Timing and Frequency Synchronization and Channel Estimation in OFDM-based Systems

Hamed Abdzadeh Ziabari

A Thesis
In the Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Electrical and Computer Engineering) at
Concordia University
Montreal, Quebec, Canada

December 2017

Concordia University

School of Graduate Studies

This is to certify that the thesis prepared

By: **Mr. Hamed Abdzadeh Ziabari**

Entitled: **Timing and Frequency Synchronization and Channel Estimation in OFDM-based Systems**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Electrical & Computer Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining commitee:

_____ Chair
Dr. Alex de Visscher

_____ External Examiner
Dr. Ioannis N. Psaromiligkos

_____ External to Program
Dr. Youmin Zhang

_____ Examiner
Dr. M. Omair Ahmad

_____ Examiner
Dr. Yousef R. Shayan

_____ Thesis Co-Supervisor
Dr. M.N.S. Swamy

_____ Thesis Co-Supervisor
Dr. Wei-Ping Zhu

Approved by _____
Dr. Wei-Ping Zhu, Graduate Program Director

Friday, February 2, 2018

_____
Dr. Amir Asif, Dean
Faculty of Engineering and Computer Science

Abstract

Timing and Frequency Synchronization and Channel Estimation in OFDM-based Systems

Hamed Abdzadeh Ziabari, Ph.D.

Concordia University, 2017

Orthogonal frequency division multiplexing (OFDM) due to its appealing features, such as robustness against frequency selective fading and simple channel equalization, is adopted in communications systems such as WLAN, WiMAX and DVB. However, OFDM systems are sensitive to synchronization errors caused by timing and frequency offsets. Besides, the OFDM receiver has to perform channel estimation for coherent detection. The goal of this thesis is to investigate new methods for timing and frequency synchronization and channel estimation in OFDM-based systems.

First, we investigate new methods for preamble-aided coarse timing estimation in OFDM systems. Two novel timing metrics using high order statistics-based correlation and differential normalization functions are proposed. The performance of the new timing metrics is evaluated using different criteria including class-separability, robustness to the carrier frequency offset, and computational complexity. It is shown that the new timing metrics can considerably increase the class-separability due to their more distinct values at correct and wrong timing instants, and thus give a significantly better detection performance than the existing timing metrics do. Furthermore, a new method for coarse estimation of the start of the frame is proposed, which remarkably reduces the probability of inter-symbol interference (ISI). The improved performances of the new schemes in multipath fading channels are shown by the probabilities of false alarm, missed-detection and ISI obtained through computer simulations. Second, a novel pilot-aided algorithm is proposed for the detection of integer frequency offset (IFO) in OFDM systems. By transforming the IFO into two new integer parameters, the proposed method can largely

reduce the number of trial values for the true IFO. The two new integer parameters are detected using two different pilot sequences, a periodic pilot sequence and an aperiodic pilot sequence. It is shown that the new scheme can significantly reduce the computational complexity while achieving almost the same performance as the previous methods do.

Third, we propose a method for joint timing and frequency synchronization and channel estimation for OFDM systems that operate in doubly selective channels. Basis expansion modeling (BEM) that captures the time variations of the channel is used to reduce the number of unknown channel parameters. The BEM coefficients along with the timing and frequency offsets are estimated by using a maximum likelihood (ML) approach. An efficient algorithm is then proposed for reducing the computational complexity of the joint estimation. The complexity of the new method is assessed in terms of the number of multiplications. The mean square estimation error of the proposed method is evaluated in comparison with previous methods, indicating a remarkable performance improvement by the new method.

Fourth, we present a new scheme for joint estimation of CFO and doubly selective channel in orthogonal frequency division multiplexing systems. In the proposed preamble-aided method, the time-varying channel is represented using BEM. CFO and BEM coefficients are estimated using the principles of particle and Kalman filtering. The performance of the new method in multipath time-varying channels is investigated in comparison with previous schemes. The simulation results indicate a remarkable performance improvement in terms of the mean square errors of CFO and channel estimates.

Fifth, a novel algorithm is proposed for timing and frequency synchronization and channel estimation in the uplink of orthogonal frequency division multiple access (OFDMA) systems by considering high-mobility situations and the generalized subcarrier assignment. By using BEM to represent a doubly selective channel, a maximum likelihood (ML) approach is proposed to jointly estimate the timing and frequency offsets of different users as well as the BEM coefficients of the time-varying channels. A space-alternating generalized expectation-maximization algorithm is then employed to transform the maximization problem for all users into several simpler maximization problems for each user. The computational complexity of the new timing and frequency offset estimator is analyzed and

its performance in comparison with that of existing methods using the mean square error is evaluated .

Finally, two novel approaches for joint CFO and doubly selective channel estimation in the uplink of multiple-input multiple-output orthogonal frequency division multiple access (MIMO-OFDMA) systems are presented. Considering high-mobility situations, where channels change within an OFDMA symbol interval, and the time varying nature of CFOs, BEM is employed to represent the time variations of the channel. Two new approaches are then proposed based on Schmidt Kalman filtering (SKF). The first approach utilizes Schmidt extended Kalman filtering for each user to estimate the CFO and BEM coefficients. The second approach uses Gaussian particle filter along with SKF to estimate the CFO and BEM coefficients of each user. The Bayesian Cramer Rao bound is derived, and performance of the new schemes are evaluated using mean square error. It is demonstrated that the new schemes can significantly improve the mean square error performance in comparison with that of the existing methods.

To my loving family!

Acknowledgments

My most sincere thanks go to my supervisors Dr. Wei-Ping Zhu and Dr. M.N.S. Swamy for their continuous support, patience and insightful guidance during different stages of my Ph.D studies and writing of this thesis.

I would like to express my gratitude to my committee members for their precious time evaluating this thesis.

I wish to thank my friends and colleagues at Concordia University: Reza Movahedinia, Ali Muhammad, Omid Saatlou, Hanif Rashtiyan, Ali Mohebbi, Hossein Kourkchi, and Dr. Xiaodong Ji who walked by my side during the ups and downs of my research, shared my moments of distress and joy, and made the past four years one of the most memorable periods of my life.

Finally, I would like to express my eternal appreciation to my family for their everlasting support and love. My special thanks go to my sister Bahareh without whose help and support this thesis could have not come into existence.

Contents

List of Figures

List of Tables

List of Symbols

$\mathbf{x}^T$:	Transpose of $\mathbf{x}$
$\mathbf{x}^H$:	Complex Conjugate Transpose of $\mathbf{x}$
$\overline{\mathbf{x}}$:	Element-wise Complex Conjugate of $\mathbf{x}$
x^*:	Complex Conjugate of x
$E\{\mathbf{x}\}$:	Expectaion of $\mathbf{x}$
diag$\{\mathbf{x}\}$:	Diagonal Matrix with Diagonal Elements $\mathbf{x}$
blkdiag$\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$:	Block Diagonal Matrix with Matrices $\mathbf{x}_1, \cdots, \mathbf{x}_N$ on the Diagonal
$\lvert\mathbf{x}\rvert$:	Absolute Value of the Elements of $\mathbf{x}$
$\lVert\mathbf{x}\rVert$:	Euclidean distance of the vector of $\mathbf{x}$
Re$\{\mathbf{x}\}$:	Real Parts of the Elements of $\mathbf{x}$
Im$\{\mathbf{x}\}$:	Imaginary Parts of the Elements of $\mathbf{x}$
$\mathbf{I}_N$:	$N \times N$ Identity Matrix
$\mathbf{0}_{N\times L}$:	$N \times L$ Zero Matrix
$\otimes$:	Kronecker Product
$\circ$:	Hadamard product
$\mathcal{N}\big(x; \mu, \sigma^2\big)$:	Univariate Gaussian Density with Mean μ and Variance σ^2.
$\mathcal{N}\big(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}\big)$:	Multivariate Gaussian Density with Mean Vector $\boldsymbol{\mu}$ and Covariance Matrix $\boldsymbol{\Sigma}$
$\mathcal{U}(a, b)$:	Uniform Distribution in the Range $[a, b]$

$\backslash x$:	Exclusion of the Parameters of the xth User
$\lvert x \rvert_N$:	x modulo N

List of Abbreviations

AWGN:	Additive White Gaussian Noise
BCRB:	Bayesian Cramer Rao Bound
BEM:	Basis Expansion Modeling
BER :	Bit Error Rate
BSEKF :	BEM Based Schmidt Extended Kalman Filtering
BSK-GPF:	BEM Based Schmidt Kalman and Gaussian Particle Filtering
CE-BEM :	Complex Exponential Basis Expansion Modeling
CFO:	Carrier Frequency Offset
CP:	Cyclic Prefix
DDN:	Double Differential Normalization
DFT:	Discrete Fourier Transform
DoM:	Difference of Magnitude
DPSS-BEM:	Discrete Prolate Spheroidal Sequences Basis Expansion Modeling
DVB-T:	Digital Video Broadcasting Terrestrial TV
EM:	Expectation-Maximization
FFO:	Fractional Frequency Offset
GCE-BEM:	Generalized Complex Exponential Basis Expansion Modeling
GPF:	Gaussian Particle Filtering
IFO:	Integer Frequency Offset
IDFT:	Inverse Discrete Fourier Transform

LTE:	Long-Term Evolution
MIMO:	Multiple Input Multiple Output
ML :	Maximum Likelihood
MMSE:	Mimimum Mean Square Error
MoD:	Magnitude of Difference
MSE :	Mean Square Error
NC-OFDM:	Non-Contiguous Orthogonal Frequency Division Multiplexing
NMSE :	Normalized Mean Square Error
OFDM:	Orthogonal Frequency Division Multiplexing
OFDMA:	Orthogonal Frequency Division Multiple Access
PAPR:	Peak-to-Average Power Ratio
P-BEM :	Polynomial Basis Expansion Modeling
PDF:	Probability Density Function
PN:	Pseudo Noise
QPSK:	Quadrature Phase Shift Keying
SAGE:	Space-Alternating Generalized Expectation-Maximization
SEKF:	Schmidt Extended Kalman Filtering
SKF:	Schmidt Kalman Filtering
SNR:	Signal to Noise Ratio
SUI:	Stanford University Interim
WiMAX:	Worldwide Interoperability for Microwave Access
WLAN:	Wireless Local Area Network

Chapter 1

Introduction

1.1 Background

The demand for multimedia and wireless communications is increasing at an extremely rapid pace and this trend is expected to continue in the near future. The common feature of many current wireless standards for high-rate multimedia transmission is the adoption of orthogonal frequency division multiplexing (OFDM) [1]. Communication systems such as digital video broadcasting terrestrial TV (DVB-T) [2], digital audio broadcasting [3], and IEEE 802.11a wireless local area network (WLAN) [4] are based on OFDM. The OFDM technique, by converting a frequency selective fading channel into a collection of flat fading channels, offers increased robustness against multipath distortions as channel equalization can be easily performed in the frequency domain through a bank of one-tap multipliers [5]. It also provides larger flexibility by allowing independent selection of the modulation parameters (like the constellation size and coding scheme) over each subcarrier [6]. Furthermore, the discrete Fourier transform (DFT) applied in OFDM systems can nowadays be efficiently implemented [7] using algorithms such as the fast Fourier transform. However, OFDM systems suffer from some disadvantages including large peak-to-average power-ratio (PAPR) as well as high sensitivity to synchronization errors (which are caused by timing and frequency offsets). Synchronization issues are of

great importance in digital communication systems, especially in the OFDM-based systems [8]. Errors associated with timing synchronization which refers to correctly finding the start of the frame, degrade the performance of an OFDM receiver and result in intercarrier interference (ICI) and inter-symbol interference (ISI) [9]. Timing offset can also have a severe impact on system performance due to its adverse effect on channel estimation [10–13]. OFDM systems are also sensitive to carrier frequency offset (CFO) which is caused by the Doppler shift and/or local oscillators' mismatch between the transmitter and the receiver. The fractional part of CFO introduces ICI and destroys the orthogonality of OFDM subcarriers [8], whereas the integer part of CFO can result in appearance of received symbols in a wrong position at the DFT output [1].

Timing estimation aims at obtaining the timing instant of start of the frame and is usually performed in two stages, coarse timing estimation and fine timing estimation. The former stage aims at finding an inter-symbol interference (ISI)-free part of the preamble and preventing the destruction of orthogonality between subcarriers. When the first stage is performed, there remains residual timing offset. As long as this residual timing offset is less than the length of CP, the orthogonality between subcarriers is maintained. However, this residual timing offset can considerably degrade the performance of the pilot-aided channel estimators and tracking methods [14–16]. Moreover, if this residual timing offset is not estimated and compensated for, each OFDM symbol needs to employ a long CP to mitigate ISI, which would significantly reduce the data throughput [17]. Consequently, the second stage aims at the estimation of the residual timing offset.

CFO estimation, on the other hand, is usually carried out by dividing it into two parts, a fractional part (FFO) and an integer part (IFO). The fractional part causes destruction of orthogonality between the subcarriers and the integer part results in the cyclic shifts of subcarriers in the frequency domain. First, FFO is estimated and compensated for, and next IFO is estimated. There are also methods that estimate CFO in two steps called coarse and fine estimation. First, a coarse estimate of CFO is obtained, and next the

residual CFO is estimated.

It is worth mentioning that coarse and fine timing or frequency synchronizations are achieved when coarse and fine timing offset or CFO estimates are obtained and the signal is corrected accordingly.

1.2 Literature Survey

1.2.1 Frequency Selective Slow Fading Channels

Most of the synchronization schemes in the literature are developed for OFDM-based systems that operate in multipath slow fading channels. Timing offset estimation methods developed for these channels can be divided into two main categories: blind methods and preamble-aided schemes. The methods in the former category usually use the periodicity of the cyclic prefix (CP) [18–27], whereas the schemes in the second category apply the characteristics of OFDM training symbols for timing synchronization.

With regard to the first catagory, timing synchronization is performed in [18] by obtaining the estimate of channel-tap powers using the correlation characteristics of the CP. In [21], a maximum likelihood approach is adopted under the assumption of an AWGN channel. In [23], a metric is defined that estimates the timing offset irrespective of the channel conditions when the signal-to-noise ratio is high. The timing estimation method in [24] makes use of the correlation between the cyclic extension and data portion of the received OFDM symbols. Although it can identify the ISI-free region in multipath channels, a large number of symbols is required to obtain an accurate estimate [18]. The algorithm in [25] achieves synchronization when an autocorrelation matrix constructed from the received signal achieves a minimum rank. This method is more complex and requires more statistics. A blind method is proposed in [28] based on invariance properties and cyclostationarity of CP. The proposed technique in [29] estimates symbol timing offsets by minimizing the power difference between the subcarriers with similar indices over two

successive OFDM symbols based on the assumption that the channel slowly changes over time. The authors in [30] investigate a CP-based synchronization method for an arbitrary size observation window.

Regarding the second category, i.e., preamble aided timing synchronization in slow fading channels, the authors of [31] proposed a robust timing metric to acheive coarse timing synchronization. This metric is designed using the correlation properties of a preamble with two identical parts in the time domain. A timing offset estimation method equivalent to [31] was proposed in [32] and was analyzed in [33]. In [34–36], the authors changed the sign of the identical parts of the preamble in [31] to make the timing metric sharper. A timing metric taking advantage of the central symmetry of a preamble was proposed in [37]. In [14], the authors presented a method to improve the false alarm probability by applying the periodic parts of more than one preamble. The application of differential normalization for making the values of the coarse timing metric at wrong and correct timing instants more distinct is proposed in [38]. The authors in [39] demonstrated that sufficient statistics for detection of a preamble composed of two identical parts do not exist and therefore, the second order statistics used in previous methods are not optimal in terms of utilizing sufficient statistics. They also proposed fourth order statistics for frame detection. Fourth order differential normalization is used in [40] to improve the synchronization performance by increasing the class-seperablity. Timing metrics that apply the cross-correlation of the received signal and pure preamble can be found in [41–49]. The cross-correlation based methods are usually used for fine timing estimation. In [41], for having a sharper timing metric, the identical parts of the preamble are multiplied by a pseudo noise (PN) sequence. A timing metric that works independent of the preamble structure is proposed in [42]. The multiplication of a correlation metric with a cross-correlation metric was proposed in [43, 44]. The authors in [45] improved the timing estimation performance by increasing the correlation length. The sensitivity of this method to the carrier frequency offset (CFO) was later removed in [46]. A cross-correlation based

method using Zadoff–Chu sequence is presented in [49] with the aim of reducing the timing estimation sensitivity to CFO. Timing estimation based on multiple signal classification is proposed in [50].

Frequency synchronization, which is to estimate and correct the CFO, similar to timing synchronization, can be accomplished either blindly or with utilization of training symbols. The authors in [51] proposed a blind method for CFO estimation based on a kurtosis-type criterion. In [52] it was pointed out that CFO destroys the orthogonality between the subcarriers leading to nondiagonal signal covariance matrices in the frequency domain, and therefore, a blind method was proposed that enforces a diagonal structure by minimizing the power of nondiagonal elements. Two methods for constant modulous signaling are proposed in [53] and [54]. The former employs the banded structure of the covariance matrix of the received signal with perfect CFO compensation, the latter uses a cost function based on the assumption that in slowly time-varying channels subcarriers having the same indexes in two consecutive OFDM symbols experience nearly the same channel effect. The authors in [55] proposed to perform CFO estimation based on covariance fitting criterion between two nearby OFDM symbols. For MIMO-OFDM systems with constant modulus constellation, a rank reduction criterion is applied in [56]. A CP-based method that is robust to the symbol timing synchronization error is introduced in [57]. Furthermore, some blind methods taking advantage of null subcarriers are presented in [58–62].

Pilot-aided CFO estimation in multipath slow fading channels is usually performed by dividing the CFO into a fractional part (FFO) and an integer part (IFO). With regard to the estimation of the FFO, in [31], the authors used a preamble composed of two identical parts in the time domain, and took advantage of the phase difference between the two parts. In [63] and [64], FFO estimation methods were proposed that utilize the phase difference between several identical parts of a preamble, and offer a wider estimation range than that in [31]. A general CFO estimator using training OFDM symbols was presented in [65] which takes advantage of the channel side information. A periodogram based approach

is presented in [66]. The methods in [67] and [68] obtain the FFO by adopting an ML approach, and the scheme in [69] resorts to the least-squares (LS) principle. Regarding IFO estimation, the integer part is estimated in [31] using a differentially-modulated PN sequence positioned between the subcarriers of two consecutive OFDM symbols. To reduce the number of training symbols, the authors in [70] proposed to put the differentially-modulated PN sequence between the adjacent subcarrier of one OFDM symbol. Later in [71], a maximum likelihood approach was presented for frequency selective channels. In [72] the adoption of cross ambiguity function was proposed to jointly estimate the timing and integer frequency offset.

There are also methods that estimate the timing offset jointly with the frequency offset or estimate these parameters jointly with frequency selective slow fading channel. In [15], joint timing offset and channel estimation is performed using ML criterion. To estimate channel jointly with CFO, an expectation-maximization (EM) based algorithm is proposed in [73]. The authors in [16] have put forward an ML algorithm for joint CFO and timing offset estimation. In [67, 68], joint timing and CFO estimation is performed based on an ML principle using preambles with repeated parts in the time domain. The work in [74] represents a method for correlation based timing estimation followed by joint timing and FFO estimation. Further, proposed in [46,66] are methods that achieve frequency estimate in two steps of estimation: coarse and fine.

Considering the above mentioned methods, pilot-aided schemes achieve synchronization faster than blind methods and therefore, are more suitable for packet-based systems. With regard to timing synchronization, it is worth mentioning that the cross-correlation metrics are not as robust as auto-correlation metrics because they are usually sensitive to CFO [75], and need to know the exact preamble signal which may not be always practical [67]. Furthermore, the existing autocorrelation-based methods do not fully utilize the characteristics of a periodic preamble. The efficient utilization of these characteristics can result in a significant reduction of the probabilities of false alarm and missed detection.

On the other hand, methods dealing with IFO estimation such as the ML approach in [71] has a high computational complexity which may not be appealing in practice.

Orthogonal frequency division multiple access (OFDMA) is a multiuser version of OFDM in which subcarriers are divided into mutually exclusive clusters that are assigned to different users [1]. Synchronization in OFDMA systems is usually performed in three steps. In the first step, synchronization is performed during the downlink transmission where each mobile terminal carries out synchronization with the pilots transmitted by the base station by using schemes in simple OFDM systems. The estimated timing and frequency offsets are used for downlink data detection and also as references for uplink transmission. The signals received in the uplink transmission are affected by residual timing and frequency errors. Thus, at the second step, the base station has to estimate these residual offsets. This task is more challenging than OFDM systems because each user has its own timing and frequency offset, and the signal received at the base station is the superposition of the signals of all users. Finally the third step aims at the correction of timing and frequency errors. The correction of the signals is made at the user's side based on instructions transmitted by the base station on a feedback control channel [76].

For estimation of timing and frequency offset in the uplink transmission in a frequency selective slow fading channel, the schemes are usually dependent on the utilized subcarrier assignment scheme. In general, there are three types of distributing subcarriers among active users [76] : subband subcarrier assignment scheme, interleaved subcarrier assignment scheme, and generalized subcarrier assignment scheme. In the first scheme, a group adjacent subcarriers is assigned to each user. The main drawback of this approach is that it does not exploit the frequency diversity offered by the multipath channel since a deep fade might hit a substantial number of subcarriers of a given user. In the second scheme, the subcarriers of each user are uniformly spaced over the signal bandwidth at a fixed distance from each other. Although this method can fully exploit the channel frequency diversity, the current trend in OFDMA favors a more flexible allocation strategy. In the

third scheme, users can select the best available subcarriers (i.e., those with the highest signal-to-noise ratios (SNRs)). This scheme allows dynamic resource allocation and provides more flexibility than the other schemes.

Synchronization methods for the subband subcarrier assignment scheme can be found in [77, 78] and for the interleaved subcarrier assignment scheme can be found in [79–86]. For the generalized subcarrier assignment scheme which is more flexible than the other schemes, joint estimation of CFO and/or timing and channel based on alternating projection and space-alternating generalized expectation-maximization (SAGE) algorithms are respectively presented in [87] and [88]. To reduce the complexity, in [89] and [90], line search and variable projection methods are utilized respectively. Some further works that have addressed synchronization in the uplink for the generalized subcarrier assignment scheme are found in [91–97]. In [98] and [99], channels and CFOs have been estimated based on Rao-Blackwellization principle for MIMO-OFDMA systems. Both of these works assume that the channels are constant within an OFDMA symbol but vary from one OFDMA symbol to another. In addition, the authors in [99] have taken into account the fact that CFO can be time varying for different OFDMA symbols.

1.2.2 Doubly Selective Channels

All the schemes presented in the previous subsection assume that the channel remains constant at least within one OFDM symbol. Recent research, however, has been focused more on the estimation of frequency offset and doubly selective (time varying frequency selective) channels. Underwater acoustic channels, future aeronautical communication channels [100], and the high data rate wireless communication channels in LTE-Advanced and WiMAX are examples of doubly selective channels. OFDM systems operating in these channels confront with channel changes within one OFDM symbol which can incur loss of orthogonality between subcarriers and inter-carrier interference [101]. This issue degrades the performance of the methods addressed in previous subsection, and makes

them unsuitable for doubly selective channels.

In order to estimate the channel, the authors in [102] proposed to utilize a basis expansion modeling (BEM) for capturing the variations of doubly selective channels. This method significantly reduces the number of unknown channel parameters to be estimated. Using a truncated Fourier series, the complex exponential BEM (CE-BEM) was proposed in [103]. Due to Gibbs phenomenon and spectral leakage, CE-BEM results in considerable phase and amplitude errors at the beginning and end of data blocks [104]. To overcome this weakness of CE-BEM, different BEMs have been proposed including Generalized CE-BEM (GCE-BEM) in [105], polynomial BEM (P-BEM) in [106], and discrete prolate spheroidal sequences BEM (DPSS-BEM) in [104]. Later, in [107], the authors addressed joint estimation of doubly selective channel and carrier frequency offset (CFO) using basis expansion modeling of the channel and recursive least squares, ML and maximum-a-posteriori approaches. In [108], joint CFO and channel estimation was accomplished using BEM for a MIMO-OFDM system, and in [109] joint estimation of CFO and BEM coefficients was performed for an OFDM system by utilizing expectation-maximization algorithm and Kalman filtering. Note that the methods in [108] and [109] deal mainly with joint channel and CFO tracking, and similar to [107], they did not consider timing estimation. It is also noted that [110–113] deal with blind estimation of timing and/or frequency offset in doubly selective channels. To my best knowledge, there have been only blind methods for joint estimation of timing and frequency offset in the case of channels changing within an OFDM symbol [110–113]. In other words, all of the previous preamble-aided methods have considered at most joint estimation of CFO and doubly selective channel in OFDM systems, and have neglected the issue of timing offset estimation. Moreover, the existing CFO estimation methods for doubly selective channels are mainly based on ML criterion. Taking into account the fact that the ML approach for CFO estimation is a grid search based algorithm and needs an exhaustive search over all possible values of

CFO, the ML method appears to be impractical, and has thus been replaced with methods such as Newton-Raphson-based scheme in [107] or extended Kalman filtering in [108]. Unfortunately, these methods suffer from the drawback that they are not guaranteed to converge to the true CFO because of the nonlinearity of the CFO estimation problem.

Synchronization in the uplink of OFDMA systems operating in doubly selective channels is recently carried out in [114–117]. In [114,115,117], the authors proposed a method for CFO and doubly selective channel estimation using SAGE algorithm. The paper [116] deals with frequency synchronization and time-varying channel estimation applying oblique projection. A method for joint CFO and time-varying channel estimation for the uplink of MIMO-OFDMA systems is presented in [99] using Kalman and particle filters. To the author's best knowledge, similar to OFDM systems, all of the previous preamble-aided synchronization methods in doubly selective channels have considered only CFO estimation, and the problem of preamble-aided timing estimation in OFDMA systems in a doubly selective channel has not been yet tackled. Furthermore, the above mentioned methods for OFDMA systems in doubly selective channels, assumed that CFO is constant for all OFDMA symbols, and moreover, did not perform CFO and channel tracking. In practice, however, CFO can be time varying due to oscillators frequency drifting or time-varying Doppler shifts [118–120] as modeled in [99,121,122]. Besides, for CFO estimation, there is no closed-form solution, and thus grid search is often needed.

1.3 Contribution

The work presented in this thesis represents a number of contributions in the field of timing and frequency synchronization and channel estimation for OFDM-based systems. The most significant contributions of this research can be summarized as follows:

1) We investigate the use of n-th order statistics for coarse timing synchronization [123]. Considering a preamble composed of two identical parts in the time domain, n-th order timing metrics using correlation and differential normalization functions are proposed. The

performance of the n-th order metrics and the effect of increasing the order are examined. We show that by increasing the order of the correlation and normalization functions, the difference between the values of the timing metrics at wrong and correct timing instants can be considerably increased, thereby improving the performance of timing estimation. To evaluate the performance of the new timing metrics, we consider two class-separablity criteria: the first one is the difference between the means of the timing metrics at the wrong and correct timing instances, and the second criterion utilizes both the means and the variances of the timing metrics to formulate a difference function. Furthermore, the computational complexity of the metrics, their robustness to the carrier frequency offset (CFO) and the performance limit on the order of the metrics are also investigated. Simulations are carried out in multipath fading channels, giving false alarm and missed detection probabilities along with the probability of ISI of the proposed methods as compared with the existing methods in the literature.

2) We propose a novel pilot-aided algorithm for the detection of IFO in OFDM systems [124]. By transforming the IFO into two new integer parameters, the proposed method can largely reduce the number of trial values for the true IFO, and consequently, the complexity of the IFO detector is significantly decreased. Furthermore, it is shown that the proposed scheme can be applied to many of existing IFO metrics. The complexity, probability of correct detection and BER of the new scheme are compared with those of previous methods in multipath fading channels. It is shown that the new scheme can significantly reduce the computational complexity while achieving almost the same performance as compared to previous methods.

3) We investigate the crucial problem of preamble-aided timing, CFO and doubly selective channel estimation and propose a novel ML algorithm for joint estimation of these parameters [125]. It is assumed that using correlation-based method, a coarse estimate of timing offset is already obtained, and thus the proposed joint estimation problem aims at estimation of residual timing offset along with CFO and doubly selective channel. To

reduce the computational complexity, we propose a new method which first performs initial joint estimation of timing, CFO, and channel. Consequently, CFO, timing and channel estimates are further refined. Moreover, the new fine CFO estimation scheme that takes advantage of fast Fourier transform (FFT) is of a remarkably reduced complexity. Finally, we evaluate the mean square error (MSE) performance of the new method.

4) We present a new preamble-aided method for joint CFO and doubly selective channel estimation [126]. We model the time-varying channel using BEM, and propose particle and Kalman filtering for CFO and channel estimation. Particle filters as an importance sampling based method provide a closed form solution, and guarantee convergence to the true CFO [127]. We demonstrate via computer simulation that the new method achieves significant performance improvement compared with previous methods.

5) A novel algorithm is proposed for timing and frequency synchronization and channel estimation in the uplink of OFDMA systems under high-mobility situations and considering generalized subcarrier assignment [128]. We adopt an ML approach, use a basis expansion model (BEM) to represent the time-variations of the channel, and apply SAGE algorithm to separate the received signals of different users. We analyze the computational complexities of the new timing and frequency offset estimators and evaluate their performance using MSE indicating a significant performance improvement.

6) We address the problem of joint time varying CFO and doubly selective channel estimation in the uplink of MIMO-OFDMA systems considering a generalized subcarrier assignment scheme [129]. We propose two novel methods using BEM of doubly selective channels and Schmidt Kalman filtering (SKF) [99,130] that are capable of estimating and tracking CFOs and channels. By utilization of SKF, we transform the joint estimation problem for all users into several simpler problems corresponding to each user. The first scheme, which is named "BSEKF", makes use of BEM and Schmidt extended Kalman filtering (SEKF) [99,131] to estimate time varying CFO and channel of each user. Since the extended Kalman filtering is susceptible to potential divergence issues [99,132,133], we

present the second scheme based on marginalized particle filters [134, 135], called "BSK-GPF", in which Gaussian particle filters (GPF) [136, 137] are used to estimate CFO and Schmidt Kalman filters [130] are employed to estimate the BEM coefficients. In comparison with [99] in which channel is assumed constant within an OFDMA symbol, our new schemes have the advantage of coping with channel changes within an OFDMA symbol. Besides, the proposed method based on GPF does not need resampling compared with the particle filtering method in [99], and thus has lower complexity [137]. Compared with the schemes in [114–117], the two proposed schemes benefit from taking the time-variation of CFO in different OFDMA symbol into account, and offer tracking capabilities with closed form solutions. The Bayesian Cramer-Rao bound (BCRB) for the estimation of CFO and doubly selective channel is also derived. It is shown that the new schemes achieve an improved performance in terms of the MSE compared with previous methods.

1.4 Organization of the Thesis

The rest of the thesis is organized as follows. In Chapter 2, characteristics of wireless communication channels along with signal models of OFDM-based systems operating in slow fading frequency selective channels and doubly selective channels are presented. Chapter 3 first deals with the proposed timing synchronization method based on utilization of high order statistics, and next presents a new IFO detection method based on reducing the number of trial values of IFO. Different criteria are exploited including class-separabiltiy, and false alarm and missed detection for timing synchronization and detection probabiltiy and BER for IFO estimation to evaluate the performance of the new methods. The proposed scheme for joint timing, CFO and channel estimation for OFDM systems is developed for doubly selective channels using an ML approach in chapter 4. First, a new approach based on ML criterion along with a reduced complexity estimator is introduced. Then, a new method for joint CFO and doubly selective channel estimation in OFDM systems using particle and Kalman filtering is presented. Next, timing and frequency synchronization

for the uplink of an OFDMA system using SAGE algorithm is discussed. In this chapter, we also address joint CFO and doubly selective channel estimation for MIMO OFDMA systems based on particle and Kalman filtering. Finally, Chapter 5 provides the conclusion and future work.

Chapter 2

OFDM-Based Systems

In this chapter, we explain the structure of a communication system using the OFDM technique. First, we briefly address the characteristics of wireless communications channels, and next present the models of signal transmission and reception over different types of wireless channels.

2.1 Wireless Communication Channels

We consider baseband representation of a wireless communication channel [138]

$$h(t,\tau) = \sum_{l} \gamma_l(t)\delta(\tau - \tau_l) \tag{1}$$

where $\gamma_l(t)$ and τ_l are the complex amplitude and corresponding delay of the lth path. The autocorrelation function of $h(t,\tau)$ is defined as

$$R_h(\Delta t, \tau_1, \tau_2) = \mathrm{E}\{h^*(t,\tau_1)h(t+\Delta t,\tau_2)\} = R_h(\Delta t,\tau_1)\delta(\tau_2 - \tau_1) \tag{2}$$

which is obtained using the assumption that the scattering at two delays is uncorrelated. In general, $R_h(\Delta t,\tau)$ gives the average power output as a function of the time delay τ and the difference Δt in observation time. The range of values of τ over which $R_h(0,\tau)$

is essentially nonzero is called the multipath spread of the channel and is denoted by T_m [139].

A completely analogous characterization of the time-variant multipath channel is obtained by taking the Fourier transform of $h(t, \tau)$ as follows [139]

$$h(t, f) = \int_{-\infty}^{\infty} h(t, \tau) e^{-j2\pi f \tau} d\tau \tag{3}$$

Under the assumption that the channel is wide-sense-stationary, we consider the autocorrelation function

$$R_h(\Delta t, f_2, f_1) = \mathrm{E}\{h^*(t, f_1) h(t + \Delta t, f_2)\} \tag{4}$$

and we have $R_h(\Delta t, f_2, f_1) = R_h(\Delta t, \Delta f) = \int_{-\infty}^{\infty} R_h(\Delta t, \tau_1) e^{-j2\pi f \tau_1} d\tau_1$

Supposeing $\Delta t = 0$ and $R_h(0, f_2, f_1) = R_h(\Delta f)$, we have

$$R_h(\Delta f) = \int_{-\infty}^{\infty} R_h(\tau) e^{-j2\pi \Delta f \tau} d\tau \tag{5}$$

Since $R_h(\Delta f)$ is an autocorrelation function in the frequency variable, it provides a measure of the frequency coherence of the channel. Because of the Fourier transform relationship between $R_h(\Delta f)$ and $R_h(\tau)$, the reciprocal of the multipath spread is a measure of the coherence bandwidth of the channel. That is, $\Delta f \approx \frac{1}{T_m}$. Thus, two sinusoids with frequency separation greater than Δf are affected differently by the channel [139].

The time variations in the channel show a Doppler broadening. In order to relate the Doppler effects to the time variations of the channel, we consider the Fourier transform of $R_h(\Delta t, \Delta f)$ with respect to the variable Δt to be the function [139]

$$S_h(\lambda, \Delta f) = \int_{-\infty}^{\infty} R_h(\Delta t, \Delta f) e^{-j2\pi \lambda \Delta t} d\Delta t \tag{6}$$

becomes

$$S_h(\lambda) = \int_{-\infty}^{\infty} R_h(\Delta t, 0) e^{-j2\pi \lambda \Delta t} d\Delta t \tag{7}$$

gives the signal intensity as a function of the Doppler frequency λ. It is observed that if the channel is time-invariant, $R_h(\Delta t, 0) = R_h(\Delta t) = 1$ and $S_h(\lambda)$ becomes equal to the delta function $\delta(\lambda)$. Therefore, when there are no time variations in the channel, there is no spectral broadening observed in the transmission of a pure frequency tone. The range of values of λ over which $S_h(\lambda)$ is essentially nonzero is called the Doppler spread B_d of the channel [139].

Since $S_h(\lambda)$ is related to $R_h(\Delta t)$ by the Fourier transform, the reciprocal of B_d is a measure of the coherence time of the channel. That is, $(\Delta t)_h \approx \frac{1}{B_d}$ where $(\Delta t)_h$ denotes the coherence time. Clearly, a slowly changing channel has a large coherence time or, equivalently, a small Doppler spread [139].

Based on the above discussion, wireless radio channels can be categorized as

- Flat fading: If a channel has a constant gain and linear phase response over a bandwidth which is greater than the signal bandwidth, then the signal will undergo flat or frequency nonselective fading. In other words, a fading channel is flat or frequency nonselective if the channel coherence bandwidth is greater than the signal bandwidth [140].

- Frequency selective fading: If the channel has a constant gain and a linear phase response over a bandwidth which is smaller than the signal bandwidth, then the signal undergoes frequency selective fading. In other words, a fading channel is frequency selective if the channel coherence bandwidth is smaller than the signal bandwidth [140].

- Slow fading: In a slow fading channel, the channel impulse response changes at a much slower rate than the symbol rate. In other words, the channel coherence time is much greater than the symbol duration, or equivalently, the Doppler spreading is much smaller than the signal bandwidth [140].

- Fast fading: If the channel impulse response changes rapidly within a signal symbol

duration, the channel is classified as a fast fading channel. In other words, when the channel coherence time is smaller than the symbol duration, or equivalently, the Doppler spreading is greater than the signal bandwidth, a signal undergoes fast fading [140].

Taking into account the above categorization of fading channels, in the next two subsections, we discuss the OFDM signal model in slow frequency selective fading channels and fast frequency selective fading channels (doubly selective channels), respectively.

2.2 OFDM-Based Signal Model in Slow Fading Frequency Selective Channels

A packet-based OFDM system is considered in which one or more preambles are transmitted at the start of each frame for the task of synchronization. Fig. 1 depicts the structure of the OFDM system which utilizes N subcarriers. Suppose $S_i(n), 0 \leq n \leq N-1$, denotes the nth element of the ith preamble vector in the frequency domain which will be transmitted on the nth subcarrier, the time domain samples of the ith training symbol can be expressed as

$$s_i(k) = \sum_{n=0}^{N-1} S_i(n) e^{j\frac{2\pi}{N}kn},\ 0 \leq k \leq N-1 \tag{8}$$

After appending the cyclic prefix (CP) of length G (which is greater than the length of the channel impulse response), the signal is transmitted through a multipath channel with L_c taps. The lth tap of the channel impulse response corresponding to the kth sample of the ith OFDM symbol is denoted by $h_i(k,l)$. Here, it is assumed that the channel remains constant within several OFDM symbols, and thus, $h_i(k,l)$ can be simplified to $h(l)$. In view of this, the received signal corresponding to the ith OFDM symbol can be expressed as

$$r_i(k) = e^{j\frac{2\pi}{N}k\epsilon} y_i(k-\theta) + z_i(k), 0 \leq k \leq N-1 \tag{9}$$

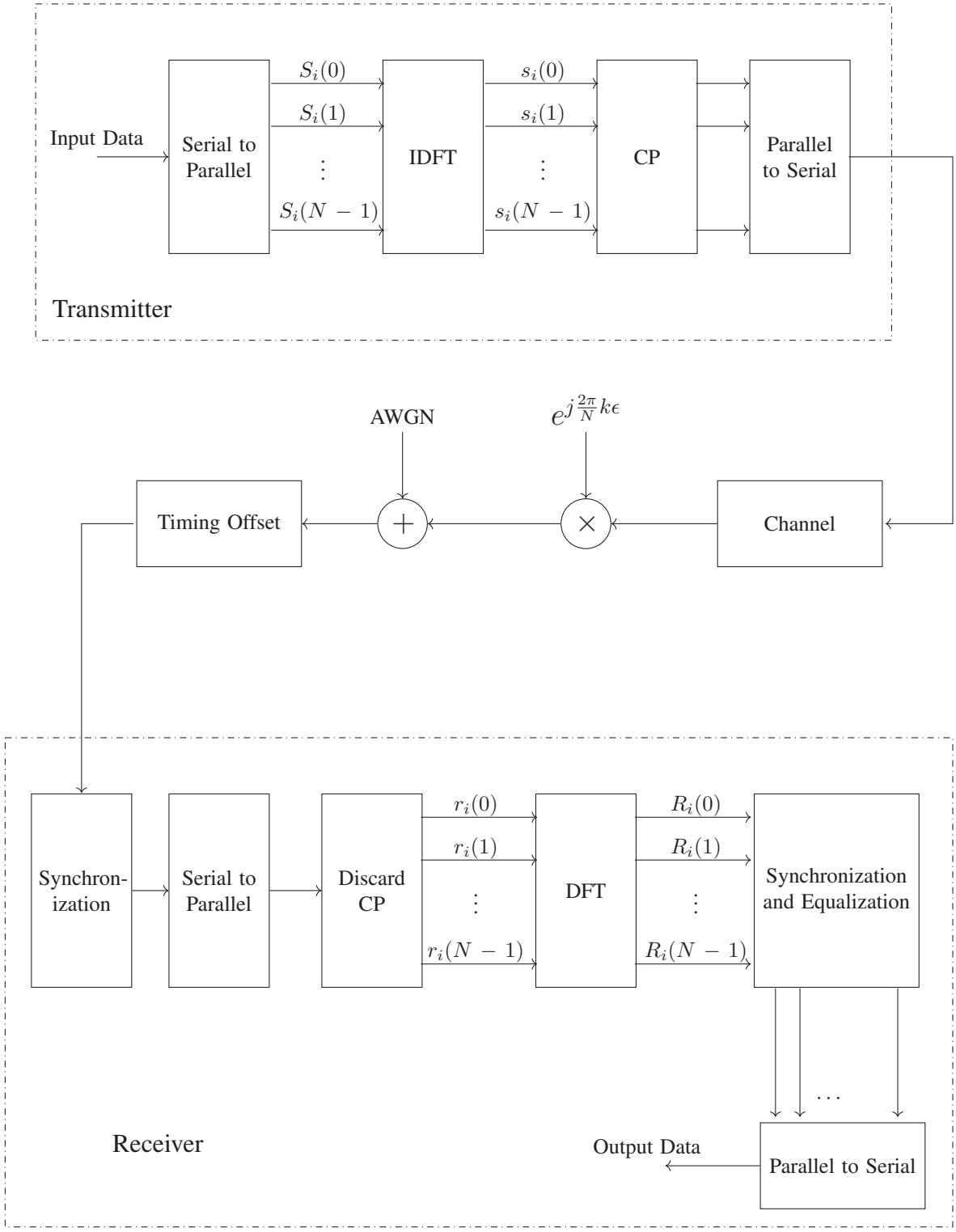

Figure 1: The structure of an OFDM system.

where

$$y_i(k) = \sum_{l=0}^{L_c-1} h(l)s(k-l), \tag{10}$$

and $z_i(k)$ is the sample of additive white Gaussian noise (AWGN) with zero mean and variance σ_z^2. Besides, θ and ϵ are the timing and normalized frequency offsets.

The receiver first estimates the timing and frequency offsets. After correction for these offsets, removing the CP, and taking discrete Fourier transform (DFT), the received signal vector can be expressed in the frequency domain as

$$R_i(n) = H(n)S_i(n) + Z_i(n), \tag{11}$$

where $H(n)$ and $Z_i(n)$ denote the N-point DFTs of $h(l), 0 \leq l \leq L_c - 1$ and $z_i(k),\ 0 \leq k \leq N-1$, respectively.

Note that correction of the timing offset means adjusting the received signal window so that the window contains the ISI free OFDM symbol, and correction of CFO means multiplying the received signal by $e^{-j\frac{2\pi}{N}k\epsilon}$.

2.3 OFDM-Based Signal Model in Doubly Selective Channels

2.3.1 OFDM Systems

In this section, we resort to a matrix representation for the signals. To this end, we equivalently express (8) as

$$\mathbf{s}_i = \mathbf{F}^H \mathbf{S}_i. \tag{12}$$

where, $\mathbf{F}$ is the DFT matrix whose element on the nth row and kth column is $[\mathbf{F}]_{n,k} = \frac{1}{\sqrt{N}}\exp(-j\frac{2\pi}{N}kn)$, $\mathbf{s}_i = [s_i(0), s_i(1), \cdots, s_i(N-1)]^T$, and $\mathbf{S}_i = [S_i(0), S_i(1), \cdots, S_i(N-1)]^T$. After inserting the cyclic prefix (CP), the OFDM symbol is transmitted through an

L_c-tap time-varying channel whose l-th tap ($l = 0, 1, \cdots, L_c - 1$) at time k and for the ith OFDM symbol is denoted by $h_i(k, l)$. In view of this, the samples of the received signal in the time domain can be represented as

$$r_i(k) = e^{j\frac{2\pi}{N}k\epsilon} \sum_{l=0}^{L_c-1} h_i(l, k)s_i(k - l - \theta) + z_i(k),\ 0 \le k \le N - 1 \tag{13}$$

Let the received signal vector of length N be denoted by $\mathbf{r}_i = [r_i(0), r_i(1), \cdots, r_i(N-1)]^T$. Using the above channel model, the received signal vector can be expressed as

$$\mathbf{r}_i = \Gamma(\varepsilon)\check{\mathbf{A}}_i(\theta)\mathbf{h}_i + \mathbf{z}_i, \tag{14}$$

with $\Gamma(\varepsilon) = \text{diag}\{e^{j\frac{2\pi}{N}n\varepsilon_k}, n = G, G+1, \cdots, G+N-1\}$, $\mathbf{h}_i = [\mathbf{h}_{i,0}^T, \mathbf{h}_{i,1}^T, \cdots, \mathbf{h}_{i,L_c-1}^T]^T$, and $\mathbf{h}_{i,l} = [h_i(0, l), h_i(1, l), \cdots, h_i(N-1, l)]^T$. Besides, we have $\check{\mathbf{A}}_i(\theta) = [\mathbf{s}_i^\theta, \mathbf{s}_i^{\theta+1}, \cdots, \mathbf{s}_i^{\theta+L_c-1}]$, and $\mathbf{s}_i^l$ is a diagonal matrix generated by cyclically shifting $\mathbf{s}_i$ by l, namely, $\mathbf{s}_i^l = \text{diag}\big\{s_i(l), s_i(l+1), \cdots, s_i(N-1), s_i(0), s_i(1), \cdots, s_i(l-1)\big\}$.

2.3.2 OFDMA Systems

We consider an OFDMA system with N subcarriers and M simultaneously active users. The block diagram of such an OFDMA system is shown in Fig. 2. The mth user transmits information bits on N_m subcarriers where $N_1 + N_2 + \cdots + N_M = N$. Thus, the data symbols of the mth user and ith OFDM symbol in the frequency domain denoted by $\mathbf{S}_{i,m} = [S_{i,m}(0), S_{i,m}(1), \cdots, S_{i,m}(N-1)]^T$ have only N_m non-zero elements. Taking the IDFT of $\mathbf{S}_{i,m}$, the time-domain signal of the mth user can be expressed as

$$\mathbf{s}_{i,m} = \mathbf{F}^H\mathbf{S}_{i,m}. \tag{15}$$

An OFDMA symbol is then generated by adding cyclic prefix (CP) of length G, and transmitted through a doubly selective channel. Thus, the kth element of the ith OFDM

symbol of the mth user $\mathbf{y}_{i,m}$ at the receiver can be expressed as

$$y_{i,m}(k) = e^{j\frac{2\pi}{N}k\varepsilon_m} \sum_{l=0}^{L_c-1} h_{i,m}(k,l)s_{m,i}(n-l-\theta_m) + z_{i,m}(k), \tag{16}$$

where ε_m and θ_m are the CFO and timing offset of the mth user, respectively, L_c is the number of channel taps, and $h_{i,m}(k,l)$ denotes the lth channel tap gain at the ith OFDMA symbol and time k for the mth user. Furthermore, $z_{i,m}(k)$ indicates the additive white Gaussian noise (AWGN) with zero mean and variance $\sigma^2_{z_m}$. At the receiver front end, the received signal is the superposition of the signals of all users, which can be written in the vector form as

$$\mathbf{r}_i = \sum_{m=1}^{M} \mathbf{y}_{i,m}, \tag{17}$$

where

$$\mathbf{y}_{i,m} = \Gamma(\varepsilon_m)\check{\mathbf{A}}_{i,m}(\theta_m)\mathbf{h}_{i,m} + \mathbf{z}_{i,m}, \tag{18}$$

with $\Gamma(\varepsilon_m) = \text{diag}\{e^{j\frac{2\pi}{N}n\varepsilon_m}, n = G, G+1, \cdots, G+N-1\}$, $\mathbf{h}_{i,m} = [\mathbf{h}^T_{i,m,0}, \mathbf{h}^T_{i,m,1}, \cdots, \mathbf{h}^T_{i,m,L_c-1}]^T$, and $\mathbf{h}_{i,m,l} = [h_{i,m}(0,l), h_{i,m}(1,l), \cdots, h_{i,m}(N-1,l)]^T$. Besides, we have $\check{\mathbf{A}}_{i,m}(\theta_m) = [\mathbf{S}^{\theta_m}_{i,m}, \mathbf{S}^{\theta_m+1}_{i,m}, \cdots, \mathbf{S}^{\theta_m+L_c-1}_{i,m}]$, and $\mathbf{S}^l_k$ is a diagonal matrix generated by cyclically shifting $\mathbf{S}_{i,m}$ by l, namely, $\mathbf{S}^l_{i,m} = \text{diag}\big\{S_{i,m}(l), S_{i,m}(l+1), \cdots, S_{i,m}(N-1), S_{i,m}(0), S_{i,m}(1), \cdots, S_{i,m}(l-1)\big\}$.

2.4 Conclusion

In this section, different types of wireless communication channels have been first discussed. Next the signal model of an OFDM system in slow fading frequency selective channels and doubly selective channels were briefly explained, and block diagrams of the transmitter and receiver were depicted. In addition, the signal model of an OFDMA system operating in a doubly selective channel was explained, and components of the transmitter and receiver were shown in a block diagram.

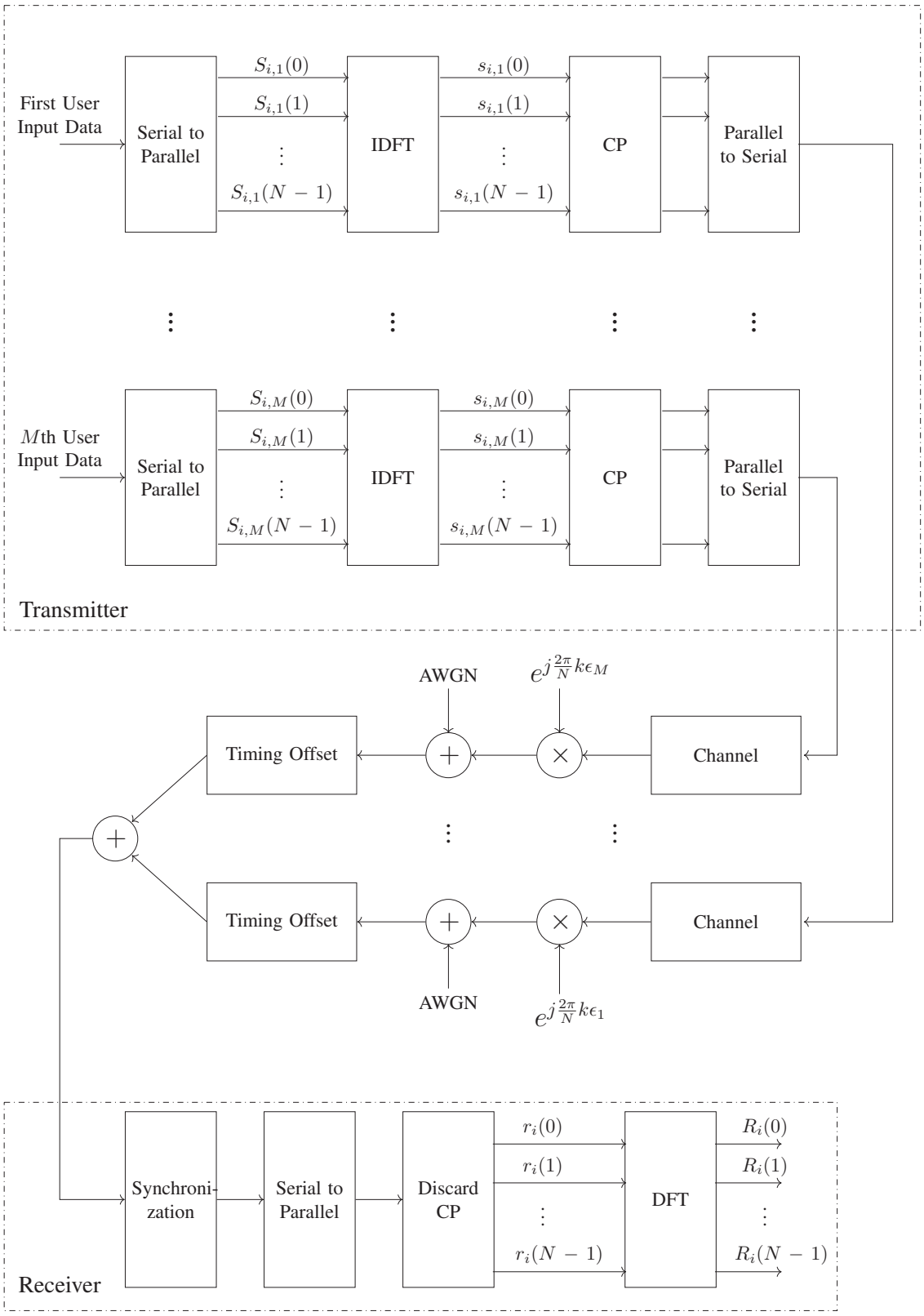

Figure 2: The structure of an OFDMA system.

Chapter 3

Synchronization in Slow Fading Frequency Selective Channels

3.1 Introduction

In this chapter, we present two new schemes for synchronization in slow fading frequency selective channels. The first method is developed for coarse timing synchronization using high order statistics. The performance of this method is evaluated using different criteria such as class-separability, false alarm and missed-detection probabilities, and robustness to CFO. The second method deals with a reduced complexity IFO detection scheme which is based on reducing the number of trial values of IFO. Computational complexity, probability of correct detection, and BER of the new method is finally compared with previous schemes.

3.2 Proposed Method: Timing Synchronization

3.2.1 Preliminaries

Coarse timing synchronization is to determine a coarse estimate of the start of the frame [1] or equivalently, determining the correct position of the preamble. It usually consists of two stages: 1) frame/preamble detection and 2) coarse timing estimation. The methods that deal with timing synchronization usually perform both of the stages using a timing metric. When the value of the timing metric goes above a predefined threshold the preamble is detected. Subsequently, the position of the maximum value of the timing metric is used to obtain a coarse estimate of the start of the frame.

In [31], the authors proposed the utilization of a preamble that is composed of two identical parts in the time domain as described by

$$s(k) = s(k + (N/2)), 0 \leq k \leq (N/2) - 1. \tag{19}$$

As long as the length of each identical part of the preamble is greater than the channel delay spread, the two identical parts of the preamble remain identical after passing through a multipath fading channel i.e.,

$$y(k) = y(k + \frac{N}{2}), 0 \leq k \leq (N/2) - 1. \tag{20}$$

Then, the received signal vector of length N at time instant d can be written as

$$\mathbf{r} = \Big[\underbrace{r(d), r(d+1), \ldots, r(d+\frac{N}{2}-1)}_{\mathbf{r}^d}, \underbrace{r(d+\frac{N}{2}), r(d+\frac{N}{2}+1), \ldots, r(d+N-1)}_{\mathbf{r}^{d+\frac{N}{2}}}\Big]^T. \tag{21}$$

By taking advantage of the correlation between the two identical parts of the preamble,

the authors of [31] proposed a timing metric as follows

$$M_{Sc}(d) = \frac{|P_{Sc}(d)|^2}{(R_{Sc}(d))^2}, \tag{22}$$

where

$$P_{Sc}(d) = \sum_{k=0}^{(N/2)-1} \overline{r}(d+k) r(d+k+(N/2)), \tag{23}$$

is the correlation function, $\overline{r}(k)$ denotes the complex conjugate of $r(k)$, and

$$R_{Sc}(d) = \sum_{k=0}^{(N/2)-1} |r\left(d+k+(N/2)\right)|^2, \tag{24}$$

is the normalization function.

In [14], an equivalent form of the metric (22) was presented and extended to the case of more than two identical parts and more than one preamble. It was shown that when the periodic parts of more than one preamble are used, the probability of false alarm improves. The idea of using differential normalization was introduced in [38]. It was shown that when the normalization function (24) is replaced by the magnitude of difference (MoD) or difference of magnitude (DoM) normalization functions

$$R_{MoD}(d) = \sum_{k=0}^{(N/2)-1} |r\left(d+k\right) - r\left(d+k+(N/2)\right)|^2, \tag{25}$$

$$R_{DoM}(d) = \sum_{k=0}^{(N/2)-1} \left(|r\left(d+k\right)| - |r\left(d+k+(N/2)\right)|\right)^2, \tag{26}$$

as in

$$M_{MoD}(d) = \frac{|P_{Sc}(d)|}{R_{MoD}(d)}, \tag{27}$$

$$M_{DoM}(d) = \frac{|P_{Sc}(d)|}{R_{DoM}(d)}, \tag{28}$$

then the probabilities of false alarm and missed detection would be reduced. This improvement was justified by the fact that at the correct timing points the correlation function

reaches a peak while the normalization function has its lowest value. Therefore, the differential timing metrics give greater peaks at the correct timing points and consequently have values at the correct and wrong timing points that are more distinct compared with previous methods.

It has been shown in [39] that sufficient statistics for the detection of a preamble composed of two identical parts does not exist, and the second order statistics used in previous methods are not optimal in terms of utilizing sufficient statistics. They also proposed fourth order statistics for frame detection. The fourth order correlation function was expressed as

$$P_4(d) = \sum_{k=0}^{L-1} \overline{B}^{r1,d}(k) B^{r1,d+\frac{N}{2}}(k), \tag{29}$$

where $\mathbf{B}^{r1,d}$ is a vector composed of subvectors whose elements are second order products expressed by

$$\begin{aligned}\mathbf{B}^{r1,d} = \{&\mathbf{B}_1^{r1,d}, \mathbf{B}_2^{r1,d}, \cdots, \mathbf{B}_{\frac{N}{2}-1}^{r1,d}\} \\ &\underbrace{\{r(d)r(d+1), r(d+1)r(d+2), \cdots, r(d+\frac{N}{2}-2)r(d+\frac{N}{2}-1)}_{\mathbf{B}_1^{r1,d}}, \\ &\underbrace{r(d)r(d+2), r(d+1)r(d+3), \cdots, r(d+\frac{N}{2}-3)r(d+\frac{N}{2}-1)}_{\mathbf{B}_2^{r1,d}}, \\ &\cdots, \underbrace{r(d)r(d+\frac{N}{2}-1)}_{\mathbf{B}_{\frac{N}{2}-1}^{r1,d}}\}.\end{aligned} \tag{30}$$

The normalization function for this correlation function is defined as

$$R_4(d) = \sum_{k=0}^{L-1} |B^{r1,d}(k)|^2 + \sum_{k=0}^{L-1} |B^{r1,d+\frac{N}{2}}(k)|^2, \tag{31}$$

resulting in the following timing metric

$$M_4(d) = 2\frac{|P_4(d)|}{R_4(d)}. \tag{32}$$

Corresponding to (29), two fourth order differential normalization functions referred to as double differential normalization (DDN) and CFO independent double differential normalization (CIDDN) are recently presented in [40], respectively, as

$$R^{DDN}(d) = \sum_{k=0}^{L-1} \left(Q^{DDN,d}(k)\right)^2. \tag{33}$$

$$R^{CIDDN}(d) = \sum_{k=0}^{L-1} \left(Q^{CIDDN,d}(k)\right)^2. \tag{34}$$

where

$$\begin{aligned}\mathbf{Q}^{DDN,d} = \{&|r(d) - r(d+\frac{N}{2})|.|r(d+1) - r(d+1+\frac{N}{2})|,\\ &|r(d) - r(d+\frac{N}{2})|.|r(d+2) - r(d+2+\frac{N}{2})|,\\ &\cdots, |r(d) - r(d+\frac{N}{2})|.|r(d+\frac{N}{2}-1) - r(d+N-1)|,\\ &|r(d+1) - r(d+1+\frac{N}{2})|.|r(d+2) - r(d+2+\frac{N}{2})|,\\ &|r(d+1) - r(d+1+\frac{N}{2})|.|r(d+3) - r(d+3+\frac{N}{2})|,\\ &\cdots, |r(d+1) - r(d+1+\frac{N}{2})|.|r(d+\frac{N}{2}-1) - r(d+N-1)|,\\ &\cdots, |r(d+\frac{N}{2}-2) - r(d+N-2)|.|r(d+\frac{N}{2}-1) - r(d+N-1)|\}\end{aligned} \tag{35}$$

and

$$\begin{aligned}\mathbf{Q}^{CIDDN,d} = \{&(|r(d)| - |r(d+\frac{N}{2})|)(|r(d+1)| - |r(d+1+\frac{N}{2})|),\\ &(|r(d)| - |r(d+\frac{N}{2})|)(|r(d+2)| - |r(d+2+\frac{N}{2})|),\\ &\cdots,(|r(d)| - |r(d+\frac{N}{2})|)(|r(d+\frac{N}{2}-1)| - |r(d+N-1)|),\\ &(|r(d+1)| - |r(d+1+\frac{N}{2})|)(|r(d+2)| - |r(d+2+\frac{N}{2})|),\\ &(|r(d+1)| - |r(d+1+\frac{N}{2})|)(|r(d+3)| - |r(d+3+\frac{N}{2})|),\\ &\cdots,(|r(d+1)| - |r(d+1+\frac{N}{2})|)(|r(d+\frac{N}{2}-1)| - |r(d+N-1)|),\\ &\cdots,(|r(d+\frac{N}{2}-2)| - |r(d+N-2)|)(|r(d+\frac{N}{2}-1)| - |r(d+N-1)|)\} \quad (36)\end{aligned}$$

It is evident that elements of $\mathbf{Q}^{DDN,d}$ and $\mathbf{Q}^{CIDDN,d}$ are second order products. The elements of $\mathbf{Q}^{DDN,d}$ are generated using the magnitude of difference of the samples of the received signal, whereas the elements of $\mathbf{Q}^{CIDDN,d}$ are formed using difference of the magnitude of the samples of the received signal. It was shown in [40] that these normalization functions can considerably increase the difference between the values of the timing metric at correct and wrong timing instants and therefore,can considerably reduce the probability of missed-detection. The timing metric with the fourth order correlation and differential normalization functions can be expressed as [40]

$$M_{DDN}(d) = \frac{|P_4(d)|}{R^{DDN}(d)}, \quad (37)$$

$$M_{CIDDN}(d) = \frac{|P_4(d)|}{R^{CIDDN}(d)}. \quad (38)$$

3.2.2 nth Order Timing Metric

Here, with the aim of making the timing metric values more distinctive at the correct and wrong timing instants, we propose to use higher order statistics for timing synchronization.

Consider the received signal vector of length N at time instant d

$$\mathbf{r} \;=\; \Big[\underbrace{r(d), r(d+1), \ldots, r(d+\frac{N}{2}-1)}_{\mathbf{r}^d}, \underbrace{r(d+\frac{N}{2}), r(d+\frac{N}{2}+1), \ldots, r(d+N-1)}_{\mathbf{r}^{d+\frac{N}{2}}}\Big]^T. \tag{39}$$

where $\mathbf{r}$ has been broken into two subvectors $\mathbf{r}^d$ and $\mathbf{r}^{d+\frac{N}{2}}$ of length $\frac{N}{2}$. Note that $r(d+k)$ is the kth element of the vector $\mathbf{r}$ at timing instant d, and $\mathbf{r}^d$ and $\mathbf{r}^{d+\frac{N}{2}}$ are the vectors containing the first and second $N/2$ elements of the vector $\mathbf{r}$, respectively. An n-th order product (where n is even) from the received signal vector $\mathbf{r}$ can be generated by multiplication of the complex conjugated version of an $n/2$-th order product from $\mathbf{r}^d$ with another $n/2$-th order product from $\mathbf{r}^{d+\frac{N}{2}}$. As an example, an n-th order product can be expressed as

$$\underbrace{\overline{r}(d)\overline{r}(d+1)\ldots\overline{r}(d+\frac{n}{2}-1)}_{\text{from } \mathbf{r}^d}\;\underbrace{r(d+\frac{N}{2})r(d+\frac{N}{2}+1)\ldots r(d+\frac{N}{2}+\frac{n}{2}-1)}_{\text{from } \mathbf{r}^{d+\frac{N}{2}}}. \tag{40}$$

Note that the samples of the received preamble that are $\frac{N}{2}$ samples apart are highly correlated. To take advantage of this correlation, for each element of $\mathbf{r}^d$ used in the product there should be an element from $\mathbf{r}^{d+\frac{N}{2}}$ with the distance of $\frac{N}{2}$ samples, i.e., if $r(d+k)$ is used, then $r(d+k+\frac{N}{2})$ should also be used.

Another form of n-th order product can be generated by applying the amplitude of the difference between elements of $\mathbf{r}^d$ and the corresponding elements of $\mathbf{r}^{d+\frac{N}{2}}$ that are $\frac{N}{2}$ samples apart. For example, we can have the following product

$$\Big|r(d)-r(d+\frac{N}{2})\Big|^2\Big|r(d+1)-r(d+1+\frac{N}{2})\Big|^2\ldots\Big|r(d+\frac{n}{2}-1)-r(d+\frac{N}{2}+\frac{n}{2}-1)\Big|^2. \tag{41}$$

We could also generate another n-th order product by using an idea similar to that in

(41), but based on the amplitudes of the elements of $\mathbf{r}^d$ and $\mathbf{r}^{d+\frac{N}{2}}$ such as

$$\Big(\big|r(d)\big| - \big|r(d+\frac{N}{2})\big|\Big)^2\Big(\big|r(d+1)\big| - \big|r(d+1+\frac{N}{2})\big|\Big)^2 \ldots \\ \times \Big(\big|r(d+\frac{n}{2}-1)\big| - \big|r(d+\frac{N}{2}+\frac{n}{2}-1)\big|\Big)^2. \quad (42)$$

Taking into account the above three forms of n-th order products, we discuss in what follows how these products are utilized to compute the new timing metrics.

First, we consider the n-th order correlation function. A vector whose elements are $\frac{n}{2}$th order products can be generated by element-wise multiplying $n/2$ vectors whose elements are first order terms. Here, we generate these $n/2$ vectors by cyclically shifting the vector $\mathbf{r}^d$ with different amounts of cyclic shifts. When the $\frac{n}{2}$th order vector thus generated is element-wise multiplied by another $\frac{n}{2}$th order vector that is generated in the same way but from the vector $\mathbf{r}^{d+\frac{N}{2}}$, a vector whose elements are n-th order products is then produced.

Let $\mathbf{r}^d_{u_m}$ be a vector obtained by cyclically shifting the elements of $\mathbf{r}^d$ in (39) in the base of $\frac{N}{2}$ by u_m samples i.e.,

$$\mathbf{r}^d_{u_m} = \Big[r(d+u_m), r(d+u_m+1), \ldots, r(d+\frac{N}{2}-1), r(d), r(d+1), \ldots, r(d+u_m-1)\Big]^T \quad (43)$$

Then, the following multiplication

$$\overline{\mathbf{r}}^d_{u_1} \circ \overline{\mathbf{r}}^d_{u_2} \circ \cdots \circ \overline{\mathbf{r}}^d_{u_{\frac{n}{2}}} \quad (44)$$

gives a vector whose elements are $\frac{n}{2}$-th order products, where $\circ$ denotes the Hadamard product (the element by element multiplication of the vectors), and $\overline{\mathbf{r}}^d_{u_1}$ is the element-wise complex conjugated version of $\mathbf{r}^d_{u_1}$. It is obvious that the cyclic shifts $(u_1, u_2, \cdots, u_{\frac{n}{2}})$ can have various combinations. In fact, we have a total of $(\frac{N}{2})^{\frac{n}{2}}$ different sets of cyclic shifts, and each of these cyclic shifts changes the resulting vector in (44). Therefore, we also have $(\frac{N}{2})^{\frac{n}{2}}$ different forms of (44). The k-th form of the vector (44) can be expressed as

$\overline{\mathbf{r}}^d_{u_{1,k}} \circ \overline{\mathbf{r}}^d_{u_{2,k}} \circ \cdots \circ \overline{\mathbf{r}}^d_{u_{\frac{n}{2},k}}$, which corresponds to the k-th set of cyclic shifts $(u_{1,k}, u_{2,k}, \cdots, u_{\frac{n}{2},k})$. For future reference, it is important to redefine (43) as

$$\mathbf{r}^d_{u_{m,k}} = \Big[r(d+u_{m,k}), r(d+u_{m,k}+1), \ldots, r(d+\frac{N}{2}-1), r(d), r(d+1), \ldots, r(d+u_{m,k}-1)\Big]^T \tag{45}$$

where according to the discussion above, $u_{m,k}$ with $0 \leq u_{m,k} \leq \frac{N}{2}-1$ denotes the value of the cyclic shift, m, $(1 \leq m \leq \frac{n}{2})$ indicates the index of the vector in a set of cyclic shifts, and k, $(0 \leq k \leq (\frac{N}{2})^{\frac{n}{2}}-1)$, denotes the set of cyclic shifts.

If we generate another vector whose elements are $\frac{n}{2}$-th order products similar to $\overline{\mathbf{r}}^d_{u_{1,k}} \circ \overline{\mathbf{r}}^d_{u_{2,k}} \circ \cdots \circ \overline{\mathbf{r}}^d_{u_{\frac{n}{2},k}}$, but from the vector $\mathbf{r}^{d+\frac{N}{2}}$, and apply the Hadamard product to them, we have the following vector whose elements are n-th order products

$$\mathbf{\Lambda}^d_k = \overline{\mathbf{r}}^d_{u_{1,k}} \circ \overline{\mathbf{r}}^d_{u_{2,k}} \circ \cdots \circ \overline{\mathbf{r}}^d_{u_{\frac{n}{2},k}} \circ \mathbf{r}^{d+\frac{N}{2}}_{u_{1,k}} \circ \mathbf{r}^{d+\frac{N}{2}}_{u_{2,k}} \circ \cdots \circ \mathbf{r}^{d+\frac{N}{2}}_{u_{\frac{n}{2},k}}. \tag{46}$$

Note that the vector $\mathbf{\Lambda}^d_k$ contains nth order products generated corresponding to the kth set of cyclic shifts $(u_{1,k}, u_{2,k}, \cdots, u_{\frac{n}{2},k})$, and as mentioned before we have a total of $(\frac{N}{2})^{\frac{n}{2}}$ different sets of cyclic shifts resulting in $(\frac{N}{2})^{\frac{n}{2}}$ different forms of $\mathbf{\Lambda}^d_k$. Out of these $(\frac{N}{2})^{\frac{n}{2}}$ forms, q forms of $\mathbf{\Lambda}^d_k$ are put in a vecotor $\mathbf{\Lambda}^d$ as follows:

$$\mathbf{\Lambda}^d = \Big[(\mathbf{\Lambda}^d_0)^T, (\mathbf{\Lambda}^d_1)^T, \cdots, (\mathbf{\Lambda}^d_{q-1})^T\Big]^T, \ 1 \leq q \leq (\frac{N}{2})^{\frac{n}{2}} \tag{47}$$

where $(\mathbf{\Lambda}^d_k)^T$ denotes the transpose of $\mathbf{\Lambda}^d_k$, and the number of elements of $\mathbf{\Lambda}^d$ is $L = \frac{N}{2}q$. Furthermore, q is a design parameter that determines the correlation length and is designed based on the expected performance and affordable complexity as will be discussed later. Then, we add elements of $\mathbf{\Lambda}^d$ in order to generate the n-th order correlation function as follows

$$P_n(d) = \sum_{m=0}^{L-1} \Lambda^d(m) = \sum_{l=0}^{\frac{N}{2}-1} \sum_{k=0}^{q-1} \Lambda^d_k(l) \tag{48}$$

where $\Lambda^d(m)$ is the m-th element of Λ^d, and $L = \frac{N}{2}q$ is the correlation length.

One can look at the generation of $P_n(d)$ from another point of view. We know from (45) that the elements of $\mathbf{r}^d_{u_{m,k}}$ can be written as

$$r^d_{u_{m,k}}(l) = r\Big(d + |l + u_{m,k}|_{\frac{N}{2}}\Big),\ 0 \le l \le \frac{N}{2} - 1 \tag{49}$$

where $|\cdot|_{\frac{N}{2}}$ denotes the modulo $N/2$ operation. Obviously, a similar relation for elements of $\mathbf{r}^{d+\frac{N}{2}}_{u_{m,k}}$ can be obtained by replacing d in (49) with $d + \frac{N}{2}$. Thus, the elements of $\mathbf{\Lambda}^d_k$ in (46) can be expressed as

$$\begin{aligned}\Lambda^d_k(l) = {} & \overline{r}\Big(d + |l + u_{1,k}|_{\frac{N}{2}}\Big)\overline{r}\Big(d + |l + u_{2,k}|_{\frac{N}{2}}\Big)\cdots\overline{r}\Big(d + |l + u_{\frac{n}{2},k}|_{\frac{N}{2}}\Big) \\ & \times r\Big(d + \frac{N}{2} + |l + u_{1,k}|_{\frac{N}{2}}\Big)r\Big(d + \frac{N}{2} + |l + u_{2,k}|_{\frac{N}{2}}\Big)\cdots r\Big(d + \frac{N}{2} + |l + u_{\frac{n}{2},k}|_{\frac{N}{2}}\Big).\end{aligned} \tag{50}$$

By summing all the elements $\Lambda^d_k(l)$ resulting from different values of cyclic shifts $u_{m,k}$, we can equivalently express $P_n(d)$ as

$$\begin{aligned}P_n(d) = {} & \sum_{l=0}^{\frac{N}{2}-1}\sum_{u_{1,k}=0}^{U_1-1}\sum_{u_{2,k}=0}^{U_2-1}\cdots\sum_{u_{\frac{n}{2},k}=0}^{U_{\frac{n}{2}}-1}\overline{r}\Big(d+|l+u_{1,k}|_{\frac{N}{2}}\Big)\overline{r}\Big(d+|l+u_{2,k}|_{\frac{N}{2}}\Big)\cdots\overline{r}\Big(d+|l+u_{\frac{n}{2},k}|_{\frac{N}{2}}\Big) \\ & \times r\Big(d + |l + u_{1,k}|_{\frac{N}{2}} + \frac{N}{2}\Big)r\Big(d + |l + u_{2,k}|_{\frac{N}{2}} + \frac{N}{2}\Big)\cdots r\Big(d + |l + u_{\frac{n}{2},k}|_{\frac{N}{2}} + \frac{N}{2}\Big)\end{aligned} \tag{51}$$

where $0 \le U_1, U_2, \ldots, U_{\frac{n}{2}} \le \frac{N}{2} - 1$ indicate the range of the cyclic shifts. Note that the correlation length is $L = \frac{N}{2}U_1U_2\cdots U_{\frac{n}{2}} = \frac{N}{2}q$.

Similar to the n-th order correlation function, we can define the n-th order normalization function. The difference is that we employ n-th order products that are similar to (41), rather than (40), to generate the normalization function. In other words, instead of using the multiplication of the correlated parts (as performed for the correlation function), we apply the difference of the correlated parts for the normalization function. For example, we use elements $\Big|r(d+k) - r(d+k+\frac{N}{2})\Big|^2,\ 0 \le k \le \frac{N}{2} - 1$, which can be written in the vector form $\Big|\mathbf{r}^d - \mathbf{r}^{d+\frac{N}{2}}\Big|^2$ where $|\cdot|^2$ means the square of the absolute value of each

element of the vector i.e.

$$\left|\mathbf{r}^d - \mathbf{r}^{d+\frac{N}{2}}\right|^2 = \Bigg[\left|r(d+k) - r(d+k+\frac{N}{2})\right|^2, \left|r(d+k+1) - r(d+k+\frac{N}{2}+1)\right|^2, \cdots, \left|r(d+k+\frac{N}{2}-1) - r(d+k+N-1)\right|^2\Bigg]^T. \quad (52)$$

Evidently, elements of (52) are second order products. To generate the nth order products, we element-wise multiply the $n/2$ vectors that are cyclicly shifted versions of (52), namely,

$$\mathbf{\Psi}_k^d = \left|\mathbf{r}_{u_{1,k}}^d - \mathbf{r}_{u_{1,k}}^{d+\frac{N}{2}}\right|^2 \circ \left|\mathbf{r}_{u_{2,k}}^d - \mathbf{r}_{u_{2,k}}^{d+\frac{N}{2}}\right|^2 \circ \cdots \circ \left|\mathbf{r}_{u_{\frac{n}{2},k}}^d - \mathbf{r}_{u_{\frac{n}{2},k}}^{d+\frac{N}{2}}\right|^2 \quad (53)$$

where $\mathbf{r}_{u_{m,k}}^d$ is defined in (45). As mentioned before, there are $(\frac{N}{2})^{\frac{n}{2}}$ possible sets of cyclic shifts which correspond to different values of k resulting in different $\mathbf{\Psi}_k^d$. We put q of these different $\mathbf{\Psi}_k^d$ in the vector

$$\mathbf{\Psi}^d = \left[(\mathbf{\Psi}_0^d)^T, (\mathbf{\Psi}_1^d)^T, \cdots, (\mathbf{\Psi}_{q-1}^d)^T\right]^T,\ 1 \le q \le (\frac{N}{2})^{\frac{n}{2}}. \quad (54)$$

The differential normalization function is then defined as the sum of the elements of Ψ^d, i.e.,

$$\Delta_n^{DC}(d) = \sum_{m=0}^{L-1} \Psi^d(m) = \sum_{l=0}^{\frac{N}{2}-1} \sum_{k=0}^{q-1} \Psi_k^d(l). \quad (55)$$

To gain further insight into how the new normalization function is generated, using (49), we get the following expression for the elements of $\mathbf{\Psi}_k^d$ as

$$\Psi_k^d(l) = \left|r\big(d+|l+u_{1,k}|_{\frac{N}{2}}\big) - r\big(d+\frac{N}{2}+|l+u_{1,k}|_{\frac{N}{2}}\big)\right|^2 \left|r\big(d+|l+u_{2,k}|_{\frac{N}{2}}\big) - r\big(d+\frac{N}{2}+|l+u_{2,k}|_{\frac{N}{2}}\big)\right|^2 \cdots \times \left|r\big(d+|l+u_{\frac{n}{2},k}|_{\frac{N}{2}}\big) - r\big(d+\frac{N}{2}+|l+u_{\frac{n}{2},k}|_{\frac{N}{2}}\big)\right|^2. \quad (56)$$

Thus, $\Delta_n^{DC}(d)$ in (55), summing nth order products corresponding to different values of

cyclic shifts, can be equivalently written as

$$\Delta_n^{DC}(d) = \sum_{l=0}^{\frac{N}{2}-1} \sum_{u_{1,k}=0}^{U_1-1} \sum_{u_{2,k}=0}^{U_2-1} \cdots \sum_{u_{\frac{n}{2},k}=0}^{U_{\frac{n}{2}}-1} \left| r\big(d + |l + u_{1,k}|_{\frac{N}{2}}\big) - r\big(d + |l + u_{1,k}|_{\frac{N}{2}} + \frac{N}{2}\big) \right|^2$$
$$\times \left| r\big(d+|l+u_{2,k}|_{\frac{N}{2}}\big) - r\big(d+|l+u_{2,k}|_{\frac{N}{2}}+\frac{N}{2}\big) \right|^2 \ldots \left| r\big(d+|l+u_{\frac{n}{2},k}|_{\frac{N}{2}}\big) - r\big(d+|l+u_{\frac{n}{2},k}|_{\frac{N}{2}}+\frac{N}{2}\big) \right|^2. \tag{57}$$

Finally, the proposed timing metric is defined as

$$M_n^{DC}(d) = \frac{|P_n(d)|}{\Delta_n^{DC}(d)}. \tag{58}$$

It is shown in Appendix A that this timing metric has the following means at the wrong timing instant $\widetilde{d}$ (when none of the elements of the received signal vector $\mathbf{r}$ belongs to the preamble) and correct timing instant Θ (when all the samples of the received signal vector $\mathbf{r}$ belong to the preamble)

$$E\{M_n^{DC}(\widetilde{d})\} = \sqrt{\frac{\pi}{2^{n+2}L}}, \tag{59}$$

$$E\{M_n^{DC}(\Theta)\} = \left(\frac{SNR}{|1 - e^{j\pi\varepsilon}|^2 SNR + 2} \right)^{\frac{n}{2}}. \tag{60}$$

It is evident from (60) that the mean of the proposed metric at the correct timing point depends on CFO, which is a consequence of the dependence of the normalization function on CFO. To make the normalization function robust to CFO, instead of applying the magnitude of the difference of each pair in (52) or (53), we propose to use the difference of the magnitudes from each correlated pair, namely, we modify (53), (54) and (55), respectively, as

$$\hat{\mathbf{\Psi}}_k^d = \left(\left|\mathbf{r}_{u_{1,k}}^d\right| - \left|\mathbf{r}_{u_{1,k}}^{d+\frac{N}{2}}\right| \right)^2 \circ \left(\left|\mathbf{r}_{u_{2,k}}^d\right| - \left|\mathbf{r}_{u_{2,k}}^{d+\frac{N}{2}}\right| \right)^2 \circ \cdots \circ \left(\left|\mathbf{r}_{u_{\frac{n}{2},k}}^d\right| - \left|\mathbf{r}_{u_{\frac{n}{2},k}}^{d+\frac{N}{2}}\right| \right)^2 \tag{61}$$

$$\hat{\mathbf{\Psi}}^d = \left[(\hat{\mathbf{\Psi}}_0^d)^T, (\hat{\mathbf{\Psi}}_1^d)^T, \cdots, (\hat{\mathbf{\Psi}}_{q-1}^d)^T \right]^T, \ 1 \le q \le (\frac{N}{2})^{\frac{n}{2}} \tag{62}$$

$$\Delta_n^{IC}(d) = \sum_{m=0}^{L-1} \hat{\Psi}^d(m) = \sum_{l=0}^{\frac{N}{2}-1} \sum_{k=0}^{q-1} \hat{\Psi}_k^d(l) \tag{63}$$

where $\Delta_n^{IC}(d)$ is the n-th order differential normalization function that is independent of CFO. Considering

$$\begin{aligned}\hat{\Psi}_k^d(l) = &\left(\left| r\big(d + |l + u_{1,k}|_{\frac{N}{2}}\big) \right| - \left| r\big(d + \frac{N}{2} + |l + u_{1,k}|_{\frac{N}{2}}\big) \right| \right)^2 \\ &\times \left(\left| r\big(d + |l + u_{2,k}|_{\frac{N}{2}}\big) \right| - \left| r\big(d + \frac{N}{2} + |l + u_{2,k}|_{\frac{N}{2}}\big) \right| \right)^2 \\ &\cdots \times \left(\left| r\big(d + |l + u_{\frac{n}{2},k}|_{\frac{N}{2}}\big) \right| - \left| r\big(d + \frac{N}{2} + |l + u_{\frac{n}{2},k}|_{\frac{N}{2}}\big) \right| \right)^2, \end{aligned} \tag{64}$$

$\Delta_n^{IC}(d)$ can be equivalently expressed as

$$\begin{aligned}\Delta_n^{IC}(d) = &\sum_{l=0}^{\frac{N}{2}-1} \sum_{u_{1,k}=0}^{U_1-1} \sum_{u_{2,k}=0}^{U_2-1} \cdots \sum_{u_{\frac{n}{2},k}=0}^{U_{n/2}-1} \left(\left| r\big(d + |l + u_{1,k}|_{\frac{N}{2}}\big) \right| - \left| r\big(d + |l + u_{1,k}|_{\frac{N}{2}} + \frac{N}{2}\big) \right| \right)^2 \\ &\times \left(\left| r\big(d + |l + u_{2,k}|_{\frac{N}{2}}\big) \right| - \left| r\big(d + |l + u_{2,k}|_{\frac{N}{2}} + \frac{N}{2}\big) \right| \right)^2 \cdots \\ &\times \left(\left| r\big(d + |l + u_{\frac{n}{2},k}|_{\frac{N}{2}}\big) \right| - \left| r\big(d + |l + u_{\frac{n}{2},k}|_{\frac{N}{2}} + \frac{N}{2}\big) \right| \right)^2. \end{aligned} \tag{65}$$

Finally, the n-th order timing metric that works independent of CFO is proposed as

$$M_n^{IC}(d) = \frac{|P_n(d)|}{\Delta_n^{IC}(d)}. \tag{66}$$

In the next section, we show that this metric is not affected by CFO. In Appendix B, we have obtained the means of this timing metric at the wrong and correct timing instants as given by

$$E\{M_n^{IC}(\widetilde{d})\} = \frac{1}{2(2-\frac{\pi}{2})^{\frac{n}{2}}} \sqrt{\frac{\pi}{L}}, \tag{67}$$

$$E\{M_n^{IC}(\Theta)\} \geq \left(\frac{SNR}{|1-e^{j\pi\varepsilon}|^2 SNR+2}\right)^{\frac{n}{2}}. \tag{68}$$

Fig. 3 illustrates the structure of an OFDM frame and the received vectors $\mathbf{r}^d$ and $\mathbf{r}^{d+\frac{N}{2}}$ corresponding to a wrong timing instant $d=\widetilde{d}$ and a correct timing instant $d=\Theta$. In Fig 4, we illustrate as to how the new correlation function $P_n(d)$ can be generated. First, $\mathbf{\Lambda}_k^d$ given in (46) are formed. Next, different forms of $\mathbf{\Lambda}_k^d$ are stacked in $\mathbf{\Lambda}^d$ as indicated in (47). Finally, the summation of elements of $\mathbf{\Lambda}^d$ generates the proposed correlation function.

The block diagram of the new differential normalization function $\Delta_n^{IC}(d)$ is depicted in Fig. 5. According to this figure, $\hat{\mathbf{\Psi}}_k^d$ in (61) is first generated. Different $\hat{\mathbf{\Psi}}_k^d$ corresponding to different values of k are then stacked in $\hat{\mathbf{\Psi}}^d$ as given in (62). Finally, summation of elements of $\hat{\mathbf{\Psi}}^d$ gives the new normalization function $\Delta_n^{IC}(d)$.

Fig. 6 shows the n-th order metric defined in (66) for different values of n together with the reference metric in [14]. All the metrics in this figure are normalized by their maximum values. Obviously, the proposed metric is sharper than the reference metric in [14] (even for $n=2$), and as the order increases, the peak of the proposed metric becomes sharper. Note that when $n=2$ and $q=1$, the new metrics $M_n^{DC}(d)$ and $M_n^{IC}(d)$, respectively, reduce to the MoD and DoM metrics in [38]. Besides, when $n=4$, the proposed metrics $M_n^{DC}(d)$ and $M_n^{IC}(d)$, respectively, reduce to an equivalent form of DDN and that of CIDDN in [40].

It is worth mentioning that the proposed algorithms require only one OFDM symbol (or one preamble) with two identical parts in the time domain as shown in (19). Such an OFDM symbol can be easily generated by transmitting a zero sequence on the odd subcarriers and a PN sequence on the even subcarriers. Furthermore, as our algorithms use a training symbol with two identical parts in the time domain, and the two identical parts of the preamble remain identical after passing through a multipath fading channel, the proposed methods work in both frequency selective and AWGN channels.

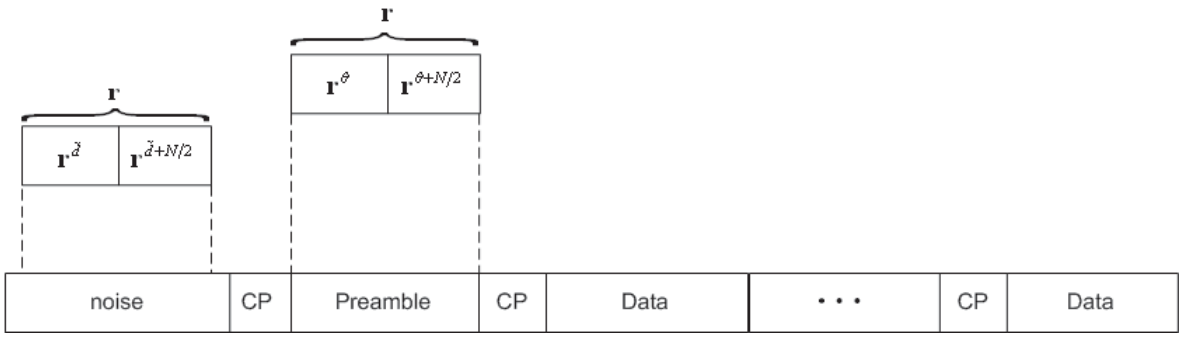

Figure 3: The structure of an OFDM frame and the received vectors $\mathbf{r}^d$ and $\mathbf{r}^{d+\frac{N}{2}}$ at timing instants $d = \widetilde{d}$ and $d = \Theta$.

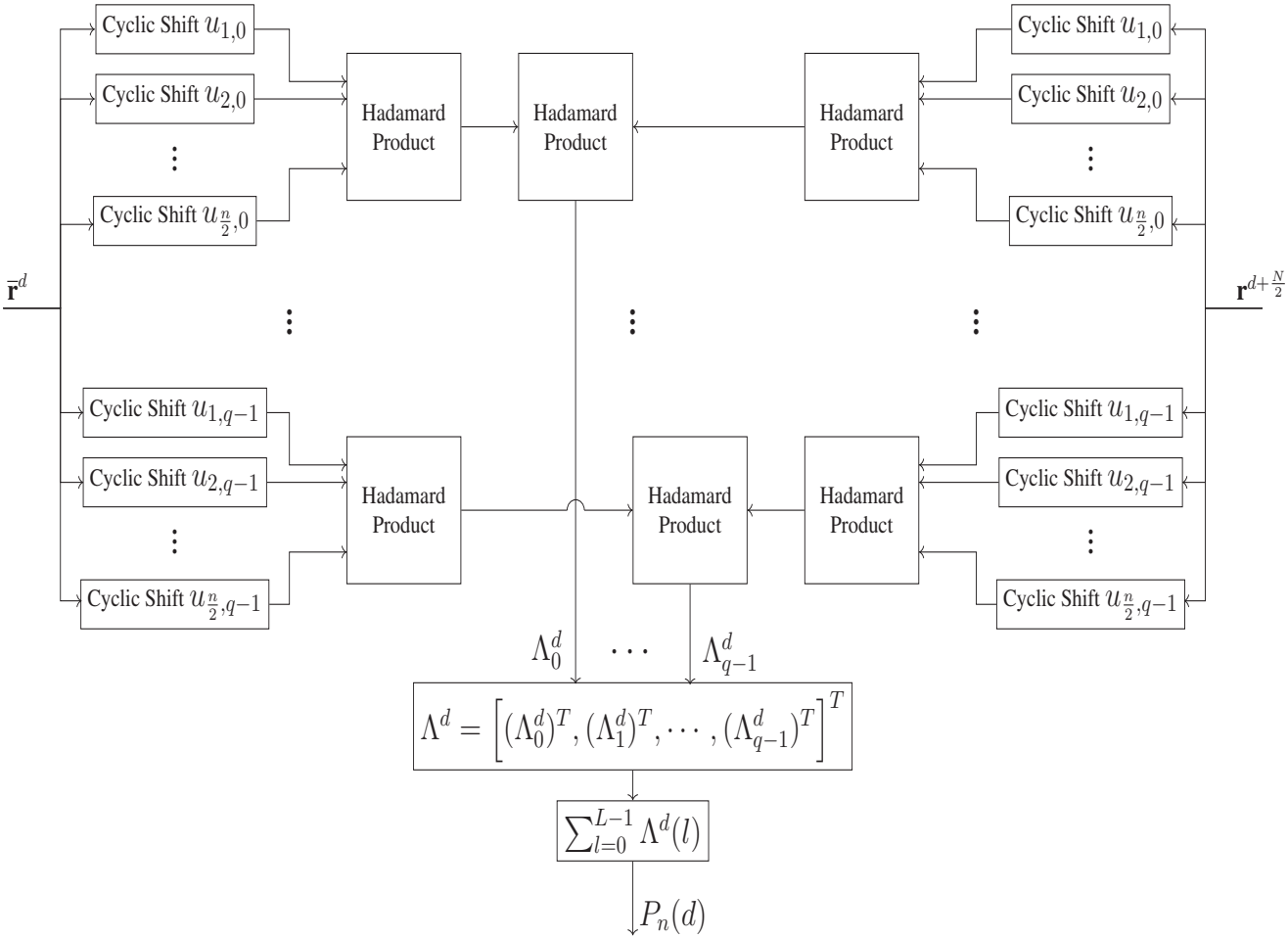

Figure 4: The block diagram of generation of the nth order correlation function $P_n(d)$.

3.2.3 Coarse Timing Synchronization

The synchronization procedure using the proposed metrics is described as follows. At the first stage, the preamble is detected when the timing metric reaches a predefined threshold,

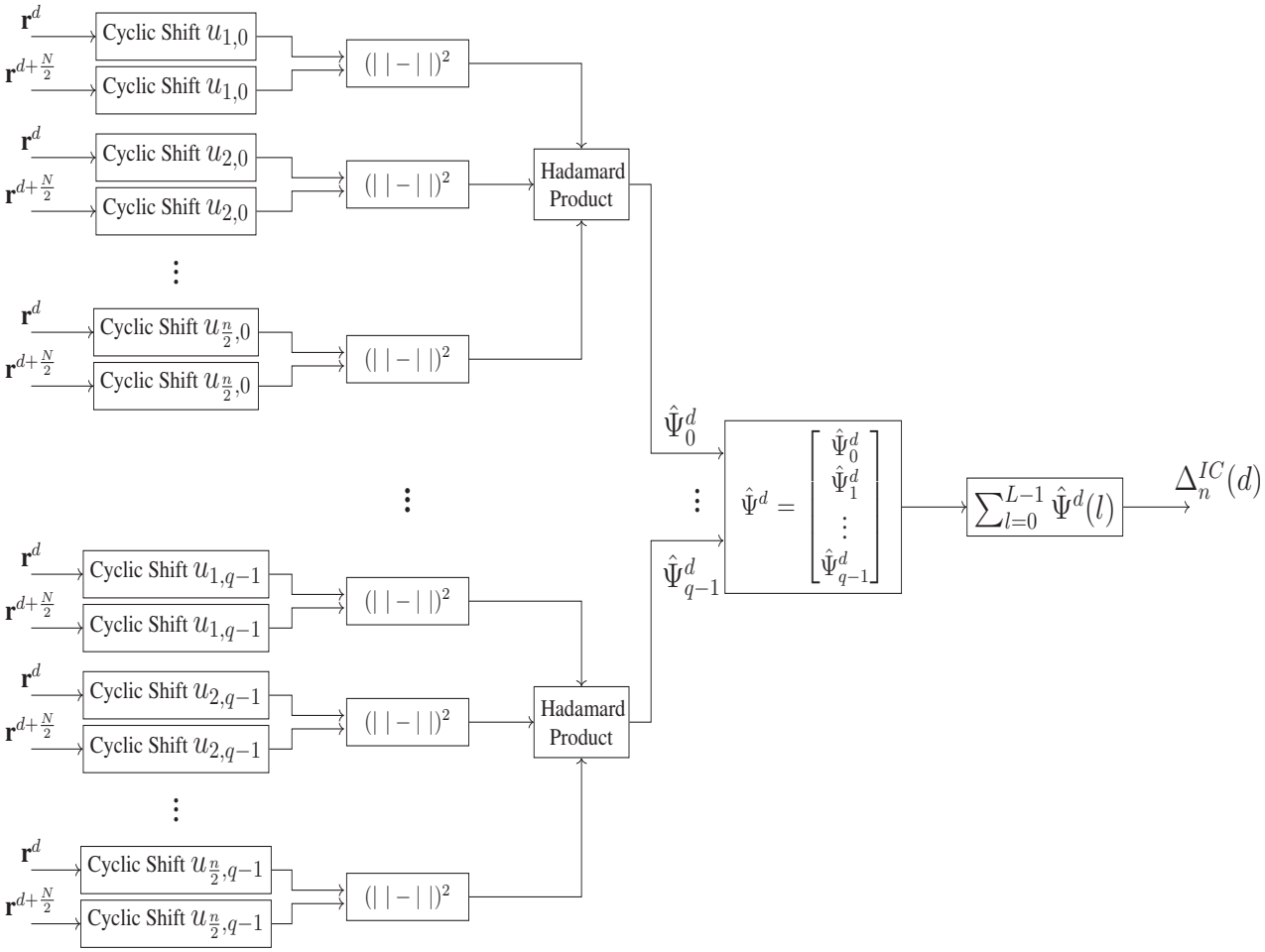

Figure 5: The block diagram of generation of the nth order normalization function $\Delta_n^{IC}(d)$.

and at the second stage, the coarse estimate of the start of a frame is obtained within the next N timing instants. In what follows, we first briefly explain the second stage and then introduce a new approach for coarse estimation of the start of a frame.

The aim of coarse timing estimation is to obtain an ISI-free region of the preamble [14]. This ISI-free region is indicated by timing instants corresponding to the plateau of the proposed timing metrics (and also metrics presented in [31] and [14,38–40]), i.e., the timing instants $-G \leq d \leq 0$ as shown in Fig. 6. In the existing works, the maximum of a timing metric within the next N samples (after detection of the preamble) determines the coarse timing estimate [14]. The disadvantage of this method is that sometimes the maximum of the timing metric (which is supposed to be within the plateau) is not in the ISI-free

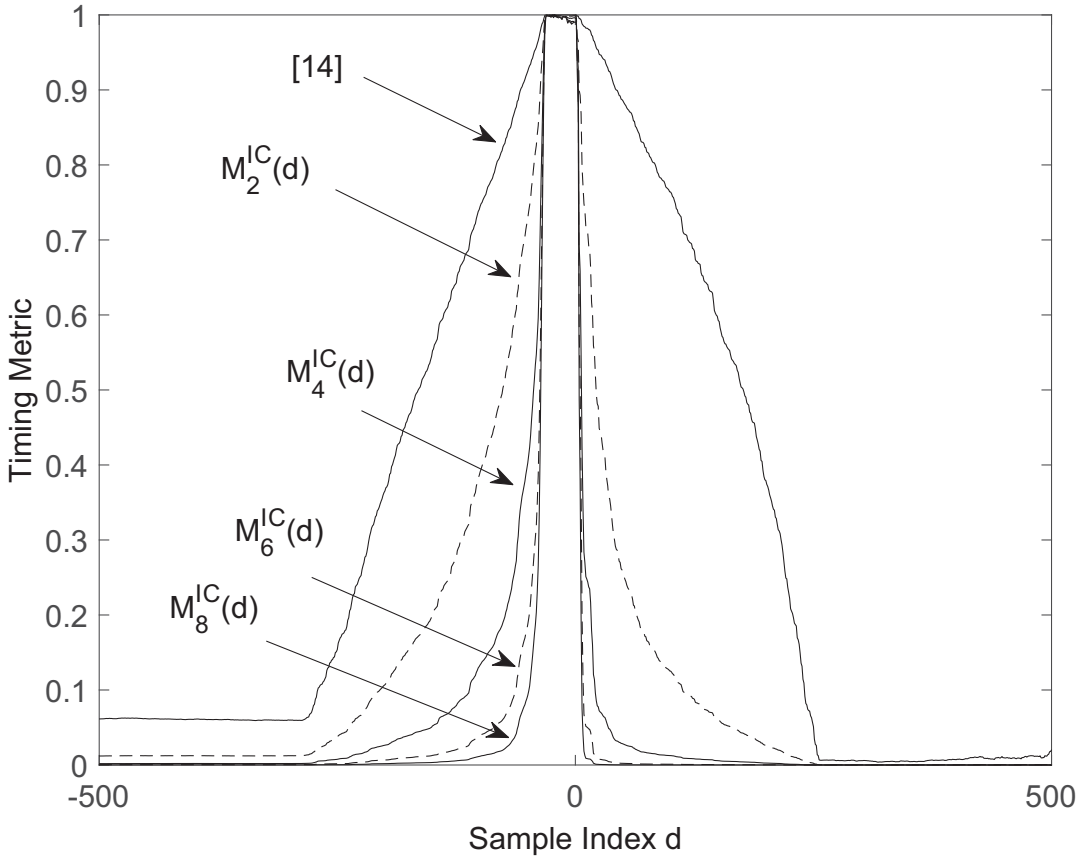

Figure 6: The value of the proposed timing metric $M_n^{IC}(d)$ along with the timing metric in [14] averaged over 10000 realizations in SUI-1 channel.

region $(-G \leq d \leq 0)$.

To overcome this weakness, instead of using only the timing instant corresponding to the maximum of the timing metric, we make use of a set of timing instants corresponding to the largest values of the timing metric. These timing instants are then averaged to produce a coarse estimate of the start of the frame. This way we can ensure that the timing estimate is in the ISI-free region $(-G \leq d \leq 0)$.

Accordingly, the proposed coarse timing estimation is performed as follows:

$$[\hat{d}_0, \hat{d}_1, \cdots, \hat{d}_{\psi-1}] = \arg\max_{d}\{M^{new}(d)\}, 1 \leq \psi \leq G \tag{69}$$

$$\hat{d} = \frac{\hat{d}_0 + \hat{d}_1 + \cdots + \hat{d}_{\psi-1}}{\psi} \tag{70}$$

where $\hat{d}_0, \hat{d}_1, \cdots, \hat{d}_{\psi-1}$ are the timing instants corresponding to the ψ largest values of the new timing metric $M^{new}(d)$ (in (58) or (66)), and ψ, $(1 \leq \psi \leq G)$ denotes the number of timing instants with the largest values (ψ is a parameter to be designed according to the expected performance). Here, $\text{args}\max_d\{M^{new}(d)\}$ finds the timing instants corresponding to the largest values of the timing metric, and $\hat{d}$ is the coarse timing estimate.

Next, as another two-stage timing synchronization method, we propose to use different timing metrics at different stages. The first stage which is preamble detection, utilizes a timing metric that has a good detection performance, and the second stage which is to obtain a coarse estimate of the start of a frame, employs a timing metric with low ISI probability. This scheme will be referred to as "combined method".

3.2.4 Performance Evaluation

The performance of timing metrics as classifiers is often assessed using the class-separablity criteria. As such, we investigate how distinct the values of the timing metrics at the correct and wrong timing points become as the order increases. To this end, we first consider the class-separablity in terms of the means of timing metrics at the wrong and correct timing points. Then, both the means and the variances of the metrics are utilized for class-separablity evaluation. We also investigate the performance limit of the timing metric on the order, as well as robustness of the new metrics to CFO. Finally, we consider the computational complexity of the proposed metrics.

3.2.4.1 Class-separability in Terms of Means

Here, the class-separability is assessed using the criterion defined in [38] that takes into account the difference between the means of timing metrics at correct and wrong timing instances. Thus, using the means derived in appendix A, the class-separability between

the n_1th order and n_2th order metrics that are dependent on CFO can be compared as

$$\alpha_{n_2}^{n_1} = \frac{E\{M_{n_1}^{DC}(\Theta)\}/E\{M_{n_1}^{DC}(\widetilde{d})\}}{E\{M_{n_2}^{DC}(\Theta)\}/E\{M_{n_2}^{DC}(\widetilde{d})\}} = \frac{\sqrt{2^{n_1-n_2}}SNR^{\frac{n_1-n_2}{2}}}{(|1-e^{j\pi\varepsilon}|^2 SNR+2)^{\frac{n_1-n_2}{2}}}. \tag{71}$$

Here, $\alpha_{n_2}^{n_1}$ compares the ratio of the means of the timing metric $M_{n_1}^{DC}(d)$ at correct and wrong timing points with that of $M_{n_2}^{DC}(d)$. Obviously, for a timing metric with more distinct means at correct and wrong timing points, the ratio of the means at the corresponding points is greater. Therefore, the metric has a better class-separability. According to this criterion, when $\alpha_{n_2}^{n_1} > 1$, the metric $M_{n_1}^{DC}(d)$ has better performance than $M_{n_2}^{DC}(d)$. When CFO is small, (71) reduces to $\alpha_{n_2}^{n_1} \simeq SNR^{\frac{n_1-n_2}{2}}$. Consequently, by increasing the order i.e., $n_1 > n_2$, the performance of $M_n^{DC}(d)$ improves for $SNR > 0$ dB.

Fig. 7 depicts $\alpha_{n_2}^{n_1}$ versus the SNR curves of $M_n^{DC}(d)$ for different values of $n_1 - n_2$ obtained using both the theoretical analysis and the Mont Carlo simulation in a multipath fading channel (the parameters of the simulation are given in the next section). Clearly, the simulation result matches very well with the theoretical analysis. Also, as $n_1 - n_2$ increases, the difference between the means significantly increases giving a better class-separability. Furthermore, the class-separability improves as SNR increases. Note that the methods presented in [38] and [40] are special cases of the proposed n-th order metrics.

The comparison of class-separability of CFO independent metrics for different values of order is given by

$$\alpha_{n_2}^{n_1} = \frac{E\{M_{n_1}^{IC}(\Theta)\}/E\{M_{n_1}^{IC}(\widetilde{d})\}}{E\{M_{n_2}^{IC}(\Theta)\}/E\{M_{n_2}^{IC}(\widetilde{d})\}} \tag{72}$$

which is shown in Fig. 8 versus SNR for different values of $n_1 - n_2$ (note that since we only have a bound for $E\{M_n^{IC}(\Theta)\}$ in (68), we cannot find an expression for $\alpha_{n_2}^{n_1}$, and therefore, theoretical result is not available). Again, it is noticed that the class-separability of the proposed timing metric improves with increasing $(n_1 - n_2)$, and with SNR. In this figure,

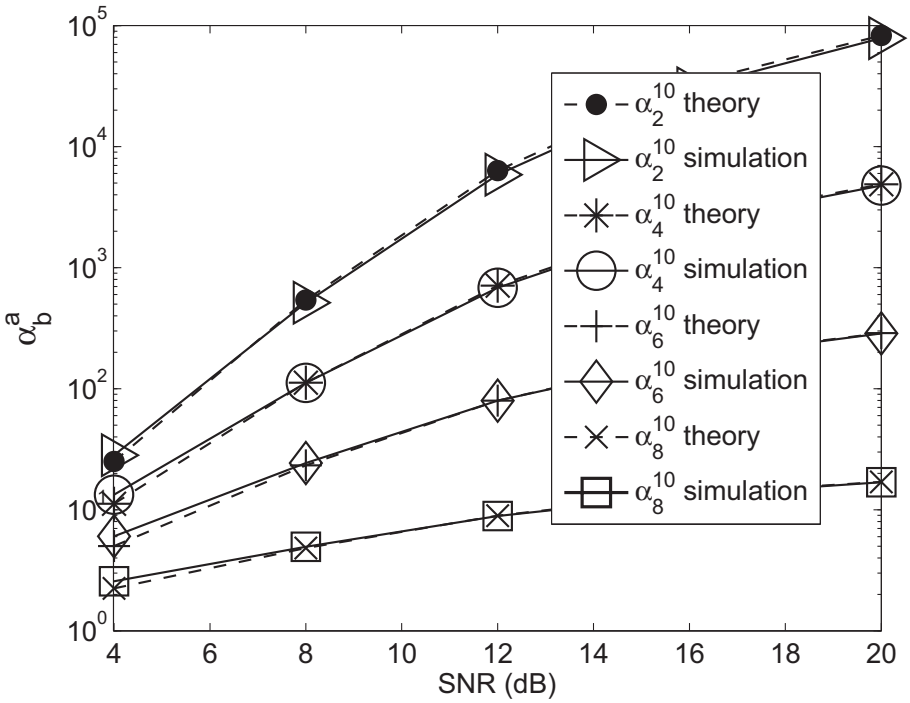

Figure 7: The class-separability using means of the timing metric $M_n^{DC}(d)$ obtained from analytical formula and Mont Carlo simulation.

we also have defined $\alpha_R^{n_1}$ as

$$\alpha_R^{n_1} = \frac{E\{M_{n_1}^{IC}(\Theta)\}/E\{M_{n_1}^{IC}(\widetilde{d})\}}{E\{M_R(\Theta)\}/E\{M_R(\widetilde{d})\}} \tag{73}$$

where $M_R(d)$ is the metric defined in [14], for which we have the means $E\{M_R(\Theta)\} = \frac{SNR}{SNR+1}$ and $E\{M_R(\widetilde{d})\} = \sqrt{\frac{\pi}{2N}}$. It is evident that $M_{n_1}^{IC}(d)$ has a considerably better performance as compared with $M_R(d)$.

3.2.4.2 Class-separability in Terms of Means and Variances

To evaluate the class-separability using both the means and the variances of a timing metric, we adopt the distance parameter defined in [40], namely,

$$D = (m_2 - \sigma_2) - (m_1 + \sigma_1) \tag{74}$$

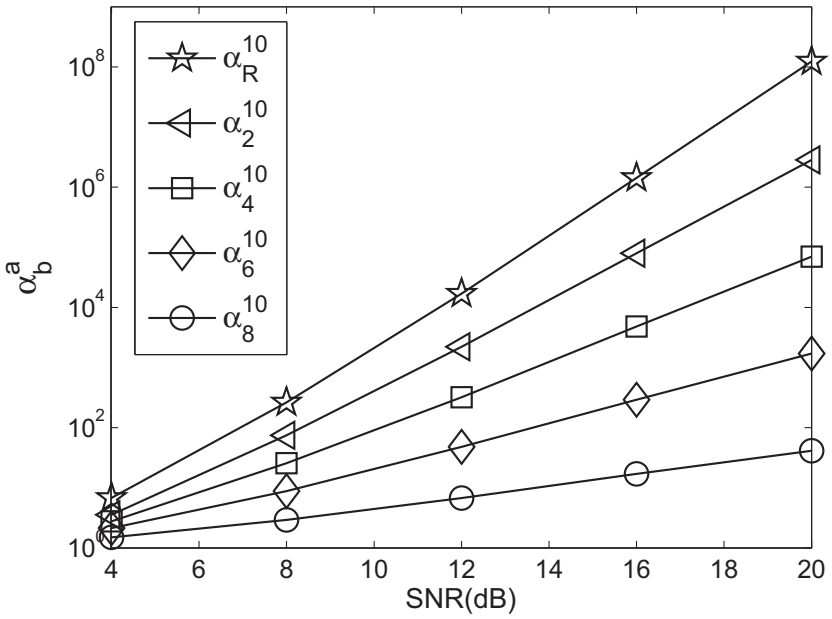

Figure 8: The class-separability using means of CFO independent timing metrics ($M_n^{IC}(d)$ and the metric in [14]) obtained from Mont Carlo simulation.

where m_1 and σ_1 are the mean and standard deviation of the timing metric at the wrong timing instant, respectively, and m_2 and σ_2 the mean and standard deviation at the correct timing points. The class-separability using distance D in (74) is demonstrated in Fig. 9. Since the theoretical derivation of the variances is often very difficult, one may resort to numerical simulation for obtaining the distance D.

Fig. 10 depicts the class-seprabilty criterion in terms of both the means and the variances for the CFO-dependent metric $M_n^{DC}(d)$. It is observed that the class-separability improves in general with increasing metric order as long as SNR is larger than 5 dB. The performance deterioration for low SNRs is due to the fact that the increase in the sum of the standard deviations becomes greater than the increase in the difference between the means, thus decreasing the class-separbility at low SNR.

Similarly, Fig. 11 shows the class-separbility of the CFO-independent metric $M_n^{IC}(d)$ in terms of the distance D together with that of the method in [14]. For low SNRs and small values of the correlation length, the performance becomes worse with increasing metric

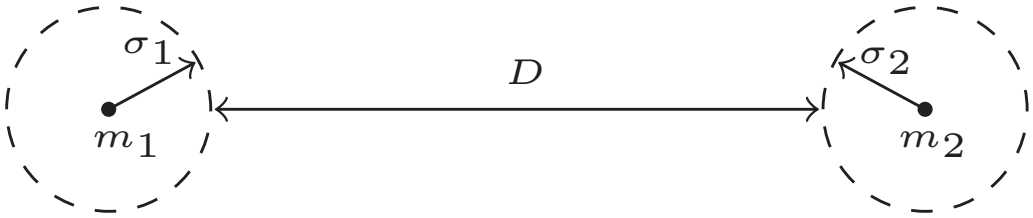

Figure 9: The illustration of class separability using the distance $D = (m_2 - \sigma_2) - (m_1 + \sigma_1)$.

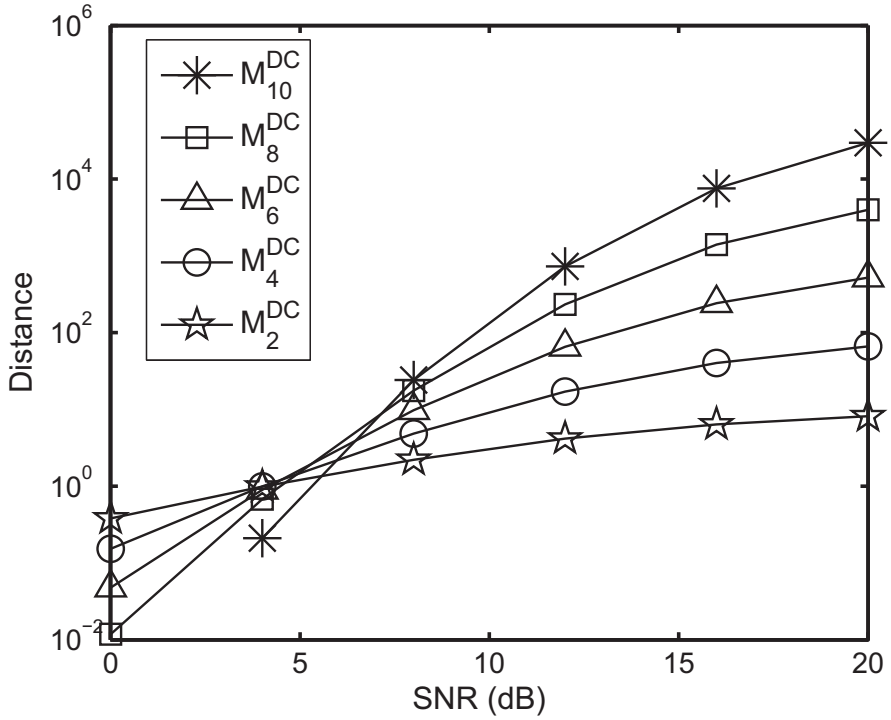

Figure 10: The class-separability using means and variances of the timing metric $M_n^{DC}(d)$.

order. For example, $M_{10}^{IC}(d)$ with $L = N$ has a worse performance compared with $M_8^{IC}(d)$ with $L = N$ at $SNR = 12dB$. However, when the correlation lengths of all the metrics and/or SNR increase, not only the class-separability of each metric is improved, but also the class-separability improves with increasing order. Furthermore, the class-separability of $M_n^{IC}(d)$ is significantly better than that of [14].

In Fig. 12 we have compared the class-separability using means and variances of

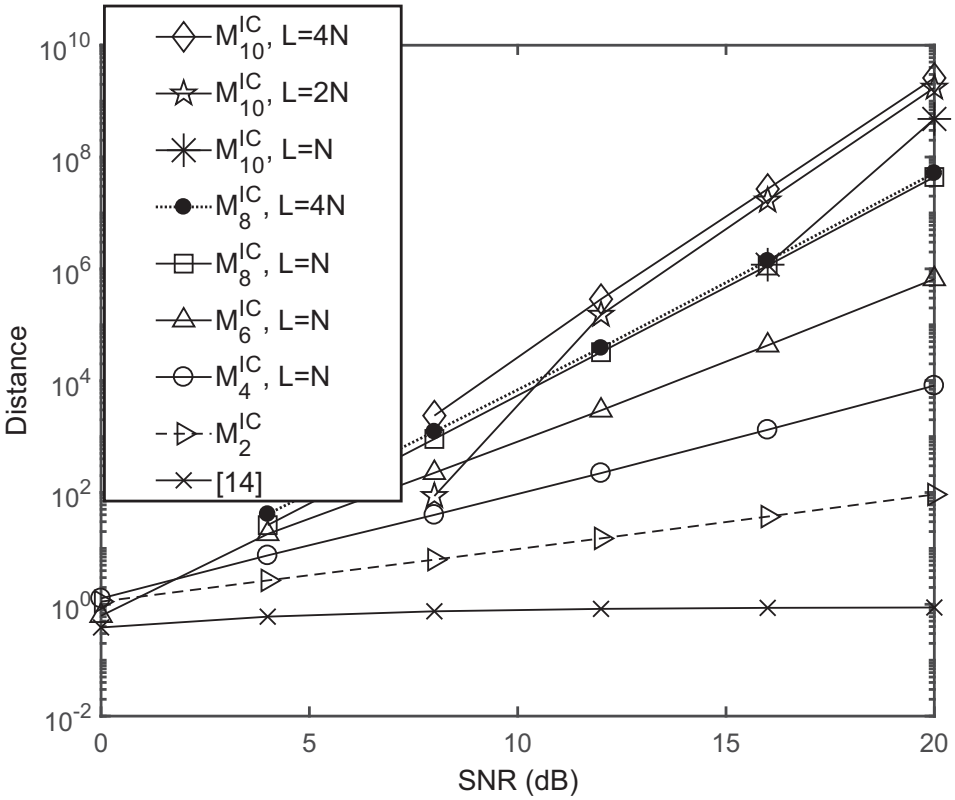

Figure 11: The class-separability using means and variances of CFO independent timing metrics ($M_n^{IC}(d)$ and the metric in [14]).

$M_n^{IC}(d)$ with those of the metrics in [35–37] which again shows that $M_n^{IC}(d)$ has a significantly better class-separability.

It is worth mentioning that although for some values of SNR the performance may deteriorate with increasing the order of the timing metric, the new method is capable of improving the performance for $SNR \geq 0\ dB$. For example, according to Figs. 9 and 10, $M_6^{IC}(d)$, $M_4^{IC}(d)$, and $M_2^{IC}(d)$ all have a better performance than the metrics in [1] and [35–37] when $SNR \geq 0\ dB$. Thus, the new scheme can provide a better performance for $SNR \geq 0\ dB$ when applied in wireless communication systems such as WLAN, WiMAX, and LTE in comparison with previous methods.

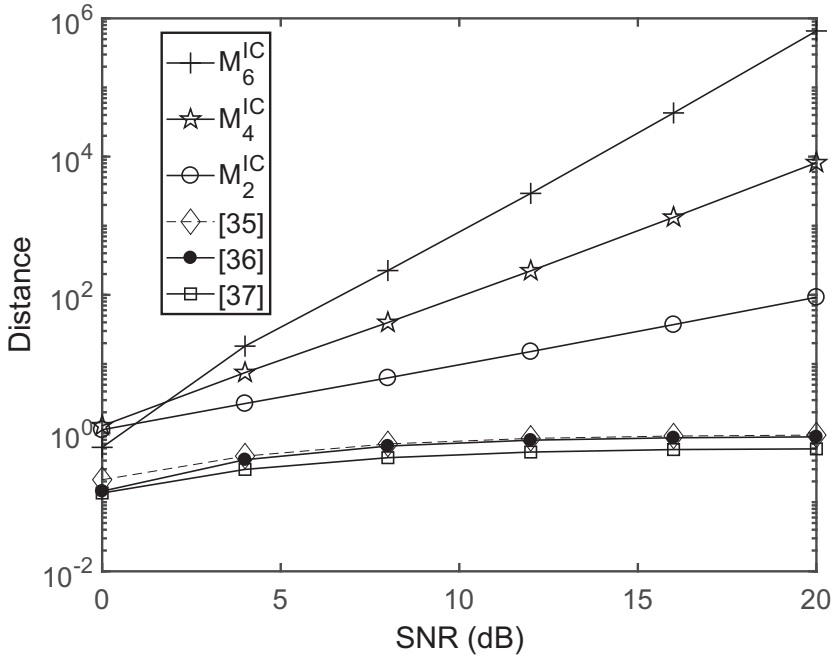

Figure 12: The class-separability using means and variances of $M_n^{IC}(d)$ and the metric in [35–37].

3.2.4.3 Performance Limit on the Order

In the previous subsection, it was observed that under low SNR and for some correlation lengths, the class-separability does not improve with increasing the order of the metric. This issue brings up the following question: how much can we increase the order so as to obtain further improved performances? To give a definite answer to this question, we need to analyze the means and the variances of the metric. Although we have obtained general expressions for the means in (59), (60), (67) and (68), deriving the variances is a very difficult task if not impossible. Thus, we have to rely on simulations that are to be performed for specific and limited values of n. According to Figs. 7 and 8, as the order increases, the difference between the means significantly increases. At the same time, as shown in Fig. 13, the standard deviation at correct timing points also increases with the order. The increase in the standard deviation for certain values of n, correlation length and SNR, can be higher than the increase in the distance between the means, thus

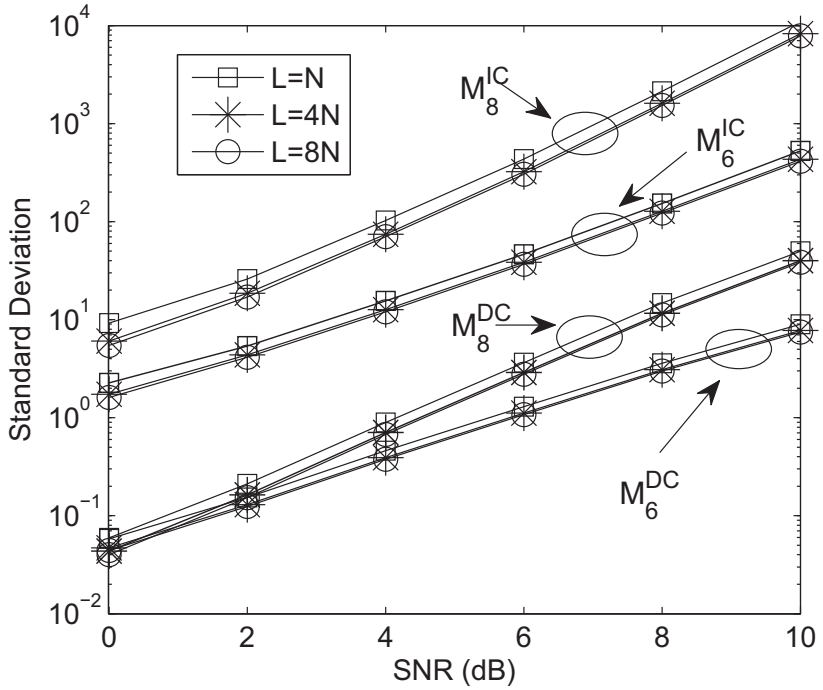

Figure 13: The standard deviation for the 6th and 8th order timing metrics ($M_n^{DC}(d)$ and $M_n^{IC}(d)$) at correct timing point.

decreasing the class-separability as demonstrated in Figs. 10 and 11. In other words, how much we can increase the order is dependent on the variance of the metric. To achieve a better class-separability, the order should be increased up to a value such that the sum of the standard deviations of the metrics at the correct and wrong timing points is smaller than the difference of the means at the corresponding points. As an example, according to Fig. 11, for correlation length $L = N$ and $SNR = 8\ dB$, the maximum is $n = 8$. It is worth mentioning that the standard deviation (as depicted in Fig. 13) is also a function of the correlation length, and as the correlation length increases the variance decreases, resulting in higher class-separability.

3.2.4.4 Robustness to CFO

Without loss of generality, we consider the proposed timing metrics in the absence of noise where the received signal can be written as

$$r(k) = e^{j\frac{2\pi}{N}k\epsilon}y(k),\ 0 \le k \le N-1. \tag{75}$$

Assuming that $\mathbf{\Lambda}^d$ has only one subvector corresponding to cyclic shifts $(u_{1,k}, u_{2,k}, \cdots, u_{\frac{n}{2},k}) = (0, 1, \cdots, \frac{n}{2}-1)$, as given by

$$\mathbf{\Lambda}^d = \mathbf{\Lambda}_0^d = \overline{\mathbf{r}}_0^d \circ \overline{\mathbf{r}}_1^d \circ\ \cdots \circ \overline{\mathbf{r}}_{\frac{n}{2}-1}^d \circ \mathbf{r}_0^{d+\frac{N}{2}} \circ \mathbf{r}_1^{d+\frac{N}{2}} \circ \cdots \circ \mathbf{r}_{\frac{n}{2}-1}^{d+\frac{N}{2}}. \tag{76}$$

Thus, according to (51), we have the following correlation function

$$\begin{aligned} P_n(d) = \sum_{l=0}^{\frac{N}{2}-1} \overline{r}(d+l)\overline{r}\Big(d+|l+1|_{\frac{N}{2}}\Big)\cdots\overline{r}\Big(d+|l+\frac{n}{2}-1|_{\frac{N}{2}}\Big)r\Big(d+l+\frac{N}{2}\Big)r\Big(d+|l+1|_{\frac{N}{2}}+\frac{N}{2}\Big) \\ \cdots r\Big(d+|l+\frac{n}{2}-1|_{\frac{N}{2}}+\frac{N}{2}\Big) = \sum_{l=0}^{\frac{N}{2}-1} e^{j\pi\epsilon\frac{n}{2}}\Big|y(l)\Big|^2\Big|y\big(|l+1|_{\frac{N}{2}}\big)\Big|^2\cdots\Big|y\big(|l+\frac{n}{2}-1|_{\frac{N}{2}}\big)\Big|^2. \end{aligned} \tag{77}$$

It can be shown that the magnitude of $P_n(d)$ can be expressed as

$$\big|P_n(d)\big| = \sum_{l=0}^{\frac{N}{2}-1} \Big|y(l)\Big|^2\Big|y\big(|l+1|_{\frac{N}{2}}\big)\Big|^2\cdots\Big|y\big(|l+\frac{n}{2}-1|_{\frac{N}{2}}\big)\Big|^2, \tag{78}$$

which clearly indicates that $|P_n(d)|$ is not affected by the frequency offset. With (75) and (76) above, one can derive the proposed normalization functions as

$$\Delta_n^{DC}(d) = |1-e^{j\pi\epsilon}|^n \sum_{k=0}^{\frac{N}{2}-1} \Big|y(k)\Big|^2\Big|y\big(|k+1|_{\frac{N}{2}}\big)\Big|^2\cdots\Big|y\big(|k+\frac{n}{2}-1|_{\frac{N}{2}}\big)\Big|^2, \tag{79}$$

$$\Delta_n^{IC}(d) = \sum_{k=0}^{\frac{N}{2}-1} \Big(\Big|y(k)\Big| - \Big|y\big(k+\frac{N}{2}\big)\Big|\Big)^2 \Big(\Big|y\big(|k+1|_{\frac{N}{2}}\big)\Big| - \Big|y\big(|k+1|_{\frac{N}{2}}+\frac{N}{2}\big)\Big|\Big)^2 \cdots \Big(\Big|y\big(|k+\frac{n}{2}-1|_{\frac{N}{2}}\big)\Big| - \Big|y\big(|k+\frac{n}{2}-1|_{\frac{N}{2}}+\frac{N}{2}\big)\Big|\Big)^2. \quad (80)$$

Evidently, the first normalization function (79) depends on CFO, whereas the second normalization function (80) does not. As a consequence, the first timing metric (58) is sensitive to CFO, whereas the second timing metric (66) is robust to CFO since both $|P_n(d)|$ and $\Delta_n^{IC}(d)$ therein are independent of CFO.

3.2.4.5 Computational Complexity

In this section, we assess the computational complexity of the proposed timing metrics. The complexity is mainly due to the computation of the correlation function. Without loss of generality, we consider again the case of $\mathbf{\Lambda}^d$ having only one subvector as in (76). Considering the vector $\mathbf{C}^d = \overline{\mathbf{r}}_0^d \circ \overline{\mathbf{r}}_1^d \circ \cdots \circ \overline{\mathbf{r}}_{\frac{n}{2}-1}^d$ at the timing instant d, it is observed that $\mathbf{C}^{d+1}$ can be expressed in terms of the elements of $\mathbf{C}^d$ as

$$\mathbf{C}^{d+1} = \Big[\overline{r}(d+1)\overline{r}(d+\frac{N}{2})\cdots\overline{r}(d+\frac{N}{2}-\frac{n}{2}+2), \overline{r}(d+2)\overline{r}(d+1)\overline{r}(d+\frac{N}{2})\cdots\overline{r}(d+\frac{N}{2}-\frac{n}{2}+3), \cdots, \overline{r}(d+\frac{n}{2}-1)\overline{r}(d+\frac{n}{2}-2)\cdots\overline{r}(d+\frac{N}{2}), C^d(\frac{n}{2}), C^d(\frac{n}{2}+1), \cdots, C^d(\frac{N}{2}-1), \overline{r}(d+\frac{N}{2})\overline{r}(d+\frac{N}{2}-1)\cdots\overline{r}(d+\frac{N}{2}-\frac{n}{2}+1)\Big]^T. \quad (81)$$

Similarly, with $\mathbf{D}^d = \mathbf{r}_0^{d+\frac{N}{2}} \circ \mathbf{r}_1^{d+\frac{N}{2}} \circ \cdots \circ \mathbf{r}_{\frac{n}{2}-1}^{d+\frac{N}{2}}$, we have $\mathbf{D}^{d+1}$ as

$$\mathbf{D}^{d+1} = \Big[r(d+\frac{N}{2}+1)r(d+N)\cdots r(d+N-\frac{n}{2}+2), r(d+\frac{N}{2}+2)r(d+\frac{N}{2}+1)r(d+N)\cdots \times r(d+N-\frac{n}{2}+3), \cdots, r(d+\frac{N}{2}+\frac{n}{2}-1)r(d+\frac{N}{2}+\frac{n}{2}-2)\cdots r(d+N), D^d(\frac{n}{2}), D^d(\frac{n}{2}+1), \cdots, D^d(\frac{N}{2}-1), r(d+N)r(d+N-1)\cdots r(d+N-\frac{n}{2}+1)\Big]^T. \quad (82)$$

Using (81) and (82), we get

$$\mathbf{\Lambda}_0^{d+1} = \mathbf{C}^{d+1} \circ \mathbf{D}^{d+1} = \Big[C^{d+1}(0)D^{d+1}(0), C^{d+1}(1)D^{d+1}(1), \cdots ,$$
$$C^{d+1}(\frac{n}{2}-2)D^{d+1}(\frac{n}{2}-2), \Lambda_0^d(\frac{n}{2}), \Lambda_0^d(\frac{n}{2}+1), \cdots , \Lambda_0^d(\frac{N}{2}-1), C^{d+1}(\frac{N}{2}-1)D^{d+1}(\frac{N}{2}-1)\Big]^T. \tag{83}$$

Therefore, the n-th order correlation function can finally be written as

$$P(d+1) = P(d) - \Lambda_0^d(0) - \Lambda_0^d(1) - \cdots - \Lambda_0^d(\frac{n}{2}-1)$$
$$+ \Lambda_0^{d+1}(0) + \Lambda_0^{d+1}(1) + \cdots + \Lambda_0^{d+1}(\frac{n}{2}-2) + \Lambda_0^{d+1}(\frac{N}{2}-1). \tag{84}$$

For $\Lambda_0^d(0)$ we need $n-1$ multiplications; $n-5$ of these multiplications are again used in $\Lambda_0^d(1)$ and only 4 new multiplications are utilized. The same behavior is observed comparing $\Lambda_0^d(1)$ and $\Lambda_0^d(2)$ and so on. Therefore, with careful design of $\mathbf{\Lambda}^d$, we need $6n-10$ multiplications for each subvector of $\mathbf{\Lambda}^d$. In general, we need $(6n-10)q$ complex multiplications for the n-th order ($n \geq 6$) correlation function, where q is the number of subvectors of $\mathbf{\Lambda}^d$ in (47). In Table 3, we summarize the computational complexities of different methods in terms of complex multiplications and additions. Obviously, for the proposed n-th order metrics, the computational complexity increases with the order of the metrics. As an example, the 6th order metric, with a correlation length of $L = N$ (or $q = 2$), needs 52 complex multiplications and 12 complex additions. In comparison with [40], which for the same correlation length needs 24 complex multiplications and 8 complex additions, the new method increases the complexity approximately by a factor of 2. It is worth mentioning that as we demonstrated by different class-separability criteria in this section, and as will be shown by false alarm and missed-detection probabilities in the next section, the 6th order metrics are capable of improving the performance by a factor greater than 2.

Table 3: Computational Complexities of Different Coarse Timing Methods

Method	**Multiplication**	**Addition**
[14]	2	2
[38]	2	2
[40]	$12q,\ 1 \leq q \leq N/2$	$4q,\ 1 \leq q \leq N/2$
Proposed n-th order metrics	$(6n-10)q,\ n \geq 6,$ $1 \leq q \leq (N/2)^{n/2-1}$	$nq,$ $1 \leq q \leq (N/2)^{n/2-1}$

3.2.4.6 Simulation Results

To evaluate the performance of the new timing estimation methods by simulation, we consider an OFDM system with 512 subcarriers and CP length of 64. The Stanford University Interim (SUI) channel modeling, namely, SUI-1 [141] is used for modeling a multipath frequency selective channel. The sampling rate is set to 5 MHz. Unless otherwise stated, in the evaluation of the performance of $M_n^{DC}(d)$ in (58), CFO is chosen as $\varepsilon = 0.1$, whereas for illustrating the performance of $M_n^{IC}(d)$ in (66) we choose $\varepsilon = 5.5$, due to the fact that $M_n^{DC}(d)$ is vulnerable to ε whereas $M_n^{IC}(d)$ is robust against ε.

It is worth mentioning that in Figs. 14-17 both x and y axes are logarithmically shown. Special attention should be paid to the threshold axis. For example, the distance between the threshold points 10^0 and 10^1 appears equal to the distance between the threshold points 10^1 and 10^2. However the latter distance is actually 9 times that of the former one. Furthermore, note that the n-th order metrics reduce to that in [38] and [40] for $n = 2$ and $n = 4$, respectively.

Fig. 14 depicts the false alarm and missed detection probabilities of the n-th order metric for different values of n at $SNR = 10dB$ and correlation length $L = N$. For a fair comparison, the error probability in this figure is only depicted for the metrics that are dependent on CFO. The probability of false alarm is the probability that at the wrong timing points (none of the samples belong to the preamble) the metric goes above a predefined threshold. The missed detection probability is the probability that at the

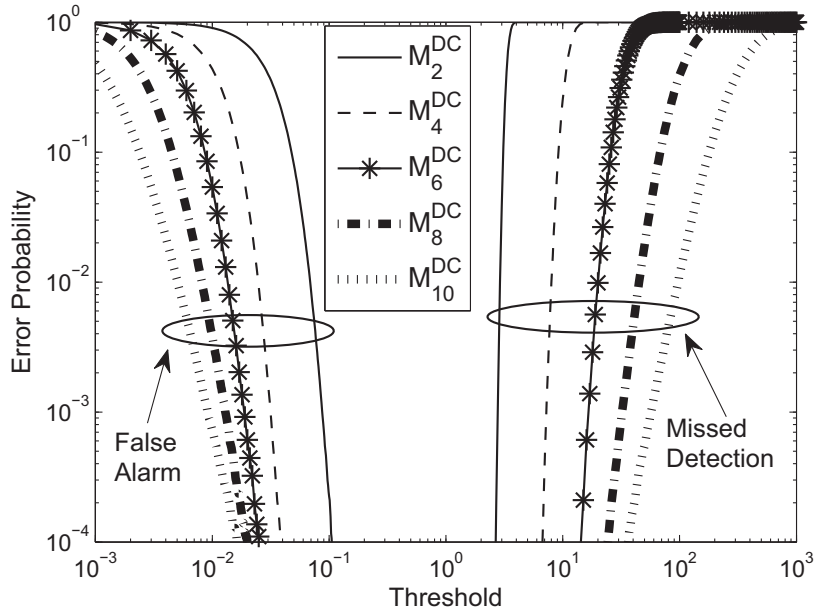

Figure 14: Probabilities of false alarm and missed detection for $M_n^{DC}(d)$ in SUI-1 channel at $SNR = 10\ dB$.

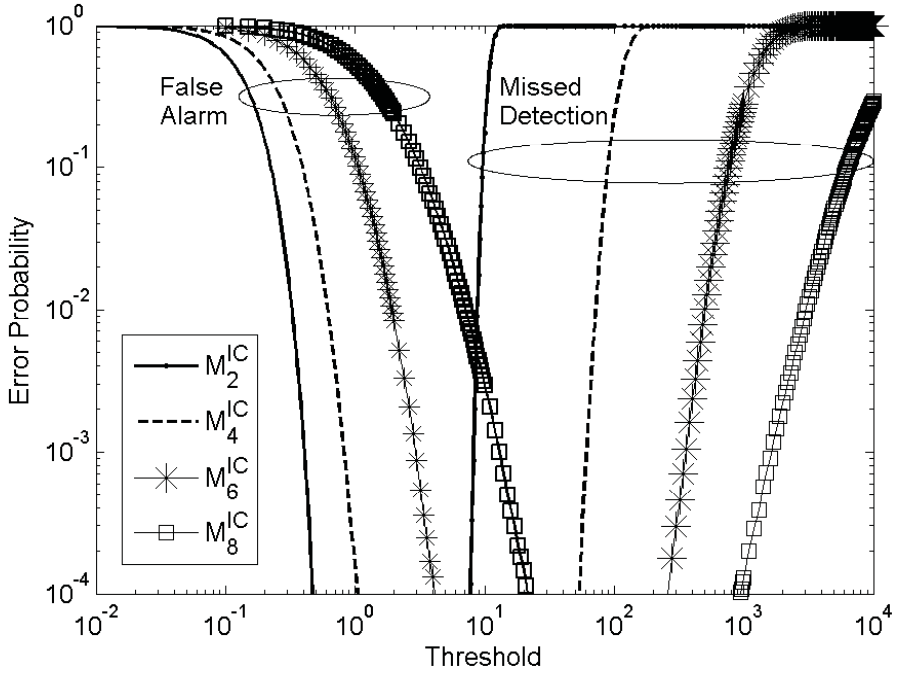

Figure 15: Probabilities of false alarm and missed detection for $M_n^{IC}(d)$ in SUI-1 channel at $SNR = 10\ dB$.

correct timing points, the metric has a value lower than the threshold. It is observed that as the order increases, both the false alarm and missed-detection probabilities improve and the range of acceptable thresholds remarkably increases. For example, for the probabilities of false alarm and missed detection equal to 10^{-3}, the threshold can be respectively chosen in the ranges $0.087 < th < 2.77$, $0.032 < th < 7.2$, $0.018 < th < 16$, $0.012 < th < 32$ and $0.009 < th < 57$ for $n = 2, 4, 6, 8, 10$.

Fig. 15 illustrates the performance of $M_n^{IC}(d)$ in (66) in terms of the false alarm and missed detection probabilities. All the metrics in this figure are robust to CFO. According to this figure, as the order increases, the missed detection probability significantly decreases and at the same time, the false alarm probability increases. However, the amount of reduction in the missed detection probability is much greater than the increase in the false alarm probability. As a result, the range of acceptable thresholds increases. For example, for $n = 2, 4, 6, 8$ and an error probability of 10^{-3}, we respectively have the following threshold ranges: $0.4 < th < 8.2$, $0.81 < th < 61$, $2.9 < th < 350$ and $13 < th < 1600$. In other words, when the order increases by two, the threshold range increases at least by a factor of 4.

At $SNR = 6\ dB$, the probabilities of the false alarm and missed detection for $M_n^{IC}(d)$ are plotted in Fig. 16. It is observed that for having these probabilities equal to 10^{-3}, the threshold ranges of the method in [14], $M_2^{IC}(d)$, $M_4^{IC}(d)$, and $M_6^{IC}(d)$ are $0.16 < th < 0.75$, $0.4 < th < 3.5$, $0.81 < th < 11.5$, and $2.9 < th < 27$, respectively. It can be concluded that $M_6^{IC}(d)$ has an acceptable threshold range that is 40.8 times larger than that of [14], 7.77 times larger than that of $M_2^{IC}(d)$, and 2.25 times larger than that of $M_4^{IC}(d)$.

False alarm and missed detection probabilities of the 10th order metric $M_{10}^{IC}(d)$ are depicted in Fig. 17 for different values of the correlation length L at $SNR = 6dB$. It is noticed that for $L = N$ and $L = 2N$, the false alarm and missed detection probabilities intersect. Thus, for error probabilities lower than the error probability at the point of intersection there is no acceptable threshold range. Comparing with Fig. 16, where the

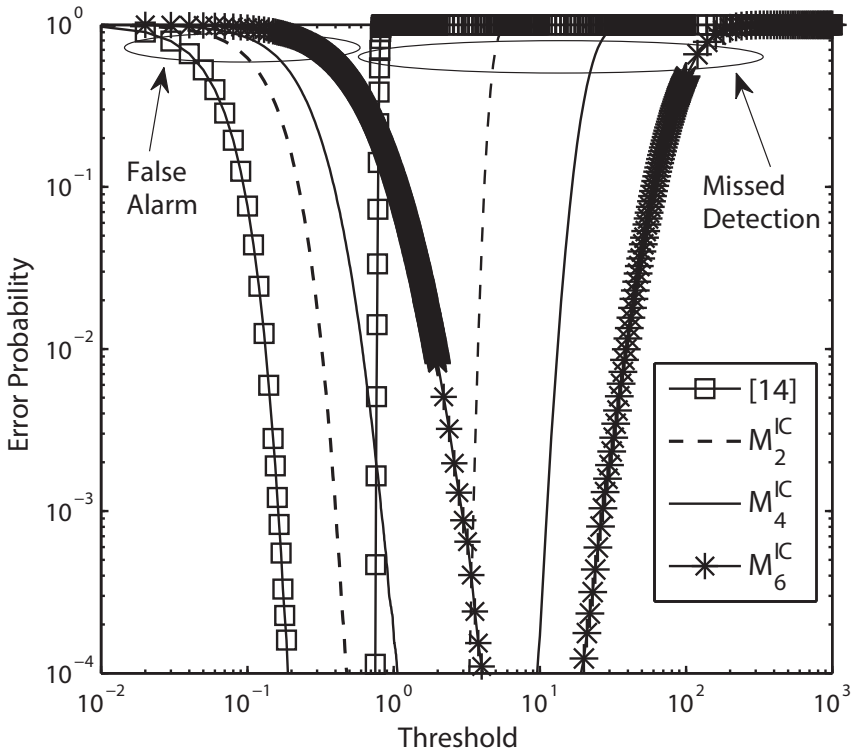

Figure 16: Probabilities of false alarm and missed detection for $M_n^{IC}(d)$ in SUI-1 channel at $SNR = 6\ dB$.

false alarm and missed detection probabilities of the 6th order metric have no intersection, it can be concluded that with increasing order, the performance may deteriorate (as discussed in the previous section), or the correlation length needs to be increased. For example, when the correlation length increases to $L = 16N$, no intersection is noticed.

In Fig. 18, we have compared the performance of $M_2^{IC}(d)$ with metrics in [35–37] at $SNR = 10\ dB$. It is observed that for having error probability equal to 10^{-2}, the acceptable threshold ranges for metrics in [35], [36], [37], and $M_2^{IC}(d)$ are $0.01 < th < 0.78$, $0.11 < th < .87$, $0.018 < th < 0.039$, and $0.32 < th < 8.9$, respectively. Therefore, the acceptable threshold range for $M_2^{IC}(d)$ is considerably larger than the metrics in [35–37]. This is the consequence of the class-separability performance of the new metric as illustrated in Fig. 12.

Fig. 19 depicts the ISI probabilities for different CFO-dependent timing estimators.

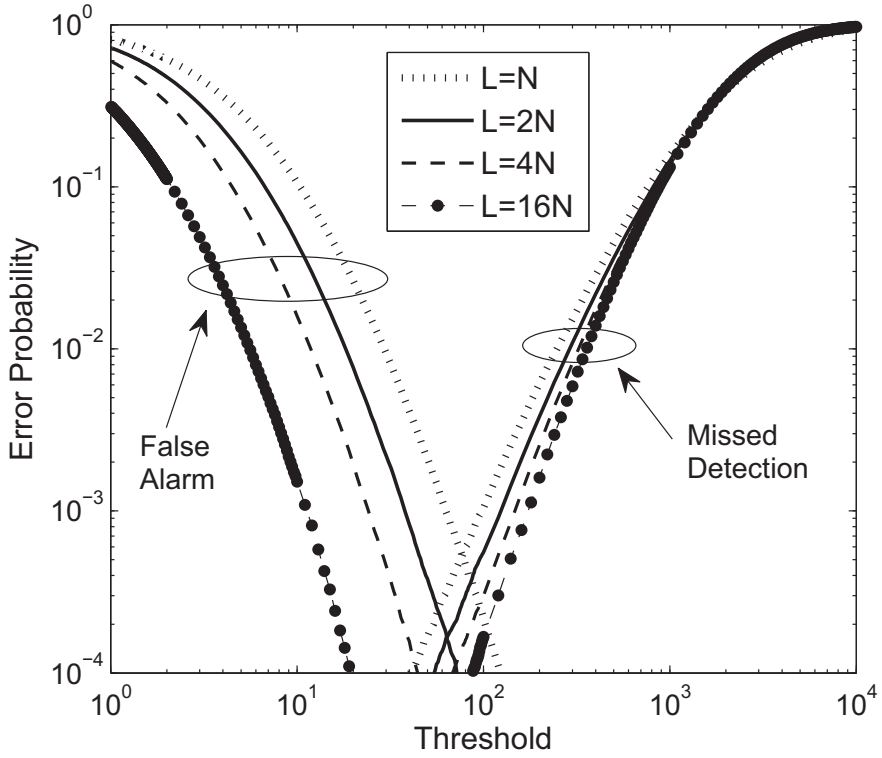

Figure 17: Probabilities of false alarm and missed detection for $M_{10}^{IC}(d)$ in SUI-1 channel at $SNR = 6\ dB$.

Probability of ISI is the probability that the timing estimate causes ISI, i.e., the FFT window includes samples from two neighboring OFDM symbols. As can be observed, when the proposed coarse timing estimation method is used along with the proposed metric $M_6^{DC}(d)$ in (58), the probability of ISI significantly improves. Furthermore, the ISI probability improves when the value of ψ increases. In this figure, we have also depicted the performance of the metric in [37] when timing back-off is performed i.e., the timing estimate corresponding to the maximum of the timing metric is shifted back towards the CP. According to this figure, the new method is capable of having similar performance as that of an impulse-like metric [37] with timing-back off. It is worth mentioning that although using timing-back off along with the metric in [37] can improve the probability of ISI of [37], our metric, as depicted in Figs. 12 and 18, has a considerably better performance than the metric in [37] in terms of class-separability and false alarm and

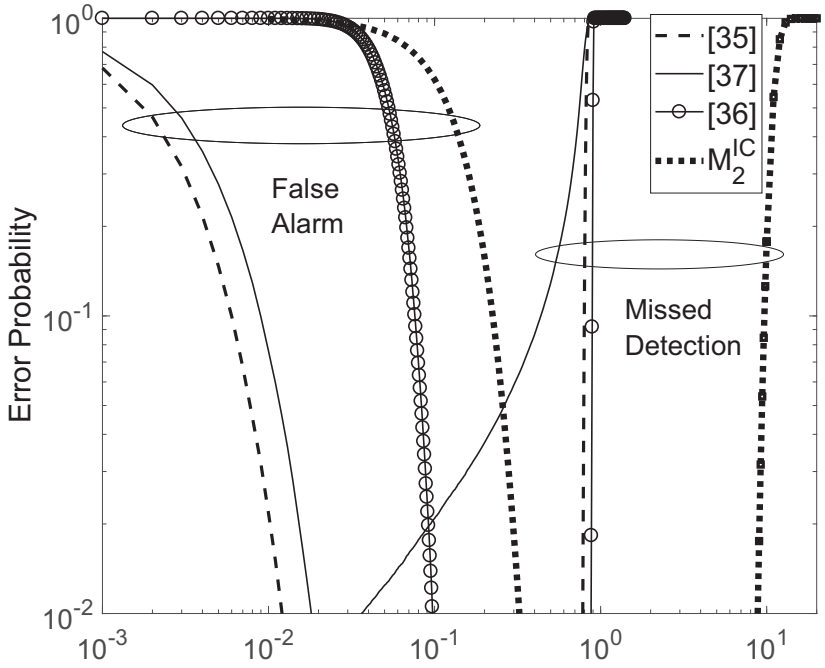

Figure 18: Probabilities of false alarm and missed detection for $M_2^{IC}(d)$ in SUI-1 channel at $SNR = 10\ dB$ in comparison with those of [35–37].

missed detection probabilities.

In Fig. 20, the ISI performance of different methods that are not affected by CFO is shown. It is evident that when the new estimation method uses the peaks of $M_6^{IC}(d)$ in (66), the performance improves for $SNR > 8dB$. In this figure, we have also plotted two curves resulting from the proposed combined method. Recall that the combined method is a two stage method where the first stage i.e. frame detection is done by the new nth order timing metric and the second stage is performed by application of the metric [14] into the new timing estimation method (69) as $M^{new}(d)$. As can be observed, the proposed combined method has a significantly lower probability of ISI.

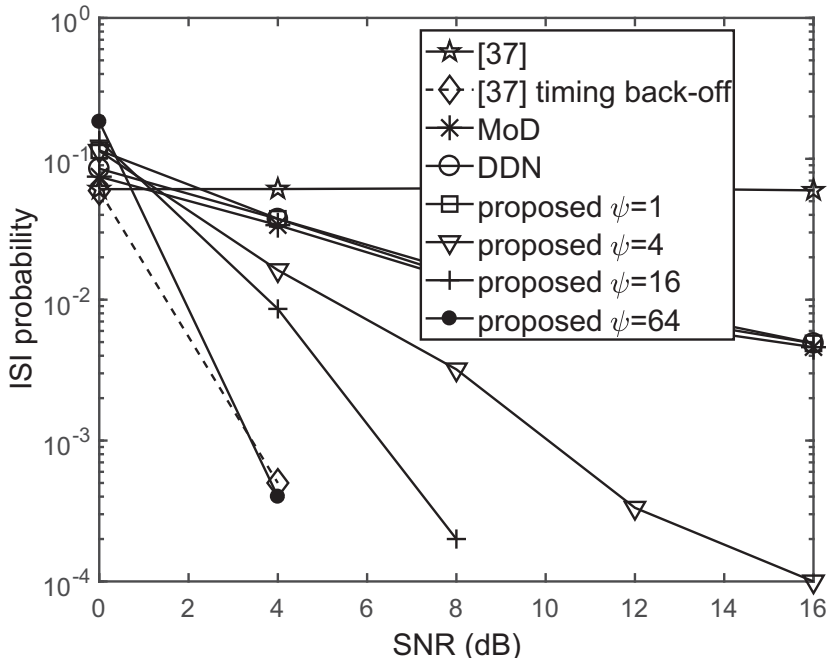

Figure 19: Probability of ISI for different timing schemes.

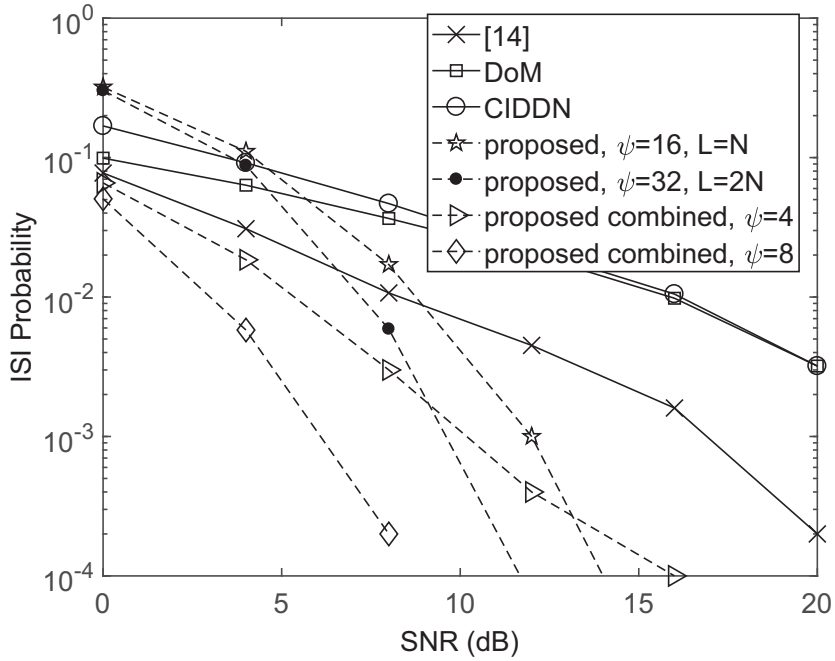

Figure 20: Probabilities of ISI for different timing schemes using CFO independent timing metrics.

3.3 Proposed Method: IFO Detection

3.3.1 Preliminaries

For frequency synchronization, CFO is usually divided into two parts, fractional frequency offset (FFO) and integer frequency offset (IFO) [71]. In other words we can write CFO as

$$\varepsilon = q(\nu + \lambda); \tag{85}$$

where $\nu, \left(-\frac{1}{2} < \nu \leq \frac{1}{2}\right)$, is FFO, λ is IFO, and q is the number of periodic parts of a preamble in the time domain. The estimation of CFO ε is usually performed in two stages [31, 71, 72, 142]: FFO estimation in the time domain using the periodic parts of a preamble with an estimation range of $(-q/2, q/2]$, and IFO detection in the frequency domain. Assuming that timing and FFO synchronization is already performed (similar to the previous works that focus on IFO detection [31, 71, 72, 142], the received signal after taking discrete Fourier transform (DFT) can be expressed as

$$R_i(n) = e^{j\frac{2\pi}{N}Gq(i-1)\lambda} H(|n - q\lambda|_N) S_i(|n - q\lambda|_N) + Z_i(n), \tag{86}$$

where $H(k)$ denotes the N-point DFT of the channel impulse response of length L_c i.e., $h(l), l = 0, 1, \cdots, L_c - 1$, and the frequency domain samples of the ith OFDM symbol $(i \geq 1)$ are denoted by $S_i(k)$. Furthermore, $Z_i(k)$ is additive white Gaussian noise with zero mean and variance σ_z^2, and $|x|_N$ denotes the modulus operation that reduces the value of x to an integer value in the interval $[0, N - 1]$.

IFO detection is usually carried out by obtaining the number of cyclic shifts of the pilots in the frequency domain. The method in [31], takes advantage of the correlation between the received differentially modulated PN sequence and the original PN sequence

to detect IFO as

$$\widehat{\lambda} = \arg\max_{\widetilde{\lambda}} B(\widetilde{\lambda}) \tag{87}$$

with

$$B(\widetilde{\lambda}) = \left| \sum_{n=0}^{N/2-1} R_1^*(qn) R_2(qn) S_1(|qn - q\widetilde{\lambda}|_N) S_2^*(|qn - q\widetilde{\lambda}|_N) \right|. \tag{88}$$

where the trial values of λ is represented by $\widetilde{\lambda}$. Later an IFO metric was presented in [71] using an ML approach. This metric given by

$$\Phi_{conv}(\widetilde{\lambda}) \;=\; \sum_{l=0}^{L_c-1} \left| \sum_{m=0}^{\frac{N}{q}-1} R_1(mq) S_1^*\Big(|mq \,-\, q\widetilde{\lambda}|_N\Big) e^{j\frac{2\pi}{N}mql} \right|^2, \;\; \widetilde{\lambda} \;=\; 0, 1, 2, ..., \frac{N}{q} \,-\, 1. \tag{89}$$

has a better detection capability at the cost of increased computational complexity. To detect IFO in the presence of residual timing offset, the authors in [72] resorted to the ambiguity function. Their method can detect both residual timing offset θ and IFO as

$$(\widehat{\theta}, \widehat{\lambda}) = \arg\max_{(\widetilde{\theta}, \widetilde{\lambda})} A(\widetilde{\theta}, \widetilde{\lambda}) \tag{90}$$

where in the frequency domain

$$A(\widetilde{\theta}, \widetilde{\lambda}) = \sum_{\tau=\widetilde{\theta}}^{\widetilde{\theta}+L_c-1} \left| \sum_{n=0}^{N-1} R(|n+\widetilde{\lambda}|_N) S_1^*(n) e^{j\frac{2\pi}{N}n\tau} \right|^2, \; \widetilde{\lambda} = 0, 1, 2, ..., \frac{N}{q}-1, \; \widetilde{\theta} = 0, 1, \cdots, G-L_c \tag{91}$$

or equivalently in the time domain

$$A(\widetilde{\theta}, \widetilde{\lambda}) = \sum_{\tau=\widetilde{\theta}}^{\widetilde{\theta}+L_c-1} \left| \sum_{k=0}^{N-1} r(k) s_1^*(k-\tau) e^{-j\frac{2\pi}{N}kq\widetilde{\lambda}} \right|^2, \; \widetilde{\lambda} = 0, 1, 2, ..., \frac{N}{q}-1, \; \widetilde{\theta} = 0, 1, \cdots, G-L_c. \tag{92}$$

Note that θ is the residual timing offset, and $\widetilde{\theta}$ and $\widehat{\theta}$ are respectively its trial and estimate

values.

3.3.2 New Method for IFO Detection

In previous works, IFO is usually obtained by detection of the number of cyclic shifts of the pilot samples in the frequency domain. The IFO detector searches among the possible trial values of IFO, and picks the trial IFO that produces the maximum correlation value. Therefore, to be able to detect IFO that is as large as the OFDM bandwidth i.e., $-\frac{N}{2q} < \lambda \leq \frac{N}{2q}$, the previous detectors need to check N/q trial values of IFO [71].

Here, we propose an IFO detection method which does not need checking the entire N/q trial values of IFO while covering the entire OFDM bandwidth. To this end, we first express the IFO as

$$\lambda = \gamma \times \alpha + \beta \tag{93}$$

where α is a parameter to be designed that takes an integer value in the range $1 < \alpha \leq N/q$, and $\beta = |\lambda|_\alpha$ denotes the modulus operation that reduces the value of λ to an integer value in the interval $[0, \alpha - 1]$. Since α is known, detection of λ is transformed into the detection of γ and β. The parameter γ takes integer values in the range $[0, \frac{N}{q\alpha} - 1]$, and the parameter β takes integer values in the range $[0, \alpha - 1]$. If γ and β can be detected separately, we only need to check $\frac{N}{q\alpha}$ trial values for detection of γ, and α trial values for detection of β. Therefore, the total number of trial values of IFO that are needed to be checked reduces to

$$T(\alpha) = \alpha + \frac{N}{q\alpha} \tag{94}$$

which is much smaller than N/q trial values as required in previous methods.

First, we explain how α is designed, and next explain how β and γ are detected. To find the value of α that results in the least number of trial values of IFO, we diffrentiate $T(\alpha)$ with respect to α and put it equal to zero: $\frac{\partial T(\alpha)}{\partial \alpha} = 1 - \frac{N}{q\alpha^2} = 0$. It immediately follows

that the optimal value of α can be expressed as

$$\alpha_{opt} = \sqrt{\frac{N}{q}}. \tag{95}$$

In practical systems, N is chosen to be a power of 2, and the parameter q is also a power of 2. Therefore, $\frac{N}{q} = 2^m$ where m is an integer. N and/or q can be designed in the way that m is an even integer and therefore $\sqrt{N/q}$ is an integer. (Note that in the case where m is an odd integer, α_{opt} can be chosen as $\alpha_{opt} = \sqrt{2N/q}$). Substituting (95) into (94), the least possible number of trial values of IFO can be expressed as

$$T(\alpha_{opt}) = 2\sqrt{N/q}, \tag{96}$$

which is again significantly smaller than N/q trial values of previous methods.

In what follows, we propose approaches for finding β and γ. It will be demonstrated that β is detected using a periodic pilot sequence in the frequency domain, and γ is detected using an aperiodic pilot sequence in the frequency domain.

We indicate the trial values of β and γ by $\widetilde{\beta}$ and $\widetilde{\gamma}$, respectively, and use the conventional IFO metric $\Phi_{conv}(\widetilde{\lambda})$ that is based on the correlation of the received pilot samples in the frequency domain with the cyclically shifted original pilot samples. Such a metric is presented in [71] as

$$\Phi_{conv}(\widetilde{\lambda}) \;=\; \sum_{l=0}^{L_c-1} \left| \sum_{m=0}^{\frac{N}{q}-1} R_1(mq) S_1^*\Big(|mq - q\widetilde{\lambda}|_N\Big) e^{j\frac{2\pi}{N}mql} \right|^2, \;\; \widetilde{\lambda} \;=\; 0, 1, 2, ..., \frac{N}{q} - 1. \tag{97}$$

To be able to detect β separately from γ, the metric used for detection of β, denoted by $\Lambda_1(\widetilde{\beta})$, should not be affected by $\widetilde{\gamma}$. In other words, we need to have a metric which transforms the value of the cyclic shift $\widetilde{\lambda}$ into $\widetilde{\beta}$, i.e. when the signal is cyclically shifted

by $\widetilde{\lambda}$, it appears as if it is cyclically shifted by $\widetilde{\beta}$:

$$\Lambda_1(\widetilde{\beta}) = \Lambda_1(\widetilde{\lambda}) = \Lambda_1(\widetilde{\beta} + \widetilde{\gamma} \times \alpha). \tag{98}$$

Condition (98) is satisfied if the metric $\Lambda_1(\widetilde{\beta})$ is periodic with period α. To make the metric in (97) periodic with period α, we use a preamble whose pilot samples in the frequency domain are periodic, i.e.,

$$S_1(mq) = S_1\Big(|mq - q\alpha|_N\Big),\ m = 0, 1, ..., N/q - 1. \tag{99}$$

Thus, the new detector is proposed as

$$\Lambda_1(\widetilde{\beta}) \quad = \quad \sum_{l=0}^{L_c-1} \left| \sum_{m=0}^{\frac{N}{q}-1} R_1(mq) S_1^*\Big(|mq - q\widetilde{\beta}|_N\Big) e^{j\frac{2\pi}{N}mql} \right|^2, \widetilde{\beta} \quad = \quad 0, 1, ..., \alpha \ - \ 1. \tag{100}$$

It is straightforward to check that when (99) holds, (100) satisfies (98). The detected value of β can then be expressed as

$$\widehat{\beta} = \arg\max_{\widetilde{\beta}} \{\Lambda_1(\widetilde{\beta})\}. \tag{101}$$

Now that β is detected, we are left with the detection of γ. We replace β in (93) with $\widehat{\beta}$, and next maximize the metric (97) over the values of $\widetilde{\gamma}$. Since the values of $\widetilde{\gamma}$ can result in trial values of λ i.e., $\widetilde{\lambda} = \widetilde{\gamma} \times \alpha + \widehat{\beta}$ that can be as large as the OFDM bandwidth, the pilot sequence used for detection of γ should be aperiodic. A periodic pilot sequence limits the IFO detection range to the value of periodicity. For example, the periodic pilot sequence in (99) results in the detection range $[0, \alpha - 1]$. Thus, for detection of γ, we make use of a second OFDM preamble generated from an aperiodic pilot sequence in the frequency

domain. The detector of γ is proposed as

$$\widehat{\gamma} = \arg\max_{\widetilde{\gamma}} \{\Lambda_2(\widetilde{\gamma})\}, \tag{102}$$

where

$$\Lambda_2(\widetilde{\gamma}) = \sum_{l=0}^{L_c-1} \left| \sum_{m=0}^{\frac{N}{q}-1} R_2(mq) \times S_2^*\Big(|mq - q \times (\widetilde{\gamma}\alpha + \widehat{\beta})|_N\Big) e^{j\frac{2\pi}{N}mql} \right|^2, \widetilde{\gamma} = 0, 1, ..., \frac{N}{q\alpha} - 1. \tag{103}$$

Finally, IFO is detected as

$$\widehat{\lambda} = \widehat{\gamma} \times \alpha + \widehat{\beta}. \tag{104}$$

The following comments are worth mentioning:

1) In the discussion above, for simplicity we have assumed that the two pilot sequences for detection of β and γ were placed in separate preambles. If we want to place both of the pilot sequences in the same preamble, and the total number of pilot samples be equal to that in previous methods, we simply should replace $S_1(n)$, $S_2(n)$ and $R_2(n)$, in (100) and (103), respectively, with $\breve{S}_1(n)$, $\breve{S}_2(n)$ and $R_1(n)$, where

$$\breve{S}_1(n) = \begin{cases} S_1(m) & n = 2mq,\ m = 0, 1, ..., \frac{N}{2q} - 1 \\ 0 & \text{otherwise} \end{cases} \tag{105}$$

$$\breve{S}_2(n) = \begin{cases} S_1(m) & n = (2m+1)q,\ m = 0, 1, ..., \frac{N}{2q} - 1 \\ 0 & \text{otherwise} \end{cases}. \tag{106}$$

In this case, it is easy to show that if we have $S_1(2mq) = S_1\Big(|2mq - q\alpha|_N\Big)$, then $\Lambda_1(\widetilde{\beta})$ in (100) becomes periodic with period α.

2) Although we designed new IFO metrics based on the metric in [71], our algorithm can be applied to a wide range of metrics. In this respect, the general procedure is as

follows: i) model the IFO as (93), ii) obtain the optimum value of α according to (95) (note that α can also take other values as long as the condition (98) is satisfied.), iii) detect β by maximizing an IFO metric that satisfies the condition (98), and iv) detect γ by using the maximum value of an IFO metric that utilizes an aperiodic pilot sequence in the frequency domain.

3) In deriving (94) and (95), we considered the general case of $-\frac{N}{2q} < \lambda \leq \frac{N}{2q}$ being the range of IFO and $\frac{N}{q}$ being the total number of trial values of IFO. If in particular applications we have a range limit for IFO, for example $-\frac{\lambda_1}{2} < \lambda \leq \frac{\lambda_1}{2}$, then $\frac{N}{q}$ is simply replaced with λ_1, and (94) and (95) are changed accordingly.

3.3.3 Performance Evaluation

3.3.3.1 Computational Complexity

Now, we evaluate the computational complexity of the new method. According to [71], the metric in (97) needs $4\frac{N}{q}(L_c - 1)$ real multiplications for each trial value, and it has to check $\frac{N}{q}$ trial values. Thus, it needs $\frac{N}{q}\left(4\frac{N}{q}(L_c - 1)\right)$ real multiplications in total for IFO detection. In contrast, when the proposed method using two preambles is applied to the metric in [71], the number of trial values can be reduced to $2\sqrt{\frac{N}{q}}$. Therefore, the new method reduces the total number of real multiplications to $2\sqrt{\frac{N}{q}}\left(4\frac{N}{q}(L_c - 1)\right)$. In case of the proposed scheme utilizing only one preamble, the number of pilot samples for both of the new metrics is one-half of the conventional metric, and therefore, for each trial value of IFO, our method needs $2\frac{N}{q}(L_c - 1)$ real multiplications. Hence, our method needs $\left(2\sqrt{\frac{N}{q}}\right)2\frac{N}{q}(L_c - 1)$ real multiplications in total. Following the same reasoning, it is also found that the number of additions for our method in cases of using two preambles and one preamble are respectively $\frac{1}{2}\sqrt{\frac{N}{q}}$ and $\sqrt{\frac{N}{q}}$ lower than that of [71]. The complexity of the new method when applied to other metrics can also be obtained following a similar manner.

In Table 4, we list the complexity of different IFO detection methods. In this table,

for [31], [71] and the proposed methods using the metrics in [31] and [71], we have given numerical examples for $N = 512$, $L_c = 5$, and $q = 2$. For [72] and the proposed methods using the metric in [72], we have used the parameters $N = 256$, $L_c = 5$, $q = 1$, and $G = 16$. It is obvious that the complexity of the new approach is significantly reduced .

3.3.3.2 Probability of Correct Detection

In this subsection, we analyze the performance of the new method in terms of the probability of correct detection and BER. First, we investigate the probability of correct detection of the proposed method that is given by

$$P_c\left(\widehat{\lambda}=\lambda\right)=P\left(\widehat{\beta}=\beta,\widehat{\gamma}=\gamma\right)= \\ P\Bigg(\Lambda_1(\widetilde{\beta}_0)<\Lambda_1(\beta),\ldots,\Lambda_1(\widetilde{\beta}_{\alpha-2})<\Lambda_1(\beta),\Lambda_2(\widetilde{\gamma}_0)<\Lambda_2(\gamma),\ldots,\Lambda_2(\widetilde{\gamma}_{\frac{N}{q\alpha}-2})<\Lambda_2(\gamma)\Bigg) \tag{107}$$

where $\widetilde{\beta}_n, 0 \leq n \leq \alpha-2$ and $\widetilde{\gamma}_n, 0 \leq n \leq \frac{N}{q\alpha}-2$ are the nth wrong trial values of β and γ, respectively. For the conventional IFO detectors, for example [71], the probability of correct detection can be expressed as

$$P_M\left(\widehat{\lambda}=\lambda\right)=P\left(\Lambda_{conv}(\widetilde{\lambda}_0)<\Lambda_{conv}(\lambda),\ldots,\Lambda_{conv}(\widetilde{\lambda}_{\frac{N}{q}-2})<\Lambda_{conv}(\lambda)\right) \tag{108}$$

where $\widetilde{\lambda}_n, 0 \leq n \leq \frac{N}{q}-2$ is the nth wrong trial value of λ. We consider an AWGN channel, and assume that $\Lambda_{conv}(\widetilde{\lambda})$, $\Lambda_1(\widetilde{\beta})$, and $\Lambda_2(\widetilde{\gamma})$ have the same number of pilot samples.

Adopting an approach similar to [142], it is found that the three detectors at correct detected values, i.e., $\Lambda_{conv}(\lambda)$, $\Lambda_1(\beta)$, and $\Lambda_2(\gamma)$ have an identical non-central Chi-squared distribution with the mean $(\frac{N}{q})^2\sigma_s^4+\frac{N}{q}\sigma_s^2\sigma_z^2$ and variance $2(\frac{N}{q})^2\sigma_s^4\sigma_z^4+4(\frac{N}{q})^3\sigma_s^6\sigma_z^2$. It is assumed that $E\{S_i(n)\}=0$ and $E\big[Re\{S_i(n)\}^2\big]=E\big[Im\{S_i(n)\}^2\big]=\frac{\sigma_s^2}{2}$, where $E\{S_i(n)\}$, $Re\{S_i(n)\}$ and $Im\{S_i(n)\}$ are the expectation, real and imaginary parts of $S_i(n)$, respectively, and σ_s^2 denotes the variance of $S_i(n)$ at pilot positions. On the other

Table 4: Complexity of IFO Detectors

Method	**Real Multiplications**	**Example**
[31]	$\frac{N}{q}(4\frac{N}{q})$	262144
Proposed using [31] (two preambles)	$2\sqrt{\frac{N}{q}}(4\frac{N}{q})$	32768
Proposed using [31] (one preamble)	$2\sqrt{\frac{N}{q}}(2\frac{N}{q})$	16384
[71]	$\frac{N}{q}\left(4\frac{N}{q}\,(L_c-1)\right)$	1048576
Proposed using [71] (two preambles)	$2\sqrt{\frac{N}{q}}\left(4\frac{N}{q}(L_c-1)\right)$	131072
Proposed using [71] (one preamble)	$2\sqrt{\frac{N}{q}}\left(2\frac{N}{q}(L_c-1)\right)$	65536
[72]	$4G(N+\frac{N}{2}\log_2 N)$	81920
Proposed using [72] (two preambles)	$4G\sqrt{N}+8L_cN+$ $(2G+2L_c)\sqrt{N}\log_2\sqrt{N}$	13952
Proposed using [72] (one preamble)	$4G(5\sqrt{N}+\frac{\sqrt{N}}{2}\log_2\frac{\sqrt{N}}{2})+$ $4L(2\sqrt{N}+\frac{\sqrt{N}}{2}\log_2\sqrt{N})$	7936

hand, at false detected values, $\Lambda_{conv}(\widetilde{\lambda}_n), 0 \leq n \leq \frac{N}{q}-2$, and $\Lambda_2(\widetilde{\gamma}_n), 0 \leq n \leq \frac{N}{q\alpha}-2$ have an identical Chi-squared distribution with the mean $\frac{N}{q}\sigma_s^4+\frac{N}{q}\sigma_s^2\sigma_z^2$ and variance $(\frac{N}{q})^2\sigma_s^8+(\frac{N}{q})^2\sigma_s^4\sigma_z^4+2(\frac{N}{q})^2\sigma_s^6\sigma_z^2$, and $\Lambda_1(\widetilde{\beta}_n), 0 \leq n \leq \alpha-2$, has a Chi-squared distribution with the mean $\frac{N}{q\alpha}\frac{N}{q}\sigma_s^4+\frac{N}{q}\sigma_s^2\sigma_z^2$ and variance $(\frac{N}{q\alpha})^2(\frac{N}{q})^2\sigma_s^8+(\frac{N}{q})^2\sigma_s^4\sigma_z^4+2\frac{N}{q\alpha}(\frac{N}{q})^2\sigma_s^6\sigma_z^2$.

According to the above analysis, the probability of correct detection of the new method is dependent on fewer events than that of [71], in view of $\alpha+\frac{N}{q\alpha}-2<\frac{N}{q}-1$. Thus, reducing the number of trial values of IFO has a favourable effect on the probability of correct detection. In other words, each wrong trial IFO, with some probability, has the potential to be detected as the true IFO. Consequently, decreasing the number of wrong trial values increases the probability of correct detection of IFO. On the other hand, using periodic pilot samples has the unfavourable effect of increasing the mean and variance of $\Lambda_1(\widetilde{\beta})$ at wrong trial IFOs.

3.3.3.3 Simulation Results

We use the Stanford University Interim (SUI) channel modeling, namely, SUI-1 [141] for modeling a multipath fading channel. The considered OFDM system has a sampling rate of 5 MHz. In the first scenario, we assume that the two pilot sequences are placed in different preambles, and compare the method in [31], the proposed method using the metric in [31], the scheme in [71] and the proposed method using the metric in [71], by using $N = 128$, $G = 16$, $q = 2$ and $\alpha = 8$. For [72] and the new method using the metric in [72], we utilize $N = 64$, $G = 8$ and $q = 1$. Furthermore, we have $\alpha = 8$ and, IFO has a uniform distribution in $(-\frac{N}{2q}, \frac{N}{2q}]$.

In Fig. 21, we show the probability of correct detection for different detectors in the absence of FFO. According to this figure, when the new method is applied to the previous IFO detectors the probability of correct detection slightly improves. As mentioned previously, this improvement is mainly due to reducing the number of wrong trial IFOs that have the potential to be wrongly detected as the correct value of IFO.

To assess the performance of the new method in the case of imperfect FFO synchronization, we consider the presence of residual FFO with a uniform distribution in $(-\frac{q}{10}, \frac{q}{10}]$. Fig. 22 depicts the probability of correct detection of different detectors in such a case. Once again, the new method slightly improves the performance as shown by the improved probability of correct detection. It is worth mentioning that, as discussed in the previous section, the new method also offers a remarkable reduction in the complexity.

In the second scenario, we consider the case of the new method utilizing only one preamble and the total number of pilots being equal to that of previous methods. In this respect, for the method in [31], the proposed method using the metric in [31], the scheme in [71] and the proposed method using the metric in [71], we have $N = 512$, $G = 64$, $q = 2$ and $\alpha = 16$. Furthermore, for [72] and the new method using the metric in [72], we utilize $N = 256$, $G = 32$ and $q = 1$.

Figs. 23 and 24 illustrate the probability of correct detection for different detectors in

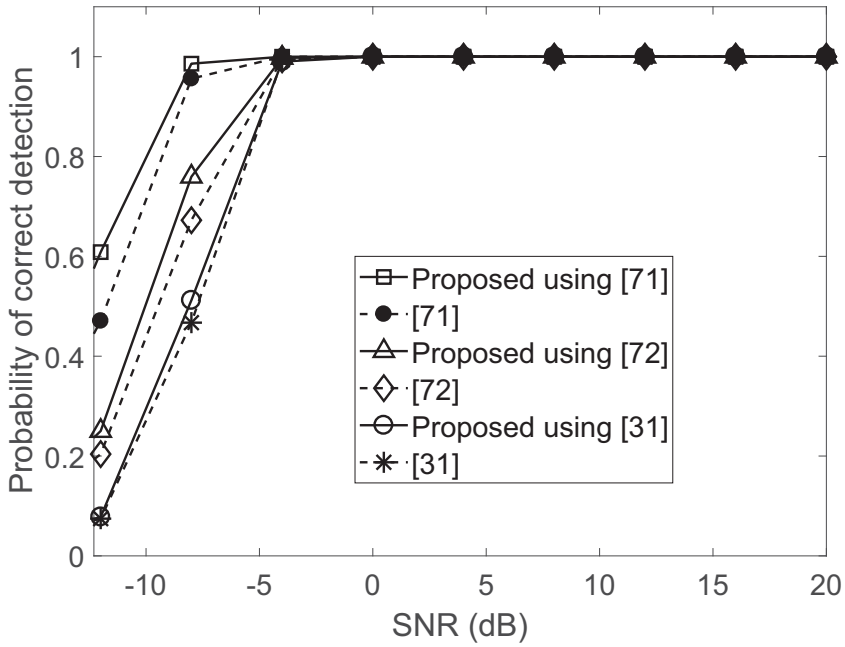

Figure 21: Probability of detection of different methods using two preambles in the absence of residual FFO.

the absence and presence of residual FFO, respectively. It is observed that the performance for $SNR \geq -4dB$ is the same as that of previous methods, and the performance slightly deteriorates for $SNR < -4dB$. Note that in a reasonable and practical SNR range, say $SNR > 0dB$, our method has the same performance but with a significant reduction of complexity.

The BER performance for quadrature phase-shift keying (QPSK) signals for the same total number of pilots is illustrated in Fig. 25. It is observed that the utilization of the new method results in the same BER as that of the previous methods for $SNR \geq 0dB$. We indeed have expected the same BER performance because as depicted in Figs. 21, 22, 23 24, the new IFO metrics have the same performance in terms of the probability of correct detection for $SNR \geq -4dB$.

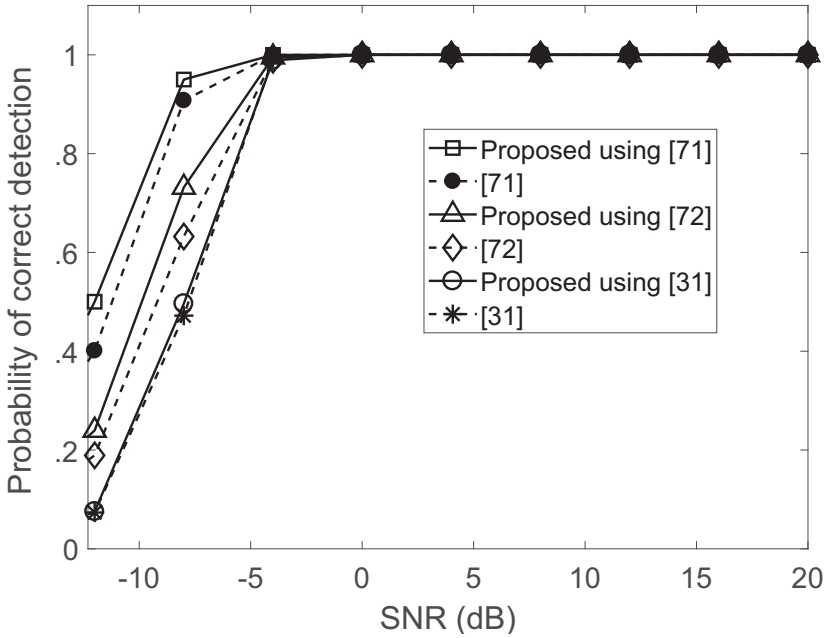

Figure 22: Probability of detection of different methods using two preambles in the presence of residual FFO.

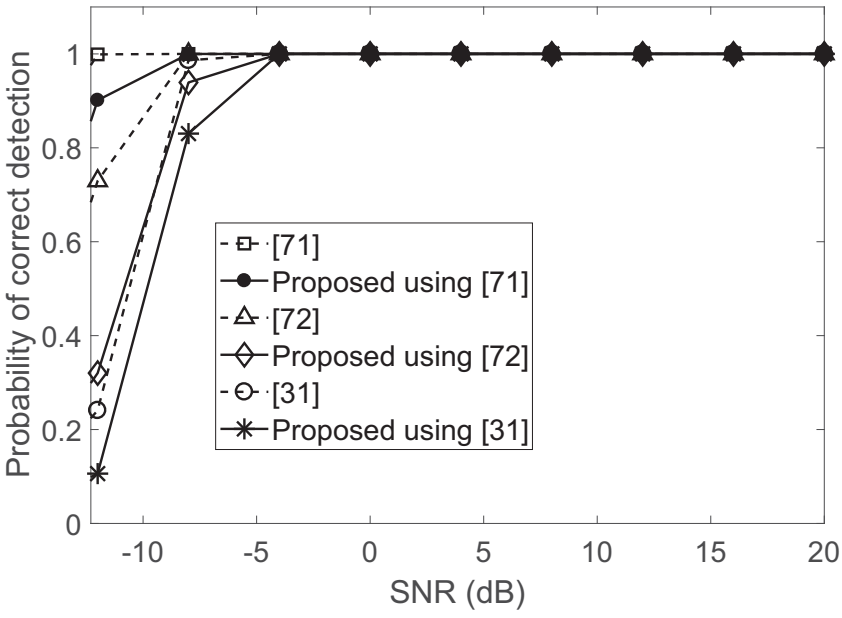

Figure 23: Probability of detection of different methods using one preamble in the absence of residual FFO.

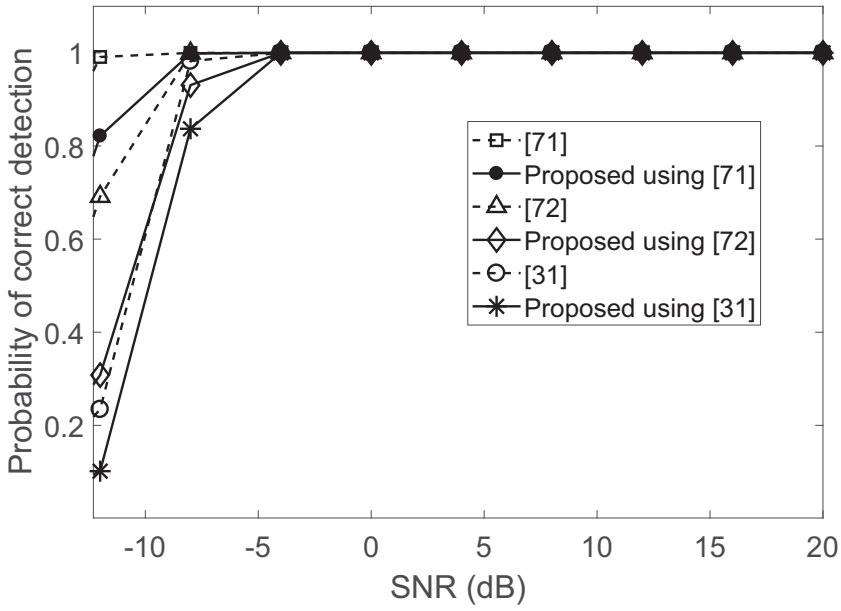

Figure 24: Probability of detection of different methods using one preamble in the presence of residual FFO.

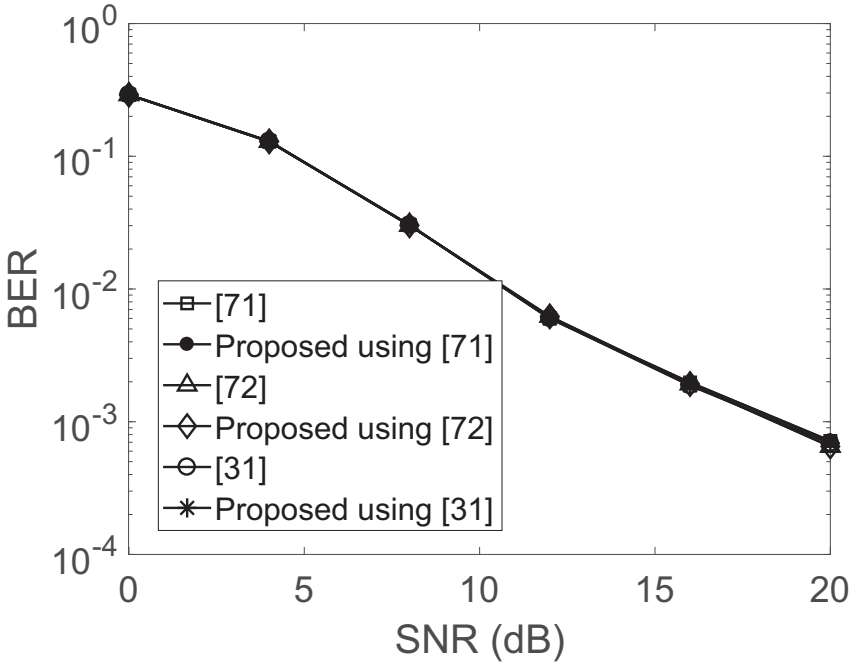

Figure 25: BER for different IFO detectors.

3.4 Conclusion

In this chpater, two novel timing metrics using n-th order statistics have been proposed and their performances have been assessed using different criteria such as class-separability, robustness to CFO, and computational complexity. The effect of increasing the order of the timing metric has also been investigated. A new method for coarse estimation of the start of the frame has been presented. The performances of the new timing schemes in multipath frequency selective fading channels have been evaluated through computer simulations. By evaluating the false alarm and missed-detection probabilities, it has been demonstrated that the new methods remarkably improve the detection performance as compared with previous methods. Finally, it has been shown that the new methods also offer a significant reduction in the ISI probability.

Moreover, we have presented a new IFO detection method for OFDM systems that significantly reduces the number of trial values of IFO. The new method transforms IFO detection into the detection of two new integer parameters. It was also indicated that the new scheme can be applied to many of previously presented IFO metrics. The computational complexity, detection probability and BER of the new scheme was studied in comparison with previous methods. It has been shown that the new method can significantly reduce the computational complexity while maintaining almost the same detection performance as compared to previous schemes.

Chapter 4

Synchronization in Doubly Selective Channels

4.1 Introduction

This chapter presents some new schemes for synchronization and channel estimation in OFDM based systems that operate in doubly selective channels. First, joint timing, CFO and channel estimation in OFDM system is performed adopting an ML approach. Second, a novel method is proposed for joint CFO and channel estimation using Kalman and particle filtering. Third, timing and frequency synchronization is addressed in OFDMA system based on ML criterion. Finally, joint CFO and channel estimation in MIMO OFDMA systems is performed using particle and Kalman filtering. It is shown that the new schemes can considerably improve the performance compared with previous methods.

4.2 Proposed Method: Joint Timing, CFO and Channel Estimation in OFDM Systems Using ML

4.2.1 Preliminaries

We consider the time-domain samples $s(k)$ of an OFDM symbol generated by taking the inverse discrete Fourier transform (IDFT) of the frequency domain samples $S(n)$, expressed in the matrix form as

$$\mathbf{s} = \mathbf{F}^H \mathbf{S}. \tag{109}$$

Here, $\mathbf{F}$ is the discrete Fourier transform (DFT) matrix whose element on the nth row and kth column is $[\mathbf{F}]_{n,k} = \frac{1}{\sqrt{N}} \exp(-j\frac{2\pi}{N}kn)$, $\mathbf{s} = [s(0), s(1), \cdots, s(N-1)]^T$, and $\mathbf{S} = [S(0), S(1), \cdots, S(N-1)]^T$. Furthermore, $(\cdot)^T$ and $(\cdot)^H$ respectively denote the transpose and the Hermitian transpose operations. After inserting the cyclic prefix (CP) of length G, the OFDM symbol is transmitted through an L_c-tap time-varying channel. The l-th channel tap ($l = 0, 1, \cdots, L_c - 1$) at time k, denoted by $h(k, l)$, can be expressed using BEM as [102]

$$h(k, l) = \sum_{q=0}^{Q-1} b_{k,q} c_{q,l} = \mathbf{b}_k^T \mathbf{c}_l, \tag{110}$$

where $b_{k,q}$ is the qth basis function at time k, $c_{q,l}$ is the qth BEM coefficient corresponding to the lth channel tap, and Q is the number of basis functions. Furthermore, $\mathbf{b}_k = [b_{k,0}, b_{k,1}, \cdots, b_{k,Q-1}]^T$ and $\mathbf{c}_l = [c_{0,l}, c_{1,l}, \cdots, c_{Q-1,l}]^T$. Defining $\mathbf{h}_l = [h(0, l), h(1, l), \cdots, h(N-1, l)]^T$, $\mathbf{B} = \mathbf{I}_{L_c} \otimes [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{N-1}]^T$ (where $\mathbf{I}_{L_c}$ indicates an $L_c \times L_c$ identity matrix, and $\otimes$ denotes Kronecker product), and $\mathbf{c} = [\mathbf{c}_0^T, \mathbf{c}_1^T, \cdots, \mathbf{c}_{L_c-1}^T]^T$, we have

$$\mathbf{h} = [\mathbf{h}_0^T, \mathbf{h}_1^T, \cdots, \mathbf{h}_{L_c-1}^T]^T = \mathbf{B}\mathbf{c}. \tag{111}$$

Let the received signal vector of length N at time instant d be denoted by $\mathbf{r}^d = [r(d), r(d+1), \cdots, r(d+N-1)]^T$. Using the above channel model, the received signal

vector at the timing instant of reception of the preamble i.e. $d = \theta$, can be expressed as

$$\mathbf{r}^{\theta} = e^{j\frac{2\pi}{N}G\varepsilon}\Gamma(\varepsilon)\mathbf{A}\mathbf{c} + \mathbf{z} \tag{112}$$

where θ and ε are respectively timing and frequency offsets, $\Gamma(\varepsilon) = \text{diag}\{e^{j\frac{2\pi}{N}k\varepsilon}, k = 0, 1, \cdots, N-1\}$, $\mathbf{A} = [\mathbf{s}^{(0)}\breve{\mathbf{B}}, \mathbf{s}^{(1)}\breve{\mathbf{B}}, \cdots, \mathbf{s}^{(L_c-1)}\breve{\mathbf{B}}]$, and $\mathbf{s}^{(l)}$ is a diagonal matrix whose elements are obtained by cyclically shifting the vector $\mathbf{s}$, i.e., $\mathbf{s}^{(l)} = \text{diag}\{[s(N-l), s(N-l+1), \cdots, s(N-1), s(0), s(1), \cdots, s(N-l-1)]\}$. Furthermore, we have $\breve{\mathbf{B}} = [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{N-1}]^T$ and $\mathbf{z}$ is the additive white Gaussian noise vector with zero mean and covariance matrix $\sigma_z^2\mathbf{I}_N$.

4.2.2 Joint Timing, Frequency Offset, and Channel Estimation

As mentioned previously, existing preamble-aided methods such as [107] have not tackled the timing estimation problem in doubly selective channels. Here, we propose a joint estimation method for obtaining the timing, frequency offset and doubly selective channel. To this end, we consider the signal model (112), and assume that $\mathbf{s}$ is a preamble, and therefore $\mathbf{A}$ is known at the receiver. We denote $\widetilde{\theta}$, $\tilde{\varepsilon}$, and $\tilde{\mathbf{c}}$ as the trial values of θ, ε, and $\mathbf{c}$, respectively. The probability density function of $\mathbf{r}^{\widetilde{\theta}}$ given $\widetilde{\theta}$, $\widetilde{\varepsilon}$, and $\tilde{\mathbf{c}}$ is given by

$$f(\mathbf{r}^{\widetilde{\theta}}|\widetilde{\theta}, \tilde{\varepsilon}, \tilde{\mathbf{c}}) = \frac{1}{(\pi\sigma_z^2)^N}\exp\left\{ - \frac{\left\|\mathbf{r}^{\widetilde{\theta}} - e^{j\frac{2\pi}{N}G\widetilde{\varepsilon}}\Gamma(\widetilde{\varepsilon})\mathbf{A}\widetilde{\mathbf{c}}\right\|^2}{\sigma_z^2} \right\}. \tag{113}$$

Thus, the log-likelihood function of $\widetilde{\theta}$, $\widetilde{\varepsilon}$, and $\widetilde{\mathbf{c}}$ after discarding the terms independent of $\widetilde{\theta}$, $\widetilde{\varepsilon}$, and $\widetilde{\mathbf{c}}$ can be expressed as

$$\Lambda(\widetilde{\theta}, \tilde{\varepsilon}, \tilde{\mathbf{c}}) = -\left\|\mathbf{r}^{\widetilde{\theta}} - e^{j\frac{2\pi}{N}G\widetilde{\varepsilon}}\Gamma(\widetilde{\varepsilon})\mathbf{A}\widetilde{\mathbf{c}}\right\|^2. \tag{114}$$

From (114), the estimates of θ, ε, and $\mathbf{c}$ denoted by $\widehat{\theta}$, $\widehat{\varepsilon}$, and $\widehat{\mathbf{c}}$, can be expressed as

$$(\hat{\theta}, \hat{\varepsilon}, \hat{\mathbf{c}}) = \arg \max_{\widetilde{\theta}, \widetilde{\varepsilon}, \widetilde{\mathbf{c}}} \Lambda(\widetilde{\theta}, \widetilde{\varepsilon}, \widetilde{\mathbf{c}}). \tag{115}$$

Unfortunately, obtaining $(\widehat{\theta}, \widehat{\varepsilon}, \widehat{\mathbf{c}})$ above needs an exhaustive search over $\widetilde{\theta}$, $\widetilde{\varepsilon}$, and $\widetilde{\mathbf{c}}$ which is not practical. Thus, we adopt the following approach. First, we keep $\widetilde{\theta}$ and $\widetilde{\varepsilon}$ fixed and maximize $\Lambda(\widetilde{\theta}, \widetilde{\varepsilon}, \widetilde{\mathbf{c}})$ with respect to $\widetilde{\mathbf{c}}$, yielding the following estimate of $\mathbf{c}$ which is dependent on $\widetilde{\theta}$ and $\widetilde{\varepsilon}$

$$\widehat{\mathbf{c}}(\widetilde{\theta}, \widetilde{\varepsilon}) = e^{-j\frac{2\pi}{N}G\widetilde{\varepsilon}}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\Gamma^H(\widetilde{\varepsilon})\mathbf{r}^{\widetilde{\theta}}. \tag{116}$$

By replacing $\tilde{\mathbf{c}}$ in (114) with $\widehat{\mathbf{c}}(\widetilde{\theta}, \widetilde{\varepsilon})$, we have

$$\Lambda(\widetilde{\theta}, \widetilde{\varepsilon}) = -\left\| \mathbf{r}^{\widetilde{\theta}} - e^{j\frac{2\pi}{N}G\widetilde{\varepsilon}}\Gamma(\widetilde{\varepsilon})\mathbf{A}\widehat{\mathbf{c}}(\widetilde{\theta}, \widetilde{\varepsilon}) \right\|^2. \tag{117}$$

At the next step, we keep $\widetilde{\theta}$ fixed and search for ε. Discarding the terms in $\Lambda(\widetilde{\theta}, \widetilde{\varepsilon})$ that are indepndent of $\widetilde{\varepsilon}$ results in

$$\Upsilon(\widetilde{\theta}, \widetilde{\varepsilon}) = \mathbf{r}^{\widetilde{\theta}^H}\Gamma(\widetilde{\varepsilon})\mathbf{T}\Gamma^H(\widetilde{\varepsilon})\mathbf{r}^{\widetilde{\theta}} \tag{118}$$

where $\mathbf{T} = \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$. Adopting an approach similar to [73], (118) can be rewritten as

$$\begin{aligned}\Upsilon(\widetilde{\theta}, \widetilde{\varepsilon}) = \sum_{m=0}^{N-1} \big(r^{\widetilde{\theta}}(m)\big)^* T(m,m) r^{\widetilde{\theta}}(m) + \sum_{m=0}^{N-2}\sum_{n=m+1}^{N-1} e^{j\frac{2\pi}{N}(m-n)\widetilde{\varepsilon}} \big(r^{\widetilde{\theta}}(m)\big)^* T(m,n) r^{\widetilde{\theta}}(n) \\ + \sum_{m=0}^{N-2}\sum_{n=m+1}^{N-1} e^{j\frac{2\pi}{N}(n-m)\widetilde{\varepsilon}} r^{\widetilde{\theta}}(m) T(m,n) \big(r^{\widetilde{\theta}}(n)\big)^*\end{aligned} \tag{119}$$

where a^* denotes the complex conjugate of a, and $T(k,n)$ denotes the element of matrix $\mathbf{T}$ on the kth row and nth column. The first term on the right hand side of (119) is not

dependent on $\widetilde{\varepsilon}$ and can be dropped. Consequently, we have

$$\Upsilon(\widetilde{\theta},\widetilde{\varepsilon}) = \mathrm{Re}\left\{ \sum_{m=0}^{N-2} \sum_{n=m+1}^{N-1} e^{j\frac{2\pi}{N}(m-n)\widetilde{\varepsilon}} \times \left(r^{\widetilde{\theta}}(m)\right)^* T(m,n) r^{\widetilde{\theta}}(n) \right\} \tag{120}$$

where $\mathrm{Re}\{a\}$ denotes the real part of a. We can express (120) equivalently as

$$\Upsilon(\widetilde{\theta},\widetilde{\varepsilon}) = \mathrm{Re}\left\{ \sum_{k=0}^{N-1} \chi^{\widetilde{\theta}}(k) e^{-j\frac{2\pi}{N}k\widetilde{\varepsilon}} \right\} \tag{121}$$

where $\chi^{\widetilde{\theta}}(k) = \sum_{m=0}^{N-1-k} \left(r^{\widetilde{\theta}}(m)\right)^* T(m,k+m) r^{\widetilde{\theta}}(k+m)$. Therefore, the estimate of ε can be written as

$$\widehat{\varepsilon}(\widetilde{\theta}) = \arg\max_{\widetilde{\varepsilon}} \Upsilon(\widetilde{\theta},\widetilde{\varepsilon}) \tag{122}$$

which can be implemented using the FFT as seen from (121). The next step is replacing $\widetilde{\varepsilon}$ with $\widehat{\varepsilon}(\widetilde{\theta})$ in (117) and omitting the terms of (117) that are independent of $\widetilde{\theta}$, leading to the following timing metric

$$\Lambda(\widetilde{\theta}) = \left(\mathbf{r}^{\widetilde{\theta}}\right)^H \Gamma\left(\hat{\varepsilon}(\widetilde{\theta})\right) \mathbf{T} \Gamma^H\left(\hat{\varepsilon}(\widetilde{\theta})\right) \mathbf{r}^{\widetilde{\theta}} - \left\|\mathbf{r}^{\widetilde{\theta}}\right\|^2. \tag{123}$$

Finally, the timing estimate is obtained as

$$\widehat{\theta} = \arg\max_{\widetilde{\theta}} \Lambda(\widetilde{\theta}). \tag{124}$$

Putting back $\widehat{\theta}$ at (122), the estimate of CFO is obtained, and using $\widehat{\theta}$ and CFO estimates in (116), the BEM coefficients and consequently channel (using (3)) is estimated.

4.2.3 Complexity Reduced Implementation

The CFO metric in (121) can be computed using FFT, which significantly reduces the computational complexity in comparison with that in [107]. However, to increase the accuracy of the estimator, we have to increase the size of FFT which may considerably

increase the computational complexity of the CFO estimator. For example, for having a resolution of $\frac{1}{M}$ where M is a power of 2, we have to conduct $MN\log_2(MN)$ multiplications for FFT calculation at each timing instant. The increase in the complexity of CFO estimator will naturally increase the overall complexity of joint timing, frequency offset and channel estimation. Taking into account this fact, in this subsection, we first propose a multi-level grid search algorithm for reduction of the complexity of CFO estimation. As a result, the computational complexity of CFO estimation with resolution of $\frac{1}{M}$ reduces from $MN\log_2(MN)$ to $N\log_2 N$. Next, we explain how this method can be incorporated in joint timing, frequency and channel estimation.

We know from [143] that the new CFO estimator gives a reliable estimate using an FFT length of multiples of N i.e. $L_{FFT} = 2^{n-1}N, n = 1, 2, \cdots$, where L_{FFT} is the FFT length. For having a resolution of $\frac{1}{M}$, instead of directly computing all the coefficients of an MN-point FFT, we first compute the coefficients of an N point FFT, and the bin of the coefficient with maximum real part is selected. We denote this CFO estimate by $\widehat{\varepsilon}_1$, which has the resolution of 1 i.e., $\widehat{\varepsilon}_1 - \frac{1}{2} \leq \varepsilon \leq \widehat{\varepsilon}_1 + \frac{1}{2}$. Next, we increase the resolution by a factor of 2 by computing a $2N$ point FFT. Since we know from the previous CFO estimation that CFO is in the region $\widehat{\varepsilon}_1 - \frac{1}{2} \leq \varepsilon \leq \widehat{\varepsilon}_1 + \frac{1}{2}$, we only need to check coefficients of the $2N$ point FFT corresponding to frequencies in the set i.e. $\widetilde{\varepsilon} = \widehat{\varepsilon}_1 - \frac{1}{2}, \widehat{\varepsilon}_1, \widehat{\varepsilon}_1 + \frac{1}{2}$. The bin of the coefficient with maximum real value indicates the CFO estimate at the second step which we denote by $\widehat{\varepsilon}_2$. The CFO estimation region will now be $\widehat{\varepsilon}_2 - \frac{1}{4} \leq \varepsilon \leq \widehat{\varepsilon}_2 + \frac{1}{4}$. Then, again we increase the length of the FFT by a factor of 2 i.e. $L_{FFT} = 4N$, and only compute the FFT coefficients corresponding to frequencies in the region $\widehat{\varepsilon}_2 - \frac{1}{4} \leq \varepsilon \leq \widehat{\varepsilon}_2 + \frac{1}{4}$, i.e., $\widetilde{\varepsilon} = \widehat{\varepsilon}_2 - \frac{1}{4}, \widehat{\varepsilon}_2, \widehat{\varepsilon}_2 + \frac{1}{4}$, and the coefficient whose real value is maximum indicates the new CFO estimate $\widehat{\varepsilon}_3$ resulting in the new estimation region $\widehat{\varepsilon}_3 - \frac{1}{8} \leq \varepsilon \leq \widehat{\varepsilon}_3 + \frac{1}{8}$. By continuing this procedure, at the $n = m$th step, we have $\widehat{\varepsilon}_n$ with the resolution $\frac{1}{M} = \frac{1}{2^m}$, where $m = \log_2 M$.

We now compute the computational complexity of the new method. At the first step,

we need to compute all N coefficients of an N-point FFT which has the complexity of $N\log_2 N$. From the second step and so on, (i.e. for nth step where $1 < n \leq m$), we only need to compute 2 new coefficients of an $2^{n-1}N$ point FFT, which needs $2\log_2(2^{n-1}N)$ multiplications. Therefore, in total we need to compute $N\log_2 N + 2\log_2(2N) + 2\log_2(4N) + \cdots + 2\log_2(2^{m-1}N) = N\log_2 N + m(m-1) + 2(m-1)\log_2 N$ multiplications where $m = \log_2 M$. Thus, as compared to the number of multiplications of the previous methods $MN\log_2(MN)$, the new method significantly reduces the computational complexity.

Now to reduce the overall complexity of joint timing, CFO and channel estimation, we propose to perform the estimation process in two steps. In the first step, initial joint estimates of timing, channel and CFO are obtained in which only ε_1 as mentioned above is estimated along with timing and BEM coefficients. After this stage is accomplished, refined estimates of CFO (ε_n for $n > 1$ using the above algorithm), channel and timing are obtained utilizing the timing, CFO, and BEM coefficients estimates of the previous step.

It is worth mentioning that in the proposed reduced complexity method, for simplicity, we proposed that $\varepsilon_n,\ n = 1$ be calculated as coarse CFO estimate. Evidently, if we want to have more accurate estimation in initial joint timing, CFO and channel estimation step, we can simply use $\varepsilon_n,\ 1 < n < m$ for coarse CFO estimation.

4.2.4 Performance Evaluation

4.2.4.1 Computational Complexity

In Table 5, we have shown the computational complexities of different estimators in terms of the number of multiplications for each timing instant, where ζ_g and ζ_i respectively denote the number of grid searches in coarse CFO estimation and the number of iterations for fine CFO estimation in [107]. It is assumed that in [107] and our method, we have precomputed $\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$. In the third column of the table, we have provided a numerical example where $N = 128$, $Q = 3$, and $L_c = 5$. For [107], we have considered the optimistic case that

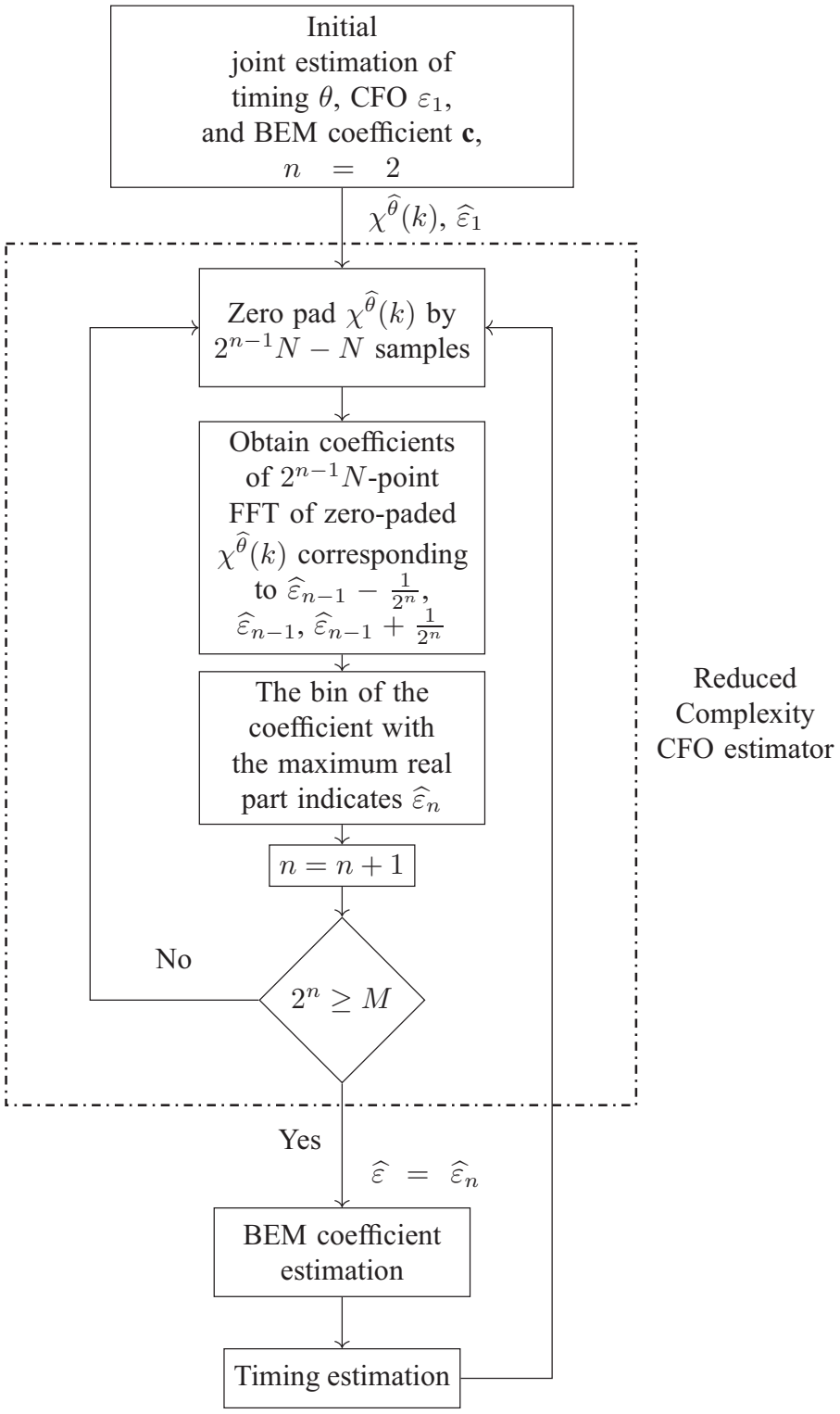

Figure 26: Block diagram of the proposed scheme.

Table 5: Complexity of Timing, Frequency and Channel Estimators in OFDM Systems

Method	**Number of Multiplications**	Example
CFO estimation [107]	$\zeta_g(3N^2+N)+\zeta_i(9N^2+3N)$	6603520
Proposed CFO estimation	$N(N+1)+(\log_2 M)(\log_2 M-1)$ $+2(N+\log_2 M-1)\log_2 N$	18108
Channel estimation [107]	$2QN+L_c^2NQ$	10368
Proposed channel estimation	$2QN+L_c^2NQ$	10368
Proposed timing estimation	$3N^2+2N$	49408

$\zeta_g = N$ and $\zeta_i = 2$. For our method, we have set the resolution to $M = N^3$ resulting in approximate frequency spacing of 5×10^{-7}. Evidently, the new CFO estimation method significantly reduces the computational complexity.

4.2.4.2 Simulation Result

In this section, we evaluate the MSE performance of the new method. As mentioned in [109], the hybrid CRB (HCRB) matrix for CFO and channel is given by

$$\mathbf{HCRB} = \begin{bmatrix} \mathbf{HCRB_h} & \mathbf{HCRB}_{12} \\ \mathbf{HCRB}_{21} & \mathrm{HCRB}_\varepsilon \end{bmatrix}$$

where $\mathbf{HCRB_h}$ and $\mathrm{HCRB}_\varepsilon$ are the bounds on the MSE of $\mathbf{h}$ and ε , respectively, and $\mathbf{HCRB}_{12}$ and $\mathbf{HCRB}_{21}$ are the cross terms. In [109], $\mathbf{HCRB}$ is derived as, $\mathbf{HCRB} = \Big(\mathrm{blkdiag}\big(\frac{1}{\sigma_z^2}\Omega^H\Omega, \frac{2}{\sigma_z^2}\mathrm{Tr}\big\{\mathbf{R_h}^{-1}\Omega'^H\Omega'\big\}\big), \mathrm{blkdiag}\big(\mathbf{R_h}^{-1}, 0\big)\Big)^{-1}$, where $\mathrm{blkdiag}\{\mathbf{x},\mathbf{y}\}$ is a block diagonal matrix with the matrices $\mathbf{x}$ and $\mathbf{y}$ on its main diagonal, $\mathrm{Tr}(\cdot)$ denotes the trace operation, $\Omega = \Gamma(\varepsilon)\breve{\mathbf{A}}$ and $\Omega' = \frac{\partial}{\partial\varepsilon}(\Omega)$. Furthermore, we have $\breve{\mathbf{A}} = [\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_{L_c-1}]$ and $\mathbf{R_h}$ is the correlation matrix of $\mathbf{h}$. Thus, the lower bound for MSE of ε and $\mathbf{h}$ are respectively $\mathrm{HCRB}_\varepsilon$ and $\overline{\mathbf{HCRB}}_\mathbf{h} = \frac{1}{L_c N}\mathrm{Tr}\{\mathbf{HCRB_h}\}$.

An OFDM system with $N = 128$ and $G = 16$ is considered. The carrier frequency

and sampling time are $f_c = 1.2$ GHz and $T_s = 1.87 \times 10^{-6}$ seconds, respectively. CFO is set to $\varepsilon = 1.3$, and the considered mobile velocity is 260 km/hr. Therefore, we have $f_d N T_s = 0.25$ where f_d is the Doppler frequency defined as $f_d = \frac{v f_c}{c}$ with c being the light speed. We have generated a multipath Rayleigh fading channel using Jake's model [144]. The channel has four taps ($L_c = 4$) with time spacings of 1.87μs. The average path gain of the lth tap is $e^{-\frac{l}{10}}$, and $Q = 2$.

In Figs. 27 and 28, we have respectively depicted the normalized MSE (NMSE) of the new channel and CFO estimation in comparison with those obtained by the conventional method [107] for different values of timing offset (TO) using GCE-BEM. In this figure, the label "Conv, SC" indicates the NMSE when coarse synchronization for [107] is achieved using [31]. It is evident that the CFO and channel estimator in [107] is sensitive to timing offset. On the other hand, the CFO and channel estimate of the new method is not affected by the timing offset, and their NMSE is considerably lower than that of [107] in the presence of TO. In this figure, we have also shown the HCRB. Note that the HCRB for CFO is not necessarily tight as the result of taking the expectation of hybrid Fisher information matrix with respect to the channel statistics in deriving **HCRB** [109].

In order to evaluate the performance of the timing estimator, in Fig. 29, we have demonstrated the MSE of the proposed timing estimation scheme using different BEMs in comparison with the ML method [16] and the threshold-based methods [48] and [49] which are presented for slow fading channels. Evidently, the new method developed for fast fading channels, has a remarkably lower MSE. Furthermore, the MSE of the methods in [48] and [49] increases with SNR, which can be explained by the increase in the unreliability of the thresholds for first channel tap detection in [48] and [49].

We have demonstrated NMSE of CFO and channel estimates of the new method for different values of SNR in Fig. 30. According to this figure, the indicated BEMs have approximately the same performance. Fig. 31 depicts the timing, CFO and channel estimates for the proposed method in different high mobility situations using DPSS-BEM.

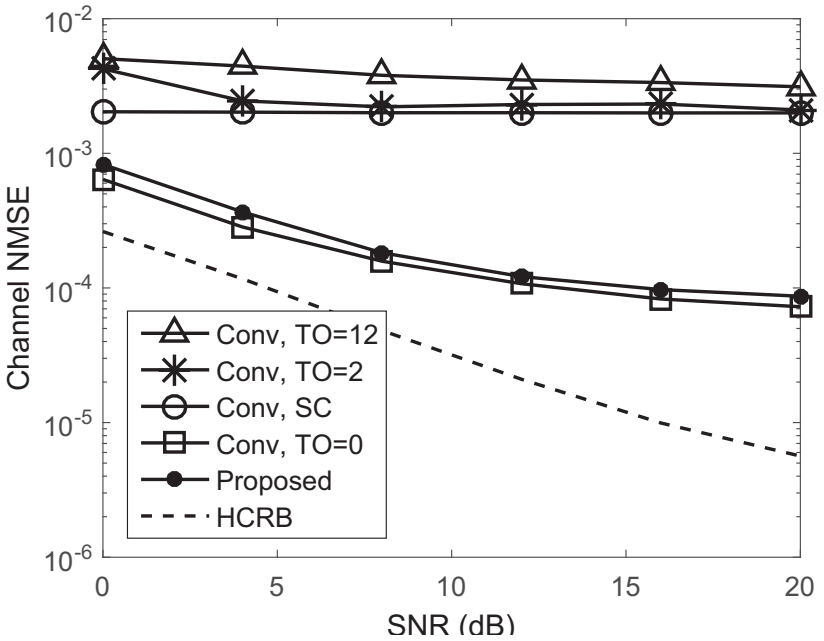

Figure 27: NMSE of different channel estimators.

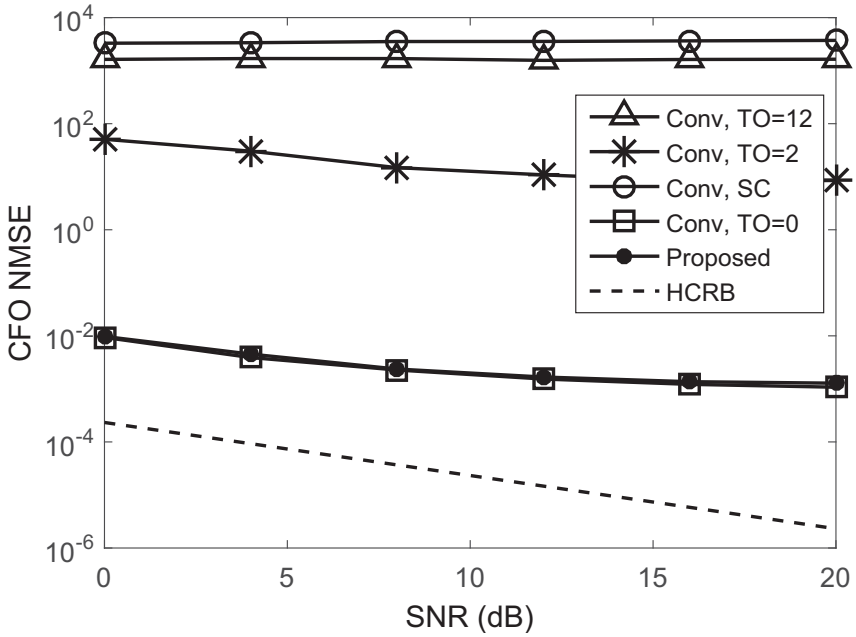

Figure 28: NMSE of different CFO estimators.

It is observed that as $f_d N T_s$ increases the NMSE improves. This is due to the fact that for lower mobility situations there are more errors because of over parametrization by BEM [145].

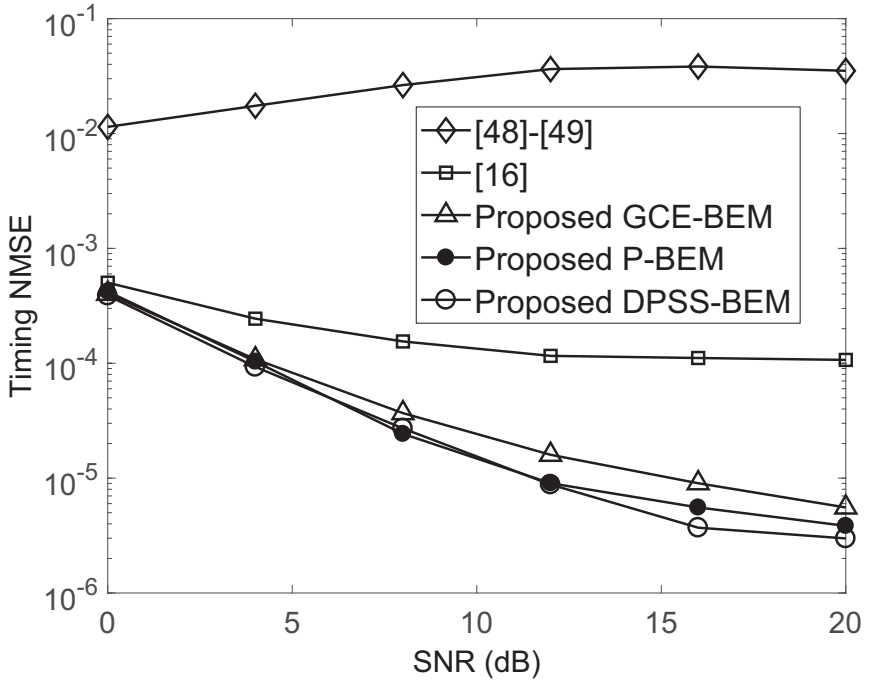

Figure 29: NMSE of timing estimates of different methods.

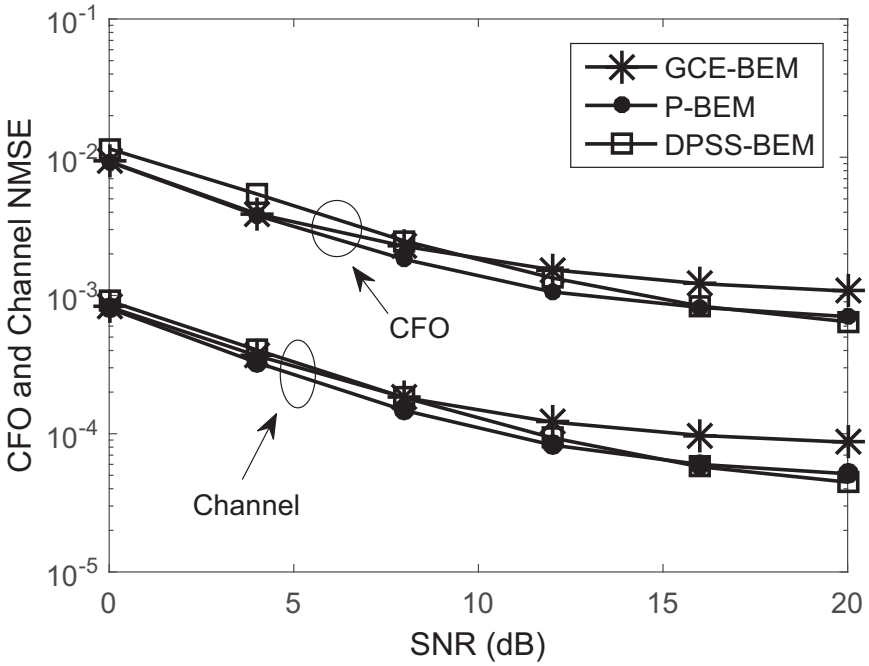

Figure 30: NMSE of CFO and channel estimators versus SNR for different BEM.

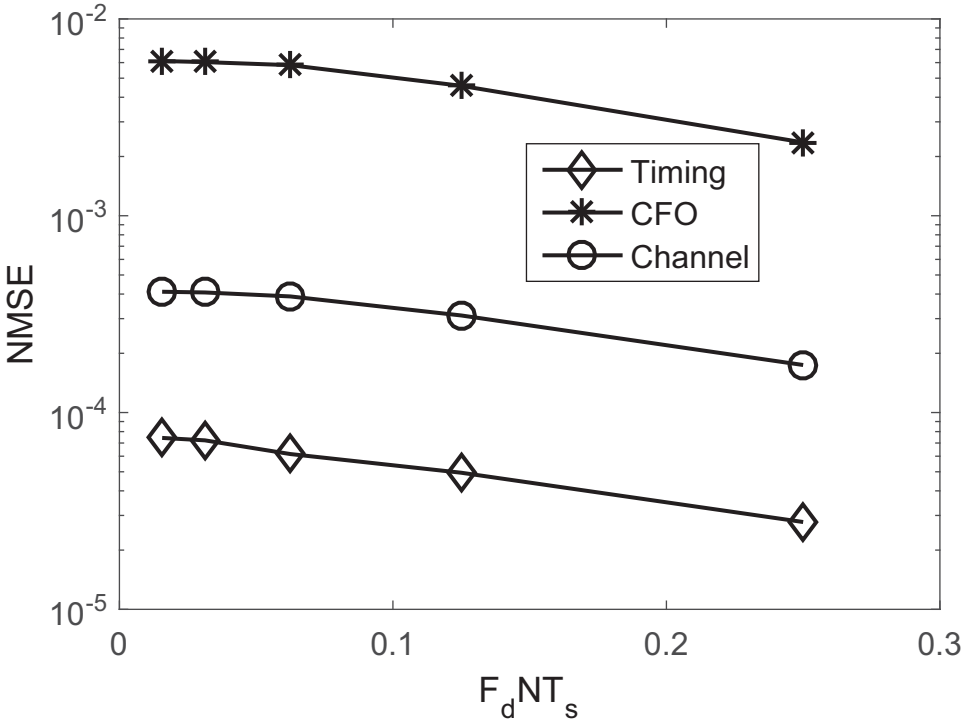

Figure 31: NMSE of timing, CFO and channel estimators versus $f_d N T_s$.

4.3 Proposed Method: Joint CFO and Channel Estimation in OFDM Systems Using Particle and Kalman Filtering

4.3.1 Preliminaries

4.3.1.1 Kalman Filtering

To concisely present the Kalman filter, we refer to [146]. The Kalman filter model assumes that the state of a system at a time t is evolved from the prior state at time $t-1$ according to

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t + \mathbf{w}_t \tag{125}$$

where $\mathbf{x}_t$ denotes the state vector containing the parameters of interest at time t, $\mathbf{u}_t$ is the vector containing any control inputs, $\mathbf{F}_t$ indicates the state transition matrix which applies the effect of each system state parameter at time $t-1$ on the system state at time t, $\mathbf{B}_t$ denotes the control input matrix which applies the effect of each control input parameter in the vector $\mathbf{u}_t$ on the state vector, and $\mathbf{w}_t$ is the vector containing the process noise terms for each parameter in the state vector. The process noise is considered to be drawn from a zero mean multivariate normal distribution with the covariance matrix Q_t. Measurements of the system can also be expressed, according to the model

$$\mathbf{y}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{v}_t \tag{126}$$

where $\mathbf{y}_t$ is the vector of measurements, $\mathbf{H}_t$ denotes the transformation matrix that maps the state vector parameters into the measurement domain, and $\mathbf{v}_t$ is the measurement noise vector for each observation in the measurement vector. The measurement noise is considered to be zero mean Gaussian white noise with covariance $\mathbf{R}_t$.

The Kalman filter provides an algorithm to determine an estimate of $\mathbf{x}_t$ by combining models of the system and noisy measurements of certain parameters or linear functions of parameters. The estimates of the parameters of interest in the state vector are provided by pdfs. The Kalman filter is based on Gaussian pdfs. To fully describe the Gaussian functions, we need to know their variances and covariances and these are stored in the covariance matrix P_t. The Kalman filter algorithm involves two stages: prediction and measurement update. The standard Kalman filter equations for the prediction stage are

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t\hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t\mathbf{u}_t \tag{127}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t\mathbf{P}_{t-1|t-1}\mathbf{F}_t^H + \mathbf{Q}_t, \tag{128}$$

and the measurement update equations are given by

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}) \tag{129}$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{H}_t\mathbf{P}_{t|t-1} \tag{130}$$

where

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^H(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^H + \mathbf{R}_t)^{-1}. \tag{131}$$

It is worth mentioning that if the signal and noise are jointly Gaussian, it is shown that the Kalman filter is an optimal mimimum mean square error (MMSE) estimator [147].

4.3.1.2 Particle Filtering

In particle filtering, the pdf $p(\varepsilon|\mathbf{r}_{1:i})$ is approximated by random measures defined by particles and weights associated to each particle [148]. Let the ηth particle corresponding to the ith OFDM symbol and its weight, respectively, be denoted by $\varepsilon_i^{(\eta)}$ and $w_i^{(\eta)}$. The particles are generated from a distribution $\pi(\cdot)$ called importance function, and their weights are obtained as

$$w_i^{(\eta)} = \frac{p(\varepsilon_{1:i}^{(\eta)}|\mathbf{r}_{1:i})}{\pi(\varepsilon_{1:i}^{(\eta)}|\mathbf{r}_{1:i})}. \tag{132}$$

where $\varepsilon_i^{(\eta)} = \{\varepsilon_1^{(\eta)}, \varepsilon_2^{(\eta)}, \cdots, \varepsilon_i^{(\eta)}\}$ and $w_{1:i}^{(\eta)} = \{w_1^{(\eta)}, w_2^{(\eta)}, \cdots, w_i^{(\eta)}\}$. The optimal importance function conditioned on both $\varepsilon_{1:i-1}^{(\eta)}$ and $\mathbf{r}_{1:i}$ is given by [148]

$$\pi(\varepsilon_{1:i}^{(\eta)}|\mathbf{r}_{1:i}) = p(\varepsilon_i|\varepsilon_{1:i-1}^{(\eta)}, \mathbf{r}_{1:i}) \tag{133}$$

and the corresponding weights can be obtained sequentially as [148]

$$w_i^{(\eta)} = w_{i-1}^{(\eta)}p(\mathbf{r}_i|\varepsilon_{i-1}^{(\eta)}). \tag{134}$$

Thus, $p(\varepsilon|\mathbf{r}_{1:i})$ is approximated as

$$p(\varepsilon_{1:i}^{(\eta)}|\mathbf{r}_{1:i}) = \sum_{\eta=0}^{N_\eta - 1} w_i^{(\eta)} \delta(\varepsilon_{1:i} - \varepsilon_{1:i}^{(\eta)}). \tag{135}$$

where N_η is the number of particles.

Using Kalman filtering and the above-mentioned particle filtering, in what follows, we describe the proposed method for joint estimation of CFO and doubly selective channel in details.

4.3.2 Joint CFO and Channel Estimation in OFDM Systems

To estimate CFO and BEM coefficients, we estimate the posterior distribution function (PDF) $p(\mathbf{c}_i, \varepsilon|\mathbf{r}_{1:i})$ which can be expressed as

$$p(\mathbf{c}_{1:i}, \varepsilon|\mathbf{r}_{1:i}) = p(\mathbf{c}_i|\mathbf{r}_{1:i}, \varepsilon) p(\varepsilon|\mathbf{r}_{1:i}) \tag{136}$$

where $\mathbf{c}_{1:i} = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_i\}$ and $\mathbf{r}_{1:i} = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_i\}$. Thus, estimation of $p(\mathbf{c}_{1:i}, \varepsilon|\mathbf{r}_{1:i})$ can be carried out by obtaining $p(\mathbf{c}_{1:i}|\mathbf{r}_{1:i}, \varepsilon)$ and $p(\varepsilon|\mathbf{r}_{1:i})$. The PDF $p(\varepsilon|\mathbf{r}_{1:i})$ is estimated using particle filtering, and $p(\mathbf{c}_{1:i}|\mathbf{r}_{1:i}, \varepsilon)$ is obtained using Kalman filtering. This approach which aims at marginalizing out the linear dynamics and estimating the non-linear parameters using particle filtering is called marginalized particle filtering [134] or Rao-Blackwellized particle filtering [98].

First, we aim at generating particles with the PDF given in (133). Adopting an approach similar to [149, 150], we approximate this PDF by a beta distribution. Since the random values generated from a beta distribution are in the range $[0,\ 1]$, and CFO is assumed to be in the range $[-0.5,\ 0.5]$, a new variable $\xi_i^{(\eta)}$ is used whose relationship with ε is $\xi_i^{(\eta)} = \varepsilon_i^{(\eta)} + 0.5$, and we show the random variables generated by the beta distribution

by $\xi_i^{(\eta)}$. The parameters of the PDF $\beta(\xi, U_i, V_i)$ are defined as

$$U_i = \bar{\xi}_i\Big(\frac{\bar{\xi}_i(1-\bar{\xi}_i)}{\sigma_{\xi_i}^2} - 1\Big) \tag{137}$$

$$V_i = (1-\bar{\xi}_i)\Big(\frac{\bar{\xi}_i(1-\bar{\xi}_i)}{\sigma_{\xi_i}^2} - 1\Big) \tag{138}$$

where $\bar{\xi}_i = \sum_{\eta=1}^{N_\eta} w_{i-1}^{(\eta)} \xi_{i-1}^{(\eta)}$ and $\sigma_{\xi_i}^2 = \sum_{\eta=1}^{N_\eta} w_{i-1}^{(\eta)} \big(\xi_{i-1}^{(\eta)} - \bar{\xi}_i\big)^2$.

Next, for each particle, the mean and covariance of the BEM coefficients are predicted utilizing the prediction step of Kalman filtering as follows

- **Prediction Step of Kalman Filtering:**

$$\hat{\mathbf{c}}_{i|i-1}^{(\eta)} = \mathbf{A}_c \hat{\mathbf{c}}_{i-1|i-1}^{(\eta)} \tag{139}$$

$$\mathbf{P}_{i|i-1}^{(\eta)} = \mathbf{A}_c \mathbf{P}_{i-1|i-1}^{(\eta)} \mathbf{A}_c^H + \mathbf{U} \tag{140}$$

where $\hat{\mathbf{c}}_{i|i-1}^{(\eta)} = \mathrm{E}\{\mathbf{c}_i | \mathbf{r}_{1:i-1}, \hat{\varepsilon}_{1:i-1}^{(\eta)}\}$, $\mathbf{P}_{i|i-1}^{(\eta)} = \mathrm{cov}\{\mathbf{c}_i | \mathbf{r}_{1:i-1}, \hat{\varepsilon}_{1:i-1}^{(\eta)}\}$ with $\mathrm{E}\{\cdot\}$ and $\mathrm{cov}\{\cdot\}$ denoting the expectation and covariance operations, respectively.

Now that we have an estimate for the BEM coefficients, the weights are updated for each particles using (134) as follows

$$w_i^{(\eta)} = w_{i-1}^{(\eta)} \mathcal{N}(\mathbf{r}_i, \mathbf{m}_i^{(\eta)}, \mathbf{f}_i^{(\eta)}) \tag{141}$$

where $\mathbf{m}_i^{(\eta)}$ and $\mathbf{f}_i^{(\eta)}$ are the estimates of the mean and variance of the received signal using particles and the BEM coefficients predicted using Kalman filtering, and we have

$$\mathbf{m}_i^{(\eta)} = \Gamma(\varepsilon_i^{(\eta)}) \mathbf{A}_i \hat{\mathbf{c}}_{i|i-1}^{(\eta)} \tag{142}$$

$$\mathbf{f}_i^{(\eta)} = \Gamma(\varepsilon_i^{(\eta)}) \mathbf{A}_i \hat{\mathbf{c}}_{i|i-1}^{(\eta)} \mathbf{P}_{i|i-1}^{(\eta)} \Big(\Gamma(\varepsilon_i^{(\eta)}) \mathbf{A}_i \hat{\mathbf{c}}_{i|i-1}^{(\eta)}\Big)^H + \sigma_z^2 \mathbf{I}. \tag{143}$$

Now that the prediction step of Kalman filtering is carried out, it is time to perform the

correction step of Kalman filtering to obtain an estimate for BEM coefficients as follows

- **Correction Step of Kalman Filtering:**

$$\mathbf{K}_i^{(\eta)} = \mathbf{P}_{i|i-1}^{(\eta)} \Big(\Gamma(\varepsilon_i^{(\eta)}) \mathbf{A}_i \hat{\mathbf{c}}_{i|i-1}^{(\eta)} \Big)^H \mathbf{f}_i^{(\eta)^{-1}} \tag{144}$$

$$\hat{\mathbf{c}}_{i|i}^{(\eta)} = \hat{\mathbf{c}}_{i|i-1}^{(\eta)} + \mathbf{K}_i^{(\eta)} (\mathbf{r}_i - \Gamma(\varepsilon_i^{(\eta)}) \mathbf{A}_i \hat{\mathbf{c}}_{i|i-1}^{(\eta)}) \tag{145}$$

$$\mathbf{P}_{i|i}^{(\eta)} = \mathbf{P}_{i|i-1}^{(\eta)} \Big(\mathbf{I} - \mathbf{K}_i^{(\eta)} \hat{\mathbf{c}}_{i|i-1}^{(\eta)} \Big) \tag{146}$$

After performing the above procedure for all particles, CFO and BEM coefficients are estimated as

$$\hat{\mathbf{c}}_i = \sum_{p=1}^{N_\eta} w_i^{(\eta)} \hat{\mathbf{c}}_{i|i}^{(\eta)}, \tag{147}$$

$$\hat{\varepsilon_i} = \sum_{p=1}^{N_\eta} w_i^{(\eta)} \varepsilon_i^{(\eta)}. \tag{148}$$

The proposed algorithm is summarized in Table 6.

4.3.3 Initialization and Resampling

The proposed algorithm is initialized as follows. The particles $\xi_0^{(\eta)}, p = 1, 2, \cdots, N_\eta$ are generated from a uniform distribution $\mathcal{U}(0, 1)$ which has values in the range $[0, 1]$, and $w_0^{(\eta)} = \frac{1}{N_\eta},\ p = 1, 2, \cdots, N_\eta$. It is also assumed that the first OFDM symbol is known at the receiver.

It is well known that particle filters suffer from degeneracy. It means that after a few iterations, the values of a large portion of the weights are very close to zero. Thus, a large computational complexity is devoted to updating the weights that almost have no contributions [151]. If it is desired to apply the proposed method over several OFDM symbols, resampling can be used to eliminate the particles with small weights and replicates particles with large weights [148]. Note that effective particle size $Neff = \frac{1}{\sum_{p=1}^{N_\eta} (w_i^{(\eta)})^2} < \frac{N_\eta}{2}$ can

Table 6: Proposed Algorithm for OFDM Systems Based on Particle and Kalman Filtering

Proposed Algorithm

Draw $\xi_i^{(\eta)} \sim \mathcal{U}(0,1)$ for $\eta = 1, 2, \cdots, N_\eta$

Set $w_i^{(\eta)} = \frac{1}{N_\eta}$ for $\eta = 1, 2, \cdots, N_\eta$

Set $\hat{\mathbf{c}}_{0|0}^{(\eta)} = \mathbf{0}_{Q\times 1}$, $\mathbf{P}_{0|0}^{(\eta)} = \mathbf{R}_{c_l}[0]$ for $\eta = 1, 2, \cdots, N_\eta$

For i=1:N_i

Compute $\bar{\xi}_i$, $\sigma_{\xi_i}^2$, U_i, and V_i

For p=1:N_η

Perform Kalman prediction: $\hat{\mathbf{c}}_{i|i-1}^{(\eta)}$, $\mathbf{P}_{i|i-1}^{(\eta)}$

Draw $\xi_i^{(\eta)} \sim \beta(\xi; U_i, V_i)$,

Update weights $w_i^{(\eta)} = w_{i-1}^{(\eta)} \mathcal{N}(\mathbf{r}_i, \mathbf{m}_i^{(\eta)}, \mathbf{f}_i^{(\eta)})$

Conduct Kalman correction: $\hat{\mathbf{c}}_{i|i}^{(\eta)}$, $\mathbf{P}_{i|i}^{(\eta)}$

End

Normalize weights $w_i^{(\eta)} = \frac{w_i^{(\eta)}}{\sum_{\eta=1}^{N_\eta} w_i^{(\eta)}}$

Resample if $Neff = \frac{1}{\sum_{p=1}^{N_\eta} (w_i^{(\eta)})^2} < \frac{N_\eta}{2}$

BEM coefficients estimate: $\hat{\mathbf{c}}_i = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)} \hat{\mathbf{c}}_{i|i}^{(\eta)}$

Channel estimate: $\hat{\mathbf{h}}_i = \mathbf{B}\hat{\mathbf{c}}_i$

CFO estimate: $\hat{\varepsilon_i} = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)} \xi_i^{(\eta)} - 0.5$

End

be used as a measure of degeneracy of particles [152]. The interested readers are referred to [148, 149, 151, 152] for further details.

4.3.4 Simulation Results

To evaluate the performance of the new method, we have considered an OFDM system with $N = 128$, $G = 16$ and sampling time $T_s = 10^{-6}$. The carrier frequency and mobile velocity are $f_c = 2$ GHz and 586 km/hr, respectively, corresponding to $f_d N T_s = 0.5$ where f_d is Doppler frequency. The channel is created using Jake's model [153] and has 3 taps with delays $\tau(l) = 3lT_s,\ l = 0, 1, 2$ and corresponding average path gains $e^{-\frac{\tau(l)}{10}}$. In the first two figures, we have chosen $Q = 4$, and in the next two figures we have set $Q = 3$.

In Fig. 32, we have depicted MSE of different CFO estimators for different values of signal to noise ratio (SNR). It is evident that the new scheme has a considerably better performance than the ML accompanied with Newton-Raphson scheme in [107] and the particle filter based algorithm in [150]. The superior performance of our new method compared with [107] is due to the fact that the Newton-Raphson-based scheme in [107] does not necessarily converge to the true CFO. Moreover, the new method has better performance than the method in [150], because the latter did not take into account channel variations within an OFDM symbol.

Fig. 33 illustrates the channel estimation performance of the proposed method in comparison with the methods in [107] and [150]. Again it is obvious that the new scheme has a significantly lower MSE as compared to previous methods.

The CFO MSEs of the proposed method are demonstrated in Fig. 34 for different BEMs. Clearly, the Pol-BEM results in lower MSE compared with other BEMs, and CE-BEM has the largest MSE. This issue is due to the fact that, as is evident from Fig. 35, Pol-BEM has the lowest MSE and CE-BEM has the largest MSE for modeling the doubly selective channel.

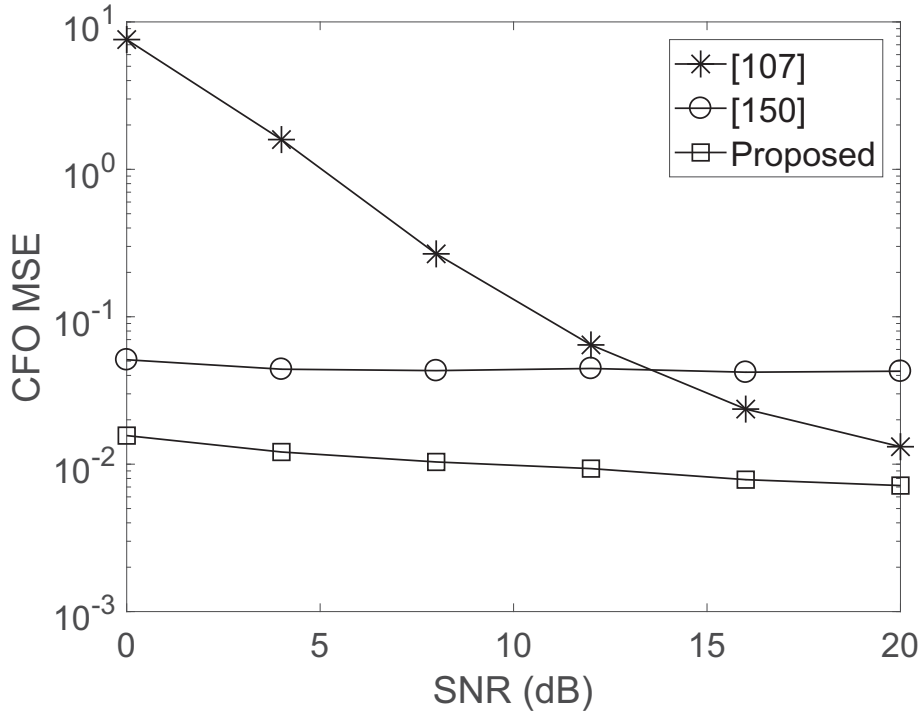

Figure 32: MSE of different CFO estimators for OFDM systems.

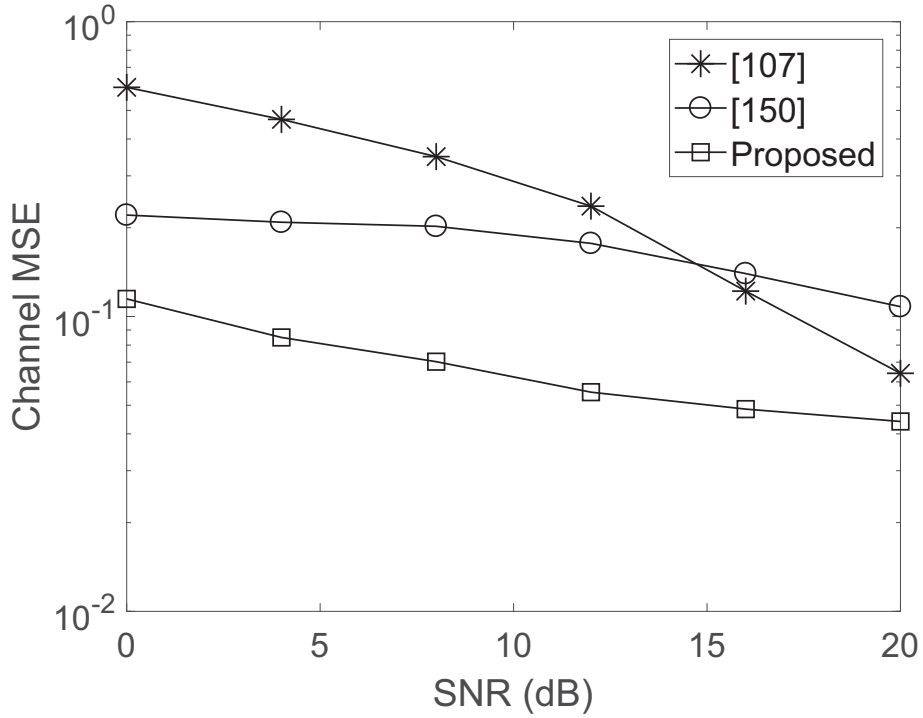

Figure 33: MSE of different channel estimators for OFDM systems.

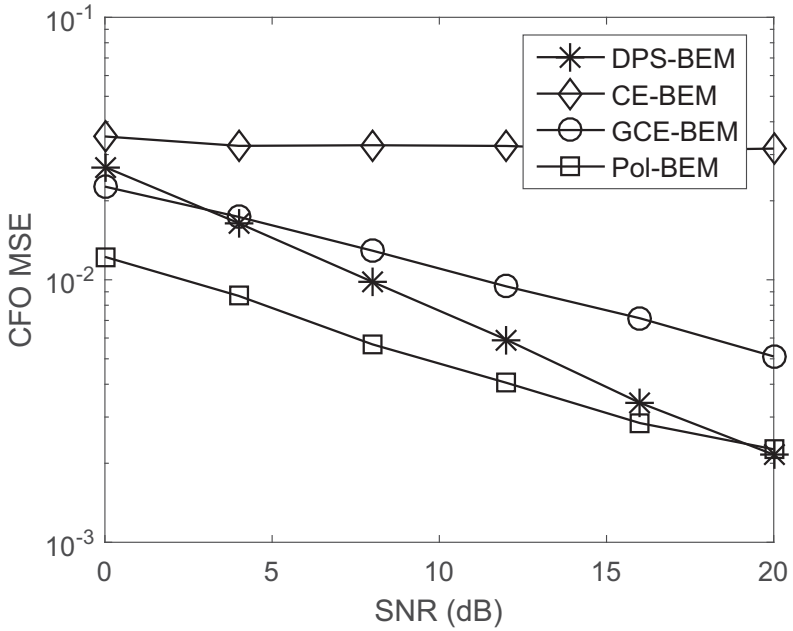

Figure 34: MSE of CFO estimates for the proposed estimator using different BEM.

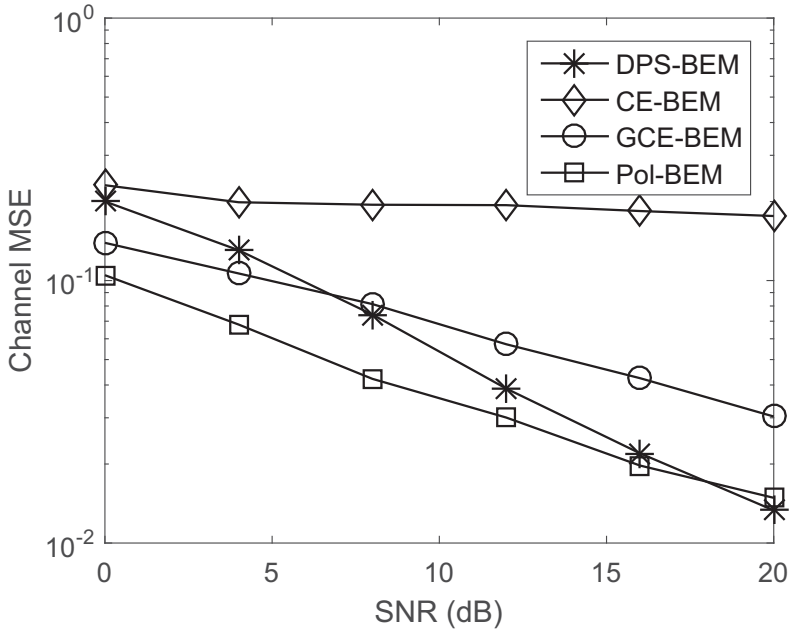

Figure 35: MSE of channel estimates for the proposed estimator using different BEM.

4.4 Proposed Method: Joint Timing, CFO and Channel Estimation in OFDMA Systems Using ML

4.4.1 Preliminaries

The doubly selective channel for each user denoted by $\mathbf{h}_{m,i}$, has NL unknown parameters, whereas each OFDM symbol has N samples. To reduce the number of unknown channel parameters, $h_{m,i}(k,l)$ can be represented using BEM as follows

$$h_{m,i}(k,l) = \sum_{q=0}^{Q-1} b_{k,q} c_{m,i,q,l} = \mathbf{b}_k^T \mathbf{c}_{m,i,l}, \tag{149}$$

where $b_{k,q}$ denotes the qth basis function at time k, $c_{m,i,q,l}$ shows the qth BEM coefficient corresponding to the lth channel tap of ith OFDM symbol of mth user. Furthermore, we have $\mathbf{b}_k = [b_{k,0}, b_{k,1}, \cdots, b_{k,Q-1}]^T$, $\mathbf{c}_{m,i,l} = [c_{m,i,0,l}, c_{m,i,1,l}, \cdots, c_{m,i,Q-1,l}]^T$, and Q is the number of basis functions. Defining $\mathbf{h}_{l,m,i} = [h_{m,i}(0,l), h_{m,i}(1,l), \cdots, h_{m,i}(N-1,l)]^T$, $\mathbf{B} = \mathbf{I}_{L_c} \otimes [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{N-1}]^T$ (where $\mathbf{I}_{L_c}$ indicates an $L_c \times L_c$ identity matrix, and $\otimes$ denotes Kronecker product), and $\mathbf{c}_{m,i} = [\mathbf{c}_{m,i,0}^T, \mathbf{c}_{m,i,1}^T, \cdots, \mathbf{c}_{m,i,L_c-1}^T]^T$, then the channel vector for mth user and ith OFDM symbol can be represented as

$$\mathbf{h}_{m,i} = \mathbf{B}\mathbf{c}_{m,i}. \tag{150}$$

Using the above channel model, we can write (17) as

$$\mathbf{r}_i = \sum_{m=1}^{M} \Gamma(\varepsilon_k) \mathbf{A}_{m,i}(\theta_m) \mathbf{c}_{m,i} + \mathbf{z}_i, \tag{151}$$

where $\mathbf{A}_{m,i}(\theta_m) = [\mathbf{s}_{m,i}^{(\theta_m)} \breve{\mathbf{B}}, \mathbf{s}_{m,i}^{(\theta_m+1)} \breve{\mathbf{B}}, \cdots, \mathbf{s}_{m,i}^{(\theta_m+L_c-1)} \breve{\mathbf{B}}]$, $\breve{\mathbf{B}} = [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{N-1}]^T$, and $\mathbf{z}_i = \sum_{m=1}^{M} \mathbf{z}_{m,i}$.

4.4.2 Joint Timing, Frequency Offset, and Channel Estimation

We assume that each user's frame is preceded by a preamble, and concentrating on the received signal corresponding to the preamble, for simplicity, we drop the subscript i in the signal model in the rest of this section. Thus, from (151), the probability distribution function of the received signal given given the trial values of $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_K]$, $\boldsymbol{\varepsilon} = [\varepsilon_1, \cdots, \varepsilon_K], \mathbf{c} = [\mathbf{c}_1^T, \cdots, \mathbf{c}_K^T]^T$ can be expressed as

$$f(\mathbf{r}|\widetilde{\theta}, \widetilde{\varepsilon}, \widetilde{\mathbf{c}}) = \frac{1}{(\pi\sigma_z^2)^N} \exp\left\{ -\frac{1}{\sigma_z^2} \left\| \mathbf{r} - \sum_{m=1}^{M} \Gamma(\widetilde{\varepsilon}_m)\mathbf{A}_m(\widetilde{\theta}_m)\widetilde{\mathbf{c}}_m \right\|^2 \right\}, \tag{152}$$

where $\widetilde{\theta}_k$, $\widetilde{\varepsilon}_k$, $\widetilde{\mathbf{c}}_k$, $\widetilde{\boldsymbol{\theta}}$, $\widetilde{\boldsymbol{\varepsilon}}$, and $\widetilde{\mathbf{c}}$ are the trial values of θ_k, ε_k, $\mathbf{c}_k$, $\boldsymbol{\theta}$, $\boldsymbol{\varepsilon}$, and $\mathbf{c}$, respectively. After discarding the irrelevant terms of (152), the ML estimates of θ, ε, and $\mathbf{c}$ can be obtained by maximizing the function

$$\Lambda(\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\varepsilon}}, \widetilde{\mathbf{c}}) = -\left\| \mathbf{r} - \sum_{m=1}^{M} \Gamma(\widetilde{\varepsilon}_m)\mathbf{A}_m(\widetilde{\theta}_m)\widetilde{\mathbf{c}}_m \right\|^2. \tag{153}$$

Direct maximization of the above equation needs an exhaustive search over the multi-dimensional space spanned by $\widetilde{\theta}_m, \widetilde{\varepsilon}_m$, and $\widetilde{\mathbf{c}}_m$ for $m = 1, 2, \cdots, M$, which is prohibitively complex in practice.

Instead of direct maximization of the likelihood function, one can utilize the EM algorithm which provides the same final result in an iterative way with reduced complexity. The EM algorithm is performed at two steps iteratively: the first step is expectation step, at which the log-likelihood of the complete data is obtained, and the second step is maximization step which maximizes the expectation with respect to unknown data [88]. However, the EM algorithm suffers from a slow convergence rate, and existence of some free parameters whose inappropriate choice can result not only in slower convergence rate but also in convergence to local stationary points. To remedy these weaknesses, SAGE algorithm is applied which has faster convergence rate with better stability [154]. SAGE

algorithm is an extension of EM algorithm in which at each iteration only one subset of unknown parameters are estimated and the rest are kept fixed [155]. For further information about the EM and SAGE algorithm, the reader is referred to [154, 156, 157]. Here, we resort to the SAGE algorithm which starts from an initial guess for θ_m, ε_m, and $\mathbf{c}_m$ denoted by $\hat{\theta}_m^0, \hat{\varepsilon}_m^0$, and $\hat{\mathbf{c}}_m^0$, respectively. We will later elaborate on how $\hat{\theta}_m^0, \hat{\varepsilon}_m^0$, and $\hat{\mathbf{c}}_m^0$ are obtained. Using these initial guesses, $\hat{\mathbf{u}}_m^0$ for $m = 1, 2, \cdots, M$ is obtained as

$$\hat{\mathbf{u}}_m^0 = \Gamma(\hat{\varepsilon}_m^0)\mathbf{A}_m(\hat{\theta}_m^0)\hat{\mathbf{c}}_m^0. \tag{154}$$

Next, the algorithm is performed iteratively, where each iteration consists of M cycles (one cycle per user), and at the end of each cycle the timing, CFO and BEM coefficient estimates of the corresponding user is updated. At the pth cycle of the ith iteration, we compute

-E-step (Expectation step):

$$\hat{\mathbf{y}}_p^i = \mathbf{r} - \sum_{m=1}^{p-1} \hat{\mathbf{u}}_m^i - \sum_{m=p+1}^{K} \hat{\mathbf{u}}_m^i \tag{155}$$

-M-step (Maximization step):

$$(\hat{\theta}_p^i, \hat{\varepsilon}_p^i, \hat{\mathbf{c}}_p^i) = \arg \max_{\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i, \widetilde{\mathbf{c}}_p^i} \Phi(\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i, \widetilde{\mathbf{c}}_p^i) \tag{156}$$

where

$$\Phi(\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i, \widetilde{\mathbf{c}}_p^i) = -\left\| \hat{\mathbf{y}}_p^i - \Gamma(\widetilde{\varepsilon}_p^i)\mathbf{A}_p(\widetilde{\theta}_p^i)\widetilde{\mathbf{c}}_p^i \right\|^2 \tag{157}$$

After obtaining $(\hat{\theta}_p^i, \hat{\varepsilon}_p^i, \hat{\mathbf{c}}_p^i)$, we update $\hat{\mathbf{u}}_p^i$ as follows

$$\hat{\mathbf{u}}_p^i = \Gamma(\hat{\varepsilon}_p^i)\mathbf{A}_p(\hat{\theta}_p^i)\hat{\mathbf{c}}_p^i. \tag{158}$$

The algorithm is continued until there is no significant change in the estimates.

Evidently, using the SAGE algorithm, maximization of (153) is transformed into M simpler maximization (157). Despite this fact, estimation of $\theta_p^i, \varepsilon_p^i$ and $\mathbf{c}_p^i$ using (157) needs an exhaustive search for values of $\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i$, and $\widetilde{\mathbf{c}}_p^i$ which suffers from high computational complexity and is thus not practical. Now, we propose the following procedure for obtaining $\hat{\theta}_p^i, \hat{\varepsilon}_p^i$ and $\hat{\mathbf{c}}_p^i$). First, we keep $\widetilde{\theta}_p^i$ and $\widetilde{\varepsilon}_p^i$ fixed, and maximize (157) over the values of $\widetilde{\mathbf{c}}_p^i$. Thus, the estimate of $\mathbf{c}_p^i$, which is dependent on $\widetilde{\theta}_p^i$ and $\widetilde{\varepsilon}_p^i$, can be written as

$$\hat{\mathbf{c}}_p^i(\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i) = \big(\mathbf{A}_p^H(\widetilde{\theta}_p^i)\mathbf{A}_p(\widetilde{\theta}_p^i)\big)^{-1}\mathbf{A}_p^H(\widetilde{\theta}_p^i)\Gamma^H(\widetilde{\varepsilon}_p^i)\hat{\mathbf{y}}_p^i. \tag{159}$$

Next, we replace $\widetilde{\mathbf{c}}_p^i$ with $\hat{\mathbf{c}}_p^i(\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i)$ in (157), and discard the terms that are irrelevant of $\widetilde{\varepsilon}_p^i$ and $\widetilde{\theta}_p^i$. Thus, for estimation of ε_p^i and θ_p^i, we maximize

$$\Phi(\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i) = (\hat{\mathbf{y}}_p^i)^H\Gamma(\widetilde{\varepsilon}_p^i)\mathbf{V}_p^i\Gamma^H(\widetilde{\varepsilon}_p^i)\hat{\mathbf{y}}_p^i \tag{160}$$

where $\mathbf{V}_p^i = \mathbf{A}_p(\widetilde{\theta}_p^i)\big(\mathbf{A}_p^H(\widetilde{\theta}_p^i)\mathbf{A}_p(\widetilde{\theta}_p^i)\big)^{-1}\mathbf{A}_p^H(\widetilde{\theta}_p^i)$.

In order to estimate ε_p^i, we keep $\widetilde{\theta}_p^i$ fixed, and maximize (160) with respect to $\widetilde{\varepsilon}_p^i$. Notice that (160) can be written as

$$\Phi(\widetilde{\theta}_p^i, \widetilde{\varepsilon}_p^i) = \sum_{m=0}^{N-1}\big(\hat{y}_p^i(m)\big)^* V_p^i(m,m)\hat{y}_p^i(m) + \sum_{m=0}^{N-2}\sum_{n=m+1}^{N-1} e^{j\frac{2\pi}{N}(n-m)\widetilde{\varepsilon}_p^i}\big(\hat{y}_p^i(m)\big)^* V_p^i(m,n)\hat{y}_p^i(n) + \sum_{m=0}^{N-2}\sum_{n=m+1}^{N-1} e^{j\frac{2\pi}{N}(m-n)\widetilde{\varepsilon}_p^i}\hat{y}_p^i(m)V_p^i(m,n)\big(\hat{y}_p^i(n)\big)^* \tag{161}$$

where $V_p^i(m,n)$ indicates the element of $\mathbf{V}_p^i$ on the mth row and nth column. Considering that $\sum_{m=0}^{N-1}\big(\hat{y}_p^i(m)\big)^* V_p^i(m,m)\hat{y}_p^i(m)$ is independent of $\widetilde{\varepsilon}_p^i$, this term can be eliminated. Also, noting that the elements of $\mathbf{V}_p^i$ are real, and the second and third terms on the right hand side of (161) are complex conjugate of each other, CFO can be estimated as

$$\hat{\varepsilon}_p^i(\widetilde{\theta}_p^i) = \arg\max_{\widetilde{\varepsilon}_p^i}\Phi(\widetilde{\varepsilon}_p^i), \tag{162}$$

where

$$\Phi(\widetilde{\varepsilon}_p^i) = \mathrm{Re}\left\{ \sum_{m=0}^{N-2} \sum_{n=m+1}^{N-1} e^{j\frac{2\pi}{N}(n-m)\widetilde{\varepsilon}_p^i} \big(\hat{y}_p^i(m)\big)^* V_p^i(m,n) \hat{y}_p^i(n) \right\}. \tag{163}$$

To implement (163) using FFT, we can equivalently express (163) as

$$\Phi(\widetilde{\varepsilon}_p^i) = \mathrm{Re}\left\{ \sum_{n=0}^{N-1} \Xi(n) e^{j\frac{2\pi}{N} n \widetilde{\varepsilon}_p^i} \right\}, \tag{164}$$

where

$$\Xi(n) = \sum_{m=0}^{N-1-n} \big(\hat{y}_p^i(m)\big)^* V_p^i(m, n+m) \hat{y}_p^i(n+m). \tag{165}$$

Now that we have an estimate for CFO, we are left with estimation of timing offset. Replacing $\widetilde{\varepsilon}_p^i$ in (160) with $\hat{\varepsilon}_p^i(\widetilde{\theta}_p^i)$, the timing estimator is obtained as

$$\hat{\theta}_p^i = \arg\max_{\widetilde{\theta}_p^i} \Phi(\widetilde{\theta}_p^i), \tag{166}$$

where

$$\Phi(\widetilde{\theta}_p^i) = (\hat{\mathbf{y}}_p^i)^H \Gamma\big(\hat{\varepsilon}_p^i(\widetilde{\theta}_p^i)\big) \mathbf{V}_p^i \Gamma^H\big(\hat{\varepsilon}_p^i(\widetilde{\theta}_p^i)\big) \hat{\mathbf{y}}_p^i. \tag{167}$$

Now, by using this timing estimate in (162), the final estimate of CFO is obtained. The BEM coefficients can also be obtained by putting the timing and CFO estimates back into (159). Finally, using the BEM coefficients along with the expression $\hat{\mathbf{h}}_{p,i} = \mathbf{B}\hat{\mathbf{c}}_{p,i}$, we obtain the doubly selective channel estimate of the pth user.

The following comments are of interest: 1) For initialization of the sage algorithm, the following approach is adopted. We set the initial CFO estimates for all users to zero. The initial timing and channel estimates for each user are obtained as if the received signal only belongs to that particular user and there is no interference from signal of other users. In other words, for the mth user we have $\hat{\varepsilon}_m^0 = 0$,

$$\hat{\theta}_m^0 = \arg\max_{\widetilde{\theta}_m^0} \left\{ \mathbf{r}^H \Gamma\big(\hat{\varepsilon}_m^0\big) \mathbf{V}_m^0(\widetilde{\theta}_m^0) \Gamma^H\big(\hat{\varepsilon}_m^0\big) \mathbf{r} \right\} \tag{168}$$

$$\hat{\mathbf{c}}_m^0(\hat{\theta}_m^0, \hat{\varepsilon}_m^0) = \left(\mathbf{A}_m^H(\hat{\theta}_m^0)\mathbf{A}_m(\hat{\theta}_m^0)\right)^{-1}\mathbf{A}_m^H(\hat{\theta}_m^0)\Gamma(\hat{\varepsilon}_m^0)\hat{\mathbf{r}}. \tag{169}$$

We have summarized the proposed method in Tabel 7, where N_S denotes the number of iterations of the SAGE algorithm.

2) According to (164), to increase the resolution of the CFO estimation, we should increase the length of the FFT which increases the complexity. As a solution to this problem, we can obtain a coarse estimate of CFO jointly with timing offset. When the timing issue is resolved, and remaining CFO is small, we can apply iterative algorithms such as the one based on Newton Raphson method presented in [114].

4.4.3 Performance Evaluation

4.4.3.1 Computational Complexity

We first evaluate the computational complexity considering that for the multiplication of an $N_1 \times N_2$ matrix with an $N_2 \times N_3$ matrix, we need $N_1N_2N_3$ complex multiplications and $N_1N_3(N_2-1)$ complex additions. For multiplication of a diagonal $N_1 \times N_1$ matrix with an $N_1 \times N_2$ matrix, we need N_1N_2 complex multiplications. Further, the inversion of an $N_1 \times N_1$ matrix needs N_1^3 complex multiplications.

At each cycle of an iteration, for each trial value of the timing offset of each user, the computational complexity is as follows: Computing (167) needs $3N^2+N$ complex multiplications and N^2-1 complex additions. Performing (164) needs $N(N-1)+L_F\log_2 L_F+L_cQN^2$ complex multiplications and $\frac{N(N-1)}{2}+L_F\log_2 L_F+(L_cQ-1)N^2$ complex additions, where L_F is the FFT length. Calculation of (159) needs $2L_cQN$ complex multiplications and $L_cQ(N-1)$ complex additions. Computation of (155) needs $K(NL_cQ+N)$ complex multiplications and $KN(L_cQ-1)+(K-1)N$ complex additions. For initialization, we also need $K(3N^2+N+2L_cQN)$ complex multiplications and $K(N^2-1+KL_cQ(N-1))$ complex additions. Furthermore, it is assumed that the matrix $\mathbf{V}_p^i$ and $\left(\mathbf{A}_p^H(\widetilde{\theta}_p^i)\mathbf{A}_p(\widetilde{\theta}_p^i)\right)^{-1}\mathbf{A}_p^H(\widetilde{\theta}_p^i)$

Table 7: Proposed SAGE-Based Algorithm for OFDMA Systems

Inputs: $\mathbf{r}$, $\mathbf{A}_p(\widetilde{\theta}_p)$, $p = 1, \cdots, K$.

Initilaization: - Set $\hat{\varepsilon}_p^0 = 0$, and obtain $\hat{\theta}_p^0$ and $\hat{\mathbf{c}}_p^0$ using (168) and (169), respectively, $p = 1, \cdots, K$.

- Compute $\hat{\mathbf{u}}_p^0 = \Gamma(\hat{\varepsilon}_p^0)\mathbf{A}_p(\hat{\theta}_p^0)\hat{\mathbf{c}}_p^0$, $p = 1, \cdots, K$.

For $i = 0 : N_S - 1$

For $p = 1 : K$

- **E-step**: Compute $\hat{\mathbf{y}}_p^i = \mathbf{r} - \sum_{k=1}^{p-1} \hat{\mathbf{u}}_k^i - \sum_{k=p+1}^{K} \hat{\mathbf{u}}_k^i$.
- **M-step**:
- Estimate the CFO for each trial value of timing offset using (164) and $\hat{\varepsilon}_p^i(\widetilde{\theta}_p^i) = \arg\max_{\widetilde{\varepsilon}_p^i} \Phi(\widetilde{\varepsilon}_p^i)$.
- Estimate the timing offset using the estimated CFOs, (167), and $\hat{\theta}_p^i = \arg\max_{\widetilde{\theta}_p^i} \Phi(\widetilde{\theta}_p^i)$.
- Estimate the CFO using the estimated timing offset $\hat{\varepsilon}_p^i = \hat{\varepsilon}_p^i(\hat{\theta}_p^i)$.
- Estimate the BEM coefficients as $\hat{\mathbf{c}}_p^i = \left(\mathbf{A}_p^H(\hat{\theta}_p^i)\mathbf{A}_p(\hat{\theta}_p^i)\right)^{-1}\mathbf{A}_p^H(\hat{\theta}_p^i)\Gamma^H(\hat{\varepsilon}_p^i)\hat{\mathbf{y}}_p^i$.
- Compute $\hat{\mathbf{u}}_p^i = \Gamma(\hat{\varepsilon}_p^i)\mathbf{A}_p(\hat{\theta}_p^i)\hat{\mathbf{c}}_p^i$.

End

Set $\hat{\mathbf{u}}_p^{i+1} = \hat{\mathbf{u}}_p^i$, $p = 1, \cdots, K$.

End

Outputs: Estimate the timing offsets, CFOs, and BEM coefficients $\hat{\theta}_p = \hat{\theta}_p^{N_S-1}$, $\hat{\varepsilon}_p = \hat{\varepsilon}_p^{N_S-1}$, $\hat{\mathbf{h}}_p = \mathbf{B}\hat{\mathbf{c}}_p^{N_S-1}$, $p = 1, \cdots, K$.

are precalculated.

In Table 8, we have demonstrated the complexity of different estimators where N_θ denotes the number of trial values of the timing offsets, and N_g denotes the number of grid searches needed in [114]. Since each multiplication counts for several additions, and consequently the computational complexity is mainly due to multiplications, in this table we have only compared the number of multiplications. Roughly speaking, the overall complexity of the new scheme is N_θ and N_S times that of [114] and [125], respectively, where $N_\theta \leq G$ and N_S is typically 5.

Table 8: Complexity of Estimators for OFDMA Systems

Estimator	**Number of Multiplications**
CFO [114]	$N_g N_i(N^2+3N)$
Channel [114]	$2N_S K L_c Q N$
Timing [125]	$N_\theta K(N^2+3N)$
CFO [125]	$N_\theta K(N(N-1)+L_c Q N^2+L_F \log_2 L_F)$
Channel [125]	$2N_\theta K L_c Q N$
Proposed Timing	$N_S N_\theta K(3N^2+N)$
Proposed CFO	$N_S N_\theta K(N(N-1)+L_c Q N^2+L_F \log_2 L_F)$
Proposed Channel	$2N_S N_\theta K L_c Q N$

4.4.3.2 Simulation Results

We have utilized computer simulation to evaluate the performance of the new estimators. An OFDMA system with $N = 256$ subcarriers, CP length of $G = 32$ and 4 users with generalized subcarrier assignment is considered. The simulated doubly selective channel has 4 taps with delay spacings of $0.81\mu s$ where the power delay profile of the lth tap is given by $\exp(-\frac{l}{L_c})$, and Jake's model is used. The carrier frequency and the mobile velocity are respectively 1.9 GHz and 240 km/hr resulting in $f_d T = 0.25$, where f_d is the Doppler frequency, and T is the OFDM useful symbol duration. Furthermore, we have made use of the generalized complex exponential BEM [105] with $Q = 2$.

Fig. 36 , 37, and 38, respectively, illustrate the MSEs of the proposed timing, CFO and channel estimators compared with previous methods. Here, the SLML method indicates an ML method developed for slow fading channels, i.e., without using the BEM such as the one in [87]. In this figure, OML refers to an estimator that treats an OFDMA symbol similar to an OFDM symbol. Obviously, the new method has a significantly lower timing, CFO and channel MSEs than the SLML and OML schemes. This is because the SLML method does not consider the time variations of the channel. Furthermore, the OML method assumes that the OFDMA symbol is from only one user. However, in reality, the received OFDMA symbol is the superposition of the signals of all users, and the signals of

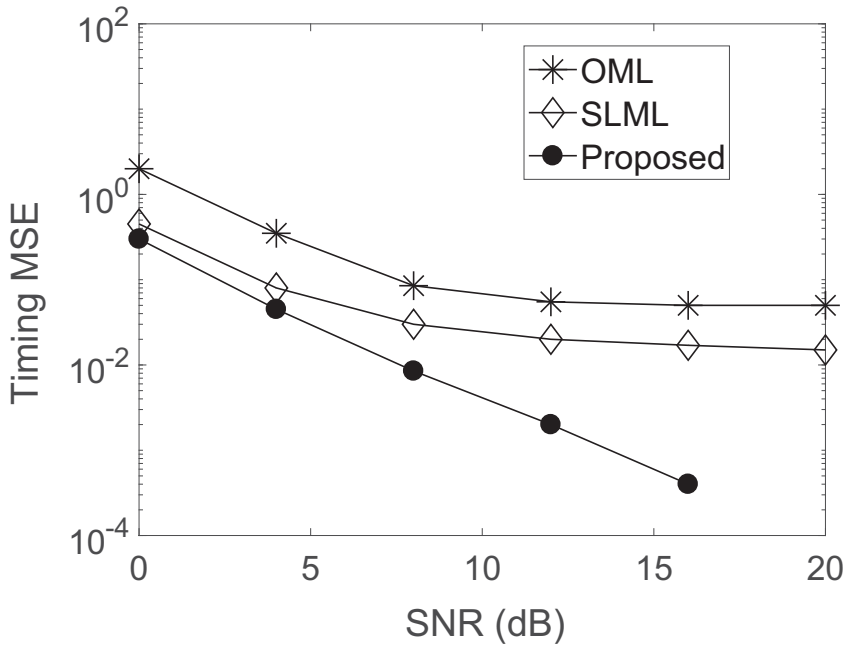

Figure 36: MSE of different timing estimators.

other users act as interference for a specific user. This issue significantly deteriorates the performance of the OML method. Also our method is better than [114] where the timing offset is not estimated.

Now, we evaluate the performance from inter carrier interference (ICI) point of view. The ICI is caused by two factors: CFO and time selectivity of the channel [158]. To show the effectiveness of the new estimation method in suppression of the ICI, we have compared the bit error rate (BER) for different methods utilizing zero-forcing equalization in Fig. 39. In this figure, the ideal curve indicates the case of perfectly knowing the timing offset, CFO, and channel. Evidently, the new scheme significantly improves the performance.

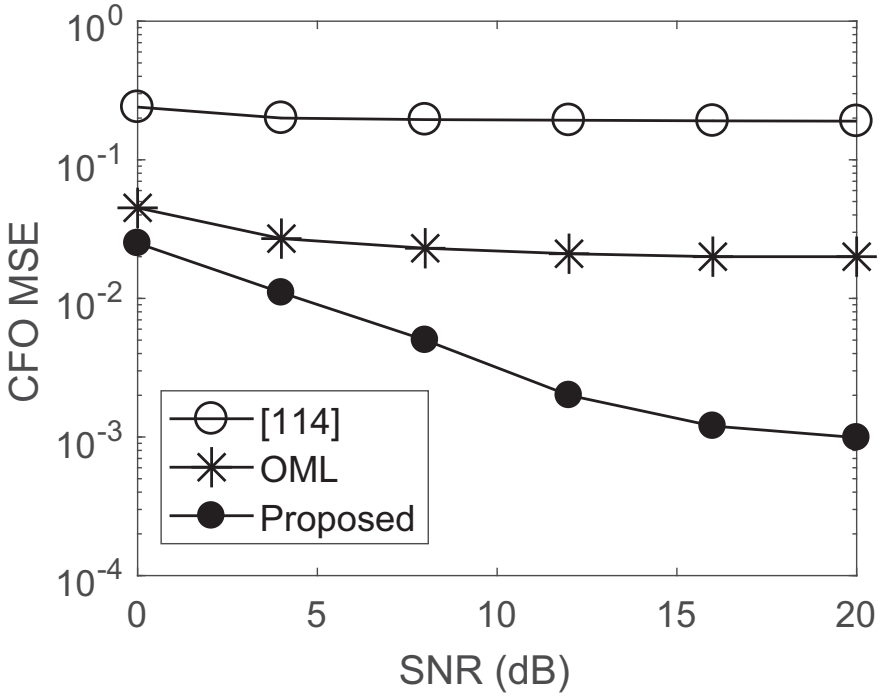

Figure 37: MSE of different CFO estimators for OFDMA systems.

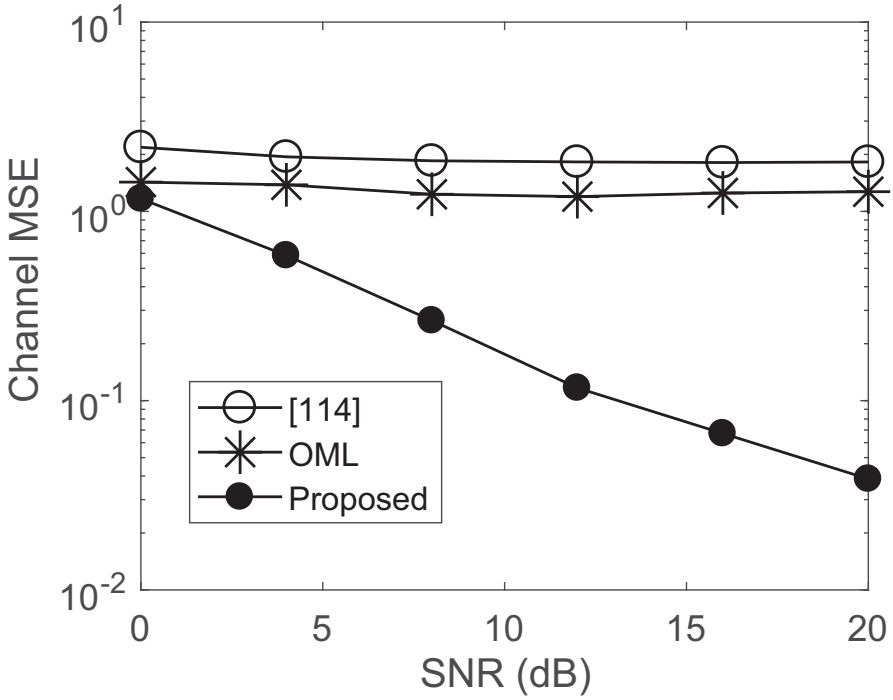

Figure 38: MSE of different channel estimators for OFDMA systems.

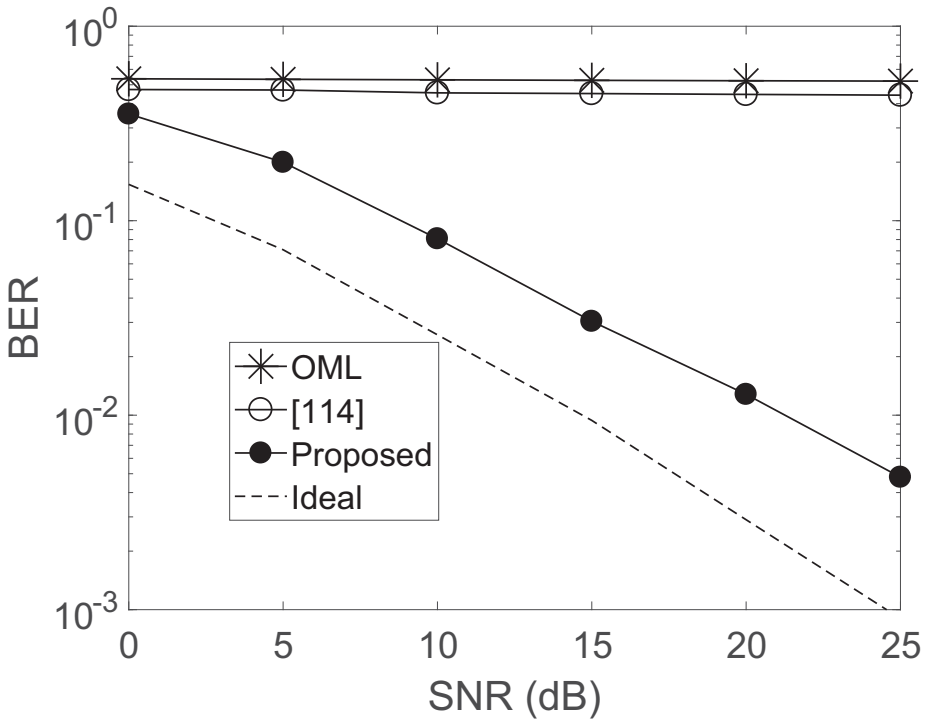

Figure 39: BER of different methods in OFDMA systems.

4.5 Proposed Method: Joint CFO and Channel Estimation in MIMO-OFDMA Systems Using Particle and Kalman Filtering

4.5.1 Preliminaries

4.5.1.1 System Model

We consider the uplink of a MIMO-OFDMA system consisting of a base station with N_R receive antennas and M active users where each user is equipped with N_T transmit antennas. The total number of subcarriers is N which are shared by all users according to a generalized subcarrier assignment scheme where the number of subcarriers assigned to the mth user is denoted by N_m. We let the ith OFDMA symbol of the mth user in

the frequency domain to be sent from the tth, $t = 1, 2, \cdots, N_T$, transmit antenna be represented by $\mathbf{S}_{i,m}^t = [S_{i,m}^t(0), S_{i,m}^t(1), \cdots, S_{i,m}^t(N-1)]^T$ where $S_{i,m}^t(n)$ denotes the nth element of $\mathbf{S}_{i,m}^t$, and is equal to zero if the nth subcarrier is not assigned to the mth user. The time domain vector corresponding to $\mathbf{S}_{i,m}^t$ is denoted by $\mathbf{s}_{i,m}^t$. The lth tap of the time-varying channel between the tth transmit antenna of the mth user and the pth receive antenna of the base station for the ith OFDM symbol at time k is indicated by $h_{i,m}^{t,p}(l,k)$. Using a vector form representation of the doubly selective channel, we have

$$\mathbf{h}_{i,m}^{t,p} = \left[(\mathbf{h}_{i,m,0}^{t,p})^T, (\mathbf{h}_{i,m,1}^{t,p})^T, \cdots, (\mathbf{h}_{i,m,L_c-1}^{t,p})^T\right]^T, \tag{170}$$

where $\mathbf{h}_{i,m,l}^{t,p} = [h_{i,m}^{t,p}(l,0), h_{i,m}^{t,p}(l,1), \cdots, h_{i,m}^{t,p}(l,N-1)]^T$. The transmitted signal will be also affected by CFO. We denote the normalized CFO affecting the ith OFDM symbol of the mth user transmitted from the tth antenna and received at the pth antenna of base station by $\varepsilon_{i,m}^{t,p}$. Since the distances between the transmit antennas of the users are much smaller than the distance between the receive antennas of the base station, it is reasonable to assume that the CFOs between all transmit antennas of a user and a specific receive antenna at the base station are the same [99]. Thus, we drop the superscript t, and indicate the CFO by $\varepsilon_{i,m}^{p}$. The received signal corresponding to the mth user at the pth receive antenna, after removing the CP, can be written as

$$r_{i,m}^p(k) = e^{j\frac{2\pi}{N}(G+k)\varepsilon_{i,m}^p} \sum_{t=1}^{N_T} \sum_{l=0}^{L_c-1} h_{i,m}^{t,p}(l,k) s_{i,m}^t(k-l) + z_{i,m}^p(k), \tag{171}$$

where $z_{i,m}^{t,p}(k)$ is additive white Gaussian noise (AWGN) with zero mean and variance $\sigma_{z_{i,m}^p}^2$. Resorting to matrix representation, we can express the received signal at the pth antenna which is the superposition of the signals from all users as

$$\mathbf{r}_i^p = \sum_{m=1}^{M} \Gamma(\varepsilon_{i,m}^p) \check{\mathbf{A}}_{i,m} \mathbf{h}_{i,m}^p + \mathbf{z}_i^p \tag{172}$$

where

$$\mathbf{r}_{i,m}^{p} = [r_{i,m}^{p}(0), r_{i,m}^{p}(1), \cdots, r_{i,m}^{p}(N-1)]^{T} \tag{173}$$

$$\Gamma(\varepsilon_{i,m}^{p}) = \text{diag}\{e^{j\frac{2\pi}{N}k\varepsilon_{i,m}^{p}}, k = G, G+1, \cdots, G+N-1\} \tag{174}$$

$$\check{\mathbf{A}}_{i,m} = [\check{\mathbf{A}}_{i,m}^{1}, \check{\mathbf{A}}_{i,m}^{2}, \cdots, \check{\mathbf{A}}_{i,m}^{N_T}] \tag{175}$$

$$\check{\mathbf{A}}_{i,m}^{t} = [\mathbf{s}_{i,m}^{t,(0)}, \mathbf{s}_{i,m}^{t,(1)}, \cdots, \mathbf{s}_{i,m}^{t,(L_c-1)}] \tag{176}$$

$$\mathbf{h}_{i,m}^{p} = \left[\left(\mathbf{h}_{i,m}^{1,p}\right)^{T}, \left(\mathbf{h}_{i,m}^{2,p}\right)^{T}, \cdots, \left(\mathbf{h}_{i,m}^{N_T,p}\right)^{T}\right]^{T} \tag{177}$$

$$\mathbf{s}_{i,m}^{t,(l)} = \text{diag}\left\{s_{i,m}^{t}(l), s_{i,m}^{t}(l+1), \cdots, s_{i,m}^{t}(N-1), s_{i,m}^{t}(0), s_{i,m}^{t}(1), \cdots, s_{i,m}^{t}(l-1)\right\} \tag{178}$$

Moreover, $\mathbf{s}_{i,m}^{t,(l)}$ is a diagonal matrix formed by cyclically shifting $\mathbf{s}_{i,m}^{t}$ by l samples, and $\mathbf{z}_{i}^{p}$ is an AWGN matrix with zero mean and covariance $\sigma_z^2\mathbf{I}_N$.

4.5.1.2 System Model Using BEM

We express $h_{i,m}^{t,p}(l,k)$ using BEM as

$$h_{i,m}^{t,p}(l,k) = \sum_{q=0}^{Q-1} b_{k,q} c_{i,m,l,q}^{t,p} \tag{179}$$

where Q is the number of BEM coefficients, $b_{k,q}$ is the qth basis function at time k, and $c_{i,m,l,q}$ is the qth BEM coefficient for the lth channel tap and ith OFDM symbol of the mth user. Employing (179), a doubly selective channel can be rewritten using BEM in the matrix form as

$$\mathbf{h}_{i,m}^{t,p} = \mathbf{B}\mathbf{c}_{i,m}^{t,p}, \tag{180}$$

where

$$\mathbf{B} = \mathbf{I}_{L_c} \otimes [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{N-1}]^{T} \tag{181}$$

$$\mathbf{c}_{i,m}^{t,p} = \left[\left(\mathbf{c}_{i,m,0}^{t,p}\right)^{T}, \left(\mathbf{c}_{i,m,1}^{t,p}\right)^{T}, \cdots, \left(\mathbf{c}_{i,m,L_c-1}^{t,p}\right)^{T}\right]^{T} \tag{182}$$

$$\mathbf{b}_k = [b_{k,0}, b_{k,1}, \cdots, b_{k,Q-1}]^{T} \tag{183}$$

$$\mathbf{c}_{i,m,l}^{t,p} = [c_{i,m,l,0}^{t,p}, c_{i,m,l,1}^{t,p}, \cdots, c_{i,m,l,Q-1}^{t,p}]^T. \tag{184}$$

Using (180), we can rewrite the received signal at the pth receive antenna as

$$\mathbf{r}_i^p = \sum_{m=1}^{M} \Gamma(\varepsilon_{i,m}^p)\mathbf{A}_{i,m}\mathbf{c}_{i,m}^p + \mathbf{z}_i^p, \tag{185}$$

where

$$\mathbf{A}_{i,m} = [\mathbf{A}_{i,m}^1, \mathbf{A}_{i,m}^2, \cdots, \mathbf{A}_{i,m}^{N_T}] \tag{186}$$

$$\mathbf{c}_{i,m}^p = \left[\left(\mathbf{c}_{i,m}^{1,p}\right)^T, \left(\mathbf{c}_{i,m}^{2,p}\right)^T, \cdots, \left(\mathbf{c}_{i,m}^{N_T,p}\right)^T\right]^T \tag{187}$$

$$\mathbf{A}_{i,m}^t = [\mathbf{s}_{i,m}^{t,(0)}\breve{\mathbf{B}}, \mathbf{s}_{i,m}^{t,(1)}\breve{\mathbf{B}}, \cdots, \mathbf{s}_{i,m}^{t,(L_c-1)}\breve{\mathbf{B}}] \tag{188}$$

$$\breve{\mathbf{B}} = [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{N-1}]^T \tag{189}$$

4.5.1.3 Space State Model of Channel and CFO

In this subsection, we present the widely used autoregressive models for time variations of channel and CFO [99,101,108,109]. With regard to the doubly selective channel, it should be mentioned that although the channel changes within an OFDMA symbol interval, the BEM coefficients are constant for an OFDMA symbol and change from one OFDMA symbol to another, according to the following first order autoregressive model [108]

$$\mathbf{c}_{i,m}^p = \mathbf{\Xi}_m^p \mathbf{c}_{i-1,m}^p + \mathbf{v}_{i,m}^p, \tag{190}$$

where $\mathbf{\Xi}_m^p = \text{blkdiag}\{\breve{\mathbf{\Xi}}_m^p, \breve{\mathbf{\Xi}}_m^p, \cdots, \breve{\mathbf{\Xi}}_m^p\}$ is an $N_T L_c Q \times N_T L_c Q$ block diagonal matrix with the $N_T L_c$ submatrices $\breve{\mathbf{\Xi}}_m^p$ on the diagonal, and $\mathbf{v}_{i,m}^p$ is a Gaussian noise vector with zero mean and covariance $\mathbf{U}_m = \text{blkdiag}\{\mathbf{U}_{0,m}^{1,p}, \mathbf{U}_{1,m}^{1,p}, \cdots, \mathbf{U}_{L_c-1,m}^{N_T,p}\}$. Furthermore, $\breve{\Xi}_m^p$ and $\mathbf{U}_{l,m}^{t,p}$ are determined by computing the following set of Yule-waker equations [108], $\breve{\mathbf{\Xi}}_m^p = \mathbf{R}_{l,m}^{t,p}[1]\left(\mathbf{R}_{l,m}^{t,p}[0]\right)^{-1}$, $\mathbf{U}_{l,m}^{t,p} = \mathbf{R}_{l,m}^{t,p}[0] + \breve{\mathbf{\Xi}}_m^p \mathbf{R}_{l,m}^{t,p}[-1]$ where $\mathbf{R}_{l,m}^{t,p}[s] = \text{E}[\mathbf{c}_{i,m,l}^{t,p}(\mathbf{c}_{i-s,m,l}^{t,p})^H] = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_{h,m,l}^{t,p}\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}$ with $\mathbf{R}_{h,m,l}^{t,p}[s] = \text{E}[\mathbf{h}_{i,m,l}^{t,p}(\mathbf{h}_{i-s,m,l}^{t,p})^H]$.

It is also assumed that CFO of a user changes from one OFDMA symbol to another according to the model [99, 108, 122]

$$\varepsilon_{i,m}^{p} = \Omega_m \varepsilon_{i-1,m}^{p} + w_i^{p}, \tag{191}$$

where w_i^p is a sample of AWGN with zero mean and variance σ_w^2, and Ω_m is typically chosen between 0.99 and 0.9999 [108].

4.5.2 A Schimdt Extended Kalman Filtering Based Approach

In this section, we propose a joint CFO and channel estimation approach based on BEM and SEKF, called BSEKF scheme. We follow the presentation of SEKF in [99] by first, expressing $\mathbf{r}_i^p$, $\Gamma(\varepsilon_{i,m}^p)\mathbf{A}_{i,m}$, and $\mathbf{c}_{i,m}^p$ in terms of their real and imaginary parts, namely,

$$\mathbf{y}_i^p = \begin{bmatrix} \mathrm{Re}\{\mathbf{r}_i^p\} \\ \mathrm{Im}\{\mathbf{r}_i^p\} \end{bmatrix} \tag{192}$$

$$\mathbf{\Lambda}_{i,m}(\varepsilon_{i,m}^p) = \begin{bmatrix} \mathrm{Re}\{\Gamma(\varepsilon_{i,m}^p)\mathbf{A}_{i,m}\} & -\mathrm{Im}\{\Gamma(\varepsilon_{i,m}^p)\mathbf{A}_{i,m}\} \\ \mathrm{Im}\{\Gamma(\varepsilon_{i,m}^p)\mathbf{A}_{i,m}\} & \mathrm{Re}\{\Gamma(\varepsilon_{i,m}^p)\mathbf{A}_{i,m}\} \end{bmatrix} \tag{193}$$

$$\boldsymbol{\zeta}_{i,m}^p = \begin{bmatrix} \mathrm{Re}\{\mathbf{c}_{i,m}^p\} \\ \mathrm{Im}\{\mathbf{c}_{i,m}^p\} \end{bmatrix} \tag{194}$$

where $\mathrm{Re}\{\mathbf{x}\}$ and $\mathrm{Im}\{\mathbf{x}\}$ denote the matrices whose elements are the real and imaginary parts of the elements of $\mathbf{x}$, respectively. Using the notation $\backslash m$ for exclusion of the parameters of the mth user, and in turn, denoting $\varepsilon_{i,\backslash m}^p$ and $\boldsymbol{\zeta}_{i,\backslash m}^p$, respectively, as the CFO and BEM coefficients of all users except the mth user, we have

$$\mathbf{\Lambda}_{i,\backslash m}(\varepsilon_{i,\backslash m}^p) = [\mathbf{\Lambda}_{i,1}(\varepsilon_{i,1}^p), \cdots, \mathbf{\Lambda}_{i,m-1}, (\varepsilon_{i,m-1}^p), \mathbf{\Lambda}_{i,m+1}(\varepsilon_{i,m+1}^p), \cdots, \mathbf{\Lambda}_{i,M}(\varepsilon_{i,M}^p)] \tag{195}$$

$$\boldsymbol{\zeta}_{i,\backslash m}^{p} = \left[(\boldsymbol{\zeta}_{i,1}^{p})^{T}, \cdots, (\boldsymbol{\zeta}_{i,m-1}^{p})^{T}, (\boldsymbol{\zeta}_{i,m+1}^{p})^{T}, \cdots, (\boldsymbol{\zeta}_{i,M}^{p})^{T}\right]^{T}. \tag{196}$$

Then, we can rewrite the dynamic and observation models, respectively, as

$$\boldsymbol{\zeta}_{i,m}^{p} = \boldsymbol{\Pi}_{m}\boldsymbol{\zeta}_{i-1,m}^{p} + \boldsymbol{V}_{i,m}^{p} \tag{197}$$

$$\varepsilon_{i,m}^{p} = \Omega_{m}\varepsilon_{i-1,m}^{p} + w_{i,m}^{p}, \tag{198}$$

$$\mathbf{y}_{i}^{p} = \boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^{p})\boldsymbol{\zeta}_{i,m}^{p} + \boldsymbol{\Lambda}_{i,\backslash m}(\varepsilon_{i,\backslash m}^{p})\boldsymbol{\zeta}_{i,\backslash m}^{p} + \boldsymbol{Z}_{i}^{p} \tag{199}$$

where

$$\boldsymbol{\Pi}_{m} = \begin{bmatrix} \mathrm{Re}\{\boldsymbol{\Xi}_{m}^{p}\} & -\mathrm{Im}\{\boldsymbol{\Xi}_{m}^{p}\} \\ \mathrm{Im}\{\boldsymbol{\Xi}_{m}^{p}\} & \mathrm{Re}\{\boldsymbol{\Xi}_{m}^{p}\} \end{bmatrix} \tag{200}$$

and $\boldsymbol{V}_{i,m}^{p} = \left[(\mathrm{Re}\{\mathbf{v}_{i,m}^{p}\})^{T}\ (\mathrm{Im}\{\mathbf{v}_{i,m}^{p}\})^{T}\right]^{T}$ and $\boldsymbol{Z}_{i}^{p} = \left[(\mathrm{Re}\{\mathbf{z}_{i}^{p}\})^{T}\ (\mathrm{Im}\{\mathbf{z}_{i}^{p}\})^{T}\right]^{T}$. Note that in (199) the term $\boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^{p})\boldsymbol{\zeta}_{i,m}^{p}$ is related to the signal of the mth user and the term $\boldsymbol{\Lambda}_{i,\backslash m}(\varepsilon_{i,\backslash m}^{p})\boldsymbol{\zeta}_{i,\backslash m}^{p}$ is the interference caused by other users.

Let the entire parameters of all users be represented by $\boldsymbol{\chi}_{i}^{p} = [(\boldsymbol{\chi}_{i,m}^{p})^{T}, (\boldsymbol{\chi}_{i,\backslash m}^{p})^{T}]^{T}$, where $\boldsymbol{\chi}_{i,m}^{p}$ is the vector containing the parameters of the mth user

$$\boldsymbol{\chi}_{i,m}^{p} = \left[\varepsilon_{i,m}^{p},\ (\boldsymbol{\zeta}_{i,m}^{p})^{T}\right]^{T}, \tag{201}$$

and $\boldsymbol{\chi}_{i,\backslash m}^{p}$ corresponds to the parameters of other users

$$\boldsymbol{\chi}_{i,\backslash m}^{p} = \left[\varepsilon_{i,\backslash m}^{p},\ (\boldsymbol{\zeta}_{i,\backslash m}^{p})^{T}\right]^{T}. \tag{202}$$

Then, considering the dynamic and observation models in (197), (198), and (199), one may estimate $\boldsymbol{\chi}_{i,m}^{p}$ by the extended Kalman filtering. However, the direct computation of the extended Kalman gain for all users is inefficient [99, 130]. Instead, we propose to make use of SEKF [99,132] that offers reduced computational complexity by transforming the problem of estimation of the parameters of all users into several simpler problems of

parameter estimation for each user. To this end, the unknown parameters are divided into two groups: essential parameters that are the unknown parameters of a specific user, and nuisance parameters that are the unknown parameters of other users. As shown in [99, 130], SEKF can reduce the complexity by omitting the computation for nuisance variables.

We now employ BSEKF for each user, and regard $\boldsymbol{\chi}_{i,m}^p$ and $\boldsymbol{\chi}_{i,\backslash m}^p$ respectively as the essential and nuisance variables for the mth user. Furthermore, we indicate the estimates of $\boldsymbol{\chi}_{i,m}^p$ and $\boldsymbol{\chi}_{i,\backslash m}^p$ by $\hat{\boldsymbol{\chi}}_{i|i,m}^p$ and $\hat{\boldsymbol{\chi}}_{i|i,\backslash m}^p$, respectively, given the observation $\mathbf{y}_1^p, \mathbf{y}_2^p, \cdots, \mathbf{y}_i^p$, and by $\hat{\boldsymbol{\chi}}_{i|i-1,m}^p$ and $\hat{\boldsymbol{\chi}}_{i|i-1,\backslash m}^p$ given the observations $\mathbf{y}_1^p, \mathbf{y}_2^p, \cdots, \mathbf{y}_{i-1}^p$, respectively. By linearizing (199) using the first order taylor series expansion around $\hat{\boldsymbol{\chi}}_{i|i-1}^p$, we have

$$\mathbf{y}_i^p \approx \boldsymbol{\Lambda}_{i,m}(\hat{\varepsilon}_{i|i-1,m}^p)\hat{\boldsymbol{\zeta}}_{i|i-1,m}^p + \boldsymbol{\Lambda}_{i,\backslash m}(\hat{\varepsilon}_{i|i-1,\backslash m}^p)\hat{\boldsymbol{\zeta}}_{i|i-1,\backslash m}^p + [\mathbf{J}_{i,m},\ \mathbf{J}_{i,\backslash m}] \begin{bmatrix} \boldsymbol{\chi}_{i,m}^p - \hat{\boldsymbol{\chi}}_{i|i-1,m}^p \\ \boldsymbol{\chi}_{i,\backslash m}^p - \hat{\boldsymbol{\chi}}_{i|i-1,\backslash m}^p \end{bmatrix} + \boldsymbol{Z}_i^p. \tag{203}$$

Note that $\mathbf{J}_{i,m}$ and $\mathbf{J}_{i,\backslash m}$ are the Jacobian matrices defined as

$$\mathbf{J}_{i,m} = \frac{\partial}{\partial \boldsymbol{\chi}_{i,m}^p} \big(\boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^p)\boldsymbol{\zeta}_{i,m}^p\big)\big|_{\hat{\boldsymbol{\chi}}_{i|i-1,m}^p} = [\mathbf{J}_1, \mathbf{J}_2] \tag{204}$$

$$\mathbf{J}_{i,\backslash m} = [\mathbf{J}_{i,1}, \cdots, \mathbf{J}_{i,m-1}, \mathbf{J}_{i,m+1}, \cdots, \mathbf{J}_{i,m}] \tag{205}$$

where

$$\mathbf{J}_1 = \begin{bmatrix} -\mathrm{Im}\{\boldsymbol{\Delta}\boldsymbol{\Gamma}(\hat{\varepsilon}_{i|i-1,m}^p)\mathbf{A}_{i,m}\} & -\mathrm{Re}\{\boldsymbol{\Delta}\boldsymbol{\Gamma}(\hat{\varepsilon}_{i|i-1,m}^p)\mathbf{A}_{i,m}\} \\ \mathrm{Re}\{\boldsymbol{\Delta}\boldsymbol{\Gamma}(\hat{\varepsilon}_{i|i-1,m}^p)\mathbf{A}_{i,m}\} & -\mathrm{Im}\{\boldsymbol{\Delta}\boldsymbol{\Gamma}(\hat{\varepsilon}_{i|i-1,m}^p)\mathbf{A}_{i,m}\} \end{bmatrix} \times \hat{\boldsymbol{\zeta}}_{i|i-1,m}^p \tag{206}$$

$$\mathbf{J}_2 = \boldsymbol{\Lambda}_{i,m}(\hat{\varepsilon}_{i|i-1,m}^p) \tag{207}$$

with $\boldsymbol{\Delta} = \frac{2\pi}{N}\mathrm{diag}\{G, G+1, \cdots, G+N-1\}$.

We can also rewrite the dynamic models as

$$\boldsymbol{\chi}_{i,m}^p = \boldsymbol{\rho}_m \boldsymbol{\chi}_{i-1,m}^p + \boldsymbol{Z}_{i,m}^p, \tag{208}$$

$$\boldsymbol{\chi}^p_{i,\backslash m} = \boldsymbol{\rho}_{\backslash m}\boldsymbol{\chi}^p_{i-1,\backslash m} + \boldsymbol{Z}^p_{i,\backslash m}, \tag{209}$$

where $\boldsymbol{\rho}_m = \text{blkdiag}\{\Omega_m, \boldsymbol{\Pi}_m\}$ and $\boldsymbol{Z}^p_{i,m}$ is an AWGN matrix with zero mean and covariance matrix $\mathbf{Q} = \text{blkdiag}\{\sigma_w^2, \mathbf{Q}_u\}$ where

$$\mathbf{Q}_u = \frac{1}{2}\begin{bmatrix} \text{Re}\{\mathbf{U}_m\} & -\text{Im}\{\mathbf{U}_m\} \\ \text{Im}\{\mathbf{U}_m\} & \text{Re}\{\mathbf{U}_m\} \end{bmatrix}.$$

Moreover, $\boldsymbol{\rho}_{\backslash m} = \text{blkdiag}\{\Omega_1, \boldsymbol{\Pi}_1, \cdots, \Omega_{m-1}, \boldsymbol{\Pi}_{m-1}, \Omega_{m+1}, \boldsymbol{\Pi}_{m+1}, \cdots, \Omega_M, \boldsymbol{\Pi}_M\}$, and $\boldsymbol{Z}^p_{i,\backslash m}$ is an AWGN matrix with zero mean and covariance corresponding to the noise of the state of all users except the mth user.

Consider the linearized observation and dynamic models in (203), (208) and (209), and the covariance matrix of $\boldsymbol{\chi}^p_i$ i.e., $\mathbf{P}_{i|i-1} = \text{E}\big\{\big(\boldsymbol{\chi}^p_i - \hat{\boldsymbol{\chi}}^p_{i|i-1}\big)\big(\boldsymbol{\chi}^p_i - \hat{\boldsymbol{\chi}}^p_{i|i-1}\big)^T\big\}$ which can be expressed as

$$\mathbf{P}_{i|i-1} = \begin{bmatrix} \text{P}_{i|i-1,m,m} & \text{P}_{i|i-1,m,\backslash m} \\ \text{P}_{i|i-1,\backslash m,m} & \text{P}_{i|i-1,\backslash m,\backslash m} \end{bmatrix}, \tag{210}$$

where $\mathbf{P}_{i|i-1,m,m} = \text{E}\big\{\big(\boldsymbol{\chi}^p_{i,m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,m}\big)\big(\boldsymbol{\chi}^p_{i,m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,m}\big)^T\big\}$, $\mathbf{P}_{i|i-1,m,\backslash m} = \text{E}\big\{\big(\boldsymbol{\chi}^p_{i,m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,m}\big)\big(\boldsymbol{\chi}^p_{i,\backslash m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,\backslash m}\big)^T\big\}$, $\mathbf{P}_{i|i-1,\backslash m,m} = \text{E}\big\{\big(\boldsymbol{\chi}^p_{i,\backslash m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,\backslash m}\big)\big(\boldsymbol{\chi}^p_{i,m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,m}\big)^T\big\}$, and $\mathbf{P}_{i|i-1,\backslash m,\backslash m} = \text{E}\big\{\big(\boldsymbol{\chi}^p_{i,\backslash m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,\backslash m}\big)\big(\boldsymbol{\chi}^p_{i,\backslash m} - \hat{\boldsymbol{\chi}}^p_{i|i-1,\backslash m}\big)^T\big\}$. Then, for the mth user, the prediction and correction steps of the BSEKF can be described as [99, 130]

- **Prediction Step of the BSEKF:**

$$\hat{\boldsymbol{\chi}}^p_{i|i-1,m} = \boldsymbol{\rho}_m\hat{\boldsymbol{\chi}}^p_{i-1|i-1,m} \tag{211}$$

$$\mathbf{P}_{i|i-1,m,m} = \boldsymbol{\rho}_m\mathbf{P}_{i-1|i-1,m,m}\boldsymbol{\rho}_m^T + \mathbf{Q} \tag{212}$$

- **Correction Step of the BSEKF:**

$$\boldsymbol{\mathcal{G}} = \Big(\mathbf{J}_{i,m}\mathbf{P}_{i|i-1,m,m}(\mathbf{J}_{i,m})^T + \mathbf{J}_{i,m}\mathbf{P}_{i|i-1,m,\backslash m}(\mathbf{J}_{i,\backslash m})^T + \mathbf{J}_{i,\backslash m}\mathbf{P}_{i|i-1,\backslash m,m}(\mathbf{J}_{i,m})^T + \mathbf{J}_{i,\backslash m}\mathbf{P}_{i|i-1,\backslash m,\backslash m}(\mathbf{J}_{i,\backslash m})^T + \frac{\sigma_z^2}{2}\mathbf{I}_{2N}\Big)^{-1} \quad (213)$$

$$\mathbf{K}_{i,m} = \Big(\mathbf{P}_{i|i-1,m,m}(\mathbf{J}_{i,m})^T + \mathbf{P}_{i|i-1,m,\backslash m}(\mathbf{J}_{i,\backslash m})^T\Big)\boldsymbol{\mathcal{G}} \quad (214)$$

$$\hat{\boldsymbol{\chi}}^p_{i|i,m} = \hat{\boldsymbol{\chi}}^p_{i|i-1,m} + \mathbf{K}^p_{i,m}\Big(\mathbf{y}^p_i - \boldsymbol{\Lambda}_{i,m}(\hat{\varepsilon}^p_{i|i-1,m})\hat{\boldsymbol{\zeta}}^p_{i|i-1,m} + \boldsymbol{\Lambda}_{i,\backslash m}(\hat{\varepsilon}^p_{i|i-1,\backslash m})\hat{\boldsymbol{\zeta}}^p_{i|i-1,\backslash m}\Big) \quad (215)$$

$$\mathbf{P}_{i|i,m,m} = \boldsymbol{\mathcal{E}}\mathbf{P}_{i|i-1,m,m}(\boldsymbol{\mathcal{E}})^T - \boldsymbol{\mathcal{E}}\mathbf{P}_{i|i,m,\backslash m}(\boldsymbol{\mathcal{F}})^T - \boldsymbol{\mathcal{F}}\mathbf{P}_{i|i,\backslash m,m}(\boldsymbol{\mathcal{E}})^T + \boldsymbol{\mathcal{F}}\mathbf{P}_{i|i-1,\backslash m,\backslash m}(\boldsymbol{\mathcal{F}})^T + \frac{\sigma_z^2}{2}\mathbf{K}_{i,m}(\mathbf{K}_{i,m})^T \quad (216)$$

$$\mathbf{P}_{i|i,m,\backslash m} = \boldsymbol{\mathcal{E}}\mathbf{P}_{i|i-1,m,\backslash m} - \boldsymbol{\mathcal{F}}\mathbf{P}_{i|i-1,\backslash m,\backslash m} \quad (217)$$

$$\mathbf{P}_{i|i,\backslash m,m} = (\mathbf{P}_{i|i,m,\backslash m})^T \quad (218)$$

$$\mathbf{P}_{i|i,\backslash m,\backslash m} = \mathbf{P}_{i|i-1,\backslash m,\backslash m} \quad (219)$$

where $\boldsymbol{\mathcal{E}} = \mathbf{I}_{2N_TQL_c+1} - \mathbf{K}_{i,m}\mathbf{J}_{i,m}$ and $\boldsymbol{\mathcal{F}} = \mathbf{K}_{i,m}\mathbf{J}_{i,\backslash m}$, and $\mathbf{K}_{i,m}$ is the BSEKF gain. Thus, at the prediction step, $\boldsymbol{\chi}^p_{i,m}$ and its covariance are predicted, and at the correction step, the estimate and covariance of $\boldsymbol{\chi}^p_{i,m}$ are calculated. In Table 9, we have summarized the proposed BSEKF method.

4.5.3 A Schmidt Kalman and Gaussian Particle Filtering Based Approach

Here, we propose a new estimation method called BSK-GPF that uses BEM in conjunction with Schmidt Kalman and Gaussian particle filtering. The basic idea is to divide the parameters of each user that are to be estimated into two groups: a nonlinear parameter CFO and linear parameters i.e., BEM coefficients. Using the probability density function

Table 9: Proposed BSEKF Algorithm

Initialization:

Set $\hat{\boldsymbol{\chi}}^p_{0|0,m} = \mathbf{0}_{(L_cN_TQ+1)\times 1}$ and $\mathbf{P}_{0|0} = \mathbf{I}_{ML_cN_TQ+M}$.

Prediction Step of the BSEKF:

Compute $\hat{\boldsymbol{\chi}}^p_{i|i-1,m} = \boldsymbol{\rho}_m \hat{\boldsymbol{\chi}}^p_{i-1|i-1,m}$.

Compute $\mathbf{P}_{i|i-1,m,m} = \boldsymbol{\rho}_m \mathbf{P}_{i-1|i-1,m,m} \boldsymbol{\rho}_m^T + \mathbf{Q}$.

Correction Step of the BSEKF:

Compute $\mathbf{K}_{i,m}$ using (214).

Compute $\hat{\boldsymbol{\chi}}^p_{i|i,m} = \hat{\boldsymbol{\chi}}^p_{i|i-1,m} + \mathbf{K}^p_{i,m}\Big(\mathbf{y}^p_i - \boldsymbol{\Lambda}_{i,m}(\hat{\varepsilon}^p_{i|i-1,m})\hat{\boldsymbol{\zeta}}^p_{i|i-1,m} + \boldsymbol{\Lambda}_{i,\backslash m}(\hat{\varepsilon}^p_{i|i-1,\backslash m})\hat{\boldsymbol{\zeta}}^p_{i|i-1,\backslash m}\Big)$.

Compute $\mathbf{P}_{i|i,m,m}$, $\mathbf{P}_{i|i,m,\backslash m}$, $\mathbf{P}_{i|i,m,\backslash m}$, $\mathbf{P}_{i|i,\backslash m,\backslash m}$ using (216), (217), (218), and (219).

(pdf) $p(\varepsilon^p_{i,m}, \boldsymbol{\zeta}^p_{i,m}|\mathbf{y}^p_i)$, and Baye's theorem

$$p(\varepsilon^p_{i,m}, \boldsymbol{\zeta}^p_{i,m}|\mathbf{y}^p_i) = p(\boldsymbol{\zeta}^p_{i,m}|\mathbf{y}^p_i, \varepsilon^p_{i,m}) p(\varepsilon^p_{i,m}|\mathbf{y}^p_i), \tag{220}$$

the linear parameters are marginalized out and estimated using Kalman filtering, whereas the non-linear parameter is obtained using particle filtering. This approach is called Marginalized particle filtering [134]. In what follows, we explain the new algorithm in detail, with a particular focus on the CFO estimation utilizing GPF and the BEM coefficient estimation using SKF (BSKF).

In particle filtering, the PDF $p(\varepsilon^p_{i,m}|\mathbf{y}^p_i)$ is estimated by discrete random measures defined by particles and the weights assigned to the particles as [148]

$$p(\varepsilon^p_{i,m}|\mathbf{y}^p_i) = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)} \delta(\varepsilon^p_{i,m} - \varepsilon^{p,(\eta)}_{i,m}) \tag{221}$$

where $\varepsilon^{p,(\eta)}_{i,m}$ and $w_i^{(\eta)}$ denote the ηth particle and its weight, respectively, N_η the number of particles and $\delta(\cdot)$ Dirac delta function. The particles are generated randomly from a

distribution $\pi(\varepsilon_{i,m}^{p})$ called importance function, and the weights are chosen as [148]

$$w_i^{(\eta)} = \frac{p\big(\varepsilon_{i,m}^{p,(\eta)}|\mathbf{y}_i^p\big)}{\pi\big(\varepsilon_{i,m}^{p,(\eta)}|\mathbf{y}_i^p\big)}. \tag{222}$$

A well known problem in particle filtering is sample degeneration that leads to performance deterioration [148]. It means that only a few particles representing the distribution have considerable weights [136]. To solve this problem, resampling can be used. However, resampling results in particle impoverishment [137]. A class of particle filters that does not undergo sample degeneration and consequently is immune to particle impoverishment, is GPF [136]. GPF also benefits from significant complexity reduction owing to avoiding resampling [137], and hence, GPF is adopted in the following. The details of the GPF can be found in [136].

According to the marginalized particle filtering, first, particles $\varepsilon_{i,m}^{p,(\eta)}$ are generated from the distribution $\pi\big(\varepsilon_{i,m}^{p}|\mathbf{y}_i^p\big)$, and next the weights are selected as

$$w_i^{(\eta)} = \frac{p\big(\mathbf{y}_i^p|\varepsilon_{i,m}^{p,(\eta)}\big)\mathcal{N}\big(\varepsilon_{i,m}^{p,(\eta)};\bar{\mu}_i^{(\eta)},(\bar{\sigma}_i^{(\eta)})^2\big)}{\pi\big(\varepsilon_{i,m}^{p,(\eta)}|\mathbf{y}_i^p\big)}, \tag{223}$$

where

$$p\big(\mathbf{y}_i^p|\varepsilon_{i,m}^{p,(\eta)}\big) = \mathcal{N}\big(\mathbf{y}_i^p;\bar{\boldsymbol{\mu}}_i^{(\eta)},\bar{\boldsymbol{\Sigma}}_i^{(\eta)}\big). \tag{224}$$

In (223), $\mathcal{N}\big(x;\mu,\sigma^2\big)$ is the univariate Gaussian density with mean μ and variance σ^2. Similarly in (224), $\mathcal{N}\big(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}\big)$ indicates the multivariate Gaussian density with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Here, $\bar{\boldsymbol{\mu}}_i^{(\eta)}$ and $\bar{\boldsymbol{\Sigma}}_i^{(\eta)}$ are the mean vector and covariance matrix of $\mathbf{y}_i^p$ and are obtained using the prediction step of BSKF as follows. Considering the mth user, by linearizing (199) with respect to nuisance parameters, we have

$$\mathbf{y}_i^p \approx \boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^{p})\boldsymbol{\zeta}_m^p + \boldsymbol{\Lambda}_{i,\backslash m}(\hat{\varepsilon}_{i|i-1,\backslash m}^{p})\hat{\boldsymbol{\zeta}}_{i|i-1,\backslash m}^{p} + \mathbf{J}_{i,\backslash m}\big(\boldsymbol{\chi}_{i,\backslash m}^p - \hat{\boldsymbol{\chi}}_{i|i-1,\backslash m}^p\big) + \boldsymbol{Z'}_i^p. \tag{225}$$

Then,

$$\bar{\boldsymbol{\mu}}_i^{(\eta)} = \boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^{p,(\eta)})\hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)} + \boldsymbol{\Lambda}_{i,\backslash m}(\hat{\varepsilon}_{i|i-1,\backslash m}^{p})\hat{\boldsymbol{\zeta}}_{i|i-1,\backslash m}^{p} \tag{226}$$

$$\bar{\boldsymbol{\Sigma}}_i^{(\eta)} \;=\; \boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^{p,(\eta)})\mathbf{P}_{i|i-1,m,m}^{(\eta)}\Big(\boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^{p,(\eta)})\Big)^T \;+\; \mathbf{J}_{i,\backslash m}\mathbf{P}_{i|i-1,\backslash m,\backslash m}\big(\mathbf{J}_{i,\backslash m}\big)^T \;+\; \frac{\sigma_z^2}{2}\mathbf{I}_{2N} \tag{227}$$

where $\hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)}$, $\mathbf{P}_{i|i-1,m,m}^{p,(\eta)}$, and $\mathbf{P}_{i|i-1,\backslash m,\backslash m}$ are obtained in the prediction step of BSKF. Representing the covariance matrix of $\boldsymbol{\vartheta}_i^p = [(\boldsymbol{\zeta}_{i,m}^p)^T, (\boldsymbol{\chi}_{i,\backslash m}^p)^T]^T$ i.e., $\mathbf{P}_{i|i-1}^{(\eta)} = \mathrm{E}\big\{\big(\boldsymbol{\vartheta}_i^p - \hat{\boldsymbol{\vartheta}}_{\boldsymbol{i}i|i-1}^{p,(\eta)}\big)\big(\boldsymbol{\vartheta}_i^p - \hat{\boldsymbol{\vartheta}}_{\boldsymbol{i}i|i-1}^{p,(\eta)}\big)^T\big\}$ as

$$\mathbf{P}_{i|i-1}^{(\eta)} = \begin{bmatrix} \mathrm{P}_{i|i-1,m,m}^{(\eta)} & \mathrm{P}_{i|i-1,m,\backslash m}^{(\eta)} \\ \mathrm{P}_{i|i-1,\backslash m,m}^{(\eta)} & \mathrm{P}_{i|i-1,\backslash m,\backslash m} \end{bmatrix} \tag{228}$$

where $\mathbf{P}_{i|i-1,m,m}^{(\eta)} = \mathrm{E}\big\{\big(\boldsymbol{\zeta}_{i,m}^p - \hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)}\big)\big(\boldsymbol{\zeta}_{i,m}^p - \hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)}\big)^T\big\}$, $\mathbf{P}_{i|i-1,m,\backslash m}^{(\eta)} = \mathrm{E}\big\{\big(\boldsymbol{\zeta}_{i,m}^p - \hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)}\big)\big(\boldsymbol{\chi}_{i,\backslash m}^p - \hat{\boldsymbol{\chi}}_{i|i-1,\backslash m}^{p}\big)^T\big\}$, $\mathbf{P}_{i|i-1,\backslash m,m}^{(\eta)} = \mathrm{E}\big\{\big(\boldsymbol{\chi}_{i,\backslash m}^p - \hat{\boldsymbol{\chi}}_{i|i-1,\backslash m}^{p}\big)\big(\boldsymbol{\zeta}_{i,m}^p - \hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)}\big)^T\big\}$, $\mathbf{P}_{i|i-1,\backslash m,\backslash m} = \mathrm{E}\big\{\big(\boldsymbol{\chi}_{i,\backslash m}^p - \hat{\boldsymbol{\chi}}_{i|i-1,\backslash m}^{p}\big)\big(\boldsymbol{\chi}_{i,\backslash m}^p - \hat{\boldsymbol{\chi}}_{i|i-1,\backslash m}^{p}\big)^T\big\}$, the prediction step of BSKF can be described as

- **Prediction Step of the BSKF:**

$$\hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)} = \boldsymbol{\Pi}_m\hat{\boldsymbol{\zeta}}_{i-1|i-1,m}^{p,(\eta)} \tag{229}$$

$$\mathbf{P}_{i|i-1,m,m}^{(\eta)} = \boldsymbol{\Pi}_m\mathbf{P}_{i-1|i-1,m,m}^{(\eta)}\boldsymbol{\Pi}_m^T + \mathbf{Q}_u \tag{230}$$

Following the prediction step of BSKF and having obtained the wights of particle filters, one can perform the correction step of BSKF as follows.

- **Correction Step of the BSKF:**

$$\boldsymbol{\mathcal{G}}^{(\eta)} = \Big(\mathbf{J}_{i,m}^{(\eta)}\mathbf{P}_{i|i-1,m,m}^{(\eta)}(\mathbf{J}_{i,m}^{(\eta)})^T + \mathbf{J}_{i,m}^{(\eta)}\mathbf{P}_{i|i-1,m,\backslash m}^{(\eta)}(\mathbf{J}_{i,\backslash m})^T + \mathbf{J}_{i,\backslash m}\mathbf{P}_{i|i-1,\backslash m,m}^{(\eta)}(\mathbf{J}_{i,m}^{(\eta)})^T + \mathbf{J}_{i,\backslash m}\mathbf{P}_{i|i-1,\backslash m,\backslash m}(\mathbf{J}_{i,\backslash m})^T + \frac{\sigma_z^2}{2}\mathbf{I}_{2N}\Big)^{-1} \quad (231)$$

$$\mathbf{K}_{i,m}^{(\eta)} = \Big(\mathbf{P}_{i|i-1,m,m}^{(\eta)}(\mathbf{J}_{i,m}^{(\eta)})^T + \mathbf{P}_{i|i-1,m,\backslash m}^{(\eta)}(\mathbf{J}_{i,\backslash m})^T\Big)\boldsymbol{\mathcal{G}}^{(\eta)} \quad (232)$$

$$\hat{\boldsymbol{\zeta}}_{i|i,m}^{p,(\eta)} = \hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)} + \mathbf{K}_{i,m}^{p}\Big(\mathbf{y}_i^p - \boldsymbol{\Lambda}_{i,m}(\varepsilon_{i,m}^{p,(\eta)})\hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)} - \boldsymbol{\Lambda}_{i,\backslash m}(\hat{\varepsilon}_{i|i-1,\backslash m}^{p})\hat{\boldsymbol{\zeta}}_{i|i-1,\backslash m}^{p}\Big) \quad (233)$$

$$\mathbf{P}_{i|i,m,m}^{(\eta)} = \boldsymbol{\mathcal{E}}^{(\eta)}\mathbf{P}_{i|i-1,m,m}^{(\eta)}(\boldsymbol{\mathcal{E}}^{(\eta)})^T - \boldsymbol{\mathcal{E}}^{(\eta)}\mathbf{P}_{i|i,m,\backslash m}^{(\eta)}(\boldsymbol{\mathcal{F}}^{(\eta)})^T - \boldsymbol{\mathcal{F}}^{(\eta)}\mathbf{P}_{i|i,\backslash m,m}^{(\eta)}(\boldsymbol{\mathcal{E}}^{(\eta)})^T + \boldsymbol{\mathcal{F}}^{(\eta)}\mathbf{P}_{i|i-1,\backslash m,\backslash m}(\boldsymbol{\mathcal{F}}^{(\eta)})^T + \frac{\sigma_z^2}{2}\mathbf{K}_{i,m}^{(\eta)}(\mathbf{K}_{i,m}^{(\eta)})^T \quad (234)$$

$$\mathbf{P}_{i|i,m,\backslash m} = \boldsymbol{\mathcal{E}}^{(\eta)}\mathbf{P}_{i|i-1,m,\backslash m} - \boldsymbol{\mathcal{F}}^{(\eta)}\mathbf{P}_{i|i-1,\backslash m,\backslash m} \quad (235)$$

$$\mathbf{P}_{i|i,\backslash m,m} = (\mathbf{P}_{i|i,m,\backslash m})^T \quad (236)$$

$$\mathbf{P}_{i|i,\backslash m,\backslash m} = \mathbf{P}_{i|i-1,\backslash m,\backslash m} \quad (237)$$

where $\mathbf{J}_{i,m}^{(\eta)} = \boldsymbol{\Lambda}_{i,m}(\hat{\varepsilon}_{i,m}^{p,(\eta)})$ is the Jacobian matrix corresponding to the ηth particle, $\boldsymbol{\mathcal{E}}^{(\eta)} = \mathbf{I}_{2N_TQL_c} - \mathbf{K}_{i,m}^{(\eta)}\mathbf{J}_{i,m}^{(\eta)}$, and $\boldsymbol{\mathcal{F}}^{(\eta)} = \mathbf{K}_{i,m}^{(\eta)}\mathbf{J}_{i,\backslash m}$.

Now, using the generated particles, BEM and CFO coefficients can be, respectively, estimated as

$$\hat{\boldsymbol{\zeta}}_{i,m}^{p} = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)}\hat{\boldsymbol{\zeta}}_{i|i,m}^{p,(\eta)}, \quad (238)$$

$$\hat{\varepsilon}_{i,m}^{p} = \mu_i = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)}\varepsilon_{i,m}^{p,(\eta)}. \quad (239)$$

Furthermore, we have

$$\mathbf{P}_{i|i,m,m} = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)}\Big(\mathbf{P}_{i|i,m,m}^{(\eta)} + \big(\hat{\boldsymbol{\zeta}}_{i|i,m}^{p,(\eta)} - \boldsymbol{\zeta}_{i,m}^{p}\big)\big(\hat{\boldsymbol{\zeta}}_{i|i,m}^{p,(\eta)} - \boldsymbol{\zeta}_{i,m}^{p}\big)^T\Big). \quad (240)$$

$$\sigma_i^2 = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)} \Big(\varepsilon_{i,m}^{p,(\eta)} - \mu_i\Big)^2. \tag{241}$$

Finally, the time update step of GPF is performed to generate particles for the next OFDM symbol as given by

$$\varepsilon_{i,m}^{p,(\eta)} \sim \mathcal{N}\big(\varepsilon_{i,m}^{p}; \mu_i, \sigma_i^2\big), \tag{242}$$

$$\varepsilon_{i+1,m}^{p,(\eta)} \sim p\big(\varepsilon_{i+1,m}^{p} | \varepsilon_{i,m}^{p,(\eta)}\big), \tag{243}$$

$$\bar{\mu}_{i+1} = \frac{1}{N_\eta} \sum_{\eta=1}^{N_\eta} \varepsilon_{i+1,m}^{p,(\eta)}, \tag{244}$$

$$\bar{\sigma}_{i+1}^2 = \frac{1}{N_\eta} \sum_{\eta=1}^{N_\eta} \Big(\varepsilon_{i+1,m}^{p,(\eta)} - \bar{\mu}_{i+1}\Big)^2. \tag{245}$$

In Table 10, we have summarized the proposed algorithm, where $\mathcal{U}(-0.5, 0.5)$ denotes the uniform distribution in the range $[-0.5, 0.5]$, and $\mathbf{0}_{L_c N_T Q \times 1}$ is an $L_c N_T Q \times 1$ zero vector.

4.5.4 Bayesian Cramer Rao Bound (BCRB)

Cramer-Rao bound provides a lower bound on the MSE of any unbiased estimator for unknown parameters [159]. In situations, where the parameters to be estimated are random and the prior information is available, BCRB is applied [160]. Here, we derive the BCRB for joint estimation of CFO and channel.

To derive the BCRB, we define $\boldsymbol{\varpi} = [\varepsilon_{i,1}^p, \mathrm{Re}\{\mathbf{h}_{i,1}^p\}^T, \mathrm{Im}\{\mathbf{h}_{i,1}^p\}^T, \cdots, \varepsilon_{i,M}^p, \mathrm{Re}\{\mathbf{h}_{i,M}^p\}^T, \mathrm{Im}\{\mathbf{h}_{i,M}^p\}^T]^T$, and consider the Baysian information matrix (BIM) as given by [107]

$$\boldsymbol{\mathcal{J}}_B = \boldsymbol{\mathcal{J}}_D + \boldsymbol{\mathcal{J}}_P. \tag{246}$$

where $\boldsymbol{\mathcal{J}}_D = \mathrm{E}_{\boldsymbol{\varpi}}\big\{[\boldsymbol{\mathcal{J}}]_{i,j}\big\}$, with $\boldsymbol{\mathcal{J}}$ being the Fisher information matrix (FIM), $[\boldsymbol{\mathcal{J}}_P]_{i,j} = -\mathrm{E}_{\boldsymbol{\varpi}}\Big\{\frac{\partial^2 \ln p(\boldsymbol{\varpi})}{\partial \boldsymbol{\varpi} \partial \boldsymbol{\varpi}^T}\Big\}$, and $\mathrm{E}_{\boldsymbol{\varpi}}\{\cdot\}$ denoting the expectation operation with respect to $\boldsymbol{\varpi}$. The

Table 10: Proposed BSK-GPF Algorithm

Initialization:

Draw $\varepsilon_{0,m}^{p,(\eta)} \sim \mathcal{U}(-0.5, 0.5)$ for $\eta = 1, 2, \cdots, N_\eta$

Compute $\bar{\mu}_0 = \frac{1}{N_\eta} \sum_{\eta=1}^{N_\eta} \varepsilon_{0,m}^{p,(\eta)}$ and $\bar{\sigma}_0^2 = \frac{1}{N_\eta} \sum_{\eta=1}^{N_\eta} \left(\varepsilon_{0,m}^{p,(\eta)} - \bar{\mu}_0 \right)^2$

Set $\hat{\boldsymbol{\zeta}}_{0|0,m}^{p,(\eta)} = \mathbf{0}_{L_c N_T Q \times 1}$ and $\mathbf{P}_{0|0}^{(\eta)} = \mathbf{I}_{ML_c N_T Q + M - 1}$ for $\eta = 1, \cdots, N_\eta$

GPF Measurement update :

For $\eta = 1 : N_\eta$

Generate $\varepsilon_{i,m}^{p,(\eta)} \sim \pi\left(\varepsilon_{i,m}^{p} | \mathbf{y}_i^p\right)$

Compute the prediction step of the BSKF: $\hat{\boldsymbol{\zeta}}_{i|i-1,m}^{p,(\eta)}$, $\mathbf{P}_{i|i-1,m,m}^{(\eta)}$

Compute $\bar{\boldsymbol{\mu}}_i^{(\eta)}$ and $\bar{\boldsymbol{\Sigma}}_i^{(\eta)}$ using (226) and (227)

Compute $p\left(\mathbf{y}_i^p | \varepsilon_{i,m}^{p,(\eta)}\right) = \mathcal{N}\left(\mathbf{y}_i^p; \bar{\boldsymbol{\mu}}_i^{(\eta)}, \bar{\boldsymbol{\Sigma}}_i^{(\eta)}\right)$

Compute $w_i^{(\eta)} = \frac{p\left(\mathbf{y}_i^p | \varepsilon_{i,m}^{p,(\eta)}\right) \mathcal{N}\left(\varepsilon_{i,m}^{p,(\eta)}; \bar{\mu}_i^{(\eta)}, (\bar{\sigma}_i^{(\eta)})^2\right)}{\pi\left(\varepsilon_{i,m}^{p,(\eta)} | \mathbf{y}_i^p\right)}$,

Compute Correction Step of the BSKF: $\mathbf{K}_{i,m}^{(\eta)}$,

$\hat{\boldsymbol{\zeta}}_{i|i,m}^{p,(\eta)}$, $\mathbf{P}_{i|i,m,m}^{(\eta)}$, $\mathbf{P}_{i|i,m,\backslash m}^{(\eta)}$, $\mathbf{P}_{i|i,\backslash m,m}^{(\eta)}$, $\mathbf{P}_{i|i,\backslash m,\backslash m}$,

End

Normalize weights $w_i^{(\eta)} = \frac{w_i^{(\eta)}}{\sum_{\eta=1}^{N_\eta} w_i^{(\eta)}}$

Estimate BEM coefficients: $\hat{\boldsymbol{\zeta}}_{i,m}^{p} = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)} \hat{\boldsymbol{\zeta}}_{i|i,m}^{p,(\eta)}$

Estimate CFO: $\mu_i = \hat{\varepsilon}_{i,m}^{p} = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)} \varepsilon_{i,m}^{p,(\eta)}$

Compute $\sigma_i^2 = \sum_{\eta=1}^{N_\eta} w_i^{(\eta)} \left(\varepsilon_{i,m}^{p,(\eta)} - \mu_i \right)^2$

Compute $\mathbf{P}_{i|i,m,m}$ according to (240)

GPF Time update :

Generate $\varepsilon_{i,m}^{p,(\eta)} \sim \mathcal{N}\left(\varepsilon_{i,m}^{p}; \mu_i, \sigma_i^2\right)$

Generate $\varepsilon_{i+1,m}^{p,(\eta)} \sim p\left(\varepsilon_{i+1,m}^{p} | \varepsilon_{i,m}^{p,(\eta)}\right)$

Compute $\bar{\mu}_{i+1} = \frac{1}{N_\eta} \sum_{\eta=1}^{N_\eta} \varepsilon_{i+1,m}^{p,(\eta)}$

Compute $\bar{\sigma}_{i+1}^2 = \frac{1}{N_\eta} \sum_{\eta=1}^{N_\eta} \left(\varepsilon_{i+1,m}^{p,(\eta)} - \bar{\mu}_{i+1} \right)^2$

FIM can be expressed in terms of its submatrices $\boldsymbol{\mathcal{J}}_{m,m'}$ as

$$\boldsymbol{\mathcal{J}}_{m,m'} = \frac{2}{\sigma_z^2} \mathrm{Re} \begin{bmatrix} \mathbb{T}_m^H \mathbb{T}_{m'} & -j \mathbb{T}_m^H \dot{\mathbf{A}}_{i,m'} & \mathbb{T}_m^H \dot{\mathbf{A}}_{i,m'} \\ j \dot{\mathbf{A}}_{i,m}^H \mathbb{T}_{m'} & \dot{\mathbf{A}}_{i,m}^H \dot{\mathbf{A}}_{i,m'} & j \dot{\mathbf{A}}_{i,m}^H \dot{\mathbf{A}}_{i,m'} \\ \dot{\mathbf{A}}_{i,m}^H \mathbb{T}_{m'} & -j \dot{\mathbf{A}}_{i,m}^H \dot{\mathbf{A}}_{i,m'} & \dot{\mathbf{A}}_{i,m}^H \dot{\mathbf{A}}_{i,m'} \end{bmatrix} \tag{247}$$

where $\mathbb{T}_m = \mathbf{\Delta\Gamma}(\varepsilon_{i,m}^p)\check{\mathbf{A}}_{i,m}\mathbf{h}_{i,m}^p$ and $\dot{\mathbf{A}}_{i,m} = \Gamma(\varepsilon_{i,m}^p)\check{\mathbf{A}}_{i,m}$. Thus, we can obtain the submatrices of $\boldsymbol{\mathcal{J}}_D$ for $m \neq m'$ as

$$\boldsymbol{\mathcal{J}}_{D,m,m'} = \frac{2}{\sigma_z^2}\mathrm{Re}\begin{bmatrix} 0 & \mathbf{0}_{1\times N_T NL_c} & \mathbf{0}_{1\times N_T NL_c} \\ \mathbf{0}_{N_T NL_c\times 1} & \check{\mathbf{A}}_{i,m}^H\check{\mathbf{\Gamma}}\check{\mathbf{A}}_{i,m'} & j\check{\mathbf{A}}_{i,m}^H\check{\mathbf{\Gamma}}\check{\mathbf{A}}_{i,m'} \\ \mathbf{0}_{N_T NL_c\times 1} & -j\check{\mathbf{A}}_{i,m}^H\check{\mathbf{\Gamma}}\check{\mathbf{A}}_{i,m'} & \check{\mathbf{A}}_{i,m}^H\check{\mathbf{\Gamma}}\check{\mathbf{A}}_{i,m'} \end{bmatrix} \tag{248}$$

where $\check{\mathbf{\Gamma}}$ is a diagonal matrix with diagonal elements $[\check{\mathbf{\Gamma}}]_{i,i} = \frac{\sin^2(2\pi na/N)}{(2\pi na/N)^2}$, and it is assumed that CFO is uniformly distributed in the range $[-a, a]$. For $m = m'$, we have

$$\boldsymbol{\mathcal{J}}_{D,m,m'} = \frac{2}{\sigma_z^2} \times \mathrm{Re}\begin{bmatrix} \mathrm{tr}(\mathbf{\Delta}\check{\mathbf{A}}_{i,m}\mathbf{R}_{i,m}^p\check{\mathbf{A}}_{i,m}^H\mathbf{\Delta}^H) & \mathbf{0}_{1\times N_T NL_c} & \mathbf{0}_{1\times N_T NL_c} \\ \mathbf{0}_{N_T NL_c\times 1} & \check{\mathbf{A}}_{i,m}^H\check{\mathbf{A}}_{i,m} & j\check{\mathbf{A}}_{i,m}^H\check{\mathbf{A}}_{i,m} \\ \mathbf{0}_{N_T NL_c\times 1} & -j\check{\mathbf{A}}_{i,m}^H\check{\mathbf{A}}_{i,m} & \check{\mathbf{A}}_{i,m}^H\check{\mathbf{A}}_{i,m} \end{bmatrix}. \tag{249}$$

To obtain $\boldsymbol{\mathcal{J}}_P$, we consider $p(\boldsymbol{\varpi}) = \prod_{m=1}^M p(\boldsymbol{\varpi}_{h,m})p(\varepsilon_{i,m}^p)$ where $\boldsymbol{\varpi}_{h,m} = [\mathrm{Re}\{\mathbf{h}_{i,m}^p\}^T, \mathrm{Im}\{\mathbf{h}_{i,m}^p\}^T]^T$, $p(\varepsilon_{i,m}^p) = \frac{1}{2a}$ and $p(\boldsymbol{\varpi}_{h,m}) = (\pi)^{-\frac{N_T NL_c}{2}}|\mathbf{R}_{\varpi_{h,m}}|^{-\frac{1}{2}}\exp\big(-\frac{1}{2}\boldsymbol{\varpi}_{h,m}^T\mathbf{R}_{\varpi_{h,m}}^{-1}\boldsymbol{\varpi}_{h,m}\big)$ with $\mathbf{R}_{\varpi_{h,m}} = E\{\boldsymbol{\varpi}_{h,m}\boldsymbol{\varpi}_{h,m}^T\}$. Therefore, $\boldsymbol{\mathcal{J}}_P$ can be represented as

$$\boldsymbol{\mathcal{J}}_P = \mathrm{blkdiag}(\boldsymbol{\mathcal{J}}_{P,1}, \cdots, \boldsymbol{\mathcal{J}}_{P,M}) \tag{250}$$

where $\boldsymbol{\mathcal{J}}_{P,m} = \mathrm{blkdiag}(0, \mathbf{R}_{\varpi_{h,m}}^{-1})$. Finally, the BCRB for the joint CFO and channel estimation can be obtained as

$$BCRB = (\boldsymbol{\mathcal{J}}_B)^{-1}. \tag{251}$$

4.5.5 Performance Evaluation

In this section, we evaluate the performance of the new methods using computer simulation. We consider a MIMO OFDMA system with $N = 128$ subcarriers, CP length of $G = 16$, and 4 users. The users and base station are respectively equipped with 3 transmit

and receive antennas. Each user transmits its data using 32 subcarriers where the sampling time and carrier frequency are $T_s = 10^{-6}$ seconds and $f_c = 2$ GHz, respectively. The channel for each antenna of each user is a multipath Rayleigh fading channel generated using Jake's model [153]. The channel has $L_c = 3$ taps with time spacings of 1 μs and exponential power delay profile with the average power of $e^{-\frac{l}{L_c}}$, $l = 0, 1, 2$. We use the generalized complex exponential BEM [105] with $Q = 2$, and unless otherwise stated, the normalized Doppler frequency is $f_d N T_s = 0.5$ which corresponds to mobile velocity of 586 km/h. For performance comparison of BSK-GPF with other schemes, the CFOs of user 1 to user 4 are set to $0.4, -0.3, 0.4, -0.3$, and for performance comparison of BSEKF approach CFOs are $0.2, -0.2, 0.2, -0.2$. Furthermore, the importance function is chosen to be $\pi\big(\varepsilon_{i,m}^p|\mathbf{y}_i^p\big) = \mathcal{N}\big(\varepsilon_{i,m}^p; \bar{\mu}_i^{(\eta)}, (\bar{\sigma}_i^{(\eta)})^2\big)$.

In Figs. 40 and 41, we have depicted, respectively, the MSE plots of CFO and channel estimates of the proposed BSK-GPF method in comparison with previous schemes using one preamble and 200 particles. The ML grid search approach followed by Newton-Rapson method in [114,117] is indicated by "ML-NR", and a particle based method similar to SK-APF in [99] that does not take into account the channel time variations during an OFDM symbol is shown by "SK-APF". We have also depicted the BCRB as a lower bound. Evidently, the new BSK-GPF method has a significantly lower MSE than other schemes considering that the Newton-Rapson scheme does not neccessarily converge to the true CFO [161], and the SK-APF cannot cope with the channel changes within one OFDMA symbol interval.

Fig. 42 demonstrates the MSE of CFO and channel estimates obtained using BSEKF approach as compared with the SEKF method in [99] using two preambles. It is oberved that the new method, by modeling channel variations within an OFDMA symbol using BEM, offers a considerably better performance.

The performances of the proposed estimators versus mobile speed at $SNR = 8\ dB$ are illustrated in Figs. 43 and 44. It is worth mentioning that the number of particles

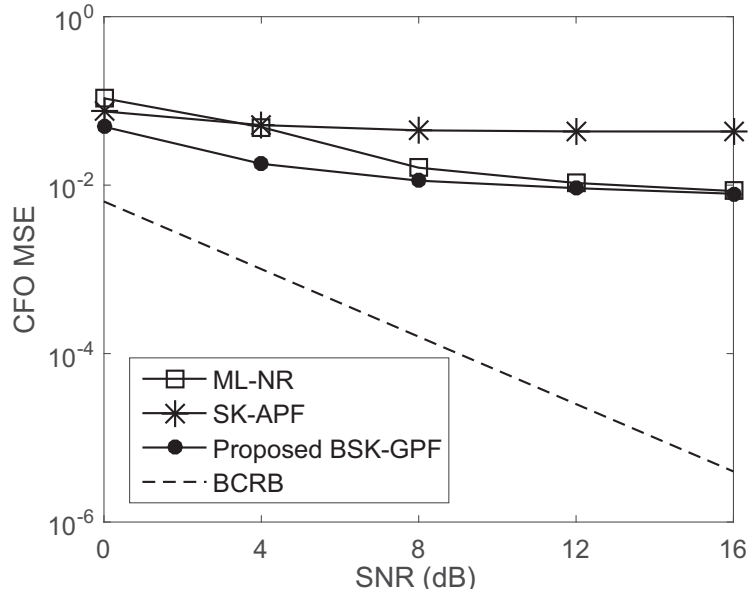

Figure 40: MSE of CFO estimates versus SNR for different particle and ML based methods

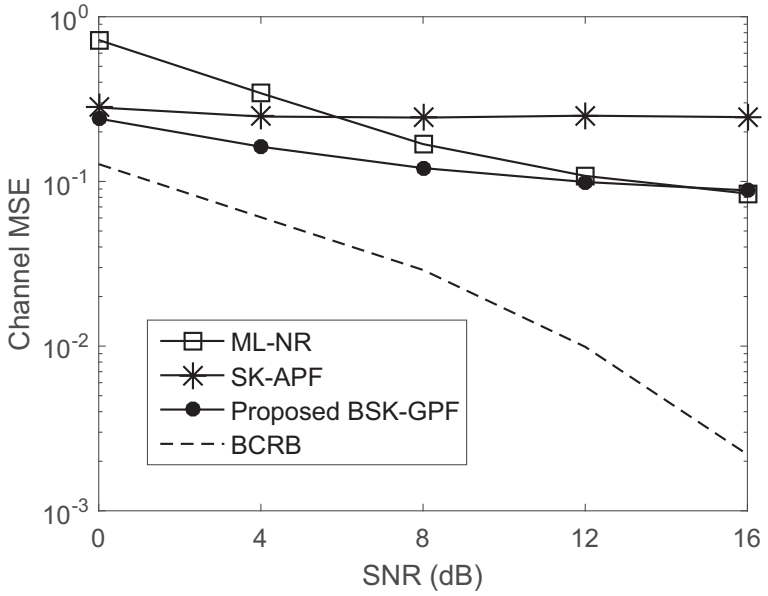

Figure 41: MSE of channel estimates versus SNR for different particle and ML based methods.

utilized for BSK-GPF and SK-APF schemes are 50. It is noticed that both of the new schemes have significantly better performance in terms of MSE in comparison with their corresponding methods in [99].

Fig. 45 shows the CFO and channel estimates of the proposed BSK-GPF algorithm

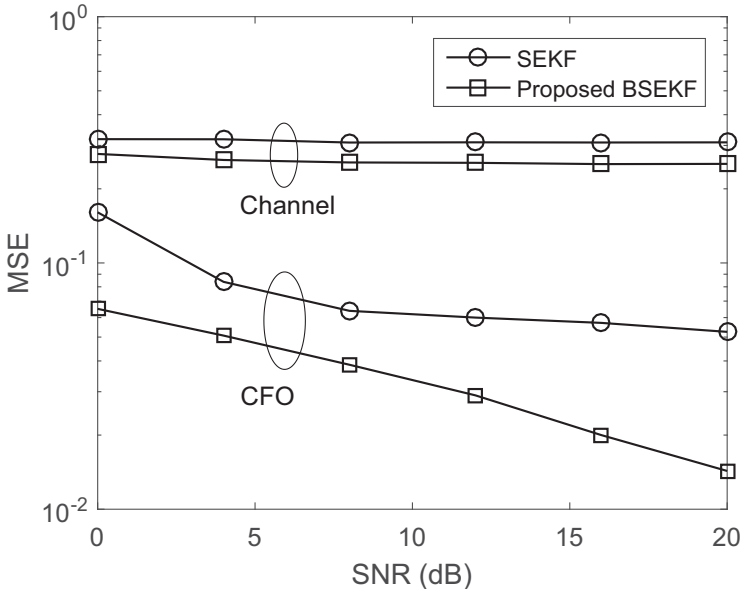

Figure 42: MSE of CFO and channel estimates versus SNR for the SEKF and BSEKF methods.

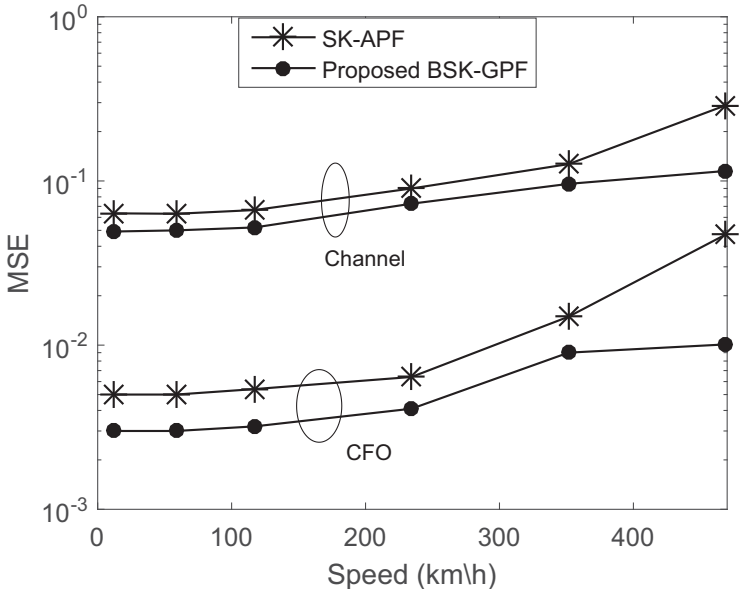

Figure 43: MSE of CFO and channel estimates versus mobile speed at $SNR = 8$ dB for particle-based methods.

versus SNR for different number of particles. It is observed that as the number of particle increases the performance improves, which is due to the fact that increasing the number of particles results in better estimation of the PDF.

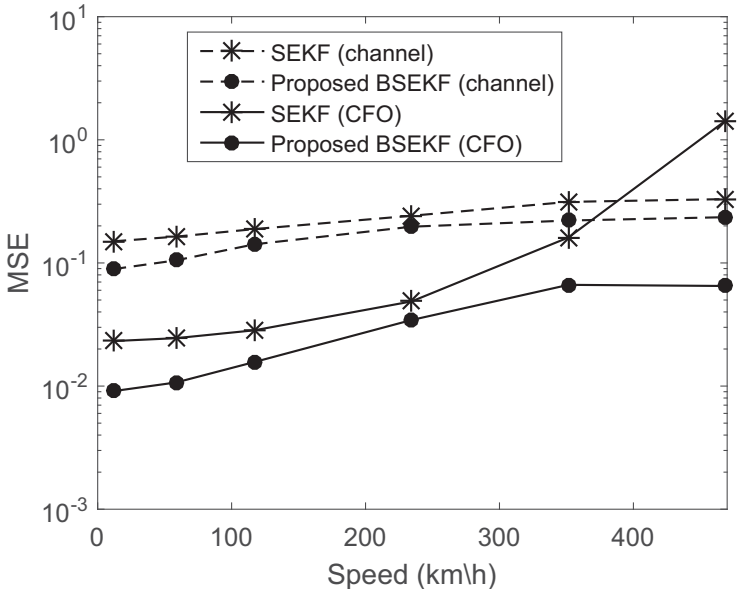

Figure 44: MSE of CFO and channel estimates versus mobile speed at $SNR = 8$ dB for SEKF based approaches.

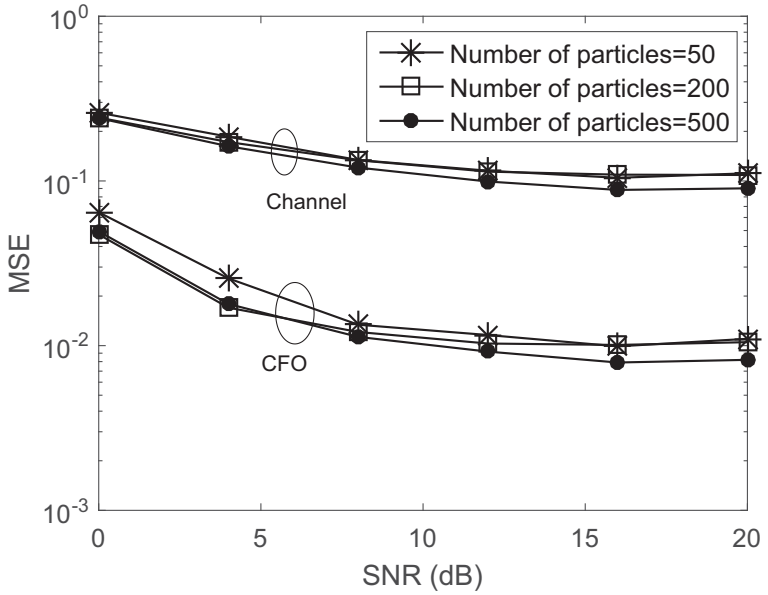

Figure 45: MSE of CFO and channel estimates versus SNR for different number of particles.

In Figs. 46 a and b, the tracking performances of the proposed approaches are depicted where $SNR = 15\ dB$ and 200 particles are used. As expected, the MSE decreases as the number of preambles increases.

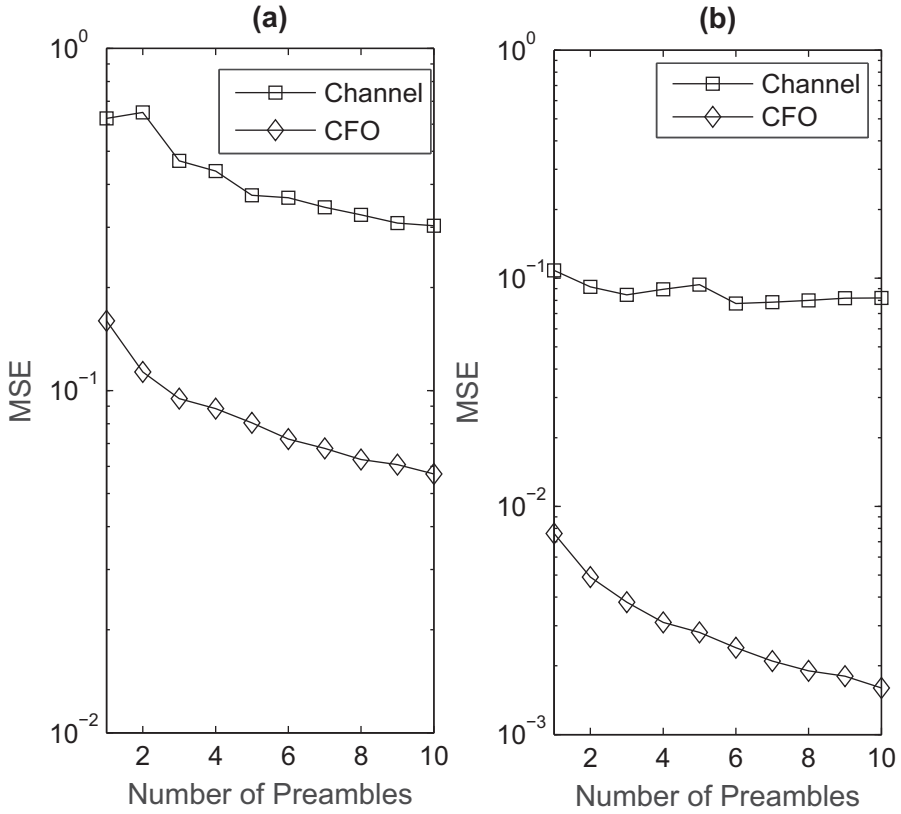

Figure 46: MSE of CFO and channel estimates for different number of preambles a) BSEKF approach, b) BSK-GPF approach.

4.6 Conclusion

In this section, first, a novel method for joint estimation of timing, frequency offset and multipath fast fading channel has been proposed for OFDM systems. The new scheme has been developed using the ML criterion, transmission of a preamble and BEM of the channel. A new scheme for complexity reduction of the joint estimation has been also introduced based on a low complexity CFO estimation algorithm. It has been demonstrated using MSE criterion that the new method significantly improves the timing, frequency offset and channel estimation performance in comparison with previous methods in the literature.

Second, we have proposed a preamble-aided method based on particle and Kalman filtering along with BEM for joint CFO and doubly selective channel estimation in OFDM

systems. The performance of the new method has been evaluated in a frequency selective fast fading channel using MSE of CFO and channel estimates. It has been shown that the new scheme can significantly improve the estimation performance in comparison with previous schemes in the literature.

Third, this section presents a new scheme for timing and frequency synchronization and channel estimation in the uplink transmission of OFDMA systems in doubly selective channels. For joint estimation of timing, frequency offsets and a doubly selective channel, a new preamble-aided estimator is derived using the ML criterion along with the BEM of channel and SAGE algorithm. Performance evaluation revealed considerable improvement by the new estimation method compared with previous schemes in terms of MSE.

Fourth, two new preamble-aided joint estimators of CFO and time varying channel have been proposed for the uplink of MIMO OFDMA systems. It is assumed that CFOs of users vary from one OFDMA symbol to another, and also the channel changes within one OFDMA symbol. The doubly selective channel is modeled using BEM to reduce the number of unknown parameters. The first estimator is based on SEKF, and the second one takes advantages of GPF and SKF by adopting a marginalized Kalman filtering approach. The BCRB has also been derived for the estimation of CFO and doubly selective channel. The performances of the new schemes for different SNRs, mobile velocities and number of preambles are evaluated. The results indicate that the new schemes have considerably lower MSEs compared with previous methods.

Chapter 5

Conclusion and Future Work

5.1 Concluding Remarks

In this dissertation, timing and frequency synchronization and channel estimation in OFDM-based systems have been investigated and several approaches for estimation of timing and frequency offsets and channel have been proposed. First, we addressed timing synchronization in an OFDM system in a frequency selective slow fading channel using a preamble composed of two identical parts in the time domain [123]. Two novel timing metrics based on n-th order correlation and differential normalization functions have been proposed and their performances have been assessed using different criteria such as class-separability, robustness to CFO, computational complexity, and false alarm and missed detection probabilities. The effect of increasing the order of the timing metric has also been investigated. A new method for coarse estimation of the start of the frame has been presented which, instead of using only the timing instant corresponding to the maximum of the timing metric (as performed in previous works), makes use of a set of timing instants corresponding to the largest values of the timing metric. The performances of the new timing schemes in multipath frequency selective fading channels have been evaluated through computer simulations. By evaluating the false alarm and missed-detection probabilities, it has been demonstrated that the new methods remarkably improve the detection

performance as compared with previous methods. Finally, it has been shown that the new methods also offer a significant reduction in the ISI probability.

Second, we have presented a new IFO detection method for OFDM systems in multipath slow fading channels [124]. The new method, by transforming IFO detection into the detection of two new integer parameters, significantly reduces the number of trial values of IFO. A general procedure for application of the new method to previously proposed IFO metrics in the literature has been presented. Different criteria including computational complexity, detection probability and BER have been used to evaluate the performance of the new method. It has been shown that the new method can significantly reduce the computational complexity while maintaining almost the same detection performance as compared to previous schemes.

Third, a novel method for joint estimation of timing, frequency offset and doubly selective channel has been proposed [125]. The new scheme is based on ML criterion, transmission of a preamble and BEM of the channel. Since the direct maximization of the ML function is not practical, first a timing and frequency offsets dependent estimate of the channel is obtained and put back in the ML function. Next, maximization is performed with respect to timing offset and CFO. Furthermore, a new scheme for complexity reduction of the joint estimation has been proposed based on a low complexity CFO estimation algorithm. It has been demonstrated using MSE criterion that the new method significantly improves the timing, frequency offset and channel estimation performance in comparison with previous methods.

Fourth, we have introduced a preamble-aided method based on particle and Kalman filtering along with BEM for joint CFO and multipath time varying channel estimation in OFDM systems [126]. The performance of the new method has been evaluated using MSE of CFO and channel estimates. It has been shown that the new scheme can considerably improve the estimation performance compared with previous schemes in the literature.

Fifth, a new scheme for timing and frequency synchronization and channel estimation

in the uplink transmission of OFDMA systems in doubly selective channels has been proposed [128]. For joint estimation of timing, frequency offsets and a doubly selective channel, a new preamble-aided ML estimator is derived based on the SAGE algorithm that transforms an M dimensional maximization problem into M simpler one dimensional maximization problems. Performance evaluation revealed considerable improvement by the new estimation method compared with previous schemes in terms of MSE.

Finally, we have presented two new preamble-aided joint estimators of CFO and time varying channel for the uplink of MIMO OFDMA systems [129]. It is assumed that CFOs of users vary from one OFDMA symbol to another, and also the channel changes within an OFDMA symbol. The doubly selective channel is modeled using BEM to reduce the number of unknown parameters. The first estimator is based on SEKF, and the second estimator makes use of GPF and SKF by adopting a marginalized Kalman filtering approach. The BCRB for the estimation of CFO and doubly selective channel has been presented, and the performances of the new methods have been evaluated using MSE criterion. It has been shown that the new schemes have an improved performance compared with previous methods.

5.2 Suggestions for Future Work

There are a number of interesting topics related to timing and frequency synchronization in OFDM-based systems that can be subjects of further research. In particular, some of the recommendations for future work are as follows.

- Synchronization in OFDM-based cooperative systems: The basic principle in cooperative systems is to construct a virtual multiple-antenna system by sharing antennas of neighboring users in a distributed manner. Thus, the same benefits of multiple-antenna systems can be achieved with proper cooperative strategies [17]. The concept of cooperative relaying is based on the fact that a transmitted signal

can be received by multiple relays which retransmit a processed version of the received signal to the destination [162]. Multiple nodes in these cooperative systems are distributed in space and have their own oscillators, which results in multiple timing offsets and CFO in cooperative transmission. These offsets can drastically undermine the potential of cooperative networks. Therefore, accurate timing and frequency synchronization is critical in OFDM-based cooperative systems [17, 163].

- Synchronization in massive multiuser MIMO OFDM systems: Large-scale MIMO or "massive MIMO" systems is potentially one of the key technologies to achieve high capacity performance in the fifth generation of mobile cellular systems [164–166]. Massive MIMO can provide high data throughput and high power efficiency along with improvement in communication reliability with fairly simple signal processing [167,168]. However, these potential benefits of massive MIMO systems are dependent heavily on synchronization. Achieving perfect synchronization in massive multiuser MIMO OFDM can be more challenging owing to the existence of multiple timing and frequency offsets between multiple users and the BS [169].

- Synchronization in NC-OFDM systems: Non-Contiguous OFDM (NC-OFDM) plays an important role in cognitive radio communications [170, 171]. In an NC-OFDM transmitter, only the subcarriers that do not coincide with the primary users' band are modulated with data symbols, and the other ones are modulated by zeros in order to limit the interference power observed at the primary users receiver [172]. In these systems, receivers have to be protected from the interference introduced by cognitive radio transmission. Thus, an important design issue in these systems is an interference-robust synchronization algorithm [171].

Bibliography

[1] M. Morelli, C.-C. Kuo, and M.-O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (ofdma): A tutorial review," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1394–1427, July 2007.

[2] "Digital video broadcasting (dvb-t); frame structure, channel coding, modulation for digital terrestrial television," *ETS 300 744, Eur. Telecommun. Standard*, 1997.

[3] "Radio broadcasting systems: Digital audio broadcasting to mobile, portable and fixed receivers," *ETS 300 401, Eur. Telecommun. Standard*, 1995.

[4] "Part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications, higher-speed physical layer extension in the 5 ghz band," *IEEE802.11a*, 1999.

[5] L. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *Communications, IEEE Transactions on*, vol. 33, no. 7, pp. 665–675, Jul 1985.

[6] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: a convenient framework for time-frequency processing in wireless communications," *Proceedings of the IEEE*, vol. 88, no. 5, pp. 611–640, May 2000.

[7] D. Toumpakaris, J. Lee, and H.-L. Lou, "Estimation of integer carrier frequency offset in ofdm systems based on the maximum likelihood principle," *Broadcasting, IEEE Transactions on*, vol. 55, no. 1, pp. 95–108, March 2009.

[8] B. Ai, Z.-X. Yang, C.-Y. Pan, J. hua Ge, Y. Wang, and Z. Lu, "On the synchronization techniques for wireless ofdm systems," *Broadcasting, IEEE Transactions on*, vol. 52, no. 2, pp. 236–244, June 2006.

[9] Y. Mostofi and D. Cox, "Mathematical analysis of the impact of timing synchronization errors on the performance of an ofdm system," *Communications, IEEE Transactions on*, vol. 54, no. 2, pp. 226–230, Feb 2006.

[10] M. Speth, S. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using ofdm. i," *Communications, IEEE Transactions on*, vol. 47, no. 11, pp. 1668–1677, Nov 1999.

[11] M.-H. Hsieh and C.-H. Wei, "Channel estimation for ofdm systems based on comb-type pilot arrangement in frequency selective fading channels," *Consumer Electronics, IEEE Transactions on*, vol. 44, no. 1, pp. 217–225, Feb 1998.

[12] Y. Mostofi and D. Cox, "Analysis of the effect of timing synchronization errors on pilot-aided ofdm systems," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1, Nov 2003, pp. 638–642 Vol.1.

[13] C. Athaudage and A. Jayalath, "Enhanced mmse channel estimation using timing error statistics for wireless ofdm systems," *Broadcasting, IEEE Transactions on*, vol. 50, no. 4, pp. 369–376, Dec 2004.

[14] M. Ruan, M. Reed, and Z. Shi, "Training symbol based coarse timing synchronization in ofdm systems," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 5, pp. 2558–2569, May 2009.

[15] E. G. Larsson, G. Liu, J. Li, and G. B. Giannakis, "Joint symbol timing and channel estimation for ofdm based wlans," *IEEE Communications Letters*, vol. 5, no. 8, pp. 325–327, Aug 2001.

[16] J. Chen, Y. c. Wu, S. Ma, and T. s. Ng, "Ml joint cfo and channel estimation in ofdm systems with timing ambiguity," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2436–2440, July 2008.

[17] Q. Huang, M. Ghogho, J. Wei, and P. Ciblat, "Practical timing and frequency synchronization for ofdm-based cooperative systems," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3706–3716, July 2010.

[18] W.-L. Chin, "Blind symbol synchronization for ofdm systems using cyclic prefix in time-variant and long-echo fading channels," *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 1, pp. 185–195, Jan 2012.

[19] ——, "Ml estimation of timing and frequency offsets using distinctive correlation characteristics of ofdm signals over dispersive fading channels," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 2, pp. 444–456, Feb 2011.

[20] J.-J. van de Beek, M. Sandell, and P. Borjesson, "Ml estimation of time and frequency offset in ofdm systems," *Signal Processing, IEEE Transactions on*, vol. 45, no. 7, pp. 1800–1805, Jul 1997.

[21] ——, "Ml estimation of time and frequency offset in ofdm systems," *Signal Processing, IEEE Transactions on*, vol. 45, no. 7, pp. 1800–1805, Jul 1997.

[22] D. Lee and K. Cheun, "Coarse symbol synchronization algorithms for ofdm systems in multipath channels," *Communications Letters, IEEE*, vol. 6, no. 10, pp. 446–448, Oct 2002.

[23] S. Ma, X. Pan, G.-H. Yang, and T.-S. Ng, "Blind symbol synchronization based on cyclic prefix for ofdm systems," *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 4, pp. 1746–1751, May 2009.

[24] K. Ramasubramanian and K. Baum, "An ofdm timing recovery scheme with inherent delay-spread estimation," in *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, vol. 5, 2001, pp. 3111–3115 vol.5.

[25] R. Negi and J. Cioffi, "Blind ofdm symbol synchronization in isi channels," *Communications, IEEE Transactions on*, vol. 50, no. 9, pp. 1525–1534, Sep 2002.

[26] C. Williams, M. Beach, and S. McLaughlin, "Robust ofdm timing synchronisation," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 4, May 2006, pp. 1947–1950.

[27] V. Krishnamurthy, C. Athaudage, and D. Huang, "Adaptive ofdm synchronization algorithms based on discrete stochastic approximation," *Signal Processing, IEEE Transactions on*, vol. 53, no. 4, pp. 1561–1574, April 2005.

[28] J. Manco-Vasquez, V. Jimenez, and M. Fernandez-Getino Garcia, "Robust timing synchronization for multicarrier systems based on rst invariance," *Communications Letters, IEEE*, vol. 17, no. 6, pp. 1244–1247, June 2013.

[29] A. Al-Dweik, S. Younis, A. Hazmi, C. Tsimenidis, and B. Sharif, "Efficient ofdm symbol timing estimator using power difference measurements," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 2, pp. 509–520, Feb 2012.

[30] P. S. Wang and D. W. Lin, "On maximum-likelihood blind synchronization over wssus channels for ofdm systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 19, pp. 5045–5059, Oct 2015.

[31] T. Schmidl and D. Cox, "Robust frequency and timing synchronization for ofdm," *Communications, IEEE Transactions on*, vol. 45, no. 12, pp. 1613–1621, Dec 1997.

[32] A. Coulson, "Maximum likelihood synchronization for ofdm using a pilot symbol: algorithms," *Selected Areas in Communications, IEEE Journal on*, vol. 19, no. 12, pp. 2486–2494, Dec 2001.

[33] ——, "Maximum likelihood synchronization for ofdm using a pilot symbol: analysis," *Selected Areas in Communications, IEEE Journal on*, vol. 19, no. 12, pp. 2495–2503, Dec 2001.

[34] H. Minn, M. Zeng, and V. Bhargava, "On timing offset estimation for ofdm systems," *Communications Letters, IEEE*, vol. 4, no. 7, pp. 242–244, July 2000.

[35] H. Minn, V. Bhargava, and K. Letaief, "A robust timing and frequency synchronization for ofdm systems," *Wireless Communications, IEEE Transactions on*, vol. 2, no. 4, pp. 822–839, July 2003.

[36] K. Shi and E. Serpedin, "Coarse frame and carrier synchronization of ofdm systems: a new metric and comparison," *Wireless Communications, IEEE Transactions on*, vol. 3, no. 4, pp. 1271–1284, July 2004.

[37] B. Park, H. Cheon, C. Kang, and D. Hong, "A novel timing estimation method for ofdm systems," *Communications Letters, IEEE*, vol. 7, no. 5, pp. 239–241, May 2003.

[38] J. Zhang and X. Huang, "Autocorrelation based coarse timing with differential normalization," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 2, pp. 526–530, February 2012.

[39] H. Abdzadeh-Ziabari and M. Shayesteh, "Sufficient statistics, classification, and a novel approach for frame detection in ofdm systems," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 6, pp. 2481–2495, July 2013.

[40] A. Mohebbi, H. Abdzadeh-Ziabari, and M. Shayesteh, "Novel coarse timing synchronization methods in ofdm systems using fourth-order statistics," *Vehicular Technology, IEEE Transactions on*, vol. 64, no. 5, pp. 1904–1917, May 2015.

[41] G. Ren, Y. Chang, H. Zhang, and H. Zhang, "Synchronization method based on a new constant envelop preamble for ofdm systems," *Broadcasting, IEEE Transactions on*, vol. 51, no. 1, pp. 139–143, March 2005.

[42] Y. Kang, S. Kim, D. Ahn, and H. Lee, "Timing estimation for ofdm systems by using a correlation sequence of preamble," *Consumer Electronics, IEEE Transactions on*, vol. 54, no. 4, pp. 1600–1608, November 2008.

[43] A. Awoseyila, C. Kasparis, and B. Evans, "Improved preamble-aided timing estimation for ofdm systems," *Communications Letters, IEEE*, vol. 12, no. 11, pp. 825–827, November 2008.

[44] ——, "Robust time-domain timing and frequency synchronization for ofdm systems," *Consumer Electronics, IEEE Transactions on*, vol. 55, no. 2, pp. 391–399, May 2009.

[45] H. Abdzadeh-Ziabari, M. Shayesteh, and M. Manaffar, "An improved timing estimation method for ofdm systems," *Consumer Electronics, IEEE Transactions on*, vol. 56, no. 4, pp. 2098–2105, November 2010.

[46] H. Abdzadeh-Ziabari and M. Shayesteh, "Robust timing and frequency synchronization for ofdm systems," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 8, pp. 3646–3656, Oct 2011.

[47] C. L. Wang and H. C. Wang, "On joint fine time adjustment and channel estimation for ofdm systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 10, pp. 4940–4944, October 2009.

[48] ——, "Optimized joint fine timing synchronization and channel estimation for mimo systems," *IEEE Transactions on Communications*, vol. 59, no. 4, pp. 1089–1098, April 2011.

[49] M. M. U. Gul, X. Ma, and S. Lee, "Timing and frequency synchronization for ofdm downlink transmissions using zadoff-chu sequences," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1716–1729, March 2015.

[50] Y. Huang, A. Hu, Y. Huang, S. Xie, D. Zhu, and M. Xue, "An integer time delay estimation algorithm based on zadoff-chu sequence in ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2941–2947, July 2014.

[51] "Blind carrier-frequency offset estimation in siso, mimo, and multiuser ofdm systems," *Communications, IEEE Transactions on*, vol. 52, no. 10, pp. 1832–1832, Oct 2004.

[52] T. Roman, S. Visuri, and V. Koivunen, "Blind frequency synchronization in ofdm via diagonality criterion," *Signal Processing, IEEE Transactions on*, vol. 54, no. 8, pp. 3125–3135, Aug 2006.

[53] J.-H. Oh, J.-G. Kim, and J.-T. Lim, "Blind carrier frequency offset estimation for ofdm systems with constant modulus constellations," *Communications Letters, IEEE*, vol. 15, no. 9, pp. 971–973, September 2011.

[54] S. Lmai, A. Bourre, C. Laot, and S. Houcke, "An efficient blind estimation of carrier frequency offset in ofdm systems," *Vehicular Technology, IEEE Transactions on*, vol. 63, no. 4, pp. 1945–1950, May 2014.

[55] A. Jayaprakash and G. R. Reddy, "Covariance-fitting-based blind carrier frequency offset estimation method for ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 101–10 105, Dec 2016.

[56] W. Zhang and Q. Yin, "Blind carrier frequency offset estimation for mimo-ofdm with constant modulus constellations via rank reduction criterion," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6809–6815, Aug 2016.

[57] T. C. Lin and S. M. Phoong, "A new cyclic-prefix based algorithm for blind cfo estimation in ofdm systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3995–4008, June 2016.

[58] D. Huang and K. Letaief, "Carrier frequency offset estimation for ofdm systems using subcarriers," *Communications, IEEE Transactions on*, vol. 54, no. 5, pp. 813–823, May 2006.

[59] J. Zhu and W. Lee, "Carrier frequency offset estimation for ofdm systems with null subcarriers," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 5, pp. 1677–1690, Sept 2006.

[60] Y. Wu, S. Attallah, and J. W. M. Bergmans, "On the optimality of the null subcarrier placement for blind carrier offset estimation in ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2109–2115, May 2009.

[61] U. Tureli, H. Liu, and M. Zoltowski, "A high efficiency carrier estimator for ofdm communications," in *Signals, Systems amp; Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*, vol. 1, Nov 1997, pp. 505–509 vol.1.

[62] W. Zhang and F. Gao, "Blind frequency synchronization for multiuser ofdm uplink with large number of receive antennas," *IEEE Transactions on Signal Processing*, vol. 64, no. 9, pp. 2255–2268, May 2016.

[63] M. Morelli and U. Mengali, "An improved frequency offset estimator for ofdm applications," *Communications Letters, IEEE*, vol. 3, no. 3, pp. 75–77, March 1999.

[64] Z. Zhang, K. Long, and Y. Liu, "Complex efficient carrier frequency offset estimation algorithm in ofdm systems," *Broadcasting, IEEE Transactions on*, vol. 50, no. 2, pp. 159–164, June 2004.

[65] D. Huang and K. Letaief, "Enhanced carrier frequency offset estimation for ofdm using channel side information," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 10, pp. 2784–2793, Oct 2006.

[66] G. Ren, Y. Chang, H. Zhang, and H. Zhang, "An efficient frequency offset estimation method with a large range for wireless ofdm systems," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 4, pp. 1892–1895, July 2007.

[67] J.-W. Choi, J. Lee, Q. Zhao, and H.-L. Lou, "Joint ml estimation of frame timing and carrier frequency offset for ofdm systems employing time-domain repeated preamble," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 1, pp. 311–317, January 2010.

[68] H.-T. Hsieh and W.-R. Wu, "Maximum likelihood timing and carrier frequency offset estimation for ofdm systems with periodic preambles," *Vehicular Technology, IEEE Transactions on*, vol. 58, no. 8, pp. 4224–4237, Oct 2009.

[69] M. Morelli and M. Moretti, "Carrier frequency offset estimation for ofdm direct-conversion receivers," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2670–2679, July 2012.

[70] Z. Liu, B. Weng, and Q. Zhu, "Frequency offset estimation for differential ofdm," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1737–1748, July 2005.

[71] M. Morelli and M. Moretti, "Integer frequency offset recovery in ofdm transmissions over selective channels," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 12, pp. 5220–5226, December 2008.

[72] D. Li, Y. Li, H. Zhang, L. Cimini, and Y. Fang, "Integer frequency offset estimation for ofdm systems with residual timing offset over frequency selective fading channels," *Vehicular Technology, IEEE Transactions on*, vol. 61, no. 6, pp. 2848–2853, July 2012.

[73] Y. Sun, Z. Xiong, and X. Wang, "Em-based iterative receiver design with carrier-frequency offset estimation for mimo ofdm systems," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 581–586, April 2005.

[74] J. C. Lin, Y. T. Sun, and H. V. Poor, "Initial synchronization exploiting inherent diversity for the lte sector search process," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1114–1128, Feb 2016.

[75] T. Strohmer and S. Beaver, "Optimal ofdm design for time-frequency dispersive channels," *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1111–1122, July 2003.

[76] M. Morelli, C. C. J. Kuo, and M. O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (ofdma): A tutorial review," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1394–1427, July 2007.

[77] J. J. van de Beek, P. O. Borjesson, M. L. Boucheret, D. Landstrom, J. M. Arenas, P. Odling, C. Ostberg, M. Wahlqvist, and S. K. Wilson, "A time and frequency synchronization scheme for multiuser ofdm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1900–1914, Nov 1999.

[78] S. Barbarossa, M. Pompili, and G. B. Giannakis, "Channel-independent synchronization of orthogonal frequency division multiple access systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 2, pp. 474–486, Feb 2002.

[79] Z. Cao, U. Tureli, and Y.-D. Yao, "Efficient structure-based carrier frequency offset estimation for interleaved ofdma uplink," in *Communications, 2003. ICC '03. IEEE International Conference on*, vol. 5, May 2003, pp. 3361–3365 vol.5.

[80] ——, "Deterministic multiuser carrier-frequency offset estimation for interleaved ofdma uplink," *IEEE Transactions on Communications*, vol. 52, no. 9, pp. 1585–1594, Sept 2004.

[81] J. Lee, S. Lee, K. J. Bang, S. Cha, and D. Hong, "Carrier frequency offset estimation using esprit for interleaved ofdma uplink systems," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, pp. 3227–3231, Sept 2007.

[82] H. T. Hsieh and W. R. Wu, "Blind maximum-likelihood carrier-frequency-offset estimation for interleaved ofdma uplink systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 160–173, Jan 2011.

[83] W. Zhang, F. Gao, Q. Yin, and A. Nallanathan, "Blind carrier frequency offset estimation for interleaved ofdma uplink," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3616–3627, July 2012.

[84] S. W. Keum, D. H. Kim, and H. M. Kim, "An improved frequency offset estimation based on companion matrix in multi-user uplink interleaved ofdma systems," *IEEE Signal Processing Letters*, vol. 21, no. 4, pp. 409–413, April 2014.

[85] P. Cheng, Z. Chen, F. de Hoog, and C. K. Sung, "Sparse blind carrier-frequency offset estimation for ofdma uplink," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 5254–5265, Dec 2016.

[86] W. C. Huang, C. H. Pan, C. P. Li, and H. J. Li, "Subspace-based semi-blind channel estimation in uplink ofdma systems," *IEEE Transactions on Broadcasting*, vol. 56, no. 1, pp. 58–65, March 2010.

[87] M. O. Pun, M. Morelli, and C. C. J. Kuo, "Maximum-likelihood synchronization and channel estimation for ofdma uplink transmissions," *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 726–736, April 2006.

[88] M. o. Pun, M. Morelli, and C. c. J. Kuo, "Iterative detection and frequency synchronization for ofdma uplink transmissions," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 629–639, Feb 2007.

[89] Y. Na and H. Minn, "Line search based iterative joint estimation of channels and frequency offsets for uplink ofdm systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 12, pp. 4374–4382, December 2007.

[90] X. N. Zeng and A. Ghrayeb, "Joint cfo and channel estimation for ofdma uplink: an application of the variable projection method," *IEEE Transactions on Wireless Communications*, vol. 8, no. 5, pp. 2306–2311, May 2009.

[91] L. Sanguinetti and M. Morelli, "A low-complexity scheme for frequency estimation in uplink ofdma systems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2430–2437, August 2010.

[92] K. Lee, S. H. Moon, S. R. Lee, and I. Lee, "Low complexity pilot assisted carrier frequency offset estimation for ofdma uplink systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 2690–2695, August 2012.

[93] H. Solis-Estrella and A. G. Orozco-Lugo, "Carrier frequency offset estimation in ofdma using digital filtering," *IEEE Wireless Communications Letters*, vol. 2, no. 2, pp. 199–202, April 2013.

[94] P. Sun and L. Zhang, "Low complexity pilot aided frequency synchronization for ofdma uplink transmission," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3758–3769, July 2009.

[95] Z. Wang, Y. Xin, and G. Mathew, "Iterative carrier-frequency offset estimation for generalized ofdma uplink transmission," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1373–1383, March 2009.

[96] H. Wang and Q. Yin, "Multiuser carrier frequency offsets estimation for ofdma uplink with generalized carrier assignment scheme," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3347–3353, July 2009.

[97] L. Bai and Q. Yin, "Frequency synchronization for the ofdma uplink based on the tile structure of ieee 802.16e," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 5, pp. 2348–2353, Jun 2012.

[98] Z. Jiang, Z. Li, X. Zhang, and D. Yang, "Joint frequency offset and channel estimation using rao-blackwellized particle filter for uplink mimo-ofdma systems," in *2008 IEEE International Conference on Communications*, May 2008, pp. 698–702.

[99] K. J. Kim, M. O. Pun, and R. A. Iltis, "Joint carrier frequency offset and channel estimation for uplink mimo-ofdma systems using parallel schmidt rao-blackwellized particle filters," *IEEE Transactions on Communications*, vol. 58, no. 9, pp. 2697–2708, September 2010.

[100] T. Hrycak, S. Das, G. Matz, and H. G. Feichtinger, "Practical estimation of rapidly varying channels for ofdm systems," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3040–3048, November 2011.

[101] H. Hijazi and L. Ros, "Joint data qr-detection and kalman estimation for ofdm time-varying rayleigh channel complex gains," *IEEE Transactions on Communications*, vol. 58, no. 1, pp. 170–178, January 2010.

[102] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct 1998.

[103] H. A. Cirpan and M. K. Tsatsanis, "Maximum likelihood blind channel estimation in the presence of doppler shifts," *IEEE Transactions on Signal Processing*, vol. 47, no. 6, pp. 1559–1569, Jun 1999.

[104] T. Zemen and C. F. Mecklenbrauker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3597–3607, Sept 2005.

[105] G. Leus, "On the estimation of rapidly time-varying channels," in *2004 12th European Signal Processing Conference*, Sept 2004, pp. 2227–2230.

[106] D. K. Borah and B. T. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 862–873, Jun 1999.

[107] H. Nguyen-Le and T. Le-Ngoc, "Pilot-aided joint cfo and doubly-selective channel estimation for ofdm transmissions," *IEEE Transactions on Broadcasting*, vol. 56, no. 4, pp. 514–522, Dec 2010.

[108] E. P. Simon, L. Ros, H. Hijazi, J. Fang, D. P. Gaillot, and M. Berbineau, "Joint carrier frequency offset and fast time-varying channel estimation for mimo-ofdm systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 955–965, March 2011.

[109] E. P. Simon, L. Ros, H. Hijazi, and M. Ghogho, "Joint carrier frequency offset and channel estimation for ofdm systems via the em algorithm in the presence of very high mobility," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 754–765, Feb 2012.

[110] J.-C. Lin, "Maximum-likelihood frame timing instant and frequency offset estimation for ofdm communication over a fast rayleigh-fading channel," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 4, pp. 1049–1062, July 2003.

[111] T. Lv, H. Li, and J. Chen, "Joint estimation of symbol timing and carrier frequency offset of ofdm signals over fast time-varying multipath channels," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4526–4535, Dec 2005.

[112] Y. Mostofi and D. C. Cox, "Robust timing synchronization design in ofdm systems - part ii: high-mobility cases," *IEEE Transactions on Wireless Communications*, vol. 6, no. 12, pp. 4340–4348, December 2007.

[113] W. L. Chin, "Blind symbol synchronization for ofdm systems using cyclic prefix in time-variant and long-echo fading channels," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 1, pp. 185–195, Jan 2012.

[114] P. Muneer and S. M. Sameer, "Pilot-aided joint estimation of doubly selective channel and carrier frequency offsets in ofdma uplink with high-mobility users," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 1, pp. 411–417, Jan 2015.

[115] ——, "Joint ml estimation of cfo and channel, and a low complexity turbo equalization technique for high mobility ofdma uplinks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3642–3654, July 2015.

[116] ——, "Iterative joint carrier frequency offset and doubly selective channel estimation in high-mobility mimo-ofdma uplink using oblique projection," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7110–7121, Sept 2016.

[117] P. Muneer and S. Sameer, "Joint estimation of doubly selective channels and carrier frequency offsets in high mobility mimo-ofdma uplink," *Wireless Personal Communications*, vol. 95, no. 4, pp. 4113–4130, 2017.

[118] S. L. Talbot and B. Farhang-Boroujeny, "Time-varying carrier offsets in mobile ofdm," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2790–2798, September 2009.

[119] M.-O. Pun, M. Morelli, and C. C. J. Kuo, "An ecm-based receiver for ofdm signals affected by carrier frequency offsets," in *VTC-2005-Fall. 2005 IEEE 62nd Vehicular Technology Conference, 2005.*, vol. 1, Sept 2005, pp. 640–644.

[120] W. Hou, X. Wang, J. Y. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1658–1667, May 2014.

[121] T. Nyblom, T. Roman, M. Enescu, and V. Koivunen, "Time-varying carrier offset tracking in ofdm systems using particle filtering," in *Proceedings of the Fourth IEEE International Symposium on Signal Processing and Information Technology, 2004.*, Dec 2004, pp. 217–220.

[122] T. Kang and R. A. Iltis, "Iterative decoding, offset and channel estimation for ofdm using the unscented kalman filter," in *2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers*, Nov 2007, pp. 1728–1732.

[123] H. Abdzadeh-Ziabari, W.-P. Zhu, and M. N. S. Swamy, "Improved coarse timing estimation in ofdm systems using high-order statistics," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 5239–5253, Dec 2016.

[124] ——, "Integer frequency offset detection with reduced complexity in ofdm systems," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2017, pp. 1–4.

[125] ——, "Joint maximum likelihood timing, frequency offset, and doubly selective channel estimation for ofdm systems," *Accepted for publication in IEEE Transactions on Vehicular Technology*, Available in IEEE Xplore.

[126] ——, "A marginalized particle filtering-based approach for carrier frequency offset and doubly selective channel estimation in ofdm systems," *to be submitted to 2018 Canadian Conference on Electrical and Computer Engineering (CCECE).*

[127] J. Chen, Y. C. Wu, S. C. Chan, and T. S. Ng, "Joint maximum-likelihood cfo and channel estimation for ofdma uplink using importance sampling," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3462–3470, Nov 2008.

[128] H. Abdzadeh-Ziabari, W.-P. Zhu, and M. N. S. Swamy, "Timing and frequency synchronization and channel estimation for ofdma uplink under high mobility conditions," *submitted to IEEE Transactions on Vehicular Technology*, under review.

[129] ——, "Joint carrier frequency offset and doubly selective channel estimation for mimo-ofdma uplink with kalman and particle filtering," *submitted to IEEE Transactions on Signal Processing*, under review.

[130] M. Grewal and A. Andrews, *Kalman Filtering: Theory and Practice.* Prentice Hall, New Jersey, 2001.

[131] K. J. Kim, M. O. Pun, T. Reid, and R. A. Iltis, "Joint frequency offset and channel estimation for ul-mimo-ofdma systems using the parallel schmidt kalman filters," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, vol. 3, April 2007, pp. III–205–III–208.

[132] K. J. Kim, R. A. Iltis, and H. V. Poor, "Frequency offset and channel estimation in cooperative relay networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3142–3155, Sept 2011.

[133] F. Lehmann, "A gaussian sum approach to blind carrier phase estimation and data detection in turbo coded transmissions," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2619–2632, September 2009.

[134] T. Schon, F. Gustafsson, and P. J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, July 2005.

[135] F. Septier, Y. Delignon, A. Menhaj-Rivenq, and C. Garnier, "Monte carlo methods for channel, phase noise, and frequency offset estimation with unknown noise variances in ofdm systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3613–3626, Aug 2008.

[136] J. H. Kotecha and P. M. Djuric, "Gaussian particle filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2592–2601, Oct 2003.

[137] J. Lim and D. Hong, "Gaussian particle filtering approach for carrier frequency offset estimation in ofdm systems," *IEEE Signal Processing Letters*, vol. 20, no. 4, pp. 367–370, April 2013.

[138] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for ofdm systems with rapid dispersive fading channels," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 902–915, Jul 1998.

[139] J. G. P. M. Salehi, "Digital communications 5e," 2008.

[140] F. Xiong, *Digital modulation techniques*. Artech House, 2006.

[141] V. Erceg, K. Hari, M. Smith, D. S. Baum, K. Sheikh, C. Tappenden, J. Costa, C. Bushue, A. Sarajedini, R. Schwartz *et al.*, "Channel models for fixed wireless applications," 2001.

[142] Y. H. Kim, I. Song, S. Yoon, and S. R. Park, "An efficient frequency offset estimator for ofdm systems and its performance characteristics," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 5, pp. 1307–1312, Sep 2001.

[143] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 591–598, Sep 1974.

[144] W. C. Jakes, *Microwave Mobile Communications*. Wiley-IEEE Press, 1974.

[145] N. Ricklin and J. R. Zeidler, "Data-aided joint estimation of carrier frequency offset and frequency-selective time-varying channel," in *2008 IEEE International Conference on Communications*, May 2008, pp. 589–593.

[146] R. Faragher, "Understanding the basis of the kalman filter via a simple and intuitive derivation [lecture notes]," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 128–132, Sept 2012.

[147] S. M. Kay, "Fundamentals of statistical signal processing, volume i: estimation theory," 1993.

[148] P. M. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 19–38, 2003.

[149] 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 *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 3496–3499.

[150] H. Hijazi, A. Dakhlallah, S. A. Hajj, A. C. A. Ghouwayel, and A. Dhayni, "Joint cfo and time-varying channel estimation by particle filtering in ofdm systems," in *2013 Third International Conference on Communications and Information Technology (ICCIT)*, June 2013, pp. 241–245.

[151] A. Doucet, N. J. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump markov linear systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 613–624, Mar 2001.

[152] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb 2002.

[153] W. C. Jakes and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE Press, 1994.

[154] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2664–2677, Oct 1994.

[155] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the sage algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 434–450, Mar 1999.

[156] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.

[157] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the em algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, Apr 1988.

[158] V. Nguyen-Duy-Nhat, H. Nguyen-Le, C. Tang-Tan, and T. Le-Ngoc, "Sir analysis for ofdm transmission in the presence of cfo, phase noise and doubly selective fading," *IEEE Communications Letters*, vol. 17, no. 9, pp. 1810–1813, September 2013.

[159] H. Godrich, A. M. Haimovich, and R. S. Blum, "Target localization accuracy gain in mimo radar-based systems," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2783–2803, June 2010.

[160] H. Hijazi and L. Ros, "Analytical analysis of bayesian cramer-rao bound for dynamical rayleigh channel complex gains estimation in ofdm system," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1889–1900, May 2009.

[161] K. Yamamura, "Simple algorithms for tracing solution curves," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 40, no. 8, pp. 537–541, Aug 1993.

[162] S. Chakraborty and D. Sen, "Joint estimation of time, frequency offsets and channel gains with icis in ef multi-relay dmimo-ofdm system," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, pp. 1–1, 2016.

[163] X. Li, F. Ng, and T. Han, "Carrier frequency offset mitigation in asynchronous cooperative ofdm transmissions," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 675–685, Feb 2008.

[164] W. Ni, X. Dong, and W. S. Lu, "Near-optimal hybrid processing for massive mimo systems via matrix decomposition," *IEEE Transactions on Signal Processing*, vol. PP, no. 99, pp. 1–1, 2017.

[165] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

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

[167] L. Liu, D. W. Matolak, C. Tao, and Y. Li, "Analysis of an upper bound on the effects of large scale attenuation on uplink transmission performance for massive mimo systems," *IEEE Access*, vol. 5, pp. 4285–4297, 2017.

[168] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for tdd/fdd massive mimo systems with spatial basis expansion model," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3170–3184, April 2017.

[169] W. Zhang, F. Gao, and H. M. Wang, "Frequency synchronization for massive mimo multi-user uplink," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.

[170] C. Zhang and K. Pang, "Synchronization sequence generated by modified park algorithm for nc-ofdm transmission," *IEEE Signal Processing Letters*, vol. 22, no. 4, pp. 385–389, April 2015.

[171] P. Kryszkiewicz and H. Bogucka, "Low complex, narrowband-interference robust synchronization for nc-ofdm cognitive radio," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3644–3654, Sept 2016.

[172] ——, "In-band-interference robust synchronization algorithm for an nc-ofdm system," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 2143–2154, May 2016.

Appendix A

Derivation of the Means of $M_n^{DC}(d)$

In this appendix, the mean of the proposed timing metric $M_n^{DC}(d)$ in (58) at a wrong timing point and that at a correct timing point are calculated. Consider the correlation function (48) first. At a wrong timing instant, we have $E\{|r(\widetilde{d}+k)|^2\} = \sigma_z^2$, and according to the central limit theorem for an M-dependent sequence [18], the real and imaginary parts of $P_n(\widetilde{d})$ as the summation of an M-dependent sequence (for $n \ll \frac{N}{2}$) have an approximately Gaussian distribution with zero mean and variance $L\sigma_z^{2n}/2$. Therefore, $|P_n(\widetilde{d})|$ has a Rayleigh distribution with mean

$$E\{|P_n(\widetilde{d})|\} = \frac{\sqrt{\pi L}}{2}\sigma_z^n. \tag{252}$$

For the normalization function (57), one can easily calculate

$$E\Big\{\big|r(\widetilde{d}+k) - r(\widetilde{d}+k+\frac{N}{2})\big|^2\Big\} = E\Big\{\big|z(k) - z(k+\frac{N}{2})\big|^2\Big\} = 2\sigma_z^2. \tag{253}$$

Taking the expectation of (57) and using (253), we obtain

$$E\{\Delta_n^{DC}(\widetilde{d})\} = L(2\sigma_z^2)^{\frac{n}{2}}. \tag{254}$$

Next, similar to [9], by considering (252) and (254), it can be shown that at a wrong timing point, we have

$$E\{M_n^{DC}(\widetilde{d})\} \simeq \frac{E\{|P_n(\widetilde{d})|\}}{E\{\Delta_n^{DC}(\widetilde{d})\}} = \sqrt{\frac{\pi}{2^{n+2}L}}. \tag{255}$$

At the correct timing point, considering (51) and [3], it is straightforward to show that

$$P_n(\Theta) = \sum_{l=0}^{\frac{N}{2}-1} \sum_{u_{1,k}=0}^{U_1-1} \sum_{u_{2,k}=0}^{U_2-1} \cdots \sum_{u_{\frac{n}{2},k}=0}^{U_{\frac{n}{2}}-1} e^{j\frac{2\pi}{N}(\frac{N}{2})\frac{n}{2}\varepsilon} \times \left|y\big(|l+u_{1,k}|_{\frac{N}{2}}\big)\right|^2 \left|y\big(|l+u_{2,k}|_{\frac{N}{2}}\big)\right|^2 \cdots \left|y\big(|l+u_{\frac{n}{2},k}|_{\frac{N}{2}}\big)\right|^2 + Z_1(l), \tag{256}$$

where $Z_1(l)$ with zero mean is the noisy term including sums of products of different samples of noise and signal. Therefore, the mean of $|P_n(\Theta)|$ can be expressed as

$$E\{|P_n(\Theta)|\} \simeq L\sigma_y^n. \tag{257}$$

Furthermore, by taking the expectation of (57), and using the following equation

$$E\Big\{\big|r(\Theta+k)-r(\Theta+k+\frac{N}{2})\big|^2\Big\} = E\Big\{\big|e^{j\frac{2\pi}{N}k\varepsilon}y(k)+z(k)-e^{j\frac{2\pi}{N}(k+N/2)\varepsilon}y(k)-z(k+\frac{N}{2})\big|^2\Big\} = E\Big\{\big|(1-e^{j\pi\varepsilon})y(k)+z(k)-z(k+\frac{N}{2})\big|^2\Big\} = |1-e^{j\pi\varepsilon}|^2\sigma_y^2+2\sigma_z^2, \tag{258}$$

it immediately follows that

$$E\{\Delta_n^{DC}(\Theta)\} = L\left(|1-e^{j\pi\varepsilon}|^2\sigma_y^2+2\sigma_z^2\right)^{n/2}. \tag{259}$$

Note that in deriving (258), we have used the fact that the multiplication of $z(k)$ by $e^{j\frac{2\pi}{N}k\varepsilon}$ does not change the characteristics of the noise $z(k)$ [10]. Finally, the mean of the timing

metric $M_n^{DC}(d)$ at the correct timing point Θ is derived as

$$E\{M_n^{DC}(\Theta)\} \simeq \frac{E\{|P_n(\Theta)|\}}{E\{\Delta_n^{DC}(\Theta)\}} = \frac{\sigma_y^n}{\left(|1-e^{j\pi\varepsilon}|^2\sigma_y^2 + 2\sigma_z^2\right)^{n/2}} = \left(\frac{SNR}{|1-e^{j\pi\varepsilon}|^2 SNR + 2}\right)^{\frac{n}{2}}. \tag{260}$$

Appendix B

Derivation of the Means of $M_n^{IC}(d)$

We now derive the means of the metric $M_n^{IC}(d)$ at wrong and correct timing instants. By taking the expectation of (65) and utilizing the equation

$$E\left\{\left(\left|r(\widetilde{d}+k)\right|-\left|r(\widetilde{d}+k+\frac{N}{2})\right|\right)^2\right\}=E\left\{\left(\left|z(k)\right|-\left|z(k+\frac{N}{2})\right|\right)^2\right\}=(2-\frac{\pi}{2})\sigma_z^2, \quad (261)$$

the mean of the normalization function $\Delta_n^{IC}(\widetilde{d})$ can be derived as

$$E\{\Delta_n^{IC}(\widetilde{d})\}=L\left((2-\frac{\pi}{2})\sigma_z^2\right)^{\frac{n}{2}}. \quad (262)$$

Now, we treat $\Delta_n^{IC}(d)$ as a factor and consider $|P_n(\widetilde{d})|$ to be a Rayleigh random variable with mean (252) (as derived in Appendix A). Thus, at a wrong timing point, the timing metric has the following mean

$$E\{M_n^{IC}(\widetilde{d})\}=\frac{1}{2(2-\frac{\pi}{2})^{\frac{n}{2}}}\sqrt{\frac{\pi}{L}}. \quad (263)$$

At a correct timing instant, by using the inequality

$$\left(\left|r(d+k)\right|-\left|r(d+k+\frac{N}{2})\right|\right)^2\leq\left|r(d+k)-r(d+k+\frac{N}{2})\right|^2, \quad (264)$$

we have

$$\Delta_n^{IC}(\Theta) \leq \Delta_n^{DC}(\Theta). \tag{265}$$

Taking the expectation of both sides in (265) and using (259), we have

$$E\{\Delta_n^{IC}(\Theta)\} \leq L\left(|1-e^{j\pi\varepsilon}|^2\sigma_y^2 + 2\sigma_z^2\right)^{n/2}. \tag{266}$$

Thus, by using (257) and (266), we can obtain a lower bound of the mean of the timing metric $M_n^{IC}(d)$ at the correct timing point Θ,

$$E\{M_n^{IC}(\Theta)\} \simeq \frac{E\{|P_n(\Theta)|\}}{E\{\Delta_n^{IC}(\Theta)\}} \geq \frac{\sigma_y^n}{\left(|1-e^{j\pi\varepsilon}|^2\sigma_y^2 + 2\sigma_z^2\right)^{n/2}} = \left(\frac{SNR}{|1-e^{j\pi\varepsilon}|^2 SNR + 2}\right)^{\frac{n}{2}}. \tag{267}$$