國 立 交 通 大 學

電信工程研究所

博 士 論 文

寬頻多入多出直接轉換傳收機之
射頻劣化效應估算與補償技術

On the Radio Impairment Estimation and Compensation Techniques for Wideband MIMO Direct-Conversion Transceivers

研 究 生：許宸睿
指導教授：沈文和　博士
王忠炫　博士

中 華 民 國 九 十 九 年 十 一 月

寬頻多入多出直接轉換傳收機之射頻劣化效應估算與補償技術

On the Radio Impairment Estimation and Compensation Techniques for Wideband MIMO Direct-Conversion Transceivers

研 究 生：許宸睿　　Student：Chen-Jui Hsu

指導教授：沈文和 博士　　Advisor：Dr. Wern-Ho Sheen

王忠炫 博士　　Dr. Chung-Hsuan Wang

國立交通大學

電信工程研究所

博 士 論 文

A Dissertation
Submitted to Institute of Communication Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in
Communication Engineering
Hsinchu, Taiwan

2010 年 11 月

寬頻多入多出直接轉換傳收機之射頻劣化效應估算與補償技術

研究生：許宸睿　　指導教授：沈文和 博士
王忠炫 博士

國 立 交 通 大 學
電 信 工 程 學 系

摘要

在無線通訊傳收機射頻架構設計中，直接轉換架構(direct-conversion radio architecture)是一個具有低成本，低功率消耗和小體積的類比前端設計。然而，此架構卻會產生額外的射頻劣化效應，諸如 I-Q 失衡(I-Q Imbalance)與直流偏移(dc offset)等。這些劣化效應加上所有射頻架構都會產生的頻率偏移(frequency offset)，若無適當的補償，對通訊系統將造成嚴重的效能損失。此篇論文旨在針對多入多出(multiple input, multiple output, MIMO)直接轉換傳收機之射頻劣化效應估算與補償技術做深入研究。基本而言，射頻劣化效應估算與補償技術可分為兩類：第一類為估算補償技術，另一類則是自我校正技術。估算補償技術是在通訊傳輸中，在接收機端去除接收訊號的射頻劣化效應之技術;而自我校正技術則是在通訊傳輸前去除本身射頻劣化效應之技術。此兩類技術皆在本論文中做深入研究；所探討的射頻劣化效應包括了不隨頻率變動的 I-Q 失衡(frequency-independent I-Q imbalance)、隨頻率變動的 I-Q 失衡(frequency-dependent I-Q imbalance)、直流偏移與頻率偏移。

本篇論文在估算補償技術研究方面，探討在多入多出通訊系統下的傳送機與接收機之射頻劣化效應估算與消除方法。首先提出了一個二階段干擾消除架構：在第一階段消除接收端產生之 I-Q 失衡、頻率偏移與直流偏移，而在第二階段消除傳送端產生之 I-Q 失衡。此

二階段消除架構能適用於各種多入多出之運作模式，包括了空間多工、時空區塊編碼與傳送端波束成形等技術，以及任意的傳送天線數與接收天線數。此外，此架構也概括了各種通訊應用下的消除架構，諸如應用於無線點對點對等(wireless peer-to-peer)通訊、上行鏈路(uplink)與下行鏈路(downlink)行動通訊。接著我們提出多種參數估算方法，第一種為在最小平方準則下最佳之聯合參數估算方法，分析指出此方法為不偏估計器（unbiased estimator）以及在有興趣範圍的訊號雜訊比(SNR)下其效能可達到 Cramér-Rao 下限(lower bounds)，但也是最高複雜度的。因此，我們設計其它多種低複雜度估測器，包含了特殊角度旋轉週期訓練設計、直流偏移與頻率偏移簡化估測器與藉由週期訓練協助之低複雜度遞迴估算法。電腦模擬顯示低複雜度設計所造成之效能損失幾乎可被忽略，與現今文獻所提之技術做比較，本論文所提之技術能擁有更低錯誤率以及較短之訓練符元長度需求。

在自我校正技術方面，本論文提出一個新的時域方法(time-domain method)，在迴路(feedback loop)不需要專門額外的類比硬體電路下，能夠同時自我校正傳送機與接收機之射頻劣化效應。此時域方法適用於各類的通訊系統並能同時校正不隨頻率變動與隨頻率變動的 I-Q 失衡和直流偏移。此外，我們亦提出最佳訓練數列化之設計方法，並由分析與模擬驗證了此方法之正確性與有效性。

On the Radio Impairment Estimation and Compensation Techniques for Wideband MIMO Direct-Conversion Transceivers

Student: Chen-Jui Hsu

Advisor: Dr. Wern-Ho Sheen

Dr. Chung-Hsuan Wang

Department of Communication Engineering

National Chiao Tung University

Abstract

Direct-conversion radio architecture is a low-cost, low-power, small-size design for the analog front-end of a wireless communication transceiver. Nevertheless, it induces extra radio impairments such as I-Q imbalance and dc offset that, along with frequency offset which is commonly encountered in all radio frequency (RF) architectures, incur severe degradation on communication performance if not accurately compensated. This dissertation investigates radio impairments estimation and compensation techniques for wideband MIMO (multiple input, multiple output) direct-conversion transceivers. Basically, there are two types of techniques: one is estimation/compensation and the other is self-calibration. The estimation/compensation technique is to remove the impairments from the received signal during communication at the receiving side, while self-calibration is a technique to remove the transceiver's own radio impairments before communication commences. Both types of techniques are studied in the dissertation with a complete set of radio impairments taken into consideration, including frequency-independent and dependent I-Q imbalances, dc offset and frequency offset.

For the estimation/compensation technique, this dissertation investigates the estimation and cancellation of the transmitter and receiver radio impairments in the MIMO communication systems. Firstly, a two-stage cancellation architecture is proposed with the receiver frequency-independent and dependent I-Q imbalances, frequency offset, and dc offsets being cancelled in the first stage and the transmitter frequency-independent and dependent I-Q imbalances cancelled in the second stage. The architecture is general to accommodate different forms of MIMO operation including spatial multiplexing, STBC (space-time block coded) and transmit beam forming, with any number of transmit and receive antennas. In addition, it generalizes the cancellation architecture for various types of application configurations such as wireless peer-to-peer communication, downlink and uplink of mobile cellular communications. Secondly, several methods of estimation of radio parameters are proposed. One is the optimum joint estimation of all radio parameters based on least squares criterion. It is shown through analysis that the estimator is unbiased and can achieve the Cramér-Rao lower bound (CRLB) for the signal-to-noise ratios (SNRs) of interest. The others are reduced-complexity methods, including the special phase-rotated periodic training design, simplified frequency and dc offset estimators and low-complexity iterative estimation aided by periodic training. Simulation results show that the proposed methods have negligible performance degradation when using the reduced-complexity designs and outperform the existing ones in error-rate performance and/or the number of training symbols required.

For the self-calibration technique, a new time-domain method is proposed to self-calibrate simultaneously the transmitter and receiver impairments without a dedicated analog circuit in the feedback loop. Thanks to the time-domain approach, the method is applicable to all types of systems and is able to calibrate jointly the frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance, and dc offset. In addition, training sequence design is investigated to optimize the performance of calibration, and analysis and simulations are conducted to confirm the effectiveness of the proposed method.

Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisor, Prof. Wern-Ho Sheen, for giving me the opportunity to join his lab. Thanks for always pointing me the right direction, for helping me analyze and attack the research problems, for patiently reading and revising my draft papers, for providing encouragement and support throughout my Ph.D. studies. I could not have completed my Ph.D. degrees so smoothly without his advice, guidance and valuable comments.

Besides, I would also like to thank Dr. Racy Cheng and my former student colleague Jin-Fa Shyu for providing me some suggestions and possible directions in my research topic. Special thanks to all the student colleagues of Broadband Radio Access Systems Lab (BRASLAB) for their kind help in many aspects.

Finally, I am deeply indebted to my family whose love has accompanied me through this long journey.

Contents

List of Figures

List of Tables

Notations

n_t : number of transmit antennas

n_r : number of receive antennas

$\boldsymbol{j}$: $\sqrt{-1}$

$\otimes$: linear convolution operator

$\delta(t)$: Dirac delta function

$\delta(n)$: Kronecker delta function

$(\cdot)^*$: complex conjugate operator

$(\cdot)^T$: transpose of a matrix or vector

$(\cdot)^H$: complex conjugate transpose of a matrix or vector

$(\cdot)^{-1}$: inverse of a square matrix

$(\cdot)^\dagger$: pseudo-inverse of a matrix or vector

$[\mathbf{A}]_{i,j}$: (i, j)-th entry of matrix $\mathbf{A}$

$tr(\cdot)$: trace of a matrix

$\det(\cdot)$: determinant of a matrix

$\text{diag}\{\mathbf{x}\}$: diagonal matrix with vector $\mathbf{x}$ on the diagonal

$\text{diag}\{\mathbf{X}_1,\ldots,\mathbf{X}_M\}$: block diagonal matrix with the submatrices $\mathbf{X}_1,\ldots,\mathbf{X}_M$ on the diagonal

$E\{\cdot\}$: expectation operator

$\text{Re}\{\cdot\}$: real parts

$\text{Im}\{\cdot\}$: imaginary parts

$\log_{10}(\cdot)$: base-10 logarithm

$\mathbf{1}_N$: all 1 vector with dimension N

$\mathbf{I}_N$: $N \times N$ identity matrix

$\mathbf{0}_N$: $N \times 1$ all zero vector

$\mathbf{O}_{M\times N}$: $M \times N$ all zero matrix

Chapter 1

Introduction

The demands for high-rate wireless communication services and cost-effective devices are getting stronger in recent years. On one hand, OFDM (orthogonal frequency-division multiplexing) equipped with multiple antennas at both transmitter and receiver has been devised as a key technology to enabling high-rate, high spectral-efficiency wireless communications. OFDM is effective in combating inter-symbol interference (ISI) in high-rate transmission [1] while using multiple transmit and receive antennas, also known as a multiple-input multiple-output (MIMO) technology, is capable of providing diversity gain, array gain (power gain), and/or degree-of-freedom gain over the single-input single-output (SISO) systems [2]-[4]. MIMO-OFDM has been adopted in the IEEE 802.11n, IEEE 802.16, and 3GPP Long Term Evolution (LTE) specifications.

On the other hand, the direct-conversion radio architecture has been widely employed in today's wireless devices [5]-[8]; without using intermediary frequency (IF) stages and the image rejection (IR) filter, it is more amenable to monolithic integration and, thus, facilitates a low cost, small form factor design. This architecture, however, is very sensitive to imperfections of the analog circuitry. In particular, mismatch between I and Q branch circuitry, called the in-phase/quadrature (I-Q) imbalance, induces mirror-frequency interference, and imperfection in

A/D and D/A converters and leakage from local oscillators induce dc offset. These radio frequency (RF) impairments, along with frequency offset between transmitter and receiver, significantly degrade communication performance if not accurately compensated [9]-[65]. This is particularly true in the next-generation high-rate systems, where wide bandwidth and high-order modulation are deemed to be employed. Furthermore, the radio impairments make the power amplifier (PA) linearization circuit at the transmitter more difficult to work satisfactorily [50], [52], [56], and that may reduce the PA's efficiency.

In this study, we are concerned with digital compensations of radio impairments in the wideband MIMO systems with direct-conversion transceiver.

1.1 Estimation and Compensation Techniques of Radio Impairments

Removal and/or compensation of the radio impairments in the direct-conversion radio architecture have been an area of extensive research. Generally speaking, two types of technique have been proposed [9]-[61]: one is estimation/compensation and the other is self-calibration. The estimation/compensation technique is to remove the impairments from the received signal during communication at the receiving side [9]-[49], whereas self-calibration is a technique to remove the transceiver's own radio impairments before communication [50]-[61]. Both techniques find their applications in real systems [9]-[61], although self-calibration has the advantage to facilitate the design of the PA's (power amplifier's) linearization circuit. This dissertation will focus on both types of digital technique.

1.1.1 Estimation/Compensation Techniques

Quite a lot of works in the literature have been devoted to the radio impairments estimation and compensation for the direct-conversion architecture [9]-[49]. In [9]-[34], different methods were

developed to estimate and compensate the receiver radio impairments in SISO (single-input, single output) and MIMO (multiple input, multiple output) receivers. In [35]-[49], the cascaded transmitter and receiver radio impairments were investigated, with [35]-[40] focusing on the SISO systems and [41]-[49] on the MIMO systems. In the application configuration such as the downlink of a mobile cellular system, only the receiver radio impairments needed to be considered because the transmitter impairments can be neglected due to the high-precision implementation of the analog front-end in the base station. For the application configuration of wireless peer-to-peer communication such as the wireless mesh networks, however, the cascaded effect of the transmitter and receiver radio impairments should be considered because it is very likely that a low-cost and, hence, a less precise analog front-end is implemented at both transmitter and receiver. In the following, estimation/compensation techniques for different system configurations are discussed in more details.

1.1.2 Estimation/Compensation Techniques for Receiver Radio Impairments

Various digital techniques have been developed for estimation and compensation of receiver radio impairments, which can be generally categorized into two types. One is the data-aided (DA) methods using known training sequences [16]-[34], and the other is the non-data-aided (NDA) methods exploiting some statistical properties of the received signal [9]-[15].

- **Non-Data-Aided Techniques**

In [9]-[15], adaptive or non-adaptive (block-based) statistical signal processing-based (blind) techniques were proposed for compensation of the receiver radio impairments. Blind source separation (BSS) techniques were introduced in [9]-[11] to cancel the self-image interference due to frequency-independent [9] and frequency-dependent [10][11] I-Q imbalance, respectively, based on the assumption that the desired signal and the image interference are statistically independent.

Such independence may be valid in the low-IF scheme but does not hold true in direct-conversion receivers. On the other hand, circularity-based compensation techniques were proposed in [12]-[15] based on the assumption that the desired signal is a complex random process with a circular-symmetric distribution. In [12]-[14], frequency-independent I-Q imbalance and/or dc offset were considered, while frequency-dependent I-Q imbalance was of interest in [15]. Convergence to the desired solution is not always guaranteed using blind approaches and no training data is needed at the expense of slow convergence rate.

- **Data-Aided Techniques**

In [16]-[20], dc offset was investigated under frequency offset, among which a low-complexity joint estimation was proposed in [20] based on periodicity of a training sequence. In [21][22], I-Q imbalances were investigated for OFDM (orthogonal frequency division multiplexing) systems in the absence of frequency offset and dc offset. In [21], I-Q imbalance was compensated by using a simple two-tap adaptive frequency-domain equalizer, while in [22], adaptive and non-adaptive post-FFT schemes were proposed for both frequency-dependent and independent I-Q imbalances at the receiver. In [24]-[29], I-Q imbalance was investigated jointly with frequency offset, with [24]-[26] focusing on frequency-independent I-Q imbalance and [27]-[29] on frequency-dependent I-Q imbalance. In [27]-[29], an FIR (finite impulse response) filter was proposed for compensating frequency-dependent I-Q imbalance and an asymmetric phase compensator for frequency-independent one. In particular, the periodic structure of a training sequence was exploited in [29] to develop a low-complexity estimation of I-Q imbalance and frequency offset. In [30], a joint least squares estimation of I-Q imbalance, dc offset and channel was proposed, and latter extended to [31], which developed a joint estimation of frequency offset, channel, (frequency-independent) I-Q imbalance and dc offset in the time domain. Under the assumption of white Gaussian noise at the output of sampler, the ML (maximum likelihood) criterion was used in [31] to derive the solution. In [32], the pilot-based estimation and compensation of

I-Q imbalance were investigated for OFDM systems under timing offset and frequency offset.

In [33], I-Q imbalance was investigated jointly with symbol detection for MIMO-OFDM systems, with no consideration of frequency offset and dc offset. Using an extended channel model that incorporates I-Q imbalance, symbol detection was performed on the extended channel without explicit estimation/compensation of I-Q imbalance. Non-adaptive (e.g., least-squares) or adaptive (e.g., recursive least squares and least mean squares) filtering can be used for estimating the extended channel prior to symbol detection. Likewise, MIMO detectors such as ML, ZF (zero-forcing), MMSE (minimum mean-squared error), etc. [4] can be employed for symbol detection. Note that using the extended channel increases system dimension, and hence the detection complexity, as compared to the receiver in which RF impairments are estimated and compensated before the symbol detection.

1.1.3 Estimation/Compensation Techniques for Cascaded Transmitter and Receiver Radio Impairments

Basically, two types of compensation for cascaded transmitter and receiver radio impairments have been investigated in the design of a digital receiver: one is cancellation based and the other is joint-detection based. In the cancellation type of techniques, radio impairments are estimated and cancelled explicitly from the received signal before a detection, which is designed with no consideration of radio impairments [35][36]. In the joint-detection type of technique, on the other hand, the detection and radio impairments compensation are performed jointly with no explicit estimation and cancellation of the radio impairments [37]-[49]. Due to the effect of mirror-frequency interference and inter-carrier interference, the dimension of signal detection in the joint-detection type of receiver is at least twice of that of the cancellation type, and that increases the detection complexity very significantly. This is especially true for the MIMO-OFDM (multiple-input, multiple-output, orthogonal frequency division multiplexing) systems, where MIMO detection has to be done for each sub-carrier.

For the MIMO-OFDM systems, which is a key enabling technology for high-rate wireless systems, the transmitter and receiver radio impairments were studied in [33],[34],[41]-[43], [45]-[49]. In particular, I-Q imbalances were investigated exclusively for the spatial-multiplexing systems in [33],[41]-[43] and for the space-time block coded (STBC) systems in [33][45][46]. In [47], the transmitter and receiver I-Q imbalances were investigated jointly with frequency offset for spatial-multiplexing and space-frequency block coded (SFBC) systems in a high mobility environment. In [48][49], the authors proposed linear per-tone equalization (PTEQ) to the spatial-multiplexing and STBC systems in the presence of transmitter and receiver I-Q imbalances and frequency offset. The PTEQ method, nevertheless, is only applicable to a linear MIMO detection and suffers from a slow convergence. In all these works, the joint-detection type of receiver design was investigated exclusively. In [34], a cancellation type of MIMO-OFDM receiver was proposed, but only receiver frequency-independent I-Q imbalance and frequency offset were considered.

1.1.4 Calibration Techniques

Many interesting works have been devoted to self-calibration of the radio impairments in the direct-conversion architecture [50]-[61]. In [50]-[53], adaptive methods were proposed to calibrate transmitter frequency-independent I-Q imbalance and dc offset by using an envelope-detector (ED) in the feedback loop. Transmitter frequency-dependent I-Q imbalance was calibrated for the continuous frequency-shift-keying (CFSK) systems in [54], with no consideration on other impairments. Frequency-dependent I-Q imbalance is particularly problematic in a wideband system, where the I-branch and Q-branch analog filters are difficult to be kept perfectly matched over the entire band. The works in [55]-[58] discussed calibration of transmitter frequency-independent and dependent I-Q imbalances by using a digital low-IF architecture in the feedback loop, as shown in Figure 1.1. And in [59], transmitter frequency-independent I-Q imbalance was calibrated by using same reference signal at both transmit and receive paths.

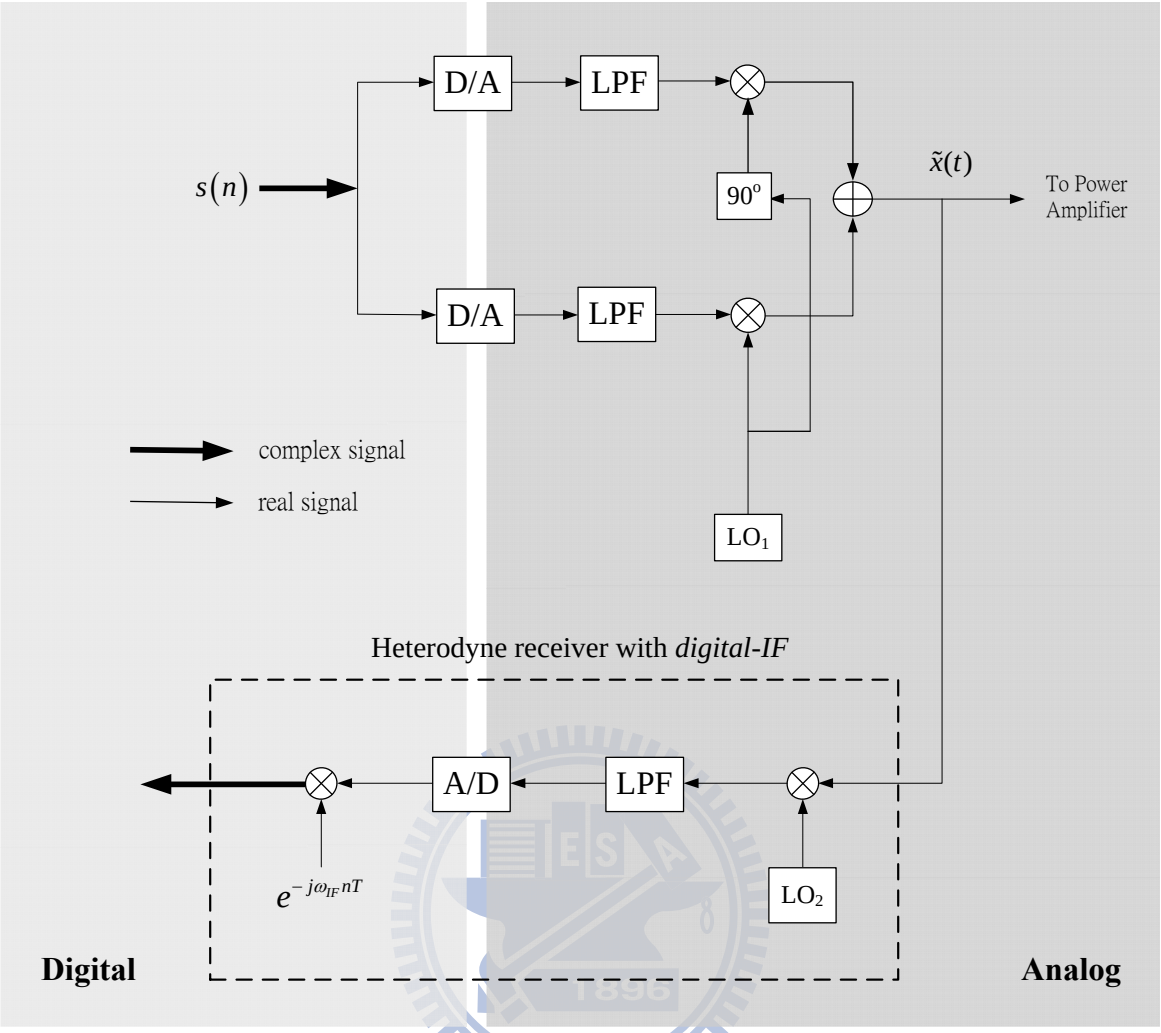


Figure 1.1: Calibration feedback loop using heterodyne receiver with *digital-IF*

So far, most self-calibration techniques in the literature have been focused on the transmitter radio impairments, either by employing an ED [50]-[53] or a digital low-IF radio architecture [55]-[58] in the feedback loop so that the receiver impairments can be safely neglected. However, using ED or low-IF radio architecture in the feedback loop increases the hardware complexity. Besides, the calibration of the receiver impairments is important in its own right. For example, if one's own receiver has been calibrated before communication, only the transmitter impairments (of the transmitting device), rather than the cascaded effect of the transmitter and receiver impairments, need to be estimated and compensated for at the receiver, and that reduces the receiver complexity. Very recently, the issue of joint self-calibration of transmitter and receiver radio im-

pairments was investigated in [60] and [61]. In [60], a two-feedback method was proposed, where the phase of the receive oscillator is shifted by exact 90 degrees in the second feedback, aiming to separate transmitter and receiver I-Q imbalances. Unfortunately, it is very difficult for analog circuitry to have an exact 90 degrees phase rotation in real systems. In [61], a new method was proposed for OFDM (orthogonal frequency-division multiplexing) type of systems with no dedicated analog circuit in the feedback loop. The method, however, can only calibrate the frequency-independent I-Q imbalance and dc offset.

1.2 Dissertation Outline and Contributions

This dissertation focuses on study and development of both types of estimation and compensation techniques of radio impairments for different MIMO communication application configurations, where the estimation/compensation techniques are investigated in Chapter 3 and Chapter 4 and the calibration technique in Chapter 5. In particular, Chapter 3 focuses on estimation/compensation technique of receiver radio impairments for the application configuration such as the downlink of a mobile cellular system, and then Chapter 4 extends to estimation/compensation technique of cascaded transmitter and receiver radio impairments for the application configurations such as the wireless peer-to-peer communication and the uplink of a mobile cellular system. The rest of this dissertation is organized as follows.

In Chapter 2, first, we describe the direct-conversion transceiver model including frequency-independent and frequency-dependent I-Q imbalances and dc offset. Then, three basic types of MIMO signal model that incorporates the transmitter and/or receiver radio impairments are introduced, which are related with different application configurations.

In Chapter 3, we propose to do joint estimation of frequency, receiver dc offset, receiver I-Q imbalance, and channel in MIMO receivers for improving its performance. Firstly, a receiver architecture that facilitates joint estimation and compensation of frequency, receiver dc offset, re-

ceiver I-Q imbalance and channel is proposed; both frequency-dependent and frequency-independent I-Q imbalances are included. Secondly, the LS criterion is applied to obtain the joint estimators, with a special training-sequence design to reduce complexity. Simplified estimators on frequency and dc offset are also proposed with almost no loss in performance. Lastly, the LS estimators are shown through analysis to be un-biased and approach to CRLB (Cramér-Rao lower bound) for signal-to-noise ratios (SNRs) of interest. The results of this chapter have been published in [62] and [63].

In Chapter 4, the cancellation technique of cascaded transmitter and receiver radio impairments is investigated for the MIMO-OFDM systems with direct conversion radio architecture. A two-stage generalized cancellation architecture is proposed, where in the first stage, the receiver radio impairments such as frequency-independent and dependent I-Q imbalances and dc offset are cancelled along with frequency offset and transmitter dc offset in the time domain, and in the second stage, the transmitter frequency-independent and dependent I-Q imbalances are cancelled in the frequency domain. The two-stage cancellation architecture generalizes the cancellation architectures for transmitter and/or receiver radio impairments. Moreover, the technique is unique in that it is effective in different forms of MIMO operations including spatial multiplexing, STBC (space-time block coded) and transmit beam forming, with any number of transmit and receive antennas. In addition, two methods of radio parameters estimation are proposed. The first is the optimum joint least-squares estimation of channel and radio impairments which can be seen as an extension of Chapter 3. The second is a low-complexity iterative estimation that exploits the periodic structure of a training sequence. The proposed methods are simulated and compared with existing methods. Simulation results confirm the superiority of the proposed methods. The results of this chapter have been presented partly in [64] and [65] and submitted as the journal paper [66].

In Chapter 5, a new self-calibration method is proposed which enables to self-calibrate the own transmitter and receiver radio impairments simultaneously, with no dedicated analog circuit

in the feedback loop. A frequency offset between transmitter and receiver is introduced purposely so as to separate the transmitter impairments from those of the receiver during calibration. Based on a time-domain approach, the new method is applicable to all types of communication systems and is able to calibrate jointly the frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance, and dc offset. In addition, optimal training sequences are devised to best the calibration performance. The calibration performance is analyzed that agrees very well with the simulations. Simulation and analytical results confirm the effectiveness of the proposed method. The results of this chapter will be presented partly in [67] and submitted as the journal paper [68].

Finally, Chapter 6 concludes the dissertation and discusses some possible extensions and topics for future research.

Chapter 2

System Models

In this chapter, direct-conversion transceiver model including frequency-independent and frequency-dependent I-Q imbalances and dc offset are described in Section 2.1. Then, three basic types of MIMO signal model that incorporates the transmitter and/or receiver radio impairments are given in Section 2.2, which are corresponding to different application configurations.

2.1 Direct-Conversion Transceiver Model

Figure 2.1 shows a mathematical model of a direct-conversion transmitter with I-Q imbalance and dc offset effects. Define $s(t) = s^I(t) + \boldsymbol{j}s^Q(t)$ be the baseband signal appearing at the input of transmitter. $f_{0,i} = f_{0,i}^I + \boldsymbol{j}f_{0,i}^Q$ characterizes the dc offset, which is due mainly to the imperfection of D/A converters. The transmit analog filter effect of I and Q branch is modeled by $c_T^I(t)$ and $c_T^Q(t)$ with unit power, if $c_T^I(t) \neq c_T^Q(t)$, it is said that there is a frequency-dependent I-Q imbalance. Taking into account the transmitter frequency-independent I-Q imbalance, the complex sinusoidal signal coming into the mixer is

$$C_T(t) = \cos(2\pi f_c t) + \boldsymbol{j}\alpha_T \sin(2\pi f_c t + \theta_T) = \gamma_T e^{j2\pi f_c t} + \varphi_T e^{-j2\pi f_c t},$$

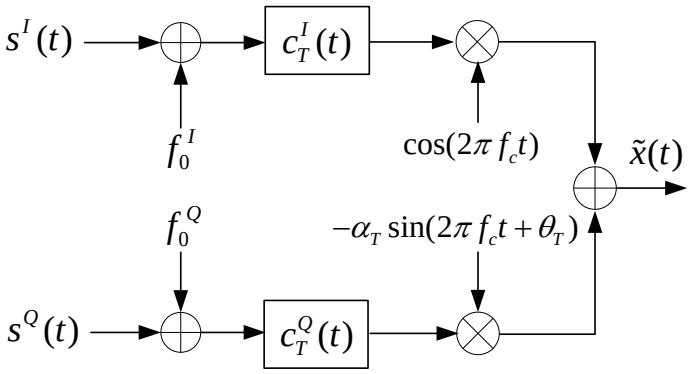


Figure 2.1: Mathematical model of a direct-conversion transmitter with I-Q imbalance and dc offset.

where α_T and θ_T are the gain and phase imbalances due to the imperfect analog circuitry of the mixer, respectively, $\gamma_T = 1/2\left(1+\alpha_T e^{j\theta_T}\right)$ and $\varphi_T = 1/2\left(1-\alpha_T e^{-j\theta_T}\right)$ characterize the corresponding transmitter frequency-independent I-Q imbalance effect, and f_c is the center frequency. Ideally, the complex sinusoidal $C_T(t)$ should contain only the positive frequency, i.e. $C_T(t) = e^{j2\pi f_c t}$. However, due to the gain and phase mismatches α_T and θ_T, there is a negative frequency component $e^{-j2\pi f_c t}$ with a magnitude of $|\varphi_T|$, which, along with frequency-dependent I-Q imbalance, cause interfering images in the transmitted signal. After up-conversion, the pass-band transmitted signal including overall transmitter radio impairments effect is $\tilde{x}_i(t) = \mathrm{Re}\left\{x_i(t)e^{j2\pi f_c t}\right\}$ with its equivalent base-band signal :

$$x(t) = c_{T+}(t) \otimes s(t) + c_{T-}(t) \otimes s^*(t) + f_1, \tag{2.1}$$

where

$$c_{T\pm}(t) = \frac{1}{2}\left(c_T^I(t) \pm \alpha_T e^{j\theta_T} c_T^Q(t)\right), \tag{2.2}$$

and

$$f_1 = f_0 \otimes c_{T+}(t) + f_0^* \otimes c_{T-}(t), \tag{2.3}$$

is the equivalent transmit dc offset. In (2.1), $s(t)$ can be viewed as being transmitted by two

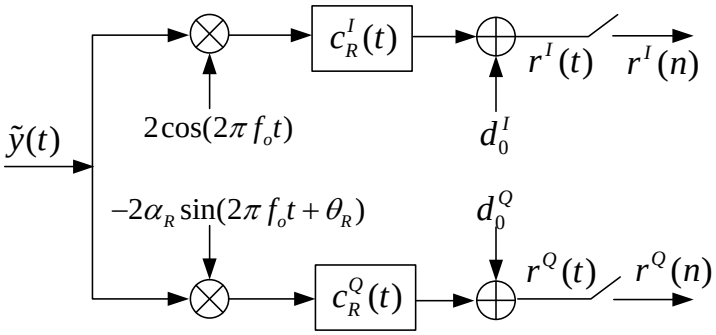


Figure 2.2: Mathematical model of a direct-conversion receiver with I-Q imbalance and dc offset.

channels with a corruption from dc offset; one is the desired channel with impulse response $c_{T+}(t)$, and the other is the mirror-frequency channel with impulse response $c_{T-}(t)$. That is, I-Q imbalances incur mirror-frequency interference in the transmitted signal, and dc offset, on the other hand, imposes a more strict specification on A/D converters at the receiver and may incur in-band interference in the presence of frequency offset. Note that with $c_T^I(t) = c_T^Q(t) = c_T(t)$, (2.2) degenerates to the case of no frequency-dependent I-Q imbalance with $c_{T+}(t) = \gamma_T c_T(t)$ and $c_{T-}(t) = \varphi_T c_T(t)$. In addition, with no I-Q imbalances and dc offset, i.e., $c_T^I(t) = c_T^Q(t) = c_T(t)$, $\alpha_T = 1$ and $\theta_T = f_0 = 0$, $x(t) = s(t) \otimes c_T(t)$, as one might expect.

Similar to the direct-conversion transmitter, the mathematical model of a direct-conversion receiver with I-Q imbalance and dc offset is shown in Figure 2.2. Taking into account the frequency-independent I-Q imbalance at the RF front-end, the complex sinusoidal signal coming into the mixer is

$$C_R(t) = 2\cos(2\pi f_c t) - \boldsymbol{j} 2\alpha_R \sin(2\pi f_c t + \theta_R) = \gamma_R e^{-\boldsymbol{j} 2\pi f_c t} + \varphi_R e^{\boldsymbol{j} 2\pi f_c t},$$

where α_R and θ_R are the gain and phase imbalances, respectively, $\gamma_R = \left(1 + \alpha_R e^{-j\theta_R}\right)$, and $\varphi_R = \left(1 - \alpha_R e^{j\theta_R}\right)$. After down conversion, $c_R^I(t)$ and $c_R^Q(t)$ are the base-band filters used to remove out-of-band noise and high-frequency components. Again, it is said to have a frequency-dependent I-Q imbalance if $c_R^I(t) \neq c_R^Q(t)$. Frequency-dependent I-Q imbalance is mostly encountered in a wide-band RF receiver because it is generally difficult to maintain base-band

filters to have the same response over a wide frequency range. $d_{0,j} = d_{0,j}^{I} + \boldsymbol{j} d_{0,j}^{Q}$ is the dc offset which is due to the imperfection of A/D converters and the self-mixing at the receiver's mixer. After sampling, the end-to-end equivalent discrete system can be modeled as (up to a constant for the case of no aliasing)

$$r(n) = r^{I}(n) + \boldsymbol{j} r^{Q}(n) = c_{R+}(n) \otimes y(n) + c_{R-}(n) \otimes y^{*}(n) + d_{0}, \tag{2.4}$$

where

$$c_{R\pm}(n) = \frac{1}{2}\left(c_{R}^{I}(n) \pm \alpha_{R} e^{\mp j\theta_{R}} c_{R}^{Q}(n)\right). \tag{2.5}$$

Again, (2.4) says that the receiver I-Q imbalances induces mirror-frequency interference in the received signal. With $c_{R}^{I}(n) = c_{R}^{Q}(n) = c_{R}(n)$, (2.5) degenerates to the case of no frequency-dependent I-Q imbalance with $c_{R+}(n) = 1/2 \cdot \gamma_{R} c_{R}(n)$, and $c_{R-}(n) = 1/2 \cdot \varphi_{R} c_{R}(n)$. In addition, if $\alpha_{R} = 1$, $\theta_{R} = d_{0} = 0$ and $c_{R}^{I}(n) = c_{R}^{Q}(n) = c_{R}(n)$, we have the no I-Q imbalance and dc offset case; that is $r(n) = c_{R}(n) \otimes y(n)$.

2.2 System Models

The radio impairments may appear in transmitter and/or receiver, which depends on different system configurations. In the following, we address three kinds of system models corresponding to different application configurations, and the associated estimation/compensation techniques will be investigated in Chapter 3 and Chapter 4.

2.2.1 MIMO Systems under Receiver Radio Impairments

In this section, we address the MIMO system model that incorporates the receiver radio impairments, which is corresponding to the application configuration such as the downlink of the mobile cellular system.

The MIMO signal model with a direct-conversion receiver is depicted in Figure 2.3. Consider a block transmission with a prefix to avoid inter-block interference. The pass-band transmit signal

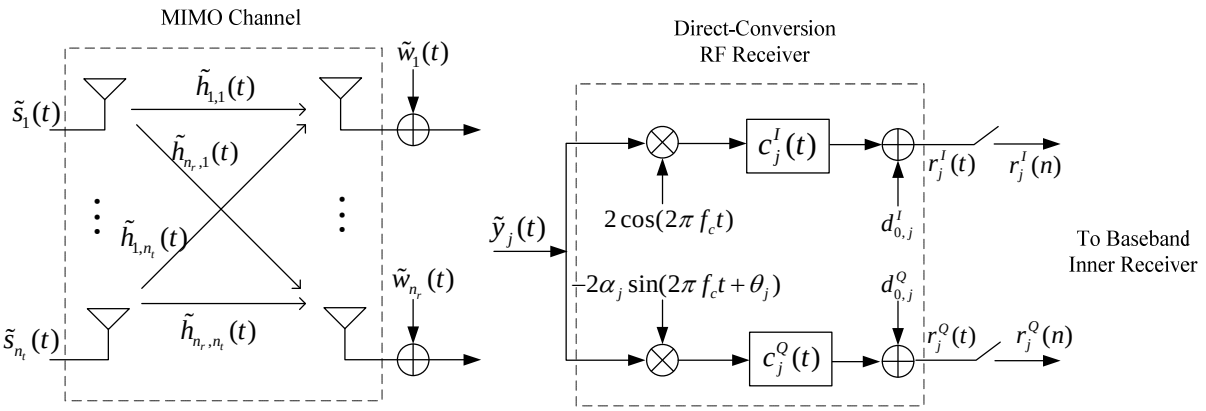


Figure 2.3: MIMO signal model with direct-conversion RF receiver.

from *i*-th transmit antenna is $\tilde{s}_i(t) = \mathrm{Re}\left\{s_i(t)e^{j2\pi f_c t}\right\}$, with the base-band signal

$$s_i(t) = \sum_k \sum_{n=-N_g}^{N-1} s_{i,k}(n)\, g_T\left(t-\left(k\left(N+N_g\right)+n\right)T_s\right), \tag{2.6}$$

where f_c is the carrier frequency, $s_{i,k}(n)$ is the transmitted *k*–th block symbols with the transmit power $\sigma_s^2 = \frac{1}{N}\sum_{i=1}^{n_t}\sum_{n=0}^{N-1}\left|s_{i,k}(n)\right|^2$, N_g is the length of prefix, N is the length of useful data, $g_T(t)$ is the transmit filter with unit power, and T_s is the symbol time.

Let $\tilde{h}_{j,i}(t) = \mathrm{Re}\left\{h_{j,i}(t)e^{j2\pi f_c t}\right\}$ denote the channel response from *i*-th transmit to *j*-th receive antenna, and $h_{j,i}(t)$ be its base-band equivalent. The band-pass signal received from *j*-th antenna is

$$\tilde{y}_j(t) = \mathrm{Re}\left\{y_j(t)e^{j2\pi(f_c+\Delta f)t}\right\} + \tilde{w}_j(t), \tag{2.7}$$

where

$$y_j(t) = \sum_{i=1}^{n_t} s_i(t) \otimes h_{j,i}(t), \tag{2.8}$$

$\tilde{w}_j(t) = \mathrm{Re}\left\{w_{0,j}(t)e^{j2\pi f_c t}\right\}$ is the pass-band additive white Gaussian noise and $w_{0,j}(t)$ is its base-band equivalent, Δf is the frequency offset which is same for all receive branches.

At the $j'th$ receiver side, α_j and θ_j are the frequency-independent gain and phase imbalances, respectively, $c_j^I(t)+\boldsymbol{j}c_j^Q(t)$ is the (unit-energy) base-band filter, and $d_{0,j}=d_{0,j}^I+\boldsymbol{j}d_{0,j}^Q$ is the dc offset. Note that the radio parameters of I-Q imbalances and dc offsets are different from one branch to another. From (2.4), the end-to-end equivalent discrete system can be modeled as (up to a constant for the case of no aliasing)

$$r_j(n)=c_{+,j}(n)\otimes\left[y_j(n)e^{\boldsymbol{j}2\pi\nu n}+w_{0,j}(n)\right]+c_{-,j}(n)\otimes\left[y_j(n)e^{\boldsymbol{j}2\pi\nu n}+w_{0,j}(n)\right]^*+d_{0,j}, \quad (2.9)$$

where $c_{\pm,j}(n)=1/2\cdot\left[c_j^I(n)\pm c_j^Q(n)\alpha_j e^{\mp j\theta_j}\right]$, $\nu=\Delta f T_S$ is the normalized frequency offset, $y_j(n)=\sum_{i=1}^{n_t}s_i(n)\otimes h_{j,i}(n)$, and $w_{0,j}(n)$ is a zero mean additive white Gaussian noise with $\sigma_w^2\doteq \mathrm{E}\left[\left|w_{0,j}(n)\right|^2\right]$, $j=1\cdots n_r$. In Chapter 3, we will investigate the estimation/compensation technique for receiver radio impairments.

2.2.2 MIMO-OFDM Systems under Cascaded Transmitter and Receiver Radio Impairments

In this section, we address the MIMO-OFDM system model that incorporates the cascaded transmitter and receiver radio impairments, which is corresponding to the application configuration of wireless peer-to-peer communication such as the wireless mesh networks.

Figure 2.4 depicts a MIMO-OFDM system that employs a direct-conversion radio transceiver as the analog front-end subsystem. The base-band data signal for transmit antenna i is

$$s_i(n)=s_i^I(n)+\boldsymbol{j}s_i^Q(n)=\sum_k\sum_{m=-N_g}^{N-1}s_{i,k}(m)\delta\left(n-k(N_g+N)-m\right), \quad (2.10)$$

where

$$s_{i,k}(m)=\frac{1}{N}\sum_{l=0}^{N-1}S_{i,k}(l)e^{\frac{j2\pi ml}{N}} \quad (2.11)$$

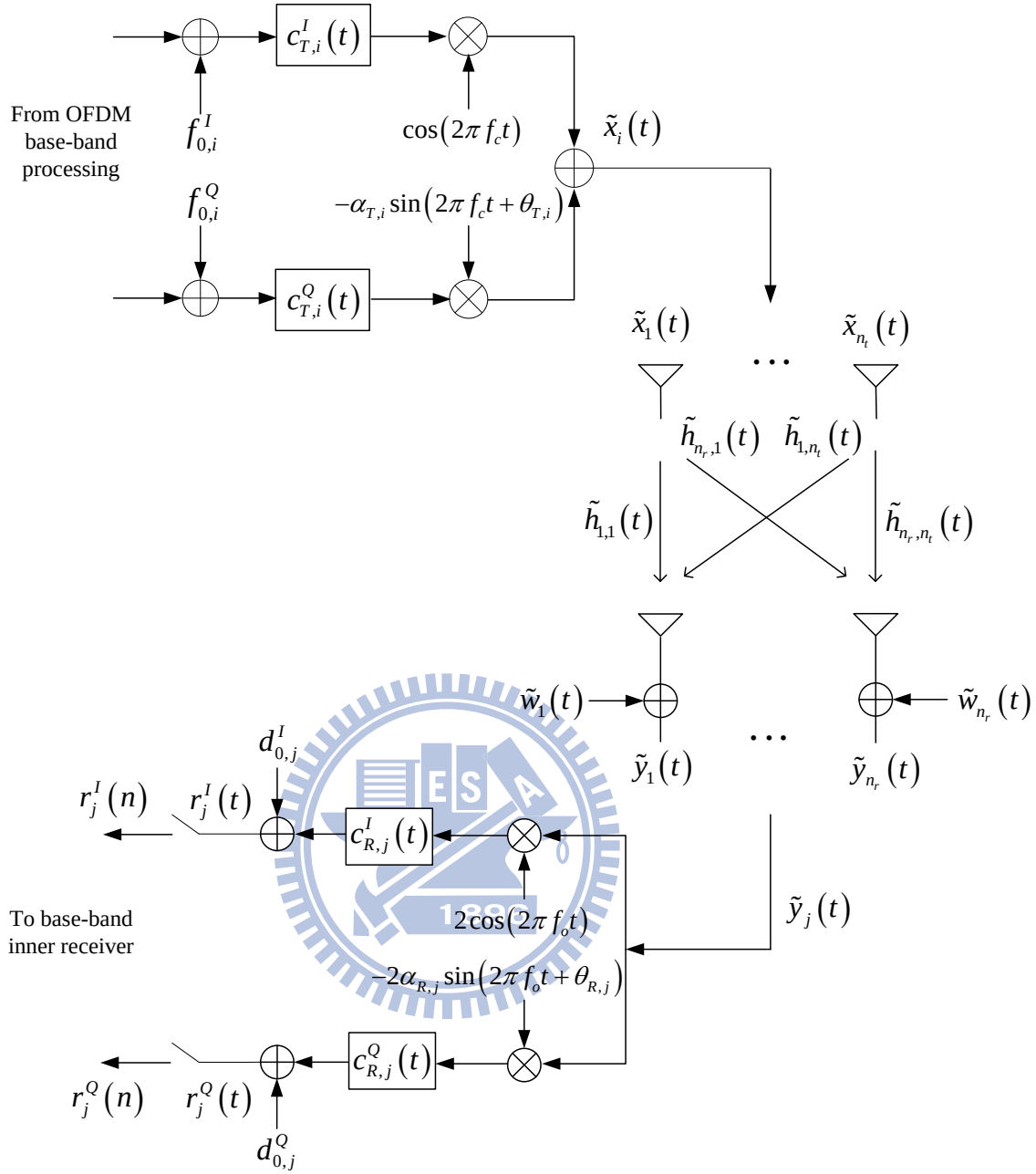


Figure 2.4: MIMO signal model with direct-conversion radio transceiver under the effects of I-Q imbalances, dc offsets, and frequency offset.

is the inverse discrete Fourier transform (IDFT) of the transmitted data $\{S_{i,k}(l)\}_{l=0}^{N-1}$ in OFDM symbol k with the transmit power $\sigma_s^2 = \frac{1}{N}\sum_{i=1}^{n_t}\sum_{m=0}^{N-1}|s_{i,k}(m)|^2$, N_g is the length of cyclic prefix

and N is the FFT size. The cyclic prefix is assumed larger than the maximum delay spread of the overall channel consisting of the transmit filter, radio channel and receiver filter, and therefore there is no inter-OFDM symbols and inter-carrier interference.

At the $i'th$ transmitter, $f_{0,i} = f_{0,i}^{I} + \boldsymbol{j} f_{0,i}^{Q}$ characterizes the dc offset, and $c_{T,i}^{I}(t) + \boldsymbol{j} c_{T,i}^{Q}(t)$ is the (unit-energy) base-band transmit filter. The frequency-independent I-Q imbalance is characterized by the parameters $\alpha_{T,i}$ and $\theta_{T,i}$, which are the gain and phase imbalances respectively due to the imperfect analog circuitry of the mixer. f_c is the carrier frequency which is same for all transmit antennas, $\tilde{h}_{j,i}(t) = \mathrm{Re}\left\{ h_{j,i}(t) e^{j2\pi f_c t} \right\}$ is the channel response from transmit antenna i to receive antenna j, and $h_{j,i}(t)$ is its equivalent base-band. $\tilde{w}_j(t) = \mathrm{Re}\left\{ w_{0,j}(t) e^{j2\pi f_o t} \right\}$ is the pass-band additive white Gaussian noise, and $w_{0,j}(t)$ is its base-band equivalent.

At the $j'th$ receiver, $\alpha_{R,j}$ and $\theta_{R,j}$ are the frequency-independent gain and phase imbalances, respectively, $c_{R,j}^{I}(t) + \boldsymbol{j} c_{R,j}^{Q}(t)$ is the (unit-energy) base-band filter, and $d_{0,j} = d_{0,j}^{I} + \boldsymbol{j} d_{0,j}^{Q}$ is the dc offset. $f_0 = f_c - \Delta f$ is the local oscillator frequency of the receiver, where Δf is the frequency offset between transmitter and receiver which is same for all receiver branches. From (2.1) and (2.4), the received discrete signal can be expressed as

$$\begin{aligned} r_j(n) &= r_j^{I}(n) + \boldsymbol{j} r_j^{Q}(n) \\ &= c_{R+,j}(n) \otimes \left(y_j(n) e^{j2\pi\nu n} + w_{0,j}(n) \right) + c_{R-,j}(n) \otimes \left(y_j(n) e^{j2\pi\nu n} + w_{0,j}(n) \right)^{*} + d_{0,j} \ , \end{aligned} \tag{2.12}$$

where

$$y_j(n) = \sum_{i=1}^{n_t} \left(s_i(n) \otimes c_{T+,i}(n) + s_i^{*}(n) \otimes c_{T-,i}(n) + f_{1,i} \right) \otimes h_{j,i}(n), \tag{2.13}$$

$$f_{1,i} = f_{0,i} \otimes c_{T+,i}(n) + f_{0,i}^{*} \otimes c_{T-,i}(n) \ , \tag{2.14}$$

$$c_{T\pm,i}(n) = 1/2 \cdot \left(c_{T,i}^{I}(n) \pm \alpha_{T,i} e^{j\theta_{T,i}} c_{T,i}^{Q}(n) \right), \tag{2.15}$$

$$c_{R\pm,j}(n) = 1/2 \cdot \left(c_{R,j}^{I}(n) \pm \alpha_{R,j} e^{\mp j\theta_{R,j}} c_{R,j}^{Q}(n) \right), \tag{2.16}$$

$\nu = \Delta f T_s$ is the normalized frequency offset, and $w_{0,j}(n)$ is a zero mean additive white Gaus-

sian noise with $\sigma_w^2 \doteq \mathrm{E}\left[\left|w_{0,j}(n)\right|^2\right]$, $j = 1 \cdots n_r$.. We have used the notation, $x(n) \doteq x(t)\big|_{t=nT_s}$ for a signal $x(t)$, where T_s is the sampling period.

2.2.3 MIMO-OFDM Systems under Transmitter Radio Impairments

For the application configuration such as uplink of mobile cellular system, the receiver radio impairments are negligible except for frequency offset due to the high-precision implementation of the analog front-end in the base station while the transmitter radio impairments needed to be considered due to the mobile station with a less precise analog front-end. As a special case in Figure 2.4, there is no receiver I-Q imbalances and receiver dc offset, i.e. $\alpha_{R,j} = 1$, $\theta_{R,j} = d_{0,j} = 0$ and $c_{R,j}^{I}(n) = c_{R,j}^{Q}(n) = c_{R,j}(n)$, $j = 1, \ldots, n_r$. Hence, the received discrete signal model in (2.12) for a MIMO-OFDM system under transmitter radio impairments and frequency offset degenerates to

$$\begin{aligned} r_j(n) &= c_{R,j}(n) \otimes \left(y_j(n) e^{j2\pi\nu n} + w_{0,j}(n) \right) \\ &= c_{R,j}(n) \otimes \left[e^{j2\pi\nu n} \sum_{i=1}^{n_t} \left(s_i(n) \otimes c_{T+,i}(n) + s_i^*(n) \otimes c_{T-,i}(n) + f_{1,i} \right) \otimes h_{j,i}(n) + w_{0,j}(n) \right]. \end{aligned} \tag{2.17}$$

In Chapter 4, we will investigate the estimation/compensation technique for cascaded transmitter and receiver radio impairments.

Chapter 3

Estimation/Compensation Technique for Receiver Radio Impairments

In this chapter, the estimation/compensation technique of receiver radio impairments is investigated for applications such as downlink MIMO communications. A receiver architecture that facilitates joint estimation and compensation of radio parameters is proposed in Section 3.1. The least squares (LS) criterion is then applied to obtain the joint estimators in Section 3.2, with a special phase-rotated periodic training-sequence design to reduce complexity. Simplified estimators on frequency and dc offset are also proposed with almost no loss in performance. The computational complexities and performances of proposed estimators are analyzed in Section 3.3 and Section 3.4, respectively. Simulation results are given in Section 3.5. Finally, Section 3.6 summarizes this chapter.

3.1 Receiver Architecture

We start from the received signal model (2.9) with receiver radio impairments effect in subsection 2.2.2. Motivated by (2.9), one way to process the received signal is to cancel out firstly the self-image interference due to I-Q imbalance. By introducing the filter $\rho_j(n)$, we have

$$\begin{aligned} r_j(n)-\rho_j(n)\otimes r_j^*(n) &= \left(c_{+,j}(n)-\rho_j(n)\otimes c_{-,j}^*(n)\right)\otimes\left(y_j(n)e^{j2\pi\nu n}+w_{0,j}(n)\right) \\ &\quad+\underbrace{\left(c_{-,j}(n)-\rho_j(n)\otimes c_{+,j}^*(n)\right)}_{=0}\otimes\left(y_j(n)e^{j2\pi\nu n}+w_{0,j}(n)\right)^* \\ &\quad+\left(d_{0,j}-\rho_j(n)\otimes d_{0,j}^*\right). \end{aligned} \tag{3.1}$$

To completely cancel out the self-image interference, $\rho_j(n)=\left(c_{+,j}^*(n)\right)^{-1}\otimes c_{-,j}(n)$, where $\left(c_{+,j}^*(n)\right)^{-1}$ is the inverse filter of $c_{+,j}^*(n)$. (For the case of no frequency-dependent I-Q imbalance, $\rho_j(n)=\varphi_j/\gamma_j^*$ as is given in [31].) Thus, (3.1) becomes

$$\begin{aligned} &r_j(n)-\rho_j(n)\otimes r_j^*(n) \\ &=\underbrace{\left[c_{+,j}(n)-\rho_j(n)\otimes c_{-,j}^*(n)\right]}_{=c_j(n)}\otimes\left(y_j(n)e^{j2\pi\nu n}+w_{0,j}(n)\right)+\left(d_{0,j}-\rho_j(n)\otimes d_{0,j}^*\right) \\ &=e^{j2\pi\nu n}\left(\sum_{i=1}^{n_t}s_i(n)\otimes g_{j,i}(n)\right)+d_j+w_j(n), \end{aligned} \tag{3.2}$$

where $c_j(n)=c_{+,j}(n)-\rho_j(n)\otimes c_{-,j}^*(n)$, $g_{j,i}(n)\doteq h_{j,i}(n)\otimes\left(c_j(n)e^{-j2\pi\nu n}\right)$, $d_j=d_{0,j}-\rho_j(n)\otimes d_{0,j}^*$, and $w_j(n)=c_j(n)\otimes w_{0,j}(n)$. $g_{j,i}(n)$ is the overall impulse response from *i*-th transmit to *j*-th receive antenna after canceling out the self-image interference, d_j is the dc offset, and $w_j(n)$ is an additive Gaussian noise but generally not white. In (3.2), $\rho_j(n)$, $g_{j,i}(n)$, ν, and d_j are the deterministic, unknown parameters to be estimated. In what follows, $\rho_j(n)$ and $g_{j,i}(n)$ will be approximated as FIR (finite impulse response) filters with large enough taps, although they are generally IIR (infinite impulse response) ones, as can be seen in (3.1).

Following (3.2), we propose the receiver architecture as in Figure 3.1; after cancelling the

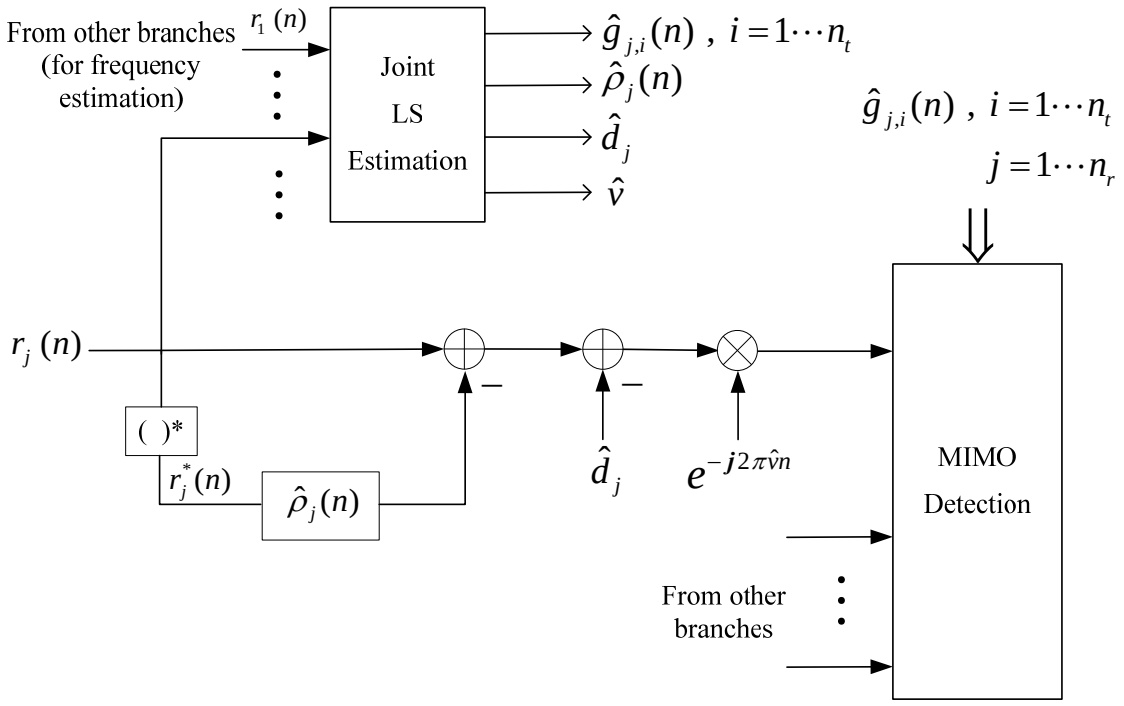


Figure 3.1: MIMO receiver with joint LS estimation/compensation of radio parameters.

self-image interference, dc offset is compensated next, and then compensation for frequency offset follows. The parameters $\rho_j(n)$, $g_{j,i}(n)$, v, and d_j will be estimated jointly in the least squares sense, as to be discussed in the next section, and the MIMO detection is done with the MMSE detector, based on the estimated channel responses $\{\hat{g}_{j,i}(n)\}$ along with the compensated received signals from all branches. Some other types of MIMO detectors can be used as well [70].

3.2 Joint Least Squares Estimation

In this section, the set of radio parameters $\rho_j(n)$, $g_{j,i}(n)$, v, and d_j are estimated jointly in the sense of least squares. In order to do that, $\rho_j(n)$ and $g_{j,i}(n)$ are approximated by the FIR filters $\boldsymbol{\rho}_j = \left[\rho_j(0), \rho_j(1), \ldots, \rho_j\left(L_{\rho_j}-1\right)\right]^T$ and $\mathbf{g}_{j,i} = \left[g_{j,i}(0), g_{j,i}, (1), \ldots, g_{j,i}\left(L_{\mathbf{g}_{j,i}}-1\right)\right]^T$ with

the length L_{ρ_j} and $L_{\mathbf{g}_{j,i}}$, respectively. In the following, $L_{\rho_j} = L_\rho$, $\forall\, j$, and $L_{\mathbf{g}_{j,i}} = L_{\mathbf{g}}$, $\forall\, i, j$. In addition, L_ρ and $L_{\mathbf{g}}$ are assumed to be large enough (to contain 99% of the energy) in this and next sections, and thus approximation error is negligible. The impact of L_ρ and $L_{\mathbf{g}}$ will be investigated by computer simulations in Section 3.4. For a data-aided estimation as considered here, $s_{i,k}(n)$ is known perfectly to the receiver. A total of $P \geq 1$ training blocks will be assumed in this paper, starting from the zero-th block $(k = 0)$.

3.2.1 Least Squares Estimators

Consider the case of $L_{\mathbf{g}} \leq N_g$; thus there will be no inter-block interference. Let $\mathbf{r}_j(k) = \left[r_{j,k}(0), r_{j,k}(1), \ldots, r_{j,k}(N-1)\right]^T$ be the useful part of the k-th received block, where $r_{j,k}(n) \doteq r_j\left(k\left(N_g + N\right) + n\right)$, $\mathbf{R}_j(k)$ be the $N \times L_\rho$ received signal matrix with $\left[\mathbf{R}_j(k)\right]_{m,l} = r_{j,k}(m-l)$, $0 \leq m \leq N-1$, $0 \leq l \leq L_\rho - 1$, and $\mathbf{S}_i(k)$ be the $N \times L_{\mathbf{g}}$ signal matrix with $\left[\mathbf{S}_i(k)\right]_{m,l} = s_{i,k}(m-l)$, $0 \leq m \leq N-1$, $0 \leq l \leq L_{\mathbf{g}} - 1$. From (3.2), the useful part of the j-th received signal can be written as the vector form

$$\mathbf{r}_j(k) - \mathbf{R}_j^*(k)\boldsymbol{\rho}_j = \boldsymbol{\Gamma}_k(\nu)\left(\sum_{i=1}^{n_t}\mathbf{S}_i(k)\mathbf{g}_{j,i}\right) + d_j\mathbf{1}_N + \mathbf{w}_j(k),\ k = 0, \ldots, P-1, \tag{3.3}$$

where $\boldsymbol{\Gamma}_k(\nu) = e^{j2\pi k\left(N_g+N\right)\nu} \cdot \mathrm{diag}\left\{1, e^{j2\pi\nu}, \cdots, e^{j2\pi\nu(N-1)}\right\}$ is a diagonal matrix, and $\mathbf{w}_j(k) = \left[w_{j,k}(0), w_{j,k}(1), \ldots, w_{j,k}(N-1)\right]^T$ with $w_{j,k}(n) \doteq w_j\left(k\left(N_g + N\right) + n\right)$. Furthermore, define

$$\mathbf{r}_j = \left[\mathbf{r}_j^T(0), \mathbf{r}_j^T(1), \ldots, \mathbf{r}_j^T(P-1)\right]^T,$$

$$\mathbf{R}_j \doteq \left[\mathbf{R}_j^T(0), \mathbf{R}_j^T(1), \ldots, \mathbf{R}_j^T(P-1)\right]^T,$$

$$\mathbf{S}(k) \doteq \left[\mathbf{S}_1(k), \mathbf{S}_2(k), \ldots, \mathbf{S}_{n_t}(k)\right],$$

$$\mathbf{g}_j = \left[\mathbf{g}_{j,1}^T, \mathbf{g}_{j,2}^T, \ldots, \mathbf{g}_{j,n_t}^T\right]^T,$$

and

$$\mathbf{w}_j = \left[\mathbf{w}_j^T(0), \mathbf{w}_j^T(1), \ldots, \mathbf{w}_j^T(P-1)\right]^T.$$

The total useful received signal for training is

$$\mathbf{r}_j - \mathbf{R}_j^* \boldsymbol{\rho}_j = \boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j + d_j \mathbf{1}_{PN} + \mathbf{w}_j, \tag{3.4}$$

where $\mathbf{S} \doteq \left[\mathbf{S}^T(0), \mathbf{S}^T(1), \ldots, \mathbf{S}^T(P-1)\right]^T$, and $\boldsymbol{\Gamma}(\nu) = \text{diag}\left\{\boldsymbol{\Gamma}_0(\nu), \boldsymbol{\Gamma}_1(\nu), \ldots, \boldsymbol{\Gamma}_{P-1}(\nu)\right\}$. From (3.4), the joint least-squares estimates of all parameters are obtained by minimizing the cost function

$$\Lambda\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_1, \cdots, \tilde{\boldsymbol{\rho}}_{n_r}, \tilde{d}_1, \cdots, \tilde{d}_{n_r}, \tilde{\mathbf{g}}_1, \cdots, \tilde{\mathbf{g}}_{n_r}\right) = \sum_{j=1}^{n_r} \Lambda_j\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_j, \tilde{d}_j, \tilde{\mathbf{g}}_j\right) \tag{3.5}$$

with

$$\Lambda_j\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_j, \tilde{d}_j, \tilde{\mathbf{g}}_j\right) = \left\|\mathbf{r}_j - \mathbf{R}_j^* \tilde{\boldsymbol{\rho}}_j - \tilde{d}_j \mathbf{1}_{PN} - \boldsymbol{\Gamma}(\tilde{\nu})\mathbf{S}\tilde{\mathbf{g}}_j\right\|^2. \tag{3.6}$$

Recall that each branch of receiver has its own I-Q imbalance, dc offset and channel response, but the frequency offset is the same for all branches. Therefore, given a fixed trial frequency offset $\tilde{\nu}$, $\boldsymbol{\rho}_j$, d_j, and $\mathbf{g}_j$ can be estimated by simply minimizing the cost function $\Lambda_j\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_j, \tilde{d}_j, \tilde{\mathbf{g}}_j\right)$. In other words, $\boldsymbol{\rho}_j$, d_j, and $\mathbf{g}_j$ can be estimated independently form one branch to another. On the other hand, equation (3.5) can be used to estimate jointly the frequency offset to increase performance by exploiting the diversity and power gain inherent in MIMO systems.

The LS solution of (3.5) or (3.6) can be obtained successively as follows [31]. Firstly, under a fixed trial $\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_j, \tilde{d}_j\right)$, the LS-estimate of channel $\hat{\mathbf{g}}_j$ is given by [71]

$$\begin{aligned}\hat{\mathbf{g}}_j\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_j, \tilde{d}_j\right) &= \left(\boldsymbol{\Gamma}(\tilde{\nu})\mathbf{S}\right)^{\dagger}\left(\mathbf{r}_j - \mathbf{R}_j^* \tilde{\boldsymbol{\rho}}_j - \tilde{d}_j \mathbf{1}_{PN}\right) \\ &= \left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H\boldsymbol{\Gamma}^H(\tilde{\nu})\left(\mathbf{r}_j - \mathbf{R}_j^* \tilde{\boldsymbol{\rho}}_j - \tilde{d}_j \mathbf{1}_{PN}\right),\end{aligned} \tag{3.7}$$

Secondly, substituting $\hat{\mathbf{g}}_j$ in (3.6),

$$\Lambda_j\left(\tilde{\nu},\tilde{\boldsymbol{\rho}}_j,\tilde{d}_j\right)=\left\|\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\left(\mathbf{r}_j-\mathbf{R}_j^{*}\tilde{\boldsymbol{\rho}}_j-\tilde{d}_j\mathbf{1}_{PN}\right)\right\|^2, \tag{3.8}$$

where

$$\mathbf{C}\left(\tilde{\nu}\right)=\boldsymbol{\Gamma}\left(\tilde{\nu}\right)\mathbf{B}\boldsymbol{\Gamma}^H\left(\tilde{\nu}\right) \tag{3.9}$$

with

$$\mathbf{B}=\mathbf{S}\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H. \tag{3.10}$$

Note that $\mathbf{C}\left(\tilde{\nu}\right)$ and $\mathbf{B}$ are projection matrices. By minimizing (3.8) with respect to $\tilde{d}_j$, given a fixed $\left(\tilde{\nu},\tilde{\boldsymbol{\rho}}_j\right)$, the LS-estimate for d_j is

$$\hat{d}_j=\mathbf{f}^H\left(\tilde{\nu}\right)\left(\mathbf{r}_j-\mathbf{R}_j^{*}\tilde{\boldsymbol{\rho}}_j\right). \tag{3.11}$$

where

$$\mathbf{f}^H\left(\tilde{\nu}\right)\doteq\left(\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\mathbf{1}_{PN}\right)^{\dagger}=\begin{cases}\dfrac{\mathbf{1}_{PN}^H\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)}{\left\|\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\mathbf{1}_{PN}\right\|^2}, & \text{if } \left\|\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\mathbf{1}_{PN}\right\|^2\neq 0\\ \mathbf{0}_{PN}^T, & \text{otherwise}\end{cases}, \tag{3.12}$$

Thirdly, substituting $\hat{d}_j$ in (3.8), one gets

$$\Lambda_j\left(\tilde{\nu},\tilde{\boldsymbol{\rho}}_j\right)=\left\|\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H\left(\tilde{\nu}\right)\right)\left(\mathbf{r}_j-\mathbf{R}^{*}\tilde{\boldsymbol{\rho}}_j\right)\right\|^2, \tag{3.13}$$

and the LS-estimate $\hat{\boldsymbol{\rho}}_j$ is

$$\hat{\boldsymbol{\rho}}_j=\mathbf{P}_j\left(\tilde{\nu}\right)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H\left(\tilde{\nu}\right)\right)\mathbf{r}_j, \tag{3.14}$$

where $\mathbf{P}_j\left(\tilde{\nu}\right)=\left[\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H\left(\tilde{\nu}\right)\right)\mathbf{R}_j^{*}\right]^{\dagger}$. After a simple manipulation, $\hat{\boldsymbol{\rho}}_j\left(\tilde{\nu}\right)$ can be further simplified as

$$\hat{\boldsymbol{\rho}}_j\left(\tilde{\nu}\right)=\left[\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{R}_j^{*}\right]^{-1}\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j, \tag{3.15}$$

where

$$\mathbf{Q}(\tilde{\nu}) = \mathbf{Q}^H(\tilde{\nu}) = \left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\tilde{\nu})\right)^H \left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\tilde{\nu})\right). \tag{3.16}$$

Finally, substituting $\hat{\boldsymbol{\rho}}_j$ in (3.13) and using (3.5), we have

$$\Lambda(\tilde{\nu}) = \sum_{j=1}^{n_r} \Lambda_j(\tilde{\nu}) \tag{3.17}$$

with

$$\Lambda_j(\tilde{\nu}) = \left\| \left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\tilde{\nu})\right)\left(\mathbf{r}_j - \mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j(\tilde{\nu})\right)\right\|^2, \tag{3.18}$$

and the LS-estimate $\hat{\nu}$ is given by

$$\hat{\nu} = \arg\min_{\tilde{\nu}} \left\{\Lambda(\tilde{\nu})\right\}. \tag{3.19}$$

Generally, no close-form is available for $\hat{\nu}$; an exhaustive search has to be performed for the solution. As discussed in [69], the exhaustive search can be implemented in two-step procedure. The first step is a coarse search which calculates $\Lambda(\tilde{\nu})$ over a grid of $\tilde{\nu}$-values, say $\{\tilde{\nu}_n\}$, and determines the location $\tilde{\nu}_M$ of the maximum. Then, the second step is a fine search where the $\Lambda(\tilde{\nu}_n)$-values are interpolated and the local maximum nearest to $\tilde{\nu}_M$ is found. Moreover, in the following, we show that the first step of grid search can be implemented with FFT (fast Fourier transform) techniques.

Using (3.15), $\Lambda_j(\tilde{\nu})$ in (3.18) can be further simplified as

$$\begin{aligned}
\Lambda_j(\tilde{\nu}) &= \left\| \left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\tilde{\nu})\right)\left(\mathbf{r}_j - \mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j(\tilde{\nu})\right)\right\|^2 \\
&= \left(\mathbf{r}_j - \mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j(\tilde{\nu})\right)^H \mathbf{Q}(\tilde{\nu})\left(\mathbf{r}_j - \mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j(\tilde{\nu})\right) \\
&= \mathbf{r}_j^H\mathbf{Q}(\tilde{\nu})\mathbf{r}_j - \left(\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{r}_j\right)^H \left[\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{r}_j
\end{aligned}. \tag{3.20}$$

Furthermore, from (3.16) and (3.12), for $\mathbf{f}(\tilde{\nu}) \neq \mathbf{0}_{PN}^T$,

$$\begin{aligned}
\mathbf{Q}(\tilde{\nu}) &= \left(\mathbf{I}_{PN} - \mathbf{f}(\tilde{\nu})\mathbf{1}_{PN}^H\right)\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\tilde{\nu})\right) \\
&= \left(\mathbf{I}_{PN} - \frac{\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\mathbf{1}_{PN}^H}{\left\|\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right\|^2}\right)\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\left(\mathbf{I}_{PN} - \frac{\mathbf{1}_{PN}\mathbf{1}_{PN}^H\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)}{\left\|\left(\mathbf{I} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right\|^2}\right)
\end{aligned}$$

$$= \left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right) - \frac{\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN} \times \left[\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right]^H}{\left\|\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right\|^2}. \tag{3.21}$$

Coincidentally, the relevant elements of $\Lambda_j(\tilde{\nu})$ in (3.20) to be calculated are in the form of $\mathbf{v}^H\mathbf{Q}(\tilde{\nu})\mathbf{u}$, where $\mathbf{v} = [v_0, v_1, \ldots, v_{PN-1}]^T$ and $\mathbf{u} = [u_0, u_1, \ldots, u_{PN-1}]^T$ are $PN \times 1$ vectors. In addition, from (3.21),

$$\mathbf{v}^H\mathbf{Q}(\tilde{\nu})\mathbf{u} = \mathbf{v}^H\mathbf{u} - \mathbf{v}^H\mathbf{C}(\tilde{\nu})\mathbf{u} - \frac{\mathbf{v}^H\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN} \times \left[\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right]^H \mathbf{u}}{\left\|\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right\|^2}. \tag{3.22}$$

Define $\mathbf{v}^T = \left[\mathbf{v}_1^T, \mathbf{v}_2^T, \ldots, \mathbf{v}_P^T\right]$, $\mathbf{u}^T = \left[\mathbf{u}_1^T, \mathbf{u}_2^T, \ldots, \mathbf{u}_P^T\right]$, and

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1P} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2P} \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_{P1} & \mathbf{B}_{P2} & \cdots & \mathbf{B}_{PP} \end{bmatrix},$$

where $\mathbf{v}_r$ and $\mathbf{u}_s$ are $N \times 1$ vectors, and $\mathbf{B}_{rs}$ is an $N \times N$ matrix with $1 \le r, s \le P$. It can be shown that

$$\mathbf{v}^H\mathbf{C}(\tilde{\nu})\mathbf{u} = \sum_{r=1}^{P}\sum_{s=1}^{P} e^{-j2\pi(s-r)(N+N_g)\tilde{\nu}} \left[-\eta_{r,s}(0) + \sum_{m=0}^{N-1}\left(\eta_{r,s}(m)e^{-j2\pi m\tilde{\nu}} + \varsigma_{r,s}(m)e^{j2\pi m\tilde{\nu}}\right)\right], \tag{3.23}$$

where

$$\eta_{r,s}(m) = \sum_{k=m}^{N-1}\left[\mathbf{B}_{rs}\right]_{k-m,k} v_{r,k-m}^{*} u_{s,k}, \tag{3.24}$$

$$\varsigma_{r,s}(m) = \sum_{k=m}^{N-1}\left[\mathbf{B}_{rs}\right]_{k,k-m} v_{r,k}^{*} u_{s,k-m}, \tag{3.25}$$

and $x_{r,k} = \left[\mathbf{x}_r\right]_k$, $\mathbf{x} = \mathbf{v}, \mathbf{u}$. Thus, the term $\mathbf{v}^H\mathbf{C}(\tilde{\nu})\mathbf{u}$ in the form of (3.23) can be calculated efficiently by using FFT (fast Fourier transform) in searching of $\hat{\nu}$. To this end, the $2N$ correlations $\{\eta_{r,s}(m)\}$ and $\{\varsigma_{r,s}(m)\}$ are first calculated from (3.24) and (3.25). Next, the sequences

$$\tilde{\eta}_{r,s}(m)=\begin{cases}\eta_{r,s}(m), & 0\le m<N-1\\ 0, & N\le m<KN-1\end{cases} \tag{3.26}$$

and

$$\tilde{\varsigma}_{r,s}(m)=\begin{cases}\varsigma_{r,s}(m), & 0\le m<N-1\\ 0, & N\le m<KN-1\end{cases} \tag{3.27}$$

are formed, K being a design parameter (pruning factor). Finally, the FFT in (3.23) is equivalently implemented as

$$\sum_{m=0}^{N-1}\eta_{r,s}(m)e^{-j2\pi m\tilde{\nu}}=\sum_{m=0}^{KN-1}\tilde{\eta}_{r,s}(m)e^{\frac{-j2\pi mn}{KN}} \tag{3.28}$$

$$\sum_{m=0}^{N-1}\varsigma_{r,s}(m)e^{j2\pi m\tilde{\nu}}=\left(\sum_{m=0}^{KN-1}\tilde{\varsigma}^{*}_{r,s}(m)e^{\frac{-j2\pi mn}{KN}}\right)^{*} \tag{3.29}$$

for KN frequency grids :

$$\tilde{\nu}_n=\frac{n}{KN},\quad -\frac{KN}{2}\le n<\frac{KN}{2}. \tag{3.30}$$

In summary, the joint LS estimates are summarized as follows :

$$\hat{\nu}=\arg\min_{\tilde{\nu}}\left\{\Lambda(\tilde{\nu})=\sum_{j=1}^{n_r}\Lambda_j(\tilde{\nu})\right\}, \tag{3.31}$$

$$\Lambda_j(\tilde{\nu})=\mathbf{r}_j^H\mathbf{Q}(\tilde{\nu})\mathbf{r}_j-\left(\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{r}_j\right)^H\left[\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{r}_j, \tag{3.32}$$

$$\hat{\boldsymbol{\rho}}_j=\left[\mathbf{R}_j^T\mathbf{Q}(\hat{\nu})\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}(\hat{\nu})\mathbf{r}_j, \tag{3.33}$$

$$\hat{d}_j=\frac{\mathbf{1}_{PN}^H\left(\mathbf{I}_{PN}-\mathbf{C}(\hat{\nu})\right)}{\left\|\left(\mathbf{I}_{PN}-\mathbf{C}(\hat{\nu})\right)\mathbf{1}_{PN}\right\|^2}\cdot\left(\mathbf{r}_j-\mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j\right), \tag{3.34}$$

and

$$\hat{\mathbf{g}}_j=\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H\boldsymbol{\Gamma}^H(\hat{\nu})\left(\mathbf{r}_j-\mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j-\hat{d}_j\mathbf{1}_{PN}\right). \tag{3.35}$$

3.2.2 Low-Complexity Implementation

From (3.32)-(3.34), it is observed that the calculation of the projection matrices $\mathbf{B}$ and $\mathbf{C}(\tilde{\nu})$ plays a key role in determining the complexity of the estimators of $\hat{\nu}$, $\hat{\boldsymbol{\rho}}_j$, and $\hat{d}_j$. Motivated by the design for SISO systems in [69], we design the following phase-rotated periodic training format :

$$\mathbf{S}=\underbrace{\left[\underbrace{\left[e^{j\phi_0}\mathbf{A}\right]^H\cdots\left[e^{j\phi_0}\mathbf{A}\right]^H}_{Q}\underbrace{\left[e^{j\phi_1}\mathbf{A}\right]^H\cdots\left[e^{j\phi_1}\mathbf{A}\right]^H}_{Q}\cdots\underbrace{\left[e^{j\phi_{P-1}}\mathbf{A}\right]^H\cdots\left[e^{j\phi_{P-1}}\mathbf{A}\right]^H}_{Q}\right]^H}_{Q\cdot P}, \tag{3.36}$$

where $\mathbf{A}$ is an $n_tL_{\mathbf{g}}\times n_tL_{\mathbf{g}}$ full-rank matrix, $N=Q\cdot n_tL_{\mathbf{g}}$, $Q\geq 1$, and $\{\phi_k\}_{k=0}^{P-1}$ are parameters to optimize the estimation performance. Example $\{\phi_k\}$ for $Q=P=2$ will be given in Section 3.5; nevertheless, the issue of optimum design of them will not be pursued any further in this chapter. In this way, the projection matrix $\mathbf{B}$ becomes

$$\mathbf{B}=\frac{1}{Q\cdot P}\begin{bmatrix}\mathbf{F} & e^{j(\phi_0-\phi_1)}\mathbf{F} & \cdots & e^{j(\phi_0-\phi_{P-1})}\mathbf{F}\\ e^{j(\phi_1-\phi_0)}\mathbf{F} & \mathbf{F} & \cdots & e^{j(\phi_1-\phi_{P-1})}\mathbf{F}\\ \vdots & \vdots & & \vdots\\ e^{j(\phi_{P-1}-\phi_0)}\mathbf{F} & e^{j(\phi_{P-1}-\phi_1)}\mathbf{F} & \cdots & \mathbf{F}\end{bmatrix}, \tag{3.37}$$

with

$$\mathbf{F}=\begin{bmatrix}\mathbf{I}_{n_tL_{\mathbf{g}}} & \mathbf{I}_{n_tL_{\mathbf{g}}} & \cdots & \mathbf{I}_{n_tL_{\mathbf{g}}}\\ \mathbf{I}_{n_tL_{\mathbf{g}}} & \mathbf{I}_{n_tL_{\mathbf{g}}} & \cdots & \mathbf{I}_{n_tL_{\mathbf{g}}}\\ \vdots & \vdots & & \vdots\\ \mathbf{I}_{n_tL_{\mathbf{g}}} & \mathbf{I}_{n_tL_{\mathbf{g}}} & \cdots & \mathbf{I}_{n_tL_{\mathbf{g}}}\end{bmatrix}_{N\times N}, \tag{3.38}$$

which contains Q^2 matrices of $\mathbf{I}_{n_tL_{\mathbf{g}}}$. Recall that $\mathbf{A}$ is a full-rank square matrix, and therefore its projection matrix $\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H$ is equal to $\mathbf{I}_{n_tL_{\mathbf{g}}}$. Since $\mathbf{B}$ is a sparse matrix now, $\mathbf{C}(\tilde{\nu})=\boldsymbol{\Gamma}(\tilde{\nu})\mathbf{B}\boldsymbol{\Gamma}^H(\tilde{\nu})$ can be calculated much efficiently. In particular, using (3.37), $\mathbf{v}^H\mathbf{C}(\tilde{\nu})\mathbf{u}$ in (3.23) becomes

$$\mathbf{v}^H\mathbf{C}(\tilde{\nu})\mathbf{u}=\sum_{r=1}^{P}\sum_{s=1}^{P}e^{-j2\pi(s-r)(N+N_g)\tilde{\nu}}\left[-\eta_{r,s}(0)+\sum_{m=0}^{Q-1}\left(\eta_{r,s}(m)e^{-j2\pi mn_tL_g\tilde{\nu}}+\varsigma_{r,s}(m)e^{j2\pi mn_tL_g\tilde{\nu}}\right)\right], \tag{3.39}$$

with

$$\begin{aligned}\eta_{r,s}(m)&=\sum_{k=mn_tL_g}^{N-1}\left[\mathbf{B}_{rs}\right]_{k-mn_tL_g,k}v^*_{r,k-mn_tL_g}u_{s,k}\\&=\sum_{k=mn_tL_g}^{N-1}\frac{e^{j(\phi_{r-1}-\phi_{s-1})}}{Q\cdot P}\left[\mathbf{F}\right]_{k-mn_tL_g,k}v^*_{r,k-mn_tL_g}u_{s,k}\\&=\frac{e^{j(\phi_{r-1}-\phi_{s-1})}}{Q\cdot P}\sum_{k=mn_tL_g}^{N-1}v^*_{r,k-mn_tL_g}u_{s,k},\end{aligned} \tag{3.40}$$

and

$$\varsigma_{r,s}(m)=\frac{e^{j(\phi_{r-1}-\phi_{s-1})}}{Q\cdot P}\sum_{k=mn_tL_g}^{N-1}v^*_{r,k}u_{s,k-mn_tL_g},\quad 0\le m\le Q-1. \tag{3.41}$$

Hence, the complexity of CFO estimation is reduced by about n_tL_g times as to be discussed for more detail in next Section. However, the frequency range that can be estimated is also reduced by the same factor with this design. For SISO systems ($n_t=n_r=1$) with $P=1$, the training format in (3.36) degenerates to the one in [69]. Furthermore, for the undesirable case of $\phi_0=\phi_1=\cdots=\phi_{P-1}$, $\left\|\left(\mathbf{I}_{PN}-\mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right\|^2\Big|_{\tilde{\nu}=0}=0$ in (3.12). In other words, with this design of training sequence, it is not able to estimate dc offset when frequency offset is zero because $d_j\mathbf{1}_{PN}$ is now located in the space spanned by the column vectors of $\mathbf{S}$, and its effect is included in the estimate $\hat{\mathbf{g}}_j$. During the real data transmission, however, the receiver needs the estimate $\hat{d}_j$ for dc offset compensation.

In addition, from the simulation results in Section 3.5, frequency-dependent I-Q imbalance has little effect on the estimation of ν and d_j, and can be neglected with almost no loss in performance. With this observation, we propose the simplified estimators for frequency and dc offset

by replacing $\mathbf{R}_j$ as $\mathbf{r}_j$ in (3.32) and (3.34) as follows, i.e. $L_\rho = 1$ case ;

$$\hat{\nu}_s = \arg\min_{\tilde{\nu}} \left\{ \Lambda_s(\tilde{\nu}) = \sum_{j=1}^{n_r} \Lambda_{j,s}(\tilde{\nu}) \right\} \tag{3.42}$$

with

$$\begin{aligned} \Lambda_{j,s}(\tilde{\nu}) &= \mathbf{r}_j^H \mathbf{Q}(\tilde{\nu}) \mathbf{r}_j - \left(\mathbf{r}_j^T \mathbf{Q}(\tilde{\nu}) \mathbf{r}_j \right)^H \left[\mathbf{r}_j^T \mathbf{Q}(\tilde{\nu}) \mathbf{r}_j^* \right]^{-1} \mathbf{r}_j^T \mathbf{Q}(\tilde{\nu}) \mathbf{r}_j \\ &= \mathbf{r}_j^H \mathbf{Q}(\tilde{\nu}) \mathbf{r}_j - \frac{\left| \mathbf{r}_j^T \mathbf{Q}(\tilde{\nu}) \mathbf{r}_j \right|^2}{\mathbf{r}_j^T \mathbf{Q}(\tilde{\nu}) \mathbf{r}_j^*} \end{aligned} , \tag{3.43}$$

where no matrix inversion is needed, and

$$\hat{d}_{j,s}(\hat{\nu}_s) = \mathbf{f}^H(\hat{\nu}_s) \left(\mathbf{r}_j - \mathbf{r}_j^* \hat{\rho}_j(\hat{\nu}_s) \right), \tag{3.44}$$

with

$$\hat{\rho}_j(\hat{\nu}_s) = \frac{\mathbf{r}_j^T \mathbf{Q}(\hat{\nu}_s) \mathbf{r}_j}{\mathbf{r}_j^T \mathbf{Q}(\hat{\nu}_s) \mathbf{r}_j^*} . \tag{3.45}$$

It will be shown in Section 3.5 that no degradation on the BER performance when using the simplified estimators $\hat{\nu}_s$ and $\hat{d}_{j,s}(\hat{\nu}_s)$.

3.3 Computational Complexity Analysis

The computational complexities of the joint LS estimation and the low-complexity implementation are analyzed in this section. Only the required number of real multiplications is compared because the computational complexity is mainly dominated by multiplication. The complexity of frequency estimator $\hat{\nu}$ in (3.31) is first analyzed in subsection 3.3.1 followed by the analyses of I-Q imbalance, dc offset and channel estimators $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$ and $\hat{\mathbf{g}}_j$ (3.33)-(3.35) in subsection 3.3.2. Finally, the complexity comparisons of proposed methods and previous method in [31] as our special case are discussed in subsection 3.3.3

3.3.1 Frequency Estimator $\hat{\nu}$

- **Complexity of $\mathbf{v}^H\mathbf{C}(\tilde{\nu})\mathbf{u}$:**

Firstly, it requires $N(N+1)$ complex products for calculation of the correlations $\{\eta_{r,s}(m)\}$ and $\{\varsigma_{r,s}(m)\}$ in (3.24) and (3.25). Secondly, it requires $\eta(KN/2)\log_2(KN)$ complex products for calculation of the KN-pt FFT in (3.28) and (3.29), where $\eta = 1-(\log_2 K + 2(1/K-1))/\log_2(KN)$. Finally, the calculation of $\mathbf{v}^H\mathbf{C}(\tilde{\nu}_n)\mathbf{u}$ in (3.23) needs $2P^2$ FFT with correlations and additional P^2 complex products. Totally, it requires $P^2N\cdot[K+2N+2+\eta K\log_2(KN)]$ complex products to calculate $\{\mathbf{v}^H\mathbf{C}(\tilde{\nu}_n)\mathbf{u}\}_{n=1}^{KN}$.

For phase-rotated periodic training, only $2Q$ correlations $\{\eta_{r,s}(m)\}$ and $\{\varsigma_{r,s}(m)\}$ in (3.40) and (3.41) are involved in computation, and the FFT in (3.39) is now performed over KQ points, not KN. Hence, the calculation of $\{\mathbf{v}^H\mathbf{C}(\tilde{\nu}_n)\mathbf{u}\}_{n=1}^{KQ}$ with phase-rotated periodic training requires $P^2(N/n_tL_g)\cdot[K+N+n_tL_g+\eta K\log_2(KN/n_tL_g)]$ complex product, which is reduce by about n_tL_g times.

- **Complexity of $\mathbf{v}^H\mathbf{Q}(\tilde{\nu})\mathbf{u}$:**

Assume that the scalars $\{1/\mathbf{1}_{PN}^H(\mathbf{I}_{PN}-\mathbf{C}(\tilde{\nu}_n))\mathbf{1}_{PN}\}_{n=1}^{KN}$ and the vectors $\{(\mathbf{I}_{PN}-\mathbf{C}(\tilde{\nu}_n))\mathbf{1}_{PN}\}_{n=1}^{KN}$ can be pre-calculated and saved. Then, the calculation of $\{\mathbf{v}^H\mathbf{Q}(\tilde{\nu}_n)\mathbf{u}\}_{n=1}^{KN}$ in (3.22) requires $P^2N\cdot[K+2N+2+\eta K\log_2(KN)]+PN\cdot(2KN+1)+2KN$ complex products.

For phase-rotated periodic training, the calculation of $\{\mathbf{v}^H\mathbf{Q}(\tilde{\nu}_n)\mathbf{u}\}_{n=1}^{KQ}$ requires $P^2(N/n_tL_g)\cdot[K+N+n_tL_g+\eta K\log_2(KN/n_tL_g)]+PN\cdot(2KN/n_tL_g+1)+2KN/n_tL_g$ complex

products.

- **Complexity of Frequency Estimator $\hat{\nu}$**

Now, the complexity of coarse search of the cost function $\left\{\Lambda_j(\tilde{\nu}_n)\right\}_{n=1}^{KN}$ in (3.20) is analyzed as follows while the interpolation based fine search is relatively easier and can be neglected. Define $\mathbf{r}_j^{(k)}$ as the $k'th$ column of $\mathbf{R}_j$. First, it requires to calculate $\left\{\mathbf{r}_j^H\mathbf{Q}(\tilde{\nu}_n)\mathbf{r}_j\right\}_{n=1}^{KN}$, $\left\{\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu}_n)\mathbf{r}_j\right\}_{n=1}^{KN}$ corresponding to $\left\{\mathbf{r}_j^{(k)T}\mathbf{Q}(\tilde{\nu}_n)\mathbf{r}_j\right\}_{n=1}^{KN}$, for $k=1\sim L_\rho$, and $\left\{\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu}_n)\mathbf{R}_j^*\right\}_{n=1}^{KN}$ corresponding to $\left\{\mathbf{r}_j^{(k)T}\mathbf{Q}(\tilde{\nu}_n)\mathbf{r}_j^{(l)*}\right\}_{n=1}^{KN}$, for $k=1\sim L_\rho$ with $l=k\sim L_\rho$. This part of complexity is $\left[\left(1+L_\rho\right)L_\rho/2+L_\rho+1\right]\cdot\left\{P^2N\cdot\left[K+2N+2+\eta K\log_2(KN)\right]+PN\cdot(2KN+1)+2KN\right\}$ complex products. Second, the inverse matrix term $\left\{\left[\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu}_n)\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu}_n)\mathbf{r}_j\right\}_{n=1}^{KN}$ can be implemented by Gaussian elimination, whose complexity is $KN\cdot\left(L_\rho^3+3L_\rho^2-L_\rho\right)/3$ complex products. Finally, it further needs $KN\cdot L_\rho$ complex products. The number of overall required complex products for frequency estimator is

$$n_r\cdot\left[\left(L_\rho^2+3L_\rho+2\right)/2\right]\cdot\left\{P^2N\cdot\left[K+2N+2+\eta K\log_2(KN)\right]+PN\cdot(2KN+1)+2KN\right\}$$
$$+n_r\cdot KN\cdot\left(L_\rho^3+3L_\rho^2+2L_\rho\right)/3. \tag{3.46}$$

For phase-rotated periodic training, the first part of complexity is $\left[\left(1+L_\rho\right)L_\rho/2+L_\rho+1\right]\cdot\left\{P^2\left(N/n_tL_{\mathbf{g}}\right)\cdot\left[K+N+n_tL_{\mathbf{g}}+\eta K\log_2\left(KN/n_tL_{\mathbf{g}}\right)\right]+PN\cdot\left(2KN/n_tL_{\mathbf{g}}+1\right)\right.$ $\left.+2KN/n_tL_{\mathbf{g}}\right\}$ complex products. The second part requires $\left(KN/n_tL_{\mathbf{g}}\right)\cdot\left(L_\rho^3+3L_\rho^2-L_\rho\right)/3$ complex products. Finally, it further needs $\left(KN/n_tL_{\mathbf{g}}\right)\cdot L_\rho$ complex products. The number of overall required complex products for frequency estimator with phase-rotated periodic training is

$$\left[\left(1+L_\rho\right)L_\rho/2+L_\rho+1\right]\cdot\left\{P^2\left(N/n_tL_{\mathbf{g}}\right)\cdot\left[K+N+n_tL_{\mathbf{g}}+\eta K\log_2\left(KN/n_tL_{\mathbf{g}}\right)\right]\right.$$
$$\left.+PN\cdot\left(2KN/n_tL_{\mathbf{g}}+1\right)+2KN/n_tL_{\mathbf{g}}\right\}+n_r\cdot\left(KN/n_tL_{\mathbf{g}}\right)\cdot\left(L_\rho^3+3L_\rho^2+2L_\rho\right)/3. \tag{3.47}$$

3.3.2 I-Q Imbalance, DC Offset and Channel Estimators $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$ and $\hat{\mathbf{g}}_j$.

Assume that $\mathbf{B}=\mathbf{S}\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H$ in (3.10) can be pre-calculated and saved. It requires extra $\left(P^2N^2-PN\right)/2$ complex products for calculation of the Hermitian matrix $\mathbf{C}\left(\hat{\nu}\right)=\boldsymbol{\Gamma}\left(\hat{\nu}\right)\mathbf{B}\boldsymbol{\Gamma}^H\left(\hat{\nu}\right)$ in (3.9) and P^2N^2+PN complex products for $\mathbf{v}^H\mathbf{C}\left(\hat{\nu}\right)\mathbf{u}$. Then, to calculate $\mathbf{v}^H\mathbf{Q}\left(\hat{\nu}\right)\mathbf{u}$, it totally requires P^2N^2+4PN complex products.

For $\hat{\boldsymbol{\rho}}_j\left(\hat{\nu}\right)$, first, it requires to calculate $\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j$ corresponding to $\mathbf{r}_j^{(k)T}\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j$, for $k=1\sim L_\rho$, and $\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{R}_j^*$ corresponding to $\mathbf{r}_j^{(k)T}\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j^{(l)*}$, for $k=1\sim L_\rho$ with $l=k\sim L_\rho$, whose complexity is $\left(P^2N^2+4PN\right)\cdot\left(L_\rho^2+3L_\rho\right)/2$ complex products. Second, the inverse matrix term $\left[\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j$ can be implemented by Gaussian elimination, whose complexity is $\left(L_\rho^3+3L_\rho^2-L_\rho\right)/3$ complex products. The number of overall required complex products for I-Q imbalance estimator $\hat{\boldsymbol{\rho}}_j\left(\hat{\nu}\right)$ is

$$P^2N^2\cdot\left(L_\rho^2+3L_\rho+1\right)/2+PN\cdot\left(4L_\rho^2+12L_\rho-1\right)/2+\left(L_\rho^3+3L_\rho^2-L_\rho\right)/3. \tag{3.48}$$

For $\hat{d}_j\left(\hat{\nu},\hat{\boldsymbol{\rho}}_j\right)$, the vector $\mathbf{f}^H\left(\hat{\nu}\right)$ in (3.12) has been calculated during the calculation of $\hat{\boldsymbol{\rho}}_j\left(\hat{\nu}\right)$. Hence, the number of overall required complex products for dc offset estimator $\hat{d}_j\left(\hat{\nu},\hat{\boldsymbol{\rho}}_j\right)$ is

$$\left(L_\rho+1\right)PN\,. \tag{3.49}$$

For $\hat{\mathbf{g}}_j\left(\hat{\nu},\hat{\boldsymbol{\rho}}_j,\hat{d}_j\right)$, assume that the matrix $\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H$ in (3.35) can be pre-calculated and saved and the term $\left(\mathbf{r}_j-\mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j\right)$ has been calculated during the calculation of dc offset estimator. Hence, to calculate $\boldsymbol{\Gamma}^H\left(\hat{\nu}\right)\left(\mathbf{r}_j-\mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j-\hat{d}_j\mathbf{1}\right)$ in (3.35), it requires only PN complex products. The number of overall required complex products for channel estimator $\hat{\mathbf{g}}_j\left(\hat{\nu},\hat{\boldsymbol{\rho}}_j,\hat{d}_j\right)$ is

$$\left(n_tL_{\mathbf{g}}+1\right)PN\,. \tag{3.50}$$

For phase-rotated periodic training, the required complex products for calculation of the Hermitian matrix $\mathbf{C}\left(\hat{\nu}\right)=\boldsymbol{\Gamma}\left(\hat{\nu}\right)\mathbf{B}\boldsymbol{\Gamma}^H\left(\hat{\nu}\right)$ reduce to $\left(PN/n_tL_{\mathbf{g}}\right)\left(PN-N\right)/2$ not $PN\left(PN-1\right)/2$ due to the sparse matrix $\mathbf{B}$ in (3.37). Similarly, the required complex products for $\mathbf{v}^H\mathbf{C}\left(\hat{\nu}\right)\mathbf{u}$ reduce to $P^2N^2/n_tL_{\mathbf{g}}+PN$ not P^2N^2+PN. Then, the complexity of $\mathbf{v}^H\mathbf{Q}\left(\hat{\nu}\right)\mathbf{u}$ becomes $P^2N^2/n_tL_{\mathbf{g}}+4PN$ complex products. Hence, the number of overall required complex products for I-Q imbalance estimator $\hat{\boldsymbol{\rho}}_j\left(\hat{\nu}\right)$ with phase-rotated periodic training becomes

$$\left(P^2N^2/n_tL_{\mathbf{g}}\right)\cdot\left(L_\rho^2+3L_\rho+1\right)/2+PN\cdot\left(4L_\rho^2+12L_\rho-Q\right)/2+\left(L_\rho^3+3L_\rho^2-L_\rho\right)/3\,. \tag{3.51}$$

For $\hat{d}_j\left(\hat{\nu},\hat{\boldsymbol{\rho}}_j\right)$ and $\hat{\mathbf{g}}_j\left(\hat{\nu},\hat{\boldsymbol{\rho}}_j,\hat{d}_j\right)$, the complexities are kept the same as the original ones.

3.3.3 Discussion

The number of required real multiplications for all estimators with arbitrary training and phase-rotated periodic training are summarized in Table 3.1. Some comparisons and special cases are addressed as follows.

As is shown, the complexities of frequency and I-Q imbalance estimators $\hat{\nu}$ and $\hat{\boldsymbol{\rho}}_j$ with

Table 3.1: Computational Complexities of Arbitrary Training and Phase-Rotated Periodic Training (Number of Real Multiplications)

	Arbitrary Training	Phase-Rotated Periodic Training
Frequency offset $\hat{\nu}$	$n_r \cdot \left(2L_\rho^2 + 6L_\rho + 4\right) \cdot \left\{P^2 N\left[K + 2N + 2 + \eta K \log_2\left(KN\right)\right] + PN \cdot \left(2KN + 1\right) + 2KN\right\} + n_r \cdot KN \cdot \left(4L_\rho^3 + 12L_\rho^2 + 8L_\rho\right)/3$	$n_r\left(2L_\rho^2 + 6L_\rho + 4\right) \cdot \left\{P^2\left(N/n_t L_{\mathbf{g}}\right) \cdot \left[K + N + n_t L_{\mathbf{g}} + \eta K \log_2\left(KN/n_t L_{\mathbf{g}}\right)\right] + PN \cdot \left(2KN/n_t L_{\mathbf{g}} + 1\right) + 2KN/n_t L_{\mathbf{g}}\right\} + n_r\left(KN/n_t L_{\mathbf{g}}\right) \cdot \left(4L_\rho^3 + 12L_\rho^2 + 8L_\rho\right)/3$
I-Q imbalance $\hat{\boldsymbol{\rho}}_j$	$P^2N^2 \cdot \left(2L_\rho^2 + 6L_\rho + 2\right) + PN \cdot \left(8L_\rho^2 + 24L_\rho - 2\right) + \left(4L_\rho^3 + 12L_\rho^2 - 4L_\rho\right)/3$	$\left(P^2N^2/n_t L_{\mathbf{g}}\right) \cdot \left(2L_\rho^2 + 6L_\rho + 2\right) + PN \cdot \left(8L_\rho^2 + 24L_\rho - 2Q\right) + \left(2L_\rho^3 + 6L_\rho^2 - 2L_\rho\right)/3$
DC offset $\hat{d}_j$	$4PN \cdot \left(L_\rho + 1\right)$	$4PN \cdot \left(L_\rho + 1\right)$
Channel $\hat{\mathbf{g}}_j$	$4PN\left(n_t L_{\mathbf{g}} + 1\right)$	$4PN\left(n_t L_{\mathbf{g}} + 1\right)$

phase-rotated periodic training are reduced by about $n_t L_{\mathbf{g}}$ times compared with those with general training. However, the frequency range that can be estimated with phase-rotated periodic training is also reduced by the same factor. For the simplified estimators of frequency and dc offset, the complexities reduce to the case of $L_\rho = 1$. Finally, the proposed estimators reduce to the ML estimators in [31] for the special case of SISO systems ($n_t = n_r = 1$) with frequency-independent I-Q imbalance ($L_\rho = 1$) and one training block ($P = 1$).

3.4 Performance Analysis

In this section, the mean and variance of the estimators $\hat{\nu}$, $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$, and $\hat{\mathbf{g}}_j$ are analyzed under the conditions of $\mathrm{SNR} \gg 1$ and $N \gg 1$, where $\mathrm{SNR} \doteq \sigma_s^2 / \sigma_w^2$ with σ_s^2 and σ_w^2 defined in subsection 2.2.1. We start from the analysis of frequency estimator $\hat{\nu}$.

3.4.1 Mean and Variance of Frequency Estimator $\hat{\nu}$

From [69] and [72], using the fact that $\hat{\nu}$ is close to true carrier frequency offset ν for $\mathrm{SNR} \gg 1$ and $N \gg 1$, one has

$$\mathrm{E}\{\hat{\nu}\} \approx \nu - \frac{\mathrm{E}\{\dot{\Lambda}(\nu)\}}{\mathrm{E}\{\ddot{\Lambda}(\nu)\}}, \tag{3.52}$$

and

$$\mathrm{E}\left\{(\hat{\nu}-\nu)^2\right\} \approx \frac{\mathrm{E}\left\{\left[\dot{\Lambda}(\nu)\right]^2\right\}}{\left[\mathrm{E}\{\ddot{\Lambda}(\nu)\}\right]^2}, \tag{3.53}$$

where $\dot{\Lambda}(\nu) = \dfrac{\partial \Lambda(\nu)}{\partial \nu}$, and $\ddot{\Lambda}(\nu) = \dfrac{\partial^2 \Lambda(\nu)}{\partial^2 \nu}$. Unfortunately, it is quite cumbersome to evaluate (3.52) and (3.53) with $\Lambda(\nu)$ given in (3.17), where ν is estimated jointly with $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$, and $\hat{\mathbf{g}}_j$. To simplify the analysis, the effects of I-Q imbalance and dc offset on frequency estimation will be neglected here, i.e., $\boldsymbol{\rho}_j = \mathbf{0}_{L_\rho}$, and $d_j = 0$. By setting $\boldsymbol{\rho}_j = \mathbf{0}_{L_\rho}$, and $\mathbf{f}(\tilde{\nu}) = \mathbf{0}_{PN}$, (3.17) becomes,

$$\begin{aligned}
\Lambda(\tilde{\nu}) &= \sum_{j=1}^{n_r} \left\| \left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{r}_j \right\|^2 \\
&= \sum_{j=1}^{n_r} \mathbf{r}_j^H \left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{r}_j \\
&= \sum_{j=1}^{n_r} \left\|\mathbf{r}_j\right\|^2 - \mathbf{r}_j^H \boldsymbol{\Gamma}(\tilde{\nu})\mathbf{B}\boldsymbol{\Gamma}^H(\tilde{\nu})\mathbf{r}_j \ .
\end{aligned} \tag{3.54}$$

It will be shown in Section 3.5 by computer simulations that the simplified analysis is very accu-

rate for the ranges of $\boldsymbol{\rho}_j$ and d_j of practical interest.

From (3.54), it can be shown that

$$\dot{\Lambda}(\nu) = \boldsymbol{j}2\pi \cdot \sum_{j=1}^{n_r} \mathbf{r}_j^H \boldsymbol{\Gamma}(\nu)\mathbf{D}\boldsymbol{\Gamma}^H(\nu)\mathbf{r}_j \,, \tag{3.55}$$

and

$$\ddot{\Lambda}(\nu) = 4\pi^2 \cdot \sum_{j=1}^{n_r} \mathbf{r}_j^H \boldsymbol{\Gamma}(\nu)\mathbf{E}\boldsymbol{\Gamma}^H(\nu)\mathbf{r}_j \,, \tag{3.56}$$

with $\mathbf{D} = \boldsymbol{\Phi}\mathbf{B} - \mathbf{B}\boldsymbol{\Phi}$, $\mathbf{E} = 2\boldsymbol{\Phi}\mathbf{B}\boldsymbol{\Phi} - \mathbf{B}\boldsymbol{\Phi}^2 - \boldsymbol{\Phi}^2\mathbf{B}$. Recall that $\mathbf{B} = \mathbf{S}\left(\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H$. Using $\mathbf{r}_j = \boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j + \mathbf{w}_j$ in (3.55) and (3.56) one gets

$$\dot{\Lambda}(\nu) = \boldsymbol{j}2\pi \cdot \sum_{j=1}^{n_r} \mathbf{g}_j^H\mathbf{S}^H\mathbf{D}\tilde{\mathbf{w}}_j + \tilde{\mathbf{w}}_j^H\mathbf{D}\mathbf{S}\mathbf{g}_j + \tilde{\mathbf{w}}_j^H\mathbf{D}\tilde{\mathbf{w}}_j \,, \tag{3.57}$$

$$\ddot{\Lambda}(\nu) = 4\pi^2 \cdot \sum_{j=1}^{n_r} \mathbf{g}_j^H\mathbf{S}^H\mathbf{E}\mathbf{S}\mathbf{g}_j + \mathbf{g}_j^H\mathbf{S}^H\mathbf{E}\tilde{\mathbf{w}}_j + \tilde{\mathbf{w}}_j^H\mathbf{E}\mathbf{S}\mathbf{g}_j + \tilde{\mathbf{w}}_j^H\mathbf{E}\tilde{\mathbf{w}}_j \,, \tag{3.58}$$

where $\tilde{\mathbf{w}}_j = \boldsymbol{\Gamma}^H(\nu)\mathbf{w}_j$ with $E\left\{\tilde{\mathbf{w}}_j\tilde{\mathbf{w}}_j^H\right\} \approx \sigma_w^2\mathbf{I}_{PN}$. In addition,

$$\begin{aligned} E\left\{\dot{\Lambda}(\nu)\right\} &= \boldsymbol{j}2\pi \cdot E\left\{\sum_{j=1}^{n_r} \tilde{\mathbf{w}}_j^H\mathbf{D}\tilde{\mathbf{w}}_j\right\} \\ &\approx \boldsymbol{j}2\pi \cdot \sum_{j=1}^{n_r} tr\left\{\mathbf{D} \cdot E\left\{\tilde{\mathbf{w}}_j\tilde{\mathbf{w}}_j^H\right\}\right\} \\ &= \boldsymbol{j}2\pi\sigma_w^2 \cdot \sum_{j=1}^{n_r} tr\{\mathbf{D}\} \\ &= 0 \end{aligned} \tag{3.59}$$

and

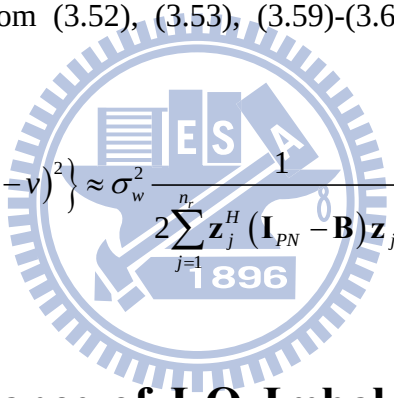
$$\begin{aligned} \mathrm{E}\left\{\ddot{\Lambda}(\nu)\right\} &= 4\pi^2 \cdot \sum_{j=1}^{n_r} \mathbf{g}_j^H\mathbf{S}^H\mathbf{E}\mathbf{S}\mathbf{g}_j + \mathrm{E}\left\{\tilde{\mathbf{w}}_j^H\mathbf{E}\tilde{\mathbf{w}}_j\right\} \\ &\approx 4\pi^2 \cdot \sum_{j=1}^{n_r} \mathbf{g}_j^H\mathbf{S}^H\mathbf{E}\mathbf{S}\mathbf{g}_j + \sigma_w^2 \cdot tr\{\mathbf{E}\} \\ &= 4\pi^2 \cdot \sum_{j=1}^{n_r} \mathbf{g}_j^H\mathbf{S}^H\mathbf{E}\mathbf{S}\mathbf{g}_j \\ &= 2 \cdot \sum_{j=1}^{n_r} \mathbf{z}_j^H\left(\mathbf{B} - \mathbf{I}_{PN}\right)\mathbf{z}_j \,, \end{aligned} \tag{3.60}$$

where $\mathbf{z}_j = \boldsymbol{j}2\pi\boldsymbol{\Phi}\mathbf{S}\mathbf{g}_j$, $\boldsymbol{\Phi} = diag\{\kappa(1),\kappa(2),\ldots,\kappa(P)\}$ and $\kappa(p) = [(N_g+N)p+1, (N_g+N)p+2,\ldots,(N_g+N)p+N]$. On the other hand,

$$\begin{aligned}
\mathrm{E}\left\{\left(\dot{\Lambda}(\nu)\right)^2\right\} &= -4\pi^2 \cdot \mathrm{E}\left\{\left(\sum_{j=1}^{n_r} \mathbf{g}_j^H \mathbf{S}^H \mathbf{D}\tilde{\mathbf{w}}_j + \tilde{\mathbf{w}}_j^H \mathbf{D}\mathbf{S}\mathbf{g}_j + \tilde{\mathbf{w}}_j^H \mathbf{D}\tilde{\mathbf{w}}_j\right)^2\right\} \\
&\approx -4\pi^2 \cdot \sum_{j=1}^{n_r} 2 \cdot \mathbf{g}_j^H \mathbf{S}^H \mathbf{D} \cdot E\left\{\tilde{\mathbf{w}}_j \tilde{\mathbf{w}}_j^H\right\} \cdot \mathbf{D}\mathbf{S}\mathbf{g}_j \qquad (3.61) \\
&= 2\sigma_w^2 \cdot \sum_{j=1}^{n_r} \mathbf{z}_j^H \left(\mathbf{I}_{PN} - \mathbf{B}\right)\mathbf{z}_j^H \ ,
\end{aligned}$$

The approximation in (3.61) is justifiable for SNR $\gg 1$, and the last equality is obtained by using $\mathbf{S}^H(\boldsymbol{\Phi}\mathbf{B}-\mathbf{B}\boldsymbol{\Phi}) = \mathbf{S}^H(\boldsymbol{\Phi}\mathbf{B}-\boldsymbol{\Phi})$, $(\boldsymbol{\Phi}\mathbf{B}-\mathbf{B}\boldsymbol{\Phi})\mathbf{S} = (\boldsymbol{\Phi}-\mathbf{B}\boldsymbol{\Phi})\mathbf{S}$, and the fact that $(\mathbf{I}_{PN}-\mathbf{B})$ is a projection matrix. Therefore, from (3.52), (3.53), (3.59)-(3.61), we have $\mathrm{E}\{(\hat{\nu}-\nu)\} \approx 0$ (an unbiased estimator) and

$$\mathrm{E}\left\{(\hat{\nu}-\nu)^2\right\} \approx \sigma_w^2 \frac{1}{2\sum_{j=1}^{n_r} \mathbf{z}_j^H \left(\mathbf{I}_{PN} - \mathbf{B}\right)\mathbf{z}_j}. \qquad (3.62)$$



3.4.2 Mean and Variance of I-Q Imbalance, DC Offset and Channel Estimators $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$ and $\hat{\mathbf{g}}_j$

Next we analyze the estimators $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$, and $\hat{\mathbf{g}}_j$ with no influence of frequency offset; that is, $\hat{\nu} = \nu$ is assumed in the analysis. It will be shown in Section 3.5 by computer simulations that the analysis predicts the MSE (mean-squared error) performance very well even when $\hat{\nu} \neq \nu$. For the estimator $\hat{\boldsymbol{\rho}}_j$, from (3.14) and (3.4), it is easy to show that

$$\hat{\boldsymbol{\rho}}_j(\nu) = \mathbf{P}_j(\nu)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\left(\mathbf{R}_j^* \boldsymbol{\rho}_j + \boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j + d_j \mathbf{1}_{PN} + \mathbf{w}_j\right). \qquad (3.63)$$

In addition, by using (3.21) and the identities,

$$\mathbf{P}_j(\nu)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{R}_j^*\boldsymbol{\rho}_j=\boldsymbol{\rho}_j \quad ,$$

$$\mathbf{P}_j(\nu)\underbrace{\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{1}_{PN}}_{=\mathbf{0}_{PN}}d_j=\mathbf{0}_{L_\rho} \quad ,$$

$$\mathbf{P}_j(\nu)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j=\left[\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\underbrace{\mathbf{Q}(\tilde{\nu})\boldsymbol{\Gamma}(\nu)\mathbf{S}}_{=\mathbf{0}_{PN}}\mathbf{g}_j=\mathbf{0}_{L_\rho},$$

one has,

$$\hat{\boldsymbol{\rho}}_j=\boldsymbol{\rho}_j+\mathbf{P}_j(\nu)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{w}_j. \tag{3.64}$$

For use in this section, $r_j(n)$ in (2.9) is rewritten as

$$r_j(n)=r_{j,d}(n)+r_{j,r}(n), \tag{3.65}$$

where $r_{j,d}(n)=c_{+,j}(n)\otimes y_j(n)e^{j2\pi\nu n}+c_{-,j}(n)\otimes y_j^*(n)e^{-j2\pi\nu n}+d_{0,j}$ is the deterministic part, and $r_{j,r}(n)=c_{+,j}(n)\otimes w_{0,j}(n)+c_{-,j}(n)\otimes w_{0,j}^*(n)$ is the random part due to noise. $\{r_{j,r}(n)\}$ is generally a colored Gaussian noise. In this way, $\mathbf{R}_j$ can be rewritten as $\mathbf{R}_j=\mathbf{R}_{j,d}+\mathbf{R}_{j,r}$, where $\mathbf{R}_{j,d}$ and $\mathbf{R}_{j,r}$ are the deterministic and random part of $\mathbf{R}_j$, respectively.

Taking expectation on both sides of (3.64),

$$\mathrm{E}\{\hat{\boldsymbol{\rho}}_j\,|\,\nu\}\approx\boldsymbol{\rho}_j+\mathbf{P}_{j,d}(\nu)\left(\mathbf{I}_{PN}-\mathbf{1}\mathbf{f}^H(\nu)\right)\mathrm{E}\{\mathbf{w}_j\}=\boldsymbol{\rho}_j \tag{3.66}$$

where $\mathbf{P}_{j,d}(\nu)$ is the deterministic part of $\mathbf{P}_j(\nu)$. The approximation is justifiable with SNR $\gg 1$. Furthermore, MSE is derived as

$$\begin{aligned}\mathrm{E}\left\{\left\|\hat{\boldsymbol{\rho}}_j-\boldsymbol{\rho}_j\right\|^2|\,\nu\right\}&=tr\left\{\mathrm{E}\left\{\left(\hat{\boldsymbol{\rho}}_j-\boldsymbol{\rho}_j\right)\left(\hat{\boldsymbol{\rho}}_j-\boldsymbol{\rho}_j\right)^H|\,\nu\right\}\right\}\\&=tr\left\{\mathrm{E}\left\{\mathbf{P}_j(\nu)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{w}_j\mathbf{w}_j^H\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)^H\mathbf{P}_j^H(\nu)\right\}\right\}\\&\approx tr\left\{\mathbf{P}_{j,d}(\nu)\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{K}_{\mathbf{w}_j}\left(\mathbf{I}_{PN}-\mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)^H\mathbf{P}_{j,d}^H(\nu)\right\},\end{aligned} \tag{3.67}$$

where $\mathbf{K}_{\mathbf{w}_j} \doteq \mathrm{E}\left\{\mathbf{w}_j \mathbf{w}_j^H\right\}$ is the correlation matrix of $\mathbf{w}_j$. The explicit expression of $\mathbf{K}_{\mathbf{w}_j}$ can be found in next subsection.

For the estimator $\hat{d}_j$, from (3.11)

$$\begin{aligned}\hat{d}_j\left(\hat{\boldsymbol{\rho}}_j, \nu\right) &= \mathbf{f}^H(\nu)\left(\mathbf{r}_j - \mathbf{R}_j^* \hat{\boldsymbol{\rho}}_j\right) \\ &= \underbrace{\mathbf{f}^H(\nu)\mathbf{1}_{PN}}_{=\mathbf{1}_{PN}} d_j + \mathbf{f}^H(\nu)\mathbf{R}_j^*\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right) + \underbrace{\mathbf{f}^H(\nu)\boldsymbol{\Gamma}(\nu)\mathbf{S}}_{=\mathbf{0}_{n_t L_{\mathbf{g}}}^T}\mathbf{g}_j + \mathbf{f}^H(\nu)\mathbf{w}_j \quad , \\ &= d_j + \mathbf{f}^H(\nu)\mathbf{R}_j^*\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right) + \mathbf{f}^H(\nu)\mathbf{w}_j\end{aligned} \tag{3.68}$$

Similar to (3.66), we have $\mathrm{E}\left\{\hat{d}_j - d_j \mid \nu\right\} \approx 0$, and MSE is given by, under SNR $\gg 1$,

$$\begin{aligned}\mathrm{E}\left\{\left|\hat{d}_j - d_j\right|^2 \mid \nu\right\} &\approx \mathrm{E}\left\{\left(\mathbf{f}^H(\nu)\mathbf{R}_{j,d}^*\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right) + \mathbf{f}^H(\nu)\mathbf{w}_j\right)\left(\mathbf{f}^H(\nu)\mathbf{R}_{j,d}^*\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right) + \mathbf{f}^H(\nu)\mathbf{w}_j\right)^H\right\} \\ &= \mathbf{f}^H(\nu)\mathbf{R}_{j,d}^* \mathrm{E}\left\{\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right)\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right)^H\right\}\mathbf{R}_{j,d}^T\mathbf{f}(\nu) + \mathbf{f}^H(\nu)\mathbf{K}_{\mathbf{w}_j}\mathbf{f}(\nu) \\ &\quad + 2\,\mathrm{Re}\left\{\mathbf{f}^H(\nu)\mathbf{R}_{j,d}^*\mathbf{P}_{j,d}(\nu)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{K}_{\mathbf{w}_j}\mathbf{f}(\nu)\right\}\end{aligned} \tag{3.69}$$

where we have used the approximation

$$\begin{aligned}\mathrm{E}\left\{\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right)\mathbf{w}_j^H\right\} &= \mathrm{E}\left\{\mathbf{P}_j(\nu)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{w}_j\mathbf{w}_j^H\right\} \\ &\approx \mathbf{P}_{j,d}(\nu)\left(\mathbf{I}_{PN} - \mathbf{1}_{PN}\mathbf{f}^H(\nu)\right)\mathbf{K}_{\mathbf{w}_j} \quad .\end{aligned} \tag{3.70}$$

Finally, for the estimator $\hat{\mathbf{g}}_j$,

$$\begin{aligned}\hat{\mathbf{g}}_j\left(\hat{d}_j, \hat{\boldsymbol{\rho}}_j, \nu\right) &= \left(\boldsymbol{\Gamma}(\nu)\mathbf{S}\right)^{\dagger}\left(\mathbf{r}_j - \mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j - \hat{d}_j\mathbf{1}_{PN}\right) \\ &= \left(\boldsymbol{\Gamma}(\nu)\mathbf{S}\right)^{\dagger}\left(\mathbf{R}_j^*\boldsymbol{\rho}_j + \boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j + d_j\mathbf{1}_{PN} + \mathbf{w}_j - \mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j - \hat{d}_j\mathbf{1}_{PN}\right) \\ &= \mathbf{g}_j + \mathbf{S}^{\dagger}\boldsymbol{\Gamma}^H(\nu)\mathbf{R}_j^*\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right) + \mathbf{S}^{\dagger}\boldsymbol{\Gamma}^H(\nu)\mathbf{1}_{PN}\left(d_j - \hat{d}_j\right) + \mathbf{S}^{\dagger}\boldsymbol{\Gamma}^H(\nu)\mathbf{w}_j \quad .\end{aligned} \tag{3.71}$$

Similarly, $\mathrm{E}\left\{\hat{\mathbf{g}}_j - \mathbf{g}_j \mid \nu\right\} \approx 0_{n_t L_{\mathbf{g}}}$, and MSE is given by

$$\begin{aligned}&\mathrm{E}\left\{\left\|\hat{\mathbf{g}}_j - \mathbf{g}_j\right\|^2 \mid \nu\right\} \\ &\quad \approx tr\left\{\mathbf{S}^{\dagger}\boldsymbol{\Gamma}^H(\tilde{\nu})\mathbf{R}_{j,d}^* \cdot \mathrm{E}\left\{\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right)\left(\boldsymbol{\rho}_j - \hat{\boldsymbol{\rho}}_j\right)^H\right\} \cdot \mathbf{R}_{j,d}^T\boldsymbol{\Gamma}(\tilde{\nu})\left(\mathbf{S}^{\dagger}\right)^H\right\} \\ &\quad\quad + tr\left\{\mathbf{S}^{\dagger}\boldsymbol{\Gamma}^H(\nu)\mathbf{1}_{PN} \cdot \mathrm{E}\left\{\left|\hat{d}_j - d_j\right|^2\right\} \cdot \mathbf{1}_{PN}^H\boldsymbol{\Gamma}(\nu)\left(\mathbf{S}^{\dagger}\right)^H\right\}\end{aligned}$$

$$
\begin{aligned}
&+tr\left\{\mathbf{S}^{\dagger}\boldsymbol{\Gamma}^{H}(\nu)\cdot\mathrm{E}\left\{\mathbf{w}_j\mathbf{w}_j^{H}\right\}\cdot\boldsymbol{\Gamma}(\nu)\left(\mathbf{S}^{\dagger}\right)^{H}\right\}\\
&+2\,\mathrm{Re}\left(tr\left\{\mathbf{S}^{\dagger}\boldsymbol{\Gamma}^{H}(\nu)\mathbf{R}_{j,d}^{*}\cdot\mathrm{E}\left\{\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)\left(d_j-\hat{d}_j\right)^{H}\right\}\cdot\mathbf{1}_{PN}^{H}\Gamma(\nu)\left(\mathbf{S}^{\dagger}\right)^{H}\right\}\right)\\
&+2\,\mathrm{Re}\left(tr\left\{\mathbf{S}^{\dagger}\boldsymbol{\Gamma}^{H}(\nu)\mathbf{R}_{j,d}^{*}\cdot\mathrm{E}\left\{\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)\mathbf{w}_j^{H}\right\}\cdot\boldsymbol{\Gamma}(\nu)\left(\mathbf{S}^{\dagger}\right)^{H}\right\}\right)\\
&+2\,\mathrm{Re}\left(tr\left\{\mathbf{S}^{\dagger}\boldsymbol{\Gamma}^{H}(\nu)\mathbf{1}_{PN}\cdot\mathrm{E}\left\{\left(d_j-\hat{d}_j\right)\mathbf{w}_j^{H}\right\}\cdot\boldsymbol{\Gamma}(\nu)\left(\mathbf{S}^{\dagger}\right)^{H}\right\}\right).
\end{aligned}
\tag{3.72}
$$

In deriving (3.72), we have used the following approximations

$$
\begin{aligned}
\mathrm{E}\left\{\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)\left(d_j-\hat{d}_j\right)^{H}\right\}&\approx\mathrm{E}\left\{\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)\left(\mathbf{f}^{H}(\nu)\mathbf{R}_{j,d}^{*}\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)+\mathbf{f}^{H}(\nu)\mathbf{w}_j\right)^{H}\right\}\\
&=\mathrm{E}\left\{\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)^{H}\right\}\mathbf{R}_{j,d}^{T}\mathbf{f}(\nu)+\mathrm{E}\left\{\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)\mathbf{w}_j^{H}\right\}\mathbf{f}(\nu)\ ,
\end{aligned}
\tag{3.73}
$$

and

$$
\mathrm{E}\left\{\left(d_j-\hat{d}_j\right)\mathbf{w}_j^{H}\right\}\approx\mathbf{f}^{H}(\nu)\mathbf{R}_{j,d}^{*}\mathrm{E}\left\{\left(\boldsymbol{\rho}_j-\hat{\boldsymbol{\rho}}_j\right)\mathbf{w}_j^{H}\right\}+\mathbf{f}^{H}(\nu)\mathbf{K}_{\mathbf{w}_j}\,.
\tag{3.74}
$$

3.4.3 Derivation of Cramér-Rao Lower Bound

In previous subsections, it is shown that $\hat{\nu}$, $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$, and $\hat{\mathbf{g}}_j$ are unbiased estimators provided that $\mathrm{SNR}\gg 1$ and $N\gg 1$. In this subsection, we derive the CRLB for those estimators. The derivation follows that given in [31][73]. From (3.4),

$$
\mathbf{r}_j-\mathbf{R}_j^{*}\boldsymbol{\rho}_j=\boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j+d_j\mathbf{1}+\mathbf{w}_j\,.
$$

Define $\boldsymbol{\omega}=\left[\nu,\boldsymbol{\lambda}_1^{T},\ldots,\boldsymbol{\lambda}_{n_r}^{T}\right]^{T}$ with $\boldsymbol{\lambda}_j^{T}=\left[\boldsymbol{\rho}_j^{T},d_j,\mathbf{g}_j^{T},\boldsymbol{\rho}_j^{H},d_j^{*},\mathbf{g}_j^{H}\right]$ of length $L_{\boldsymbol{\lambda}}$. Since $\mathbf{w}_j$, $j=1,\ldots,n_r$ are independent of each other, the probability density function of the observation vector $\mathbf{r}_j$, given $\boldsymbol{\omega}$, is

$$
\begin{aligned}
f(\boldsymbol{\omega})\doteq f\left(\mathbf{r}_1,\cdots,\mathbf{r}_{n_r}\,|\,\boldsymbol{\omega}\right)&=\prod_{j=1}^{n_r}p\left(\mathbf{r}_j-\mathbf{R}_j^{*}\boldsymbol{\rho}_j\,|\,\mathbf{g}_j,d_j,\boldsymbol{\rho}_j,\nu\right)\\
&=\prod_{j=1}^{n_r}\frac{1}{\pi^{N}\det\left(\mathbf{K}_{\mathbf{w}_j}\right)}\exp\left\{-\left(\mathbf{r}_j-\mathbf{R}_j^{*}\boldsymbol{\rho}_j-d_j\mathbf{1}_{PN}-\boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j\right)^{H}\mathbf{K}_{\mathbf{w}_j}^{-1}\left(\mathbf{r}_j-\mathbf{R}_j^{*}\boldsymbol{\rho}_j-d_j\mathbf{1}_{PN}-\boldsymbol{\Gamma}(\nu)\mathbf{S}\mathbf{g}_j\right)\right\},
\end{aligned}
$$

(3.75)

where $\mathbf{K}_{\mathbf{w}_j} \doteq \mathrm{E}\left\{\mathbf{w}_j \mathbf{w}_j^H\right\}$ is the correlation matrix of $\mathbf{w}_j$. Recall that

$$\mathbf{w}_j = \left[\mathbf{w}_j^T(0), \mathbf{w}_j^T(1), \ldots, \mathbf{w}_j^T(P-1)\right]^T$$

with $w_j(n) = c_j(n) \otimes w_{0,j}(n)$, we have $\mathbf{w}_j(k) = \mathbf{\Phi}_{g_j} \mathbf{w}_{0,j}(k)$ and

$$\mathbf{w}_{0,j}(k) = \left[w_{0,j,k}\left(-L_{g_j}+1\right) \cdots w_{0,j,k}(0)\ w_{0,j,k}(1) \cdots w_{0,j,k}(N-1)\right]^T$$

with $w_{0,j,k}(n) \doteq w_{0,j}\left(k\left(N_g+N\right)+n\right)$ and

$$\mathbf{\Phi}_{g_j} = \begin{bmatrix} g_j\left(L_{g_j}-1\right) & \cdots & g_j(0) & 0 & \cdots & 0 \\ 0 & g_j\left(L_{g_j}-1\right) & \cdots & g_j(0) & 0 & \cdots & 0 \\ & \ddots & & \ddots & & \ddots & \\ 0 & \cdots & 0 & g_j\left(L_{g_j}-1\right) & \cdots & & g_j(0) \end{bmatrix}_{N \times \left(N+L_{g_j}-1\right)}, \tag{3.76}$$

where L_{g_j} is the length of $g_j(n)$. Therefore,

$$\mathbf{w}_j = \tilde{\mathbf{\Phi}}_j \mathbf{w}_{0,j} \tag{3.77}$$

with $\mathbf{w}_{0,j} = \left[\mathbf{w}_{0,j}^T(0), \mathbf{w}_{0,j}^T(1), \ldots, \mathbf{w}_{0,j}^T(P-1)\right]^T$ and $\tilde{\mathbf{\Phi}}_j = diag\underbrace{\left\{\mathbf{\Phi}_{g_j}, \mathbf{\Phi}_{g_j}, \ldots, \mathbf{\Phi}_{g_j}\right\}}_{P}$. Finally, we can get

$$\mathbf{K}_{\mathbf{w}_j} = \sigma_w^2 \tilde{\mathbf{\Phi}}_j \tilde{\mathbf{\Phi}}_j^H. \tag{3.78}$$

Let $\mathbf{M} = E\left\{\left(\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T}\right)^H \left(\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T}\right)\right\}$ be the Fisher information matrix. From [73], the CRLB for each respective parameter is given by,

$$\mathrm{var}(\hat{v}) \geq \left[\mathbf{M}^{-1}\right]_{11}, \tag{3.79}$$

$$\mathrm{E}\left\{\left\|\hat{\boldsymbol{\rho}}_j - \boldsymbol{\rho}_j\right\|^2\right\} \geq \sum_{k=1+(j-1)L_\lambda}^{1+(j-1)L_\lambda+L_\rho} \left[\mathbf{M}^{-1}\right]_{kk}, \tag{3.80}$$

$$\mathrm{var}\left(\hat{d}_j\right) \geq \left[\mathbf{M}^{-1}\right]_{kk\left|k=1+(j-1)L_\lambda+L_\rho+1\right.}, \tag{3.81}$$

and

$$\mathrm{E}\left\{\left\|\hat{\mathbf{g}}_j - \mathbf{g}_j\right\|^2\right\} \geq \sum_{k=1+(j-1)L_\lambda+L_\rho+2}^{1+(j-1)L_\lambda+L_\rho+1+L_g} \left[\mathbf{M}^{-1}\right]_{kk}. \tag{3.82}$$

By definition,

$$\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^T} = \left(\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \nu} \;\; \frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\lambda}_1^T} \;\; \cdots \;\; \frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\lambda}_{n_r}^T} \right), \tag{3.83}$$

and

$$\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\lambda}_j^T} = \left(\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\rho}_j^T} \;\; \frac{\partial \ln f(\boldsymbol{\omega})}{\partial d_j} \;\; \frac{\partial \ln f(\boldsymbol{\omega})}{\partial \mathbf{g}_j^T} \;\; \frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\rho}_j^H} \;\; \frac{\partial \ln f(\boldsymbol{\omega})}{\partial d_j^*} \;\; \frac{\partial \ln f(\boldsymbol{\omega})}{\partial \mathbf{g}_j^H} \right). \tag{3.84}$$

After some derivation, the partial derivative of $\ln f(\boldsymbol{\omega})$ with respect to each parameter is shown to be

$$\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \nu} = \sum_{j=1}^{n_r} \left(\mathbf{w}_j^H \mathbf{K}_{\mathbf{w}_j}^{-1} \boldsymbol{\Gamma}(\nu) \mathbf{z}_j + \mathbf{z}_j^H \boldsymbol{\Gamma}^H(\nu) \mathbf{K}_{\mathbf{w}_j}^{-1} \mathbf{w}_j \right), \tag{3.85}$$

$$\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\rho}_j^T} = \left(\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \boldsymbol{\rho}_j^H} \right)^* = \mathbf{w}_j^H \mathbf{K}_{\mathbf{w}_j}^{-1} \mathbf{R}_j^*, \tag{3.86}$$

$$\frac{\partial \ln f(\boldsymbol{\omega})}{\partial d_j} = \left(\frac{\partial \ln f(\boldsymbol{\omega})}{\partial d_j^*} \right)^* = \mathbf{w}_j^H \mathbf{K}_{\mathbf{w}_j}^{-1} \mathbf{1}, \tag{3.87}$$

and

$$\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \mathbf{g}_j^T} = \left(\frac{\partial \ln f(\boldsymbol{\omega})}{\partial \mathbf{g}_j^H} \right)^* = \mathbf{w}_j^H \mathbf{K}_{\mathbf{w}_j}^{-1} \boldsymbol{\Gamma}(\nu) \mathbf{S}. \tag{3.88}$$

From (3.83)-(3.88), the matrix $\mathbf{M}$ can be evaluated.

Table 3.2: System Parameters and Radio Impairments

System Parameter and Radio Impairments	Parameter Value
Channel Bandwidth	20 MHz
FFT length (N), cyclic prefix length (N_g)	$N=64$, $N_g=16$
OFDM-Symbol Time (T_{OFDM}), Symbol Time (T_s)	$T_{OFDM}=4\ \mu s$, $T_s=50\ ns$
Sub-carrier Spacing $\left(1/NT_s\right)$	0.3125 MHz
Number of Transmit and Receive Antenna (n_t,n_r)	$n_t=2$, $n_r=2,3$
Frequency independent I-Q Imbalance (α_j,θ_j)	$(\alpha_1=1.08,\theta_1=5^o),(\alpha_2=1.09,\theta_2=6^o)$ $(\alpha_3=1.1,\ \theta_3=7^o)$
$\{h_j^I(n),h_j^Q(n)\}$: 2rd order Butterworth with cut-off frequency (f_j^I, f_j^Q) MHz	$(f_1^I=8, f_1^Q=8.3),(f_2^I=7.9, f_2^Q=8.2)$ $(f_3^I=8.1, f_3^Q=8.4)$
Frequency offset v	$(0.25/N)$ for static channel, uniform over $\left(-\frac{0.5}{N},\frac{0.5}{N}\right)$ for fading channel
DC offset $(d_{0,j})$, with signal power normalized to 1	$d_{0,1}=\frac{0.2}{\sqrt{2}}(1+j)$, $d_{0,2}=\frac{0.15}{\sqrt{2}}(1+j)$, $d_{0,3}=\frac{0.1}{\sqrt{2}}(1+j)$

3.5 Simulation Results

The performance of the proposed estimators is evaluated for an un-coded MIMO OFDM system. Table 3.2 gives the system parameters and radio impairments which are typical values in real systems [5]-[7],[50]-[52],[74]. The transmission is done on a packet-by-packet basis with the training portion consisting of two OFDM symbols at the beginning of each packet. A wide-sense

stationary uncorrelated scattering (WSSUS) discrete channel is considered, with the impulse response $h_{j,i}(\tau)=\sum_{l=0}^{L} h_{j,i}(l)\delta(\tau-lT_s)$, where $L+1$ is the length of the channel, and $\{h_{j,i}(l)\}$ are tap gains which are mutually independent complex Gaussian random variables with zero mean and variance σ_l^2 . Exponential multi-path intensity profile is employed with $\sigma_l^2=\sigma_0^2\cdot\exp(-lT_s/T_{RMS})$, where T_{RMS} is the root-mean square delay spread, and to maintain unit power gain, $\sigma_0^2=1-\exp(-T_s/T_{RMS})$. The channel remains unchanged during a packet. $T_{RMS}=50\ ns$, $L=10$, and $L_{\mathbf{g}}=16$. In Figures 3.2-3.6, the training sequence is the one given in [75] for the case of $n_t=2$. In Figure 8, the low-complexity training sequence of the square matrix $\mathbf{A}$ is designed as $5230F641_H$ given in [76] for the first transmit antenna and its circular shift by $L_{\mathbf{g}}$ for the second transmit antenna. In addition, $K=P=2$, $\phi_0=0$, and $\phi_1=\pi/2$. The frequency resolutions for coarse and fine frequency estimation are 0.1 and 0.001 sub-carrier spacing, respectively.

In Figure 3.2, simulations are given to verify the MSE analysis with the system $n_t=n_r=2$ under the static channel $h_{j,i}(l)=\sqrt{(1-\exp(-1))(1-\exp(-l))},\ \forall j,i,\ 0\le l\le 10$. Only the results of the 1st receive antenna are shown; similarity is observed for the 2nd receive antenna. As is shown, the analysis predicts the MSE performance very well for all the estimators in the SNRs of interest. (Note that the simulations for the estimators $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$, and $\hat{\mathbf{g}}_j$ have used the real estimated frequency $\hat{v}$ which may not be equal to the true frequency v.) In addition, the variance of estimators approaches to respective CRLB. One observation worthy noting is that the analysis of frequency estimation is done under the perfect condition of no dc offset and I-Q imbalance. This explains why the analytical MSE of frequency estimator is smaller slightly than its corresponding CRLB in Figure 3.2. In Figure 3.3, the MSE performance is evaluated for the Rayleigh fading channel. In this case, frequency offset is set to vary uniformly between -0.5 and 0.5 of

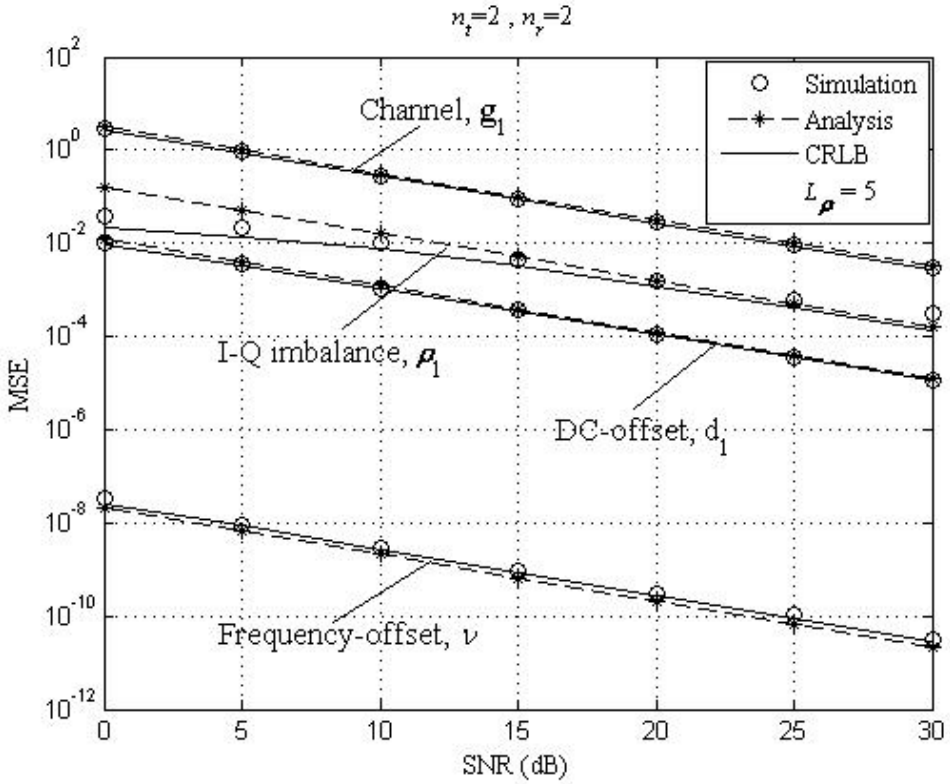


Figure 3.2: MSE performance in a static channel

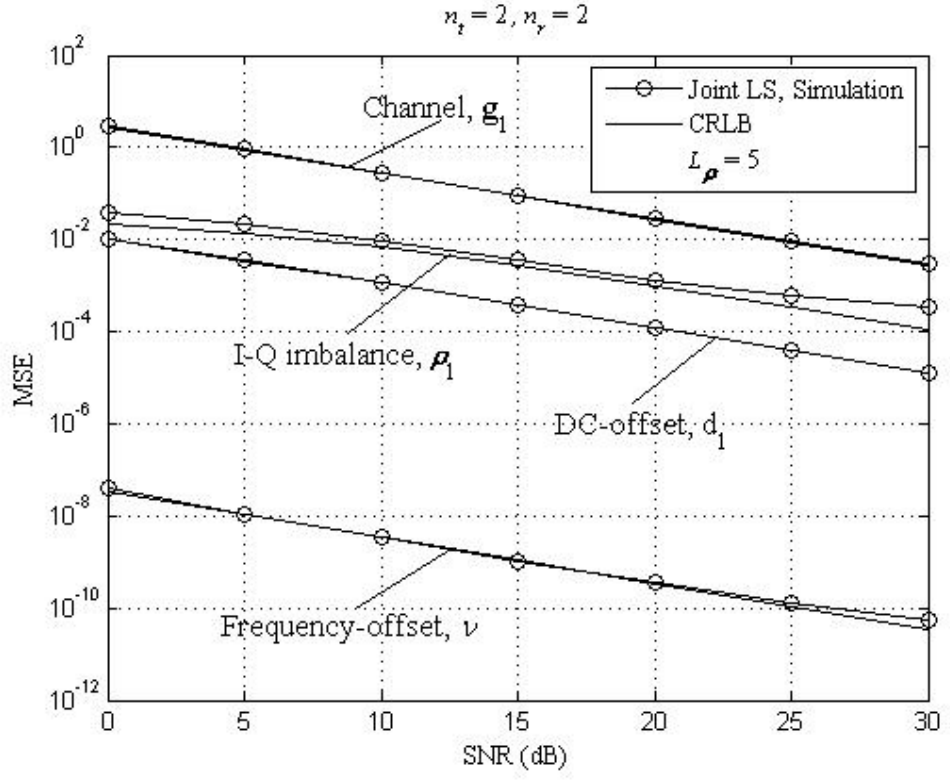


Figure 3.3: MSE performance of the joint estimators in Rayleigh fading channels

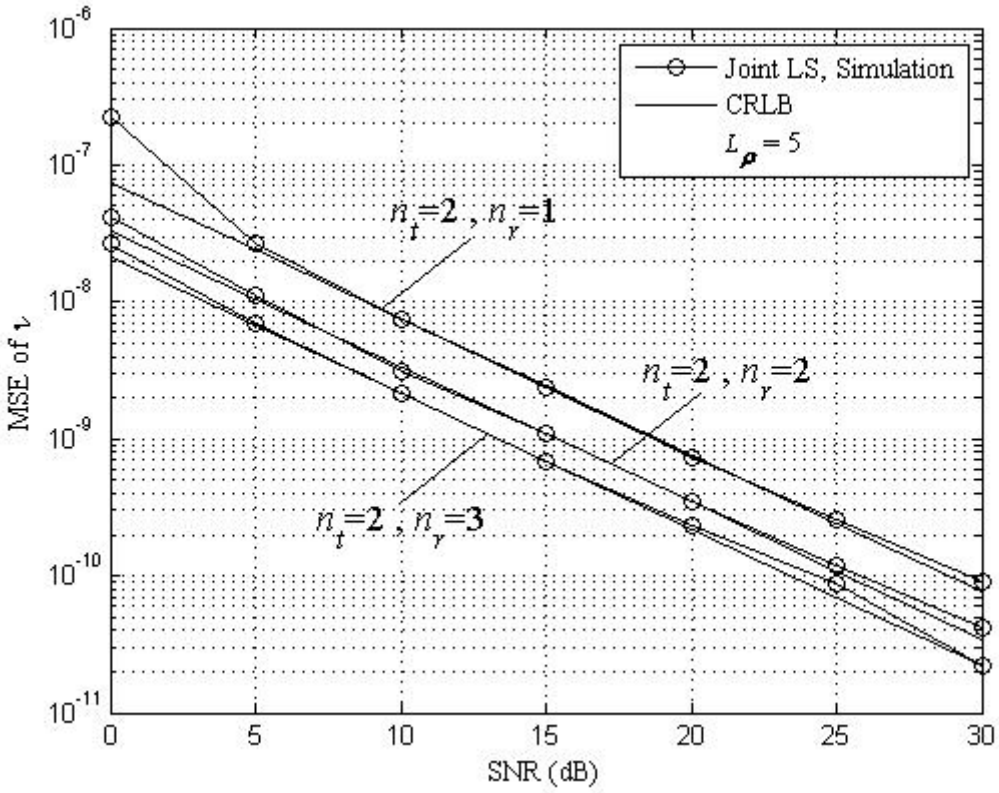


Figure 3.4: MSE performance of the frequency estimator in Rayleigh fading channels

sub-carrier spacing. Again, the estimators performed very closely to the CRLB bounds. Moreover, the variance of $\hat{\rho}$ tends to have a floor at high SNR region. This may be attributed to having modeling error by using $L_\rho = 5$ in this case. The error, however, does not cause too much loss in BER performance, as to be shown in Figure 3.6. The performance of frequency offset estimation is shown in Figure 3.4 with different number of received antennas. It is clearly shown that more than one receive antenna branch can be used in the estimation to improve performance by exploiting the power and diversity gain offered by multiple receive antennas.

Figure 3.5 investigates the effect of L_ρ on the MSE performance of the joint estimators by computer simulations. As can be seen, L_ρ has a significant impact on the channel and I-Q imbalance estimation, especially at high SNR region, and that causes error floor in BER performance as can be seen in Figure 3.6. Nevertheless, it affects the estimation of dc offset and fre

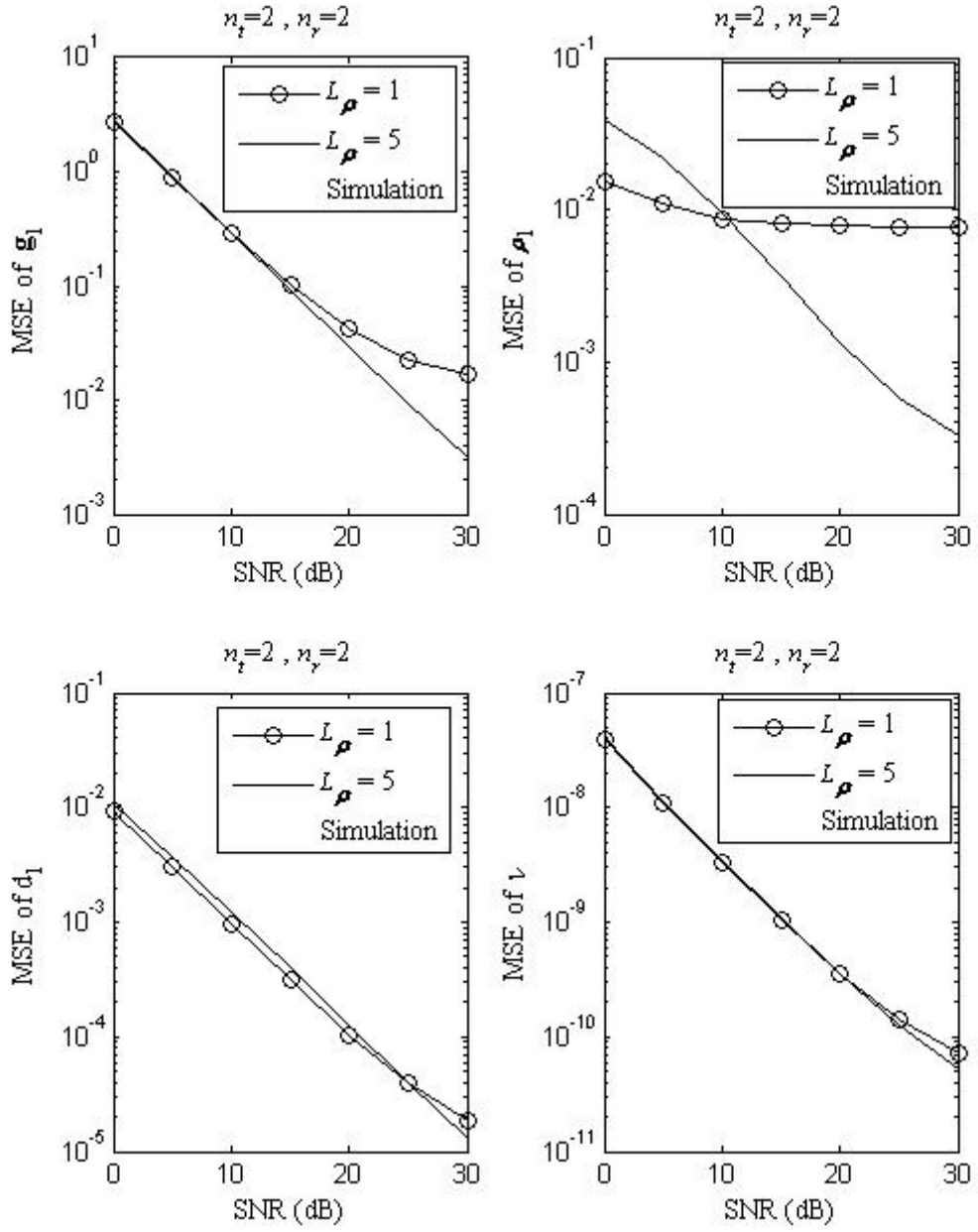


Figure 3.5: The effects of L_ρ on MSE in Rayleigh fading channels

quency offset in a very insignificant way; this motivates us to use the simplified estimators proposed in (3.42) and (3.44).

Figure 3.6 shows the impact of L_ρ on BER performance in the Rayleigh fading channel. 64-QAM is the modulation scheme. The receiver is the one given in Figure 3.1; after

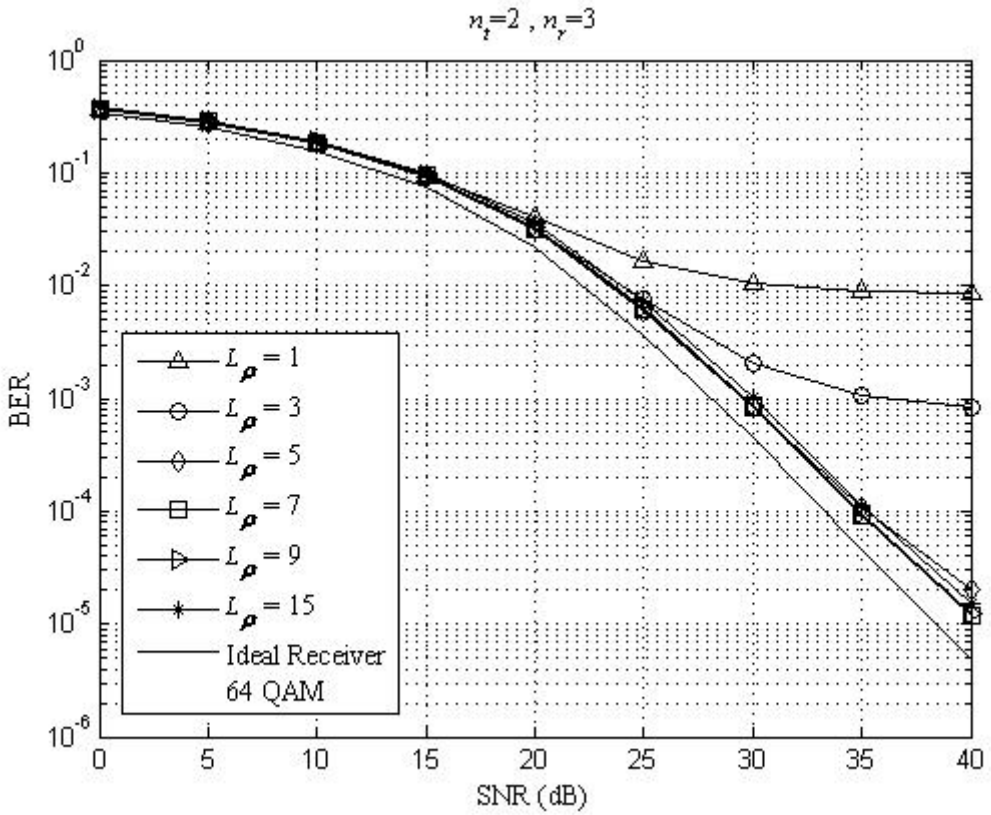


Figure 3.6: The effects of L_ρ on BER performance in Rayleigh fading channels

compensating I-Q imbalance, dc and frequency offset, MMSE MIMO detection is performed based on the channel estimation coming out of joint LS estimators. Clearly, the modeling error due to using a small L_ρ incurs error floor in BER performance, as predicted in Figure 3.6, where a small L_ρ results in large MSE for estimating the channel and I-Q imbalance. On the other hand, too large a L_ρ, e.g., $L_\rho = 15$ degrades slightly the BER performance as can be seen in the figure due to the extra noise induced by using a large filter length. In the figure, ideal receiver is the one with perfect RF compensation.

Finally, in Figure 3.7, we show the BER performance by using low-complexity training sequence and/or simplified frequency and dc offset estimators. The low-complexity training works very well and almost no performance loss is observed with the low-complexity implementations.

In fact, the low-complexity training performs a little better than the training used in [75] for $L_\rho = 3$.

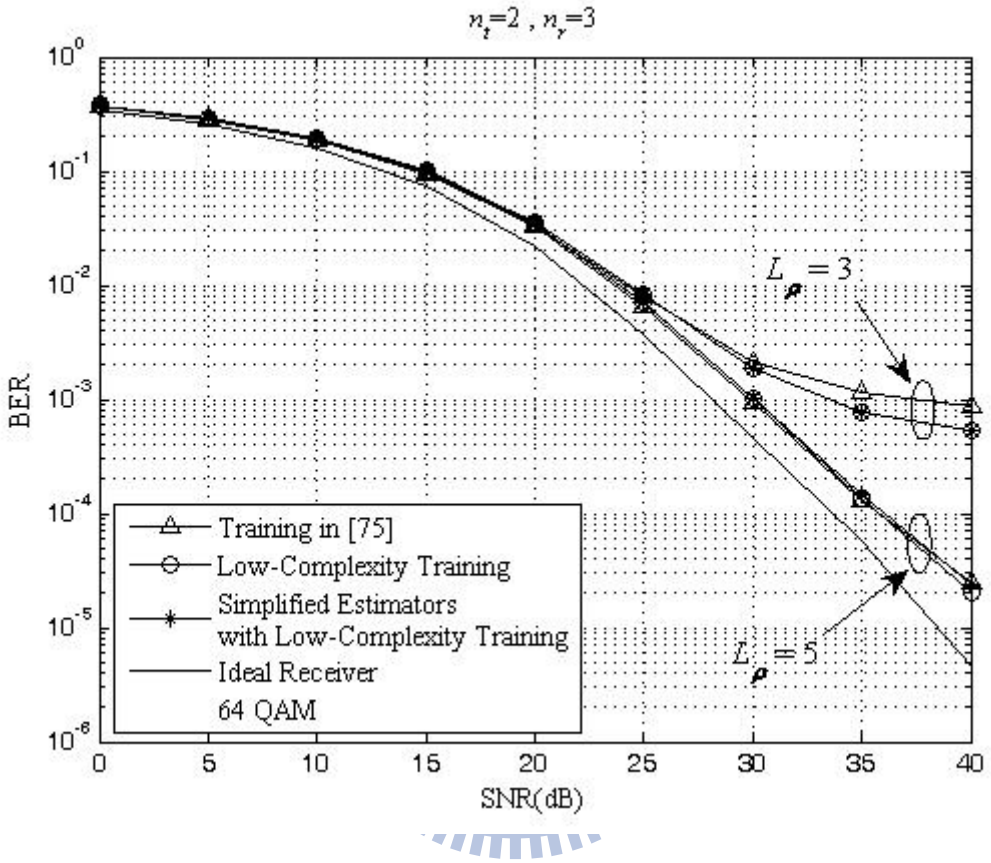


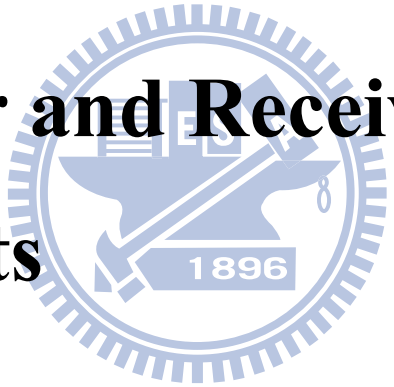
Figure 3.7: BER performance with low-complexity and/or simplified estimators in Rayleigh fading channels.

3.6 Summary

In Chapter 3, the theory of joint least-squares estimation of frequency, I-Q imbalance, dc offset and channel is developed for MIMO receivers with direct-conversion RF architecture. Both frequency-independent and dependent I-Q imbalances are included. Previously, RF parameters were estimated separately and that leads an inferior performance. The estimators are shown through analysis to be unbiased and approach to CRLB for the signal-to-noise ratios of interest. Special attention is paid to the implementation complexity issue; several measures are proposed including a special phase-rotated periodic training-sequence design and low-complexity estimators for frequency and dc offset. Simulation results show that the performance degradation is negligible when using the low-complexity designs.

Chapter 4

Estimation/Compensation Technique for Cascaded Transmitter and Receiver Radio Impairments



In this chapter, the cancellation technique of joint transmitter and receiver radio impairments is investigated for the MIMO-OFDM systems with direct conversion radio architecture. A two-stage cancellation architecture is proposed in Section 4.1. The cancellation technique is unique in that it is effective in different forms of MIMO operations including spatial multiplexing, STBC (space-time block coded) and transmit beam forming, with any number of transmit and receive antennas. In addition, two methods of radio parameters estimation are proposed in Section 4.2. The first is the optimum joint least-squares estimation of channel and radio impairments which can be seen as an extension of previous chapter. The second is a low-complexity iterative estimation that exploits the periodic structure of a training sequence. The proposed methods are simu-

lated and compared with existing methods in Section 4.3. Finally, conclusions are given in Section 4.4.

4.1 Two-Stage Cancellation of Radio Impairments

We propose a two-stage architecture for canceling the transmitter and receiver radio impairments. In the first stage, the receiver impairments such as frequency-independent and dependent I-Q imbalances and dc offset are cancelled along with frequency offset and transmitter dc offset in the time domain, whereas the transmitter frequency-independent and dependent I-Q imbalances are cancelled in the frequency domain in the second stage. The parameters used for cancellation are assumed to be perfect in this discussion, with their estimation to be discussed in the next section.

4.1.1 Time-Domain Cancellation (1st Stage)

We start from the received signal model (2.12) including cascaded effect of transmitter and receiver radio impairments in subsection 2.2.2. First, a time-domain filter $\rho_j(n)$ is introduced in *j*-th receive branch to cancel the receiver mirror-frequency interference as follows.

$$r_j(n)-\rho_j(n)\otimes r_j^*(n)=\tilde{c}_{R+,j}(n)\otimes\left[y_j(n)e^{j2\pi\nu n}+w_{0,j}(n)\right]+\tilde{c}_{R-,j}(n)\otimes\left[y_j(n)e^{j2\pi\nu n}+w_{0,j}(n)\right]^*+d_j, \tag{4.1}$$

where $\tilde{c}_{R\pm,j}(n)=c_{R\pm,j}(n)-\rho_j(n)\otimes c_{R\mp,j}^*(n)$, and $d_j=d_{0,j}-\rho_j(n)\otimes d_{0,j}^*$. Under a perfect cancellation, $\tilde{c}_{R-,j}(n)=0$, that is, $\rho_j(n)=\left(c_{R+,j}^*(n)\right)^{-1}\otimes c_{R-,j}(n)$, where $\left(c_{R+,j}^*(n)\right)^{-1}$ is the inverse filter of $c_{R+,j}^*(n)$. Then, (4.1) becomes

$$\begin{aligned} r_j(n)-\rho_j(n)\otimes r_j^*(n)&=\tilde{c}_{R+,j}(n)\otimes\left(y_j(n)e^{j2\pi\nu n}+w_{0,j}(n)\right)+d_j \\ &=e^{j2\pi\nu n}\left(\sum_{i=1}^{n_t}s_i(n)\otimes h_{+,j,i}(n)+s_i^*(n)\otimes h_{-,j,i}(n)+f_j\right)+d_j+w_j(n), \end{aligned} \tag{4.2}$$

where

$$h_{\pm,j,i}(n) = c_{T\pm,i}(n) \otimes h_{j,i}(n) \otimes \left(\tilde{c}_{R+,j}(n) e^{-j2\pi\nu n} \right), \tag{4.3}$$

$$f_j = \sum_{i=1}^{n_t} f_{1,i} \otimes h_{j,i}(n) \otimes \left(\tilde{c}_{R+,j}(n) e^{-j2\pi\nu n} \right), \tag{4.4}$$

and

$$w_j(n) = \tilde{c}_{R+,j}(n) \otimes w_{0,j}(n). \tag{4.5}$$

Here, $h_{+,j,i}(n)$ is the overall impulse response from transmit branch i to receive branch j consisting of transmit filter, channel, and receive filter after canceling out the receiver mirror-frequency interference, $h_{-,j,i}(n)$ is the overall impulse response due to the transmitter I-Q imbalance, f_j and d_j are the equivalent transmitter and receiver dc offsets, and $w_j(n)$ is the Gaussian noise after receiver I-Q cancellation. Following (4.2), the first stage time-domain cancellation is straightforward as shown in Figure 4.1: the receiver I-Q imbalance is cancelled first, followed by the receiver dc offset, the frequency offset, and lastly the transmitter dc offset. Clearly, $\{\rho_j(n)\}$, $\{f_j\}$, $\{d_j\}$, ν, are the parameters needed to be estimated for the first stage cancellation.

After cancellation, the received signal is

$$\begin{aligned} z_j(n) &= e^{-j2\pi\nu n}\left[r_j(n) - \rho_j(n) \otimes r_j^*(n) - d_j \right] - f_j \\ &= \sum_{i=1}^{n_t} \left[s_i(n) \otimes h_{+,j,i}(n) + s_i^*(n) \otimes h_{-,j,i}(n) \right] + \omega_j(n) \end{aligned}, \tag{4.6}$$

where $\omega_j(n) = e^{-j2\pi\nu n} w_j(n)$. (4.6) says that at this point the transmitter mirror-frequency interference is the only impairment left to be compensated. For later use, we denote $z_{j,k}(m)$, $m = 0, \cdots, N-1$ be the useful part of k-th OFDM symbol within the sequence $z_j(n)$, $n = \cdots, -1, 0, 1, \cdots$, after removing the cyclic prefix.

4.1.2 Frequency-Domain Cancellation (2nd Stage)

In this part, we consider the general MIMO structure of linear dispersion (LD) codes, which subsumes spatial multiplexing, STBC and transmit beam-forming as special cases and is applicable to any number of transmit and receive antennas [77]. Without loss of generality, the first code block that starts from the zero-th OFDM symbol is considered for notation simplicity. From [77], a set of n_s data symbols $\{V_m(l)\}_{m=1}^{n_s}$, which are to be transmitted on sub-carrier l over κ consecutive OFDM symbols, are encoded as a $\kappa \times n_t$ LD code matrix $\mathbf{S}(l)$ as follows

$$\begin{aligned}\mathbf{S}(l) &= \sum_{m=1}^{n_s}\left(\mathrm{Re}\{V_m(l)\}\mathbf{A}_m + \boldsymbol{j}\,\mathrm{Im}\{V_m(l)\}\mathbf{B}_m\right) \\ &\doteq \begin{bmatrix} S_{1,0}(l) & \cdots & S_{n_t,0}(l) \\ \vdots & \ddots & \vdots \\ S_{1,\kappa-1}(l) & \cdots & S_{n_t,\kappa-1}(l) \end{bmatrix},\end{aligned} \tag{4.7}$$

where $\mathbf{A}_m$ and $\mathbf{B}_m$ are $\kappa \times n_t$ complex-valued dispersion matrices that are designed to reap the diversity and/or degree of freedom gains of the MIMO channel. As shown in [77], with a proper selection of the dispersion matrices, spatial-multiplexing, STBC or transmit beam-forming can be viewed as a special case of (4.7). Recall that in our notation in (2.11), $\{S_{i,k}(l)\}_{l=0}^{N-1}$ is the data symbols input to the IDFT of transmit branch i at OFDM k. Denote $\{Z_{j,k}(l)\}_{l=0}^{N-1} = \mathrm{DFT}\left[\{z_{j,k}(n)\}_{n=0}^{N-1}\right]$. From (4.6),

$$Z_{j,k}(l) = \sum_{i=1}^{n_t}\left[H_{+,j,i}(l)S_{i,k}(l) + H_{-,j,i}(l)S_{i,k}^{*}(-l)\right] + \Omega_{j,k}(l),\ j = 1\cdots n_r,\ k = 0\cdots\kappa-1, \tag{4.8}$$

where $\{H_{\pm,j,i}(l)\}_{l=0}^{N-1} = \mathrm{DFT}\left[\{h_{\pm,j,i}(n)\}\right]$, and $\{\Omega_{j,k}(l)\}_{l=0}^{N-1} = \mathrm{DFT}\left[\{\omega_{j,k}(n)\}\right]$. In addition, using a matrix form for those $Z_{j,k}(l)$ corresponding to the LD code matrix $\mathbf{S}(l)$, we have

$$\mathbf{Z}(l) = \mathbf{S}(l)\mathbf{G}_{+}(l) + \mathbf{S}^{*}(-l)\mathbf{G}_{-}(l) + \boldsymbol{\Omega}(l), \tag{4.9}$$

where

$$\mathbf{Z}(l)=\begin{bmatrix} Z_{1,0}(l) & \cdots & Z_{n_r,0}(l) \\ \vdots & \ddots & \vdots \\ Z_{1,\kappa-1}(l) & \cdots & Z_{n_r,\kappa-1}(l) \end{bmatrix},$$

$$\mathbf{G}_{\pm}(l)=\begin{bmatrix} H_{\pm,1,1}(l) & \cdots & H_{\pm,n_r1}(l) \\ \vdots & \cdots & \vdots \\ H_{\pm,1,n_t}(l) & \cdots & H_{\pm,n_r,n_t}(l) \end{bmatrix},$$

and

$$\mathbf{\Omega}(l)=\begin{bmatrix} \Omega_{1,0}(l) & \cdots & \Omega_{n_r,0}(l) \\ \vdots & \ddots & \vdots \\ \Omega_{1,\kappa-1}(l) & \cdots & \Omega_{n_r,\kappa-1}(l) \end{bmatrix}.$$

Furthermore, it is convenient to rewrite (4.9) in the form of real matrices and vectors. Let $\mathbf{z}_p(l)$, $\mathbf{g}_{+,p}(l)$, $\mathbf{g}_{-,p}(l)$ and $\boldsymbol{\omega}_p(l)$ denote the p-th column of $\mathbf{Z}(l)$, $\mathbf{G}_+(l)$, $\mathbf{G}_-(l)$ and $\mathbf{\Omega}(l)$, respectively, where $p=1,\ldots,n_r$, and define

$$\tilde{\mathbf{z}}(l)=\left[\mathrm{Re}\left\{\mathbf{z}_1^T(l)\right\}\mathrm{Im}\left\{\mathbf{z}_1^T(l)\right\}\ldots\mathrm{Re}\left\{\mathbf{z}_{n_r}^T(l)\right\}\mathrm{Im}\left\{\mathbf{z}_{n_r}^T(l)\right\}\right]^T,$$

$$\tilde{\mathbf{v}}(l)=\left[\mathrm{Re}\left\{V_1(l)\right\}\mathrm{Im}\left\{V_1(l)\right\}\ldots\mathrm{Re}\left\{V_{n_s}(l)\right\}\mathrm{Im}\left\{V_{n_s}(l)\right\}\right]^T,$$

$$\tilde{\mathbf{g}}_{\pm,j}(l)=\left[\mathrm{Re}\left\{\mathbf{g}_{\pm,j}^T(l)\right\}\mathrm{Im}\left\{\mathbf{g}_{\pm,j}^T(l)\right\}\right]^T,$$

$$\widetilde{\mathbf{A}}_{\pm,m}=\begin{bmatrix} \mathrm{Re}\{\mathbf{A}_m\} & \mp\mathrm{Im}\{\mathbf{A}_m\} \\ \pm\mathrm{Im}\{\mathbf{A}_m\} & \mathrm{Re}\{\mathbf{A}_m\} \end{bmatrix},$$

$$\widetilde{\mathbf{B}}_{\pm,m}=\begin{bmatrix} -\mathrm{Im}\{\mathbf{B}_m\} & \mp\mathrm{Re}\{\mathbf{B}_m\} \\ \pm\mathrm{Re}\{\mathbf{B}_m\} & -\mathrm{Im}\{\mathbf{B}_m\} \end{bmatrix},$$

and

$$\tilde{\boldsymbol{\omega}}(l)=\left[\mathrm{Re}\left\{\boldsymbol{\omega}_1^T(l)\right\}\mathrm{Im}\left\{\boldsymbol{\omega}_1^T(l)\right\}\ldots\mathrm{Re}\left\{\boldsymbol{\omega}_{n_r}^T(l)\right\}\mathrm{Im}\left\{\boldsymbol{\omega}_{n_r}^T(l)\right\}\right]^T.$$

It can be shown that

$$\tilde{\mathbf{z}}(l)=\widetilde{\mathbf{G}}_+(l)\tilde{\mathbf{v}}(l)+\widetilde{\mathbf{G}}_-(l)\tilde{\mathbf{v}}(-l)+\tilde{\boldsymbol{\omega}}(l), \tag{4.10}$$

and

$$\tilde{\mathbf{z}}(-l)=\widetilde{\mathbf{G}}_{+}(-l)\tilde{\mathbf{v}}(-l)+\widetilde{\mathbf{G}}_{-}(-l)\tilde{\mathbf{v}}(l)+\tilde{\boldsymbol{\omega}}(-l)\ , \tag{4.11}$$

where

$$\widetilde{\mathbf{G}}_{\pm}(l)=\begin{bmatrix}\widetilde{\mathbf{A}}_{\pm,1}\tilde{\mathbf{g}}_{\pm,1}(l) & \widetilde{\mathbf{B}}_{\pm,1}\tilde{\mathbf{g}}_{\pm,1}(l) & \cdots & \widetilde{\mathbf{A}}_{\pm,n_s}\tilde{\mathbf{g}}_{\pm,1}(l) & \widetilde{\mathbf{B}}_{\pm,n_s}\tilde{\mathbf{g}}_{\pm,1}(l)\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \widetilde{\mathbf{A}}_{\pm,1}\tilde{\mathbf{g}}_{\pm,n_r}(l) & \widetilde{\mathbf{B}}_{\pm,1}\tilde{\mathbf{g}}_{\pm,n_r}(l) & \cdots & \widetilde{\mathbf{A}}_{\pm,n_s}\tilde{\mathbf{g}}_{\pm,n_r}(l) & \widetilde{\mathbf{B}}_{\pm,n_s}\tilde{\mathbf{g}}_{\pm,n_r}(l)\end{bmatrix}. \tag{4.12}$$

From (4.10) and (4.11), the mirror-frequency interference cancellation is proposed by introducing the cancellation matrix $\boldsymbol{\Phi}(l)$ as follows

$$\begin{aligned}\tilde{\mathbf{z}}(l)-\boldsymbol{\Phi}(l)\tilde{\mathbf{z}}(-l)&=\underbrace{\left(\widetilde{\mathbf{G}}_{+}(l)-\boldsymbol{\Phi}(l)\widetilde{\mathbf{G}}_{-}(-l)\right)}_{\text{effective channel after cancellation}}\tilde{\mathbf{v}}(l)+\underbrace{\left(\widetilde{\mathbf{G}}_{-}(l)-\boldsymbol{\Phi}(l)\widetilde{\mathbf{G}}_{+}(-l)\right)}_{\substack{=\mathbf{0}_{2\kappa n_r\times 2n_s}\\ \text{to cancel mirror-frequency interference}}}\tilde{\mathbf{v}}(-l)+\tilde{\boldsymbol{\omega}}(l)-\boldsymbol{\Phi}(l)\tilde{\boldsymbol{\omega}}(-l)\\ &=\mathbf{H}(l)\tilde{\mathbf{v}}(l)+\boldsymbol{\omega}(l)\ ,\end{aligned} \tag{4.13}$$

where

$$\mathbf{H}(l)=\widetilde{\mathbf{G}}_{+}(l)-\boldsymbol{\Phi}(l)\widetilde{\mathbf{G}}_{-}(-l) \tag{4.14}$$

and

$$\boldsymbol{\omega}(l)=\tilde{\boldsymbol{\omega}}(l)-\boldsymbol{\Phi}(l)\tilde{\boldsymbol{\omega}}(-l) \tag{4.15}$$

are the effective channel and noise after cancellation, respectively. Note that the dimension of $\widetilde{\mathbf{G}}_{+}(l)$ and $\widetilde{\mathbf{G}}_{-}(l)$ is $2\kappa n_r\times 2n_s$, which is a tall or square matrix, i.e. $\kappa n_r\geq n_s$ so as to avoid a underdetermined case. Clearly, to completely cancel out the mirror-frequency interference, we need to have

$$\widetilde{\mathbf{G}}_{-}(l)-\boldsymbol{\Phi}(l)\widetilde{\mathbf{G}}_{+}(-l)=\mathbf{O}_{2\kappa n_r\times 2n_s}\ , \tag{4.16}$$

For the case $\kappa n_r=n_s$, (4.16) has a unique solution of $\hat{\boldsymbol{\Phi}}(l)=\widetilde{\mathbf{G}}_{-}(l)\widetilde{\mathbf{G}}_{+}^{-1}(-l)$. For the case $\kappa n_r>n_s$, however, there are infinite solutions. Naturally, the one with minimum noise power

$\mathrm{E}\left\{\left\|\boldsymbol{\omega}(l)\right\|^2\right\}$ is the desired one. That is, the optimal $\widehat{\boldsymbol{\Phi}}(l)$ is obtained by solving the following constrained optimization problem.

$$\widehat{\boldsymbol{\Phi}}(l)=\underset{\widetilde{\boldsymbol{\Phi}}(l)}{\arg\min}\,\mathrm{E}\left\{\left\|\boldsymbol{\omega}(l)\right\|^2\right\}\quad,\quad \text{subject to } \widetilde{\mathbf{G}}_{-}(l)-\widetilde{\boldsymbol{\Phi}}(l)\widetilde{\mathbf{G}}_{+}(-l)=\mathbf{O}_{2\kappa n_r\times 2n_s}. \tag{4.17}$$

From (4.15),

$$\begin{aligned}\mathrm{E}\left\{\left\|\boldsymbol{\omega}(l)\right\|^2\right\}&=tr\left\{\mathrm{E}\left\{\left(\tilde{\boldsymbol{\omega}}(l)-\boldsymbol{\Phi}(l)\tilde{\boldsymbol{\omega}}(-l)\right)\left(\tilde{\boldsymbol{\omega}}(l)-\boldsymbol{\Phi}(l)\tilde{\boldsymbol{\omega}}(-l)\right)^T\right\}\right\}\\&\approx 2\kappa n_r\sigma^2+\sigma^2\cdot tr\left\{\boldsymbol{\Phi}^T(l)\boldsymbol{\Phi}(l)\right\}.\end{aligned} \tag{4.18}$$

Here we have used the approximations, $\mathrm{E}\left\{\tilde{\boldsymbol{\omega}}(l)\tilde{\boldsymbol{\omega}}^T(l)\right\}\approx\sigma^2\mathbf{I}_{2\kappa n_r}$ and $\mathrm{E}\left\{\tilde{\boldsymbol{\omega}}(l)\tilde{\boldsymbol{\omega}}^T(-l)\right\}$ $\approx\mathbf{O}_{2\kappa n_r\times 2\kappa n_r}$. Therefore, the constrained optimization problem (4.17) becomes

$$\widehat{\boldsymbol{\Phi}}(l)=\underset{\widetilde{\boldsymbol{\Phi}}(l)}{\arg\min}\,tr\left\{\widetilde{\boldsymbol{\Phi}}^T(l)\widetilde{\boldsymbol{\Phi}}(l)\right\},\quad \text{subject to } \widetilde{\mathbf{G}}_{-}(l)-\widetilde{\boldsymbol{\Phi}}(l)\widetilde{\mathbf{G}}_{+}(-l)=\mathbf{O}_{2\kappa n_r\times 2n_s}. \tag{4.19}$$

The problem amounts to solve the minimum norm solution of the linear equations of $\widetilde{\mathbf{G}}_{+}^T(-l)\widetilde{\boldsymbol{\Phi}}^T(l)=\widetilde{\mathbf{G}}_{-}^T(l)$ for $\widetilde{\boldsymbol{\Phi}}^T(l)$ which is given by [78]

$$\widehat{\boldsymbol{\Phi}}(l)=\left[\widetilde{\mathbf{G}}_{+}(-l)\left(\widetilde{\mathbf{G}}_{+}^T(-l)\widetilde{\mathbf{G}}_{+}(-l)\right)^{-1}\widetilde{\mathbf{G}}_{-}^T(l)\right]^T. \tag{4.20}$$

After the transmitter mirror-frequency interference is cancelled, (4.13) is used to detect $\tilde{\mathbf{v}}(l)$. Any type of MIMO detectors can be used, although the linear MMSE (minimum mean-squared error) detector is considered exclusively in Section 4.3. The overall two-stage cancellation architecture is presented in Figure 4.1. From (4.9) and (4.12), it is noted that $\left\{h_{\pm,j,i}(n)\right\}$ are the parameters needed to be estimated for the second stage cancellation.

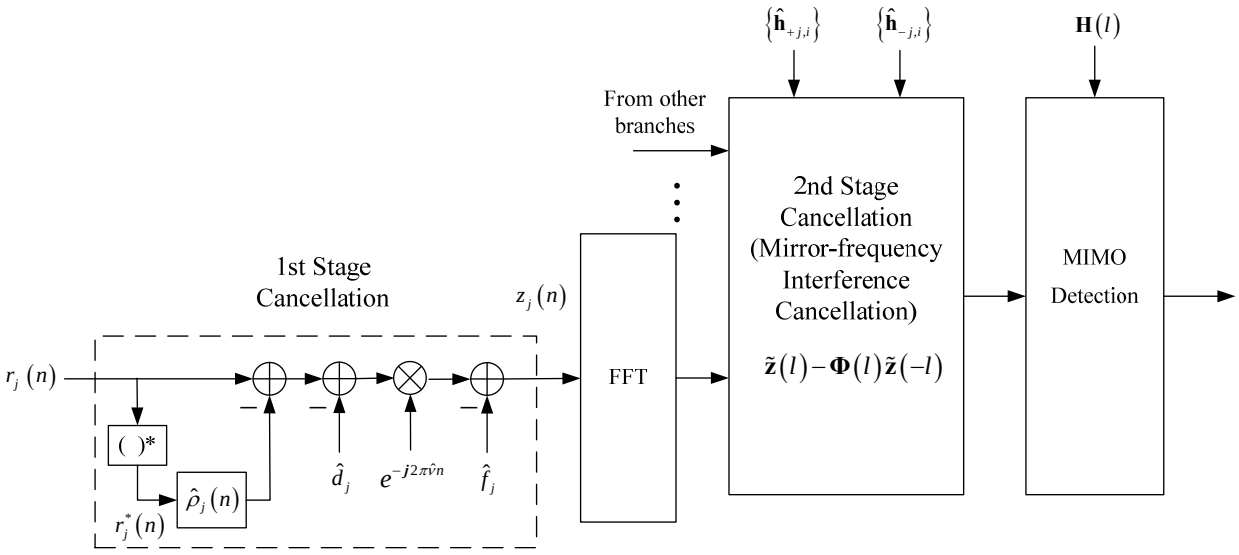


Figure 4.1: The two-stage cancellation architecture for cascaded transmitter and receiver radio impairments.

4.1.3 Summary of Cancellation Architectures for Different Application Configurations

Now, we summarize the cancellation architectures for different configurations. In this chapter, the generalized two-stage cancellation architecture is developed as shown in Figure 4.1 for cancellation of the cascaded transmitter and receiver radio impairments in the application of wireless peer-to-peer MIMO-OFDM communications. For the application configuration such as the downlink of mobile cellular MIMO systems as considered in Chapter 3, the transmitter radio impairments are absence as seen in (2.9), and the cancellation architecture is a special case of the generalized two-stage cancellation architecture, which degenerates to first-stage cancellation without transmitter dc offset cancellation as shown in Figure 4.2. For the application configuration such as the uplink of mobile cellular MIMO-OFDM systems, on the other hand, the receiver radio impairments are negligible except for frequency offset as seen in (2.17), and the cancellation architecture degenerates to a special case of the generalized two-stage cancellation architec-

ture, where the first-stage cancellation reduces to frequency offset and transmitter dc offset cancellation as shown in Figure 4.3.

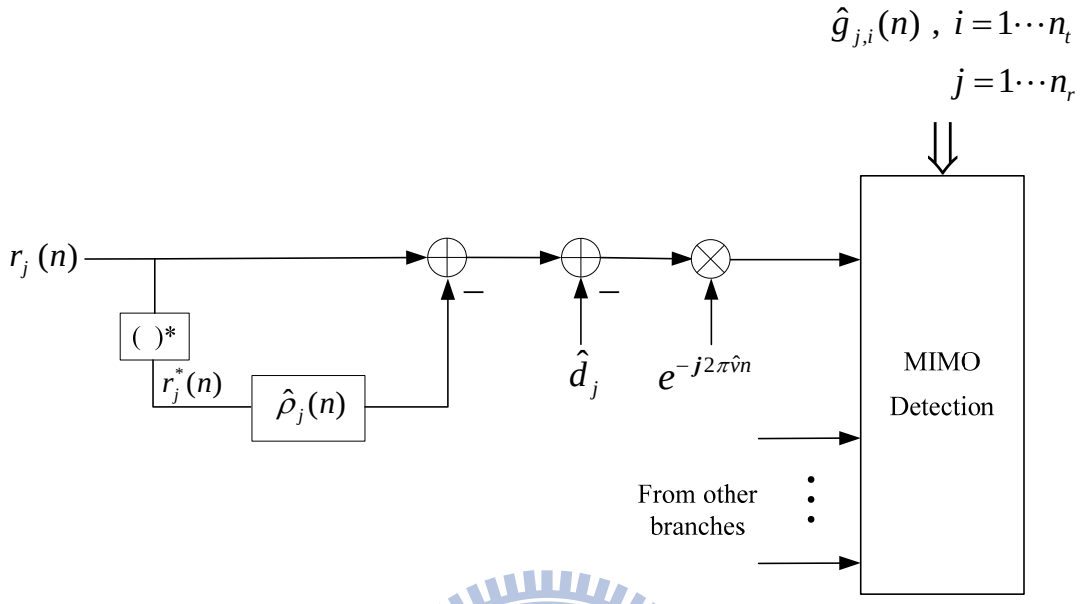


Figure 4.2: The first-stage cancellation architecture for receiver radio impairments.

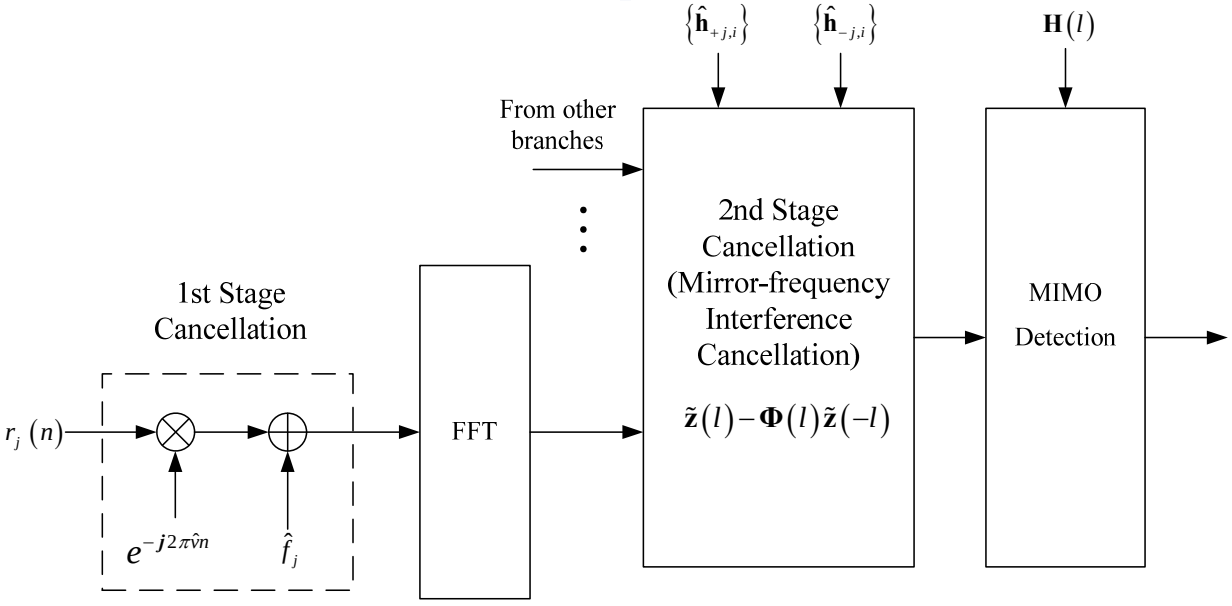


Figure 4.3: The two-stage cancellation architecture for transmitter radio impairments.

4.2 Estimation of Radio Parameters

In this section, two methods are developed for the estimation of $\{\rho_j(n)\}$, $\{f_j\}$, $\{d_j\}$, ν and $\{h_{\pm,j,i}(n)\}$. The first one is an optimum joint estimation of all parameters under the least-squares (LS) principle, which can be seen as an extension of previous chapter, and the second is a low-complexity estimation by exploiting a periodic structure of training sequence.

4.2.1 Joint LS Estimation (JLSE)

Here, a joint estimation scheme is developed based on the LS principle. That is, the optimum estimates $\{\hat{\rho}_j(n)\}$, $\{\hat{f}_j\}$, $\{\hat{d}_j\}$, $\hat{\nu}$ and $\{\hat{h}_{\pm,j,i}(n)\}$ are sought to minimize the square error of $\{w_j(n)\}$ in (4.2). To begin with, $\rho_j(n)$ and $h_{\pm,j,i}(n)$ are modeled as finite impulse response (FIR) filters; $\boldsymbol{\rho}_j = [\rho_j(0),\ldots,\rho_j(L_\rho-1)]^T$, and $\mathbf{h}_{\pm,j,i} = [h_{\pm,j,i}(0),\ldots,h_{\pm,j,i}(L_\pm-1)]^T$, where L_ρ and $L_\pm$ are the corresponding filter's lengths.

Assuming a total of P OFDM training symbols, indexed from $k=-P,\ldots,-1$, are available for the parameters estimation. Let $\mathbf{r}_j(k) = [r_{j,k}(0), r_{j,k}(1),\ldots,r_{j,k}(N-1)]^T$ with $r_{j,k}(n) = r_j(k(N_g+N)+n)$ be the useful received data part of OFDM symbol k, $\mathbf{R}_j(k)$ be the $N\times L_\rho$ received signal matrix with $[\mathbf{R}_j(k)]_{p,q} = r_{j,k}(p-q)$, $0\le p\le N-1$, $0\le q\le L_\rho-1$, and $\mathbf{T}_{\pm,i}(k)$ be the $N\times L_\pm$ signal matrices with $[\mathbf{T}_{\pm,i}(k)]_{p,q} = s_{i,k}(p-q)$, $0\le p\le N-1$, $0\le q\le L_{\boldsymbol{h}_\pm}-1$. Then for $k=-P,\ldots,-1$, (4.2) can be written in the vector-matrix form

$$\mathbf{r}_j(k) - \mathbf{R}^*_j(k)\boldsymbol{\rho}_j = \boldsymbol{\Gamma}_k(\nu)\left[\left(\sum_{i=1}^{n_t}\mathbf{T}_{+,i}(k)\mathbf{h}_{+,j,i} + \mathbf{T}^*_{-,i}(k)\mathbf{h}_{-,j,i}\right) + f_j\mathbf{1}_N\right] + d_j\mathbf{1}_N + \mathbf{w}_j(k), \quad (4.21)$$

where $\boldsymbol{\Gamma}_k(\nu) = e^{j2\pi k(N_g+N)\nu}\cdot\mathrm{diag}\{1, e^{j2\pi\nu},\cdots,e^{j2\pi\nu(N-1)}\}$ and $\mathbf{w}_j(k) = [w_{j,k}(0)\ldots w_{j,k}(N-1)]^T$

with $w_{j,k}(n) = w_j\left(k\left(N_g + N\right) + n\right)$. Furthermore, define

$$\mathbf{r}_j = \left[\mathbf{r}_j^T(-P), \mathbf{r}_j^T(-P+1), \ldots, \mathbf{r}_j^T(-1)\right]^T,$$

$$\mathbf{R}_j = \left[\mathbf{R}_j^T(-P), \mathbf{R}_j^T(-P+1), \ldots, \mathbf{R}_j^T(-1)\right]^T,$$

$$\mathbf{T}(k) = \left[\mathbf{T}_{+,1}(k), \ldots, \mathbf{T}_{+,n_t}(k), \mathbf{T}_{-,1}^*(k), \ldots, \mathbf{T}_{-,n_t}^*(k), \mathbf{1}_N\right],$$

$$\mathbf{h}_{\pm,j} = \left[\mathbf{h}_{\pm,j,1}^T, \mathbf{h}_{\pm,j,2}^T, \ldots, \mathbf{h}_{\pm,j,n_t}^T\right]^T,$$

$$\mathbf{h}_j = \left[\mathbf{h}_{+,j}^T, \mathbf{h}_{-,j}^T, f_j\right]^T,$$

and

$$\mathbf{w}_j = \left[\mathbf{w}_j^T(-P), \mathbf{w}_j^T(-P+1), \ldots, \mathbf{w}_j^T(-1)\right]^T.$$

Then, we have

$$\mathbf{r}_j - \mathbf{R}_j^* \boldsymbol{\rho}_j = \boldsymbol{\Gamma}(\nu)\mathbf{T}\mathbf{h}_j + \hat{d}_j \mathbf{1}_{PN} + \mathbf{w}_j, \tag{4.22}$$

where

$$\mathbf{T} = \left[\mathbf{T}^T(-P), \mathbf{T}^T(-P+1), \ldots, \mathbf{T}^T(-1)\right]^T,$$

and

$$\boldsymbol{\Gamma}(\nu) = \operatorname{diag}\left\{\boldsymbol{\Gamma}_{-P}(\nu), \boldsymbol{\Gamma}_{-P+1}(\nu), \ldots, \boldsymbol{\Gamma}_{-1}(\nu)\right\}.$$

From (4.22), the joint least-squares estimates are given by

$$\left\{\hat{\nu}, \{\hat{\boldsymbol{\rho}}_j\}, \{\hat{d}_j\}, \{\hat{\mathbf{h}}_j\}\right\} = \arg \min_{\tilde{\nu}, \{\tilde{\boldsymbol{\rho}}_j\}, \{\tilde{d}_j\}, \{\tilde{\mathbf{h}}_j\}} \Lambda\left(\tilde{\nu}, \{\tilde{\boldsymbol{\rho}}_j\}, \{\tilde{d}_j\}, \{\tilde{\mathbf{h}}_j\}\right), \tag{4.23}$$

where

$$\Lambda\left(\tilde{\nu}, \{\tilde{\boldsymbol{\rho}}_j\}, \{\tilde{d}_j\}, \{\tilde{\mathbf{h}}_j\}\right) = \sum_{j=1}^{n_r} \Lambda_j\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_j, \tilde{d}_j, \tilde{\mathbf{h}}_j\right) \tag{4.24}$$

and

$$\Lambda_j\left(\tilde{\nu}, \tilde{\boldsymbol{\rho}}_j, \tilde{d}_j, \tilde{\mathbf{h}}_j\right) = \left\|\mathbf{r}_j - \mathbf{R}_j^* \tilde{\boldsymbol{\rho}}_j - \tilde{d}_j \mathbf{1}_{PN} - \boldsymbol{\Gamma}(\tilde{\nu})\mathbf{T}\tilde{\mathbf{h}}_j\right\|^2. \tag{4.25}$$

Note that each receive branch has its own I-Q imbalance, dc offset and channel response. There-

fore, given a fixed trial frequency offset $\tilde{\nu}$, $\boldsymbol{\rho}_j$, d_j, and $\mathbf{h}_j$ can be estimated by minimizing respective cost function $\Lambda_j\left(\tilde{\nu},\tilde{\boldsymbol{\rho}}_j,\tilde{d}_j,\tilde{\mathbf{h}}_j\right)$. In other words, $\boldsymbol{\rho}_j$, d_j, and $\mathbf{h}_j$ can be estimated separately form one branch to another, and that reduces complexity. On the other hand, (4.23) can be used to estimate jointly the frequency offset to increase performance by exploiting the diversity and power gains that are inherent in MIMO systems. Since the optimization problem given in (4.23)-(4.25) carries the same form as that in Section 3.2.1 of previous chapter, the recursive optimization procedure is omitted here (refer to Section 3.2.1 for detail) and the joint LS estimates are summarized as follows.

$$\hat{\nu}=\arg\min_{\tilde{\nu}}\left\{\Lambda\left(\tilde{\nu}\right)=\sum_{j=1}^{n_r}\Lambda_j\left(\tilde{\nu}\right)\right\}, \tag{4.26}$$

$$\Lambda_j\left(\tilde{\nu}\right)=\mathbf{r}_j^H\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j-\left(\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j\right)^H\left[\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}\left(\tilde{\nu}\right)\mathbf{r}_j, \tag{4.27}$$

$$\mathbf{Q}\left(\tilde{\nu}\right)=\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)-\frac{\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\mathbf{1}_{PN}\cdot\left[\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\mathbf{1}_{PN}\right]^H}{\left\|\left(\mathbf{I}_{PN}-\mathbf{C}\left(\tilde{\nu}\right)\right)\mathbf{1}_{PN}\right\|^2}, \tag{4.28}$$

$$\hat{\boldsymbol{\rho}}_j=\left[\mathbf{R}_j^T\mathbf{Q}\left(\hat{\nu}\right)\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}\left(\hat{\nu}\right)\mathbf{r}_j, \tag{4.29}$$

$$\hat{d}_j=\frac{\mathbf{1}_{PN}^H\left(\mathbf{I}_{PN}-\mathbf{C}\left(\hat{\nu}\right)\right)}{\left\|\left(\mathbf{I}_{PN}-\mathbf{C}\left(\hat{\nu}\right)\right)\mathbf{1}_{PN}\right\|^2}\cdot\left(\mathbf{r}_j-\mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j\right), \tag{4.30}$$

and

$$\hat{\mathbf{h}}_j=\left(\mathbf{T}^H\mathbf{T}\right)^{-1}\mathbf{T}^H\boldsymbol{\Gamma}^H\left(\hat{\nu}\right)\left(\mathbf{r}_j-\mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j-\hat{d}_j\mathbf{1}_{PN}\right), \tag{4.31}$$

where $\mathbf{C}\left(\tilde{\nu}\right)=\boldsymbol{\Gamma}\left(\tilde{\nu}\right)\mathbf{T}\left(\mathbf{T}^H\mathbf{T}\right)^{-1}\mathbf{T}^H\boldsymbol{\Gamma}^H\left(\tilde{\nu}\right)$ is projection matrix. This part of development can be considered as an extension of Section 3.2.1, where only the receiver-side impairments are considered.

4.2.2 Low-Complexity Estimation with Periodic Training (LCE-PT)

As is to be shown in the next subsection, the complexity of the joint estimation is rather high due to that all parameters need to be estimated jointly, and an exhaustive search has to be performed for the estimation of frequency offset in (4.26). In this subsection, a low-complexity LS estimation is proposed by exploiting the periodic structure of the training sequence.

Consider a periodic training sequence consisting of $Q+2$ periods with J samples in each period, i.e., $s_i(n)=s_i(n+J)$, $n=-J,\ldots,M-1$, where $M=JQ$ From (4.2), define $r_{c,j}(n)$ be the signal after receiver I-Q imbalance and dc offset compensation:

$$\begin{aligned} r_{c,j}(n) &= r_j(n)-\rho_j(n)\otimes r_j^*(n)-d_j \\ &= e^{j2\pi\nu n}u(n)+w_j(n) \end{aligned} \tag{4.32}$$

where

$$u(n)=\sum_{i=1}^{n_t} s_i(n)\otimes h_{+,j,i}(n)+s_i^*(n)\otimes h_{-,j,i}(n)+f_j\,. \tag{4.33}$$

Assuming that the equivalent channel length $L_{\pm}$ is shorter than the period J, i.e. $L_{\pm}<J$, then $u(n)$ is also a periodic signal with $u(n)=u(n+J)$, $n=0,\cdots,M-1$. And, without considering noise,

$$r_{c,j}(n+J)=e^{j2\pi\nu J}r_{c,j}(n),\quad n=0,\ldots,M-1. \tag{4.34}$$

Define $\mathbf{r}_{1,j}=\left[r_j(0),\ldots,r_j(M-1)\right]^T$, $\mathbf{r}_{2,j}=\left[r_j(J),\ldots,r_j(M+J-1)\right]^T$, $\mathbf{R}_{1,j}$ and $\mathbf{R}_{2,j}$ as the $M\times L_\rho$ received signal matrices with the entry $\left[\mathbf{R}_{1,j}\right]_{p,q}=r_j(p-q)$ and $\left[\mathbf{R}_{2,j}\right]_{p,q}=r_j(p-q+J)$, $0\le p\le M-1$, $0\le q\le L_\rho-1$. Using (4.32), (4.34) can be written in vector-matrix form as

$$\mathbf{r}_{2,j}-\mathbf{R}_{2,j}^*\boldsymbol{\rho}_j-d_j\mathbf{1}_M=e^{j2\pi\nu J}\left(\mathbf{r}_{1,j}-\mathbf{R}_{1,j}^*\boldsymbol{\rho}_j-d_j\mathbf{1}_M\right), \tag{4.35}$$

and

$$\mathbf{r}_{2,j} - e^{j2\pi\nu J}\mathbf{r}_{1,j} - \left(\mathbf{R}^*_{2,j} - e^{j2\pi\nu J}\mathbf{R}^*_{1,j}\right)\boldsymbol{\rho}_j - \left(1 - e^{j2\pi\nu J}\right)\mathbf{1}_M d_j = \mathbf{0}_M , \tag{4.36}$$

Based on (4.35) and (4.36), a suboptimal iterative estimation is proposed as follows. The iterative principle is that given a frequency offset estimate $\hat{\nu}$, (4.36) can be used to estimate $\boldsymbol{\rho}_j$ and d_j, while given estimates $\hat{\boldsymbol{\rho}}_j$ and $\hat{d}_j$, (4.35) can be used to estimate the frequency offset ν.

First, an initial estimation of ν is obtained by

$$\hat{\nu}_1 = \frac{\arg\left\{\sum_{j=1}^{n_r} \mathbf{r}^H_{2,j}\mathbf{r}_{1,j}\right\}}{2\pi J}. \tag{4.37}$$

The estimate $\hat{\nu}_1$ is then used in (4.36) for the joint estimation of $\boldsymbol{\rho}_j$ and d_j. From (4.36) and given the initial estimate $\hat{\nu}_1$, $\boldsymbol{\rho}_j$ and d_j can be estimated jointly by minimizing the following cost function:

$$\Lambda_j\left(\tilde{\boldsymbol{\rho}}_j, \tilde{d}_j \middle| \hat{\nu}_1\right) = \left\| \mathbf{r}_{2,j} - e^{j2\pi\hat{\nu}_1 J}\mathbf{r}_{1,j} - \left(\mathbf{R}^*_{2,j} - e^{j2\pi\hat{\nu}_1 J}\mathbf{R}^*_{1,j}\right)\tilde{\boldsymbol{\rho}}_j - \left(1 - e^{j2\pi\hat{\nu}_1 J}\right)\mathbf{1}_M \tilde{d}_j \right\|^2 . \tag{4.38}$$

Define $\boldsymbol{\chi}_j \doteq \left[\tilde{\boldsymbol{\rho}}_j^T \quad \tilde{d}_j\right]^T$, and $\boldsymbol{\Psi}_j\left(\hat{\nu}_1\right) = \left[\left(\mathbf{R}^*_{2,j} - e^{j2\pi\hat{\nu}_1 J}\mathbf{R}^*_{1,j}\right) \quad \mathbf{1}_M\left(1 - e^{j2\pi\hat{\nu}_1 J}\right)\right]$. The cost function (4.38) can be rewritten as

$$\Lambda_j\left(\boldsymbol{\chi}_j \middle| \hat{\nu}_1\right) = \left\| \left(\mathbf{r}_{2,j} - e^{j2\pi\hat{\nu}_1 J}\mathbf{r}_{1,j}\right) - \boldsymbol{\Psi}_j\left(\hat{\nu}_1\right)\boldsymbol{\chi}_j \right\|^2 , \tag{4.39}$$

and the LS estimate of $\boldsymbol{\chi}_j$ is given by

$$\hat{\boldsymbol{\chi}}_j\left(\hat{\nu}_1\right) = \left(\boldsymbol{\Psi}^H_j\left(\hat{\nu}_1\right)\boldsymbol{\Psi}_j\left(\hat{\nu}_1\right)\right)^{-1}\boldsymbol{\Psi}^H_j\left(\hat{\nu}_1\right)\left(\mathbf{r}_{2,j} - e^{j2\pi\hat{\nu}_1 J}\mathbf{r}_{1,j}\right). \tag{4.40}$$

Then, the frequency offset ν can be estimated using (4.35) with the estimates $\hat{\boldsymbol{\rho}}_j$ and $\hat{d}_j$:

$$\hat{\nu}\left(\hat{\boldsymbol{\rho}}_j, \hat{d}_j\right) = \frac{\arg\left\{\sum_{j=1}^{n_r}\left(\mathbf{r}_{2,j} - \mathbf{R}^*_{2,j}\hat{\boldsymbol{\rho}}_j - \hat{d}_j\mathbf{1}_M\right)^H\left(\mathbf{r}_{1,j} - \mathbf{R}^*_{1,j}\hat{\boldsymbol{\rho}}_j - \hat{d}_j\mathbf{1}_M\right)\right\}}{2\pi J}, \tag{4.41}$$

which in turns is used for the new estimation of $\hat{\boldsymbol{\chi}}_j\left(\hat{\nu}\right)$. The algorithm begins from (4.37) and iterates between (4.38)-(4.41) which is summarized as follows, where N_I is the number of it-

erations. Lastly, the channels $h_{\pm,j,i}(n)$ and transmitter dc offset f_j can be estimated using (4.31) with the estimates $\hat{\boldsymbol{\rho}}_j$, $\hat{d}_j$, and frequency offset $\hat{\nu}$, as shown in step 9.

1. $\hat{\nu} = \arg\left\{\sum_{j=1}^{n_r} \mathbf{r}_{2,j}^H \mathbf{r}_{1,j}\right\} \Big/ 2\pi J$
2. for $m = 1$ to N_I do
3. for $j = 1$ to n_r do
4. $\boldsymbol{\Psi}_j(\hat{\nu}) = \left[\left(\mathbf{R}_{2,j}^* - e^{j2\pi\hat{\nu}J}\mathbf{R}_{1,j}^*\right) \quad \mathbf{1}_M\left(1 - e^{j2\pi\hat{\nu}J}\right)\right]$
5. $\hat{\boldsymbol{\chi}}_j \doteq \begin{bmatrix}\hat{\boldsymbol{\rho}}_j^T & \hat{d}_j\end{bmatrix}^T = \left(\boldsymbol{\Psi}_j^H(\hat{\nu})\boldsymbol{\Psi}_j(\hat{\nu})\right)^{-1}\boldsymbol{\Psi}_j^H(\hat{\nu})\left(\mathbf{r}_{2,j} - e^{j2\pi\hat{\nu}J}\mathbf{r}_{1,j}\right)$
6. end for
7. $\hat{\nu} = \arg\left\{\sum_{j=1}^{n_r}\left(\mathbf{r}_{2,j} - \mathbf{R}_{2,j}^*\hat{\boldsymbol{\rho}}_j - \hat{d}_j\mathbf{1}_M\right)^H\left(\mathbf{r}_{1,j} - \mathbf{R}_{1,j}^*\hat{\boldsymbol{\rho}}_j - \hat{d}_j\mathbf{1}_M\right)\right\} \Big/ 2\pi J$
8. end for
9. $\hat{\mathbf{h}}_j = \left(\mathbf{T}^H\mathbf{T}\right)^{-1}\mathbf{T}^H\boldsymbol{\Gamma}^H(\hat{\nu})\left(\mathbf{r}_j - \mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j - \hat{d}_j\mathbf{1}_{PN}\right)$

4.2.3 Computational Complexity Analysis

The complexity of the joint LS estimation (JLSE) and the low-complexity estimation with periodic training (LCE-PT) is compared in this subsection. Only the required number of real multiplications is compared because the computational complexity is mainly dominated by multiplication. In addition, we only focus on the comparison of the estimators of $\hat{\nu}$, $\boldsymbol{\rho}_j$, and d_j because the estimators for channels $h_{\pm,j,i}(n)$ and f_j are the same for the two estimation schemes.

The computational complexity of JLSE in (4.26)-(4.30) can be analyzed following the procedure of Section 3.3. Assume that the scalars $1\Big/\left\|\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}\right\|^2$, the vectors $\left(\mathbf{I}_{PN} - \mathbf{C}(\tilde{\nu})\right)\mathbf{1}_{PN}$ and the matrix $\mathbf{T}\left(\mathbf{T}^H\mathbf{T}\right)^{-1}\mathbf{T}^H$ can be pre-calculated and saved. For the frequency estimator $\hat{\nu}$, an exhaustive search of the minimum cost function (4.26) has to be performed. For evaluation of (4.27), first, it needs to calculate the relevant elements which are in the

form of $\mathbf{v}^H\mathbf{Q}(\tilde{\nu})\mathbf{u}$, where $\mathbf{v}$ and $\mathbf{u}$ are $PN\times 1$ vectors. As shown in Section 3.2, the evaluation of the element $\mathbf{v}^H\mathbf{C}(\tilde{\nu})\mathbf{u}$ in $\mathbf{v}^H\mathbf{Q}(\tilde{\nu})\mathbf{u}$ can be implemented through FFT. Then the total number of real multiplications required for calculation of $\mathbf{v}^H\mathbf{Q}(\tilde{\nu})\mathbf{u}$ is

$$4P^2N\cdot\left[K+2N+2+\eta K\log_2(KN)\right]+4PN\cdot(2KN+1)+8KN,$$

where $\eta=1-\left(\log_2 K+2(1/K-1)\right)/\log_2(KN)$ with the design parameter K called pruning factor. Second, the inverse matrix term $\left[\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{R}_j^*\right]^{-1}\mathbf{R}_j^T\mathbf{Q}(\tilde{\nu})\mathbf{r}_j$ can be implemented by Gaussian elimination, whose complexity is $4KN\cdot\left(L_\rho^3+3L_\rho^2-L_\rho\right)/3$ complex products. For the other estimators $\hat{\boldsymbol{\rho}}_j$ and $\hat{d}_j$, refer to Section 3.3.

For frequency estimator of LCE-PT in (4.41), it requires $n_r\cdot J(Q+1)\cdot L_\rho$ complex products to calculate $\left\{\mathbf{r}_j-\mathbf{R}_j^*\hat{\boldsymbol{\rho}}_j-\hat{d}_j\mathbf{1}_{M+J}\right\}_{j=1}^{n_r}$, where $\mathbf{r}_j$ and $\mathbf{R}_j^*$ are of size $J(Q+1)\times 1$ and $J(Q+1)\times L_\rho$, respectively. Totally, it requires $n_r\left[JQ(L_\rho+1)+JL_\rho\right]$ complex products to calculate $\hat{\nu}\left(\hat{\boldsymbol{\rho}}_j,\hat{d}_j\right)$. For estimators $\hat{\boldsymbol{\rho}}_j$ and $\hat{d}_j$ of LCE-PT in (4.40), first, it requires $JQ+L_\rho-1$ complex products to form $\boldsymbol{\Psi}_j(\hat{\nu}_1)$. Second, it requires $\left[(L_\rho+2)(L_\rho+1)/2+(L_\rho+2)\right]\cdot JQ$ complex products to calculate $\mathbf{A}_j^H(\hat{\nu})\mathbf{A}_j(\hat{\nu})$ and $\mathbf{A}_j^H(\hat{\nu})\left(\mathbf{r}_{2,j}-e^{j2\pi\hat{\nu}K}\mathbf{r}_{1,j}\right)$. Finally, the inverse matrix term $\left(\mathbf{A}_j^H(\hat{\nu})\mathbf{A}_j(\hat{\nu})\right)^{-1}\mathbf{A}_j^H(\hat{\nu})\left(\mathbf{r}_{2,j}-e^{j2\pi\hat{\nu}K}\mathbf{r}_{1,j}\right)$ can be implemented by Gaussian elimination, whose complexity is $\left[(L_\rho+1)^3+3(L_\rho+1)^2-(L_\rho+1)\right]/3$ complex products.

The complexities of the two methods are summarized in Table 4.1. In order for a fair comparison, we assume there is the same useful training length M for two estimation schemes, i.e. $M=PN=JQ$. It can be shown that the complexity order of JLSE is $O(M^2)$ while that of LCE-PT is $O(M)$. However, it can be seen from (4.41) that the frequency estimation range of

Table 4.1: Computational Complexities of JLSE and LCE-PT (Number of Real Multiplications)

	JLSE	LCE-PT
$\hat{\nu}$	$n_r \cdot \left(2L_\rho^2 + 6L_\rho + 4\right) \cdot \left\{P^2 N\left[K + 2N + 2 + \eta K \log_2\left(KN\right)\right] + PN \cdot \left(2KN + 1\right) + 2KN\right\} + n_r \cdot KN \cdot \left(4L_\rho^3 + 12L_\rho^2 + 8L_\rho\right)/3$	$n_r\left[JQ\left(N_I L_\rho + N_I + 1\right) + N_I JL_\rho\right]$
$\hat{\rho}_j$ and $\hat{d}_j$	$P^2 N^2 \cdot \left(2L_\rho^2 + 6L_\rho + 2\right) + PN \cdot \left(8L_\rho^2 + 28L_\rho + 2\right) + \left(4L_\rho^3 + 12L_\rho^2 - 4L_\rho\right)/3$	$N_I\left(4L_\rho^3 + 24L_\rho^2 + 44L_\rho\right)/3 + N_I JQ\left(2L_\rho^2 + 10L_\rho + 16\right)$

LCE-PT is reduced by J times of that of JLSE. Moreover, the performance of JLSE is better than LCE-PT as can be seen in simulation results section 4.3.

4.3 Simulation Results

The performance of the proposed methods is evaluated for un-coded SISO and MIMO-OFDM systems. Table 4.2 gives the system parameters and radio impairments. The transmission is done on a packet-by-packet basis beginning with the training sequence. A wide-sense stationary un-correlated scattering (WSSUS) discrete channel is employed, with the impulse response $h_{j,i}(\tau) = \sum_{m=0}^{L} h_{j,i}(m)\delta(\tau - mT_s)$, where $L+1$ is the length of the channel, and $\{h_{j,i}(m)\}$ are tap gains which are mutually independent complex Gaussian random variables with zero mean and variance σ_m^2. Exponential multi-path intensity profile is used with $\sigma_m^2 = \sigma_0^2 \cdot \exp(-mT_s/T_{RMS})$, where T_{RMS} is the root-mean square delay spread, and to maintain unit power gain, $\sigma_0^2 = 1 - \exp(-T_s/T_{RMS})$. The channel remains unchanged during a packet. The parameters are set as $T_{RMS} = 50\ ns$, $L = 4$, $L_+ = 11$ and $L_- = 8$. SNR $\doteq \sigma_s^2/\sigma_w^2$ with σ_s^2 and σ_w^2 defined in subsection 2.2.2. The periodic and OFDM training symbols are generated from the

Table 4.2: System Parameters and Radio Impairments

System Parameter	Parameter Value
Signal Bandwidth	20 MHz
FFT length (N), cyclic prefix length (N_g)	$N=64$, $N_g=16$
Symbol Time (T_s)	$T_s=50\ ns$
Subcarrier Spacing $(1/NT_s)$	0.3125 MHz
Frequency independent I-Q Imbalance $(\alpha_{T,i},\theta_{T,i})$, $(\alpha_{R,j},\theta_{R,j})$	$(\alpha_{T,1}=1.05,\theta_{T,1}=-5^o)$, $(\alpha_{T,2}=0.94,\theta_{T,2}=-6^o)$ $(\alpha_{R,1}=1.08,\theta_{R,1}=5^o)$, $(\alpha_{R,2}=0.91,\theta_{R,2}=6^o)$, $(\alpha_{R,3}=0.92,\theta_{R,3}=-5^o)$
$\{c_{T,i}^I(n),c_{T,i}^Q(n)\}$, $\{c_{R,j}^I(n),c_{R,j}^Q(n)\}$	I part: $[1 \quad 0.2 \quad -0.05]$, Q part: $[1.08 \quad 0.15 \quad 0.02]$
Frequency offset ν	uniform over $(-0.95,0.95)$ $(1/NT_s)$
DC offset $f_{0,i}$ and $d_{0,j}$, with signal power normalized to 1	$d_{0,1}=f_{0,1}=0.1\times(1+j)/\sqrt{2}$, $d_{0,2}=f_{0,2}=-0.1\times(1+j)/\sqrt{2}$, $d_{0,3}=0.08\times(1+j)/\sqrt{2}$

sequence whose tones are unit norm with random phase, which are similar to the short and long training sequences of 802.11a spec [79], respectively.

Figure 4.4 shows the mean-squared error (MSE) comparison of the proposed two estimation methods JLSE and LCE-PT in SISO systems. For fair comparison, the same periodic training sequence of total 8 periods ($Q=6$) with period $J=32$ is used for the two methods. The JLSE method employs the last 7 periods for training while discarding the first period as the prefix. The normalized mean-squared error (NMSE) is the mean-square error of the frequency offset

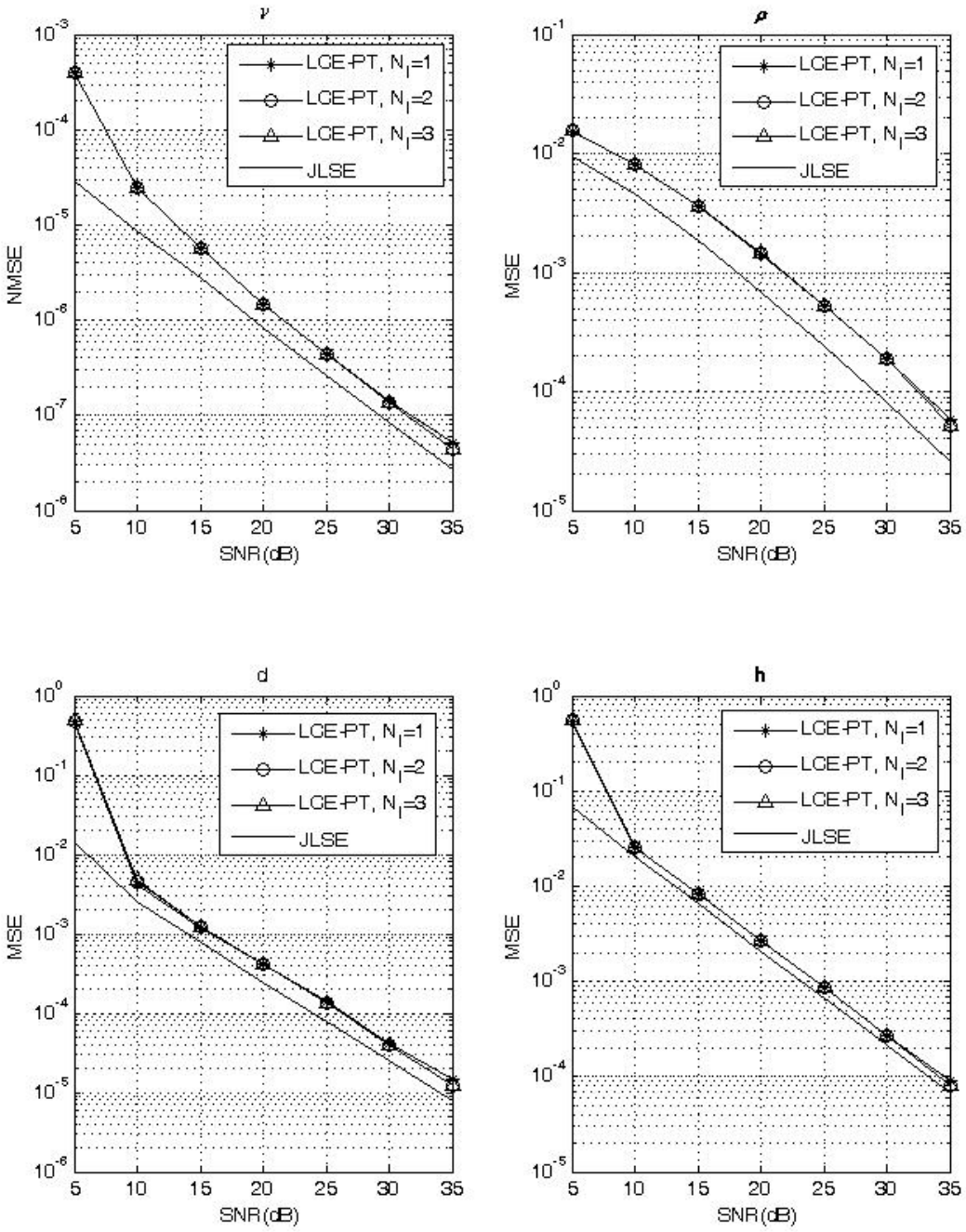


Figure 4.4: MSE performance comparison between JLSE and LCE-PT schemes.

normalized by subcarrier spacing. It can be seen that JLSE outperforms LCE-PT by a margin of 1-3 dB in MSE performance. Somewhat surprising, LCE-PT only suffers from a very slight

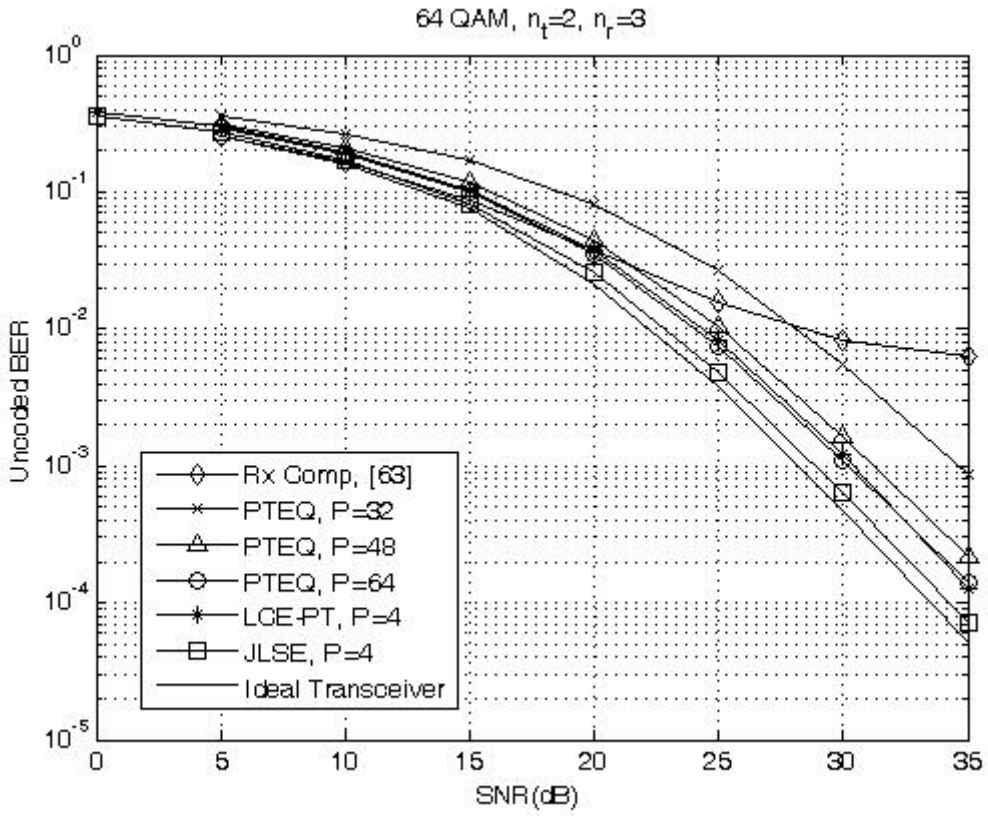


Figure 4.5: Performance comparisons between the proposed methods, receiver radio compensation only [63] and per-tone equalization (PTEQ) [48] (spatial-multiplexing MIMO-OFDM).

performance loss for using $N_I = 1$.

Figures 4.5 and 4.6 compare the un-coded bit error-rate (BER) performance of the receiver based on the proposed radio impairments cancellation with the existing receivers. For a fair comparison, MMSE detection is employed in all receivers and the same training length is used for the JLSE and LCE-PT methods. The JLSE method employs $P = 4$ OFDM training symbols for training. For the LCE-PT method, the training structure begins with a periodic training sequence of total 5 periods ($Q = 3$) with period $J = 32$ for estimation of ρ_j, d_j, and frequency offset ν followed by 2 OFDM training symbols for estimation of $\hat{\mathbf{h}}_j$, with overall training length $P = 4$ (OFDM symbol) equal to that used for the JLSE method. In Figure 4.5, the comparison is

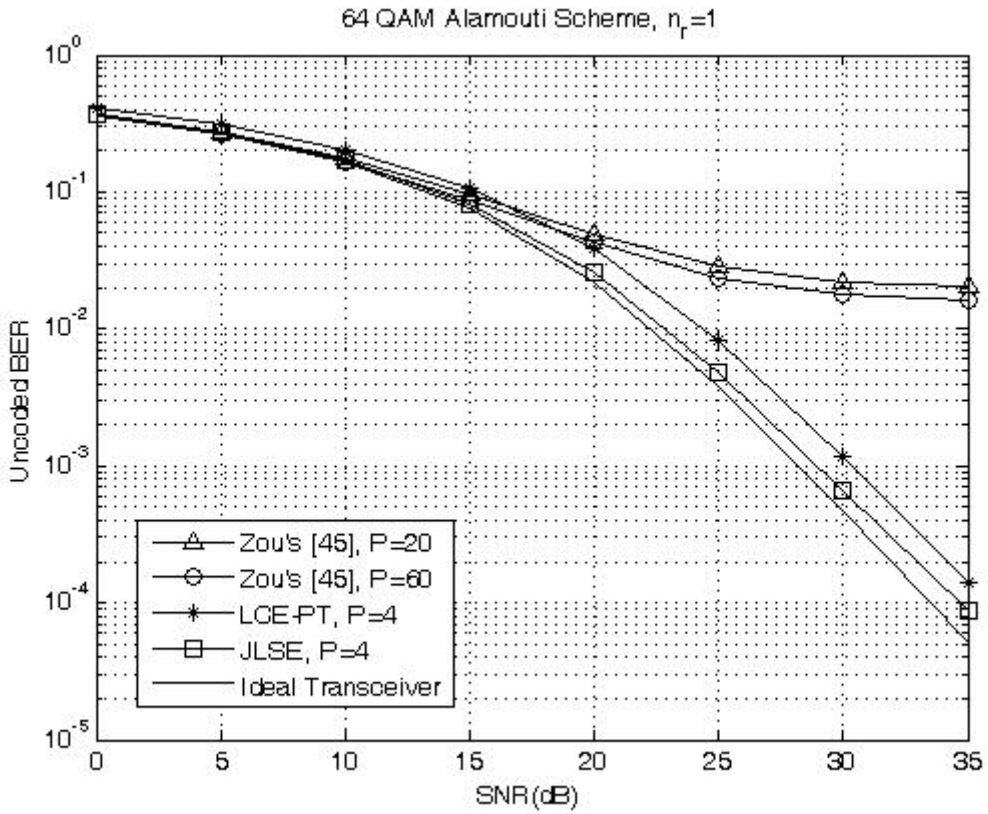


Figure 4.6: Performance comparisons between the proposed methods and Zou's method [45] (STBC MIMO-OFDM).

made with the one in [63], where only the receiver radio impairments are compensated, and the one in [48], where PTEQ is employed, for a spatial-multiplexing MIMO-OFDM system. Since there is no method for frequency estimation in [48], ideal frequency estimation is assumed in PTEQ method [48]. Clearly, the transmitter radio impairments incur error floor if left not compensated in [63], as one can expect. Also, our design significantly outperforms PTEQ in terms of BER and/or the required training symbols. In PTEQ, $P \geq 48$ is needed for a satisfactory performance, and $P \geq 64$ is comparable with $P = 4$ in our methods. In comparison of proposed two methods, it can be seen that JLSE outperforms LCE-PT by 1 dB in un-coded BER performance with the same training length $P = 4$ (OFDM symbol).

In Figure 4.6, comparison is made with the one in [45] as an example of STBC MI-

MO-OFDM systems. Since no frequency offset and dc offset was considered in [45], no dc offset is assumed in our simulations. In addition, a phase de-rotation based frequency offset compensation is done directly from the received signal with a perfect frequency estimation. Clearly in the figure, the method [45] does not work in the presence of frequency offset even with a perfect estimation at the receiver; this can be attributed to that in the presence of I-Q imbalance, with direct compensation of frequency offset from the received signal, the mirror frequency part still suffers from the effect of frequency offset. Again the BER performance loss caused by using the low-complexity LCE-PT is about 1 dB in this case compared with JLSE.

4.4 Summary

In Chapter 4, new estimation and cancellation techniques are proposed for the cascaded transmitter and receiver radio impairments in the linear-dispersion coded MIMO-OFDM systems. A two-stage cancellation architecture is developed, which enables to explicitly cancel the cascaded radio impairments without increasing the dimension of signal detection. In addition, two methods are proposed for the radio parameters estimation, including the optimal joint estimation of all radio parameters and a low-complexity estimation aided by periodic training. The optimal method has a much better estimation performance while the low-complexity one is simpler but with a narrower frequency estimation range. The new methods are simulated and compared with the existing ones. Simulation results show that significant performance gain is observed with our new methods both in STBC and spatial multiplexing MIMO-OFDM systems in error-rate performance and/or the number of training symbols required.

Chapter 5

Calibration Technique

In this chapter, a new method is proposed to self-calibrate the transmitter and receiver impairments simultaneously, with no dedicated analog circuit in the feedback loop. Based on a time-domain approach, the new method is applicable to all types of communication systems and is able to calibrate jointly the frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance, and dc offset. In addition, optimal training sequences are devised to best the calibration performance. Simulation and analytical results confirm the effectiveness of the proposed method.

The rest of this chapter is organized as follows. Section 5.1 describes the models of radio impairments and digital calibration circuits. Section 5.2 develops the joint estimator under the nonlinear least-squares principle with optimal training design. The proposed method is then analyzed in Section 5.3, and simulation results and summary are given in Section 5.4 and Section 5.5.

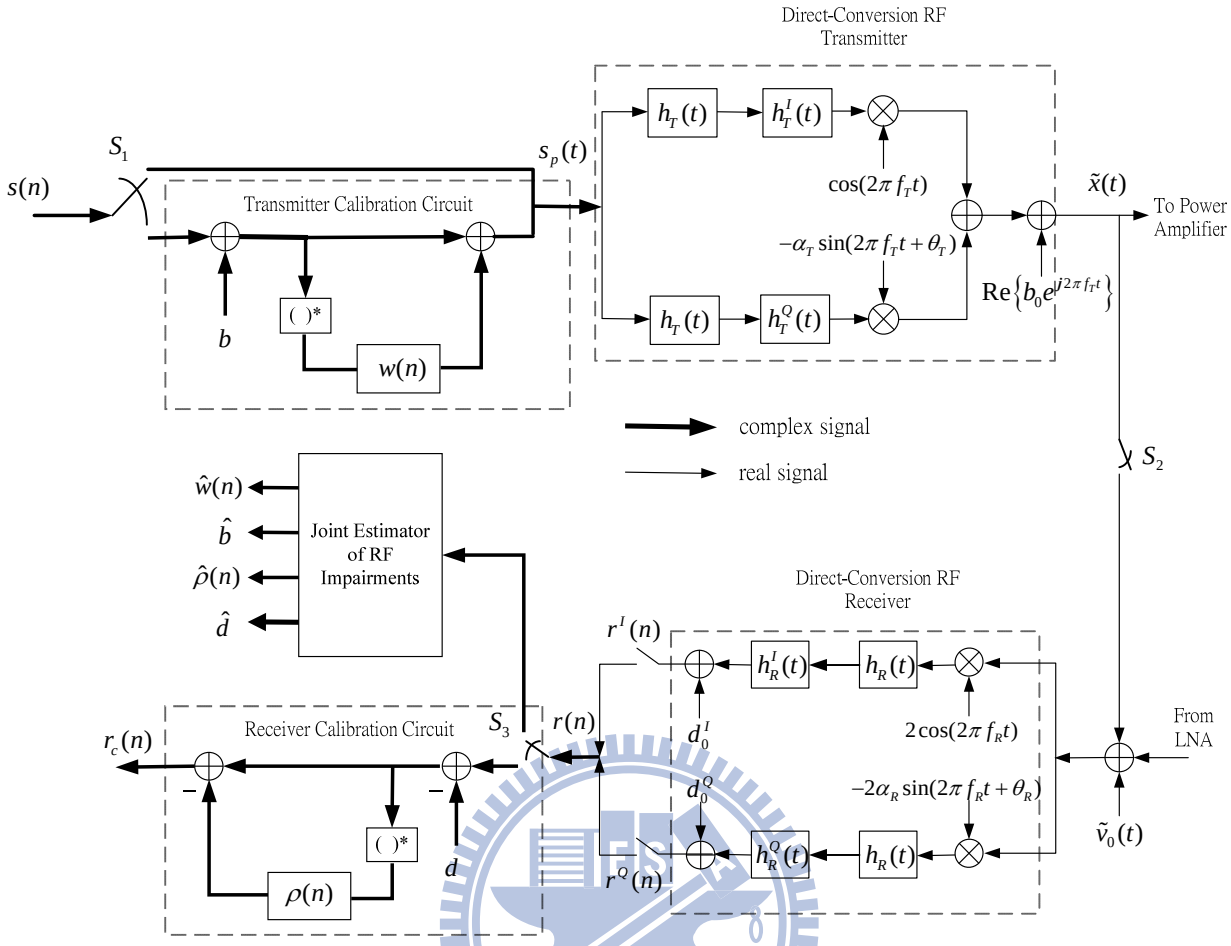


Figure 5.1: The calibration system consists of a direct-conversion RF transceiver, calibration circuits and a joint estimator of the calibration parameters.

5.1 Radio Impairments and Calibration Circuits

Figure 5.1 depicts the calibration system that consists of calibration circuits, direct-conversion radio transceiver, and a joint estimator of the transmitter and receiver radio impairments. The switches S_1, S_2 and S_3 are to control the flow of the signal path; during the calibration training, S_1 and S_3 are at the upper positions, and S_2 is closed, whereas during the normal communication, S_1 and S_3 are at the lower positions and S_2 is open. In practice, S_2 may be implemented with parasitic coupling or others [61].

5.1.1 Radio Impairments

At the transmitter, the considered radio impairments include frequency-dependent I-Q imbalance, frequency-independent I-Q imbalance and carrier feed-through. Frequency-dependent I-Q imbalance is due to mismatch between the in-phase (I) and quadrature-phase (Q) analog filters, which are denoted by $h_T(t) \otimes h_T^I(t)$ and $h_T(t) \otimes h_T^Q(t)$, respectively. Here, $h_T(t)$ is the common part of the filters, and $\otimes$ is the operation of linear convolution. Frequency-independent I-Q imbalance is due to gain and phase mismatches between the I and Q branches of the mixer circuitry and are denoted by α_T and θ_T, respectively. And, carrier feed-through induces dc offset at the receiving side [5]-[8] and is characterized by $\mathrm{Re}\left\{b_0 e^{j2\pi f_T t}\right\}$, where $b_0 = b_0^I + \boldsymbol{j} b_0^Q$, f_T is the transmit center frequency. In the following, b_0 is called transmitter dc offset.

Define

$$s_p(t) \doteq s_p^I(t) + \boldsymbol{j} s_p^Q(t) = \sum_n s_p(n)\delta(t - nT_s) \tag{5.1}$$

be the signal appearing at the input of transmitter, where $s_p(n)$ is its discrete-time signal, T_s is the symbol duration. During the calibration training, $s_p(n) = s(n)$ is the training sequence.

Under the effects of the radio impairments, the pass-band transmit signal is $\tilde{x}(t) = \mathrm{Re}\left\{x(t) e^{j2\pi f_T t}\right\}$, with its equivalent base-band signal given by,

$$x(t) = h_{T,+}(t) \otimes s_p(t) + h_{T,-}(t) \otimes s_p^*(t) + b_0, \tag{5.2}$$

where

$$h_{T,\pm}(t) = 1/2 \cdot \left[h_T^I(t) \pm \alpha_T e^{j\theta_T} h_T^Q(t)\right] \otimes h_T(t). \tag{5.3}$$

In (5.2), $s_p(t)$ can be viewed as being transmitted by two channels with a corruption from dc offset; one is the desired channel with impulse response $h_{T,+}(t)$, and the other is the mir-

ror-frequency channel with impulse response $h_{T,-}(t)$. That is, I-Q imbalances incur mirror-frequency interference in the transmitted signal. Dc-offset, on the other hand, imposes a more strict specification on A/D converters at the receiver and may incur in-band interference in the presence of frequency offset. Note that with no I-Q imbalances and dc offset, i.e., $h_T^I(t) = h_T^Q(t) = \delta(t)$, $\alpha_T = 1$ and $\theta_T = b_0 = 0$, $x(t) = s_p(t) \otimes h_T(t)$, as one might expect.

Similar to the transmitter, the considered radio impairments in the receiver include frequency-independent I-Q imbalance, characterized by α_R and θ_R, frequency-dependent I-Q imbalance, characterized by the filters $h_R(t) \otimes h_R^I(t)$ and $h_R(t) \otimes h_R^Q(t)$, and dc offset, $d_0 = d_0^I + \boldsymbol{j} d_0^Q$. $h_R(t)$ is the common part of filters, $f_R = f_T - \Delta f$ is the receive center frequency, and Δf is a frequency offset between transmitter and receiver which is introduced purposely to separate the transmitter radio impairments from those of the receiver; the introduction of the frequency offset is essential in this method, as to be detailed in Section III. $\tilde{v}_0(t) = \mathrm{Re}\left\{ v_0(t) e^{j2\pi f_R t} \right\}$ is the pass-band additive white Gaussian noise with $v_0(t)$ being its base-band equivalent signal.

After the internal loop-back with a proper gain adjustment, the received base-band signal is given by

$$r(t) = h_{R,+}(t) \otimes \left[e^{j2\pi\Delta f t} x(t) + v_0(t) \right] + h_{R,-}(t) \otimes \left[e^{j2\pi\Delta f t} x(t) + v_0(t) \right]^* + d_0 \ , \tag{5.4}$$

where

$$h_{R,\pm}(t) = 1/2 \cdot \left[h_R^I(t) \pm \alpha_R e^{\mp j\theta_R} h_R^Q(t) \right] \otimes h_R(t). \tag{5.5}$$

Again, (5.4) says that the receiver I-Q imbalances induces mirror-frequency interference in the received signal, and in the absence of I-Q imbalances and dc offset,

$$r(t) = h_R(t) \otimes \left[e^{j2\pi\Delta f t} x(t) + v_0(t) \right]. \tag{5.6}$$

5.1.2 Calibration Circuits

At the transmitter, we propose to use a pre-distortion filter, $w(n)$, and a dc correction term, b, to calibrate I-Q imbalances and dc offset, as in Figure 1. Using a pre-distortion filter to remove the mirror-frequency interference has been reported in [57]. After calibration, $s_p(n)$ is given by

$$s_p(n) = \left[s(n) + b \right] + w(n) \otimes \left[s(n) + b \right]^* , \tag{5.7}$$

where $s(n)$ is the transmitted symbol. For convenience, the equivalent discrete-time signal model will be used throughout the rest of the paper. In this way, (5.2) is rewritten as

$$x(n) = g_{T,+}(n) \otimes s(n) + g_{T,-}(n) \otimes s^*(n) + \Delta b , \tag{5.8}$$

where $u(n) = u(t)\big|_{t=nT_s} \quad u \in \{ r, x, h_{T,+}, h_{T,-}, v_0 \}$,

$$g_{T,+}(n) = h_{T,+}(n) + w^*(n) \otimes h_{T,-}(n), \tag{5.9}$$

$$g_{T,-}(n) = h_{T,-}(n) + w(n) \otimes h_{T,+}(n) , \tag{5.10}$$

and

$$\Delta b = g_{T,+}(n) \otimes b + g_{T,-}(n) \otimes b^* + b_0 . \tag{5.11}$$

Note that $g_{T,+}(n)$ is regarded as the overall impulse response of the desired channel, $g_{T,-}(n)$ is that of the mirror-frequency channel, and Δb is the residual dc offset after calibration. Ideally, $g_{T,-}(n) = 0$ and $\Delta b = 0$ which give

$$w(n) = -h_{T,+}^{-1}(n) \otimes h_{T,-}(n), \tag{5.12}$$

and

$$b = -b_0 \otimes g_{T,+}^{-1}(n), \tag{5.13}$$

where $h_{T,+}^{-1}(n)$ and $g_{T,+}^{-1}(n)$ are the inverse filters of $h_{T,+}(n)$ and $g_{T,+}(n)$, respectively. Fur-

thermore, it is interesting to see that $w(n)=b=0$ for the case of no I-Q imbalance and dc offset. Note that $w(n)$ and b are estimated during the calibration training.

As in most of the relevant literature, for example in [57] and [58], the image-rejection-ratio (IRR) will be adopted as the performance measure for the I-Q imbalance calibration, which is defined as

$$IRR_T(f)=10\log_{10}\frac{|G_{T,+}(f)|^2}{|G_{T,-}(f)|^2}\text{dB}, \tag{5.14}$$

where $U(f)$ is the Fourier transform (FT) of $u(n)$. In addition, the ratio

$$\varepsilon_T=10\log_{10}\frac{|\Delta b|^2}{|b_0|^2}\text{dB} \tag{5.15}$$

will be adopted as the performance index for the dc offset calibration.

Similarly, at the receiver a time-domain compensation filter, $\rho(n)$, is employed to remove the receiver mirror-frequency interference, and a dc correction term, d, is to remove the dc offset (see Figure 1). Thus, the received signal after calibration is given by

$$\begin{aligned} r_c(n)&=\left(r(n)-d\right)-\rho(n)\otimes\left(r(n)-d\right)^* \\ &=g_{R,+}(n)\otimes\left[e^{j2\pi\mu n}x(n)+v_0(n)\right]+g_{R,-}(n)\otimes\left[e^{j2\pi\mu n}x(n)+v_0(n)\right]^*+\Delta d \end{aligned} \tag{5.16}$$

where

$$r(n)=h_{R,+}(n)\otimes\left[e^{j2\pi\mu n}x(n)+v_0(n)\right]+h_{R,-}(n)\otimes\left[e^{j2\pi\mu n}x(n)+v_0(n)\right]^*+d_0, \tag{5.17}$$

$\mu=\Delta fT_s$ is the normalized frequency offset,

$$g_{R,\pm}(n)=h_{R,\pm}(n)-\rho(n)\otimes h_{R,\mp}^*(n), \tag{5.18}$$

and

$$\Delta d = (d_0 - d) - \rho(n) \otimes (d_0 - d)^* . \tag{5.19}$$

$\{v_0(n)\}$ are i.i.d. zero-mean Gaussian noise with variance of σ_0^2. Clearly, it is desirable to have $g_{R,-}(n) = 0$ and $\Delta d = 0$; or equivalently,

$$\rho(n) = \left(h_{R+}^*(n)\right)^{-1} \otimes h_{R-}(n) \ , \tag{5.20}$$

and

$$d = d_0 , \tag{5.21}$$

where $\left(h_{R,+}^*(n)\right)^{-1}$ is the inverse filters of $h_{R,+}^*(n)$.

Like the transmitter side, the receiver calibration performance is evaluated by

$$IRR_R(f) = 10\log_{10}\frac{|G_{R,+}(f)|^2}{|G_{R,-}(f)|^2} \text{dB} \tag{5.22}$$

for I-Q imbalance and by

$$\varepsilon_R = 10\log_{10}\frac{|\Delta d|^2}{|d_0|^2} \ \text{dB} \tag{5.23}$$

for the dc offset, respectively. Likewise, $\rho(n) = d = 0$ for the case of no I-Q imbalance and dc offset. Also, $\rho(n)$ and d are estimated during the calibration training and used over the normal communication.

5.2 Joint Estimation of Calibration Parameters

In this section, a joint estimation of the calibration parameters, $w(n)$, b, $\rho(n)$ and d, is developed first, followed by the design of training sequence and frequency offset to optimize the calibration performance. Recall that the estimation is done during the calibration training,

where $s_p(n) = s(n)$.

5.2.1 Non-linear Least-Squares Estimation

Using the equivalent discrete-time model, the received signal in (5.4) can be rewritten as

$$\begin{aligned} r(n) = {} & e^{j2\pi\mu n}\left[f_{1,+}(n) \otimes s(n) + f_{1,-}(n) \otimes s^*(n) + b_1 \right] \\ & + e^{-j2\pi\mu n}\left[f_{2,+}(n) \otimes s(n) + f_{2,-}(n) \otimes s^*(n) + b_2 \right] + d_0 + v(n), \end{aligned} \tag{5.24}$$

where

$$f_{1,\pm}(n) = \left(h_{R,+}(n) e^{-j2\pi\mu n} \right) \otimes h_{T,\pm}(n), \tag{5.25}$$

$$f_{2,\pm}(n) = \left(h_{R,-}(n) e^{j2\pi\mu n} \right) \otimes h^*_{T,\mp}(n), \tag{5.26}$$

$$b_1 = \left(h_{R,+}(n) e^{-j2\pi\mu n} \right) \otimes b_0, \tag{5.27}$$

$$b_2 = \left(h_{R,-}(n) e^{j2\pi\mu n} \right) \otimes b_0^*, \tag{5.28}$$

and

$$v(n) = h_{R,+}(n) \otimes v_0(n) + h_{R,-}(n) \otimes v_0^*(n) \tag{5.29}$$

is the noise component. Our goal is to estimate $w(n)$, b, $\rho(n)$ and d, given the training sequence $s(n)$ and the frequency offset μ. Obviously, one possible way to do this is to estimate $h_{T,\pm}(n)$, $h_{R,\pm}(n)$, b_0 and d_0 directly from (5.24) and apply them to (5.12), (5.13), (5.20) and (5.21). Direct estimation of $h_{T,\pm}(n)$, $h_{R,\pm}(n)$ and b_0, however, is very complicated as can be seen from (5.24)-(5.28). Instead, a simpler estimation method is proposed based on the following observations.

$$\begin{aligned} w(n) &= -h^{-1}_{T,+}(n) \otimes h_{T,-}(n) \\ &= -f^{-1}_{1,+}(n) \otimes f_{1,-}(n), \end{aligned} \tag{5.30}$$

$$
\begin{aligned}
b &= -b_0 \otimes g_{T,+}^{-1}(n) \\
&= -b_0 \otimes \left(h_{R,+}(n)e^{-j2\pi\mu n}\right) \otimes \left[\left(h_{R,+}(n)e^{-j2\pi\mu n}\right) \otimes \left(h_{T,+}(n) + w^*(n) \otimes h_{T,-}(n)\right)\right]^{-1} \\
&= -b_1 \otimes \left[f_{1,+}(n) + w^*(n) \otimes f_{1,-}(n)\right]^{-1},
\end{aligned}
\tag{5.31}
$$

and

$$
\begin{aligned}
\rho(n) &= \left(h_{R,+}^*(n)\right)^{-1} \otimes h_{R,-}(n) \\
&= \left(f_{1,+}^*(n)e^{-j2\pi\mu n}\right)^{-1} \otimes \left(f_{2,-}(n)e^{-j2\pi\mu n}\right).
\end{aligned}
\tag{5.32}
$$

Therefore, $w(n)$, b, and $\rho(n)$ can be calculated through $f_{1,\pm}(n)$, $f_{2,-}(n)$, and b_1 which, along with d_0, can be estimated from (5.24) in a much easier way, as to be discussed below.

In the proposed method, $f_{1,\pm}(n)$, $f_{2,\pm}(n)$, b_1, b_2, and d_0 will be estimated under the least-squares principle, although $f_{2,+}(n)$ and b_2 only serve as auxiliary variables and are not needed in the final evaluation. To this end, firstly let $f_{1,\pm}(n)$ and $f_{2,\pm}(n)$ be modeled as FIR (finite impulse response) filters,

$$
\mathbf{f}_{i,\pm} = \left[f_{i,\pm}(0), f_{i,\pm}(1), \ldots, f_{i,\pm}(L_f - 1)\right]^T, \quad i = 1, 2,
\tag{5.33}
$$

where L_f is the filters' length and usually is not known in advance. In Section V, it will be shown that the estimation performance is quite insensitive to the value of L_f if it is selected not too small. Consider a training sequence $\{s(n)\}_{n=-N_g}^{N-1}$, where $N_g \geq L_f$, and $s(n) = s(n+N)$, $n = -N_g, \ldots, -1$ is the cyclic-prefix. Define $\mathbf{S}$ be the $N \times L_f$ signal matrix with $[\mathbf{S}]_{i,j} = s(i-j)$, $0 \leq i \leq N-1$, $0 \leq j \leq L_f - 1$. Then, (5.24) can be rearranged into the following vector-matrix form

$$
\mathbf{r} = \mathbf{\Phi f} + \mathbf{v},
\tag{5.34}
$$

where

$$
\mathbf{r} = \left[r(0), r(1), \ldots, r(N-1)\right]^T,
\tag{5.35}
$$

$$\mathbf{T} = \begin{bmatrix} \mathbf{S} & \mathbf{S}^* & \mathbf{1}_N \end{bmatrix}, \tag{5.36}$$

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Gamma}_N(\mu)\mathbf{T} & \mathbf{\Gamma}_N(-\mu)\mathbf{T} & \mathbf{1}_N \end{bmatrix}, \tag{5.37}$$

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}_{1,+}^T, \mathbf{f}_{1,-}^T, b_1, \mathbf{f}_{2,+}^T, \mathbf{f}_{2,-}^T, b_2, d_0 \end{bmatrix}^T, \tag{5.38}$$

$\mathbf{v} = \left[v(0), v(1), \ldots, v(N-1) \right]^T$, and $\mathbf{\Gamma}_N(\mu) = \text{diag}\left\{ 1, e^{j2\pi\mu}, \ldots, e^{j2\pi\mu(N-1)} \right\}$.

From (5.34), the least-squares estimate $\hat{\mathbf{f}}$ is given by

$$\hat{\mathbf{f}} = \Upsilon \mathbf{r}, \tag{5.39}$$

where $\Upsilon = \left(\mathbf{\Phi}^H \mathbf{\Phi} \right)^{-1} \mathbf{\Phi}^H$ is the pseudo inverse of $\mathbf{\Phi}$. Note that $\mathbf{\Phi}$ has to have full-rank in order to assure identifiability. After obtaining $\hat{\mathbf{f}}$, $w(n)$, b, and $\rho(n)$ can be evaluated as in (5.30)-(5.32). Substitute (5.34) into (5.39), one has

$$\hat{\mathbf{f}} = \mathbf{f} + \Upsilon \mathbf{v}, \tag{5.40}$$

which is an unbiased estimate of $\mathbf{f}$ with the mean-square error (MSE) given below,

$$\mathrm{E}\left[\left\| \hat{\mathbf{f}} - \mathbf{f} \right\|^2 \right] = tr\left\{ \Upsilon \mathrm{E}\left[\mathbf{v}\mathbf{v}^H \right] \Upsilon^H \right\} = tr\left\{ \Upsilon \mathbf{C}_{\mathbf{v}} \Upsilon^H \right\}, \tag{5.41}$$

where $\mathbf{C}_{\mathbf{v}} = \mathrm{E}\left[\mathbf{v}\mathbf{v}^H \right]$ is the noise correlation matrix. Notice that with no frequency offset, i.e., $\mu = 0$, and $\mathbf{\Gamma}(\mu) = \mathbf{I}_N$, where $\mathbf{I}_N$ is the identity matrix with size $N \times N$, (5.34) becomes

$$\mathbf{r} = \mathbf{S}\left(\mathbf{f}_{1,+} + \mathbf{f}_{2,+} \right) + \mathbf{S}^*\left(\mathbf{f}_{1,-} + \mathbf{f}_{2,-} \right) + \left(b_1 + b_2 + d_0 \right)\mathbf{1}_N. \tag{5.42}$$

In such an un-desirable case, $\mathbf{f}_{1,+}$ and $\mathbf{f}_{2,+}$ are not identifiable, so are $\mathbf{f}_{1,-}$ and $\mathbf{f}_{2,-}$, and b_1, b_2, and d_0. In other words, the transmitter impairments and receiver impairments cannot be estimated separately. This explains the necessity of introducing the frequency offset μ in our method.

5.2.2 Training Sequence Design

Theoretically, the optimal training sequence is the one that minimizes MSE $\mathrm{E}\left[\left\|\hat{\mathbf{f}}-\mathbf{f}\right\|^2\right]=tr\left\{\Upsilon\mathbf{C}_{\mathrm{v}}\Upsilon^H\right\}$. As seen in (5.29), however, $\mathbf{C}_{\mathrm{v}}$ is a function of $h_{R,+}(n)$ and $h_{R,-}(n)$, and therefore the optimal training sequence differs from a transceiver to another. Here, the simplified measure

$$tr\left\{\Upsilon\Upsilon^H\right\}=tr\left\{\left(\mathbf{\Phi}^H\mathbf{\Phi}\right)^{-1}\right\} \tag{5.43}$$

is adopted for the search of the optimum training sequence. The measure is optimal if $v(n)$, $n=0,\ldots,N-1$ are white Gaussian noises.

Assuming that $\frac{1}{N}\sum_{n=0}^{N-1}\left|s(n)\right|^2=1$, then from (5.36) and (5.37), one has

$$tr\left\{\mathbf{\Phi}^H\mathbf{\Phi}\right\}=\left(4L_f+3\right)\cdot N. \tag{5.44}$$

For this case, it was derived in [80] that the minimum MSE in (5.43) is achieved provided that

$$\mathbf{\Phi}^H\mathbf{\Phi}=N\cdot\mathbf{I}_{4L_f+3} \tag{5.45}$$

which, following (5.37), gives

$$\mathbf{T}^H\mathbf{T}=N\cdot\mathbf{I}_{2L_f+1}, \tag{5.46}$$

$$\mathbf{T}^H\mathbf{\Gamma}\left(-2\mu\right)\mathbf{T}=\mathbf{O}_{2L_f+1\times 2L_f+1}, \tag{5.47}$$

and

$$\mathbf{T}^H\mathbf{\Gamma}\left(\pm\mu\right)\mathbf{1}_N=\mathbf{0}_{2L_f+1}\ . \tag{5.48}$$

In (5.46)-(5.48), it is clear that $s(n)$ needs to be devised jointly with frequency offset μ in order to have the best performance, and that complicates the design very significantly. In the following, a simpler method is proposed.

Consider a periodic training sequence that consists of $P+1$ periods with K samples in

each period, i.e., $s(n)=s(n+K)$, $n=-K,\ldots,0,\ldots,N-1$, where $N=KP$, and $N_g=K$. Define $\mathbf{S}_1$ be the signal matrix for one period, with $[\mathbf{S}_1]_{i,j}=s(i-j)$, $0\le i\le K-1$, $0\le j\le L_f-1$, $\mathbf{T}_1=\begin{bmatrix}\mathbf{S}_1 & \mathbf{S}_1^* & \mathbf{1}_K\end{bmatrix}$, and $\mathbf{T}=\underbrace{\begin{bmatrix}\mathbf{T}_1^T & \cdots & \mathbf{T}_1^T\end{bmatrix}}_{P}^T$, then the matrix $\mathbf{\Phi}$ can be decomposed as follows.

$$\begin{aligned}\mathbf{\Phi}&=\begin{bmatrix}\mathbf{\Gamma}_N(\mu)\mathbf{T} & \mathbf{\Gamma}_N(-\mu)\mathbf{T} & \mathbf{1}_N\end{bmatrix}\\&=\begin{bmatrix}\mathbf{\Gamma}_K(\mu)\mathbf{T}_1 & \mathbf{\Gamma}_N(-\mu)\mathbf{T}_1 & \mathbf{1}_K\\ e^{j2\pi\mu K}\mathbf{\Gamma}_K(\mu)\mathbf{T}_1 & e^{-j2\pi\mu K}\mathbf{\Gamma}_K(-\mu)\mathbf{T}_1 & \mathbf{1}_K\\ \vdots & \vdots & \vdots\\ e^{j2\pi\mu(P-1)K}\mathbf{\Gamma}_K(\mu)\mathbf{T}_1 & e^{-j2\pi\mu(P-1)K}\mathbf{\Gamma}_K(-\mu)\mathbf{T}_1 & \mathbf{1}_K\end{bmatrix},\\&=\mathbf{\Phi}_2\mathbf{\Phi}_1\end{aligned}\tag{5.49}$$

where

$$\mathbf{\Phi}_1=\begin{bmatrix}\mathbf{\Gamma}_K(\mu)\mathbf{T}_1 & \mathbf{O}_{K\times 2L_f+1} & \mathbf{0}_K\\ \mathbf{O}_{K\times 2L_f+1} & \mathbf{\Gamma}_K(-\mu)\mathbf{T}_1 & \mathbf{0}_K\\ \mathbf{0}_{2L_f+1}^T & \mathbf{0}_{2L_f+1}^T & \sqrt{K}\end{bmatrix},\tag{5.50}$$

and

$$\mathbf{\Phi}_2=\begin{bmatrix}\mathbf{I}_K & \mathbf{I}_K & 1/\sqrt{K}\cdot\mathbf{1}_K\\ e^{j2\pi\mu K}\mathbf{I}_K & e^{-j2\pi\mu K}\mathbf{I}_K & 1/\sqrt{K}\cdot\mathbf{1}_K\\ \vdots & \vdots & \vdots\\ e^{j2\pi\mu(P-1)K}\mathbf{I}_K & e^{-j2\pi\mu(P-1)K}\mathbf{I}_K & 1/\sqrt{K}\cdot\mathbf{1}_K\end{bmatrix}.\tag{5.51}$$

Thus, $\mathbf{\Phi}_1^H\mathbf{\Phi}_1=K\cdot\mathbf{I}_{4L_f+3}$ and $\mathbf{\Phi}_2^H\mathbf{\Phi}_2=P\cdot\mathbf{I}_{2K+1}$ constitute a sufficient condition of (5.45). Furthermore, the condition $\mathbf{\Phi}_1^H\mathbf{\Phi}_1=K\cdot\mathbf{I}_{4L_f+3}$, called Condition A, splits into the following three sub-conditions:

Condition A.1: $\mathbf{S}_1^H\mathbf{S}_1=K\cdot\mathbf{I}_{L_f}$, (5.52)

Condition A.2: $\mathbf{S}_1^T\mathbf{S}_1=\mathbf{O}_{L_f\times L_f}$, (5.53)

and

Condition A.3 : $\mathbf{S}_1^T \mathbf{1}_N = \mathbf{0}_{L_f}$. (5.54)

In [81], methods were given to design sequences that satisfy Conditions (A.1) and (A.2), whereas (A.3) just says that the designed sequence has a zero mean. As an example, using the frequency-domain nulling (FDN) method in [81], the sequence ($K = 64$)

$$s(n) = \frac{1}{N} \sum_{k=0}^{63} S(k) e^{\frac{j2\pi nk}{64}}, \quad n = 0, \cdots, 63 \tag{5.55}$$

with

$$S(k) = \begin{cases} e^{j\phi_k}, \text{ arbitrary } \phi_k, \text{ for } k \in J, \text{ and } J = [1,5,9,\ldots,61] \\ 0 \quad , \text{ for } k \notin J \end{cases}.$$

can be shown to satisfy Conditions (A.1), (A.2) and (A.3).

The condition $\mathbf{\Phi}_2^H \mathbf{\Phi}_2 = P \cdot \mathbf{I}_{2K+1}$, called Condition B, amounts to

$$\begin{bmatrix} P \cdot \mathbf{I}_K & \gamma_1(-\mu) \cdot \mathbf{I}_K & \gamma_2(-\mu) \cdot \mathbf{1}_K \\ \gamma_1(\mu) \cdot \mathbf{I}_K & P \cdot \mathbf{I}_K & \gamma_2(\mu) \cdot \mathbf{1}_K \\ \gamma_2(\mu) \cdot \mathbf{1}_K^H & \gamma_2(-\mu) \cdot \mathbf{1}_K^H & P \end{bmatrix} = P \cdot \mathbf{I}_{2K+1}, \tag{5.56}$$

where

$$\gamma_1(\mu) = \frac{1 - \left(e^{j4\pi\mu K}\right)^P}{1 - e^{j4\pi\mu K}} = 0, \tag{5.57}$$

and

$$\gamma_2(\mu) = \frac{1}{\sqrt{K}} \cdot \frac{1 - \left(e^{j2\pi\mu K}\right)^P}{1 - e^{j2\pi\mu K}} = 0 \tag{5.58}$$

which lead to

$$\mu = \frac{k}{PK}, \ \{k \in Z \,|\, k \notin iP/2, \ i \in Z\}. \tag{5.59}$$

From (5.52)-(5.54) and (5.59), μ is designed separately from $s(n)$ in our method.

5.3 Performance Analysis

In this section, the mean and variance of the calibrated $IRR_T(f)$, $IRR_R(f)$, ε_T and ε_R are analyzed, and numerical results in Section V confirms the accuracy of the analysis. To begin with, define $\Upsilon \doteq \left[\Upsilon_{\mathbf{f}_{1,+}}^T \quad \Upsilon_{\mathbf{f}_{1,-}}^T \quad \Upsilon_{b_1}^T \quad \Upsilon_{\mathbf{f}_{2,+}}^T \quad \Upsilon_{\mathbf{f}_{2,-}}^T \quad \Upsilon_{b_2}^T \quad \Upsilon_{d_0}^T \right]^T$. Then, from (5.40), the estimates $\hat{\mathbf{f}}_{1,+}$, $\hat{\mathbf{f}}_{1,-}$, $\hat{\mathbf{f}}_{2,-}$, $\hat{b}_1$, and $\hat{d}_0$ are expressed as

$$\hat{\mathbf{f}}_{1,+} = \mathbf{f}_{1,+} + \Upsilon_{\mathbf{f}_{1,+}} \mathbf{v} \ , \tag{5.60}$$

$$\hat{\mathbf{f}}_{1,-} = \mathbf{f}_{1,-} + \Upsilon_{\mathbf{f}_{1,-}} \mathbf{v} \ , \tag{5.61}$$

$$\hat{b}_1 = b_1 + \Upsilon_{b_1} \mathbf{v} \, , \tag{5.62}$$

$$\hat{\mathbf{f}}_{2,-} = \mathbf{f}_{2,-} + \Upsilon_{\mathbf{f}_{2,-}} \mathbf{v} \, , \tag{5.63}$$

and

$$\hat{d}_0 = d_0 + \Upsilon_{d_0} \mathbf{v} \, . \tag{5.64}$$

During the internal loop-back, the signal-to-noise ratio (SNR) is usually very high ($\mathrm{SNR} \gg 1$), and, hence, it is reasonable to assume that $\hat{\mathbf{f}}_{1,+} \approx \mathbf{f}_{1,+}$. Recall that $\mathbf{f}_{1,+}$ is the desired channel response from transmitter to receiver, as is given in (5.25). The approximation is accurate for the SNRs of interest as to be discussed in Section 5.4. Using this approximation, the estimated calibration filters $\hat{w}(n)$ and $\hat{\rho}(n)$ in (5.30) and (5.32) are rewritten as

$$\begin{aligned} \hat{w}(n) &= -\hat{f}_{1,+}^{-1}(n) \otimes \hat{f}_{1,-}(n) \\ &\approx -f_{1,+}^{-1}(n) \otimes \hat{f}_{1,-}(n), \end{aligned} \tag{5.65}$$

and

$$\hat{\rho}(n) \approx \left(f_{1,+}^{*}(n) e^{-j2\pi\mu n} \right)^{-1} \otimes \left(\hat{f}_{2,-}(n) e^{-j2\pi\mu n} \right), \tag{5.66}$$

respectively. In addition, $\hat{w}(n)$ and $\hat{\rho}(n)$ will be modeled as FIR filters by truncating the in-

verse filters $f_{1,+}^{-1}(n)$ and $\left(f_{1,+}^{*}(n)e^{-j2\pi\mu n}\right)^{-1}$; that is $\hat{\mathbf{w}}=\left[\hat{w}(0),\hat{w}(1),\ldots,\hat{w}(L-1)\right]^{T}$, and $\hat{\boldsymbol{\rho}}=\left[\hat{\rho}(0),\hat{\rho}(1),\ldots,\hat{\rho}(L-1)\right]^{T}$, where L is the filter length that can be selected as long as one wishes for the desirable accuracy. Using this modeling, $\hat{w}(n)$ and $\hat{\rho}(n)$ can be rearranged in the following vector-matrix form

$$\begin{aligned}\hat{\mathbf{w}} &\approx -\mathbf{F}_{1,+}(0)\times\hat{\mathbf{f}}_{1,-}\\ &=-\mathbf{F}_{1,+}(0)\times\left(\mathbf{f}_{1,-}+\boldsymbol{\Upsilon}_{\mathbf{f}_{1,-}}\mathbf{v}\right)\\ &=\mathbf{w}+\mathbf{v}_{w},\end{aligned}\tag{5.67}$$

and

$$\begin{aligned}\hat{\boldsymbol{\rho}} &\approx \mathbf{F}_{1,+}^{*}(\mu)\times\boldsymbol{\Gamma}_{L_f}(-\mu)\hat{\mathbf{f}}_{2,-}\\ &=\mathbf{F}_{1,+}^{*}(\mu)\times\boldsymbol{\Gamma}_{L_f}(-\mu)\left(\mathbf{f}_{2,-}+\boldsymbol{\Upsilon}_{\mathbf{f}_{2,-}}\mathbf{v}\right)\\ &=\boldsymbol{\rho}+\mathbf{v}_{\rho},\end{aligned}\tag{5.68}$$

where

$$\mathbf{w}=-\mathbf{F}_{1,+}(0)\mathbf{f}_{1,-},\tag{5.69}$$

$$\mathbf{v}_{w}=\left[v_{w}(0),v_{w}(1),\ldots,v_{w}(L-1)\right]^{T}=-\mathbf{F}_{1,+}(0)\boldsymbol{\Upsilon}_{\mathbf{f}_{1,-}}\mathbf{v},\tag{5.70}$$

$$\boldsymbol{\rho}=\mathbf{F}_{1,+}^{*}(\mu)\boldsymbol{\Gamma}_{L_f}(-\mu)\mathbf{f}_{2,-},\tag{5.71}$$

$$\begin{aligned}\mathbf{v}_{\rho} &=\left[v_{\rho}(0),v_{\rho}(1),\ldots,v_{\rho}(L-1)\right]^{T}\\ &=\mathbf{F}_{1,+}^{*}(\mu)\boldsymbol{\Gamma}_{L_f}(-\mu)\boldsymbol{\Upsilon}_{\mathbf{f}_{2,-}}\mathbf{v},\end{aligned}\tag{5.72}$$

and $\mathbf{F}_{1,+}(\mu)$ is the $L\times L_f$ sized convolution matrix constructed by the truncation inverse filter $\left(f_{1,+}(n)e^{j2\pi\mu n}\right)^{-1}$.

To analyze the calibrated $IRR_T(f)$, firstly, $g_{T,+}(n)$ and $g_{T,-}(n)$ are approximated as

$$\begin{aligned}g_{T,+}(n) &=h_{T,+}(n)+\hat{w}^{*}(n)\otimes h_{T,-}(n)\\ &\approx h_{T,+}(n)\end{aligned}\tag{5.73}$$

and

$$
\begin{aligned}
g_{T,-}(n) &= \underbrace{h_{T,-}(n) + w(n) \otimes h_{T,+}(n)}_{\approx 0} + v_w(n) \otimes h_{T,+}(n) \\
&\approx v_w(n) \otimes h_{T,+}(n),
\end{aligned}
\tag{5.74}
$$

respectively. The approximation in (5.73) is good because $|h_{T,+}(n)| \gg |h_{T,-}(n)|$ for the radio impairments in real systems (see Section 5.4), and (5.74) is good because ideally $w(n)$ is sought to make $h_{T,-}(n) + w(n) \otimes h_{T,+}(n) = 0$ in our method. Next, using the approximations in (5.73) and (5.74), the calibrated $IRR_T(f)$ in (5.14) is approximated as

$$
\begin{aligned}
IRR_T(f) &\approx 10 \log_{10} \frac{|H_{T,+}(f)|^2}{|H_{T,+}(f)|^2 |V_w(f)|^2} \\
&= -10 \log_{10} |V_w(f)|^2 \ \text{(dB)}
\end{aligned}
\tag{5.75}
$$

where $V_w(f)$ is the Fourier transform of $v_w(n)$. Furthermore, according to (5.70), we have

$$
\begin{aligned}
V_w(f) &\approx -\boldsymbol{\psi}^H(f) \mathbf{F}_{1,+}(0) \boldsymbol{\Upsilon}_{\mathbf{f}_{1,-}} \mathbf{G} \mathbf{v}_0 \\
&= \boldsymbol{\chi}_w^H(f) \mathbf{v}_0
\end{aligned}
\tag{5.76}
$$

where

$$
\mathbf{v}_0 = \left[v_0(0), v_0(1), \ldots, v_0(N-1) \right]^T,
$$

$$
\boldsymbol{\psi}(f) = \left[1, e^{j2\pi f}, \ldots, e^{j2\pi f(L-1)} \right]^T,
$$

$$
\boldsymbol{\chi}_w^H(f) = -\boldsymbol{\psi}^H(f) \mathbf{F}_{1,+}(0) \boldsymbol{\Upsilon}_{\mathbf{f}_{1,-}} \mathbf{G},
$$

and $\mathbf{G}$ is the $N \times N$ sized convolution matrix constructed by $h_{R,+}(n)$. In (5.76), we have used the approximation $v(n) \approx h_{R,+}(n) \otimes v_0(n)$ because $|h_{R,+}(n)| \gg |h_{R,-}(n)|$ in real systems. Since $\{v_0(n)\}_{n=0}^{N-1}$ are zero mean i.i.d complex circular symmetric Gaussian variables with variance σ_0^2, $|V_w(f)|^2$ is an exponentially distributed random variable with the probability density function (pdf) given by

$$p\left(\left|V_w(f)\right|^2\right)=\frac{1}{\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2}e^{-\frac{1}{\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2}\left|V_w(f)\right|^2}. \tag{5.77}$$

Using (5.77) and table of integral [82] (see 4.331.1 and 4.335.1 in pages 571 and 572),

$$\int_0^\infty e^{-\mu x}\ln x dx=-\frac{1}{\mu}(\mathbf{C}+\ln\mu),$$

and

$$\int_0^\infty e^{-\mu x}(\ln x)^2 dx=\frac{1}{\mu}\left[\frac{\pi^2}{6}+(\mathbf{C}+\ln\mu)^2\right],$$

where $\mathbf{C}=0.577215...$ and $\mu>0$, one gets

$$\begin{aligned}\mathrm{E}\left\{IRR_T(f)\right\}&=-10\cdot\mathrm{E}\left\{\log_{10}\left|V_w(f)\right|^2\right\}\\&=-10\log_{10}\frac{\boldsymbol{\chi}_w^H(f)\boldsymbol{\chi}_w(f)\sigma_0^2}{e^{\mathbf{C}}}\mathrm{dB},\end{aligned} \tag{5.78}$$

$$\begin{aligned}\mathrm{E}\left\{\left(IRR_T(f)\right)^2\right\}&=100\cdot\mathrm{E}\left\{\left(\log_{10}\left|V_w(f)\right|^2\right)^2\right\}\\&=\left(10\log_{10}e^{\frac{\pi}{\sqrt{6}}}\right)^2+\left(\mathrm{E}\left\{IRR_T(f)\right\}\right)^2,\end{aligned} \tag{5.79}$$

and $\mathrm{VAR}\left\{IRR_T(f)\right\}=\left(10\log_{10}e^{\frac{\pi}{\sqrt{6}}}\right)^2$, where $e^{\mathbf{C}}=1.781072...$ is the Euler's constant.

Similarly, to analyze the calibrated $IRR_R(f)$, $g_{R,+}(n)$ and $g_{R,-}(n)$ are approximated as

$$\begin{aligned}g_{R,+}(n)&=h_{R,+}(n)-\hat{\rho}(n)\otimes h_{R,-}^*(n)\\&\approx h_{R,+}(n)\end{aligned} \tag{5.80}$$

and

$$\begin{aligned}g_{R,-}(n)&=\underbrace{h_{R,-}(n)-\rho(n)\otimes h_{R,+}^*(n)}_{\approx 0}-v_\rho(n)\otimes h_{R,+}^*(n)\\&\approx -v_\rho(n)\otimes h_{R,+}^*(n).\end{aligned} \tag{5.81}$$

Thus, the calibrated $IRR_R(f)$ in (5.22) is estimated as

$$IRR_R(f) \approx 10\log_{10}\frac{\left|H_{R,+}(f)\right|^2}{\left|H_{R,+}^*(-f)\right|^2} - 10\log_{10}\left|V_\rho(f)\right|^2 \text{ dB} \tag{5.82}$$

where

$$\begin{aligned} V_\rho(f) &\approx \boldsymbol{\psi}^H(f)\mathbf{F}_{1,+}^*(\mu)\boldsymbol{\Gamma}_{L_f}(-\mu)\boldsymbol{\Upsilon}_{\mathbf{f}_{2,-}}\mathbf{G}\mathbf{v}_0 \\ &= \boldsymbol{\chi}_\rho^H(f)\mathbf{v}_0, \end{aligned} \tag{5.83}$$

and

$$\boldsymbol{\chi}_\rho^H(f) = \boldsymbol{\psi}^H(f)\mathbf{F}_{1,+}^*(\mu)\boldsymbol{\Gamma}_{L_f}(-\mu)\boldsymbol{\Upsilon}_{\mathbf{f}_{2,-}}\mathbf{G}\,.$$

Consequently, it can be shown that

$$\mathrm{E}\left\{IRR_R(f)\right\} = 10\log_{10}\frac{\left|H_{R,+}(f)\right|^2}{\left|H_{R,+}^*(-f)\right|^2} - 10\cdot\mathrm{E}\left\{\log_{10}\left|V_\rho(f)\right|^2\right\} \tag{5.84}$$

with

$$\mathrm{E}\left\{\log_{10}\left|V_\rho(f)\right|^2\right\} = \log_{10}\frac{\boldsymbol{\chi}_\rho^H(f)\boldsymbol{\chi}_\rho(f)\sigma_0^2}{e^{\mathbf{c}}}\,. \tag{5.85}$$

and $\mathrm{VAR}\left\{IRR_R(f)\right\} = \left(10\log_{10}e^{\frac{\pi}{\sqrt{6}}}\right)^2 = \mathrm{VAR}\left\{IRR_T(f)\right\}$.

Next, we analyze ε_T and ε_R. From (5.31) and (5.62), $\hat{b}$ can be approximated as

$$\begin{aligned} \hat{b} &\approx -\hat{b}_1 \otimes f_{1,+}^{-1}(n) \\ &= \frac{-b_1 - \boldsymbol{\Upsilon}_{b_1}\mathbf{v}}{\sum_n f_{1,+}(n)} \\ &\approx b + v_b \end{aligned} \tag{5.86}$$

where

$$v_b = -\frac{\boldsymbol{\Upsilon}_{b_1}\mathbf{v}}{\sum_n f_{1,+}(n)}\quad . \tag{5.87}$$

In addition, from (5.11), (5.86), and (5.73), Δb is approximated as

$$\begin{aligned} \Delta b &= \underbrace{g_{T,+}(n)\otimes b + b_0}_{\approx 0} + g_{T,-}(n)\otimes\left(b^* + v_b^*\right) + g_{T,+}(n)\otimes v_b \\ &\approx g_{T,+}(n)\otimes v_b \approx h_{T,+}(n)\otimes v_b \end{aligned}$$

$$
\begin{aligned}
&= -\frac{\left(\sum_n h_{T,+}(n)\right)\cdot \Upsilon_{b_1}\mathbf{v}}{\sum_n f_{1,+}(n)} \\
&\approx \boldsymbol{\chi}_{\Delta b}^H \mathbf{v}_0,
\end{aligned}
\tag{5.88}
$$

where

$$
\boldsymbol{\chi}_{\Delta b}^H = -\frac{\left(\sum_n h_{T,+}(n)\right)\cdot \Upsilon_{b_1}\mathbf{G}}{\sum_n f_{1,+}(n)}. \tag{5.89}
$$

Again, $|\Delta b|^2$ forms an exponentially distributed random variable with pdf

$$
p\left(|\Delta b|^2\right) = \frac{1}{\boldsymbol{\chi}_{\Delta b}^H \boldsymbol{\chi}_{\Delta b}\sigma_0^2} e^{-\frac{1}{\boldsymbol{\chi}_{\Delta b}^H \boldsymbol{\chi}_{\Delta b}\sigma_0^2}|\Delta b|^2}, \tag{5.90}
$$

and the expectation of ε_T in (5.15) is given by

$$
\begin{aligned}
\mathrm{E}\{\varepsilon_T\} &= \mathrm{E}\left\{10\log_{10}\frac{|\Delta b|^2}{|b_0|^2}\right\} \\
&\approx 10\log_{10}\frac{\boldsymbol{\chi}_{\Delta b}^H \boldsymbol{\chi}_{\Delta b}\sigma_0^2}{|b_0|^2 e^{\mathbf{C}}}.
\end{aligned}
\tag{5.91}
$$

For ε_R, using (5.64), Δd in (5.19) can be approximated as

$$
\begin{aligned}
\Delta d &= \left(d_0 - \hat{d}\right) - \hat{\rho}(n) \otimes \left(d_0 - \hat{d}\right)^* \\
&\approx d_0 - \hat{d} \\
&= -\Upsilon_{d_0}\mathbf{v} \\
&\approx \boldsymbol{\chi}_{\Delta d}^H \mathbf{v}_0
\end{aligned}
\tag{5.92}
$$

where

$$
\boldsymbol{\chi}_{\Delta d}^H = -\Upsilon_{d_0}\mathbf{G}. \tag{5.93}
$$

Then, the expectation of ε_R in (5.23) is given by

$$
\begin{aligned}
\mathrm{E}\{\varepsilon_R\} &= \mathrm{E}\left\{10\log_{10}\frac{|\Delta d|^2}{|d_0|^2}\right\} \\
&\approx 10\log_{10}\frac{\boldsymbol{\chi}_{\Delta d}^H \boldsymbol{\chi}_{\Delta d}\sigma_0^2}{|d_0|^2 e^{\mathbf{C}}}.
\end{aligned}
\tag{5.94}
$$

Similarly, it can be shown that $\mathrm{VAR}\{\varepsilon_T\} = \mathrm{VAR}\{\varepsilon_R\} = \left(10\log_{10} e^{\frac{\pi}{\sqrt{6}}}\right)^2$.

Some observations are in order based on our simplified analysis. First, it is observed that the estimates $\hat{\mathbf{f}}_{1,+}$, $\hat{\mathbf{f}}_{1,-}$, $\hat{\mathbf{f}}_{2,-}$, $\hat{b}_1$, and $\hat{d}_0$ in (5.60)-(5.64) are all zero mean with different variances, and so are the estimated calibration parameters $\hat{\mathbf{w}}$, $\hat{\boldsymbol{\rho}}$ and $\hat{b}$ in (5.67), (5.68) and (5.86). Second, for the calibration performance indices $IRR_T(f)$, $IRR_R(f)$, ε_T and ε_R, they have different means as functions of operating SNRs in (5.78), (5.84), (5.91) and (5.94) but with the same variances equal to a constant $\left(10\log_{10} e^{\frac{\pi}{\sqrt{6}}}\right)^2 = 5.57^2$ independent of operating SNRs. This is an interesting analysis result which indeed agrees well with the simulation results as seen in next section.

Table 5.1: The RF impairments

RF Impairments	Parameter Value
Frequency independent I-Q imbalance (α_T, θ_T), (α_R, θ_R).	$(\alpha_T = 1.05,\ \theta_T = -5^o)$, $(\alpha_R = 1.08,\ \theta_R = 5^o)$.
Frequency dependent I-Q imbalance $\{h_T^I(n), h_T^Q(n)\}$, $\{h_R^I(n), h_R^Q(n)\}$.	I part : $[1 \quad 0.2 \quad 0.1 \quad 0.05]$ Q part : $[0.9 \quad 0.1 \quad 0.08 \quad 0.12]$
DC offset b_0 and d_0, with signal power normalized to 1	$b_0 = -0.1\times(1+\boldsymbol{j})/\sqrt{2}$, $d_0 = 0.1\times(1+\boldsymbol{j})/\sqrt{2}$.

5.4 Simulation Results

In this section, the performance of the proposed method is evaluated through analysis and computer simulations. Table 5.1 summarizes the transmitter and receiver RF impairments. In all the results, $1/T_s = 20\text{MHz}$, SNR $\doteq \frac{1}{N}\sum_{n=0}^{N}|s(n)|^2 \Big/ \sigma_0^2$, and the optimal training employs the sequence in (5.55) with $K = 64$, $P = 3$, and $\mu = 23/(3\cdot 64)$. Simulation results are obtained with 10^6 realizations

Figures 5.2 and 5.3 investigate the effects of L_f on the performance of the calibrated $\mathrm{E}[IRR_T(f)]$ and $\mathrm{E}[IRR_R(f)]$ by computer simulations, respectively. Recall that L_f is the length of the filters $f_{i,\pm}(n)$, $i = 1, 2$. As can be seen, the calibration performances are insensitive to the values of L_f as long as it is selected larger than 6 in this case; similarly results are observed for the dc-offset calibration. In practice, since L_f may not be exactly known in advance,

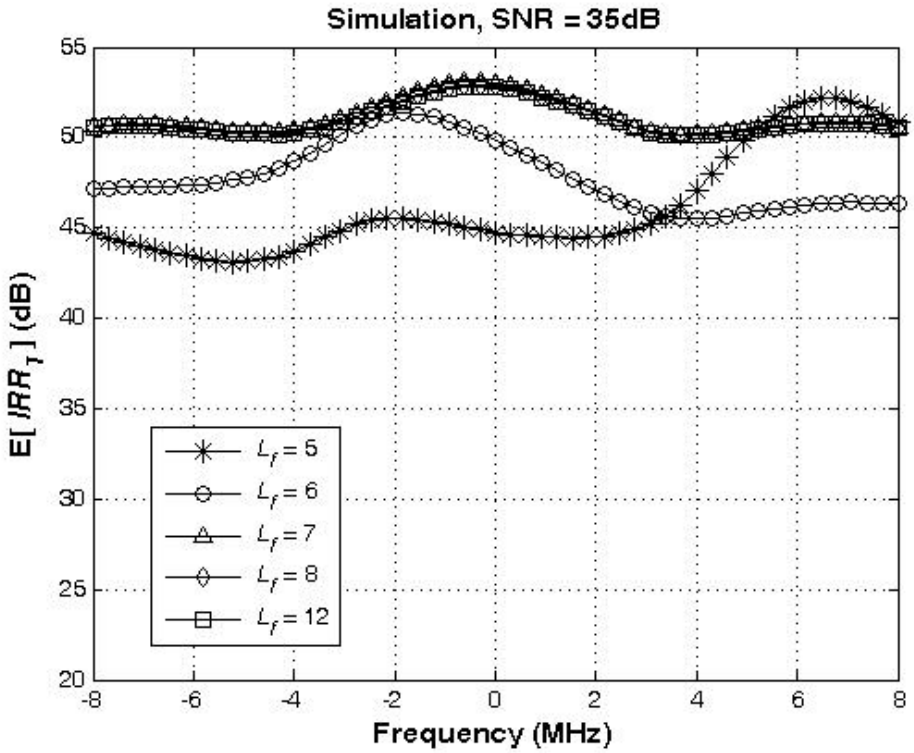


Figure 5.2: Performance of the calibrated $\mathrm{E}[IRR_T]$ with different L_f's under optimal training.

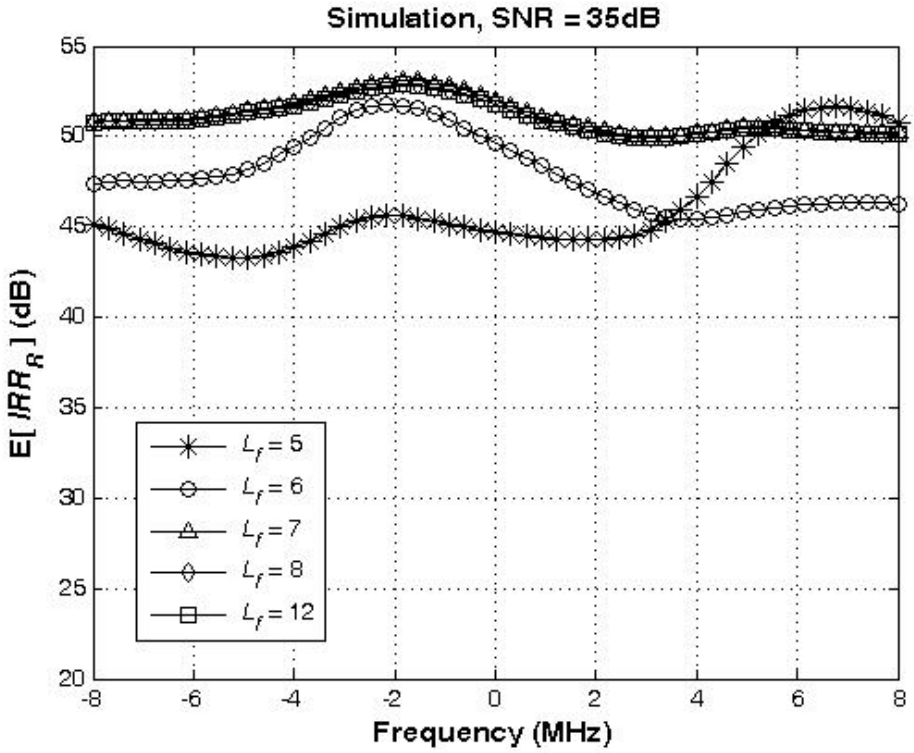


Figure 5.3: Performance of the calibrated $\mathrm{E}[IRR_R]$ with different L_f's under optimal training.

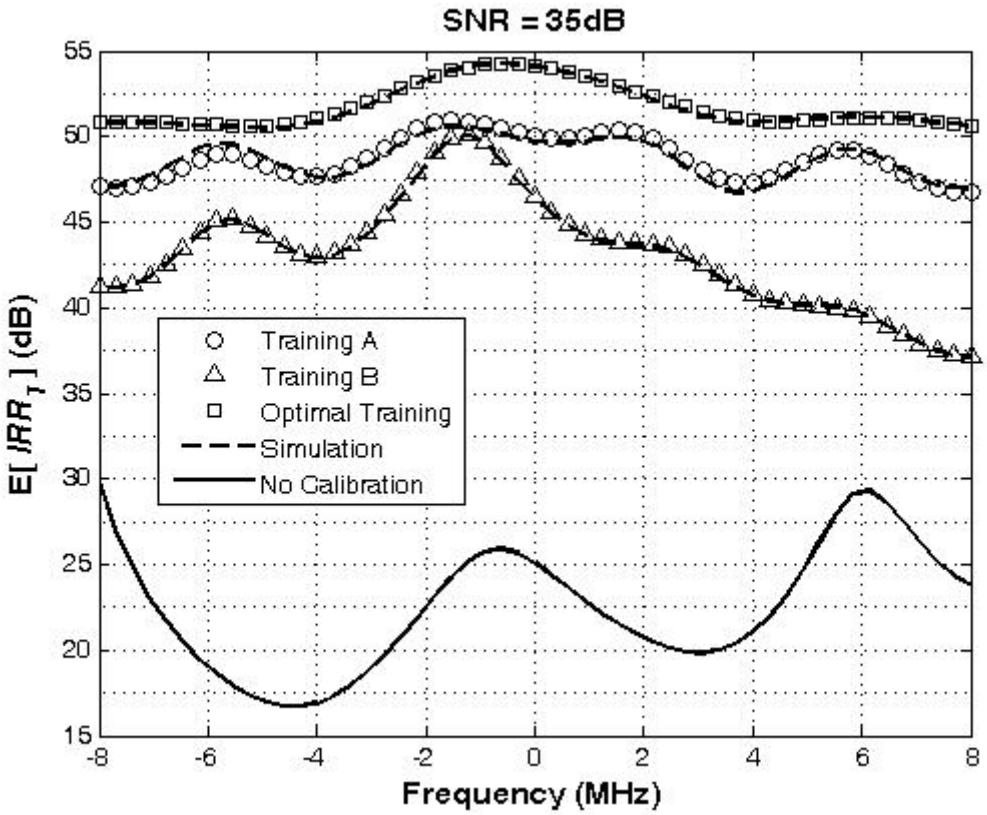


Figure 5.4: Performance of the calibrated $\mathrm{E}\left[IRR_T\right]$ with different training designs.

it is advisable to use a sufficiently large L_f to avoid performance degradation. In the rest of this section, $L_f = 7$ is used.

In Figures 5.4 and 5.5, the calibrated $\mathrm{E}\left[IRR_T(f)\right]$ and $\mathrm{E}\left[IRR_R(f)\right]$ are investigated with three periodic training designs ($K = 64$, $P = 3$): the optimal training, Training A and Training B. Training A uses the sequence in (5.55) but with $S(k) = e^{j\phi_k}\ \forall k$ and $\mu = 23/(3\cdot 64)$, and Training B uses the optimal sequence in (5.55) but with $\mu = 8/64$. Here, for comparison purpose, Training A is selected to violate Conditions A.2 and A.3 and Training B to violate Condition B. As is shown, the optimal training indeed gives a superior performance than the other two. Compared to the case of no calibration, the proposed method provides around 20-35 dB performance improvement over the whole frequency band. Shown in the figures also include the simulation

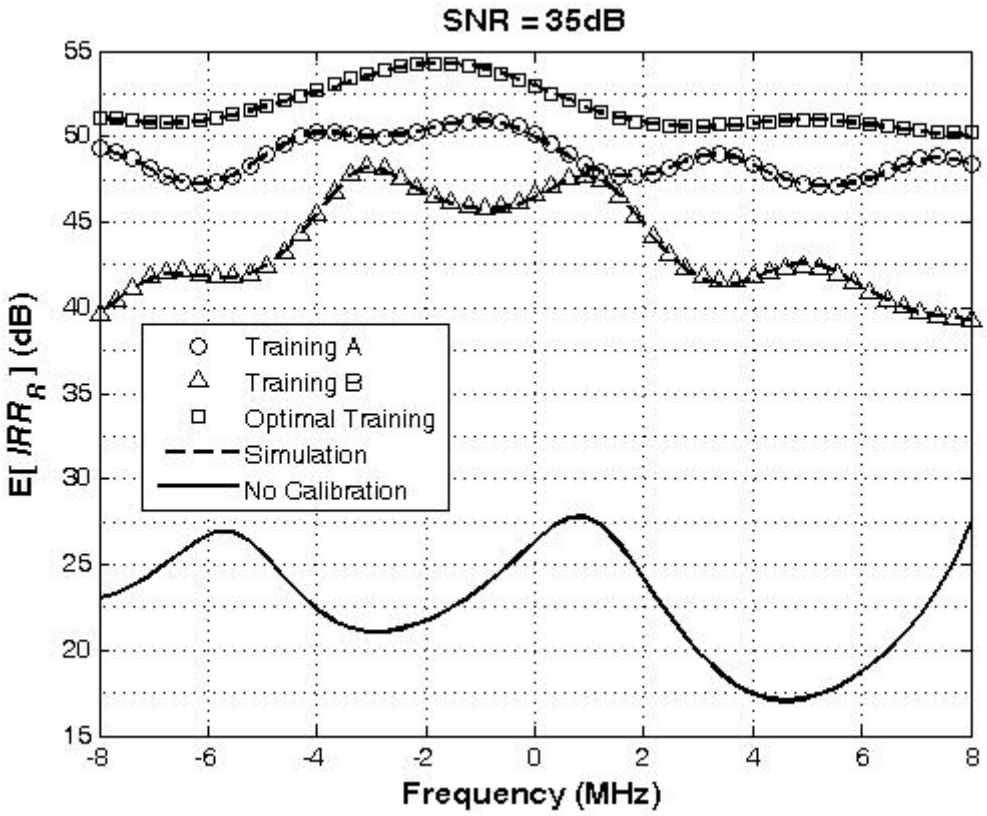


Figure 5.5: Performance of the calibrated $\mathrm{E}[IRR_R]$ with different training designs.

results which show perfect match to those with analysis.

Figure 5.6 shows the empirical cumulative distribution functions (CDFs) of the calibrated $IRR_T(f)$ and $IRR_R(f)$ at f=4 MHz, which are constructed with 10^6 realizations. The empirical and analytical means and standard deviations are given in Table 2 where it shows very good agreement between simulation and analysis. The smallest IRR_T and IRR_R observed are both 37 dB which are around 16 dB and 20 dB better than the cases of no calibration, respectively. Figure 5.7 shows the empirical CDFs of the calibrated ε_T and ε_R. The largest ε_T and ε_R observed are -26 dB and -24 dB, respectively. The empirical and analytical means and standard deviations of the calibrated ε_T and ε_R are also shown in Table 5.2. Again, it shows very good match between simulation and analysis. Furthermore, Table 5.2 confirms that the variances of the calibrated IRR_T, IRR_R, ε_T, and ε_R are same and independent of the operating SNRs.

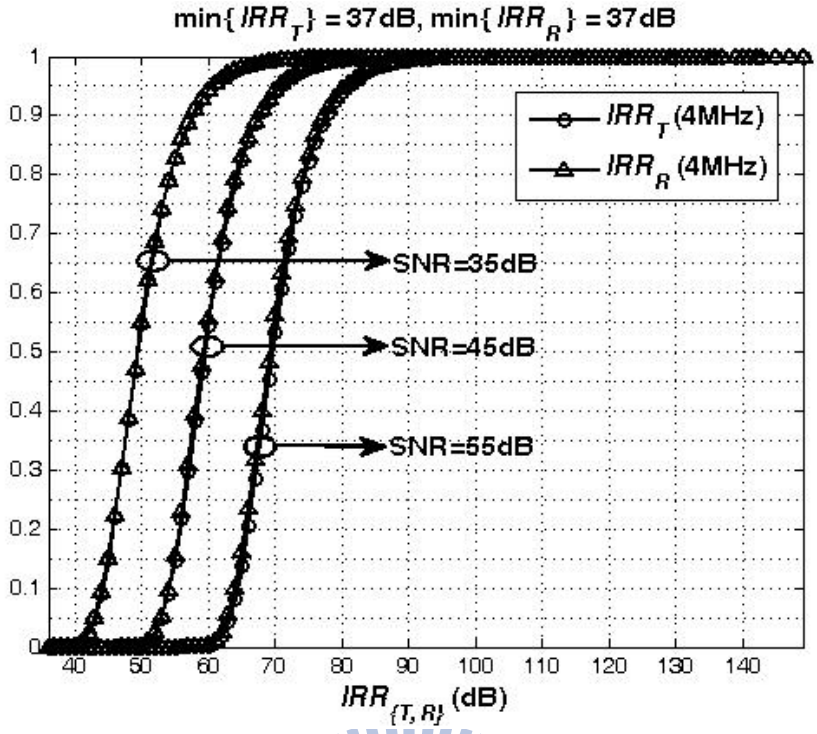


Figure 5.6: Empirical cumulative distribution functions of the calibrated IRR_T and IRR_R constructed with 10^6 realizations.

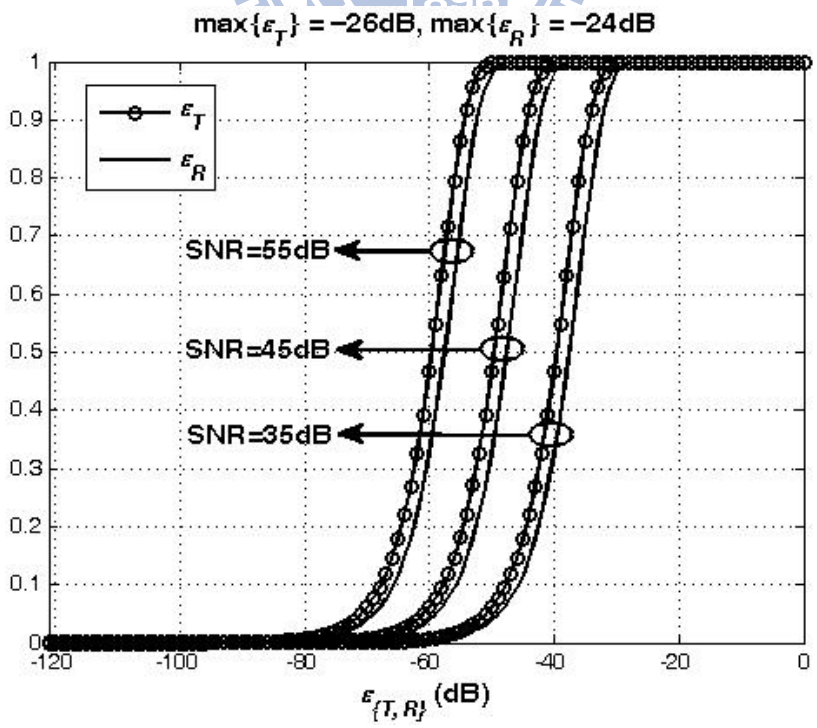


Figure 5.7: Empirical cumulative distribution functions of the calibrated ε_T and ε_R constructed with 10^6 realizations.

Table 5.2: Example mean and standard deviation of the calibrated IRR_T, IRR_R, ε_T, and ε_R.

Parameters	SNR(dB)	Mean (dB)		Standard deviation (dB)	
		Simulation	Analysis	Simulation	Analysis
IRR_T (4MHz)	35	50.8	50.9	5.59	5.57
	45	60.8	60.9	5.57	5.57
	55	70.8	70.9	5.56	5.57
IRR_R (4MHz)	35	50.8	50.8	5.53	5.57
	45	60.8	60.8	5.58	5.57
	55	70.8	70.8	5.58	5.57
ε_T	35	-40	-40.3	5.57	5.57
	45	-50	-50.3	5.57	5.57
	55	-60	-60.3	5.56	5.57
ε_R	35	-38	-38	5.6	5.57
	45	-48	-48	5.57	5.57
	55	-58	-58	5.59	5.57

Figures 8 and 9 show a sample received signal constellation and bit error rate performance respectively for an un-coded 64-QAM OFDM (orthogonal frequency-division multiplexing) system with and without calibration. A one-tap equalizer is employed at the receiver for the simulated OFDM system that uses 64-point FFT (fast Fourier transform) with 52 subcarriers carrying data. As are shown, the adverse effects due to radio impairments are removed almost completely by the proposed calibration.

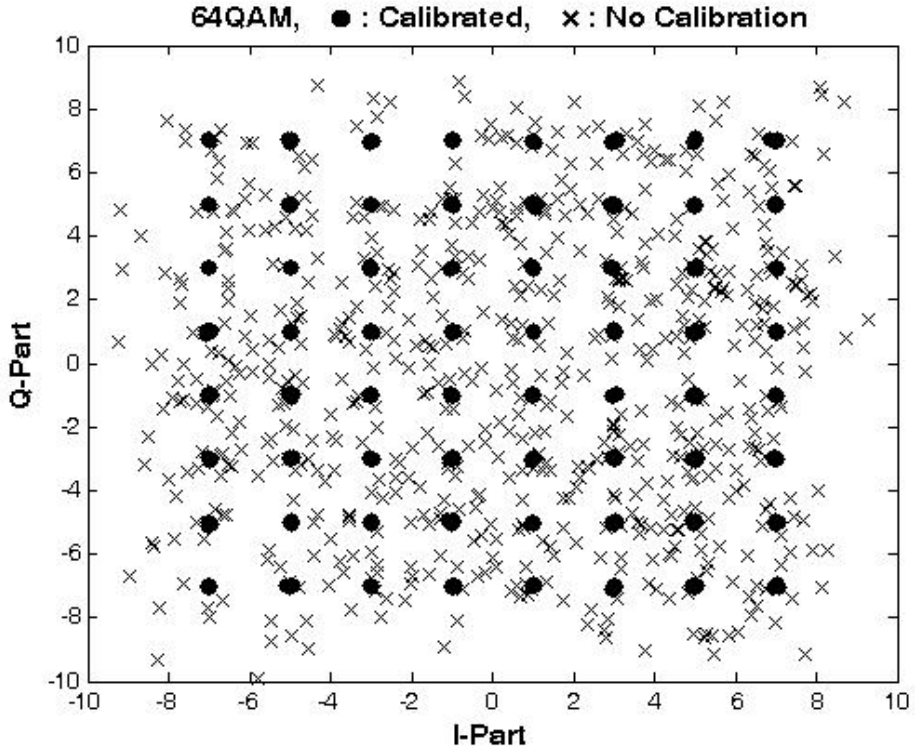


Figure 5.8: Sample signal constellation with and without calibrations ($\sigma_0^2 = 0$).

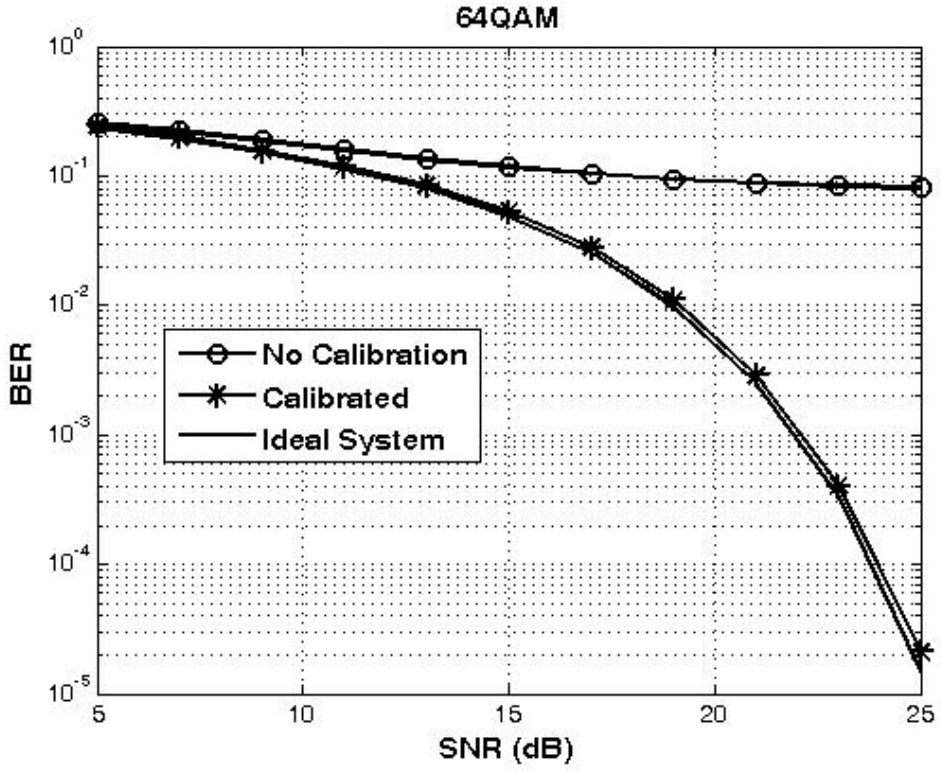
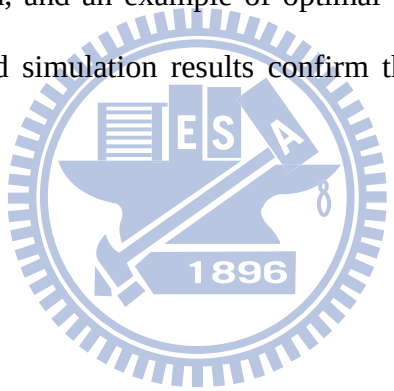


Figure 5.9: Bit error rate performance with and without calibration (64QAM).

5.5 Summary

In Chapter 5, a digital self-calibration method is proposed for the direct-conversion radio transceiver to calibrate its own transmitter and receiver radio impairments, including frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance, and dc offset. By introducing a shift between transmit and receive frequencies, the radio impairments appearing at the transmitter and receiver can be calibrated simultaneously without a dedicated analog circuitry in the feedback loop. The calibration parameters are estimated based on the non-linear least-squares principle, and the calibration performance is analyzed that agrees very well with the simulations. In addition, the issue of optimal training design is investigated; sufficient conditions for optimal training are provided, and an example of optimal training is given for the periodic training structure. Analytical and simulation results confirm the effectiveness of the proposed method.



Chapter 6

Conclusions

In this dissertation, the radio impairment estimation and compensation techniques are investigated for the wideband MIMO communication systems with direct-conversion radio architecture. A complete set of radio impairments is taken into consideration, including frequency-independent and dependent I-Q imbalances, dc offset and frequency offset. Both estimation/compensation and self-calibration techniques are studied in this dissertation.

The estimation/compensation technique is to remove the impairments from the received signal during communication at the receiving side. In Chapter 3 and Chapter 4, the estimation and cancellation techniques for MIMO systems are investigated, with Chapter 3 focusing on receiver radio impairments for any types of MIMO systems and extended to cascaded transmitter and receiver radio impairments for MIMO-OFDM systems in Chapter 4. First, a two-stage cancellation architecture is developed, which enables to explicitly cancel the radio impairments without increasing the dimension of signal detection. The two-stage architecture generalizes the cancellation architectures for various types of system configurations such as wireless peer-to-peer communication, downlink and uplink of a mobile cellular system. Moreover, it is general to accommodate different forms of MIMO operation including spatial multiplexing, STBC (space-time

block coded) and transmit beam forming, with any number of transmit and receive antennas. Second, several methods of estimation of radio parameters are proposed. The optimum method is the joint least squares estimation of all radio parameters, which is shown to be unbiased and approaches to CRLB for the SNRs of interest, but with the highest complexity. Several other methods are then developed aiming to reduce the estimation complexity, including the special phase-rotated periodic training design, simplified frequency and dc offset estimators and low-complexity iterative estimation aided by periodic training. Simulation results show that the un-coded BER performance degradation is negligible when using the reduced-complexity designs, and the proposed techniques outperform the existing ones in error-rate performance and/or the number of training symbols required.

On the other hand, the self-calibration is a technique to remove the transceiver's own radio impairments before communication. In Chapter 5, we propose to calibrate its own transmitter and receiver radio impairments, including frequency-independent I-Q imbalance, frequency-dependent I-Q imbalance, and dc offset. By introducing a shift between transmit and receive frequencies, the radio impairments appearing at the transmitter and receiver can be calibrated simultaneously without a dedicated analog circuitry in the feedback loop. Based on the time-domain approach, the proposed method is applicable to all types of communication systems. In addition, the issue of optimal training design is investigated; sufficient conditions for optimal training are provided with an example of optimal periodic training design. Analytical and simulation results confirm the effectiveness of the proposed self-calibration technique.

Some possible extensions and future research topics are addressed in the following. Both types of techniques can be further extended to combine with PA linearization and other synchronization issues such as signal detection, timing and clock offset estimation. The estimation methods can be investigated in the pilot-based systems without initial acquisition or training, and non-data-aided estimation technique can also be explored. Training sequence design to optimize the performance of the estimation/compensation technique might be another issue.

Bibliography

[1] R. V. Nee and R. Prasad, OFDM for Wireless Multimedia Communications. Artech House Publisher, 2000.

[2] G. J. Foschini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311-335, Mar. 1998.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. on Telecommunications*, vol. 10, pp. 585-595, Nov. 1999.

[4] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, 2005.

[5] A. A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE J. Solid-State Circuits*, vol. 30, pp. 1399–1410, Dec. 1995.

[6] B. Razavi, "Design consideration for direct-conversion radio receivers," *IEEE Trans. on Circuit Systems*, vol. 44, no. 6, pp. 428–435, Jun. 1997.

[7] B. Razavi, RF Microelectronics, Prentice-Hall, 1998.

[8] S. Mirabbasi and K. Martin, "Classical and modern receiver architectures," *IEEE Communications Magazine*, vol. 38, pp.132–139, Nov. 2000.

[9] M. Valkama, M. Renfors, and V. Koivunen, "Advanced methods for I/Q imbalance compensation in communication receivers," *IEEE Trans. on Signal Processing*, vol. 49, pp. 2335–2344, Oct. 2001.

[10] M. Valkama, M. Renfors, and V. Koivunen, "Compensation of frequency-selective I/Q imbalances in wideband receivers: models and algorithms," in *Proc. IEEE 3rd Workshop Signal Processing Advances in Wireless Communications (SPAWC)*, Taoyuan, Taiwan, pp. 42–45, Mar. 2001.

[11] K. P. Pun, J. E. Franca, C. Azeredo-Leme, C. F. Chan, and C. S. Choy, “Correction of frequency-dependent I/Q mismatches in quadrature receivers,” *Electronic Letters*, vol. 37, pp. 1415–1417, Nov. 2001.

[12] M. Valkama, M. Renfors, and V. Koivunen, “Blind signal estimation in conjugate signal models with application to I/Q imbalance compensation,” *IEEE Signal Processing Letters*, vol. 12, pp. 733-736, Nov. 2005.

[13] G. T. Gil, Y. D. Kim, and Y. H. Lee, “Non-data-aided approach to I/Q mismatch compensation in low-IF receivers,” *IEEE Trans. on Signal Processing*, vol. 55, no. 7, pp. 3360–3365, Jul. 2007.

[14] G. T. Gil, “Nondata-aided I/Q mismatch and DC offset compensation for direct-conversion receivers,” *IEEE Trans. on Signal Processing*, vol. 56, no. 7, pp. 2662–2668, Jul. 2008.

[15] L. Anttila, M. Valkama, and M. Renfors, “Circularity-based I/Q imbalance compensation in wideband direct-conversion receivers,” *IEEE Trans. on Vehicular Technology*, vol. 57, no. 4, pp. 2099–2113, Jul. 2008.

[16] B. Lindoff, “Using a direct conversion receiver in EDGE terminals: a new DC offset compensation algorithm,” in *Proc. PIMRC 2000*, pp. 959–963, Sep. 2000.

[17] D. Hui, B. Lindoff, and K. Zangi, “Enhanced DC estimation via sequence-specific frequency offset,” in *Proc. IEEE VTC 2002*, pp. 161–165, Sep. 2002.

[18] C. K. Ho, S. Sun, and P. He, “Low complexity frequency offset estimation in the presence of DC offset,” in *Proc. IEEE ICC’03*, pp. 2051–2055, May 2003.

[19] S. Marsili, “DC offset estimation in OFDM based WLAN application,” in *Proc. IEEE GLOBECOM’04*, pp. 3531–3535, Dec. 2004.

[20] H. Lin, X. Wang, and K. Yamashita, “A low-complexity carrier frequency offset estimator independent of DC offset, ” *IEEE Communications Letters*, vol. 12, no. 7, pp. 520-522, Jul. 2008.

[21] A. Schuchert, R. Hasholzner, and P. Antoine, “A novel IQ imbalance compensation scheme for the reception of OFDM signals,” *IEEE Trans. on Consumer Electronics*, vol. 47, no. 3, pp. 313–318, Aug. 2001.

[22] A. Tarighat, R. Bagheri, and A. H. Sayed, “Compensation schemes and performance analysis of IQ imbalance in OFDM receivers,” *IEEE Trans. on Signal Processing*, vol. 53, no. 8, pp. 3257–3268, Aug. 2005.

[23] J. Tubbax, B. Come, L. V. der Perre, S. Donnay, M. Engels, H. D. Man, and M. Moonen,

"Compensation of IQ imbalance and phase noise in OFDM systems," *IEEE Trans. on Wireless Communications*, vol. 4, no. 3, pp. 872–877, May 2005.

[24] J. Tubbax, B. Come, L. V. der Perre, S. Donnay, M. Engels, M. Moonen, and H. D. Man, "Joint compensation of IQ imbalance and frequency offset in OFDM systems," in *Proc. IEEE Globecom*, vol. 4, pp. 2365–2369, Dec. 2003.

[25] I. Barhumi and M. Moonen, "IQ-Imbalance compensation for OFDM in the presence of IBI and carrier-frequency offset," *IEEE Trans. on Signal Processing*, vol. 55, pp. 256–266, Jan. 2007.

[26] F. Horlin, A. Bourdoux, and L. Van der Perre, "Low-complexity EM-based joint acquisition of the carrier frequency offset and IQ imbalance,'' *IEEE Trans. on Wireless Communications*, vol . 7, no. 6, pp. 2212–2220, Jun. 2008.

[27] G. Xing, M. Shen, and H. Liu, "Frequency offset and I/Q Imbalance compensation for direct conversion receivers," *IEEE Trans. on Wireless Communications*, vol. 4, no. 2, pp. 673–680, Mar. 2005.

[28] H. Lin and K. Yamashita, "Subcarrier allocation based compensation for carrier frequency offset and I/Q imbalances in OFDM systems," *IEEE Trans. on Wireless Communications*, vol. 8, no. 1, pp. 18–23, Jan. 2009.

[29] H. Lin, X. Zhu, and K. Yamashita, "Low-complexity pilot-aided compensation for carrier frequency offset and I/Q imbalance," *IEEE Trans. on Communications*, vol. 58, no. 2, pp. 448–452, Feb. 2010.

[30] I. H. Sohn, E. R. Jeong, and Y. H. Lee, "Data-aided approach to I/Q mismatch and DC offset compensation in communication receivers," *IEEE Communications Letters*, vol. 6, pp. 547–549, Dec. 2002.

[31] G. T. Gil, I. H. Sohn, J. K. Park, and Y. H. Lee, "Joint ML estimation of carrier frequency, channel, I/Q mismatch, and DC offset in communication receivers," *IEEE Trans. on Vehicular Technology*, vol. 54, no. 1, pp. 338–349, Jan. 2005.

[32] K.-Y. Sung and C.-C. Chao, "Estimation and compensation of I/Q Imbalance in OFDM direct-conversion receivers," *IEEE J. Select. Topics Signal Processing*, vol. 3, pp. 438–453, Jun. 2009.

[33] A. Tarighat and A. H. Sayed, "MIMO OFDM receivers for systems with IQ imbalances," *IEEE Trans. on Signal Processing*, vol. 53, no. 9, pp. 3583 - 3596, September 2005.

[34] Y.-H. Chung and S.-M. Phoong, "Joint estimation of I/Q imbalance, CFO and channel response for MIMO OFDM systems," *IEEE Trans. on Communications*, vol. 58, pp.

1485–1492, May 2010.

[35] J. Feigin and D. Brady, “Joint transmitter/receiver I/Q imbalance compensation for direct conversion OFDM in packet-switched multipath environments,” *IEEE Trans. on Signal Processing*, vol. 56, pp. 4588-4593, Nov. 2009.

[36] D. Tandur, C.-Y. Lee, and M. Moonen, “Efficient compensation of RF impairments for OFDM systems,” in *Proc. IEEE WCNC 2009*, Apr. 2009.

[37] A. Tarighat and A. H. Sayed, “Joint compensation of transmitter and receiver impairments in OFDM systems, ” *IEEE Trans. on Wireless Communications*, vol. 6, pp. 240-246, Jan. 2007.

[38] H. Lin, X. Zhu, and K. Yamashita, “Hybrid domain compensation for analog impairments in OFDM systems,” in *Proc. IEEE GLOBECOM’08*, Nov. 2008.

[39] D. Tandur and M. Moonen, “Joint adaptive compensation of transmitter and receiver IQ Imbalance under carrier frequency offset in OFDM based systems,” *IEEE Trans. on Signal Processing*, vol. 55, no. 11, pp. 5246 – 5252, Nov. 2007.

[40] B. Narasimhan, D. Wang, S. Narayanan, H. Minn, and N. Al-Dhahir, “Digital compensation of frequency-dependent joint Tx/Rx I/Q imbalance in OFDM Systems under high mobility,” *IEEE J. Select. Topics Signal Processing*, vol. 3, pp. 405–417, Jun. 2009.

[41] H. Kamata, K. Sakaguchi, and K. Araki, “An effective IQ imbalance compensation scheme for MIMO-OFDM communication system," in *Proc. IEEE PIMRC’05*, Berlin, Germany, pp. 181-185, Sep. 2005.

[42] T. C. W. Schenk, P. F. M. Smulders, and E. R. Fledderus, “Estimation and compensation of frequency selective TX/RX IQ imbalance in MIMO OFDM systems,” in *Proc. ICC 2006*, pp. 251-256, June 2006.

[43] R. M. Rao and B. Daneshrad, “Analog impairments in MIMO-OFDM systems,” *IEEE Trans. on Wireless Communications*, vol. 5, pp. 3382–3387, Dec. 2006.

[44] Y. Zou, M. Valkama, and M. Renfors, “Digital compensation of IQ imbalance effects in space-time coded transmit diversity systems,” *IEEE Trans. on Signal Processing*, vol. 56, pp. 2496-2508, Jun. 2008.

[45] Y. Zou, M. Valkama, and M. Renfors, “Compensation of frequency-selective I/Q imbalances in space-time coded multi-antenna OFDM systems,” in *Proc. IEEE ISCCSP 2008*, pp. 123–128, Mar. 2008.

[46] M. Cao and H. Ge, “IQ imbalance mitigation for STBC MIMO-OFDM communication sys-

tems," in *Proc. IEEE ICASSP*, pp. 3093 – 3096, Apr. 2008.

[47] B. Narasimhan, S. Narayanan, H. Minn, and N. Al-Dhahir, "Reduced-complexity baseband compensation of joint Tx/Rx I/Q-imbalance in mobile MIMO-OFDM," *IEEE Trans. on Wireless Communications*, vol. 9, pp. 1720–1728, May 2010.

[48] D. Tandur and M. Moonen, "Compensation of RF impairments in MIMO OFDM systems," in *Proc. IEEE ICASSP*, pp. 3097–3100, Apr. 2008.

[49] D. Tandur and M. Moonen, "STBC MIMO OFDM systems with implementation impairments," in *Proc. IEEE VTC'2008-Fall*, Sep. 2008.

[50] J. K. Cavers and M. W. Liao, "Adaptive compensation for imbalance and offset losses in direct conversion transceivers," *IEEE Trans. on Vehicular Technology*, vol. 42, pp. 581–588, Nov. 1993.

[51] D. Hilborn, S. Stapleton, and J. Cavers, "An adaptive direct conversion transmitter," *IEEE Trans. on Vehicular Technology*, vol. 43, no. 2, pp. 223–233, May 1994.

[52] J. K. Cavers, "New methods for adaptation of quadrature modulators and demodulators in amplifier linearization circuits," *IEEE Trans. on Vehicular Technology*, vol. 46, no. 3, pp. 707–716, Aug. 1997.

[53] R. Marchesani, "Digital pre-compensation of imperfections in quadrature modulators," *IEEE Trans. on Communications*, vol. 48, no. 4, pp. 552–556, Apr. 2000.

[54] J. Tuthill and A. Cantoni, "Efficient compensation for frequency-dependent errors in analog reconstruction filters used in *I/Q* modulators," *IEEE Trans. on Communications*, vol. 53, pp. 489–496, Mar. 2005.

[55] M. Windisch and G. Fettweis, "Adaptive I/Q imbalance compensation in low-IF transmitter architectures," in *Proc. IEEE Veh. Technol. Conf.* (VTC-2004), Los Angeles, CA, pp. 2096–2100, Sep. 2004.

[56] L. Ding, Z. Ma, D. R.Morgan, M. Zierdt, and G. T. Zhou, "Frequency dependent modulator imbalance in pre-distortion linearization systems: Modeling and compensation," in *Proc. Asilomar Conf. Signals, Syst., Comp.*, Pacific Grove, CA, pp. 688–692, Nov. 2003.

[57] L. Anttila, M. Valkama, and M. Renfors, "Frequency-selective I/Q mismatch calibration of wideband direct-conversion transmitters," *IEEE Trans. on Circuits and Systems II*, vol. 55, pp. 359-363, Apr. 2008.

[58] Y. Zou, M. Valkama, and M. Renfors, "Pilot-based compensation of frequency-selective I/Q imbalances in direct-conversion OFDM transmitters," in *IEEE Veh. Technol. Conf.*

(VTC-F'08), Calgary, Canada, Sep. 2008.

[59] C-H. Liu, "Joint Tx and Rx IQ imbalance compensation of OFDM transceiver in mesh network," in *Proc. IEEE GLOBECOM*, New Orleans, Louisiana, Nov. 2008.

[60] S. A. Bassam, S. Boumaiza, and F. M. Ghannouchi, "Block-wise estimation of and compensation for IQ imbalance in direct-conversion transmitters," *IEEE Trans. on Signal Processing*, vol. 57, pp. 4970–4973, Dec. 2009.

[61] B. Debaillie, P. V. Wesemael, G. Vandersteen, and J. Craninckx, "Calibration of direct-conversion transceivers," *IEEE J. Select. Topics Signal Processing*, vol. 3, pp. 488–498, Jun. 2009.

[62] C.-J. Hsu, R. Cheng, and W.-H. Sheen, "Joint estimation and compensation of frequency, DC-offset, I-Q imbalance and channel in MIMO receivers," in *Proc. IEEE ICASSP*, pp. 2533–2536, Apr. 2009.

[63] C.-J. Hsu, R. Cheng, and W.-H. Sheen, "Joint least-squares estimation of frequency, DC-offset, I-Q imbalance and channel in MIMO receivers," *IEEE Trans. on Vehicular Technology*, vol. 58, pp. 2201–2213, Jun. 2009.

[64] C.-J. Hsu and W.-H. Sheen, "Joint estimation and compensation of transmitter and receiver radio impairments in MIMO-OFDM receivers," in *Proc. IEEE PIMRC'09*, Tokyo, Japan, pp. 2489–2493, Sep. 2009.

[65] C.-J. Hsu and W.-H. Sheen, "Compensation of cascaded radio impairments in MIMO-OFDM systems with direct-conversion architecture," to be presented in *Global Mobile Congress* (GMC'10), Shanghai, China, Oct. 2010.

[66] C.-J. Hsu and W.-H. Sheen, "Estimation and cancellation of transmitter and receiver radio impairments in MIMO-OFDM systems with direct-conversion architecture" submitted to *IEEE Trans. on Wireless Communications*, Sep. 2010.

[67] C.-J. Hsu and W.-H. Sheen, "'A new self-calibration method for transmitter and receiver radio impairments in direct-conversion architecture," to be presented in *Proc. IEEE GLOBECOM'10*, Dec. 2010.

[68] C.-J. Hsu and W.-H. Sheen, "A joint calibration of transmitter and receiver impairments in direct-conversion radio architecture," submitted to *IEEE Trans. on Communications*, Aug. 2010.

[69] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. on Communications*, vol. 48, pp. 1580–1589, Sep. 2000.

[70] T. Kailath, H. Vikalo, and B. Hassibi, MIMO Receiver Algorithms, in Space-Time Wireless Systems, edited by H. Bolcskei, D. Gesbert, C, Papadias, and A. van der Veen, Cambridge, 2006.

[71] J. Mendel, Lessons in Estimation Theory for Signal Processing, Communications and Control, Prentice-Hall, 1995.

[72] M. H. Meyers and L. E. Franks, "Joint carrier and timing recovery for PAM systems," *IEEE Trans. on Communications*, vol. 28, pp. 1121–1128, Aug. 1980.

[73] A. van den Bos, "A Cramer-Rao lower bound for complex parameters," *IEEE Trans. on Signal Processing*, vol.42, p. 2859, Oct. 1994.

[74] M. Mailand, R. Richter, and H.-J. Jentschel, "IQ-imbalance and its compensation for non-ideal analog receivers comprising frequency-selective components," *Advances in Radio Science*, Vol. 4, pp. 189–195, 2006.

[75] TGn Sync Group, *IEEE P802.11 Wireless LAN - TGn Sync Proposal Technical Specification*, Proposal of IEEE802.11n, IEEE Document 802.11-04/889r4, January 2005.

[76] C. Tellambura, M. G. Parker, Y. J. Guo, S. J. Shepherd, and S. K. Barton, "Optimal sequences for channel estimation using discrete Fourier transform techniques," *IEEE Trans. on Communications*, vol. 47, pp. 230-238, Feb. 1999.

[77] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. on Information Theory*, vol. 48, pp. 1804–1824, Jul. 2002.

[78] S. Haykin, Adaptive Filter Theory, fourth edition. Prentice-Hall, 2002.

[79] IEEE Std. 802.11a-1999, Part 11: Wireless LAN Medium Control (MAC) and Physical Layer (PHY) Specifications: High-speed Physical Layer in the 5GHz Band. LAN/MAN Standards Committee of the IEEE computer society, Sep. 16, 1999.

[80] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. on Wireless Communications*, vol. 5, pp. 1158–1168, May 2006.

[81] H. Minn and D. Munoz, "Pilot designs for channel estimation of OFDM systems with frequency-dependent I/Q imbalances," in *Proc. IEEE WCNC'09*, pp. 713–717, Apr. 2006.

[82] I. S. Gradshteyn, I. M. Ryzhik; A. Jeffrey, and A. Zwillinger, editors. Table of Integrals, Series, and Products, seventh edition. Academic Press, 2007.

Vita

Chen-Jui Hsu was born in Hsinchu, Taiwan, in April 1982. He received the B.S. degree in electronics engineering and the M.S. degree in communication engineering in 2004 and 2006, respectively, from the National Chiao Tung University, Hsinchu, Taiwan, where he is currently working toward the Ph.D. degree in communication engineering.

His current research interests include multiple-input multiple-output transceiver design, particularly in radio-frequency front-end impairment estimation and compensation algorithms.