in (2.29), can be calculated as

$$b_i(\tau_k) \approx b_i\left(\hat{\tau}_k^{[m]}\right) + (\tau_k - \hat{\tau}_k^{[m]}) b_i'(\tau_k)\big|_{\tau_k = \hat{\tau}_k^{[m]}} + \frac{1}{2}(\tau_k - \hat{\tau}_k^{[m]})^2 b_i''(\tau_k)\big|_{\tau_k = \hat{\tau}_k^{[m]}}. \tag{2.34}$$

In (2.34), $b_i'(\tau_k)$ and $b_i''(\tau_k)$ are the first and second order derivatives of the function $b_i(\tau_k)$ with respect to τ_k and are given by

$$b_i'(\tau_k) = \sum_{\ell=-Lg}^{Lg} t_k^{[r]}(\ell + \lfloor i/Q \rfloor)\; g'(\text{mod}(i/Q)T_s - \ell T - \tau_k T) \tag{2.35a}$$

$$b_i''(\tau_k) = \sum_{\ell=-Lg}^{Lg} t_k^{[r]}(\ell + \lfloor i/Q \rfloor)\; g''(\text{mod}(i/Q)T_s - \ell T - \tau_k T) \tag{2.35b}$$

where $g'(t)|_{\tau_k} = \frac{\partial}{\partial \tau_k} g(t)$ and $g''(t)|_{\tau_k} = \frac{\partial^2}{(\partial \tau_k)^2} g(t)$ are the first and second order derivatives of transmitted pulse shaping function, $g(t)$, evaluated at $\tau = \tau_k$. Using forward difference approximation of derivatives given in [113], $g'(t)$ and $g''(t)$ can be determined, respectively, as

$$g'(t)|_{\tau_k} \simeq \frac{g(t)|_{\tau_k+\delta} - g(t)|_{\tau_k}}{\delta}, \qquad g''(t)|_{\tau_k} \simeq \frac{g(t)|_{\tau_k+2\delta} - 2g(t)|_{\tau_k+\delta} + g(t)|_{\tau_k}}{\delta^2}, \tag{2.36}$$

Table 2.1: Proposed ECM algorithm for the DF cooperative systems

Initialization

Obtain $\hat{\nu}_k^{[0]}, \hat{\tau}_k^{[0]}$, and $\hat{\eta}_k^{[0]}$ for $k = 1, \ldots, K$ using alternating projection and (2.20) with coarser step size i.e., 0.01

ECM for DF systems

for $m = 0, 1, \ldots$

for $k = 1, 2, \ldots, K$

$\hat{\mathbf{z}}_k^{[m]} = \hat{\eta}_k^{[m]} \hat{\mathbf{\Lambda}}_k^{[m]} \hat{\mathbf{G}}_k^{[m]} \mathbf{t}_k^{[r]} + \gamma_k \big(\mathbf{y} - \sum_{k=1}^{K} \hat{\eta}_k^{[m]} \hat{\mathbf{\Lambda}}_k^{[m]} \hat{\mathbf{G}}_k^{[m]} \mathbf{t}_k^{[r]} \big)$

end

for $k = 1, 2, \ldots, K$

$$\hat{\nu}_k^{[m+1]} = \hat{\nu}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} (\frac{2\pi i}{Q}) \Im \Big\{ \Big(\hat{z}_k^{[m]}(i) \Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m]}/Q} b_i(\hat{\tau}_k^{[m]}) \Big\}}{\sum_{i=0}^{LQ-1} (\frac{2\pi i}{Q})^2 \Re \Big\{ \Big(\hat{z}_k^{[m]}(i) \Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m]}/Q} b_i(\hat{\tau}_k^{[m]}) \Big\}}$$

$$\hat{\tau}_k^{[m+1]} = \hat{\tau}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} \Re \Big\{ \Big(\hat{z}_k^{[m]}(i) \Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i'(\hat{\tau}_k^{[m]}) \Big\}}{\sum_{i=0}^{LQ-1} \Re \Big\{ \Big(\hat{z}_k^{[m]}(i) \Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i''(\hat{\tau}_k^{[m]}) \Big\}}$$

$$\hat{\eta}_k^{[m+1]} = \frac{1}{\sum_{i=0}^{LQ-1} |b_i(\hat{\tau}_k^{[m+1]})|^2} \sum_{i=0}^{LQ-1} \frac{\hat{z}_k^{[m]}(i) \Big(b_i(\hat{\tau}_k^{[m+1]}) \Big)^*}{e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q}}.$$

end

$\hat{\nu}_k^{[m]} = \hat{\nu}_k^{[m+1]}$, $\hat{\tau}_k^{[m]} = \hat{\tau}_k^{[m+1]}$, $\hat{\eta}_k^{[m]} = \hat{\eta}_k^{[m+1]}$.

end

as $\delta \to 0^+$. Using (2.34), (2.33) can be rewritten as

$$\begin{aligned} \hat{\tau}_k^{[m+1]} = \arg\max_{\tau_k} & \sum_{i=0}^{LQ-1} \Re \left\{ \left(\hat{z}_k^{[m]}(i) \right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i \left(\hat{\tau}_k^{[m]} \right) \right\}, \\ & + (\tau_k - \hat{\tau}_k^{[m]}) \sum_{i=0}^{LQ-1} \Re \left\{ \left(\hat{z}_k^{[m]}(i) \right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i' \left(\hat{\tau}_k^{[m]} \right) \right\}, \\ & + \frac{1}{2} (\tau_k - \hat{\tau}_k^{[m]})^2 \sum_{i=0}^{LQ-1} \Re \left\{ \left(\hat{z}_k^{[m]}(i) \right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i'' \left(\hat{\tau}_k^{[m]} \right) \right\}. \end{aligned} \quad (2.37)$$

By taking the derivative of (2.37) with respect to τ_k and equating the result to zero, the estimate of $\hat{\tau}_k$ at the $(m+1)^{\text{th}}$ iteration is given by

$$\hat{\tau}_k^{[m+1]} = \hat{\tau}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} \Re \left\{ \left(\hat{z}_k^{[m]}(i) \right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i' \left(\hat{\tau}_k^{[m]} \right) \right\}}{\sum_{i=0}^{LQ-1} \Re \left\{ \left(\hat{z}_k^{[m]}(i) \right)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i'' \left(\hat{\tau}_k^{[m]} \right) \right\}}. \quad (2.38)$$

Finally, in the third step, by setting ν_k and τ_k to $\hat{\nu}_k^{[m+1]}$ and $\hat{\tau}_k^{[m+1]}$, respectively,

the estimate of η_k, at the $(m+1)^{\text{th}}$ iteration is calculated as

$$\begin{aligned}\hat{\eta}_k^{[m+1]} &= \arg\min_{\eta_k} \left\|\hat{\mathbf{z}}_k^{[m]} - \eta_k \mathbf{\Lambda}_k \mathbf{G}_k \mathbf{t}_k^{[r]}\right\|^2 \Big|_{\substack{\nu_k=\hat{\nu}_k^{[m+1]},\\ \tau_k=\hat{\tau}_k^{[m+1]}}} \\ &= \underset{\eta_k}{\text{argmin}} \sum_{i=0}^{LQ-1} \left|\hat{z}_k^{[m]}(i) - \eta_k e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i(\hat{\tau}_k^{[m+1]})\right|^2. \end{aligned} \tag{2.39}$$

By taking the derivative of (2.39) with respect to η_k and setting the result to zero, $\hat{\eta}_k^{[m+1]}$ can be determined as

$$\hat{\eta}_k^{[m+1]} = \frac{1}{\sum_{i=0}^{LQ-1}\left|b_i\left(\hat{\tau}_k^{[m+1]}\right)\right|^2} \sum_{i=0}^{LQ-1} \frac{\hat{z}_k^{[m]}(i)\left(b_i\left(\hat{\tau}_k^{[m+1]}\right)\right)^*}{e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q}}. \tag{2.40}$$

The proposed ECM estimator for DF relaying is summarized in Table 2.1. By reapplying the above algorithm, for $k = \{1 \cdots, K\}$, estimates of the MCFOs, MTOs, and multiple channel gains for all the relays can be obtained at $\mathbb{D}$. The iterations stop when the difference between LLFs of the two iterations is smaller than a threshold χ, i.e.,

$$\left|\left\|\mathbf{y} - \sum_{k=1}^{K} \hat{\eta}_k^{[m+1]} \hat{\mathbf{\Lambda}}_k^{[m+1]} \hat{\mathbf{G}}_k^{[m+1]} \mathbf{t}_k^{[r]}\right\|^2 - \left\|\mathbf{y} - \sum_{k=1}^{K} \hat{\eta}_k^{[m]} \hat{\mathbf{\Lambda}}_k^{[m]} \hat{\mathbf{G}}_k^{[m]} \mathbf{t}_k^{[r]}\right\|^2\right| \leq \chi. \tag{2.41}$$

2.3.1.3 SAGE Estimator for DF Relaying

As shown in [114] the SAGE algorithm can be applied to improve the convergence rate of the ECM approach. Using SAGE, the parameter $\boldsymbol{\lambda}$ is divided into K groups denoted by $\boldsymbol{\lambda}_k$, for $k = \{1 \cdots, K\}$. During the estimation process each group is updated while keeping the remaining groups fixed at their latest updated values. In addition, for each group a *hidden data set* is selected [114]. In this case, the hidden data set denoted by $\mathbf{x}_k$ for $\boldsymbol{\lambda}_k$ is given by

$$\mathbf{x}_k = \eta_k \mathbf{\Lambda}_k \mathbf{G}_k \mathbf{t}_k^{[r]} + \mathbf{w}. \tag{2.42}$$

The updating process for $\boldsymbol{\lambda}_k$ at the m^{th} iteration in the proposed SAGE estimator consists of *E*-and *M-steps*, which are derived using the steps outlined for the proposed ECM algorithm.

E-step While setting $\boldsymbol{\lambda}_\ell = \hat{\boldsymbol{\lambda}}_\ell^{[m]}\ \forall \ell \neq k$, the expectation of the LLF of the hidden data set for the parameter $\boldsymbol{\lambda}_k$, $\mathbf{N}\left(\boldsymbol{\lambda}_k|\hat{\boldsymbol{\lambda}}^{[m]}\right)$, is determined as

$$\mathbf{N}\left(\boldsymbol{\lambda}_k|\hat{\boldsymbol{\lambda}}^{[m]}\right) \triangleq \mathbb{E}\left\{\log f\left(\mathbf{x}_k\middle|\boldsymbol{\lambda}_k, \{\hat{\boldsymbol{\lambda}}_\ell^{[m]}\}_{\ell\neq k}\right)\ \middle|\ \mathbf{y}, \hat{\boldsymbol{\lambda}}^{[m]}\right\}, \tag{2.43}$$

where

$$f\left(\mathbf{x}_k\middle|\boldsymbol{\lambda}_k, \{\hat{\boldsymbol{\lambda}}_\ell^{[m]}\}_{\ell\neq k}\right) = f\left(\mathbf{x}_k|\boldsymbol{\lambda}_k\right) = \frac{1}{(\pi\sigma_w^2)^{LQ}} \exp\left\{\frac{\|\mathbf{x}_k - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\|^2}{\sigma_w^2}\right\}. \tag{2.44}$$

Substituting (2.44) into (2.43), we obtain

$$\begin{aligned}\mathbf{N}\left(\boldsymbol{\lambda}_k|\hat{\boldsymbol{\lambda}}^{[m]}\right) &= E_2 - \frac{1}{\sigma_w^2}\mathbb{E}\left\{\left\|\mathbf{x}_k - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\right\|^2\middle|\mathbf{y}, \hat{\boldsymbol{\lambda}}^{[m]}\right\}\\ &= E_2 - \frac{1}{\sigma_w^2}\left\|\hat{\mathbf{x}}_k^{[m]} - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\right\|^2,\end{aligned} \tag{2.45}$$

where

$$\begin{aligned}\hat{\mathbf{x}}_k^{[m]} \triangleq \mathbb{E}\left\{\mathbf{x}_k\middle|\mathbf{y}, \hat{\boldsymbol{\lambda}}^{[m]}\right\} &= \hat{\eta}_k^{[m]}\hat{\boldsymbol{\Lambda}}_k^{[m]}\hat{\mathbf{G}}_k^{[m]}\mathbf{t}_k^{[r]} + \left(\mathbf{y} - \sum_{k=1}^{K}\hat{\eta}_k^{[m]}\hat{\boldsymbol{\Lambda}}_k^{[m]}\hat{\mathbf{G}}_k^{[m]}\mathbf{t}_k^{[r]}\right)\\ &= \mathbf{y} - \sum_{\substack{\ell=1,\\ \ell\neq k}}^{K}\hat{\eta}_\ell^{[m]}\hat{\boldsymbol{\Lambda}}_\ell^{[m]}\hat{\mathbf{G}}_\ell^{[m]}\mathbf{t}_\ell^{[r]},\end{aligned} \tag{2.46}$$

and $E_2 = -(LQ)\log(\pi\sigma_w^2)$ is a constant independent of $\boldsymbol{\lambda}_k$.

M-step In this step, the estimate of $\boldsymbol{\lambda}_k$ in the $(m+1)^{\text{th}}$ iteration, $\hat{\boldsymbol{\lambda}}_k^{[m+1]}$, is determined as

$$\hat{\boldsymbol{\lambda}}_k^{[m+1]} = \arg\max_{\boldsymbol{\lambda}_k}\mathbf{N}\left(\boldsymbol{\lambda}_k|\hat{\boldsymbol{\lambda}}^{[m]}\right) = \arg\min_{\boldsymbol{\lambda}_k}\left\|\hat{\mathbf{x}}_k^{[m]} - \eta_k\boldsymbol{\Lambda}_k\mathbf{G}_k\mathbf{t}_k^{[r]}\right\|^2. \tag{2.47}$$

Using similar steps as the *M-step* of the ECM algorithm in Section 2.3.1.2, the computational complexity associated with the estimation of $\hat{\boldsymbol{\lambda}}^{[m+1]}$ can be further reduced. Subsequently, for the proposed SAGE estimator, by replacing $\hat{z}_k^{[m]}(i)$ with $\hat{x}_k^{[m]}(i)$, Eqs. (2.32), (2.38), and (2.40) can be applied to estimate $\hat{\nu}_k^{[m+1]}$, $\hat{\tau}_k^{[m+1]}$, and $\hat{\eta}_k^{[m+1]}$, respectively. Table 2.2 summarizes the proposed SAGE estimator for DF relaying.

Remark 2.4 *Even though it cannot be analytically shown that the proposed ECM and SAGE algorithms converge to a global maximum, in [112, page 1] and [114,*

Table 2.2: Proposed SAGE algorithm for the DF cooperative systems

Initialization

Obtain $\hat{\nu}_k^{[0]}, \hat{\tau}_k^{[0]}$, and $\hat{\eta}_k^{[0]}$ for $k = 1, \ldots, K$ using alternating projection and (2.20) with coarser step size i.e., 0.01

SAGE for DF systems

for $m = 0, 1, \ldots$

for $k = 1, 2, \ldots, K$

$$\hat{\mathbf{x}}_k^{[m]} = \mathbf{y} - \sum_{l=1, l\neq k}^{K} \hat{\eta}_l^{[m]} \hat{\mathbf{\Lambda}}_l^{[m]} \hat{\mathbf{G}}_l^{[m]} \mathbf{t}_l^{[r]}$$

$$\hat{\nu}_k^{[m+1]} = \hat{\nu}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} (\frac{2\pi i}{Q}) \Im\Big\{ \Big(\hat{x}_k^{[m]}(i)\Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m]}/Q} b_i(\hat{\tau}_k^{[m]}) \Big\}}{\sum_{i=0}^{LQ-1} (\frac{2\pi i}{Q})^2 \Re\Big\{ \Big(\hat{x}_k^{[m]}(i)\Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m]}/Q} b_i(\hat{\tau}_k^{[m]}) \Big\}}$$

$$\hat{\tau}_k^{[m+1]} = \hat{\tau}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} \Re\Big\{ \Big(\hat{x}_k^{[m]}(i)\Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i'(\hat{\tau}_k^{[m]}) \Big\}}{\sum_{i=0}^{LQ-1} \Re\Big\{ \Big(\hat{x}_k^{[m]}(i)\Big)^* \hat{\eta}_k^{[m]} e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q} b_i''(\hat{\tau}_k^{[m]}) \Big\}}$$

$$\hat{\eta}_k^{[m+1]} = \frac{1}{\sum_{i=0}^{LQ-1} |b_i(\hat{\tau}_k^{[m+1]})|^2} \sum_{i=0}^{LQ-1} \frac{\hat{x}_k^{[m]}(i) \Big(b_i(\hat{\tau}_k^{[m+1]})\Big)^*}{e^{j2\pi i \hat{\nu}_k^{[m+1]}/Q}}.$$

end

$\hat{\nu}_k^{[m]} = \hat{\nu}_k^{[m+1]}$, $\hat{\tau}_k^{[m]} = \hat{\tau}_k^{[m+1]}$, $\hat{\eta}_k^{[m]} = \hat{\eta}_k^{[m+1]}$.

end

page 4] it is established that in general ECM and SAGE algorithms monotonically increase the LLF at every iteration and converge to a local maximum. Moreover, if the algorithms are initialized in a region suitably close to the global maximum, then the sequence of estimates converge monotonically to the global maximum [114, page 4]. In our simulations, initial rough estimates, $\hat{\nu}_k^{[0]}, \hat{\tau}_k^{[0]}, \hat{\eta}_k^{[0]}$, *are obtained using alternating projection via* (2.20) *while using a coarse step size, e.g.,* 10^{-2}. *Finally, simulation results in Section 2.5 investigate the performance of the proposed ECM and SAGE algorithms for different initialization step size values and illustrate that the proposed algorithms converge to the true estimates with the initialization step size of* 10^{-2}.

2.3.2 Estimation Algorithms for AF Cooperative Networks

In this subsection the LS, ECM, and SAGE algorithms for the joint estimation of MCFOs, MTOs, and multiple channel gains in AF cooperative networks are derived.

2.3.2.1 LS Estimator for AF Model

Based on the training signal model at $\mathbb{D}$ for AF relaying in (2.4), the LS estimate of the parameters $\boldsymbol{\alpha}, \bar{\boldsymbol{\nu}}$, and $\bar{\boldsymbol{\tau}}$ can be determined by minimizing the cost function

$$\bar{\mathbf{J}}(\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\tau}}, \boldsymbol{\alpha}) = \|\bar{\mathbf{y}} - \bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}\|^2. \tag{2.48}$$

Given $\bar{\boldsymbol{\nu}}$ and $\bar{\boldsymbol{\tau}}$, the LS estimate of $\boldsymbol{\alpha}$ can be straightforwardly shown to be

$$\hat{\boldsymbol{\alpha}} = (\bar{\boldsymbol{\Omega}}^H \bar{\boldsymbol{\Omega}})^{-1} \bar{\boldsymbol{\Omega}}^H \bar{\mathbf{y}}. \tag{2.49}$$

By substituting (2.49) into (2.48), estimates of MCFOs and MTOs, $\hat{\bar{\boldsymbol{\nu}}}, \hat{\bar{\boldsymbol{\tau}}}$, respectively, are obtained via

$$\hat{\bar{\boldsymbol{\nu}}}, \hat{\bar{\boldsymbol{\tau}}} = \arg\max_{\bar{\boldsymbol{\nu}}, \bar{\boldsymbol{\tau}}} \bar{\mathbf{y}}^H \bar{\boldsymbol{\Omega}} (\bar{\boldsymbol{\Omega}}^H \bar{\boldsymbol{\Omega}})^{-1} \bar{\boldsymbol{\Omega}}^H \bar{\mathbf{y}}. \tag{2.50}$$

The maximization in (2.50) needs to be carried out using a multi-dimensional exhaustive search over the set of possible timing and frequency offsets. Thus, in order to reduce the computational complexity associated with obtaining these estimates the maximization in (2.50) is carried out using AP, where the AP algorithm is initialized using a similar approach as that of DF relaying in Section 2.3.1.1. In addition, the numerical investigation in Section VI indicates that the MCFOs and MTOs estimates obtained using (2.49) and (2.50) are close to the true estimates at mid-to-high SNR. Finally, the LS estimates of the channel gains $\boldsymbol{\alpha}$ can be obtained by substituting $\hat{\bar{\boldsymbol{\nu}}}$ and $\hat{\bar{\boldsymbol{\tau}}}$ into (2.49).

2.3.2.2 ECM Estimator for AF Relaying

The entries of the vector, $\bar{\boldsymbol{\theta}}$ in (2.8) can be rearranged to obtain a new parameter vector of interest $\bar{\boldsymbol{\lambda}} \triangleq \left[\bar{\boldsymbol{\lambda}}_1^T, \cdots, \bar{\boldsymbol{\lambda}}_K^T\right]^T$, where $\bar{\boldsymbol{\lambda}}_k \triangleq [\bar{\nu}_k, \bar{\tau}_k, \alpha_k]^T$ is a vector of three parameters corresponding to $\mathbb{R}_k$. Since the observed signal, $\bar{\mathbf{y}}$, in (2.4) is the incomplete data set in the case of AF relaying, we define the complete or hidden data set as $\bar{\mathbf{z}} \triangleq \left[\bar{\mathbf{z}}_1^T, \cdots, \bar{\mathbf{z}}_K^T\right]^T$, where

$$\bar{\mathbf{z}}_k = \alpha_k \left(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_k^{[r]}(\tau_k^{[rd]}) + \mathbf{n}_k. \tag{2.51}$$

In (2.51), $\mathbf{n}_k \sim \mathcal{CN}(0, \gamma_k \tilde{\sigma}_n^2 \mathbf{I}_{LQ})$ (γ_k is defined below (2.21)) is obtained by decomposing the overall noise vector, $\mathbf{n} \triangleq \bar{\boldsymbol{\Psi}}\boldsymbol{\beta} + \mathbf{w}$, into K components, such that $\sum_{k=1}^K \mathbf{n}_k = \mathbf{n}$. Based on the derivation in Appendix A, the covariance matrix for $\mathbf{n}$ is $\boldsymbol{\Sigma}_{\mathbf{n}} = \left(\sigma_u^2 \sum_{k=1}^K |\beta_k|^2 + \sigma_w^2\right) \mathbf{I}_{LQ}$. Given that $\beta_k = \zeta_k \eta_k$ where η_k is the fading

channel gain with variance $\sigma_{\eta_k}^2$, the covariance matrix for $\mathbf{n}$ can be modified to be $\tilde{\sigma}_n^2\mathbf{I}_{LQ}$, where $\tilde{\sigma}_n^2 \triangleq \sigma_u^2\sum_{k=1}^K \sigma_{\eta_k}^2\zeta_k^2 + \sigma_w^2$. Thus, the relationship between the complete and incomplete data sets is given by $\sum_{k=1}^K \bar{\mathbf{z}}_k = \bar{\mathbf{y}}$. For further iterative processing, let us denote $\hat{\bar{\boldsymbol{\lambda}}}^{[m]} \triangleq \left[\left(\hat{\bar{\boldsymbol{\lambda}}}_1^{[m]}\right)^T, \cdots, \left(\hat{\bar{\boldsymbol{\lambda}}}_K^{[m]}\right)^T\right]^T$ as the estimated value of $\bar{\boldsymbol{\lambda}}$ at the m^{th} iteration, where $\hat{\bar{\boldsymbol{\lambda}}}_k^{[m]} \triangleq \left[\hat{\bar{\nu}}_k^{[m]}, \hat{\bar{\tau}}_k^{[m]}, \hat{\alpha}_k^{[m]}\right]^T$. Note that for AF relaying, the rough initial estimates, $\hat{\bar{\nu}}_k^{[0]}, \hat{\bar{\tau}}_k^{[0]}, \hat{\alpha}_k^{[0]}$, are obtained by applying AP on the proposed LS estimator in (2.49), and (2.50) with a coarse step size. The *E*-and *M-steps* of the proposed ECM estimator at the m^{th} iteration are derived in the following subsections.

E-step Using the received signal, $\bar{\mathbf{y}}$, and the current estimate $\hat{\bar{\boldsymbol{\lambda}}}^{[m]}$, the expectation of the complete LLF given the parameter $\bar{\boldsymbol{\lambda}}$, $\bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}|\hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right)$, is computed as

$$\bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}|\hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right) \triangleq \mathbb{E}\left\{\log f(\bar{\mathbf{z}}|\bar{\boldsymbol{\lambda}})\middle|\bar{\mathbf{y}}, \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right\}, \tag{2.52}$$

where the probability density function of $\bar{\mathbf{z}}$ as a function of $\bar{\boldsymbol{\lambda}}$ is given by

$$f\left(\bar{\mathbf{z}}|\bar{\boldsymbol{\lambda}}\right) = \prod_{k=1}^K f\left(\bar{\mathbf{z}}_k|\bar{\boldsymbol{\lambda}}_k\right) = \prod_{k=1}^K \frac{1}{(\pi\gamma_k\tilde{\sigma}_n^2)^{LQ}} \exp\left\{\frac{\|\bar{\mathbf{z}}_k - \alpha_k(\bar{\boldsymbol{\Lambda}}_k\bar{\mathbf{G}}_k\mathbf{t}^{[s]}) \odot \bar{\mathbf{t}}_k^{[r]}\|^2}{\gamma_k\tilde{\sigma}_n^2}\right\}. \tag{2.53}$$

where $\bar{\mathbf{t}}_k^{[r]}$ refers to $\bar{\mathbf{t}}_k^{[r]}(\hat{\bar{\tau}}_k^{[0]})$ in Section 2.3.2. Substituting (2.53) into (2.52), we obtain

$$\bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}|\hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right) = E_3 - \sum_{k=1}^K \frac{1}{\gamma_k\tilde{\sigma}_n^2}\left\|\hat{\bar{\mathbf{z}}}_k^{[m]} - \alpha_k\left(\bar{\boldsymbol{\Lambda}}_k\bar{\mathbf{G}}_k\mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_k^{[r]}\right\|^2, \tag{2.54}$$

where $E_3 = -LQ\sum_{k=1}^K \log(\pi\gamma_k\tilde{\sigma}_n^2)$ is a constant independent of $\bar{\boldsymbol{\lambda}}$. Given that $\bar{\mathbf{z}}_k$ and $\bar{\mathbf{y}}$ are jointly Gaussian distributed, $\forall\ k$, we have

$$\begin{aligned}\hat{\bar{\mathbf{z}}}_k^{[m]} &\triangleq \mathbb{E}\left\{\bar{\mathbf{z}}_k|\bar{\mathbf{y}}, \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right\} \\ &= \hat{\alpha}_k^{[m]}\left(\hat{\bar{\boldsymbol{\Lambda}}}_k^{[m]}\hat{\bar{G}}_k^{[m]}\mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_k^{[r]} + \gamma_k\left(\bar{\mathbf{y}} - \sum_{k=1}^K \hat{\alpha}_k^{[m]}\left(\hat{\bar{\boldsymbol{\Lambda}}}_k^{[m]}\hat{\bar{G}}_k^{[m]}\mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_k^{[r]}\right),\end{aligned} \tag{2.55}$$

where $\hat{\bar{\boldsymbol{\Lambda}}}_k^{[m]} \triangleq \bar{\boldsymbol{\Lambda}}_k|_{\bar{\nu}_k=\hat{\bar{\nu}}_k^{[m]}}$ and $\hat{\bar{G}}_k^{[m]} \triangleq \bar{\mathbf{G}}_k|_{\bar{\tau}_k=\hat{\bar{\tau}}_k^{[m]}}$.

M-step The updated value of $\bar{\boldsymbol{\lambda}}$, $\hat{\bar{\boldsymbol{\lambda}}}^{[m+1]}$, is determined as

$$\hat{\bar{\boldsymbol{\lambda}}}^{[m+1]} = \arg\max_{\bar{\boldsymbol{\lambda}}} \bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}|\hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right) = \arg\min_{\bar{\boldsymbol{\lambda}}} \sum_{k=1}^{K} \left\| \hat{\bar{z}}_k^{[m]} - \alpha_k\big(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}^{[s]}\big) \odot \bar{\mathbf{t}}_k^{[r]} \right\|^2. \quad (2.56)$$

From (2.56) it can be straightforwardly observed that the updating process of $\bar{\boldsymbol{\lambda}}$ can be decoupled into K updating processes of $\bar{\boldsymbol{\lambda}}_k$ for $k = \{1, \cdots, K\}$. Thus, the update equation to determine $\hat{\bar{\boldsymbol{\lambda}}}_k^{[m+1]}$ is given by

$$\hat{\bar{\boldsymbol{\lambda}}}_k^{[m+1]} = \arg\min_{\bar{\boldsymbol{\lambda}}_k} \left\| \hat{\bar{z}}_k^{[m]} - \alpha_k\big(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}^{[s]}\big) \odot \bar{\mathbf{t}}_k^{[r]} \right\|^2. \quad (2.57)$$

Similar to DF relaying, the proposed ECM estimator for AF relaying minimizes (2.57) in three steps. Following the same steps as in Section 2.3.1.2, the updated value of $\bar{\nu}_k$, $\hat{\bar{\nu}}_k^{[m+1]}$, is obtained as

$$\hat{\bar{\nu}}_k^{[m+1]} = \hat{\bar{\nu}}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} \left(\frac{2\pi i}{Q}\right) \Im\left\{ \left(\hat{\bar{z}}_k^{[m]}(i)\right)^* \hat{\alpha}_k^{[m]} e^{j2\pi i \hat{\bar{\nu}}_k^{[m]}/Q} \bar{t}_k^{[r]}(i) \bar{b}_i\left(\hat{\bar{\tau}}_k^{[m]}\right) \right\}}{\sum_{i=0}^{LQ-1} \left(\frac{2\pi i}{Q}\right)^2 \Re\left\{ \left(\hat{\bar{z}}_k^{[m]}(i)\right)^* \hat{\alpha}_k^{[m]} e^{j2\pi i \hat{\bar{\nu}}_k^{[m]}/Q} \bar{t}_k^{[r]}(i) \bar{b}_i\left(\hat{\bar{\tau}}_k^{[m]}\right) \right\}}, \quad (2.58)$$

where $\bar{b}_i(\bar{\tau}_k) = \sum_{\ell=-Lg}^{Lg} t^{[s]}(\ell + \lfloor i/Q \rfloor)\, g(\mathrm{mod}(i/Q)T_s - \ell T - \bar{\tau}_k T)$. Similarly the updated value of $\bar{\tau}_k$, $\hat{\bar{\tau}}_k^{[m+1]}$, is obtained as

$$\hat{\bar{\tau}}_k^{[m+1]} = \hat{\bar{\tau}}_k^{[m]} - \frac{\sum_{i=0}^{LQ-1} \Re\left\{ \left(\hat{\bar{z}}_k^{[m]}(i)\right)^* \hat{\alpha}_k^{[m]} e^{j2\pi i \hat{\bar{\nu}}_k^{[m+1]}/Q} \bar{t}_k^{[r]}(i) \bar{b}_i'\left(\hat{\bar{\tau}}_k^{[m]}\right) \right\}}{\sum_{i=0}^{LQ-1} \Re\left\{ \left(\hat{\bar{z}}_k^{[m]}(i)\right)^* \hat{\alpha}_k^{[m]} e^{j2\pi i \hat{\bar{\nu}}_k^{[m+1]}/Q} \bar{t}_k^{[r]}(i) \bar{b}_i''\left(\hat{\bar{\tau}}_k^{[m]}\right) \right\}}, \quad (2.59)$$

where $\bar{b}_i'(\bar{\tau}_k)$ and $\bar{b}_i''(\bar{\tau}_k)$ can be obtained using (2.35), by replacing the sequence $t_k^{[r]}$ with $t^{[s]}$. Finally, the updated value of α_k, $\hat{\alpha}_k^{[m+1]}$ is given by

$$\hat{\alpha}_k^{[m+1]} = \frac{1}{\sum_{i=0}^{LQ-1} |\bar{b}_i(\hat{\bar{\tau}}_k^{[m+1]})|^2 |\bar{t}_k^{[r]}(i)|^2} \sum_{i=0}^{LQ-1} \frac{\hat{\bar{z}}_k^{[m]}(i) \left(\bar{t}_k^{[r]}(i)\right)^* \left(\bar{b}_i(\hat{\bar{\tau}}_k^{[m+1]})\right)^*}{e^{j2\pi i \hat{\bar{\nu}}_k^{[m+1]}/Q}}. \quad (2.60)$$

The tabular form of the proposed ECM estimator for AF relaying can be obtained by modifying the initialization in Table 2.1 using (2.49) and (2.50) and by replacing $\hat{\mathbf{z}}_k^{[m]}$, $\hat{\nu}_k^{[m+1]}$, $\hat{\tau}_k^{[m+1]}$, and $\hat{\eta}_k^{[m+1]}$ in Table 2.1 by their counterparts, $\hat{\bar{\mathbf{z}}}_k^{[m]}$, $\hat{\bar{\nu}}_k^{[m+1]}$, $\hat{\bar{\tau}}_k^{[m+1]}$, and $\hat{\alpha}_k^{[m+1]}$, given in (2.55), (2.58), (2.59), and (2.60), respectively. By reapplying the above process, for $k = \{1, \cdots, K\}$, the estimates of the system parameters for all relays can be determined at $\mathbb{D}$. Similar to the DF case, the iterative process is

terminated when the difference between the LLF of two iterations is smaller than χ.

2.3.2.3 SAGE Estimator for AF Relaying

The parameter $\bar{\boldsymbol{\lambda}}$ is divided into K groups of $\bar{\boldsymbol{\lambda}}_k$ and the hidden data set, $\bar{\mathbf{x}}_k$, for $\bar{\boldsymbol{\lambda}}_k$ is defined as

$$\bar{\mathbf{x}}_k = \alpha_k\big(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}^{[s]}\big) \odot \bar{\mathbf{t}}_k^{[r]} + \mathbf{n}. \tag{2.61}$$

E-step While setting $\bar{\boldsymbol{\lambda}}_\ell = \hat{\bar{\boldsymbol{\lambda}}}_\ell^{[m]}$, $\forall \ell \neq k$, the expectation of the LLF of $\bar{\mathbf{x}}_k$, given $\bar{\boldsymbol{\lambda}}_k$, $\bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}_k | \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right)$, is determined as

$$\bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}_k | \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right) \triangleq \mathbb{E}\left\{\log f\left(\bar{\mathbf{x}}_k \middle| \bar{\boldsymbol{\lambda}}_k, \{\hat{\bar{\boldsymbol{\lambda}}}_\ell^{[m]}\}_{\ell \neq k}\right) \;\middle|\; \bar{\mathbf{y}}, \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right\}. \tag{2.62}$$

where

$$f\left(\bar{\mathbf{x}}_k \middle| \bar{\boldsymbol{\lambda}}_k, \{\hat{\bar{\boldsymbol{\lambda}}}_\ell^{[m]}\}_{\ell \neq k}\right) = f\left(\bar{\mathbf{x}}_k | \bar{\boldsymbol{\lambda}}_k\right) = \frac{1}{(\pi \tilde{\sigma}_n^2)^{LQ}} \exp\left\{\frac{\|\bar{\mathbf{x}}_k - \alpha_k(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}^{[s]}) \odot \bar{\mathbf{t}}_k^{[r]}\|^2}{\tilde{\sigma}_n^2}\right\}. \tag{2.63}$$

Substituting (2.63) into (2.62), we obtain

$$\bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}_k | \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right) = E_4 - \frac{1}{\tilde{\sigma}_n^2}\left\|\hat{\bar{\mathbf{x}}}_k^{[m]} - \alpha_k\big(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}^{[s]}\big) \odot \bar{\mathbf{t}}_k^{[r]}\right\|^2, \tag{2.64}$$

where

$$\hat{\bar{\mathbf{x}}}_k^{[m]} \triangleq \mathbb{E}\left\{\bar{\mathbf{x}}_k \middle| \bar{\mathbf{y}}, \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right\} = \bar{\mathbf{y}} - \sum_{\ell=1, \ell \neq k}^{K} \hat{\alpha}_\ell^{[m]}\left(\hat{\bar{\boldsymbol{\Lambda}}}_\ell^{[m]} \hat{\bar{G}}_\ell^{[m]} \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_\ell^{[r]}, \tag{2.65}$$

and $E_4 = -LQ \log(\pi \tilde{\sigma}_n^2)$ is a constant independent of $\bar{\boldsymbol{\lambda}}_k$.

M-step The estimate of $\bar{\boldsymbol{\lambda}}_k$ in the $(m+1)^{\text{th}}$ iteration, $\hat{\bar{\boldsymbol{\lambda}}}_k^{[m+1]}$, is determined as

$$\hat{\bar{\boldsymbol{\lambda}}}^{[m+1]} = \arg\max_{\bar{\boldsymbol{\lambda}}_k} \bar{\mathbf{N}}\left(\bar{\boldsymbol{\lambda}}_k \middle| \hat{\bar{\boldsymbol{\lambda}}}^{[m]}\right) = \arg\min_{\bar{\boldsymbol{\lambda}}_k} \left\|\hat{\bar{\mathbf{x}}}_k^{[m]} - \alpha_k\big(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}^{[s]}\big) \odot \bar{\mathbf{t}}_k^{[r]}\right\|^2. \tag{2.66}$$

For the proposed SAGE algorithm, by replacing $\hat{\bar{z}}_k^{[m]}(i)$ with $\hat{\bar{x}}_k^{[m]}(i)$, Eqs. (2.58), (2.59), and (2.60) can be used to estimate $\hat{\bar{\nu}}_k^{[m+1]}$, $\hat{\bar{\tau}}_k^{[m+1]}$, and $\hat{\alpha}_k^{[m+1]}$, respectively. The tabular form of the proposed SAGE estimator for AF relaying can be obtained

by modifying the initialization in Table 2.2, using (2.49) and (2.50), and by replacing $\hat{\mathbf{x}}_k^{[m]}$, $\hat{\nu}_k^{[m+1]}$, $\hat{\tau}_k^{[m+1]}$, and $\hat{\eta}_k^{[m+1]}$ by their counterparts, $\hat{\bar{\mathbf{x}}}_k^{[m]}$, $\hat{\bar{\nu}}_k^{[m+1]}$, $\hat{\bar{\tau}}_k^{[m+1]}$, and $\hat{\alpha}_k^{[m+1]}$, given in (2.65), (2.58), (2.59), and (2.60), respectively.[6]

2.3.3 Complexity Analysis of the Proposed Estimators

Throughout this chapter, computational complexity is defined as the number of additions plus multiplications. In this subsection the computational complexity of ML estimation in [40] and the proposed ECM and SAGE algorithms for DF relaying are analyzed. In order to avoid repetition, the case of AF relaying has been omitted, since in the AF scenario the computational complexity of the proposed estimators can be determined by using the number of iterations required by each algorithm and by adding the additional multiplications required due to the factor $\bar{t}_k^{[r]}(i)$ in (2.58), (2.59), and (2.60). The computational complexity of the ML algorithm, denoted by C_{ML} is calculated as

$$C_{\mathrm{ML}} = \underbrace{2K^2LQ + K^3 + KLQ}_{\boldsymbol{\nu}\,,\,\boldsymbol{\tau} \text{ in (2.20)}} + \underbrace{(\mathcal{N}2K/\kappa)\left[(2K^2+1)LQ + (1+K)(LQ)^2 + K^3\right]}_{\boldsymbol{\eta} \text{ in (2.20)}}, \tag{2.67}$$

where $\mathcal{N}$ denotes the number of alternating projection cycles used [45], and κ denotes the step size in the ML search in (2.20). The computational complexity of the proposed ECM algorithm, denoted by C_{ECM} is calculated as

$$\begin{aligned} C_{\mathrm{ECM}} = C_{\mathrm{I}} + K\varsigma[&\underbrace{12LQ + 4L_gLQ + 2}_{(2.32)} + \underbrace{10LQ + 8L_gLQ + 2}_{(2.38)} + \underbrace{6LQ + 4L_gLQ + 1}_{(2.40)} \\ &+ \underbrace{(K+2)LQ + (K+1)L^2Q + (K+1)(LQ)^2 + 1}_{(2.25)} + \underbrace{2Q(2L_g+1) + 2}_{(2.36)}], \end{aligned} \tag{2.68}$$

where C_{I} is the computational cost associated with determining the initial rough estimates given by $C_I = C_{\mathrm{ML}}|_{\kappa=\mu}$, μ is the coarse step size used to calculate the initial estimates for the proposed ECM algorithm, and ς denotes the average total number of iterations required. The value of ς for the ECM and SAGE algorithms have been determined through numerical simulations in Section 2.5 as illustrated in Fig. 2.10, e.g., in the case of DF relaying $\varsigma = 29$ and 16 for ECM and SAGE algorithms, respectively, with SNR = 15 and $K = 2$ relays. Similarly, the computa-

[6]Note that *Remark 2.4* above Section 2.3.2 also holds for AF relaying.

Table 2.3: CPU processing time for ML, ECM and SAGE with 4 Relays DF system at SNR = 20 dB using Intel Core 2 Quad 2.66 GHz processor.

Implemented Algorithm	CPU Processing Time (minutes)
ML [20]	179.725
ECM	0.2698
SAGE	0.1945

tional complexity of the proposed SAGE algorithm, denoted by C_{SAGE}, is calculated as

$$C_{\text{SAGE}} = C_{\text{I}} + K\varsigma[\underbrace{12LQ + 4L_gLQ + 2}_{\hat{\nu}_k^{[m+1]} \text{ in Table 2.2}} + \underbrace{10LQ + 8L_gLQ + 2}_{\hat{\tau}_k^{[m+1]} \text{ in Table 2.2}} + \underbrace{6LQ + 4L_gLQ + 1}_{\hat{\eta}_k^{[m+1]} \text{ in Table 2.2}}$$
$$+ \underbrace{KLQ + (K-1)L^2Q + (K-1)(LQ)^2}_{(2.46)} + \underbrace{2Q(2L_g+1)+2}_{(2.36)}]. \quad (2.69)$$

Based on (2.67), (2.68) and (2.69), the following remarks are in order:

Remark 2.5 *In order to reach the CRLB for the estimation of MCFOs, MTOs, and channel gains (See Fig. 2.6 and 2.7 in Section 2.5), the step size, κ for the ML in [40] and the proposed LS estimator needs to be very small, e.g., $\kappa = 10^{-5}$. This in turn significantly increases the computational complexity of these estimators given that the maximizations in* (2.20) *and* (2.50) *for DF and AF relaying, respectively, need to be carried out over a significantly larger set of possible values. However, a step size of say, $\mu = 10^{-2}$, suffices to obtain rough initial estimates for the proposed ECM and SAGE algorithms.*

Remark 2.6 *In order to quantitatively compare the computational complexity of the ML in [40] and the proposed ECM and SAGE estimators for DF relaying, we have evaluated C_{ML}, C_{ECM}, and C_{SAGE} in (2.67), (2.68) and (2.69), respectively, for $K = 4$ relays at an SNR of 20 dB. It is observed that even by considering the complexity associated to the initialization step of the proposed ECM and SAGE estimators, these schemes are 772 and 946 times more computationally efficient than the ML estimator in [40] carried out using AP, respectively. The computational complexity of the proposed algorithms is also evaluated using CPU execution time [115]. For the case of DF relaying networks, Table 2.3 depicts the execution times for the ML estimator in [40], and for the proposed ECM and SAGE estimators with SNR = 20 dB and $K = 4$ relays when an Intel Core 2 Quad 2.66 GHz processor processor with 4 GB byte of RAM is used. It can be observed from Table*

2.3 that compared to the ML estimator in [40] the proposed ECM and SAGE estimators are capable of estimating the overall network's synchronization parameters and channel gains approximately 666 *and* 924 *times more quickly.*

Remark 2.7 *The proposed ECM and SAGE algorithms need to apply the ML and LS estimators for initialization once only at system start-up. Afterwards, the estimates of previously transmitted frames may be used to update the new estimates since the timing and carrier frequency offsets do not rapidly change from frame to frame. This is due to the fact that the oscillator properties are mainly affected by temperature and other physical phenomena that do not rapidly fluctuate with time [116]. The CPU processing times shown in Table 2.3 for the proposed algorithms represent the worst case time required only for the first frame at system startup only. Later, ECM or SAGE algorithm can be initialized with the previous estimates and can quickly converge to the new estimates.*

2.4 ML Decoding

In order to decode the received signal at $\mathbb{D}$ in the presence of multiple impairments, an ML decoder for both DF and AF multi-relay cooperative systems is proposed.

2.4.1 Decoding in DF-Relaying Networks

During the DTP, the decoder at $\mathbb{D}$ evaluates the metric, $d_p(i)$, according to[7]

$$d_p(i) = \sum_{k=1}^{K} \hat{\eta}_k e^{j2\pi i \hat{\nu}_k/Q} \sum_{\ell=-D_g}^{D_g} c_p(\ell, i) g(\lfloor i/Q \rfloor T_s - \ell T - \hat{\tau}_k T), \qquad p = 1, \cdots, M^{D_g+1} \tag{2.70}$$

where $\mathbf{c}_p(i) \triangleq [c_p(-D_g, i), \cdots, c_p(D_g, i)]^T = [\hat{s}(\lfloor i/Q \rfloor - D_g), \cdots, \hat{s}(\lfloor i/Q \rfloor - 1), a_p(0), \cdots, a_p(D_g)]^T$ is a $(2D_g + 1) \times 1$ vector and D_g is a constant pulse shaping filter lag during the DTP. Note that $\mathbf{a}_p \triangleq [a_p(0), \cdots, a_p(D_g)]^T$ is a $(D_g + 1) \times 1$ vector of the p^{th} permutation of the M symbols within the constellation. Next, the error between the metric, $d_p(i)$, and the received signal, $y(i)$ at $\mathbb{D}$ is determined as

$$e_p(i) = |y(i) - d_p(i)|^2, \qquad p = 1, \cdots, M^{D_g+1}. \tag{2.71}$$

Thereafter, the summation of the errors, $e_p(i)$, in (2.71) that correspond to the n^{th} symbol, $\epsilon_p(n)$, is given by

[7] Subscript p refers to the p^{th} permutation, for $p = 1, \cdots, M^{D_g+1}$.

Table 2.4: The proposed ML decoder for DF relaying

for $i = 0, 1, \ldots, LQ-1$
 $d_p(i)=\sum_{k=1}^{K} \hat{\eta}_k e^{j2\pi i\hat{\nu}_k/Q} \sum_{\ell=-D_g}^{D_g} c_p(\ell, i) g(\lfloor i/Q \rfloor T_s - \ell T - \hat{\tau}_k T)$
 $e_p(i) = |y(i) - d_p(i)|^2$
 if $(\text{mod}(\frac{i+1}{Q}) == 0)$
  $\epsilon_p = \sum_{q=0}^{Q-1} e_p(i-q)$
  $p_{\text{sel}} = \arg\min_p \ \epsilon_p$
  $\hat{s}(\frac{i+1}{Q} - 1) = a_{p_{\text{sel}}}(0)$
 end
end

$$\epsilon_p(n) = \sum_{q=0}^{Q-1} e_p(nQ+q), \qquad p = 1, \cdots, M^{D_g+1}. \tag{2.72}$$

The permutation that results in the smallest error is denoted by p_{sel}, and is selected as

$$p_{\text{sel}} = \arg\min_{p} \epsilon_p(n), \qquad p = 1, \cdots, M^{D_g+1}. \tag{2.73}$$

Finally, using p_{sel} the n^{th} received symbol is decoded as $\hat{s}(n) = a_{p_{\text{sel}}}(0)$, where $a_{p_{\text{sel}}}(0)$ is the zero$^{\text{th}}$ element of the vector $\mathbf{a}_{p_{\text{sel}}}$ corresponding to p_{sel}. The proposed ML decoder for DF relaying is summarized in Table 2.4.

2.4.2 Decoding in AF-Relaying Networks

Similar to the DF case, the estimates $\hat{\bar{\boldsymbol{\tau}}}$, $\hat{\bar{\boldsymbol{\nu}}}$ and $\hat{\boldsymbol{\alpha}}$ are used to decode the received signal. The decoder at $\mathbb{D}$ in the AF system evaluates the metric

$$\bar{d}_p(i) = \sum_{k=1}^{K} \hat{\alpha}_k e^{j2\pi i\hat{\bar{\nu}}_k/Q} \sum_{\ell=-D_g}^{D_g} c_p(\ell, i) g(\lfloor i/Q \rfloor T_s - \ell T - \hat{\bar{\tau}}_k T). \tag{2.74}$$

Using the metric in (2.74) and the same steps as outlined in Section 2.4.1, the n^{th} source symbol, $\hat{s}(n)$, can be decoded. The tabular form of the ML decoder for AF relaying is given by replacing $d_p(i)$ with $\bar{d}_p(i)$ in Table 2.4.

Remark 2.8 *The ML decoder outlined above is derived to showcase that the estimates obtained during the TP using the proposed estimators can be applied to effectively decode the received signal in multi-relay cooperative networks. Since, it is a well-known that the complexity of ML decoding increases exponentially with*

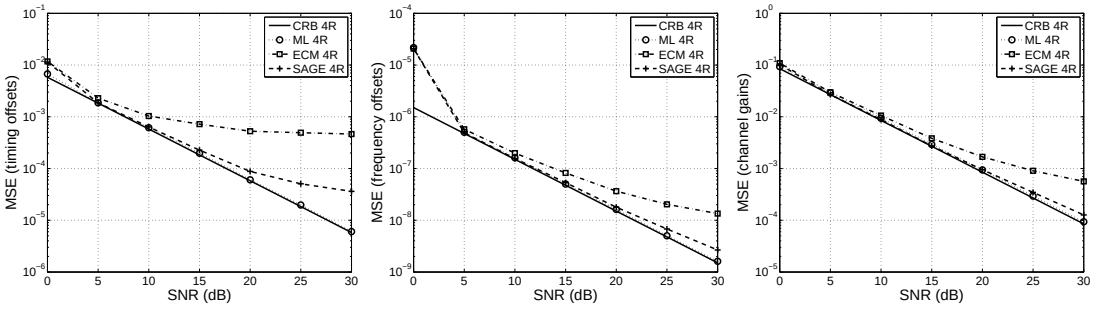
constellation size, the design of more computationally efficient decoders for cooperative networks in the presence of MCFOs, MTOs, and unknown channel gains is addressed in Chapter 3.

2.5 Simulation Results

In this section, we present the simulation results to evaluate the performance of our estimators. We use *quadrature phase-shift keying* (QPSK) modulation. Without loss of generality, we assume $\sigma_u^2 = \sigma_w^2 = 1/\text{SNR}$. The propagation loss is modeled as $(d/d_0)^{-m}$, where d is the distance between transmitter and receiver, d_0 is the reference distance, and m is the path loss exponent [7]. The following simulations are based on TS length, $L = 64$, $\sigma_\rho^2 = \sigma_\eta^2 = 1$, $d_0 = 1$km, and $m = 2.7$, which corresponds to urban area cellular networks. The timing offsets at $\mathbb{D}$, $\boldsymbol{\tau}^{[rd]}$ are assumed to be uniformly distributed over the range $(-0.5, 0.5)$. Based on the methodology in [1, 18], the timing offset estimation errors from $\mathbb{S} - \mathbb{R}$, $\epsilon_k^{[sr]}$, is assumed to follow a Gaussian distribution, i.e, $\epsilon_k^{[sr]} \sim \mathcal{N}(0, \sigma_\tau^2)$, where σ_τ^2 is set to the lower bound on the variance of timing offset estimation error in point-to-point systems [7, p. 328]. Carrier frequency offsets for DF relaying at $\mathbb{D}$, $\boldsymbol{\nu}^{[rd]}$, are uniformly distributed in the full acquisition range $(-0.5, 0.5)$. For AF relaying, since carrier frequency offsets from source to relays, $\boldsymbol{\nu}^{[sr]}$, are carried over to the destination, $\boldsymbol{\nu}^{[sr]}$ and $\boldsymbol{\nu}^{[rd]}$ have the range $(-0.25, 0.25)$ in order to limit the total frequency offset from source to destination $\boldsymbol{\nu}^{[sd]}$ to the range $(-0.5, 0.5)$, i.e., full acquisition. Distinct phase shift keying training sequences are generated at $\mathbb{S}$ and all relays similar to [25]. $d^{[sr]}$ and $d^{[rd]}$ are used to denote the $\mathbb{S}$-$\mathbb{R}$ and $\mathbb{R}$-$\mathbb{D}$ distances, respectively. Finally, the *mean-square error* (MSE) performance of various estimators and the *bit error rate* (BER) performance of the overall multi-relay cooperative network is detailed in the following subsections.

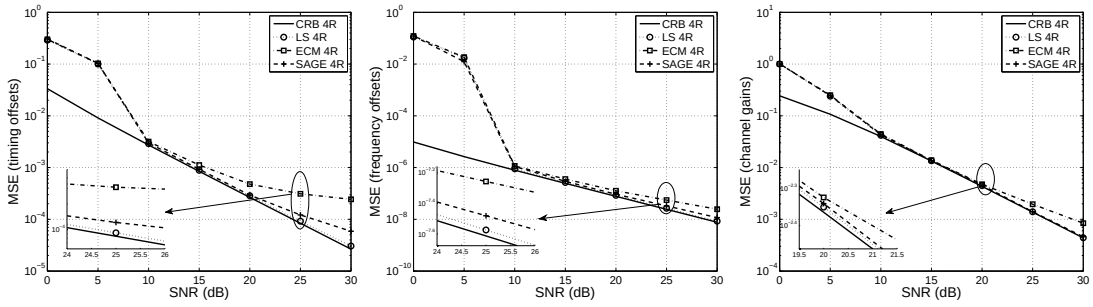
2.5.1 Estimator Performance

Specific channels are used for the following simulations, i.e., $\boldsymbol{\rho} = [.279 - .9603j,\ .8837 + .4681j,\ -.343 + .732i,\ -.734 - .451i]^T$ and $\boldsymbol{\eta} = [.7820 + .6233j, .9474 - .3203j, -.2413 + .724i, .5141 - .893i]^T$ similar to [2,12,21]. Unless otherwise specified, $K = 4$ relays, $Q = 2$ in the TP, and $d^{[sr]} = d^{[rd]} = 1$ km are used. The remaining parameters are set as $\chi = 0.001$, $\delta = 0.0001$, $\kappa = 10^{-5}$, $\mathcal{N} = 2$ and 3 for DF and AF relaying, respectively, $\mu = 10^{-2}$, and $L_g = 10$. Finally, the MSE for the estimation of a parameter say, frequency offset, $\boldsymbol{\nu}$, is defined as the average MSE



(a) (b) (c)

Figure 2.6: MSE and CRLBs of (a) MTOs, (b) MCFOs and (c) channel coefficients estimation as a function of SNR (dB) for DF relaying.



(a) (b) (c)

Figure 2.7: MSE and CRLBs of (a) MTOs, (b) MCFOs and (c) channel coefficients estimation as a function of SNR (dB) for AF system

over all the simulations runs, i.e., $\text{MSE}(\boldsymbol{\nu}) = \sum_{m=1}^{10^4} \sum_{k=1}^{K} \left(\hat{\nu}_k^{[m]} - \nu_k^{[m]} \right)^2 / 10^4$.

Figs. 2.6–2.7 (a), (b), and (c) show the CRLB and the MSE for the estimation of MTOs, MCFOs, and channel gains for DF and AF relaying, respectively. It is shown that the MSEs of the ML and proposed LS estimators for both DF and AF relaying are close to their CRLBs at mid-to-high SNRs. In comparison, the proposed ECM and SAGE estimators are close to the CRLB at mid-SNR values but exhibit some small performance degradation with respect to the CRLB when estimating MCFOs, MTOs, and channel gains at high SNR. In addition, Figs. 2.6(a) and 2.7(a) indicate that while estimating MTOs at high SNR, the MSEs of the proposed ECM and SAGE estimators exhibit an error floor. This error floor is caused by the Taylor series approximations in Eqs. (2.30) and (2.34) as well as the approximation in Eq. (2.36), which are used to linearize the LLF under consideration. However, as shown in Section 2.3.3, compared to the ML and proposed LS schemes, the proposed ECM and SAGE estimators significantly reduce the computational complexity associated with estimating impairments in cooperative networks. Moreover, at low SNR for AF relaying, the proposed LS, ECM, and SAGE estimators demonstrate poor performance due to the considerable timing

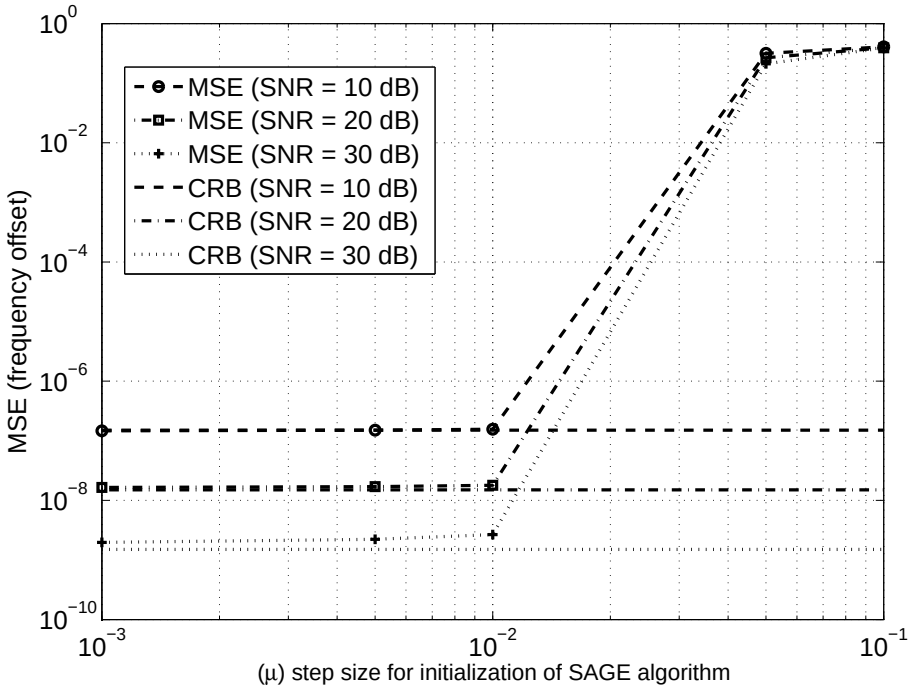


Figure 2.8: MSE of the frequency offset estimation vs. initialization of SAGE algorithm with different values of coarse step size, μ, in DF cooperative networks.

offset estimation error from the source to relays and the noise at the relays which is amplified and forwarded to the destination. Finally, Figs. 2.6 and 2.7 show that the proposed SAGE estimator outperforms the ECM algorithm for all SNR values.

Fig. 2.8 presents the impact of the initialization of the SAGE algorithm on the estimator's performance for different step size values, μ. It is shown that decreasing the step size from $\mu = 10^{-2}$ results in diminishing returns in the frequency offset estimation accuracy since the MSE of the proposed SAGE estimator for $\mu = 10^{-2}$ is already close to the CRLB. Thus, it can be concluded that for $\mu = 10^{-2}$ the proposed SAGE estimator is initialized in the region of a local maximum, which turns out to be the global maximum and converges to the true estimates. Consequently, in all the simulations in this section the step size, $\mu = 10^{-2}$, is used to initialize the proposed ECM and SAGE estimators[8]. Note that results similar to that of Fig. 2.8 are observed in the case of AF relaying cooperative networks and are omitted from this chapter to avoid repetition.

Fig. 2.9 compares the frequency offset estimation MSE of the proposed SAGE estimator against the MSE of the initial estimates. This result further shows that obtaining the frequency offsets using the proposed SAGE algorithm significantly improves estimation accuracy. In addition, unlike the ML and LS estimators, this

[8]Note that the ML estimator in [40] requires an exhaustive search with very small step size values, e.g., 10^{-5}, to reach the CRLB as explained in *Remark 2.5*.

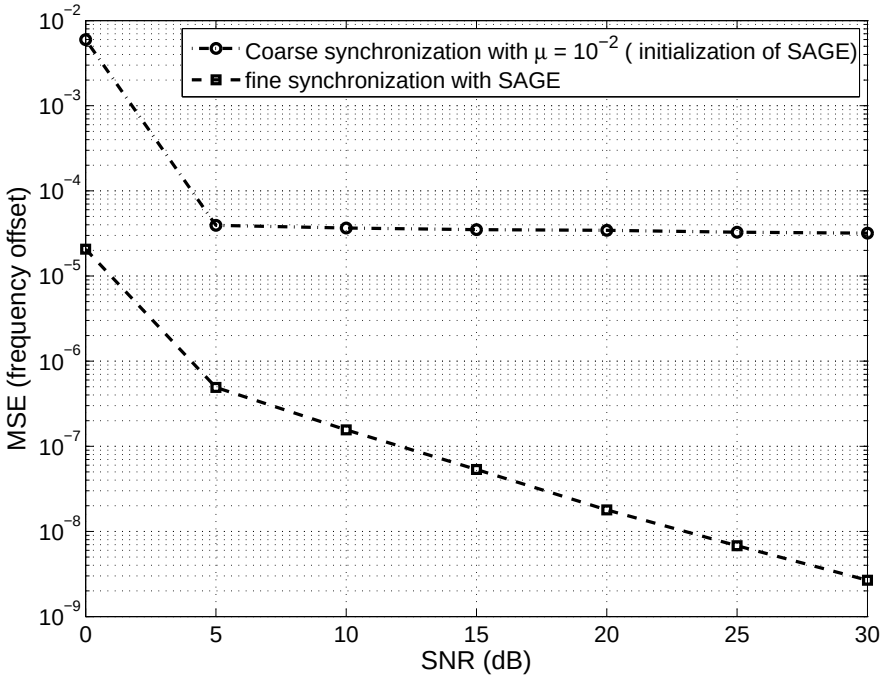


Figure 2.9: MSE of the frequency offset estimation with coarse initialization ($\mu = 10^{-2}$) and fine estimation using SAGE algorithm in DF cooperative networks.

improvement in estimation accuracy is achieved without performing an exhaustive search over a large set of possible frequency offset values with small step size values, e.g., 10^{-5}.

Fig. 2.10 shows the average number of iterations required by the proposed ECM and SAGE algorithms to converge in DF-relaying networks. It can be observed that at an SNR of 20 dB, the average number of iterations required by the SAGE estimator is 2 and 3.4 times fewer than that of ECM algorithm for networks with 2 and 4 relays, respectively. Note that similar results are obtained for the case of AF relaying.

Fig. 2.11 shows the CRLB for the frequency and timing offset estimation for AF relaying when the relays are located at different physical locations: estimation performance slightly improves by moving the relays closer to $\mathbb{D}$, i.e., $d^{[sr]} = 1.3$km and $d^{[rd]} = 0.7$km, due to lesser propagation loss from relays to $\mathbb{D}$. However, the performance degrades by moving the relays closer to $\mathbb{S}$, i.e., $d^{[sr]} = 0.7$km and $d^{[rd]} = 1.3$km, due to larger propagation loss from the relays to $\mathbb{D}$. Note that the improvement in the estimation performance is lesser, while moving the relays closer to $\mathbb{D}$, due to the additional amplification of noise at $\mathbb{D}$.

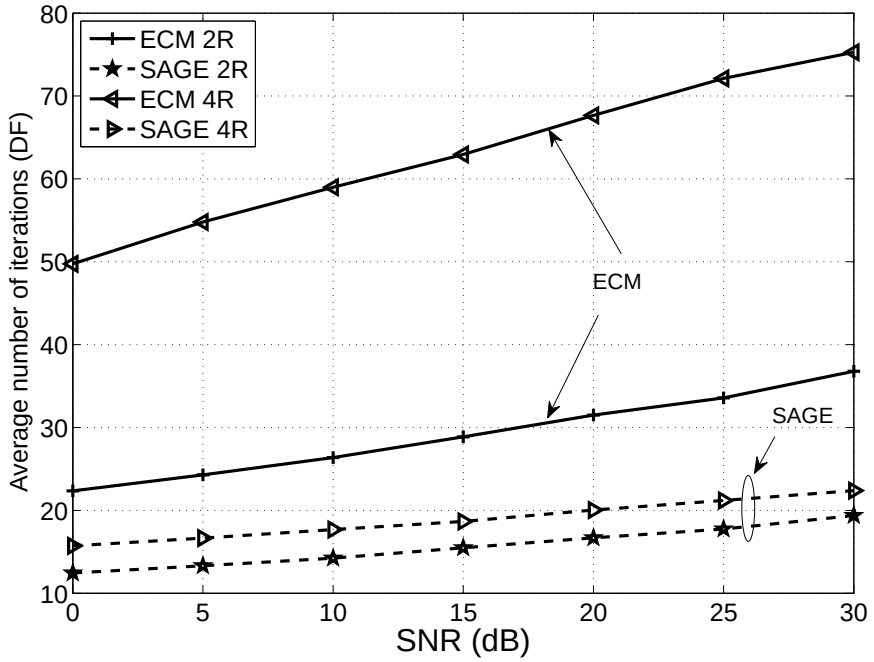


Figure 2.10: Average number of iterations for ECM and SAGE in DF cooperative networks for $K = 2$ and 4 relays.

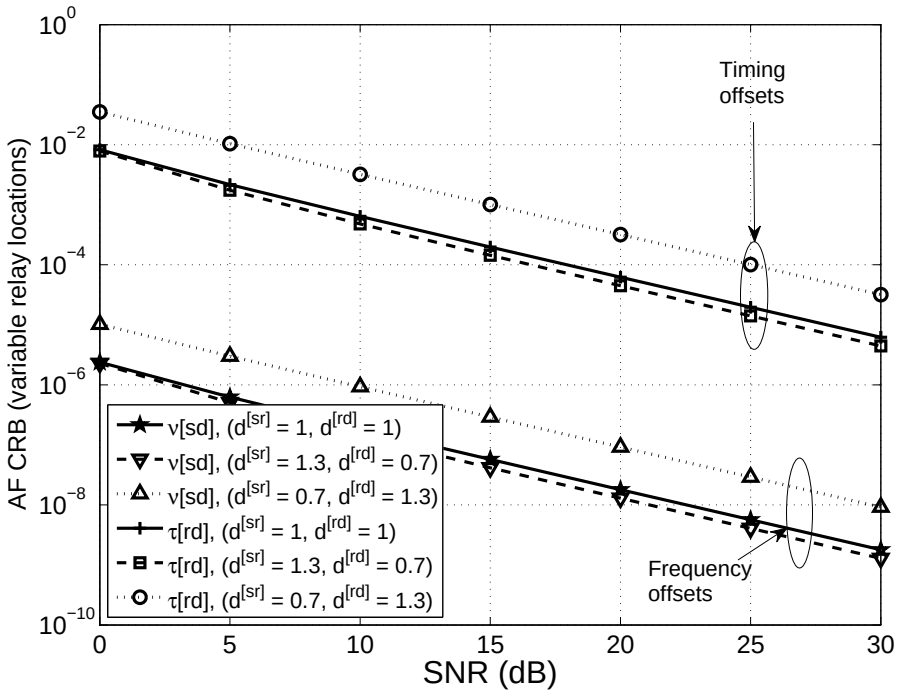
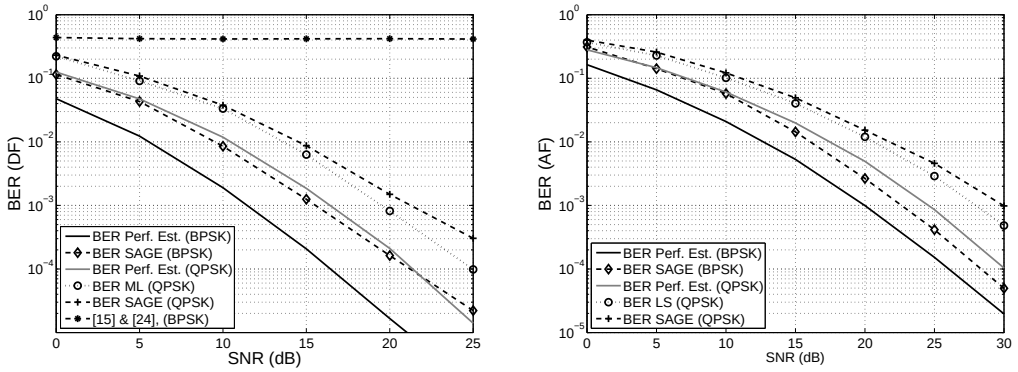


Figure 2.11: CRLB for the MCFOs and MTOs with AF relays at different locations.

2.5.2 Cooperative System Performance

The channel gains from the source to relays and from the relays to destination are modeled as *independent and identically distributed (i.i.d)* complex Gaussian ran-



(a) (b)

Figure 2.12: (a) BER performance for a DF cooperative system with $K = 2$. (b) BER performance for an AF cooperative system with $K = 2$.

dom variables with $\mathcal{CN}(0, 1)$. We use $Q = 4$ for the DTP and $D_g = 4$ in (2.70) and (2.74) for the proposed ML decoder. Cooperative communication networks with 2 relays are considered, where the relays are distributed throughout the network. Fixed gain relaying is applied for AF relaying as shown in Section 2.1.1.1. We assume a training sequence length of 80 symbols and a frame length of $L = 450$ symbols.

Figs. 2.12(a) and 2.12(b) show the BER performance with 2 relays for an uncoded DF and AF cooperative networks for *binary phase-shift keying* (BPSK) and QPSK, respectively. The results show that the BER performance of a DF relaying cooperative network using QPSK is within 2 dB of the ideal case of perfect impairment estimation when using the ML and SAGE estimator in combination with the proposed ML decoder. However, at high SNR the BER plot corresponding to the SAGE estimator deviates from that of the ML estimator due to the error floor of the proposed SAGE estimator as also depicted in Fig. 2.6.

Similar results are obtained for the case of AF relaying, where the the gap between the cases of imperfect and perfect impairment estimation for BPSK and QPSK is in the range of 2-2.5 dB for both the proposed LS and SAGE estimators at low-to-mid SNRs[9]. This gap increases at high SNR for the proposed SAGE estimator due to the error floor of this estimator at high SNR as illustrated in Fig. 2.7. In Fig. 2.12(a) we also plot the BER results for a cooperative system that first employs the re-synchronization filter in [25] to compensate MTOs and then attempts to remove MCFOs by employing the algorithm in [48]. This plot, which is denoted by "[14] & [17] 2R" shows that such an approach fails to decode the received signal at $\mathbb{D}$ since the re-synchronization filter in [25] fails to compensate

[9]Similar results are obtained for the proposed ECM estimator.

MTOs in the presence of MCFOs. Subsequently, the algorithm in [48] fails to nullify MCFOs, since the input signal is corrupted by MTOs. This corroborates our claim that previously proposed algorithms cannot decode the received signal in the presence of both MCFOs and MTOs. Finally, we note that in the case of BPSK the application of the proposed SAGE estimators and ML decoders results in an overall cooperative network BER of below 10^{-3} for SNRs greater than 16dB and 24dB for the DF and AF systems, respectively.

2.6 Conclusions

In this chapter, the training and data transmission methods for both DF-and AF-relaying multi-relay cooperative networks affected by MCFOs, MTOs, and unknown channel gains are presented. New closed-form FIM and CRLB expressions for the multiple parameter estimation problem are derived. The derived FIM shows that there exists coupling between the estimation errors of MCFOs, MTOs, and channel gains, which establishes that these parameters must be jointly estimated at the destination. In order to reduce overhead and complexity, two iterative estimators based on the ECM and SAGE algorithms are derived and their performance is compared against the CRLBs. Though global convergence of the proposed ECM and SAGE algorithm cannot be shown analytically, numerical simulations indicate that through proper initialization using an LS estimator the proposed estimators can obtain MCFOs, MTOs, and unknown channel gains jointly at the destination. In addition, it is established through computational complexity analyses that at SNR of 20 dB for a 4-relay cooperative network, the proposed ECM and SAGE estimators are each over two orders of magnitude more computationally efficient than the previously proposed ML estimator in [40]. Next, an ML approach is proposed to decode the received signal at the destination for both DF and AF systems. Simulation results show that the combination of proposed estimators and ML decoder result in BER performance that is within 2-2.5 dB of that of a perfectly synchronized cooperative system.

Chapter 3

Data-Aided Synchronization in STBC based AF Systems

This chapter seeks to investigate the application of distributed space time block codes (DSTBCs) in AF cooperative communication systems in order to achieve full order diversity in the presence of the channel and synchronization impairments. In Chapter 2, we addressed the problem of timing and carrier synchronization in AF and DF cooperative networks. However, the transceiver designs and the estimation algorithms proposed in Chapter 2 cannot be applied to DSTBC-AF cooperative networks due to the particular processing required at the relays and destination to enable the transmission and decoding of DSTBC in AF cooperative networks. Consequently, the AF transceiver design in Chapter 2 does not achieve full cooperative diversity. Moreover, the main focus of Chapter 2 was on the design of the estimation algorithms for the joint estimation of multiple impairments. The design of computationally efficient compensation algorithms is addressed in this chapter. Chapter 3 is organized as follows:

In Section 3.1, taking into account the multiple synchronization and channel impairments, a new transceiver structure at the relays and a novel receiver design at the destination in *distributed space-time block code (DSTBC)* based *amplify-and-forward (AF)* cooperative networks are proposed. In Section 3.2, the CRLBs for the multiple parameter estimation problem are derived. In Section 3.3, a *least squares (LS)* estimator for the multi-parameter estimation problem is derived. In order to significantly reduce the receiver complexity at the destination, a *differential evolution (DE)* based estimation algorithm is applied and the initialization and constraints for the convergence of the proposed DE algorithm are investigated. Next, the computational complexity of the proposed LS and DE based algorithms is analyzed. In Section 3.4, novel optimal and sub-optimal minimum mean-square

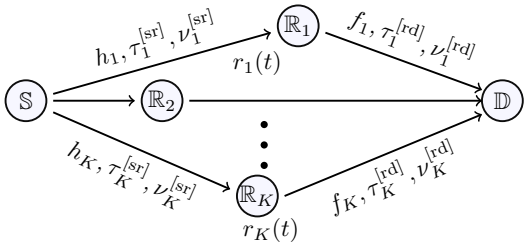


Figure 3.1: System model for the AF cooperative network.

error (MMSE) receiver designs at the destination node are proposed in order to detect the signal from the multiple relays in the presence of unknown channels, MTOs, and MCFOs. In order to reduce overhead, a low complexity version of an MMSE compensation algorithm, denoted by *L-MMSE*, is proposed that can detect the received signals using only the overall channel gains, MTOs, and MCFOs estimates obtained at the destination. In Section 3.5, numerical and simulation results that investigate the MSE and BER performances of the proposed estimators and compensation algorithms, respectively, are presented. Finally, Section 3.6 concludes the chapter and summarizes its key findings.

3.1 System Model and Transceiver Design

As shown in Fig. 3.1, a half-duplex AF cooperative system with one source node, $\mathbb{S}$, K relays, $\mathbb{R}_1, \ldots, \mathbb{R}_K$, and a single destination node, $\mathbb{D}$ is considered. Each node is equipped with a single omnidirectional antenna. The channels from the source to the kth relay and the kth relay to the destination are denoted by h_k and f_k, respectively. The index $k = 1, \ldots, K$ is used for the K relays. In Fig. 3.1, τ_k and ν_k are used to denote timing and carrier frequency offsets, respectively. Throughout this chapter, the following set of system design assumptions is considered:

A1. Each transmission frame from source to relays and relays to destination is comprised of two periods: a *training period (TP)* followed by a *data transmission period (DTP)*. Without loss of generality, it is assumed that during the TP, unit-amplitude phase shift keying (PSK) *training signals (TSs)* are transmitted from the source to the kth relay and from the kth relay to the destination, $\forall\ k$. The TSs from all the relays are linearly independent. Such TSs have also been considered previously, e.g., in [1, 2, 25].

A2. Quasi-static and frequency flat fading channels are considered, i.e., the channel gains do not change over the length of a frame but change from frame to frame according to a complex Gaussian distribution, $\mathcal{CN}(0, \sigma^2)$. The use of

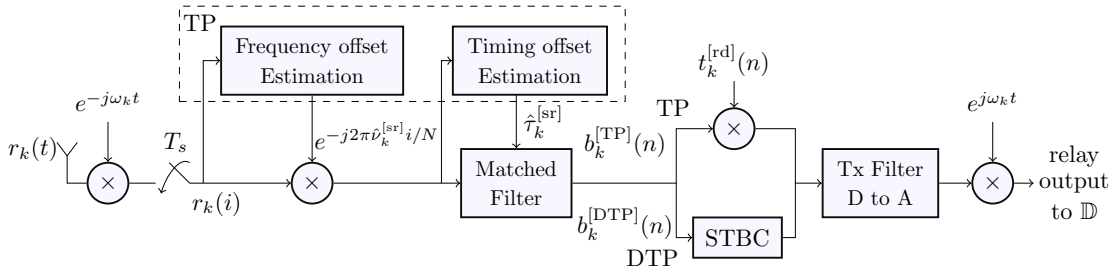


Figure 3.2: Proposed AF transceiver design at the kth relay transceiver.

such channels is motivated by prior research in this field [1,2,12,21–25]. Moreover, the assumption of frequency flat channels can be broadened to frequency selective channels by employing OFDM.

A3. The timing and carrier frequency offsets are modeled as *unknown* deterministic parameters over the frame length, which is similar to the approach adopted in [25], [2], and [21].

The proposed system model at the relays and destination is detailed below.

3.1.1 Proposed Transceiver Model at the Relays

The block diagram of the AF transceiver at the kth relay is shown in Fig. 3.2. During the first time slot, the source transmits the training and data symbols to all relays. The received signal at the kth relay, $r_k(t)$, is down converted by oscillator frequency ω_k. The received signal is oversampled by a factor N, such that $T = T_s N$, where T is the symbol duration and T_s is the sampling period. L denotes the number of training symbols during the TP and the number of data symbols during DTP. For clarity, here, the indices $n = 0, 1, \ldots, L-1$ and $i = 0, 1, \ldots, LN-1$ are used to denote the symbols and the T_s-spaced samples, respectively. Each relay uses the training signal from the source node to estimate the source-to-relay carrier timing and frequency offsets. The proposed transceiver design and signal model corresponding to the TP and DTP are detailed in the following two subsections.

3.1.1.1 Training Model at the Relays

During the TP, the sampled received signal vector at the kth relay, $\mathbf{r}_k \triangleq [r_k(0), \ldots, r_k(LN-1)]^T$, is given by

$$\mathbf{r}_k = h_k \mathbf{\Lambda}_k^{[sr]} \mathbf{G}_k^{[sr]} \mathbf{t}^{[sr]} + \mathbf{u}_k, \qquad k = 1, \ldots, K \tag{3.1}$$

where,

- h_k denotes the *unknown* channel gain from the source to the kth relay, which changes from one frame to another frame following the distribution, $h_k \sim \mathcal{CN}(0, \sigma_h^2)$,
- $\mathbf{\Lambda}_k^{[\mathrm{sr}]} \triangleq \mathrm{diag}\left(\left[e^{j2\pi\nu_k^{[\mathrm{sr}]}(0)/N}, \ldots, e^{j2\pi\nu_k^{[\mathrm{sr}]}(LN-1)/N}\right]\right)$ is a diagonal $LN \times LN$ matrix, $\nu_k^{[\mathrm{sr}]}$ denotes the *unknown* carrier frequency offset, normalized by the symbol duration T, between the source and the kth relay,
- $\mathbf{G}_k^{[\mathrm{sr}]}$ is the $LN \times L$ matrix of the samples of the pulse shaping filter such that $\left[\mathbf{G}_k^{[\mathrm{sr}]}\right]_{i,n} \triangleq g_{rrc}(iT_s - nT - \tau_k^{[\mathrm{sr}]}T)$, $\tau_k^{[\mathrm{sr}]}$ denotes the normalized *unknown* timing offset between the source and the kth relay, $g_{rrc}(t)$ stands for the root raised-cosine pulse shaping function,
- $\mathbf{t}^{[\mathrm{sr}]} \triangleq \left[t^{[\mathrm{sr}]}(0), \ldots, t^{[\mathrm{sr}]}(L-1)\right]^T$ is the TS transmitted from source to relays, and
- $\mathbf{u}_k \triangleq [u_k(0), \ldots, u_k(LN-1)]^T$, $u_k(i)\ \forall\ i$, denotes the zero-mean complex additive white Gaussian noise (AWGN) at the ith sample of the received signal, i.e., $u_k(i) \sim \mathcal{CN}(0, \sigma_{u_k}^2)$.

Note that the noise terms at the relays are assumed to be mutually uncorrelated, i.e., $\mathbb{E}\{\mathbf{u}_k^H \mathbf{u}_{\bar{k}}\} = 0$, for $k \neq \bar{k}$.

The estimates of the timing and frequency offsets at the kth relay, $\hat{\tau}_k^{[\mathrm{sr}]}$ and $\hat{\nu}_k^{[\mathrm{sr}]}$, respectively, are used for matched filtering and frequency offset correction at the kth relay, as shown in the Fig. 3.2. Standard estimation and synchronization techniques for carrier and timing synchronization in point-to-point communication systems, e.g., [7, 117], can be applied to obtain $\hat{\tau}_k^{[\mathrm{sr}]}$ and $\hat{\nu}_k^{[\mathrm{sr}]}$, $\forall k$, and will not be discussed here.

Let the timing and frequency offset estimation errors between the source and the kth relay be denoted by $\delta_{\tau_k^{[\mathrm{sr}]}} = \tau_k^{[\mathrm{sr}]} - \hat{\tau}_k^{[\mathrm{sr}]}$ and $\delta_{\nu_k^{[\mathrm{sr}]}} = \nu_k^{[\mathrm{sr}]} - \hat{\nu}_k^{[\mathrm{sr}]}$, respectively. According to the results in [1], $\delta_{\tau_k^{[\mathrm{sr}]}}$ and $\delta_{\nu_k^{[\mathrm{sr}]}}$ are Gaussian distributed, i.e., $\delta_{\tau_k^{[\mathrm{sr}]}} \sim \mathcal{N}(0, \sigma^2_{\tau_k^{[\mathrm{sr}]}})$ and $\delta_{\nu_k^{[\mathrm{sr}]}} \sim \mathcal{N}(0, \sigma^2_{\nu_k^{[\mathrm{sr}]}})$ [1], where $\sigma^2_{\tau_k^{[\mathrm{sr}]}}$ and $\sigma^2_{\nu_k^{[\mathrm{sr}]}}$ are timing and frequency offset estimation error variances, respectively. In this chapter, $\sigma^2_{\tau_k^{[\mathrm{sr}]}}$ and $\sigma^2_{\nu_k^{[\mathrm{sr}]}}$ are set to their respective lower bounds in point-to-point systems given in [7, Ch. 6]. Taking into account the timing and frequency offset estimation errors, the vector of imperfectly synchronized received signals at the kth relay, $\mathbf{b}_k^{[\mathrm{TP}]} \triangleq [b_k^{[\mathrm{TP}]}(0), \ldots, b_k^{[\mathrm{TP}]}(L-1)]^T$, as shown in Fig. 3.2, can be written as [7], [118]

$$\mathbf{b}_k^{[\mathrm{TP}]} = h_k \mathbf{x}_k^{[\mathrm{TP}]} + \mathbf{n}_k, \tag{3.2}$$

where

- $\mathbf{x}_k^{[\mathrm{TP}]} \triangleq \mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} \mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}} \mathbf{t}^{[\mathrm{sr}]}$, $\mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} \triangleq \mathrm{diag}\left(\left[e^{j2\pi\delta_{\nu_k^{[\mathrm{sr}]}}(0)}, \ldots, e^{j2\pi\delta_{\nu_k^{[\mathrm{sr}]}}(L-1)}\right]\right)$ is an $L \times L$ diagonal matrix, $\mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}$ is an $L \times L$ matrix, such that $\left[\mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\right]_{n,\bar{n}} \triangleq g_{rc}(nT - \bar{n}T - \delta_{\tau_k^{[\mathrm{sr}]}}T)$, for $n, \bar{n} = 0, \cdots, L-1$, $g_{rc}(t)$ stands for the raised-cosine function, which results from the convolution of $g_{rrc}(t)$ and root raised-cosine matched filter,
- $\mathbf{n}_k \triangleq \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]} \hat{\mathbf{\Lambda}}_k^{[\mathrm{sr}]} \mathbf{u}_k$ is the matched filtered noise, $\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}$ is an $L \times LN$ matrix such that $\left[\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}\right]_{n,i} \triangleq g_{rrc}(nT - iT_s - \hat{\tau}_k^{[\mathrm{sr}]}T)$, and $\hat{\mathbf{\Lambda}}_k^{[\mathrm{sr}]}$ is an $LN \times LN$ matrix such that $\hat{\mathbf{\Lambda}}_k^{[\mathrm{sr}]} \triangleq \mathbf{\Lambda}_k^{[\mathrm{sr}]}\Big|_{\nu_k^{[\mathrm{sr}]} = \hat{\nu}_k^{[\mathrm{sr}]}}$.

Subsequently, the kth relay applies its unique training signal, $\mathbf{t}_k^{[\mathrm{rd}]} \triangleq \left[t_k^{[\mathrm{rd}]}(0), \ldots, t_k^{[\mathrm{rd}]}(L-1)\right]^T$ to the matched filtered output, $\mathbf{b}_k^{[\mathrm{TP}]}$, performs the pulse shaping operation, and up converts the analog signal by the oscillator frequency ω_k (see Fig. 3.2). Note that the proposed AF transceiver in Fig. 3.2 is less complex than a DF relaying transceiver since there is no channel estimation, equalization, and decoding blocks after the matched filtering block.

3.1.1.2 Data Transmission Model at the Relays

During the DTP, the modulated data symbol vector, $\mathbf{s} \triangleq [s(0), s(1), \ldots, s(L-1)]^T$, is transmitted from the source to the relays. The sampled received signal vector at the kth relay during the DTP is given by (3.1), where $\mathbf{t}^{[\mathrm{sr}]}$ is replaced by the data symbol vector, $\mathbf{s}$. After performing timing and frequency synchronization, the matched filtered signal vector at the kth relay in the DTP, $\mathbf{b}_k^{[\mathrm{DTP}]} \triangleq \left[b_k^{[\mathrm{DTP}]}(0), \ldots, b_k^{[\mathrm{DTP}]}(L-1)\right]^T$, is given by [7], [118]

$$\mathbf{b}_k^{[\mathrm{DTP}]} = h_k \mathbf{x}_k^{[\mathrm{DTP}]} + \mathbf{n}_k, \tag{3.3}$$

where $\mathbf{x}_k^{[\mathrm{DTP}]} \triangleq \mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} \mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}} \mathbf{s}$ and $\mathbf{n}_k$ are defined below (3.2). During the DTP, before forwarding its signal to the destination, each relay applies a DSTBC to the corresponding matched filter output, performs the pulse shaping operation, and up converts the analog signal by the oscillator frequency ω_k (see Fig. 3.2).

3.1.2 Proposed Receiver Model at the Destination

The block diagram for the proposed receiver at the destination is shown in Fig. 3.3. As described in Section 3.1.1, each relay applies its independent training and

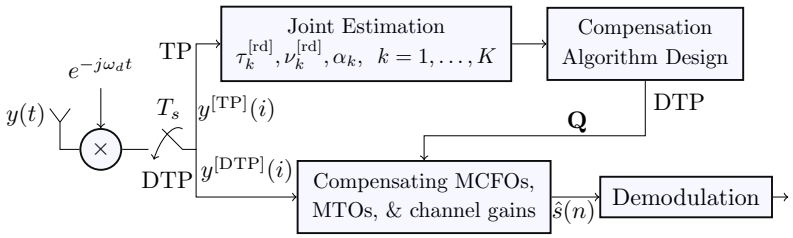


Figure 3.3: Proposed AF receiver design at the destination.

DSTBC. Consequently, the received signal at the destination, $y(t)$, is a superposition of signals from the different relays. At the destination, $y(t)$ is down converted by the oscillator frequency ω_d and then oversampled by the factor N. Furthermore, the multiple channel gains, MTOs, and MCFOs are estimated at the receiver using the known and linearly independent TSs transmitted from the relays. The proposed receiver design and system model at the destination during the TP and the DTP is outlined in the following two subsections.

3.1.2.1 Training Model at the Destination

During the TP, the sampled received signal vector at the destination, $\mathbf{y}^{[\mathrm{TP}]} \triangleq [y^{[\mathrm{TP}]}(1), \ldots, y^{[\mathrm{TP}]}(LN-1)]^T$, is given by

$$\mathbf{y}^{[\mathrm{TP}]} = \sum_{k=1}^{K} \zeta_k f_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \left(\mathbf{b}_k^{[\mathrm{TP}]} \odot \mathbf{t}_k^{[\mathrm{rd}]} \right) + \mathbf{w}, \tag{3.4}$$

where

- f_k denotes the *unknown* channel gain from the kth relay to the destination, which changes from frame to frame following the distribution $f_k \sim \mathcal{CN}(0, \sigma_f^2)$,
- $\zeta_k = 1/\sqrt{\sigma_h^2 + \sigma_{u_k}^2}$ satisfies the kth relay's power constraint,
- $\mathbf{\Lambda}_k^{[\mathrm{rd}]} \triangleq \mathrm{diag}\left([e^{j2\pi\nu_k^{[\mathrm{rd}]}(0)/N}, \ldots, e^{j2\pi\nu_k^{[\mathrm{rd}]}(LN-1)/N}] \right)$ is an $LN \times LN$ matrix, $\nu_k^{[\mathrm{rd}]}$ denotes the normalized *unknown* frequency offset from the kth relay to the destination,
- $\left[\mathbf{G}_k^{[\mathrm{rd}]}\right]_{i,n} \triangleq g_{rrc}(iT_s - nT - \tau_k^{[\mathrm{rd}]}T)$ is an $LN \times L$ matrix, $\tau_k^{[\mathrm{rd}]}$ denotes the normalized fractional *unknown* timing offset between the kth relay and destination, and
- $\mathbf{w} \triangleq [w(0), \ldots, w(LN-1)]^T$ and $w(i)\ \forall\ i$, denotes the zero-mean complex AWGN at the ith sample of the received signal, i.e., $w(i) \sim \mathcal{CN}(0, \sigma_w^2)$.

Substituting (3.1) into (3.4), $\mathbf{y}^{[\mathrm{TP}]}$ can be written as

$$\mathbf{y}^{[\mathrm{TP}]} = \sum_{k=1}^{K} \alpha_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \mathbf{p}_k + \sum_{k=1}^{K} \beta_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \mathbf{v}_k + \mathbf{w}, \tag{3.5}$$

where $\mathbf{p}_k \triangleq \mathbf{x}_k^{[\mathrm{TP}]} \odot \mathbf{t}_k^{[\mathrm{rd}]}$, $\mathbf{v}_k \triangleq \mathbf{n}_k \odot \mathbf{t}_k^{[\mathrm{rd}]}$, $\alpha_k \triangleq \zeta_k f_k h_k$ and $\beta_k \triangleq \zeta_k f_k$. The general model in (3.5) can be written compactly in matrix form as

$$\mathbf{y}^{[\mathrm{TP}]} = \mathbf{\Xi A p} + \mathbf{\Xi B v} + \mathbf{w}. \tag{3.6}$$

where

- $\mathbf{\Xi} \triangleq [\mathbf{\Xi}_1, \ldots, \mathbf{\Xi}_K]$ is an $LN \times LK$ matrix, $\mathbf{\Xi}_k \triangleq \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]}$ is the $LN \times L$ matrix of the kth relay's frequency offset, $\nu_k^{[\mathrm{rd}]}$, and timing offset, $\tau_k^{[\mathrm{rd}]}$,
- $\mathbf{A} \triangleq \mathrm{diag}(\alpha_1, \ldots, \alpha_K) \otimes \mathbf{I}_L$ is the $LK \times LK$ matrix of the overall channel gains from source-to-relays-to-destination,
- $\mathbf{B} \triangleq \mathrm{diag}(\beta_1, \ldots, \beta_K) \otimes \mathbf{I}_L$ is the $LK \times LK$ matrix of relay-to-destination channels,
- $\mathbf{p} \triangleq \left[\mathbf{p}_1^T, \ldots, \mathbf{p}_K^T\right]^T$, and $\mathbf{v} \triangleq \left[\mathbf{v}_1^T, \ldots, \mathbf{v}_K^T\right]^T$.

Using the training received signal, $\mathbf{y}^{[\mathrm{TP}]}$, and known TSs, $\mathbf{t}^{[\mathrm{sr}]}$ and $\mathbf{t}_k^{[\mathrm{rd}]}$, $\forall k$, the unknown parameters α_k, $\tau_k^{[\mathrm{rd}]}$ and $\nu_k^{[\mathrm{rd}]}$, $\forall\ k$, can be jointly estimated at the destination (see Section 3.3).

3.1.2.2 Data Transmission Model at the Destination

During the DTP, the sampled received signal at the destination receiver, $\mathbf{y}^{[\mathrm{DTP}]} \triangleq \left[y^{[\mathrm{DTP}]}(1), \ldots, y^{[\mathrm{DTP}]}(LN-1)\right]^T$, is given by

$$\mathbf{y}^{[\mathrm{DTP}]} = \sum_{k=1}^{K} \beta_k \mathbf{\Lambda}_k^{[\mathrm{rd}]} \mathbf{G}_k^{[\mathrm{rd}]} \mathbf{\Omega}_k \mathbf{b}_k^{[\mathrm{DTP}]} + \mathbf{w}, \tag{3.7}$$

where $\mathbf{\Omega}_k$ is the predefined $L \times L$ STBC matrix at the kth relay. Substituting (3.3) into (3.7), the signal model in (3.7) can be written in matrix form as

$$\mathbf{y}^{[\mathrm{DTP}]} = \mathbf{\Xi A \Omega x}^{[\mathrm{DTP}]} + \mathbf{\Xi B \Omega n} + \mathbf{w}, \tag{3.8}$$

where $\boldsymbol{\Omega} \triangleq \text{diag}(\boldsymbol{\Omega}_1, \ldots, \boldsymbol{\Omega}_K)$, $\mathbf{x}^{[\text{DTP}]} \triangleq \left[\left(\mathbf{x}_1^{[\text{DTP}]}\right)^T, \ldots, \left(\mathbf{x}_K^{[\text{DTP}]}\right)^T\right]^T$, and $\mathbf{n} \triangleq \left[\mathbf{n}_1^T, \ldots, \mathbf{n}_K^T\right]^T$. The estimates obtained during the TP are used to design matrix, $\mathbf{Q}$, which is used to compensate the effect of multiple channel gains, MTOs, and MCFOs and enable the detection of the signals from relays at the destination (see Section 3.4).

Note that unlike [1], the proposed processing structure at the relays in Fig. 3.2 is not based on the assumption of perfect timing and frequency offset estimation and perfect matched filtering at the relays [105]. More importantly, the authors in [1] did not consider the presence of MCFOs in their system.

3.2 Cramer-Rao Lower Bounds

In this section, *Fisher information matrix (FIM)* and CRLBs for the joint estimation of multiple channel gains, MTOs, and MCFOs for AF-DSTBC cooperative systems are derived in closed-form.

Based on the assumptions in Section 3.1, the AWGN at the relays, $\mathbf{u}_k$, $\forall k$, and destination, $\mathbf{w}$ in (3.4), are mutually independent. Accordingly, the received training signal at the destination, $\mathbf{y}^{[\text{TP}]}$ in (3.6), is a circularly symmetric complex Gaussian random variable, $\mathbf{y}^{[\text{TP}]} \sim \mathcal{CN}\left(\boldsymbol{\mu}_{\mathbf{y}^{[\text{TP}]}}, \boldsymbol{\Sigma}_{\mathbf{y}^{[\text{TP}]}}\right)$, with mean $\boldsymbol{\mu}_{\mathbf{y}^{[\text{TP}]}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}^{[\text{TP}]}}$ that are given by (see Appendix B.1)

$$\boldsymbol{\mu}_{\mathbf{y}^{[\text{TP}]}} = \sum_{k=1}^{K} \alpha_k \boldsymbol{\Xi}_k \mathbf{t}_k = \boldsymbol{\Xi}\mathbf{A}\mathbf{t}, \qquad \text{and} \tag{3.9a}$$

$$\boldsymbol{\Sigma}_{\mathbf{y}^{[\text{TP}]}} = \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{B}^H\boldsymbol{\Xi}^H + \boldsymbol{\Sigma}_{\mathbf{w}}, \tag{3.9b}$$

respectively. In (3.9), $\mathbf{t}_k \triangleq \left(\mathbf{t}^{[\text{sr}]} \odot \mathbf{t}_k^{[\text{rd}]}\right)$, $\mathbf{t} \triangleq \left[\mathbf{t}_1^T, \ldots, \mathbf{t}_K^T\right]^T$, $\boldsymbol{\Sigma}_{\mathbf{w}} \triangleq \sigma_w^2 \mathbf{I}_{LN}$, $\boldsymbol{\Sigma}_{\mathbf{p}} \triangleq \mathbb{E}\left\{(\mathbf{p}-\mathbf{t})(\mathbf{p}-\mathbf{t})^H\right\}$, $\boldsymbol{\Sigma}_{\mathbf{v}} \triangleq \mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\} \odot \mathbf{t}^{[\text{rd}]}\left(\mathbf{t}^{[\text{rd}]}\right)^H$, $\mathbf{t}^{[\text{rd}]} \triangleq \left[\left(\mathbf{t}_1^{[\text{rd}]}\right)^T, \ldots, \left(\mathbf{t}_K^{[\text{rd}]}\right)^T\right]^T$, and $\mathbf{A}$, $\mathbf{B}$, and $\boldsymbol{\Xi}$ are defined below (3.6).

The first step in determining the FIM is to formulate the parameter vector of interest. The destination node must estimate the MTOs and MCFOs from the relays to the destination, $\boldsymbol{\tau}^{[\text{rd}]} \triangleq \left[\tau_1^{[\text{rd}]}, \ldots, \tau_K^{[\text{rd}]}\right]^T$ and $\boldsymbol{\nu}^{[\text{rd}]} \triangleq \left[\nu_1^{[\text{rd}]}, \ldots, \nu_K^{[\text{rd}]}\right]^T$, respectively, and the overall channel gains from source to relays to destination, $\boldsymbol{\alpha} \triangleq [\alpha_1, \ldots, \alpha_K]^T$. As a result, the parameter vector of interest, $\boldsymbol{\lambda}$, is given by

$$\boldsymbol{\lambda} \triangleq \left[\Re\{\boldsymbol{\alpha}\}^T, \Im\{\boldsymbol{\alpha}\}^T, \left(\boldsymbol{\nu}^{[\text{rd}]}\right)^T, \left(\boldsymbol{\tau}^{[\text{rd}]}\right)^T\right]^T. \tag{3.10}$$

Note that $\mathbf{B}$ in (3.6) does not need to be estimated since it only scales the AWGN forwarded from the relays.

Theorem 3.1 *Based on the proposed AF-DSTBC relaying scheme, the FIM for the estimation of* $\boldsymbol{\lambda}$ *is given by*

$$\mathbf{F} = \begin{bmatrix} 2\Re\{\boldsymbol{\Psi}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & -2\Im\{\boldsymbol{\Psi}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & -2\Im\{\boldsymbol{\Psi}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \mathbf{D}\boldsymbol{\Psi}\boldsymbol{\mathcal{A}}\} & 2\Re\{\boldsymbol{\Psi}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Gamma}\boldsymbol{\mathcal{A}}\} \\ 2\Im\{\boldsymbol{\Psi}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & 2\Re\{\boldsymbol{\Psi}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & 2\Re\{\boldsymbol{\Psi}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \mathbf{D}\boldsymbol{\Psi}\boldsymbol{\mathcal{A}}\} & 2\Im\{\boldsymbol{\Psi}^H \Sigma^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Gamma}\boldsymbol{\mathcal{A}}\} \\ 2\Im\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Psi}^H \mathbf{D}\boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & 2\Re\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Psi}^H \mathbf{D}\boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & 2\Re\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Psi}^H \mathbf{D}\Sigma^{-1}_{\mathbf{y}^{[TP]}} \mathbf{D}\boldsymbol{\Psi}\boldsymbol{\mathcal{A}}\} & 2\Im\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Psi}^H \mathbf{D}\boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Gamma}\boldsymbol{\mathcal{A}}\} \\ 2\Re\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Gamma}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & -2\Im\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Gamma}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Psi}\} & -2\Im\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Gamma}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \mathbf{D}\boldsymbol{\Psi}\boldsymbol{\mathcal{A}}\} & 2\Re\{\boldsymbol{\mathcal{A}}^H \boldsymbol{\Gamma}^H \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[TP]}} \boldsymbol{\Gamma}\boldsymbol{\mathcal{A}}\} \end{bmatrix} + \boldsymbol{\Pi}, \tag{3.11}$$

where

- $\boldsymbol{\Psi} \triangleq [\boldsymbol{\Xi}_1 \mathbf{t}_1, \ldots, \boldsymbol{\Xi}_K \mathbf{t}_K]$ is an $LN \times K$ matrix,
- $\boldsymbol{\Gamma} \triangleq \left[\boldsymbol{\Lambda}_1 \left(\mathbf{G}_K^{[\mathrm{rd}]}\right)' \mathbf{t}_1, \ldots, \boldsymbol{\Lambda}_K \left(\mathbf{G}_K^{[\mathrm{rd}]}\right)' \mathbf{t}_K\right]$ is an $LN \times K$ matrix,
- $\left(\mathbf{G}_k^{[\mathrm{rd}]}\right)' \triangleq \frac{\partial \mathbf{G}_k^{[\mathrm{rd}]}}{\partial \tau_k^{[\mathrm{rd}]}}$, $\forall k$, is an $LN \times L$ matrix,
- $\mathbf{D} \triangleq 2\pi \times \mathrm{diag}\{0, 1, 2, \ldots, LN-1\}$ is an $LN \times LN$ matrix,
- $\boldsymbol{\mathcal{A}} \triangleq \mathrm{diag}\{\alpha_1, \ldots, \alpha_K\}$ is a $K \times K$ matrix,
- $\boldsymbol{\Pi}$ is a $4K \times 4K$ matrix, such that $[\boldsymbol{\Pi}]_{m,\bar{m}} = \mathrm{Tr}\left\{\boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[\mathrm{TP}]}} \boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m} \boldsymbol{\Sigma}^{-1}_{\mathbf{y}^{[\mathrm{TP}]}} \boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},\bar{m}}\right\}$ for $m, \bar{m} = 1, \ldots, 4K$,
- $\boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m}$ is an $LN \times LN$ matrix such that

$$\boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m} \triangleq \begin{cases} \boldsymbol{\Xi}\mathbf{U}_{\alpha_m} \boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{A}^H \boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{U}^H_{\alpha_m} \boldsymbol{\Xi}^H, & m = 1, \ldots, K \\ j\boldsymbol{\Xi}\mathbf{U}_{\alpha_{m-K}} \boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{A}^H \boldsymbol{\Xi}^H - j\boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{U}^H_{\alpha_{m-K}} \boldsymbol{\Xi}^H, & m = K+1, \ldots, 2K \\ \mathbf{U}_{\nu^{[\mathrm{rd}]}_{m-2K}} \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{A}^H \boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{A}^H \mathbf{U}^H_{\nu^{[\mathrm{rd}]}_{m-2K}} + \mathbf{U}_{\nu^{[\mathrm{rd}]}_{m-2K}} \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}} \mathbf{B}^H \boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}} \mathbf{B}^H \mathbf{U}^H_{\nu^{[\mathrm{rd}]}_{m-2K}}, & m = 2K+1, \ldots, 3K \\ \mathbf{U}_{\tau^{[\mathrm{rd}]}_{m-3K}} \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{A}^H \boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{A}^H \mathbf{U}^H_{\tau^{[\mathrm{rd}]}_{m-3K}} + \mathbf{U}_{\tau^{[\mathrm{rd}]}_{m-3K}} \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}} \mathbf{B}^H \boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}} \mathbf{B}^H \mathbf{U}^H_{\tau^{[\mathrm{rd}]}_{m-3K}}, & m = 3K+1, \ldots, 4K \end{cases}$$

- $\mathbf{U}_{\alpha_k} \triangleq \mathrm{diag}\left(\mathbf{0}_{L \times L(k-1)}, \mathbf{I}_L, \mathbf{0}_{L \times L(K-k)}\right)$ is an $L \times LK$ matrix, and $\mathbf{U}_{\nu_k^{[\mathrm{rd}]}} \triangleq \left[\mathbf{0}_{LN \times L(k-1)}, j\mathbf{D}\boldsymbol{\Xi}_k, \mathbf{0}_{LN \times L(K-k)}\right]$ and $\mathbf{U}_{\tau_k^{[\mathrm{rd}]}} \triangleq \left[\mathbf{0}_{LN \times L(k-1)}, \boldsymbol{\Lambda}_k^{[\mathrm{rd}]} \left(\mathbf{G}_k^{[\mathrm{rd}]}\right)', \mathbf{0}_{LN \times L(K-k)}\right]$ are $LN \times LK$ matrices.

Proof: See Appendix B.1.

Combining the real and imaginary parts of the channel, the new set of parameters, $\bar{\boldsymbol{\lambda}} \triangleq \left[\boldsymbol{\alpha}^T, \left(\boldsymbol{\nu}^{[\mathrm{rd}]}\right)^T, \left(\boldsymbol{\tau}^{[\mathrm{rd}]}\right)^T\right]^T$, is given by

$$\bar{\boldsymbol{\lambda}} = \underbrace{\begin{bmatrix} \mathbf{I}_K & j\mathbf{I}_K & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{I}_K & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \mathbf{I}_K \end{bmatrix}}_{\triangleq \mathbf{M}} \boldsymbol{\lambda}. \tag{3.12}$$

The CRLB matrix, $\mathbf{C}$, for the complex-valued estimation vector, $\bar{\boldsymbol{\lambda}}$ is evaluated as $\mathbf{C} = \mathbf{M}\mathbf{F}^{-1}\mathbf{M}^H$, and can be written in block matrix form as

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\nu}^{[\mathrm{rd}]}} & \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\tau}^{[\mathrm{rd}]}} \\ \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} & \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}} \\ \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\alpha}} & \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} & \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}} \end{bmatrix}, \tag{3.13}$$

where $\mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$, $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$ are $K \times K$ matrices and the CRLBs for the estimation of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\tau}^{[\mathrm{rd}]}$, and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, are given by the diagonal elements of $\mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$, $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, respectively.

Remark 3.1 *Eq.* (3.11) *shows that the FIM for the estimation of* $\boldsymbol{\lambda}$ *is not block diagonal. Thus, there exists coupling between the estimation errors of channel gains, MTOs, and MCFOs, i.e., the estimation performance of one parameter is affected by the presence of the remaining parameters. This shows the importance of jointly estimating channel gains, MTOs, and MCFOs in AF-DSTBC relaying cooperative systems. More importantly, this result indicates that the previously proposed methods that assume perfect timing or frequency synchronization while estimating MCFOs and MTOs in [2] and [1], respectively, cannot be applied to jointly estimate and compensate these impairments in AF-DSTBC relaying cooperative networks. This finding is also corroborated by the simulation results in Section 3.5.2, where it is shown that application of the algorithms in [1, 2] cannot achieve synchronization in the presence of both MTOs and MCFOs in AF-DSTBC cooperative networks.*

3.3 Joint Parameter Estimation at the Destination

In this section, the LS estimator for the joint estimation of multiple channel gains, MTOs and MCFOs is derived. Subsequently, it is shown that using DE, the com-

putational complexity for obtaining these impairments in DSTBC-AF relaying networks can be significantly reduced.

3.3.1 LS Estimation

Based on the signal model in (3.6), the LS estimates of the parameters, $\boldsymbol{\alpha}, \boldsymbol{\tau}^{[\mathrm{rd}]}$, and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, can be determined by minimizing the cost function, $\mathcal{J}_{\boldsymbol{\alpha},\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, according to

$$\mathcal{J}_{\boldsymbol{\alpha},\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} = \left\| \mathbf{y}^{[\mathrm{TP}]} - \left[\Lambda_1^{[\mathrm{rd}]} \mathbf{G}_1^{[\mathrm{rd}]}, \ldots, \Lambda_K^{[\mathrm{rd}]} \mathbf{G}_K^{[\mathrm{rd}]} \right] \mathbf{At} \right\|^2 = \left\| \mathbf{y}^{[\mathrm{TP}]} - \boldsymbol{\Psi}\boldsymbol{\alpha} \right\|^2. \quad (3.14)$$

Given $\boldsymbol{\tau}^{[\mathrm{rd}]}$ and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, it is straightforward to show that the LS estimate of $\boldsymbol{\alpha}$, denoted by $\hat{\boldsymbol{\alpha}}$, can be determined as

$$\hat{\boldsymbol{\alpha}} = \left(\boldsymbol{\Psi}^H \boldsymbol{\Psi} \right)^{-1} \boldsymbol{\Psi}^H \mathbf{y}^{[\mathrm{TP}]}. \quad (3.15)$$

Substituting (3.15) in (3.14), the estimates of MTOs and MCFOs, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, respectively, are obtained via

$$\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}, \hat{\boldsymbol{\nu}}^{[\mathrm{rd}]} = \arg \min_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}} \underbrace{ - \left(\mathbf{y}^{[\mathrm{TP}]} \right)^H \boldsymbol{\Psi} \left(\boldsymbol{\Psi}^H \boldsymbol{\Psi} \right)^{-1} \boldsymbol{\Psi}^H \mathbf{y}^{[\mathrm{TP}]} }_{\triangleq \chi\left(\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}\right)}, \quad (3.16)$$

where arg min denotes the arguments, $\boldsymbol{\tau}^{[\mathrm{rd}]}$ and $\boldsymbol{\nu}^{[\mathrm{rd}]}$, that minimize the expression $\chi\left(\boldsymbol{\tau}^{[\mathrm{rd}]}, \boldsymbol{\nu}^{[\mathrm{rd}]}\right)$ and $\mathbf{y}^{[\mathrm{TP}]}$, defined in (3.6), is a function of the timing and frequency offset estimation errors from the source-to-relay link. The channel estimates, $\hat{\boldsymbol{\alpha}}$, are obtained by substituting $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$ back into (3.15).

The minimization in (3.16) requires a multidimensional exhaustive search over the discretized set of the possible timing and frequency offset values, which is inherently very computationally complex. Furthermore, to reach the CRLB (see Fig. 3.6 in Section 3.5), the exhaustive search in (3.16) needs to be carried out with very high resolution[1], which significantly increases the sets of possible values for both the frequency and timing offsets, $\boldsymbol{\nu}^{[\mathrm{rd}]}$ and $\boldsymbol{\tau}^{[\mathrm{rd}]}$, respectively, and in turn, further increases the complexity of the proposed LS estimator.

Note that in order to reduce the complexity associated with the exhaustive search in (3.16), alternating projection (AP) has been used to transform the multidimensional minimization problem into a series of one-dimensional optimizations that are carried out sequentially [108]. However, in our extensive simulations we have observed that AP does not converge to the true estimates. This may be at-

[1]Step sizes of 10^{-2} and 10^{-4} for MTOs and MCFOs, respectively.

tributed to the source-to-relay timing and frequency offset estimation errors, which are not known at the destination and significantly affect the initialization of AP. Therefore, in this chapter, *differential evolution (DE)* is employed as a computationally efficient algorithm to carry out the minimization in (3.16) [119]. The initialization and parameterization of the proposed DE estimator are outlined below.

3.3.2 Differential Evolution based Estimation

DE and genetic algorithms are considered as a subclass of *evolutionary algorithms* since they attempt to evolve the solution for a problem through recombination, mutation, and survival of the fittest. More specifically, DE is an optimization algorithm, where a number of parameter vectors are generated and updated at each iteration in order to reach the solution [119]. Let d, for $d = 1, \ldots, D_I$, denote the number of iterations[2]. The detailed steps of the proposed DE algorithm for carrying out the minimization in (3.16) are as follows:

Step 1 *Initialization*: A *population* comprising of D_P *target vectors* is generated. The first iteration's ℓth target vector, $\boldsymbol{\theta}_{\ell,1} \triangleq \left[\tau_{1,\ell,1}^{[\mathrm{rd}]}, \ldots, \tau_{K,\ell,1}^{[\mathrm{rd}]}, \nu_{1,\ell,1}^{[\mathrm{rd}]}, \ldots, \nu_{K,\ell,1}^{[\mathrm{rd}]}\right]^T$, for $\ell = 1, \ldots, D_P$, is generated by selecting $\nu_{k,\ell,1}^{[\mathrm{rd}]}$ and $\tau_{k,\ell,1}^{[\mathrm{rd}]}$, $\forall k, \ell$, uniformly and randomly from the set of possible values for the frequency and timing offsets, respectively, i.e., $\nu_{k,\ell,1}^{[\mathrm{rd}]}$, $\tau_{k,\ell,1}^{[\mathrm{rd}]} \sim \mathcal{U}(-0.5, 0.5)$, $\forall k, \ell$. Subsequently, at each iteration, the D_P target vectors in the population are updated according to Steps 2-4 below.

Step 2 *Mutation*: At the dth iteration for each individual target vector, $\boldsymbol{\theta}_{\ell,d}$, three distinct and different target vectors denoted by $\boldsymbol{\theta}_{r1,d}$, $\boldsymbol{\theta}_{r2,d}$, and $\boldsymbol{\theta}_{r3,d}$, are randomly selected from the population. Subsequently, a *mutant* vector corresponding to the ℓth target vector and dth iteration, $\boldsymbol{\varpi}_{\ell,d} = [\varpi_{1,\ell,d}, \ldots, \varpi_{2K,\ell,d}]^T$, is created via

$$\boldsymbol{\varpi}_{\ell,d} = \boldsymbol{\theta}_{r1,d} + (\boldsymbol{\theta}_{r2,d} - \boldsymbol{\theta}_{r3,d}) \cdot D_F, \tag{3.17}$$

where D_F is a real positive scaling factor that controls the rate at which the population evolves.

Step 3 *Crossover*: In this step, the DE employs a uniform crossover between each target vector, $\boldsymbol{\theta}_{\ell,d}$, and the mutant vector, $\boldsymbol{\varpi}_{\ell,d}$, to create a *trial* vector,

[2]Also called "*generation*" in DE terminology.

$\boldsymbol{\vartheta}_{\ell,d} \triangleq [\vartheta_{1,\ell,d}, \ldots, \vartheta_{2K,\ell,d}]^T$, such that

$$\vartheta_{q,\ell,d} = \begin{cases} \varpi_{q,\ell,d}, & \iota \leq D_R \\ \theta_{q,\ell,d}, & \text{otherwise} \end{cases}, \qquad q = 1, \ldots, 2K \tag{3.18}$$

where $\iota \sim \mathcal{U}(0,1)$ and the crossover probability, $D_R \in [0,1]$ controls the fraction of parameter values that are copied from the mutant vector, $\boldsymbol{\varpi}_{\ell,d}$, to the trial vector, $\boldsymbol{\vartheta}_{\ell,d}$, $\forall \ell$.

Step 4 *Selection*: If compared to the ℓth target vector, $\boldsymbol{\theta}_{\ell,d}$, the ℓth trial vector, $\boldsymbol{\vartheta}_{\ell,d}$, results in a smaller objective function, χ, in (3.16), it replaces the target vector in the next iteration, i.e.,

$$\boldsymbol{\theta}_{\ell,d+1} = \begin{cases} \boldsymbol{\vartheta}_{\ell,d}, & \chi\left(\boldsymbol{\vartheta}_{\ell,d}\right) \leq \chi\left(\boldsymbol{\theta}_{\ell,d}\right) \\ \boldsymbol{\theta}_{\ell,d}, & \text{otherwise.} \end{cases} \tag{3.19}$$

Step 5 *Stopping Criteria*: There are various stopping criteria for the DE. However, our extensive simulations show that to reach the CRLB and obtain accurate estimates for different channel realizations, the algorithm can be terminated if the target vector that results in the lowest objective function, χ, i.e., the estimates of the timing and frequency offsets, remains unchanged for a predefined number, ς, of consecutive iterations.

By executing Steps 1-5 above, the estimates of the timing and frequency offsets, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, respectively, can be obtained. Substituting these estimates in (3.15) also generates the desired channel estimates.

The performance and convergence of the proposed DE algorithm to the true estimates is highly dependent on the values of the DE parameters, e.g., the population size D_P, the scaling factor D_F, the crossover probability D_R, and the stopping criteria ς. As a matter of fact, each unique optimization problem requires the DE parameters to be initialized differently [119], where finding the appropriate values for these parameters can be a difficult and a non-trivial task [120]. However, by applying the results in Section 3.2 and by following the general guidelines in [119,120], we are able to determine an appropriate parameterization for the DE algorithm.

Note that D_P is typically selected to be 10 times the dimension of the target vector. Since, we are solving a $2K$-dimensional minimization problem in (3.16), we set $D_P = 20K$. D_F is generally selected between 0.5 and 1.0 and extensive simulations in Section 3.5.1 show that $D_F = 0.85$ results in the fastest convergence while maintaining estimation performance close to that of the CRLB. Moreover,

from the FIM in Section 3.2, we know that the estimation of parameters of vector, $\boldsymbol{\lambda}$, is mutually coupled. As a result, D_R is set to $D_R = 0.9$ [119]. Finally, in order to reach the CRLB, the stopping criteria is set to $\varsigma = 80$.

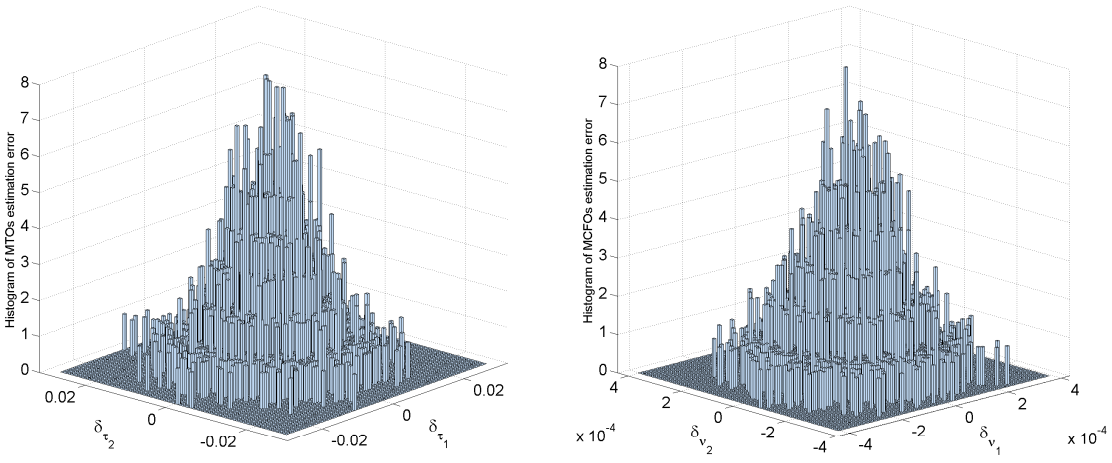
In DE terminology, the above procedure can be classified as "DE-rand", since to obtain the mutant vector, $\boldsymbol{\varpi}_{\ell,d}$, the target vector, $\boldsymbol{\theta}_{r1,d}$ in (3.17), is randomly selected from the population. In another variant of the DE, denoted by "DE-best", at each iteration, the target vector that minimizes the objective function, χ, is selected as $\boldsymbol{\theta}_{r1,d}$ in (3.17). However, in our problem formulation, "DE-best" does not always converge to the true values of the timing and frequency offsets. Similar findings are also observed in [119], where it is shown that "DE-rand" is more effective at reaching the global solution than "DE-best" for different optimization problems [119, p. 154].

Remark 3.2 *Even though the general conditions for global convergence of evolutionary algorithms are established in [121], it cannot be analytically shown that DE meets these conditions [122] and converges to the global solution of* (3.16)*. Moreover, in [121], no specific algorithm that meets these conditions is proposed. Although in [122], a variation of DE is proposed, it is indicated in [49] that this new approach also does not guarantee the global convergence of the DE algorithm. Nevertheless, by appropriately selecting the DE parameters, in our extensive simulations, we have observed the proposed DE estimator to always converge to the true values of the timing and frequency offsets as described in Sections IV-C and VI-A.*

3.3.3 Statistical Properties of the Estimates

In order to compare the performance of the proposed estimators against the CRLB, it needs to be shown that they are unbiased estimators. Moreover, as shown in Section 3.4, the statistical properties of estimates can be used to design the destination receiver such that it is more robust to the MTO, MCFO, and multiple channel estimation errors. Even though it is difficult to analytically determine the statistics of the estimates obtained using the proposed LS and DE estimators, in this subsection we numerically show that the proposed estimators are unbiased and their second order statistics can be accurately approximated by the CRLB derived in Section 3.2.

Figs. 3.4(a) and 3.4(b) show the histogram of the relay-to-destination timing and frequency offset estimation errors, i.e., $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} = \boldsymbol{\tau}^{[\mathrm{rd}]} - \hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$ and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} = \boldsymbol{\nu}^{[\mathrm{rd}]} - \hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, for 3600 realizations, $K = 2$ relays, SNR = 20 dB, and $\boldsymbol{\tau}^{[rd]}, \boldsymbol{\nu}^{[rd]} \sim \mathcal{U}(-0.5, -0.5)$. It can be observed that the means of the timing and frequency



(a) (b)

Figure 3.4: Histogram of the estimation errors of the proposed DE algorithm for estimation of (a) MTOs and (b) MCFO, K=2.

offset estimation errors, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$, respectively, are very close to zero. As a matter of fact, the average values of the timing and frequency offset estimates are observed to be $\mathbb{E}\{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}\} = [2.25 \times 10^{-5}, -8.85 \times 10^{-5}]^T$ and $\mathbb{E}\{\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}\} = [2.49 \times 10^{-5}, 1.09 \times 10^{-6}]^T$. Thus, it can be empirically concluded that the proposed DE and LS algorithms for estimation of MTOs and MCFOs are asymptotically unbiased estimators. Subsequently, it straightforwardly follows from (3.15) that the estimates of channel gains $\boldsymbol{\alpha}$ are also unbiased. More importantly, the numerical results in Figs. 3.4(a) and 3.4(b) demonstrate that the estimation errors for the timing and frequency offsets can be modeled as a multi-variate Gaussian distribution, i.e., $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}\right)$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}\right)$, where $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$, and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, are given in (3.13)[3]. More specifically, the average MSEs between the numerically evaluated covariance matrices of $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$, and the analytical covariance matrices, $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, are 1.264×10^{-10} and 7.2×10^{-18}, respectively.

3.3.4 Complexity Analysis

In this subsection, the computational complexity of the LS and the DE based estimation techniques are presented. Computational complexity is defined as the number of additions plus multiplications [3]. In addition, comparing two values or generating a random number is considered as a single arithmetic operation.

The computational complexity of the LS algorithm, denoted by C_{LS} is calculated

[3]Similar results are obtained for the channel gains, $\boldsymbol{\alpha}$, i.e., the distribution of $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}$ is given by $\boldsymbol{\delta}_{\boldsymbol{\alpha}} \sim \mathcal{CN}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}\right)$.

as

$$C_{\mathrm{LS}} = \underbrace{\left(\frac{1}{\kappa_\tau \kappa_\nu}\right)^{K} \left[2K^3 + (4LN-1)K^2 + (2(LN)^2 - LN)K + (LN)^2 + LN - 1\right]}_{\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}\,,\,\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}\ \text{in (3.16)}} + C_\alpha, \tag{3.20}$$

where κ_τ and κ_ν denote the step sizes or resolutions for timing and frequency offsets, respectively, in the exhaustive search in (3.16) and $C_\alpha = 2K^3 + (4LN - 1)K^2 + (LN - 1)K$ is the complexity for obtaining $\hat{\boldsymbol{\alpha}}$ via (3.15).

The computational complexity of the proposed DE algorithm, denoted by C_{DE} is calculated as

$$C_{\mathrm{DE}} = D_I \left\{ \underbrace{6D_P K}_{(3.17)} + \underbrace{4D_P K}_{(3.18)} + \underbrace{D_P(C_\chi + 1)}_{(3.19)} + \underbrace{D_P(C_\chi + 1)}_{\text{stopping criteria}} \right\} + C_\alpha, \tag{3.21}$$

where D_I is the total number of iterations required for the DE to converge and $C_\chi = 2K^3+(4LN-1)K^2+(2(LN)^2-LN)K+(LN)^2+LN-1$ is the computational complexity to evaluate the objective function χ in (3.16). Note that extensive simulations demonstrate that the proposed DE algorithm on average converges to the true estimates after approximately $D_I = 220$ and $D_I = 530$ iterations for $K = 2$ and $K = 4$, respectively, for a wide range of SNRs.

Remark 3.3 *By evaluating C_{LS} and C_{DE} in (3.20) and (3.21), it is observed that impairment estimation using the DE is 5.6×10^7 and 1.2×10^{19} times more computationally efficient compared to the proposed LS algorithm, for $K = 2$ and $K = 4$ relays, respectively. Note that this comparison is carried out by evaluating the LS estimates using exhaustive search in* (3.16) *and by setting $L = 80$, $\kappa_\tau = 10^{-2}$, $\kappa_\nu = 10^{-4}$, $D_P = 20K$, $D_R = 0.9$, $D_F = 0.85$, and $D_I = 220$ and 530 for $K = 2$ and $K = 4$, respectively. These values are selected to ensure that both estimators reach the CRLB over a wide range of SNR values. This comparison serves to illustrate the significantly lower complexity of the proposed DE based estimator.*

3.4 Compensation Algorithms at the Destination

The received signal at the destination, $\mathbf{y}^{[\mathrm{DTP}]}$ in (3.7), is the superposition of the relays' transmitted signals that are attenuated differently, are no longer aligned with each other in time, and are experiencing phase rotations at different rates due to the different channels, MTOs, and MCFOs, respectively. In this section, we derive

the MMSE and the low complexity MMSE (L-MMSE) receivers to compensate the effect of these impairments.

3.4.1 MMSE Compensation Design

The goal of the compensation algorithm is to design an $L \times LN$ matrix, $\mathbf{Q}$, as shown in Fig. 3.3, which is applied to the received signal, $\mathbf{y}^{[\mathrm{DTP}]}$, to remove the effect of MTOs, MCFOs, and multiple channels. The MMSE compensation matrix, $\mathbf{Q}$, is given by minimizing the cost function $\mathcal{J}_{\mathbf{s}}$ below

$$\mathcal{J}_{\mathbf{s}} = \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}},\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{u},\mathbf{w},\mathbf{s}} \left\{ \left\| \mathbf{Q}\mathbf{y}^{[\mathrm{DTP}]} - \mathbf{s} \right\|^2 \right\}, \tag{3.22}$$

where $\mathbf{s}$ is defined below (3.2), and the expectation in (3.22) is taken with respect to the statistics of the estimation errors $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} = \boldsymbol{\tau}^{[\mathrm{sr}]} - \hat{\boldsymbol{\tau}}^{[\mathrm{sr}]}$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} = \boldsymbol{\nu}^{[\mathrm{sr}]} - \hat{\boldsymbol{\nu}}^{[\mathrm{sr}]}$, $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$, and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$ and $\mathbf{s}$, $\mathbf{u}$, and $\mathbf{w}$. Furthermore, to meet the power constraints at the destination, it is assumed that $\mathrm{Tr}\left\{\mathbb{E}_{\mathbf{s}}\left\{\mathbf{s}\mathbf{s}^H\right\}\right\} \leq L$.

Assuming that the estimators used to obtain these synchronization and channel impairments are unbiased and asymptotically optimal, i.e., reach the CRLB for high SNR or very large number of observation symbols, the estimation errors, $\boldsymbol{\delta}_{\boldsymbol{\alpha}}$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$, and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$, can be modeled by $\boldsymbol{\delta}_{\boldsymbol{\alpha}} \sim \mathcal{CN}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}\right)$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{sr}]},\boldsymbol{\tau}^{[\mathrm{sr}]}}\right)$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{sr}]},\boldsymbol{\nu}^{[\mathrm{sr}]}}\right)$, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}\right)$, and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}\right)$, respectively, [18]. Note that $\mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}$, $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{rd}]},\boldsymbol{\tau}^{[\mathrm{rd}]}}$, and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}}$, are given in (3.13), and $\mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{sr}]},\boldsymbol{\tau}^{[\mathrm{sr}]}}$ and $\mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{sr}]},\boldsymbol{\nu}^{[\mathrm{sr}]}}$ can be determined from the results in [41]. In the following, knowledge of the statistics of the estimation errors for the channels and synchronization parameters are applied to design a compensation algorithm at the destination receiver that is robust to imperfect channel and synchronization parameter estimation.

Theorem 3.2 *The average MSE of the recovered data with respect to the source signal, $\mathcal{J}_{\mathbf{s}}$, is given by*

$$\begin{aligned}\mathcal{J}_{\mathbf{s}} = Tr\Big\{ &\mathbf{Q}\hat{\boldsymbol{\Xi}}\mathbf{R}\hat{\boldsymbol{\Xi}}^H\mathbf{Q}^H + \mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{w}}\mathbf{Q}^H + \boldsymbol{\Phi}_{\mathbf{s}} - \mathbf{Q}\hat{\boldsymbol{\Xi}}\hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[DTP]},\mathbf{s}} - \boldsymbol{\Phi}^H_{\mathbf{x}^{[DTP]},\mathbf{s}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H\hat{\boldsymbol{\Xi}}^H\mathbf{Q}^H \Big\} \\ &+ \mathcal{O}_p\left(\sigma^2_{\tau_k^{[sr]}}\right) + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[rd]}}\right), \end{aligned} \tag{3.23}$$

where

- $\hat{\boldsymbol{\Xi}} \triangleq \boldsymbol{\Xi}\big|_{\boldsymbol{\tau}^{[\mathrm{rd}]}=\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]},\boldsymbol{\nu}^{[\mathrm{rd}]}=\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}}$ is an $LN \times LK$ matrix, $\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}} = \Big[\underbrace{\boldsymbol{\Phi}_{\mathbf{s}},\ldots,\boldsymbol{\Phi}_{\mathbf{s}}}_{K}\Big]^T$ is an $LK \times L$ matrix, $\boldsymbol{\Phi}_{\mathbf{s}} = \mathbf{I}_L$,

- $\mathbf{R} = \hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H + (\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \otimes \mathbf{I}_L) \odot (\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}}\boldsymbol{\Omega}^H) + \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H$ is an $LK \times LK$ matrix,
- $\hat{\mathbf{A}} \triangleq \mathrm{diag}(\hat{\boldsymbol{\alpha}}) \otimes \mathbf{I}_L$ is an $LK \times LK$ matrix, $\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \triangleq \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}\big|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}}$ is a $K \times K$ matrix,
- $\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]}} = \mathbf{1}_{K\times K} \otimes \boldsymbol{\Phi}_{\mathbf{s}}$ is an $LK \times LK$ matrix,
- $\boldsymbol{\Sigma}_{\mathbf{n}}$ is the $LK \times LK$ covariance matrix of the matched filtered noise at the relays, $\mathbf{n}$, with its kth $L \times L$ block diagonal matrix, $[\boldsymbol{\Sigma}_{\mathbf{n}}]_{(k-1)L:kL,(k-1)L:kL}$, given by $[\boldsymbol{\Sigma}_{\mathbf{n}}]_{(k-1)L:kL,(k-1)L:kL} \triangleq \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\boldsymbol{\Sigma}_{\mathbf{u}_k}\left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H$, $\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \triangleq \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}\big|_{\hat{\tau}_k^{[\mathrm{sr}]}=\tau_k^{[\mathrm{sr}]}}$ is an $L \times LN$ matrix, and $\boldsymbol{\Sigma}_{\mathbf{u}_k} \triangleq \sigma_{u_k}^2\mathbf{I}_{LN}$, and
- $\mathcal{O}_p(\cdot)$ denotes the big omicron function for stochastic parameters [123].

Proof: See Appendix B.2.

Proposition 1: Using (3.23), the MMSE compensation matrix, $\mathbf{Q}^{[\mathrm{MMSE}]}$, is given by

$$\mathbf{Q}^{[\mathrm{MMSE}]} = \boldsymbol{\Phi}^H_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H\hat{\boldsymbol{\Xi}}^H\left(\hat{\boldsymbol{\Xi}}\mathbf{R}\hat{\boldsymbol{\Xi}}^H + \boldsymbol{\Sigma}_{\mathbf{w}}\right)^{-1}, \tag{3.24}$$

where $\mathbf{R}$, $\boldsymbol{\Phi}_{\mathbf{x}^{[\mathrm{DTP}]},\mathbf{s}}$, $\hat{\mathbf{A}}$, and $\hat{\boldsymbol{\Xi}}$ are defined below (3.23), and $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}_{\mathbf{w}}$ are defined below (3.8) and (3.9), respectively.

Proof: Eq. (3.24) follows from Theorem 3.1 by taking the derivative of $\mathcal{J}_{\mathbf{s}}$ in (3.23) with respect to $\mathbf{Q}$, setting the results to zero, and carrying out straightforward algebraic manipulations.

Remark 3.4 *Evaluation of the MMSE compensation matrix, $\mathbf{Q}^{[MMSE]}$, requires the knowledge of the overall channel gains, $\hat{\boldsymbol{\alpha}}$, relay to destination timing and frequency offsets, $\hat{\boldsymbol{\tau}}^{[rd]}$ and $\hat{\boldsymbol{\nu}}^{[rd]}$, respectively, relay to destination channel gains $\boldsymbol{\beta}$, and source to relay timing offsets, $\boldsymbol{\tau}^{[sr]} \triangleq [\hat{\tau}_1^{[sr]}, \ldots, \tau_K^{[sr]}]^T$, at the destination. Even though, $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\tau}}^{[rd]}$, and $\hat{\boldsymbol{\nu}}^{[rd]}$ can be jointly estimated via the estimators proposed in Section 3.3 and are known at the destination, $\boldsymbol{\beta}$ and $\boldsymbol{\tau}^{[sr]}$ are not known at the destination. Thus, the proposed compensation algorithm requires the relays to feed forward the estimates of the source to relay timing offsets and the channel gains, $\hat{\tau}_k^{[sr]}$ and $\hat{h}_k$, respectively, to the destination. Note that using $\hat{h}_k$, an estimate of the relay to destination channels, $\hat{\beta}_k$, can be obtained according to $\hat{\beta}_k = \frac{\hat{\alpha}_k}{\hat{h}_k}$.*

Remark 3.5 *According to the results in Appendix B.2, it can be concluded that due to the presence of channel, timing, and frequency offset estimation errors, $\mathbf{Q}^{[MMSE]}$ is not the optimal MMSE compensation matrix at the destination. However, one*

can derive this optimal MMSE matrix by assuming perfect knowledge of channels, timing, and frequency offsets from the source to relays and the relays to destination, i.e., $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[rd]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[rd]}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[sr]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[sr]}} = \mathbf{0}_{K\times 1}$.

Proposition 2*: Based on the assumption of perfect synchronization and channel estimation, the optimal MMSE compensation matrix,* $\mathbf{Q}^{[OPT]}$*, can be determined as*

$$\mathbf{Q}^{[OPT]} = \boldsymbol{\Phi}^{H}_{\mathbf{x}^{[DTP]},\mathbf{s}} \boldsymbol{\Omega}^{H} \mathbf{A}^{H} \boldsymbol{\Xi}^{H} \left(\boldsymbol{\Xi} \mathbf{R}^{[OPT]} \boldsymbol{\Xi}^{H} + \boldsymbol{\Sigma}_{\mathbf{w}} \right)^{-1}, \tag{3.25}$$

where $\mathbf{R}^{[OPT]} = \mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[DTP]}}\boldsymbol{\Omega}^{H}\mathbf{A}^{H} + \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^{H}\mathbf{B}^{H}$.

Proof*: By assuming* $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[rd]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[rd]}} = \boldsymbol{\delta}_{\boldsymbol{\tau}^{[sr]}} = \boldsymbol{\delta}_{\boldsymbol{\nu}^{[sr]}} = \mathbf{0}_{K\times 1}$ *in* (3.22) *and carrying out the steps in Appendix B.2, the result in* (3.25) *follows.*
Even though the proposed optimal compensation approach has limited practical applications, it can be used to benchmark the performance of DSTBC-AF cooperative networks. As a matter of fact, $\mathbf{Q}^{[OPT]}$ *is applied in Section 3.5.2 to determine a lower bound on the BER performance of DSTBC-AF cooperative networks in the presence of channel and synchronization impairments.*

3.4.2 L-MMSE Compensation Design

In order to reduce the synchronization overhead, in this subsection, a low complexity approach that does not require the relays to feed forward the synchronization and channel estimates to the destination is proposed.

The compensation matrix $\mathbf{Q}^{[\text{MMSE}]}$ in (3.24) is a function of the source to relay channels and timing offsets since the matrix $\mathbf{R}$ defined below (3.23) is a function of $\mathbf{B}$ and $\boldsymbol{\Sigma}_{\mathbf{n}}$, respectively. However, over a large number of channel realizations, the second order channel statistics can be used to approximate the source to relay channel gains on average. As a result, $\mathbf{B}$ in (3.24) can be replaced by $\mathbf{B}^{[\text{L-MMSE}]} \approx \text{diag}\left(\sigma_f\zeta_1, \ldots, \sigma_f\zeta_K\right) \otimes \mathbf{I}_L$. Moreover, the dependence on the timing offset estimates from the source to relay can be eliminated by approximating the block diagonal matrix $\boldsymbol{\Sigma}_{\mathbf{n}}$, defined below (3.23), with the diagonal matrix $\boldsymbol{\Sigma}_{\mathbf{n}}^{[\text{L-MMSE}]}$ with its kth diagonal block given by $\left[\boldsymbol{\Sigma}_{\mathbf{n}}^{[\text{L-MMSE}]}\right]_{(k-1)L:kL,(k-1)L:kL} = \sigma^2_{u_k}\mathbf{I}_L$ at high SNR. The latter follows from the definition of $\boldsymbol{\Sigma}_{\mathbf{n}}$ and extensive simulations that show that the off-diagonal elements of $\boldsymbol{\Sigma}_{\mathbf{n}}$ that are dependent on the source to relay timing offsets vanish much more quickly compared to its diagonal elements as SNR increases. Using the above approximations and (3.24), a low complexity compensation matrix, $\mathbf{Q}^{[\text{L-MMSE}]}$, for mitigating the effect of impairments can be found as

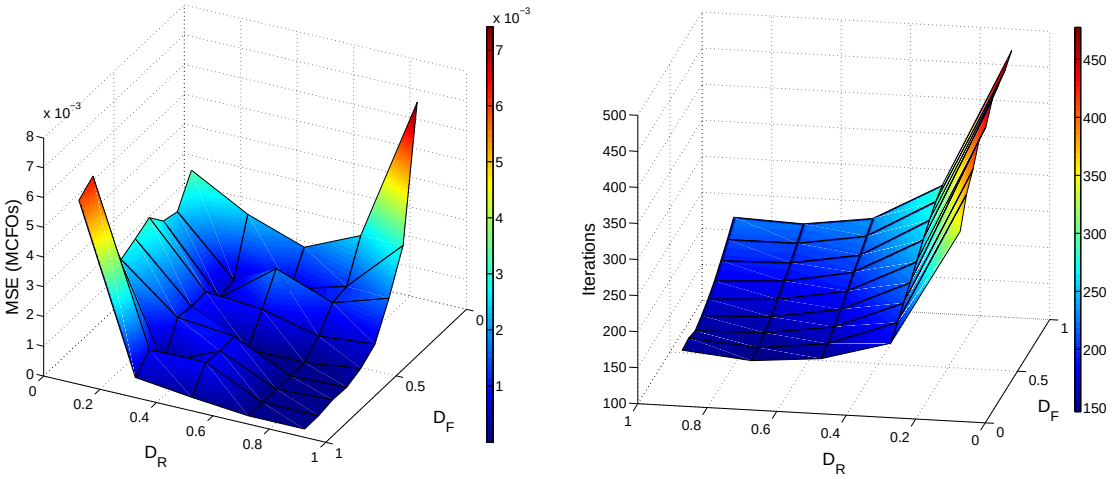
$$\mathbf{Q}^{[\text{L-MMSE}]} = \boldsymbol{\Phi}^{H}_{\mathbf{x}^{[\text{DTP}]},\mathbf{s}} \boldsymbol{\Omega}^{H} \hat{\mathbf{A}}^{H} \hat{\boldsymbol{\Xi}}^{H} \left(\hat{\boldsymbol{\Xi}} \mathbf{R}^{[\text{L-MMSE}]} \hat{\boldsymbol{\Xi}}^{H} + \boldsymbol{\Sigma}_{\mathbf{w}} \right)^{-1}, \tag{3.26}$$

where $\mathbf{R}^{[\text{L-MMSE}]} = \hat{\mathbf{A}}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\text{DTP}]}}\boldsymbol{\Omega}^H\hat{\mathbf{A}}^H + (\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \otimes \mathbf{I}_L) \odot (\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}^{[\text{DTP}]}}\boldsymbol{\Omega}^H) + \mathbf{B}^{[\text{L-MMSE}]}\boldsymbol{\Omega}$ $\boldsymbol{\Sigma}_{\mathbf{n}}^{[\text{L-MMSE}]}\boldsymbol{\Omega}^H \left(\mathbf{B}^{[\text{L-MMSE}]}\right)^H$.

Simulation results in Section 3.5.2 show that the BER performance of a multi-relay DSTBC-AF cooperative network using the proposed L-MMSE approach in (3.26) is very close to that of the MMSE receiver in (3.24) over a wide range of SNRs. This result indicates that the approximations used to arrive at (3.26) hold for practical scenarios of interest. Thus, the proposed L-MMSE approach can be effectively applied to compensate the effect of MTOs, MCFOs, and channels in DSTBC-AF cooperative networks, while significantly reducing synchronization overhead.

3.5 Simulation Results

In this section, we investigate the receiver performance at the destination, where multiple channel gains, MTOs, and MCFOs are jointly estimated and compensated in order to decode the received signal. In our simulation setup, we consider $K = 2$ and 4 relays in DSTBC-AF cooperative systems. Quadrature phase-shift keying modulation is employed for data transmission. Length of the training signals, $\mathbf{t}^{[\text{sr}]}$ and $\mathbf{t}_k^{[\text{rd}]}$, $\forall k$, are set to $L = 80$ symbols during the TP and length of the source data vector $\mathbf{s}$, is set to $L = 400$ symbols during the DTP, resulting in a synchronization overhead of 16%. Oversampling factor is set to $N = 2$ and a root-raised cosine filter with a roll-off factor of 0.22 is employed. At each relay, the DSTBC is generated randomly based on an isotropic distribution on the space of $L \times L$ unitary matrices, which is a benchmark method for generating DSTBC in AF cooperative networks [51, 52]. The propagation loss is modeled as $\eta = (\varepsilon/\varepsilon_0)^{-\rho}$, where ε is the distance between transmitter and receiver, ε_0 is the reference distance, and ρ is the path loss exponent [7]. We set $\varepsilon_0 = 1$km, and $\rho = 2.7$, which corresponds to an urban area cellular network [7]. The timing and frequency offsets at the relays and the destination, $\boldsymbol{\tau}^{[\text{sr}]}$, $\boldsymbol{\tau}^{[\text{rd}]}$, $\boldsymbol{\nu}^{[\text{sr}]}$, and $\boldsymbol{\nu}^{[\text{rd}]}$ are uniformly drawn from the full acquisition range, $(-0.5, 0.5)$. Without loss of generality, it is assumed that the AWGN at all relays has the same variance, i.e., $\sigma_u^2 = \sigma_{u_1}^2 = \ldots = \sigma_{u_K}^2$ and $\sigma_u^2 = \sigma_w^2 = 1/\text{SNR}$. The MSE performance of the proposed estimators and the end-to-end BER of DSTBC-AF cooperative networks using the proposed compensation algorithms are detailed below.



(a) (b)

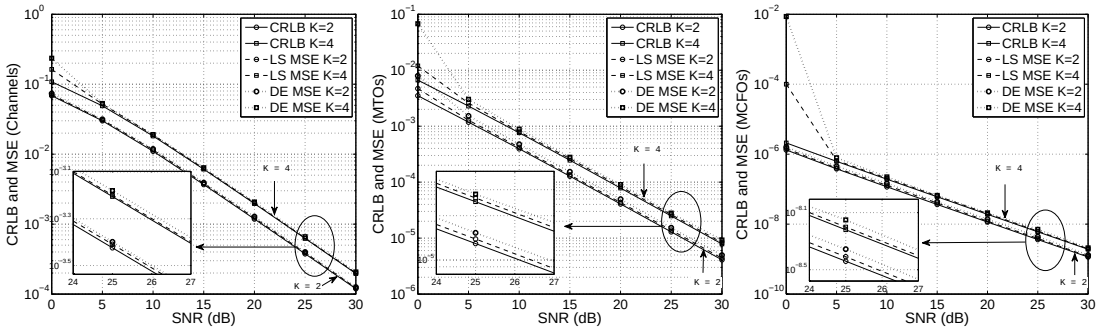
Figure 3.5: (a) CRLB and MSE performance of the MCFOs' estimation for DE based algorithm with different parameter settings, as a function of SNR (dB). (b) Number of generations (iterations) required for the convergence of DE algorithm with different parameter settings.

3.5.1 MSE performance

Following the approach in [2, 3, 12, 21] and for ease of reproducing the CRLBs and MSE curves in Fig. 3.5 and Fig. 3.6, specific channels are used. For $K = 2$, $\mathbf{h} = [.279 - .9603j, .8837 + .4681j]^T$ and $\mathbf{f} = [.7820 + .6233j, .9474 - .3203j]^T$ and for $K = 4$, $\mathbf{h} = [.279 - .9603j, .8837 + .4681j, -.343 + .732i, -.734 - .451i]^T$ and $\mathbf{f} = [.7820 + .6233j, .9474 - .3203j, -.2413 + .724i, .5141 - .893i]^T$. Note that in the next subsection, we employ random Rayleigh fading channels for analyzing the end-to-end BER performance of the overall system. As stated in Section 3.3.4 and to reach the CRLB, the step sizes, κ_τ and κ_ν, for the exhaustive search in (3.16) for the LS estimator are set to 10^{-2} and 10^{-4}, respectively. Without loss of generality, the CRLBs and the MSE estimation performance for the first relay node is presented only, where similar results are observed for the other relays.

Fig. 3.5(a) shows the MSE of frequency offset estimation[4] as a function of the DE parameters: the crossover probability, D_R, and the scaling factor, D_F. It is shown that the lowest MSE is achieved for $D_F = 0.85$ and $D_R = 0.9$. Fig. 3.5(b) shows the number of iterations required for the convergence of the DE algorithm for different D_R and D_F values. It can be seen that the DE algorithm converges more quickly as D_R increases from 0.1 to 0.9. As a matter of fact, Fig. 3.5(b) demonstrates that the proposed DE algorithm converges with the fewest number of iterations when $D_F = 0.25$ and $D_R = 0.9$. However, as shown in Fig. 3.5(a),

[4]Similar results can be obtained for timing offset estimation.



(a) (b) (c)

Figure 3.6: CRLBs in (3.13) and MSE for the estimation of (a) channel gains, (b) MTOs, and (c) MCFOs as a function of SNR (dB).

this fast convergence comes at the cost of MSE performance. The results in Figs. 3.5(a) and 3.5(b) demonstrate that the proposed DE algorithm provides an effective trade-off between synchronization performance and complexity that can be used to meet the performance requirements of cooperative communication systems. In this scenario, faster convergence can be achieved by selecting a smaller value for D_F, e.g., $D_F = 0.25$, while compromising estimation accuracy. In order to achieve the best estimation accuracy, in the remainder of this section, we set $D_F = 0.85$ and $D_R = 0.9$.

Figs. 3.6(a), 3.6(b), and 3.6(c) show the CRLBs and the MSEs for the estimation of channel gains, MTOs, and MCFOs for different SNRs and $K = 2$ and 4 relays. Fig. 3.6 shows that the CRLBs and the MSEs for a 2-relay DSTBC-AF cooperative network are lower compared to that of a network equipped with 4 relays. This is because as the number of relays increases, more parameters need to be jointly estimated at the destination and the effect of amplified and forwarded AWGN from the relays is also more prominent at the destination. Moreover, it is demonstrated that at low SNR (SNR < 5 dB), the proposed estimators exhibit poor performance due to the noise at the relays, which is amplified and forwarded to the destination. However at moderate-to-high SNRs (SNR ≥ 5 dB), Fig. 3.6 shows that the MSEs of the proposed LS and DE estimators are close to the CRLB. Note while showing comparable estimation accuracy, the proposed DE algorithm is significantly less complex than the LS estimator as shown in *Remark 3.3*.

3.5.2 BER performance

In this subsection the BER performance of a DSTBC-AF cooperative system that employs the proposed DE estimator and compensation algorithms is investigated.

The channels from the source to relays and from the relays to destination are modeled as independent and identically distributed complex Gaussian random variables with $\mathcal{CN}(0,1)$.

Fig. 3.7 shows the end-to-end BER performance of a 4-relay DSTBC-AF cooperative system, using the proposed MMSE, L-MMSE, and optimal compensation algorithms in (3.24), (3.26), and (3.25), respectively. The performances of the proposed compensation algorithms are also compared to that of the zero-forcing (ZF) receiver, which obtains an estimate of the transmitted signal via $\hat{\mathbf{s}} = (\boldsymbol{\Upsilon}^H \boldsymbol{\Upsilon})^{-1} \boldsymbol{\Upsilon}^H \mathbf{y}^{[\mathrm{DTP}]}$, where $\boldsymbol{\Upsilon} \triangleq \sum_{k=1}^{K} \hat{\alpha}_k \hat{\boldsymbol{\Xi}}_k \boldsymbol{\Omega}_k$. For the MMSE, L-MMSE, and ZF methods, the compensation matrices, $\mathbf{Q}^{[\mathrm{MMSE}]}$ in (3.24), $\mathbf{Q}^{[\text{L-MMSE}]}$ in (3.26), and $\boldsymbol{\Upsilon}$ defined above are evaluated using the estimated values, $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}$, and $\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}$, obtained in the training period using the DE algorithm[5] while the optimal compensation matrix $\mathbf{Q}^{[\mathrm{OPT}]}$ in (3.25) is evaluated based on perfect knowledge of channels and synchronization parameters. Fig. 3.7 shows that both the proposed MMSE and L-MMSE receivers outperform the ZF receiver with a performance gain of 2 dB or more at moderate-to-high SNRs. Furthermore, Fig. 3.7 shows that the BER of a DSTBC-AF cooperative system based on the estimated impairments is close that of the ideal scenario with perfect knowledge of impairments and optimal compensation (a gap of 4-5 dBs). This result demonstrates the robustness of the proposed transceiver designs to the estimation errors. Finally, Fig. 3.7 demonstrates that even though the proposed L-MMSE receiver does not require relays to feed forward the source to relay parameter estimates to the destination, its BER is very close to that of the MMSE receiver with a performance gap of only 0.1-0.3 dB.

Fig. 3.8 depicts the BER results for a cooperative system that first employs the re-synchronization filter in [1] to compensate MTOs and then attempts to remove MCFOs by employing the algorithm in [2]. This plot, which is labeled as " [1] & [2]", shows that such an approach fails to equalize the effect of impairments at the destination since the re-synchronization filter in [1] fails to compensate MTOs in the presence of MCFOs. Subsequently, the algorithm in [2] fails to nullify MCFOs, since the input signal is corrupted by MTOs. In addition, even though Fig. 3.8 illustrates the BER performance of a cooperative system applying the transceiver design in Chapter 2 for 2 and 4 relays, it is important to consider that the proposed DSTBC-AF transceiver design in this chapter and the scheme in Chapter 2 vastly differ. In Chapter 2, a less complex AF transceiver design is used that does not allow for the application of DSTBCs, and as a result, sacrifices spatial diversity to achieve lower complexity. On the other hand, as shown in Fig. 3.8, the DSTBC-AF transceiver design proposed in this paper achieves full diversity gain at high SNR,

[5]Estimation using LS estimator results in similar BER performance as obtained using the DE algorithm.

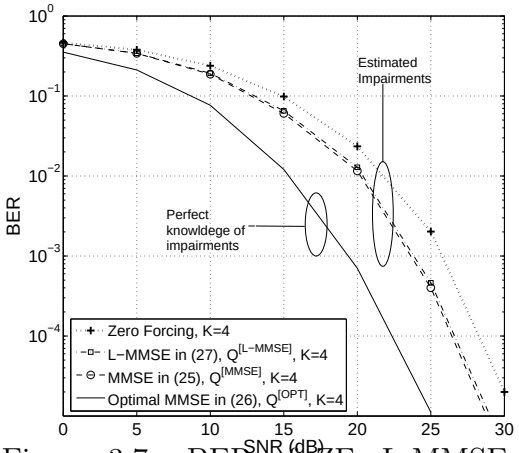


Figure 3.7: BER of ZF, L-MMSE, MMSE, and optimal receivers for 4-relay AF-DSTBC cooperative systems.

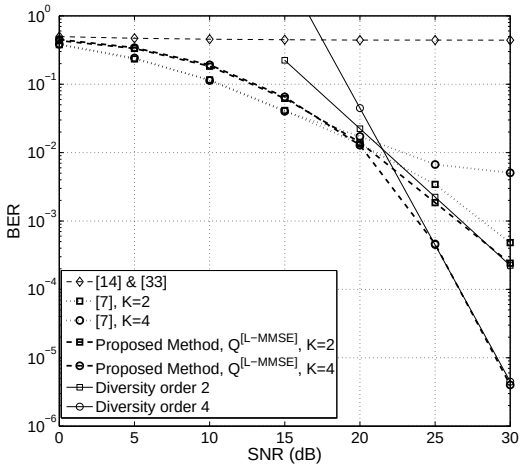


Figure 3.8: BER of the proposed L-MMSE receiver while using the DE-based estimator against the approaches in [1], [2], and [3].

while requiring additional synchronization overhead at the relays to allow for the application of DSTBCs.

Fig. 3.8 shows that at low SNR, the proposed algorithm exhibits a higher BER performance compared to that of [3]. This follows from the inherent relay transceiver structure for the DSTBC-AF scheme proposed here, where the relays need to estimate and compensate the effect of timing and frequency offsets before applying DSTBCs and forwarding their signals to the destination. As a result, the forwarded signals from the relays are affected by the timing and frequency offset estimation errors that negatively affect cooperative performance at the destination node. This effect is particularly more prominent at low SNR, where the timing and frequency offset estimation errors at the relays have larger variances. On the other hand, the algorithm in [3] does not apply a matched filter at each relay, cannot apply DSTBCs, and does not require the relays to achieve carrier synchronization before forwarding their signal to the destination. Hence, the approach in [3] is not significantly influenced by the timing and frequency offset estimation errors at the relays at low SNR. At high SNR, the proposed algorithm exhibits large performance gain compared to the results in [3], e.g., at SNR = 30 dB, the BER of the proposed algorithm is 4×10^{-6} for $K = 4$ relays while the BER of the algorithm in [3] is 5×10^{-2}. Moreover, due to the application of DSTBCs in the proposed cooperative system, the BER of a 4-relay system achieves a large diversity gain over that of a 2-relay network while the algorithm in [3] provides little to no spatial diversity gain as the number of relays increases.

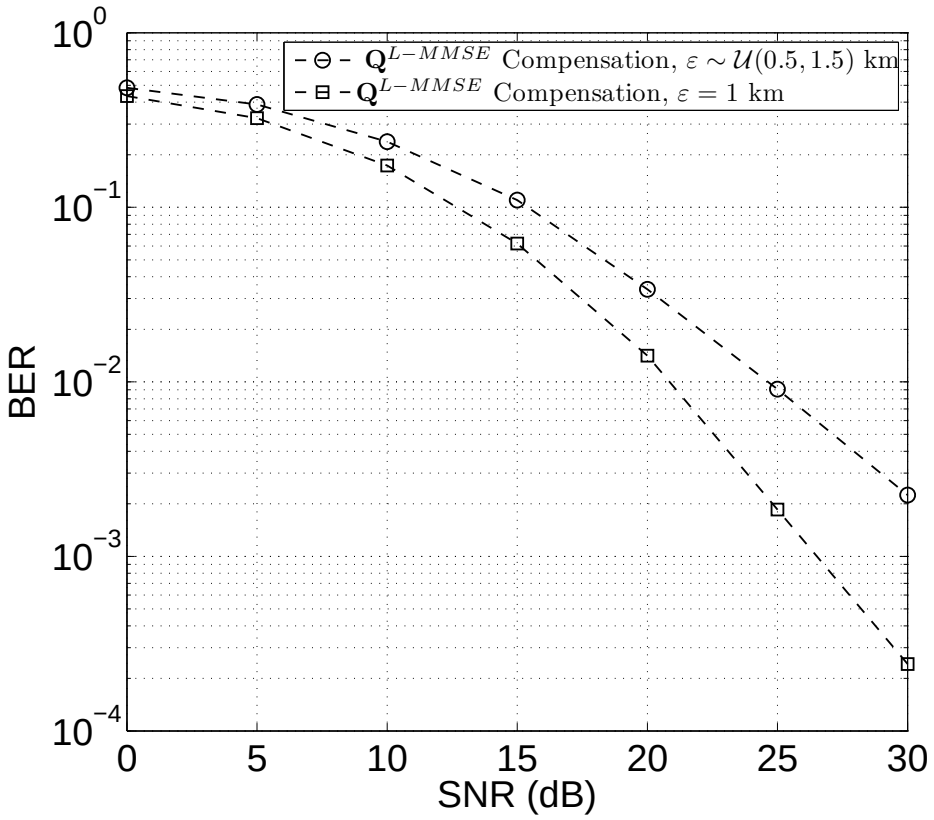


Figure 3.9: BER of the proposed L-MMSE receiver employing the DE-based estimator, when relays are uniformly distributed in the wireless network, i.e., $\varepsilon \sim \mathcal{U}(0.5, 1.5)$ km.

Fig. 3.9 plots the BER of a 2-relay cooperative network employing the proposed L-MMSE receiver, where the relays are uniformly distributed in the wireless network, i.e., $\varepsilon \sim \mathcal{U}(0.5, 1.5)$ km. Fig. 3.9 compares the BER obtained from the variable relay location setup with the BER achieved from the fixed relay location setup, $\varepsilon = 1$ km. It can be observed from the figure that BER of the variable relay location setup is close to the BER of the fixed relay setup with a performance margin of 5 dB, at BER $= 2 \times 10^{-3}$. The higher BER with the variable relay location setup is to be expected because of the large path loss encountered, when relays get closer to the source or the destination.

3.6 Conclusions

This chapter introduced a novel DSTBC based AF relaying cooperative network in the presence of MTOs, MCFOs, and unknown channels. New CRLBs for the multi-parameter estimation problem are derived and LS and DE based estimators

for joint estimation of multiple channel gains, MTOs, and MCFOs are proposed. The convergence, parametrization, and complexity of the proposed DE algorithm are analyzed and it is empirically shown that the proposed DE estimator is unbiased and many orders of magnitude more computationally efficient than the proposed LS estimator when considering 2 or 4 relay AF-DSTBC networks, respectively. Simulation results show that the MSEs of the proposed estimators are close to the CRLB at moderate-to-high SNRs (SNR $\geq$ 5dB). Finally, a novel MMSE receiver is proposed to decode the received signal at the destination by compensating for the effect of multiple channel gains, MTOs and MCFOs. To reduce the overhead associated with the synchronization in AF-DSTBC networks, an L-MMSE receiver is also proposed that does not require the relays to feed forward any estimates to the destination. Simulations show that the BER plots of the proposed MMSE and L-MMSE receivers are only 0.1-0.3 dB apart while showing a performance gap of 4-5 dB from the derived optimal MMSE receiver that has access to perfect knowledge of impairments. More importantly, these results demonstrate that the combination of the proposed estimators and the compensation algorithms ensure that a multi-relay AF-DSTBC cooperative network can achieve full spatial diversity in the presence of MTOs, MCFOs, and multiple unknown channels.

Chapter 4

Training Sequence Design

This chapter seeks to focus on the design of optimal training sequences (TSs) for the efficient estimation of MTOs and multiple channel parameters. In Chapter 2 and Chapter 3, our focus was on the design of estimation and compensation algorithms for the joint estimation of multiple impairments in DA cooperative communication systems. However, considering DA or training-based transmission, the design of optimal training sequences for efficient estimation of multiple impairments is also critical. In Chapter 3 and Chapter 4, we arbitrarily used unit amplitude phase shift keying (PSK) training signals. In this chapter, we will consider the design guidelines to develop optimal training sequences for efficient and joint estimation of MTOs and multiple channel parameters. Chapter 4 is organized as follows:

Section 4.1 outlines the system model and the general set of assumptions. In Section 4.2, a new *hybrid Cramér-Rao lower bound* (HCRB) for the joint estimation of MTOs and channels is derived. Next, in order to show that the proposed TS design guidelines also improve estimation accuracy, the *conditional Cramér-Rao lower bound* (ECRB), which is a tighter lower bound on the estimation accuracy of the parameters of interest at high signal-to-noise ratios, is also derived. In Section 4.3, by minimizing the derived HCRB as a function of training sequences, three TS design guidelines for designing optimal TSs are proposed. Next, according to these guidelines, two training sequences are proposed. In Section 4.4, the MAP estimator for the joint estimation of MTOs and channels is derived. In Section 4.5, simulation results are presented. Numerical results show that the proposed training sequence design guidelines not only lower the HCRB, but they also lower the ECRB and the mean-square error of the proposed maximum a posteriori estimator. Moreover, extensive simulations demonstrate that the application of the proposed training sequences significantly lowers the bit-error rate performance of the multi-relay cooperative networks when compared to training sequences that

violate the proposed design guidelines. Finally, Section 4.6 concludes the paper's findings.

4.1 System Model

We consider a communication system where multiple nodes communicate with a single destination node. Multi-relay cooperative or multi-user distributed networks are examples of the system model under consideration. All nodes are assumed to be equipped with a single antenna. In addition to being affected by multiple channel parameters, the received signal is affected by MTOs due to the random propagation delays at each node.

Signal transmission consists of a *training period* and a *data transmission period*. In the training period, the estimates of MTOs and channel parameters are obtained by employing different TSs of length L that are transmitted from the K distributed nodes. In the data transmission period, the estimates obtained during the training period are used at the receiver to detect the signals from multiple nodes. In order to improve the overall end-to-end system performance in the data transmission period, the goal of this paper is to design optimal TSs in the training period to efficiently estimate MTOs and channel parameters. Throughout this paper, the following set of assumptions and system design parameters are considered:

A1. Quasi-static and frequency-flat fading channels are considered, i.e., the channel parameters do not change over the length of a frame but they change from frame to frame. The assumption of frequency-flat channels can be broadened to frequency-selective channels by employing orthogonal frequency division multiple access. Moreover, the use of such channels is motivated by prior research in this field [1, 3, 25, 94, 124, 125].

A2. Over a frame, the timing offsets are modeled as deterministic but *unknown* parameters [1, 3, 25].

A3. The effect of CFO on the received signal is not considered, since the topic of TS design for MCFO estimation has been extensively addressed in the literature, e.g., see [101] and references therein.

The sampled baseband received signal, $\mathbf{y} \triangleq [y(0), \ldots, y(QL-1)]^T$, at the receiver, prior to matched filtering, is given by

$$\mathbf{y} = \mathbf{\Psi}\mathbf{h} + \mathbf{w}, \tag{4.1}$$

where:

- T denotes the symbol duration, $T_s = T/Q$ is the sampling interval, Q is the oversampling factor,
- $\boldsymbol{\Psi} \triangleq [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_K]$ is a $QL \times K$ matrix, $\boldsymbol{\xi}_k \triangleq \mathbf{G}_k \mathbf{t}_k$ is a $QL \times 1$ vector for $k = 1, \ldots, K$,
- $\mathbf{G}_k \triangleq [\mathbf{g}_0(\tau_k), \ldots, \mathbf{g}_{L-1}(\tau_k)]$ is a $QL \times L$ matrix of the samples of the pulse shaping filter such that $\mathbf{g}_n(\tau_k) \triangleq [g(-nT + \tau_k T), \ldots, g(-nT + iT_s + \tau_k T), \ldots, g(-nT + (QL-1)T_s + \tau_k T)]^T$, for $n = 0, \ldots, L-1$, $\mathbf{t}_k = [t_k(0), \ldots, t_k(L-1)]^T$ denotes the kth node's TS, $\tau_k \in (-0.5, 0.5)$ is the fractional *unknown* timing offset between the kth node and the receiver that is normalized by T,
- $\mathbf{h} \triangleq [h_1, \ldots, h_K]^T$ is the channel vector, h_k denotes the *unknown* channel gain from the kth node to the receiver that changes from frame to frame according to $h_k \sim \mathcal{CN}(0, \sigma_h^2)$, and
- $\mathbf{w} \triangleq [w(0), \ldots, w(QL-1)]^T$ and $w(i)$, for $i = 0, \ldots, QL-1$, denotes the zero-mean complex additive white Gaussian noise (AWGN) at the ith sample of the received signal, i.e., $w(i) \sim \mathcal{CN}(0, \sigma_w^2)$.

Since the root-raise cosine (RRC) waveform is widely applied in communication systems, in this paper, it is used for the pulse shaping filter, $g(t)$ [1,3,25]. Furthermore, it is assumed that a coarse synchronizer is first applied and signals from the different nodes are within the same symbol period, i.e., the difference between the timing offsets of any two nodes is given by $|\tau_k - \tau_{\bar{k}}| < 1$, for $k, \bar{k} = 1, \ldots, K$ and $k \neq \bar{k}$. Most communication systems use such coarse synchronizers before applying a fine synchronization algorithm to estimate the timing offsets within a symbol period [7,25,26]. For clarity, the indices $n, \bar{n} = 0, 1, \ldots, L-1$, $i = 0, 1, \ldots, LQ-1$, and $k, \bar{k} = 1, \ldots, K$ are reserved to denote symbols, T_s-spaced samples, and the nodes, respectively.

4.2 Estimation Lower Bounds

In this section, closed-form expressions for the HCRB for the joint estimation of deterministic MTOs and random channels are derived. Next, to find a tighter lower bound on the variance of the estimation error for these parameters, the ECRB is obtained.

4.2.1 Hybrid Cramér-Rao Lower Bounds

The HCRB is a lower bound on the joint estimation of the random and deterministic parameters and unlike the CRB is not a function of the random parameters [19,126]. In this context, the derived HCRB does not depend on the particular channels. Thus, by minimizing the HCRB the design guidelines obtained here are general and are applicable to all random realizations of the Rayleigh fading channel coefficients. In the following analysis, we make no specific assumption on the distribution of the timing offsets. For example, depending on the receiver design, the timing offsets can take on either uniform [25] or normal distributions [92], respectively. To ensure generality, we assume timing offsets to be *deterministic* and unknown parameters that can assume any value within the symbol period. Even though due to this assumption the resulting HCRB is a function of timing offsets, it is numerically shown later that this bound does not vary for different timing offset values. Furthermore, it is also numerically shown in [72] that optimality of TSs is independent of the timing offset values. Hence, the TS design guidelines obtained in the later sections are applicable to the whole range of timing offsets.

The first step in determining the HCRB is to formulate the parameter vector of interest, $\boldsymbol{\theta} \triangleq [\boldsymbol{\theta}_r^T, \boldsymbol{\theta}_d^T]^T$, which is given by

$$\boldsymbol{\theta} \triangleq [\Re\{\mathbf{h}\}^T, \Im\{\mathbf{h}\}^T, \boldsymbol{\tau}^T]^T, \tag{4.2}$$

where $\boldsymbol{\theta}_r \triangleq [\Re\{\mathbf{h}\}^T, \Im\{\mathbf{h}\}^T]^T$ is the random vector of channel parameters and $\boldsymbol{\theta}_d \triangleq \boldsymbol{\tau} \triangleq [\tau_1, \ldots, \tau_K]^T$ is the deterministic vector of MTOs. Note that according to the assumption in Section 4.1, the complex channel vector is distributed as $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}_{K\times 1}, \sigma_h^2 \mathbf{I}_{K\times K})$. In the following, the *hybrid information matrix* (HIM) and the HCRB for the estimation of $\boldsymbol{\theta}$ are formulated.

Theorem 4.1 *The HIM for the estimation of the parameters of interest, $\boldsymbol{\theta}$, given the observation vector, $\mathbf{y}$, is a $3K \times 3K$ matrix given by*

$$\mathbf{HIM} = \underbrace{\frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \Re\{\mathbf{U}\} \end{bmatrix}}_{=\mathbb{E}_{\boldsymbol{\theta}_r|\boldsymbol{\theta}_d}[\mathbf{FIM}]} + \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1} & \mathbf{0}_{2K\times K} \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times K} \end{bmatrix}, \tag{4.3}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1} \triangleq diag\Big(\underbrace{\frac{2}{\sigma_h^2}, \ldots, \frac{2}{\sigma_h^2}}_{2K}\Big)$ *is the* $2K\times 2K$ *covariance matrix of* $\boldsymbol{\theta}_r \triangleq [\Re\{\mathbf{h}\}^T, \Im\{\mathbf{h}\}^T]^T$, $\mathbf{U} \triangleq \sigma_h^2 diag\big(\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K^H\boldsymbol{\delta}_K\big)$ *is a* $K \times K$ *diagonal matrix,* $\boldsymbol{\delta}_k \triangleq \mathbf{R}_k\mathbf{t}_k \;\forall\; k$, *and*

$\mathbf{R}_k \triangleq \frac{\partial \mathbf{G}_k}{\partial \tau_k}$ *is* $QL \times L$ *matrix.*

Proof: See Appendix C.1.

HCRB is given by the inverse of the **HIM** in (4.3). In order to ensure that the **HIM** in (4.3) is full rank, the HCRB does not approach infinity, and the parameters of interest can be accurately estimated, it is essential to transmit linearly independent TSs from all the nodes. Moreover, since the off-diagonal blocks of the HIM in (4.3), i.e., upper right $2K \times K$ and lower left $K \times 2K$ submatrices of **HIM**, are zero, the HCRB matrix for the estimation of MTOs, $\mathbf{HCRB}(\boldsymbol{\tau})$, is given by

$$\mathbf{HCRB}(\boldsymbol{\tau}) = \frac{\sigma_w^2}{2}\text{diag}\left([\Re\{\mathbf{U}\}]^{-1}\right) = \frac{\sigma_w^2}{2\sigma_h^2}\left[\frac{1}{\Re\{\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1\}},\ldots,\frac{1}{\Re\{\boldsymbol{\delta}_K^H\boldsymbol{\delta}_K\}}\right]^T. \quad (4.4)$$

Similarly, by using the inverse of the upper left $2K \times 2K$ submatrix of the **HIM**, the HCRB for the estimation of the combined real and imaginary parts of the channel vector is given by

$$\begin{aligned}\mathbf{HCRB}(\mathbf{h}) &= \text{diag}\left(\mathbf{J}\left[\frac{2}{\sigma_w^2}\begin{bmatrix}\Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\}\\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\}\end{bmatrix} + \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1}\right]^{-1}\mathbf{J}^H\right)\\ &= \frac{\sigma_w^2}{2}\text{diag}\left(\mathbf{J}\underbrace{\begin{bmatrix}\Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} + \frac{\sigma_w^2}{\sigma_h^2}\mathbf{I}_{K\times K} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\}\\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} + \frac{\sigma_w^2}{\sigma_h^2}\mathbf{I}_{K\times K}\end{bmatrix}^{-1}}_{\triangleq\boldsymbol{\Gamma}}\mathbf{J}^H\right),\end{aligned} \quad (4.5)$$

where $\mathbf{J} = [\mathbf{I}_{K\times K} \quad j\mathbf{I}_{K\times K}]$ is $K \times 2K$ matrix used to obtain the HCRB of $\mathbf{h}$ from the HCRB of $\boldsymbol{\theta}_r$ [19].

4.2.2 Conditional Cramér-Rao Lower Bounds

The HCRB derived in Section 4.2.1 may not be a tight lower bound on the estimation error variance of an estimator since the Fisher's information matrix, **FIM**, in (C.3), depends on $\boldsymbol{\theta}$ [19, page 7]. Thus, in this section, the ECRB, which is a tighter lower bound is derived. Following [19, page 6], ECRB is given by

$$\mathbf{ECRB} = \mathbb{E}_{\boldsymbol{\theta}_r}\left[\mathbf{FIM}^{-1}\right], \quad (4.6)$$

In (4.6), the expectation is taken with respect to the priori distribution of $\boldsymbol{\theta}_r$, $p(\boldsymbol{\theta}_r) = \frac{\exp\{-\mathbf{h}^H\boldsymbol{\Sigma}_\mathbf{h}^{-1}\mathbf{h}\}}{(\pi)^K\det(\boldsymbol{\Sigma}_\mathbf{h})}$ and $\boldsymbol{\Sigma}_\mathbf{h} \triangleq \text{diag}(\underbrace{\sigma_h^2,\ldots,\sigma_h^2}_{K})$ is the covariance matrix of

$\mathbf{h}$. Using the closed-form results for the inverse of the Fisher's information matrix, $\mathbf{FIM}^{-1}$, given in [25], the ECRB for joint estimation of MTOs and multiple channel gains can be determined as

$$\mathbf{ECRLB}(\boldsymbol{\tau}) = \frac{\sigma_w^2}{2}\mathbb{E}_{\boldsymbol{\theta}}\left[\left(\Re\left\{\mathbf{D}^H\boldsymbol{\Delta}^H\left(\mathbf{I}_{LQ} - \boldsymbol{\Psi}(\boldsymbol{\Psi}^H\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^H\right)\boldsymbol{\Delta}\mathbf{D}\right\}\right)^{-1}\right], \tag{4.7a}$$

$$\mathbf{ECRLB}(\mathbf{h}) = \frac{\sigma_w^2}{2}\mathbb{E}_{\boldsymbol{\theta}}\Big[2(\boldsymbol{\Psi}^H\boldsymbol{\Psi})^{-1} + (\boldsymbol{\Psi}^H\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D} \left(\Re\left\{\mathbf{D}^H\boldsymbol{\Delta}^H\left(\mathbf{I}_{LQ} - \boldsymbol{\Psi}(\boldsymbol{\Psi}^H\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^H\right)\boldsymbol{\Delta}\mathbf{D}\right\}\right)^{-1}\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}(\boldsymbol{\Psi}^H\boldsymbol{\Psi})^{-1}\Big], \tag{4.7b}$$

where $\mathbf{D} \triangleq \text{diag}(h_1, \ldots, h_K)$ is a $K \times K$ diagonal matrix and $\boldsymbol{\Delta} \triangleq \frac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\tau}} \triangleq [\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K]$ is a $QL \times K$ matrix. It is not mathematically tractable to find closed-form expressions for $\mathbf{ECRB}(\boldsymbol{\tau})$ and $\mathbf{ECRB}(\mathbf{h})$ in (4.7). Therefore, in this paper, the expectation with respect to $\boldsymbol{\theta}$ in (4.7) is numerically calculated over a large number of simulated realizations. Using Jensen's inequality [19] and the fact that the matrix, $\mathbf{FIM}$, in (C.3), depends on $\boldsymbol{\theta}$, the relation between ECRB and HCRB can be determined as [19, page 7]

$$\mathbf{ECRB}(\boldsymbol{\tau}) > \mathbf{HCRB}(\boldsymbol{\tau}) \qquad \text{and} \qquad \mathbf{ECRB}(\mathbf{h}) > \mathbf{HCRB}(\mathbf{h}). \tag{4.8}$$

The above relationships can also be numerically observed through simulations in Fig. 4.1, where HCRB and ECRB for MTOs estimation are evaluated using (4.4) and (4.7a), respectively, for different signal-to-noise-ratios (SNRs). The results in Fig. 4.1 are obtained for $K = 4$ nodes, TS length $L = 64$, oversampling factor $Q = 2$, and random TSs, i.e., $\mathbf{t}_k = [\exp(-j\phi_0), \ldots, \exp(-j\phi_{L-1})]^T$, $\forall\ k$, and $\phi_n \sim \mathcal{U}(-\pi, \pi)$, $\forall\ n$. Without loss of generality, only the lower bounds for the first node are plotted for different values of the timing offsets, $\tau = [-0.5 - 0.4, \ldots, 0.5]$. Moreover, the HCRB and the ECRB values are averaged over 500 trials, where for each trial, the timing offsets for the remaining nodes are assumed to be uniformly distributed over $(-0.5, 0.5)$. It can be observed from Fig. 4.1 that the HCRB is a lower bound compared to the ECRB for both SNR = 10 dB and SNR = 20 dB. This outcome is also confirmed for the different TSs in Section 4.5. Fig. 4.1 also shows that both the HCRB and the ECRB do not vary for different timing offset values, which is also confirmed later for all the proposed TSs in Section 4.5. Consequently, although the HCRB is a function of the timing offsets, the TS design guidelines obtained by minimizing the HCRB in Section 4.3 are independent of the timing offset values.

Recall that even though the ECRB is a tighter bound compared to the HCRB,

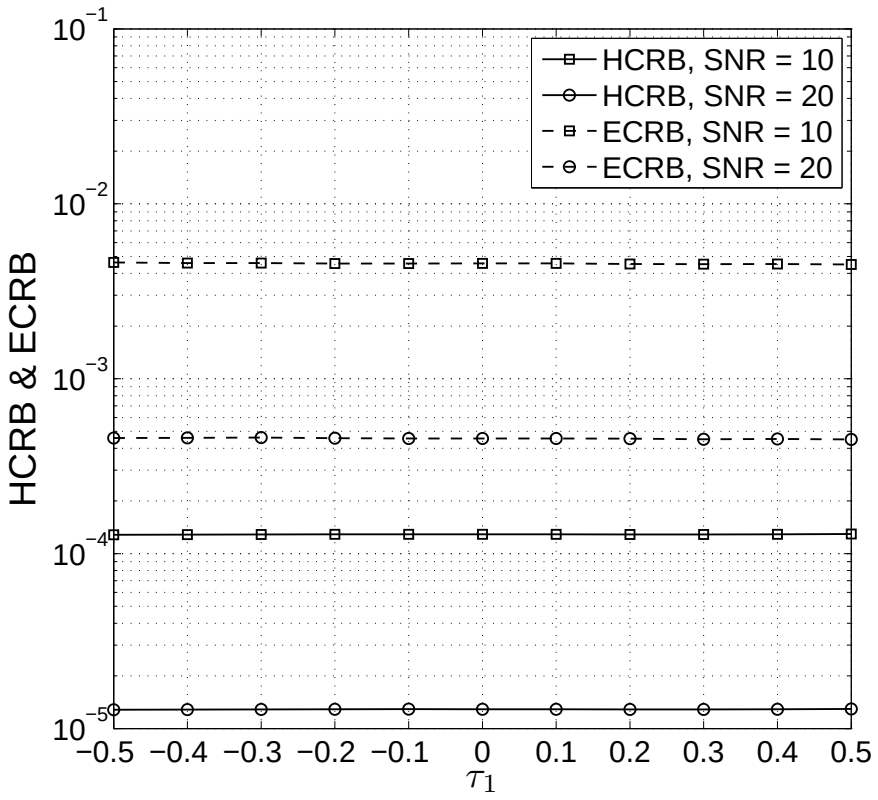


Figure 4.1: HCRB and ECRB for the estimation of τ_1 for the different values of τ_1 with $K = 4$ nodes.

it cannot be derived in closed-form. Thus, analytical solutions for the optimal TSs cannot be obtained by minimizing the ECRB. As a result, in Section 4.3, guidelines for optimal TS design are obtained by minimizing the HCRB instead. Nevertheless, in Section 4.5, through numerical simulations it is demonstrated that the TSs that minimize the HCRB also minimize the ECRB and the MSE of the derived MAP estimator.

4.3 Training Sequence Design

In this paper, the optimal TS is defined as the TS that jointly minimizes the HCRBs of $\boldsymbol{\tau}$ and $\mathbf{h}$. The following subsections present criteria required for minimizing the HCRBs of $\boldsymbol{\tau}$ and $\mathbf{h}$.

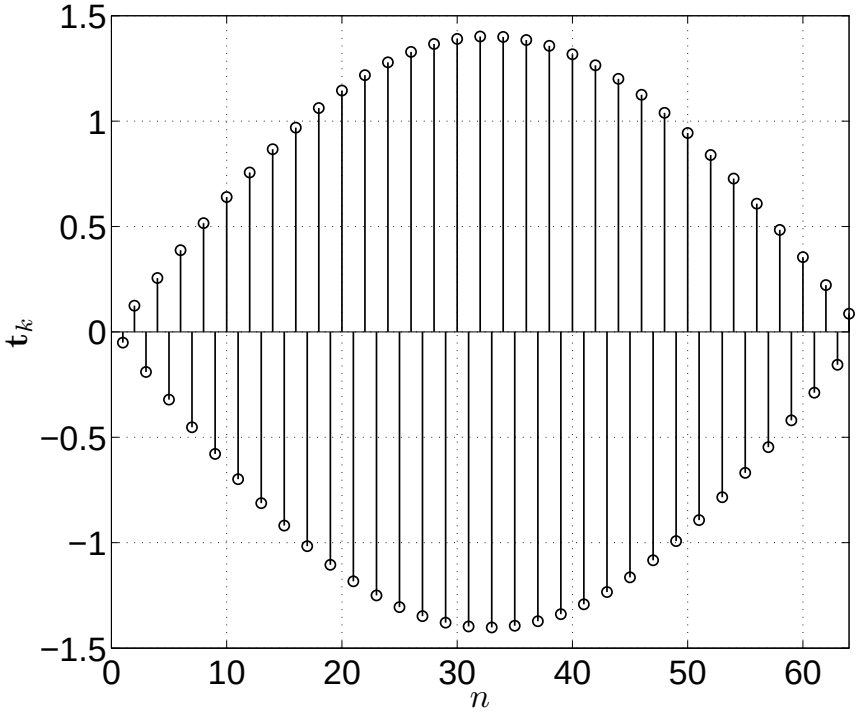


Figure 4.2: Optimal solution to (4.9) for the minimization of $\mathbf{HCRB}(\boldsymbol{\tau})$, where TS length $L = 64$ and $\mathbf{R}_k$ is evaluated at $\tau_k = 0.4$.

4.3.1 Minimization of HCRB($\boldsymbol{\tau}$)

To minimize $\mathbf{HCRB}(\boldsymbol{\tau})$ in (4.4), we have to maximize each of $\Re\{\boldsymbol{\delta}_1^H \boldsymbol{\delta}_1\}, \ldots, \Re\{\boldsymbol{\delta}_K^H \boldsymbol{\delta}_K\}$. Using the definition of $\boldsymbol{\delta}_k \triangleq \mathbf{R}_k \mathbf{t}_k$ below (4.3), for $k = 1, \ldots, K$, the optimal TS, $\mathbf{t}_k$, that minimizes the HCRB of $\boldsymbol{\tau}$, is the solution to

$$\arg\max_{\mathbf{t}_k} \quad \Re\{\mathbf{t}_k^H \mathbf{R}_k^T \mathbf{R}_k \mathbf{t}_k\}, \qquad \texttt{s.t.} \quad \mathbf{t}_\mathtt{k}^\mathtt{H} \mathbf{t}_\mathtt{k} \leq \mathtt{L}, \tag{4.9}$$

where $\mathbf{R}_k \triangleq \frac{\partial \mathbf{G}_k}{\partial \tau_k}$ is a matrix of real numbers, since elements of matrix $\mathbf{G}_k$ are samples of the real RRC waveform. Given that $\mathbf{x}^H \mathbf{R}_k^T \mathbf{R}_k \mathbf{x} = \|\mathbf{R}_k \mathbf{x}\|^2 \geq 0$ for any $L \times 1$ vector $\mathbf{x}$ and all eigenvalues of $\mathbf{R}_k^T \mathbf{R}_k$ are greater than zero, $\mathbf{R}_k^T \mathbf{R}_k$ is a symmetric positive definite matrix and the optimization problem in (4.9) is convex [127]. Thus, the optimal solution to (4.9) is given by

$$\mathbf{t}_k = \sqrt{L} \lambda_{\max} \left(\mathbf{R}_k^T \mathbf{R}_k \right), \tag{4.10}$$

where $\lambda_{\max} \left(\mathbf{R}_k^T \mathbf{R}_k \right)$ is the eigenvector corresponding to the maximum eigenvalue of $\mathbf{R}_k^T \mathbf{R}_k$. The resulting TS is shown in Fig. 4.2 where $\mathbf{R}_k$ is evaluated at $\tau_k = 0.4$ and TS length $L = 64$. It is important to note that the TS in Fig. 4.2 undergoes a

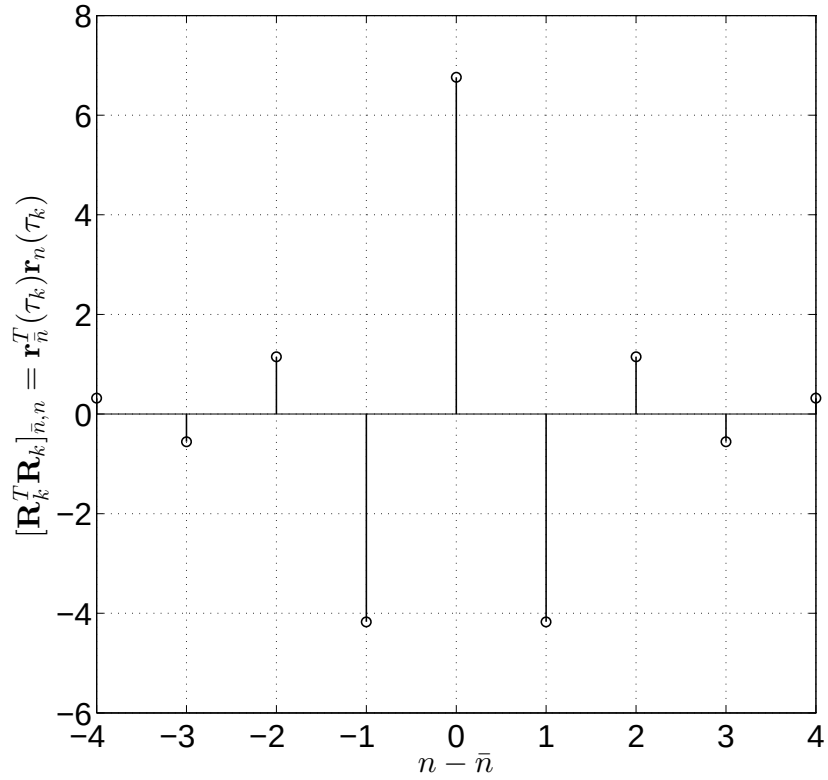


Figure 4.3: $[\mathbf{R}_k^T\mathbf{R}_k]_{\bar{n},n} = \mathbf{r}_{\bar{n}}^T(\tau_k)\mathbf{r}_n(\tau_k)$ for the different values of $n-\bar{n}$ with $\tau_k = 0.4$, $n = L/2$ and $L = 64$.

sign change from symbol to symbol and similar TSs, as shown in Fig. 4.2, are also obtained by evaluating (4.10) for different values of τ_k.

Remark 4.1 *To develop more insight into the optimal solution in* (4.10) *and establish a more comprehensive TS design guideline that maximizes the cost function in* (4.9)*, we numerically study the structure of the matrix* $\mathbf{R}_k$*. Let us write* $\mathbf{R}_k$ *in terms of its column vectors,* $\mathbf{R}_k = [\mathbf{r}_0(\tau_k), \ldots, \mathbf{r}_{L-1}(\tau_k)]$*, where* $\mathbf{r}_n(\tau_k) = [r(-nT + \tau_k T), \ldots, r(-nT + iT_s + \tau_k T), \ldots, r(-nT + (QL-1)T_s + \tau_k T)]^T$*,* $\forall\ n$*, and* $r(t) \triangleq \frac{\partial g(t)}{\partial \tau_k}$*.* $\mathbf{R}_k^T\mathbf{R}_k$ *is an* $L \times L$ *matrix such that*

$$\left[\mathbf{R}_k^T\mathbf{R}_k\right]_{\bar{n},n} = \mathbf{r}_{\bar{n}}^T(\tau_k)\mathbf{r}_n(\tau_k). \tag{4.11}$$

Fig. 4.3 plots the elements of matrix $\mathbf{R}_k^T\mathbf{R}_k$*, e.g., diagonal elements* $(n-\bar{n}=0)$*, the first diagonal above* $(n-\bar{n}=1)$ *and below the main diagonal* $(n-\bar{n}=-1)$*, the second diagonal above* $(n-\bar{n}=2)$ *and below the main diagonal* $(n-\bar{n}=-2)$*, and so on for* $n = L/2$*,* $L = 64$*, and* $\tau_k = 0.4$*. It can be observed from Fig. 4.3 that for the matrix* $\mathbf{R}_k^T\mathbf{R}_k$*, the elements corresponding to even values of* $(n-\bar{n})$ *are positive*

while the elements for the odd values of $(n-\bar{n})$ *are negative*[1]*. Similar results are also observed* $\forall\ n, \tau_k$*. Consequently, the cost function in* (4.9)*,* $\Re\{\mathbf{t}_k^H \mathbf{R}_k^T \mathbf{R}_k \mathbf{t}_k\}$*, is maximized, when the transmitted TS alternates in sign every symbol period since based on the structure of* $\mathbf{R}_k^T \mathbf{R}_k$*, opposite-signed TS symbols are multiplied by the negative-valued elements of* $\mathbf{R}_k^T \mathbf{R}_k$*. Similarly, it can be concluded that for complex TSs, the cost function,* $\Re\{\mathbf{t}_k^H \mathbf{R}_k^T \mathbf{R}_k \mathbf{t}_k\}$ *is maximized by transmitting the TS that exhibits a phase shift of* π *radians from symbol to symbol. Fig. 4.3 also shows that the magnitude of elements* $\mathbf{R}_k^T \mathbf{R}_k$ *decay as* $|n-\bar{n}|$ *increases, since the RRC function decays with every sample.*

The above analysis shows that TSs that undergo a π *radian phase shift every symbol period are consistent with the optimal solution that minimizes the HCRB for MTO estimation. Such TSs with sign changes are also reported as optimal based on intuition or simulation in [12, 72]. However, no analytical results were included to support these claims in [72] and [12].*

In [72], it is shown that the TS in Fig. 4.2 is optimal for the estimation of a *single* timing offset in point-to-point MIMO systems. In this paper, our focus is to design optimal TSs for joint estimation of MTO and multiple channel parameters, which will be obtained by jointly minimizing $\mathbf{HCRB}(\boldsymbol{\tau})$ and $\mathbf{HCRB}(\mathbf{h})$.

4.3.2 Minimization of HCRB(h)

To minimize $\mathbf{HCRB}(\mathbf{h})$ in (4.5), a closed-form expression for the matrix inverse, $\mathbf{\Gamma}$, in (4.5) needs to be determined. For asymptotically large values of TS length, L,

$$\operatorname{diag}(\Re\{\mathbf{\Psi}^H \mathbf{\Psi}\}) \gg \operatorname{diag}\left(\frac{\sigma_w^2}{\sigma_h^2}\mathbf{I}_{K\times K}\right), \tag{4.12}$$

since for large values of L the contribution of the prior information matrix to the HCRB is considerably smaller than $\mathbb{E}_{\boldsymbol{\theta}}[\mathbf{FIM}]$ [19]. This is also numerically shown in Fig. 4.4, where $\frac{\sigma_w^2/\sigma_h^2}{[\mathbf{\Psi}^H\mathbf{\Psi}]_{1,1}}$ is plotted versus TS length, L, for oversampling factor $Q=2$, SNR $=10$ dB ($\sigma_w^2=0.1$), and $\sigma_h^2=1$. Random TSs are used similar to those of Section 4.2.2.[2]

[1] This is also because the slope of the RRC function, $r(t)$, changes after every time period T.

[2] Though results for only first node are presented, similar results are observed for the remaining nodes.

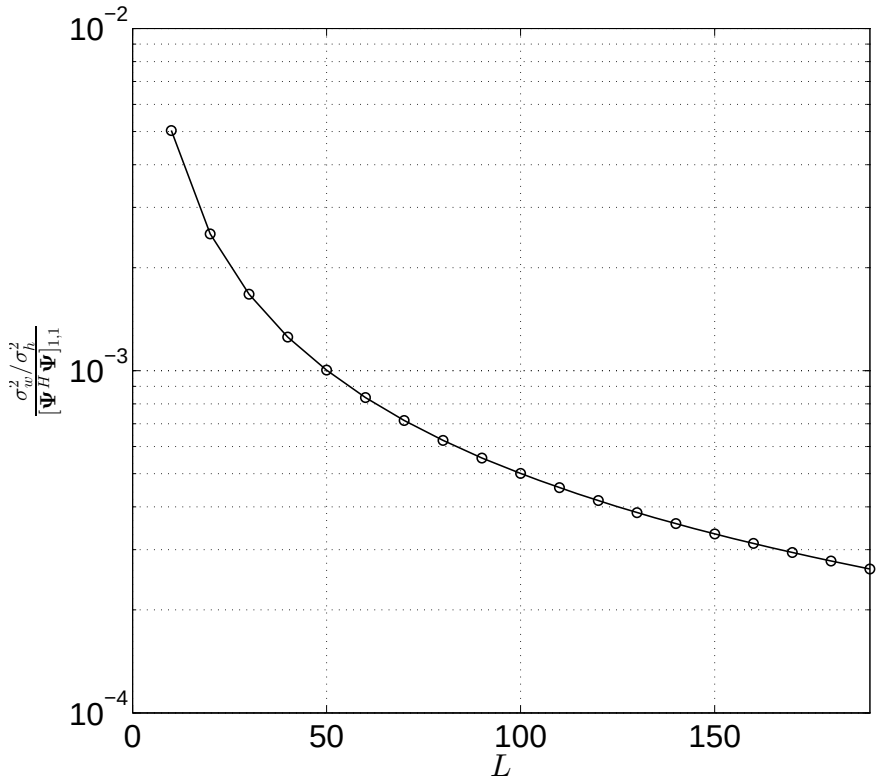


Figure 4.4: $\frac{\sigma_w^2/\sigma_h^2}{[\mathbf{\Psi}^H\mathbf{\Psi}]_{1,1}}$ versus TS length, L, for oversampling factor $Q = 2$, SNR = 10 dB ($\sigma_w^2 = 0.1$), and $\sigma_h^2 = 1$.

Using (4.12), (4.5) can be approximated as

$$\begin{aligned}\widetilde{\mathbf{HCRB}}(\mathbf{h}) &= \frac{\sigma_w^2}{2}\mathrm{diag}\left(\mathbf{J}\begin{bmatrix} \Re\{\mathbf{\Psi}^H\mathbf{\Psi}\} & -\Im\{\mathbf{\Psi}^H\mathbf{\Psi}\} \\ \Im\{\mathbf{\Psi}^H\mathbf{\Psi}\} & \Re\{\mathbf{\Psi}^H\mathbf{\Psi}\} \end{bmatrix}^{-1}\mathbf{J}^H\right) \\ &= \sigma_w^2\mathrm{diag}\Big([\mathbf{\Psi}^H\mathbf{\Psi}]^{-1}\Big), \end{aligned} \tag{4.13}$$

where $\widetilde{\mathbf{HCRB}}(\mathbf{h})$ is the asymptotic HCRB for the estimation of channel, $\mathbf{h}$. According to (4.13), minimizing $\widetilde{\mathbf{HCRB}}(\mathbf{h})$ is equivalent to minimizing $\mathrm{Tr}\left([\mathbf{\Psi}^H\mathbf{\Psi}]^{-1}\right)$. Based on the results in [18, page 65], the following lemma applies.

Lemma 4.1 *For a $K \times K$ positive definite matrix $\mathbf{X}$, $Tr(\mathbf{X}^{-1}) \geq \sum_{k=1}^{K} \frac{1}{[\mathbf{X}]_{k,k}}$, with equality sign applying if $\mathbf{X}$ is diagonal.*

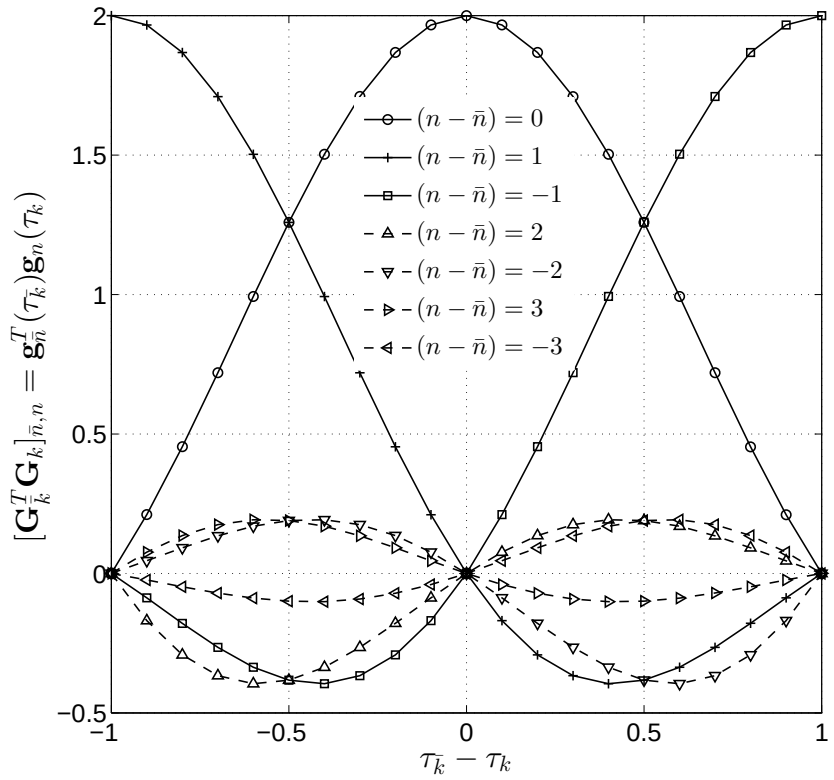


Figure 4.5: $[\mathbf{G}_{\bar{k}}^T\mathbf{G}_k]_{\bar{n},n}=\mathbf{g}_{\bar{n}}^T(\tau_{\bar{k}})\mathbf{g}_n(\tau_k)$ for different values of τ_k and $\tau_{\bar{k}}$, such that $|\tau_{\bar{k}}-\tau_k|<1$, $n=L/2$ and $L=64$.

Using Lemma 4.1, for the positive definite matrix $\mathbf{\Psi}^H\mathbf{\Psi}$, we have

$$\mathrm{Tr}\left([\mathbf{\Psi}^H\mathbf{\Psi}]^{-1}\right)\geq\sum_{k=1}^{K}\frac{1}{[\mathbf{\Psi}^H\mathbf{\Psi}]_{k,k}}. \tag{4.14}$$

Proposition 4.1 $\mathbf{\Psi}^H\mathbf{\Psi}$ *is a positive definite matrix.*

Proof: See Appendix C.2.

Thus, using (4.14) and Proposition 4.1, HCRB of $\mathbf{h}$ can be minimized by ensuring that the off diagonal elements of the $K\times K$ matrix $\mathbf{\Psi}^H\mathbf{\Psi}$, whose elements are given by

$$[\mathbf{\Psi}^H\mathbf{\Psi}]_{\bar{k},k}=\mathbf{t}_{\bar{k}}^H\mathbf{G}_{\bar{k}}^T\mathbf{G}_k\mathbf{t}_k,\qquad k,\bar{k}=1,\ldots,K, \tag{4.15}$$

are zero, i.e., $\mathbf{\Psi}^H\mathbf{\Psi}$ is a diagonal matrix. In order to minimize the off-diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$, the structure of the $L\times L$ matrix $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, which depends on the values of the timing offsets from different nodes needs to be analyzed. Let us

consider the diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$ first. Since after a symbol period T, the RRC function decays very quickly, it is observed numerically that the matrix $\mathbf{G}_k^T\mathbf{G}_k$ is diagonally dominant for any value of τ_k. Consequently, the diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$ that are given by (4.15), with $k = \bar{k}$, are maximized if $\mathbf{t}_k^H\mathbf{t}_k = L$.

The off-diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$ are given by (4.15), for $k \neq \bar{k}$, which depend on the structure of the matrix $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, for $k \neq \bar{k}$. It is numerically observed that $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$ is a tri-diagonally dominant matrix for any value of τ_k and $\tau_{\bar{k}}$ with $|\tau_{\bar{k}} - \tau_k| < 1$. This is also numerically demonstrated in Fig. 4.5, where the elements of $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, $[\mathbf{G}_{\bar{k}}^T\mathbf{G}_k]_{n,\bar{n}} = \mathbf{g}_n(\tau_{\bar{k}})^T\mathbf{g}_{\bar{n}}(\tau_k)$, are plotted for different values of $|\tau_{\bar{k}} - \tau_k|$ and $n - \bar{n}$. Fig. 4.5 shows that given $|\tau_{\bar{k}} - \tau_k| < 1$, the diagonal elements, $n - \bar{n} = 0$, the first diagonal elements above the main diagonal, $n - \bar{n} = 1$, and the first diagonal elements below the main diagonal, $n - \bar{n} = -1$, of matrix $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$ are dominant. Consequently, to minimize the off-diagonal elements of $\mathbf{\Psi}^H\mathbf{\Psi}$, i.e., $[\mathbf{\Psi}^H\mathbf{\Psi}]_{\bar{k},k} = \mathbf{t}_{\bar{k}}^H\mathbf{G}_{\bar{k}}^T\mathbf{G}_k\mathbf{t}_k$, for $k \neq \bar{k}$, the following two conditions can be developed:

1. The TSs from different nodes need to be mutually orthogonal.

2. The TSs from any node should also be mutually orthogonal to $+T$-shifted TSs and $-T$-shifted TSs from every other node to minimize effects on the first diagonals above and below the main diagonal elements of $\mathbf{G}_{\bar{k}}^T\mathbf{G}_k$, respectively.

4.3.3 Conditions for the Optimal TSs and the Proposed Training Sequences

In summary, from Section 4.3.1 and Section 4.3.2, the following three conditions can be obtained for the design of optimal TSs:

C1: The TSs from all nodes exhibit a phase shift of π radians every symbol period, which corresponds to (4.10).

C2: The TSs from all nodes are mutually orthogonal.

C3: The TSs from any node are orthogonal to $\pm$ T-shifted TSs from every other node.

As highlighted in Chapter 1, the design guidelines in [12] are not complete. In this paper, according to the proposed analytical framework, a new criterion, C3, is proposed. The results in Section 4.5 demonstrate that neglecting C3 in the C3 in the design of TSs significantly degrades joint channel and MTO estimation performance. Moreover, guideline C1 is generalized for complex TSs. It is not easy to find the TSs that jointly satisfy C1-C3, since the TSs that satisfy C2 and C3 do

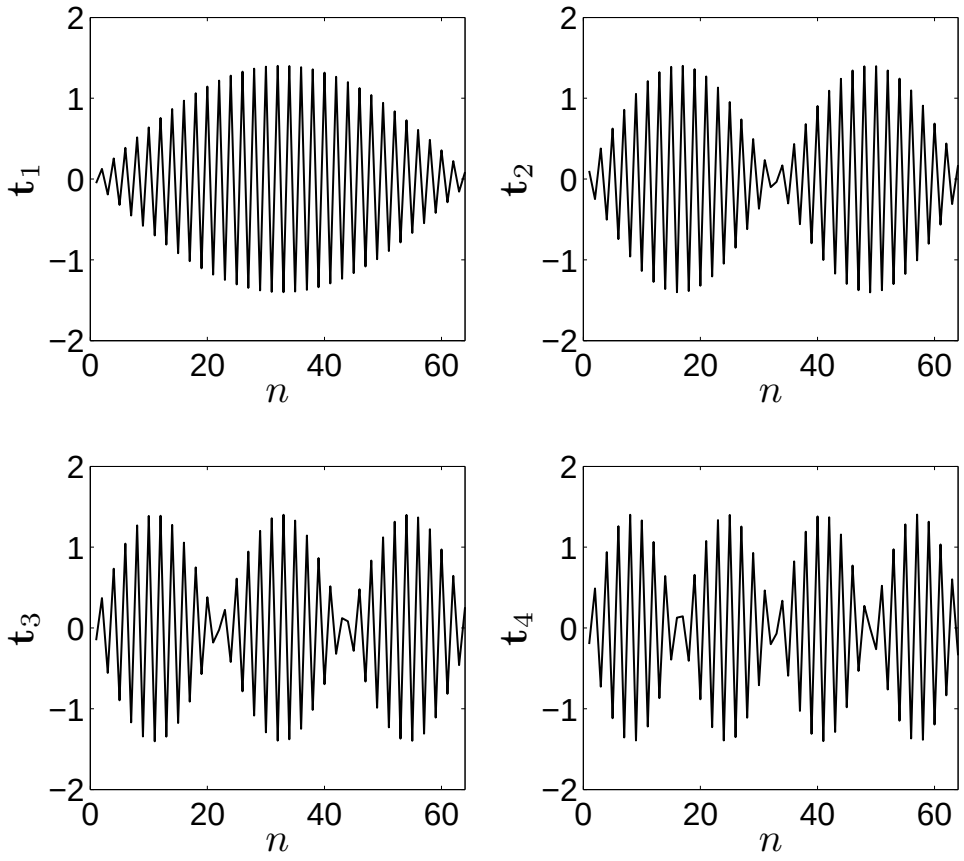


Figure 4.6: TSs in Proposed Example 1 for $K = 4$ relays and TS length $L = 64$.

not exactly satisfy C1 and vice versa. However, in what follows, we propose two techniques for generating TSs that nearly satisfy the above conditions and improve the estimation and relaying performance.

4.3.3.1 Proposed Technique 1 (Proposed-1)

Although the TSs obtained by optimizing the cost function in (4.9) exhibit a π-radian phase shift every symbol period, they do not meet conditions C2 and C3. To satisfy C2, we propose to exploit the K eigenvectors of $\mathbf{R}_k^T\mathbf{R}_k$ since they are mutually orthogonal. Consequently, K TSs can be obtained by multiplying the K eigenvectors corresponding to the first K maximum eigenvalues of $\left(\mathbf{R}^{[0]}\right)^T\mathbf{R}^{[0]}$, where $\mathbf{R}^{[0]} \triangleq \frac{\partial \mathbf{G}}{\partial \tau}\Big|_{\tau=0}$, by $\sqrt{L}$. This technique can be used to construct TSs of any desired length. For example, the TSs for $K = 4$ nodes and TS length $L = 64$ are shown in Fig. 4.6. It can be observed from Fig. 4.6 that all 4 TSs exhibit π radian phase shift every symbol period satisfying C1. In addition, C2 is also satisfied because eigenvectors of $\left(\mathbf{R}^{[0]}\right)^T\mathbf{R}^{[0]}$ are mutually orthogonal. Moreover,

even though C3 is not exactly satisfied, every TS is nearly orthogonal to $\pm T$ shifted TSs from every other node and the numerical results in Section 4.5 show that the proposed TSs significantly enhance estimation and system performance. Note that the Proposed-1 TS matches with the TS proposed in [72] for single timing offset estimation in MIMO systems.

4.3.3.2 Proposed Technique 2 (Proposed-2)

We propose to use the Walsh-Hadamard codes as a second approach for generating the optimal TSs since they always meet condition C2 [128]. The Walsh-Hadamard codes here are generated using the initial matrix $\mathbf{W} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, which is repeated $\log_2(L) - 1$ times according to $\mathbf{W} = \begin{bmatrix} \mathbf{W} & \mathbf{W} \\ \mathbf{W} & -\mathbf{W} \end{bmatrix}$ to get an $L \times L$ Hadamard matrix, $\mathbf{W}$, where the TS length in this case is $L = 2^q$, for any positive integer q. To satisfy C1, we propose to select the columns of $\mathbf{W}$ as TSs that exhibit maximum phase shift from symbol to symbol, i.e., $L-1, \ldots, L-K$ sign changes for K nodes. Similar to the technique above, the proposed TSs do not fully satisfy conditions C1 and C3 but they are nearly optimal and the results in Section 4.5 show that they considerably enhance system performance.

In (4.9), the necessary and sufficient condition for the design of training sequences (TSs) that minimize $\mathbf{HCRB}(\boldsymbol{\tau})$ is provided. In (4.10) the solution for meeting this condition is also derived. Moreover, in (4.13) and Lemma 1, for asymptotically large TS length, L, the necessary and sufficient condition for minimizing $\mathbf{HCRB}(\mathbf{h})$ is also established. Although based on these conditions the TS design guidelines C1–C3 are derived, it cannot be analytically established that C1–C3 are also necessary conditions to minimize $\mathbf{HCRB}(\boldsymbol{\tau})$ and $\mathbf{HCRB}(\mathbf{h})$. Nevertheless, our simulation results in Section VI indicate that the proposed conditions C1–C3 are necessary conditions for optimal TS design because the proposed TSs based on C1–C3 achieve a BER performance that is close to the benchmark BER. Moreover, the TSs that violate these conditions demonstrate poor BER performance.

4.4 Maximum-A-Posteriori Estimator

In this section, a *data-aided* MAP estimator for the joint estimation of MTOs and multiple channels is derived. The MAP estimate of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}$ maximizes the log of the

posterior probability density function (PDF), given by

$$\log p(\boldsymbol{\theta}|\mathbf{y}) = \log \left\{ \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} \right\}, \tag{4.16}$$

where $p(\mathbf{y}|\boldsymbol{\theta})$ is the likelihood function given by

$$p(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{(\pi\sigma_w^2)^{LQ}} \exp \left\{ -\frac{\|\mathbf{y} - \boldsymbol{\Psi}\mathbf{h}\|^2}{\sigma_w^2} \right\}, \tag{4.17}$$

and $p(\boldsymbol{\theta})$ is the priori distribution of $\boldsymbol{\theta}$, given by

$$p(\boldsymbol{\theta}) = \frac{1}{(\pi)^K \det(\boldsymbol{\Sigma}_\mathbf{h})} \exp \left\{ -\mathbf{h}^H \boldsymbol{\Sigma}_\mathbf{h}^{-1} \mathbf{h} \right\}. \tag{4.18}$$

In (4.18), $\boldsymbol{\Sigma}_\mathbf{h} \triangleq \text{diag}\big(\underbrace{\sigma_h^2, \ldots, \sigma_h^2}_{K} \big)$ is the covariance matrix of $\mathbf{h}$ and $\boldsymbol{\Sigma}_\mathbf{h}^{-1} \triangleq \text{diag}\Big(\underbrace{\frac{1}{\sigma_h^2}, \ldots, \frac{1}{\sigma_h^2}}_{K} \Big)$. Substituting (4.17) and (4.18) into (4.16), the posterior PDF in (4.16) can be written as

$$\log p(\boldsymbol{\theta}|\mathbf{y}) = -\log(\pi\sigma_h^2)^K (\pi\sigma_w^2)^{LQ} - \frac{\|\mathbf{y} - \boldsymbol{\Psi}\mathbf{h}\|^2}{\sigma_w^2} - \mathbf{h}^H \boldsymbol{\Sigma}_\mathbf{h}^{-1} \mathbf{h} - \log p(\mathbf{y}), \tag{4.19}$$

where $\log(\pi\sigma_h^2)^K (\pi\sigma_w^2)^{LQ}$ and $\log p(\mathbf{y})$ are independent of $\boldsymbol{\theta}$. Thus, the MAP estimate of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\{ \frac{\|\mathbf{y} - \boldsymbol{\Psi}\mathbf{h}\|^2}{\sigma_w^2} + \mathbf{h}^H \boldsymbol{\Sigma}_\mathbf{h}^{-1} \mathbf{h} \right\}. \tag{4.20}$$

Taking the derivative of the above cost function with respect to $\mathbf{h}^H$ and equating the result to zero, the MAP estimate of channels, $\hat{\mathbf{h}}$, is determined as

$$\hat{\mathbf{h}} = \left(\boldsymbol{\Psi}^H \boldsymbol{\Psi} + \boldsymbol{\Sigma}_\mathbf{h}^{-1} \right)^{-1} \boldsymbol{\Psi}^H \mathbf{y}. \tag{4.21}$$

Substituting (4.21) back into the cost function in (4.20), the MAP estimate of MTOs, $\hat{\boldsymbol{\tau}}$, is determined as

$$\hat{\boldsymbol{\tau}} = \arg\max_{\boldsymbol{\tau}} \left\{ \mathbf{y}^H \boldsymbol{\Psi} \left(\boldsymbol{\Psi}^H \boldsymbol{\Psi} + \boldsymbol{\Sigma}_\mathbf{h}^{-1} \right)^{-1} \boldsymbol{\Psi}^H \mathbf{y} \right\}. \tag{4.22}$$

Using the estimate of $\hat{\boldsymbol{\tau}}$ from (4.22) and by evaluating $\boldsymbol{\Psi}$ at $\boldsymbol{\tau} = \hat{\boldsymbol{\tau}}$, the channel estimates, $\hat{\mathbf{h}}$, can be obtained using (4.21). Since $\mathbf{y}$ and $\boldsymbol{\theta}$ are not jointly Gaussian, it cannot be analytically shown that the proposed MAP estimator in (4.21) and

Table 4.1: Summary of TSs used in the simulations with TS length $L = 64$ and $K = 4$ nodes.

Proposed-1	$\sqrt{L}$ times the eigenvectors corresponding to the first 4 maximum eigenvalues of $\left(\mathbf{R}^{[0]}\right)^T \mathbf{R}^{[0]}$, $\left(\mathbf{R}^{[0]}\right.$ is defined in Sec. 4.3.3.1 $\left.\right)$
Proposed-2	Walsh-Hadamard Matrix $\mathbf{W}$ given in Sec. 4.3.3.2 (Column No. 2, 18, 34, and 50)
Violate-C1	Walsh-Hadamard Matrix $\mathbf{W}$ given in Sec. 4.3.3.2 (Column No. 1, 17, 33, and 49)
Violate-C2	$\mathbf{t}_1 = [1.00 + j0.00, -0.92 + j0.38, 0.70 - j0.70, -0.38 + j0.92, \ldots]^T$ $\mathbf{t}_2 = [0.92 + 0.38, -0.70 - j0.70, 0.38 + j0.92, 0.00 - j1.00, \ldots]^T$ $\mathbf{t}_3 = [0.70 + j0.70, 0.00 - 1.00, 0.70 - j0.70, 0.00 + j1.00, \ldots]^T$ $\mathbf{t}_4 = [0.00 + j1.00, 0.00 - j1.00, 0.00 + j1.00, 0.00 - j1.00, \ldots]^T$
Violate-C3	Walsh-Hadamard Matrix $\mathbf{W}$ given in Sec. 4.3.3.2 (Column No. 5, 6, 7, and 8)

(4.22) is unbiased [18]. However, through simulations, we have found that the estimated parameters in (21) and (22) have a very small bias, e.g., at SNR = 10 dB, it is found that $\mathbb{E}\{\boldsymbol{\tau} - \hat{\boldsymbol{\tau}}\} = -0.0027$ and $\mathbb{E}\{\mathbf{h} - \hat{\mathbf{h}}\} = 0.0009 - j0.0013$, when averaged over 500 simulations. In the following section, the estimation accuracy of the proposed MAP estimator will be assessed by the ECRBs, $\mathbf{ECRB}(\boldsymbol{\tau})$ and $\mathbf{ECRB}(\mathbf{h})$, given in (4.7).

4.5 Simulation Results

In order to demonstrate the advantage of the proposed TSs for improving the estimation and system performance, they are compared to TSs that violate the proposed design guidelines, C1-C3. The two proposed TSs and three non-optimal TSs, which violate C1, C2, and C3, respectively, are given in Table 4.1 for TS length $L = 64$. Throughout this section, the proposed TSs in Sections 4.3.3.1 and 4.3.3.2 are referred to as Proposed-1 and Proposed-2. The three non-optimal TSs are referred to as Violate-C1, Violate-C2, and Violate-C3 and are chosen as follows.

Walsh-Hadamard codes are used for the non-optimal TSs denoted by Violate-C1. Even though these TSs satisfy C2, they are selected in a fashion to not satisfy C1, i.e., have the least frequent sign changes. As shown in Section 4.3.1 this choice of TSs is expected to increase the HCRB for the estimation of $\boldsymbol{\tau}$. The next TSs referred to as Violate-C2, are generated using 16-phase-shift keying (16-PSK) modulated symbols. It is not possible to completely satisfy C1 by having symbol-to-symbol phase shifts of π radians for all the TSs since the resulting TSs will be linearly dependent and the **HIM** in (4.3) will be ill-conditioned. However, in order

to comply with C1 as closely as possible, the TSs are selected to have maximum phase shift, close to π radians, e.g., π, $7\pi/8$, $9\pi/8$, $6\pi/8$ radians, from symbol to symbol. Moreover, the resulting TSs are correlated and violate C2 to a large extent, which according to the results in Section 4.3.2, increases the HCRB for the channel estimation. Finally, for Violate-C3 TSs, the Walsh-Hadamard codes are applied again, satisfying C2. However, these codes are selected to ensure that C3 is mainly violated, i.e., the inner product of $\mathbf{t}_k$ and ± 1 shifted $\mathbf{t}_{\bar{k}}$, for $k \neq \bar{k}$, is large. As shown in Section 4.3.2, this also results in a larger HCRB for estimation of channels, $\mathbf{h}$.

Without loss of generality, sum of HCRBs, ECRBs, and MSEs for the estimation of timing offsets or channels from all the nodes are evaluated in each simulation run. In all simulations, the network is assumed to be equipped with $K = 4$ nodes. The oversampling factor is set to $Q = 2$ and a RRC filter with a roll-off factor of 0.22 is employed. The normalized timing offsets from nodes to destination, τ_k, $\forall\ k$, are uniformly distributed over the whole symbol period duration, i.e., $\tau_k \sim \mathcal{U}(-0.5, 0.5)$. The channels, h_k, $\forall\ k$, are modeled as independent and identically distributed complex Gaussian random variables with $\mathcal{CN}(0, \sigma_h^2 = 1)$.

4.5.1 Estimation Performance of the different TSs

Fig. 4.7 plots the HCRB and ECRB for MTOs estimation versus SNR for the TSs in Table 4.1. Since the HCRB for MTO estimation is not affected by violating C2 and C3 (see Section 4.3.1), the results for TSs Violate-C2 and Violate-C3 are not plotted. Fig. 4.7 demonstrates that the HCRB and ECRB for MTO estimation are the lowest for the Proposed-1 and Proposed-2 TSs and worst for the Violate-C1 TS.

Fig. 4.8 plots the HCRB and ECRB for the multiple channel estimation versus SNR for the Proposed-1, Proposed-2, Violate-C2, and Violate-C3 TSs. The results for the Violate-C1 TS is not plotted since the HCRB for the multiple channel estimation is only affected by violating C2 and C3 (see Section 4.3.2). It can be observed from Fig. 4.8 that the HCRB and ECRB for the multiple channel estimation are the lowest for the TSs, Proposed-1 and Proposed-2, and worst for the TSs Violate-C2 and Violate-C3. Further, numerical results show that the HCRB and ECRB for the Proposed-1 and Proposed-2 TSs are very close. Thus, to avoid repetition, only a single curve is plotted for both the HCRB and ECRB of the Proposed-1 and Proposed-2 TSs. Finally, as mentioned in Section 4.2.2, Fig. 4.7 and Fig. 4.8 show that the TSs that minimize the HCRB also minimize the ECRB.

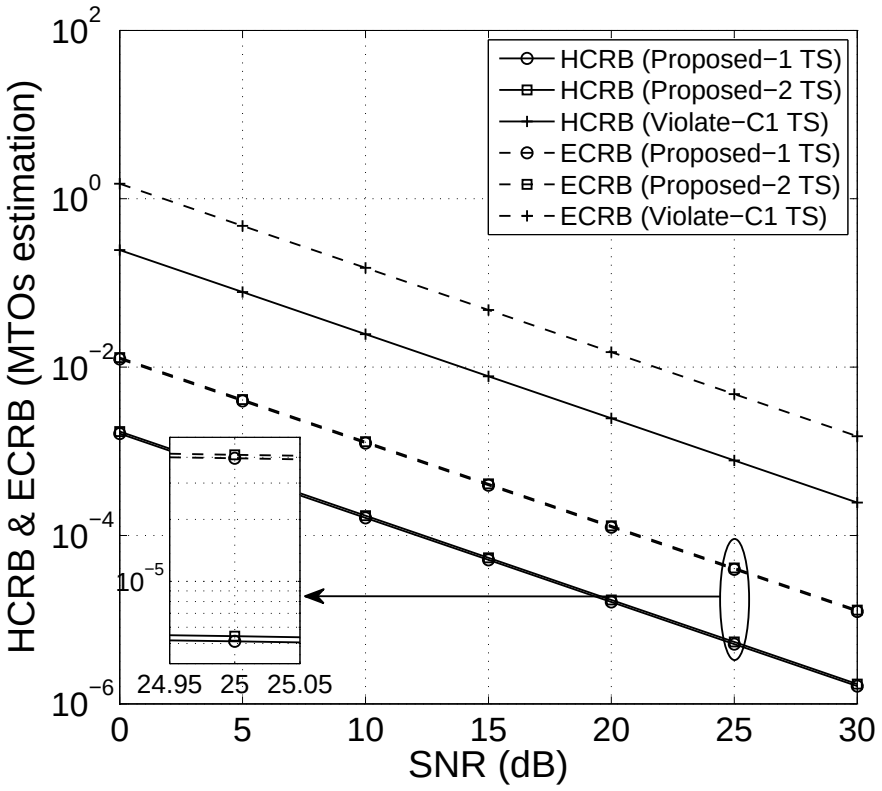


Figure 4.7: HCRB and ECRB versus SNR for the estimation of $\boldsymbol{\tau}$ for the TSs given in Table 4.1.

Fig. 4.9 plots the MAP estimator's MSE for MTOs estimation versus SNR for the Proposed-1 and Violate-C1 TSs. For clarity, the results for the Proposed-2 TS are not plotted since they are very close to that of the Proposed-1 TS. It can be observed from Fig. 4.9 that the MAP estimator's MSE is close to the ECRB over a wide range of SNR values for both the Proposed-1 and Violate-C1 TSs. Fig. 4.9 also shows that for the Violate-C1 TS, the MSE of the MAP estimator is lower than the ECRB at low SNR since the MAP estimator's estimation range is limited to $(-0.5, 0.5)$, given that the timing offset values, τ_k, are assumed to be $\tau_k \in (-0.5, 0.5)$, $\forall k$. However, in its inherent structure, ECRB does not take the range of the possible timing offset values into account and grows without bound as the SNR decreases (for more information, refer to [25] and [27]).

Fig. 4.10 plots the MAP estimator's MSE for the multiple channel estimation versus SNR for the Proposed-1, Violate-C2, and Violate-C3 TSs. It can be observed from Fig. 4.10 that the MSE performance of the MAP estimator is close to the ECRB at moderate-to-high SNR values for all TSs. Similar to Fig. 4.9, the MAP estimator's MSE and the ECRB for the Proposed-2 TS are not plotted since the results are similar to that of Proposed-1 TS. Moreover, since the MAP estimator

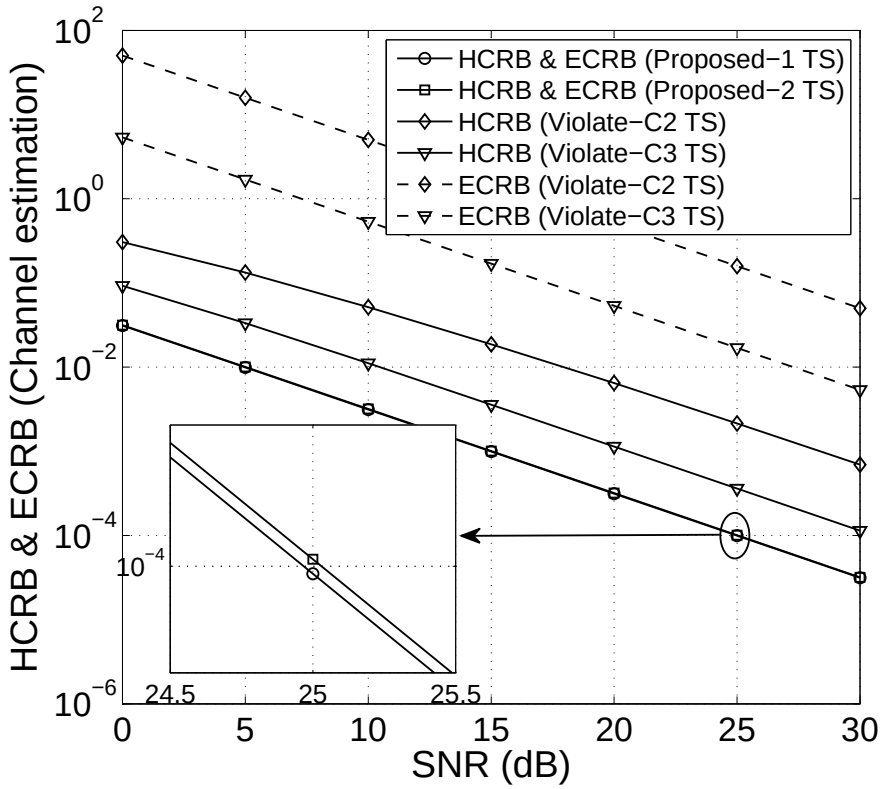


Figure 4.8: HCRB and ECRB versus SNR for the estimation of $\mathbf{h}$ for the TSs given in Table 4.1.

takes the prior information on the distribution of channels and the range of timing offset values into account (see (4.20)) and the ECRB does not (see (4.6)), the estimator's MSE is lower than the ECRB at low SNR for the Violate-C2 and Violate-C3 TSs.

From Figs. 4.7–4.10 it can be concluded that the Proposed-1 and Proposed-2 TSs can significantly improve estimation accuracy since they result in the lowest HCRB, ECRB, and MAP estimation MSE. These results also demonstrate the large performance gain in terms of the estimation accuracy for the proposed TSs over non-optimal TSs that violate the proposed conditions, which numerically validates the proposed conditions for the design of optimal TSs in Section 4.3.3.

4.5.2 System BER performance of the different TSs

Fig. 4.11 shows the end-to-end BER performance of a DF 4-relay cooperative network versus SNR for all the TSs given in Table 4.1. The proposed MAP estimator is applied for the joint estimation of MTOs and multiple channels during the training period. 256-quadrature amplitude modulation (256-QAM) is employed for

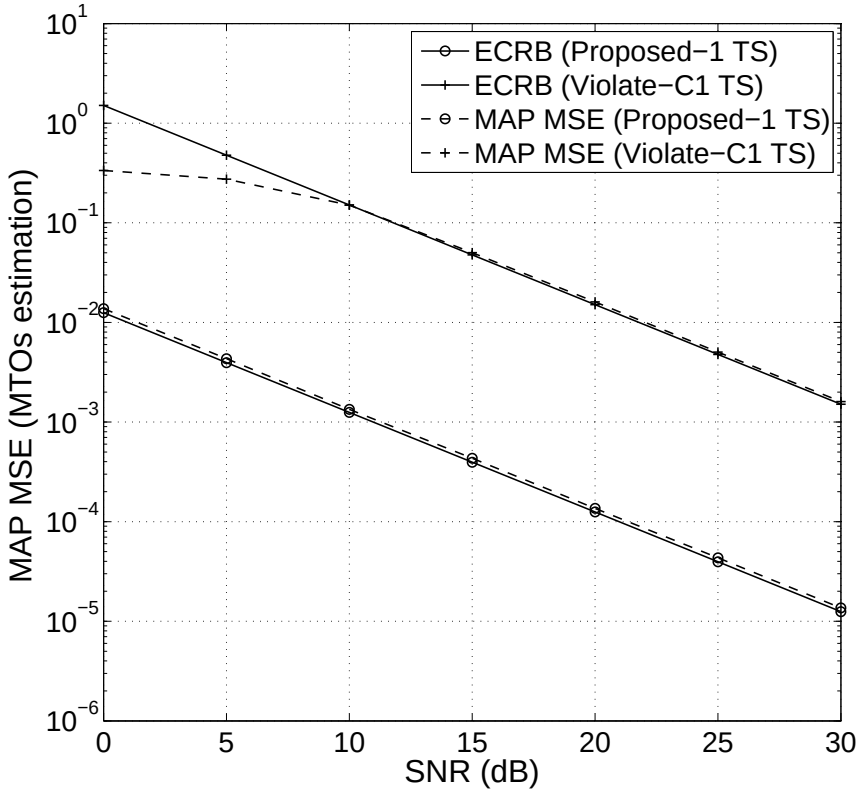


Figure 4.9: MSE of MAP estimator and ECRB versus SNR for the estimation of $\boldsymbol{\tau}$ for the TSs given in Table 4.1.

data transmission and length of the source data vector is set to 576 symbols (frame length $= 576+64 = 640$ symbols) during the data transmission period, resulting in a synchronization overhead of 10%. Distributed space time block codes are applied at all relays to exploit the spatial diversity [51, 52]. To decode the source signal $\mathbf{s}$, a minimum mean-square error linear receiver given by

$$\hat{\mathbf{s}} = (\boldsymbol{\Lambda}^H \boldsymbol{\Lambda} + \sigma_w^2 \mathbf{I}_{L\times L})^{-1}\mathbf{y},$$

is employed at the destination node, where $\boldsymbol{\Lambda} \triangleq \sum_{k=1}^{K} \hat{h}_k \hat{\mathbf{G}}_k$, $\hat{\mathbf{G}}_k = \mathbf{G}_k|_{\tau_k=\hat{\tau}_k}$. The BER performance of the overall system using different TSs is compared with the benchmark BER, which assumes perfect knowledge of MTO and channels. Fig. 4.11 shows that for the Proposed-1 and Proposed-2 TSs, the BER performance of the overall system is very close to the benchmark BER plot over a wide range of SNR values, i.e., performance gap of only 0.4 dB lies between the benchmark and proposed system's BER at moderate-to-high SNR values. Fig. 4.11 also compares the BER performance of the proposed TSs with the TSs proposed in [12]. Fig. 4.11

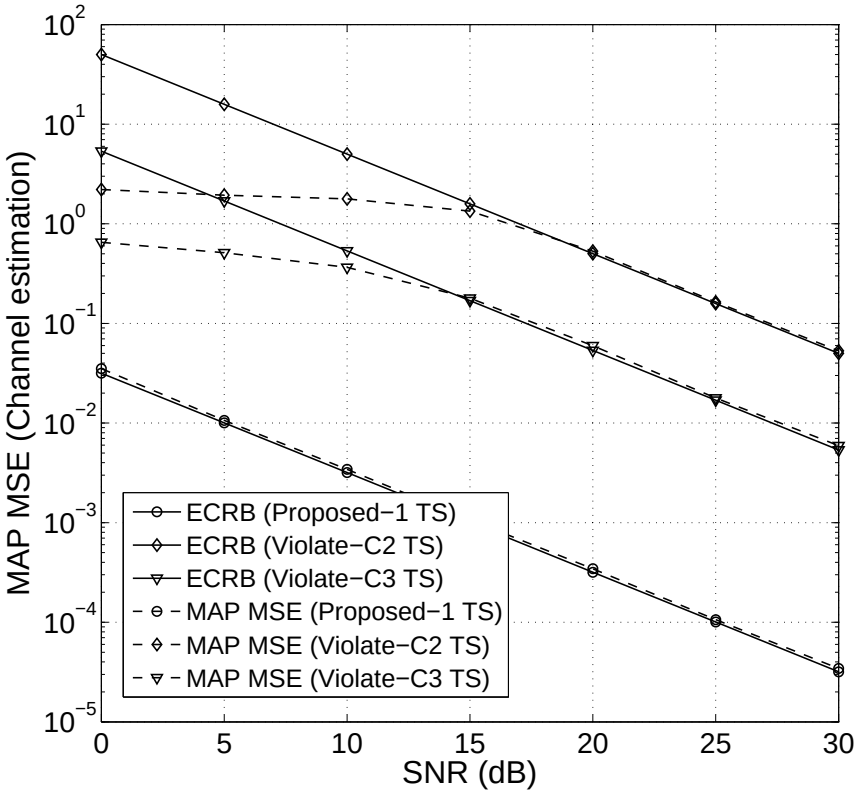


Figure 4.10: MSE of MAP estimator and ECRB versus SNR for the estimation of $\mathbf{h}$ for the TSs given in Table 4.1.

shows that our proposed TSs significantly outperform the TS in [12]. Similarly, for the Violate-C1, Violate-C2, and Violate-C3 TSs, the BER results show poor performance. Specifically, for the TSs, Violate-C1 and Violate-C2, receiver almost fails to decode the received signal at the destination for SNR < 30 dB. It has been found through simulations that for $K = 2$ users, satisfying C1 is more critical than satisfying C2 and C3, while for $K > 2$ users, satisfying C2 is more crucial than satisfying C1 and C3.[3] Fig. 4.11 demonstrates that the large BER performance gain can be achieved by employing the proposed TSs compared to non-optimal TSs that violate the proposed conditions C1-C3.

Finally, we have observed through simulations that increasing the TS length for a given fixed frame length, improves the BER performance of the overall system and the results asymptotically converge to the BER with perfect synchronization. Thus, the TS length needs to be carefully selected by system designers to achieve the desired system performance at a specific SNR.

[3]This is observed through simulations but the results for $K = 2$ and 3 users are not included here due to space limitations.

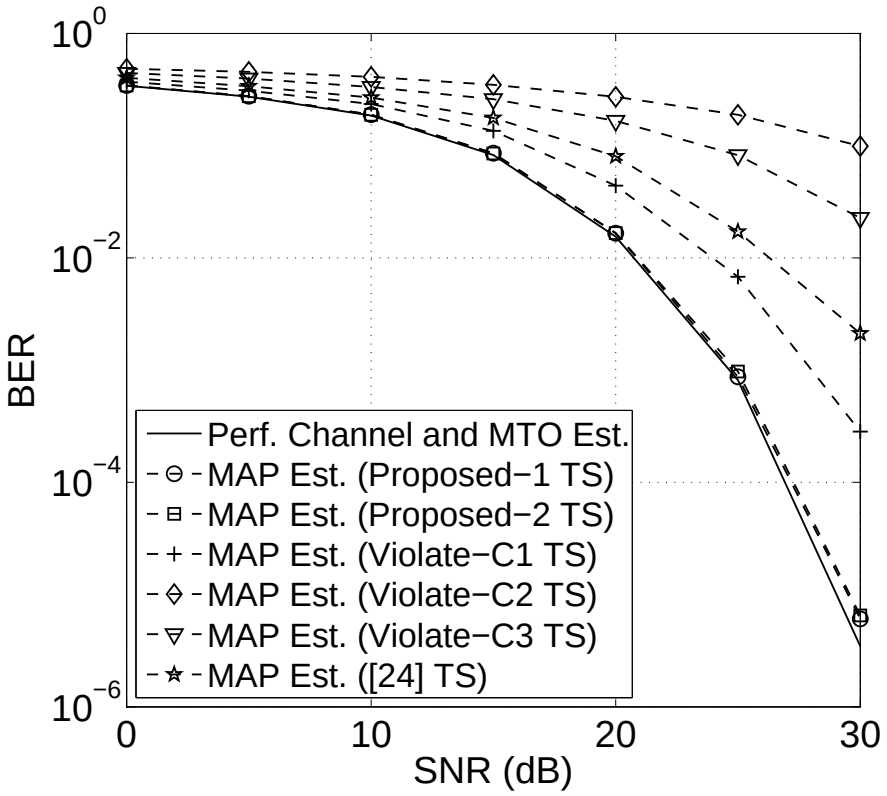


Figure 4.11: BER performance of overall system versus SNR for the TSs given in Table 4.1 and existing TS in [24], with TS length = 64, data transmission length = 576, 256-QAM modulation, and synchronization overhead = 10%.

4.6 Conclusions

In this chapter, the optimal TS design for the efficient and joint estimation of MTOs and multiple channels in distributed multi-user and multi-relay cooperative networks was addressed. The HCRB, ECRB, and MAP estimator for the estimation of parameters of interest are derived. Next, using these results, guidelines for the design of TSs that minimize the HCRB for joint estimation of timing offsets and channels are proposed. It has been observed through numerical simulations that the ECRB serves as a tighter lower bound for the MAP estimator's MSE at moderate-to-high SNR. Our proposed training guidelines show that the optimal TSs that jointly minimize the HCRB of MTO and multiple channel estimation satisfy three conditions: C1. the optimal TSs from the different nodes exhibit π radian phase shift every symbol period, C2. they are mutually orthogonal, and C3. they are orthogonal to $\pm$ T-shifted TSs from every other node. Numerical results show that the proposed optimal TS design conditions not only lower the HCRB, but also lower the tighter bound, ECRB and the MSE of the derived MAP

estimator. Moreover, by applying the proposed guidelines, two TSs are proposed and the estimation MSE and BER performance of the proposed TSs are compared against non-optimal TSs. Simulation results demonstrated large performance gain in terms of estimation accuracy and end-to-end BER performance when applying the proposed TSs compared to other TS choices.

The proposed TSs, which are summarized in Table 4.1, can also be used for other network topologies, e.g., multi-hop systems and star networks. In multi-hop systems, the proposed TSs can be applied to estimate MTOs and channel parameters in each hop. In star networks, there exist point-to-point communication links between all nodes and the proposed TS with $K = 1$ can be applied to estimate the timing offset and channel parameters corresponding to each link.

Chapter 5

Non-Data Aided Synchronization

This chapter seeks to focus on blind synchronization and channel estimation in DF cooperative communication systems, where no training signals are used. Note that Chapters 2-4 address synchronization for DA cooperative systems and assume the transmission of training sequences to assist the joint estimation of multiple channels, MTOs, and MCFOs. Chapter 5 is organized as follows:

Section 5.1 describes the overall system model and the assumptions for achieving joint blind synchronization and channel estimation. Section 5.2 details the system model for blind synchronization and channel estimation in broadcasting phase. For the broadcasting phase, the synchronization problem is identical to the one in single-input-single-output (SISO) systems and existing techniques in the literature can be used. Section 5.3 explains the system model for blind source separation and synchronization in the relaying phase. For the relaying phase, we propose to exploit blind source separation at the destination to convert the difficult problem of jointly estimating the multiple synchronization parameters in the relaying phase into more tractable sub-problems of estimating many individual timing and carrier offsets for the independent relay. This allows timing and carrier offsets to be estimated using the same algorithms as in the broadcasting phase. Next, we also propose a criteria for best relay selection at the destination. Section 5.4 presents numerical and simulation results that study the MSE and BER performances of the proposed blind receivers. Simulation results demonstrate the excellent end-to-end BER performance of the proposed blind scheme with relay selection, which is shown to achieve the full diversity order at the destination. Finally, Section 5.5 concludes the chapter and summarizes its key findings.

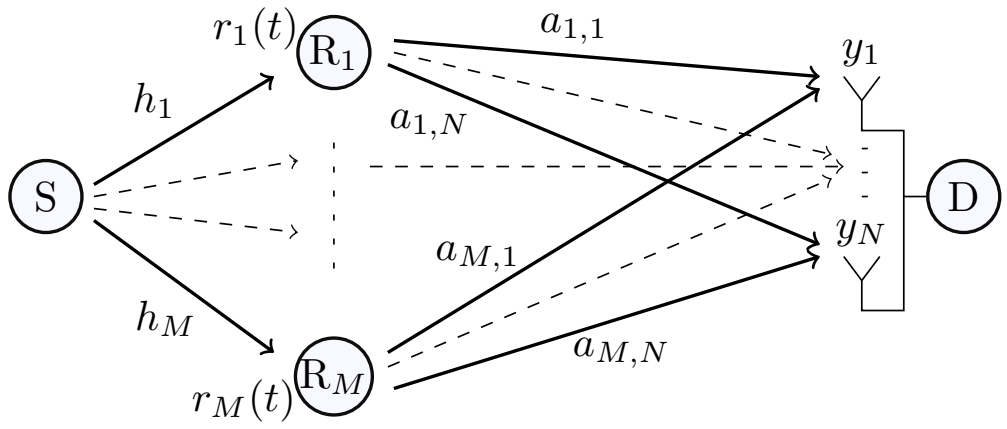


Figure 5.1: System model for the blind cooperative communication.

5.1 System Model

We consider a two hop decode-and-forward cooperative communication system with one source node, one destination node and M relay nodes. All nodes operate in half-duplex mode. The source and the relay nodes are equipped with a single antenna while the destination node is equipped with an array of N antenna elements as shown in Fig. 5.1. The channel coefficients for the source to relay and relay to destination links are denoted by h_m and $a_{m,j}$ respectively, where $m = 1, 2, \ldots, M$ is the relay index and $j = 1, 2, \ldots, N$ is the antenna index. These coefficients are modeled as zero-mean complex Gaussian random variables with unit variance, leading to a Rayleigh fading channel model.

Communication between the source node and the destination node takes place in two phases. In the *broadcasting phase*, the source node broadcasts a frame of D modulated symbols to the relay nodes. Each relay node operates independently and blindly estimates its channel, timing and carrier offsets before decoding the incoming information. In the *relaying phase*, each relay interleaves and then transmits the decoded data to the destination. The multiple antennas at the destination node help to achieve Blind Source Separation (BSS) of multiple relay signals (cf. Section 5.3). After BSS, the best relay is selected (cf. Section 5.3.2.2) and its timing and carrier offsets are blindly estimated. Finally the synchronized best relay signal is de-interleaved and the transmitted information is decoded. Note that training sequences are not required for synchronization in any of the communication phases. Throughout this chapter, the following set of system design assumptions is considered:

A1. We assume quasi-static and flat fading channel, i.e., all the fading coefficients are constant for a frame of transmitted bits but are independent and identically

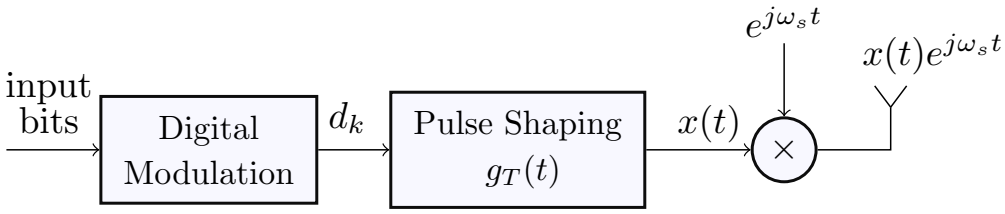


Figure 5.2: Source Transmitter for blind cooperative communication.

(i.i.d.) distributed from frame to frame.

A2. Over a frame, the timing and frequency offsets, as shown in Fig. 3.1, are modeled as deterministic but *unknown* parameters.

5.2 Blind Synchronization in the Broadcasting Phase

The synchronization problem in the broadcasting phase is to estimate the channel coefficient h_m, timing offset $\tau_{s,m}$ and carrier offset $f_{s,m}$ at each relay without pilots or training sequences, where subscript s denotes the source and m is the relay index. Note that since all the relays act independently, the synchronization problem in the broadcasting phase is the same as in Single Input Single Output (SISO) systems, for which blind solutions already exist in the literature. See, for example, [96, 129, 130] and references therein.

Let $x(t)$ be the pulse modulated baseband signal which is transmitted by the source and is given by

$$x(t) = \sum_{k=0}^{D-1} d(k)\, g_T(t - kT), \tag{5.1}$$

where $d(k)$ denotes the complex valued modulation symbol k taken from an M-ary PSK (Phase Shift Keying) constellation, $g_T(t)$ is the transmitter pulse shape, T is the symbol period and D is the transmission block length. The signal is up converted to the carrier frequency ω_s, as shown in the Fig. 5.2, before transmission.

At each relay, the received signal $r_m(t)$ is down converted by the oscillator frequency, $-\omega_s + \omega_{s,m}$, where $\omega_{s,m}$ is the analog frequency offset between the source and relay m. The baseband received signal at relay m is given by

$$\tilde{r}_{m,\tau_{s,m}}(t) = h_m \sum_{k=0}^{D-1} d(k)\, g_T(t - kT - \tau_{s,m}T) e^{j\omega_{s,m}t} + v_{r,m}(t) \tag{5.2}$$

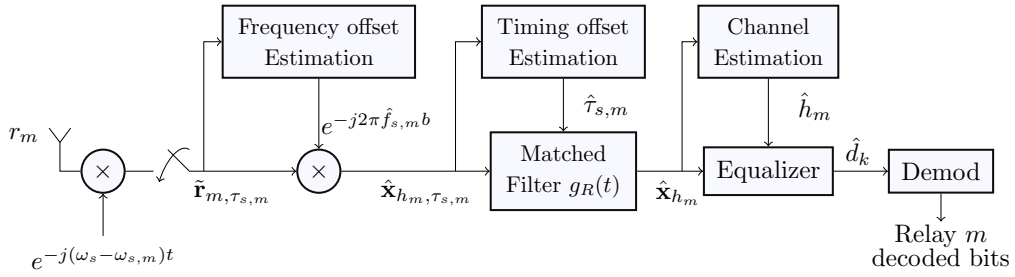


Figure 5.3: Relay m receiver.

where ' $\tilde{\cdot}$ ' over $r_{m,\tau_{s,m}}(t)$ indicates the received signal with frequency offset corruption, h_m is the complex channel coefficient from source to relay m, $v_{r,m}(t) = n_{r,m}(t)e^{-j(\omega_s-\omega_{s,m})t}$ is the noise term for relay m, $n_{r,m}(t)$ is the zero mean additive white Gaussian noise (AWGN), $\tau_{s,m}$, normalized by the symbol duration T, is the fractional unknown timing offset, $|\tau_{s,m}| \leq 1/2$, between the source and relay m.

After sampling, the received signal at each relay can be written as

$$\tilde{r}_{m,\tau_{s,m}}(bT_s) = h_m \sum_{k=0}^{D-1} d(k)\, g_T(bT_s - kT - \tau_{s,m}T)e^{j2\pi f_{s,m}b} + v_{r,m}(bT_s) \qquad (5.3)$$

where b is the sampling index, $\omega_{s,m}t = \omega_{s,m}bT_s = 2\pi(F_{s,m}/F_s)b$, $F_{s,m}$ is the frequency offset in Hz and $f_{s,m} = F_{s,m}/F_s$ is the digital frequency offset in cycles/sample between source and relay m. Thus, the blind timing and carrier synchronization problem in the broadcasting phase can be expressed as determining $\tau_{s,m}$ and $f_{s,m}$ at each relay, given the received signal, $\tilde{r}_{m,\tau_{s,m}}(t)$ at relay.

5.2.1 Proposed Relay Receiver

The block diagram of the proposed blind relay receiver is shown in Fig. 5.3. We propose to do frequency offset estimation first since carrier offset is a very sensitive parameter and must be estimated very accurately. Later, frequency offset correction is followed by timing offset estimation, matched filtering, channel estimation, and equalization, as explained in the following subsections.

5.2.1.1 Frequency Offset Estimation and Correction

Many well known techniques are available in the literature for blind or Non-Data Aided (NDA) estimation of a single carrier frequency offset [118, 131–134]. These methods can be classified into two main types: (i) algorithms that operate in feedback mode and employ Automatic Frequency Control (AFC), which has the purpose of tracking the frequency offset in a closed loop system [131, 132] and (ii)

algorithms which operate in feed forward mode and are based on open loop frequency estimation [118, 133, 134]. Closed-loop schemes are more suitable for continuous mode transmission. However, in applications, where data are transmitted in short bursts or frames, or in applications where a fast reacquisition after a deep fade is required, the acquisition of an AFC loop may possibly last too long [135]. Therefore, open-loop schemes are more appealing for frame based transmission schemes because of their short estimation times [136]. In [134], a thorough analysis of the statistical properties of open loop Non-linear Least Squares (NLS) carrier synchronizers is provided and a blind carrier offset estimator with improved performance is proposed. It is important to note that this algorithm can work even in the presence of an unknown timing offset.

We choose to estimate the frequency offset using non-linear least square approach by exploiting the fourth-order conjugate cyclostationary statistics of the received signal [134]. The frequency offset estimate $\hat{f}_{s,m}$ is given by

$$\hat{f}_{s,m} = \frac{1}{Q} \arg \max_{|\dot{\alpha}|<1/(2Q)} J(\dot{\alpha}) \tag{5.4}$$

where Q is the oversampling factor, $\dot{\alpha}$ is a trial value of the frequency and objective function $J(\cdot)$ is given by

$$J(\dot{\alpha}) = \sum_{j=0}^{Q-1} \left| \frac{1}{R} \sum_{b=0}^{R-1} \tilde{r}^4_{m,\tau_{s,m}}(b)\, e^{-j2\pi(\dot{\alpha}+q/Q)b} \right|^2 \tag{5.5}$$

where $q = 0, 1, \ldots, Q-1$, R is the length of the signal $\tilde{\mathbf{r}}_{m,\tau_{s,m}} \triangleq [\tilde{r}_{m,\tau_{s,m}}(0), \ldots, \tilde{r}_{m,\tau_{s,m}}(R-1)]^T$, which is corrupted by the carrier offset. Note that the above estimator can handle frequency offsets in the range ($|f_{s,m}| \leq 1/8QT$) where T is the symbol time. The frequency offset correction is performed as shown in Fig. 5.3 with output

$$\hat{x}_{h_m,\tau_{s,m}}(bT_s) = h_m \sum_{k=0}^{D-1} d(k)\, g_T(bT_s - kT - \tau_{s,m}T) + \hat{v}_{r,m}(bT_s) \tag{5.6}$$

which is an estimate of baseband source output signal, $x(t)$, corrupted by timing offset and the channel, $\hat{v}_{r,m}(bT_s) = v_{r,m}(bT_s)e^{-j2\pi\hat{f}_{s,m}b}$.

5.2.1.2 Timing Offset Estimation and Matched Filtering

Blind or Non-Data Aided (NDA) estimation of a single unknown timing offset is also a well established area of research [137–139]. The proposed solutions can be

classified in two categories: (i) feedback algorithms which derive an estimate of the timing error and feed the corrective signal back to an interpolator [138, 139] and (ii) feed forward algorithms in which the timing offset estimate is derived from the received signal before it is corrected in the interpolator [139]. In frame based transmission systems, where fast timing recovery is needed, feed forward recovery schemes are more appealing because of faster acquisition time [139].

The next step is to estimate the timing offset, for which any blind SISO timing estimation algorithm can be used. In this work, we choose the blind feedforward timing offset estimator using Square-Law Nonlinearity [139]. The timing offset estimate $\hat{\tau}$ is given by

$$\hat{\tau}_{s,m} = E_r \left\{ -\frac{1}{2\pi} \arg \left(\sum_{\ell=rQL}^{(r+1)LQ-1} |\hat{x}_{h_m,\tau_{s,m}}(\ell)|^2 \, e^{-j2\pi\ell/Q} \right) \right\} \tag{5.7}$$

where $E\{\cdot\}$ denotes expectation operator and r corresponds to different sections of length QL of the squared input signal $\hat{\mathbf{x}}_{h_m,\tau_{s,m}} \triangleq [\hat{x}_{h_m,\tau_{s,m}}(0), \ldots, \hat{x}_{h_m,\tau_{s,m}}(R-1)]^T$, which is corrupted by the timing offset. Note that the above estimator can handle timing offsets in the range $|\tau_{s,m}| \leq 1/2$. The timing offset estimate $\hat{\tau}_{s,m}$ is then used by the matched filter and the signal $\hat{x}_{h_m,\tau_{s,m}}(t)$ is filtered with impulse response $g_R(t + \hat{\tau}_{s,m})$, which is matched to the transmit filter $g_T(t)$. Sampled matched filter output, $\hat{\mathbf{x}}_{h_m} \triangleq [\hat{x}_{h_m}(0), \ldots, \hat{x}_{h_m}(R-1)]^T$, is given by

$$\hat{x}_{h_m}(bT_s) = h_m \sum_{k=0}^{D-1} d(k) g_T(bT_s - kT) + \hat{v}'_{r,m}(bT_s) \tag{5.8}$$

where the sequence $\hat{\mathbf{x}}_{h_m}$ at the symbol rate is an estimate of $x(t)$, which is free from offsets but corrupted by channel.

5.2.1.3 Blind Channel Estimation and Equalization

Finally, we estimate the channel using cumulant interference subspace cancelation method. The blind channel estimation is described in detail in [129] and has three main steps. The first step is to compute a series of fourth order cumulant matrices of the received signal. The second step is to evaluate Singular Value Decomposition of the concatenated cumulant matrix and select left singular vectors associated with the smallest singular values. The singular vectors are in the left null space of the concatenated cumulant matrix. Finally, applying these singular vectors on the cumulant matrix extracts the channel information by canceling the interference subspace of the channel convolution matrix [129].

Equalization is performed on the basis of the channel estimate and demodulator generates the decoded bits, as illustrated in Fig. 5.3. Note that the blind channel estimation is performed last since the algorithm assumes perfect timing and carrier synchronization for its proper operation.

5.3 Blind Synchronization in the Relaying Phase

The synchronization problem in the relaying phase is more difficult and complicated than in the broadcasting phase. This is because of the need to estimate the multiple timing and carrier offsets at the destination node. In this work, we propose to exploit blind source separation at the destination node to decouple the timing and carrier offsets from relay to relay. This converts the difficult problem of estimating multiple offsets into more tractable sub-problem of estimating many independent timing and carrier offsets.

Blind separation of multiple sources with multiple sensors or antennas at the receiver has been extensively explored in the literature [140–150]. These algorithms can be categorized into three different classes: (i) algorithms based on information theory (including the maximum likelihood approach) [140–142], (ii) algorithms based on cumulants (including correlation) [143–147] and (iii) algorithms based on constant modulus constraints [148–150]. The algorithms based on information theory include the maximum likelihood algorithms and the infomax algorithms [140–142]. The output of the separation system is achieved through minimizing the mutual information or maximization the entropy of the output signals. These algorithms have many tunable parameters and achieving good quality separation is highly dependent on the tuning of their parameters [151]. The algorithms based on constant modulus constraints depend on multi-stage signal separation and/or gradient descent optimization. Depending on the initialization of gradient descent algorithms, constant modulus algorithms sometimes fail to achieve global convergence resulting in poor quality source separation. Thus, their performance depends on initial tunable parameters such as step sizes [149]. The main advantage of cumulant based algorithms is that they can work off-the-shelf (no parameter tuning is required). The algorithms based directly on cumulants include Second-Order Statistic (SOS)-based algorithms [143,144] and Higher-Order Statistic based algorithms [145–147]. For these algorithms, the minimization of cross-cumulants or the maximization of autocumulants are usually used to achieve source separation. Among cumulant based algorithms, the Joint Approximate Diagonalization of Eigen-Matrices (JADE) by Cardoso et. al. [145] is the most renowned and com-

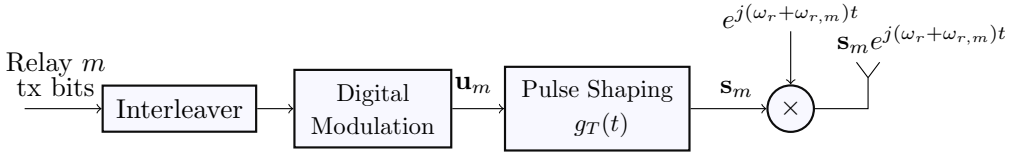


Figure 5.4: Relay m transmitter.

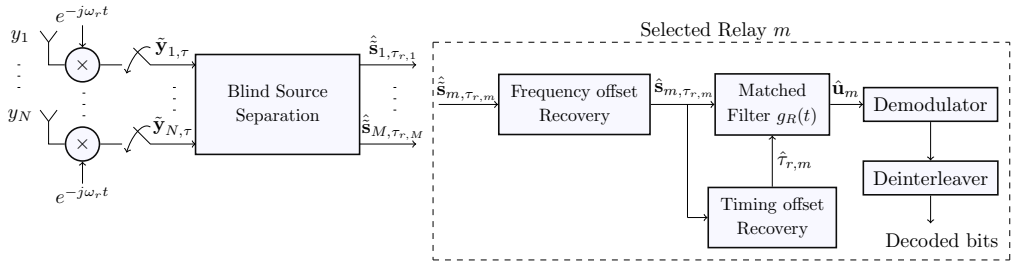


Figure 5.5: Destination Receiver.

putationally efficient algorithm [97, 147]. Thus, we implement JADE algorithm in [145] to achieve blind source separation at the destination node to decouple the timing and carrier offsets from relay to relay.

5.3.1 Proposed Relay Transmitter

The block diagram of the proposed relay m transmitter is shown in Fig. 5.4. Each relay m interleaves the decoded bits using a matrix interleaver with m rows and $(D \times K)/m$ columns, where K is the number of bits per symbol, m is the relay index and D is the frame length. After digital modulation and pulse shaping, the transmitted baseband signal of relay m is given as

$$s_m(t) = \sum_{k=0}^{D-1} u_m(k)\, g_T(t - kT), \quad \text{for} \quad m = 1, \ldots, M \tag{5.9}$$

where $u_m(k)$ denotes the complex valued data symbol k of relay m. The signal is up converted to the carrier frequency $\omega_r + \omega_{r,m}$, where ω_r is the demodulator carrier frequency at the destination and $\omega_{r,m}$ is the analog frequency offset between the relay m and the destination.

5.3.2 Proposed Destination Receiver

The block diagram of the proposed destination receiver is shown in Fig. 5.5. The frequency demodulators associated with all the antennas are fed by the same oscillator with frequency ω_r. Moreover, it is assumed that all the samplers at each

destination antenna operate at the same time instant. The sampled signal at antenna i is given as

$$\tilde{y}_{i,\tau_{r,m}}(bT_s) = \sum_{m=1}^{M} a_{m,i} \sum_{k=0}^{D-1} u_m(k)\, g_T(bT_s - kT - \tau_{r,m}T) e^{j2\pi f_{r,m} b} + v_{d,i}(bT_s) \quad (5.10)$$

where $\tilde{\mathbf{y}}_{i,\tau_{r,m}} \triangleq [\tilde{y}_{i,\tau_{r,m}}(0), \ldots, \tilde{y}_{i,\tau_{r,m}}(R-1)]^T$ is the sampled baseband received signal at the antenna i, for $i = 1, \ldots, N$, subscript '$\tau_{r,m}$' denotes that received signal is affected by the time delay between the relay m and the destination receiver, '$\tilde{\ }$' over the signal denotes that received signal is distorted by the digital frequency offset $f_{r,m}$, b is the sampling index, $\tau_{r,m}$ is the timing offset between relay m and the destination, $a_{m,i}$ is the channel coefficient between relay m and antenna i and $v_{d,i}(t)$ is the noise term at antenna i. The inner summation in (5.10) does not depend on i so we can define this factor separately as

$$\tilde{s}_{m,\tau_{r,m}}(bT_s) = \sum_{k=0}^{D-1} u_m(k)\, g_T(bT_s - kT - \tau_{r,m}T)\, e^{j2\pi f_{r,m} b} \quad (5.11)$$

where $\tilde{s}_{m,\tau_{r,m}}(t)$ is the oversampled pulse modulated signal for relay m with timing and frequency offset $\tau_{r,m}$ and $f_{r,m}$, respectively. Thus, (5.10) can be written as

$$\tilde{y}_{i,\tau_{r,m}}(b) = \sum_{m=1}^{M} a_{m,i} \tilde{s}_{m,\tau_{r,m}}(b) + v_i(b) \quad (5.12)$$

where the index b corresponds to the samples at oversampling interval T_s. Using (5.12), we can write the system model in compact matrix form as

$$\mathbf{Y} = \mathbf{A}\, \tilde{\mathbf{S}}_{\tau_{r,m}} + \mathbf{N} \quad (5.13)$$

where the matrices are defined as

$$\mathbf{Y} \triangleq [\tilde{\mathbf{y}}_{1,\tau_{r,m}}, \tilde{\mathbf{y}}_{2,\tau_{r,m}}, \ldots, \tilde{\mathbf{y}}_{N,\tau}]^T \quad (5.14)$$

$$\mathbf{A} \triangleq [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_M] \quad (5.15)$$

$$\tilde{\mathbf{S}}_{\tau_{r,m}} \triangleq [\tilde{\mathbf{s}}_{1,\tau_{r,1}}, \tilde{\mathbf{s}}_{2,\tau_{r,2}}, \ldots, \tilde{\mathbf{s}}_{M,\tau_{r,M}}]^T \quad (5.16)$$

$$\mathbf{N} \triangleq [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N]^T \quad (5.17)$$

and the vectors are defined as

$$\tilde{\mathbf{y}}_{i,\tau_{r,m}} \triangleq \big[\tilde{y}_{i,\tau_{r,m}}(0), \tilde{y}_{i,\tau_{r,m}}(T_s), \ldots, \tilde{y}_{i,\tau_{r,m}}\big((R-1)T_s\big)\big] \quad (5.18)$$

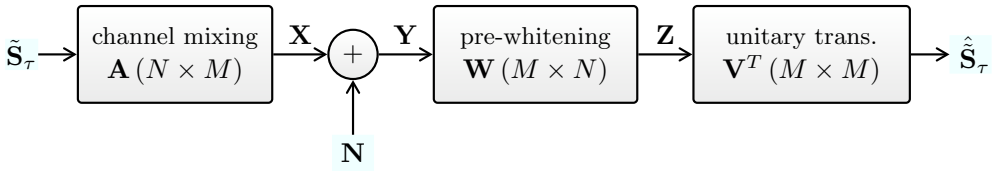


Figure 5.6: Block diagram of the blind source separation based on JADE Algorithm.

$$\mathbf{a}_m \triangleq [a_{m,1}, a_{m,2}, \ldots, a_{m,N}]^T \tag{5.19}$$

$$\tilde{\mathbf{s}}_{m,\tau_{r,m}} \triangleq \big[\tilde{s}_{m,\tau_{r,m}}(0), \tilde{s}_{m,\tau_{r,m}}(T_s), \ldots, \tilde{s}_{m,\tau_{r,m}}\big((R-1)T_s\big)\big] \tag{5.20}$$

$$\mathbf{v}_i \triangleq \big[v_i(0), v_i(T_s), \ldots, v_i\big((R-1)T_s\big)\big]^T \tag{5.21}$$

where superscript $(\cdot)^T$ denotes the transpose of a vector, $R \triangleq (D + 2L_g)\,Q$, $\tilde{\mathbf{y}}_{i,\tau_{r,m}}$ is the demodulated output vector of antenna i, $\tilde{\mathbf{s}}_{m,\tau_{r,m}}$ is the signal vector for relay m corrupted by timing and carrier offsets, $\mathbf{a}_m$ is the channel vector for relay m and $\mathbf{v}_i$ is the filtered noise vector at antenna i.

As shown in Fig. 5.5, the receiver can be divided into four separate blocks: (i) blind source separation, (ii) relay selection (ii) carrier offset recovery and (iii) timing recovery.

5.3.2.1 Blind Source Separation

The blind source separation block performs the channel estimation to decouple the source mixing and provides an estimate of each relay's data corrupted by timing and frequency offsets $\hat{\tilde{\mathbf{s}}}_{m,\tau_{r,m}}$. Note that $\hat{\tilde{\mathbf{s}}}_{m,\tau_{r,m}} \triangleq \big[\hat{\tilde{s}}_{m,\tau_{r,m}}(0), \hat{\tilde{s}}_{m,\tau_{r,m}}(T_s), \ldots, \hat{\tilde{s}}_{m,\tau_{r,m}}\big((R-1)T_s\big)\big]$ is an estimate of $\tilde{\mathbf{s}}_{m,\tau_{r,m}}$ which is defined in (5.20). For BSS, we use the JADE (Joint Approximate Diagonalization of Eigen matrices) algorithm [152]. The JADE algorithm requires some form of diversity in the received mixture of signals for its proper operation [145, 153]. The diversity order should be greater than or equal to the number of independent relay signals that need to be separated. In this work, we use N antennas at the destination, such that $N \geq M$, to provide the required diversity in the received mixture of signals for JADE algorithm.

The JADE algorithm assumes that the input signals are statistically independent and have non-gaussian distribution [154]. According to the Central Limit Theorem, the channel mixing of the independent (non-gaussian) signals $\mathbf{u}_m$, from independent relays, results in signals at each antenna, $\tilde{\mathbf{y}}_{i,\tau_{r,m}}$, having a distribution which is close to Gaussian [154]. Thus, the blind source separation objective is to apply some linear transformation to $\mathbf{Y}$ to maximize the non-gaussianity of the output $\hat{\tilde{\mathbf{S}}}_\tau$. Achieving this objective requires three main steps:

Step 1 The first step is pre-Whitening which makes the observed components $\tilde{\mathbf{y}}_{i,\tau_{r,m}}$ uncorrelated. The utility of whitening resides in the fact that it transforms the mixing matrix $\mathbf{A}$ into a unitary matrix, which eases the further estimation procedure. For whitening, we use Eigen-Value Decomposition (EVD) of the covariance matrix of received signal $\mathbf{R_Y} = E\{\mathbf{Y}\mathbf{Y}^H\} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^H$, where $\mathbf{E}$ is the unitary matrix of eigenvectors of $\mathbf{R_Y}$ and $\mathbf{\Lambda}$ is the diagonal matrix of its eigenvalues. The whitening matrix $\mathbf{W}$ is computed as

$$\mathbf{W} = \mathbf{R}_\mathbf{Y}^{-1/2} = \mathbf{E}\mathbf{\Lambda}^{-1/2}\mathbf{E}^H \tag{5.22}$$

Applying pre-whitening the whitened matrix $\mathbf{Z}$, as shown in the 5.6, is given as

$$\mathbf{Z} = \mathbf{W}\mathbf{Y} = \tilde{\mathbf{A}}\,\tilde{\mathbf{S}}_\tau + \mathbf{W}\mathbf{N} \tag{5.23}$$

where $\tilde{\mathbf{A}} = \mathbf{W}\mathbf{A}$ is a new mixing matrix. The whitened matrix $\mathbf{Z}$ accommodates the resulting uncorrelated components $\mathbf{z}_m$, where $\mathbf{Z} \triangleq [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_M]^T$ and $\mathbf{z}_m \triangleq [z_m(0), z_m(T_s), \ldots, z_m((R-1)T_s)]$. Note that the signal part of $\mathbf{Z}$ is a unitary mixture of input signal components $\tilde{\mathbf{s}}_{m,\tau_{r,m}}$.

Step 2 The next step is to determine the unitary transformation matrix $\mathbf{V}$. This is determined according to the Joint Approximate Diagonalization of Eigenmatrices (JADE) criterion. To determine $\mathbf{V}$ using the whitened received signal matrix $\mathbf{Z}$, we proceed as follows.

Let $\mathcal{Q}_{k,l} = \mathrm{Cum}(\mathbf{z}_i, \mathbf{z}_j, \mathbf{z}_k, \mathbf{z}_l)$ for $1 \leq i, j \leq M$ represents the (k,l)th cumulant matrix such that $1 \leq k, l \leq M$ and $\mathrm{Cum}(\cdot)$ denotes the cumulant operation [147]. Note that in this case there will be M^2 cumulant matrices in total. We stack all the cumulant matrices $\mathcal{Q}_{k,l}$ into a single $M^2 \times M^2$ matrix $\mathcal{Q}_{a,b}$, where $a = i + (j-1)M$ and $b = l + (k-1)M$. Taking EVD of $\mathcal{Q}_{a,b}$, we get $\mathcal{Q}_{a,b}\mathcal{M} = \mathcal{M}\mathcal{V}$, where $\mathcal{M}$ accommodates the eigenvectors of $\mathcal{Q}_{a,b}$ and $\mathcal{V}$ is the diagonal matrix containing its eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_{M^2})$. Then we unstack each column of $\mathcal{M}$ into an $M \times M$ matrix by inserting the entries columnwise, resulting in M^2 eigenmatrices $\{\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_{M^2}\}$. Let $\mathcal{N}^e = \{\lambda_d \mathcal{E}_d \,|1 \leq d \leq M\}$ be the eigen-set of M most significant eigen-pairs of the un-stacked matrices, corresponding to M most significant eigenvalues. We perform joint diagonalization of this set $\mathcal{N}^e$ under the unitary constraint $\mathbf{V}$, resulting in JADE maximization criteria $C(\mathbf{V}, \mathcal{N}^e) = \sum_{d=1}^{M} \left|\, \mathrm{diag}(\mathbf{V}^H \lambda_d \mathcal{E}_d \mathbf{V})\right|^2$. Finally $\mathbf{V}$ is obtained by performing Joint diagonalization of $M \times M$ matrices $\mathcal{R}_d = \lambda_d \mathcal{E}_d,\ d = 1, \ldots, M$ using the above maximization criteria, as detailed in [155]. Note that with correct operation, the

above procedure will converge the output towards $\mathbf{V} = \tilde{\mathbf{A}}$.

Step 3 The third and final step is to use $\mathbf{V}$ to separate the sources. The output of the blind source separation block is given by

$$\begin{aligned}\hat{\tilde{\mathbf{S}}}_\tau &= \mathbf{V}^H \mathbf{Z} \\ &= \mathbf{V}^H \tilde{\mathbf{A}} \tilde{\mathbf{S}}_\tau + \mathbf{V}^H \mathbf{W} \mathbf{N}\end{aligned} \tag{5.24}$$

It can be seen that if $\mathbf{V} = \tilde{\mathbf{A}}$, then $\mathbf{V}^H \tilde{\mathbf{A}} = \mathbf{I}_M$ and the components of $\hat{\tilde{\mathbf{S}}}_\tau$ will be free from channel mixing, though with embedded timing and frequency offsets and noise corruption. Note that $\hat{\mathbf{A}} = \mathbf{W}^{-1} \tilde{\mathbf{A}}$ is evaluated upto a permutation and scaling of its columns. This is a well known fundamental limitation common to all blind schemes [145, 153, 156]. In this work, we have resolved this ambiguity by assuming a known single symbol transmission for each source, occupying M time slots in total [156].

After BSS, the separated relay signals can be written as

$$\hat{\tilde{s}}_{m,\tau_{r,m}}(bT_s) = \sum_{k=0}^{D-1} u_m(k)\, g_T(bT_s - kT - \tau_{r,m}T)\, e^{j2\pi f_{r,m} b} + \breve{v}_{d,m}(bT_s) \tag{5.25}$$

where $\hat{\tilde{s}}_{m,\tau_{r,m}}(bT_s)$ is the baseband transmitted signal of relay m corrupted by timing and frequency offset and $\breve{v}_{d,m}(bT_s)$ is the noise element after BSS.

We assume that the destination selects the best relay after blind source separation. The relay selection procedure is discussed as follows.

5.3.2.2 Relay Selection Algorithm

For DF relaying, various decoder strategies have been proposed in the literature that have been shown to achieve the full diversity order M, where M is the number of relays, in terms of the Bit Error Rate (BER) performance [157–159]. In this regard, relay selection has emerged as a simple, powerful technique with comparatively lower implementation complexity. Various criteria for "best" relay section in relaying phase have been proposed in the literature assuming destination has single antenna. Recently, a method is proposed in [160] in which the destination chooses the best relay based on the minimum of the source-to-relay and relay-to-destination Signal to Noise Ratios. The implementation of this scheme requires the Channel State Information (CSI) of source to relay channels at the destination, which can be easily done in practice through a feedforward channel from the relay to the destination [160]. It is shown that the algorithm has superior performance

compared to existing methods and is able to extract the full diversity without requiring any feedback CSI or error detection mechanism such as cyclic redundancy check at relay nodes.

In our work, we modify the criterion in [160] for the case where destination has multiple antennas. In our scheme, after blind source separation, the destination chooses the best relay based on the following criterion

$$R_{\text{sel}} = \arg\max_{m}\{\min\{h_m, \min_{j}\{\hat{a}_{m,j}\}\}\} \tag{5.26}$$

where "sel" denotes the selected relay m. Note that the relay to destination channels $\hat{a}_{m,j}$ are blindly estimated at the destination using JADE, while information about estimated source to relay channels h_m is assumed to be available at the destination via a feedforward channel from the relay to the destination. For comparison, we also consider the case of "no relay selection" in which the destination randomly selects one of the relays.

5.3.2.3 Timing and Carrier Offset Recovery

For carrier offset and timing offset estimation on the selected relay signal, we use the same algorithms as in the broadcasting phase. The selected relay signal (say relay m) is passed to the frequency offset recovery block which estimates the carrier offset and outputs each relay's signal, $\hat{\mathbf{s}}_{m,\tau_{r,m}}$, corrupted only by the timing offset. The timing offset recovery block estimates the timing offset, $\hat{\tau}_{r,m}$, for relay m, and the signal at the input of the timing estimation block is match filtered with an impulse response $g_R(t + \hat{\tau}_{r,m})$ matched to the transmit filter $g_T(t)$. The output of matched filter block yields an estimate of interleaved data symbols at the symbol rate, $\hat{u}_m(k)$, for the selected relay m from which the transmitted data can be decoded.

5.4 Simulation Results

In this section, simulation results are presented to validate the robustness of the proposed blind cooperative communication system. We consider $M = 4$ relays and a destination equipped with $N = 5$ antennas[1]. Each transmission from the source or a relay employs Quadrature Phase Shift Keying (QPSK) modulation and transmits data in frames of length $D = 400$ symbols. The propagation channels are i.i.d. block

[1]For M relay, $N = M + 1$ antennas provides close to optimum blind source separation performance [104].

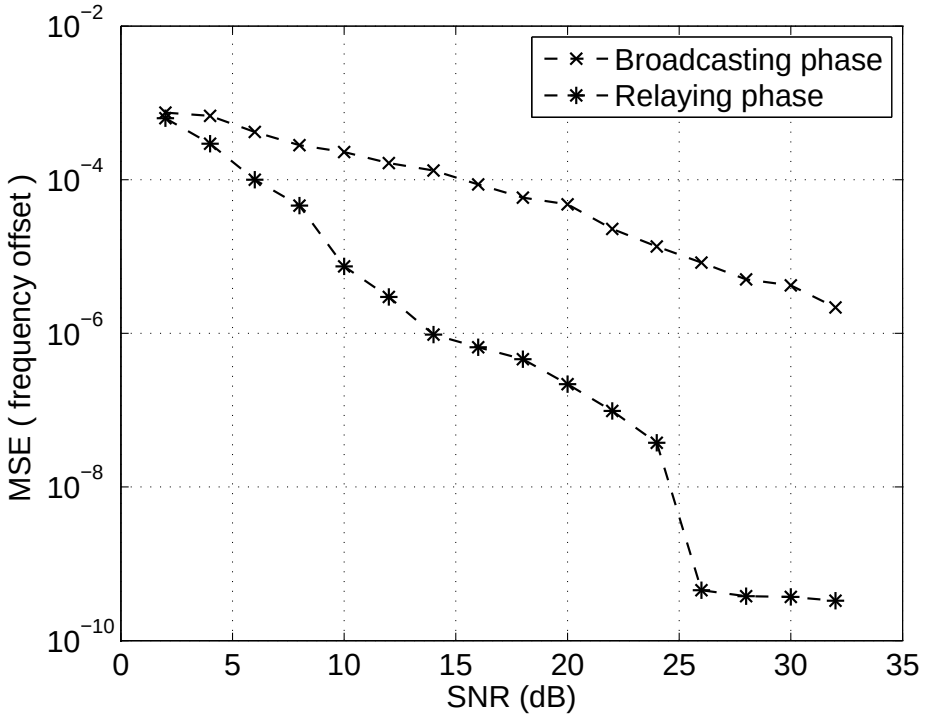


Figure 5.7: MSE of frequency offset estimation as a function of SNR (dB), with $M = 4$ Relays and $N = 5$ antennas at destination.

Rayleigh flat-fading channels. The root raised cosine filters are used for transmitter pulse shaping and receiver matched filtering, with roll-off factor, $\beta = 0.25$ and filter tap delay length set to 5 symbols. The oversampling factor used is $Q = 4$. The unknown timing offsets are assumed to be uniformly distributed as $|\tau| \leq 1/2$. The unknown digital frequency offsets are assumed to be uniformly distributed as $|f| \leq 1/32$. The timing offset estimation parameter $L = 16$ (see Section 5.2.1.2). The figures of merit used are the end-to-end (i.e., source to destination) Bit Error Rate (BER) and the Mean Square Error (MSE) of frequency and timing offsets, respectively.

5.4.1 MSE of the timing and frequency offsets

Figs. 5.7 and 5.8 show the MSE performance of frequency offset and timing offset estimation, respectively. The results are shown for both the broadcasting and relaying phase and are averaged out over the $M = 4$ relays. It can be seen from the two figures that the MSE improves with increasing Signal to Noise Ratio (SNR). Compared to the broadcasting phase, the MSE is better for the relaying phase due to the use of multiple antennas at the destination. The error floor at high SNR in the MSE of frequency offset for relaying phase is due to using a finite step size in

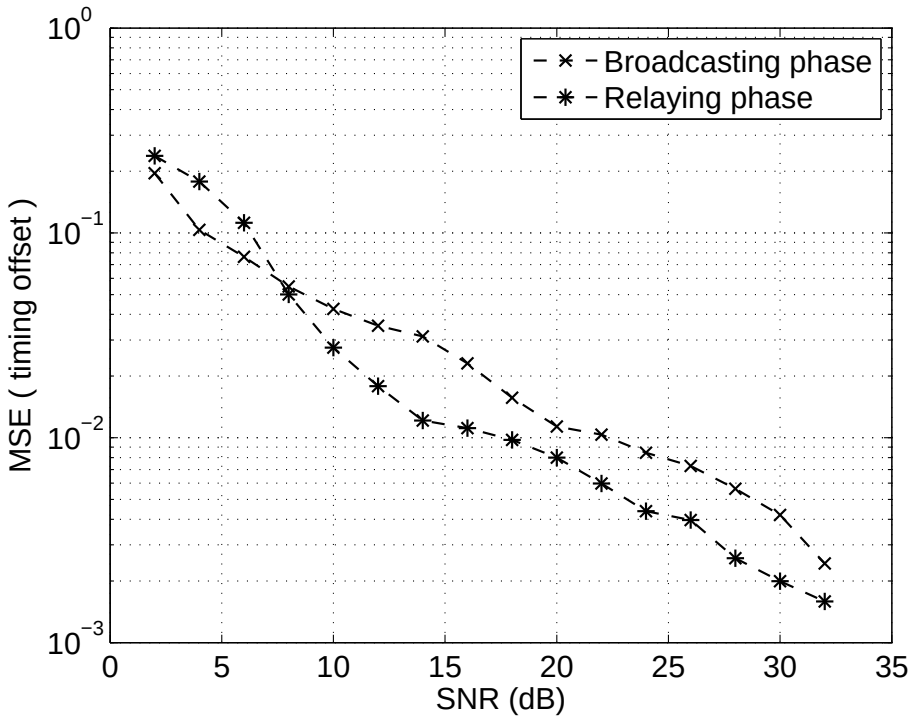


Figure 5.8: MSE of timing offset estimation as a function of SNR (dB), with $M = 4$ Relays and $N = 5$ antennas at destination.

the frequency offset estimation algorithm (see Section 5.2.1.1).

5.4.2 End-to-End BER

Fig. 5.9 shows the end-to-end (i.e., source to destination) BER of blind DF cooperative communication as a function of SNR. The following systems are considered for comparison:

1. A blind cooperative DF system operating in absence of timing and carrier offsets and endowed with perfect BSS and employing relay selection, since its performance represents a lower bound (labelled by "Lower bound").

2. A blind cooperative DF system operating in absence of timing and carrier offsets but with BSS and employing relay selection, since its performance allows effect of BSS to be studied (labelled "BSS with no offsets").

3. Proposed cooperative DF system estimating channel, timing and carrier offsets with BSS and relay selection (labelled by "Proposed - with relay selection").

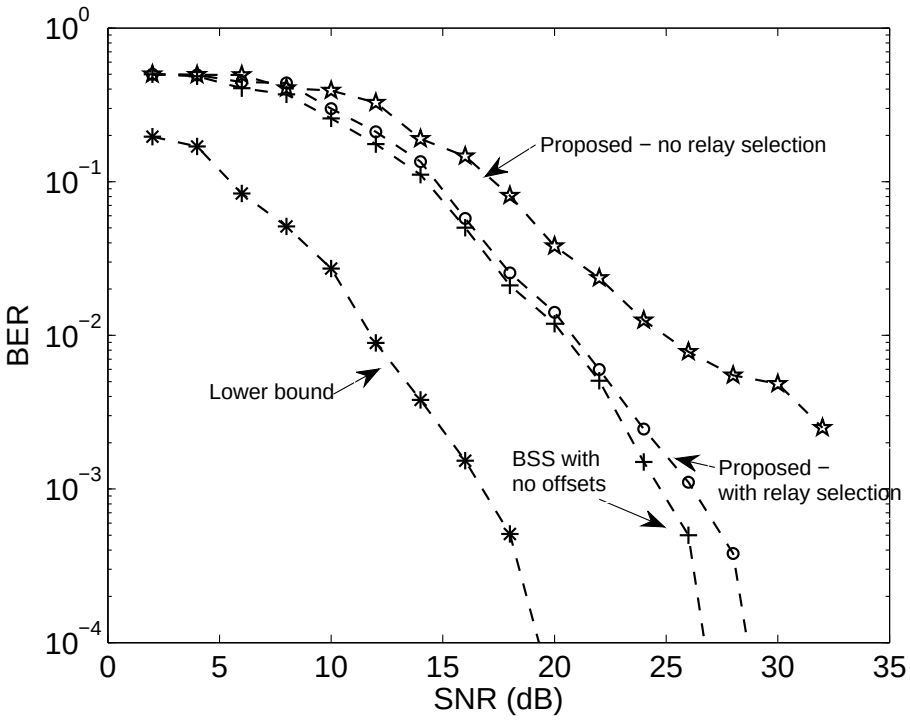


Figure 5.9: End-to-end BER of blind DF cooperative communication as a function of SNR (dB), with $M = 4$ Relays and $N = 5$ antennas at destination.

4. Proposed cooperative DF system estimating channel, timing and carrier offsets with BSS but no relay selection (labelled by "Proposed - no relay selection").

It can be seen that the slope of the "Proposed - no relay selection" BER result achieves the theoretical diversity order of 1, which confirms the correct working of the blind synchronization and channel estimation algorithms in the proposed scheme. The other three BER results which use relay selection, including the "Proposed - no relay selection" (which is the actual system performance), all have the same slope and achieve the theoretical diversity order of 4 for medium to high SNR, without any noticeable error floor in the considered SNR range. This confirms the correct working of the proposed relay selection algorithm. The "BSS with no offsets" and "Proposed - with relay selection" curves are very close to each other, indicating that the performance loss due to synchronization errors is very small. Finally, comparing the "Lower bound" and the "BSS with no offsets" curves, we can see that they have the same slope and the gap between them is due to self noise of the JADE algorithm. Similar results have also been found for 2, 3 and 5 relays.

5.5 Conclusions

In this chapter, we have proposed a scheme for blind synchronization and channel estimation in DF cooperative communication systems. Blind source separation using JADE algorithm has been proposed to decouple the timing and carrier offsets from relay to relay in the relaying phase. This allows the timing and carrier offsets to be estimated using blind SISO square-law and cyclostationary algorithms respectively in both the broadcasting and relaying phases. The channels in the broadcasting phase are estimated using blind SISO cumulant interference subspace cancellation algorithm. In addition, a relay selection algorithm has been proposed to select the best relay for decoding source information at the destination. Results have demonstrated that the proposed blind scheme works satisfactorily in terms of end-to-end BER and is able to extract the maximum diversity order due to M relays with $N = M + 1$ antennas at the destination.

Chapter 6

Conclusions and Future Research Directions

In this chapter, we summarize the general conclusions drawn from this thesis. We also outline some future research directions arising from this work.

6.1 Conclusions

The open research questions were highlighted in Section 1.2.1. The major contribution of the thesis is the addressing of those research questions.

Addressing Q1 in Section 2.1:

- We proposed a new transceiver structure at the relays and a novel receiver design at the destination which allow for the decoding of the received signal in the presence of unknown channel gains, MTOs, and MCFOs.

Addressing Q2 in Section 2.3:

- We proposed computationally efficient ECM and SAGE algorithms for the joint estimation of unknown channel gains, MTOs, and MCFOs in data-aided DF and AF cooperative communication systems.

Addressing Q3 in Section 2.4:

- We derived a maximum likelihood based detector that allows for the signal received from the multiple relays to be successfully decoded in the presence of unknown channels, MTOs, and MCFOs.

Addressing Q4 in Section 2.2:

- We derived closed-form CRLBs for the multiple parameter estimation problem for both AF and DF cooperative networks and used them to assess the performance of the proposed estimators.

Addressing Q5 in Sections 3.1, 3.3, 3.4, and 3.5:

- We proposed a new transceiver structure at the relays and a receiver design at the destination for achieving synchronization in DSTBC-AF relaying cooperative network in Section 3.1. We found that unlike the simple AF relay, described in Chapter 2, a DSTBC based AF relay has to first estimate and compensate the source to relay timing and frequency offsets before applying DSTBC. Consequently, the received signal at the destination is the function of the source to relay timing and frequency offset estimation errors. We proposed a DE based algorithm for the joint estimation of multiple impairments in Section 3.3. Next, we derived an MMSE receiver for compensating the effect of impairments and detecting the source signal at the destination in Section 3.4. Finally in Section 3.5, we observed through numerical simulations that the overall proposed transceiver design in DSTBC-AF cooperative networks achieves full spatial diversity gain for 2 and 4 relay networks and outperforms the existing cooperative strategies in the presence of unknown channels, MTOs, and MCFOs.

Addressing Q6 in Section 4.3:

- We formulated three criteria for the design of optimal TSs that minimize the derived HCRLB, ECRLB, and MSE of the MAP estimator for joint estimation of multiple channels and MTOs. Based on the proposed TS design criteria, we proposed two TSs and demonstrated that these TSs result in a lower MSE for the joint estimation of MTO and channel parameters when compared to the TSs that violate the proposed guidelines.

Addressing Q7 in Sections 5.2 and 5.3:

- We proposed a novel transceiver structure at the relays and a complete receiver design at the destination for achieving synchronization in blind DF relaying cooperative network. Blind synchronization and channel estimation in the broadcasting and the relaying phase are proposed in Section 5.2 and Section 5.3, respectively.

6.2 Future Research Directions

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

- The estimation and compensation algorithms presented in this thesis only apply to frequency flat and time non-selective channels. By employing *orthogonal frequency division multiple access (OFDMA)*, this work can be extended to incorporate doubly selective fading channels.

- To meet the demand for bandwidth-hungry multimedia and internet-related wireless services, MIMO cooperative wireless networks are an attractive solution [161]. In this context, the design of accurate synchronization techniques has been largely unaddressed. The use of multiple antennas in cooperative networks introduces new challenges and opportunities for achieving physical layer synchronization in cooperative networks. On the one hand, there are potentially more synchronization parameters and channel gains that need to be estimated making the task of synchronization more difficult. On the other hand, multiple antennas at the nodes can potentially be exploited to enhance the accuracy of estimators for the multiple synchronization parameters.

- In this thesis, we proposed an optimal training sequence design for joint estimation of multiple impairments in frequency flat fading channels. This work can be extended to provide the training sequence design guidelines for the joint estimation of multiple impairments in time varying and frequency selective fading channels.

- In Chapter 5 of this thesis, algorithms for NDA joint estimation of the synchronization and channel parameters have been studied for DF cooperative communication systems. However, the CRLBs on the estimation performance of NDA estimators have not been derived and can be the subject of future research.

Appendix A

This appendix derives the $\mathbf{F}^{[\mathrm{AF}]}$ in (2.9), which is used in Chapter 2 to obtain CLRB for the joint estimation of multiple channel parameters, MTOs, and MCFOs.

A.1 Proof of Theorem 2.1

The $(\ell, q)^{\text{th}}$ element of $4K \times 4K$ FIM in (2.9) is given by [18]

$$\left[\mathbf{F}^{[\mathrm{AF}]}(\bar{\boldsymbol{\theta}})\right]_{\ell,q} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}_{\bar{\mathbf{y}}}^H}{\partial \bar{\theta}_\ell}\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}^{-1}\frac{\partial \boldsymbol{\mu}_{\bar{\mathbf{y}}}}{\partial \bar{\theta}_q}\right\} + \mathrm{Tr}\left(\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\bar{\mathbf{y}}}}{\partial \bar{\theta}_\ell}\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\bar{\mathbf{y}}}}{\partial \bar{\theta}_q}\right). \tag{A.1}$$

where $\boldsymbol{\mu}_{\bar{\mathbf{y}}} = \bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}$ and

$$[\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}]_{\ell,i} = \mathbb{E}\Bigg\{\left(\sum_{k=1}^{R} u_k(\ell)\bar{t}_k^{[r]}(\ell T_s - \tau_k^{[rd]}T)\beta_k e^{j2\pi\ell\nu_k^{[rd]}/Q} + w(\ell)\right) \left(\sum_{m=1}^{R} u_m(i)\bar{t}_m^{[r]}(iT_s - \tau_m^{[rd]}T)\beta_k e^{j2\pi i\nu_m^{[rd]}/Q} + w(i)\right)^*\Bigg\}, \tag{A.2a}$$

$$= \sum_{k=1}^{R}\sum_{m=1}^{R}\beta_k\beta_m^*\mathbb{E}\left[e^{j2\pi\ell\nu_k^{[rd]}/Q}\bar{t}_k^{[r]}(\ell T_s - \tau_k^{[rd]}T)\left(e^{j2\pi i\nu_m^{[rd]}/Q}\bar{t}_m^{[r]}(iT_s - \tau_m^{[rd]}T)\right)^*\right] \mathbb{E}\left[u_k(\ell)u_m^*(i)\right] + \mathbb{E}\left[w(\ell)w^*(i)\right], \tag{A.2b}$$

$$= \begin{cases}\sigma_u^2\sum_{k=1}^{K}|\beta_k|^2 + \sigma_w^2, & i = \ell\,,\ k = m\\ 0, & i \neq \ell\,,\ k \neq m\end{cases} \tag{A.2c}$$

where (A.2c) follows from (A.2b) due to the assumptions of unit amplitude PSK TSs and mutual independence of the noise at the relays and destination. Accordingly the derivatives in (A.1) can be derived as

$$\frac{\partial \boldsymbol{\mu}_{\bar{\mathbf{y}}}}{\partial \Re\{\alpha_k\}} = -j\frac{\partial \boldsymbol{\mu}_{\bar{\mathbf{y}}}}{\partial \Im\{\alpha_k\}} = (\bar{\boldsymbol{\Lambda}}_k\bar{\mathbf{G}}_k\mathbf{t}_s)\odot\bar{\mathbf{t}}_k^{[r]}(\bar{\tau}_k), \tag{A.3}$$

$$\frac{\partial \boldsymbol{\mu}_{\bar{\mathbf{y}}}}{\partial \bar{\tau}_k} = (\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{R}}_k \mathbf{t}_s) \odot \bar{\mathbf{t}}_k^{[r]}(\bar{\tau}_k)\alpha_k, \qquad \text{and} \qquad \frac{\partial \boldsymbol{\mu}_{\bar{\mathbf{y}}}}{\partial \bar{\nu}_k} = j\mathbf{D}(\bar{\boldsymbol{\Lambda}}_k \bar{\mathbf{G}}_k \mathbf{t}_s) \odot \bar{\mathbf{t}}_k^{[r]}(\bar{\tau}_k)\alpha_k, \quad \text{(A.4)}$$

where $\bar{\mathbf{R}}_k \triangleq \partial \bar{\mathbf{G}}_k / \partial \bar{\tau}_k$ and $\partial \bar{t}_k^{[r]}(t)/\partial \bar{\tau}_k = 0$. Since, $\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}$ is not a function of MCFOs, MTOs, and channel gains $\boldsymbol{\alpha}$, we have

$$\frac{\partial \boldsymbol{\Sigma}_{\bar{\mathbf{y}}}}{\partial \bar{\nu}_k} = \frac{\partial \boldsymbol{\Sigma}_{\bar{\mathbf{y}}}}{\partial \bar{\tau}_k} = \frac{\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}}{\partial \Re\{\alpha_k\}} = \frac{\boldsymbol{\Sigma}_{\bar{\mathbf{y}}}}{\partial \Im\{\alpha_k\}} = 0. \quad \text{(A.5)}$$

After substituting the derivatives in (A.3), (A.4), and (A.5) into (A.1) and carrying out straightforward algebraic manipulations, the upper triangular elements of FIM, for $\ell, q = 1, 2, \cdots, K$, can be obtained as

$$\mathbf{F}_{\ell,q}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Re\{(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{G}}_q \bar{\mathbf{t}}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\}, \quad \text{(A.6)}$$

$$\mathbf{F}_{\ell,q+K}^{[\mathrm{AF}]} = -(2/\sigma_n^2)\ \Im\{(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{G}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\}, \quad \text{(A.7)}$$

$$\mathbf{F}_{\ell+K,q+K}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Re\{(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{G}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\}, \quad \text{(A.8)}$$

$$\mathbf{F}_{\ell,q+2K}^{[\mathrm{AF}]} = -(2/\sigma_n^2)\ \Im\{(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \mathbf{D}\bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{G}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\alpha_q\}, \quad \text{(A.9)}$$

$$\mathbf{F}_{\ell+K,q+2K}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Re\{(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \mathbf{D}\bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{G}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\alpha_q\}, \quad \text{(A.10)}$$

$$\mathbf{F}_{\ell+2K,q+2K}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Re\{\alpha_\ell^*(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \mathbf{D}^2 \bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{G}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\alpha_q\}, \quad \text{(A.11)}$$

$$\mathbf{F}_{\ell,q+3K}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Re\{(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{R}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\alpha_q\}, \quad \text{(A.12)}$$

$$\mathbf{F}_{\ell+K,q+3K}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Im\{(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{R}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\alpha_q\}, \quad \text{(A.13)}$$

$$\mathbf{F}_{\ell+2K,q+3K}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Im\{\alpha_\ell^*(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{G}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \mathbf{D}\bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{R}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\alpha_q\}, \quad \text{(A.14)}$$

$$\mathbf{F}_{\ell+3K,q+3K}^{[\mathrm{AF}]} = (2/\sigma_n^2)\ \Re\{\alpha_\ell^*(\bar{\mathbf{t}}_\ell^{[r]}(\bar{\tau}_l))^H \odot \left((\mathbf{t}^{[s]})^H \bar{\mathbf{R}}_\ell^H \bar{\boldsymbol{\Lambda}}_\ell^H \bar{\boldsymbol{\Lambda}}_q \bar{\mathbf{R}}_q \mathbf{t}^{[s]}\right) \odot \bar{\mathbf{t}}_q^{[r]}(\bar{\tau}_q)\alpha_q\}. \quad \text{(A.15)}$$

Note that the lower triangular elements of $\mathbf{F}^{[\mathrm{AF}]}$ can be easily obtained by simple manipulation of (A.6)-(A.15). Using (A.6)-(A.15) and lower triangular elements, we can obtain $\mathbf{F}^{\mathrm{AF}}$ given in (2.9).

Appendix B

In this appendix, we derive the FIM, $\mathbf{F}$ in (3.11), and mean-square error cost function, $\mathcal{J}_{\mathbf{s}}$ in (3.23), in Section B.1 and Section B.2, respectively.

B.1 Proof of Theorem 3.1

The elements of the $4K \times 4K$ FIM, $\mathbf{F}$ in (3.11), are given by [18]

$$[\mathbf{F}]_{m,\bar{m}} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}^{H}}{\partial \lambda_m}\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \lambda_{\bar{m}}}\right\} + \mathrm{Tr}\left\{\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \lambda_m}\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \lambda_{\bar{m}}}\right\}, \quad m,\bar{m} = 1,\ldots,4K. \tag{B.1}$$

The mean of the received signal in (3.6) during TP, $\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}$, is calculated as

$$\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}} = \mathbb{E}\left\{\boldsymbol{\Xi}\mathbf{A}\mathbf{p} + \boldsymbol{\Xi}\mathbf{B}\mathbf{v} + \mathbf{w}\right\} = \mathbb{E}\left\{\boldsymbol{\Xi}\mathbf{A}(\mathbf{x}^{[\mathrm{TP}]} \odot \mathbf{t}^{[\mathrm{rd}]}) + \boldsymbol{\Xi}\mathbf{B}(\mathbf{n} \odot \mathbf{t}^{[\mathrm{rd}]}) + \mathbf{w}\right\} \tag{B.2a}$$

$$= \boldsymbol{\Xi}\mathbf{A}(\mathbb{E}\{\mathbf{x}^{[\mathrm{TP}]}\} \odot \mathbf{t}^{[\mathrm{rd}]}) + \boldsymbol{\Xi}\mathbf{B}(\mathbb{E}\{\mathbf{n}\} \odot \mathbf{t}^{[\mathrm{rd}]}) + \mathbb{E}\{\mathbf{w}\} \tag{B.2b}$$

$$= \boldsymbol{\Xi}\mathbf{A}\mathbf{t}, \tag{B.2c}$$

where $\mathbf{x}^{[\mathrm{TP}]} \triangleq \left[\left(\mathbf{x}_1^{[\mathrm{TP}]}\right)^T, \ldots, \left(\mathbf{x}_K^{[\mathrm{TP}]}\right)^T\right]^T$ and $\mathbf{x}_k^{[\mathrm{TP}]}$ is defined below (3.2). (B.2b) follows from (B.2a) because $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\nu}^{[\mathrm{rd}]}$ and $\boldsymbol{\tau}^{[\mathrm{rd}]}$ are assumed to be deterministic parameters. Furthermore, (B.2c) follows from (B.2b) since according to the assumptions in Section 3.1.1, timing and frequency offset estimation errors from source to kth relay, $\delta_{\tau_k^{[\mathrm{sr}]}}$ and $\delta_{\nu_k^{[\mathrm{sr}]}}$, respectively, are modeled as zero-mean Gaussian random variables. Consequently, $\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\big|_{\delta_{\tau_k^{[\mathrm{sr}]}}=\delta_{\nu_k^{[\mathrm{sr}]}}=0} = \mathbf{I}_L$, $\mathbb{E}\{\mathbf{w}\} = \mathbf{0}_{LN\times 1}$, $\mathbb{E}\{\mathbf{n}_k\} = \hat{\mathbf{G}}_k^{[\mathrm{sr}]}\hat{\boldsymbol{\Lambda}}_k^{[\mathrm{sr}]}\mathbb{E}\{\mathbf{u}_k\} = \mathbf{0}_{L\times 1}$, and $\mathbb{E}\left\{\mathbf{x}_k^{[\mathrm{TP}]}\right\} = \mathbb{E}\left\{\boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\right\}\mathbf{t}^{[\mathrm{sr}]} = \mathbf{t}^{[\mathrm{sr}]}$.

Finally $\mathbf{t}$ in (B.2c) is defined as $\mathbf{t} \triangleq \Big[\underbrace{\left(\mathbf{t}^{[\mathrm{sr}]}\right)^T, \ldots, \left(\mathbf{t}^{[\mathrm{sr}]}\right)^T}_{K}\Big]^T \odot \mathbf{t}^{[\mathrm{rd}]}$. The $LN \times LN$ received signal's covariance matrix during TP, $\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}$, can be determined as

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}} &= \boldsymbol{\Xi}\mathbf{A}\mathbb{E}\left\{(\mathbf{p}-\mathbf{t})(\mathbf{p}-\mathbf{t})^H\right\}\mathbf{A}^H\boldsymbol{\Xi}^H \\ &\qquad \boldsymbol{\Xi}\mathbf{B}\left(\mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\} \odot \mathbf{t}^{[\mathrm{rd}]}\left(\mathbf{t}^{[\mathrm{rd}]}\right)^H\right)\mathbf{B}^H\boldsymbol{\Xi}^H + \mathbb{E}\left\{\mathbf{w}\mathbf{w}^H\right\} \quad \text{(B.3a)} \\ &= \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{B}^H\boldsymbol{\Xi}^H + \boldsymbol{\Sigma}_{\mathbf{w}}, \quad \text{(B.3b)}\end{aligned}$$

where $\boldsymbol{\Sigma}_{\mathbf{p}} \triangleq \mathbb{E}\left\{(\mathbf{p}-\mathbf{t})(\mathbf{p}-\mathbf{t})^H\right\}$ and $\boldsymbol{\Sigma}_{\mathbf{v}} \triangleq \mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\} \odot \mathbf{t}^{[\mathrm{rd}]}\left(\mathbf{t}^{[\mathrm{rd}]}\right)^H$ are $LK \times LK$ matrices, and $\boldsymbol{\Sigma}_{\mathbf{w}} \triangleq \mathbb{E}\left\{\mathbf{w}\mathbf{w}^H\right\}$ is an $LN \times LN$ matrix.

The derivatives of $\boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}$ in (B.1) can be calculated as

$$\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Re\{\alpha_k\}} = -j\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Im\{\alpha_k\}} = \boldsymbol{\Xi}_k\mathbf{t}_k, \quad \text{(B.4a)}$$

$$\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \tau_k^{[\mathrm{rd}]}} = \alpha_k\boldsymbol{\Lambda}_k^{[\mathrm{rd}]}\left(\mathbf{G}_k^{[\mathrm{rd}]}\right)'\mathbf{t}_k, \quad \text{(B.4b)}$$

$$\frac{\partial \boldsymbol{\mu}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \nu_k^{[\mathrm{rd}]}} = j\alpha_k\mathbf{D}\boldsymbol{\Xi}_k\mathbf{t}_k, \quad \text{(B.4c)}$$

where $\left(\mathbf{G}_k^{[\mathrm{rd}]}\right)' \triangleq \frac{\partial \mathbf{G}_k^{[\mathrm{rd}]}}{\partial \tau_k^{[\mathrm{rd}]}}$ and $\mathbf{D} \triangleq 2\pi \times \mathrm{diag}\{0, 1, 2, \ldots, LN-1\}$. Moreover, the derivative of $\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}$ in (B.1) with respect to the parameters of interest is given by

$$\boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k} \triangleq \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Re\{\alpha_k\}} = \boldsymbol{\Xi}\mathbf{U}_{\alpha_k}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{U}_{\alpha_k}^H\boldsymbol{\Xi}^H, \quad \text{(B.5a)}$$

$$\boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k+K} \triangleq \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \Im\{\alpha_k\}} = j\boldsymbol{\Xi}\mathbf{U}_{\alpha_k}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\boldsymbol{\Xi}^H - j\boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{U}_{\alpha_k}^H\boldsymbol{\Xi}^H, \quad \text{(B.5b)}$$

$$\begin{aligned}\boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k+2K} \triangleq \frac{\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \nu_k^{[\mathrm{rd}]}} &= \mathbf{U}_{\nu_k^{[\mathrm{rd}]}}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{U}_{\nu_k^{[\mathrm{rd}]}}^H \\ &\qquad + \mathbf{U}_{\nu_k^{[\mathrm{rd}]}}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{B}^H\boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{U}_{\nu_k^{[\mathrm{rd}]}}^H, \quad \text{(B.5c)}\end{aligned}$$

$$\begin{aligned}\boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},k+3K} \triangleq \frac{\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}}{\partial \tau_k^{[\mathrm{rd}]}} &= \mathbf{U}_{\tau_k^{[\mathrm{rd}]}}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{p}}\mathbf{A}^H\mathbf{U}_{\tau_k^{[\mathrm{rd}]}}^H \\ &\qquad + \mathbf{U}_{\tau_k^{[\mathrm{rd}]}}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{B}^H\boldsymbol{\Xi}^H + \boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{B}^H\mathbf{U}_{\tau_k^{[\mathrm{rd}]}}^H, \quad \text{(B.5d)}\end{aligned}$$

where $\mathbf{U}_{\alpha_k} \triangleq \mathrm{diag}\left(\mathbf{0}_{L\times L(k-1)}, \mathbf{I}_L, \mathbf{0}_{L\times L(K-k)}\right)$ is an $LK \times LK$ matrix, and $\mathbf{U}_{\nu_k^{[\mathrm{rd}]}} \triangleq \left[\mathbf{0}_{LN\times L(k-1)}, j\mathbf{D}\boldsymbol{\Xi}_k, \mathbf{0}_{LN\times L(K-k)}\right]$ and $\mathbf{U}_{\tau_k^{[\mathrm{rd}]}} \triangleq \left[\mathbf{0}_{LN\times L(k-1)}, \boldsymbol{\Lambda}_k^{[\mathrm{rd}]}\left(\mathbf{G}_k^{[\mathrm{rd}]}\right)', \mathbf{0}_{LN\times L(K-k)}\right]$ are $LN \times LK$ matrices. After substituting (B.4a), (B.5a), and (B.5b) in (B.1), the

upper triangular elements of FIM, for $k, \bar{k} = 1, \ldots, K$, can be obtained as

$$\mathbf{F}_{k,\bar{k}} = 2\Re\left\{\mathbf{t}_k^H \boldsymbol{\Xi}_k^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Xi}_{\bar{k}} \mathbf{t}_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k,\bar{k}}, \tag{B.6a}$$

$$\mathbf{F}_{k,\bar{k}+K} = -2\Im\left\{\mathbf{t}_k^H \boldsymbol{\Xi}_k^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Xi}_{\bar{k}} \mathbf{t}_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k,\bar{k}+K}, \tag{B.6b}$$

$$\mathbf{F}_{k+K,\bar{k}+K} = 2\Re\left\{\mathbf{t}_k^H \boldsymbol{\Xi}_k^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Xi}_{\bar{k}} \mathbf{t}_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k+K,\bar{k}+K}, \tag{B.6c}$$

$$\mathbf{F}_{k,\bar{k}+2K} = -2\Im\left\{\mathbf{t}_k^H \boldsymbol{\Xi}_k^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \mathbf{D}\boldsymbol{\Xi}_{\bar{k}} \mathbf{t}_{\bar{k}} \alpha_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k,\bar{k}+2K}, \tag{B.6d}$$

$$\mathbf{F}_{k+K,\bar{k}+2K} = 2\Re\left\{\mathbf{t}_k^H \boldsymbol{\Xi}_k^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \mathbf{D}\boldsymbol{\Xi}_{\bar{k}} \mathbf{t}_{\bar{k}} \alpha_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k+K,\bar{k}+2K}, \tag{B.6e}$$

$$\mathbf{F}_{k+2K,\bar{k}+2K} = 2\Re\left\{\alpha_k^* \mathbf{t}_k^H \boldsymbol{\Xi}_k^H \mathbf{D}\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \mathbf{D}\boldsymbol{\Xi}_{\bar{k}} \mathbf{t}_{\bar{k}} \alpha_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k+2K,\bar{k}+2K}, \tag{B.6f}$$

$$\mathbf{F}_{k,\bar{k}+3K} = 2\Re\left\{\mathbf{t}_k^H \boldsymbol{\Xi}_k^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{rd}]} \left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)' \mathbf{t}_{\bar{k}} \alpha_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k,\bar{k}+3K}, \tag{B.6g}$$

$$\mathbf{F}_{k+K,\bar{k}+3K} = 2\Im\left\{\mathbf{t}_k^H \boldsymbol{\Xi}_k^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{rd}]} \left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)' \mathbf{t}_{\bar{k}} \alpha_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k+K,\bar{k}+3K}, \tag{B.6h}$$

$$\mathbf{F}_{k+2K,\bar{k}+3K} = 2\Im\left\{\alpha_k^* \mathbf{t}_k^H \boldsymbol{\Xi}_k^H \mathbf{D}\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{rd}]} \left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)' \mathbf{t}_{\bar{k}} \alpha_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k+2K,\bar{k}+3K}, \tag{B.6i}$$

$$\mathbf{F}_{k+3K,\bar{k}+3K} = 2\Re\left\{\alpha_k^* \mathbf{t}_k^H \left(\left(\mathbf{G}_k^{[\mathrm{rd}]}\right)'\right)^H \left(\boldsymbol{\Lambda}_k^{[\mathrm{rd}]}\right)^H \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{rd}]} \left(\mathbf{G}_{\bar{k}}^{[\mathrm{rd}]}\right)' \mathbf{t}_{\bar{k}} \alpha_{\bar{k}}\right\} + \boldsymbol{\Pi}_{k+3K,\bar{k}+3K}, \tag{B.6j}$$

where $[\boldsymbol{\Pi}]_{m,\bar{m}}$ for $m, \bar{m} = 1, \ldots, 4K$ are the elements of $4K \times 4K$ matrix, $\boldsymbol{\Pi}$, such that $[\boldsymbol{\Pi}]_{m,\bar{m}} = \mathrm{Tr}\left\{\boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},m} \boldsymbol{\Sigma}_{\mathbf{y}^{[\mathrm{TP}]}}^{-1} \boldsymbol{\Sigma}'_{\mathbf{y}^{[\mathrm{TP}]},\bar{m}}\right\}$. Note that the lower triangular elements of $\mathbf{F}$ can be easily obtained by simple manipulation of (B.6a)-(B.6j) and are not included here for brevity. Using (B.6a)-(B.6j), $\mathbf{F}$ can be determined as given in (3.11).

B.2 Proof of Theorem 3.2

For notational brevity, we denote $\mathbf{x}^{[\mathrm{DTP}]}$ by $\mathbf{x}$ throughout this section. Substituting (3.8) in (3.22), we get

$$\mathcal{J}_{\mathbf{s}} = \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}, \mathbf{u}, \mathbf{w}, \mathbf{s}} \left\{\left\|\mathbf{Q}\left(\boldsymbol{\Xi}\mathbf{A}\boldsymbol{\Omega}\mathbf{x} + \boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Omega}\mathbf{n} + \mathbf{w}\right) - \mathbf{s}\right\|^2\right\}. \tag{B.7}$$

Since, $\mathbf{s}$, $\mathbf{w}$ and $\mathbf{u}$ are uncorrelated, (B.7) can be expanded as

$$\begin{aligned}\mathcal{J}_{\mathbf{s}} = \mathrm{Tr}\Big\{&\mathbf{Q}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\left\{\boldsymbol{\Xi}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x}}\boldsymbol{\Omega}^H\mathbf{A}^H\right\}\boldsymbol{\Xi}^H\right\}\mathbf{Q}^H + \mathbf{Q}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\left\{\boldsymbol{\Xi}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Sigma}_{\mathbf{n}}\boldsymbol{\Omega}^H\mathbf{B}^H\boldsymbol{\Xi}^H\right\}\mathbf{Q}^H\\ &- \mathbf{Q}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\left\{\boldsymbol{\Xi}\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A}\boldsymbol{\Omega}\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}\right\}\right\} - \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}}\left\{\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}^H\boldsymbol{\Omega}^H\mathbf{A}^H\right\}\boldsymbol{\Xi}^H\right\}\mathbf{Q}^H + \boldsymbol{\Phi}_{\mathbf{s}} + \mathbf{Q}\boldsymbol{\Sigma}_{\mathbf{w}}\mathbf{Q}^H\Big\},\end{aligned} \tag{B.8}$$

where $\mathbf{\Phi_x} \triangleq \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{s}}\left\{\mathbf{x}\mathbf{x}^H\right\}$, $\mathbf{\Sigma_n} \triangleq \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{u}}\left\{\mathbf{n}\mathbf{n}^H\right\}$, $\mathbf{\Phi_{x,s}} \triangleq \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{s}}$ $\left\{\mathbf{x}\mathbf{s}^H\right\}$, $\mathbf{\Phi_s} \triangleq \mathbb{E}_{\mathbf{s}}\left\{\mathbf{s}\mathbf{s}^H\right\}$ and $\mathbf{\Sigma_w} \triangleq \mathbb{E}_{\mathbf{w}}\left\{\mathbf{w}\mathbf{w}^H\right\}$. These expectations are calculated as follows.

Calculation of $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A\Omega\Phi_x\Omega}^H\mathbf{A}^H\right\}$ and $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A\Omega\Phi_{x,s}}\right\}$

Since $\boldsymbol{\delta}_{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}$, the channel matrix $\mathbf{A}$, given below (3.5), can be written as $\mathbf{A} = \hat{\mathbf{A}} + \mathbf{\Delta_A}$, where $\hat{\mathbf{A}} \triangleq \mathrm{diag}\left(\hat{\boldsymbol{\alpha}}\right) \otimes \mathbf{I}_L$ and $\mathbf{\Delta_A} \triangleq \mathrm{diag}\left(\boldsymbol{\delta}_{\boldsymbol{\alpha}}\right) \otimes \mathbf{I}_L$ are $LK \times LK$ matrices. According to the assumption in Section 3.4, $\boldsymbol{\delta}_{\boldsymbol{\alpha}}$ is distributed as $\boldsymbol{\delta}_{\boldsymbol{\alpha}} \sim \mathcal{CN}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\alpha},\boldsymbol{\alpha}}\right)$. Thus, $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A\Omega\Phi_x\Omega}^H\mathbf{A}^H\right\}$ and $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A\Omega\Phi_{x,s}}\right\}$ are evaluated as

$$\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A\Omega\Phi_x\Omega}^H\mathbf{A}^H\right\} = \hat{\mathbf{A}}\mathbf{\Omega\Phi_x\Omega}^H\hat{\mathbf{A}}^H + \left(\mathbf{C}_{\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\alpha}}} \otimes \mathbf{I}_L\right) \odot \left(\mathbf{\Omega\Phi_x\Omega}^H\right), \tag{B.9a}$$

$$\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}}\left\{\mathbf{A\Omega\Phi_{x,s}}\right\} = \hat{\mathbf{A}}\mathbf{\Omega\Phi_{x,s}}. \tag{B.9b}$$

Calculation of $\mathbf{\Phi_s}$ & $\mathbf{\Phi_x}$

Let us assume that the transmitted data symbols are uncorrelated such that $\mathbf{\Phi_s} \triangleq \mathbb{E}_{\mathbf{s}}\left\{\mathbf{s}\mathbf{s}^H\right\} = \mathbf{I}_L$. Furthermore, $\mathbf{\Phi_x}$ is an $LK \times LK$ correlation matrix, whose $L \times L$ blocks are given by

$$\left[\mathbf{\Phi_x}\right]_{(k-1)L+1:kL,(\bar{k}-1)L+1:\bar{k}L} = \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}}\left\{\mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}}\mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\mathbf{\Phi_s}\mathbf{\Delta}^H_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}\mathbf{\Delta}^H_{\mathbf{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}\right\}, \quad k,\bar{k} = 1,\ldots,K, \tag{B.10}$$

Based on the assumption in Section 3.4, $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\tau}^{[\mathrm{sr}]},\boldsymbol{\tau}^{[\mathrm{sr}]}}\right)$, $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}} \sim \mathcal{N}\left(\mathbf{0}_{K\times 1}, \mathbf{C}_{\boldsymbol{\nu}^{[\mathrm{sr}]},\boldsymbol{\nu}^{[\mathrm{sr}]}}\right)$. Thus, by applying Taylor series expansion around their means and by ignoring the second and higher order terms, the non-linear terms $\mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}}$ and $\mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}$, which are functions of $\delta_{\nu_k^{[\mathrm{sr}]}}$ and $\delta_{\tau_k^{[\mathrm{sr}]}}$, respectively, can be written as

$$\mathbf{\Delta}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} = \mathbf{\Delta}^{[0]}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}} + \left(\mathbf{\Delta}^{[0]}_{\mathbf{\Lambda}_k^{[\mathrm{sr}]}}\right)' \delta_{\nu_k^{[\mathrm{sr}]}} + \mathcal{O}_p\left(\sigma^2_{\nu_k^{[\mathrm{sr}]}}\right), \tag{B.11a}$$

$$\mathbf{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}} = \mathbf{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}} + \left(\mathbf{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}}\right)' \delta_{\tau_k^{[\mathrm{sr}]}} + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{sr}]}}\right), \tag{B.11b}$$

where $\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} \triangleq \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\Big|_{\delta_{\nu_k^{[\mathrm{sr}]}}=0}$, $\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}} \triangleq \boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}\Big|_{\delta_{\tau_k^{[\mathrm{sr}]}}=0}$, $\left(\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\right)' = \frac{\partial \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}}{\partial \delta_{\nu_k^{[\mathrm{sr}]}}}\Big|_{\delta_{\nu_k^{[\mathrm{sr}]}}=0}$, $\left(\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}}\right)' = \frac{\partial \boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}}}{\partial \delta_{\tau_k^{[\mathrm{sr}]}}}\Big|_{\delta_{\tau_k^{[\mathrm{sr}]}}=0}$. Substituting (B.11a) into (B.10), we have

$$
\begin{aligned}
[\boldsymbol{\Phi}_{\mathbf{x}}]_{(k-1)L+1:kL,(\bar{k}-1)L+1:\bar{k}L} &= \boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} \boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}} \boldsymbol{\Phi}_{\mathbf{s}} \left(\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \left(\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \\
&\quad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}} \left\{ \left(\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}}\right)' \delta_{\nu_k^{[\mathrm{sr}]}} \boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}} \boldsymbol{\Phi}_{\mathbf{s}} \left(\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \left(\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \right\} \\
&\quad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}} \left\{ \boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} \left(\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}}\right)' \delta_{\tau_k^{[\mathrm{sr}]}} \boldsymbol{\Phi}_{\mathbf{s}} \left(\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \left(\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \right\} \\
&\quad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}} \left\{ \boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} \boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}} \boldsymbol{\Phi}_{\mathbf{s}} \left(\left(\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}\right)'\right)^H \delta_{\tau_k^{[\mathrm{sr}]}} \left(\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \right\} \\
&\quad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}} \left\{ \boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} \boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}} \boldsymbol{\Phi}_{\mathbf{s}} \left(\boldsymbol{\Delta}^{[0]}_{\mathbf{G}_{\bar{k}}^{[\mathrm{sr}]}}\right)^H \left(\left(\boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_{\bar{k}}^{[\mathrm{sr}]}}\right)'\right)^H \delta_{\nu_{\bar{k}}^{[\mathrm{sr}]}} \right\} \\
&\quad + \mathcal{O}_p\left(\sigma^2_{\tau_k^{[\mathrm{sr}]}}\right) \qquad \text{(B.12a)} \\
&\approx \boldsymbol{\Phi}_{\mathbf{s}}, \qquad k,\bar{k} = 1,\ldots,K, \qquad \text{(B.12b)}
\end{aligned}
$$

where (B.12b) follows from (B.12a) because the expectations in (B.12b) are equal to $\mathbf{0}_{L\times L}$ according to the distributions of $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}$ and $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}$.

Calculation of $\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}$

Using (B.11a) and (B.12), the $L \times L$ blocks of the $LK \times LK$ matrix $\boldsymbol{\Phi}_{\mathbf{x}}$ are given by

$$
\begin{aligned}
[\boldsymbol{\Phi}_{\mathbf{x},\mathbf{s}}]_{(k-1)L+1:kL,:} &= \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}} \left\{ \boldsymbol{\Delta}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} \boldsymbol{\Delta}_{\mathbf{G}_k^{[\mathrm{sr}]}} \boldsymbol{\Phi}_{\mathbf{s}} \right\} \\
&\approx \boldsymbol{\Delta}^{[0]}_{\boldsymbol{\Lambda}_k^{[\mathrm{sr}]}} \boldsymbol{\Delta}^{[0]}_{\mathbf{G}_k^{[\mathrm{sr}]}} \boldsymbol{\Phi}_{\mathbf{s}} = \boldsymbol{\Phi}_{\mathbf{s}}, \qquad k = 1,\ldots,K. \qquad \text{(B.13)}
\end{aligned}
$$

Calculation of $\boldsymbol{\Sigma}_{\mathbf{n}}$

The $L \times L$ blocks of the $LK \times LK$ matrix $\boldsymbol{\Sigma}_{\mathbf{n}}$ can be calculated as

$$
\begin{aligned}
[\boldsymbol{\Sigma}_{\mathbf{n}}]_{(k-1)L+1:kL,(\bar{k}-1)L+1:\bar{k}L} &= \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}},\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}},\mathbf{u}} \left\{ \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]} \hat{\boldsymbol{\Lambda}}_k^{[\mathrm{sr}]} \mathbf{u}_k \mathbf{u}_{\bar{k}}^H \left(\hat{\boldsymbol{\Lambda}}_{\bar{k}}^{[\mathrm{sr}]}\right)^H \left(\hat{\boldsymbol{\mathcal{G}}}_{\bar{k}}^{[\mathrm{sr}]}\right)^H \right\} \qquad \text{(B.14a)} \\
&= \begin{cases} \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}} \left\{ \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]} \boldsymbol{\Sigma}_{\mathbf{u}_k} \left(\hat{\boldsymbol{\mathcal{G}}}_{\bar{k}}^{[\mathrm{sr}]}\right)^H \right\}, & k = \bar{k} \\ \mathbf{0}_{L\times L}, & k \neq \bar{k} \end{cases}
\end{aligned}
$$

$$k, \bar{k} = 1, \ldots, K, \quad \text{(B.14b)}$$

where $\mathbf{\Sigma}_{\mathbf{u}_k} \triangleq \sigma_{u_k}^2 \mathbf{I}_{LN}$ and (B.14b) follows from (B.14a) since the noise at different relays, $\mathbf{u}_k$ and $\mathbf{u}_{\bar{k}}$, $\forall\, k \neq \bar{k}$, are uncorrelated. Note that $\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}$ is a non-linear function of $\hat{\tau}_k^{[\mathrm{sr}]} \triangleq \tau_k^{[\mathrm{sr}]} + \delta_{\tau_k^{[\mathrm{sr}]}}$. Thus, in order to evaluate (B.14b), the Taylor series expansion of the non-linear term $\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}$ around $\tau_k^{[\mathrm{sr}]}$ is evaluated as

$$\hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]} = \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} + \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)' \delta_{\tau_k^{[\mathrm{sr}]}} + \mathcal{O}_p\left(\sigma_{\tau_k^{[\mathrm{sr}]}}^2\right), \quad \text{(B.15)}$$

where $\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \triangleq \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}\big|_{\hat{\tau}_k^{[\mathrm{sr}]}=\tau_k^{[\mathrm{sr}]}}$, $\left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)' = \frac{\partial \hat{\boldsymbol{\mathcal{G}}}_k^{[\mathrm{sr}]}}{\partial \hat{\tau}_k^{[\mathrm{sr}]}}\Big|_{\hat{\tau}_k^{[\mathrm{sr}]}=\tau_k^{[\mathrm{sr}]}}$, and since the noise at the relays are uncorrelated, $\mathbf{\Sigma}_{\mathbf{n}}$ is an $LK \times LK$ block diagonal matrix. Using (B.15) in (B.14), the $L \times L$ diagonal matrices of $\mathbf{\Sigma}_{\mathbf{n}}$ are given by

$$\begin{aligned}
[\mathbf{\Sigma}_{\mathbf{n}}]_{(k-1)L+1:kL,(k-1)L+1:kL} &= \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \mathbf{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}} \left\{ \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)' \delta_{\tau_k^{[\mathrm{sr}]}} \mathbf{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H \right\} \\
&\quad + \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}} \left\{ \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \mathbf{\Sigma}_{\mathbf{u}_k} \left(\left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)' \right)^H \delta_{\tau_k^{[\mathrm{sr}]}} \right\} + \mathcal{O}_p\left(\sigma_{\tau_k^{[\mathrm{sr}]}}^2\right) \\
&= \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \mathbf{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H + \mathcal{O}_p\left(\sigma_{\tau_k^{[\mathrm{sr}]}}^2\right) \approx \boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]} \mathbf{\Sigma}_{\mathbf{u}_k} \left(\boldsymbol{\mathcal{G}}_k^{[\mathrm{sr}]}\right)^H.
\end{aligned} \quad \text{(B.16)}$$

Calculation of $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}} \left\{ \mathbf{\Xi} \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}} \left\{ \mathbf{A} \mathbf{\Omega} \mathbf{\Phi}_{\mathbf{x}} \mathbf{\Omega}^H \mathbf{A}^H \right\} \mathbf{\Xi}^H \right\}$ and $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}} \left\{ \mathbf{\Xi} \mathbf{B} \mathbf{\Omega} \mathbf{\Sigma}_{\mathbf{n}} \mathbf{\Omega}^H \mathbf{B}^H \mathbf{\Xi}^H \right\}$

Using similar steps as in (B.11a), (B.12), (B.15), and (B.16), the expectations over $\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}$ and $\boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}$ in (B.8) can be approximated by

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{rd}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{rd}]}}} \left\{ \mathbf{\Xi} \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}} \left\{ \mathbf{A} \mathbf{\Omega} \mathbf{\Phi}_{\mathbf{x}} \mathbf{\Omega}^H \mathbf{A}^H \right\} \mathbf{\Xi}^H \right\} &= \hat{\mathbf{\Xi}} \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}} \left\{ \mathbf{A} \mathbf{\Omega} \mathbf{\Phi}_{\mathbf{x}} \mathbf{\Omega}^H \mathbf{A}^H \right\} \hat{\mathbf{\Xi}}^H + \mathcal{O}_p\left(\sigma_{\tau_k^{[\mathrm{rd}]}}^2\right) \\
&\approx \hat{\mathbf{\Xi}} \mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}} \left\{ \mathbf{A} \mathbf{\Omega} \mathbf{\Phi}_{\mathbf{x}} \mathbf{\Omega}^H \mathbf{A}^H \right\} \hat{\mathbf{\Xi}}^H, \quad \text{(B.17a)} \\
\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\tau}^{[\mathrm{sr}]}}, \boldsymbol{\delta}_{\boldsymbol{\nu}^{[\mathrm{sr}]}}} \left\{ \mathbf{\Xi} \mathbf{B} \mathbf{\Omega} \mathbf{\Sigma}_{\mathbf{n}} \mathbf{\Omega}^H \mathbf{B}^H \mathbf{\Xi}^H \right\} &= \hat{\mathbf{\Xi}} \mathbf{B} \mathbf{\Omega} \mathbf{\Sigma}_{\mathbf{n}} \mathbf{\Omega}^H \mathbf{B}^H \hat{\mathbf{\Xi}}^H + \mathcal{O}_p\left(\sigma_{\tau_k^{[\mathrm{rd}]}}^2\right) \\
&\approx \hat{\mathbf{\Xi}} \mathbf{B} \mathbf{\Omega} \mathbf{\Sigma}_{\mathbf{n}} \mathbf{\Omega}^H \mathbf{B}^H \hat{\mathbf{\Xi}}^H, \quad \text{(B.17b)}
\end{aligned}$$

where $\hat{\mathbf{\Xi}} \triangleq \mathbf{\Xi}\big|_{\boldsymbol{\nu}^{[\mathrm{rd}]}=\hat{\boldsymbol{\nu}}^{[\mathrm{rd}]}, \boldsymbol{\tau}^{[\mathrm{rd}]}=\hat{\boldsymbol{\tau}}^{[\mathrm{rd}]}}$ and $\mathbb{E}_{\boldsymbol{\delta}_{\boldsymbol{\alpha}}} \left\{ \mathbf{A} \mathbf{\Omega} \mathbf{\Phi}_{\mathbf{x}} \mathbf{\Omega}^H \mathbf{A}^H \right\}$ is given in (B.9a).

By combining the results derived in (B.9), (B.12), (B.13), (B.16), (B.17a), and (B.17b) in (B.8), the cost function in (3.23) is obtained.

Appendix C

In this appendix, a closed-form expression for the **HIM** in (4.3) is derived in Section C.1 and positive definiteness of a matrix, $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$, is proved in Section C.2.

C.1 Proof of Theorem 4.1

The **HIM** for the estimation of the parameters of interest, $\boldsymbol{\theta}$, given the observation vector, $\mathbf{y}$, is given by [19, page 12]

$$\mathbf{HIM} = \mathbb{E}_{\boldsymbol{\theta}_r|\boldsymbol{\theta}_d}[\mathbf{FIM}] + \mathbf{PIM}, \tag{C.1}$$

where **FIM** is the Fisher's information matrix for the estimation of $\boldsymbol{\theta}$ and **PIM** is the prior information matrix (PIM), such that

$$\mathbb{E}_{\boldsymbol{\theta}}[\mathbf{FIM}] = \frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \mathbf{0}_{K\times K} \\ \mathbf{0}_{K\times K} & \mathbf{0}_{K\times K} & \Re\{\mathbf{U}\} \end{bmatrix}, \tag{C.2a}$$

$$\mathbf{PIM} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathfrak{h}}^{-1} & \mathbf{0}_{2K\times K} \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times K} \end{bmatrix}. \tag{C.2b}$$

Since $\boldsymbol{\theta}_r$ and $\boldsymbol{\theta}_d$, defined below (4.2), are independent parameters, $p(\boldsymbol{\theta}_r|\boldsymbol{\theta}_d) = p(\boldsymbol{\theta}_r)$ and $\mathbb{E}_{\boldsymbol{\theta}_r|\boldsymbol{\theta}_d}[\mathbf{FIM}] = \mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$ in (C.2) [19, page 12]. The detailed derivation of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$ and **PIM** is given in the following subsections.

C.1.1 Derivation of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$

FIM for the joint estimation of MTOs and channel gains is given by [25]

$$\mathbf{FIM} = \frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}\} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}\} \\ \Re\{\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}\} & -\Im\{\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}\} & \Re\{\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}\} \end{bmatrix}, \tag{C.3}$$

where $\mathbf{D} \triangleq \text{diag}(h_1, \ldots, h_K)$ is a $K \times K$ diagonal matrix and $\boldsymbol{\Delta} \triangleq \frac{\partial \boldsymbol{\Psi}}{\partial \boldsymbol{\tau}} \triangleq [\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K]$ is a $QL \times K$ matrix. To find $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$, we have to find the expected value of all the elements of **FIM** w.r.t. $\boldsymbol{\theta}_r$. Using the fact that $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are linear operators, $\mathbb{E}[\Re\{\cdot\}] = \Re\{\mathbb{E}[\cdot]\}$ and $\mathbb{E}[\Im\{\cdot\}] = \Im\{\mathbb{E}[\cdot]\}$. Thus

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}] = \frac{2}{\sigma_w^2}\begin{bmatrix} \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & -\Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}]\} \\ \Im\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Re\{\boldsymbol{\Psi}^H\boldsymbol{\Psi}\} & \Im\{\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}]\} \\ \Re\{\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}]\} & -\Im\{\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}]\} & \Re\{\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}]\} \end{bmatrix}. \tag{C.4}$$

In order to calculate the individual elements of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$, we have to find the distribution of $\boldsymbol{\theta}_r$. The channels from the different nodes to the receiver are modeled as independent and identically distributed random variables, i.e., $h_k \sim \mathcal{CN}(0, \sigma_h^2)$, $\forall k$. Thus, $\boldsymbol{\theta}_r$ is multivariate normal distributed with mean zero and covariance $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}$, i.e., $\boldsymbol{\theta}_r \sim \mathcal{N}(\mathbf{0}_{2K\times 1}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r} = \text{diag}\Big(\underbrace{\frac{\sigma_h^2}{2}, \ldots, \frac{\sigma_h^2}{2}}_{2K}\Big)$. Thus, the PDF of $\boldsymbol{\theta}_r$, $p(\boldsymbol{\theta}_r)$, is given by

$$p(\boldsymbol{\theta}_r) = \frac{1}{(2\pi)^K \det(\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r})^{1/2}} \exp\left\{-\frac{\boldsymbol{\theta}_r^T \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1} \boldsymbol{\theta}_r}{2}\right\}. \tag{C.5}$$

Using (C.5), the submatrices of $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{FIM}]$ can be determined. Thus, $\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}]$ is given by

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}] = \begin{bmatrix} \mathbb{E}_{\boldsymbol{\theta}_r}[h_1\boldsymbol{\xi}_1^H\boldsymbol{\delta}_1] & \ldots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_K\boldsymbol{\xi}_1^H\boldsymbol{\delta}_K] \\ \vdots & \ddots & \vdots \\ \mathbb{E}_{\boldsymbol{\theta}_r}[h_1\boldsymbol{\xi}_K^H\boldsymbol{\delta}_1] & \ldots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_K\boldsymbol{\xi}_K^H\boldsymbol{\delta}_K] \end{bmatrix} = \mathbf{0}_{K\times K}. \tag{C.6}$$

The equality in (C.6) follows from the fact the channel gains, h_k, $\forall\, k$, are zero-mean random variables. It can be similarly concluded that in (C.4)

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\boldsymbol{\Psi}^H\boldsymbol{\Delta}\mathbf{D}] = \mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}] = \mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Psi}] = \mathbf{0}_{K\times K}. \tag{C.7}$$

Finally, $\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}]$ is given by

$$\mathbb{E}_{\boldsymbol{\theta}_r}[\mathbf{D}^H\boldsymbol{\Delta}^H\boldsymbol{\Delta}\mathbf{D}] = \begin{bmatrix} \mathbb{E}_{\boldsymbol{\theta}_r}[|h_1|^2\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1] & \mathbb{E}_{\boldsymbol{\theta}_r}[h_1^*\boldsymbol{\delta}_1^H\boldsymbol{\delta}_2 h_2] & \dots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_1^*\boldsymbol{\delta}_1^H\boldsymbol{\delta}_K h_K] \\ \mathbb{E}_{\boldsymbol{\theta}_r}[h_2^*\boldsymbol{\delta}_2^H\boldsymbol{\delta}_1 h_1] & \mathbb{E}_{\boldsymbol{\theta}_r}[|h_2|^2\boldsymbol{\delta}_2^H\boldsymbol{\delta}_2] & \dots & \mathbb{E}_{\boldsymbol{\theta}_r}[h_2^*\boldsymbol{\delta}_2^H\boldsymbol{\delta}_K h_K] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}_{\boldsymbol{\theta}_r}[h_K^*\boldsymbol{\delta}_K^H\boldsymbol{\delta}_1 h_1] & \mathbb{E}_{\boldsymbol{\theta}_r}[h_K^*\boldsymbol{\delta}_K^H\boldsymbol{\delta}_2 h_2] & \dots & \mathbb{E}_{\boldsymbol{\theta}_r}[|h_K|^2\boldsymbol{\delta}_K^H\boldsymbol{\delta}_K] \end{bmatrix}$$
$$= \sigma_h^2 \mathrm{diag}\left(\boldsymbol{\delta}_1^H\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_K^H\boldsymbol{\delta}_K\right), \tag{C.8}$$

where the off-diagonal elements of (C.8) are zero since the channels are uncorrelated, i.e., $\mathbb{E}_{\boldsymbol{\theta}_r}[h_{\bar{k}}^* h_k] = 0$ for $k, \bar{k} = 1, \dots, K$, and $k \neq \bar{k}$. The diagonal elements of (C.8) for $k = 1, \dots, K$ are given by

$$\mathbb{E}_{\boldsymbol{\theta}_r}[|h_k|^2\boldsymbol{\delta}_k^H\boldsymbol{\delta}_k] = \boldsymbol{\delta}_k^H\boldsymbol{\delta}_k\Big(\mathbb{E}_{\Re\{h_{\bar{k}}\}}\left[\Re\{h_{\bar{k}}\}^2\right] + \mathbb{E}_{\Im\{h_{\bar{k}}\}}\left[\Im\{h_{\bar{k}}\}^2\right]\Big) = \sigma_h^2\boldsymbol{\delta}_k^H\boldsymbol{\delta}_k. \tag{C.9}$$

Substituting (C.6), (C.7), and (C.8) into (C.4), the final result in (C.2) can be obtained.

C.1.2 Derivation of PIM

Given the fact that $\mathbf{h}$ and $\boldsymbol{\tau}$ are independent and using the definition of PIM in [19, page 12] and [126, eq. (3)], the **PIM** for the random channels and deterministic timing offsets can be written as

$$\mathbf{PIM} = \begin{bmatrix} \mathbb{E}_{\boldsymbol{\theta}_r}\left[-\Delta_{\boldsymbol{\theta}_r}^{\boldsymbol{\theta}_r}\log p(\boldsymbol{\theta}_r)\right] & \mathbf{0}_{2K\times K} \\ \mathbf{0}_{K\times 2K} & \mathbf{0}_{K\times K} \end{bmatrix}, \tag{C.10}$$

where $p(\boldsymbol{\theta}_r)$ is given in (C.5). The negative Hessian of the log likelihood function of $\boldsymbol{\theta}_r$, $-\Delta_{\boldsymbol{\theta}_r}^{\boldsymbol{\theta}_r}\log p(\boldsymbol{\theta}_r)$ is given by

$$-\Delta_{\boldsymbol{\theta}_r}^{\boldsymbol{\theta}_r}\log p(\boldsymbol{\theta}_r) = \frac{1}{2}\Delta_{\boldsymbol{\theta}_r}^{\boldsymbol{\theta}_r}\left({\boldsymbol{\theta}_r}^T\boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1}\boldsymbol{\theta}_r\right) = \boldsymbol{\Sigma}_{\boldsymbol{\theta}_r}^{-1}. \tag{C.11}$$

By using the result in (C.11) in (C.10), the final result in (C.2) follows.

C.2 Proof of Proposition 4.1

In order to show that $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ is a positive definite matrix, it is required to show that for any $K \times 1$ vector $\mathbf{x}$, $\mathbf{x}^H\boldsymbol{\Psi}^H\boldsymbol{\Psi}\mathbf{x} > 0$. Clearly

$$\mathbf{x}^H\boldsymbol{\Psi}^H\boldsymbol{\Psi}\mathbf{x} = \|\boldsymbol{\Psi}\mathbf{x}\|^2 \geq 0, \tag{C.12}$$

which demonstrates that $\mathbf{\Psi}^H\mathbf{\Psi}$ is a positive *semidefinite* matrix and its eigenvalues are nonnegative [107]. It is known that if a matrix $\mathbf{\Psi}^H\mathbf{\Psi}$ is full rank, no eigenvalue of $\mathbf{\Psi}^H\mathbf{\Psi}$ is equal to zero [107]. Since linearly independent TSs are transmitted from all the nodes (see Section 4.2.1), it can be concluded that $\mathbf{\Psi}^H\mathbf{\Psi}$ is a full rank matrix. Thus, using (C.12) and the full rank nature of $\mathbf{\Psi}^H\mathbf{\Psi}$, it can be concluded that all eigenvalues of $\mathbf{\Psi}^H\mathbf{\Psi}$ are strictly positive and $\mathbf{\Psi}^H\mathbf{\Psi}$ is a positive definite matrix.

Bibliography

[1] X. Li, C. Xing, Y-C. Wu, and S. C. Chan, "Timing estimation and resynchronization for amplify-and-forward communication systems," *IEEE Trans. Signal Processing*, vol. 58, no. 4, pp. 2218–2229, Apr. 2010.

[2] 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.

[3] 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 Processing*, vol. 60, no. 2, pp. 793–811, Feb. 2012.

[4] X. Tao, X. Xu, and Q. Cui, "An overview of cooperative communications," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 65–71, Jun. 2012.

[5] A. Nosratinia, T. E. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 74–80, Oct. 2004.

[6] Z. Sheng, K. K. Leung, and Z. Ding, "Cooperative wireless networks: from radio to network protocol designs," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 64–69, May 2011.

[7] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*, Wiley Series in Telecom. and Signal Processing, 1998.

[8] S. Barbarossa and G. Scutari, "Bio-inspired sensor network design," *IEEE Signal Processing Mag.*, vol. 24, pp. 26–35, May 2007.

[9] J. Chen, F. Jonsson, and L.-R. Zheng, "A fast and accurate phase noise measurement of free running oscillators using a single spectrum analyzer," in *Proc. IEEE NORCHIP*, 2010.

[10] O. Tipmongkolsilp, S. Zaghloul, and A. Jukan, "The evolution of cellular backhaul technologies: Current issues and future trends," vol. 13, no. 1, pp. 97–113, 2011.

[11] A. Magee, "Synchronization in next-generation mobile backhaul networks," *IEEE Commun. Mag.*, vol. 48, no. 10, pp. 110–116, Oct. 2010.

[12] H. Mehrpouyan and S. Blostein, "Estimation, training, and effect of timing offsets in distributed cooperative networks," in *Proc. IEEE GLOBECOM*, 2010.

[13] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperative diversity - part I: system description; part II: implementation and performance analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1948, Nov. 2003.

[14] J. Yindi and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 3, pp. 473–476, Apr. 2007.

[15] Y. Zheng, H. Mehrpouyan, and S. D. Blostein, "Application of phase shift in coherent multi-relay MIMO communications," in *Proc IEEE ICC*, 2009.

[16] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[17] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, NJ: Prentice Hall, 1993.

[19] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering / Tracking*, Wiley interscience, 2007.

[20] C. Andrieu, A. Doucet, S. Singh, and V. B. Tadic, "Particle methods for change detection, system identification, and control," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 423–438, Mar. 2004.

[21] O. Besson and P. Stoica, "On parameter estimation of MIMO flat-fading channels with frequency offsets," *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 602–613, Mar. 2003.

[22] T. Pham, A. Nallanathan, and Y. Liang, "Joint channel and frequency offset estimation in distributed MIMO flat-fading channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 648–656, Feb. 2008.

[23] Y. Yao and T. Ng, "Correlation-based frequency offset estimation in MIMO system," in *Proc. IEEE VTC*, 2003.

[24] Z. Lu, J. Li, L. Zhao, and J. Pang, "Iterative parameter estimation in MIMO flat-fading channels with frequency offsets," in *Proc. IEEE International Conference on Advanced Information Networking and Applications*, Apr. 2006.

[25] X. Li, Y. C. Wu, and E. Serpedin, "Timing synchronization in decode-and-forward cooperative communication systems," *IEEE Trans. Signal Processing*, vol. 57, no. 4, pp. 1444–1455, Apr. 2009.

[26] A. Kannan, T. P. Krauss, and M. D. Zoltowski, "Separation of cochannel signals under imperfect timing and carrier synchronization," *IEEE Trans. Veh. Technol.*, vol. 50, no. 1, pp. 79–96, Jan. 2001.

[27] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Particle filters for joint timing and carrier estimation: Improved resampling guidelines and weighted Bayesian Cramer-Rao bounds," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1407–1419, May 2012.

[28] 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 ICASSP*, 2011.

[29] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Particle filter for joint blind carrier frequency offset estimation and data detection," in *Proc IEEE ICSPCS*, 2010.

[30] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Performance of coarse and fine timing synchronization in OFDM receivers," in *Proc. IEEE ICFCC*, 2010.

[31] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Blind fractionally spaced equalization and timing synchronization in wireless fading channels," in *Proc. International Conference on Future Computer and Communication (ICFCC)*, May 2010.

[32] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Modified constant modulus algorithm for joint blind equalization and synchronization," in *Proc IEEE AusCTW*, 2010.

[33] Q. Jiang, K. Zhang, J. Liu, and G. Shen, "Joint carrier frequency offset and channel estimation for AF cooperative OFDM systems," *Int. Journal Wireless Personal Commun.*, vol. 55, pp. 1–27, Oct. 2010.

[34] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Achieving cooperative diversity with multiple frequency offset estimation," in *Proc. IEEE ICC*, 2011.

[35] M. T. Hossain, D. B. Smith, and S. Kandeepan, "Timing synchronization for cooperative communications with detect and forward relaying," *Springer Journal of Wireless Personal Communications*, vol. 62, no. 2, pp. 363–378, 2010.

[36] 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 Processing*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.

[37] H. Mehrpouyan, A. A. Nasir, T. Eriksson, S. D. Blostein, G. K. Karagiannidis, and T. Svensson, "Time-varying phase noise and channel estimation in MIMO systems," in *Proc IEEE SPAWC*, 2012.

[38] A. A. Nasir, H. Mehrpouyan, R. Schober, and Y. Hua, "Phase noise in MIMO systems: Bayesian Cramer-Rao bounds and soft-input estimation," *accepted for publication IEEE Trans. Signal Processing*, 2013.

[39] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify-and-forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 1907–1916, 2008.

[40] Y. Tian, X. Lei, Y. Xiao, and S. Li, "ML synchronization algorithm and estimation bounds for cooperative systems," in *Proc IEEE Pacific Asia Conference on Circuits, Communications and Systems*, 2010.

[41] Y. Tian, X. Lei, Y. Xiao, and S. Li, "SAGE based joint timing-frequency offsets and channel estimation in distributed MIMO systems," *Elsevier Journal Computer Commun.*, vol. 33, no. 17, pp. 2125–2131, Jul. 2010.

[42] L. Dai, Z. Wang, J. Wang, and Z. Yang, "Joint channel estimation and time-frequency synchronization for uplink TDS-OFDMA systems," *IEEE Trans. Consumer Electron.*, vol. 56, no. 2, pp. 494–500, May 2010.

[43] J.-Ho Lee and S.-Cheol Kim, "Time and frequency synchronization for OFDMA uplink system using the SAGE algorithm," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1176–1181, Apr. 2007.

[44] M. Morelli, "Time and frequency synchronization for the uplink of an OFDMA system," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 296–396, Feb. 2004.

[45] M.-O. Pun, M. Morelli, and C.-C. Jay Kuo, "Maximum likelihood synchronization and channel estimation for OFDMA uplink transmissions," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 726–736, Apr. 2006.

[46] H. Wang, X-G. Xia, and Q. Yin, "Computationally efficient equalization for asynchronous cooperative communications with multiple frequency offsets," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 648–655, Feb. 2009.

[47] F. Tian, X-G. Xia, and P. C. Ching, "Signal detection in a cooperative communication system with multiple CFOs by exploiting the properties of space-frequency codes," in *Proc. IEEE ICC*, 2008.

[48] D. Veronesi and D. L. Goeckel, "Multiple frequency offset compensation in cooperative wireless systems," in *Proc. IEEE GLOBECOM*, 2006.

[49] Shichuan Ma, Yaoqing Yang, and H. Sharif, "Distributed MIMO technologies in cooperative wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 78–82, May 2011.

[50] M.-K. Chang and S.-Y. Lee, "Performance analysis of cooperative communication system with hierarchical modulation over Rayleigh fading channel," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2848–2852, Jun. 2009.

[51] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[52] Y. Jing and H. Jafarkhani, "Using orthogonal and quasi-orthogonal designs in wireless relay networks," *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 4106–4118, Nov. 2007.

[53] F. Khan, Y. Chen, and M. Alouini, "Novel receivers for AF relaying with distributed STBC using cascaded and disintegrated channel estimation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1370–1379, Apr. 2012.

[54] T. Q. Duong, G. C. Alexandropoulos, H. Zepernick, and T. A. Tsiftsis, "Orthogonal space-time block codes with CSI-assisted amplify-and-forward relaying in correlated Nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 882–889, Mar. 2011.

[55] H. Mheidat, M. Uysal, and N. Al-Dhahir, "Equalization techniques for distributed space-time block codes with amplify-and-forward relaying," *IEEE Trans. Signal Processing*, vol. 55, no. 5, pp. 1839–1852, May 2007.

[56] J. Harshan and B. Sundar Rajan, "Distributed space-time block codes for two-hop wireless relay networks," *Journal of Communications*, vol. 5, no. 4, pp. 282–296, Apr. 2010.

[57] H. M. Wang and X. G. Xia, "Asynchronous cooperative communication systems: A survey on signal designs," *Sci. China Inf. Sci.*, vol. 54, no. 8, pp. 1547–1561, Apr. 2011.

[58] M. Elkashlan, P. L. Yeoh, R. H. Y. Louie, and I. B. Collings, "On the exact and asymptotic SER of receive diversity with multiple amplify-and-forward relays," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4602–4608, Nov. 2010.

[59] A. Abdaoui, S. S. Ikki, and M. H. Ahmed, "Performance analysis of MIMO cooperative relaying system based on alamouti STBC and amplify-and-forward schemes," in *Proc. IEEE ICC*, 2010.

[60] Z. Zhong, S. Zhu, and A. Nallanathan, "Delay-tolerant distributed linear convolutional space-time code with minimum memory length under frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3944–3949, Aug. 2009.

[61] M. O. Damen and A. R. Hammons, "Delay-tolerant distributed-TAST codes for cooperative diversity," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3755–3773, Oct. 2007.

[62] H. Wang, Q. Yin, and X.-G. Xia, "Full diversity space-frequency codes for frequency asynchronous cooperative relay networks with linear receivers," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 236–247, Jan. 2011.

[63] D. Wang and S. Fu, "Asynchronous cooperative communications with STBC coded single carrier block transmission," in *Proc. IEEE GLOBECOM*, 2007.

[64] Z. Li, X.-G. Xia, and M. H. Lee, "A simple orthogonal space-time coding scheme for asynchronous cooperative systems for frequency selective fading channels," *IEEE Trans. Commun.*, vol. 58, no. 8, pp. 2219–2224, Aug. 2010.

[65] K. Raghunath and A. Chockalingam, "Cooperative OFDM with amplify-and-forward relaying with timing offset," in *Proc. IEEE GLOBECOM*, 2008.

[66] F. T. Alotaibi and J. A. Chambers, "Full-rate and full-diversity extended orthogonal space-time block coding in cooperative relay networks with imperfect synchronization," in *Proc. IEEE ICASSP*, 2010.

[67] Y. Yao and X. Dong, "On the detection of distributed STBC AF cooperative OFDM signal in the presence of multiple CFOs," in *Proc. IEEE ICC*, 2010.

[68] Z. Li and X.-G. Xia, "An alamouti coded OFDM transmission for cooperative systems robust to both timing errors and frequency offsets," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1839–1844, May 2008.

[69] O-S. Shin, A. M. Chan, H. T. Kung, and V. Tarokh, "Design of an OFDM cooperative space-time diversity system," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 2203–2215, Jul. 2007.

[70] Q. Huang, M. Ghogho, J. Wei, and P. Ciblat, "Practical timing and frequency synchronization for OFDM based cooperative systems," *IEEE Trans. Signal Processing*, vol. 58, no. 7, pp. 3706–3716, Jul. 2010.

[71] H. Mheidat and M. Uysal, "Non-coherent and mismatched-coherent receivers for distributed STBCs with amplify-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 11, pp. 4060–4070, Nov. 2007.

[72] Y-C Wu and E. Serpedin, "Training sequences design for symbol timing estimation in MIMO correlated fading channels," in *Proc IEEE Global Commun. Conf. (GLOBECOM)*, 2004.

[73] M. Dong and L. Tong, "Optimal design and placement of pilot symbols for channel estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.

[74] L. Yang, X. Ma, and G. B. Giannakis, "Optimal training for MIMO fading channels with time- and frequency-selectivity," in *Proc. IEEE International Conference on Accoustics, Speech and Signal Processing (ICASSP)*, 2004.

[75] S.-A. Yang and J. Wu, "Optimal binary training sequence design for multiple-antenna systems over dispersive fading channels," *IEEE Trans. Veh. Technol.*, vol. 51, no. 5, pp. 1271–1276, Sep. 2002.

[76] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.

[77] Y. Chi, A. Gomaa, N. Al-Dhahir, and R. Calderbank, "Training signal design and tradeoffs for spectrally-efficient multi-user MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 2234–2245, 2011.

[78] E. Bjornson and Bjorn Ottersten, "A framework for training-based estimation in arbitrarily correlated rician MIMO channels with rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, Mar. 2010.

[79] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4438–4446, Sep. 2011.

[80] X. Zhou, P. Sadeghi, T. A. Lamahewa, and S. Durrani, "Design guidelines for training-based MIMO systems with feedback," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4014–4026, Oct. 2009.

[81] C. Tellambura, M. Parker, Y. J. Guo, S. Sheperd, and S. Barton, "Optimal sequence for channel estimation using discrete Fourier transform techniques," *IEEE Trans. Commun.*, vol. 47, pp. 230–238, Feb. 1999.

[82] W. Chen and U. Mirta, "Training sequence optimization: Comparisons and an alternative criterion," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 1987–1991, Dec. 2000.

[83] S. Zhang, F. Gao, and C.-X. Pei, "Optimal training design for individual channel estimation in two-way relay networks," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4987–4991, 2012.

[84] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 384–391, Mar. 2003.

[85] F. Gao, T. Cui, and A. Nallanathan, "Optimal training design for channel estimation in decode-and-forward relay networks with individual and total power constraints," *IEEE Trans. Signal Processing*, vol. 56, no. 12, pp. 5937–5949, Dec. 2008.

[86] X. Zhou, T. Lamahewa, P. Sadeghi, and S. Durrani, "Optimizing antenna configuration for mimo systems with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1177–1181, Mar. 2009.

[87] X. Zhou, S. Durrani, and H. Jones, "Connectivity analysis of wireless ad hoc networks with beamforming," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 5247–5257, Nov. 2009.

[88] X. Zhou, T. Lamahewa, P. Sadeghi, and S. Durrani, "Two-way training: Optimal power allocation for pilot and data transmission," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 564–569, Mar. 2010.

[89] T. Cui and C. Tellambura, "Joint channel and frequency offset estimation and training sequence design for MIMO systems over frequency selective channels," in *Proc IEEE Global Commun. Conf. (GLOBECOM)*, 2004.

[90] F. Simoens and M. Moeneclaey, "Computationally efficient frequency offset estimation for flat-fading MIMO channels: Performance analysis and training sequence design," in *Proc IEEE Global Commun. Conf. (GLOBECOM)*, 2004.

[91] M. Ghogho and A. Swami, "Training design for channel and CFO estimation in MIMO systems," in *Proc. IEEE International Conference on Accoustics, Speech and Signal Processing (ICASSP)*, 2006.

[92] P. A. Parker, D. W. Bliss, P. Mitran, and V. Tarokh, "Adaptive frequency synchronization for collaborative communication systems," in *Proc. IEEE ICDCSW*, 2007.

[93] P. Stoica and O. Besson, "Training sequence design for frequency offset and frequency-selective channel estimation," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1910–1917, Nov. 2003.

[94] Y.-C. Wu, S. C. Chan, and E. Serpedin, "Symbol-timing estimation in space-time coding systems based on orthogonal training sequences," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 603–613, Mar. 2005.

[95] Z. Ding and Y. Li, *Blind Equalization and Identification*, Signal Processing and Communication Series, 2001.

[96] Y. Yu, A. P. Petropulu, H. V. Poor, and V. Koivunen, "Blind estimation of multiple carrier frequency offsets," in *Proc. IEEE PIMRC*, 2007.

[97] X. Liu, J. Kountouriotis, A. P. Petropulu, and K. R. Dandekar, "Aloha with collision resolution (ALOHA-CR): Theory and software defined radio implementation," *IEEE Trans. Signal Processing*, vol. 58, no. 8, pp. 4396–4410, Aug. 2010.

[98] X. Liu, A. Petropulu, H. Poor, and V. Koivunen, "Blind separation of two users based on user delays and optimal pulse shape design," *EURASIP Journal on Wireless Communications and Networking*, 2010.

[99] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Estimation of synchronization paramters in AF cooperative networks," in *Proc IEEE ICC*, 2012.

[100] 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," *accepted for publication IEEE Trans. Signal Processing*, 2013.

[101] P. A. Parker, P. Mitran, D. W. Bliss, and V. Tarokh, "On bounds and algorithms for frequency synchronization for collaborative communication systems," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3742–3752, Aug. 2008.

[102] 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," *submitted to IEEE Trans. Commun.*, 2012.

[103] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Blind timing and carrier synchronization in decode-and-forward cooperative systems," in *Proc IEEE ICC*, 2011.

[104] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Blind timing and carrier synchronization in distributed MIMO communication systems," *IET Communications*, vol. 50, no. 7, pp. 1028–1037, May 2011.

[105] H. Mehrpouyan and S. D. Blostein, "Comments on "timing estimation and resynchronization for amplify-and-forward communication systems," *IEEE Signal Processing Lett.*, vol. 59, pp. 4047–4048, Aug. 2011.

[106] X. Li, C. Xing, Y. C. Wu, and S. C. Chan, "Authors reply to comments on "timing estimation and resynchronization for amplify-and-forward communication systems"," *IEEE Signal Processing Lett.*, vol. 59, pp. 4048–4049, Aug. 2011.

[107] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.

[108] I. Ziskand and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 10, pp. 1553–1560, Oct. 1988.

[109] T. K. Moon, "The expectation-maximazation algorithm," *IEEE Signal Processing Mag.*, vol. 13, pp. 47–60, Nov. 1996.

[110] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 4, pp. 477–489, Apr. 1988.

[111] G. J. McLachlan and T. Krishnan, *The EM algorithm and Extensions*, John Wiley and Sons, 2008.

[112] X. L. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: A general framework," *Biometrika*, vol. 8, no. 2, pp. 267–278, Jun. 1993.

[113] R. J. Leveque, *Finite difference methods for ordinary and partial differntial equations*, Society for Industrial and Applied Mathematics (SIAM), 2007.

[114] T. A. Fesler and A. O. Hero, "Space-alternating generalized expectation maximization algorithm," *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2664–2677, Oct. 1994.

[115] N. Moller, "On Schoonhage's algorithm and subquadratic integer gcd computation," *Mathematics of Computation*, vol. 77, pp. 589–607, Jan. 2008.

[116] B. P. Singh and R. Singh, *Electronic devices and integrated circuits*, Pearson Education India, 2006.

[117] G. N. Tavares, L. M. Tavares, and A. Petrolino, "On the true Cramer-Rao lower bound for data-aided carrier phase-independent frequency offset and symbol timing estimation," *IEEE Trans. Commun.*, vol. 58, pp. 442–447, Feb. 2010.

[118] F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat-fading channels: A cyclostationary approach," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 400–411, Mar. 1998.

[119] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A practical approach to global optimization.*, Springer-Verlag Berlin Heidelberg, 2005.

[120] F. Neri and V. Tirronen, "Recent advances in differential evolution: A survey and experimental analysis," *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 61–106, Oct. 2009.

[121] G. Rudolph, "Convergence of evolutionary algorithms in general search space," in *Proc. IEEE Int. Conf. Evol. Computat.*, May 1996, pp. 50–54.

[122] M. Vasile, E. Minisci, and M. Locatelli, "An inflationary differential evolution algorithm for space trajectory optimization," *IEEE Trans. Evolution. Computat.*, vol. 15, no. 2, Apr. 2011.

[123] Y. M. Bishop, S. E. Fienberg, and P. W. Holland, *Discrete Multivariate Analysis: Theory and Practice*, Springer, 2007.

[124] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A spacetime coding modem for high-data-rate wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1459–1478, Oct. 1998.

[125] K. Rajawat and A. K. Chaturvedi, "A low complexity symbol timing estimator for MIMO systems using two samples per symbol," *IEEE Commun. Lett.*, vol. 10, no. 7, pp. 525–527, Jul. 2006.

[126] S. Bay, B. Geller, A. Renaux, J.-P. Barbot, and J.-M. Brossier, "On the hybrid cramr rao bound and its application to dynamical phase estimation," *IEEE Signal Processing Lett.*, vol. 15, pp. 453–456, May 2008.

[127] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[128] G. L. Stuber, *Principles of mobile communication*, Springer, 2011.

[129] J. Fang, A. R. Leyman, Y. Chew, and Y. C. Liang, "A cumulant interference subspace cancellation method for blind SISO channel estimation," *IEEE Trans. Signal Process.*, vol. 54, pp. 784790, 2006.

[130] Y.-C. Wu and E. Serpedin, "Unified analysis of a class of blind feedforward symbol timing estimators employing second-order statistics," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 737–742, 2006.

[131] F. D. Natali, "AFC tracking algorithms," *IEEE Trans. Commun.*, vol. 32, pp. 935–947, Aug. 1984.

[132] A. N. Andrea and U. Mengal, "Performance of a quradricorrelator driven by modulated signals," *IEEE Trans. Commun.*, vol. 38, no. 11, pp. 1952–1957, Nov. 1990.

[133] M. Ghogho, A. Swami, and T. Durrani, "On blind carrier recovery in tme-selective fading channels," in *Proc. 33rd Conf. on Signals, Systems and Computers*, 1999, vol. 1, pp. 243–247.

[134] Y. Wang, K. Shi, and E. Serpedin, "Non data-aided feedforward carrier frequency offset estimators for QAM constellations: A nonlinear least-squares approach," *EURASIP Journal on Applied Signal Processing*, vol. 13, pp. 1993–2001, 2004.

[135] M. K. Simon and D. Divsalar, "Doppler corrected differential detection in MPSK," *IEEE Trans. Commun.*, vol. 37, no. 2, pp. 99–109, Feb. 1989.

[136] U. Mengali and A. N. Andrea, *Synchronization Techniques for Digital Receivers*, Plenum Press, New York, 1997.

[137] K. H. Mueller and M. Muller, "Timing recovery in digital synchronous data receivers," *IEEE Trans. Commun.*, vol. COM-14, pp. 516–530, May 1976.

[138] F. M. Gardner, "A BPSK/QPSK timing error detector for sampled receivers," *IEEE Trans. Commun.*, vol. 34, pp. 423–429, May 1986.

[139] M. Oerder and H. Meyr, "Digital filter and square timing recovery," *IEEE Trans. Commun.*, vol. 36, no. 5, pp. 605–612, May 1988.

[140] J. F. Cardoso, "Infomax and maximum likelihood for blind source separation," *IEEE Signal Processing Lett.*, vol. 4, pp. 112–114, Apr. 1997.

[141] N. Vlassis and Y. Motomura, "Efficient source adaptivity in independent component analysis," *IEEE Trans. Neural Networks*, vol. 12, no. 3, pp. 559–566, May 2001.

[142] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind convolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.

[143] C. Chang and Z. Ding, "A matrix-pencil approach to blind separation of colored nonstationary signals," *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 900–907, Mar. 2000.

[144] F. Yin, T. Mei, and J. Wang, "Blind source separation based on decorrelation and nonstationarity," *IEEE Trans. Circuits Syst.*, vol. 54, no. 5, pp. 1150–1158, 2007.

[145] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," in *Proc. IEEE Radar and Signal Process.*, 1993, vol. 140, pp. 362–370.

[146] C. B. Papadias, "Globally convergent blind source separation based on a multiuser kurtosis maximization criterion," *IEEE Signal Processing Lett.*, vol. 48, no. 12, pp. 3508–3519, Dec. 2000.

[147] T. Mei, F. Yin, and J. Wang, "Blind source separation based on cumulants with time and frequency non-properites," *IEEE Trans. Audio, Speech & Language Process.*, vol. 17, no. 6, pp. 1099–1108, Aug. 2009.

[148] C. B. Papadias and A. J. Paulraj, "A constant modulus algorithm for multiuser signal separation in presence of delay spread using antenna arrays," *IEEE Signal Processing Lett.*, vol. 4, pp. 178–181, Jun. 1997.

[149] A.-J. Van Der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[150] Y. Yao and G. B. Giannakis, "On regularity and identifiability of blind source separation under constant-modulus constraints," *IEEE Trans. Signal Processing*, vol. 4, no. 4, pp. 1272–1281, Apr. 2005.

[151] Y. Li, D. Powers, and J. Peach, "Comparison of blind source separation algorithms," in *Proc. WSES, Advances in Neural Networks and Applications*, 2000, pp. 18–21.

[152] J. F. Cardoso, "Blind signal separation: Statistical principles," in *Proc. IEEE*, 1998, vol. 86, pp. 2009–2025.

[153] A. P. Petropulu, M. Olivieri, Y. Yu, L. Dong, and A. Lackpour, "Pulse-shaping for blind multi-user separation in distributed MISO configurations," in *Proc. IEEE ICASSP*, Apr. 2008, pp. 2741–2744.

[154] A. Hyvarinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, vol. 13, pp. 411–430, 2004.

[155] J. F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 17, no. 1, pp. 161–164, 1996.

[156] E. Beres and R. Adve, "Blind channel estimation for orthogonal STBC in MISO systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 2042–2050, Jul. 2007.

[157] Y. Jing and H. Jafarkhani, "Single and multiple relay selection schemes and their achievalbe diversity orders," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 14141423, Mar. 2009.

[158] Y. Fan, J. Thompson, A. Adinoyi, and H. Yanikomeroglu, "Space diversity for mulit-antenna multi-relay channels," in *Proc. 12th European Wireless Conference*, 2006.

[159] H. Muhaidat and M. Uysal, "Cooperative diversity with multiple antenna nodes in fading relay channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3036–3046, Aug. 2008.

[160] M. Fareed and M. Uysal, "On relay selection for decode-and-forward relaying," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3341–3346, Jul. 2009.

[161] D. Gesbert, S. Hanly, H. Huang, S. S. Shitz, and W. Yu, "Cooperative communications in MIMO cellular networks," *IEEE J. Select. Areas Commun.*, vol. 28, no. 9, pp. 1377–1379, Dec. 2010.