UNIVERSITY OF CALIFORNIA

Los Angeles

Signal Processing for RF Distortion Compensation in Wireless Communication Systems

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical Engineering

by

Qiyue Zou

2008

UMI Number: 3349457

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

The dissertation of Qiyue Zou is approved.
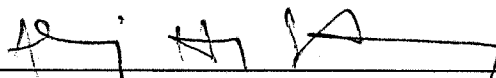
_____
Alan J. Laub

_____
Behzad Razavi

_____
Tsu-Chin Tsao

_____
Ali H. Sayed, Committee Chair

University of California, Los Angeles

2008

To my family.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my deep respect and heartfelt gratitude to my advisor, Professor Ali H. Sayed, for his valuable supervision, inspiring guidance and continuous encouragement throughout the course of my Ph.D. research. In particular, I thank him for offering me this opportunity to work with him in the areas of signal processing and communications. It is certainly a great learning experience for me to have had opportunities to discuss with him several technical topics and learn from him essential research skills and professional expertise. I have tremendously benefited from his countless efforts and dedication to the academic growth of his graduate students.

I would also like to thank my senior Dr. Alireza Tarighat for sharing his deep insights and contributing constructive opinions and broad perspectives that helped me refine the ideas in this dissertation.

I appreciate the help from my committee. Their lucid remarks were helpful to the completion of this work.

I wish to acknowledge the love, support and understanding of my family through these many years. Particularly, I thank my parents for inspiring me to pursue this challenge and my brother for his kindness.

VITA

1978   Born, Hanzhong, Shaanxi Province, China

2001   B.Eng. in Electrical and Electronic Engineering
                Nanyang Technological University
                Singapore

2001–2004   Research Engineer
                Centre for Signal Processing
                Nanyang Technological University
                Singapore

2004   M.Eng. in Electrical and Electronic Engineering
                Nanyang Technological University
                Singapore

2004–2008   Research Assistant
                Adaptive Systems Laboratory
                Department of Electrical Engineering
                University of California, Los Angeles
                Los Angeles, CA, USA

2005–2008   Teaching Assistant
                Department of Electrical Engineering
                University of California, Los Angeles
                Los Angeles, CA, USA

2008   M.A. in Mathematics

       University of California, Los Angeles

       Los Angeles, CA, USA

PUBLICATIONS

[1] Q. Zou, A. Tarighat, and A. H. Sayed, "Performance analysis and range improvement in multiband-OFDM UWB communications," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, pp. 629-632, Toulouse, France, May 2006.

[2] Q. Zou, A. Tarighat, N. Khajehnouri, and A. H. Sayed, "A phase noise compensation scheme for OFDM wireless systems," in *Proc. 14th European Signal Processing Conference (EUSIPCO)*, Florence, Italy, Sep. 2006.

[3] Q. Zou, A. Nosratinia, and A. H. Sayed, "Precoding for broadcasting with linear network codes," in *Proc. 44th Annual Allerton Conference on Communication, Control, and Computing*, Allerton, Illinois, Sep. 2006.

[4] Q. Zou, A. Tarighat, and A. H. Sayed, "Joint compensation of IQ imbalance and phase noise in OFDM systems," in *Proc. 40th Asilomar Conference on Signals, Systems and Computers*, pp. 1435–1439, Pacific Grove, California, Oct. 2006.

[5] Q. Zou, A. Tarighat, K. Y. Kim, and A. H. Sayed, "OFDM channel estimation in the presence of frequency offset, IQ imbalance, and phase noise," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, pp. 273–276, Honolulu, Hawaii, Apr. 2007.

[6] Q. Zou, A. Tarighat, and A. H. Sayed, "On the joint compensation of IQ imbalances and phase noise in MIMO-OFDM systems," in *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 37–40, New Orleans, Louisiana, May 2007.

[7] Q. Zou, A. Tarighat, and A. H. Sayed, "Performance analysis of multiband-OFDM UWB communications with application to range improvement," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3864–3878, Nov. 2007.

[8] Q. Zou, A. Tarighat, and A. H. Sayed, "Compensation of phase noise in OFDM wireless systems," *IEEE Trans. Signal Processing*, vol. 55, no. 11, pp. 5407–5424, Nov. 2007.

[9] Q. Zou, M. Mikhemar, and A. H. Sayed, "Digital compensation of RF nonlinearities in software-defined radios," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2921–2924, Las Vegas, Nevada, Apr. 2008.

[10] K. Y. Kim, Q. Zou, H. J. Choi, and A. H. Sayed, "An efficient carrier phase synchronization technique for high-order M-QAM-OFDM," *IEEE Trans. Signal Processing*, 2008, accepted for publication.

[11] Q. Zou, A. Tarighat, and A. H. Sayed, "Joint compensation of IQ imbalance and phase noise in OFDM wireless systems," *IEEE Trans. Commun.*, 2008, accepted for publication.

ABSTRACT OF THE DISSERTATION

Signal Processing for RF Distortion Compensation in Wireless Communication Systems

by

Qiyue Zou

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2008

Professor Ali H. Sayed, Chair

This dissertation addresses some challenging problems emerging from practical implementations of high-speed physical-layer communication schemes. The challenges are caused by the underlying physical-layer analog/RF (radio frequency) processing, which inevitably introduces distortions into the baseband demodulation stage. Such distortions arise, for example, in Orthogonal Frequency Division Multiplexing (OFDM) modulation, which has been adopted by several wireless standards. OFDM systems are known to be sensitive to the IQ imbalance and phase noise distortions introduced in the signal down-conversion process. These distortions seriously limit the applicable constellation size and reduce the attainable transmission data rate. Traditionally, the problem has been alleviated by increasing the transmission power of the carrier signal. In this dissertation, an alternative approach is pursued by exploiting signal processing techniques in the digital domain.

Chapter 2 develops a two-stage scheme to compensate for phase noise distortion in wireless OFDM systems. In the first stage, block-type pilot symbols

are transmitted and the channel coefficients are jointly estimated with the phase noise. In the second stage, comb-type OFDM symbols are transmitted such that the receiver can jointly estimate the data symbols and the phase noise. The joint effects of IQ imbalance and phase noise on OFDM systems are analyzed in Chapter 3, where compensation schemes are proposed and their performance is studied. In Chapter 4, the nonlinear effects caused by the front-end analog processing are investigated in the context of wideband software-defined radio systems. In the presence of strong blocker (interference) signals, such nonlinearities introduce severe cross modulation over the desired signal. The proposed digital solution scans the wide spectrum and locates the desired signal and blocker signals. After down-converting these signals separately to the baseband, they are jointly processed to mitigate the cross-modulation interferences.

It is shown that the proposed schemes can effectively mitigate these impairments, consequently, simplifying the RF and analog circuitry design in terms of implementation cost and power consumption. The proposed approach demonstrates how mixed-signal, i.e., joint analog and digital, processing techniques play a critical role in emerging radio technologies.

CHAPTER 1

Introduction

The demand for high-speed data communications through wireless links has experienced a rapid growth in recent years. Intensive efforts are being pursued to meet the increasing need for high data rates. New technologies continuously emerge to change the way we are sending and receiving information bits. Some representative examples of these new technologies are the multi-input multi-output (MIMO) transmission and reception schemes, ultra-wide band (UWB) communications, opportunistic and cooperative communications, and cognitive radio networks.

This dissertation addresses some challenging problems emerging from practical implementations of these high-speed physical-layer communication schemes. The challenges are caused by the underlying physical-layer analog/RF (radio frequency) processing, which inevitably introduces distortions into the baseband demodulation stage. Such distortions arise, for example, in Orthogonal Frequency Division Multiplexing (OFDM) modulation, which has been adopted by several wireless standards, including the IEEE 802.11a wireless local area network (WLAN) in the 5 GHz band, the IEEE 802.11g WLAN in the 2.4 GHz band, and the European digital video broadcasting system (DVB-T). Despite its popularity, OFDM systems are known to be sensitive to the IQ imbalance and phase noise distortions introduced in the signal down-conversion process. These distortions limit the applicable constellation size and reduce the attainable trans-

mission data rate. Traditionally, this problem has been alleviated by increasing the transmission power level of the carrier signal. In this dissertation, an alternative approach is pursued by exploiting signal processing techniques in the digital domain. Three common impairments in RF front-end are studied, namely, phase noise, IQ imbalance and nonlinearities. The next two sections give a brief introduction to the basic building blocks of an RF receiver and the mechanisms of these distortions.

1.1   RF Transmitter and Receiver

The block diagram of a generic RF transmitter is displayed in Fig. 1.1(a). The information bits are first mapped to proper channel codewords, which enable the receiver to correct certain transmission errors. The resulting bit sequence is mapped into constellation symbols, and the sequence of symbols is converted into a continuous-time baseband signal. The baseband signal is then upconverted to the passband by being mixed with the oscillator signal. Finally, the passband signal is amplified by the power amplifier and transmitted by the antenna.

At the receiver shown in Fig. 1.1(b), the RF signal acquired by the antenna is first processed by the band-selection surface acoustic wave (SAW) filter, whose purpose is to suppress interferences and noise outside the band of interest. After being amplified by the low-noise amplifier (LNA), the signal is down-converted to the baseband by the mixer. The analog-to-digital converter (ADC) samples the baseband signal. The transmitted information bits are eventually recovered by the baseband demodulation and decoding blocks.
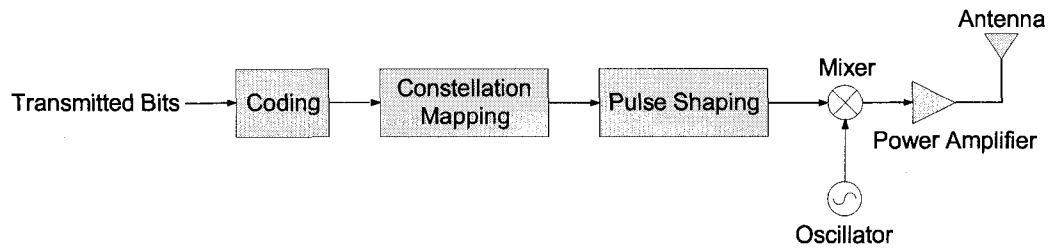
(a) Transmitter.
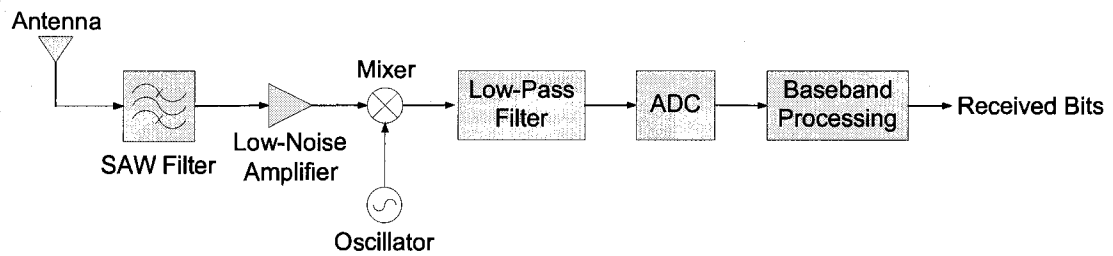
(b) Receiver.

Figure 1.1: Block diagram of an RF transmitter and receiver.

1.2 RF Distortions

According to the pre-defined functionality of each building block in the transmitter and receiver, RF circuit engineers implement the design to ensure that the system operates as desired. However, analog impairments often arise from unpredictable fabrication variations. Without cautious control over these impairments, the system performance deteriorates or even fails. This dissertation studies three common RF impairments that are considered critical limiting factors in real systems. The three impairments are phase noise, IQ imbalance, and nonlinearities.

1.2.1 Phase Noise

Oscillators used for signal down-conversion are susceptible to noise. Noise injected into an oscillator causes random deviation from its nominal frequency. Phase noise is the phase difference between the phase of the carrier signal and the phase of the local oscillator. There are mainly two types of oscillators used in practice, depending on whether or not they are used in a phase-locked loop (PLL). The so-called free-running oscillators operate without a PLL, and the generated phase noise is modeled as the accumulation of random frequency deviations and, hence, has unbounded variance. On the other hand, in a PLL oscillator, the closed-loop control mechanism tracks the phase variations of the carrier signal, and consequently, the generated phase noise has finite variance.

In a free-running oscillator, the frequency deviation $\epsilon(t)$ can be modeled as a zero-mean white Gaussian random process with single-sided power spectral density (PSD), say,

$$N_0 = \frac{\xi}{\pi},$$

where ξ is called the oscillator linewidth. The phase noise $\phi(t)$ generated by the oscillator is given by integrating $\epsilon(t)$, i.e.,

$$\phi(t) = 2\pi \int_0^t \epsilon(\lambda)d\lambda,$$

which turns out to be a Wiener process. The single-sided PSD of $\phi(t)$ is

$$S(f) = \frac{\xi}{\pi f^2}.$$

The spectrum of the phase noise centers around the carrier frequency and exhibits "narrowband" and "low-frequency" characteristics.

1.2.2 IQ Imbalances

IQ imbalance is critical to direct-conversion receivers that use the I and Q-branch mixing. This requires shifting the oscillator output by $90°$. In an ideal receiver, the sinusoidal signals used for I and Q-branch mixing have the same amplitude and are orthogonal to each other. Also, their phase is perfectly aligned with the carrier signal. However, the actual oscillator signals in the I and Q branches can be slightly different from the ideal oscillator signals due to imperfectness in the mixer and $90°$ phase shifter. The errors in the phase shifter and imbalances between the amplitudes of the I and Q-branch signals corrupt the received signal constellation. The effects of IQ imbalance can be characterized by two parameters α and θ that usually vary slowly with respect to the symbol rate. As illustrated in Fig. 1.2, the phase difference between the I and Q-branch signals is $90° - \theta$, while the amplitude of the I and Q-branch signals are $1 + \alpha$ and $1 - \alpha$, respectively.

1.2.3 Nonlinearities

Ideally, the input-output relation of the amplifiers at the transmitter and receiver is expected to be linear. However, this holds only when the signal amplitude is

5
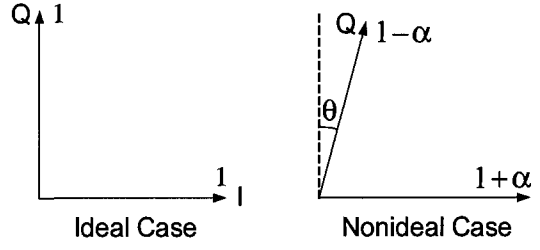
Figure 1.2: IQ Imbalance.

small. Since the dynamic range of input signals can be quite large in wireless communications, the linear relationship is usually invalid. This can cause significant distortion to the received signals. The nonlinearity can be modeled by a third-order polynomial:

$$y(t) \approx \gamma_1 x(t) + \gamma_2 [x(t)]^2 + \gamma_3 [x(t)]^3,$$

where $x(t)$ and $y(t)$ are the input and output signals, and γ_1, γ_2, γ_3 are constants. In the expression, $\gamma_1 x(t)$ represents the linear component in the output, while $\gamma_2 [x(t)]^2$ and $\gamma_3 [x(t)]^3$ are the second and third-order nonlinear components in the output. One effect of the nonlinearity is the cross modulation that occurs when the desired signal is simultaneously acquired with a much stronger interference signal of different center frequency. The resulting distortion can seriously hurt the quality of the desired signal.

1.3   Organization of this Dissertation

Chapter 2 develops a two-stage scheme to compensate for phase noise distortions in wireless OFDM systems. In the first stage, block-type pilot symbols are transmitted and the channel coefficients are jointly estimated with the phase noise in the time domain. In the second stage, comb-type OFDM symbols are transmitted

6

such that the receiver can jointly estimate the data symbols and the phase noise. It is shown both by theory and computer simulations that the proposed scheme can effectively mitigate the inter-carrier interference caused by phase noise and improve the bit error rate of OFDM systems. Another benefit of the proposed scheme is that the sensitivity of OFDM receivers to phase noise can be significantly lowered, which helps simplify the oscillator and circuitry design in terms of implementation cost and power consumption.

In Chapter 3, the joint effects of IQ imbalance and phase noise on OFDM systems are analyzed, and a compensation scheme is proposed to improve the system performance in the presence of IQ imbalance and phase noise. The scheme consists of a joint estimation of channel and impairment parameters and a joint data symbol estimation algorithm. In the proposed channel estimation algorithm, the channel coefficients are jointly estimated with the IQ imbalance parameters and the phase noise components. The joint estimation obtains a more accurate channel estimate than the conventional methods that either ignore the impairments or simply treat them as additive Gaussian noise. Its performance is also demonstrated to be close to the associated Cramer-Rao lower bound. In the proposed data symbol estimation algorithm, the joint compensation is decomposed into IQ imbalance compensation and phase noise compensation.

Chapter 4 studies the nonlinear effects caused by the wideband front-end analog processing in a software-defined radio (SDR) system. In the presence of strong blocker (interference) signals, such nonlinearities introduce severe cross modulation over the desired signals. We investigate how the nonlinear distortions can be compensated for by using digital signal processing techniques. In the proposed solution, the SDR scans the wide spectrum and locates the desired signal and strong blocker signals. After down-converting these signals separately

to the baseband, the baseband processor processes them jointly to mitigate the cross-modulation interferences. The proposed approach demonstrates how mixed-signal, i.e., joint analog and digital, processing techniques play a critical role in the emerging SDR and cognitive radio technologies.

At the end of this dissertation, Chapter 5 proposes some new ideas for future research. It is recommended to compare the digital approach with the analog approach in terms of cost and performance. This can provide system designers with more options for RF distortion compensation. Moreover, by using joint analog and digital design strategies, some analog challenges can be passed to digital designers who might have an elegant digital solution. The latest radio technologies rely more and more on complicated signal processing algorithms. The rapid advances in signal processing and digital IC technologies create many opportunities for the next generation of wireless communications.

CHAPTER 2

Phase Noise Compensation in OFDM Wireless Systems

2.1 Introduction

Orthogonal Frequency Division Multiplexing (OFDM) is a widely recognized modulation technique for high data rate communications over wireless links [BSE04, SL05]. Because of its capability to capture multipath energy and eliminate inter-symbol interference, OFDM has been chosen as the transmission method for several standards, including the IEEE 802.11a wireless local area network (WLAN) standard in the 5 GHz band, the IEEE 802.11g WLAN standard in the 2.4 GHz band, and the European digital video broadcasting system (DVB-T). Also, the OFDM-based physical layer is being considered by several standardization groups, such as the IEEE 802.15.3 wireless personal area network (WPAN) and the IEEE 802.20 mobile broadband wireless access (MBWA) groups. The heightened interest in OFDM has resulted in tremendous research activities in this field to make the real systems more reliable and less costly.

One limitation of OFDM systems is that they are highly sensitive to the phase noise introduced by local oscillators. Phase noise is the phase difference between the phase of the carrier signal and the phase of the local oscillator, and its effect on OFDM receivers has been investigated in many previous works,

such as [Mus95, PBM95, Tom98, PM02]. The distortion caused by phase noise is characterized by a common phase error (CPE) term and an inter-carrier interference (ICI) term. The CPE term represents the common rotation of all constellation points in the complex plane, while the ICI term behaves like additive Gaussian noise, as illustrated in Fig. 2.1. Compared to the single-carrier modulation methods that can track the fast variation in phase noise adaptively in a decision-directed manner, e.g., [THS06], OFDM transmits data symbols over many low-rate subcarriers, which makes it more difficult to track and compensate for phase noise.

(a) Without phase noise.   (b) With phase noise.
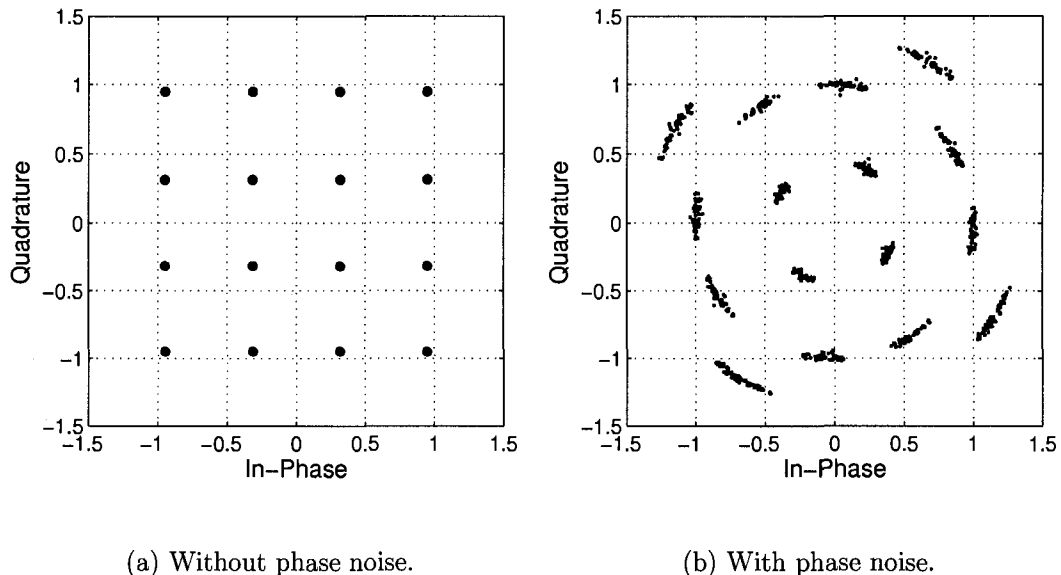
Figure 2.1: Scatter plot of 16-QAM received symbols in the absence/presence of phase noise.

2.1.1 Prior Work

The high sensitivity of OFDM receivers to phase noise imposes a stringent constraint on the design and fabrication of oscillators and the supplementary cir-

cuitry, such that the generated phase noise level will not cause the system to fail [LH00, KL05, ACP05]. This requirement increases the implementation cost of OFDM receivers because impairments associated with fabrication variations are usually either unpredictable or uncontrollable. There have been works in the literature to mitigate the effects of phase noise in the digital domain. This approach provides an efficient, low-cost and reliable solution to the phase noise problem. Some authors have proposed methods to compensate for the CPE term, in which the constellation rotation is estimated by using pilot tones embedded in OFDM symbols and then corrected by the demodulator [RK95]. Since the ICI effect is either ignored or treated as additive noise in these schemes, they perform poorly if the phase noise varies fast in comparison to the OFDM symbol rate. To overcome this difficulty, some ICI compensation schemes have been proposed. The method in [EWH01] utilizes the pilot tones, which are sufficiently separated away from the data tones, to transmit pilot symbols for phase noise estimation. In the self-cancellation method presented in [ZH01] and [RLP05], each data symbol is transmitted using two adjacent subcarriers, and the received symbols are linearly combined to suppress ICI by exploiting the fact that the ICI coefficients change slowly over adjacent subcarriers. This technique has the advantage of low implementation complexity, but it reduces the spectral efficiency by one half. In [GNE03], an FIR-type equalizer is employed to compensate for phase noise, and the filter coefficients are determined by the method of least squares. Since the filter length is limited by the number of pilot tones, it can only compensate for the ICI that is from adjacent subcarriers. A time-domain phase noise estimation and correction scheme is proposed in [CBY02], where the phase noise process is parameterized by using sinusoidal waveforms as the bases and the parameters are estimated by the method of least squares. However, using sinusoidal waveforms to approximate phase noise may not be optimal, and

11

how to jointly estimate the model parameters and the transmitted symbols is not addressed in [CBY02]. Similarly, the work [LZ04] approximates phase noise by using a small number of sinusoidal components, and it suggests inserting some pilot tones outside the spectrum occupied by data transmission and estimating the model parameters of phase noise by using the received pilot signals. This scheme requires extra bandwidth and can only correct the ICI from adjacent subcarriers because of the approximation made in modeling. A method that utilizes *a priori* information about phase noise spectrum to suppress phase noise is presented in [LZL05]. All these ICI compensation schemes assume that the receiver has perfect channel state information; however, in wireless communications, the channel is time-varying and the receiver has to estimate the channel in the presence of phase noise, which makes the scenario more complicated than what has been studied before. In [LYK05], joint channel estimation and phase noise suppression are achieved by using the expectation-maximization (EM) algorithm; however, it simply models the ICI term as additive white noise. Also, the work [WB03] proposes a channel estimation method with the aid of cyclic prefix symbols, while [KK05] proposes a joint channel estimation scheme using soft decision decoding. In [LPL06], the maximum *a posteriori* (MAP) channel estimator is derived for the case when both frequency offset and phase noise are present.

2.1.2 This Work

In this chapter, we propose a new phase noise compensation scheme for OFDM-based wireless communications with improved performance. The scheme consists of a channel estimation stage and a data transmission stage. In the channel estimation stage, block-type pilot symbols are transmitted so that the receiver

can jointly estimate the channel coefficients and the phase noise.[1] Instead of estimating the channel coefficients and phase noise in the frequency domain, we estimate them in the *time domain* by using interpolation techniques to reduce the number of unknowns. The joint channel and phase noise estimation procedure gives a more accurate estimate of the channel than other conventional methods that either ignore phase noise or simply treat it as additive Gaussian noise. The slowly-varying nature of wireless channels allows us to use the channel estimate in the subsequent data transmission stage for estimating the data symbols. In the data transmission stage, comb-type pilot symbols, which contain both data symbols and pilot symbols, are transmitted, and the data symbols and the phase noise components are jointly estimated at the receiver in order to mitigate both the CPE and ICI effects of phase noise. The proposed approach outperforms other methods for two main reasons. First, instead of approximating the ICI term as additive noise, we use the exact data model given in Section 2.2 and treat the phase noise components as unknown parameters to be determined over two stages. Second, the correlation property of phase noise is exploited to reduce the number of unknowns. This is similar to the idea of modeling the phase noise process by the sum of several sinusoidal waveforms, but the use of linear interpolation in the time domain makes our approach less sensitive to variations in phase noise models.

The chapter is organized as follows. Section 2.2 briefly describes the system model with phase noise and formulates the effects of phase noise on OFDM receivers. The proposed algorithm is presented in Section 2.3 and analyzed in Section 2.4 in terms of the effective signal-to-noise ratio. Simulation results and performance comparison of different algorithms are given in Section 2.5.

---

[1]All standardized OFDM systems today provide such full pilot symbols at the beginning of every packet. Therefore, the proposed scheme does not require any modification to the packet structure and can be applied to existing standards.

Throughout this chapter, we adopt the following notations: $(\cdot)^T$ denotes the matrix transpose, $(\cdot)^*$ denotes the matrix conjugate transpose, diag $\{\cdot\}$ represents the diagonal matrix whose diagonal entries are determined by its argument, $\mathrm{Tr}\{\cdot\}$ returns the trace of a matrix, $\mathrm{Re}\{\cdot\}$ denotes the real part of its argument, $\mathbf{I}_N$ is the identity matrix of size $N \times N$, and $\mathbf{E}\{\cdot\}$ is the expected value with respect to the underlying probability measure.

2.2 System Model

The OFDM modulation and demodulation are illustrated in Fig. 2.2. At the transmitter, the bits from information sources are first mapped into constellation symbols, and then converted into a block of N symbols $X[k]$, $k = 0, 1, \ldots, N-1$, by a serial-to-parallel converter. The N symbols are the frequency components to be transmitted using the N subcarriers of the OFDM modulator, and they are converted to OFDM symbols $x[n]$, $n = 0, 1, \ldots, N-1$, by the unitary inverse Fast Fourier Transform (IFFT), i.e.,

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi nk}{N}}, \quad n = 0, 1, \ldots, N-1.$$

A cyclic prefix of length P is added to the IFFT output in order to eliminate the inter-symbol interference caused by multipath propagation. The resulting $N + P$ symbols are converted into a continuous-time baseband signal $x(t)$ for transmission. If the continuous-time impulse response function of the baseband channel is $h(t)$ and the phase noise at the local oscillator is $\phi(t)$ (see Fig. 2.3),[2]

---

[2]In Appendix A, we give a brief review of the existing models for phase noise, and derive some useful statistical measures that will be used in assessing the performance improvement of the proposed compensation scheme in Section 2.4.
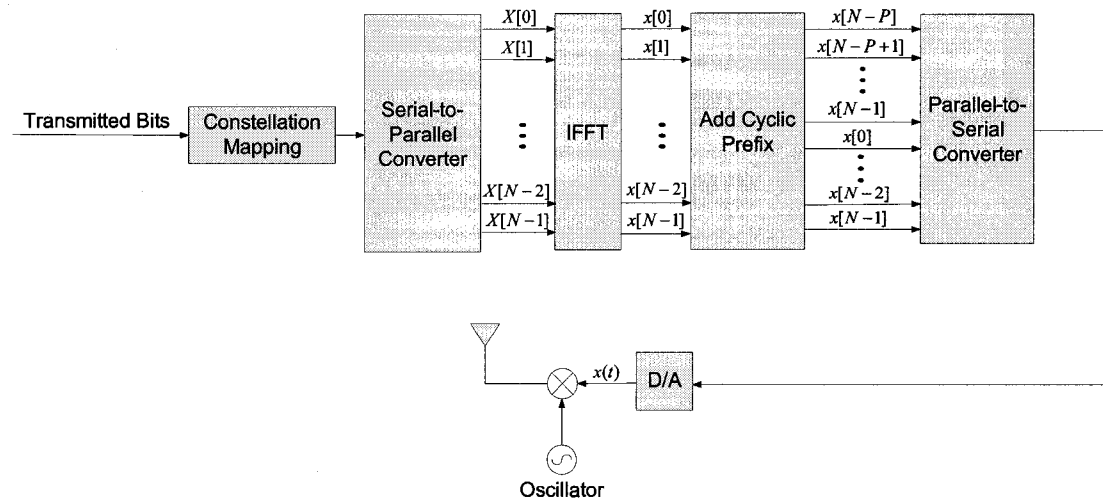
(a) OFDM Modulator.
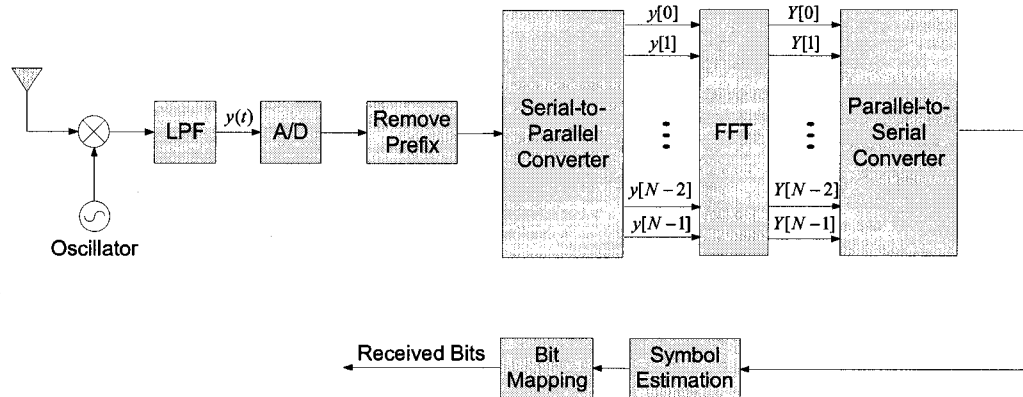
(b) OFDM Demodulator.

Figure 2.2: OFDM modulation and demodulation.
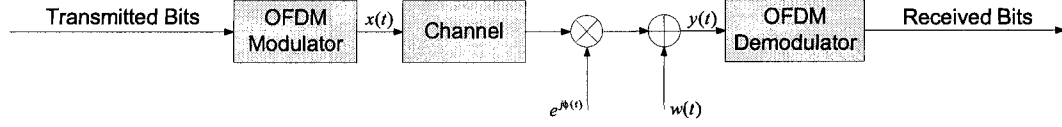
Figure 2.3: Block diagram of the OFDM system with phase noise.

then the received baseband signal is given by

$$y(t) = e^{j\phi(t)} \int_{-\infty}^{\infty} h(t - \tau)x(\tau)d\tau + w(t),$$

where $w(t)$ is additive white Gaussian noise. Let T_s be the symbol time of the system. At the demodulator, $y(t)$ is sampled at period T_s. After removing the cyclic prefix, a block of N symbols $y[n]$, $n = 0, 1, \ldots, N - 1$, is obtained, whose elements are related to $x[n]$, $n = 0, 1, \ldots, N - 1$, by

$$y[n] = e^{j\phi(nT_s)} \left(h[n] \circledast x[n] \right) + w[n]$$

$$= e^{j\phi(nT_s)} \sum_{r=0}^{N-1} h[(n - r)_N]x[r] + w[n], \tag{2.1}$$

where $\circledast$ denotes circular convolution, $h[n]$ is the discrete-time baseband channel impulse response, $(n - r)_N$ stands for

$$(n - r) \bmod N$$

and $w[n]$ is the additive noise component. Assume that $h[n]$ has length L, i.e.,

$$h[n] = 0 \text{ if } n \notin \{0, 1, \ldots, L - 1\}.$$

OFDM modulation requires $L - 1 \leq P$ in order to eliminate the inter-symbol interference. The unitary Fast Fourier Transform (FFT) is then performed on $y[n]$, $n = 0, 1, \ldots, N - 1$, to obtain $Y[k]$, $k = 0, 1, \ldots, N - 1$. Let

$$A[k] = \frac{1}{N} \sum_{n=0}^{N-1} e^{j\phi(nT_s)} e^{-j\frac{2\pi kn}{N}}, \quad k = 0, 1, \ldots, N - 1, \tag{2.2}$$

16

and

$$H[k] = \sum_{n=0}^{L-1} h[n]e^{-j\frac{2\pi kn}{N}}, \quad k = 0, 1, \ldots, N-1.$$

It follows from (2.1) that

$$Y[k] = A[k] \circledast (H[k]X[k]) + W[k]$$

$$= \sum_{r=0}^{N-1} A[r]H\left[(k-r)_N\right]X\left[(k-r)_N\right] + W[k], \quad k = 0, 1, \ldots, N-1, \quad (2.3)$$

because multiplication in the time domain implies convolution in the frequency domain and convolution in the time domain implies multiplication in the frequency domain. Here, $H[k]$ is the channel response in the kth subcarrier and $W[k]$ is the additive noise component in the kth subcarrier. Expression (2.3) can be rewritten as

$$Y[k] = A[0]H[k]X[k] + \sum_{r=1}^{N-1} A[r]H\left[(k-r)_N\right]X\left[(k-r)_N\right] + W[k] \qquad (2.4)$$

where the term

$$A[0]H[k]X[k]$$

is called the common phase error (CPE) and

$$\sum_{r=1}^{N-1} A[r]H[(k-r)_N]X[(k-r)_N]$$

is called the inter-carrier interference (ICI). In the absence of phase noise, we have the traditional relation

$$Y[k] = H[k]X[k] + W[k],$$

which follows by setting

$$A[0] = 1$$

17

and

$$A[r] = 0 \text{ for } r \neq 0.$$

Using matrix notation, expression (2.3) can be represented as

$$\mathbf{y} = \mathbf{A}\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{2.5}$$

where

$$\mathbf{y} = \begin{bmatrix} Y[0] \\ Y[1] \\ \vdots \\ Y[N-1] \end{bmatrix} (N \times 1), \quad \mathbf{x} = \begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[N-1] \end{bmatrix} (N \times 1),$$

$$\mathbf{w} = \begin{bmatrix} W[0] \\ W[1] \\ \vdots \\ W[N-1] \end{bmatrix} (N \times 1),$$

$$\mathbf{A} = \begin{bmatrix} A[0] & A[N-1] & \dots & A[1] \\ A[1] & A[0] & \dots & A[2] \\ \vdots & \vdots & \ddots & \vdots \\ A[N-1] & A[N-2] & \dots & A[0] \end{bmatrix} (N \times N), \tag{2.6}$$

$$\mathbf{H} = \begin{bmatrix} H[0] & 0 & \dots & 0 \\ 0 & H[1] & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & H[N-1] \end{bmatrix} (N \times N). \tag{2.7}$$

Note that $\mathbf{A}$ is an $N \times N$ circulant matrix. When phase noise is not present, matrix $\mathbf{A}$ in (2.5) is replaced by the identity matrix.

2.3 Proposed Algorithm

If both $\mathbf{A}$ and $\mathbf{H}$ in (2.5) were known, then the data vector $\mathbf{x}$ could be recovered, e.g., by solving [Say03]:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \| \mathbf{y} - \mathbf{A}\mathbf{H}\mathbf{x} \|^2 .$$

However, in practice, neither the channel matrix $\mathbf{H}$ nor the phase noise matrix $\mathbf{A}$ are known to the receiver. In this section, we propose a solution to deal with the situation when both $\mathbf{A}$ and $\mathbf{H}$ are unknown at the receiver. The proposed algorithm consists of two stages: one is the channel estimation stage and the other is the data transmission stage. In the channel estimation stage, we use block-type pilot symbols to jointly estimate $\mathbf{H}$ and $\mathbf{A}$. In the data transmission stage, comb-type symbols are transmitted such that $\mathbf{x}$ can be jointly estimated with $\mathbf{A}$ by using the $\mathbf{H}$ estimated in the channel estimation stage. The motivation for this algorithm is based on the fact that wireless channels are usually slowly time-varying compared to phase noise. Since the phase noise components may change significantly from one OFDM symbol to another, it is harmful to use the previous estimate of phase noise to help detect the data symbols in the subsequently received OFDM symbols. However, we can use the channel estimate for a few subsequent OFDM symbols due to the slowly time-varying nature of wireless channels. This motivates our approach to compensate for phase noise by using the joint channel estimation (with phase noise) first and then followed by the joint data symbol estimation (with phase noise). The algorithm is illustrated in Fig. 2.4.
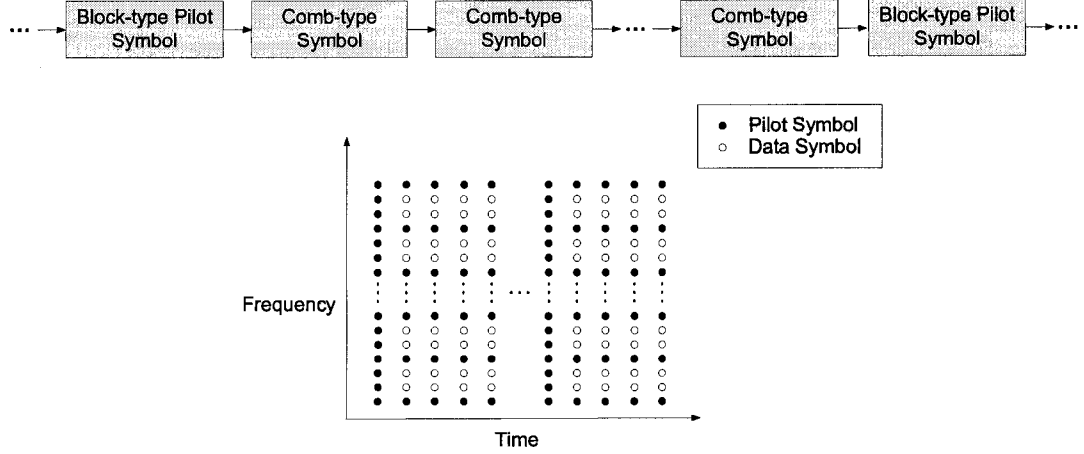
Figure 2.4: The proposed phase noise compensation scheme.

2.3.1 Joint Channel and Phase Noise Estimation

In the block-type pilot symbols, all subcarriers are used to transmit pilot symbols. For convenience of exposition, we assume that each time only one OFDM symbol is used as the block-type pilot symbol for channel estimation. Since there are only N pilot tones in each block-type pilot symbol, it is underdetermined to directly estimate the $2N$ unknowns, i.e.,

$$A[k] \text{ and } H[k], \quad k = 0, 1, \ldots, N - 1.$$

To overcome this difficulty, we can reduce the number of unknowns by properly modeling the channel and the phase noise process with fewer parameters as follows. Since the length L of the discrete-time baseband channel impulse response is normally less than the OFDM symbol size N, we can relate $H[k]$, $k = 0, 1, \ldots, N - 1$, to $h[n]$, $n = 0, 1, \ldots, L - 1$, through

$$\mathbf{h} = \mathbf{F_h h'}, \tag{2.8}$$

20

where

$$\mathbf{h} = \begin{bmatrix} H[0] \\ H[1] \\ \vdots \\ H[N-1] \end{bmatrix} (N \times 1), \quad \mathbf{h'} = \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[L-1] \end{bmatrix} (L \times 1),$$

and

$$\mathbf{F_h} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j\frac{2\pi}{N}} & \cdots & e^{-j\frac{2\pi(L-1)}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\frac{2\pi(N-1)}{N}} & \cdots & e^{-j\frac{2\pi(N-1)(L-1)}{N}} \end{bmatrix} (N \times L)$$

consists of the leading L columns of the discrete Fourier transform matrix. Instead of estimating $\mathbf{h}$, we can estimate $\mathbf{h'}$. This reduces the number of unknown channel coefficients from N (in the frequency domain) to L (in the time domain).

For the phase noise, instead of estimating

$$A[k], \quad k = 0, 1, \dots, N-1,$$

we can estimate the phase noise components in the time domain, i.e.,

$$e^{j\phi(nT_s)}, \quad n = 0, 1, \dots, N-1.$$

In order to reduce the number of unknowns, we can estimate

$$e^{j\phi(m(N-1)T_s/(M-1))}, \quad m = 0, 1, \dots, M-1 \ (M < N),$$

and then obtain the approximation of

$$e^{j\phi(nT_s)}, \quad n = 0, 1, \dots, N-1,$$

by interpolation. Let

$$\mathbf{c} = \begin{bmatrix} e^{j\phi(0)} \\ e^{j\phi(T_s)} \\ \vdots \\ e^{j\phi((N-1)T_s)} \end{bmatrix} \ (N \times 1), \quad \mathbf{c}' = \begin{bmatrix} e^{j\phi(0)} \\ e^{j\phi\left(\frac{(N-1)T_s}{M-1}\right)} \\ \vdots \\ e^{j\phi((N-1)T_s)} \end{bmatrix} \ (M \times 1).$$

Then, we can write

$$\mathbf{c} \approx \mathbf{P}\mathbf{c}',$$

where $\mathbf{P}$ is an interpolation matrix to be determined in Subsection 2.3.3. Using (2.2), we have

$$\mathbf{a} = \frac{1}{N}\mathbf{F_a}\mathbf{c} \approx \frac{1}{N}\mathbf{F_a}\mathbf{P}\mathbf{c}', \tag{2.9}$$

where

$$\mathbf{a} = \begin{bmatrix} A[0] \\ A[1] \\ \vdots \\ A[N-1] \end{bmatrix} \ (N \times 1) \tag{2.10}$$

and $\mathbf{F_a}$ is the discrete Fourier transform matrix, i.e.,

$$\mathbf{F_a} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{-j\frac{2\pi}{N}} & \dots & e^{-j\frac{2\pi(N-1)}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-j\frac{2\pi(N-1)}{N}} & \dots & e^{-j\frac{2\pi(N-1)^2}{N}} \end{bmatrix} \ (N \times N).$$

Consequently, knowing $\mathbf{x}$ during training, we can estimate $\mathbf{A}$ and $\mathbf{H}$ by solving

$$\min_{\mathbf{c}',\mathbf{h}'} \| \mathbf{y} - \mathbf{A}\mathbf{H}\mathbf{x} \|^2 \tag{2.11}$$

where $\mathbf{A}$ is related to $\mathbf{c}'$ by (2.6) and (2.9) and $\mathbf{H}$ is related to $\mathbf{h}'$ by (2.7) and (2.8).

The estimates of $\mathbf{A}$ and $\mathbf{H}$ obtained from (2.11) will have an ambiguity of a scaling factor. To resolve this ambiguity, we estimate

$$\mathbf{A}'' = \frac{1}{A[0]}\mathbf{A} \quad (A[0] \neq 0)$$

and

$$\mathbf{H}'' = A[0]\mathbf{H} \tag{2.12}$$

instead of $\mathbf{A}$ and $\mathbf{H}$. Here we assume that $A[0] \neq 0$, since $A[0]$ represents the common phase rotation of the received N symbols and is usually nonzero. By (2.6) and (2.7), $\mathbf{A}''$ is circulant and $\mathbf{H}''$ is diagonal, i.e.,

$$\mathbf{A}'' = \begin{bmatrix} A''[0] & A''[N-1] & \dots & A''[1] \\ A''[1] & A''[0] & \dots & A''[2] \\ \vdots & \vdots & \ddots & \vdots \\ A''[N-1] & A''[N-2] & \dots & A''[0] \end{bmatrix} (N \times N) \tag{2.13}$$

with

$$A''[0] = 1 \text{ and } A''[k] = \frac{1}{A[0]}A[k], \text{ for } k = 1, 2, \dots, N-1,$$

and

$$\mathbf{H}'' = \begin{bmatrix} H''[0] & 0 & \dots & 0 \\ 0 & H''[1] & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & H''[N-1] \end{bmatrix} (N \times N) \tag{2.14}$$

with

$$H''[k] = A[0]H[k], \text{ for } k = 0, 1, \dots, N-1.$$

Correspondingly, we define

$$\mathbf{c}'' = \frac{1}{A[0]}\mathbf{c}'$$

and

$$\mathbf{h}'' = A[0]\mathbf{h}'.$$

It follows from (2.9) and (2.8) that

$$\overline{\mathbf{a}} \approx \frac{1}{N}\mathbf{F_aPc}'' \qquad (2.15)$$

and

$$\overline{\mathbf{h}} = \mathbf{F_h}\mathbf{h}'' \qquad (2.16)$$

where

$$\overline{\mathbf{a}} = \begin{bmatrix} A''[0] \\ A''[1] \\ \vdots \\ A''[N-1] \end{bmatrix} (N \times 1), \quad \overline{\mathbf{h}} = \begin{bmatrix} H''[0] \\ H''[1] \\ \vdots \\ H''[N-1] \end{bmatrix} (N \times 1). \qquad (2.17)$$

Note that $A''[0]$ is 1. Problem (2.11) is thus reformulated as

$$\min_{\mathbf{c}'',\mathbf{h}''} \| \mathbf{y} - \mathbf{A}''\mathbf{H}''\mathbf{x} \|^2 \text{ subject to } A''[0] = 1, \qquad (2.18)$$

where $\mathbf{A}''$ is related to $\mathbf{c}''$ by (2.13) and (2.15) and $\mathbf{H}''$ is related to $\mathbf{h}''$ by (2.14) and (2.16). If $L + M \leq N$, then (2.18) may have a unique solution for $\mathbf{c}''$ and $\mathbf{h}''$. Finding the necessary and sufficient conditions for which (2.18) has a unique solution in general belongs to the class of problems concerned with the global or local identifiability of a system [Lju99].

The optimization problem given by (2.18) is nonlinear and nonconvex. Suppose that $\mathbf{A}''$ is known. We can reformulate (2.18) into a least-squares problem in $\mathbf{h}''$ as

$$\min_{\mathbf{h}''} \| \mathbf{y} - \mathbf{A}''\mathbf{X}\mathbf{F_h}\mathbf{h}'' \|^2,$$

where $\mathbf{X} = \mathrm{diag}\,\{\mathbf{x}\}$ and the $\mathrm{diag}\{\cdot\}$ notation refers to a diagonal matrix whose diagonal entries are the elements of its argument. The optimal $\mathbf{h}''$ is then given by

$$\mathbf{h}''_o = \left[\mathbf{F}_\mathbf{h}^* \mathbf{X}^* \left(\mathbf{A}''\right)^* \mathbf{A}'' \mathbf{X} \mathbf{F}_\mathbf{h}\right]^{-1} \mathbf{F}_\mathbf{h}^* \mathbf{X}^* (\mathbf{A}'')^* \mathbf{y}.$$

By substituting $\mathbf{H}'' = \mathrm{diag}\,\{\mathbf{F}_\mathbf{h} \mathbf{h}''_o\}$ into (2.18), the optimal $\mathbf{c}''$ is given by the solution to the following nonconvex optimization problem:

$$\mathbf{c}''_o = \arg\min_{\mathbf{c}''} \| \mathbf{y} - \mathbf{A}'' \cdot \mathrm{diag}\,\{\mathbf{F}_\mathbf{h} \mathbf{h}''_o\} \cdot \mathbf{x}\|^2$$

$$= \arg\min_{\mathbf{c}''} \| \mathbf{y} - \mathbf{A}'' \cdot \mathrm{diag}\{\mathbf{F}_\mathbf{h} \left[\mathbf{F}_\mathbf{h}^* \mathbf{X}^* \left(\mathbf{A}''\right)^* \mathbf{A}'' \mathbf{X} \mathbf{F}_\mathbf{h}\right]^{-1} \mathbf{F}_\mathbf{h}^* \mathbf{X}^* \left(\mathbf{A}''\right)^* \mathbf{y}\} \cdot \mathbf{x}\|^2$$

subject to $A''[0] = 1$. In implementation, we use the two-step iterative algorithm presented in Algorithm 2.1 to find a sub-optimal solution to (2.18). In Step 3, given a previously estimated $\mathbf{c}''$, we find an optimal $\mathbf{h}''$. Expression (2.19) can be rewritten as

$$\widehat{\mathbf{h}}''_{i-1} = \arg\min_{\mathbf{h}''} \| \mathbf{y} - \widehat{\mathbf{A}}''_{i-1} \mathbf{H}'' \mathbf{x}\|^2$$

$$= \arg\min_{\mathbf{h}''} \| \mathbf{y} - \widehat{\mathbf{A}}''_{i-1} \mathbf{X} \overline{\mathbf{h}}\|^2$$

$$= \arg\min_{\mathbf{h}''} \| \mathbf{y} - \widehat{\mathbf{A}}''_{i-1} \mathbf{X} \mathbf{F}_\mathbf{h} \mathbf{h}''\|^2.$$

This is because

$$\mathbf{H}'' = \mathrm{diag}\{\overline{\mathbf{h}}\} \text{ and } \overline{\mathbf{h}} = \mathbf{F}_\mathbf{h} \mathbf{h}''.$$

The problem is now in the form of a standard least-squares problem, and its solution is given by

$$\widehat{\mathbf{h}}''_{i-1} = \left[\mathbf{F}_\mathbf{h}^* \mathbf{X}^* \left(\widehat{\mathbf{A}}''_{i-1}\right)^* \widehat{\mathbf{A}}''_{i-1} \mathbf{X} \mathbf{F}_\mathbf{h}\right]^{-1} \mathbf{F}_\mathbf{h}^* \mathbf{X}^* \left(\widehat{\mathbf{A}}''_{i-1}\right)^* \mathbf{y}.$$

Algorithm 2.1 Joint Channel and Phase Noise Estimation

0: Start with an initial guess $\widehat{\mathbf{c}}_0''$. For example,

$$\widehat{\mathbf{c}}_0'' = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T.$$

1: $i = 1$

2: **repeat**

3:      Let $\bar{\mathbf{a}}_{i-1} = \frac{1}{N}\mathbf{F_a}\mathbf{P}\widehat{\mathbf{c}}_{i-1}''$, and find the associated optimal $\widehat{\mathbf{h}}_{i-1}''$ by solving the following least-squares problem:

$$\widehat{\mathbf{h}}_{i-1}'' = \arg\min_{\mathbf{h}''} \| \mathbf{y} - \widehat{\mathbf{A}}_{i-1}''\mathbf{H}''\mathbf{x} \|^2, \tag{2.19}$$

     where $\mathbf{x}$, $\mathbf{y}$ are known and $\widehat{\mathbf{A}}_{i-1}''$ is determined by $\bar{\mathbf{a}}_{i-1}$ according to (2.13) and (2.17). The expression for $\widehat{\mathbf{h}}_{i-1}''$ is given by

$$\widehat{\mathbf{h}}_{i-1}'' = \left[\mathbf{F_h^*}\mathbf{X^*}\left(\widehat{\mathbf{A}}_{i-1}''\right)^* \widehat{\mathbf{A}}_{i-1}''\mathbf{X}\mathbf{F_h} \right]^{-1} \mathbf{F_h^*}\mathbf{X^*}\left(\widehat{\mathbf{A}}_{i-1}''\right)^* \mathbf{y}.$$

4:      Let $\bar{\mathbf{h}}_{i-1} = \mathbf{F_h}\widehat{\mathbf{h}}_{i-1}''$, and find the associated optimal $\widehat{\mathbf{c}}_i''$ by solving the following least-squares problem:

$$\widehat{\mathbf{c}}_i'' = \arg\min_{\mathbf{c}''} \| \mathbf{y} - \mathbf{A}''\widehat{\mathbf{H}}_{i-1}''\mathbf{x} \|^2 \text{ subject to } \mathbf{g}\mathbf{c}'' = 1, \tag{2.20}$$

     where $\widehat{\mathbf{H}}_{i-1}'' = \text{diag}\{\bar{\mathbf{h}}_{i-1}\}$ and $\mathbf{g}$ is the first row of $\frac{1}{N}\mathbf{F_a}\mathbf{P}$. The expression for $\widehat{\mathbf{c}}_i''$ is given by

$$\widehat{\mathbf{c}}_i'' = (\mathbf{O^*O})^{-1}\left(\mathbf{O^*y} - \lambda\mathbf{g^*}\right),$$

     where

$$\mathbf{O} = \frac{1}{N}\mathbf{T}\mathbf{F_a}\mathbf{P}, \quad \lambda = \frac{\mathbf{g}(\mathbf{O^*O})^{-1}\mathbf{O^*y} - 1}{\mathbf{g}(\mathbf{O^*O})^{-1}\mathbf{g^*}},$$

     and $\mathbf{T}$ is the circulant matrix formed by the elements of $\widehat{\mathbf{H}}_{i-1}''\mathbf{x}$.

5:      $i = i + 1$

6: **until** there is no significant improvement in the objective function $\| \mathbf{y} - \widehat{\mathbf{A}}_i''\widehat{\mathbf{H}}_{i-1}''\mathbf{x} \|^2$.

In Step 4, given a previously estimated $\mathbf{h}''$, we find an optimal $\mathbf{c}''$. Expression (2.20) can be reformulated as

$$\widehat{\mathbf{c}}_i'' = \arg \min_{\mathbf{c}''} \| \mathbf{y} - \mathbf{A}'' \widehat{\mathbf{H}}_{i-1}'' \mathbf{x} \|^2$$

$$= \arg \min_{\mathbf{c}''} \| \mathbf{y} - \mathbf{T}\overline{\mathbf{a}} \|^2$$

$$= \arg \min_{\mathbf{c}''} \| \mathbf{y} - \frac{1}{N} \mathbf{T} \mathbf{F_a} \mathbf{P} \mathbf{c}'' \|^2,$$

where $\mathbf{T}$ is the circulant matrix whose elements are given by the vector $\widehat{\mathbf{H}}_{i-1}'' \mathbf{x}$, i.e., let

$$\widehat{\mathbf{H}}_{i-1}'' \mathbf{x} = \begin{bmatrix} \zeta_0 \\ \zeta_1 \\ \vdots \\ \zeta_{N-1} \end{bmatrix} \quad (N \times 1) \tag{2.21}$$

and define the circulant matrix

$$\mathbf{T} = \begin{bmatrix} \zeta_0 & \zeta_{N-1} & \cdots & \zeta_1 \\ \zeta_1 & \zeta_0 & \cdots & \zeta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \zeta_{N-1} & \zeta_{N-2} & \cdots & \zeta_0 \end{bmatrix} \quad (N \times N). \tag{2.22}$$

Let $\mathbf{g}$ be the first row of $\frac{1}{N} \mathbf{F_a} \mathbf{P}$. Since $\overline{\mathbf{a}} = \frac{1}{N} \mathbf{F_a} \mathbf{P} \mathbf{c}''$, the constraint $A''[0] = 1$ is equivalent to

$$\mathbf{g} \mathbf{c}'' = 1.$$

Solving the problem

$$\widehat{\mathbf{c}}_i'' = \arg \min_{\mathbf{c}''} \| \mathbf{y} - \frac{1}{N} \mathbf{T} \mathbf{F_a} \mathbf{P} \mathbf{c}'' \|^2 \text{ subject to } \mathbf{g} \mathbf{c}'' = 1$$

leads to

$$\widehat{\mathbf{c}}_i'' = (\mathbf{O}^* \mathbf{O})^{-1} (\mathbf{O}^* \mathbf{y} - \lambda \mathbf{g}^*),$$

where

$$\mathbf{O} = \frac{1}{N}\mathbf{TF_aP}, \quad \lambda = \frac{\mathbf{g(O^*O)^{-1}O^*y} - 1}{\mathbf{g(O^*O)^{-1}g^*}}.$$

It can be seen that Step 3 finds the optimal $\widehat{\mathbf{h}}''_{i-1}$ by setting $\mathbf{c}''$ to be the predetermined $\widehat{\mathbf{c}}''_{i-1}$, and hence

$$\| \mathbf{y} - \widehat{\mathbf{A}}''_{i-1}\widehat{\mathbf{H}}''_{i-1}\mathbf{x} \|^2 \leq \| \mathbf{y} - \widehat{\mathbf{A}}''_{i-1}\widehat{\mathbf{H}}''_{i-2}\mathbf{x} \|^2.$$

Step 4 finds the optimal $\widehat{\mathbf{c}}''_i$ by setting $\mathbf{h}''$ to be the predetermined $\widehat{\mathbf{h}}''_{i-1}$, and hence

$$\| \mathbf{y} - \widehat{\mathbf{A}}''_i\widehat{\mathbf{H}}''_{i-1}\mathbf{x} \|^2 \leq \| \mathbf{y} - \widehat{\mathbf{A}}''_{i-1}\widehat{\mathbf{H}}''_{i-1}\mathbf{x} \|^2.$$

Therefore, the objective function is decreasing with $i = 1, 2, \ldots$, and eventually converges to a local minimum. The obtained $\widehat{\mathbf{H}}''$ will be used in the data transmission stage when comb-type symbols are transmitted. In the next subsection, we assume that $\mathbf{H}''$ is known to be $\widehat{\mathbf{H}}''$.

2.3.2 Joint Data Symbol and Phase Noise Estimation

In each comb-type OFDM symbol, assume that Q carriers are used for pilot tones and $N - Q$ carriers for data symbols. The joint data symbol and phase noise estimation problem can be formulated according to expression (2.5) as

$$\min_{\mathbf{c}', \mathbf{x}_{\mathrm{data}}} \| \mathbf{y} - \mathbf{AH}''\mathbf{x} \|^2. \tag{2.23}$$

Note that $\mathbf{H}''$ differs from $\mathbf{H}$ by a scaling factor according to (2.12), and hence the estimates of $\mathbf{c}'$ and $\mathbf{A}$ by solving (2.23) differ from their true values by a scaling factor. This ambiguity will not affect the estimates of the data symbols because $\mathbf{x}$ also contains some pilot symbols. Expression (2.23) can be rewritten as

$$\min_{\mathbf{c}', \mathbf{x}_{\mathrm{data}}} \| \mathbf{y} - \mathbf{A}_{\mathrm{pilot}}\mathbf{H}''_{\mathrm{pilot}}\mathbf{x}_{\mathrm{pilot}} - \mathbf{A}_{\mathrm{data}}\mathbf{H}''_{\mathrm{data}}\mathbf{x}_{\mathrm{data}} \|^2, \tag{2.24}$$

28

where $\mathbf{x}_{\text{pilot}}$ is the sub-vector of $\mathbf{x}$ that consists of all the pilot symbols and $\mathbf{A}_{\text{pilot}}$, $\mathbf{H}''_{\text{pilot}}$ are the associated sub-matrices from $\mathbf{A}$ and $\mathbf{H}''$, and $\mathbf{x}_{\text{data}}$ is the sub-vector of $\mathbf{x}$ that consists of all the data symbols and $\mathbf{A}_{\text{data}}$, $\mathbf{H}''_{\text{data}}$ are the associated sub-matrices from $\mathbf{A}$ and $\mathbf{H}''$. Since there are $N - Q$ unknown data symbols in $\mathbf{x}_{\text{data}}$ and M unknown phase noise components in $\mathbf{c}'$, then (2.24) may have a unique solution if $Q \geq M$.

Since the optimization problem given by (2.24) is similar to the joint channel estimation problem we just solved, we can follow the same procedure to solve it. If $\mathbf{A}$ is known, the optimal $\mathbf{x}_{\text{data}}$ is given by

$$\mathbf{x}_{\text{data},o} = \left(\mathbf{H}''_{\text{data}}\right)^{-1} \left(\mathbf{A}^*_{\text{data}}\mathbf{A}_{\text{data}}\right)^{-1} \mathbf{A}^*_{\text{data}} \left(\mathbf{y} - \mathbf{A}_{\text{pilot}}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}}\right).$$

Then the optimal $\mathbf{c}'$ is given by the following nonconvex optimization problem:

$$\mathbf{c}'_o = \arg \min_{\mathbf{c}'} \left\| \mathbf{y} - \mathbf{A}_{\text{pilot}}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}} - \mathbf{A}_{\text{data}}\mathbf{H}''_{\text{data}}\mathbf{x}_{\text{data},o} \right\|^2$$

$$= \arg \min_{\mathbf{c}'} \left\| \mathbf{y} - \mathbf{A}_{\text{pilot}}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}} - \mathbf{A}_{\text{data}}\left(\mathbf{A}^*_{\text{data}}\mathbf{A}_{\text{data}}\right)^{-1}\mathbf{A}^*_{\text{data}} \right.$$

$$\left. \cdot \left(\mathbf{y} - \mathbf{A}_{\text{pilot}}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}}\right) \right\|^2.$$

An iterative method for finding a sub-optimal solution is presented in Algorithm 2.2, in which we first estimate $A[0]$ to initialize the iteration. The CPE coefficient $A[0]$ represents the common rotation of all the received constellation symbols, and it can be estimated by measuring the angle rotated by the pilot symbols [RK95].

2.3.3 Selection of the Interpolation Matrix

2.3.3.1 PSD of the Phase Noise is Known

If the power spectral density (PSD) of the phase noise is known, the optimal interpolation matrix $\mathbf{P}_o$ can be obtained by minimizing the mean-square-error of

Algorithm 2.2 Joint Data Symbol and Phase Noise Estimation

0: Assume that $k_{\text{pilot},j}$, $j = 1, 2, \ldots, Q$, are the subcarrier indices of the Q pilot tones. The CPE coefficient $A[0]$ is estimated by [RK95]:

$$\widehat{A}[0] = \frac{\sum_{j=1}^{Q} \left(H''[k_{\text{pilot},j}]\right)^* \left(X[k_{\text{pilot},j}]\right)^* Y[k_{\text{pilot},j}]}{\sum_{j=1}^{Q} \left|H''[k_{\text{pilot},j}]\right|^2 \left|X[k_{\text{pilot},j}]\right|^2}.$$

1: Let

$$\widehat{\mathbf{c}}_0' = \left[\begin{array}{cccc} \widehat{A}[0] & \widehat{A}[0] & \ldots & \widehat{A}[0] \end{array}\right]^T.$$

2: $i = 1$

3: **repeat**

4:   Let $\widehat{\mathbf{a}}_{i-1} = \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\widehat{\mathbf{c}}_{i-1}'$, and find the associated optimal $\widehat{\mathbf{x}}_{\text{data},i-1}$ by solving the following least-squares problem:

$$\widehat{\mathbf{x}}_{\text{data},i-1} = \arg \min_{\mathbf{x}_{\text{data}}} \left\| \mathbf{y} - \widehat{\mathbf{A}}_{\text{pilot},i-1}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}} - \widehat{\mathbf{A}}_{\text{data},i-1}\mathbf{H}''_{\text{data}}\mathbf{x}_{\text{data}} \right\|^2,$$

where $\widehat{\mathbf{A}}_{\text{pilot},i-1}$ and $\widehat{\mathbf{A}}_{\text{data},i-1}$ are determined by $\widehat{\mathbf{a}}_{i-1}$ according to (2.6) and (2.10). The expression for $\widehat{\mathbf{x}}_{\text{data},i-1}$ is given by

$$\widehat{\mathbf{x}}_{\text{data},i-1} = \left(\mathbf{H}''_{\text{data}}\right)^{-1} \left(\widehat{\mathbf{A}}_{\text{data},i-1}^* \widehat{\mathbf{A}}_{\text{data},i-1}\right)^{-1} \widehat{\mathbf{A}}_{\text{data},i-1}^* \left(\mathbf{y} - \widehat{\mathbf{A}}_{\text{pilot},i-1}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}}\right).$$

5:   Find the optimal $\widehat{\mathbf{c}}_i'$ by solving the following least-squares problem:

$$\widehat{\mathbf{c}}_i' = \arg \min_{\mathbf{c}'} \left\| \mathbf{y} - \mathbf{A}_{\text{pilot}}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}} - \mathbf{A}_{\text{data}}\mathbf{H}''_{\text{data}}\widehat{\mathbf{x}}_{\text{data},i-1} \right\|^2.$$

The expression for $\widehat{\mathbf{c}}_i'$ is given by

$$\widehat{\mathbf{c}}_i' = N\left(\mathbf{P}^*\mathbf{F}_{\mathbf{a}}^*\mathbf{T}^*\mathbf{T}\mathbf{F}_{\mathbf{a}}\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{F}_{\mathbf{a}}^*\mathbf{T}^*\mathbf{y},$$

where $\mathbf{T}$, similar to (2.21) and (2.22), is the circulant matrix formed by the elements of $\mathbf{H}''\widehat{\mathbf{x}}_{i-1}$ with $\widehat{\mathbf{x}}_{i-1}$ formed by $\mathbf{x}_{\text{pilot}}$ and $\widehat{\mathbf{x}}_{\text{data},i-1}$.

6:   $i = i + 1$

7: **until** there is no significant improvement in the objective function $\left\| \mathbf{y} - \widehat{\mathbf{A}}_{\text{pilot},i}\mathbf{H}''_{\text{pilot}}\mathbf{x}_{\text{pilot}} - \widehat{\mathbf{A}}_{\text{data},i}\mathbf{H}''_{\text{data}}\widehat{\mathbf{x}}_{\text{data},i-1} \right\|^2$.

interpolating $\mathbf{c}$ from $\mathbf{c}'$, i.e.,

$$\mathbf{P}_o = \arg \min_{\mathbf{P}} \mathbf{E} \parallel \mathbf{c} - \mathbf{P} \, \mathbf{c}' \parallel^2,$$

from which the optimal $\mathbf{P}_o$ is given by

$$\mathbf{P}_o = \mathbf{R}_{\mathbf{cc}'}\mathbf{R}_{\mathbf{c}'}^{-1}, \tag{2.25}$$

where $\mathbf{R}_{\mathbf{cc}'} = \mathbf{E}\{\mathbf{cc}'^*\}$ and $\mathbf{R}_{\mathbf{c}'} = \mathbf{E}\{\mathbf{c}'\mathbf{c}'^*\}$. Here, $\mathbf{R}_{\mathbf{cc}'}$ and $\mathbf{R}_{\mathbf{c}'}$ are determined by the auto-correlation function of the phase noise process and can be obtained from the PSD of phase noise, as explained in Appendix A.

2.3.3.2 PSD of the Phase Noise is Unknown

In this case, an interpolation matrix $\mathbf{P}_L \in \mathbb{R}^{N \times M}$ is constructed from linear interpolation, as illustrated in Fig. 2.5, and its element at the nth row and m^{th} column is given by

$$\mathbf{P}_L(n, m) =$$
$$\begin{cases} m - \frac{(n-1)(M-1)}{N-1}, & \text{if } \frac{(m-1)(N-1)}{M-1} \leq n - 1 < \frac{m(N-1)}{M-1}, \\ \frac{(n-1)(M-1)}{N-1} - (m-2), & \text{if } \frac{(m-2)(N-1)}{M-1} \leq n - 1 < \frac{(m-1)(N-1)}{M-1}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.26}$$

where $n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$.

It can be shown that for free-running oscillators with linewidth ξ, the optimal interpolator $\mathbf{P}_o$ is approximately equal to the linear interpolator $\mathbf{P}_L$ if $\xi T_s \ll 1$ (see Appendix B for a proof). In wireless systems, $T_s \sim 1\ \mu s$ and ξ ranges from 10 Hz to at most 10 kHz, for which $\xi T_s \leq 0.01$ and the assumption $\xi T_s \ll 1$ is valid.

31
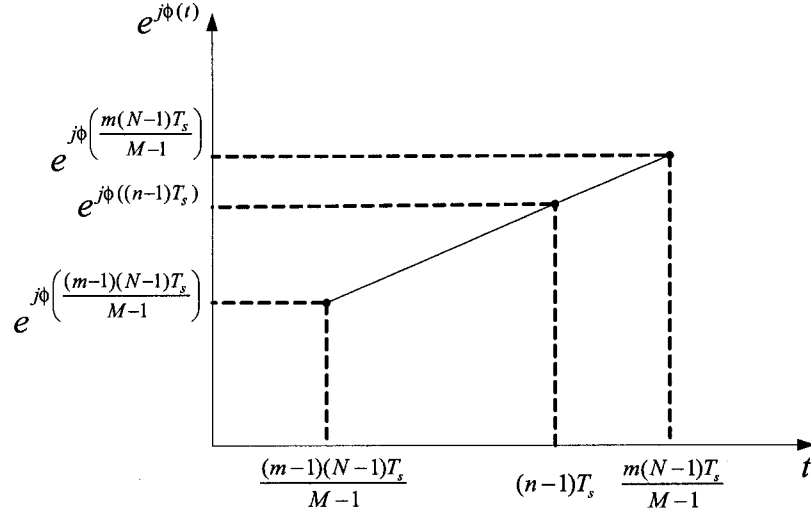
Figure 2.5: Linear interpolation.

2.4 Performance Analysis

In this section, we analyze the effect of phase noise on OFDM receivers and the performance of different compensation schemes in terms of signal-to-noise ratio degradation. The effective signal-to-noise ratio at the receiver is an important measure to characterize the system performance and can be used to predict the bit error rate (BER) [Tom98, PM02]. The expressions of the effective signal-to-noise ratio are derived by assuming that the receiver has perfect information about the channel matrix **H**. The following is also assumed:

1.) The data symbols $X[k]$ are independent and identically distributed with zero mean.

2.) The data symbols, the phase noise, the channel coefficients and the additive noise are independent of each other.

Moreover, let

$$\sigma_X^2 = \mathbf{E}\left\{|X[k]|^2\right\}, \quad \sigma_H^2 = \mathbf{E}\left\{|H[k]|^2\right\}, \quad \sigma_W^2 = \mathbf{E}\left\{|W[k]|^2\right\}.$$

2.4.1 No Compensation for Phase Noise

In this scenario, the receiver does not compensate for phase noise. Expression (2.4) can be rewritten as

$$Y[k] = H[k]X[k] + (A[0] - 1)H[k]X[k] + \sum_{r=1}^{N-1} A[r]H[(k-r)_N]X[(k-r)_N]$$

$$+ W[k],$$

where $H[k]X[k]$ is the desired signal component and the other terms are regarded as noise. The variance of $H[k]X[k]$ is given by

$$\mathbf{E}\left\{|H[k]X[k]|^2\right\} = \mathbf{E}\left\{|H[k]|^2\right\}\mathbf{E}\left\{|X[k]|^2\right\} = \sigma_H^2\sigma_X^2.$$

Let

$$\sigma_{A,k}^2 = \mathbf{E}\left\{|A[k]|^2\right\}, \quad k = 0, 1, \ldots, N-1.$$

By using the facts that (see Appendix A.3)

$$\sum_{k=0}^{N-1} \sigma_{A,k}^2 = 1 \quad \text{and} \quad \mathbf{E}\left\{A[0]\right\} = \mathbf{E}\left\{c(t)\right\}$$

where

$$c(t) = e^{j\phi(t)},$$

the variance of the noise term is computed as

$$\mathbf{E}\left\{\left|(A[0] - 1)H[k]X[k] + \sum_{r=1}^{N-1} A[r]H[(k-r)_N]X[(k-r)_N] + W[k]\right|^2\right\}$$

$$= \mathbf{E}\left\{|(A[0] - 1)H[k]X[k]|^2\right\} + \sum_{r=1}^{N-1} \mathbf{E}\left\{|A[r]H[(k-r)_N]X[(k-r)_N]|^2\right\}$$

33

$$+ \mathbf{E}\left\{|W[k]|^2\right\}$$

$$= \mathbf{E}\left\{|A[0] - 1|^2\right\}\sigma_H^2\sigma_X^2 + \sum_{k=1}^{N-1}\sigma_{A,k}^2\sigma_H^2\sigma_X^2 + \sigma_W^2$$

$$= \sum_{k=0}^{N-1}\sigma_{A,k}^2\sigma_H^2\sigma_X^2 - 2\mathrm{Re}\left\{\mathbf{E}\left\{A[0]\right\}\right\}\sigma_H^2\sigma_X^2 + \sigma_H^2\sigma_X^2 + \sigma_W^2$$

$$= 2\sigma_H^2\sigma_X^2 - 2\mathrm{Re}\left\{\mathbf{E}\left\{c(t)\right\}\right\}\sigma_H^2\sigma_X^2 + \sigma_W^2.$$

Hence, the effective signal-to-noise ratio at the receiver is given by

$$\mathrm{SNR}_{\mathrm{no}} = \frac{\sigma_H^2\sigma_X^2}{2\sigma_H^2\sigma_X^2 - 2\mathrm{Re}\left\{\mathbf{E}\left\{c(t)\right\}\right\}\sigma_H^2\sigma_X^2 + \sigma_W^2}$$

$$= \frac{\mathrm{SNR}_0}{\left(2 - 2\mathrm{Re}\left\{\mathbf{E}\left\{c(t)\right\}\right\}\right)\mathrm{SNR}_0 + 1} \tag{2.27}$$

where

$$\mathrm{SNR}_0 = \frac{\sigma_H^2\sigma_X^2}{\sigma_W^2}$$

is the effective signal-to-noise ratio when there is no phase noise in the system.

2.4.2 CPE Compensation Scheme

In this case, assume that the receiver has perfect information about $A[0]$ but has no information about $A[k]$, $k = 1, 2, \ldots, N-1$. With the CPE correction scheme, $A[0]H[k]X[k]$ in (2.4) represents the signal component and the other terms are regarded as noise. The variance of $A[0]H[k]X[k]$ is given by

$$\mathbf{E}\left\{|A[0]H[k]X[k]|^2\right\} = \mathbf{E}\left\{|A[0]|^2\right\}\mathbf{E}\left\{|H[k]|^2\right\}\mathbf{E}\left\{|X[k]|^2\right\} = \sigma_{A,0}^2\sigma_H^2\sigma_X^2,$$

while the variance of the noise term is computed as

$$\mathbf{E}\left\{\left|\sum_{r=1}^{N-1}A[r]H[(k-r)_N]X[(k-r)_N] + W[k]\right|^2\right\}$$

$$= \sum_{r=1}^{N-1}\mathbf{E}\left\{|A[r]H[(k-r)_N]X[(k-r)_N]|^2\right\} + \mathbf{E}\left\{|W[k]|^2\right\}$$

$$= \sum_{k=1}^{N-1} \sigma_{A,k}^2 \sigma_H^2 \sigma_X^2 + \sigma_W^2.$$

Hence, the effective signal-to-noise ratio is given by

$$
\begin{aligned}
\mathrm{SNR_{CPE}} &= \frac{\sigma_{A,0}^2 \sigma_H^2 \sigma_X^2}{\sum_{k=1}^{N-1} \sigma_{A,k}^2 \sigma_H^2 \sigma_X^2 + \sigma_W^2} \\
&= \frac{\sigma_{A,0}^2 \mathrm{SNR_0}}{\left(\sum_{k=1}^{N-1} \sigma_{A,k}^2\right) \mathrm{SNR_0} + 1} \\
&= \frac{\sigma_{A,0}^2 \mathrm{SNR_0}}{\left(1 - \sigma_{A,0}^2\right) \mathrm{SNR_0} + 1}
\end{aligned}
\tag{2.28}
$$

where $\sigma_{A,0}^2$ is given by (A.7), i.e.,

$$\sigma_{A,0}^2 = \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} R_c((n_1 - n_2)T_s)$$

and $R_c(\tau)$ is the auto-correlation function of $c(t)$.

2.4.3 Ideal Compensation for both CPE and ICI

In expression (2.5), if both $\mathbf{A}$ and $\mathbf{H}$ are known at the receiver, $\mathbf{AHx}$ is the desired signal component and $\mathbf{w}$ is the noise component. Hence,

$$
\begin{aligned}
\mathrm{SNR_{ideal}} &= \frac{\mathbf{E}\{\mathbf{x}^*\mathbf{H}^*\mathbf{A}^*\mathbf{AHx}\}}{\mathbf{E}\{\mathbf{w}^*\mathbf{w}\}} \\
&= \frac{\sigma_H^2 \sigma_X^2 \mathrm{Tr}\{\mathbf{E}\{\mathbf{AA}^*\}\}}{N\sigma_W^2} \\
&= \mathrm{SNR_0}
\end{aligned}
\tag{2.29}
$$

where

$$\mathrm{Tr}\{\mathbf{E}\{\mathbf{AA}^*\}\} = N \left(\sum_{k=0}^{N-1} \sigma_{A,k}^2\right) = N.$$

35

2.4.4 Proposed Algorithm

Using the proposed algorithm, we can get an estimate $\widehat{\mathbf{A}}$ of the phase noise matrix $\mathbf{A}$. Expression (2.5) can be rewritten as

$$\mathbf{y} = \widehat{\mathbf{A}}\mathbf{H}\mathbf{x} + (\mathbf{A} - \widehat{\mathbf{A}})\mathbf{H}\mathbf{x} + \mathbf{w}.$$

Recall the definition of $\mathbf{a}$ given by (2.10). Let $\widehat{\mathbf{a}}$ be the corresponding estimate of $\mathbf{a}$. It is easy to verify that

$$\mathrm{Tr}\{\mathbf{A}\mathbf{A}^*\} = N\|\mathbf{a}\|^2, \quad \mathrm{Tr}\{\widehat{\mathbf{A}}\widehat{\mathbf{A}}^*\} = N\|\widehat{\mathbf{a}}\|^2,$$

and

$$\mathrm{Tr}\{(\mathbf{A} - \widehat{\mathbf{A}})(\mathbf{A} - \widehat{\mathbf{A}})^*\} = N\|\mathbf{a} - \widehat{\mathbf{a}}\|^2.$$

Then the effective signal-to-noise ratio can be expressed as

$$\begin{aligned}
\mathrm{SNR}_{\mathrm{prop}} &= \frac{\mathbf{E}\{\|\widehat{\mathbf{A}}\mathbf{H}\mathbf{x}\|^2\}}{\mathbf{E}\{\|(\mathbf{A} - \widehat{\mathbf{A}})\mathbf{H}\mathbf{x} + \mathbf{w}\|^2\}} \\
&= \frac{\mathbf{E}\{\mathbf{x}^*\mathbf{H}^*\widehat{\mathbf{A}}^*\widehat{\mathbf{A}}\mathbf{H}\mathbf{x}\}}{\mathbf{E}\{\mathbf{x}^*\mathbf{H}^*(\mathbf{A} - \widehat{\mathbf{A}})^*(\mathbf{A} - \widehat{\mathbf{A}})\mathbf{H}\mathbf{x}\} + \mathbf{E}\{\mathbf{w}^*\mathbf{w}\}} \\
&= \frac{\sigma_H^2 \sigma_X^2 \mathrm{Tr}\{\mathbf{E}\{\widehat{\mathbf{A}}\widehat{\mathbf{A}}^*\}\}}{\sigma_H^2 \sigma_X^2 \mathrm{Tr}\{\mathbf{E}\{(\mathbf{A} - \widehat{\mathbf{A}})(\mathbf{A} - \widehat{\mathbf{A}})^*\}\} + N\sigma_W^2} \\
&= \frac{N\sigma_H^2 \sigma_X^2 \mathbf{E}\{\|\widehat{\mathbf{a}}\|^2\}}{N\sigma_H^2 \sigma_X^2 \mathbf{E}\{\|\mathbf{a} - \widehat{\mathbf{a}}\|^2\} + N\sigma_W^2} \\
&= \frac{\mathbf{E}\{\|\widehat{\mathbf{a}}\|^2\} \cdot \mathrm{SNR}_0}{\mathbf{E}\{\|\mathbf{a} - \widehat{\mathbf{a}}\|^2\} \cdot \mathrm{SNR}_0 + 1}.
\end{aligned} \tag{2.30}$$

Since

$$\widehat{\mathbf{a}} = \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\widehat{\mathbf{c}}',$$

we have

$$\begin{aligned}
\mathbf{E}\{\|\widehat{\mathbf{a}}\|^2\} &= \mathbf{E}\left\{\frac{1}{N^2}\widehat{\mathbf{c}}'^*\mathbf{P}^*\mathbf{F}_{\mathbf{a}}^*\mathbf{F}_{\mathbf{a}}\mathbf{P}\widehat{\mathbf{c}}'\right\} \\
&= \frac{1}{N}\mathbf{E}\{\widehat{\mathbf{c}}'^*\mathbf{P}^*\mathbf{P}\widehat{\mathbf{c}}'\}
\end{aligned}$$

$$= \frac{1}{N}\mathbf{E}\left\{\mathrm{Tr}\left\{\mathbf{P}\widehat{\mathbf{c}}'\widehat{\mathbf{c}}'^{*}\mathbf{P}^{*}\right\}\right\}$$

$$= \frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'}\mathbf{P}^{*}\right\}, \tag{2.31}$$

where

$$\mathbf{R}_{\widehat{\mathbf{c}}'} = \mathbf{E}\left\{\widehat{\mathbf{c}}'\widehat{\mathbf{c}}'^{*}\right\}.$$

Moreover,

$$\mathbf{E}\left\{\|\mathbf{a} - \widehat{\mathbf{a}}\|^{2}\right\} = \mathbf{E}\left\{\left\|\frac{1}{N}\mathbf{F_a}\mathbf{c} - \frac{1}{N}\mathbf{F_a}\mathbf{P}\widehat{\mathbf{c}}'\right\|^{2}\right\}$$

$$= \frac{1}{N}\mathbf{E}\left\{\|\mathbf{c} - \mathbf{P}\widehat{\mathbf{c}}'\|^{2}\right\}$$

$$= \frac{1}{N}\mathbf{E}\left\{\mathbf{c}^{*}\mathbf{c} - 2\mathrm{Re}\left\{\mathbf{c}^{*}\mathbf{P}\widehat{\mathbf{c}}'\right\} + (\widehat{\mathbf{c}}')^{*}\mathbf{P}^{*}\mathbf{P}\widehat{\mathbf{c}}'\right\}$$

$$= \frac{1}{N}\mathbf{E}\left\{N - 2\mathrm{Re}\left\{\mathbf{c}^{*}\mathbf{P}\widehat{\mathbf{c}}'\right\} + (\widehat{\mathbf{c}}')^{*}\mathbf{P}^{*}\mathbf{P}\widehat{\mathbf{c}}'\right\}$$

$$= 1 - \frac{2}{N}\mathrm{Re}\left\{\mathbf{E}\left\{\mathrm{Tr}\left\{\mathbf{P}\widehat{\mathbf{c}}'\mathbf{c}^{*}\right\}\right\}\right\} + \frac{1}{N}\mathbf{E}\left\{\mathrm{Tr}\left\{\mathbf{P}\widehat{\mathbf{c}}'(\widehat{\mathbf{c}}')^{*}\mathbf{P}^{*}\right\}\right\}$$

$$= 1 - \frac{2}{N}\mathrm{Re}\left\{\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'\mathbf{c}}\right\}\right\} + \frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'}\mathbf{P}^{*}\right\}, \tag{2.32}$$

where

$$\mathbf{R}_{\widehat{\mathbf{c}}'\mathbf{c}} = \mathbf{E}\left\{\widehat{\mathbf{c}}'\mathbf{c}^{*}\right\}.$$

With (2.31) and (2.32), expression (2.30) becomes

$$\mathrm{SNR}_{\mathrm{prop}} = \frac{\frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'}\mathbf{P}^{*}\right\} \cdot \mathrm{SNR}_0}{\left(1 - \frac{2}{N}\mathrm{Re}\left\{\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'\mathbf{c}}\right\}\right\} + \frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'}\mathbf{P}^{*}\right\}\right) \cdot \mathrm{SNR}_0 + 1}. \tag{2.33}$$

To further simplify (2.33), we note that for each particular $\mathbf{c}$, the optimal $\widehat{\mathbf{c}}'$ that minimizes

$$\|\mathbf{c} - \mathbf{P}\widehat{\mathbf{c}}'\|^{2}$$

is given by

$$\widehat{\mathbf{c}}'_o = (\mathbf{P}^{*}\mathbf{P})^{-1}\mathbf{P}^{*}\mathbf{c}.$$

We thus make the following approximation:

$$\widehat{\mathbf{c}}' \approx \widehat{\mathbf{c}}'_o = (\mathbf{P}^{*}\mathbf{P})^{-1}\mathbf{P}^{*}\mathbf{c}.$$

It follows immediately that

$$\mathbf{R}_{\widehat{\mathbf{c}}'} \approx \mathrm{E}\left\{\widehat{\mathbf{c}}'_o\left(\widehat{\mathbf{c}}'_o\right)^*\right\}$$

$$= \mathrm{E}\left\{\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{cc}^*\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\right\}$$

$$= \left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}$$

and

$$\mathbf{R}_{\widehat{\mathbf{c}}'\mathbf{c}} \approx \mathrm{E}\left\{\widehat{\mathbf{c}}'_o\mathbf{c}^*\right\}$$

$$= \mathrm{E}\left\{\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{cc}^*\right\}$$

$$= \left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}.$$

We then have

$$\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'}\mathbf{P}^*\right\} \approx \mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\right\}$$

$$= \mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\right\}$$

and

$$\mathrm{Re}\left\{\mathrm{Tr}\left\{\mathbf{P}\mathbf{R}_{\widehat{\mathbf{c}}'\mathbf{c}}\right\}\right\} \approx \mathrm{Re}\left\{\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\right\}\right\}$$

$$= \mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\right\}.$$

Hence, (2.33) is approximated by

$$\mathrm{SNR}_{\mathrm{prop}} \approx \frac{\frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\right\}\cdot\mathrm{SNR}_0}{\left(1-\frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\right\}\right)\cdot\mathrm{SNR}_0+1}.$$

If the interpolation matrix is given by $\mathbf{P} = \mathbf{P}_o = \mathbf{R}_{\mathbf{cc}'}\mathbf{R}_{\widehat{\mathbf{c}}'}^{-1}$, then

$$\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R}_{\mathbf{c}}\right\} = \mathrm{Tr}\left\{\mathbf{R}_{\mathbf{cc}'}\left(\mathbf{R}_{\mathbf{cc}'}^*\mathbf{R}_{\mathbf{cc}'}\right)^{-1}\mathbf{R}_{\mathbf{cc}'}^*\mathbf{R}_{\mathbf{c}}\right\}$$

and hence

$$\mathrm{SNR}_{\mathrm{prop}} \approx \frac{\frac{1}{N}\mathrm{Tr}\left\{\mathbf{R}_{\mathbf{cc}'}\left(\mathbf{R}_{\mathbf{cc}'}^*\mathbf{R}_{\mathbf{cc}'}\right)^{-1}\mathbf{R}_{\mathbf{cc}'}^*\mathbf{R}_{\mathbf{c}}\right\}\cdot\mathrm{SNR}_0}{\left(1-\frac{1}{N}\mathrm{Tr}\left\{\mathbf{R}_{\mathbf{cc}'}\left(\mathbf{R}_{\mathbf{cc}'}^*\mathbf{R}_{\mathbf{cc}'}\right)^{-1}\mathbf{R}_{\mathbf{cc}'}^*\mathbf{R}_{\mathbf{c}}\right\}\right)\cdot\mathrm{SNR}_0+1}. \qquad (2.34)$$

In Figs. 2.6-2.8, the theoretical effective signal-to-noise ratio is plotted for different compensation schemes and for different types of phase noise by using expressions (2.27)-(2.29) and (2.34) (see Appendix A for a brief introduction of the different types of phase noise). Fig. 2.6(a) plots the effective signal-to-noise ratio vs. SNR_0 for the free-running oscillators with linewidth $\xi = 5$ kHz, and Fig. 2.6(b) plots the effective signal-to-noise ratio vs. the oscillator linewidth ξ when $SNR_0 = 25$ dB. It is seen in Figs. 2.6(a) and 2.6(b) that SNR_{no} is about -3.0 dB for free-running oscillators. This is because

$$2 - 2\mathrm{Re}\left\{\mathbf{E}\left\{c_{\mathrm{Free}}(t)\right\}\right\} = 2.$$

From (2.27),

$$SNR_{no} = \frac{SNR_0}{2SNR_0 + 1} \approx 0.5 = -3.0 \text{ dB}$$

if $SNR_0 \gg 1$. It can be also interpreted as follows. The power of the carrier signal is

$$\mathbf{E}\left\{|c_{\mathrm{Free}}(t)|^2\right\} = 1$$

and the variance of the carrier noise is

$$\mathbf{E}\left\{|c_{\mathrm{Free}}(t) - 1|^2\right\} = 2 - 2\mathrm{Re}\left\{\mathbf{E}\left\{c_{\mathrm{Free}}(t)\right\}\right\} = 2,$$

which implies the -3.0 dB uncompensated signal-to-noise ratio. In Figs. 2.7(b) and 2.8(b), the phase noise variance is determined by the linewidth ξ and the loop bandwidth f_L or f_n according to expressions (A.2) and (A.4), respectively, and is measured in dB with respect to the carrier power, namely, dBc. In the next section, we will compare the theoretical system performance with the simulated system performance and discuss the observations from the plots.
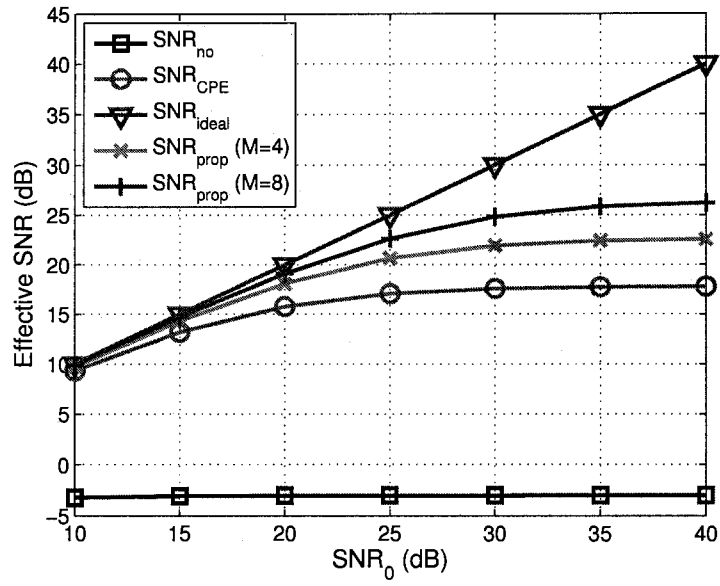
(a) $\xi = 5$ kHz.
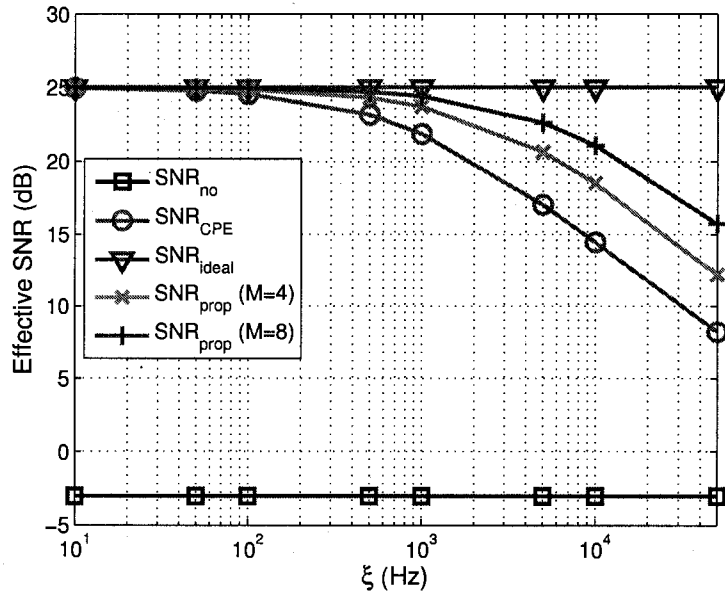
(b) $\text{SNR}_0 = 25$ dB.

Figure 2.6: Plots of theoretical effective signal-to-noise ratio for the free-running oscillators by using expressions (2.27)-(2.29) and (2.34).
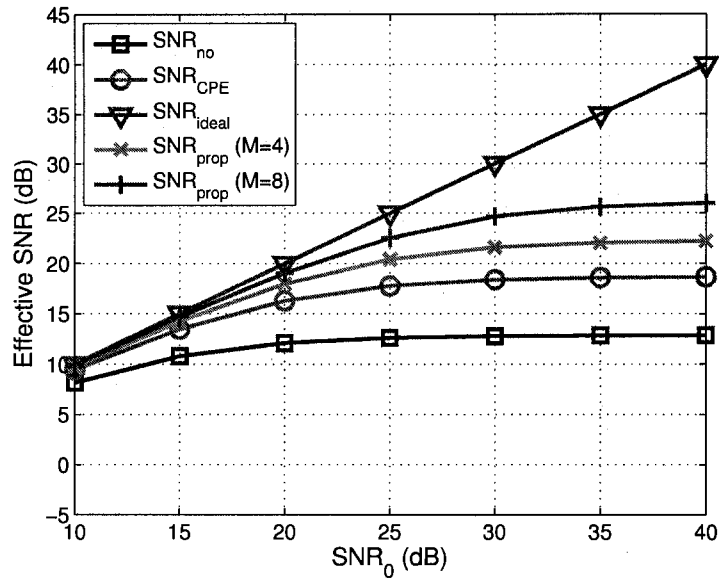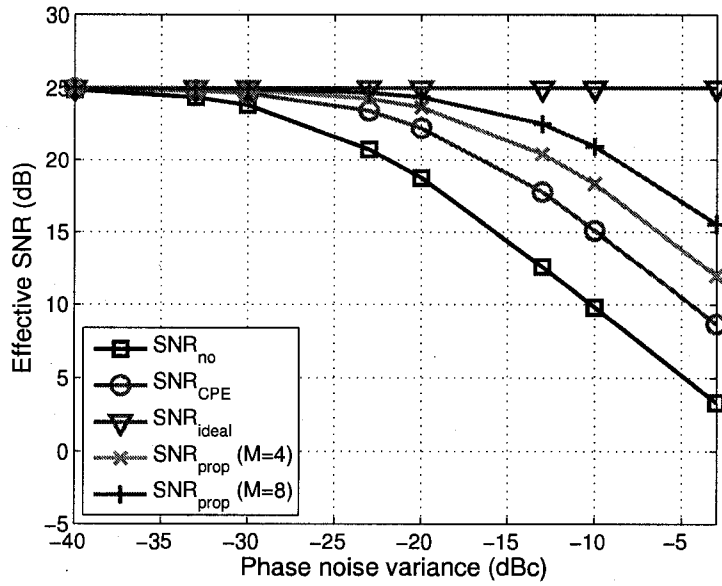
(a) $\xi = 5$ kHz.

(b) $SNR_0 = 25$ dB.

Figure 2.7: Plots of theoretical effective signal-to-noise ratio for the 1st-order PLL oscillators with $f_L = 50$ kHz by using expressions (2.27)-(2.29) and (2.34).
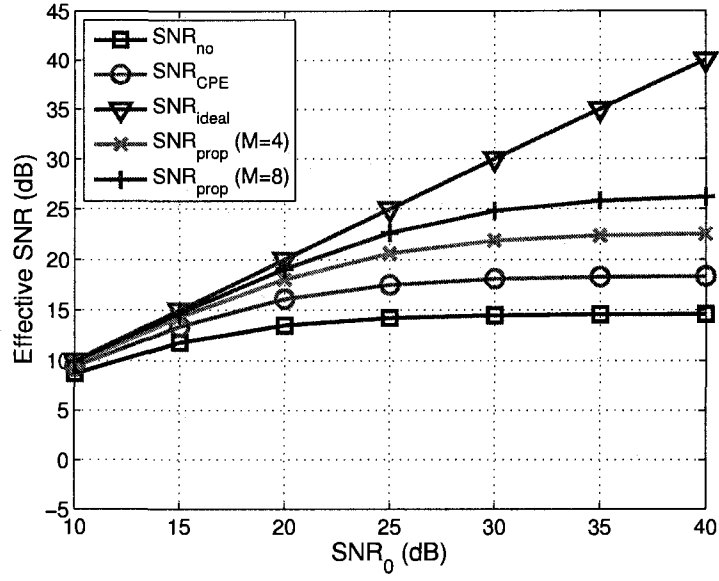
(a) $\xi = 5$ kHz.
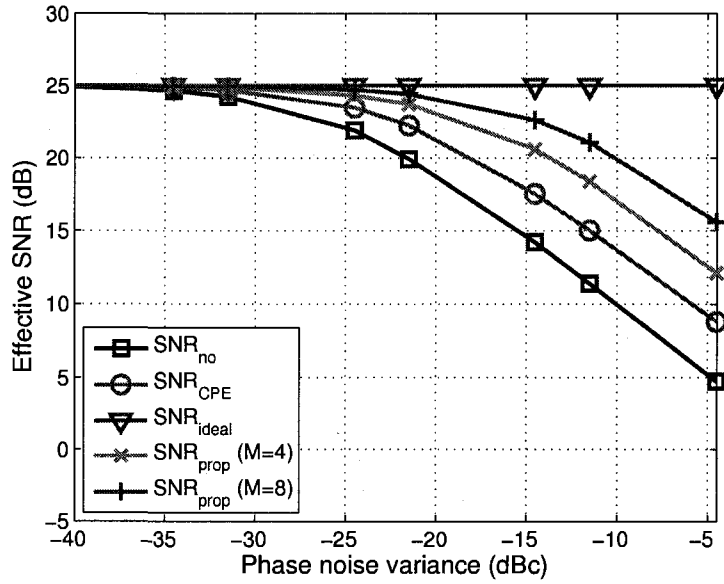
(b) $SNR_0 = 25$ dB.

Figure 2.8: Plots of theoretical effective signal-to-noise ratio for the 2nd-order PLL oscillators with $f_n = 50$ kHz by using expressions (2.27)-(2.29) and (2.34).

2.5   Computer Simulations

The proposed scheme is simulated in comparison with the ideal OFDM receiver with perfect phase noise compensation, the CPE correction scheme of [RK95], and the ICI compensation schemes proposed in [ZH01, GNE03, LZ04]. The system bandwidth is 20 MHz, i.e., $T_s = 0.05$ μs, and the constellation used for symbol mapping is 16-QAM. The OFDM symbol size is $N = 64$ and the prefix length is $P = 16$. The channel length is 6, and each tap is independently Rayleigh distributed with the power profile specified by 3 dB decay per tap. Two scenarios are studied in the simulations. One scenario assumes that the receiver has perfect channel information, while the other scenario assumes that the receiver has to estimate the channel using pilot symbols. Only one block-type pilot symbol is used in the proposed scheme for each time of channel estimation, while two are used in the CPE correction scheme. The assumed channel length in the time domain for channel estimation is $L = 12$, the length of the phase noise vector to be estimated is $M = 4$ or 8, and the number of pilot tones in the comb-type symbols is $Q = 8$ or 16. The phase noise generated by the free-running oscillators and the phase-loop locked oscillators is simulated according to the models given in Appendix A.

Figs. 2.9-2.12 show the simulated system performance when the phase noise is generated by a free-running oscillator. The phase noise spectrum for $\xi = 5$ kHz is plotted in Fig. 2.9. Figs. 2.10-2.11 compare different schemes in terms of the effective signal-to-noise ratio and the uncoded bit error rate (BER) for the phase noise spectrum with $\xi = 5$ kHz. Compared to the CPE correction scheme, the proposed scheme can achieve 5–8 dB more improvement in the effective signal-to-noise ratio for the high signal-to-noise ratio regime (more than 20 dB). Fig. 2.12 shows the system performance in terms of the effective signal-to-noise ratio for

different oscillator linewidth. The larger the linewidth, the more random the phase noise is, and consequently, the more difficult it is to compensate for the phase noise. It can be seen that compared to the CPE correction scheme, the proposed algorithm with $Q = 16$ and $M = 8$ can improve the signal-to-noise ratio by about 5 dB if perfect channel information is available (i.e., no channel estimation) and 8 dB if the receiver has to estimate the channel. The comparison between the cases with and without perfect channel information shows that the joint channel and phase noise estimation further improves the overall system performance. In other words, the proposed scheme can reduce the sensitivity of OFDM receivers to phase noise by about 8 dB. Moreover, the simulated system performance displayed here is consistent with the theoretical performance shown in Fig. 2.6. It is also demonstrated in Figs. 2.13-2.16 and 2.17-2.20 that the proposed scheme can significantly improve the effective signal-to-noise ratio and uncoded BER for the phase-loop locked oscillators. The improvement in the effective signal-to-noise ratio can be more than 8 dB. In order to further improve the effective signal-to-noise ratio and uncoded BER, we can add more pilot tones to increase Q and M; however, this reduces the data rate and spectral efficiency.

In Fig. 2.21, the proposed phase noise compensation scheme is compared with the self-cancellation scheme proposed in [ZH01], the frequency-domain FIR-type equalizer proposed in [GNE03], and the ICI suppression method using sinusoidal approximation proposed in [LZ04]. For fairness, we assume that the receivers have perfect channel information, without worrying about the overhead introduced by channel estimation.[3] The phase noise is generated by a 1st-order phase-loop locked oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz. It is shown in the figure

---

[3]The block-type pilot symbols used for channel estimation are only required at the beginning of each packet, because wireless channels are usually slowly time-varying. If the packets are long enough, then the overhead caused by the block-type symbols is negligible.
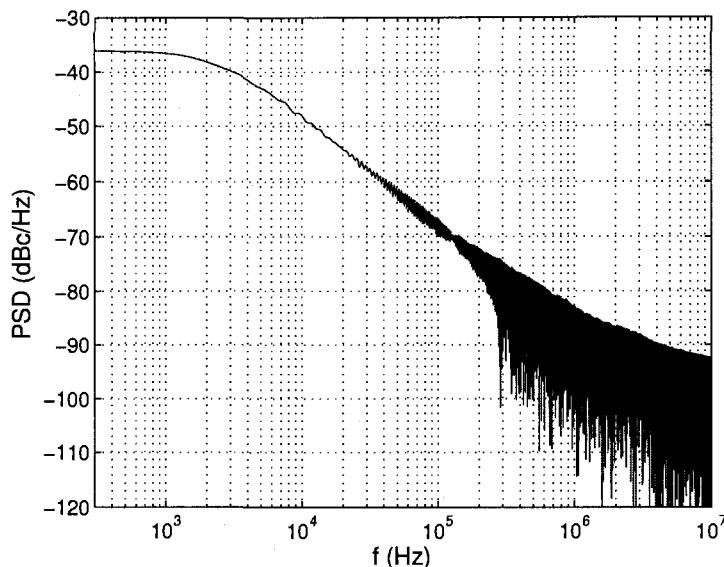
Figure 2.9: PSD of the phase noise generated by a free-running oscillator with $\xi = 5$ kHz.

that the FIR-type equalizer[4] and the ICI suppression method using sinusoidal approximation do not perform well for this type of fast-changing phase noise because they can only compensate for the ICI from adjacent subcarriers. The self-cancellation scheme works as well as the proposed method since it uses two subcarriers to transmit one data symbol, which brings the benefit of diversity gain as in MIMO communications. However, the self-cancellation method reduces the spectral efficiency by one half. In the simulation, each OFDM symbol can transmit $N - Q = 48$ data symbols in the proposed scheme but only $N/2 = 32$ symbols in the self-cancellation scheme, which demonstrates that the proposed method can achieve 50% more spectral efficiency than the self-cancellation method.

---

[4]In the simulations, the length of the FIR-type equalizer is 9.

2.6 Conclusions

In this chapter, a phase noise compensation scheme is proposed for OFDM-based wireless communications. The proposed scheme consists of two phases. One phase is the joint channel and phase noise estimation, and the other phase is the joint data symbol and phase noise estimation. The simulations show that the proposed scheme can effectively improve the system performance in terms of the effective signal-to-noise ratio and the uncoded bit error rate. It is demonstrated that the proposed algorithm can reduce the sensitivity of OFDM receivers to phase noise by about 8 dB. Since oscillators with ultra-low phase noise usually have the disadvantage of high implementation cost and high power consumption, the improvement will significantly reduce the cost and power consumption from the perspective of hardware designers.
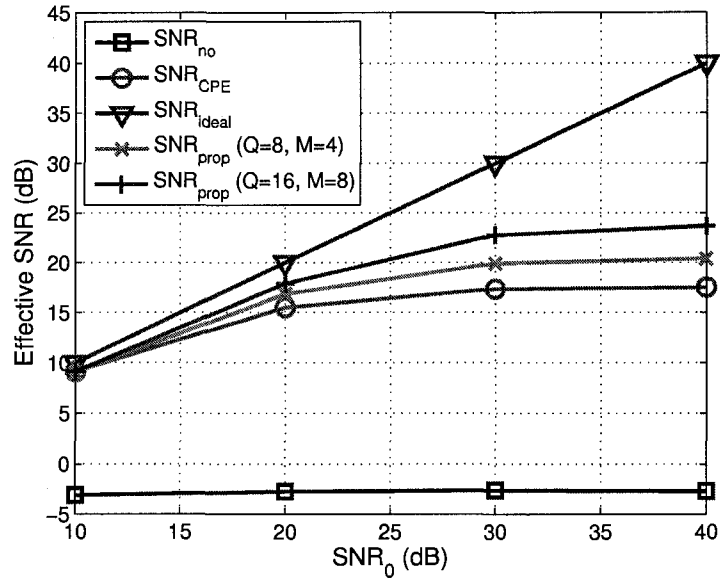
(a) Effective SNR.
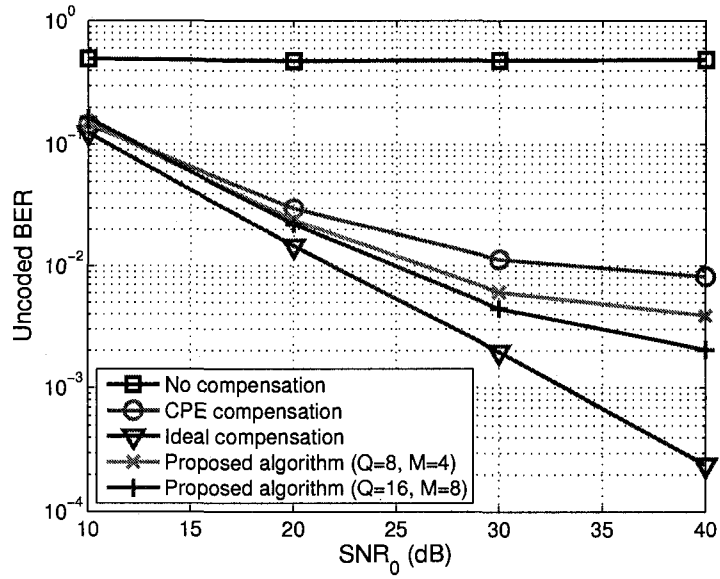
(b) Uncoded BER.

Figure 2.10: Simulation results with perfect channel information when the phase noise is generated by a free-running oscillator with $\xi = 5$ kHz.
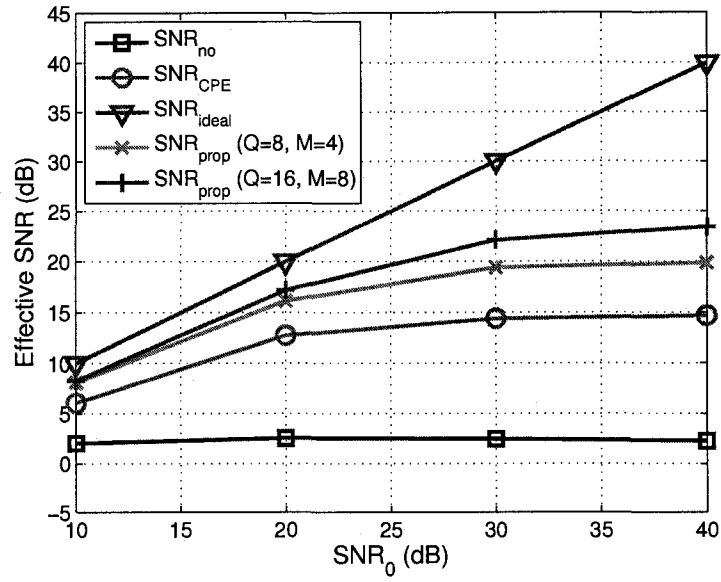
(a) Effective SNR.
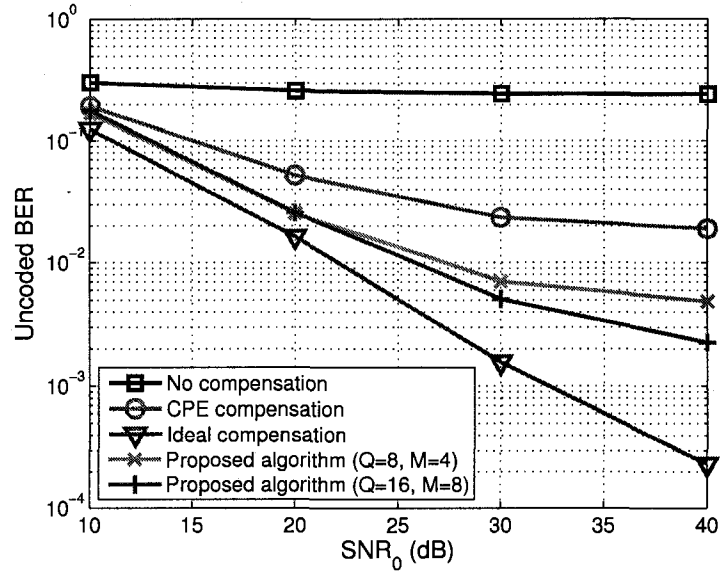
(b) Uncoded BER.

Figure 2.11: Simulation results with estimated channel information when the phase noise is generated by a free-running oscillator with $\xi = 5$ kHz.
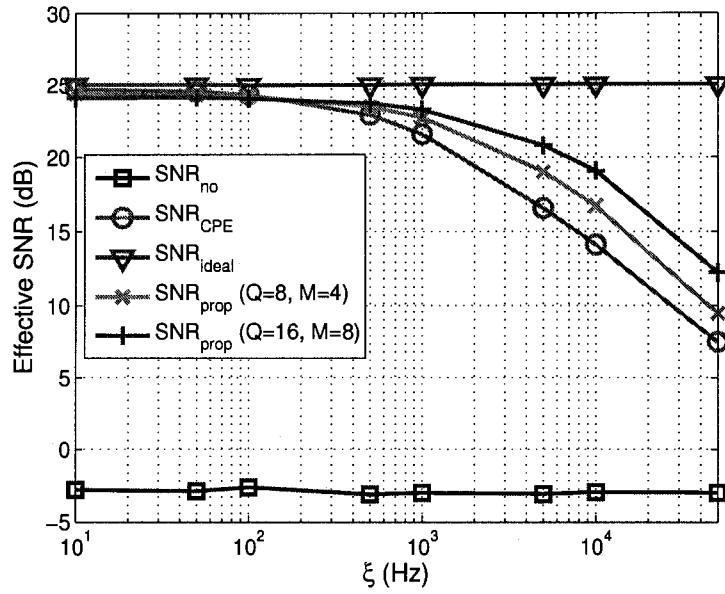
(a) With perfect channel information.
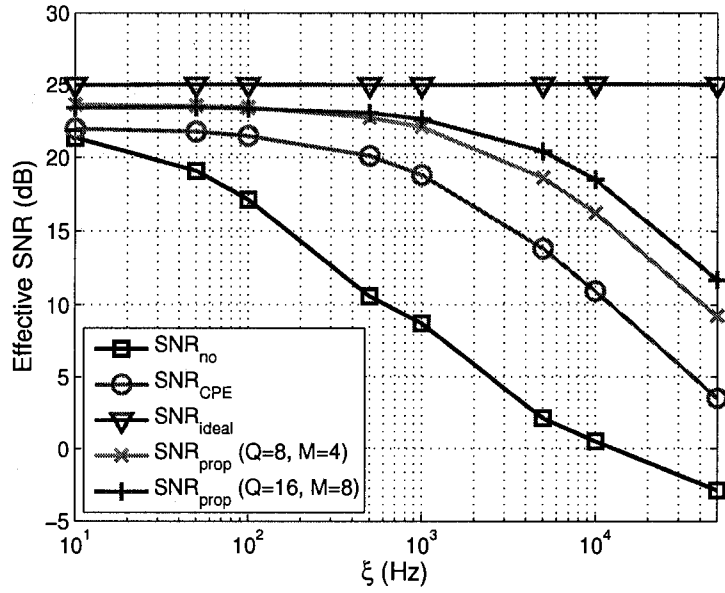
(b) With estimated channel information.

Figure 2.12: Effective signal-to-noise ratio vs. ξ for $SNR_0 = 25$ dB when the phase noise is generated by free-running oscillators.
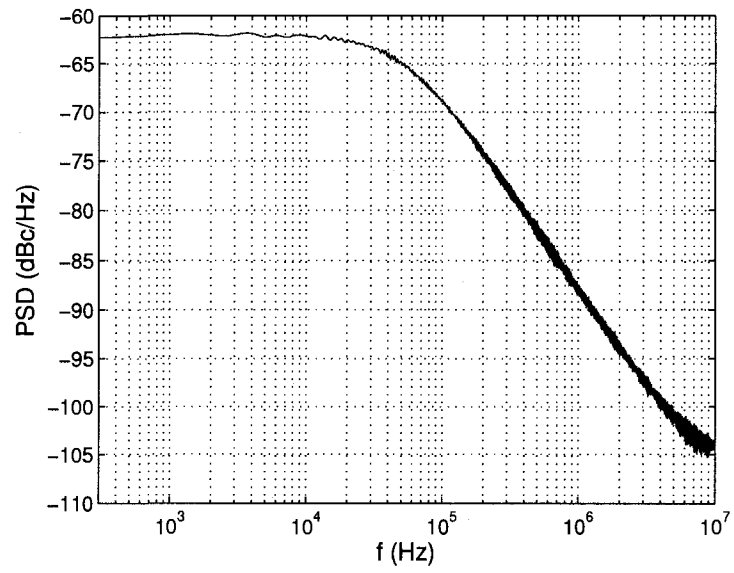
Figure 2.13: PSD of the phase noise generated by a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.
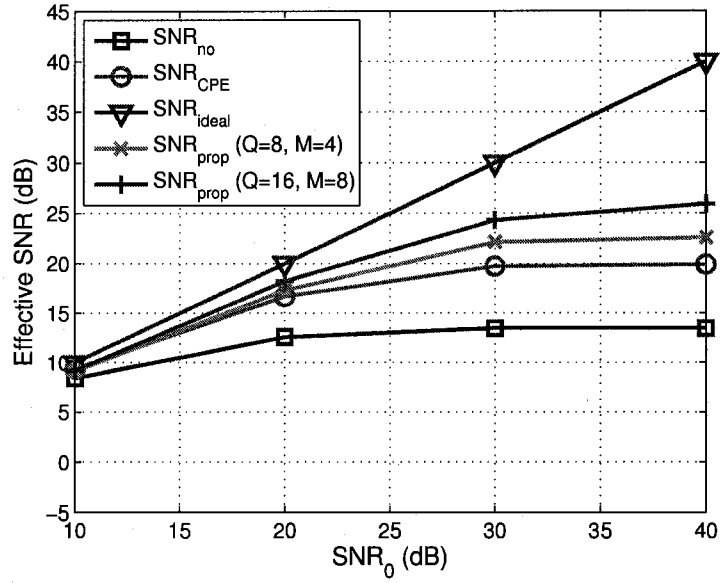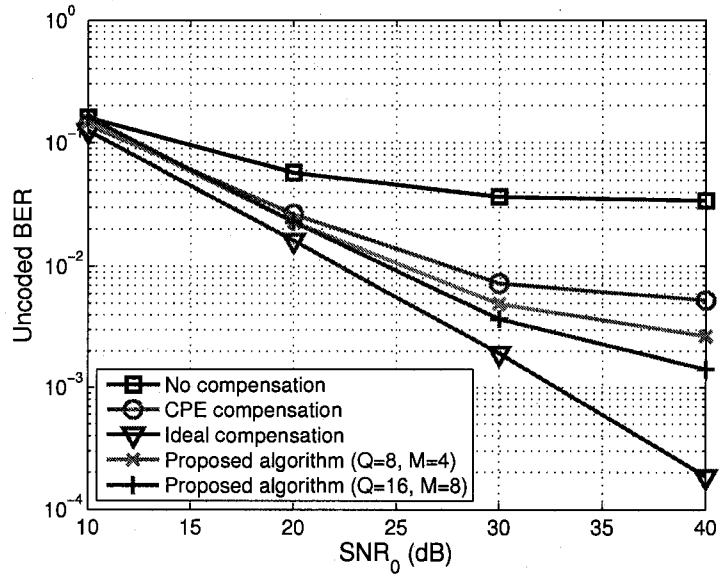
(a) Effective SNR.

(b) Uncoded BER.

Figure 2.14: Simulation results with perfect channel information when the phase noise is generated by a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.
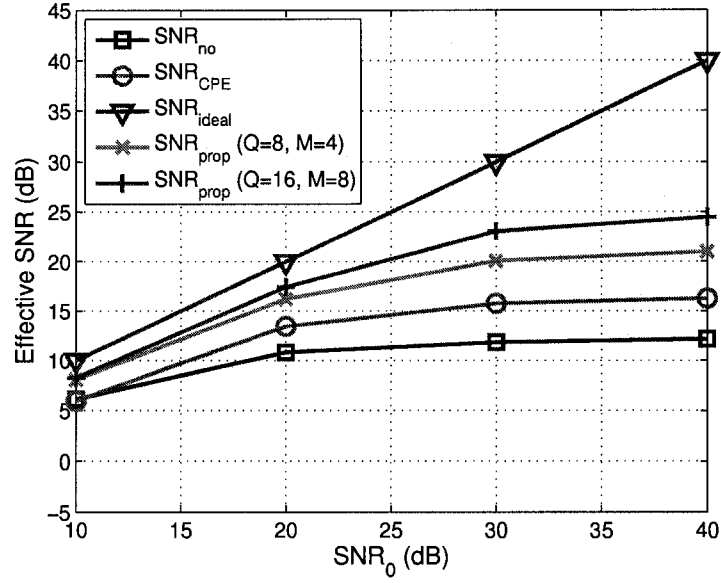
(a) Effective SNR.
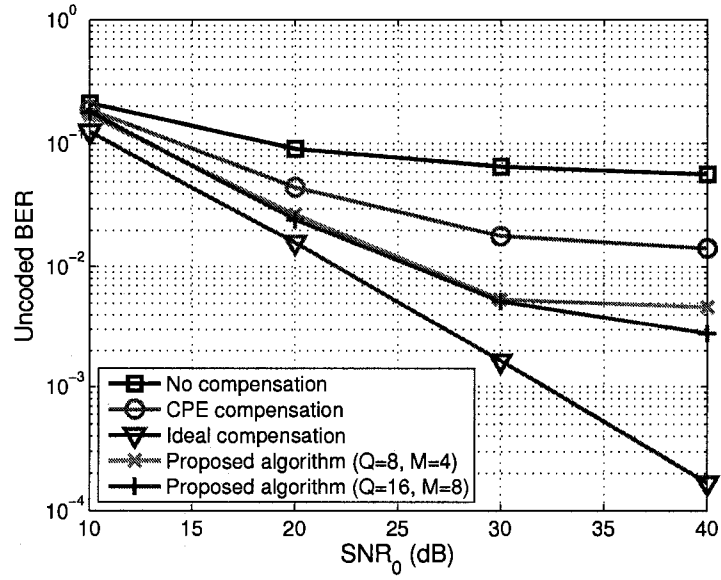
(b) Uncoded BER.

Figure 2.15: Simulation results with estimated channel information when the phase noise is generated by a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.
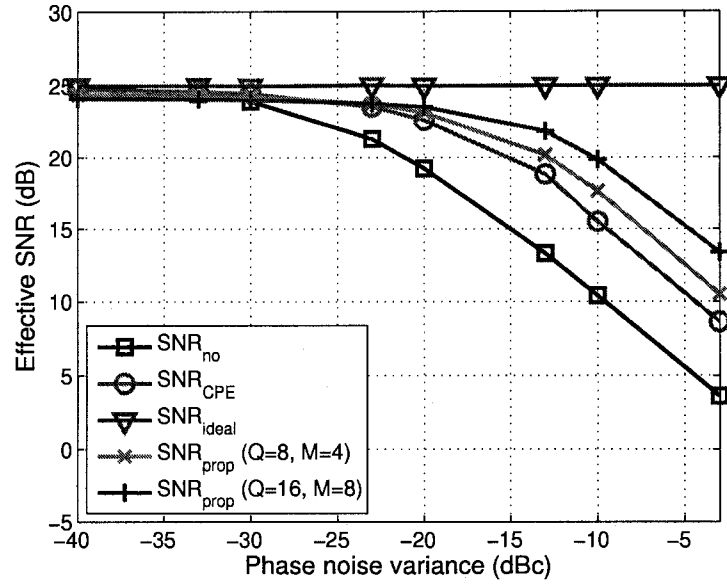
(a) With perfect channel information.
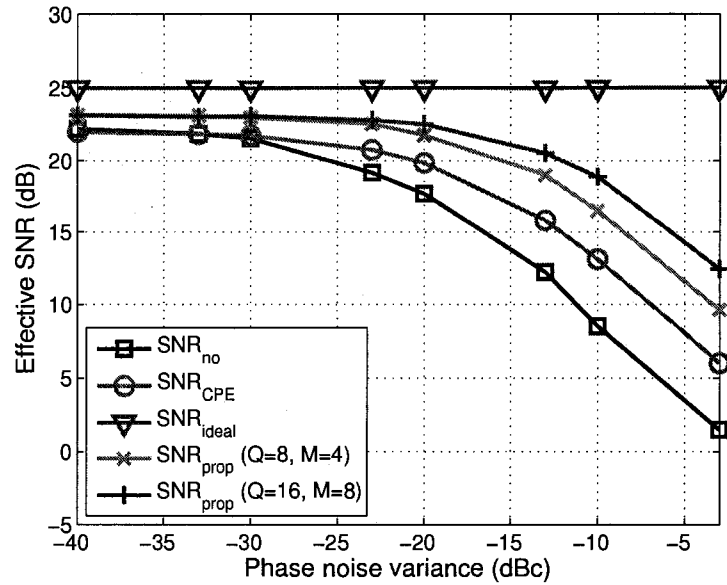
(b) With estimated channel information.

Figure 2.16: Effective signal-to-noise ratio vs. phase noise variance when the phase noise is generated by 1st-order PLL oscillators with $f_L = 50$ kHz.
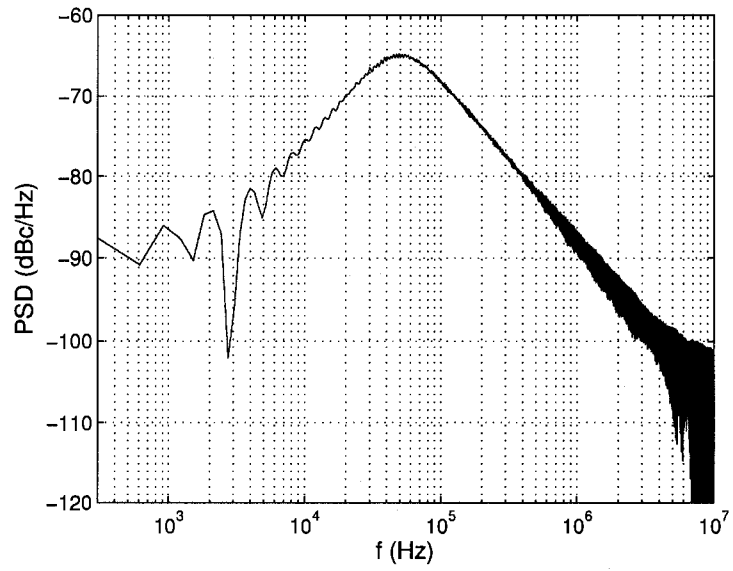
Figure 2.17: PSD of the phase noise generated by a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.
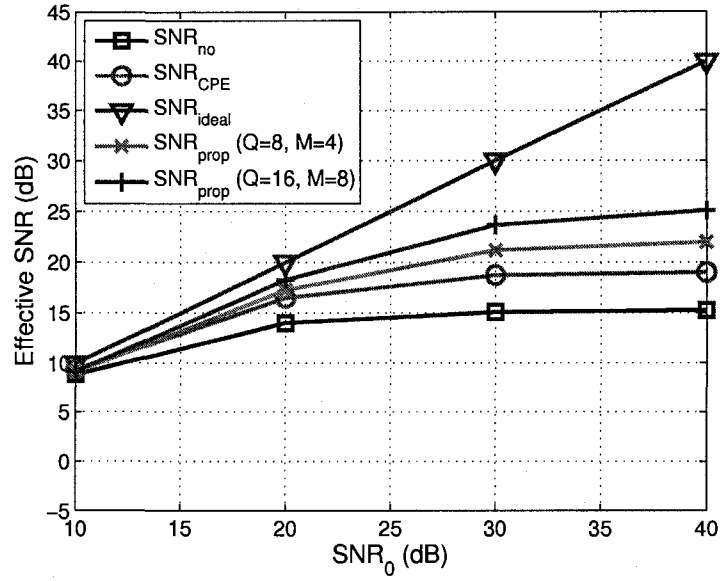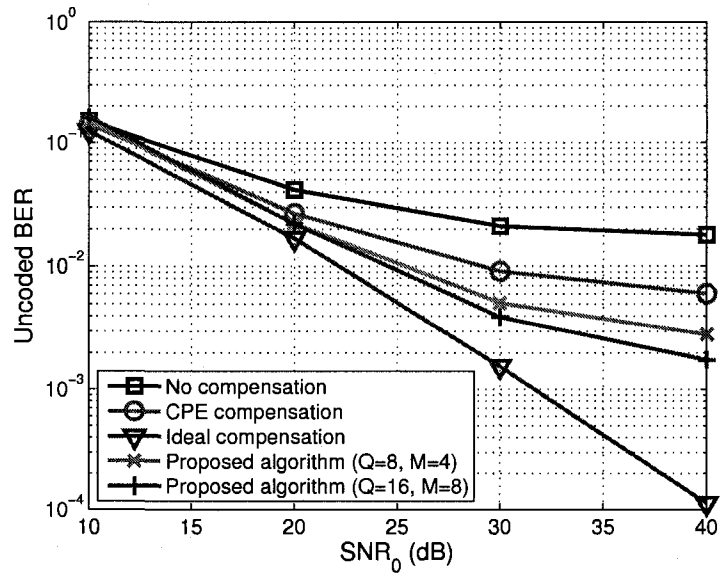
(a) Effective SNR.

(b) Uncoded BER.

Figure 2.18: Simulation results with perfect channel information when the phase noise is generated by a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.
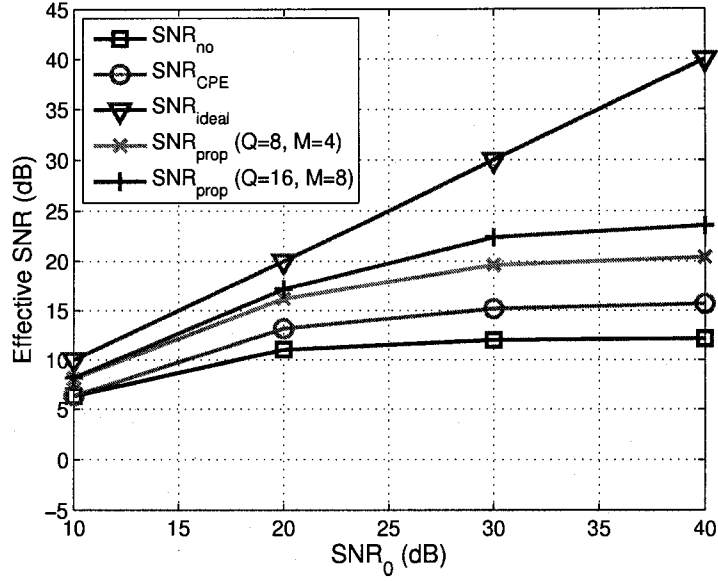
(a) Effective SNR.
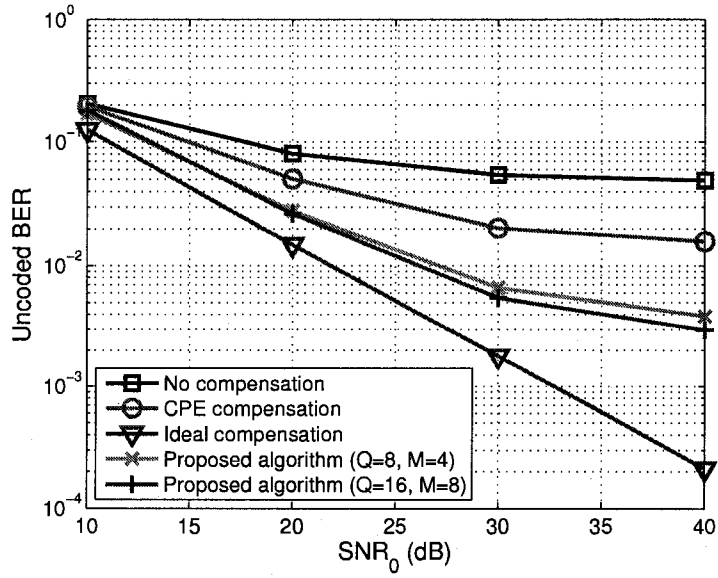
(b) Uncoded BER.

Figure 2.19: Simulation results with estimated channel information when the phase noise is generated by a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.
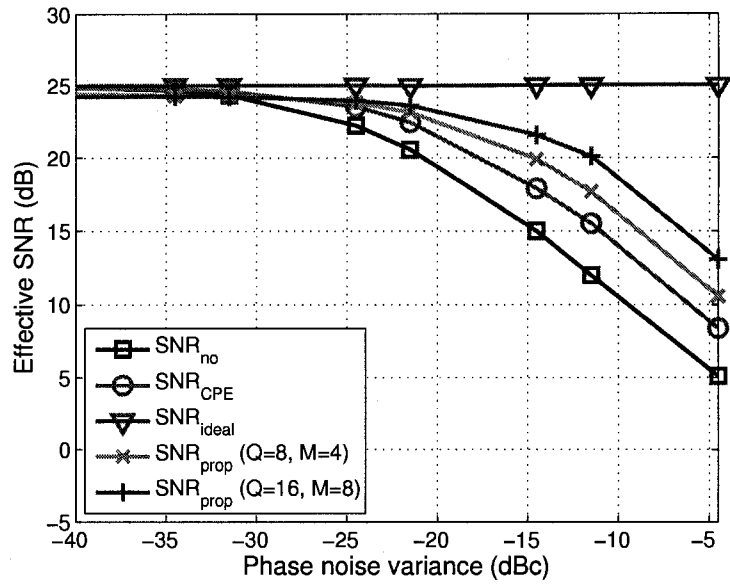
(a) With perfect channel information.
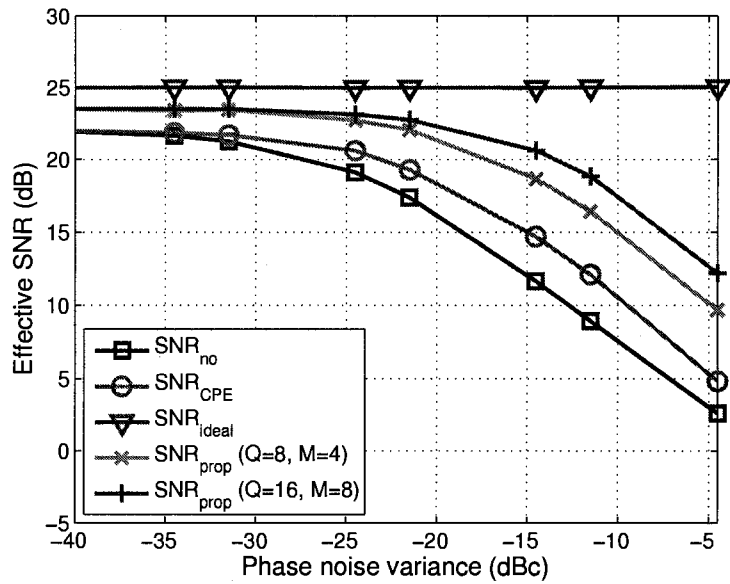
(b) With estimated channel information.

Figure 2.20: Effective signal-to-noise ratio vs. phase noise variance when the phase noise is generated by 2nd-order PLL oscillators with $f_n = 50$ kHz.
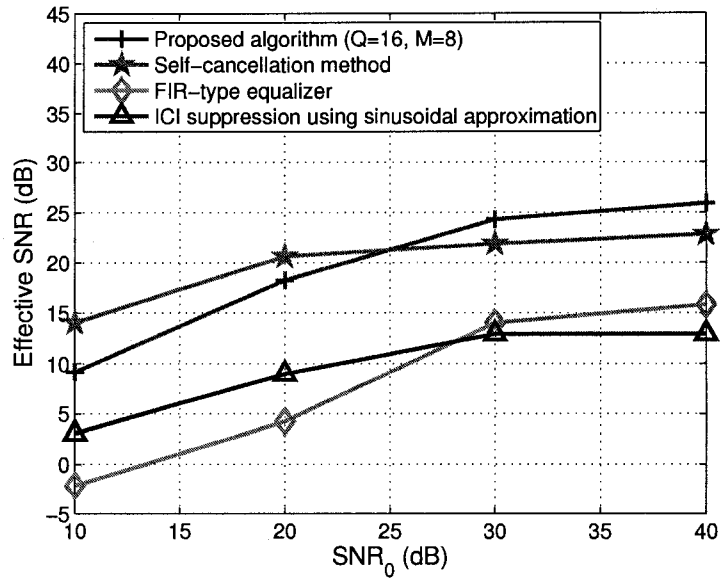
(a) Effective SNR.
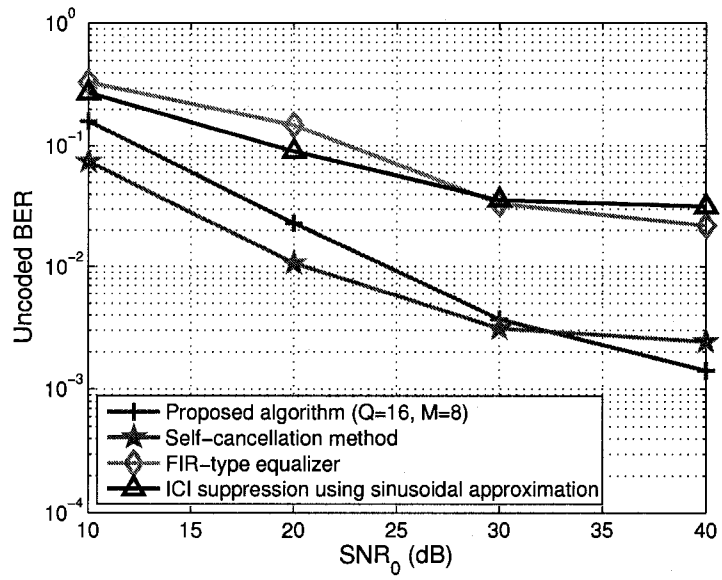
(b) Uncoded BER.

Figure 2.21: Comparison of different ICI compensation methods.

CHAPTER 3

Joint Compensation of IQ Imbalance and Phase Noise

3.1 Introduction

There are mainly two types of receiver structures for OFDM systems: one is the direct-conversion receiver and the other is the Heterodyne receiver [Raz98]. Compared to the Heterodyne receivers, the direct-conversion receivers have the advantages of low cost, low power consumption and easy integration, but they suffer more severely from analog domain impairments. One such impairment is caused by the imperfectness in the process of RF signal down-conversion to baseband. The effects of this impairment have been modeled as IQ imbalance and phase noise in the literature [Abi95, Raz98]. IQ imbalance is the mismatch in amplitude and phase between the I and Q branches in the receiver chain, while phase noise is the random unknown phase difference between the phase of the carrier signal and the phase of the local oscillator (as was explained in Chapter 2). The effects of IQ imbalance and phase noise on OFDM receivers have been investigated in previous works, such as [Liu98, BSH00, PBM95, Mus95, Tom98, PM02]. Some algorithms have also been proposed for the compensation of IQ imbalance [SHA01, VRK01, TBS05, TS05] or the compensation of phase noise [RK95, ZH01, CBY02, GNE03, LZ04, RLP05, LPL06], separately. In [TCP05],

59

the joint effects of IQ imbalance and phase noise on OFDM systems are studied, but the analysis and proposed compensation scheme are based on the concatenation model of IQ imbalance and phase noise, where only the common error term of phase noise is considered. It appears that there is still no thorough work on analyzing the system performance degradation caused by the coexistence of IQ imbalance and phase noise, as well as the optimal estimation of channel response and data symbols, in the presence of both impairments.

In this chapter, we pursue an explicit formulation for the joint effects of IQ imbalance and phase noise, and propose a joint compensation scheme with performance analysis. The scheme consists of a joint channel estimation algorithm and a joint data symbol estimation algorithm. The joint estimation technique achieves a more accurate channel estimate than other conventional methods that either ignore the impairments or simply model them as additive Gaussian noise. The mean-square-error of the channel estimation step is compared with the associated Cramer-Rao lower bound to conclude that our scheme works well with performance close to the ideal case without the impairments. In the proposed data symbol estimation algorithm, it is shown that the joint compensation can be decomposed into the IQ imbalance compensation followed by the phase noise compensation. During the payload portion of OFDM packets, which contains both data tones and pilot tones, the data symbols and the phase noise components are jointly estimated at the receiver. The performance of the proposed algorithm is analyzed in terms of the improvements in the effective signal-to-noise ratio, and is compared with other compensation methods.

This chapter is organized as follows. The next section describes the system model and formulates the joint effects of IQ imbalance and phase noise. The proposed joint channel estimation algorithm is presented in Section 3.3, and the

Cramer-Rao lower bounds for estimating the channel response are also derived. The joint data estimation algorithm is presented in Section 3.4, and its performance is analyzed in terms of the effective signal-to-noise ratio degradation. Section 3.5 discusses the complexity of the proposed algorithms and presents an efficient implementation scheme. Simulation results and performance comparison of different algorithms are discussed in Section 3.6.

Throughout this chapter, we adopt the following notation. $(\cdot)^T$ denotes the matrix transpose, $(\cdot)^*$ represents the matrix conjugate transpose, and $\text{conj}\{\cdot\}$ takes the complex conjugate of its argument elementwisely. $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ return the real and imaginary parts of its argument, respectively. $\text{diag}\{\cdot\}$ represents the diagonal matrix whose diagonal entries are determined by its argument. $\text{Tr}\{\cdot\}$ returns the trace of a matrix. $\text{E}\{\cdot\}$ is the expected value with respect to the underlying probability measure. $\mathbf{I}_K$ is the identity matrix of size $K \times K$, and $\mathbf{I}_\theta$ is the Fisher information matrix associated with the parameter vector $\boldsymbol{\theta}$.

3.2   System Model

Fig. 3.1 shows the block diagrams of a direct-conversion receiver with and without IQ imbalance and phase noise. We first formulate the effects of IQ imbalance and phase noise on the received continuous-time baseband signals in Subsection 3.2.1, and then discuss their effects on the received OFDM symbols in Subsection 3.2.2.
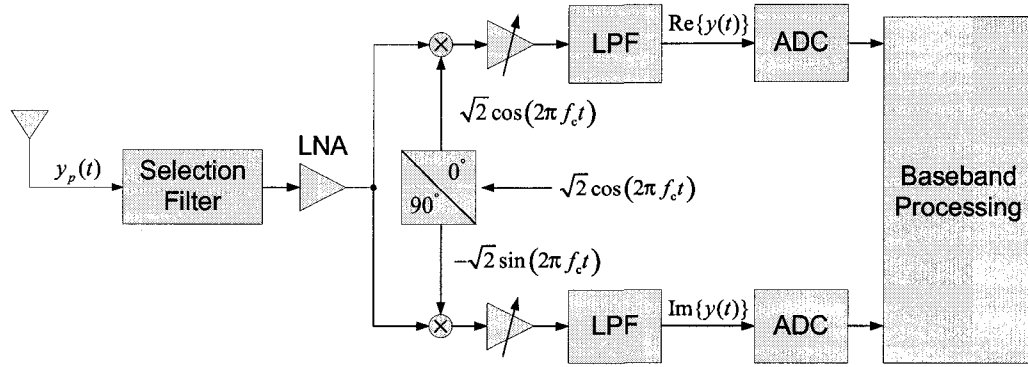
(a) A direct-conversion receiver with perfect oscillator signals.
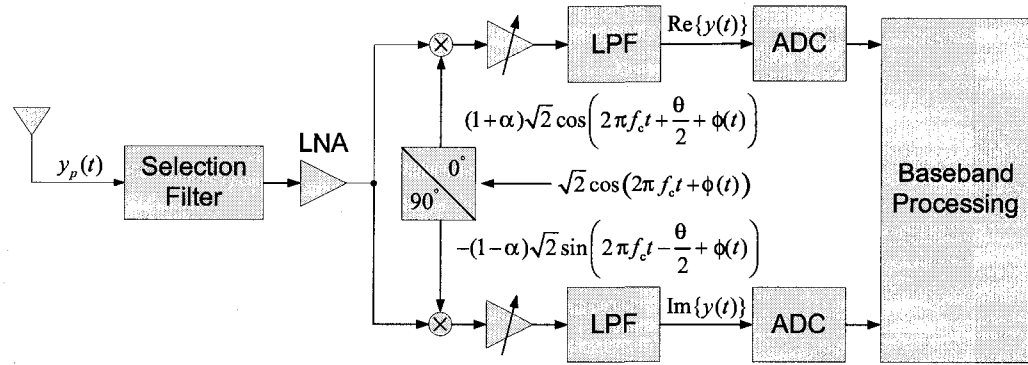
(b) A direct-conversion receiver with IQ imbalance and phase noise.

Figure 3.1: Block diagrams of a direct-conversion receiver with and without IQ imbalance and phase noise.

3.2.1 IQ Imbalance and Phase Noise

Let $x(t)$ be the transmitted continuous-time baseband signal. The transmitted passband signal, namely, the radio-frequency signal $x_p(t)$ is given by

$$x_p(t) = \mathrm{Re}\big\{\sqrt{2}x(t)e^{j2\pi f_c t}\big\} \tag{3.1}$$
$$= \sqrt{2}\mathrm{Re}\big\{x(t)\big\}\cos(2\pi f_c t) - \sqrt{2}\mathrm{Im}\big\{x(t)\big\}\sin(2\pi f_c t),$$

where f_c is the carrier frequency and the normalization factor $\sqrt{2}$ ensures that $x(t)$ and $x_p(t)$ have the same average power. Let $h_p(t)$ be the continuous-time impulse response function of the passband channel. Then the received passband signal $y_p(t)$ is given by the convolutional integral of $x_p(t)$ and $h_p(t)$ plus additive passband white Gaussian noise $w_p(t)$, i.e.,

$$y_p(t) = \int_{-\infty}^{\infty} h_p(t-\tau)x_p(\tau)d\tau + w_p(t). \tag{3.2}$$

Let

$$h(t) = h_p(t)e^{-j2\pi f_c t} \tag{3.3}$$

represent the continuous-time impulse response function of the equivalent baseband channel. We use

$$y_0(t) = \int_{-\infty}^{\infty} h(t-\tau)x(\tau)d\tau \tag{3.4}$$

to represent the received baseband signal in the absence of the impairments and noise. Note that the passband signals $x_p(t)$, $y_p(t)$ and the passband channel response $h_p(t)$ are all real functions, while the baseband signal $x(t)$ and the baseband channel response $h(t)$ are complex. It then follows from (3.1)-(3.4)

that $y_p(t)$ can be expressed as

$$
\begin{aligned}
y_p(t) &= \int_{-\infty}^{\infty} h_p(t-\tau) x_p(\tau) d\tau + w_p(t) \\
&= \int_{-\infty}^{\infty} h_p(t-\tau) \operatorname{Re}\left\{\sqrt{2} x(\tau) e^{j2\pi f_c \tau}\right\} d\tau + w_p(t) \\
&= \sqrt{2} \int_{-\infty}^{\infty} \operatorname{Re}\left\{h_p(t-\tau) x(\tau) e^{j2\pi f_c \tau}\right\} d\tau + w_p(t) \\
&= \sqrt{2} \int_{-\infty}^{\infty} \operatorname{Re}\left\{h(t-\tau) e^{j2\pi f_c (t-\tau)} x(\tau) e^{j2\pi f_c \tau}\right\} d\tau + w_p(t) \\
&= \sqrt{2} \int_{-\infty}^{\infty} \operatorname{Re}\left\{h(t-\tau) x(\tau) e^{j2\pi f_c t}\right\} d\tau + w_p(t) \\
&= \sqrt{2} \operatorname{Re}\left\{\left(\int_{-\infty}^{\infty} h(t-\tau) x(\tau) d\tau\right) e^{j2\pi f_c t}\right\} + w_p(t) \\
&= \operatorname{Re}\left\{\sqrt{2} y_0(t) e^{j2\pi f_c t}\right\} + w_p(t).
\end{aligned}
\tag{3.5}
$$

In an ideal direct-conversion receiver, as shown in Fig. 3.1(a), the sinusoidal signals used for I and Q-branch mixing have the same amplitude and are orthogonal to each other. Also, their phase is perfectly aligned with the carrier signal. Rewriting (3.5) as

$$
\begin{aligned}
y_p(t) &= \operatorname{Re}\left\{\sqrt{2} y_0(t) e^{j2\pi f_c t}\right\} + w_p(t) \\
&= \operatorname{Re}\left\{\sqrt{2}\left[y_0(t) + \frac{\sqrt{2}}{2} w_p(t) e^{-j2\pi f_c t}\right] e^{j2\pi f_c t}\right\},
\end{aligned}
$$

the received baseband signal after down-conversion is

$$
y(t) = y_0(t) + w(t).
$$

Here, $w(t)$ is the additive Gaussian noise in the baseband and is given by

$$
w(t) = \operatorname{LP}\left\{\frac{\sqrt{2}}{2} w_p(t) e^{-j2\pi f_c t}\right\}
$$

where $\operatorname{LP}\{\cdot\}$ stands for the operation of low-pass filtering.

However, as shown in Fig. 3.1(b), the actual oscillator signals in the I and Q branches are slightly different from the ideal oscillator signals due to the imperfectness of the local oscillator and 90° phase shifter. The constants α and θ model the amplitude and phase imbalances between the I and Q branches, while the phase noise term $\phi(t)$ models the phase difference between the carrier signal and the local oscillator. Based on this model, it can be shown that the received baseband signal $y(t)$ is related to the transmitted baseband signal $x(t)$ by (see Appendix C for a derivation):

$$y(t) = \left[\cos\left(\frac{\theta}{2}\right) - j\alpha\sin\left(\frac{\theta}{2}\right) \right] e^{j\phi(t)} y_0(t) \tag{3.6}$$
$$+ \left[\alpha\cos\left(\frac{\theta}{2}\right) + j\sin\left(\frac{\theta}{2}\right) \right] e^{-j\phi(t)} y_0^*(t) + w(t),$$

where $y_0(t)$ is given by (3.4). Letting

$$\mu = \cos\left(\frac{\theta}{2}\right) - j\alpha\sin\left(\frac{\theta}{2}\right),$$
$$\nu = \alpha\cos\left(\frac{\theta}{2}\right) + j\sin\left(\frac{\theta}{2}\right),$$

expression (3.6) can be represented as

$$y(t) = \mu e^{j\phi(t)} y_0(t) + \nu e^{-j\phi(t)} y_0^*(t) + w(t). \tag{3.7}$$

In the absence of IQ imbalance and phase noise, i.e., when $\alpha = 0$, $\theta = 0$, and $\phi(t) = 0$, we have the traditional relation

$$y(t) = y_0(t) + w(t).$$

3.2.2 OFDM Modulation and Demodulation

At the OFDM transmitter, the bits from information sources are first mapped into constellation symbols, and then converted into a block of N symbols $X[k]$,

$k = 0, 1, \ldots, N - 1$, by a serial-to-parallel converter. The N symbols are the frequency components to be transmitted using the N subcarriers of the OFDM modulator, and are converted to OFDM symbols $x[n]$, $n = 0, 1, \ldots, N - 1$, by the unitary inverse Fast Fourier Transform (IFFT), i.e.,

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi nk}{N}}, \quad n = 0, 1, \ldots, N - 1.$$

A cyclic prefix of length P is added to the IFFT output in order to eliminate the inter-symbol interference caused by multipath propagation. The resulting $N + P$ symbols are converted into a baseband signal $x(t)$ for transmission. Let T_s be the sampling time of the system. At the demodulator, the received baseband signal $y(t)$ is sampled at period T_s. After removing the cyclic prefix, a block of N symbols, $y[n]$, $n = 0, 1, \ldots, N - 1$, is obtained, whose elements are related to $x[n]$, $n = 0, 1, \ldots, N - 1$, through (3.7) by

$$y[n] = y(nT_s) = \mu e^{j\phi(nT_s)} y_0(nT_s) + \nu e^{-j\phi(nT_s)} y_0^*(nT_s) + w(nT_s). \tag{3.8}$$

In the continuous-time domain, $y_0(t)$ is equal to the convolutional integral of the channel impulse response and the channel input, as shown in expression (3.4). Let $h[n]$ represent the equivalent discrete-time baseband channel impulse response. Assume that $h[n]$ has length L, i.e., $h[n] = 0$ if $n \notin \{0, 1, \ldots, L - 1\}$. With the aid of the cyclic prefix and provided that $L - 1 \leq P$, linear convolution becomes circular convolution in the discrete-time domain (recall (2.1)), i.e.,

$$y_0(nT_s) = h[n] \circledast x[n] = \sum_{r=0}^{N-1} h[(n - r)_N] x[r],$$

where $\circledast$ denotes circular convolution and $(n - r)_N$ stands for

$$(n - r) \bmod N.$$

Expression (3.8) can then be written as

$$y[n] = \mu e^{j\phi(nT_s)} \left(h[n] \circledast x[n] \right) + \nu e^{-j\phi(nT_s)} \left(h[n] \circledast x[n] \right)^* + w[n],$$

66

where $w[n]$ is additive Gaussian noise. The unitary Fast Fourier Transform (FFT) is then performed on $y[n]$, $n = 0, 1, \ldots, N-1$, to obtain $Y[k]$, $k = 0, 1, \ldots, N-1$. Let

$$A[k] = \frac{1}{N} \sum_{n=0}^{N-1} e^{j\phi(nT_s)} e^{-j\frac{2\pi kn}{N}}, \quad k = 0, 1, \ldots, N-1,$$

and

$$H[k] = \sum_{n=0}^{L-1} h[n] e^{-j\frac{2\pi kn}{N}}, \quad k = 0, 1, \ldots, N-1.$$

Also, we denote

$$Z[k] = A[k] \circledast (H[k]X[k]) = \sum_{r=0}^{N-1} A[(k-r)_N] H[r] X[r]. \tag{3.9}$$

It then follows from (3.8) that the output symbols $Y[k]$, $k = 0, 1, \ldots, N-1$, after OFDM demodulation are related to the data symbols $X[k]$ by

$$Y[k] = \mu Z[k] + \nu Z^*[(N-k)_N] + W[k]$$

$$= \mu \sum_{r=0}^{N-1} A[(k-r)_N] H[r] X[r] + \nu \sum_{r=0}^{N-1} A^*[(N-k-r)_N] H^*[r] X^*[r]$$

$$+ W[k], \tag{3.10}$$

where $W[k]$ is the additive noise in the kth subcarrier and is given by the discrete Fourier transform of $w[n]$. Using matrix notation, (3.9) can be represented as

$$\mathbf{z} = \mathbf{AHx}, \tag{3.11}$$

where

$$\mathbf{z} = \begin{bmatrix} Z[0] \\ Z[1] \\ \vdots \\ Z[N-1] \end{bmatrix} (N \times 1), \quad \mathbf{x} = \begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[N-1] \end{bmatrix} (N \times 1),$$

67

$$\mathbf{A} = \begin{bmatrix} A[0] & A[N-1] & \ldots & A[1] \\ A[1] & A[0] & \ldots & A[2] \\ \vdots & \vdots & \ddots & \vdots \\ A[N-1] & A[N-2] & \ldots & A[0] \end{bmatrix} \quad (N \times N \text{ circulant}), \qquad (3.12)$$

$$\mathbf{H} = \begin{bmatrix} H[0] & 0 & \ldots & 0 \\ 0 & H[1] & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & H[N-1] \end{bmatrix} \quad (N \times N \text{ diagonal}).$$

Then, expression (3.10) can be represented as

$$\mathbf{y} = \mu \mathbf{z} + \nu \widetilde{\mathbf{z}} + \mathbf{w}, \qquad (3.13)$$

where

$$\mathbf{y} = \begin{bmatrix} Y[0] \\ Y[1] \\ \vdots \\ Y[N-1] \end{bmatrix} \quad (N \times 1), \qquad \widetilde{\mathbf{z}} = \begin{bmatrix} Z^*[0] \\ Z^*[N-1] \\ \vdots \\ Z^*[1] \end{bmatrix} \quad (N \times 1),$$

$$\mathbf{w} = \begin{bmatrix} W[0] \\ W[1] \\ \vdots \\ W[N-1] \end{bmatrix} \quad (N \times 1).$$

Combining (3.11) and (3.13), we have

$$\mathbf{y} = \mu \mathbf{A}\mathbf{H}\mathbf{x} + \nu \widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}, \qquad (3.14)$$

where

$$
\widetilde{\mathbf{A}} = \begin{bmatrix} A^*[0] & A^*[N-1] & \dots & A^*[1] \\ A^*[N-1] & A^*[N-2] & \dots & A^*[0] \\ \vdots & \vdots & \ddots & \vdots \\ A^*[1] & A^*[0] & \dots & A^*[2] \end{bmatrix} \quad (N \times N \text{ circulant}),
$$

$$
\text{conj}\{\mathbf{H}\} = \begin{bmatrix} H^*[0] & 0 & \dots & 0 \\ 0 & H^*[1] & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & H^*[N-1] \end{bmatrix} \quad (N \times N \text{ diagonal}),
$$

$$
\text{conj}\{\mathbf{x}\} = \begin{bmatrix} X^*[0] \\ X^*[1] \\ \vdots \\ X^*[N-1] \end{bmatrix} \quad (N \times 1).
$$

Expressions (3.10) and (3.14) model the effects of IQ imbalance and phase noise on the received symbols after OFDM demodulation. Based on this model, we now develop a compensation scheme and analyze the system performance with and without compensation.

3.3 Channel Estimation

In the proposed channel estimation algorithm, block-type pilot symbols are transmitted, in which all subcarriers are used for the pilot symbols known to the receiver. For convenience of exposition, we assume that at each time, only *one* OFDM symbol is used as the block-type pilot symbol for channel estimation. Since the OFDM demodulation output $\mathbf{y}$ is related to the training symbol $\mathbf{x}$

through expression (3.14), the proposed algorithm is based on the following optimization problem:

$$\min_{\mu,\nu,\mathbf{A},\mathbf{H}} \left\| \mathbf{y} - \mu\mathbf{A}\mathbf{H}\mathbf{x} - \nu\widetilde{\mathbf{A}} \cdot \mathrm{conj}\{\mathbf{H}\} \cdot \mathrm{conj}\{\mathbf{x}\} \right\|^2 . \tag{3.15}$$

We notice that there are N unknowns in $\mathbf{H}$, N unknowns in $\mathbf{A}$, plus two additional unknowns μ and ν. Thus, the solution to this problem is not unique, since we have less observations (in $\mathbf{y}$) than unknowns. To overcome this difficulty, we can reduce the number of unknowns by modeling the channel and the phase noise process with fewer parameters, as was already explained in Subsection 2.3.1. Moreover, we realize that in (3.15):

1.) There exists an ambiguity of a scaling factor in the estimates of μ, $\mathbf{A}$ and $\mathbf{H}$.

2.) There exists an ambiguity of a scaling factor in the estimates of ν, $\widetilde{\mathbf{A}}$ and $\mathrm{conj}\{\mathbf{H}\}$.

To resolve the ambiguities, instead of estimating μ, ν, $\mathbf{A}$ and $\mathbf{H}$, we estimate the following quantities:

$$\nu'' = \frac{\nu}{\mu}, \quad \mathbf{A}'' = \frac{1}{A[0]}\mathbf{A}, \quad \mathbf{H}'' = \mu A[0]\mathbf{H},$$

where

$$\mathbf{A}'' = \begin{bmatrix} A''[0] & A''[N-1] & \dots & A''[1] \\ A''[1] & A''[0] & \dots & A''[2] \\ \vdots & \vdots & \ddots & \vdots \\ A''[N-1] & A''[N-2] & \dots & A''[0] \end{bmatrix} (N \times N) \tag{3.16}$$

with

$$A''[0] = 1 \text{ and } A''[k] = \frac{1}{A[0]}A[k], \text{ for } k = 1, 2, \dots, N-1,$$

and

$$\mathbf{H}'' = \begin{bmatrix} H''[0] & 0 & \cdots & 0 \\ 0 & H''[1] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H''[N-1] \end{bmatrix} \quad (N \times N) \tag{3.17}$$

with

$$H''[k] = \mu A[0]H[k], \text{ for } k = 0, 1, \ldots, N-1.$$

Recall that in Subsection (2.3.1), we propose to estimate the *normalized* time-domain channel response $\mathbf{h}''$ of length L and the *selected normalized* phase noise samples $\mathbf{c}''$ of length M. Define

$$\overline{\mathbf{a}} = \begin{bmatrix} A''[0] \\ A''[1] \\ \vdots \\ A''[N-1] \end{bmatrix} (N \times 1), \quad \overline{\mathbf{h}} = \begin{bmatrix} H''[0] \\ H''[1] \\ \vdots \\ H''[N-1] \end{bmatrix} (N \times 1).$$

Then, $\overline{\mathbf{a}}$ and $\overline{\mathbf{h}}$ are related to $\mathbf{c}''$ and $\mathbf{h}''$ according to

$$\overline{\mathbf{a}} \approx \frac{1}{N} \mathbf{F_a P c}'' \tag{3.18}$$

and

$$\overline{\mathbf{h}} = \mathbf{F_h h}''. \tag{3.19}$$

Consequently, knowing $\mathbf{x}$ and $\mathbf{y}$, we can estimate $\mathbf{H}$ by solving

$$\min_{\nu'', \mathbf{c}'', \mathbf{h}''} \left\| \mathbf{y} - \mathbf{A}''\mathbf{H}''\mathbf{x} - \nu'' \widetilde{\mathbf{A}''} \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} \right\|^2 \tag{3.20}$$

$$\text{subject to } A''[0] = 1,$$

71

where

$$\widetilde{\mathbf{A}''} = \begin{bmatrix} (A''[0])^* & (A''[N-1])^* & \ldots & (A''[1])^* \\ (A''[N-1])^* & (A''[N-2])^* & \ldots & (A''[0])^* \\ \vdots & \vdots & \ddots & \vdots \\ (A''[1])^* & (A''[0])^* & \ldots & (A''[2])^* \end{bmatrix} \quad (N \times N \text{ circulant}),$$

(3.21)

and

$$\text{conj}\{\mathbf{H}''\} = \begin{bmatrix} (H''[0])^* & 0 & \ldots & 0 \\ 0 & (H''[1])^* & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & (H''[N-1])^* \end{bmatrix} \quad (N \times N \text{ diagonal}).$$

This optimization problem (3.20) is nonlinear and nonconvex. An iterative method for finding a sub-optimal solution is presented in Algorithm 3.1, in which we improve the estimates of ν'', $\mathbf{c}''$ and $\mathbf{h}''$ recursively by allowing small perturbations in them and then finding the optimal values for these perturbation terms. It can also be viewed as local linearization of a nonlinear system by using a first-order approximation. Since the amplitude of IQ imbalances and phase noise is usually small, the true values of ν'' and $\mathbf{c}''$ are close to their nominal values. The nominal values of ν'' and $\mathbf{c}''$ assume the absence of the impairments, i.e.,

$$\alpha = 0, \ \theta = 0, \ \phi(t) = 0,$$

which gives

$$\mu = 1, \ \nu = 0, \ \mathbf{A} = \mathbf{I}_N.$$

Hence, the nominal values of ν'' and $\mathbf{c}''$ are

$$\nu'' = 0, \ \mathbf{c}'' = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^T.$$

72

Step 0 of Algorithm 3.1 finds an initial estimate for ν'', $\mathbf{c}''$ and $\mathbf{h}''$ to start the iteration. Specifically, ν'' and $\mathbf{c}''$ are initialized to their nominal values, and $\mathbf{h}''$ is initialized to the corresponding solution to (3.20). That is, when

$$\nu'' = 0, \ \mathbf{c}'' = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T,$$

we have

$$\mathbf{A}'' = \mathbf{I}_N$$

and hence (3.20) becomes

$$\min_{\mathbf{h}''} \left\| \mathbf{y} - \mathbf{H}''\mathbf{x} \right\|^2. \tag{3.23}$$

Let $\mathbf{X} = \text{diag}\{\mathbf{x}\}$. By (3.19), problem (3.23) can be rewritten as

$$\min_{\mathbf{h}''} \left\| \mathbf{y} - \mathbf{X}\mathbf{F}_{\mathbf{h}}\mathbf{h}'' \right\|^2$$

whose solution is given by

$$\widehat{\mathbf{h}}_0'' = (\mathbf{F}_{\mathbf{h}}^*\mathbf{X}^*\mathbf{X}\mathbf{F}_{\mathbf{h}})^{-1}\mathbf{F}_{\mathbf{h}}^*\mathbf{X}^*\mathbf{y}.$$

Step 4 and Step 5 update the old estimates $\widehat{\nu}''_{i-1}$, $\widehat{\mathbf{c}}''_{i-1}$ and $\widehat{\mathbf{h}}''_{i-1}$ by

$$\widehat{\nu}''_i = \widehat{\nu}''_{i-1} + \Delta\nu'', \tag{3.24}$$

$$\widehat{\mathbf{c}}''_i = \widehat{\mathbf{c}}''_{i-1} + \Delta\mathbf{c}'', \tag{3.25}$$

$$\widehat{\mathbf{h}}''_i = \widehat{\mathbf{h}}''_{i-1} + \Delta\mathbf{h}''. \tag{3.26}$$

Expressions (3.25) and (3.26) imply that $\mathbf{A}''_{i-1}$, $\widetilde{\mathbf{A}''}_{i-1}$ and $\mathbf{H}''_{i-1}$ are updated according to

$$\mathbf{A}''_i = \mathbf{A}''_{i-1} + \Delta\mathbf{A}'', \tag{3.27}$$

$$\widetilde{\mathbf{A}''}_i = \widetilde{\mathbf{A}''}_{i-1} + \Delta\widetilde{\mathbf{A}''}, \tag{3.28}$$

$$\mathbf{H}''_i = \mathbf{H}''_{i-1} + \Delta\mathbf{H}'', \tag{3.29}$$

73

Algorithm 3.1 Joint Channel Estimation

0: Let $\widehat{\nu}_0'' = 0$ and $\widehat{\mathbf{c}}_0'' = [1 \ 1 \ \ldots \ 1]^T$. Find the initial $\widehat{\mathbf{h}}_0''$ by solving

$$\widehat{\mathbf{h}}_0'' = \arg\min_{\mathbf{h}''} \| \mathbf{y} - \mathbf{H}''\mathbf{x} \|^2.$$

The expression for $\widehat{\mathbf{h}}_0''$ is given by

$$\widehat{\mathbf{h}}_0'' = (\mathbf{F}_\mathbf{h}^* \mathbf{X}^* \mathbf{X} \mathbf{F}_\mathbf{h})^{-1} \mathbf{F}_\mathbf{h}^* \mathbf{X}^* \mathbf{y},$$

where $\mathbf{X} = \text{diag}\{\mathbf{x}\}$.

1: $i = 1$

2: **repeat**

3: Let $\overline{\mathbf{a}}_{i-1} = \frac{1}{N}\mathbf{F}_\mathbf{a}\mathbf{P}\widehat{\mathbf{c}}_{i-1}''$ and $\overline{\mathbf{h}}_{i-1} = \mathbf{F}_\mathbf{h}\widehat{\mathbf{h}}_{i-1}''$.

4: Find $\Delta\nu_i''$, $\Delta\mathbf{c}_i''$ and $\Delta\mathbf{h}_i''$ by solving the following optimization problem:

$$
\min_{\Delta\nu'', \Delta\mathbf{c}'', \Delta\mathbf{h}''} \Big\| \mathbf{y} - \big(\mathbf{A}_{i-1}'' \mathbf{H}_{i-1}'' \mathbf{x} + \widehat{\nu}_{i-1}'' \widetilde{\mathbf{A}''}_{i-1} \cdot \text{conj}\{\mathbf{H}_{i-1}''\} \cdot \text{conj}\{\mathbf{x}\}\big) \tag{3.22}
$$

$$
- \big[(\Delta\mathbf{A}'')\mathbf{H}_{i-1}''\mathbf{x} + \mathbf{A}_{i-1}''(\Delta\mathbf{H}'')\mathbf{x}\big] - \big[(\Delta\nu'')\widetilde{\mathbf{A}''}_{i-1} \cdot \text{conj}\{\mathbf{H}_{i-1}''\} \cdot \text{conj}\{\mathbf{x}\}
$$

$$
+ \widehat{\nu}_{i-1}''(\Delta\widetilde{\mathbf{A}''}) \cdot \text{conj}\{\mathbf{H}_{i-1}''\} \cdot \text{conj}\{\mathbf{x}\} + \widehat{\nu}_{i-1}'' \widetilde{\mathbf{A}''}_{i-1} \cdot \text{conj}\{\Delta\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\}\big] \Big\|^2
$$

$$
\text{subject to } \mathbf{g}\Delta\mathbf{c}'' = 0,
$$

where $\mathbf{g}$ is the first row of $\frac{1}{N}\mathbf{F}_\mathbf{a}\mathbf{P}$, and $\mathbf{A}_{i-1}''$, $\widetilde{\mathbf{A}''}_{i-1}$, $\mathbf{H}_{i-1}''$ are determined from $\overline{\mathbf{a}}_{i-1}$ and $\overline{\mathbf{h}}_{i-1}$ according to (3.16), (3.21) and (3.17). The constraint $\mathbf{g}\Delta\mathbf{c}'' = 0$ guarantees that $A''[0]$ is equal to 1. Problem (3.22) can be formulated as a standard least-squares problem (see Appendix D for details).

5: Update the estimates of ν'', $\mathbf{c}''$ and $\mathbf{h}''$ according to

$$\widehat{\nu}_i'' = \widehat{\nu}_{i-1}'' + \Delta\nu_i'', \quad \widehat{\mathbf{c}}_i'' = \widehat{\mathbf{c}}_{i-1}'' + \Delta\mathbf{c}_i'', \quad \widehat{\mathbf{h}}_i'' = \widehat{\mathbf{h}}_{i-1}'' + \Delta\mathbf{h}_i''.$$

6: $i = i + 1$

7: **until** there is no significant improvement in the objective function $\| \mathbf{y} - \mathbf{A}_i''\mathbf{H}_i''\mathbf{x} - \widehat{\nu}_i''\widetilde{\mathbf{A}''}_i \cdot \text{conj}\{\mathbf{H}_i''\} \cdot \text{conj}\{\mathbf{x}\} \|^2$.

where $\Delta \mathbf{A}''$, $\Delta \widetilde{\mathbf{A}''}$ and $\Delta \mathbf{H}''$ depend on $\Delta \mathbf{c}''$ and $\Delta \mathbf{h}''$. By substitution with (3.24) and (3.27)-(3.29), the objective function becomes

$$\left\| \mathbf{y} - \mathbf{A}''_i \mathbf{H}''_i \mathbf{x} - \widehat{\nu}''_i \widetilde{\mathbf{A}''}_i \cdot \text{conj}\{\mathbf{H}''_i\} \cdot \text{conj}\{\mathbf{x}\} \right\|^2$$

$$= \left\| \mathbf{y} - \left(\mathbf{A}''_{i-1} + \Delta \mathbf{A}''\right)\left(\mathbf{H}''_{i-1} + \Delta \mathbf{H}''\right)\mathbf{x} - \left(\widehat{\nu}''_{i-1} + \Delta \nu''\right)\left(\widetilde{\mathbf{A}''}_{i-1} + \Delta \widetilde{\mathbf{A}''}\right) \right.$$

$$\left. \cdot \text{conj}\{\mathbf{H}''_{i-1} + \Delta \mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} \right\|^2$$

$$\approx \left\| \mathbf{y} - \left(\mathbf{A}''_{i-1}\mathbf{H}''_{i-1}\mathbf{x} + \widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \text{conj}\{\mathbf{H}''_{i-1}\} \cdot \text{conj}\{\mathbf{x}\}\right) \right. \tag{3.30}$$

$$- \left[(\Delta \mathbf{A}'')\mathbf{H}''_{i-1}\mathbf{x} + \mathbf{A}''_{i-1}(\Delta \mathbf{H}'')\mathbf{x}\right] - \left[(\Delta \nu'')\widetilde{\mathbf{A}''}_{i-1} \cdot \text{conj}\{\mathbf{H}''_{i-1}\} \cdot \text{conj}\{\mathbf{x}\} \right.$$

$$\left. + \widehat{\nu}''_{i-1}(\Delta \widetilde{\mathbf{A}''}) \cdot \text{conj}\{\mathbf{H}''_{i-1}\} \cdot \text{conj}\{\mathbf{x}\} + \widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \text{conj}\{\Delta \mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\}\right] \left. \right\|^2,$$

where the approximation ignores the second-order and third-order small terms. The near-optimal incremental terms $\Delta \nu''$, $\Delta \mathbf{c}''$ and $\Delta \mathbf{h}''$ are found by minimizing (3.30). To ensure that $A''[0] = 1$ in the updating, given that

$$A''_{i-1}[0] = \mathbf{g}\widehat{\mathbf{c}}''_{i-1} = 1$$

where $\mathbf{g}$ is the first row of $\frac{1}{N}\mathbf{F_a}\mathbf{P}$, we have

$$A''_i[0] = \mathbf{g}\widehat{\mathbf{c}}''_i = \mathbf{g}\left(\widehat{\mathbf{c}}''_{i-1} + \Delta \mathbf{c}''\right) = 1$$

and hence

$$\mathbf{g}\Delta \mathbf{c}'' = 0.$$

It is shown by computer simulations that the objective function decreases with $i = 1, 2, \ldots$, and eventually converges to a local minimum. The obtained estimates of ν'' and $\mathbf{H}''$ will be used in the data transmission stage for estimating data symbols.

3.3.1 Cramer-Rao Lower Bound

To evaluate the proposed algorithm, we compare its performance with the Cramer-Rao lower bound (CRLB) that gives a lower bound on the covariance

matrix of any unbiased estimator of unknown parameters [Kay93]. Three scenarios are considered here:

1.) There is no impairment in the system.

2.) There are IQ imbalance and phase noise, but the receiver assumes no physical impairment.

3.) Both IQ imbalance and phase noise are present, and the proposed channel estimation algorithm is used.

The scenarios can be either exactly or approximately modeled by

$$\mathbf{y} = \mathbf{s}_{\theta} + \mathbf{w}, \tag{3.31}$$

where $\mathbf{y}$ is the observed data vector of length N, $\mathbf{s}_{\theta}$ is the noise-free data vector that depends on the parameter vector $\boldsymbol{\theta}$, and $\mathbf{w}$ is the vector of circularly symmetric Gaussian noise with covariance matrix $\mathbf{E}\{\mathbf{w}\mathbf{w}^*\} = \sigma_W^2 \mathbf{I}_N$.

By the CRLB [Kay93], any unbiased estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ has a covariance matrix that satisfies

$$\text{var}\{\widehat{\boldsymbol{\theta}}\} = \mathbf{E}\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^*\} \geq \mathbf{I}_{\theta}^{-1},$$

where $\text{var}\{\widehat{\boldsymbol{\theta}}\} \geq \mathbf{I}_{\theta}^{-1}$ is interpreted as meaning that the matrix $\text{var}\{\widehat{\boldsymbol{\theta}}\} - \mathbf{I}_{\theta}^{-1}$ is positive semidefinite. Here, $\mathbf{I}_{\theta}$ is the Fisher information matrix of size $|\boldsymbol{\theta}| \times |\boldsymbol{\theta}|$ and its element at the i_1^{th} row and i_2^{th} column is given by

$$[\mathbf{I}_{\theta}]_{i_1, i_2} = -\mathbf{E}\left\{\frac{\partial^2 \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_{i_1} \partial \theta_{i_2}}\right\}, \quad 1 \leq i_1 \leq |\boldsymbol{\theta}|, \ 1 \leq i_2 \leq |\boldsymbol{\theta}|,$$

where $p(\mathbf{y}; \boldsymbol{\theta})$ is the probability density function of the observations and $|\boldsymbol{\theta}|$ is the dimension of $\boldsymbol{\theta}$. Let

$$\mathbf{s}_{\theta} = \begin{bmatrix} S_{\theta}[0] & S_{\theta}[1] & \dots & S_{\theta}[N-1] \end{bmatrix}^T.$$

It is shown in Appendix E that the Fisher information matrix for the data model (3.31) is given by

$$\mathbf{I}_\theta = \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \text{Re}\left\{ \frac{\partial S_\theta[k]}{\partial\boldsymbol{\theta}} \left(\frac{\partial S_\theta[k]}{\partial\boldsymbol{\theta}}\right)^* \right\}, \tag{3.32}$$

where

$$\frac{\partial S_\theta[k]}{\partial\boldsymbol{\theta}} = \left[\frac{\partial S_\theta[k]}{\partial\theta_1} \quad \frac{\partial S_\theta[k]}{\partial\theta_2} \quad \cdots \quad \frac{\partial S_\theta[k]}{\partial\theta_{|\boldsymbol{\theta}|}} \right]^T.$$

In the following derivation, we assume:

1.) The pilot symbols $X[k]$ are independent and identically distributed with zero mean. They also have the same power and let $\sigma_P^2 = |X[k]|^2$.

2.) The pilot symbols, the phase noise, the channel coefficients and the additive noise are independent of each other.

3.) The channel coefficients $H[k]$ are circularly symmetric Gaussian distributed with mean zero and variance $\sigma_H^2 = \mathbf{E}\{|H[k]|^2\}$.

3.3.1.1   Scenario 1: No Impairments

In this scenario, there is no analog impairment in the system. We consider two cases: one estimates $\mathbf{h}$ and the other estimates $\mathbf{h}'$. Recall that in Subsection 2.3.1, $\mathbf{h}$ and $\mathbf{h}'$ are defined as

$$\mathbf{h} = \begin{bmatrix} H[0] \\ H[1] \\ \vdots \\ H[N-1] \end{bmatrix} (N \times 1), \quad \mathbf{h}' = \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[L-1] \end{bmatrix} (L \times 1).$$

Case 1: $\mathbf{h}$ is estimated

If **h** is estimated, the system is modeled as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \qquad (3.33)$$

where

$$\mathbf{s}_\theta = \mathbf{H}\mathbf{x}$$

and

$$\boldsymbol{\theta} = \left[\ \mathrm{Re}\{H[0]\} \ \ \ldots \ \ \mathrm{Re}\{H[N-1]\} \ \ \mathrm{Im}\{H[0]\} \ \ \ldots \ \ \mathrm{Im}\{H[N-1]\} \ \right]^T.$$

To compute $\mathbf{I}_\theta$, we need to compute $\frac{\partial S_\theta[k]}{\partial \theta}$ according to (3.32). Since

$$S_\theta[k] = H[k]X[k],$$

we have

$$\frac{\partial S_\theta[k]}{\partial \mathrm{Re}\{H[k']\}} = \begin{cases} X[k], & \text{if } k = k', \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\frac{\partial S_\theta[k]}{\partial \mathrm{Im}\{H[k']\}} = \begin{cases} jX[k], & \text{if } k = k', \\ 0, & \text{otherwise}. \end{cases}$$

Hence,

$$\frac{\partial S_\theta[k]}{\partial \boldsymbol{\theta}} = \left[\ 0 \ \ \ldots \ \ 0 \ \ X[k] \ \ 0 \ \ \ldots \ \ 0 \ \ jX[k] \ \ 0 \ \ \ldots \ \ 0 \ \right]^T$$

$$= X[k]\mathbf{e}_{k+1} + jX[k]\mathbf{e}_{N+k+1}.$$

Here, $\mathbf{e}_l$ represents the lth standard basis vector whose elements are all zeros except that the lth element is 1. By (3.32),

$$\mathbf{I}_\theta = \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \mathrm{Re}\left\{ \frac{\partial S_\theta[k]}{\partial \boldsymbol{\theta}} \left(\frac{\partial S_\theta[k]}{\partial \boldsymbol{\theta}} \right)^* \right\}$$

$$= \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \mathrm{Re}\left\{ (X[k]\mathbf{e}_{k+1} + jX[k]\mathbf{e}_{N+k+1}) (X[k]\mathbf{e}_{k+1} + jX[k]\mathbf{e}_{N+k+1})^* \right\}$$

$$= \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \text{Re} \left\{ |X[k]|^2 \mathbf{e}_{k+1} \mathbf{e}_{k+1}^T - j|X[k]|^2 \mathbf{e}_{k+1} \mathbf{e}_{N+k+1}^T + j|X[k]|^2 \mathbf{e}_{N+k+1} \mathbf{e}_{k+1}^T \right.$$

$$\left. + |X[k]|^2 \mathbf{e}_{N+k+1} \mathbf{e}_{N+k+1}^T \right\}$$

$$= \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \left[|X[k]|^2 \mathbf{e}_{k+1} \mathbf{e}_{k+1}^T + |X[k]|^2 \mathbf{e}_{N+k+1} \mathbf{e}_{N+k+1}^T \right]$$

$$= \frac{2\sigma_P^2}{\sigma_W^2} \mathbf{I}_{2N}.$$

By the CRLB,

$$\mathbf{E}\{ (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^* \} \geq \mathbf{I}_{\boldsymbol{\theta}}^{-1} = \frac{\sigma_W^2}{2\sigma_P^2} \mathbf{I}_{2N},$$

which implies

$$\mathbf{E} \left\{ \left| \text{Re}\{\widehat{H}[k]\} - \text{Re}\{H[k]\} \right|^2 \right\} \geq \frac{\sigma_W^2}{2\sigma_P^2}$$

and

$$\mathbf{E} \left\{ \left| \text{Im}\{\widehat{H}[k]\} - \text{Im}\{H[k]\} \right|^2 \right\} \geq \frac{\sigma_W^2}{2\sigma_P^2}.$$

Therefore, the mean-square-error of $\widehat{H}[k]$ is bounded by

$$\mathbf{E} \left\{ \left| \widehat{H}[k] - H[k] \right|^2 \right\} = \mathbf{E} \left\{ \left| \text{Re}\{\widehat{H}[k]\} - \text{Re}\{H[k]\} \right|^2 \right\}$$

$$+ \mathbf{E} \left\{ \left| \text{Im}\{\widehat{H}[k]\} - \text{Im}\{H[k]\} \right|^2 \right\}$$

$$\geq \frac{\sigma_W^2}{\sigma_P^2}. \tag{3.34}$$

Case 2: $\mathbf{h}'$ is estimated

If $\mathbf{h}'$ is estimated instead of $\mathbf{h}$, the system model becomes

$$\mathbf{y} = \mathbf{X}\mathbf{F_h}\mathbf{h}' + \mathbf{w},$$

where $\mathbf{X} = \text{diag}\{\mathbf{x}\}$. In this case,

$$\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{X}\mathbf{F_h}\mathbf{h}'$$

and

$$\boldsymbol{\theta} = \left[\begin{array}{ccccccc} \mathrm{Re}\{h[0]\} & \ldots & \mathrm{Re}\{h[L-1]\} & \mathrm{Im}\{h[0]\} & \ldots & \mathrm{Im}\{h[L-1]\} \end{array} \right]^T .$$

For $n = 0, 1, \ldots, L-1$,

$$\frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \mathrm{Re}\{h[n]\}} = \mathbf{X}\mathbf{F}_{\mathbf{h}} \cdot \frac{\partial \mathbf{h}'}{\partial \mathrm{Re}\{h[n]\}} = \mathbf{X}\mathbf{F}_{\mathbf{h}}\mathbf{e}_{n+1},$$

$$\frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \mathrm{Im}\{h[n]\}} = \mathbf{X}\mathbf{F}_{\mathbf{h}} \cdot \frac{\partial \mathbf{h}'}{\partial \mathrm{Im}\{h[n]\}} = j\mathbf{X}\mathbf{F}_{\mathbf{h}}\mathbf{e}_{n+1}.$$

It then follows that

$$\frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}} = \left[\begin{array}{ccccccc} \frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \mathrm{Re}\{h[0]\}} & \cdots & \frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \mathrm{Re}\{h[L-1]\}} & \frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \mathrm{Im}\{h[0]\}} & \cdots & \frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \mathrm{Im}\{h[L-1]\}} \end{array} \right]$$

$$= \left[\begin{array}{cc} \mathbf{X}\mathbf{F}_{\mathbf{h}} & j\mathbf{X}\mathbf{F}_{\mathbf{h}} \end{array} \right]$$

and

$$\mathbf{I}_{\boldsymbol{\theta}} = \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \mathrm{Re} \left\{ \frac{\partial S_{\boldsymbol{\theta}}[k]}{\partial \boldsymbol{\theta}} \left(\frac{\partial S_{\boldsymbol{\theta}}[k]}{\partial \boldsymbol{\theta}} \right)^* \right\}$$

$$= \frac{2}{\sigma_W^2} \mathrm{Re} \left\{ \left(\frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}} \right)^T \left[\left(\frac{\partial \mathbf{s}_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}} \right)^T \right]^* \right\}$$

$$= \frac{2}{\sigma_W^2} \mathrm{Re} \left\{ \left[\begin{array}{cc} \mathbf{X}\mathbf{F}_{\mathbf{h}} & j\mathbf{X}\mathbf{F}_{\mathbf{h}} \end{array} \right]^T \left(\left[\begin{array}{cc} \mathbf{X}\mathbf{F}_{\mathbf{h}} & j\mathbf{X}\mathbf{F}_{\mathbf{h}} \end{array} \right]^T \right)^* \right\}$$

$$= \frac{2}{\sigma_W^2} \mathrm{Re} \left\{ \left[\begin{array}{c} \mathbf{F}_{\mathbf{h}}^T \mathbf{X}^T \\ j\mathbf{F}_{\mathbf{h}}^T \mathbf{X}^T \end{array} \right] \left[\begin{array}{cc} \mathrm{conj}\{\mathbf{X}\mathbf{F}_{\mathbf{h}}\} & \mathrm{conj}\{j\mathbf{X}\mathbf{F}_{\mathbf{h}}\} \end{array} \right] \right\}$$

$$= \frac{2}{\sigma_W^2} \mathrm{Re} \left\{ \left[\begin{array}{cc} N\sigma_P^2 \mathbf{I}_L & -jN\sigma_P^2 \mathbf{I}_L \\ jN\sigma_P^2 \mathbf{I}_L & N\sigma_P^2 \mathbf{I}_L \end{array} \right] \right\}$$

$$= \frac{2N\sigma_P^2}{\sigma_W^2} \mathbf{I}_{2L}.$$

By the CRLB,

$$\mathrm{E}\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^*\} \geq \mathbf{I}_{\boldsymbol{\theta}}^{-1} = \frac{\sigma_W^2}{2N\sigma_P^2} \mathbf{I}_{2L},$$

which implies

$$\mathbf{E}\left\{\left|\mathrm{Re}\{\widehat{h}[n]\} - \mathrm{Re}\left\{h[n]\right\}\right|^2\right\} \geq \frac{\sigma_W^2}{2N\sigma_P^2}$$

and

$$\mathbf{E}\left\{\left|\mathrm{Im}\{\widehat{h}[n]\} - \mathrm{Im}\left\{h[n]\right\}\right|^2\right\} \geq \frac{\sigma_W^2}{2N\sigma_P^2}.$$

Hence, the mean-square-error of $\widehat{h}[n]$ is bounded by

$$\mathbf{E}\{|\widehat{h}[n] - h[n]|^2\} \geq \frac{\sigma_W^2}{N\sigma_P^2}.$$

Recall that

$$\mathbf{h} = \mathbf{F_h h'} \quad \text{and} \quad \widehat{\mathbf{h}} = \mathbf{F_h}\widehat{\mathbf{h}}'.$$

Since

$$\|\mathbf{h} - \widehat{\mathbf{h}}\|^2 = \|\mathbf{F_h h'} - \mathbf{F_h}\widehat{\mathbf{h}}'\|^2 = N\|\mathbf{h'} - \widehat{\mathbf{h}}'\|^2,$$

the mean-square-error of $\widehat{H}[k]$ is bounded by

$$
\begin{aligned}
\mathbf{E}\{|\widehat{H}[k] - H[k]|^2\} &= \frac{1}{N}\mathbf{E}\left\{\|\mathbf{h} - \widehat{\mathbf{h}}\|^2\right\} \\
&= \mathbf{E}\left\{\|\mathbf{h'} - \widehat{\mathbf{h}}'\|^2\right\} \\
&= L\mathbf{E}\{|\widehat{h}[n] - h[n]|^2\} \\
&\geq \frac{L\sigma_W^2}{N\sigma_P^2}.
\end{aligned}
\tag{3.35}
$$

By comparing (3.35) with (3.34), we can see that estimating $\mathbf{h'}$ instead of $\mathbf{h}$ improves the mean-square-error performance by a factor L/N ($L < N$). This is because given the N observations in $\mathbf{y}$, it is more accurate to estimate L variables instead of N variables.

3.3.1.2 Scenario 2: Without Compensation

In the presence of the impairments, the system model (3.14) can be rewritten as

$$\mathbf{y} = \mu\mathbf{A}\mathbf{H}\mathbf{x} + \nu\widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}$$

$$= \mu A[0]\mathbf{H}\mathbf{x} + \mu(\mathbf{A} - A[0]\mathbf{I}_N)\mathbf{H}\mathbf{x} + \nu\widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}.$$

We treat

$$\mathbf{H}'' = \mu A[0]\mathbf{H}$$

as the "true" channel response to be estimated because of the scalar ambiguity. The term

$$\mathbf{w}'' = \mu(\mathbf{A} - A[0]\mathbf{I}_N)\mathbf{H}\mathbf{x} + \nu\widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}$$

can be *approximately* regarded as additive white Gaussian noise with covariance matrix

$$\mathbf{E}\{\mathbf{w}''(\mathbf{w}'')^*\} = \mathbf{E}\left\{ \left[\mu(\mathbf{A} - A[0]\mathbf{I}_N)\mathbf{H}\mathbf{x} + \nu\widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}\right] \right.$$

$$\left. \cdot \left[\mu(\mathbf{A} - A[0]\mathbf{I}_N)\mathbf{H}\mathbf{x} + \nu\widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}\right]^* \right\}$$

$$= \mathbf{E}\left\{ [\mu(\mathbf{A} - A[0]\mathbf{I}_N)\mathbf{H}\mathbf{x}][\mu(\mathbf{A} - A[0]\mathbf{I}_N)\mathbf{H}\mathbf{x}]^* \right\}$$

$$+ \mathbf{E}\left\{ \left[\nu\widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\}\right] \left[\nu\widetilde{\mathbf{A}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\}\right]^* \right\}$$

$$+ \mathbf{E}\{\mathbf{w}\mathbf{w}^*\}$$

$$= \mathbf{E}\left\{ [\mu(\mathbf{A} - A[0]\mathbf{I}_N)][\mu(\mathbf{A} - A[0]\mathbf{I}_N)]^* \right\} \sigma_H^2 \sigma_P^2$$

$$+ \mathbf{E}\left\{ \left(\nu\widetilde{\mathbf{A}}\right) \left(\nu\widetilde{\mathbf{A}}\right)^* \right\} \sigma_H^2 \sigma_P^2 + \sigma_W^2 \mathbf{I}_N$$

$$\approx \left\{ \left[\left(\sum_{k=1}^{N-1} \sigma_{A,k}^2\right) |\mu|^2 + \left(\sum_{k=0}^{N-1} \sigma_{A,k}^2\right) |\nu|^2 \right] \sigma_H^2 \sigma_P^2 + \sigma_W^2 \right\} \cdot \mathbf{I}_N$$

$$= \left\{ \left[(1 - \sigma_{A,0}^2)|\mu|^2 + |\nu|^2 \right] \sigma_H^2 \sigma_P^2 + \sigma_W^2 \right\} \cdot \mathbf{I}_N$$

where the approximation assumes that

$$\mathbf{E}\left\{ (\mathbf{A} - A[0]\mathbf{I}_N)(\mathbf{A} - A[0]\mathbf{I}_N)^* \right\} \approx \left(\sum_{k=1}^{N-1} \sigma_{A,k}^2\right) \mathbf{I}_N$$

and

$$\mathbf{E}\left\{\widetilde{\mathbf{A}}\widetilde{\mathbf{A}}^{*}\right\} \approx \left(\sum_{k=0}^{N-1} \sigma_{A,k}^{2}\right) \mathbf{I}_{N}.$$

Note that the model

$$\mathbf{y} = \mathbf{H}''\mathbf{x} + \mathbf{w}''$$

with

$$\sigma_{W''}^{2} = \left[\left(1 - \sigma_{A,0}^{2}\right)|\mu|^{2} + |\nu|^{2}\right]\sigma_{H}^{2}\sigma_{P}^{2} + \sigma_{W}^{2}$$

has the same form as (3.33). It then follows immediately from (3.34) and (3.35) that

1.) if $\mathbf{h}$ is estimated,

$$\mathbf{E}\left\{\left|\mu A[0]\widehat{H}[k] - \mu A[0]H[k]\right|^{2}\right\} \geq \frac{\sigma_{W''}^{2}}{\sigma_{P}^{2}}$$

$$= \left[\left(1 - \sigma_{A,0}^{2}\right)|\mu|^{2} + |\nu|^{2}\right]\sigma_{H}^{2} + \frac{\sigma_{W}^{2}}{\sigma_{P}^{2}}; \qquad (3.36)$$

2.) if $\mathbf{h}'$ is estimated,

$$\mathbf{E}\left\{\left|\mu A[0]\widehat{H}[k] - \mu A[0]H[k]\right|^{2}\right\} \geq \frac{L\sigma_{W''}^{2}}{N\sigma_{P}^{2}}$$

$$= \frac{L}{N}\left[\left(1 - \sigma_{A,0}^{2}\right)|\mu|^{2} + |\nu|^{2}\right]\sigma_{H}^{2} + \frac{L\sigma_{W}^{2}}{N\sigma_{P}^{2}}. \qquad (3.37)$$

3.3.1.3 Scenario 3: With the Proposed Joint Estimation

In this case, the CRLB for estimating $\mathbf{H}$ is computed based on the following model:

$$\mathbf{y} = \mathbf{A}''\mathbf{H}''\mathbf{x} + \nu''\widetilde{\mathbf{A}''} \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}$$

$$= \mathbf{A}''_{\text{appro}}\mathbf{H}''\mathbf{x} + \nu''\widetilde{\mathbf{A}''}_{\text{appro}} \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} + (\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\mathbf{H}''\mathbf{x} \quad (3.38)$$

$$+ \nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w},$$

where $\mathbf{A}''_{\text{appro}}$ is determined by the vector

$$\overline{\mathbf{a}}_{\text{appro}} = \frac{1}{N}\mathbf{F_a}\mathbf{P}\mathbf{c}''$$

according to the construction of $\mathbf{A}''$. Note that $\mathbf{A}'' - \mathbf{A}''_{\text{appro}}$ represents the modeling error existing in the approximation given by (3.18). The parameter vector to be estimated is

$$\boldsymbol{\theta} = \begin{bmatrix} \text{Re}\{\nu''\} & \text{Im}\{\nu''\} & \text{Re}\{\mathbf{c}'''^T\} & \text{Im}\{\mathbf{c}'''^T\} & \text{Re}\{\mathbf{h}''^T\} & \text{Im}\{\mathbf{h}''^T\} \end{bmatrix}^T,$$

where the vector

$$\mathbf{c}''' = \begin{bmatrix} c''[1] & c''[2] & \ldots & c''[M-1] \end{bmatrix}^T$$

contains all elements of $\mathbf{c}''$ except its first element $c''[0]$. Note that $c''[0]$ is determined by $c''[m]$, $m = 1, 2, \ldots, M-1$, because of the constraint $A''[0] = 1$. Hence,

$$\mathbf{s}_{\boldsymbol{\theta}} = \mathbf{A}''_{\text{appro}}\mathbf{H}''\mathbf{x} + \nu''\widetilde{\mathbf{A}''}_{\text{appro}} \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\},$$

and the noise term in expression (3.38) is

$$\mathbf{w}'' = (\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\mathbf{H}''\mathbf{x} + \nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}.$$

The covariance matrix of $\mathbf{w}''$ is computed as

$$
\begin{aligned}
&\mathbf{E}\{\mathbf{w}''(\mathbf{w}'')^*\} \\
&= \mathbf{E}\left\{\left[(\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\mathbf{H}''\mathbf{x} + \nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}\right]\right. \\
&\quad \left. \cdot \left[(\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\mathbf{H}''\mathbf{x} + \nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}\right]^*\right\} \\
&= \mathbf{E}\left\{\left[(\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\mathbf{H}''\mathbf{x}\right]\left[(\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\mathbf{H}''\mathbf{x}\right]^*\right\} \\
&\quad + \mathbf{E}\left\{\left[\nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\}\right]\right. \\
&\quad \left. \cdot \left[\nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\} \cdot \text{conj}\{\mathbf{x}\}\right]^*\right\} + \mathbf{E}\{\mathbf{w}\mathbf{w}^*\} \\
&= \mathbf{E}\left\{\left((\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\,\mathbf{H}''\right)\left((\mathbf{A}'' - \mathbf{A}''_{\text{appro}})\,\mathbf{H}''\right)^*\right\} \cdot \sigma_P^2 \\
&\quad + \mathbf{E}\left\{\left(\nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\}\right)\left(\nu''(\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}''\}\right)^*\right\} \cdot \sigma_P^2
\end{aligned}
$$

$$+ \sigma_W^2 \mathbf{I}_N$$

$$\approx \mathbf{E}\{\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\mathrm{appro}}\|^2 |H''[k]|^2\} \cdot \sigma_P^2 \mathbf{I}_N + |\nu''|^2 \cdot \mathbf{E}\{\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\mathrm{appro}}\|^2 |H''[k]|^2\} \cdot \sigma_P^2 \mathbf{I}_N$$

$$+ \sigma_W^2 \mathbf{I}_N$$

$$= \left\{ \left(1 + |\nu''|^2\right) \cdot \mathbf{E}\{\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\mathrm{appro}}\|^2 |H''[k]|^2\} \cdot \sigma_P^2 + \sigma_W^2 \right\} \cdot \mathbf{I}_N$$

where the approximation assumes that the modeling error components are independently and identically distributed with zero mean and hence

$$\mathbf{E}\left\{ \left(\left(\mathbf{A}'' - \mathbf{A}''_{\mathrm{appro}}\right)\mathbf{H}''\right)\left(\left(\mathbf{A}'' - \mathbf{A}''_{\mathrm{appro}}\right)\mathbf{H}''\right)^*\right\} \approx \mathbf{E}\{\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\mathrm{appro}}\|^2 |H''[k]|^2\} \cdot \mathbf{I}_N,$$

and

$$\mathbf{E}\left\{ \left((\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\mathrm{appro}}) \cdot \mathrm{conj}\{\mathbf{H}''\}\right)\left((\widetilde{\mathbf{A}''} - \widetilde{\mathbf{A}''}_{\mathrm{appro}}) \cdot \mathrm{conj}\{\mathbf{H}''\}\right)^*\right\}$$

$$\approx \mathbf{E}\{\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\mathrm{appro}}\|^2 |H''[k]|^2\} \cdot \mathbf{I}_N.$$

Since

$$\overline{\mathbf{a}}_{\mathrm{appro}} = \frac{1}{N}\mathbf{F_a}\mathbf{P}\mathbf{c}'',$$

the optimal $\mathbf{c}''$ that minimizes

$$\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\mathrm{appro}}\|^2$$

is given by

$$\mathbf{c}'' = (\mathbf{P}^*\mathbf{P})^{-1}\mathbf{P}^*\mathbf{F_a}^*\overline{\mathbf{a}}.$$

Thus,

$$\min \|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\mathrm{appro}}\|^2 = \left\| \overline{\mathbf{a}} - \frac{1}{N}\mathbf{F_a}\mathbf{P}(\mathbf{P}^*\mathbf{P})^{-1}\mathbf{P}^*\mathbf{F_a}^*\overline{\mathbf{a}} \right\|^2$$

$$= \left[\overline{\mathbf{a}} - \frac{1}{N}\mathbf{F_a}\mathbf{P}(\mathbf{P}^*\mathbf{P})^{-1}\mathbf{P}^*\mathbf{F_a}^*\overline{\mathbf{a}} \right]^* \left[\overline{\mathbf{a}} - \frac{1}{N}\mathbf{F_a}\mathbf{P}(\mathbf{P}^*\mathbf{P})^{-1}\mathbf{P}^*\mathbf{F_a}^*\overline{\mathbf{a}} \right]$$

$$= \|\overline{\mathbf{a}}\|^2 - \frac{1}{N}\overline{\mathbf{a}}^*\mathbf{F_a}\mathbf{P}(\mathbf{P}^*\mathbf{P})^{-1}\mathbf{P}^*\mathbf{F_a}^*\overline{\mathbf{a}}$$

$$= \|\overline{\mathbf{a}}\|^2 - \frac{1}{N}\mathrm{Tr}\left\{ \mathbf{F_a}\mathbf{P}(\mathbf{P}^*\mathbf{P})^{-1}\mathbf{P}^*\mathbf{F_a}^* \left(\overline{\mathbf{a}}\overline{\mathbf{a}}^*\right) \right\}$$

$$
= \frac{1}{|A[0]|^2} \|\mathbf{a}\|^2 - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^*} (\overline{\mathbf{a a^*}}) \right\}
$$

$$
= \frac{1}{|A[0]|^2} - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^*} (\overline{\mathbf{a a^*}}) \right\} \tag{3.39}
$$

because $\|\mathbf{a}\|^2 = 1$. Using (3.39), we make the following approximation:

$$
\mathbf{E}\{\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\text{appro}}\|^2 |H''[k]|^2\}
$$

$$
\approx \mathbf{E} \left\{ \left(\frac{1}{|A[0]|^2} - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^*} (\overline{\mathbf{a a^*}}) \right\} \right) |H''[k]|^2 \right\}
$$

$$
= \mathbf{E} \left\{ \frac{1}{|A[0]|^2} |H''[k]|^2 - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^*} (\overline{\mathbf{a a^*}}) |H''[k]|^2 \right\} \right\}
$$

$$
= \mathbf{E} \left\{ |\mu|^2 |H[k]|^2 - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^*} (\mathbf{a a^*}) \right\} |\mu|^2 |H[k]|^2 \right\}
$$

$$
= \mathbf{E} \left\{ |\mu|^2 |H[k]|^2 \right\} \mathbf{E} \left\{ 1 - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^*} (\mathbf{a a^*}) \right\} \right\}
$$

$$
= |\mu|^2 \sigma_H^2 \left(1 - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^* R_a} \right\} \right)
$$

$$
= |\mu|^2 \sigma_H^2 \left(1 - \frac{1}{N} \text{Tr} \left\{ \mathbf{F_a P (P^* P)^{-1} P^* F_a^*} \left(\frac{1}{N^2} \mathbf{F_a R_c F_a^*} \right) \right\} \right)
$$

$$
= |\mu|^2 \sigma_H^2 \left(1 - \frac{1}{N} \text{Tr} \left\{ \mathbf{P (P^* P)^{-1} P^* R_c} \right\} \right)
$$

where

$$
\mathbf{R_a} = \mathbf{E}\{\mathbf{a a^*}\} = \frac{1}{N^2} \mathbf{F_a R_c F_a^*}.
$$

The covariance matrix of $\mathbf{w}''$ is *approximately* equal to $\sigma_{W''}^2 \mathbf{I}_N$, where

$$
\sigma_{W''}^2 = \left(1 + |\nu''|^2 \right) \cdot \mathbf{E}\{\|\overline{\mathbf{a}} - \overline{\mathbf{a}}_{\text{appro}}\|^2 |H''[k]|^2\} \cdot \sigma_P^2 + \sigma_W^2
$$

$$
= \left(1 + |\nu''|^2 \right) |\mu|^2 \sigma_H^2 \left(1 - \frac{1}{N} \text{Tr} \left\{ \mathbf{P (P^* P)^{-1} P^* R_c} \right\} \right) \sigma_P^2 + \sigma_W^2
$$

$$
= \left(|\mu|^2 + |\nu|^2 \right) \left(1 - \frac{1}{N} \text{Tr} \left\{ \mathbf{P (P^* P)^{-1} P^* R_c} \right\} \right) \sigma_H^2 \sigma_P^2 + \sigma_W^2.
$$

Consequently, $\mathbf{I}_\theta$ and the associated CRLB for $\mathbf{h}''$ can be computed (see Appendix F for more details). In the computation, the covariance matrix $\mathbf{R_c}$ of the

phase noise vector $\mathbf{c}$ depends on the phase noise spectral characteristics. Given a specific phase noise model, $\mathbf{R_c}$ can be computed analytically. By using the relation $\overline{\mathbf{h}} = \mathbf{F_h} \mathbf{h}''$, we have

$$\mathbf{E}\{|\mu A[0]\widehat{H}[k] - \mu A[0]H[k]|^2\} = \mathbf{E}\{\|\widehat{\mathbf{h}}'' - \mathbf{h}''\|^2\}. \qquad (3.40)$$

The lower bound for $H[k]$ can then be derived from the lower bound for $\mathbf{h}''$. It is noted that the estimation performance depends on M, and the CRLB allows us to select an appropriate M. A trade-off exists here, because a large M gives better interpolation performance but at the cost of the degree of freedom, while a small M reduces the number of unknowns but causes larger interpolation error. In Section 3.6, we compare the mean-square-errors of channel estimation with the derived CRLB, and show that the CRLB is a good theoretical measure for the estimation accuracy.

3.4   Data Symbol Estimation

Assume that the receiver has acquired the channel response $\mathbf{H}''$ and the IQ imbalance parameter ν''. Given the system model

$$\mathbf{y} = \mathbf{z}'' + \nu'' \widetilde{\mathbf{z}''} + \mathbf{w}$$

where $\mathbf{z}'' = \mathbf{A}''\mathbf{H}''\mathbf{x}$,

$$\mathbf{z}'' = \begin{bmatrix} Z''[0] \\ Z''[1] \\ \vdots \\ Z''[N-1] \end{bmatrix} (N \times 1), \quad \widetilde{\mathbf{z}''} = \begin{bmatrix} (Z''[0])^* \\ (Z''[N-1])^* \\ \vdots \\ (Z''[1])^* \end{bmatrix} (N \times 1),$$

we are now interested in estimating the transmitted vector $\mathbf{x}$. By inspecting the model, it is noticed that the problem can be decomposed into two separate compensation problems: IQ imbalance compensation and phase noise compensation,

as illustrated in Fig. 3.2. First, $\mathbf{z}''$ is estimated from $\mathbf{y}$ by using any IQ imbalance compensation method; then, $\mathbf{x}$ is estimated from $\widehat{\mathbf{z}}''$ by using any phase noise compensation method. Here, we apply the post-FFT IQ compensation technique developed in [TBS05] and the phase noise compensation technique developed in Chapter 2 of this dissertation.
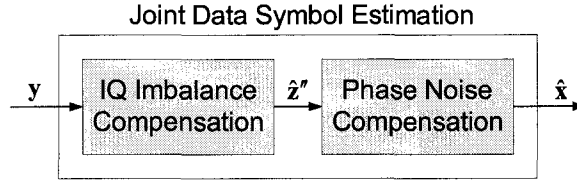
Figure 3.2: Block diagram of the data symbol estimation algorithm.

To perform the IQ compensation, i.e., estimating $\mathbf{z}''$ from $\mathbf{y}$, we notice that

$$Y[k] = Z''[k] + \nu''(Z''[(N-k)_N])^* + W[k], \tag{3.41}$$

$$Y[(N-k)_N] = Z''[(N-k)_N] + \nu''(Z''[k])^* + W[(N-k)_N]. \tag{3.42}$$

Taking the complex conjugate of (3.42) gives

$$Y^*[(N-k)_N] = (Z''[(N-k)_N])^* + (\nu'')^* Z''[k] + W^*[(N-k)_N]. \tag{3.43}$$

Multiplying (3.43) by ν'' and subtracting the result from (3.41) get

$$Y[k] - \nu'' Y^*[(N-k)_N] = (1 - |\nu''|^2) Z''[k] + W[k] - \nu'' W^*[(N-k)_N],$$

from which $Z''[k]$ can be estimated by

$$\widehat{Z}''[k] = \frac{Y[k] - \nu'' Y^*[(N-k)_N]}{1 - |\nu''|^2}.$$

To compensate for phase noise, we also assume that comb-type OFDM symbols are transmitted in the payload portion of each packet. In each comb-type

symbol, some subcarriers are used for pilot symbols, while the others are used for data symbols. The two-stage algorithm is summarized in Algorithm 3.2.

3.4.1 Performance Analysis

In this subsection, we analyze the effects of IQ imbalance and phase noise on OFDM systems in terms of the signal-to-noise ratio degradation. The expressions of the effective signal-to-noise ratio at the receiver are derived by assuming that

1.) the data symbols $X[k]$ are independent and identically distributed with mean zero and variance $\sigma_X^2 = \mathbf{E}\{|X[k]|^2\}$;

2.) the data symbols, the phase noise, the channel coefficients and the additive noise are independent of each other;

3.) the channel coefficients $H[k]$ are independently identically distributed and circularly symmetric Gaussian with mean zero and variance $\sigma_H^2 = \mathbf{E}\{|H[k]|^2\}$.

3.4.1.1 No Impairments

In this case, there is no impairment in the system. Since

$$Y[k] = H[k]X[k] + W[k],$$

the effective signal-to-noise ratio is given by

$$\mathrm{SNR}_0 = \frac{\mathbf{E}\{|H[k]X[k]|^2\}}{\mathbf{E}\{|W[k]|^2\}} = \frac{\sigma_H^2 \sigma_X^2}{\sigma_W^2}. \tag{3.46}$$

Algorithm 3.2 Data Symbol Estimation

1: Estimate $\mathbf{z}''$ from $\mathbf{y}$. This can be done by

$$\widehat{Z}''[k] = \frac{Y[k] - \nu'' Y^*[(N-k)_N]}{1 - |\nu''|^2}, \quad k = 0, 1, \ldots, N-1.$$

2: Assume that $k_{\text{pilot},j}$, $j = 1, 2, \ldots, Q$, are the subcarrier indices of the Q pilot tones. The CPE coefficient $A[0]$ is estimated by

$$\widehat{A}[0] = \frac{\sum_{j=1}^{Q} \left(H''[k_{\text{pilot},j}]\right)^* \left(X[k_{\text{pilot},j}]\right)^* \widehat{Z}''[k_{\text{pilot},j}]}{\sum_{j=1}^{Q} \left|H''[k_{\text{pilot},j}]\right|^2 \left|X[k_{\text{pilot},j}]\right|^2}.$$

3: Let $\widehat{\mathbf{c}}_0' = \left[\begin{array}{cccc} \widehat{A}[0] & \widehat{A}[0] & \ldots & \widehat{A}[0] \end{array}\right]^T$.

4: $i = 1$

5: **repeat**

6: Let $\widehat{\mathbf{a}}_{i-1} = \frac{1}{N} \mathbf{F_a P} \widehat{\mathbf{c}}_{i-1}'$ and construct $\mathbf{A}_{i-1}$ from $\widehat{\mathbf{a}}_{i-1}$ according to (3.12). Find the associated optimal $\widehat{\mathbf{x}}_{\text{data},i-1}$ by solving the following least-squares problem:

$$\widehat{\mathbf{x}}_{\text{data},i-1} = \arg \min_{\mathbf{x}_{\text{data}}} \left\| \widehat{\mathbf{z}}'' - \mathbf{A}_{\text{pilot},i-1} \mathbf{H}''_{\text{pilot}} \mathbf{x}_{\text{pilot}} - \mathbf{A}_{\text{data},i-1} \mathbf{H}''_{\text{data}} \mathbf{x}_{\text{data}} \right\|^2 \qquad (3.44)$$

where $\mathbf{x}_{\text{pilot}}$ is the sub-vector of $\mathbf{x}$ that consists of all the pilot symbols and $\mathbf{A}_{\text{pilot},i-1}$ and $\mathbf{H}''_{\text{pilot}}$ are the associated sub-matrices of $\mathbf{A}_{i-1}$ and $\mathbf{H}''$. Moreover, $\mathbf{x}_{\text{data}}$ is the sub-vector of $\mathbf{x}$ that consists of all the data symbols and $\mathbf{A}_{\text{data},i-1}$ and $\mathbf{H}''_{\text{data}}$ are the associated sub-matrices of $\mathbf{A}_{i-1}$ and $\mathbf{H}''$. The expression for $\widehat{\mathbf{x}}_{\text{data},i-1}$ is given by

$$\widehat{\mathbf{x}}_{\text{data},i-1} = \left(\mathbf{H}''_{\text{data}}\right)^{-1} \left(\mathbf{A}^*_{\text{data},i-1} \mathbf{A}_{\text{data},i-1}\right)^{-1} \mathbf{A}^*_{\text{data},i-1} \left(\widehat{\mathbf{z}}'' - \mathbf{A}_{\text{pilot},i-1} \mathbf{H}''_{\text{pilot}} \mathbf{x}_{\text{pilot}}\right).$$

7: Find the optimal $\widehat{\mathbf{c}}_i'$ by solving the following least-squares problem:

$$\widehat{\mathbf{c}}_i' = \arg \min_{\mathbf{c}'} \left\| \widehat{\mathbf{z}}'' - \mathbf{A}_{\text{pilot}} \mathbf{H}''_{\text{pilot}} \mathbf{x}_{\text{pilot}} - \mathbf{A}_{\text{data}} \mathbf{H}''_{\text{data}} \widehat{\mathbf{x}}_{\text{data},i-1} \right\|^2 \qquad (3.45)$$

where $\mathbf{A}_{\text{pilot}}$ and $\mathbf{A}_{\text{data}}$ are determined from $\mathbf{c}'$ according to (2.9) and (3.12). The expression for $\widehat{\mathbf{c}}_i'$ is given by

$$\widehat{\mathbf{c}}_i' = N \left(\mathbf{P}^* \mathbf{F_a^*} \mathbf{T}^* \mathbf{T} \mathbf{F_a} \mathbf{P}\right)^{-1} \mathbf{P}^* \mathbf{F_a^*} \mathbf{T}^* \widehat{\mathbf{z}}'',$$

where $\mathbf{T}$ is the circulant matrix formed by the elements of $\mathbf{H}'' \widehat{\mathbf{x}}_{i-1}$ with $\widehat{\mathbf{x}}_{i-1}$ formed by $\mathbf{x}_{\text{pilot}}$ and $\widehat{\mathbf{x}}_{\text{data},i-1}$.

8: $i = i + 1$

9: **until** there is no significant improvement in the objective function $\| \widehat{\mathbf{z}}'' - \mathbf{A}_{\text{pilot},i} \mathbf{H}''_{\text{pilot}} \mathbf{x}_{\text{pilot}} - \mathbf{A}_{\text{data},i} \mathbf{H}''_{\text{data}} \widehat{\mathbf{x}}_{\text{data},i-1} \|^2$.

3.4.1.2 No Compensation

In the presence of the IQ and phase noise impairments, the receiver does not perform any compensation. The system model (3.10) can be rewritten as

$$Y[k] = H[k]X[k] + (\mu A[0] - 1)H[k]X[k] + \mu \sum_{r=0, r \neq k}^{N-1} A[(k-r)_N]H[r]X[r]$$

$$+ \nu \sum_{r=0}^{N-1} A^*[(N-k-r)_N]H^*[r]X^*[r] + W[k],$$

where $H[k]X[k]$ is the desired signal component and the other terms are regarded as additive noise. The variance of $H[k]X[k]$ is

$$\mathbf{E}\{|H[k]X[k]|^2\} = \sigma_H^2 \sigma_X^2.$$

The variance of the noise term is given by

$$\mathbf{E}\left\{\left|(\mu A[0] - 1)H[k]X[k] + \mu \sum_{r=0, r \neq k}^{N-1} A[(k-r)_N]H[r]X[r]\right.\right.$$

$$\left.\left. + \nu \sum_{r=0}^{N-1} A^*[(N-k-r)_N]H^*[r]X^*[r] + W[k]\right|^2\right\}$$

$$= \mathbf{E}\left\{\left|(\mu A[0] - 1)H[k]X[k] + \mu \sum_{r=0, r \neq k}^{N-1} A[(k-r)_N]H[r]X[r]\right.\right.$$

$$\left.\left. + \nu \sum_{r=0}^{N-1} A^*[(N-k-r)_N]H^*[r]X^*[r] + W[k]\right|^2\right\}$$

$$= \mathbf{E}\left\{|(\mu A[0] - 1)H[k]X[k]|^2\right\} + |\mu|^2 \sum_{r=0, r \neq k}^{N-1} \mathbf{E}\left\{|A[(k-r)_N]H[r]X[r]|^2\right\}$$

$$+ |\nu|^2 \sum_{r=0}^{N-1} \mathbf{E}\left\{|A^*[(N-k-r)_N]H^*[r]X^*[r]|^2\right\} + \sigma_W^2$$

$$= \mathbf{E}\left\{|\mu A[0] - 1|^2\right\} \sigma_H^2 \sigma_X^2 + |\mu|^2 \sum_{k=1}^{N-1} \sigma_{A,k}^2 \sigma_H^2 \sigma_X^2 + |\nu|^2 \sum_{k=0}^{N-1} \sigma_{A,k}^2 \sigma_H^2 \sigma_X^2 + \sigma_W^2$$

$$= \left(1 - 2\mathrm{Re}\left\{\mu A[0]\right\} + |\mu|^2 + |\nu|^2\right) \sigma_H^2 \sigma_X^2 + \sigma_W^2,$$

because

$$\sum_{k=0}^{N-1} \sigma_{A,k}^2 = 1.$$

Thus, the effective signal-to-noise ratio is given by

$$\begin{aligned}
\text{SNR}_{\text{no}} &= \frac{\sigma_H^2 \sigma_X^2}{\left(1 - 2\text{Re}\left\{\mu A[0]\right\} + |\mu|^2 + |\nu|^2\right)\sigma_H^2 \sigma_X^2 + \sigma_W^2} \\
&= \frac{\text{SNR}_0}{\left(1 - 2\text{Re}\left\{\mu A[0]\right\} + |\mu|^2 + |\nu|^2\right)\text{SNR}_0 + 1}.
\end{aligned} \tag{3.47}$$

3.4.1.3   IQ and Common Phase Error (CPE) Compensation

In this case, the IQ imbalance and the CPE term are compensated for at the receiver, as proposed in [TCP05]. The system model (3.10) can now be rewritten as

$$\begin{aligned}
Y[k] &= \mu A[0]H[k]X[k] + \nu A^*[0]H^*[(N-k)_N]X^*[(N-k)_N] \\
&\quad + \mu \sum_{r=0, r \neq k}^{N-1} A[(k-r)_N]H[r]X[r] + \nu \sum_{\substack{r=0, \\ r \neq (N-k)_N}}^{N-1} A^*[(N-k-r)_N]H^*[r]X^*[r] \\
&\quad + W[k],
\end{aligned}$$

where

$$\mu A[0]H[k]X[k] + \nu A^*[0]H^*[(N-k)_N]X^*[(N-k)_N]$$

is the desired signal component and the other terms are regarded as additive noise. The variance of the desired signal is

$$\begin{aligned}
&\mathbf{E}\left\{\left|\mu A[0]H[k]X[k] + \nu A^*[0]H^*[(N-k)_N]X^*[(N-k)_N]\right|^2\right\} \\
&= \mathbf{E}\left\{\left|\mu A[0]H[k]X[k]\right|^2\right\} + \mathbf{E}\left\{\left|\nu A^*[0]H^*[(N-k)_N]X^*[(N-k)_N]\right|^2\right\} \\
&= (|\mu|^2 + |\nu|^2)\sigma_{A,0}^2 \sigma_H^2 \sigma_X^2,
\end{aligned}$$

and the variance of the noise term is

$$
\mathbf{E}\Biggl\{\Biggl| \mu \sum_{r=0,r\neq k}^{N-1} A[(k-r)_N]H[r]X[r]
$$

$$
+ \nu \sum_{\substack{r=0,\\ r\neq(N-k)_N}}^{N-1} A^*[(N-k-r)_N]H^*[r]X^*[r] + W[k] \Biggr|^2 \Biggr\}
$$

$$
= |\mu|^2 \sum_{r=0,r\neq k}^{N-1} \mathbf{E}\left\{ |A[(k-r)_N]H[r]X[r]|^2 \right\}
$$

$$
+ |\nu|^2 \sum_{\substack{r=0,\\ r\neq(N-k)_N}}^{N-1} \mathbf{E}\left\{ |A^*[(N-k-r)_N]H^*[r]X^*[r]|^2 \right\} + \mathbf{E}\left\{ |W[k]|^2 \right\}
$$

$$
= |\mu|^2 \sum_{k=1}^{N-1} \sigma_{A,k}^2 \sigma_H^2 \sigma_X^2 + |\nu|^2 \sum_{k=1}^{N-1} \sigma_{A,k}^2 \sigma_H^2 \sigma_X^2 + \sigma_W^2
$$

$$
= (|\mu|^2 + |\nu|^2)(1-\sigma_{A,0}^2)\sigma_H^2\sigma_X^2 + \sigma_W^2.
$$

Hence, the effective signal-to-noise ratio is given by

$$
\begin{aligned}
\mathrm{SNR}_{\mathrm{IQ+CPE}} &= \frac{(|\mu|^2+|\nu|^2)\sigma_{A,0}^2\sigma_H^2\sigma_X^2}{(|\mu|^2+|\nu|^2)(1-\sigma_{A,0}^2)\sigma_H^2\sigma_X^2 + \sigma_W^2} \\
&= \frac{(|\mu|^2+|\nu|^2)\sigma_{A,0}^2\mathrm{SNR}_0}{(|\mu|^2+|\nu|^2)(1-\sigma_{A,0}^2)\mathrm{SNR}_0 + 1}.
\end{aligned} \tag{3.48}
$$

3.4.1.4 Proposed Joint Compensation Scheme

With the proposed algorithm, we rewrite the system model (3.14) in the matrix form as

$$
\begin{aligned}
\mathbf{y} = \mu\mathbf{A}_{\mathrm{appro}}\mathbf{H}\mathbf{x} + \nu\widetilde{\mathbf{A}}_{\mathrm{appro}} \cdot \mathrm{conj}\{\mathbf{H}\} \cdot \mathrm{conj}\{\mathbf{x}\} + \mu(\mathbf{A} - \mathbf{A}_{\mathrm{appro}})\mathbf{H}\mathbf{x} \\
+ \nu(\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}_{\mathrm{appro}}) \cdot \mathrm{conj}\{\mathbf{H}\} \cdot \mathrm{conj}\{\mathbf{x}\} + \mathbf{w},
\end{aligned} \tag{3.49}
$$

where $\mathbf{A}_{\mathrm{appro}}$ is determined by the vector

$$
\mathbf{a}_{\mathrm{appro}} = \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{c}'
$$

according to the construction of $\mathbf{A}$. Note that $\mathbf{A} - \mathbf{A}_{\text{appro}}$ represents the modeling error existing in the approximation given by (2.9). In (3.49),

$$\mu \mathbf{A}_{\text{appro}} \mathbf{H} \mathbf{x} + \nu \widetilde{\mathbf{A}}_{\text{appro}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\}$$

is the desired signal component and

$$\mu (\mathbf{A} - \mathbf{A}_{\text{appro}}) \mathbf{H} \mathbf{x} + \nu (\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w}$$

is the noise component. The variance of the signal component is given by

$$\mathbf{E} \left\{ \left\| \mu \mathbf{A}_{\text{appro}} \mathbf{H} \mathbf{x} + \nu \widetilde{\mathbf{A}}_{\text{appro}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} \right\|^2 \right\}$$

$$= \mathbf{E} \left\{ \left(\mu \mathbf{A}_{\text{appro}} \mathbf{H} \mathbf{x} + \nu \widetilde{\mathbf{A}}_{\text{appro}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} \right)^* \left(\mu \mathbf{A}_{\text{appro}} \mathbf{H} \mathbf{x} + \nu \widetilde{\mathbf{A}}_{\text{appro}} \right. \right.$$

$$\left. \left. \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} \right) \right\}$$

$$= \mathbf{E} \left\{ (\mu \mathbf{A}_{\text{appro}} \mathbf{H} \mathbf{x})^* (\mu \mathbf{A}_{\text{appro}} \mathbf{H} \mathbf{x}) \right\} + \mathbf{E} \left\{ \left(\nu \widetilde{\mathbf{A}}_{\text{appro}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} \right)^* \right.$$

$$\left. \cdot \left(\nu \widetilde{\mathbf{A}}_{\text{appro}} \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} \right) \right\}$$

$$= N |\mu|^2 \mathbf{E} \left\{ \|\mathbf{a}_{\text{appro}}\|^2 \right\} \sigma_H^2 \sigma_X^2 + N |\nu|^2 \mathbf{E} \left\{ \|\mathbf{a}_{\text{appro}}\|^2 \right\} \sigma_H^2 \sigma_X^2$$

$$= N \left(|\mu|^2 + |\nu|^2 \right) \mathbf{E} \left\{ \|\mathbf{a}_{\text{appro}}\|^2 \right\} \sigma_H^2 \sigma_X^2$$

and the variance of the noise component is given by

$$\mathbf{E} \left\{ \left\| \mu (\mathbf{A} - \mathbf{A}_{\text{appro}}) \mathbf{H} \mathbf{x} + \nu (\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w} \right\|^2 \right\}$$

$$= \mathbf{E} \left\{ \left(\mu (\mathbf{A} - \mathbf{A}_{\text{appro}}) \mathbf{H} \mathbf{x} + \nu (\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w} \right)^* \right.$$

$$\left. \cdot \left(\mu (\mathbf{A} - \mathbf{A}_{\text{appro}}) \mathbf{H} \mathbf{x} + \nu (\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} + \mathbf{w} \right) \right\}$$

$$= \mathbf{E} \left\{ (\mu (\mathbf{A} - \mathbf{A}_{\text{appro}}) \mathbf{H} \mathbf{x})^* (\mu (\mathbf{A} - \mathbf{A}_{\text{appro}}) \mathbf{H} \mathbf{x}) \right\} + \mathbf{E} \left\{ \left(\nu (\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}_{\text{appro}}) \right. \right.$$

$$\left. \left. \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} \right)^* \left(\nu (\widetilde{\mathbf{A}} - \widetilde{\mathbf{A}}_{\text{appro}}) \cdot \text{conj}\{\mathbf{H}\} \cdot \text{conj}\{\mathbf{x}\} \right) \right\} + \mathbf{E} \left\{ \mathbf{w}^* \mathbf{w} \right\}$$

$$= N |\mu|^2 \mathbf{E} \left\{ \|\mathbf{a} - \mathbf{a}_{\text{appro}}\|^2 \right\} \sigma_H^2 \sigma_X^2 + N |\nu|^2 \mathbf{E} \left\{ \|\mathbf{a} - \mathbf{a}_{\text{appro}}\|^2 \right\} \sigma_H^2 \sigma_X^2 + N \sigma_W^2$$

$$= N \left(|\mu|^2 + |\nu|^2 \right) \mathbf{E} \left\{ \|\mathbf{a} - \mathbf{a}_{\text{appro}}\|^2 \right\} \sigma_H^2 \sigma_X^2 + N \sigma_W^2.$$

Recall the following approximation:

$$\mathbf{E}\left\{\|\mathbf{a}_{\mathrm{appro}}\|^2\right\} = \mathbf{E}\left\{\mathbf{a}_{\mathrm{appro}}^* \mathbf{a}_{\mathrm{appro}}\right\}$$

$$\approx \mathbf{E}\left\{\left(\frac{1}{N}\mathbf{F_a}\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{c}\right)^* \left(\frac{1}{N}\mathbf{F_a}\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{c}\right)\right\}$$

$$= \mathbf{E}\left\{\frac{1}{N}\mathbf{c}^*\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{c}\right\}$$

$$= \mathbf{E}\left\{\frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{c}\mathbf{c}^*\right\}\right\}$$

$$= \frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R_c}\right\}$$

and

$$\mathbf{E}\left\{\|\mathbf{a} - \mathbf{a}_{\mathrm{appro}}\|^2\right\} = \mathbf{E}\left\{\|\mathbf{a} - \mathbf{a}_{\mathrm{appro}}\|^2\right\}$$

$$= \mathbf{E}\left\{(\mathbf{a} - \mathbf{a}_{\mathrm{appro}})^*(\mathbf{a} - \mathbf{a}_{\mathrm{appro}})\right\}$$

$$= \mathbf{E}\left\{\left(\frac{1}{N}\mathbf{F_a}\mathbf{c} - \frac{1}{N}\mathbf{F_a}\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{c}\right)^*\right.$$

$$\left. \cdot \left(\frac{1}{N}\mathbf{F_a}\mathbf{c} - \frac{1}{N}\mathbf{F_a}\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{c}\right)\right\}$$

$$= 1 - \frac{1}{N}\mathbf{E}\left\{\mathbf{c}^*\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{c}\right\}$$

$$= 1 - \frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R_c}\right\}.$$

Hence, the effective signal-to-noise ratio is given by

$$\mathrm{SNR}_{\mathrm{prop}} = \frac{N\left(|\mu|^2 + |\nu|^2\right)\cdot\mathbf{E}\{\|\mathbf{a}_{\mathrm{appro}}\|^2\}\cdot\sigma_H^2\sigma_X^2}{N\left(|\mu|^2 + |\nu|^2\right)\cdot\mathbf{E}\{\|\mathbf{a} - \mathbf{a}_{\mathrm{appro}}\|^2\}\cdot\sigma_H^2\sigma_X^2 + N\sigma_W^2}$$

$$\approx \frac{\left(|\mu|^2 + |\nu|^2\right)\cdot\frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R_c}\right\}\cdot\sigma_H^2\sigma_X^2}{\left(|\mu|^2 + |\nu|^2\right)\cdot\left(1 - \frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R_c}\right\}\right)\cdot\sigma_H^2\sigma_X^2 + \sigma_W^2}$$

$$= \frac{\left(|\mu|^2 + |\nu|^2\right)\cdot\frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R_c}\right\}\cdot\mathrm{SNR}_0}{\left(|\mu|^2 + |\nu|^2\right)\cdot\left(1 - \frac{1}{N}\mathrm{Tr}\left\{\mathbf{P}\left(\mathbf{P}^*\mathbf{P}\right)^{-1}\mathbf{P}^*\mathbf{R_c}\right\}\right)\cdot\mathrm{SNR}_0 + 1}. \tag{3.50}$$

Fig. 3.3 plots the effective signal-to-noise ratio for different compensation scenarios by using expressions (3.46)-(3.48) and (3.50) when $\alpha = 0.1$, $\theta = 10°$, and the phase noise is generated by an oscillator with $\xi = 2.5$ kHz.

95
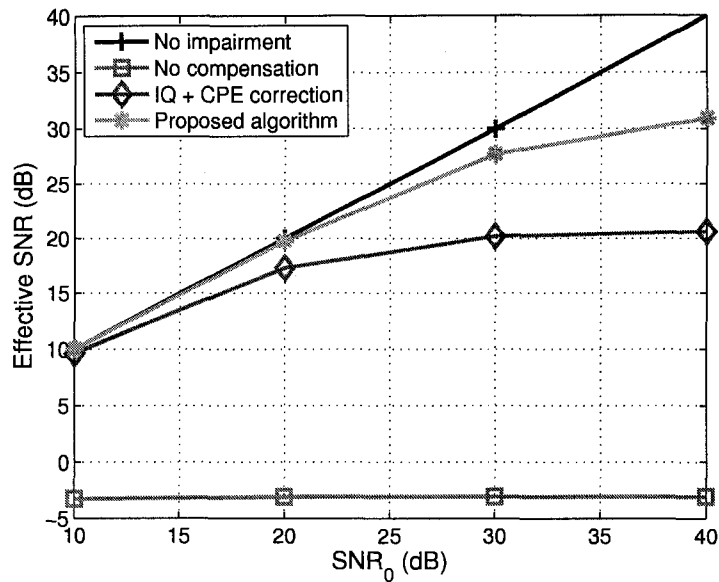
Figure 3.3: Plots of the effective signal-to-noise ratio at the receiver by using
(3.46)-(3.48) and (3.50) when $\alpha = 0.1$, $\theta = 10°$ and $\xi = 2.5$ kHz.

3.5 Complexity Analysis and Efficient Implementation

In this section, we analyze the complexity of the proposed algorithm and present
an efficient implementation scheme. The proposed channel estimation algorithm
needs to solve a linear least-squares problem, i.e., (3.22), iteratively, while the
proposed data symbol estimation algorithm needs to solve two least-squares prob-
lems, i.e., (3.44) and (3.45), iteratively. The simulations presented in the next
section suggest that about 10 iterations for the channel estimation and 20 itera-
tions for the data symbol estimation are generally sufficient to guarantee conver-
gence, as illustrated by Fig. 3.6. Solving a general least-squares problem of size
N has computational complexity $O(N^3)$. Thus, the complexity of the proposed
scheme is $O(KN^3)$, where K denotes the number of iterations.

The number of arithmetic operations required by the proposed algorithms is

estimated as follows. In the joint channel estimation algorithm (Algorithm 3.1), Step 0 solves the least-squares problem:

$$\widehat{\mathbf{h}}_0'' = \arg\min_{\mathbf{h}''} \| \mathbf{y} - \mathbf{H}''\mathbf{x} \|^2.$$

Since the lengths of $\mathbf{y}$ and $\mathbf{h}''$ are N and L, respectively, the number of arithmetic operations required to solve this problem via QR factorization is about $2NL^2 - \frac{2}{3}L^3$ (see page 83 of [TI97]). Step 4 needs to iteratively solve a least-squares problem with $M + L + 1$ unknowns and a linear constraint. Because of the constraint, the number of independent unknowns is $M + L$, and hence the required number of operations is about $2N(M + L)^2 - \frac{2}{3}(M + L)^3$. Assume that K_1 iterations are needed to reach convergence. The total number of operations required by the channel estimation algorithm is thus given by

$$2NL^2 - \frac{2}{3}L^3 + 2K_1 N(M + L)^2 - \frac{2}{3}K_1(M + L)^3.$$

In the proposed joint data symbol estimation algorithm (Algorithm 3.2), Step 6 solves a least-squares problem with $N - Q$ unknowns, and its required number of operations is about $2N(N - Q)^2 - \frac{2}{3}(N - Q)^3$. Step 7 solves a least-squares problem with M unknowns, and its required number of operations is $2NM^2 - \frac{2}{3}M^3$. Therefore, the total number of operations required by the data symbol estimation algorithm is about

$$K_2 \left[2N(N - Q)^2 - \frac{2}{3}(N - Q)^3 + 2NM^2 - \frac{2}{3}M^3 \right],$$

where K_2 is the number of iterations.

The aforementioned complexity can be reduced by the following efficient implementation. First, the channel estimation algorithm is only exploited occasionally, e.g., once per several packets, because the channel and IQ parameters are usually slowly time-varying. For data symbol estimation, on the other hand we

propose the following efficient algorithm with complexity $O(N \log_2 N)$; the algorithm exploits the circularly symmetric structure of $\mathbf{A}$ and the sparse structure of $\mathbf{P}$. In Step 6 of Algorithm 3.2, we solve the following least-squares problem:

$$\min_{\mathbf{x}_{\text{data}}} \left\| \widehat{\mathbf{z}}'' - \mathbf{A}_{i-1} \mathbf{H}'' \mathbf{x} \right\|^2 ,$$

which is equivalent to

$$\min_{\mathbf{x}_{\text{data}}} \left\| \mathbf{F}_{\mathbf{a}} \widehat{\mathbf{z}}'' - \frac{1}{N} \mathbf{F}_{\mathbf{a}} \mathbf{A}_{i-1} \mathbf{F}_{\mathbf{a}}^* \mathbf{F}_{\mathbf{a}} \mathbf{H}'' \mathbf{x} \right\|^2 , \qquad (3.51)$$

because $\frac{1}{\sqrt{N}} \mathbf{F}_{\mathbf{a}}$ is unitary. Note that $\mathbf{F}_{\mathbf{a}} \mathbf{A}_{i-1} \mathbf{F}_{\mathbf{a}}^*$ is diagonal because $\mathbf{A}_{i-1}$ is circulant. Then we solve the following *approximate* problem (with similar performance):

$$\min_{\mathbf{x}_{\text{data}}} \left\| \left(\frac{1}{N} \mathbf{F}_{\mathbf{a}} \mathbf{A}_{i-1} \mathbf{F}_{\mathbf{a}}^* \right)^{-1} \mathbf{F}_{\mathbf{a}} \widehat{\mathbf{z}}'' - \mathbf{F}_{\mathbf{a},\text{pilot}} \mathbf{H}''_{\text{pilot}} \mathbf{x}_{\text{pilot}} - \mathbf{F}_{\mathbf{a},\text{data}} \mathbf{H}''_{\text{data}} \mathbf{x}_{\text{data}} \right\|^2 , \quad (3.52)$$

which leads to

$$\mathbf{x}_{\text{data},i} = \frac{1}{N} \left(\mathbf{H}''_{\text{data}} \right)^{-1} \mathbf{F}^*_{\mathbf{a},\text{data}} \qquad (3.53)$$

$$\cdot \left[\left(\frac{1}{N} \mathbf{F}_{\mathbf{a}} \mathbf{A}_{i-1} \mathbf{F}_{\mathbf{a}}^* \right)^{-1} \mathbf{F}_{\mathbf{a}} \widehat{\mathbf{z}}'' - \mathbf{F}_{\mathbf{a},\text{pilot}} \mathbf{H}''_{\text{pilot}} \mathbf{x}_{\text{pilot}} \right] .$$

Problem (3.52) is equivalent to the original problem (3.51) if $\mathbf{A}_{i-1}$ is diagonal (with all diagonal elements identical). Consider that when $\phi(t) = 0$, we have $\mathbf{A} = \mathbf{I}_N$. When $\phi(t)$ is small, $\mathbf{A}_{i-1}$ is close to a diagonal matrix with small off-diagonal elements. By using the Fast Fourier transform (FFT) to compute (3.53), this approximate solution requires computational complexity $O(N \log_2 N)$. This is because $\frac{1}{N} \mathbf{F}_{\mathbf{a}} \mathbf{A}_{i-1} \mathbf{F}_{\mathbf{a}}^*$ can be computed by FFT and the result is diagonal. Also, $\mathbf{F}_{\mathbf{a}} \widehat{\mathbf{z}}''$ and $\mathbf{F}_{\mathbf{a},\text{pilot}} \mathbf{H}''_{\text{pilot}} \mathbf{x}_{\text{pilot}}$ can be computed by FFT with complexity at most $O(N \log_2 N)$. Thus, the total complexity is $O(N \log_2 N)$.

In Step 7 of Algorithm 3.2, we solve the following least-squares problem:

$$\min_{\mathbf{c}'} \left\| \widehat{\mathbf{z}}'' - \mathbf{A} \mathbf{H}'' \widehat{\mathbf{x}}_{i-1} \right\|^2 ,$$

which is equivalent to

$$\min_{\mathbf{c}'} \left\| \widehat{\mathbf{z}}'' - \frac{1}{N}\mathbf{T}\mathbf{F_a}\mathbf{P}\mathbf{c}' \right\|^2, \qquad (3.54)$$

where $\mathbf{T}$ is the circularly symmetric matrix formed from the elements of $\mathbf{H}''\widehat{\mathbf{x}}_{i-1}$.
Solving (3.54) is equivalent to solving

$$\min_{\mathbf{c}'} \left\| \mathbf{F_a^*}\widehat{\mathbf{z}}'' - \frac{1}{N}\mathbf{F_a^*}\mathbf{T}\mathbf{F_a}\mathbf{P}\mathbf{c}' \right\|^2.$$

Since $\frac{1}{N}\mathbf{F_a^*}\mathbf{T}\mathbf{F_a}$ is diagonal, letting

$$\overline{\mathbf{z}} = \mathbf{F_a^*}\widehat{\mathbf{z}}''$$

and

$$\overline{\mathbf{P}} = \frac{1}{N}\mathbf{F_a^*}\mathbf{T}\mathbf{F_a}\mathbf{P}$$

leads to

$$\min_{\mathbf{c}'} \left\| \overline{\mathbf{z}} - \overline{\mathbf{P}}\mathbf{c}' \right\|^2.$$

If $\mathbf{P}$ is constructed by using linear interpolation (2.26), $\mathbf{P}$ and hence $\overline{\mathbf{P}}$ are sparse and have only $2N$ nonzero elements. Solving this sparse least-squares problem requires computational complexity $O(N)$ [GH80, GHP81, Saa96]. Thus, the total computational cost is $O(N \log_2 N)$. The above analysis shows that the proposed scheme can be implemented in either a standard manner with complexity $O(N^3)$ or a more efficient manner with complexity $O(N \log_2 N)$.

3.6 Computer Simulations

In the simulations, the system bandwidth is 20 MHz, i.e., $T_s = 0.05$ μs, and the constellation used for symbol mapping is 64-QAM. The OFDM symbol size is $N = 64$ and the prefix length is $P = 16$.[1] The channel length is 6, and each tap

---

[1]This is the same as in the IEEE 802.11a standard.
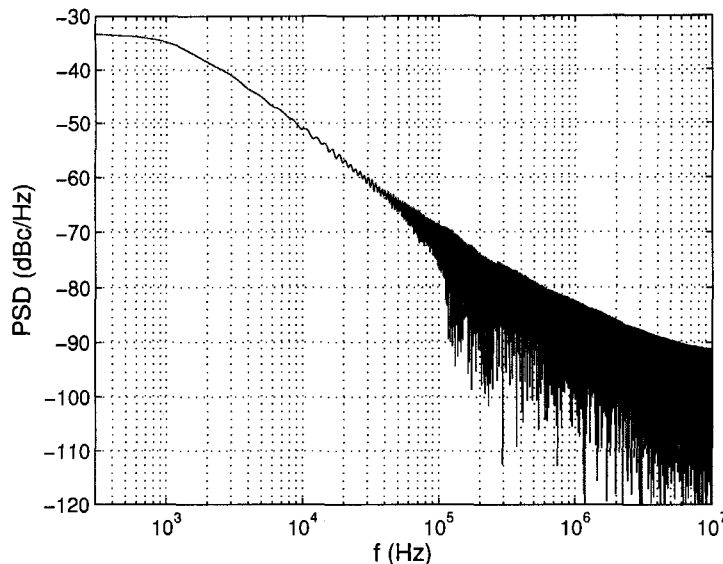
Figure 3.4: PSD of the phase noise generated by a free-running oscillator with $\xi = 2.5$ kHz.

is independently Rayleigh distributed with the power profile specified by 3 dB decay per tap. The average power of the channel response is normalized to 1, i.e., $\sigma_H^2 = 1$. We simulate an OFDM receiver with the IQ imbalance specified by $\alpha = 0.1$ and $\theta = 10°$. The phase noise is generated according to the free-running oscillator model given in Appendix A with linewidth $\xi = 2.5$ kHz, and its spectrum is shown in Fig. 3.4.

We first examine the performance of different channel estimation algorithms for different scenarios. In the simulations, only one block-type pilot symbol is used for each time of channel estimation. The assumed channel length in the time domain is $L = 16$ and the length of the phase noise vector to be estimated is $M = 8$.[2] Fig. 3.5(a) plots the mean-square-errors (MSE) of different channel estimation algorithms vs. the normalized signal-to-noise ratio at the receiver, i.e., SNR =

---

[2]L is the assumed maximum channel length and equal to the cyclic prefix length P.

σ_P^2/σ_W^2. It is shown that estimating $\mathbf{h}'$ rather than $\mathbf{h}$ can improve the accuracy in terms of MSE by a factor of $L/N = 16/64 = -6.02$ dB. The proposed joint channel estimation algorithm performs better than the conventional methods that simply treat the impairments as additive noise. In Fig. 3.5(b), the CRLB is plotted in dotted lines by using the expressions (3.34)-(3.37) and (3.40). By comparing Fig. 3.5(a) and Fig. 3.5(b), it can be seen that the CRLB gives a good measure about the accuracy of different algorithms. Moreover, Fig. 3.6 demonstrates that the proposed channel estimation algorithm requires about 10 iterations to ensure convergence.

The proposed data symbol estimation algorithm is simulated in comparison with the ideal OFDM receiver with no impairment and the IQ+CPE (common phase error) correction scheme proposed in [TCP05]. During the payload portion of OFDM packets, 16 out of the 64 subcarriers are used for pilot tones, i.e., $Q = 16$. Fig. 3.7(a) shows the uncoded bit error rate (BER) performance of the system when the receiver has the *perfect* channel information, while Fig. 3.7(b) shows the uncoded BER performance when the receiver only has the *estimated* channel information. It is demonstrated by Fig. 3.7(a) that the proposed algorithm achieves better performance in phase noise compensation even if the receiver has perfect channel information. Compared to the IQ+CPE scheme, the proposed method achieves lower BERs, because it not only corrects the common phase rotation of the received constellation but also suppresses part of the inter-carrier interference caused by phase noise. In other words, the proposed algorithm can reduce the sensitivity of OFDM receivers to the analog impairments effectively. Fig. 3.7(b) shows that if the receivers have to estimate the channel response, the proposed channel estimation algorithm obtains better channel estimates and thus improves the system performance.
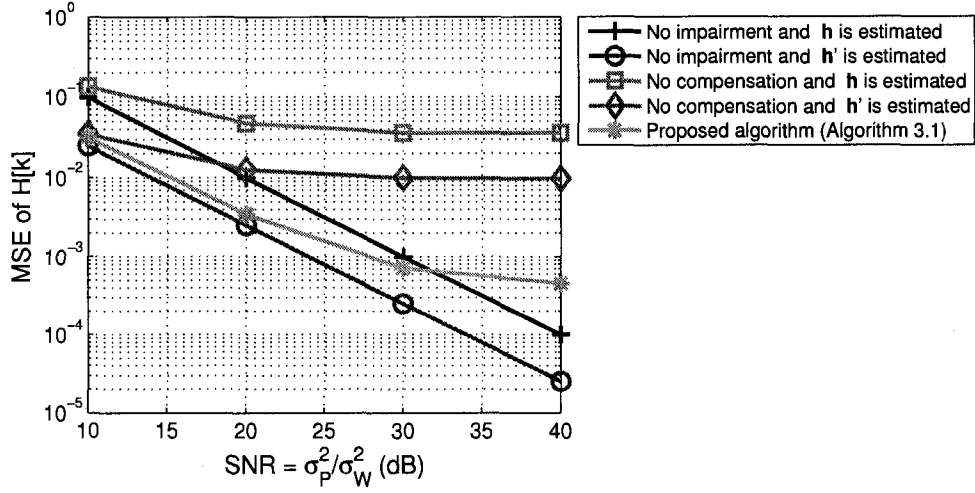
(a) MSE obtained by computer simulations.
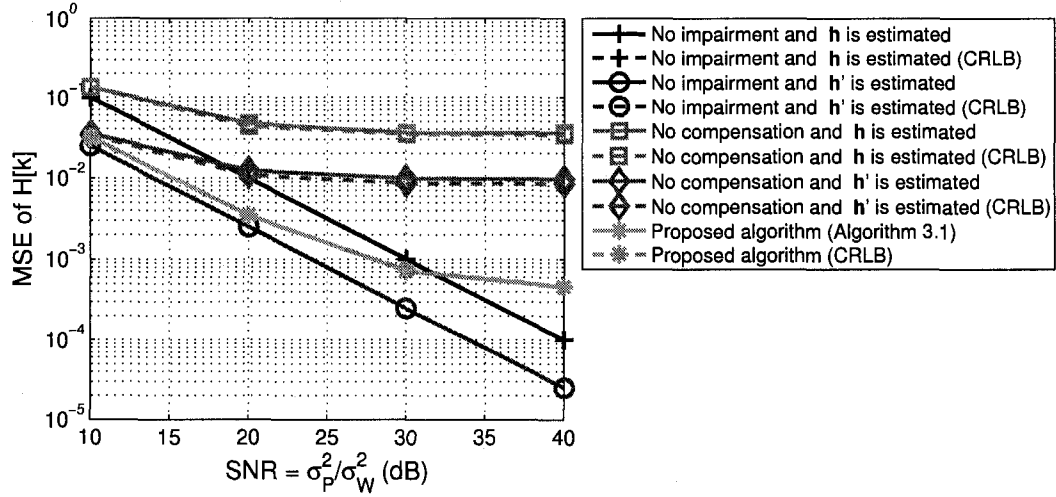
(b) CRLB computed by using the formulas derived in Subsection 3.3.1.

Figure 3.5: Plots of the MSE and CRLB for channel estimation when $\alpha = 0.1$, $\theta = 10°$ and $\xi = 2.5$ kHz.
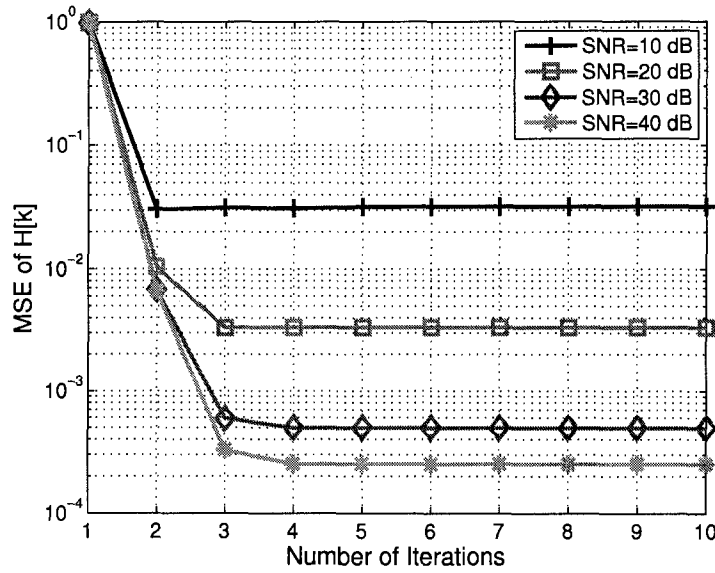
Figure 3.6: Plots of the MSE of channel estimation vs. the number of iterations when $\alpha = 0.1$, $\theta = 10°$ and $\xi = 2.5$ kHz.

As a summary of the above observations, we highlight the following points:

1. As shown in Fig. 3.3, the performance of the IQ+CPE algorithm proposed in [TCP05] saturates at about 21 dB for the impairment parameters $\alpha = 0.1$, $\theta = 10°$ and $\xi = 2.5$ kHz. On the other hand, the proposed algorithm can achieve an effective signal-to-noise ratio of 27-30 dB, which is suitable for many high data rate scenarios.

2. The CRLB analysis depicted in Fig. 3.5 shows that in principle, the channel estimation in the presence of severe IQ imbalance and phase noise can be performed with a performance close to that of an ideal system with no impairments. This emphasizes the importance and promise of compensating for RF impairments in the digital domain since the loss compared to an ideal system can be minimal based on the CRLB results.

103
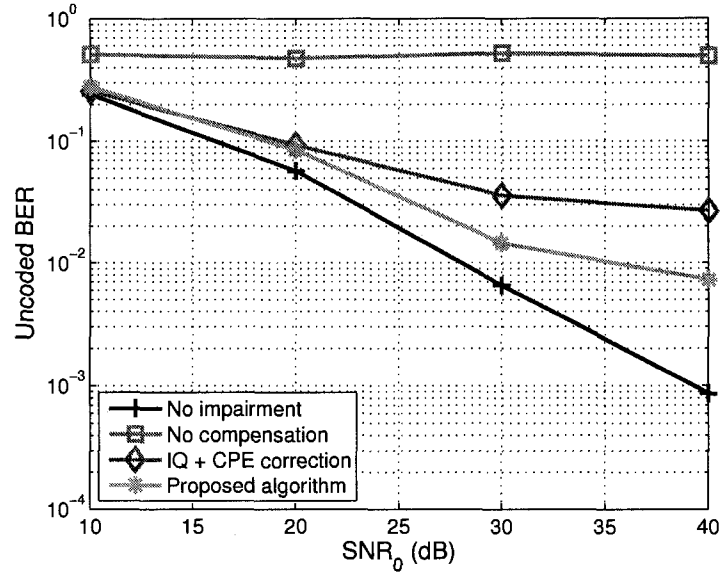
(a) With perfect channel information.
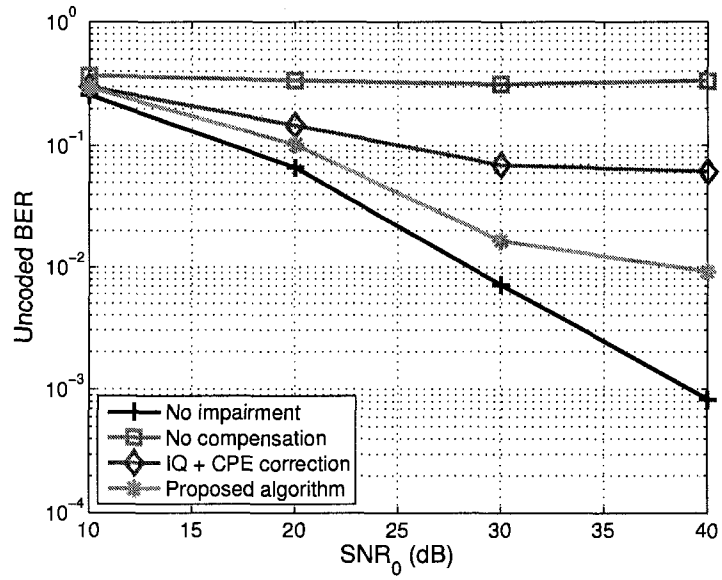
(b) With estimated channel information.

Figure 3.7: Plots of uncoded BER vs. SNR_0 for $\alpha = 0.1$, $\theta = 10°$ and $\xi = 2.5$ kHz.

3. Although the proposed solution is more complicated than other solutions available in the literature, e.g., the IQ+CPE algorithm has complexity $O(N)$, it provides a full-digital approach with better performance to solve problems from the analog domain. With the rapid advances of VLSI technologies, most of the design constraints and burdens will be shifted from the analog domain to the digital domain, for which the solution is promising.

3.7   Conclusions

In this chapter, the joint effects of IQ imbalance and phase noise on OFDM systems are studied. A compensation scheme is proposed that consists of two stages. One stage is the joint channel estimation, and the other is the joint data symbol estimation. The proposed channel estimation algorithm performs close to the derived Cramer-Rao lower bound in the presence of the impairments. Also, the analysis and simulations show that the compensation scheme can effectively improve the system performance and reduce the sensitivity of OFDM receivers to the analog impairments. This work can be further extended to include other analog distortions, e.g., carrier frequency offset. Since receivers with less analog impairments usually have the disadvantage of high implementation cost, our technique enables the use of low-cost receivers for OFDM communications.

CHAPTER 4

Compensation of RF Nonlinearities

4.1 Introduction

A software-defined radio (SDR) system is a radio communication system that can tune to any frequency band and receive any modulation across a large frequency spectrum by means of programmable hardware [Ree02, DMA03, Tut04, Ken05, BK07]. SDR systems allow the feasibility of different wireless services by using just a single reconfigurable chipset. Moreover, SDR systems offer a platform for the newly emerging cognitive radio technology, which demands high controllability and programmability for radio transmission and reception [Mit00, Hay05]. The convenience and promise of SDR face numerous challenges. Traditionally, in order to simultaneously communicate over different frequency bands, the receiver uses several RF front-end modules so that signals in different bands can be received and processed separately. Fig. 4.1 shows an RF receiver dedicated to a communication channel with carrier frequency ω_c. Because of the high out-of-band rejection characteristic of the band-selection surface acoustic wave (SAW) filter, interferences at other frequencies are suppressed and they cause little distortion to the desired signal (as demonstrated in Fig. 4.2), even in the presence of considerable front-end analog distortions such as nonlinearities and IQ imbalances [Raz98, Abi95]. Unlike conventional RF receivers, an SDR uses a wideband RF front-end module with several GHz bandwidth. A *tunable* synthesizer and mixer

106
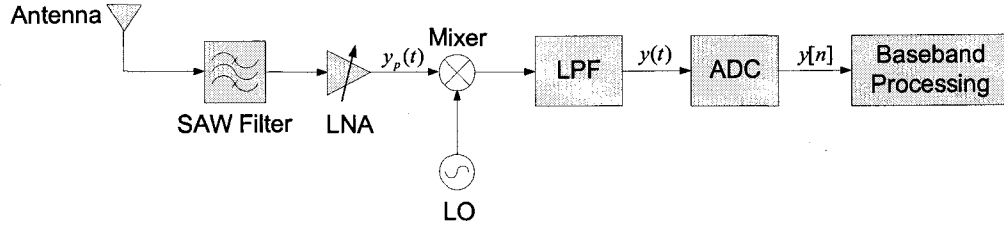
Figure 4.1: A traditional receiver dedicated to the frequency band with carrier frequency ω_c.
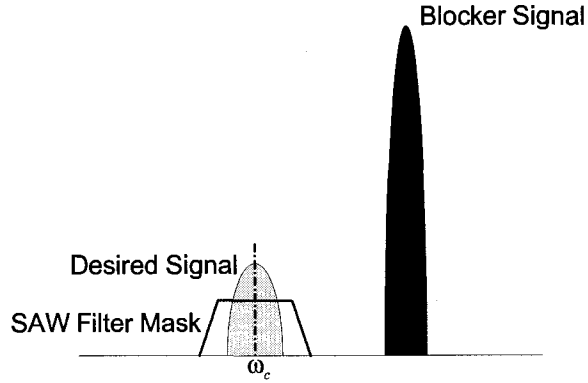
Figure 4.2: A SAW filter can effectively remove the strong blocker signals in other frequency bands.

are used to lock in the desired frequency band and down-convert the signal to the baseband [BMC06, Abi07]. Without the SAW filter,[1] all the signals and interferences existing in the wide band range are amplified and down-converted. Due to the unavoidable nonlinearity in the low-noise amplifier (LNA), the presence of strong blocker (interference) signals causes cross modulation over the desired signal. This threat becomes significantly harmful, especially when the desired signal is weak.

---

[1] A SAW filter cannot be used here because it is application-specific with a fixed center frequency and bandwidth. Until now, there is no tunable SAW filter with sufficiently good performance. Furthermore, the SAW filter cannot be integrated on-chip with the receiver circuitry, which means that a multi-standard receiver with many SAW filters will be bulky and expensive.

While analog/RF designers are striving to improve the linearity of RF receivers, there have been works in the literature to mitigate this impairment by using digital domain techniques [BBC88, NP90, HLH04, DS06, VGA06]. These digital solutions provide a flexible alternative approach to combat nonlinearities, which is particularly appropriate for SDRs that have an extremely wide bandwidth and a reconfigurable hardware/software structure. In this chapter, we propose a nonlinearity compensation scheme for the SDR structure proposed in [BMC06, Abi07]. As shown in Fig. 4.3, this SDR system has a wideband RF front-end (0.8-6.0 GHz), and is able to selectively down-convert and sample the signals in desired frequency bands. Our scheme will require two RF signal paths - one is used for capturing the signal in the desired band, while the other is used to locate and acquire the blocker signal.[2] The baseband processor then jointly processes the two discretized signals to alleviate the effects of cross modulation.

The chapter is organized as follows. The next section describes the system model and formulates the effects of RF nonlinearities when there exists only one blocker signal. The proposed compensation scheme is presented in Section 4.3, and its performance is analyzed in Section 4.4 in terms of the Cramer-Rao lower bound. Section 4.5 extends the system model and the proposed scheme to include multiple blocker signals as well as the Orthogonal Frequency Division Multiplexing (OFDM) transmission scheme. Simulation results are presented and discussed in Section 4.6.

Throughout this chapter, we adopt the following notations. $(\cdot)^T$ denotes the matrix transpose and $(\cdot)^*$ represents the matrix conjugate transpose. $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ return the real and imaginary parts of its argument, respectively. $\mathbf{E}\{\cdot\}$ is the expected value with respect to the underlying probability measure.

---

[2]The secondary RF path, used to acquire the blocker signal, can be implemented with smaller area and less power compared to the main path.
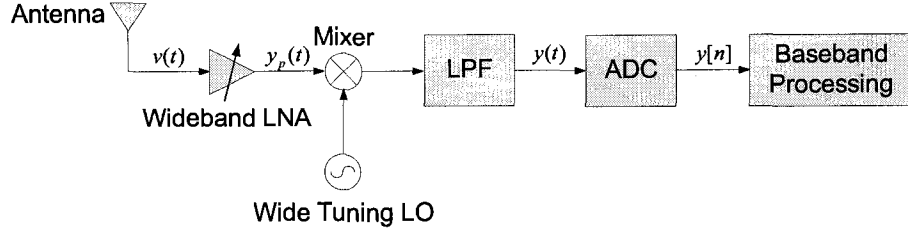
Figure 4.3: SDR with a wideband front-end RF receiver [BMC06, Abi07].

4.2 System Model

Referring to Fig. 4.3, in the presence of nonlinearities, the input-output characteristic of the receiver front-end is modeled as

$$y_p(t) = \gamma_1 v(t) + \gamma_2 \left[v(t)\right]^2 + \gamma_3 \left[v(t)\right]^3 + w_p(t), \qquad (4.1)$$

where $v(t)$ and $y_p(t)$ are the real input and output signals, $w_p(t)$ is additive white Gaussian noise, and γ_1, γ_2, γ_3 are real constants. In this expression, $\gamma_1 v(t)$ represents the linear component in the output, while $\gamma_2[v(t)]^2$ and $\gamma_3[v(t)]^3$ are the second and third-order nonlinear components in the output. The dynamic range of $v(t)$ depends on the sensitivity of the receive antenna and is assumed to be within the range $[-0.05, 0.05]$ throughout the chapter. The values of γ_1, γ_2 and γ_3 are related to circuit specification parameters:

(1) γ_1 is the small signal gain and its typical value is 35 dB, i.e., $\gamma_1 = 56.23$.

(2) γ_2 is the coefficient representing the second-order nonlinearity, usually expressed in terms of the so-called IP2 (second-order intercept point) coefficient, i.e.,

$$IP2 = 20 \log_{10} \frac{|\gamma_1|}{|\gamma_2|} + 10 \ (\text{dBm}),$$
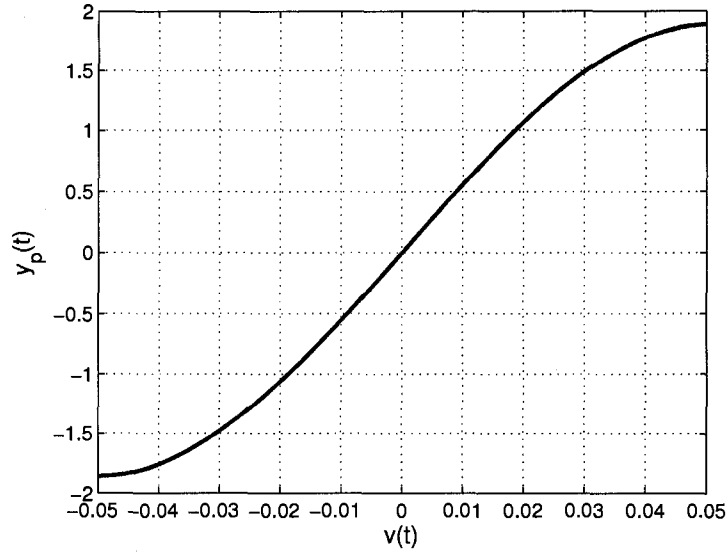
Figure 4.4: Plot of the input-output relation - $y_p(t)$ vs. $v(t)$ for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$ and $\gamma_3 = -7497.33$.

where dBm is the power ratio in decibels (dB) referenced to one milliwatt (mW). The typical value of IP2 is 30 dBm, implying that $\gamma_2 = \pm 5.623$ if $\gamma_1 = 56.23$.

(3) γ_3 is the coefficient representing the third-order nonlinearity, usually expressed in terms of the IP3 (third-order intercept point) coefficient, i.e.,

$$\text{IP3} = 20 \log_{10} V_{\text{IP3}} + 10 \text{ (dBm)}$$

where

$$V_{\text{IP3}} = \sqrt{\frac{4|\gamma_1|}{3|\gamma_3|}}.$$

The typical value of IP3 is -10 dBm, implying that $\gamma_3 = \pm 7497.33$ if $\gamma_1 = 56.23$.

For example, Fig. 4.4 plots $y_p(t)$ vs. $v(t)$ for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$ and $\gamma_3 = -7497.33$ according to (4.1).

110

Assume that the acquired signal $v(t)$ contains a desired signal around frequency ω_1 and a blocker signal around frequency ω_2.[3] Then, $v(t)$ can be represented by

$$v(t) = \text{Re}\left\{ \sqrt{2}z_1(t)e^{j\omega_1 t} + \sqrt{2}z_2(t)e^{j\omega_2 t} \right\}, \qquad (4.2)$$

where $z_1(t)$ and $z_2(t)$ are the corresponding baseband signals at ω_1 and ω_2. Taking the channel response of the desired channel into account, $z_1(t)$ is given by the convolution of the transmitted baseband signal $x_1(t)$ and the continuous-time baseband channel response $h(t)$, i.e.,

$$z_1(t) = \int_{-\infty}^{\infty} h(t-\tau)x_1(\tau)d\tau. \qquad (4.3)$$

Substituting (4.2) into (4.1) gives

$$
\begin{aligned}
y_p(t) =\ & \gamma_1 \text{Re}\left\{ \sqrt{2}z_1(t)e^{j\omega_1 t} + \sqrt{2}z_2(t)e^{j\omega_2 t} \right\} \\
& + \gamma_2 \left(\text{Re}\left\{ \sqrt{2}z_1(t)e^{j\omega_1 t} + \sqrt{2}z_2(t)e^{j\omega_2 t} \right\} \right)^2 \\
& + \gamma_3 \left(\text{Re}\left\{ \sqrt{2}z_1(t)e^{j\omega_1 t} + \sqrt{2}z_2(t)e^{j\omega_2 t} \right\} \right)^3 + w_p(t) \\
=\ & \gamma_2|z_1(t)|^2 + \gamma_2|z_2(t)|^2 \\
& + \text{Re}\left\{ \sqrt{2}\left(\gamma_1 z_1(t) + \frac{3\gamma_3}{2}z_1(t)|z_1(t)|^2 + 3\gamma_3 z_1(t)|z_2(t)|^2 \right) e^{j\omega_1 t} \right\} \\
& + \text{Re}\left\{ \sqrt{2}\left(\gamma_1 z_2(t) + \frac{3\gamma_3}{2}z_2(t)|z_2(t)|^2 + 3\gamma_3 |z_1(t)|^2 z_2(t) \right) e^{j\omega_2 t} \right\} \\
& + \text{Re}\left\{ \sqrt{2}\left(\frac{\sqrt{2}\gamma_2}{2}[z_1(t)]^2 \right) e^{j2\omega_1 t} \right\} + \text{Re}\left\{ \sqrt{2}\left(\frac{\sqrt{2}\gamma_2}{2}[z_2(t)]^2 \right) e^{j2\omega_2 t} \right\} \\
& + \text{Re}\left\{ \sqrt{2}\left(\frac{\gamma_3}{2}[z_1(t)]^3 \right) e^{j3\omega_1 t} \right\} + \text{Re}\left\{ \sqrt{2}\left(\frac{\gamma_3}{2}[z_2(t)]^3 \right) e^{j3\omega_2 t} \right\} \\
& + \text{Re}\left\{ \sqrt{2}\left(\sqrt{2}\gamma_2 z_1(t)z_2(t) \right) e^{j(\omega_1+\omega_2)t} \right\}
\end{aligned}
$$

---

[3]We will discuss the presence of multiple blocker signals in Section 4.5, which turns out to be a direct extension of this simple case.

$$+ \operatorname{Re}\left\{ \sqrt{2} \left(\sqrt{2}\gamma_2 z_1(t) z_2^*(t) \right) e^{j(\omega_1 - \omega_2)t} \right\}$$

$$+ \operatorname{Re}\left\{ \sqrt{2} \left(\frac{3\gamma_3}{2} z_1(t)[z_2(t)]^2 \right) e^{j(\omega_1 + 2\omega_2)t} \right\}$$

$$+ \operatorname{Re}\left\{ \sqrt{2} \left(\frac{3\gamma_3}{2} z_1(t)[z_2^*(t)]^2 \right) e^{j(\omega_1 - 2\omega_2)t} \right\}$$

$$+ \operatorname{Re}\left\{ \sqrt{2} \left(\frac{3\gamma_3}{2} [z_1(t)]^2 z_2(t) \right) e^{j(2\omega_1 + \omega_2)t} \right\}$$

$$+ \operatorname{Re}\left\{ \sqrt{2} \left(\frac{3\gamma_3}{2} [z_1(t)]^2 z_2^*(t) \right) e^{j(2\omega_1 - \omega_2)t} \right\} + w_p(t).$$

The produced signal components at different frequencies are listed in Table 4.1. With proper down-conversion and low-pass filtering, the received baseband signal corresponding to the carrier frequency ω_1 is given by

$$y(t) = \gamma_1 z_1(t) + \frac{3\gamma_3}{2} z_1(t)|z_1(t)|^2 + 3\gamma_3 z_1(t)|z_2(t)|^2 + w(t),$$

where $w(t)$ is the additive Gaussian noise in the baseband. This shows that $y(t)$ is distorted by the third-order harmonics

$$\frac{3\gamma_3}{2} z_1(t)|z_1(t)|^2$$

and the cross-modulation term

$$3\gamma_3 z_1(t)|z_2(t)|^2.$$

If the amplitude of the desired signal is small, i.e., $|z_1(t)| \ll 1$, then the amplitude of $\frac{3\gamma_3}{2} z_1(t)|z_1(t)|^2$ is much smaller than that of $\gamma_1 z_1(t)$ and can be neglected. Its effect becomes significant only when the amplitude of $z_1(t)$ is close to 1, but can be mitigated by properly limiting the dynamic range of the input at little cost of loosing reception sensitivity. In a wideband SDR system, however, the cross modulation term $3\gamma_3 z_1(t)|z_2(t)|^2$ is more dangerous. In real radio environments,

Table 4.1: Table of the Signal Components Generated by the Nonlinearity Model (4.1).

Frequency	Signal Components
0	$\gamma_2 |z_1(t)|^2 + \gamma_2 |z_2(t)|^2$
ω_1	$\gamma_1 z_1(t) + \frac{3\gamma_3}{2} z_1(t) |z_1(t)|^2 + 3\gamma_3 z_1(t) |z_2(t)|^2$
ω_2	$\gamma_1 z_2(t) + \frac{3\gamma_3}{2} z_2(t) |z_2(t)|^2 + 3\gamma_3 |z_1(t)|^2 z_2(t)$
$2\omega_1$	$\frac{\sqrt{2}\gamma_2}{2} [z_1(t)]^2$
$2\omega_2$	$\frac{\sqrt{2}\gamma_2}{2} [z_2(t)]^2$
$3\omega_1$	$\frac{\gamma_3}{2} [z_1(t)]^3$
$3\omega_2$	$\frac{\gamma_3}{2} [z_2(t)]^3$
$\omega_1 + \omega_2$	$\sqrt{2}\gamma_2 z_1(t) z_2(t)$
$\omega_1 - \omega_2$	$\sqrt{2}\gamma_2 z_1(t) z_2^*(t)$
$\omega_1 + 2\omega_2$	$\frac{3\gamma_3}{2} z_1(t) [z_2(t)]^2$
$\omega_1 - 2\omega_2$	$\frac{3\gamma_3}{2} z_1(t) [z_2^*(t)]^2$
$2\omega_1 + \omega_2$	$\frac{3\gamma_3}{2} [z_1(t)]^2 z_2(t)$
$2\omega_1 - \omega_2$	$\frac{3\gamma_3}{2} [z_1(t)]^2 z_2^*(t)$

the power of the blocker signal can be as much as 60–70 dB more than that of the desired signal. Since the receiver front-end has to maintain a minimum sensitivity level for the desired signal $z_1(t)$, then the simultaneously acquired blocker signal can be quite large, making the interference term $3\gamma_3 z_1(t)|z_2(t)|^2$ comparable to the desired signal component $\gamma_1 z_1(t)$. To measure the effects quantitatively, the effective signal-to-noise ratio in $y(t)$ is computed as follows:

$$
\begin{aligned}
\text{SNR}_{\text{effective}} &= \frac{\mathbf{E}\left\{\gamma_1^2 |z_1(t)|^2\right\}}{\mathbf{E}\left\{\left|\frac{3\gamma_3}{2}z_1(t)|z_1(t)|^2 + 3\gamma_3 z_1(t)|z_2(t)|^2 + w(t)\right|^2\right\}} \\
&= \frac{\gamma_1^2 \mathbf{E}\left\{|z_1(t)|^2\right\}}{\left\{\frac{9\gamma_3^2}{4}\mathbf{E}\left\{|z_1(t)|^6\right\} + 9\gamma_3^2 \mathbf{E}\left\{|z_1(t)|^4\right\}\mathbf{E}\left\{|z_2(t)|^2\right\}\right.} \\
&\qquad\qquad \left. + 9\gamma_3^2 \mathbf{E}\left\{|z_1(t)|^2\right\}\mathbf{E}\left\{|z_2(t)|^4\right\} + \sigma_w^2\right\}
\end{aligned}
\tag{4.4}
$$

where $z_1(t)$, $z_2(t)$ and $w(t)$ are assumed to be zero-mean and independent of each other, and

$$
\sigma_w^2 = \mathbf{E}\left\{|w(t)|^2\right\}
$$

is the noise variance. In order to obtain a more explicit expression of $\text{SNR}_{\text{effective}}$, we consider the following two cases:

1.) *Uniform Distribution:* Assume that the real and imaginary parts of $z_1(t)$ are i.i.d. uniformly distributed with mean zero, and the same for $z_2(t)$. This assumption approximates the case that $z_1(t)$ and $z_2(t)$ are PAM or QAM modulated in a single-carrier system, i.e.,

$$
z_i(t) = \sum_{n=-\infty}^{\infty} S_{i,n} p_i(t - nT_s), \quad i = 1, 2,
$$

where $S_{i,n}$, $i = 1, 2$, are the PAM or QAM symbols and $p_i(t)$, $i = 1, 2$, are the responses of the pulse shaping filters. Let the variances of $z_1(t)$ and

114

$z_2(t)$ be $\sigma_{z,1}^2$ and $\sigma_{z,2}^2$, respectively. It can be shown that

$$\mathbf{E}\left\{|z_1(t)|^4\right\} = \frac{7}{5}\sigma_{z,1}^4, \quad \mathbf{E}\left\{|z_1(t)|^6\right\} = \frac{81}{35}\sigma_{z,1}^6, \tag{4.5}$$

$$\mathbf{E}\left\{|z_2(t)|^4\right\} = \frac{7}{5}\sigma_{z,2}^4, \quad \mathbf{E}\left\{|z_2(t)|^6\right\} = \frac{81}{35}\sigma_{z,2}^6.$$

Also, let

$$\mathrm{SNR}_0 = \frac{\gamma_1^2 \sigma_{z,1}^2}{\sigma_w^2}$$

be the effective signal-to-noise ratio in the absence of nonlinearity, i.e., $\gamma_2 = \gamma_3 = 0$. It then follows from (4.4) that

$$\mathrm{SNR}_{\mathrm{effective}} = \frac{\gamma_1^2 \sigma_{z,1}^2}{\frac{729}{140}\gamma_3^2\sigma_{z,1}^6 + \frac{63}{5}\gamma_3^2\sigma_{z,1}^4\sigma_{z,2}^2 + \frac{63}{5}\gamma_3^2\sigma_{z,1}^2\sigma_{z,2}^4 + \sigma_w^2}$$
$$= \frac{\mathrm{SNR}_0}{\frac{\gamma_3^2}{\gamma_1^2}\left(\frac{729}{140}\sigma_{z,1}^4 + \frac{63}{5}\sigma_{z,1}^2\sigma_{z,2}^2 + \frac{63}{5}\sigma_{z,2}^4\right)\mathrm{SNR}_0 + 1}. \tag{4.6}$$

2.) *Gaussian Distribution:* We now assume that $z_1(t)$ and $z_2(t)$ are circularly symmetric Gaussian distributed with mean zero and variances $\sigma_{z,1}^2$ and $\sigma_{z,2}^2$, respectively. This assumption approximates the scenario that $z_1(t)$ is an OFDM signal and $z_2(t)$ is a Gaussian interference. In this case,

$$\mathbf{E}\left\{|z_1(t)|^4\right\} = 2\sigma_{z,1}^4, \quad \mathbf{E}\left\{|z_1(t)|^6\right\} = 6\sigma_{z,1}^6, \tag{4.7}$$

$$\mathbf{E}\left\{|z_2(t)|^4\right\} = 2\sigma_{z,2}^4, \quad \mathbf{E}\left\{|z_2(t)|^6\right\} = 6\sigma_{z,2}^6,$$

and hence

$$\mathrm{SNR}_{\mathrm{effective}} = \frac{\mathrm{SNR}_0}{\frac{\gamma_3^2}{\gamma_1^2}\left(\frac{27}{2}\sigma_{z,1}^4 + 18\sigma_{z,1}^2\sigma_{z,2}^2 + 18\sigma_{z,2}^4\right)\mathrm{SNR}_0 + 1}. \tag{4.8}$$

Fig. 4.5 plots $\mathrm{SNR}_{\mathrm{effective}}$ vs. SNR_0 when $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$. It is seen that when SNR_0 is low, e.g., less than 0 dB, $\mathrm{SNR}_{\mathrm{effective}}$ is dominated by the additive noise, and hence, is approximately
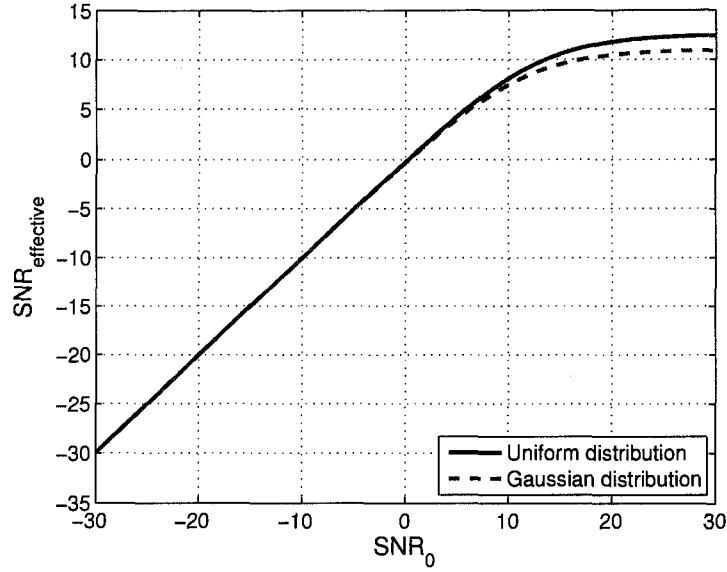
Figure 4.5: Plot of $\mathrm{SNR}_{\mathrm{effective}}$ vs. SNR_0 by (4.6) and (4.8) for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$.

equal to SNR_0. If SNR_0 is high, $\mathrm{SNR}_{\mathrm{effective}}$ is dominated by cross modulation, which saturates at 12.5 dB for the uniform distribution case and at 11.0 dB for the Gaussian distribution case. If the power of $z_2(t)$ changes to $\sigma_{z,2}^2 = 10^{-5}$ (17.0 dB less) and the other conditions remain the same, Fig. 4.6 shows that the resulting $\mathrm{SNR}_{\mathrm{effective}}$ is almost equal to SNR_0 and the nonlinearity causes no performance degradation. This observation demonstrates that a strong blocker signal can cause significant distortion to the desired signal. To overcome this problem, however, it is not feasible to limit the amplitude of the input signal $v(t)$ such that $z_2(t)$ is small, because this will also weaken the desired signal and, consequently, reduce the reception sensitivity. In the next section, we propose a compensation scheme to mitigate this effect by using digital signal processing techniques.
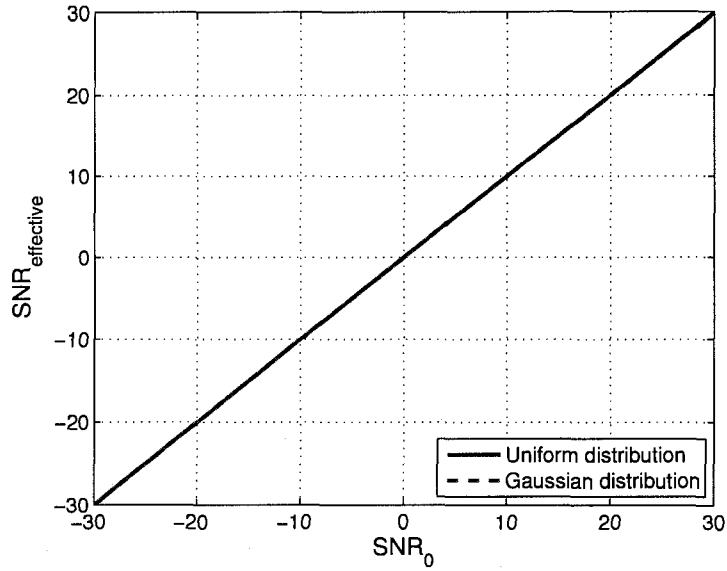
Figure 4.6: Plot of $SNR_{effective}$ vs. SNR_0 by (4.6) and (4.8) for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 10^{-5}$.

4.3 Proposed Compensation Scheme

In the proposed scheme, the SDR uses two separate RF signal paths. One path is used to capture the signal in the desired band, while the other path is used to acquire the blocker signal, as illustrated by Fig. 4.7. The two-channel signals are jointly processed in the baseband to alleviate the nonlinear effect. There are two stages in this scheme. In the first stage, the SDR exploits the pilot sequence in the desired signal to estimate the channel response and the nonlinearity parameters. These estimates are then used in the second stage to recover the transmitted data symbols. In the following discussion, we focus on the most problematic case of $\sigma_{z,1}^2 \ll \sigma_{z,2}^2$.

Before the channel and parameter estimation procedure, a robust synchronization procedure is performed in the desired channel for timing recovery. It is crucial for the receiver to be able to correctly sample each pilot or data symbol.
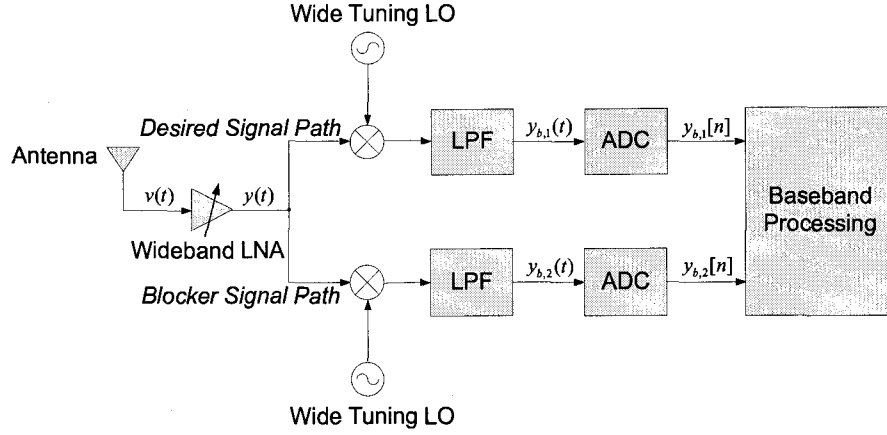
Figure 4.7: A software-defined radio with two signal paths: One is used to capture the signal in the desired band, while the other is used to acquire the blocker signal.

The length of the synchronization sequence sent by the transmitter depends on the effective signal-to-noise ratio in the worst scenario. The lower the signal-to-noise ratio, the longer the training sequence is required [Men97]. In the presence of cross modulation, the effective signal-to-noise ratio is given by (4.6) and (4.8). This degradation thus requires a longer-than-normal training sequence for the system to operate in an adverse environment. After successful synchronization and sampling, we obtain the discrete-time version of the received baseband signals at the carrier frequencies ω_1 and ω_2:

$$y_1[n] = \gamma_1 z_1[n] + \frac{3\gamma_3}{2} z_1[n]|z_1[n]|^2 + 3\gamma_3 z_1[n]|z_2[n]|^2 + w_1[n], \qquad (4.9)$$

$$y_2[n] = \gamma_1' z_2[n] + \frac{3\gamma_3'}{2} z_2[n]|z_2[n]|^2 + 3\gamma_3'|z_1[n]|^2 z_2[n] + w_2[n], \qquad (4.10)$$

where γ_1, γ_3 are the model parameters associated with the signal path of $y_1(t)$, and γ_1', γ_3' are the model parameters associated with the signal path of $y_2(t)$. Recall that in (4.3), $z_1(t)$ is given by the convolution of the baseband signal $x_1(t)$ and the continuous-time channel impulse response function $h(t)$. In the

discrete-time domain, we have

$$z_1[n] = \sum_{l=0}^{L-1} x_1[n-l]h[l], \tag{4.11}$$

where L is the length of $h[n]$, i.e.,

$$h[n] = 0 \text{ if } n \notin \{0, 1, \ldots, L-1\}.$$

In $y_1[n]$, since $|z_1[n]| \ll 1$, the third-order harmonics $\frac{3\gamma_3}{2}z_1[n]|z_1[n]|^2$ is negligible compared to the desired signal component $\gamma_1 z_1[n]$ and the cross-modulation term $3\gamma_3 z_1[n]|z_2[n]|^2$. Hence,

$$y_1[n] \approx \gamma_1 z_1[n] + 3\gamma_3 z_1[n]|z_2[n]|^2 + w_1[n]. \tag{4.12}$$

Since $\sigma_{z,1}^2 \ll \sigma_{z,2}^2$ and $|z_2[n]| \ll 1$, the secondary-path signal $y_2(t)$ is dominated by $\gamma_1' z_2(t)$, i.e.,

$$y_2[n] \approx \gamma_1' z_2[n] + w_2[n], \tag{4.13}$$

where $\frac{3\gamma_3'}{2}z_2[n]|z_2[n]|^2$ and $3\gamma_3'|z_1[n]|^2 z_2[n]$ are negligible compared to $\gamma_1' z_2[n]$.

4.3.1 Channel and Nonlinearity Parameter Estimation

In this stage, the receiver utilizes the pilot symbols transmitted along with the desired signal to estimate the channel response and the nonlinearity parameters. Thus, $x_1[n]$, $n = 0, 1, \ldots, N-1$, are known to the receiver, where N is the length of the pilot sequence and $N > L$. By (4.12), γ_1, γ_3 and $h[n]$, $n = 0, 1, \ldots, L-1$, can be estimated by solving the following optimization problem:

$$\min_{\gamma_1, \gamma_3, h[n]} \sum_{n=0}^{N-1} \left| y_1[n] - \gamma_1 z_1[n] - 3\gamma_3 z_1[n]|\widehat{z}_2[n]|^2 \right|^2$$

where $z_1[n]$ is related to $x_1[n]$ and $h[n]$ through (4.11) and $\widehat{z}_2[n]$ is given by

$$\widehat{z}_2[n] = \frac{1}{\gamma_1'}\, y_2[n]$$

according to (4.13). In this formulation, we have the product of γ_1 and $z_1[n]$ and the product of γ_3 and $z_1[n]|\widehat{z}_2[n]|^2$, which causes an ambiguity of a scaling factor in the estimate of γ_1, γ_3 and $h[n]$. To resolve this ambiguity, we estimate the following parameters instead:

$$\gamma_3'' = \frac{\gamma_3}{\gamma_1\gamma_1'^2}$$

and

$$h''[n] = \gamma_1 h[n], \quad n = 0, 1, \ldots, L-1.$$

The original problem thus becomes

$$\min_{\gamma_3'', h''[n]} \sum_{n=0}^{N-1} \left| \, y_1[n] - z_1''[n] - 3\gamma_3'' z_1''[n]|y_2[n]|^2 \, \right|^2,$$

where

$$z_1''[n] = \sum_{l=0}^{L-1} x_1[n-l]h''[l]. \tag{4.14}$$

This problem is nonlinear and nonconvex. For every *fixed* γ_3'', the associated optimal $h[n]$ can be obtained by solving

$$\min_{h''[n]} \sum_{n=0}^{N-1} \left| \, y_1[n] - z_1''[n] - 3\gamma_3'' z_1''[n]|y_2[n]|^2 \, \right|^2$$

$$= \min_{h''[n]} \sum_{n=0}^{N-1} \left| y_1[n] - \left(1 + 3\gamma_3''|y_2[n]|^2\right) \left(\sum_{l=0}^{L-1} x_1[n-l]h''[l] \right) \right|^2,$$

which can be formulated as a linear least-squares problem [Say03]:

$$\min_{\mathbf{h}} \| \, \mathbf{y} - \mathbf{A X h} \, \|^2, \tag{4.15}$$

where $\mathbf{y}$, $\mathbf{h}$, $\mathbf{A}$ and $\mathbf{X}$ are defined as follows:

$$\mathbf{y} = \begin{bmatrix} y_1[0] \\ y_1[1] \\ \vdots \\ y_1[N-1] \end{bmatrix} (N \times 1), \quad \mathbf{h} = \begin{bmatrix} h''[0] \\ h''[1] \\ \vdots \\ h''[L-1] \end{bmatrix} (L \times 1),$$

$$\mathbf{A} = \begin{bmatrix} 1 + 3\gamma_3''|y_2[0]|^2 & 0 & \cdots & 0 \\ 0 & 1 + 3\gamma_3''|y_2[1]|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 + 3\gamma_3''|y_2[N-1]|^2 \end{bmatrix} (N \times N),$$

$$\mathbf{X} = \begin{bmatrix} x_1[0] & 0 & \cdots & 0 \\ x_1[1] & x_1[0] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_1[n+L-1] & x_1[n+L-2] & \cdots & x_1[n] \\ x_1[n+L] & x_1[n+L-1] & \cdots & x_1[n+1] \\ \vdots & \vdots & \ddots & \vdots \\ x_1[N-2] & x_1[N-3] & \cdots & x_1[N-L-1] \\ x_1[N-1] & x_1[N-2] & \cdots & x_1[N-L] \end{bmatrix} (N \times L).$$

The closed-form solution of (4.15) is

$$\mathbf{h}_o = (\mathbf{X}^*\mathbf{A}^*\mathbf{A}\mathbf{X})^{-1}\mathbf{X}^*\mathbf{A}^*\mathbf{y},$$

and its associated residual error is

$$\|\mathbf{y} - \mathbf{A}\mathbf{X}\mathbf{h}_o\|^2 = \mathbf{y}^*\mathbf{y} - \mathbf{y}^*\mathbf{A}\mathbf{X}(\mathbf{X}^*\mathbf{A}^*\mathbf{A}\mathbf{X})^{-1}\mathbf{X}^*\mathbf{A}^*\mathbf{y}.$$

Note that the above residual error is a function of γ_3'' because $\mathbf{A}$ depends on γ_3''. Let

$$f(\gamma_3'') = \mathbf{y}^*\mathbf{y} - \mathbf{y}^*\mathbf{A}\mathbf{X}(\mathbf{X}^*\mathbf{A}^*\mathbf{A}\mathbf{X})^{-1}\mathbf{X}^*\mathbf{A}^*\mathbf{y}.$$

Then a one-dimensional search is conducted to find the optimal γ_3'' that minimizes $f(\gamma_3'')$. That is, the estimate of γ_3'' is given by

$$\widehat{\gamma}_3'' = \arg\min_{\gamma_3''} f(\gamma_3'') = \arg\min_{\gamma_3''} \left\{ \mathbf{y}^*\mathbf{y} - \mathbf{y}^*\mathbf{A}\mathbf{X}(\mathbf{X}^*\mathbf{A}^*\mathbf{A}\mathbf{X})^{-1}\mathbf{X}^*\mathbf{A}^*\mathbf{y} \right\}.$$

The optimal $\mathbf{h}_o$ associated with $\widehat{\gamma}_3''$ gives an estimate of $h''[n]$, $n = 0, 1, \ldots, L-1$, that is denoted by $\widehat{h}''[n]$, $n = 0, 1, \ldots, L-1$. The obtained $\widehat{\gamma}_3''$ and $\widehat{h}''[n]$, $n = 0, 1, \ldots, L-1$, are used in the data transmission stage to recover data symbols. In the next subsection, we assume that γ_3'' and $h''[n]$ are known to be $\widehat{\gamma}_3''$ and $\widehat{h}''[n]$.

4.3.2 Data Symbol Estimation

In this stage, data symbols from a known constellation are transmitted over the channel. Since

$$y_1[n] \approx z_1''[n] + 3\widehat{\gamma}_3'' z_1''[n]|y_2[n]|^2 + w_1[n],$$

$z_1''[n]$ is estimated by

$$\widehat{z}_1''[n] = \frac{y_1[n]}{1 + 3\widehat{\gamma}_3''|y_2[n]|^2}, \quad n = 0, 1, \ldots, M-1,$$

where M is the length of the data symbol block. It then follows from (4.14) that the estimate of $x_1[n]$ is given by

$$\widehat{x}_1[n] = \frac{1}{\widehat{h}''[0]} \left[\widehat{z}_1''[n] - \sum_{l=1}^{L-1} \widehat{x}_1[n-l]\widehat{h}''[l] \right],$$

where $\widehat{x}_1[n-l]$, $l = 1, 2, \ldots, L-1$, are the previously estimated data symbols. This is similar to a decision-directed method. We can also recover $x_1[n]$ from $\widehat{z}_1''[n]$ by using a channel equalizer. If $x_1[n]$ are the symbols from a known constellation like QPSK or 16-QAM, the Viterbi algorithm can be exploited to obtain a more accurate estimate of $x_1[n]$ [TV05].

To gain an interpretation about the performance of the proposed algorithm, the data model can be rewritten as

$$y_1[n] = \left(\gamma_1 + 3\gamma_3|z_2[n]|^2\right) z_1[n] + \frac{3\gamma_3}{2} z_1[n]|z_1[n]|^2 + w_1[n],$$

where

$$\left(\gamma_1 + 3\gamma_3|z_2[n]|^2\right) z_1[n]$$

is the desired signal component[4] and

$$\frac{3\gamma_3}{2} z_1[n]|z_1[n]|^2 + w_1[n]$$

is the noise/interference term. The effective signal-to-noise ratio after compensation is approximately equal to

$$\begin{aligned}
\mathrm{SNR}_{\text{effective}} &\approx \frac{\mathbf{E}\left\{\left|\left(\gamma_1 + 3\gamma_3|z_2[n]|^2\right) z_1[n]\right|^2\right\}}{\mathbf{E}\left\{\left|\frac{3\gamma_3}{2} z_1[n]|z_1[n]|^2 + w_1[n]\right|^2\right\}} \\
&= \frac{\left(\gamma_1^2 + 6\gamma_1\gamma_3\mathbf{E}\left\{|z_2[n]|^2\right\} + 9\gamma_3^2\mathbf{E}\left\{|z_2[n]|^4\right\}\right)\mathbf{E}\left\{|z_1[n]|^2\right\}}{\frac{9}{4}\gamma_3^2\mathbf{E}\left\{|z_1[n]|^6\right\} + \sigma_w^2}.
\end{aligned}$$

If both $z_1[n]$ and $z_2[n]$ are uniformly distributed, by (4.5) we have

$$\begin{aligned}
\mathrm{SNR}_{\text{effective}} &= \frac{\left(\gamma_1^2 + 6\gamma_1\gamma_3\sigma_{z,2}^2 + \frac{63}{5}\gamma_3^2\sigma_{z,2}^4\right)\sigma_{z,1}^2}{\frac{729}{140}\gamma_3^2\sigma_{z,1}^6 + \sigma_w^2} \\
&= \frac{\left(1 + \frac{6\gamma_3}{\gamma_1}\sigma_{z,2}^2 + \frac{63\gamma_3^2}{5\gamma_1^2}\sigma_{z,2}^4\right)\mathrm{SNR}_0}{\frac{729\gamma_3^2}{140\gamma_1^2}\sigma_{z,1}^4\mathrm{SNR}_0 + 1}.
\end{aligned} \tag{4.16}$$

---

[4]It assumes that $z_2[n]$ can be ideally estimated and compensated for.
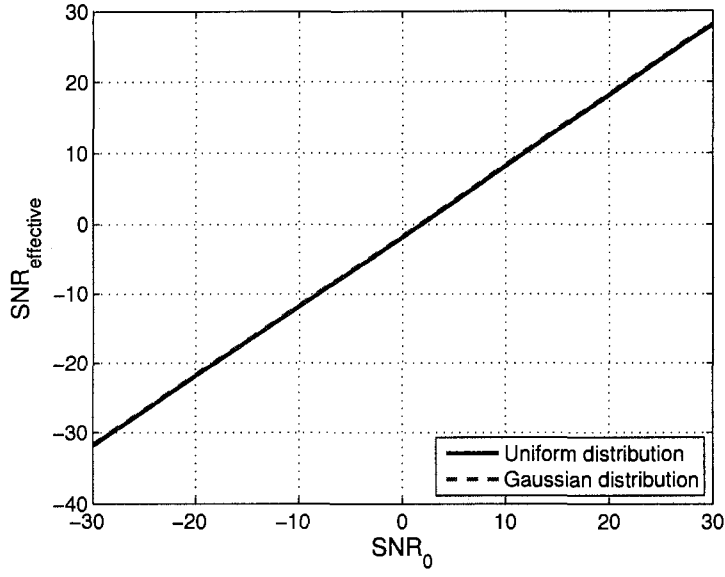
Figure 4.8: Plot of $\text{SNR}_{\text{effective}}$ after ideal compensation vs. SNR_0 by (4.16) and (4.17) for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$.

Also, if both $z_1[n]$ and $z_2[n]$ are Gaussian distributed, by (4.7) we have

$$\text{SNR}_{\text{effective}} = \frac{\left(\gamma_1^2 + 6\gamma_1\gamma_3\sigma_{z,2}^2 + 18\gamma_3^2\sigma_{z,2}^4\right)\sigma_{z,1}^2}{\frac{27}{2}\gamma_3^2\sigma_{z,1}^6 + \sigma_w^2}$$

$$= \frac{\left(1 + \frac{6\gamma_3}{\gamma_1}\sigma_{z,2}^2 + \frac{18\gamma_3^2}{\gamma_1^2}\sigma_{z,2}^4\right)\text{SNR}_0}{\frac{27\gamma_3^2}{2\gamma_1^2}\sigma_{z,1}^4\text{SNR}_0 + 1}. \tag{4.17}$$

Fig. 4.8 shows the plot of $\text{SNR}_{\text{effective}}$ vs. SNR_0 after ideal compensation for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$. Compared to Fig. 4.5, it demonstrates that the compensation technique can achieve significant improvement. In the next section, we compute the Cramer-Rao lower bounds for the channel and data symbol estimation errors, which serve as a benchmark for evaluating the performance of the proposed algorithm.

4.4 Performance Bounds

To evaluate the proposed algorithm, we compare its performance with the Cramer-Rao lower bound (CRLB) that gives a lower bound on the covariance matrix of any unbiased estimator of unknown parameters [Kay93]. Consider again a generic data model

$$\mathbf{y} = \mathbf{s}_{\boldsymbol{\theta}} + \mathbf{w}, \tag{4.18}$$

where $\mathbf{y}$ is the observed data vector with length N, $\mathbf{s}_{\boldsymbol{\theta}}$ is the noise-free data vector that depends on the parameter vector $\boldsymbol{\theta}$, and $\mathbf{w}$ is the vector of circularly symmetric Gaussian noise with covariance matrix

$$\mathbf{E}\{\mathbf{w}\mathbf{w}^*\} = \begin{bmatrix} \sigma_{w,0}^2 & 0 & \dots & 0 \\ 0 & \sigma_{w,1}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{w,N-1}^2 \end{bmatrix}.$$

Let

$$\mathbf{s}_{\boldsymbol{\theta}} = \begin{bmatrix} s_{\boldsymbol{\theta}}[0] & s_{\boldsymbol{\theta}}[1] & \dots & s_{\boldsymbol{\theta}}[N-1] \end{bmatrix}^T.$$

The Fisher information matrix for this data model is given by (see Appendix G for a derivation)

$$\mathbf{I}_{\boldsymbol{\theta}} = \sum_{n=0}^{N-1} \frac{2}{\sigma_{w,n}^2} \mathrm{Re}\left\{ \frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \boldsymbol{\theta}} \left(\frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \boldsymbol{\theta}} \right)^* \right\}, \tag{4.19}$$

where

$$\frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \theta_1} & \frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \theta_2} & \dots & \frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \theta_{|\boldsymbol{\theta}|}} \end{bmatrix}^T$$

and $|\boldsymbol{\theta}|$ is the dimension of $\boldsymbol{\theta}$. By the CRLB, any unbiased estimator $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ has a covariance matrix that satisfies

$$\mathrm{var}\{\widehat{\boldsymbol{\theta}}\} = \mathbf{E}\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^*\} \geq \mathbf{I}_{\boldsymbol{\theta}}^{-1}. \tag{4.20}$$

4.4.1 CRLB for Estimating Channel Response

The data model described by (4.9) and (4.10) is equivalent to

$$y_1[n] = z_1''[n] + 3\gamma_3'' z_1''[n]|z_2''[n]|^2 + \frac{3\gamma_3}{2}z_1[n]|z_1[n]|^2 + w_1[n], \qquad (4.21)$$

$$y_2[n] = z_2''[n] + \frac{3\gamma_3'}{2}z_2[n]|z_2[n]|^2 + 3\gamma_3'|z_1[n]|^2 z_2[n] + w_2[n], \qquad (4.22)$$

where $z_2''[n] = \gamma_1' z_2[n]$. The model can be expressed as

$$y_1[n] = s_{\boldsymbol{\theta},1}[n] + w_1'[n],$$

$$y_2[n] = s_{\boldsymbol{\theta},2}[n] + w_2'[n],$$

where the noise-free signal components $s_{\boldsymbol{\theta},1}[n]$ and $s_{\boldsymbol{\theta},2}[n]$ are given by

$$s_{\boldsymbol{\theta},1}[n] = z_1''[n] + 3\gamma_3'' z_1''[n]|z_2''[n]|^2,$$

$$s_{\boldsymbol{\theta},2}[n] = z_2''[n],$$

and the noise components $w_1'[n]$ and $w_2'[n]$ are given by

$$w_1'[n] = \frac{3\gamma_3}{2}z_1[n]|z_1[n]|^2 + w_1[n],$$

$$w_2'[n] = \frac{3\gamma_3'}{2}z_2[n]|z_2[n]|^2 + 3\gamma_3'|z_1[n]|^2 z_2[n] + w_2[n].$$

We assume that $w_1'[n]$ and $w_2'[n]$ are approximately Gaussian distributed with variances

$$\begin{aligned}
\sigma_{w_1'}^2 &= \mathbf{E}\left\{|w_1'[n]|^2\right\} \\
&= \mathbf{E}\left\{\left|\frac{3\gamma_3}{2}z_1[n]|z_1[n]|^2 + w_1[n]\right|^2\right\} \\
&= \frac{9\gamma_3^2}{4}\mathbf{E}\left\{|z_1[n]|^6\right\} + \sigma_w^2 \\
&= \begin{cases} \frac{729\gamma_3^2}{140}\sigma_{z,1}^6 + \sigma_w^2 & \text{(uniform distri.)}, \\ \frac{27\gamma_3^2}{2}\sigma_{z,1}^6 + \sigma_w^2 & \text{(Gaussian distri.)}, \end{cases}
\end{aligned}$$

126

and

$$\sigma_{w_2'}^2 = \mathbf{E}\left\{|w_2'[n]|^2\right\}$$

$$= \mathbf{E}\left\{\left|\frac{3\gamma_3'}{2}z_2[n]|z_2[n]|^2 + 3\gamma_3'|z_1[n]|^2 z_2[n] + w_2[n]\right|^2\right\}$$

$$= \frac{9}{4}(\gamma_3')^2 \mathbf{E}\left\{|z_2[n]|^6\right\} + 9(\gamma_3')^2 \mathbf{E}\left\{|z_1[n]|^4\right\}\mathbf{E}\left\{|z_2[n]|^2\right\}$$

$$+ 9(\gamma_3')^2 \mathbf{E}\left\{|z_1[n]|^2\right\}\mathbf{E}\left\{|z_2[n]|^4\right\} + \sigma_w^2$$

$$= \begin{cases} \frac{729(\gamma_3')^2}{140}\sigma_{z,2}^6 + \frac{63(\gamma_3')^2}{5}\sigma_{z,1}^4\sigma_{z,2}^2 + \frac{63(\gamma_3')^2}{5}\sigma_{z,1}^2\sigma_{z,2}^4 + \sigma_w^2 & \text{(uniform distri.)}, \\ \frac{27(\gamma_3')^2}{2}\sigma_{z,2}^6 + 18(\gamma_3')^2\sigma_{z,1}^4\sigma_{z,2}^2 + 18(\gamma_3')^2\sigma_{z,1}^2\sigma_{z,2}^4 + \sigma_w^2 & \text{(Gaussian distri.)}. \end{cases}$$

The unknown parameter vector to be estimated is

$$\boldsymbol{\theta} = \left[\begin{array}{ccccc} \gamma_3'' & \mathrm{Re}\{h''[0]\} & \mathrm{Im}\{h''[0]\} & \ldots & \mathrm{Re}\{h''[L-1]\} & \mathrm{Im}\{h''[L-1]\} \end{array}\right.$$

$$\left.\begin{array}{ccccc} \mathrm{Re}\{z_2''[0]\} & \mathrm{Im}\{z_2''[0]\} & \ldots & \mathrm{Re}\{z_2''[N-1]\} & \mathrm{Im}\{z_2''[N-1]\} \end{array}\right]^T.$$

By (4.19), the Fisher information matrix is given by

$$\mathbf{I_\theta} = \sum_{m=1}^{2}\sum_{n=0}^{N-1}\frac{2}{\sigma_{w_m'}^2}\mathrm{Re}\left\{\frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\boldsymbol{\theta}}\left(\frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\boldsymbol{\theta}}\right)^*\right\}, \qquad (4.23)$$

where

$$\frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\boldsymbol{\theta}} = \left[\begin{array}{cccccc} \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\gamma_3''} & \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Re}\{h''[0]\}} & \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Im}\{h''[0]\}} & \cdots & \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Re}\{h''[L-1]\}} & \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Im}\{h''[L-1]\}} \end{array}\right.$$

$$\left.\begin{array}{ccccc} \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Re}\{z_2''[0]\}} & \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Im}\{z_2''[0]\}} & \cdots & \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Re}\{z_2''[N-1]\}} & \frac{\partial s_{\boldsymbol{\theta},m}[n]}{\partial\mathrm{Im}\{z_2''[N-1]\}} \end{array}\right]^T$$

for $m = 1, 2$. It can be shown that for $n = 0, 1, \ldots, N-1$,

$$\frac{\partial s_{\boldsymbol{\theta},1}[n]}{\partial\gamma_3''} = 3z_1''[n]|z_2''[n]|^2,$$

$$\frac{\partial s_{\boldsymbol{\theta},1}[n]}{\partial\mathrm{Re}\{h''[l]\}} = x_1[n-l] + 3\gamma_3''x_1[n-l]|z_2''[n]|^2,$$

$$\frac{\partial s_{\boldsymbol{\theta},1}[n]}{\partial\mathrm{Im}\{h''[l]\}} = jx_1[n-l] + j3\gamma_3''x_1[n-l]|z_2''[n]|^2,$$

$$\frac{\partial s_{\theta,1}[n]}{\partial \mathrm{Re}\left\{z_2''[n']\right\}} = \begin{cases} 3\gamma_3'' z_1''[n'] \left[z_2''[n'] + (z_2''[n'])^*\right], & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases}$$

$$\frac{\partial s_{\theta,1}[n]}{\partial \mathrm{Im}\left\{z_2''[n']\right\}} = \begin{cases} j3\gamma_3'' z_1''[n'] \left[-z_2''[n'] + (z_2''[n'])^*\right], & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases}$$

and

$$\frac{\partial s_{\theta,2}[n]}{\partial \gamma_3''} = 0, \quad \frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Re}\left\{h''[l]\right\}} = 0, \quad \frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Im}\left\{h''[l]\right\}} = 0,$$

$$\frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Re}\left\{z_2''[n']\right\}} = \begin{cases} 1, & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases} \qquad \frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Im}\left\{z_2''[n']\right\}} = \begin{cases} j, & \text{if } n = n', \\ 0, & \text{if } n \neq n'. \end{cases}$$

For each particular γ_1, γ_3, $h[n]$, $n = 0, 1, \ldots, L-1$, and training sequence $x_1[n]$, $n = 0, 1, \ldots, N-1$, the associated Fisher information matrix and the CRLB can be computed by using (4.23) and (4.20), respectively. An average is then taken over the ensemble of all possible channel realizations to get the average CRLB for the estimation errors.

4.4.2 CRLB for Estimating Data Symbols

In this case, the data model is still given by (4.21) and (4.22), but the unknown parameter vector to be estimated changes to

$$\boldsymbol{\theta} = \Big[\ \mathrm{Re}\left\{z_1''[0]\right\} \quad \mathrm{Im}\left\{z_1''[0]\right\} \quad \ldots \quad \mathrm{Re}\left\{z_1''[M-1]\right\} \quad \mathrm{Im}\left\{z_1''[M-1]\right\}$$
$$\mathrm{Re}\left\{z_2''[0]\right\} \quad \mathrm{Im}\left\{z_2''[0]\right\} \quad \ldots \quad \mathrm{Re}\left\{z_2''[M-1]\right\} \quad \mathrm{Im}\left\{z_2''[M-1]\right\} \ \Big]^T .$$

The Fisher information matrix is given by

$$\mathbf{I}_{\boldsymbol{\theta}} = \sum_{m=1}^{2} \sum_{n=0}^{N-1} \frac{2}{\sigma_{w_m'}^2} \mathrm{Re}\left\{ \frac{\partial s_{\theta,m}[n]}{\partial \boldsymbol{\theta}} \left(\frac{\partial s_{\theta,m}[n]}{\partial \boldsymbol{\theta}} \right)^* \right\},$$

where $\frac{\partial s_{\theta,m}[n]}{\partial \theta}$ is now defined as

$$\frac{\partial s_{\theta,m}[n]}{\partial \theta} = \left[\begin{array}{ccccccc} \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Re}\{z_1''[0]\}} & \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Im}\{z_1''[0]\}} & \cdots & \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Re}\{z_1''[M-1]\}} & \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Im}\{z_1''[M-1]\}} & \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Re}\{z_2''[0]\}} \end{array} \right.$$
$$\left. \begin{array}{cccc} \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Im}\{z_2''[0]\}} & \cdots & \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Re}\{z_2''[M-1]\}} & \frac{\partial s_{\theta,m}[n]}{\partial \mathrm{Im}\{z_2''[M-1]\}} \end{array} \right]^T, \quad m = 1, 2.$$

For $n = 0, 1, \ldots, M-1$, we have

$$\frac{\partial s_{\theta,1}[n]}{\partial \mathrm{Re}\{z_1''[n']\}} = \begin{cases} 1 + 3\gamma_3'' |z_2''[n]|^2, & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases}$$

$$\frac{\partial s_{\theta,1}[n]}{\partial \mathrm{Im}\{z_1''[n']\}} = \begin{cases} j + j3\gamma_3'' |z_2''[n]|^2, & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases}$$

$$\frac{\partial s_{\theta,1}[n]}{\partial \mathrm{Re}\{z_2''[n']\}} = \begin{cases} 3\gamma_3'' z_1''[n'] \left(z_2''[n'] + (z_2''[n'])^* \right), & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases}$$

$$\frac{\partial s_{\theta,1}[n]}{\partial \mathrm{Im}\{z_2''[n']\}} = \begin{cases} j3\gamma_3'' z_1''[n'] \left(-z_2''[n'] + (z_2''[n'])^* \right), & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases}$$

and

$$\frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Re}\{z_1''[n']\}} = 0, \quad \frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Im}\{z_1''[n']\}} = 0,$$

$$\frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Re}\{z_2''[n']\}} = \begin{cases} 1, & \text{if } n = n', \\ 0, & \text{if } n \neq n', \end{cases} \quad \frac{\partial s_{\theta,2}[n]}{\partial \mathrm{Im}\{z_2''[n']\}} = \begin{cases} j, & \text{if } n = n', \\ 0, & \text{if } n \neq n'. \end{cases}$$

In Section 4.6, we compare the simulated estimation errors with the computed CRLB, and show that the CRLB provides a good theoretical measure of the estimation accuracy. The next section extends the current discussion to the cases of multiple blocker signals and OFDM modulated transmission.

4.5 Some Extensions

4.5.1 Multiple Blocker Signals

In the presence of multiple blocker signals, the received passband signal is represented by

$$v(t) = \text{Re} \left\{ \sum_{k=1}^{K} \sqrt{2} z_k(t) e^{j\omega_k t} \right\},$$

where $z_1(t)$ is the desired signal and $z_k(t)$, $k = 2, 3, \ldots, K$, are the blocker signals with

$$\mathbf{E} \left\{ |z_1(t)|^2 \right\} \ll \mathbf{E} \left\{ |z_k(t)|^2 \right\}, \quad k = 2, 3, \ldots, K.$$

Assume that the SDR has K RF paths with each dedicated to one of the carrier frequencies ω_k, $k = 1, 2, \ldots, K$. Similar to (4.9) and (4.10), it can be shown that the received baseband signals are

$$y_1[n] = \gamma_1 z_1[n] + \frac{3\gamma_3}{2} z_1[n]|z_1[n]|^2 + 3\gamma_3 \sum_{k'=2}^{K} z_1[n]|z_{k'}[n]|^2 + w_1[n],$$

and

$$y_k[n] = \gamma_{1,k}' z_k[n] + \frac{3\gamma_{3,k}'}{2} z_k[n]|z_k[n]|^2 + 3\gamma_{3,k}' \sum_{k'=1, \, k' \neq k}^{K} z_k[n]|z_{k'}[n]|^2 + w_k[n],$$

$$k = 2, 3, \ldots, K,$$

where $\gamma_{1,k}'$ and $\gamma_{3,k}'$ are the nonlinear model parameters associated with the signal path of $y_k(t)$, $k = 2, 3, \ldots, K$.

To estimate the model parameters and the channel response in the training stage, we use the following approximations:

$$y_1[n] \approx \gamma_1 z_1[n] + 3\gamma_3 \sum_{k'=2}^{K} z_1[n]|z_{k'}[n]|^2 + w_1[n], \tag{4.24}$$

$$y_k[n] \approx \gamma_{1,k}' z_k[n] + w_k[n], \quad k = 2, 3, \ldots, K. \tag{4.25}$$

Expression (4.25) leads to

$$z_k[n] \approx \frac{1}{\gamma'_{1,k}} y_k[n], \quad k = 2, 3, \ldots, K.$$

By substitution, (4.24) becomes

$$y_1[n] \approx \gamma_1 z_1[n] + 3 \sum_{k'=2}^{K} \frac{\gamma_3}{\gamma'^2_{1,k'}} z_1[n] |y_{k'}[n]|^2 + w_1[n].$$

Let $z''_1[n] = \gamma_1 z_1[n]$ and $\gamma''_{3,k} = \frac{\gamma_3}{\gamma_1 \gamma'^2_{1,k}}$. Hence,

$$y_1[n] \approx z''_1[n] + 3 \sum_{k'=2}^{K} \gamma''_{3,k'} z''_1[n] |y_{k'}[n]|^2 + w_1[n].$$

The same technique presented in Section 4.3.1 can be used to estimate $\gamma''_{3,k}$ and $h''[n]$, $n = 0, 1, \ldots, L - 1$, where the matrix $\mathbf{A}$ is now given by

$$\mathbf{A} = \begin{bmatrix} 1 + 3 \sum_{k'=2}^{K} \gamma''_{3,k'} |y_{k'}[0]|^2 & & \\ & \ddots & \\ & & 1 + 3 \sum_{k'=2}^{K} \gamma''_{3,k'} |y_{k'}[N-1]|^2 \end{bmatrix}.$$

In the data transmission stage, $z''_1[n]$ can be similarly estimated by

$$\widehat{z''_1}[n] = \frac{y_1[n]}{1 + 3 \sum_{k'=2}^{K} \widehat{\gamma''_{3,k'}} |y_{k'}[n]|^2}.$$

4.5.2 OFDM Systems

In OFDM systems, the pilot and data symbols are transmitted in the frequency domain. At the transmitter, the bits from information sources are first mapped into constellation symbols, and then converted into a block of N symbols $X_1[k]$, $k = 0, 1, \ldots, N - 1$, by a serial-to-parallel converter. The N symbols are the frequency components to be transmitted using the N subcarriers of the OFDM modulator, and are converted to OFDM symbols $x_1[n]$, $n = 0, 1, \ldots, N - 1$, by

the unitary inverse Fast Fourier Transform (IFFT), i.e.,

$$x_1[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_1[k] e^{j\frac{2\pi nk}{N}}, \quad n = 0, 1, \ldots, N-1.$$

A cyclic prefix of length P is added to the IFFT output in order to eliminate the inter-symbol interference caused by multipath propagation. The resulting $N + P$ symbols are converted into a baseband signal for transmission. With the aid of the cyclic prefix and provided that $L - 1 \leq P$, linear convolution becomes circular convolution in the discrete-time domain, i.e.,

$$z_1[n] = \sum_{l=0}^{L-1} x_1[(n-l)_N] h[l], \quad n = 0, 1, \ldots, N-1, \tag{4.26}$$

where $(n - l)_N$ stands for

$$(n - l) \bmod N.$$

At the receiver, the received time-domain symbols are given by (4.9), i.e.,

$$y_1[n] = \gamma_1 z_1[n] + \frac{3\gamma_3}{2} z_1[n] |z_1[n]|^2 + 3\gamma_3 z_1[n] |z_2[n]|^2 + w_1[n]$$

for $n = 0, 1, \ldots, N-1$. The unitary Fast Fourier Transform (FFT) is performed on $y_1[n]$, $n = 0, 1, \ldots, N-1$, to obtain $Y_1[k]$, $k = 0, 1, \ldots, N-1$. In the absence of nonlinearity, i.e., $\gamma_3 = 0$, we have

$$Y_1[k] = H''[k] X_1[k] + W_1[k], \quad k = 0, 1, \ldots, N-1, \tag{4.27}$$

where $H''[k]$ is the channel response in the kth subcarrier and $W_1[k]$ is the additive noise in the kth subcarrier. Note that $H''[k]$, $k = 0, 1, \ldots, N-1$, are the Fourier transform coefficients of $h''[n] = \gamma_1 h[n]$, $n = 0, 1, \ldots, L-1$, i.e.,

$$H''[k] = \sum_{n=0}^{L-1} h''[n] e^{-j\frac{2\pi kn}{N}}, \quad k = 0, 1, \ldots, N-1.$$

From (4.27), $X_1[k]$ can be estimated from

$$\widehat{X}_1[k] = \frac{Y_1[k]}{H''[k]}, \quad k = 0, 1, \ldots, N-1.$$

The obtained symbols $\widehat{X}_1[k]$ are then mapped into information bits. In the presence of cross modulation, expression (4.27) becomes invalid. To compensate for the distortion, the proposed time-domain algorithm in Section 4.3 is applied before the OFDM demodulation. In the channel estimation stage, the frequency-domain pilot symbols $X_1[k]$, $k = 0, 1, \ldots, N-1$, and hence $x_1[n]$, $n = 0, 1, \ldots, N-1$, are known to the receiver, and because of the circular convolution (4.26), the channel estimation algorithm presented in Subsection 4.3.1 is applied with $\mathbf{X}$ modified to:

$$\mathbf{X} = \begin{bmatrix} x_1[0] & x_1[N-1] & \ldots & x_1[N-L+1] \\ x_1[1] & x_1[0] & \ldots & x_1[N-L+2] \\ \vdots & \vdots & \ddots & \vdots \\ x_1[N-2] & x_1[N-3] & \ldots & x_1[N-L-1] \\ x_1[N-1] & x_1[N-2] & \ldots & x_1[N-L] \end{bmatrix}.$$

The obtained time-domain channel response $\widehat{h}''[n]$, $n = 0, 1, \ldots, L-1$, is converted to the frequency-domain response $\widehat{H}''[k]$, $k = 0, 1, \ldots, N-1$, for OFDM demodulation. In the data transmission stage, we first estimate $z_1''[n]$ in the time domain by

$$\widehat{z}_1''[n] = \frac{y_1[n]}{1 + 3\widehat{\gamma}_3'' |y_2[n]|^2}.$$

We then apply the FFT on $\widehat{z}_1''[n]$ to obtain $\widehat{Z}_1''[k]$ and the frequency-domain transmitted symbols $X_1[k]$ are estimated by

$$\widehat{X}_1[k] = \frac{\widehat{Z}_1''[k]}{\widehat{H}''[k]}, \quad k = 0, 1, \ldots, N-1.$$

4.6   Computer Simulations

In the simulations, the bandwidth of the desired signal and blocker signals is 20 MHz, and the constellation used for the desired signal is QPSK. We first simulate the single-carrier system when there is only one blocker signal. The channel response has length 4, and its taps are independently Rayleigh distributed with the total power normalized to be 1. The average received signal power is set to be $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$. The model parameters of the desired channel and the blocker signal channels are specified as $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\gamma_1' = 0.5623$, $\gamma_2' = 0.0005623$, and $\gamma_3' = -0.00749733$. Compared to the desired channel, this is equivalent to a 40 dB attenuation in the blocker signal channel. The analysis in Section 4.2 shows that the desired channel is distorted by the cross modulation, but the secondary channel is still dominated by the blocker signal. The proposed channel estimation method uses pilot sequences of length 64, and its normalized mean-square-error (MSE) and CRLB are plotted vs. the normalized signal-to-noise ratio at the receiver, i.e., $\mathrm{SNR}_0 = \gamma_1^2 \sigma_{z,1}^2 / \sigma_w^2$, in Fig. 4.9. It can be seen that the CRLB provides a good measure of the estimation accuracy. Fig. 4.10 shows the MSE and uncoded bit error rate (BER) performance of the proposed data symbol estimation algorithm when assuming that the receiver has perfect knowledge of the model parameters and the channel response. The Viterbi algorithm is used to demodulate the estimated QPSK symbols into information bits. Fig. 4.11 compares the BER performance of the whole proposed scheme, i.e., including both channel estimation and data symbol estimation, with that of a distorted receiver without any compensation. Fig. 4.12 shows the BER performance of the proposed scheme when there are two blocker signals with $\sigma_{z,2}^2 = \sigma_{z,3}^2 = 2.5 \times 10^{-4}$.

In Fig. 4.13, we demonstrate the performance of an OFDM system with 64

subcarriers and cyclic prefix length 16. The channel is assumed to be additive Gaussian. Without any compensation, the system performs poorly and saturates at a high BER. The proposed compensation scheme improves the performance significantly.

4.7 Conclusions

In this chapter, the effects of RF nonlinearities in a software-defined radio (SDR) receiver are studied. A compensation scheme is proposed that consists of two stages. One stage is the joint channel and nonlinearity parameter estimation, and the other is the data symbol estimation via distortion compensation. The proposed channel estimation algorithm performs close to the derived Cramer-Rao lower bound. Also, the analysis and simulations show that the compensation scheme can effectively improve the system performance and reduce the sensitivity of SDR receivers to the nonlinearity impairment. Since receivers with less analog impairments usually have the disadvantage of high implementation cost and power consumption, our techniques enable the use of low-cost receivers for the next-generation wireless communications that are built on the platform of SDRs.
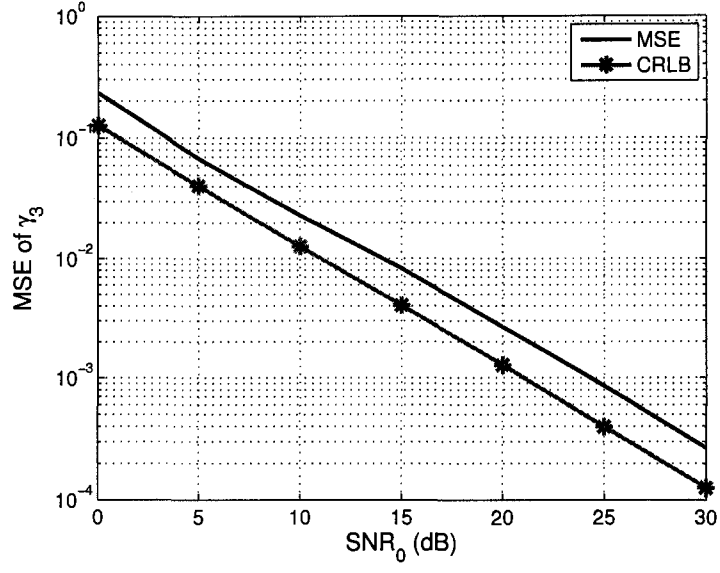
(a) MSE and CRLB of γ_3.
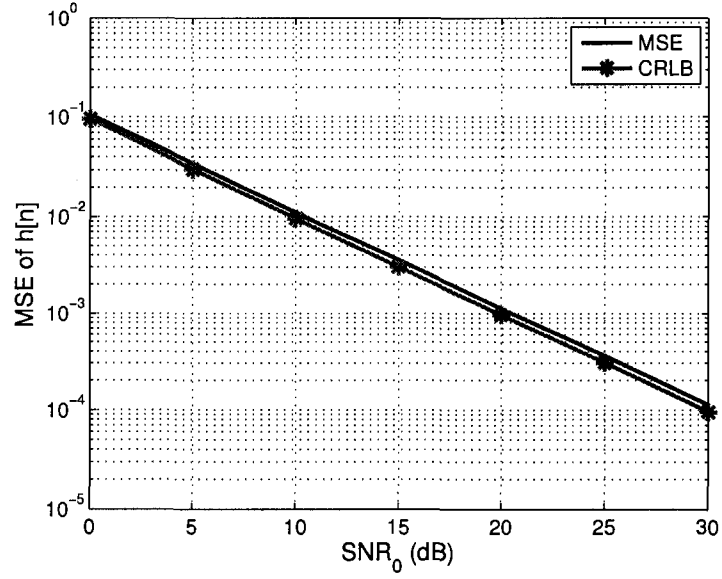
(b) MSE and CRLB of **h**.

Figure 4.9: Plots of the normalized MSE and CRLB for channel estimation when $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$.
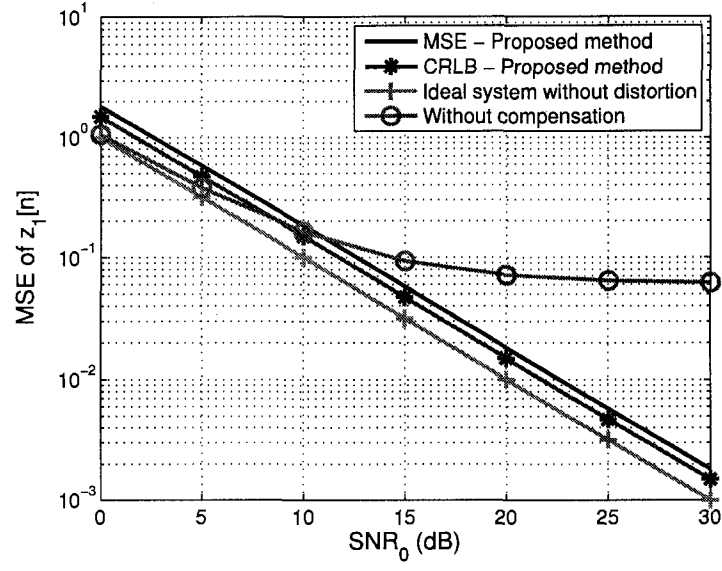
(a) MSE and CRLB of z_1.
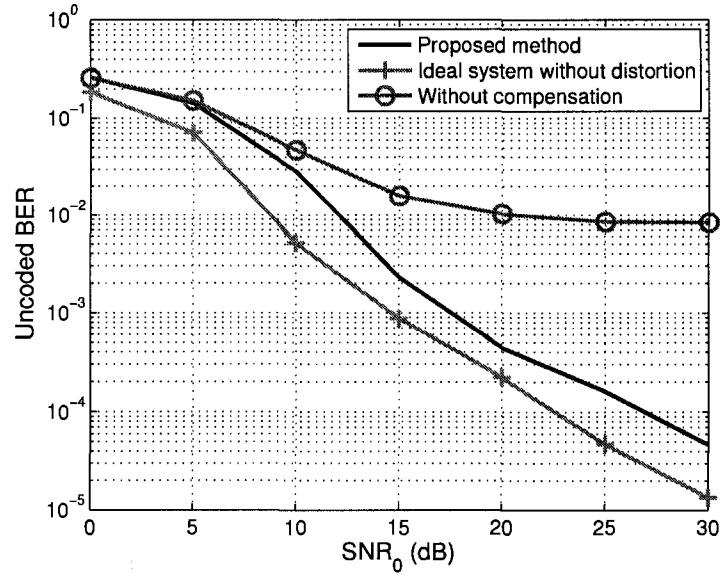
(b) Uncoded BER.

Figure 4.10: Plots of the MSE and uncoded BER for data symbol estimation when $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$.
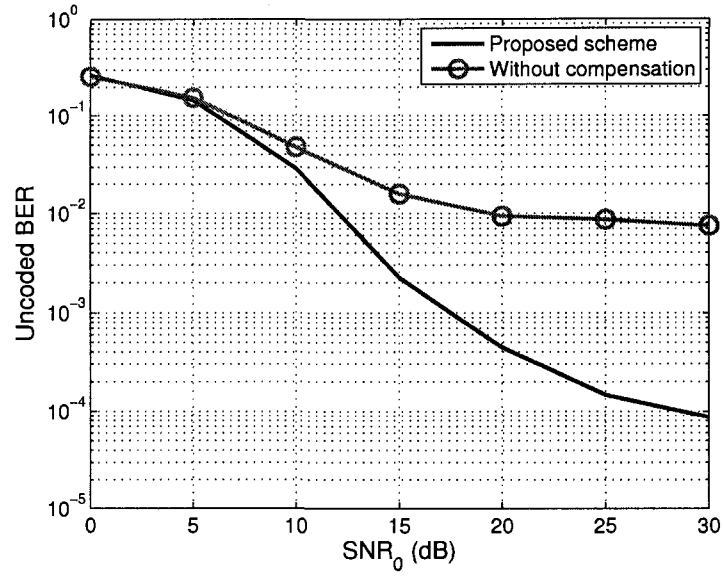
Figure 4.11: Uncoded BER of the proposed scheme for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$.
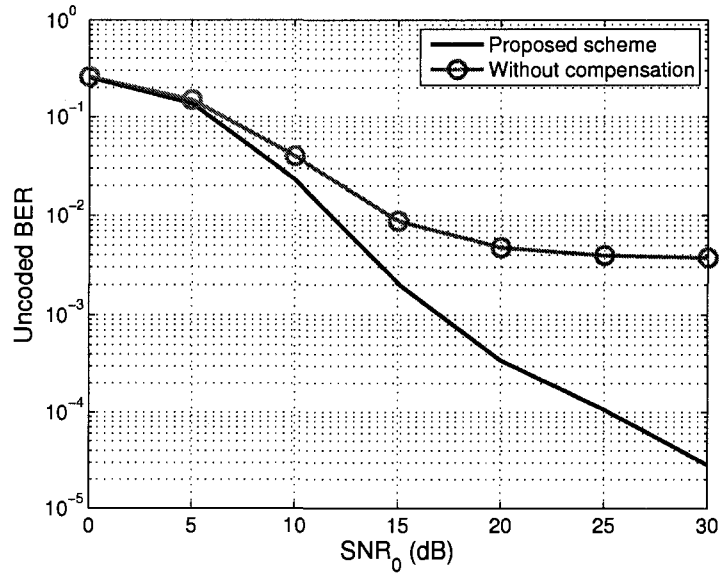
Figure 4.12: Uncoded BER in the presence of two blocker signals when $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = \sigma_{z,3}^2 = 2.5 \times 10^{-4}$.
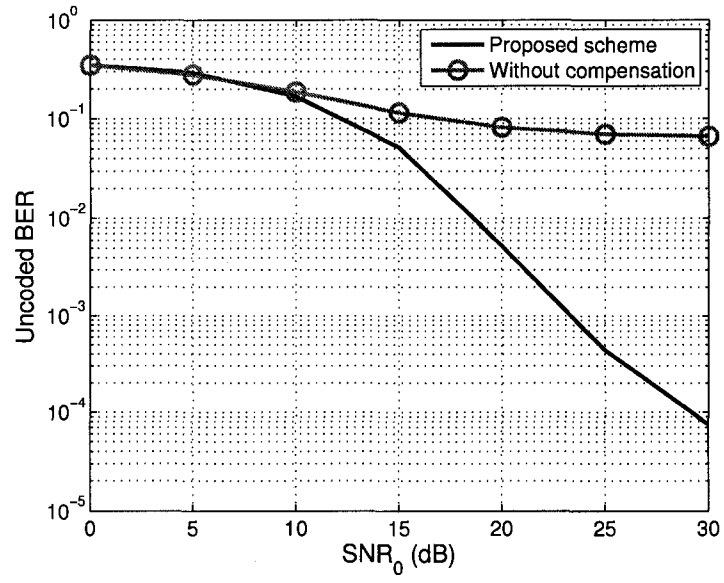
Figure 4.13: Uncoded BER of OFDM systems for $\gamma_1 = 56.23$, $\gamma_2 = 5.623$, $\gamma_3 = -7497.33$, $\sigma_{z,1}^2 = 10^{-12}$ and $\sigma_{z,2}^2 = 5 \times 10^{-4}$.

CHAPTER 5

Conclusions and Future Research

The dissertation studies the effects of RF front-end distortions on wireless communication systems. Three different distortions are considered in the work, i.e., phase noise, IQ imbalance and nonlinearity. A digital approach is proposed to compensate for these distortions. Analysis and computer simulations show that the sensitivity of the communication systems to these distortions can be significantly reduced by using our techniques. Despite of their higher complexity than other available methods, the rapid advances of digital IC technologies make the proposed algorithms feasible and implementable in real time. Consequently, the contribution of this work is to provide an alternative approach for RF wireless designers to achieve their goals. Instead of concentrating only on fine-tuning analog circuits, it allows designers to rely more on baseband digital processing. This point of view also demonstrates how mixed-signal (joint analog and digital) processing techniques play a critical role in emerging radio technologies.

Although some effort has been spent on demonstrating the effectiveness of the proposed digital schemes, there are still remaining issues. We recommend the following research directions for further study.

5.1 Analog vs. Digital

In order to reduce the impairments, analog approaches usually suggest compli-cated circuit architectures that lead to more implementation cost and power consumption. The alternative digital approach demands more complex baseband processing, also resulting in an increase in the computational cost and power consumption. It is thus interesting to compare the various expenses incurred by the two approaches. Such a comparison can serve as a valuable reference when engineers make decisions on selecting proper tools to achieve their design objectives.

5.2 Phase Noise Compensation

Chapter 2 proposes a compensation scheme to mitigate the effects of phase noise in OFDM systems. The approach utilizes the statistical property of the phase noise process, i.e., the fact that adjacent phase noise samples exhibit some corre-lation. Thus, the performance of the proposed algorithm depends on the phase noise spectrum. For example, it is expected that the performance can be poor if the phase noise is absolutely white. A study that shows how the performance depends on the phase noise spectrum may guide analog/RF engineers to shape the phase noise spectrum in favorite of the digital compensation approach.

5.3 Joint Analog and Digital Design

Traditionally, when designing a communication system, the specifications and du-ties of the RF front-end and the baseband signal processing are separated. The work in this dissertation demonstrates that analog impairments can be compen-

sated for by using both analog and digital approaches. The results suggest that closer interaction between RF designers and baseband digital designers is useful in order to optimize the system in terms of cost and performance. This could happen when a design specification is easier to achieve in the digital domain but is harder in the analog domain, and vice versa.

5.4 Software-Defined Radio and Cognitive Radio

Software-defined radios (SDR) are able to configure their transmission and reception in a more flexible manner, e.g., by changing carrier frequencies and modulation methods. The SDR technology is going to play an important role in the next generation of wireless communication systems, e.g., cognitive radios. Because of its high flexibility, SDR poses numerous design challenges for the state-of-the-art analog/RF technologies, like ultra-wideband transmission and reception that are susceptible to interferences and noises. Fast and powerful digital signal processing techniques may be able to alleviate some of the design difficulties and accomplish design objectives.

APPENDIX A

Modeling of Phase Noise

There are mainly two types of oscillators used in practice, depending on whether or not they are used in a phase-locked loop (PLL) [PM02]. The so-called free-running oscillators operate without a PLL, and the generated phase noise is modeled as the accumulation of random frequency deviations and hence has unbounded variance. On the other hand, in a PLL oscillator, the closed-loop control mechanism tracks the phase variations of the carrier signal, and consequently, the generated phase noise has finite variance.

A.1 Free-Running Oscillator

The frequency deviation $\epsilon(t)$ of an oscillator is modeled as a zero-mean white Gaussian random process with single-sided power spectral density (PSD)

$$N_0 = \frac{\xi}{\pi},$$

where ξ is called the oscillator linewidth. The phase noise $\phi_{\text{Free}}(t)$ generated by a free-running oscillator is modeled by integrating $\epsilon(t)$, i.e.,

$$\phi_{\text{Free}}(t) = 2\pi \int_0^t \epsilon(\lambda) d\lambda,$$

which turns out to be a Wiener process as well as a Gaussian process. The single-sided PSD of $\phi_{\text{Free}}(t)$ is

$$S_{\phi,\text{Free}}(f) = \frac{\xi}{\pi f^2}$$

for $f \neq 0$. The carrier noise is

$$c_{\text{Free}}(t) = e^{j\phi_{\text{Free}}(t)},$$

and its mean value satisfies

$$\mathbf{E}\left\{c_{\text{Free}}(t)\right\} \to 0$$

as $t \to \infty$. The auto-correlation function of $c_{\text{Free}}(t)$ is given by

$$R_{c,\text{Free}}(\tau) = e^{-\pi\xi|\tau|} \tag{A.1}$$

and the single-sided PSD of $c_{\text{Free}}(t)$ is Lorenzian:

$$S_{c,\text{Free}}(f) = \frac{4\xi}{\pi\left(4f^2 + \xi^2\right)}.$$

Figs. A.1–A.3 plot $S_{\phi,\text{Free}}(f)$, $R_{c,\text{Free}}(\tau)$ and $S_{c,\text{Free}}(f)$ for a free-running oscillator with $\xi = 5$ kHz.

A.2   Oscillator with PLLs

Figure A.4 shows the block diagram of a PLL system [Bes03]. The phase error $\phi(t)$ between the carrier signal and the local oscillator is filtered by a low-pass loop filter that is represented by $K_d F(s)$. The resulting signal V_c is used as the input to the oscillator such that its output can actively track the phase variations of the carrier signal. It can be shown that the closed-loop transfer function between the carrier phase signal θ_i (input) and the oscillator phase signal θ_o (output) is given by

$$H(s) = \frac{\theta_o(s)}{\theta_i(s)} = \frac{K_0 K_d F(s)}{s + K_0 K_d F(s)},$$
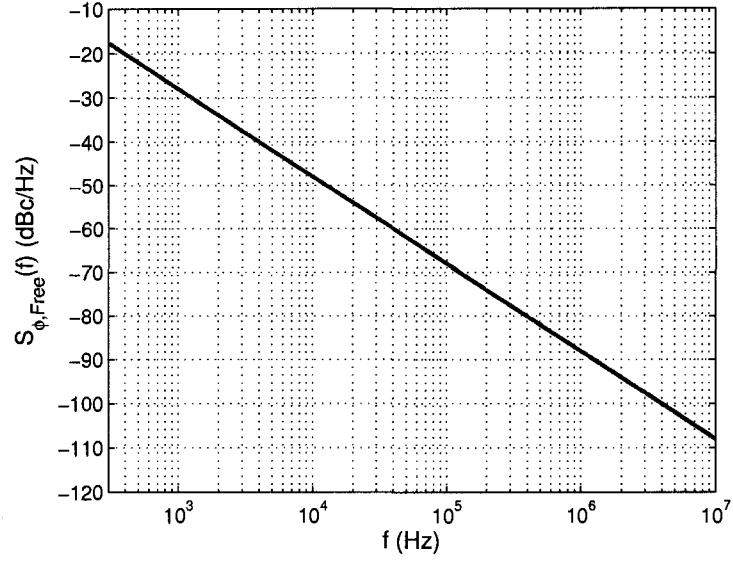
Figure A.1: Plot of $S_{\phi,\text{Free}}(f)$ for a free-running oscillator with $\xi = 5$ kHz.
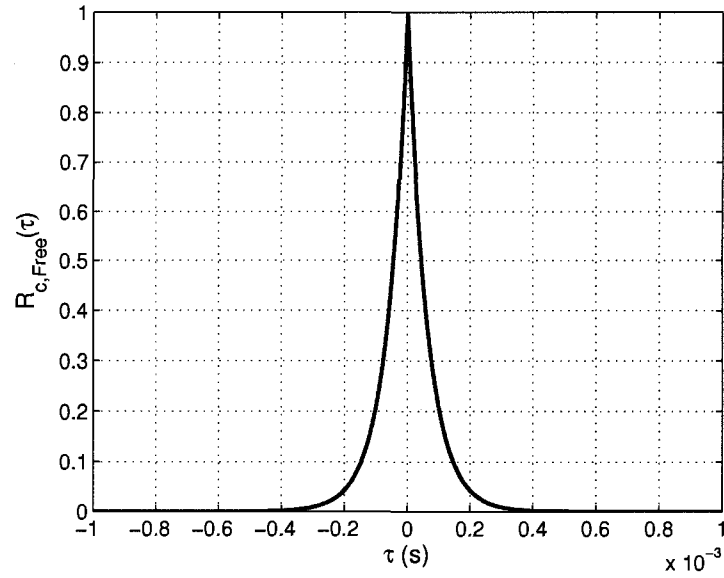
Figure A.2: Plot of $R_{c,\text{Free}}(\tau)$ for a free-running oscillator with $\xi = 5$ kHz.
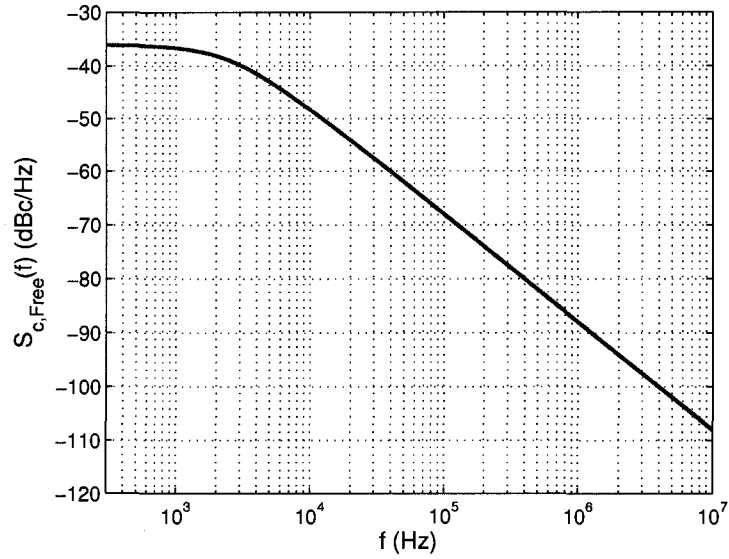
Figure A.3: Plot of $S_{c,\mathrm{Free}}(f)$ for a free-running oscillator with $\xi = 5$ kHz.

where $K_0 K_d$ is the open loop gain. Thus, the transfer function between the input phase signal θ_i and the phase noise ϕ is

$$G(s) = \frac{\phi(s)}{\theta_i(s)} = \frac{\theta_i(s) - \theta_o(s)}{\theta_i(s)} = 1 - H(s) = \frac{s}{s + K_0 K_d F(s)}.$$

Different loop filters will generate phase noise with different PSDs. In the following, we give the models for the 1st-order and 2nd-order PLLs that are commonly used in practice.
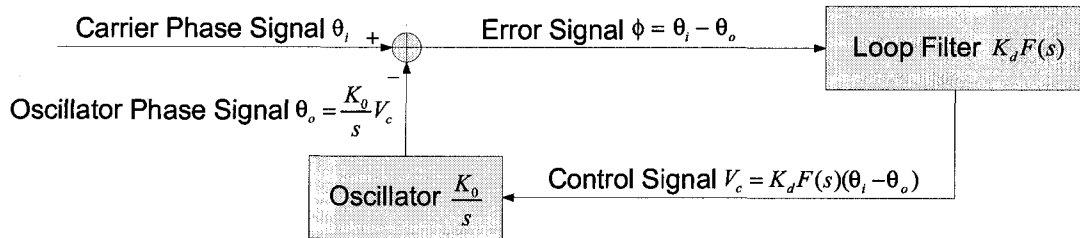
Figure A.4: Block diagram of a PLL system.

146

A.2.1 First-order PLL

For the 1st-order PLL, we have $F(s) = 1$, and the closed-loop transfer function is given by

$$H(s) = \frac{\omega_L}{s + \omega_L},$$

where $\omega_L = K_0 K_d$ is called the angular frequency of zero decibel loop gain. Since the input phase signal is generated by a free-running oscillator, we let $\theta_i(t) = \phi_{\text{Free}}(t)$. As discussed in Subsection A.1, $\phi_{\text{Free}}(t)$ is a Gaussian process and hence the resulting phase noise from PLL oscillators is also Gaussian. After correction by the PLL, the single-sided PSD of the phase noise $\phi_{\text{PLL}_1}(t)$ can be expressed as

$$
\begin{aligned}
S_{\phi,\text{PLL}_1}(f) &= |G(j2\pi f)|^2 \, S_{\phi,\text{Free}}(f) \\
&= |1 - H(j2\pi f)|^2 \, \frac{\xi}{\pi f^2} \\
&= \frac{\xi}{\pi(f^2 + f_L^2)},
\end{aligned}
$$

where

$$f_L = \frac{\omega_L}{2\pi}$$

is a measure of the loop bandwidth. The variance of $\phi_{\text{PLL}_1}(t)$ is given by

$$\sigma_{\phi,\text{PLL}_1}^2 = \int_0^\infty S_{\phi,\text{PLL}_1}(f)df = \frac{\xi}{2f_L}, \tag{A.2}$$

and the auto-correlation function of $\phi_{\text{PLL}_1}(t)$ can be calculated as

$$R_{\phi,\text{PLL}_1}(\tau) = \int_{-\infty}^\infty \frac{\xi}{2\pi(f^2 + f_L^2)} \, e^{j2\pi f \tau}df. \tag{A.3}$$

Although there is no simple closed-form expression for $R_{\phi,\text{PLL}_1}(\tau)$, relation (A.3) can be evaluated numerically and the result is useful in analyzing the effects of phase noise on OFDM systems as well as the performance of different compensation schemes.

The associated carrier noise is

$$c_{\mathrm{PLL}_1}(t) = e^{j\phi_{\mathrm{PLL}_1}(t)}.$$

Since $\phi_{\mathrm{PLL}_1}(t)$ is Gaussian distributed with mean zero and variance $\sigma^2_{\phi,\mathrm{PLL}_1} = \xi/(2f_L)$, the mean value of $c_{\mathrm{PLL}_1}(t)$ is given by

$$
\begin{aligned}
\mathbf{E}\left\{c_{\mathrm{PLL}_1}(t)\right\} &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2_{\phi,\mathrm{PLL}_1}}} e^{-\frac{z^2}{2\sigma^2_{\phi,\mathrm{PLL}_1}}} e^{jz} dz \\
&= e^{-\frac{\sigma^2_{\phi,\mathrm{PLL}_1}}{2}} \\
&= e^{-\frac{\xi}{4f_L}}.
\end{aligned}
$$

By definition, the auto-correlation function of $c_{\mathrm{PLL}_1}(t)$ is given by

$$
\begin{aligned}
R_{c,\mathrm{PLL}_1}(\tau) &= \mathbf{E}\left\{c_{\mathrm{PLL}_1}(t)c^*_{\mathrm{PLL}_1}(t-\tau)\right\} \\
&= \mathbf{E}\left\{e^{j\left[\phi_{\mathrm{PLL}_1}(t)-\phi_{\mathrm{PLL}_1}(t-\tau)\right]}\right\},
\end{aligned}
$$

where $\phi_{\mathrm{PLL}_1}(t) - \phi_{\mathrm{PLL}_1}(t-\tau)$ is Gaussian distributed with mean zero and variance

$$
\begin{aligned}
\mathbf{E}\left\{\left[\phi_{\mathrm{PLL}_1}(t) - \phi_{\mathrm{PLL}_1}(t-\tau)\right]^2\right\} &= 2\sigma^2_{\phi,\mathrm{PLL}_1} - 2R_{\phi,\mathrm{PLL}_1}(\tau) \\
&= \frac{\xi}{f_L} - 2R_{\phi,\mathrm{PLL}_1}(\tau).
\end{aligned}
$$

Taking the expectation with respect to the distribution of $\phi_{\mathrm{PLL}_1}(t) - \phi_{\mathrm{PLL}_1}(t-\tau)$ gives

$$
\begin{aligned}
R_{c,\mathrm{PLL}_1}(\tau) &= e^{R_{\phi,\mathrm{PLL}_1}(\tau)-\sigma^2_{\phi,\mathrm{PLL}_1}} \\
&= e^{R_{\phi,\mathrm{PLL}_1}(\tau)-\frac{\xi}{2f_L}}.
\end{aligned}
$$

It then follows that the single-sided PSD of $c_{\mathrm{PLL}_1}(t)$ can be calculated by

$$
\begin{aligned}
S_{c,\mathrm{PLL}_1}(f) &= 2\int_{-\infty}^{\infty} R_{c,\mathrm{PLL}_1}(\tau)e^{-j2\pi f\tau} d\tau \\
&= 2e^{-\frac{\xi}{2f_L}} \int_{-\infty}^{\infty} e^{R_{\phi,\mathrm{PLL}_1}(\tau)-j2\pi f\tau} d\tau.
\end{aligned}
$$

Figs. A.5–A.8 plot $R_{\phi,\mathrm{PLL}_1}(\tau)$, $S_{\phi,\mathrm{PLL}_1}(f)$, $R_{c,\mathrm{PLL}_1}(\tau)$ and $S_{c,\mathrm{PLL}_1}(f)$ for a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.

Figure A.5: Plot of $R_{\phi,\mathrm{PLL}_1}(\tau)$ for a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.

Figure A.6: Plot of $S_{\phi,\mathrm{PLL}_1}(f)$ for a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.
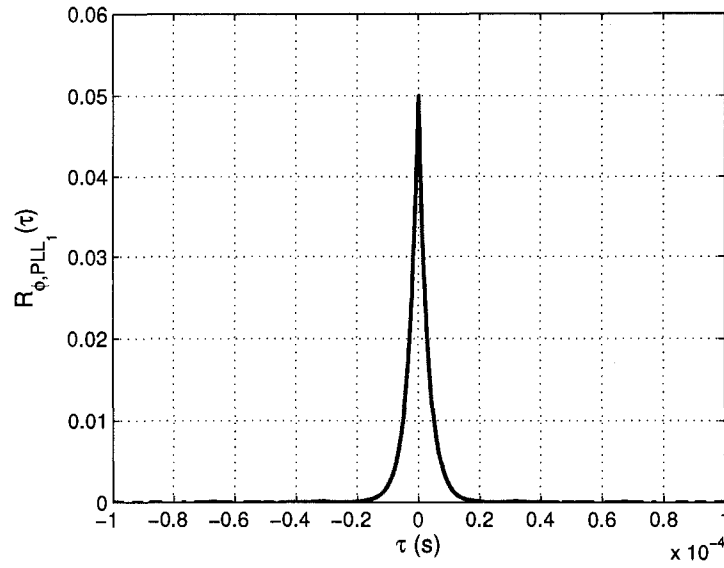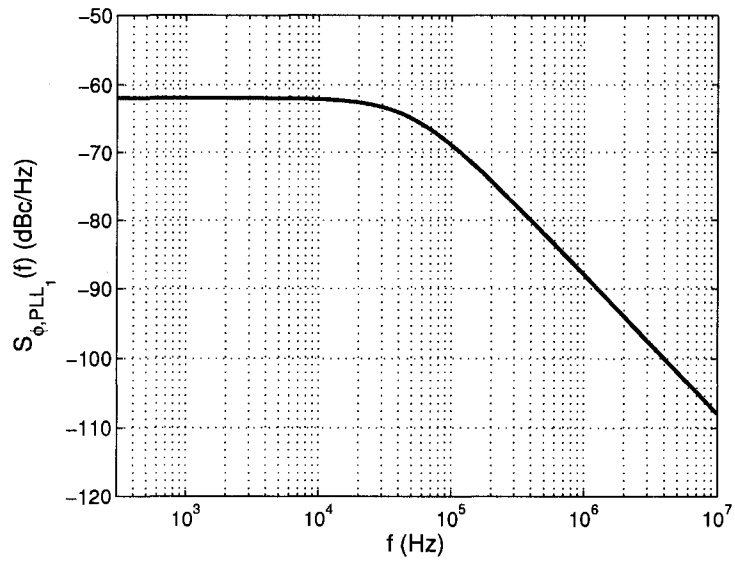
Figure A.7: Plot of $R_{c,\text{PLL}_1}(\tau)$ for a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.

Figure A.8: Plot of $S_{c,\text{PLL}_1}(f)$ for a 1st-order PLL oscillator with $\xi = 5$ kHz and $f_L = 50$ kHz.

A.2.2 Second-order PLL

For the 2nd-order PLL, the closed-loop transfer function is

$$H(s) = \frac{2\eta\omega_n s + \omega_n^2}{s^2 + 2\eta\omega_n s + \omega_n^2},$$

where ω_n is the loop natural frequency and η is the damping factor. In this dissertation, we assume that $\eta = 0.707$. Then the single-sided PSD of the phase noise $\phi_{\mathrm{PLL}_2}(t)$ is given by

$$
\begin{aligned}
S_{\phi,\mathrm{PLL}_2}(f) &= |G(j2\pi f)|^2 \, S_{\phi,\mathrm{Free}}(f) \\
&= |1 - H(j2\pi f)|^2 \frac{\xi}{\pi f^2} \\
&= \frac{\xi f^2}{\pi(f^4 + f_n^4)},
\end{aligned}
$$

where

$$f_n = \frac{\omega_n}{2\pi}$$

is a measure of the loop bandwidth. The variance of $\phi_{\mathrm{PLL}_2}(t)$ can be calculated as

$$\sigma_{\phi,\mathrm{PLL}_2}^2 = \int_0^\infty S_{\phi,\mathrm{PLL}_2}(f)df = \frac{\xi}{2\sqrt{2}f_n}. \tag{A.4}$$

The auto-correlation function of $\phi_{\mathrm{PLL}_2}(t)$ can be calculated by

$$R_{\phi,\mathrm{PLL}_2}(\tau) = \int_{-\infty}^\infty \frac{\xi f^2}{2\pi(f^4 + f_n^4)} \, e^{j2\pi f\tau} df.$$

The associated carrier noise is $c_{\mathrm{PLL}_2}(t) = e^{j\phi_{\mathrm{PLL}_2}(t)}$, and its mean value is

$$
\begin{aligned}
\mathbf{E}\left\{c_{\mathrm{PLL}_2}(t)\right\} &= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma_{\phi,\mathrm{PLL}_2}^2}} e^{-\frac{z^2}{2\sigma_{\phi,\mathrm{PLL}_2}^2}} e^{jz} dz \\
&= e^{-\frac{\sigma_{\phi,\mathrm{PLL}_2}^2}{2}} \\
&= e^{-\frac{\xi}{4\sqrt{2}f_n}}.
\end{aligned}
$$

151
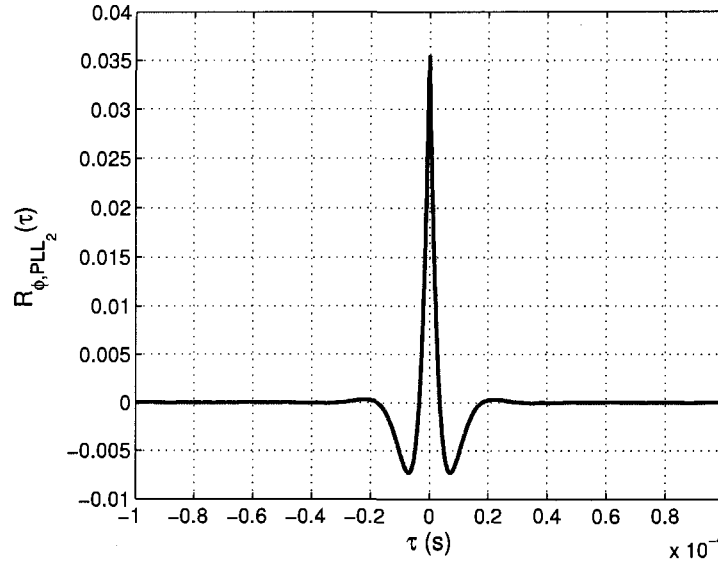
Figure A.9: Plot of $R_{\phi,\mathrm{PLL}_2}(\tau)$ for a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.

The auto-correlation function of $c_{\mathrm{PLL}_2}(t)$ is given by

$$R_{c,\mathrm{PLL}_2}(\tau) = e^{R_{\phi,\mathrm{PLL}_2}(\tau)-\sigma^2_{\phi,\mathrm{PLL}_2}}$$

$$= e^{R_{\phi,\mathrm{PLL}_2}(\tau)-\frac{\xi}{2\sqrt{2}f_n}},$$

and the single-sided PSD of $c_{\mathrm{PLL}_2}(t)$ can be calculated by

$$S_{c,\mathrm{PLL}_2}(f) = 2\int_{-\infty}^{\infty} R_{c,\mathrm{PLL}_2}(\tau)e^{-j2\pi f\tau}d\tau$$

$$= 2e^{-\frac{\xi}{2\sqrt{2}f_n}}\int_{-\infty}^{\infty} e^{R_{\phi,\mathrm{PLL}_2}(\tau)-j2\pi f\tau}d\tau.$$

Figs. A.9–A.12 plot $R_{\phi,\mathrm{PLL}_2}(\tau)$, $S_{\phi,\mathrm{PLL}_2}(f)$, $R_{c,\mathrm{PLL}_2}(\tau)$ and $S_{c,\mathrm{PLL}_2}(f)$ for a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.

In Table A.1, we summarize the statistical measures that will be used to compute the effective signal-to-noise ratio for different compensation schemes. For convenience of notation, if the phase noise model is not specified, we use $\phi(t)$

152
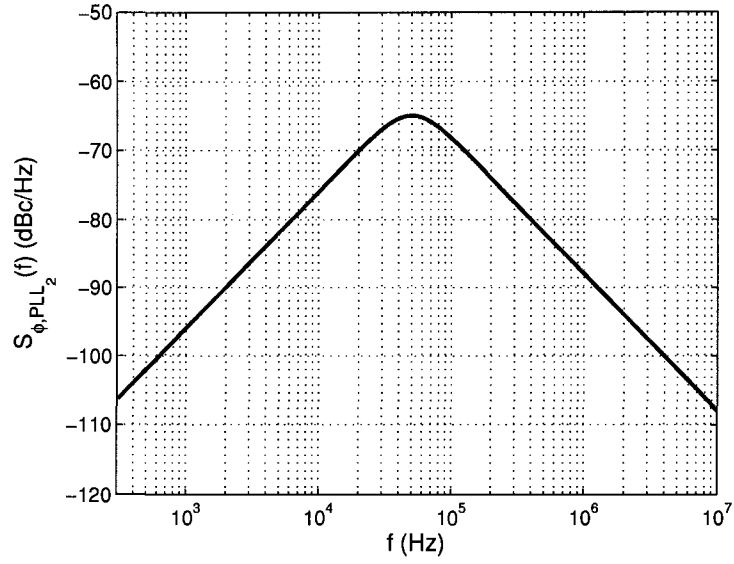
Figure A.10: Plot of $S_{\phi,\text{PLL}_2}(f)$ for a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.
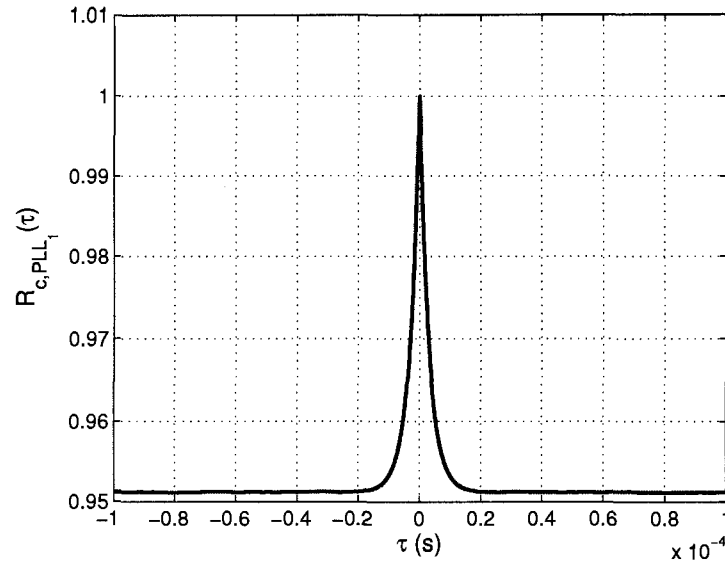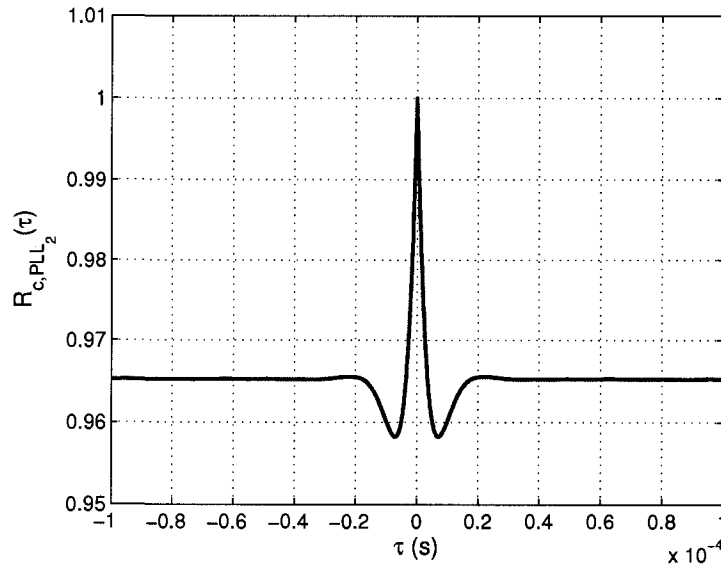

Figure A.11: Plot of $R_{c,\text{PLL}_2}(\tau)$ for a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.
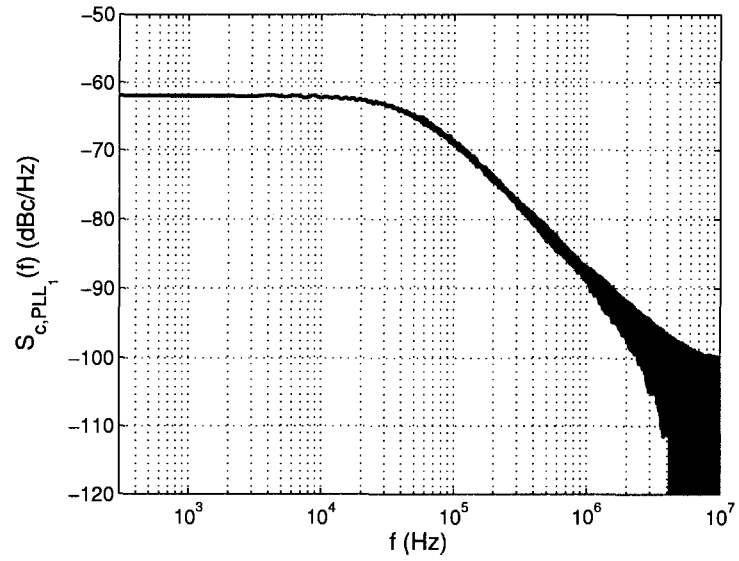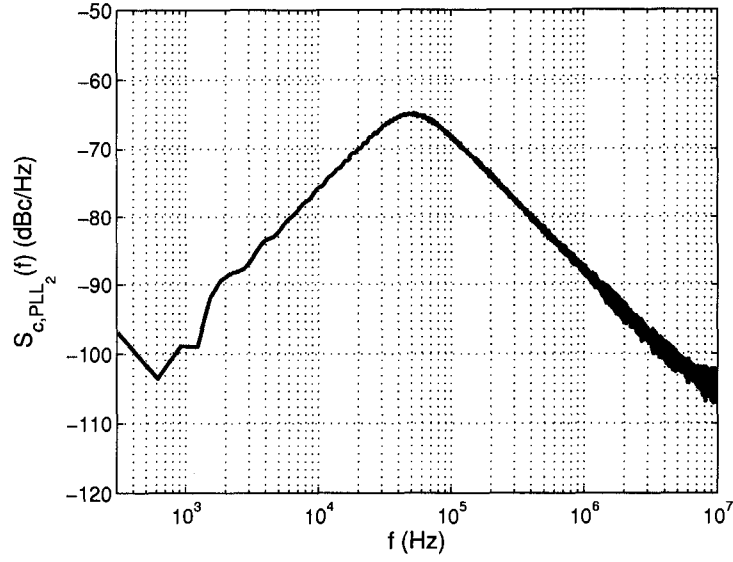
Figure A.12: Plot of $S_{c,\text{PLL}_2}(f)$ for a 2nd-order PLL oscillator with $\xi = 5$ kHz and $f_n = 50$ kHz.

and $c(t)$ to represent the phase noise and carrier noise by ignoring the subscripts. It is the same for σ_ϕ^2, $R_\phi(\tau)$, $R_c(\tau)$, $S_\phi(f)$ and $S_c(f)$.

Remark: For phase-loop locked oscillators, the phase noise $\phi(t)$ is usually small. If $|\phi(t)| \ll 1$, the carrier noise $c(t)$ can be approximated by $c(t) = e^{j\phi(t)} \approx 1 + j\phi(t)$. In this case, $R_c(\tau) \approx 1 + R_\phi(\tau)$ and $S_c(f) \approx S_\phi(f)$ for $f \neq 0$.

Table A.1: Table of statistical measures for different types of phase noise.

	Free-running	1st-order PLL	2nd-order PLL
σ_ϕ^2	—	$\frac{\xi}{2f_L}$	$\frac{\xi}{2\sqrt{2}f_n}$
$R_\phi(\tau)$	—	$\int_{-\infty}^{\infty} \frac{\xi}{2\pi(f^2+f_L^2)} e^{j2\pi f\tau} df$	$\int_{-\infty}^{\infty} \frac{\xi f^2}{2\pi(f^4+f_n^4)} e^{j2\pi f\tau} df$
$S_\phi(f)$	$\frac{\xi}{\pi f^2}$	$\frac{\xi}{\pi(f^2+f_L^2)}$	$\frac{\xi f^2}{\pi(f^4+f_n^4)}$
$\mathbf{E}\{c(t)\}$	$\to 0$	$e^{-\frac{\xi}{4f_L}}$	$e^{-\frac{\xi}{4\sqrt{2}f_n}}$
$R_c(\tau)$	$e^{-\pi\xi|\tau|}$	$e^{R_{\phi,\text{PLL}_1}(\tau)-\frac{\xi}{2f_L}}$	$e^{R_{\phi,\text{PLL}_2}(\tau)-\frac{\xi}{2\sqrt{2}f_n}}$
$S_c(f)$	$\frac{4\xi}{\pi(4f^2+\xi^2)}$	$2e^{-\frac{\xi}{2f_L}}\int_{-\infty}^{\infty} e^{R_{\phi,\text{PLL}_1}(\tau)-j2\pi f\tau} d\tau$	$2e^{-\frac{\xi}{2\sqrt{2}f_n}}\int_{-\infty}^{\infty} e^{R_{\phi,\text{PLL}_2}(\tau)-j2\pi f\tau} d\tau$

A.3 Statistical Characteristics of $A[k]$

Given the above phase noise models, the mean and auto-correlation of $A[k]$ can be derived from expression (2.2) as follows:

$$\mathbf{E}\{A[k]\} = \begin{cases} \mathbf{E}\{c(t)\}, & k = 0, \\ 0, & k = 1, 2, \ldots, N-1, \end{cases} \tag{A.5}$$

$$\mathbf{E}\{A[k_1](A[k_2])^*\} = \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} R_c((n_1 - n_2)T_s) e^{-j\frac{2\pi(k_1 n_1 - k_2 n_2)}{N}}, \tag{A.6}$$

for $k_1 = 0, 1, \ldots, N-1$ and $k_2 = 0, 1, \ldots, N-1$. In particular, the following expressions derived from (A.5) and (A.6) are useful in evaluating the performance of different phase noise compensation schemes:

$$\sigma_{A,0}^2 = \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} R_c((n_1 - n_2)T_s), \tag{A.7}$$

$$\mathbf{E}\{|A[0] - 1|^2\} = \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} R_c((n_1 - n_2)T_s) - 2\mathrm{Re}\{\mathbf{E}\{c(t)\}\} + 1, \tag{A.8}$$

$$\sum_{k=1}^{N-1} \sigma_{A,k}^2 = 1 - \sigma_{A,0}^2 = 1 - \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} R_c((n_1 - n_2)T_s), \tag{A.9}$$

where $\sigma_{A,k}^2 = \mathbf{E}\{|A[k]|^2\}$ is the average power of $A[k]$, $k = 0, 1, \ldots, N-1$. Here, (A.7) follows from (A.6) by setting $k_1 = k_2 = 0$, (A.8) follows from (A.7) and (A.5), and (A.9) follows from (A.7) by noting that

$$\sum_{k=0}^{N-1} \sigma_{A,k}^2 = \sum_{k=0}^{N-1} \mathbf{E}\{|A[k]|^2\} = \mathbf{E}\left\{\sum_{k=0}^{N-1} |A[k]|^2\right\} = \mathbf{E}\left\{\frac{1}{N} \sum_{n=0}^{N-1} |e^{j\phi(nT_s)}|^2\right\} = 1.$$

APPENDIX B

$\mathbf{P}_o \approx \mathbf{P}_L$ for Free-running Oscillators if $\xi T_s \ll 1$

Proof: Recall that

$$\mathbf{c} = \begin{bmatrix} e^{j\phi_{\mathrm{Free}}(0)} \\ e^{j\phi_{\mathrm{Free}}(T_s)} \\ \vdots \\ e^{j\phi_{\mathrm{Free}}((N-1)T_s)} \end{bmatrix} \quad (N \times 1) \quad \text{and} \quad \mathbf{c}' = \begin{bmatrix} e^{j\phi_{\mathrm{Free}}(0)} \\ e^{j\phi_{\mathrm{Free}}\left(\frac{(N-1)T_s}{M-1}\right)} \\ \vdots \\ e^{j\phi_{\mathrm{Free}}((N-1)T_s)} \end{bmatrix} \quad (M \times 1).$$

Let

$$a = e^{-\pi\xi T_s}$$

and

$$b = e^{-\frac{\pi\xi(N-1)T_s}{M-1}}.$$

It follows from expression (A.1) that the element of $\mathbf{R}_{\mathbf{cc}'}$ at the nth row and m^{th} column is given by

$$\begin{aligned} \mathbf{R}_{\mathbf{cc}'}(n,m) &= \mathbf{E}\left\{\mathbf{c}(n)\mathbf{c}'(m)^*\right\} \\ &= \mathbf{E}\left\{e^{j\phi_{\mathrm{Free}}((n-1)T_s)}e^{-j\phi_{\mathrm{Free}}\left(\frac{(m-1)(N-1)T_s}{M-1}\right)}\right\} \\ &= e^{-\pi\xi\left|(n-1)T_s - \frac{(m-1)(N-1)T_s}{M-1}\right|} \\ &= \begin{cases} a^{n-1}b^{-(m-1)}, & \text{if } n-1 \geq \frac{(m-1)(N-1)}{M-1}, \\ a^{-(n-1)}b^{m-1}, & \text{if } n-1 < \frac{(m-1)(N-1)}{M-1}, \end{cases} \end{aligned}$$

where $n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$. Also, we have

$$\mathbf{R}_{c'} = \mathbf{E}\left\{\mathbf{c}'\mathbf{c}'^*\right\} = \begin{bmatrix} 1 & b & b^2 & \ldots & b^{M-1} \\ b & 1 & b & \ldots & b^{M-2} \\ b^2 & b & 1 & \ldots & b^{M-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b^{M-1} & b^{M-2} & b^{M-3} & \ldots & 1 \end{bmatrix} \quad (M \times M).$$

Consider an arbitrary integer p with $1 \leq p \leq N$, and let q be the associated integer such that

$$\frac{(q-1)(N-1)}{M-1} \leq p - 1 < \frac{q(N-1)}{M-1}.$$

The pth row of $\mathbf{R}_{cc'}$ is given by

$$\mathbf{r}_{cc',p} = \begin{bmatrix} a^{p-1} & a^{p-1}b^{-1} & \ldots & a^{p-1}b^{-(q-1)} & a^{-(p-1)}b^q & a^{-(p-1)}b^{q+1} & \ldots \\ & & a^{-(p-1)}b^{M-1} \end{bmatrix} \quad (1 \times M).$$

Denote the qth and $(q+1)^{th}$ rows of $\mathbf{R}_{c'}$ by $\mathbf{r}_{c',q}$ and $\mathbf{r}_{c',q+1}$, respectively, i.e.,

$$\mathbf{r}_{c',q} = \begin{bmatrix} b^{q-1} & b^{q-2} & \ldots & b & 1 & b & b^2 & \ldots & b^{M-q} \end{bmatrix} \quad (1 \times M),$$

$$\mathbf{r}_{c',q+1} = \begin{bmatrix} b^q & b^{q-1} & \ldots & b^2 & b & 1 & b & \ldots & b^{M-q-1} \end{bmatrix} \quad (1 \times M).$$

It is easy to verify that

$$\begin{aligned} \mathbf{r}_{cc',p} &= \frac{a^{p-1}}{(1-b^2)b^{q-1}}\left(\mathbf{r}_{c',q} - b\mathbf{r}_{c',q+1}\right) + \frac{b^q}{(1-b^2)a^{p-1}}\left(\mathbf{r}_{c',q+1} - b\mathbf{r}_{c',q}\right) \\ &= \left[\frac{a^{p-1}}{(1-b^2)b^{q-1}} - \frac{b^{q+1}}{(1-b^2)a^{p-1}}\right]\mathbf{r}_{c',q} \qquad (\text{B.1}) \\ &\quad + \left[\frac{b^q}{(1-b^2)a^{p-1}} - \frac{a^{p-1}}{(1-b^2)b^{q-2}}\right]\mathbf{r}_{c',q+1}, \end{aligned}$$

which implies that each row of $\mathbf{R}_{cc'}$ can be expressed as a linear combination of two consecutive rows of $\mathbf{R}_{c'}$. Hence,

$$\mathbf{R}_{cc'} = \mathbf{P}'\mathbf{R}_{c'},$$

157

where the elements of $\mathbf{P}'$ can be determined from expression (B.1), i.e.,

$$\mathbf{P}'(n,m) = \begin{cases} \frac{a^{n-1}}{(1-b^2)b^{m-1}} - \frac{b^{m+1}}{(1-b^2)a^{n-1}}, & \text{if } \frac{(m-1)(N-1)}{M-1} \le n-1 < \frac{m(N-1)}{M-1}, \\ \frac{b^{m-1}}{(1-b^2)a^{n-1}} - \frac{a^{n-1}}{(1-b^2)b^{m-3}}, & \text{if } \frac{(m-2)(N-1)}{M-1} \le n-1 < \frac{(m-1)(N-1)}{M-1}, \\ 0, & \text{otherwise}, \end{cases}$$

for $n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$. It then follows from (2.25) that

$$\mathbf{P}_o = \mathbf{R}_{\mathbf{cc'}}\mathbf{R}_{\mathbf{c'}}^{-1} = \mathbf{P}'\mathbf{R}_{\mathbf{c'}}\mathbf{R}_{\mathbf{c'}}^{-1} = \mathbf{P}'.$$

In the following, we are going to show that $\mathbf{P}_o \approx \mathbf{P}_L$ if $\xi T_s \ll 1$. Using the following approximations:

$$\frac{a^{n-1}}{b^{m-1}} = e^{-\pi\xi(n-1)T_s + \frac{\pi\xi(m-1)(N-1)T_s}{M-1}} \approx 1 - \pi\xi(n-1)T_s + \frac{\pi\xi(m-1)(N-1)T_s}{M-1},$$

$$\frac{b^{m-1}}{a^{n-1}} = e^{\pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1}} \approx 1 + \pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1},$$

$$1 + b^2 = 1 + e^{-\frac{2\pi\xi(N-1)T_s}{M-1}} \approx 2,$$

$$1 - b^2 = 1 - e^{-\frac{2\pi\xi(N-1)T_s}{M-1}} \approx \frac{2\pi\xi(N-1)T_s}{M-1},$$

we then have

$$\frac{a^{n-1}}{(1-b^2)b^{m-1}} - \frac{b^{m+1}}{(1-b^2)a^{n-1}}$$

$$= \frac{1}{1-b^2} \cdot \frac{a^{n-1}}{b^{m-1}} - \frac{b^2}{1-b^2} \cdot \frac{b^{m-1}}{a^{n-1}}$$

$$\approx \frac{1}{1-b^2}\left(1 - \pi\xi(n-1)T_s + \frac{\pi\xi(m-1)(N-1)T_s}{M-1}\right)$$

$$\quad - \frac{b^2}{1-b^2}\left(1 + \pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1}\right)$$

$$= 1 - \frac{1+b^2}{1-b^2}\left(\pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1}\right)$$

$$\approx 1 - \frac{2}{\frac{2\pi\xi(N-1)T_s}{M-1}}\left(\pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1}\right)$$

$$= m - \frac{(n-1)(M-1)}{N-1},$$

and

$$\frac{b^{m-1}}{(1-b^2)a^{n-1}} - \frac{a^{n-1}}{(1-b^2)b^{m-3}}$$

$$= \frac{1}{1-b^2} \cdot \frac{b^{m-1}}{a^{n-1}} - \frac{b^2}{1-b^2} \cdot \frac{a^{n-1}}{b^{m-1}}$$

$$= \frac{1}{1-b^2} \left(1 + \pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1} \right)$$

$$\quad - \frac{b^2}{1-b^2} \left(1 - \pi\xi(n-1)T_s + \frac{\pi\xi(m-1)(N-1)T_s}{M-1} \right)$$

$$= 1 + \frac{1+b^2}{1-b^2} \left(\pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1} \right)$$

$$\approx 1 + \frac{2}{\frac{2\pi\xi(N-1)T_s}{M-1}} \left(\pi\xi(n-1)T_s - \frac{\pi\xi(m-1)(N-1)T_s}{M-1} \right)$$

$$= \frac{(n-1)(M-1)}{N-1} - (m-2).$$

Thus,

$$\mathbf{P}_o(n,m) = \mathbf{P}'(n,m)$$

$$\approx \begin{cases} m - \frac{(n-1)(M-1)}{N-1}, & \text{if } \frac{(m-1)(N-1)}{M-1} \le n-1 < \frac{m(N-1)}{M-1}, \\ \frac{(n-1)(M-1)}{N-1} - (m-2), & \text{if } \frac{(m-2)(N-1)}{M-1} \le n-1 < \frac{(m-1)(N-1)}{M-1}, \\ 0, & \text{otherwise,} \end{cases}$$

which is identical to (2.26). Therefore, $\mathbf{P}_o \approx \mathbf{P}_L$. ∎

APPENDIX C

Derivation of Expression (3.6)

Let LP$\{\cdot\}$ represent the lowpass filtering (LPF) operation. Note that

$$y_p(t) = \mathrm{Re}\left\{\sqrt{2}y_0(t)e^{j2\pi f_c t}\right\} + w_p(t).$$

The I-component of the received baseband signal is given by

$$\begin{aligned}
y_I(t) &= \mathrm{LP}\left\{y_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\\
&= \mathrm{LP}\left\{\left[\mathrm{Re}\left\{\sqrt{2}y_0(t)e^{j2\pi f_c t}\right\} + w_p(t)\right](1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\\
&= \mathrm{LP}\left\{\mathrm{Re}\left\{\sqrt{2}y_0(t)e^{j2\pi f_c t}\right\}(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\\
&\quad + \mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\\
&= \mathrm{LP}\left\{\mathrm{Re}\left\{2y_0(t)e^{j2\pi f_c t}(1+\alpha)\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\right\}\\
&\quad + \mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\\
&= \mathrm{LP}\left\{\mathrm{Re}\left\{y_0(t)e^{j2\pi f_c t}(1+\alpha)\left(e^{j\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)} + e^{-j\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)}\right)\right\}\right\}\\
&\quad + \mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\\
&= \mathrm{LP}\left\{\mathrm{Re}\left\{(1+\alpha)\,y_0(t)e^{j\left(4\pi f_c t + \frac{\theta}{2} - \phi(t)\right)} + (1+\alpha)\,y_0(t)e^{-j\left(\frac{\theta}{2} - \phi(t)\right)}\right\}\right\}\\
&\quad + \mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\\
&= \mathrm{Re}\left\{(1+\alpha)\,y_0(t)e^{-j\left(\frac{\theta}{2} - \phi(t)\right)}\right\}
\end{aligned}$$

$$+ \mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}$$

$$= (1+\alpha)\cos\left(\frac{\theta}{2} - \phi(t)\right)\mathrm{Re}\left\{y_0(t)\right\} + (1+\alpha)\sin\left(\frac{\theta}{2} - \phi(t)\right)$$

$$\cdot \mathrm{Im}\left\{y_0(t)\right\} + \mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}, \qquad \text{(C.1)}$$

where

$$\mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}$$

is the additive Gaussian noise in the I branch.

Similarly, it can be shown that the Q-component of the received baseband signal is given by

$$y_Q(t) = (1-\alpha)\sin\left(\frac{\theta}{2} + \phi(t)\right)\mathrm{Re}\left\{y_0(t)\right\} + (1-\alpha)\cos\left(\frac{\theta}{2} + \phi(t)\right)$$

$$\cdot \mathrm{Im}\left\{y_0(t)\right\} + \mathrm{LP}\left\{-w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right)\right\}, \qquad \text{(C.2)}$$

where

$$\mathrm{LP}\left\{-w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right)\right\}$$

is the additive Gaussian noise in the Q branch.

Combining (C.1) and (C.2), we have

$$y(t) = y_I(t) + jy_Q(t)$$

$$= \left[(1+\alpha)\cos\left(\frac{\theta}{2} - \phi(t)\right)\mathrm{Re}\left\{y_0(t)\right\} + (1+\alpha)\sin\left(\frac{\theta}{2} - \phi(t)\right)\mathrm{Im}\left\{y_0(t)\right\}\right.$$

$$\left. + \mathrm{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\}\right]$$

$$+ j\left[(1-\alpha)\sin\left(\frac{\theta}{2} + \phi(t)\right)\mathrm{Re}\left\{y_0(t)\right\} + (1-\alpha)\cos\left(\frac{\theta}{2} + \phi(t)\right)\right.$$

$$\left. \cdot \mathrm{Im}\left\{y_0(t)\right\} + \mathrm{LP}\left\{-w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right)\right\}\right]$$

$$= \left[(1+\alpha)\cos\left(\frac{\theta}{2} - \phi(t)\right) + j(1-\alpha)\sin\left(\frac{\theta}{2} + \phi(t)\right) \right] \mathrm{Re}\left\{y_0(t)\right\}$$

$$+ \left[(1+\alpha)\sin\left(\frac{\theta}{2} - \phi(t)\right) + j(1-\alpha)\cos\left(\frac{\theta}{2} + \phi(t)\right) \right] \mathrm{Im}\left\{y_0(t)\right\}$$

$$+ \mathrm{LP}\left\{ w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right) \right\}$$

$$+ j\mathrm{LP}\left\{ -w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right) \right\}$$

$$= \frac{1}{2}\left[(1+\alpha)\cos\left(\frac{\theta}{2} - \phi(t)\right) + j(1-\alpha)\sin\left(\frac{\theta}{2} + \phi(t)\right) \right] [y_0(t) + y_0^*(t)]$$

$$+ \frac{1}{2j}\left[(1+\alpha)\sin\left(\frac{\theta}{2} - \phi(t)\right) + j(1-\alpha)\cos\left(\frac{\theta}{2} + \phi(t)\right) \right] [y_0(t) - y_0^*(t)]$$

$$+ \mathrm{LP}\left\{ w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right) \right\}$$

$$+ j\mathrm{LP}\left\{ -w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right) \right\}$$

$$= \left[\frac{1+\alpha}{2}\cos\left(\frac{\theta}{2} - \phi(t)\right) + j\frac{1-\alpha}{2}\sin\left(\frac{\theta}{2} + \phi(t)\right) - j\frac{1+\alpha}{2} \right.$$

$$\left. \cdot \sin\left(\frac{\theta}{2} - \phi(t)\right) + \frac{1-\alpha}{2}\cos\left(\frac{\theta}{2} + \phi(t)\right) \right] y_0(t) + \left[\frac{1+\alpha}{2} \right.$$

$$\left. \cdot \cos\left(\frac{\theta}{2} - \phi(t)\right) + j\frac{1-\alpha}{2}\sin\left(\frac{\theta}{2} + \phi(t)\right) + j\frac{1+\alpha}{2}\sin\left(\frac{\theta}{2} - \phi(t)\right) \right.$$

$$\left. - \frac{1-\alpha}{2}\cos\left(\frac{\theta}{2} + \phi(t)\right) \right] y_0^*(t)$$

$$+ \mathrm{LP}\left\{ w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right) \right\}$$

$$+ j\mathrm{LP}\left\{ -w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right) \right\}$$

$$= \left[\cos\left(\frac{\theta}{2}\right) - j\alpha\sin\left(\frac{\theta}{2}\right) \right] e^{j\phi(t)} y_0(t) + \left[\alpha\cos\left(\frac{\theta}{2}\right) + j\sin\left(\frac{\theta}{2}\right) \right]$$

$$\cdot e^{-j\phi(t)} y_0^*(t) + \mathrm{LP}\left\{ w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right) \right\}$$

$$+ j\mathrm{LP}\left\{ -w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right) \right\}.$$

Letting $w(t) = w_I(t) + jw_Q(t)$ with

$$w_I(t) = \text{LP}\left\{w_p(t)(1+\alpha)\sqrt{2}\cos\left(2\pi f_c t + \frac{\theta}{2} - \phi(t)\right)\right\},$$

$$w_Q(t) = \text{LP}\left\{-w_p(t)(1-\alpha)\sqrt{2}\sin\left(2\pi f_c t - \frac{\theta}{2} - \phi(t)\right)\right\},$$

we have expression (3.6), i.e.,

$$y(t) = \left[\cos\left(\frac{\theta}{2}\right) - j\alpha\sin\left(\frac{\theta}{2}\right)\right]e^{j\phi(t)}y_0(t)$$

$$+ \left[\alpha\cos\left(\frac{\theta}{2}\right) + j\sin\left(\frac{\theta}{2}\right)\right]e^{-j\phi(t)}y_0^*(t) + w(t).$$

Given that $w_p(t)$ is white Gaussian, $w_I(t)$ and $w_Q(t)$ are also Gaussian, but not necessarily independent of each other, because the oscillator signals in the I and Q branches may not be orthogonal to each other due to the distortions. However, since the magnitudes of these distortions are usually small, $w_I(t)$ and $w_Q(t)$ are approximately independent of each other with equal variance. For this reason, we treat $w(t)$ approximately as circularly symmetric Gaussian.

APPENDIX D

Formulate (3.22) as a Standard Least-Squares Problem

In the following, we illustrate how to formulate the problem given by (3.22) into a standard least-squares problem. Let

$$
\mathbf{b} = \mathbf{y} - \left(\mathbf{A}''_{i-1}\mathbf{H}''_{i-1}\mathbf{x} + \widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\} \right).
$$

The objective function can be expanded as

$$
\begin{aligned}
& \left\| \mathbf{b} - \left[(\Delta\mathbf{A}'')\mathbf{H}''_{i-1}\mathbf{x} + \mathbf{A}''_{i-1}(\Delta\mathbf{H}'')\mathbf{x} \right] - \left[(\Delta\nu'')\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\} \right. \right. \\
& \left. \left. + \widehat{\nu}''_{i-1}(\Delta\widetilde{\mathbf{A}''}) \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\} + \widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\Delta\mathbf{H}''\} \cdot \mathrm{conj}\{\mathbf{x}\} \right] \right\|^2 \\
= & \left\| \mathrm{Re}\{\mathbf{b}\} - \left[\mathrm{Re}\{(\Delta\mathbf{A}'')\mathbf{H}''_{i-1}\mathbf{x}\} + \mathrm{Re}\{\mathbf{A}''_{i-1}(\Delta\mathbf{H}'')\mathbf{x}\} \right. \right. \\
& \quad + \mathrm{Re}\{(\Delta\nu'')\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\}\} \\
& \quad + \mathrm{Re}\{\widehat{\nu}''_{i-1}(\Delta\widetilde{\mathbf{A}''}) \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\}\} \\
& \left. \quad + \mathrm{Re}\{\widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\Delta\mathbf{H}''\} \cdot \mathrm{conj}\{\mathbf{x}\}\} \right] \Big\|^2 \\
& + \left\| \mathrm{Im}\{\mathbf{b}\} - \left[\mathrm{Im}\{(\Delta\mathbf{A}'')\mathbf{H}''_{i-1}\mathbf{x}\} + \mathrm{Im}\{\mathbf{A}''_{i-1}(\Delta\mathbf{H}'')\mathbf{x}\} \right. \right. \\
& \quad + \mathrm{Im}\{(\Delta\nu'')\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\}\} \\
& \quad + \mathrm{Im}\{\widehat{\nu}''_{i-1}(\Delta\widetilde{\mathbf{A}''}) \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\}\} \\
& \left. \quad + \mathrm{Im}\{\widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\Delta\mathbf{H}''\} \cdot \mathrm{conj}\{\mathbf{x}\}\} \right] \Big\|^2,
\end{aligned}
\tag{D.1}
$$

where

$$\Delta\mathbf{A}'' = \begin{bmatrix} \Delta A''[0] & \Delta A''[N-1] & \dots & \Delta A''[1] \\ \Delta A''[1] & \Delta A''[0] & \dots & \Delta A''[2] \\ \vdots & \vdots & \ddots & \vdots \\ \Delta A''[N-1] & \Delta A''[N-2] & \dots & \Delta A''[0] \end{bmatrix},$$

$$\Delta\mathbf{H}'' = \begin{bmatrix} \Delta H''[0] & 0 & \dots & 0 \\ 0 & \Delta H''[1] & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Delta H''[N-1] \end{bmatrix},$$

$$\Delta\widetilde{\mathbf{A}''} = \begin{bmatrix} (\Delta A''[0])^* & (\Delta A''[N-1])^* & \dots & (\Delta A''[1])^* \\ (\Delta A''[N-1])^* & (\Delta A''[N-2])^* & \dots & (\Delta A''[0])^* \\ \vdots & \vdots & \ddots & \vdots \\ (\Delta A''[1])^* & (\Delta A''[0])^* & \dots & (\Delta A''[2])^* \end{bmatrix}.$$

Let

$$\Delta\bar{\mathbf{a}} = \begin{bmatrix} \Delta A''[0] \\ \Delta A''[1] \\ \vdots \\ \Delta A''[N-1] \end{bmatrix}, \quad \Delta\bar{\mathbf{h}} = \begin{bmatrix} \Delta H''[0] \\ \Delta H''[1] \\ \vdots \\ \Delta H''[N-1] \end{bmatrix},$$

and

$$\Delta\mathbf{c}'' = \begin{bmatrix} \Delta c''[0] \\ \Delta c''[1] \\ \vdots \\ \Delta c''[M-1] \end{bmatrix}, \quad \Delta\mathbf{h}'' = \begin{bmatrix} \Delta h''[0] \\ \Delta h''[1] \\ \vdots \\ \Delta h''[L-1] \end{bmatrix}.$$

By (3.18) and (3.19),

$$\mathrm{Re}\{\Delta\overline{\mathbf{a}}\} = \mathrm{Re}\left\{\frac{1}{N}\mathbf{F_aP}\Delta\mathbf{c}''\right\} = \frac{1}{N}\mathrm{Re}\{\mathbf{F_aP}\}\mathrm{Re}\{\Delta\mathbf{c}''\} - \frac{1}{N}\mathrm{Im}\{\mathbf{F_aP}\}\mathrm{Im}\{\Delta\mathbf{c}''\},$$

$$\mathrm{Im}\{\Delta\overline{\mathbf{a}}\} = \mathrm{Im}\left\{\frac{1}{N}\mathbf{F_aP}\Delta\mathbf{c}''\right\} = \frac{1}{N}\mathrm{Re}\{\mathbf{F_aP}\}\mathrm{Im}\{\Delta\mathbf{c}''\} + \frac{1}{N}\mathrm{Im}\{\mathbf{F_aP}\}\mathrm{Re}\{\Delta\mathbf{c}''\},$$

and

$$\mathrm{Re}\{\Delta\overline{\mathbf{h}}\} = \mathrm{Re}\{\mathbf{F_h}\Delta\mathbf{h}''\} = \mathrm{Re}\{\mathbf{F_h}\}\mathrm{Re}\{\Delta\mathbf{h}''\} - \mathrm{Im}\{\mathbf{F_h}\}\mathrm{Im}\{\Delta\mathbf{h}''\},$$

$$\mathrm{Im}\{\Delta\overline{\mathbf{h}}\} = \mathrm{Im}\{\mathbf{F_h}\Delta\mathbf{h}''\} = \mathrm{Re}\{\mathbf{F_h}\}\mathrm{Im}\{\Delta\mathbf{h}''\} + \mathrm{Im}\{\mathbf{F_h}\}\mathrm{Re}\{\Delta\mathbf{h}''\}.$$

Thus, we have

$$
\begin{bmatrix}
\mathrm{Re}\{\Delta\nu''\} \\
\mathrm{Im}\{\Delta\nu''\} \\
\mathrm{Re}\{\Delta\overline{\mathbf{h}}\} \\
\mathrm{Im}\{\Delta\overline{\mathbf{h}}\} \\
\mathrm{Re}\{\Delta\overline{\mathbf{a}}\} \\
\mathrm{Im}\{\Delta\overline{\mathbf{a}}\}
\end{bmatrix}
= \mathbf{R_1}
\begin{bmatrix}
\mathrm{Re}\{\Delta\nu''\} \\
\mathrm{Im}\{\Delta\nu''\} \\
\mathrm{Re}\{\Delta\mathbf{h}''\} \\
\mathrm{Im}\{\Delta\mathbf{h}''\} \\
\mathrm{Re}\{\Delta\mathbf{c}''\} \\
\mathrm{Im}\{\Delta\mathbf{c}''\}
\end{bmatrix}.
\tag{D.2}
$$

where

$$
\mathbf{R_1} =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & \mathrm{Re}\{\mathbf{F_h}\} & -\mathrm{Im}\{\mathbf{F_h}\} & 0 & 0 \\
0 & 0 & \mathrm{Im}\{\mathbf{F_h}\} & \mathrm{Re}\{\mathbf{F_h}\} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{N}\mathrm{Re}\{\mathbf{F_aP}\} & -\frac{1}{N}\mathrm{Im}\{\mathbf{F_aP}\} \\
0 & 0 & 0 & 0 & \frac{1}{N}\mathrm{Im}\{\mathbf{F_aP}\} & \frac{1}{N}\mathrm{Re}\{\mathbf{F_aP}\}
\end{bmatrix}.
$$

Let

$$
\Delta\mathbf{c}'' =
\begin{bmatrix}
\Delta c''[0] \\
\Delta c''[1] \\
\vdots \\
\Delta c''[M-1]
\end{bmatrix}
$$

166

and

$$\Delta \mathbf{c}''' = \begin{bmatrix} \Delta c''[1] \\ \Delta c''[2] \\ \vdots \\ \Delta c''[M-1] \end{bmatrix}.$$

Denote the first row of $\frac{1}{N}\mathbf{F_a P}$ by $\mathbf{g}$. Since $\mathbf{g}\Delta \mathbf{c}'' = 0$, we have

$$\Delta c''[0] = -\frac{1}{g[1]}\mathbf{g}[2:M]\Delta \mathbf{c}''',$$

where $g[1]$ is the first element of $\mathbf{g}$, and $\mathbf{g}[2:M]$ is the sub-vector of $\mathbf{g}$ that contains all elements of $\mathbf{g}$ except $g[1]$. It then follows that

$$\Delta \mathbf{c}'' = \begin{bmatrix} -\frac{1}{g[1]}\mathbf{g}[2:M] \\ \mathbf{I}_{M-1} \end{bmatrix} \Delta \mathbf{c}'''.$$

Hence,

$$\begin{bmatrix} \mathrm{Re}\{\Delta \mathbf{c}''\} \\ \mathrm{Im}\{\Delta \mathbf{c}''\} \end{bmatrix} = \mathbf{R}_2 \begin{bmatrix} \mathrm{Re}\{\Delta \mathbf{c}'''\} \\ \mathrm{Im}\{\Delta \mathbf{c}'''\} \end{bmatrix},$$

where

$$\mathbf{R}_2 = \begin{bmatrix} -\mathrm{Re}\left\{\frac{1}{g[1]}\mathbf{g}[2:M]\right\} & \mathrm{Im}\left\{\frac{1}{g[1]}\mathbf{g}[2:M]\right\} \\ \mathbf{I}_{M-1} & 0 \\ -\mathrm{Im}\left\{\frac{1}{g[1]}\mathbf{g}[2:M]\right\} & -\mathrm{Re}\left\{\frac{1}{g[1]}\mathbf{g}[2:M]\right\} \\ 0 & \mathbf{I}_{M-1} \end{bmatrix}.$$

Moreover,

$$\begin{bmatrix} \mathrm{Re}\{\Delta \nu''\} \\ \mathrm{Im}\{\Delta \nu''\} \\ \mathrm{Re}\{\Delta \mathbf{h}''\} \\ \mathrm{Im}\{\Delta \mathbf{h}''\} \\ \mathrm{Re}\{\Delta \mathbf{c}''\} \\ \mathrm{Im}\{\Delta \mathbf{c}''\} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{2L+2} & 0 \\ 0 & \mathbf{R}_2 \end{bmatrix} \begin{bmatrix} \mathrm{Re}\{\Delta \nu''\} \\ \mathrm{Im}\{\Delta \nu''\} \\ \mathrm{Re}\{\Delta \mathbf{h}''\} \\ \mathrm{Im}\{\Delta \mathbf{h}''\} \\ \mathrm{Re}\{\Delta \mathbf{c}'''\} \\ \mathrm{Im}\{\Delta \mathbf{c}'''\} \end{bmatrix}. \tag{D.3}$$

Consequently, the original problem is formulated as the following least-squares problem:

$$\min_{\overline{\mathbf{x}}} \| \overline{\mathbf{y}} - \overline{\mathbf{A}}\overline{\mathbf{x}} \|^2, \qquad (D.4)$$

where

$$\overline{\mathbf{y}} = \begin{bmatrix} \mathrm{Re}\{\mathbf{b}\} \\ \mathrm{Im}\{\mathbf{b}\} \end{bmatrix}, \quad \overline{\mathbf{x}} = \begin{bmatrix} \mathrm{Re}\{\Delta\nu''\} \\ \mathrm{Im}\{\Delta\nu''\} \\ \mathrm{Re}\{\Delta\mathbf{h}''\} \\ \mathrm{Im}\{\Delta\mathbf{h}''\} \\ \mathrm{Re}\{\Delta\mathbf{c}'''\} \\ \mathrm{Im}\{\Delta\mathbf{c}'''\} \end{bmatrix},$$

and matrix $\overline{\mathbf{A}}$ is formed according to expressions (D.1)-(D.3), i.e.,

$$\overline{\mathbf{A}} = \begin{bmatrix} \overline{\mathbf{A}}_{11} & \overline{\mathbf{A}}_{12} & \overline{\mathbf{A}}_{13} & \overline{\mathbf{A}}_{14} & \overline{\mathbf{A}}_{15} & \overline{\mathbf{A}}_{16} \\ \overline{\mathbf{A}}_{21} & \overline{\mathbf{A}}_{22} & \overline{\mathbf{A}}_{23} & \overline{\mathbf{A}}_{24} & \overline{\mathbf{A}}_{25} & \overline{\mathbf{A}}_{26} \end{bmatrix} \mathbf{R}_1 \begin{bmatrix} \mathbf{I}_{2L+2} & 0 \\ 0 & \mathbf{R}_2 \end{bmatrix}$$

where

$$\overline{\mathbf{A}}_{11} = \mathrm{Re}\big\{ \widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\} \big\},$$

$$\overline{\mathbf{A}}_{12} = - \,\mathrm{Im}\big\{ \widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\} \big\},$$

$$\overline{\mathbf{A}}_{13} = \mathrm{Re}\big\{ \mathbf{A}''_{i-1}\mathbf{X} \big\} + \mathrm{Re}\big\{ \widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{X}\} \big\},$$

$$\overline{\mathbf{A}}_{14} = - \,\mathrm{Im}\big\{ \mathbf{A}''_{i-1}\mathbf{X} \big\} + \mathrm{Im}\big\{ \widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{X}\} \big\},$$

$$\overline{\mathbf{A}}_{15} = \mathrm{Re}\big\{ \mathcal{T}\big(\mathbf{H}''_{i-1}\mathbf{x}\big) \big\} + \mathrm{Re}\big\{ \widetilde{\mathcal{T}}\big(\widehat{\nu}''_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\}\big) \big\},$$

$$\overline{\mathbf{A}}_{16} = - \,\mathrm{Im}\big\{ \mathcal{T}\big(\mathbf{H}''_{i-1}\mathbf{x}\big) \big\} + \mathrm{Im}\big\{ \widetilde{\mathcal{T}}\big(\widehat{\nu}''_{i-1} \cdot \mathrm{conj}\{\mathbf{H}''_{i-1}\} \cdot \mathrm{conj}\{\mathbf{x}\}\big) \big\},$$

and

$$\overline{\mathbf{A}}_{21} = \mathrm{Im}\big\{\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\big\{\mathbf{H}''_{i-1}\big\} \cdot \mathrm{conj}\{\mathbf{x}\}\big\},$$

$$\overline{\mathbf{A}}_{22} = \mathrm{Re}\big\{\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\big\{\mathbf{H}''_{i-1}\big\} \cdot \mathrm{conj}\{\mathbf{x}\}\big\},$$

$$\overline{\mathbf{A}}_{23} = \mathrm{Im}\big\{\mathbf{A}''_{i-1}\mathbf{X}\big\} + \mathrm{Im}\big\{\widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{X}\}\big\},$$

$$\overline{\mathbf{A}}_{24} = \mathrm{Re}\big\{\mathbf{A}''_{i-1}\mathbf{X}\big\} - \mathrm{Re}\big\{\widehat{\nu}''_{i-1}\widetilde{\mathbf{A}''}_{i-1} \cdot \mathrm{conj}\{\mathbf{X}\}\big\},$$

$$\overline{\mathbf{A}}_{25} = \mathrm{Im}\big\{\mathcal{T}\big(\mathbf{H}''_{i-1}\mathbf{x}\big)\big\} + \mathrm{Im}\big\{\widetilde{\mathcal{T}}\big(\widehat{\nu}''_{i-1} \cdot \mathrm{conj}\big\{\mathbf{H}''_{i-1}\big\} \cdot \mathrm{conj}\{\mathbf{x}\}\big)\big\},$$

$$\overline{\mathbf{A}}_{26} = \mathrm{Re}\big\{\mathcal{T}\big(\mathbf{H}''_{i-1}\mathbf{x}\big)\big\} - \mathrm{Re}\big\{\widetilde{\mathcal{T}}\big(\widehat{\nu}''_{i-1} \cdot \mathrm{conj}\big\{\mathbf{H}''_{i-1}\big\} \cdot \mathrm{conj}\{\mathbf{x}\}\big)\big\}.$$

Here, $\mathcal{T}(\cdot)$ and $\widetilde{\mathcal{T}}(\cdot)$ are defined as follows. Let $\mathbf{v}$ be an arbitrary vector of length N that is represented by

$$\mathbf{v} = \begin{bmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_{N-1} \end{bmatrix}.$$

$\mathcal{T}(\mathbf{v})$ and $\widetilde{\mathcal{T}}(\mathbf{v})$ are $N \times N$ matrices that are given by

$$\mathcal{T}(\mathbf{v}) = \begin{bmatrix} \xi_0 & \xi_{N-1} & \cdots & \xi_1 \\ \xi_1 & \xi_0 & \cdots & \xi_2 \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{N-1} & \xi_{N-2} & \cdots & \xi_0 \end{bmatrix}, \quad \widetilde{\mathcal{T}}(\mathbf{v}) = \begin{bmatrix} \xi_0 & \xi_{N-1} & \cdots & \xi_1 \\ \xi_{N-1} & \xi_{N-2} & \cdots & \xi_0 \\ \vdots & \vdots & \ddots & \vdots \\ \xi_1 & \xi_0 & \cdots & \xi_2 \end{bmatrix}.$$

The solution of (D.4) is

$$\overline{\mathbf{x}}_o = (\overline{\mathbf{A}}^T\overline{\mathbf{A}})^{-1}\overline{\mathbf{A}}^T\overline{\mathbf{y}}.$$

APPENDIX E

Cramer-Rao Lower Bound (CRLB)

Consider the data model given by (3.31). The joint probability density function (PDF) of $\mathbf{y}$ is

$$p(\mathbf{y}; \boldsymbol{\theta}) = \prod_{k=0}^{N-1} \frac{1}{\pi \sigma_W^2} \exp \left\{ -\frac{1}{\sigma_W^2} \left[(\text{Re}\{Y[k]\} - \text{Re}\{S_{\boldsymbol{\theta}}[k]\})^2 \right. \right.$$
$$\left. \left. + (\text{Im}\{Y[k]\} - \text{Im}\{S_{\boldsymbol{\theta}}[k]\})^2 \right] \right\}.$$

The Log-Likelihood function is then given by

$$\ln p(\mathbf{y}; \boldsymbol{\theta}) = -N \ln \left(\pi \sigma_W^2 \right) - \frac{1}{\sigma_W^2} \sum_{k=0}^{N-1} \left[(\text{Re}\{Y[k]\} - \text{Re}\{S_{\boldsymbol{\theta}}[k]\})^2 + (\text{Im}\{Y[k]\} \right.$$
$$\left. - \text{Im}\{S_{\boldsymbol{\theta}}[k]\})^2 \right].$$

Taking the derivative of $\ln p(\mathbf{y}; \boldsymbol{\theta})$ with respect to the elements of $\boldsymbol{\theta}$ gives

$$\frac{\partial \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_{i_1}} = \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \left[(\text{Re}\{Y[k]\} - \text{Re}\{S_{\boldsymbol{\theta}}[k]\}) \frac{\partial \text{Re}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_1}} \right.$$
$$\left. + (\text{Im}\{Y[k]\} - \text{Im}\{S_{\boldsymbol{\theta}}[k]\}) \frac{\partial \text{Im}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_1}} \right],$$

and

$$\frac{\partial^2 \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_{i_1} \partial \theta_{i_2}} = \frac{2}{\sigma_W^2} \sum_{k=0}^{N-1} \left[(\text{Re}\{Y[k]\} - \text{Re}\{S_{\boldsymbol{\theta}}[k]\}) \frac{\partial^2 \text{Re}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_1} \partial \theta_{i_2}} \right.$$
$$- \frac{\partial \text{Re}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_1}} \frac{\partial \text{Re}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_2}} + (\text{Im}\{Y[k]\} - \text{Im}\{S_{\boldsymbol{\theta}}[k]\})$$
$$\left. \cdot \frac{\partial^2 \text{Im}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_1} \partial \theta_{i_2}} - \frac{\partial \text{Im}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_1}} \frac{\partial \text{Im}\{S_{\boldsymbol{\theta}}[k]\}}{\partial \theta_{i_2}} \right].$$

Taking the expectation obtains

$$\mathbf{E}\left\{\frac{\partial^2 \ln p(\mathbf{y};\boldsymbol{\theta})}{\partial\theta_{i_1}\partial\theta_{i_2}}\right\} = -\frac{2}{\sigma_W^2}\sum_{k=0}^{N-1}\left[\frac{\partial\mathrm{Re}\left\{S_{\boldsymbol{\theta}}[k]\right\}}{\partial\theta_{i_1}}\frac{\partial\mathrm{Re}\left\{S_{\boldsymbol{\theta}}[k]\right\}}{\partial\theta_{i_2}}\right.$$
$$\left.+\frac{\partial\mathrm{Im}\left\{S_{\boldsymbol{\theta}}[k]\right\}}{\partial\theta_{i_1}}\frac{\partial\mathrm{Im}\left\{S_{\boldsymbol{\theta}}[k]\right\}}{\partial\theta_{i_2}}\right].$$

It then follows that

$$-\mathbf{E}\left\{\frac{\partial^2 \ln p(\mathbf{y};\boldsymbol{\theta})}{\partial\theta_{i_1}\partial\theta_{i_2}}\right\} = \frac{2}{\sigma_W^2}\sum_{k=0}^{N-1}\mathrm{Re}\left\{\frac{\partial S_{\boldsymbol{\theta}}[k]}{\partial\theta_{i_1}}\left(\frac{\partial S_{\boldsymbol{\theta}}[k]}{\partial\theta_{i_2}}\right)^*\right\}.$$

Thus, the Fisher information matrix is given by

$$\mathbf{I}_{\boldsymbol{\theta}} = \frac{2}{\sigma_W^2}\sum_{k=0}^{N-1}\mathrm{Re}\left\{\frac{\partial S_{\boldsymbol{\theta}}[k]}{\partial\boldsymbol{\theta}}\left(\frac{\partial S_{\boldsymbol{\theta}}[k]}{\partial\boldsymbol{\theta}}\right)^*\right\}.$$

APPENDIX F

Computation of the CRLB for Joint Channel Estimation

To compute $\mathbf{I}_\theta$, we need to calculate the derivatives of $\mathbf{s}_\theta$ with respect to the elements of $\boldsymbol{\theta}$. Recall that

$$\boldsymbol{\theta} = [\ \mathrm{Re}\{\nu''\}\ \ \mathrm{Im}\{\nu''\}\ \ \mathrm{Re}\{c''[1]\}\ \ \ldots\ \ \mathrm{Re}\{c''[M-1]\}\ \ \mathrm{Im}\{c''[1]\}\ \ \ldots$$

$$\mathrm{Im}\{c''[M-1]\}\ \ \mathrm{Re}\{h''[0]\}\ \ \ldots\ \ \mathrm{Re}\{h''[L-1]\}\ \ \mathrm{Im}\{h''[0]\}\ \ \ldots$$

$$\mathrm{Im}\{h''[L-1]\}\ ]^T.$$

By the chain rule,

$$\frac{\partial \mathbf{s}_\theta}{\partial \theta_i} = \frac{\partial \mathbf{A}''_{\mathrm{appro}}}{\partial \theta_i}\mathbf{H}''\mathbf{x} + \mathbf{A}''_{\mathrm{appro}}\frac{\partial \mathbf{H}''}{\partial \theta_i}\mathbf{x} + \frac{\partial \nu''}{\partial \theta_i}\widetilde{\mathbf{A}}''_{\mathrm{appro}} \cdot \mathrm{conj}\{\mathbf{H}''\} \cdot \mathrm{conj}\{\mathbf{x}\} \quad \text{(F.1)}$$

$$+ \nu''\frac{\partial \widetilde{\mathbf{A}}''_{\mathrm{appro}}}{\partial \theta_i} \cdot \mathrm{conj}\{\mathbf{H}''\} \cdot \mathrm{conj}\{\mathbf{x}\} + \nu''\widetilde{\mathbf{A}}''_{\mathrm{appro}} \cdot \frac{\partial\, \mathrm{conj}\{\mathbf{H}''\}}{\partial \theta_i} \cdot \mathrm{conj}\{\mathbf{x}\},$$

where the derivative terms are computed as follows:

1.

$$\frac{\partial \nu''}{\partial \mathrm{Re}\,\{\nu''\}} = 1, \quad \frac{\partial \nu''}{\partial \mathrm{Im}\,\{\nu''\}} = j, \quad \frac{\partial \nu''}{\partial \theta_i} = 0 \text{ for other } \theta_i.$$

2. Denote the first row of matrix $\frac{1}{N}\mathbf{F}_\mathbf{a}\mathbf{P}$ by $\mathbf{g}$. Recall that the first element of $\overline{\mathbf{a}}_{\mathrm{appro}}$ is restricted to be 1 to resolve the scalar ambiguity, i.e., $\mathbf{g}\mathbf{c}'' = 1$. The first element of $\mathbf{c}''$, $c''[0]$, is given by

$$c''[0] = \frac{1}{g[1]} - \frac{1}{g[1]}\mathbf{g}[2:M]\mathbf{c}''',$$

172

where $\mathbf{c}'''$ is the vector consisting of all elements of $\mathbf{c}''$ except $c''[0]$. Hence,

$$\overline{\mathbf{a}}_{\text{appro}} = \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{c}'' = \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}(\mathbf{g}_1 + \mathbf{G}_2\mathbf{c}'''),$$

where

$$\mathbf{g}_1 = \begin{bmatrix} \frac{1}{g[1]} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ (M \times 1), \quad \mathbf{G}_2 = \begin{bmatrix} -\frac{1}{g[1]}\mathbf{g}[2:M] \\ \mathbf{I}_{M-1} \end{bmatrix} \ (M \times (M-1)).$$

Here, $g[1]$ is the first element of $\mathbf{g}$, and $\mathbf{g}[2:M]$ is the vector consisting of all elements of $\mathbf{g}$ except $g[1]$. Thus, for $m = 1, 2, \ldots, M-1$,

$$\frac{\partial \overline{\mathbf{a}}_{\text{appro}}}{\partial \mathrm{Re}\,\{c''[m]\}} = \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{G}_2 \cdot \frac{\partial \mathbf{c}'''}{\partial \mathrm{Re}\,\{c''[m]\}}$$

$$= \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{G}_2\mathbf{e}_m$$

$$= \text{the } m^{th} \text{ column of } \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{G}_2,$$

$$\frac{\partial \overline{\mathbf{a}}_{\text{appro}}}{\partial \mathrm{Im}\,\{c''[m]\}} = \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{G}_2 \cdot \frac{\partial \mathbf{c}'''}{\partial \mathrm{Im}\,\{c''[m]\}}$$

$$= \frac{j}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{G}_2\mathbf{e}_m$$

$$= j \cdot \text{the } m^{th} \text{ column of } \frac{1}{N}\mathbf{F}_{\mathbf{a}}\mathbf{P}\mathbf{G}_2,$$

$$\frac{\partial \overline{\mathbf{a}}_{\text{appro}}}{\partial \theta_i} = 0 \text{ for other } \theta_i.$$

3. Since $\overline{\mathbf{h}} = \mathbf{F_h}\mathbf{h}''$, then for $n = 0, 1, \ldots, L-1$,

$$\frac{\partial \overline{\mathbf{h}}}{\partial \mathrm{Re}\,\{h''[n]\}} = \mathbf{F_h} \cdot \frac{\partial \mathbf{h}''}{\partial \mathrm{Re}\,\{h''[n]\}}$$

$$= \mathbf{F_h}\mathbf{e}_{n+1}$$

$$= \text{the } (n+1)^{th} \text{ column of } \mathbf{F_h},$$

$$\frac{\partial \overline{\mathbf{h}}}{\partial \mathrm{Im}\,\{h''[n]\}} = \mathbf{F_h} \cdot \frac{\partial \mathbf{h}''}{\partial \mathrm{Im}\,\{h''[n]\}}$$

$$= j\mathbf{F_h}\mathbf{e}_{n+1}$$

$$= j \cdot \text{the } (n+1)^{th} \text{ column of } \mathbf{F_h},$$

$$\frac{\partial \overline{\mathbf{h}}}{\partial \theta_i} = 0 \text{ for other } \theta_i.$$

Consequently, $\mathbf{I}_\theta$ can be computed using (3.32) and (F.1).

APPENDIX G

Derivation of Expression (4.19)

Consider the data model given by (4.18). The joint probability density function (PDF) of $\mathbf{y}$ is

$$p(\mathbf{y}; \boldsymbol{\theta}) = \prod_{n=0}^{N-1} \frac{1}{\pi \sigma_{w,n}^2}$$

$$\cdot \exp \left\{ -\frac{1}{\sigma_{w,n}^2} \left[(\text{Re}\{y[n]\} - \text{Re}\{s_{\boldsymbol{\theta}}[n]\})^2 + (\text{Im}\{y[n]\} - \text{Im}\{s_{\boldsymbol{\theta}}[n]\})^2 \right] \right\}.$$

The Log-Likelihood function is then given by

$$\ln p(\mathbf{y}; \boldsymbol{\theta}) = -\sum_{n=0}^{N-1} \ln \left(\pi \sigma_{w,n}^2 \right) - \sum_{n=0}^{N-1} \frac{1}{\sigma_{w,n}^2} \left[(\text{Re}\{y[n]\} - \text{Re}\{s_{\boldsymbol{\theta}}[n]\})^2 + (\text{Im}\{y[n]\} \right.$$

$$\left. - \text{Im}\{s_{\boldsymbol{\theta}}[n]\})^2 \right].$$

Taking the derivative of $\ln p(\mathbf{y}; \boldsymbol{\theta})$ with respect to the elements of $\boldsymbol{\theta}$ gives

$$\frac{\partial \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_{i_1}} = \sum_{n=0}^{N-1} \frac{2}{\sigma_{w,n}^2} \left[(\text{Re}\{y[n]\} - \text{Re}\{s_{\boldsymbol{\theta}}[n]\}) \frac{\partial \text{Re}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_1}} \right.$$

$$\left. + (\text{Im}\{y[n]\} - \text{Im}\{s_{\boldsymbol{\theta}}[n]\}) \frac{\partial \text{Im}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_1}} \right],$$

and

$$\frac{\partial^2 \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_{i_1} \partial \theta_{i_2}} = \sum_{n=0}^{N-1} \frac{2}{\sigma_{w,n}^2} \left[(\text{Re}\{y[n]\} - \text{Re}\{s_{\boldsymbol{\theta}}[n]\}) \frac{\partial^2 \text{Re}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_1} \partial \theta_{i_2}} \right.$$

$$- \frac{\partial \text{Re}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_1}} \frac{\partial \text{Re}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_2}} + (\text{Im}\{y[n]\} - \text{Im}\{s_{\boldsymbol{\theta}}[n]\})$$

$$\left. \cdot \frac{\partial^2 \text{Im}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_1} \partial \theta_{i_2}} - \frac{\partial \text{Im}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_1}} \frac{\partial \text{Im}\{s_{\boldsymbol{\theta}}[n]\}}{\partial \theta_{i_2}} \right].$$

Taking the expectation obtains

$$
\mathbf{E} \left\{ \frac{\partial^2 \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_{i_1} \partial \theta_{i_2}} \right\} = - \sum_{n=0}^{N-1} \frac{2}{\sigma_{w,n}^2} \left[\frac{\partial \mathrm{Re} \left\{ s_{\boldsymbol{\theta}}[n] \right\}}{\partial \theta_{i_1}} \frac{\partial \mathrm{Re} \left\{ s_{\boldsymbol{\theta}}[n] \right\}}{\partial \theta_{i_2}} \right.
$$
$$
\left. + \frac{\partial \mathrm{Im} \left\{ s_{\boldsymbol{\theta}}[n] \right\}}{\partial \theta_{i_1}} \frac{\partial \mathrm{Im} \left\{ s_{\boldsymbol{\theta}}[n] \right\}}{\partial \theta_{i_2}} \right] .
$$

It then follows that

$$
-\mathbf{E} \left\{ \frac{\partial^2 \ln p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_{i_1} \partial \theta_{i_2}} \right\} = \sum_{n=0}^{N-1} \frac{2}{\sigma_{w,n}^2} \mathrm{Re} \left\{ \frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \theta_{i_1}} \left(\frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \theta_{i_2}} \right)^* \right\} .
$$

Thus, the Fisher information matrix is given by

$$
\mathbf{I}_{\boldsymbol{\theta}} = \sum_{n=0}^{N-1} \frac{2}{\sigma_{w,n}^2} \mathrm{Re} \left\{ \frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \boldsymbol{\theta}} \left(\frac{\partial s_{\boldsymbol{\theta}}[n]}{\partial \boldsymbol{\theta}} \right)^* \right\} .
$$

REFERENCES

[Abi95]   A. A. Abidi. "Direct-conversion radio transceivers for digital communications." *IEEE J. Solid-State Circuits*, **30**(12):1399–1410, Dec. 1995.

[Abi07]   A. A. Abidi. "The path to the software-defined radio receiver." *IEEE J. Solid-State Circuits*, **42**(5):954–966, May 2007.

[ACP05]   G. D. Astis, D. Cordeau, J.-M. Paillot, and L. Dascalescu. "A 5-GHz fully integrated full PMOS low-phase-noise LC VCO." *IEEE J. Solid-State Circuits*, **40**(10):2087–2091, Oct. 2005.

[BBC88]   E. Biglieri, S. Barberis, and M. Catena. "Analysis and compensation of nonlinearities in digital transmission systems." *IEEE J. Sel. Areas Commun.*, **6**(1):42–51, Jan. 1988.

[Bes03]   R. E. Best. *Phase-Locked Loops: Design, Simulation, and Applications.* McGraw-Hill, New York, 5th edition, 2003.

[BK07]   J. Bard and V. J. Kovarik. *Software Defined Radio: The Software Communications Architecture.* Wiley, New Jersey, 2007.

[BMC06]   R. Bagheri, A. Mirzaei, S. Chehrazi, M. E. Heidari, M. Lee, M. Mikhemar, W. Tang, and A. A. Abidi. "An 800-MHz–6-GHz software-defined wireless receiver in 90-nm CMOS." *IEEE J. Solid-State Circuits*, **41**(12):2860–2876, Dec. 2006.

[BSE04]   A. R. S. Bahai, B. R. Saltzberg, and M. Ergen. *Multi-Carrier Digital Communications: Theory and Applications of OFDM.* Springer, New York, 2nd edition, 2004.

[BSH00]   M. Buchholz, A. Schuchert, and R. Hasholzner. "Effects of tuner IQ imbalance on multicarrier-modulation systems." In *Proc. IEEE International Caracas Conference on Devices, Circuits and Systems*, pp. 15–17, 2000.

[CBY02]   R. A. Casas, S. L. Biracree, and A. E. Youtz. "Time domain phase noise correction for OFDM signals." *IEEE Trans. Broadcast.*, **48**(3):230–236, Sep. 2002.

[DMA03]   M. Dillinger, K. Madani, and N. Alonistioti. *Software Defined Radio: Architectures, Systems and Functions.* Wiley, New Jersey, 2003.

[DS06]   Y. Ding and A. Sano. "Adaptive nonlinearity compensation for power amplifiers based on model-matching approach." In *Proc. 25th Chinese Control Conference*, pp. 894–899, Aug. 2006.

[EWH01]   M. S. El-Tanany, Y. Wu, and L. Házy. "Analytical modeling and simulation of phase noise interference in OFDM-based digital television terrestrial broadcasting systems." *IEEE Trans. Broadcast.*, **47**(1):20–31, Mar. 2001.

[GH80]   A. George and M. T. Heath. "Solution of sparse linear least squares problems using Givens rotations." *Linear Algebra and Its Applications*, **34**:69–83, 1980.

[GHP81]   A. George, M. T. Heath, and R. J. Plemmons. "Solution of large-scale sparse least squares problems using auxiliary storage." *SIAM Journal on Scientific Computing*, **2**(4):416–429, 1981.

[GNE03]   M. R. Gholami, S. Nader-Esfahani, and A. A. Eftekhar. "A new method of phase noise compensation in OFDM." In *Proc. IEEE Int. Conf. on Communications (ICC)*, pp. 3443–3446, May 2003.

[Hay05]   S. Haykin. "Cognitive radio: Brain-empowered wireless communications." *IEEE J. Sel. Areas Commun.*, **23**(2):201–220, Feb. 2005.

[HLH04]   D. Huang, H. Leung, and X. Huang. "A rational function based predistorter for high power amplifier." In *Proc. International Symposium on Circuits and Systems (ISCAS)*, pp. 1040–1043, May 2004.

[Kay93]   S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, New Jersey, 1993.

[Ken05]   P. Kenington. *RF and Baseband Techniques for Software Defined Radio*. Artech House Publishers, Massachusetts, 2005.

[KK05]   Y.-H. Kim and S.-C. Kim. "Joint channel estimation with phase noise suppression and soft decision decoding scheme for OFDM-based WLANs." In *Proc. IEEE 62nd Vehicular Technology Conference (VTC)*, pp. 161–165, Sep. 2005.

[KL05]   K. Kwok and H. C. Luong. "Ultra-low-voltage high-performance CMOS VCOs using transformer feedback." *IEEE J. Solid-State Circuits*, **40**(3):652–660, Mar. 2005.

[LH00]   T. H. Lee and A. Hajimiri. "Oscillator phase noise: A tutorial." *IEEE J. Solid-State Circuits*, **35**(3):326–336, Mar. 2000.

[Liu98]   C.-L. Liu. "Impacts of I/Q imbalance on QPSK-OFDM-QAM detection." *IEEE Trans. Consum. Electron.*, **44**(3):984–989, Aug. 1998.

[Lju99]   L. Ljung. *System Identification: Theory for the User*. Prentice Hall, New Jersey, 2nd edition, 1999.

[LPL06]   D. D. Lin, R. A. Pacheco, T. J. Lim, and D. Hatzinakos. "Joint estimation of channel response, frequency offset, and phase noise in OFDM." *IEEE Trans. Signal Process.*, **54**(9):3542–3554, Sep. 2006.

[LYK05]   J.-H. Lee, J.-S. Yang, S.-C. Kim, and Y.-W. Park. "Joint channel estimation and phase noise suppression for OFDM systems." In *Proc. IEEE 61st Vehicular Technology Conference (VTC)*, pp. 467–470, May 2005.

[LZ04]   G. Liu and W. Zhu. "Compensation of phase noise in OFDM systems using an ICI reduction scheme." *IEEE Trans. Broadcast.*, **50**(4):399–407, Dec. 2004.

[LZL05]   D. D. Lin, Y. Zhao, and T. J. Lim. "OFDM phase noise cancellation via approximate probabilistic inference." In *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 27–32, Mar. 2005.

[Men97]   U. Mengali. *Synchronization Techniques for Digital Receivers*. Springer, New York, 1997.

[Mit00]   J. Mitola. *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*. Ph.D. Dissertation, Royal Institute of Technology (KTH), Sweden, May 2000.

[Mus95]   C. Muschallik. "Influence of RF oscillators on an OFDM signal." *IEEE Trans. Consumer Electron.*, **41**(3):592–603, Aug. 1995.

[NP90]   S. W. Nam and E. J. Powers. "On the linearization of Volterra nonlinear systems using third-order inverses in the digital frequency-domain." In *Proc. International Symposium on Circuits and Systems (ISCAS)*, pp. 407–410, May 1990.

[PBM95]   T. Pollet, M. V. Bladel, and M. Moeneclaey. "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise." *IEEE Trans. Commun.*, **43**(2/3/4):191–193, Feb./Mar./Apr. 1995.

[PM02]   L. Piazzo and P. Mandarini. "Analysis of phase noise effects in OFDM modems." *IEEE Trans. Commun.*, **50**(10):1696–1705, Oct. 2002.

[Raz98]   B. Razavi. *RF Microelectronics*. Prentice Hall, New Jersey, 1998.

[Ree02]   J. H. Reed. *Software Radio: A Modern Approach to Radio Engineering*. Prentice Hall, New Jersey, 2002.

[RK95]   P. Robertson and S. Kaiser. "Analysis of the effects of phase-noise in Orthogonal Frequency Division Multiplex (OFDM) systems." In *Proc. IEEE Int. Conf. on Communications (ICC)*, pp. 1652–1657, Jun. 1995.

[RLP05]   H.-G. Ryu, Y. Li, and J.-S. Park. "An improved ICI reduction method in OFDM communication system." *IEEE Trans. Broadcast.*, **51**(3):395–400, Sep. 2005.

[Saa96]   Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Co., Boston, 1996.

[Say03]   A. H. Sayed. *Fundamentals of Adaptive Filtering*. Wiley, New Jersey, 2003.

[SHA01]   A. Schuchert, R. Hasholzner, and P. Antoine. "A novel IQ imbalance compensation scheme for the reception of OFDM signals." *IEEE Trans. Consum. Electron.*, **47**(3):313–318, Aug. 2001.

[SL05]   H. Schulze and C. Lueders. *Theory and Applications of OFDM and CDMA: Wideband Wireless Communications*. Wiley, New Jersey, 2005.

[TBS05]   A. Tarighat, R. Bagheri, and A. H. Sayed. "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers." *IEEE Trans. Signal Process.*, **53**(8):3257–3268, Aug. 2005.

[TCP05]   J. Tubbax, B. Côme, 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. Wireless Commun.*, 4(3):872–877, May 2005.

[THS06]   A. Tarighat, R. C. J. Hsu, A. H. Sayed, and B. Jalali. "Digital adaptive phase noise reduction in coherent optical links." *J. Lightwave Technol.*, **24**(3):1269–1276, Mar. 2006.

[TI97]   L. N. Trefethen and D. Bau III. *Numerical Linear Algebra*. SIAM, Philadelphia, 1997.

[Tom98]   L. Tomba. "On the effect of Wiener phase noise in OFDM systems." *IEEE Trans. Commun.*, **46**(5):580–583, May 1998.

[TS05]   A. Tarighat and A. H. Sayed. "MIMO OFDM receivers for systems with IQ imbalances." *IEEE Trans. Signal Process.*, **53**(9):3583–3596, Sep. 2005.

[Tut04]   W. H. W. Tuttlebee. *Software Defined Radio: Baseband Technologies for 3G Handsets and Basestations.* Wiley, New Jersey, 2004.

[TV05]   D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, New York, 2005.

[VGA06]   M. Valkama, A. S. H. Ghadam, L. Anttila, and M. Renfors. "Advanced digital signal processing techniques for compensation of nonlinear distortion in wideband multicarrier radio receivers." *IEEE Trans. Microw. Theory Tech.*, **54**(6):2356–2366, Jun. 2006.

[VRK01]   M. Valkama, M. Renfors, and V. Koivunen. "Advanced methods for I/Q imbalance compensation in communication receivers." *IEEE Trans. Signal Process.*, **49**(10):2335–2344, Oct. 2001.

[WB03]   S. Wu and Y. Bar-Ness. "OFDM channel estimation in the presence of frequency offset and phase noise." In *Proc. IEEE Int. Conf. on Communications (ICC)*, pp. 3366–3370, May 2003.

[ZH01]   Y. Zhao and S.-G. Häggman. "Intercarrier interference self-cancellation scheme for OFDM mobile communication systems." *IEEE Trans. Commun.*, **49**(7):1185–1191, Jul. 2001.

[ZTS06a]   Q. Zou, A. Tarighat, and A. H. Sayed. "Performance analysis and range improvement in multiband-OFDM UWB communications." In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pp. 629–632, Toulouse, France, May 2006.

[ZTK06]   Q. Zou, A. Tarighat, N. Khajehnouri, and A. H. Sayed. "A phase noise compensation scheme for OFDM wireless systems." In *Proc. 14th European Signal Processing Conference (EUSIPCO)*, Florence, Italy, Sep. 2006.

[ZNS06]   Q. Zou, A. Nosratinia, and A. H. Sayed. "Precoding for broadcasting with linear network codes." In *Proc. 44th Annual Allerton Conference on Communication, Control, and Computing*, Allerton, Illinois, Sep. 2006.

[ZTS06b] Q. Zou, A. Tarighat, and A. H. Sayed. "Joint compensation of IQ imbalance and phase noise in OFDM systems." In *Proc. 40th Asilomar Conference on Signals, Systems and Computers*, pp. 1435–1439, Pacific Grove, California, Oct. 2006.

[ZTK07] Q. Zou, A. Tarighat, K. Y. Kim, and A. H. Sayed. "OFDM channel estimation in the presence of frequency offset, IQ imbalance, and phase noise." In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pp. 273–276, Honolulu, Hawaii, Apr. 2007.

[ZTS07a] Q. Zou, A. Tarighat, and A. H. Sayed. "On the joint compensation of IQ imbalances and phase noise in MIMO-OFDM systems." In *Proc. of IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 37–40, New Orleans, Louisiana, May 2007.

[ZTS07b] Q. Zou, A. Tarighat, and A. H. Sayed. "Compensation of phase noise in OFDM wireless systems." *IEEE Trans. Signal Processing*, **55**(11):5407–5424, Nov. 2007.

[ZTS07c] Q. Zou, A. Tarighat, and A. H. Sayed. "Performance analysis of multiband-OFDM UWB communications with application to range improvement." *IEEE Trans. Veh. Technol.*, **56**(6):3864–3878, Nov. 2007.

[ZMS08] Q. Zou, M. Mikhemar, and A. H. Sayed. "Digital compensation of RF nonlinearities in software-defined radios." In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2921–2924, Las Vegas, Nevada, Apr. 2008.

[KZC08] K. Y. Kim, Q. Zou, H. J. Choi, and A. H. Sayed. "An efficient carrier phase synchronization technique for high-order M-QAM-OFDM." *IEEE Trans. Signal Processing*, 2008. Accepted for publication.

[ZTS08] Q. Zou, A. Tarighat, and A. H. Sayed. "Joint compensation of IQ imbalance and phase noise in OFDM wireless systems." *IEEE Trans. Commun.*, 2008. Accepted for publication.