ADVANCED DSP FOR BROADBAND COMMUNICATIONS:

FROM SPARSITY CONSTRAINTS TO DIRTY RF

by

Ahmad Abdulrahman Gomaa Mohammed

APPROVED BY SUPERVISORY COMMITTEE:

Dr. Naofal Al-Dhahir, Chair

Dr. Murat Torlak

Dr. John P. Fonseka

Dr. Carlos Busso

*To my Mother, Father and Sister - for always being there for me*

# ADVANCED DSP FOR BROADBAND COMMUNICATIONS: FROM SPARSITY CONSTRAINTS TO DIRTY RF

by

AHMAD ABDULRAHMAN GOMAA MOHAMMED, B.Sc., M.Sc.

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS

December 2012

## ACKNOWLEDGMENTS

First of all, I would like to express my praise, thankfulness and gratefulness to ALLAH, my Lord, for guiding me throughout my life and my study and giving patience to my family while I am staying abroad.

I am utterly grateful to my mother, father and sister for always being there for me and providing me with ubiquitous and spiritual support and inspiration.

I would like to thank my adviser, Prof. Naofal Al-Dhahir, for providing the opportunity to work under his supervision as a research student. I have learned both a great deal of technical and ethical aspects from him which is certainly going to be beneficial throughout my life.

I would also like to acknowledge my Ph.D. supervisory committee members, Dr. Murat Torlak, Dr. John Fonseka and Dr. Carlos Busso, for their experienced suggestions supplementing this research.

I also appreciate Zahidul Islam, Payam, Bala, M. Mokhtar, A. ElSamadouny and A. Salem from BITS Lab for their continuous help and suggestions. I also thank A. Hesham, M. Aboelsoud, Essam, M. Zinati, M. Salem Deddah, A. Omar and A. Mostafa for their constant help in providing home-like comfort abroad.

August 2012

# PREFACE

This dissertation was produced in accordance with guidelines which permit the inclusion as part of the dissertation the text of an original paper or papers submitted for publication. The dissertation must still conform to all other requirements explained in the Guide for the Preparation of Masters Theses and Doctoral Dissertations at The University of Texas at Dallas. It must include a comprehensive abstract, a full introduction and literature review and a final overall conclusion. Additional material (procedural and design data as well as descriptions of equipment) must be provided in sufficient detail to allow a clear and precise judgment to be made of the importance and originality of the research reported.

It is acceptable for this dissertation to include as chapters authentic copies of papers already published, provided these meet type size, margin and legibility requirements. In such cases, connecting texts which provide logical bridges between different manuscripts are mandatory. Where the student is not the sole author of a manuscript, the student is required to make an explicit statement in the introductory material to that manuscript describing the students contribution to the work and acknowledging the contribution of the other author(s). The signatures of the Supervising Committee which precede all other material in the dissertation attest to the accuracy of this statement.

ADVANCED DSP FOR BROADBAND COMMUNICATIONS:

FROM SPARSITY CONSTRAINTS TO DIRTY RF

Publication No. _____

Ahmad Abdulrahman Gomaa Mohammed, Ph.D.
The University of Texas at Dallas, 2012

Supervising Professor: Dr. Naofal Al-Dhahir

We propose a sparsity-aware approach to estimate and mitigate asynchronous narrow-band interference (NBI) in orthogonal frequency division multiplexing systems with multiple transmit and/or multiple receive antennas. We consider the practical scenarios where one or multiple asynchronous NBI signals experience fast fading and/or frequency-selective fading channels. Furthermore, we propose a novel technique for estimating the desired signal's channel in the presence of unknown NBI. Our approach does not require any prior information about the NBI. Simulation results demonstrate the effectiveness of our proposed techniques in mitigating NBI over practical ranges of NBI power levels, spectral widths, and mobility levels. Furthermore, we propose a new framework for the design of sparse finite impulse response (FIR) equalizers. We formulate greedy and convex-optimization-based solutions for sparse FIR linear equalizer tap vectors given a maximum allowable loss in the decision-point signal-to-noise ratio. Then, we extend our formulation to decision-feedback equalizers and multiple-antenna systems. This is followed by further generalization to the channel

shortening setup where we propose a novel approach to design a sparse target impulse response and a sparse FIR channel shortening equalizer. As an application of current practical interest, we consider self far-end crosstalk cancellation on vectored very high-speed digital subscriber line systems. Finally, We derive a new analytical expression for the error vector magnitude (EVM) in asynchronous single-carrier frequency-division-multiple access systems under carrier frequency offset (CFO) and joint transmit-receive phase noise (PN). The derived EVM expression reveals an interesting cross-layer relationship between the subcarrier mapping scheme at the medium-access control layer and the immunity to CFO and PN at the physical layer. Furthermore, we propose an iterative reduced-complexity joint decoding and PN compensation scheme which does not require any pilots in PN tracking and exploits the low-pass nature of the PN process without assuming a specific PN model. Simulation results show the effectiveness of our proposed digital baseband compensation scheme in PN mitigation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF FREQUENTLY USED ACRONYMS

| Acronym | Description |
|---|---|
| NBI | Narrow-band interference |
| OFDM | Orthogonal frequency division multiplexing |
| SISO | Single-input single-output |
| MIMO | Multiple-input multiple-output |
| CS | Compressive sensing |
| TD(FD) | Time-domain (Frequency-domain) |
| CIR | channel impulse response |
| FFT | Fast Fourier Transform |
| SNR | Signal-to-noise ratio |
| BER | Bit error rate |
| OMP | Orthogonal matching pursuit |
| OLS | Orthogonal least squares |
| LE | Linear equalizer |
| DFE | Decision-Feedback equalizer |
| FIR | Finite impulse response |
| MMSE | Minimum mean square error |
| TIR | Target impulse response |
| CSE | Channel shortening equalizer |
| CFO | Carrier frequency offset |
| PN | Phase noise |
| SC-FDMA | single-carrier frequency-division-multiple access |
| EVM | Error vector magnitude |
| LTE | Long-term evolution |
| AWGN | Additive white Gaussian noise |

# CHAPTER 1

# INTRODUCTION

## 1.1 Introduction

We propose various advanced digital signal processing (DSP) algorithms for estimating sparse signals, designing sparse filters and compensating the radio-frequency (RF) analog front-end (AFE) impairments. All the proposed algorithms and techniques are intended for the base-band. We propose a sparsity-aware approach for estimating narrow-band interference (NBI) signals exploiting their inherent sparsity in frequency-domain. Furthermore, we present a new design framework for sparse filters and equalizers design. Moreover, we study the impact of the RF impairments on the performance of wireless communication systems and, in addition, we propose a low-complexity digital baseband compensation technique.

## 1.2 Narrow-band Interference Estimation

Narrow-band interference (NBI) arises in various wireline and wireless communications systems employing orthogonal frequency division multiplexing (OFDM). For example, non-intentional NBI impairs the performance of multi-band OFDM ultra wide bandwidth (UWB) systems where the other licensed systems operating in the same band cause interference to the UWB system in parts of the operating bandwidth. Furthermore, in the recently-released wireless local area networks (WLANs) IEEE 802.11n standard [2], which employs multiple-input multiple-output (MIMO) and OFDM technologies, NBI is generated by Bluetooth devices operating in the same band [3]. The same problem also arises in other WLAN

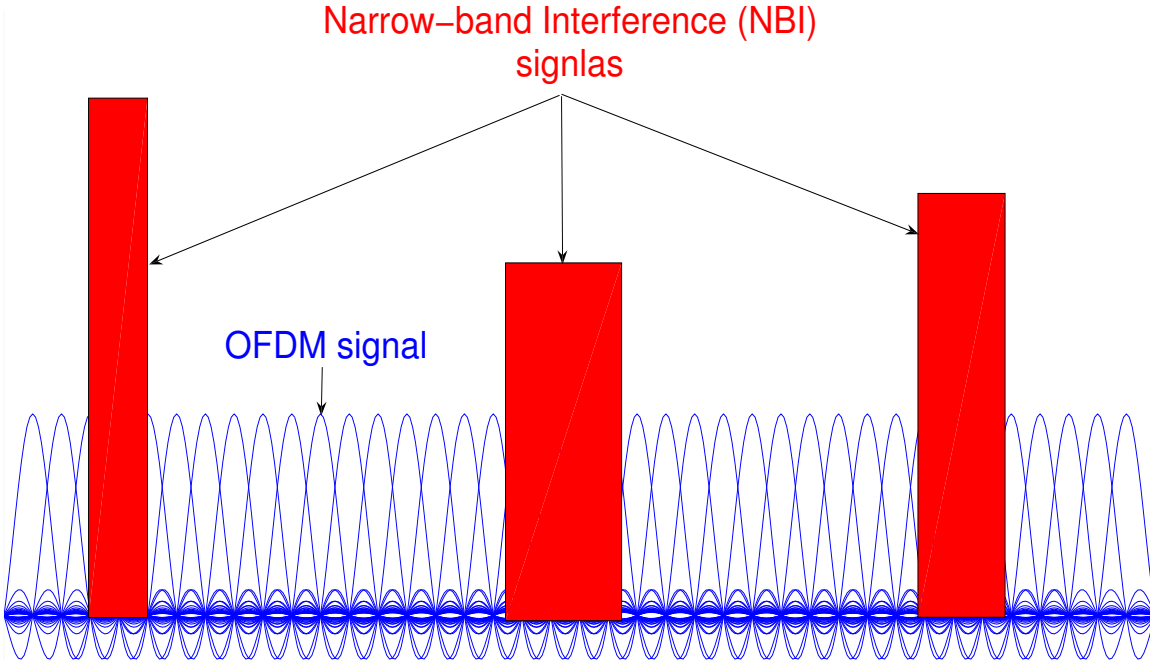


Figure 1.1. NBI signals jamming the OFDM subcarriers in frequency-domain

standards such as IEEE 802.11g. In addition, NBI impairs wired systems such as digital subscriber lines and power line communications in the form of radio frequency interference from AM and amateur radio. Moreover, intentional NBI (also known as jamming) affects wireless military networks. We use the terms NBI and jammer interchangeably in this dissertation. Figure 1.1 depicts the frequency domain where NBI signals jam the OFDM signal.

Among the NBI mitigation techniques, the one in [4] uses a finite impulse response (FIR) prediction error filter (PEF) to whiten the interference narrow-band spectrum. This method assumes that the NBI is an auto-regressive process; otherwise the PEF length must be very long to whiten the NBI spectrum. The PEF is also used in [5] as an erasure insertion mechanism which localizes the erasures to the subcarriers surrounding the NBI

without affecting the remaining subcarriers. In [4] and [5], only a single-tone NBI is assumed. However, in this dissertation, we consider NBI affecting several OFDM subcarriers. In [6], the first subcarrier is assumed to be interference free and the error term between the received and decoded signals of the first subcarrier is used to predict the error term in the next subcarrier. One drawback of this method is that any error in the interference estimate of one subcarrier is propagated to all subsequent subcarriers. Channel estimation in the presence of unknown NBI is addressed in [7] where the NBI is treated as colored noise in addition to the additive white Gaussian noise (AWGN). Then, the expectation maximization (EM) algorithm is used to estimate the channel with the expectation made over the unknown variance of the noise plus interference term. The distribution of the interference is assumed to be known in [7]. In [8], jammed pilot tones are detected and excised by averaging the noise power at each pilot location over a number of consecutive OFDM symbols. Then, the average noise power is compared to a threshold to detect jammed pilots. However, only one jammed pilot is assumed within the OFDM symbol and its location is assumed to be fixed over eight consecutive OFDM symbols; but these may not be valid assumptions in practice. In [9], a few silent (unmodulated) subcarriers are used to measure the interference signal and, then, the linear minimum mean square error (LMMSE) technique is used to estimate the whole NBI signal. The constrained minimum mean-output-energy (CMMOE) is proposed in [10] for NBI suppression where the receiver is designed to minimize the output energy while preserving the desired symbol and canceling the inter-carrier interference. Both LMMSE and CMMOE approaches require prior knowledge about the interference statistics.

Recent results from compressive sensing (CS) theory [11] showed how to reconstruct a sparse vector from insufficient noisy measurements. In this dissertation, we exploit the

inherent sparsity of the NBI signals in the frequency-domain (FD) and use CS theory convex optimization techniques to estimate them.

### 1.3 Sparse Filters Design

Broadband communication channels are characterized by long channel impulse responses (CIRs) that could span tens and even hundreds of symbol periods. Consequently, very long equalizers have to be employed to mitigate the resulting severe inter-symbol interference (ISI). This increases the complexity of computing and implementing (i.e., filtering the received signal) finite impulse response (FIR) equalizers which grows proportional to the number of taps. To reduce complexity at the expense of a tolerable performance loss, long equalizers with few nonzero taps (i.e., sparse equalizers) have been proposed in the literature. In [12], the number of equalizer taps is fixed but the tap spacings are allowed to be nonuniform and the branch-and-bound algorithm is used to search for the locations of the nonzero taps. In [13], the number of nonzero taps is reduced by choosing only the strongest taps of the minimum mean square error (MMSE) solution. However, a main disadvantage of both techniques is that the whole equalizer tap vector must be calculated which is computationally complex. The decision feedback equalizer (DFE) is considered in [14] where an ad-hoc algorithm is proposed to determine the locations of the feedforward filter (FFF) taps. This algorithm computes a measure for the signal-to-noise ratio (SNR) of each channel tap and assigns the FFF taps to the channel taps having the highest SNRs. Given a fixed total number of DFE taps, the numbers of FFF and feedback filter (FBF) taps are optimized in [15]. Sparse equalization becomes even more critical for multiple-input multiple-output (MIMO) systems [16] where the number of equalizer taps grows proportional to the product of the

number of input and output data streams. In practice, many dispersive channels have sparse power-delay profiles (PDPs), i.e., have few dominant paths over the whole CIR duration. The work in [17] shows that the infinite-length symbol-spaced zero-forcing (ZF) linear equalizers (LEs) for sparse channels are inherently nonuniformly spaced. Greedy and sequential search algorithms are proposed in [18] and [19], respectively, for finding the taps locations for single-input single-output (SISO) LEs and DFEs. In this dissertation, we present a new framework for designing sparse SISO and MIMO FIR equalizers. Specifically, we formulate both greedy algorithms and $l_1$-norm minimization programs to determine the locations and weights of the nonzero equalizer taps subject to a given maximum tolerable performance loss. Also, our formulation is not restricted to a specific greedy algorithm as in [18] and [19]. In the context of FIR filter design, $l_1$-norm minimization was recently proposed in [20] for the problem of sparse digital filters design subject to a number of constraints on the frequency domain response. However, our sparse equalization problem differs from that in [20] in the formulation, constraints, and applications.

Since equalization complexity increases with the CIR length, a front-end equalizer, commonly known as a channel shortening equalizer (CSE), is designed such that the cascade of the long CIR and the CSE is equivalent to a short target impulse response (TIR). This channel shortening technique enables the implementation of maximum-likelihood (ML) or maximum-a-posteriori (MAP) detectors at practical complexity levels. Several TIR design criteria have been investigated in the literature (see [21] and the references therein). In [21], the unit-tap constraint (UTC) and the unit-energy constraint (UEC) criteria were unified under a single framework that lends itself to other constraints as well. In [22], the CSE designs were generalized to MIMO channels. In [23], the TIR taps were chosen to be noncontiguous

and were designed to coincide with the significant multipath arrivals of the original channel. To the best of our knowledge, in all of the previous work, with the exception of [23], the TIR taps were assumed to be contiguous and the CSE is designed using the MMSE criterion. However, more degrees of freedom and, hence, performance improvement can be obtained by designing the TIR taps to be noncontiguous. In this dissertation, we allow both of the TIR taps and the CSE taps to be noncontiguous and sparse and formulate both greedy algorithms and $l_1$-norm minimization programs to determine the locations and values of the sparse taps. In addition, we prove that the resulting channel shortening MMSE is reduced or at least remains the same as in the conventional contiguous case while reducing the implementation complexity significantly.

Self far-end crosstalk is a major performance-limiting impairment in vectored very high-speed digital subscriber line (VDSL) systems [24]. The concept of vectored DSL transmission was introduced in [25] where joint signal processing/detection techniques were employed to cancel self FEXT between the vectored DSL lines at each tone affected by self FEXT. The linear ZF crosstalk canceler was shown in [26] to be near-optimal for VDSL systems due to the diagonally-dominant structure of the MIMO channel matrix; hence, there is no need to design the canceler based on the more-complex MMSE criterion. Furthermore, it is well-known that most of the self FEXT power emanates from few neighboring lines in the binder. This property was utilized in [27] to reduce the number of nonzero self FEXT cancellation taps whose locations were assigned to the neighboring lines with the largest average crosstalk power. In this dissertation, we formulate both greedy algorithms and $l_1$-norm minimization programs to design a sparse self FEXT canceler subject to a given maximum performance loss from the highly-complex full self FEXT canceler.

## 1.4 Dirty RF and Analog Front-End Impairments

Two main architectures have been proposed for the radio-frequency analog front-end of communication transceivers [28]. The first one is the super-heterodyne architecture where the signal is converted from (to) the RF band to (from) the baseband in two stages where the signal is first translated to an intermediate frequency (IF) and, then, translated to the RF band or the baseband. The second one is the direct-conversion architecture (DCA) where the signal is translated directly from (to) the RF band to (from) the baseband in one stage, i.e., without using the IF stage, hence it is also called the zero-IF architecture. The DCA has received a lot of attention over the past few years thanks to the removal of the bulky IF-stage analog components, especially SAW filters.

However, the DCA encounters several problems resulting from the imperfections and inaccuracies accompanying the manufacturing process. Some of these problems are [29] DC offset, carrier frequency offset (CFO), phase noise (PN), power amplifier non-linearity, and Inphase-Quadrature imbalance (IQI). Several techniques have been proposed to estimate and compensate for these RF impairments, see [30, 31] and the references therein. In [32], the generalized input-output model for mobile systems under joint transmit-receive IQI was developed and a reduced-complexity digital baseband compensation technique was developed. Joint compensation of multiple RF impairments was also proposed in the literature [33, 34]. Furthermore, experimental performance evaluation under RF impairments was performed in [35]. The impact of PN and IQI on multiple-input-multiple-output orthogonal-frequency-division-multiplexing systems was recently investigated in [36]. Furthermore, the impact of CFO and PN on channel estimation was recently studied in [37].

In this dissertation, we focus on the PN problem as one of the major RF AFE impairments [28,29] resulting from the imperfections and inaccuracies of the crystal oscillator's fabrication process. Several techniques [38–44] have been proposed to estimate and compensate for PN in OFDM systems. For example, in [40], joint channel estimation and PN compensation is proposed using training symbols and pilot subcarriers multiplexed with data subcarriers. The impact of CFO and PN on channel estimation is investigated in [45]. In [41–44], iterative detection and PN compensation techniques are proposed with pilot subcarriers used to estimate the PN-induced common phase-error (CPE) as an initialization for the PN compensation process. In [42], each OFDM symbol is divided into sub-blocks where the PN process is assumed quasi-static over each sub-block. A Kalman filter for PN tracking is proposed in [43] where knowledge of the PN statistical model is assumed. In [44], the PN process is approximated by a periodic signal with few Fourier series coefficients.

A common assumption in all of the above-cited approaches is the availability of pilot subcarriers multiplexed in each data OFDM symbol. Unfortunately, this assumption is not valid for single-carrier frequency-division-multiple access (SC-FDMA) systems which have received a lot of attention since the adoption of SC-FDMA in the uplink of the 3GPP LTE standard [46]. SC-FDMA exhibits significantly lower cubic metric and peak-to-average power ratio (PAPR) than OFDM [47]. This feature increases the cost-efficiency of the power amplifier used in the RF front-end of the user terminal. To maintain the desirable low-PAPR property of the SC-FDMA signal, pilot subcarriers are not multiplexed with data subcarriers in the same SC-FDMA symbol due to their higher power. Instead, at the beginning of each uplink frame, the pilots are transmitted in a dedicated training symbol followed by six data SC-FDMA symbols void of pilots, as in the LTE-uplink [46]. These training symbols are

shown to be sufficient for channel tracking even under high-mobility scenarios. However, they may not be sufficient for tracking other time-varying impairments such as the PN impairing the transmit and receive crystal oscillators. The PN samples vary from one time instant to another within a single SC-FDMA symbol. Hence, PN tracking is a challenging task for the LTE uplink where consecutive training symbols are six symbols apart with no pilots inserted in the data symbols in between. In OFDM systems, the PN tracking task is easier thanks to the pilot subcarriers multiplexed with data subcarriers.

The PN maximum-a-priori (MAP) estimate is obtained in [1] for single-carrier single-user systems employing turbo equalization without pilots. The MAP solution was shown in [1] to be a second-order phase-locked loop (PLL) for constant-amplitude modulation (such as BPSK and QPSK). Furthermore, the PN statistical model was assumed available in [1] and the extension to higher-order modulation is not straightforward. In this dissertation, we tackle the problem of non-pilot-aided PN compensation in SC-FDMA systems without knowledge of the PN model. Our main contributions are summarized next. We propose an iterative fast-converging joint data decoding and PN compensation scheme which exploits the low-pass nature of the PN process without assuming a priori knowledge about the exact PN model to avoid model mismatch problems. In addition, no restrictions on the constellation size or shape are imposed and several accurate approximations are made to reduce the compensation complexity significantly. In addition, we derive a new analytical expression for the EVM measure under joint transmit-receive PN and carrier frequency offset (CFO). EVM is widely adopted in most of the modern wireless communication standards as a performance metric. Although CFO and PN are compensated in most practical systems, residual CFO and PN still exist. Hence, the derived EVM expression enables accurate performance evaluation

in the presence of these residual impairments. Furthermore, based on the derived EVM expression, we compare the robustness of both the localized and the distributed subcarrier mapping schemes under CFO and PN. The impact of PN on the EVM was investigated in [48] for single-carrier single-user systems. However, the derived expression in [48] assumed no channel equalization and had functional dependence on the instantaneous realization of the channel which varies randomly over time. However, in this dissertation, we include channel equalization effects and take the statistical expectation over the channel realization to replace the channel realization by its autocorrelation function.

## 1.5 Orthogonal Frequency Division Multiplexing

OFDM has been adopted in several communications standards, e.g. the wireless local area networks (WLANs) IEEE 802.11n standard [2] and the long-term evolution (LTE) standard [49]. OFDM converts the frequency selective fading channels into several flat fading subchannels using the fast Fourier transform (FFT) processing. Then, the equalization of each subchannel is implemented using a single-tap equalizer. The weight of each tap is easily computed and updated based on the channel frequency response. Figure 1.2 shows a 64-KHz frequency selective channel and its corresponding sixty-four flat fading subchannels each with 1 kHz-bandwidth. In the sequel, we describe two main flavors of the OFDM, namely, cyclic-prefix OFDM and zero-padding OFDM.

### 1.5.1 Cyclic-Prefix OFDM

First, information bits are mapped into generally complex symbols using a specified constellation, e.g. quadrature amplitude modulation (QAM), where the number of bits represented

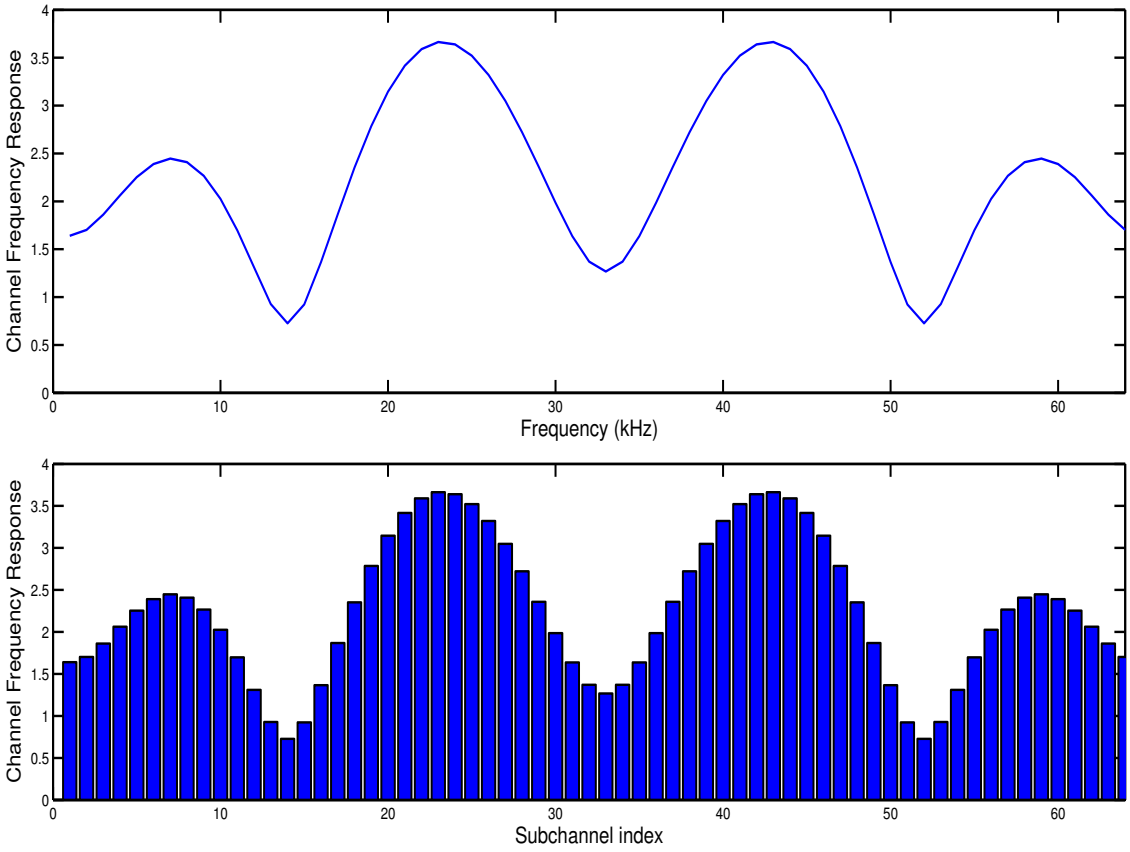


Figure 1.2. Frequency-selective fading channel (up) versus its discretized flat fading subchannels (bottom)

by the symbol depends on the constellation size. Then, $N$ symbols are collected into a block whose $N$-point inverse fast Fourier transform (IFFT) is computed. This conceptually implies that the constellation symbols are in frequency-domain. In cyclic-prefix OFDM (CP-OFDM), the last $\nu$ samples of the IFFT output block are copied and appended to the block beginning increasing the block length to $N + \nu$. Then, the resulting block is passed to the digital-to-analog (D/A) conversion unit for transmission. Since the first $\nu$ samples of the transmitted block are exactly the same as the last $\nu$ ones, we can say that this artificially introduces some periodicity to the transmitted block. The cyclic prefix performs two tasks. First, it separates consecutive OFDM blocks and, hence, prevents inter-block interference

assuming that the CP length is longer than that of the channel impulse response (CIR). Second, it converts the linear convolution between the transmitted signal and the channel into circular convolution which can be converted into a point-wise multiplication using FFT at the receiver. This assumes that the channel impulse response does not vary over the OFDM block interval. For mathematical illustration of the CP-OFDM process, we denote the $N$ constellation symbols by $\mathbf{X} \triangleq [X_0\ X_1\ \ldots\ X_{N-1}]^T$ and the size-$N$ FFT matrix by $\mathbf{F}_N$ where $()^T$ denotes the transpose operation. Hence, the $(N+\nu)\times 1$ transmitted signal is given by

$$\mathbf{x} = \begin{bmatrix} \mathbf{0}_{\nu\times(N-\nu)} & \mathbf{I}_\nu \\ & \mathbf{I}_N \end{bmatrix} \mathbf{F}_N^H \mathbf{X} \tag{1.1}$$

where $\mathbf{I}_N$ and $\mathbf{I}_\nu$ denote the identity matrices of sizes $N$ and $\nu$, respectively. Furthermore, $\mathbf{0}_{\nu\times(N-\nu)}$ denotes an all-zero matrix of size $\nu\times(N-\nu)$. Assuming non-time varying channel, the noise-free received $(N+\nu)\times 1$ signal is given by

$$\mathbf{y} = \mathbf{H}_{\text{Toep}}\mathbf{x} \tag{1.2}$$

where $\mathbf{H}_{\text{Toep}}$ is the $(N+\nu)\times(N+\nu)$ Toeplitz channel matrix with the first column given by the length-$L$ CIR padded by $(N+\nu-L)$ zeros. In the receiver, the $\nu$ samples corresponding to the cyclic prefix are removed and, then, the FFT is applied as follows

$$\begin{aligned} \mathbf{Y} &= \mathbf{F}_N \begin{bmatrix} \mathbf{0}_{N\times\nu} & \mathbf{I}_N \end{bmatrix} \mathbf{y} = \mathbf{F}_N \underbrace{\begin{bmatrix} \mathbf{0}_{N\times\nu} & \mathbf{I}_N \end{bmatrix} \mathbf{H}_{\text{Toep}} \begin{bmatrix} \mathbf{0}_{\nu\times(N-\nu)} & \mathbf{I}_\nu \\ & \mathbf{I}_N \end{bmatrix}}_{\triangleq \mathbf{H}} \mathbf{F}_N^H \mathbf{X} \\ &= \underbrace{\mathbf{F}_N \mathbf{H} \mathbf{F}_N^H}_{\triangleq \mathbf{\Lambda}} \mathbf{X} = \mathbf{\Lambda}\mathbf{X} \end{aligned} \tag{1.3}$$

where $\mathbf{H}$ is an $N\times N$ circulant channel matrix thanks to the cyclic prefix. The first column of $\mathbf{H}$ is given by the length-$L$ CIR padded by $(N-L)$ zeros. The $N\times N$ circulant matrix is the matrix whose $n^{\text{th}}$ column is obtained by circularly shifting the $((n-1) \bmod N)^{\text{th}}$ column

down by one entry. One property of circulant matrices is that they are diagonalized by FFT matrices. Hence, the matrix $\mathbf{\Lambda}$ in (1.3) is diagonal and its diagonal entries are the $N$-point FFT of the zero-padded CIR. In other words, the diagonal entries of the matrix $\mathbf{\Lambda}$ are the channel frequency response (CFR). Then, the constellation symbols are simply recovered as follows

$$\widehat{X}_k = Y_k/\lambda_k \tag{1.4}$$

where $Y_k$ is the $k^{\text{th}}$ entry of the vector $\mathbf{Y}$ and $\lambda_k$ is the $k^{\text{th}}$ diagonal entry of the matrix $\mathbf{\Lambda}$. One disadvantage of CP is the wasted energy since the CP samples are removed in the receiver.

### 1.5.2 Zero-Padding OFDM

Instead of using the last $\nu$ samples as the guard sequence in CP-OFDM, we use zeros as the guard sequence between OFDM blocks in ZP-OFDM. Zero-padding overcomes the problem of wasted energy inherent in CP-OFDM since no energy is used in the zero padding. The $(N+\nu)\times 1$ transmitted signal is given by

$$\mathbf{x} = \begin{bmatrix} \mathbf{I}_N \\ \mathbf{0}_{\nu\times N} \end{bmatrix} \mathbf{F}_N^H \mathbf{X} = \underbrace{\begin{bmatrix} \mathbf{F}_N \\ \mathbf{0}_{\nu\times N} \end{bmatrix}}_{\triangleq \mathbf{F}_{\text{zp}}} \mathbf{X} \tag{1.5}$$

Assuming non-time varying channel, the noise-free received $(N+\nu)\times 1$ signal is given by

$$\mathbf{y} = \mathbf{H}_{\text{Toep}}\mathbf{x} = \mathbf{H}_{\text{Toep}}\mathbf{F}_{\text{zp}}\mathbf{X} \tag{1.6}$$

where $\mathbf{H}_{\text{Toep}}$ is the $(N+\nu)\times(N+\nu)$ Toeplitz matrix defined as in Section 1.5.1. However, thanks to the all-zero submatrix in $\mathbf{F}_{\text{zp}}$, we can replace $\mathbf{H}_{\text{Toep}}$ with an $(N+\nu)\times(N+\nu)$ circulant matrix, $\mathbf{H}$, without changing the received OFDM signal [50]. The first column of this

circulant matrix is the zero-padded CIR. Observing that circulant matrices are diagonalized by FFT matrices, we apply the $P$-point FFT to $\mathbf{y}$ where $P = N + \nu$ and write

$$\mathbf{Y} = \mathbf{F}_P \mathbf{y} = \mathbf{F}_P \mathbf{H} \underbrace{\mathbf{F}_P^H \mathbf{F}_P}_{=\mathbf{I}_P} \mathbf{F}_{\text{zp}} \mathbf{X} = \mathbf{\Lambda} \mathbf{F}_P \mathbf{F}_{\text{zp}} \mathbf{X} \tag{1.7}$$

where $\mathbf{\Lambda} = \mathbf{F}_P \mathbf{H} \mathbf{F}_P^H$ is a diagonal matrix whose diagonal entries represent the $P$-point FFT of the channel impulse response. Noting that $\mathbf{V}^{-1} = \mathbf{V}^H$, the symbols vector $\mathbf{X}$ can be recovered as follows

$$\widehat{\mathbf{X}} = \begin{bmatrix} \mathbf{F}_N & \mathbf{0}_{N \times \nu} \end{bmatrix} \mathbf{F}_P^H \mathbf{\Lambda}^{-1} \mathbf{Y} \tag{1.8}$$

where $\mathbf{\Lambda}^{-1}$ is a diagonal matrix denoting the inverse of $\mathbf{\Lambda}$. Note that $\mathbf{\Lambda}^{-1}$ can be easily computed since $\mathbf{\Lambda}$ is diagonal.

### 1.5.3 DFT-precoded OFDM

Another flavor of OFDM transmission is the discrete Fourier transform (DFT)-precoded OFDM where the data symbol are first precoded using a DFT matrix before applying the inverse DFT. One can argue that the DFT precoding cancels the effect of the inverse DFT and then we just have a single-carrier transmission. In fact, this argument is true and the multiple access version of this transmission scheme is called single-carrier frequency-division multiple access (SC-FDMA) where the data symbols of each user is first DFT-precoded and then assigned to a specific set of subcarriers while the other subcarriers are set to zeros since they will be used by other users. Next, a larger-size inverse DFT is applied where its size is bigger than the number of data symbols. SC-FDMA is adopted in the uplink transmission in the Long-Term Evolution (LTE) standard due to its low peak-to-average power ratio (PAPR) which is significantly lower than that of OFDM.

## 1.6 Multiple-Input Multiple-Output Systems

MIMO systems have received a lot of attention in the past two decades and have been adopted in recent communication standards, e.g. IEEE802.11n and LTE-Advanced. The transmitter can use its multiple antennas in several ways according to the application and goal. Among these ways are transmit diversity, spatial multiplexing, and antenna beamforming. Transmit diversity is used to increase the system reliability by transmitting correlated symbols over multiple antennas. Space-time block codes (STBC) have been proposed in [51] and have been extended in the literature to space-frequency block codes. Several STBC designs have been proposed to trade-off rate, diversity order, and detection complexity. When transmit diversity is employed, only one receive antenna is enough for linear detection. Spatial multiplexing aims at increasing the transmission rate by transmitting different uncorrelated streams over different antennas. Herein, the number of receive antennas needs to be at least equal to the number of transmit antennas for linear detection. In transmit beamforming, the antenna beam pattern is physically shaped to maximized the transmitted signal power in one direction and diminish the transmitted power in other directions. This can be useful if the desired receiver direction is known at the transmitter and, moreover, does not vary quickly. Transmit beamforming is achieved by spatially precoding the signal before transmission.

## 1.7 Dissertation Contributions

Recent results from compressive sensing (CS) theory [11] showed how to reconstruct a sparse vector from insufficient noisy measurements. We exploit the inherent sparsity of the NBI signals in the frequency-domain (FD) and use CS theory convex optimization techniques

to estimate them. The main contributions of our NBI estimation work are summarized as follows

1. NBI estimation in SISO and MIMO systems for both ZP-OFDM and CP-OFDM transmission techniques without requiring any prior information about the NBI statistics, location, or power.
2. Channel estimation in the presence of NBI.
3. Consideration of a number of practical issues such as scenarios where the NBI signal experiences a time-varying frequency-selective (FS) fading channel and where the desired signal is jammed by multiple NBI signals.
4. Reduced-complexity implementation of our proposed approach.

In regards to our work on sparse finite impulse response filters design, we summarize our main contributions as follows

1. Proposing a novel approach for sparse FIR filter design.
2. Applying our approach to three broadband communication scenarios:
   (a) SISO and MIMO equalization including both linear and decision-feedback architectures.
   (b) Channel shortening including both target impulse response design and channel shortening equalizer.
   (c) Self FEXT crosstalk cancelation for vectored VDSL systems.

3. Significant complexity reduction in all the three scenarios under study with minimal performance loss.

Finally, regarding our work on the dirty RF and analog front-end impairments, we summarize our contributions as follows

1. Derivation of a new EVM expression for SC-FDMA uplink transmissions under CFO and joint transmit-receive PN taking into account channel equalization effects.

2. Based on the derived EVM expression, we compared the localized and distributed subcarrier mapping schemes from the viewpoint of their immunity to CFO and PN.

3. Proposing an iterative fast-converging reduced-complexity algorithm for joint decoding and PN compensation in SC-FDMA systems.

# CHAPTER 2

# SPARSITY-AWARE APPROACH FOR NBI ESTIMATION IN MIMO OFDM SYSTEMS

## 2.1 Introduction

In this chapter, we present a novel approach based on compressive sensing theory to estimate and mitigate asynchronous narrow-band interference (NBI) in orthogonal frequency division multiplexing (OFDM) systems with multiple transmit and/or multiple receive antennas. We consider the practical scenarios where one or multiple asynchronous NBI signals experience fast fading and/or frequency-selective fading channels. Furthermore, we propose a novel technique for estimating the desired signal's channel in the presence of unknown NBI. Our approach does not require any prior information about the NBI. Simulation results demonstrate the effectiveness of our proposed techniques in mitigating NBI and approaching the interference-free performance limit over practical ranges of NBI power levels, spectral widths, and mobility levels.

The rest of this chapter is organized as follows. In Section 2.2, we give a brief background about compressive sensing (CS) theory. In the same section, we describe the system model and the narrow-band interference model. We apply our proposed NBI estimation approach to single-input single-output (SISO) zero-padded (ZP) OFDM systems in Section 2.3. The extension to MIMO systems is presented in Section 2.4. In Section 2.5, we address the problem of channel estimation in the presence of NBI. In Section 2.6, we consider a

number of practical issues such as scenarios where the NBI signal experiences a time-varying frequency-selective (FS) fading channel and where the desired signal is jammed by multiple NBI signals. In the same section, we propose a reduced-complexity implementation of our proposed approach and extend it to cyclic-prefix (CP)-based OFDM systems which are widely used in practice. Finally, simulation results are presented in Sections 2.7. The contents of this chapter are adapted and reprinted from [52][1].

*Notation*: Unless otherwise stated, lower and upper case bold letters denote vectors and matrices, respectively, and their subscripts denote their sizes. The matrices $\mathbf{F}$ and $\mathbf{I}$ denote the Fast Fourier Transform (FFT) matrix and the identity matrix, respectively, and their subscripts, when written, denote theirs sizes. The matrix $\mathbf{0}_{m\times n}$ denotes the all-zero matrix of size $m \times n$ and the notation $\text{diag}\,(x_1, \ldots, x_N)$ denotes an $N \times N$ diagonal matrix whose diagonal elements are $\{x_1, \ldots, x_N\}$. Also, $(\ )^H$, $(\ )^*$, $(\ )^T$ and $(\ )^{-1}$ denote the matrix complex-conjugate transpose, complex conjugate, transpose, and inverse operations, respectively. The $j^{th}$ element of $\mathbf{a}$ is denoted by $a(j)$. $\mathbf{P}_P^n$ denotes the permutation matrix defined in [53] where the $k^{th}$ element in the vector $\mathbf{P}_P^n\mathbf{a}$ is $a((P-k+n) \bmod P)$.

## 2.2 Compressive Sensing Theory Background and System Model

### 2.2.1 Compressive Sensing Theory Background

CS theory [11, 54] asserts that we can recover a sparse vector $\mathbf{x} \in \mathbb{C}^N$ from a measurement vector $\mathbf{y} \in \mathbb{C}^M$ efficiently and uniquely where $M \ll N$. In other words, the exact solution of the under-determined system of equations $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be computed under some conditions

[1]Copyright [2011] IEEE. Reprinted, with permission, from A. Gomaa and N. Al-Dhahir, A Sparsity-Aware Approach for NBI Estimation in MIMO-OFDM, IEEE Transactions on Wireless Communications, June 2011

on $\mathbf{A}$ where $\mathbf{A}$ denotes the $M \times N$ measurement matrix. The word "sparse" means that $\mathbf{x}$ contains few nonzero elements. The sparse vector $\mathbf{x}$ is computed by solving the following $l_1$-norm constrained minimization problem

$$\min_{\tilde{\mathbf{x}} \in \mathbb{C}^N} \| \tilde{\mathbf{x}} \|_1 \quad \text{subject to} \quad \mathbf{A}\tilde{\mathbf{x}} = \mathbf{y} \tag{2.1}$$

where, $\| \mathbf{x} \|_1 \triangleq \sum_k \mid x_k \mid$ denotes the $l_1$-norm of $\mathbf{x}$. The problem formulation in (2.1) was extended to the case of noisy measurements as follows

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{z} \tag{2.2}$$

where $\mathbf{z} \in \mathbb{C}^M$ is a zero-mean random noise vector with covariance matrix $\mathbf{C}_\mathbf{z} \triangleq E(\mathbf{z}\mathbf{z}^H)$. In the noisy case, $\mathbf{x}$ can be recovered by solving the following convex optimization program [55]

$$\min_{\tilde{\mathbf{x}} \in \mathbb{C}^N} \| \tilde{\mathbf{x}} \|_1 \quad \text{subject to} \quad \| \mathbf{y} - \mathbf{A}\tilde{\mathbf{x}} \|_2^2 \leq \epsilon \tag{2.3}$$

where $\| \ . \ \|_2$ denotes the $l_2$-norm and the constant $\epsilon$ is chosen such that it bounds the amount of noise in the measurements. In fact, the convex optimization problem in (2.3) is a second-order cone program and can be solved efficiently [56].

### 2.2.2 Model Description

Assuming perfect frequency synchronization between the transmitted desired signal and the receiver, we write the received signal vector in the time-domain (TD), including the all-zeroes guard sequence, for zero-padded SISO-OFDM systems as follows

$$\mathbf{y} = \mathbf{H} \underbrace{\begin{bmatrix} \mathbf{F}_N^H \\ \mathbf{0}_{\nu \times N} \end{bmatrix}}_{:=\mathbf{F}_{\text{zp}}} \mathbf{X} + \mathbf{\Lambda}_{\text{fo}} \mathbf{H}_{\text{J}} \mathbf{j} + \mathbf{z} \tag{2.4}$$

where $\mathbf{X}$, $\mathbf{j}$ and $\mathbf{z}$ denote, respectively, the transmitted data vector, the TD transmitted NBI vector, and the TD zero-mean complex AWGN vector whose elements are independent and identically distributed with variance $N_o$. Furthermore, $\mathbf{H}$ and $\mathbf{H}_{\text{J}}$ denote the $P \times P$ TD FS fading channel matrices of the desired signal and the interference signal, respectively, where $P = N + \nu$ with $N$ denoting the length of the data vector and $\nu$ denoting the guard sequence length. Instead of cyclic prefixing, we use zero-padding [50] (implied by the precoding matrix $\mathbf{F}_{\text{zp}}$) to eliminate inter-block interference (IBI) in the guard sequence. ZP-OFDM is adopted in multi-band OFDM ultra-wide-band (UWB) [57] systems. Assuming that the channel of the desired signal is quasi-static over one OFDM symbol, then $\mathbf{H}$ becomes a Toeplitz lower-triangular matrix. However, thanks to the all-zero submatrix in $\mathbf{F}_{\text{zp}}$, we can replace it with a circulant matrix without changing the received OFDM signal [50]. The first column of this circulant matrix is the zero-padded channel impulse response (CIR) coefficients vector denoted by $\begin{bmatrix} h_0 & h_1 & .... & h_{L-1} & \mathbf{0}_{1\times(P-L)} \end{bmatrix}^T$ where $h_k$ is the $k^{\text{th}}$ complex CIR coefficient and $L$ is the number of CIR taps such that $L \leq \nu + 1$. Also, we assume the NBI channel to be quasi-static and, hence, $\mathbf{H}_{\text{J}}$ is a $P \times P$ Toeplitz matrix whose first column is the zero-padded NBI CIR vector. Furthermore, $\mathbf{\Lambda}_{\text{fo}} = \text{diag}\left(1, \exp\left(i\frac{2\pi\alpha}{P}\right), \ldots, \exp\left(i\frac{2\pi\alpha(P-1)}{P}\right)\right)$ where $i \triangleq \sqrt{-1}$, $\exp(.)$ is the exponential function, and $\alpha$ is a uniformly-distributed random variable [5] over the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. The variable $\alpha$ is used, as in [58] and [5], to model the frequency offset (FO) between the transmitted signal and the NBI signal since the NBI tones are not guaranteed to lie on the receiver FFT grid. This NBI model is more accurate than the one used in [59] and [60] where the NBI is assumed to be synchronous with the transmitted signal causing the NBI tones to lie exactly on the receiver FFT grid. The transmitted NBI signal $\mathbf{j}$ can be expressed as $\mathbf{j} = \mathbf{F}_P^H \mathbf{J}$ where $\mathbf{J}$ is a length-$P$ vector whose

$k^{\text{th}}$ element, $J(k)$, represents the NBI signal at the $k^{\text{th}}$ subcarrier on the $P$-point FFT grid. By definition, only few elements of $\mathbf{J}$ are nonzero and their indices (known as NBI support) are given by $I_J = \{k : J(k) \neq 0\}$ whose cardinality is denoted by $r$. In other words, $\mathbf{J}$ is a sparse vector with $r$ nonzero elements. We assume the NBI signal to be quasi-static over one OFDM symbol and to vary independently from OFDM symbol to another. However, in many practical systems, the NBI signal is fixed over several successive OFDM symbols thanks to the high data rates supported by Today's communications systems. These high rates make the OFDM symbol duration very small over which the NBI signal does not vary significantly. This fact can be exploited to provide a more accurate estimate of the NBI frequency support.

### 2.2.3 Receiver Windowing

Our CS-based approach for NBI estimation (as will be explained later) is based on the NBI sparsity in FD. In practical systems, however, this sparsity is destroyed by both the non-circulant structure of $\mathbf{H}_{\text{J}}$ and the FO [5] between the transmitted signal and the NBI signal. These two factors cause the NBI energy to spread over all subcarriers after taking the FFT of the received signal. One well-known technique to spectrally contain the signal in FD is windowing. When applied to the received signal before taking the FFT, the window reduces the FFT side lobes of the NBI signal and, hence, enhances its sparsity in FD. Applying the window and taking the FFT, we get

$$\mathbf{Y} = \mathbf{F}_P \mathbf{\Lambda}_{\text{W}} \mathbf{y} = \mathbf{F}_P \mathbf{\Lambda}_{\text{W}} \mathbf{H} \mathbf{F}_{\text{zp}} \mathbf{X} + \mathbf{F}_P \mathbf{\Lambda}_{\text{W}} \mathbf{\Lambda}_{\text{fo}} \mathbf{H}_{\text{J}} \mathbf{F}_P^H \mathbf{J} + \mathbf{F}_P \mathbf{\Lambda}_{\text{W}} \mathbf{z} \tag{2.5}$$

where $\mathbf{\Lambda}_{\mathrm{W}} = \mathrm{diag}\left(w(0), w(1), \ldots, w(P-1)\right)$ and $w(n)$ is the $n^{\mathrm{th}}$ sample of the window function. Since $\mathbf{H}$ is circulant, it can be expressed as $\mathbf{H} = \mathbf{F}_P^H \mathbf{\Lambda} \mathbf{F}_P$ where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal is the $P$-point discrete Fourier transform (DFT) of the first column of $\mathbf{H}$. For large values of $P$, $\mathbf{H}_{\mathrm{J}}$ can be approximated as a circulant[2] matrix [61] and, hence, expressed as $\mathbf{H}_{\mathrm{J}} = \mathbf{F}_P^H \mathbf{\Lambda}_{\mathrm{J}} \mathbf{F}_P$ where $\mathbf{\Lambda}_{\mathrm{J}}$ is a diagonal matrix whose diagonal is the $P$-point DFT of the first column of $\mathbf{H}_{\mathrm{J}}$. Therefore, we re-write (2.5) as follows

$$\mathbf{Y} = \underbrace{\mathbf{H}_{\mathrm{W}} \mathbf{\Lambda} \mathbf{V}}_{\triangleq \tilde{\mathbf{\Lambda}}} \mathbf{X} + \underbrace{\mathbf{H}_{\mathrm{eqv}} \mathbf{\Lambda}_{\mathrm{J}} \mathbf{J}}_{\triangleq \mathbf{J}_{\mathrm{eqv}}} + \underbrace{\mathbf{F}_P \mathbf{\Lambda}_{\mathrm{W}} \mathbf{z}}_{\mathbf{Z}} = \tilde{\mathbf{\Lambda}}_{P \times N} \mathbf{X} + \mathbf{J}_{\mathrm{eqv}} + \mathbf{Z} \tag{2.6}$$

where $\mathbf{H}_{\mathrm{W}} \triangleq \mathbf{F}_P \mathbf{\Lambda}_{\mathrm{W}} \mathbf{F}_P^H$ and $\mathbf{H}_{\mathrm{eqv}} \triangleq \mathbf{F}_P \mathbf{\Lambda}_{\mathrm{W}} \mathbf{\Lambda}_{\mathrm{fo}} \mathbf{F}_P^H$ are circulant matrices, and $\mathbf{V} \triangleq \mathbf{F}_P \mathbf{F}_{\mathrm{zp}}$. Since $\mathbf{\Lambda}_{\mathrm{J}}$ is a diagonal matrix, it does not affect the sparsity of $\mathbf{J}$; hence, $\mathbf{\Lambda}_{\mathrm{J}} \mathbf{J}$ is also a sparse vector with the same support as $\mathbf{J}$ but with different amplitudes for its non-zero entries. The effect of windowing is to diminish the off-diagonal elements of the circulant matrix $\mathbf{H}_{\mathrm{eqv}}$ such that the equivalent NBI vector, $\mathbf{J}_{\mathrm{eqv}}$, is spectrally contained, i.e., sparse.

## 2.3 CS-Based NBI Estimation for SISO-OFDM

To illustrate the main ideas, we start with the simplest case of a SISO ZP-OFDM system impaired by a single NBI signal and assume perfect knowledge of the desired signal's channel. Starting from (2.6), the sparse FD NBI vector $\mathbf{J}_{\mathrm{eqv}}$ can be recovered from $\mathbf{Y}$ using CS-based convex optimization techniques. However, being unknown, $\mathbf{X}$ and $\mathbf{Z}$ have to be modelled as noise which reduces the accuracy of NBI estimation. Instead, we reduce the noise level by canceling the unknown data term in (2.6) first. This technique is known as signal-blocking in the context of array processing. Assuming knowledge of $\mathbf{\Lambda}$ and, hence, $\tilde{\mathbf{\Lambda}}$ at the receiver, we filter $\mathbf{Y}$ by $\mathbf{W}$ such that $\mathbf{W}\tilde{\mathbf{\Lambda}} = \mathbf{0}$. To this end, $\mathbf{W}$ is designed to be the projection

[2]However, in simulations, we generate $\mathbf{H}_{\mathrm{J}}$ as a Toeplitz matrix.

matrix on the left null-subspace of $\tilde{\mathbf{\Lambda}}$ as follows [3]

$$\mathbf{W} = \mathbf{I}_P - \tilde{\mathbf{\Lambda}}\tilde{\mathbf{\Lambda}}^{\dagger} \tag{2.7}$$

where $\tilde{\mathbf{\Lambda}}^{\dagger} \triangleq \left(\tilde{\mathbf{\Lambda}}^H \tilde{\mathbf{\Lambda}}\right)^{-1} \tilde{\mathbf{\Lambda}}^H$ is the Moore-Penrose pseudo-inverse [62]. Since $\tilde{\mathbf{\Lambda}}$ is a *tall* matrix, it has a nontrivial (i.e., non-zero) left null-subspace. Now, it is clear that we kept the received guard sequence to make $\tilde{\mathbf{\Lambda}}$ a tall matrix. Multiplying $\mathbf{Y}$ by $\mathbf{W}$, we get

$$\widetilde{\mathbf{Y}} = \mathbf{W}\mathbf{Y} = \mathbf{W}\mathbf{J}_{\text{eqv}} + \underbrace{\mathbf{W}\mathbf{Z}}_{:=\tilde{\mathbf{Z}}}. \tag{2.8}$$

Note that $\mathbf{W}$ is a rank-$\nu$ matrix; hence, the linear system in (2.8) is under-determined and falls under CS theory framework. Therefore, we formulate and solve the following convex program to estimate $\mathbf{J}_{\text{eqv}}$

$$\min_{\tilde{\mathbf{J}}_{\text{eqv}} \in \mathbb{C}^P} \| \tilde{\mathbf{J}}_{\text{eqv}} \|_1 \quad \text{subject to} \quad \| \widetilde{\mathbf{Y}} - \mathbf{W}\tilde{\mathbf{J}}_{\text{eqv}} \|_2^2 \leq \epsilon \tag{2.9}$$

where $\epsilon$ is a design parameter which is chosen to be a slightly greater than $E\left[\| \tilde{\mathbf{Z}} \|_2^2\right] \triangleq \text{tr}\left(\tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^H\right) = N_o \text{tr}(\mathbf{W}\mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{\Lambda}_{\text{W}}^H\mathbf{F}_P^H\mathbf{W}^H)$ where $E[.]$ denotes the statistical expectation and tr(.) denotes the matrix trace operation. Note that knowledge of $N_o$ is required for NBI estimation and data detection. Next, we cancel $\tilde{\mathbf{J}}_{\text{eqv}}$ from $\mathbf{Y}$, remove the window, and proceed to the data detection and decoding stage.

Our CS-based approach assumes that multiplying $\mathbf{Y}$ by $\mathbf{W}$ does not cancel the NBI signal. To derive the condition for $\mathbf{W}$ not to cancel the NBI term, we consider the simple case where the NBI is exactly sparse with $r$ nonzero elements, i.e., there is no leakage and, hence, no windowing. Ignoring the noise term, we write $\mathbf{\Lambda}^{-1}\mathbf{Y} = \mathbf{V}\mathbf{X} + \tilde{\mathbf{J}}$ where $\tilde{\mathbf{J}} = \mathbf{\Lambda}^{-1}\mathbf{\Lambda}_{\text{J}}\mathbf{J}$

[3] Note that if $\mathbf{P}$ projects on the column subspace of a matrix, then $\mathbf{I}-\mathbf{P}$ projects on its orthogonal complement which is the left null-subspace.

is still an $r$-sparse vector. The signal-blocking matrix in this case is $\mathbf{W} = \mathbf{I}_P - \mathbf{V}\mathbf{V}^H$ where $\mathbf{V}^H = \mathbf{V}^\dagger$. Now, the NBI term will be cancelled if $\mathbf{W}\tilde{\mathbf{J}} = \mathbf{0}$ which is a set of $\nu$ equations with $r$ unknowns because $\mathbf{W}$ is a rank-$\nu$ matrix and $\tilde{\mathbf{J}}$ is an $r$-sparse vector. For this set of equations to have a non-trivial (i.e., nonzero) solution, $r$ (number of unknowns) has to be greater than or equal to $\nu$ (number of independent equations). In words, the NBI spectral support (in subcarriers) has to be less than or equal to the guard sequence length, and the NBI has to lie in the null-space of $\mathbf{W}$. In fact, if $r > \nu$, our CS-based approach will fail even if the NBI signal is not cancelled because the number of measurements will be smaller than the sparsity level, and this violates CS-based sparse signal recovery. This analysis assumes that the coefficients matrix of this set of equations has no column with all-zero entries, so all nonzero entries of $\tilde{\mathbf{J}}$ are excited regardless of their locations. In ZP-OFDM, this is a valid assumption thanks to the inherent precoding matrix $\mathbf{V}$. If $\mathbf{V}$ is replaced, for example, by $\begin{bmatrix} \mathbf{I}_N \\ \mathbf{0}_{\nu\times N} \end{bmatrix}$, then this assumption is violated as the coefficients matrix would have $N$ columns with all-zero entries.

## 2.4 Extension to MIMO-OFDM Systems

For concreteness, we describe our approach for spatial diversity techniques using two transmit antennas. However, it can be generalized using the theory of orthogonal designs [63] to larger numbers of transmit antennas; for example four transmit antennas as in the IEEE 802.11n standard [64]. We start by developing a space time block coding (STBC) scheme for ZP-OFDM systems.

### 2.4.1 STBC for ZP-OFDM

Inspired by the Alamouti STBC scheme [65] and the STBC scheme of ZP single-carrier transmission [53], we propose the following transmit diversity scheme for ZP-OFDM with 2 transmit antennas. First, the coded symbols are divided into $N\times 1$ blocks $\{\mathbf{X}(k), k = 0, 1, ..\}$. Denoting the transmitted OFDM block at time $k$ from the $l^{\text{th}}$ antenna by $\mathbf{x}_l(k)$, we write the outputs of the 2 transmit antennas at times 0 and 1 as follows

$$\mathbf{x}_1(0) = \mathbf{F}_{\text{zp}}\mathbf{X}(0),\ \ \mathbf{x}_2(0) = \mathbf{F}_{\text{zp}}\mathbf{X}(1),\ \ \mathbf{x}_1(1) = -\mathbf{F}_{\text{zp}}\mathbf{P}_{tx}\mathbf{X}^*(1),\ \ \mathbf{x}_2(1) = \mathbf{F}_{\text{zp}}\mathbf{P}_{tx}\mathbf{X}^*(0) \tag{2.10}$$

where $\mathbf{P}_{tx}$ is a precoding matrix introduced to make the equivalent MIMO channel matrix in the form of the Quaternion matrix[4] which enables simple linear-complexity maximum-likelihood detection at the receiver with maximum spatial diversity order. Assuming that the channel is quasi-static over two consecutive OFDM blocks, we write the un-windowed NBI-free received OFDM blocks at times 0 and 1 at the $q^{\text{th}}$ receive antenna as follows

$$\begin{aligned}\mathbf{y}^q(0) &= \mathbf{H}_1^q\mathbf{F}_{\text{zp}}\mathbf{X}(0) + \mathbf{H}_2^q\mathbf{F}_{\text{zp}}\mathbf{X}(1) + \mathbf{z}^q(0),\\ \mathbf{y}^q(1) &= -\mathbf{H}_1^q\mathbf{F}_{\text{zp}}\mathbf{P}_{tx}\mathbf{X}^*(1) + \mathbf{H}_2^q\mathbf{F}_{\text{zp}}\mathbf{P}_{tx}\mathbf{X}^*(0) + \mathbf{z}^q(1)\end{aligned} \tag{2.11}$$

where $1 \le q \le N_r$ with $N_r$ denoting the number of receive antennas, $\mathbf{z}^q(k)$ is the complex AWGN vector of the $q^{\text{th}}$ receive antenna at time $k$, and $\mathbf{H}_l^q$ is the $P \times P$ circulant channel matrix between the the $l^{\text{th}}$ transmit antenna and the $q^{\text{th}}$ receive antenna. Note that we remove the window after estimating the NBI; hence, we do not consider it while designing the STBC scheme. It was shown in [53] that $\mathbf{P}_P^{(n)}\mathbf{H}^*\mathbf{P}_P^{(n)} = \mathbf{H}^H$ for any $n$ and any circulant matrix $\mathbf{H}$ where $\mathbf{P}_P^{(n)}$ was defined in Section 2.1. Negating, conjugating, and multiplying

[4] A Quaternion matrix has the form $\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}$

$\mathbf{y}^q(1)$ by $\mathbf{P}_P^{(n)}$, we get

$$\begin{aligned}\hat{\mathbf{y}}^q(1) &\triangleq -\mathbf{P}_P^{(n)}(\mathbf{y}^q(1))^* \\ &= \mathbf{P}_P^{(n)}(\mathbf{H}_1^q)^*\mathbf{F}_{\text{zp}}^*\mathbf{P}_{tx}^*\mathbf{X}(1) - \mathbf{P}_P^{(n)}(\mathbf{H}_2^q)^*\mathbf{F}_{\text{zp}}^*\mathbf{P}_{tx}^*\mathbf{X}(0) - \mathbf{P}_P^{(n)}(\mathbf{z}^q(1))^*.\end{aligned} \tag{2.12}$$

Setting $\mathbf{F}_{\text{zp}}^*\mathbf{P}_{tx}^* = \mathbf{P}_P^{(n)}\mathbf{F}_{\text{zp}}$ and observing that $\mathbf{F}_{\text{zp}}^H\mathbf{F}_{\text{zp}} = \mathbf{I}_N$, we get $\mathbf{P}_{tx} = \mathbf{F}_{\text{zp}}^H\mathbf{P}_P^{(n)}\mathbf{F}_{\text{zp}}^*$ and

$$\hat{\mathbf{y}}^q(1) = -(\mathbf{H}_2^q)^H\mathbf{F}_{\text{zp}}\mathbf{X}(0) + (\mathbf{H}_1^q)^H\mathbf{F}_{\text{zp}}\mathbf{X}(1) - \mathbf{P}_P^{(n)}(\mathbf{z}^q(1))^*. \tag{2.13}$$

Finally, we choose $n$ in $\mathbf{P}_P^{(n)}$ such that the multiplication by $\mathbf{P}_{tx}$ at the transmitter is easily implemented. This can be achieved if $\mathbf{P}_{tx}$ is designed to be a diagonal matrix, and it can be easily shown that setting $n = N$ achieves this goal. Stacking the $P$-point DFTs of $\mathbf{y}^q(0)$ and $\hat{\mathbf{y}}^q(1)$ in a single vector, we get

$$\mathbf{Y}^q \triangleq \begin{bmatrix}\mathbf{F}_P\mathbf{y}^q(0) \\ \mathbf{F}_P\hat{\mathbf{y}}^q(1)\end{bmatrix} = \underbrace{\begin{bmatrix}\mathbf{\Lambda}_1^q & \mathbf{\Lambda}_2^q \\ -(\mathbf{\Lambda}_2^q)^H & (\mathbf{\Lambda}_1^q)^H\end{bmatrix}}_{\triangleq\mathbf{\Lambda}^q}\begin{bmatrix}\mathbf{V}\mathbf{X}(0) \\ \mathbf{V}\mathbf{X}(1)\end{bmatrix} + \begin{bmatrix}\mathbf{Z}^q(0) \\ \mathbf{Z}^q(1)\end{bmatrix} \tag{2.14}$$

where $\mathbf{\Lambda}_l^q \triangleq \mathbf{F}_p\mathbf{H}_l^q\mathbf{F}_p^H$ is a diagonal matrix whose diagonal is the $P$-point DFT of the first column of $\mathbf{H}_l^q$. Also, $\mathbf{Z}^q(0) \triangleq \mathbf{F}_P\mathbf{z}^q(0)$ and $\mathbf{Z}^q(1) \triangleq -\mathbf{F}_P\mathbf{P}_P^N(\mathbf{z}^q(1))^*$. Note that the equivalent MIMO channel matrix, $\mathbf{\Lambda}^q$, is a Quaternion as desired.

### 2.4.2 CS-Based NBI Estimation for MIMO-OFDM

Now, consider the $2 \times N_r$ MIMO ZP-OFDM system. The receive antennas are placed far enough from each other to ensure that their received NBI signals are uncorrelated. Hence, we only need to describe our NBI estimation technique for one receive antenna since the same technique can be applied at each receive antenna to estimate and cancel its NBI signal. Again, we apply windowing before taking the DFT to enhance the sparsity of the NBI signal. Taking NBI and windowing into consideration, assuming that the NBI is quasi-static over

two consecutive OFDM symbols, and after some algebra, we get

$$\begin{aligned}\mathbf{Y}(0) &= \mathbf{F}_P\mathbf{\Lambda}_{\mathrm{W}}\mathbf{y}(0) = \mathbf{H}_{\mathrm{W}}\mathbf{\Lambda}_1\mathbf{V}\mathbf{X}(0) + \mathbf{H}_{\mathrm{W}}\mathbf{\Lambda}_2\mathbf{V}\mathbf{X}(1) + \mathbf{J}_{\mathrm{eqv}} + \mathbf{Z}(0)\\ \mathbf{Y}(1) &= -\mathbf{F}_P\mathbf{P}_P^{(N)}\mathbf{\Lambda}_{\mathrm{W}}\mathbf{y}^*(1) = -\mathbf{H}'_{\mathrm{W}}\mathbf{\Lambda}_2^H\mathbf{V}\mathbf{X}(0) + \mathbf{H}'_{\mathrm{W}}\mathbf{\Lambda}_1^H\mathbf{V}\mathbf{X}(1) + \tilde{\mathbf{L}}\mathbf{J}^*_{\mathrm{eqv}} + \mathbf{Z}(1)\end{aligned} \tag{2.15}$$

with

$$\begin{aligned}\tilde{\mathbf{L}} &= -\mathbf{F}_P\mathbf{P}_P^{(N)}\mathbf{F}_P^T\\ \mathbf{H}'_{\mathrm{W}} &= \mathbf{F}_P\mathbf{P}_P^{(N)}\mathbf{\Lambda}_{\mathrm{W}}\mathbf{P}_P^{(N)}\mathbf{F}_P^H\\ \mathbf{Z}(0) &= \mathbf{F}_P\mathbf{\Lambda}_{\mathrm{W}}\mathbf{z}(0)\\ \mathbf{Z}(1) &= -\mathbf{F}_P\mathbf{P}_P^{(N)}\mathbf{\Lambda}_{\mathrm{W}}\mathbf{z}^*(1)\end{aligned}$$

From the structure of $\mathbf{P}_P^{(N)}$, we observe that $\tilde{\mathbf{L}}$ is diagonal and $\mathbf{H}'_{\mathrm{W}}$ is circulant. Stacking $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$ into a single vector, we write

$$\begin{aligned}\mathbf{Y} &\triangleq \begin{bmatrix}\mathbf{Y}(0)\\ \mathbf{Y}(1)\end{bmatrix}\\ &= \underbrace{\begin{bmatrix}\mathbf{H}_{\mathrm{W}}\mathbf{\Lambda}_1\mathbf{V} & \mathbf{H}_{\mathrm{W}}\mathbf{\Lambda}_2\mathbf{V}\\ -\mathbf{H}'_{\mathrm{W}}\mathbf{\Lambda}_2^H\mathbf{V} & \mathbf{H}'_{\mathrm{W}}\mathbf{\Lambda}_1^H\mathbf{V}\end{bmatrix}}_{\triangleq\tilde{\mathbf{\Lambda}}}\underbrace{\begin{bmatrix}\mathbf{X}(0)\\ \mathbf{X}(1)\end{bmatrix}}_{\triangleq\mathbf{X}} + \underbrace{\begin{bmatrix}\mathbf{I}_P & \mathbf{0}_{P\times P}\\ \mathbf{0}_{P\times P} & \tilde{\mathbf{L}}\end{bmatrix}}_{\triangleq\mathbf{L}}\begin{bmatrix}\mathbf{J}_{\mathrm{eqv}}\\ \mathbf{J}^*_{\mathrm{eqv}}\end{bmatrix} + \underbrace{\begin{bmatrix}\mathbf{Z}(0)\\ \mathbf{Z}(1)\end{bmatrix}}_{\triangleq\mathbf{Z}}\end{aligned} \tag{2.16}$$

For a better CS-based estimate of $\mathbf{J}$, we must cancel the unknown data term $\mathbf{X}$ in (2.16) as we did in Section 2.3. Assuming knowledge[5] of $\tilde{\mathbf{\Lambda}}$ at the receiver, we filter $\mathbf{Y}$ by $\mathbf{W}$ such that $\mathbf{W}\tilde{\mathbf{\Lambda}} = \mathbf{0}$. To this end, $\mathbf{W}$ is chosen to be the projection matrix on the left null-subspace of $\tilde{\mathbf{\Lambda}}$ given by

$$\mathbf{W} = \mathbf{I}_{2P} - \tilde{\mathbf{\Lambda}}\tilde{\mathbf{\Lambda}}^{\dagger} \tag{2.17}$$

Note that $\tilde{\mathbf{\Lambda}}$ is a *tall* matrix of size $2P \times 2N$; hence, it has a nontrivial (i.e, non-zero) left null-subspace. We keep the received guard sequences to make $\tilde{\mathbf{\Lambda}}$ a tall matrix. Multiplying

[5]Channel estimation is investigated in Section 2.5.

$\mathbf{Y}$ by $\mathbf{W}$, we get

$$\tilde{\mathbf{Y}} \triangleq \mathbf{WY} = \underbrace{\mathbf{WL}}_{\triangleq\widetilde{\mathbf{W}}} \underbrace{\begin{bmatrix}\mathbf{I}_P \\ \mathbf{I}_P\end{bmatrix}}_{\triangleq\mathbf{I}_1} \mathbf{J}_{\text{eqv},R} + i\,\mathbf{WL} \underbrace{\begin{bmatrix}\mathbf{I}_P \\ -\mathbf{I}_P\end{bmatrix}}_{\triangleq\mathbf{I}_2} \mathbf{J}_{\text{eqv},I} + \underbrace{\mathbf{WZ}}_{\triangleq\tilde{\mathbf{Z}}} \tag{2.18}$$

where the subscripts $R$ and $I$ denote the real and imaginary parts, respectively. Separating the real and imaginary parts of $\tilde{\mathbf{Y}}$, we get

$$\begin{aligned}\tilde{\mathbf{Y}}_R &= \begin{bmatrix}\widetilde{\mathbf{W}}_R\mathbf{I}_1 & -\widetilde{\mathbf{W}}_I\mathbf{I}_2\end{bmatrix}\begin{bmatrix}\mathbf{J}_{\text{eqv},R} \\ \mathbf{J}_{\text{eqv},I}\end{bmatrix} + \tilde{\mathbf{Z}}_R \\ \tilde{\mathbf{Y}}_I &= \begin{bmatrix}\widetilde{\mathbf{W}}_I\mathbf{I}_1 & \widetilde{\mathbf{W}}_R\mathbf{I}_2\end{bmatrix}\begin{bmatrix}\mathbf{J}_{\text{eqv},R} \\ \mathbf{J}_{\text{eqv},I}\end{bmatrix} + \tilde{\mathbf{Z}}_I.\end{aligned} \tag{2.19}$$

Vertically stacking $\tilde{\mathbf{Y}}_R$ and $\tilde{\mathbf{Y}}_I$, we get

$$\tilde{\mathbf{Y}}_T = \underbrace{\begin{bmatrix}\widetilde{\mathbf{W}}_R\mathbf{I}_1 & -\widetilde{\mathbf{W}}_I\mathbf{I}_2 \\ \widetilde{\mathbf{W}}_I\mathbf{I}_1 & \widetilde{\mathbf{W}}_R\mathbf{I}_2\end{bmatrix}}_{\triangleq\mathbf{A}} \underbrace{\begin{bmatrix}\mathbf{J}_{\text{eqv},R} \\ \mathbf{J}_{\text{eqv},I}\end{bmatrix}}_{\triangleq\mathbf{J}_T} + \underbrace{\begin{bmatrix}\tilde{\mathbf{Z}}_R \\ \tilde{\mathbf{Z}}_I\end{bmatrix}}_{\triangleq\tilde{\mathbf{Z}}_T}. \tag{2.20}$$

Since $\mathbf{W}$ is a rank-deficient matrix (its rank is $2\nu$), the linear system described by (2.20) represents an under-determined linear system of equations in the same form of (2.2). Hence, a convex optimization program similar to those in (2.3) and (2.9) is formulated to estimate $\mathbf{J}_T$ and cancel it from the received signal. finally, we remove the window, and forward the signal to the detection and decoding units.

## 2.5 Channel Estimation in the Presence of NBI

So far we assumed that the receiver has perfect knowledge of the desired signal's channel for the computation of the $\mathbf{W}$ matrices in (2.7) and (2.17). In this subsection, we relax this assumption and propose a novel technique to estimate the channel in the presence of NBI. We investigate channel estimation for $2 \times 1$ systems; however, our approach can also be extended to other MIMO configurations. The idea is to first estimate and cancel the NBI

prior to channel estimation by exploiting knowledge of the training matrix at the receiver. In the channel estimation phase, we transmit the following training sequences

$$\mathbf{x}_1(0) = \mathbf{x}_2(0) = \mathbf{F}_{\text{zp}}\mathbf{m}_0, \qquad \mathbf{x}_1(1) = -\mathbf{x}_2(1) = \mathbf{F}_{\text{zp}}\mathbf{m}_1, \tag{2.21}$$

where the sequences $\mathbf{m}_0$ and $\mathbf{m}_1$ are known at the receiver. Note that we need not follow the transmission scheme in (2.10) because our goal is not to detect the transmitted training symbols since they are already known at the receiver. The corresponding received OFDM symbols at times 0 and 1 are

$$\begin{aligned} \mathbf{y}(0) &= \mathbf{H}_1\mathbf{F}_{\text{zp}}\mathbf{m}_0 + \mathbf{H}_2\mathbf{F}_{\text{zp}}\mathbf{m}_0 + \mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j} + \mathbf{z}(0) \\ \mathbf{y}(1) &= \mathbf{H}_1\mathbf{F}_{\text{zp}}\mathbf{m}_1 - \mathbf{H}_2\mathbf{F}_{\text{zp}}\mathbf{m}_1 + \mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j} + \mathbf{z}(1). \end{aligned} \tag{2.22}$$

Since $\{\mathbf{H}_k,\ k = 1, 2\}$ are circulant matrices, we only need to estimate their first columns. However, the channel length most probably will not exceed the guard sequence length $\nu$. Then, we only need to estimate $\mathbf{h}_k$ which denotes the first $\nu$ elements of the first column of $\mathbf{H}_k$. In addition, since the channel is assumed time invariant over two consecutive OFDM symbols, it can be modeled as a linear time-invariant system whose input and impulse response can be commuted (commutative property of linear convolution). Then, (2.22) can be re-written as follows

$$\mathbf{y}(0) = \mathbf{M}_0(\mathbf{h}_1 + \mathbf{h}_2) + \mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j} + \mathbf{z}(0), \quad \mathbf{y}(1) = \mathbf{M}_1(\mathbf{h}_1 - \mathbf{h}_2) + \mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j} + \mathbf{z}(1) \tag{2.23}$$

where $\mathbf{M}_k$, $k = 0, 1$ are $P \times \nu$ tall Toeplitz matrices with their first columns equal to $\mathbf{F}_N^H\mathbf{m}_k$, $k = 0, 1$, respectively. Windowing the received signals and taking the DFT, we get

$$\begin{aligned} \mathbf{Y}(0) &\triangleq \mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{y}(0) = \tilde{\mathbf{M}}_0(\mathbf{h}_1 + \mathbf{h}_2) + \mathbf{J}_{\text{eqv}} + \mathbf{Z}(0) \\ \mathbf{Y}(1) &\triangleq \mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{y}(1) = \tilde{\mathbf{M}}_1(\mathbf{h}_1 - \mathbf{h}_2) + \mathbf{J}_{\text{eqv}} + \mathbf{Z}(1) \end{aligned} \tag{2.24}$$

where $\tilde{\mathbf{M}}_0 = \mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{M}_0$, $\tilde{\mathbf{M}}_1 = \mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{M}_1$, $\mathbf{Z}(0) = \mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{z}(0)$, $\mathbf{Z}(1) = \mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{z}(1)$, and $\mathbf{J}_{\text{eqv}} = \mathbf{F}_P\mathbf{\Lambda}_{\text{W}}\mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j}$. Note that $\tilde{\mathbf{M}}_0$ and $\tilde{\mathbf{M}}_1$ are tall matrices and known at the receiver.

Then, we multiply $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$ by the left null-space projection matrices of $\tilde{\mathbf{M}}_0$ and $\tilde{\mathbf{M}}_1$, respectively, to cancel the unknown terms containing the unknown CIR vectors as follows

$$\begin{aligned}\tilde{\mathbf{Y}}(0) &= \left(\mathbf{I}_P - \tilde{\mathbf{M}}_0\tilde{\mathbf{M}}_0^{\dagger}\right)\mathbf{Y}(0) \\ &= \left(\mathbf{I}_P - \tilde{\mathbf{M}}_0\tilde{\mathbf{M}}_0^{\dagger}\right)\mathbf{J}_{\text{eqv}} + \left(\mathbf{I}_P - \tilde{\mathbf{M}}_0\tilde{\mathbf{M}}_0^{\dagger}\right)\mathbf{Z}(0) \triangleq \mathbf{A}_0\mathbf{J}_{\text{eqv}} + \tilde{\mathbf{Z}}(0) \qquad (2.25)\\ \tilde{\mathbf{Y}}(1) &= \left(\mathbf{I}_P - \tilde{\mathbf{M}}_1\tilde{\mathbf{M}}_1^{\dagger}\right)\mathbf{Y}(1) \\ &= \left(\mathbf{I}_P - \tilde{\mathbf{M}}_1\tilde{\mathbf{M}}_1^{\dagger}\right)\mathbf{J}_{\text{eqv}} + \left(\mathbf{I}_P - \tilde{\mathbf{M}}_1\tilde{\mathbf{M}}_1^{\dagger}\right)\mathbf{Z}(1) \triangleq \mathbf{A}_1\mathbf{J}_{\text{eqv}} + \tilde{\mathbf{Z}}(1) \qquad (2.26)\end{aligned}$$

Observing that both (2.25) and (2.26) represent under-determined linear systems of equations, we recover $\mathbf{J}_{\text{eqv}}$ from either one by formulating and solving a convex $l_1$-norm minimization program as before. Note that the transmission scheme in (2.21) makes both $\mathbf{h}_1$ and $\mathbf{h}_2$ multiplied by the same matrix in (2.23) and, hence, both of them are cancelled in a single shot through multiplication by its left null-space projection matrix. Furthermore, $\mathbf{A}_0$ and $\mathbf{A}_1$ are computed only once offline since $\mathbf{M}_0$ and $\mathbf{M}_1$ are known and fixed. Next, we cancel the NBI estimate from $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$, convert back to TD, and remove the window to get

$$\begin{aligned}\mathbf{y}_c(0) &= \mathbf{M}_0(\mathbf{h}_1 + \mathbf{h}_2) + \tilde{\mathbf{z}}(0) \triangleq \mathbf{M}_0\mathbf{h}_a + \tilde{\mathbf{z}}(0) \\ \mathbf{y}_c(1) &= \mathbf{M}_1(\mathbf{h}_1 - \mathbf{h}_2) + \tilde{\mathbf{z}}(1) \triangleq \mathbf{M}_1\mathbf{h}_s + \tilde{\mathbf{z}}(1)\end{aligned} \qquad (2.27)$$

where $\tilde{\mathbf{z}}(k) = \mathbf{z}(k) + \mathbf{j}_{\text{res}}$, and $\mathbf{j}_{\text{res}}$ is the residual NBI after cancellation. The linear least-squares (LLS) estimators [66] of $\mathbf{h}_a$ and $\mathbf{h}_s$ have the following closed forms

$$\hat{\mathbf{h}}_a = \left(\mathbf{M}_0^H\mathbf{M}_0\right)^{-1}\mathbf{M}_0^H\mathbf{y}_c(0), \quad \hat{\mathbf{h}}_s = \left(\mathbf{M}_1^H\mathbf{M}_1\right)^{-1}\mathbf{M}_1^H\mathbf{y}_c(1) \qquad (2.28)$$

where the channel statistics are not required. Finally, $\hat{\mathbf{h}}_a$ and $\hat{\mathbf{h}}_s$ are used to estimate $\mathbf{h}_1$ and $\mathbf{h}_2$ as follows

$$\hat{\mathbf{h}}_1 = (\hat{\mathbf{h}}_a + \hat{\mathbf{h}}_s)/2, \qquad \hat{\mathbf{h}}_2 = (\hat{\mathbf{h}}_a - \hat{\mathbf{h}}_s)/2. \qquad (2.29)$$

We conclude this section with the following remark. Our CS-based approach can also be applied to other MIMO schemes such as spatial multiplexing provided that the number of receive antennas is greater than or equal to the number of transmit antennas. This condition is necessary to make the overall channel matrix tall with a nontrivial left-null subspace. With this condition satisfied, the application of our technique becomes straightforward.

## 2.6 Extensions to Practical Scenarios

In this section, we investigate several extensions of our CS-based approach to more advanced NBI scenarios including mobile and multiple NBI signals. In addition, we propose a reduced-complexity design of the $\mathbf{W}$ matrices used to cancel out the unknown data terms. Furthermore, we extend our approach to CP-OFDM.

### 2.6.1 Mobile NBI

In many practical scenarios, the NBI source is mobile. Hence, the NBI signal not only experiences a FS fading channel but also a fast-fading channel. Consequently, the NBI channel matrix $\mathbf{H}_{\mathrm{J}}$ is no longer Toeplitz and cannot be approximated as circulant. Therefore, $\mathbf{\Lambda}_{\mathrm{J}} = \mathbf{F}_P \mathbf{H}_{\mathrm{J}} \mathbf{F}_P^H$ is no longer diagonal; however, as shown in [67], it can be well approximated by a banded matrix with $2D+1$ significant diagonals, where $D$ depends on the NBI Doppler spread. When multiplied by $\mathbf{J}$, $\mathbf{\Lambda}_{\mathrm{J}}$ causes the FD NBI signal to spill over the $D$ adjacent subcarriers from each side. However, the window, which we apply to the received signal, reduces this spectral leakage such that the overall NBI vector, $\mathbf{J}_{\mathrm{eqv}}$, is still sparse. Nevertheless, $\mathbf{J}_{\mathrm{eqv}}$ will have a larger spectral width than that of the original NBI vector $\mathbf{J}$. Observe that both the NBI mobility and asynchronicity with the transmitted signal spread the NBI energy over

all the subcarriers. It is also worth mentioning that we resort to the windowing technique to enhance the NBI sparsity because we can not estimate the NBI channel and the FO between the NBI and the desired signal. However, if this information was somehow available at the receiver, we could simply reverse their effects or lump them into the measurement matrix when solving the convex optimization program. Furthermore, we can no longer assume that the NBI is quasi-static over two consecutive OFDM symbols as in Section 2.4.2 since the NBI channel actually varies within each OFDM symbol. For mobile NBI, (2.16) becomes

$$\mathbf{Y} = \tilde{\mathbf{\Lambda}}\mathbf{X} + \mathbf{Z} + \mathbf{L}\underbrace{\begin{bmatrix}\mathbf{J}_{\text{eqv},0}^T & \mathbf{J}_{\text{eqv},1}^H\end{bmatrix}^T}_{\triangleq \mathbf{J}_T} \tag{2.30}$$

where $\mathbf{L}$ was defined in (2.16) and $\mathbf{J}_{\text{eqv},0}$ and $\mathbf{J}_{\text{eqv},1}$ are the equivalent NBI vectors over OFDM symbols 0 and 1, respectively. Designing $\mathbf{W}$ as in (2.17) and applying it to $\mathbf{Y}$, we get

$$\tilde{\mathbf{Y}} = \mathbf{W}\mathbf{Y} = \underbrace{\mathbf{W}\mathbf{L}}_{\triangleq \mathbf{A}}\mathbf{J}_T + \underbrace{\mathbf{W}\mathbf{Z}}_{\triangleq \tilde{\mathbf{Z}}}. \tag{2.31}$$

Again, (2.31) has the same form as (2.2), and CS-based techniques can be used to estimate $\mathbf{J}_T$.

### 2.6.2 Multiple NBI Signals

Here, we investigate the scenario of multiple NBI signals coming from multiple sources where each NBI signal experiences its own (generally, time-varying) channel and FO with the desired signal. In this case, (2.16) becomes

$$\mathbf{Y} = \tilde{\mathbf{\Lambda}}\mathbf{X} + \mathbf{Z} + \mathbf{L}\underbrace{\left[\sum_{l=1}^{N_J}\left(\mathbf{J}_{\text{eqv},0}^l\right)^T \quad \sum_{l=1}^{N_J}\left(\mathbf{J}_{\text{eqv},1}^l\right)^H\right]^T}_{\triangleq \mathbf{J}_T} \tag{2.32}$$

where $N_J$ denotes the number of NBI sources and $\mathbf{J}_{\text{eqv},0}^l$ and $\mathbf{J}_{\text{eqv},1}^l$ are the equivalent NBI vectors from the $l^{\text{th}}$ source over OFDM symbols 0 and 1, respectively. From (2.32), we

observe that the multiple NBI scenario still falls under our CS-based framework unless the number and/or the spectral widths of the NBI signals are excessively large such that $\mathbf{J}_T$ is no longer sparse. However, this is not a practical situation as far as *narrow*-band interference is concerned. Note that we need not know the number of NBI signals nor their channels.

### 2.6.3 Reduced-Complexity Design

In this section, we propose a reduced-complexity design of the $\mathbf{W}$ matrix which we apply to the windowed received signal to cancel out the unknown data term. The proposed designs in (2.7) and (2.17) require the computation of the pseudo-inverse matrix which is computationally complex. The idea is to replace $\tilde{\mathbf{\Lambda}}^{\dagger}$ in (2.7) and (2.17) by a product of easy-to-compute matrices such that this product is equal to the identity matrix when multiplied by $\tilde{\mathbf{\Lambda}}$ which is the same role of the matrix $\tilde{\mathbf{\Lambda}}^{\dagger}$. Investigating the structure of $\tilde{\mathbf{\Lambda}}$ defined in (2.16), we propose the following alternative $\mathbf{W}$ design for the MIMO[6] case

$$\mathbf{W} = \mathbf{I}_{2P} - \tilde{\mathbf{\Lambda}} \begin{bmatrix} \mathbf{V}^H & \mathbf{0}_{N\times P} \\ \mathbf{0}_{N\times P} & \mathbf{V}^H \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_t^{-1} & \mathbf{0}_{P\times P} \\ \mathbf{0}_{P\times P} & \mathbf{\Lambda}_t^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_1^H & -\mathbf{\Lambda}_2 \\ \mathbf{\Lambda}_2^H & \mathbf{\Lambda}_1 \end{bmatrix} \begin{bmatrix} \mathbf{H}_{\mathrm{W}}^{-1} & \mathbf{0}_{P\times P} \\ \mathbf{0}_{P\times P} & (\mathbf{H}'_{\mathrm{W}})^{-1} \end{bmatrix} \tag{2.33}$$

where $\mathbf{\Lambda}_t \triangleq \mathbf{\Lambda}_1^H\mathbf{\Lambda}_1 + \mathbf{\Lambda}_2^H\mathbf{\Lambda}_2$ is a diagonal matrix whose inverse is easily computed. Furthermore, $\mathbf{H}_{\mathrm{W}}^{-1} = \mathbf{F}_P\mathbf{\Lambda}_{\mathrm{W}}^{-1}\mathbf{F}_P^H$ and $(\mathbf{H}'_{\mathrm{W}})^{-1} = \mathbf{F}_P\mathbf{P}_P^{(N)}\mathbf{\Lambda}_{\mathrm{W}}^{-1}\mathbf{P}_P^{(N)}\mathbf{F}_P^H$ where $\mathbf{\Lambda}_{\mathrm{W}}^{-1}$ is easily computed since $\mathbf{\Lambda}_{\mathrm{W}}$ is diagonal. Instead of storing $\mathbf{W}$ and multiplying it by the vector $\mathbf{Y}$, we exploit the structure of $\mathbf{W}$ and filter $\mathbf{Y}$ by $\mathbf{W}$ to obtain $\mathbf{WY}$ directly. Furthermore, the $k^{\text{th}}$ column of $\mathbf{W}$ is computed by filtering the $k^{\text{th}}$ column of $\mathbf{I}_{2P}$ by $\mathbf{W}$. In Tables 2.1 and 2.2, we list the number of real multiplications and additions needed to filter a length-$2P$ vector by the

[6]The design of $\mathbf{W}$ for SISO systems is even simpler.

Table 2.1. Complexity analysis of the original design of $\mathbf{W}$ in (2.17)

| | Original design of $\mathbf{W}$ in (2.17) |
|---|---|
| Real multiplications | $\sim \mathrm{O}\big(40P + 24P\mathrm{Log}_2 P\big)$ |
| Real additions | $\sim \mathrm{O}\big(28P + 24P\mathrm{Log}_2 P\big)$ |

Table 2.2. Complexity analysis of the reduced-complexity design of $\mathbf{W}$ in (2.33)

| | Reduced-complexity design of $\mathbf{W}$ in (2.33) |
|---|---|
| Real multiplications | $\sim \mathrm{O}\big(8N^3 + 16N^2 + 36P + 24PLog_2P + 8NLog_2N\big)$ |
| Real additions | $\sim \mathrm{O}\big(28P + 16N^2 + 24P\mathrm{Log}_2 P + 8N\mathrm{Log}_2 N\big)$ |

original and reduced-complexity designs of $\mathbf{W}$, respectively. We assume that $P$ is a power of 2 to simplify the complexity analysis.

### 2.6.4 Extension to CP-OFDM

So far, we used ZP-OFDM instead of CP-OFDM to avoid IBI in the guard sequence needed to get a tall matrix in front of the unknown data term to ensure a non-trivial left null-subspace. In this section, we present a method to obtain a tall matrix in CP-OFDM without keeping the guard sequence. This method is based on the fact that the FD data vector $\mathbf{X}$ contains elements known at the receiver. In the IEEE802.11n standard [2], for example, $\mathbf{X}$ contains 8 (14) null subcarriers and 4 (6) pilot subcarriers in the high-throughput (HT) 20 (40) MHz modes, respectively. In the non-HT mode, more null subcarriers are used. We denote the number of the known elements of $\mathbf{X}$ as $u$ and denote the indices of the known and unknown elements of $\mathbf{X}$ by $I_K$ and $I_U$, respectively. We explain our approach for SISO systems to simplify the presentation. Following Section 2.4.2, the extension to MIMO systems becomes straightforward with the only difference being that the STBC scheme would be the conventional Alamouti [65] scheme. The received CP-OFDM symbol after CP removal

can be written as follows

$$\mathbf{y} = \mathbf{H}\mathbf{F}_N^H\mathbf{Q}\mathbf{X} + \mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j} + \mathbf{z} \tag{2.34}$$

where $\mathbf{H}$ and $\mathbf{H}_{\text{J}}$ are the $N \times N$ channel matrices of the desired and NBI signals, respectively, $\mathbf{Q}$ is a unitary precoding matrix, $\mathbf{\Lambda}_{\text{fo}} = \text{diag}\left(1, \exp\left(i\frac{2\pi\alpha}{N}\right), \ldots, \exp\left(i\frac{2\pi\alpha(N-1)}{N}\right)\right)$, and $\mathbf{z}$ and $\mathbf{j}$ are the noise and NBI vectors, respectively. To get a tall matrix in front of the unknown data term, we perform the following operation

$$\mathbf{y}_1 = \mathbf{y} - \mathbf{H}\mathbf{F}_N^H\mathbf{Q}_K\mathbf{X}_K = \mathbf{H}\mathbf{F}_N^H\mathbf{Q}_U\mathbf{X}_U + \mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j} + \mathbf{z} \tag{2.35}$$

where $\mathbf{Q}_K$ and $\mathbf{Q}_U$ are sub-matrices of $\mathbf{Q}$ containing the columns whose indices are those in $I_K$ and $I_U$, respectively. Similarly, $\mathbf{X}_K$ and $\mathbf{X}_U$ are sub-vectors of $\mathbf{X}$ containing the known and unknown elements of $\mathbf{X}$, respectively. Applying the window to $\mathbf{y}_1$ and taking the DFT, we get

$$\begin{aligned}\mathbf{Y}_1 &\triangleq \mathbf{F}_N\mathbf{\Lambda}_{\text{W}}\mathbf{y}_1 \\ &= \underbrace{\mathbf{F}_N\mathbf{\Lambda}_{\text{W}}\mathbf{H}\mathbf{F}_N^H\mathbf{Q}_U}_{\triangleq\tilde{\mathbf{\Lambda}}}\mathbf{X}_U + \underbrace{\mathbf{F}_N\mathbf{\Lambda}_{\text{W}}\mathbf{\Lambda}_{\text{fo}}\mathbf{H}_{\text{J}}\mathbf{j}}_{\triangleq\mathbf{J}_{\text{eqv}}} + \underbrace{\mathbf{F}_N\mathbf{\Lambda}_{\text{W}}\mathbf{z}}_{\triangleq\mathbf{Z}} \triangleq \tilde{\mathbf{\Lambda}}\mathbf{X}_U + \mathbf{J}_{\text{eqv}} + \mathbf{Z}\end{aligned} \tag{2.36}$$

where $\tilde{\mathbf{\Lambda}}$ is an $N \times (N-u)$ *tall* matrix in front of the unknown vector $\mathbf{X}_U$. Consequently, we multiply $\mathbf{Y}_1$ by $\mathbf{W} \triangleq \mathbf{I}_N - \tilde{\mathbf{\Lambda}}\tilde{\mathbf{\Lambda}}^\dagger$ to cancel the unknown data term as follows

$$\tilde{\mathbf{Y}}_1 = \mathbf{W}\mathbf{Y}_1 = \mathbf{W}\mathbf{J}_{\text{eqv}} + \underbrace{\mathbf{W}\mathbf{Z}}_{\triangleq\tilde{\mathbf{Z}}} \tag{2.37}$$

Since the rank of $\mathbf{W}$ is $u < N$, the linear system in (2.37) is under-determined. Hence, we formulate and solve an $l_1$-norm minimization program similar to those in (2.3) and (2.9). Then, we estimate $\mathbf{J}_{\text{eqv}}$ and proceed as before. Using a precoding matrix prevents the cancellation of the NBI signal in the data cancellation step. It can be easily shown that without $\mathbf{Q}$, the NBI signal will be cancelled if it does not lie on the known tones even if

the NBI and the desired signals experience different channels. In ZP-OFDM, precoding is implicitly done through the matrix $\mathbf{V}$. Similar to the ZP-OFDM case, the reduced-complexity design is $\mathbf{W} \triangleq \mathbf{I}_N - \tilde{\mathbf{\Lambda}}\mathbf{Q}_U^H\mathbf{F}_N\mathbf{H}^{-1}\mathbf{\Lambda}_{\mathrm{W}}^{-1}\mathbf{F}_N^H$ where $\mathbf{H}^{-1}$ is easily computed since $\mathbf{H}$ is a circulant matrix. Note that the $(\nu - L + 1)$ IBI-free samples of the CP sequence can be used as in Section 2.4.2 to increase the number of measurements and improve the performance.

## 2.7 Simulation Results

We simulate the performance of a $2 \times 2$ MIMO-OFDM[7] system with $N = 128$ and $\nu = 32$. Selection receive diversity is used in the receiver. Although the channel realizations of the desired signal and the NBI signal are different, they both follow an exponentially-decaying power delay profile with $L = 8$ complex Gaussian taps. We employ a nonsystematic rate-1/2 convolutional code (CC) with octal generator (133,171) and constraint length = 7. Coded bits are quadrature phase shift keying (QPSK) modulated. We use the Hamming window to spectrally contain the NBI signal. The NBI temporal duty factor is assumed to be unity, i.e., the NBI signal affects every OFDM signal, which is a stringent assumption because the NBI typically goes on and off in most practical scenarios. The NBI carrier frequency is randomly generated in each OFDM symbol, so it can lie anywhere in the spectrum. Furthermore, unless otherwise stated, $\alpha$ is randomly generated in each OFDM symbol from the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$.

Denoting the information bit energy by $E_b$, we define the signal-to-noise ratio (SNR) as $\frac{E_b}{N_o}$ and the signal-to-interference ratio (SIR) as $\frac{E_b}{\mathbf{J}^H\mathbf{J}}$. Figure 2.1 shows the bit error rate

[7]Except for Figure 2.7 where we simulate a SISO system.

(BER) performance of our CS-based approach versus SNR for $r = 1$ and 3 with known and estimated channel information. In fact, NBI with $r = 3$ is of special importance to the IEEE 802.11g/n networks where the OFDM subcarrier spacing is 312.5 kHz [2] and the frequency-hopping Bluetooth signal occupies up to 1 MHz at a time. Hence, the Bluetooth signal plays the role of NBI with $r \approx 3$. Furthermore, we consider NBI signals with $r = 1$, i.e., single-tone jammers, since they are shown in [58] to cause the most degradation to the BER performance. Note that $r$ is the spectral width of the *transmitted* NBI signal; however the received NBI signal has a larger spectral width due to energy leakage as explained in Section 2.2.3. We also show the performance of the interference-free case and the case where NBI is present but ignored at the receiver. On the same figure, we compare our scheme with the threshold-excision (THE) [68] scheme where the energy of all the frequency components are computed and compared with a preset threshold. Then, the frequency subcarriers whose energy exceed this threshold are nulled out. In the simulations, this threshold is optimized numerically to get the smallest BER values over all SNR values. The superiority of our CS-based scheme to the THE scheme is evident in Figure 2.1 for both $r = 1$ and 3. Furthermore, our scheme achieves substantial performance gains compared to the case when NBI is ignored. We also observe that the performance loss due to channel estimation errors is insignificant for both $r = 1$ and 3. When ignoring the NBI, Figure 2.1 shows that single-tone NBI signals cause more degradation than multiple-tone NBI as reported in [58]. The opposite is true when applying our CS-based approach since single-tone NBI signals are more sparse in FD than multiple-tone ones; hence, our CS-based NBI estimates are more accurate for the former than the latter.

The case of NBI experiencing a time-varying FS channel is investigated in Figure 2.2

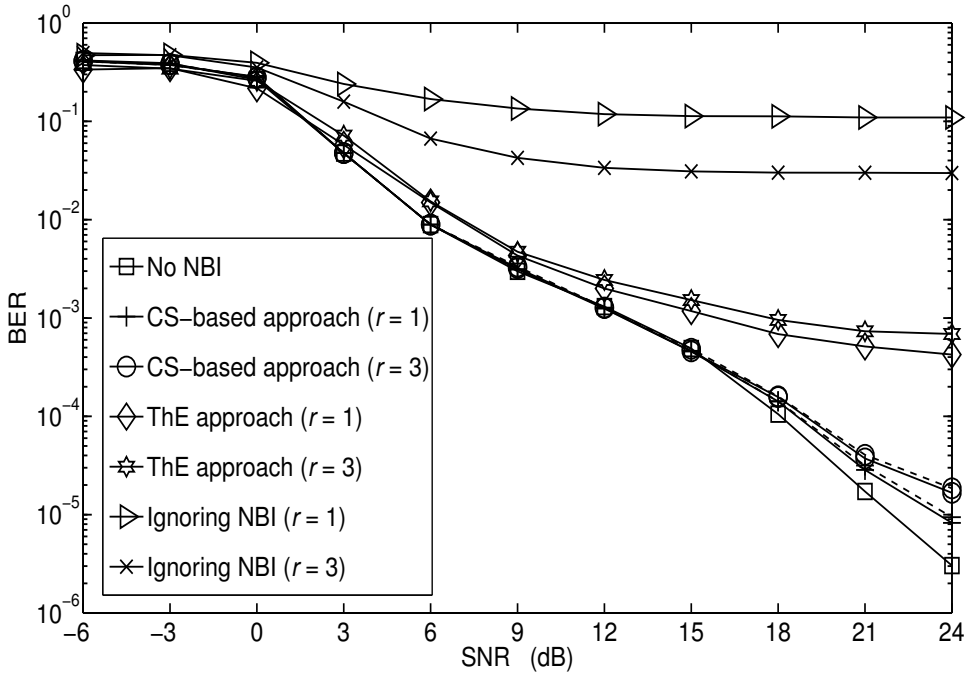


Figure 2.1. BER Performance for ZP-OFDM with known (solid lines) and estimated (dashed lines) channel for $r = 1$ and 3 with SIR = -10 dB.

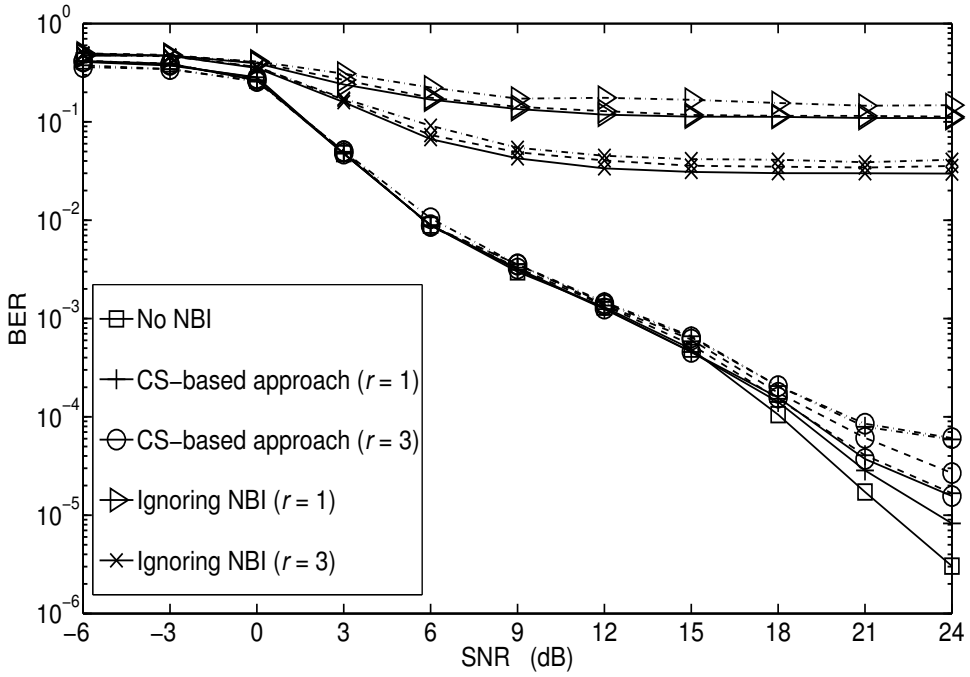


Figure 2.2. BER Performance for ZP-OFDM with NBI experiencing quasi-static (i.e., $D = 0$) (solid lines), time-varying with $D = 1$ (dashed lines), and time-varying with $D = 2$ (dash-dotted lines) FS channels for $r = 1$ and 3 with SIR = -10 dB.

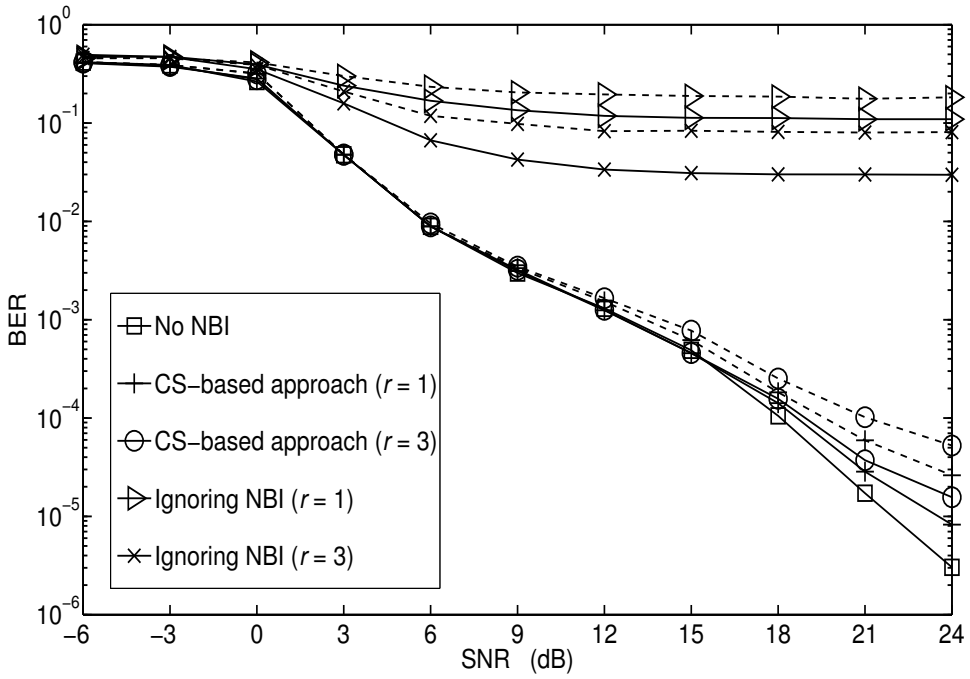


Figure 2.3. BER Performance for ZP-OFDM with one (solid lines) and two (dashed lines) NBI signals for $r = 1$ and 3 with SIR = -10 dB.

with $D = 1$ and 2, i.e., the NBI FD channel matrix has 3 and 5 main diagonals, respectively. The time variation of the NBI channel causes the NBI signal to further spread over all subcarriers; hence, additional performance degradation is expected as shown in Figure 2.2. However, the performance gains over the case where NBI is ignored are still substantial. In Figure 2.3, we show the performance with two NBI signals having the same $r$ but each of them experiences its own independent FS channel. The two NBI signals can be anywhere in the spectrum so they can, in general, overlap. For comparison, we also show the performance of a single NBI signal where it is clear that our CS-based approach retains its robustness to multiple NBI signals.

Figure 2.4 compares the performance of the reduced-complexity design of the $\mathbf{W}$ matrix to that of the original (more-complex) design of the $\mathbf{W}$ matrix. The reduced-complexity

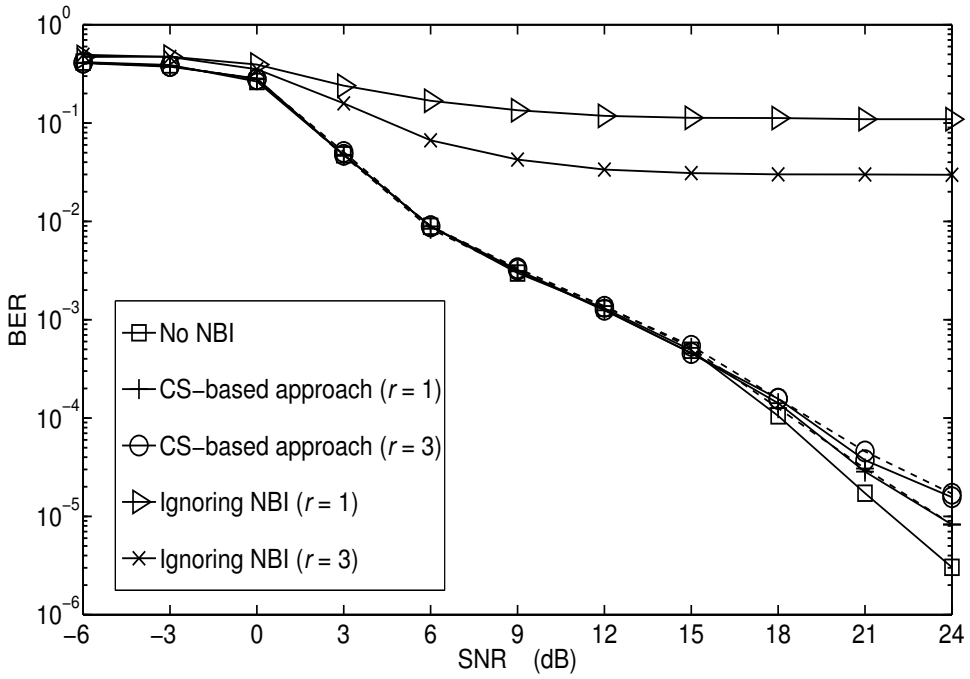


Figure 2.4. Performance Comparison between the reduced-complexity design (dashed lines) of $\mathbf{W}$ in (2.33) and the original more-complex design (solid lines) in (2.17) for $r = 1$ and 3 with SIR = -10 dB.

design virtually achieves the same performance of the more-complex design while avoiding the inversion of a channel-dependent matrix. In Figure 2.5, we investigate the effect of increasing the NBI power where it can be seen that our CS-based approach can mitigate strong NBI with SIR as low as -20 dB without significant performance loss. The performance of our CS-based approach for CP-OFDM systems is investigated in Figure 2.6 where we follow the OFDM symbol structure of the IEEE802.11n HT 40 MHz mode [2], i.e., 20 known subcarriers out of 128 total subcarriers in each OFDM symbol. In fact, even better performance is expected for the non-HT mode where more subcarriers are known at the receiver resulting in an under-determined system with more independent measurements (equations) and, hence, enhancing the sparse NBI recovery. In Figure 2.7, we compare the performances of our CS-based approach and two other approaches from the literature; namely, the LMMSE and

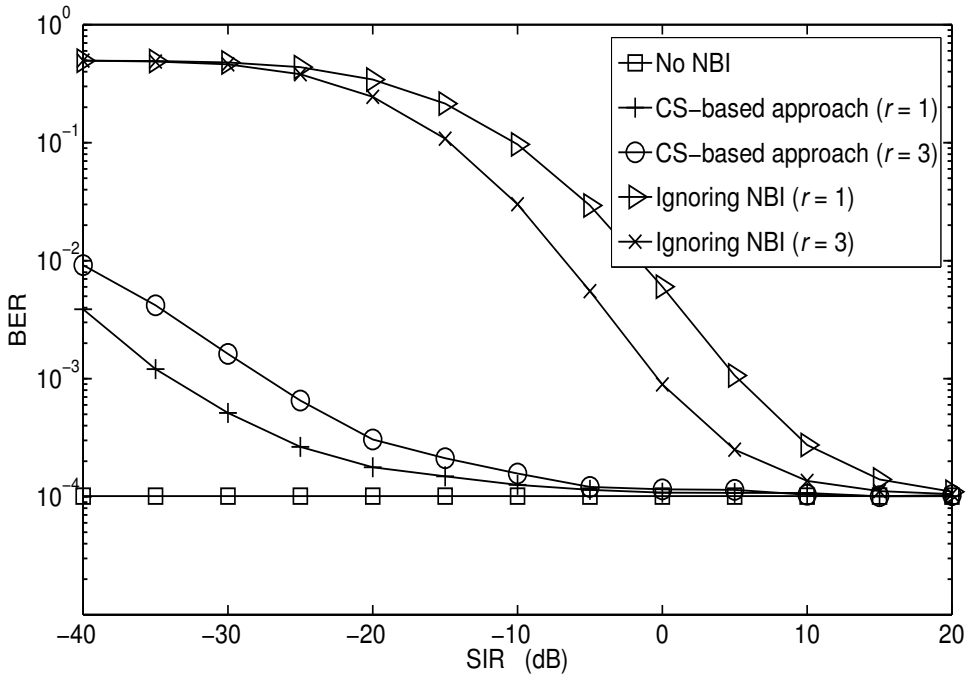


Figure 2.5. BER versus SIR for ZP-OFDM systems with $r = 1$ and 3 with SNR = 18 dB.

CMMOE approaches proposed in [9] and [10], respectively. We provided both approaches with the exact auto-correlation matrix of the NBI signal which involved the profile of the fading channel experienced by the NBI, the FO distribution, and the transmitted NBI vector. Furthermore, the LMMSE approach is genie-aided, *in each simulation run*, with perfect knowledge of locations of the 6 subcarriers with the highest interference power where these subcarriers are kept silent (unmodulated). Clearly, this is not a practical assumption because the transmitted NBI vector is not known and the silent subcarriers locations are kept fixed over all OFDM symbols regardless of the NBI spectral location in each symbol. We find that our CS-based approach, *without* any prior information, outperforms the LMMSE approach at the whole simulated SNR range and outperforms the CMMOE approach at high SNR levels.

Next, we investigate the impact of system parameters on the performance of our

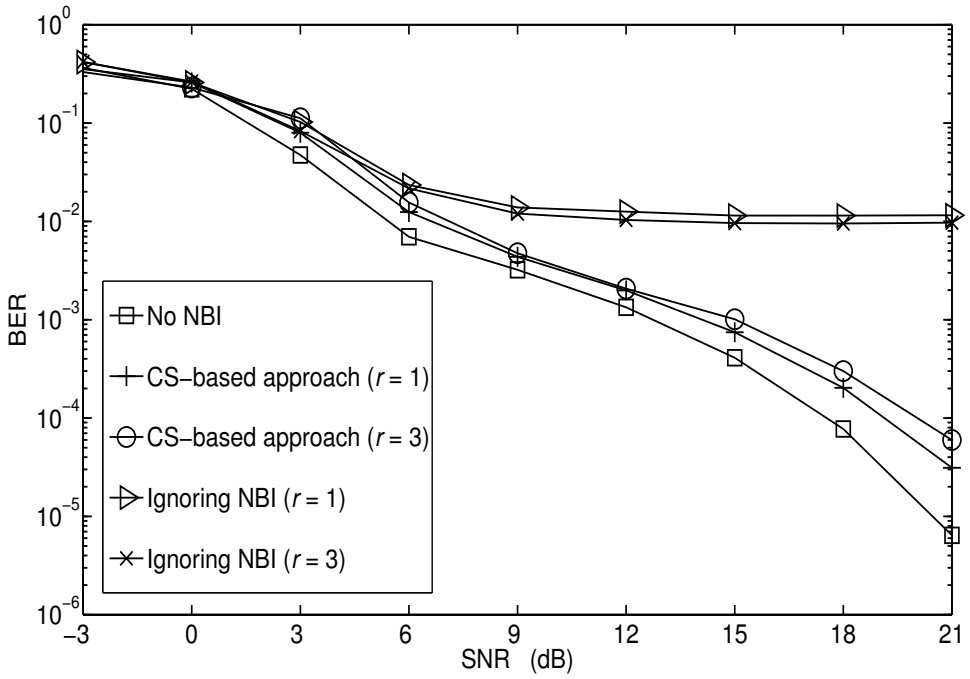


Figure 2.6. BER performance for CP-OFDM systems with the HT 40MHz mode for $r = 1$ and 3 with SIR = -10 dB.

approach for ZP-OFDM systems through the metric $\eta_{\mathrm{J}} \triangleq \frac{\sum_{m=1}^{M_s} \|\mathbf{J}_{\mathrm{eqv}} - \tilde{\mathbf{J}}_{\mathrm{eqv}}\|_2^2}{\sum_{m=1}^{M_s} \|\mathbf{J}_{\mathrm{eqv}}\|_2^2}$ where $M_s = 10^3$ is the number of realizations. The metric $\eta_{\mathrm{J}}$ is the normalized average error vector magnitude squared (AEVMS) of the NBI; therefore, small values of $\eta_{\mathrm{J}}$ indicate accurate NBI estimation, and vice verse. In Figure 2.8, we investigate the effect of the FO $\alpha$ on the performance with and without windowing. The NBI estimation accuracy degrades as $\alpha$ increases because the NBI signal becomes more spread in FD; hence, its sparsity is reduced.

However, using windowing reduces the amount of degradation significantly thanks to its role in spectrally containing the NBI signal. The impact of increasing the spectral width of the transmitted NBI signal $r$ is studied in Figure 2.9 for 1 and 2 transmit antennas. Increasing $r$ reduces the sparsity of the NBI signal in FD and, hence, reduces the NBI estimation accuracy. However, employing transmit diversity through multiple transmit antennas

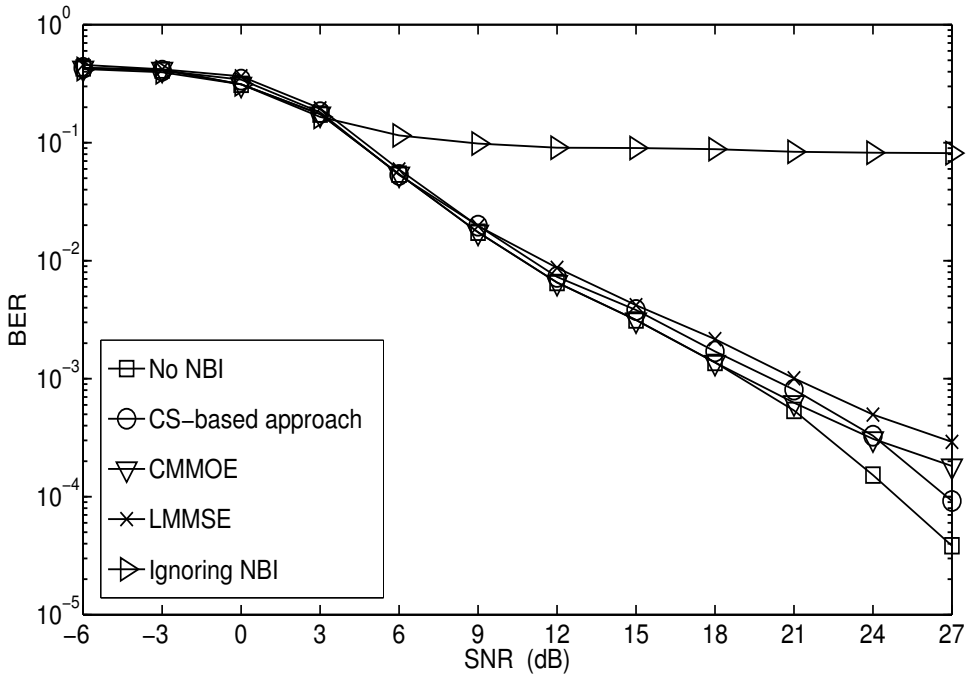


Figure 2.7. Performance comparison between CS, LMMSE, and CMMOE approaches for SISO CP-OFDM systems with the HT 40MHz mode for $r = 3$ with SIR = -10 dB.

increases the rank[8] of the measurement matrix $\mathbf{W}$ and, hence, the number of equations, so the NBI estimation accuracy is improved.

In Figure 2.10, we show that increasing the guard sequence length $\nu$ improves the performance as a result of increasing the number of equations. Finally, we show the performance of our approach at very high SNR levels in Figure 2.11 with and without NBI spectrum leakage. For the impractical case with no NBI leakage, our CS-based approach without windowing is able to recover the NBI signal almost perfectly at high SNR values since the NBI signal is exactly sparse. However, for practical cases, the NBI leakage becomes the limiting factor even at very high SNR values.

[8]According to Sections 2.3 and 2.4.2, rank of $\mathbf{W} = \nu$ and $2\nu$ for 1 and 2 transmit antennas.

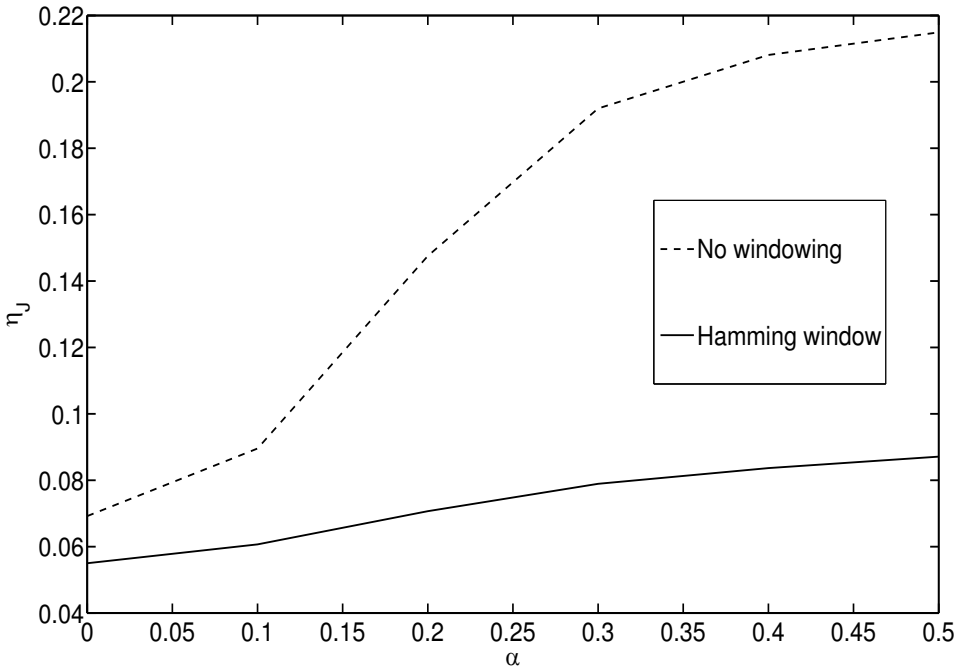


Figure 2.8. Impact of frequency-offset between desired signal and NBI signal on the performance with SNR = 20 dB, SIR = -10 dB, and $r = 3$.

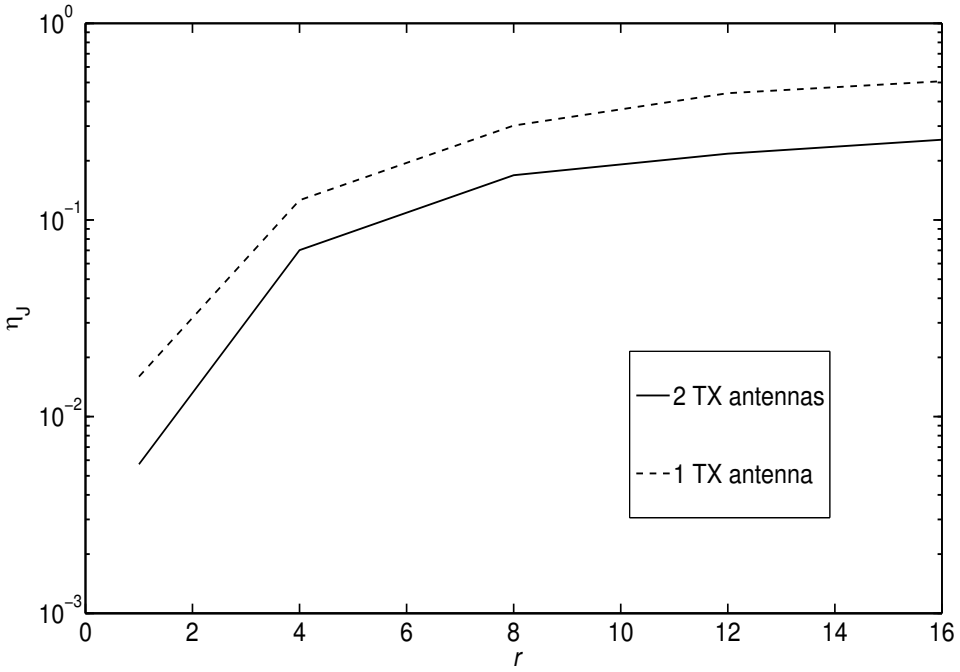


Figure 2.9. Impact of transmitted NBI spectral width $r$ on the performance with SNR = 25 dB, SIR = -10 dB, $\alpha = 0.1$, and 2 receive antennas with selection receive diversity.

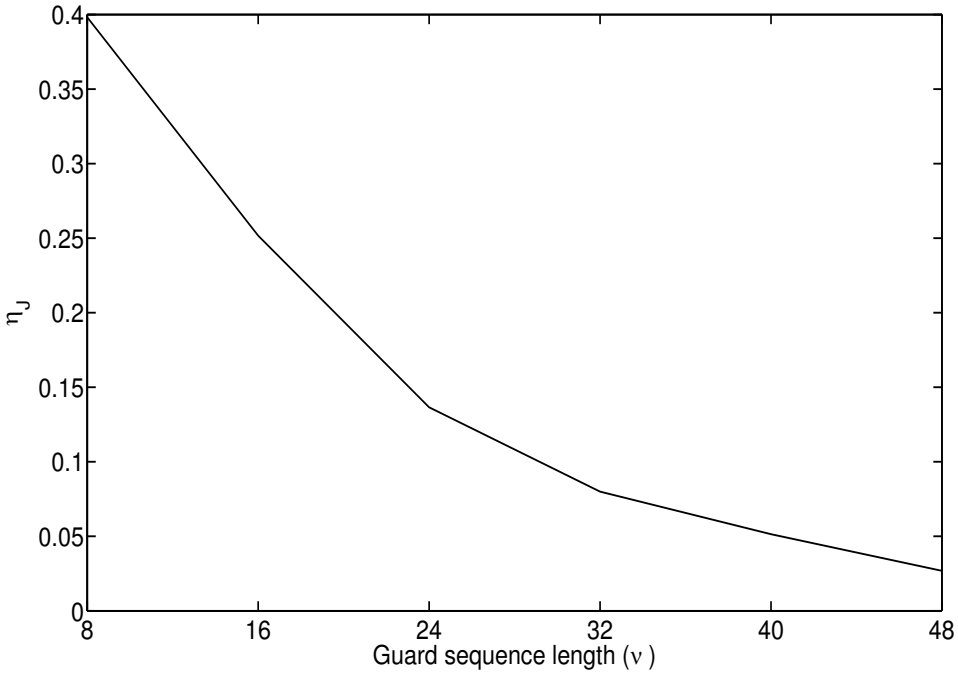


Figure 2.10. Impact of Guard sequence length $\nu$ on the performance with N = 128, SNR = 20 dB, SIR = -10 dB, $r = 3$, and $\alpha = 0.2$

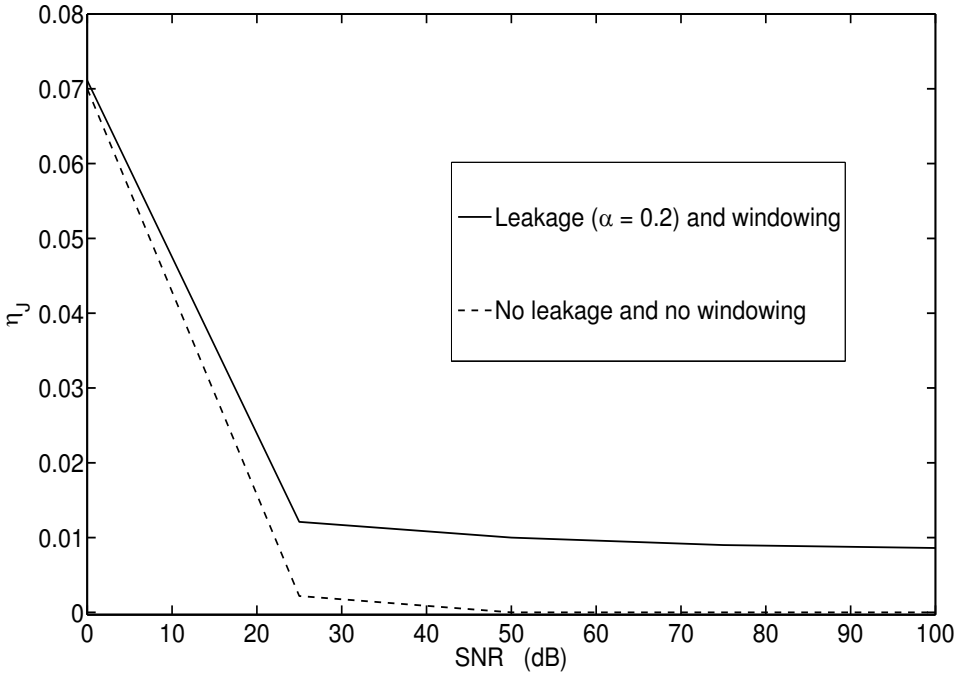


Figure 2.11. Performance at high SNR values with SIR = -10 dB and $r = 1$.

# CHAPTER 3

# DESIGN FRAMEWORK FOR SPARSE FIR MIMO EQUALIZERS

## 3.1 Introduction

In this chapter, we present a new framework for designing sparse SISO and MIMO FIR linear and non-linear equalizers. Specifically, we formulate both greedy algorithms and $l_1$-norm minimization programs to determine the locations and weights of the nonzero equalizer taps subject to a given maximum tolerable performance loss. Also, our formulation is not restricted to a specific greedy algorithm as in [18] and [19]. Furthermore, we generalize our sparse design framework to the channel shortening setup. Channel shortening is important for communication systems operating over broadband channels with long channel impulse responses. We propose a novel approach to design a sparse target impulse response and a sparse FIR channel shortening equalizer. Moreover, we consider self far-end crosstalk (FEXT) cancellation on vectored very high-speed digital subscriber line systems for cellular backhaul networks. We design a sparse self FEXT canceler subject to a given maximum performance loss from the highly-complex full self FEXT canceler.

The rest of this chapter is organized as follows. In Section 3.2, we provide a brief overview of sparse signals recovery techniques. The sparse equalization problem is formulated in Section 3.3. Our sparsity-aware approach for channel shortening is presented in Section 3.4. In Section 3.5, we describe our approach for partial self FEXT cancellation for vectored

VDSL. Finally, simulation results are presented in Sections 3.6. The contents of this chapter are adapted and reprinted from [69][1].

*Notations*: Unless otherwise stated, lower and upper case bold letters denote vectors and matrices, respectively, and $\mathbf{A}(:, i)$ denotes the $i^{\text{th}}$ column of $\mathbf{A}$. The matrix $\mathbf{I}$ denotes the identity matrix and its size is denoted by the subscript. The matrix $\mathbf{0}_{m\times n}$ denotes the all-zero matrix of size $m \times n$. Also, $(\ )^H$, $(\ )^*$, $(\ )^{-1}$, and $(\ )^\dagger$ denote the matrix complex-conjugate transpose, the complex conjugate, the matrix inverse, and the matrix pseudo-inverse operations, respectively. The portion of the vector $\mathbf{x}$ starting from the index $a$ and ending at the index $b$ is denoted by $\mathbf{x}_{a:b}$. The notation $\text{diag}\,(S_1, S_2, \ldots, S_L)$ denotes an $L \times L$ diagonal matrix whose diagonal elements are $\{S_1, S_2, \ldots, S_L\}$. The operators $\| \ . \ \|_0$, $\| \ . \ \|_1$, and $\| \ . \ \|_2$ denote the $l_0$-norm, $l_1$-norm, and the $l_2$-norm, respectively. The operators $E\,[.]$ and $|.|$ denote the statistical expectation and the absolute value, respectively.

### 3.2 Sparse Signals Recovery Background

Consider the system of equations $\mathbf{y} = \mathbf{A}\mathbf{x}+\mathbf{z}$ where $\mathbf{y} \in \mathbb{C}^M$ is a known measurement vector, $\mathbf{x} \in \mathbb{C}^N$ is an unknown vector, $\mathbf{A}$ denotes the $M \times N$ measurement matrix, and $\mathbf{z} \in \mathbb{C}^M$ is a bounded noise (error) vector. To obtain the sparsest solution to this system of equations, the following optimization problem is solved

$$\min_{\tilde{\mathbf{x}}\in\mathbb{C}^N} \| \tilde{\mathbf{x}} \|_0 \quad \text{subject to} \quad \| \mathbf{y} - \mathbf{A}\tilde{\mathbf{x}} \|_2^2 \leq \epsilon \tag{3.1}$$

where $\epsilon$ is chosen such that it bounds the amount of noise in the measurements and $\| \tilde{\mathbf{x}} \|_0$ is the number of nonzero entries of $\tilde{\mathbf{x}}$. However, in general, finding the optimal solution to this

[1]Copyright [2011] IEEE. Reprinted, with permission, from A. Gomaa and N. Al-Dhahir, A new Design Framework for Sparse FIR MIMO Equalizers, IEEE Transactions on Communications, August 2011

problem is not computationally feasible. Hence, two main approaches have been proposed in the literature to compute a sparse suboptimal solution to this system of equations; specifically, $l_1$-norm minimization and greedy algorithms. The $l_1$-norm minimization approach is formulated by replacing $\| \tilde{\mathbf{x}} \|_0$ in (3.1) by $\| \tilde{\mathbf{x}} \|_1$. However, the resulting solution is not exactly sparse because many small but nonzero entries will appear in $\tilde{\mathbf{x}}$. One method to enforce a finite number of nonzero entries is to apply an additional heuristic optimization step as in [20]. Alternatively, the greedy algorithms provide more control on the set of nonzero elements whose indices and values are determined iteratively. We describe one of the most widely-used greedy algorithms which is the orthogonal matching pursuit (OMP) [70] algorithm. It takes $\mathbf{y}$, $\mathbf{A}$, and a certain stopping criterion as its inputs and computes a sparse solution $\tilde{\mathbf{x}}$ for the unknown vector $\mathbf{x}$ as its output. Hence, we denote the OMP operation by $\tilde{\mathbf{x}} = \text{OMP}\ (\mathbf{y}, \mathbf{A}, \text{stopping criterion})$. The stopping criterion can be a predefined sparsity level (number of nonzero entries) of $\tilde{\mathbf{x}}$ or an upperbound on the norm of the residual error term $\| \mathbf{y} - \mathbf{A}\tilde{\mathbf{x}} \|_2$. The OMP algorithm is summarized in the following steps:

**Initialization**: Define an empty index set $I_0 = \phi$, set the initial residual $\mathbf{r}_0 = \mathbf{y}$, initialize $\tilde{\mathbf{x}} = \mathbf{0}$, and set $k = 1$.

**The $k^{\text{th}}$ iteration**:

1. Compute $\delta_i = \left| \mathbf{r}_{k-1}^H \mathbf{A}(:, i) \right|$ for all $i \notin I_{k-1}$.

2. Choose $c_k = \underset{i}{\arg\max}\ \delta_i$ where $c_k$ is the index of the nonzero entry computed at the $k^{\text{th}}$ iteration.

3. Update $I_k = I_{k-1} \cup c_k$. In this step, the indices of the nonzero elements are augmented by the index computed at the $k^{\text{th}}$ iteration.

4. Compute $\tilde{\mathbf{x}}(I_k) = (\mathbf{A}(:, I_k))^{\dagger}\, \mathbf{y}$ where $\tilde{\mathbf{x}}(I_k)$ holds the elements of $\tilde{\mathbf{x}}$ indexed by $I_k$.

5. Compute $\mathbf{r}_k = \mathbf{y} - \mathbf{A}(:, I_k)\tilde{\mathbf{x}}(I_k)$ where $\mathbf{r}_k$ is the residual error term at the $k^{\text{th}}$ iteration.

6. Check the stopping criterion. If met, exit the algorithm, else set $k = k + 1$ and go to Step 1.

In words, the OMP algorithm tries to find the columns (atoms) of the matrix $\mathbf{A}$ (dictionary) whose linear combination is close (matched) to $\mathbf{y}$. The orthogonal least squares (OLS) [71] is another greedy algorithm whose initialization step is the same as that of the OMP but the $k^{\text{th}}$ iteration is described as follows

1. Implement Steps 3, 4, and 5 in the OMP algorithm for each possible $c_k = i$ such that $i \notin I_{k-1}$ and find $c_k$ and $\tilde{\mathbf{x}}(I_k)$ that minimize the resulting residual error term $\| \mathbf{r}_k \|_2$.

2. Use $\tilde{\mathbf{x}}(I_k)$ from the previous step and implement Step 6 in the OMP algorithm.

Another greedy algorithm that can be used is the recently-proposed CoSaMP algorithm [72]. Furthermore, it is worth mentioning that the sparse equalization problems considered in this chapter are different from the well-known compressed sensing (CS) problem [11]. In CS theory, the measurement matrix $\mathbf{A}$ has much fewer rows than columns while, in our problems, as will be shown later, $\mathbf{A}$ is either square or tall with full column rank.

## 3.3 Sparse FIR Equalization

### 3.3.1 Signal Model

We consider a linear, time-invariant, dispersive, and noisy communication channel with $n_i$ input (transmit) antennas and $n_o$ output (receive) antennas. We use the complex-valued equivalent baseband signal model. Assuming an oversampling factor of $l$, the received samples at the $j^{\text{th}}$ output antenna ($1 \leq j \leq n_o$) at time $k$ have the form

$$\mathbf{y}_k^{(j)} = \sum_{i=1}^{n_i} \sum_{m=0}^{\nu^{(i,j)}} \mathbf{h}_m^{(i,j)} x_{k-m}^{(i)} + \mathbf{n}_k^{(j)} \tag{3.2}$$

where $\mathbf{y}_k^{(j)}$ is the $j^{\text{th}}$ channel output vector, $\mathbf{h}_m^{(i,j)}$ is the CIR between the $i^{\text{th}}$ input and the $j^{\text{th}}$ output whose memory is denoted by $\nu^{(i,j)}$, and $\mathbf{n}_k$ is the noise vector at the $j^{\text{th}}$ output antenna. All of these three quantities are $l \times 1$ column vectors corresponding to the $l$ time samples per symbol in the assumed temporally-oversampled channel model. Furthermore, $x_{k-m}^{(i)}$ is the transmitted symbol from the $i^{\text{th}}$ input antenna ($1 \leq i \leq n_i$). By grouping $\mathbf{y}_k^{(j)}, 1 \leq j \leq n_o$, into a single $ln_o \times 1$ vector $\mathbf{y}_k$, we can relate $\mathbf{y}_k$ to the corresponding $n_i \times 1$ vector of input samples as follows

$$\mathbf{y}_k = \sum_{m=0}^{\nu} \mathbf{H}_m \mathbf{x}_{k-m} + \mathbf{n}_k \tag{3.3}$$

where $\mathbf{H}_m$ is the $(ln_o \times n_i)$ $m^{\text{th}}$ MIMO channel matrix, $\mathbf{x}_{k-m}$ is the $n_i \times 1$ input vector at time $k-m$, and $\nu = \max\limits_{i,j} \ \nu^{(i,j)}$. Grouping $\{\mathbf{y}_k\}$ over a block of $N_f$ symbol periods, the input-output relation in (3.3) can be expressed as [16]

$$\mathbf{y}_{k:k-N_f+1} = \mathbf{H}\mathbf{x}_{k:k-N_f-\nu+1} + \mathbf{n}_{k:k-N_f+1} \tag{3.4}$$

where $\mathbf{y}_{k:k-N_f+1}$, $\mathbf{x}_{k:k-N_f-\nu+1}$, and $\mathbf{n}_{k:k-N_f+1}$ are column vectors grouping the received, transmitted, and noise samples over that block. Furthermore, $\mathbf{H}$ is a block Toeplitz

matrix whose first block row is constructed by the matrices $\{\mathbf{H}_m\}_{m=0}^{m=\nu}$ followed by zero matrices. The $n_i\,(N_f+\nu) \times n_i\,(N_f+\nu)$ input correlation matrix is defined by $\mathbf{R}_{xx} \equiv E\left[\mathbf{x}_{k:k-N_f-\nu+1}\mathbf{x}_{k:k-N_f-\nu+1}^H\right]$, and the $n_o\,(lN_f) \times n_o\,(lN_f)$ noise correlation matrix is defined by $\mathbf{R}_{nn} \equiv E\left[\mathbf{n}_{k:k-N_f+1}\mathbf{n}_{k:k-N_f+1}^H\right]$. Both $\mathbf{R}_{xx}$ and $\mathbf{R}_{nn}$ are assumed to be positive-definite correlation matrices. Furthermore, the output-input cross-correlation and the output auto-correlation matrices are given, respectively, by $\mathbf{R}_{yx} \equiv E\left[\mathbf{y}_{k:k-N_f+1}\mathbf{x}_{k:k-N_f-\nu+1}^H\right] = \mathbf{H}\mathbf{R}_{xx}$ and $\mathbf{R}_{yy} \equiv E\left[\mathbf{y}_{k:k-N_f+1}\mathbf{y}_{k:k-N_f+1}^H\right] = \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{R}_{nn}$.

In the sequel, we show how to design sparse FIR equalizers such that the performance loss does not exceed a predefined limit. In Sections 3.3.2 and 3.3.3, we investigate LEs and DFEs, respectively, for SISO systems while in Section 3.3.4, we consider DFEs for MIMO systems with $n_o \geq n_i$.

### 3.3.2 Sparse FIR SISO-LE

In the FIR LE model, the received samples $\{\mathbf{y}_k\}$ are applied to a fractionally-spaced FIR equalizer with $lN_f$ taps. The $k^{\text{th}}$ error sample is given by

$$e_k = x_{k-\Delta} - \mathbf{W}^H\mathbf{y}_{k:k-N_f+1} \tag{3.5}$$

where $\mathbf{W}$ is an $lN_f \times 1$ vector of stacked tap weights and $\Delta$ is an integer representing the decision delay where $0 \leq \Delta \leq N_f+\nu-1$. A widely-used criterion for evaluating the equalizer performance is the MMSE criterion. The decision-point mean square error (MSE) can be expressed as

$$\text{MSE} \equiv \xi \equiv E\left[|e_k|^2\right] = E\left[e_k e_k^H\right]. \tag{3.6}$$

Writing $x_{k-\Delta}$ as $x_{k-\Delta} = \mathbf{1}_\Delta^H \mathbf{x}_{k:k-N_f-\nu+1}$ where $\mathbf{1}_\Delta^H = [\ \underbrace{0\ \ldots\ 0}_{\Delta}\quad 1\quad \underbrace{0\ \ldots\ 0}_{N_f+\nu-\Delta-1}\ ]$ and substituting for $e_k$ from (3.5) into (3.6), we get

$$\xi = \varepsilon_x - \mathbf{W}^H \mathbf{r}_\Delta - \mathbf{r}_\Delta^H \mathbf{W} + \mathbf{W}^H \mathbf{R}_{yy} \mathbf{W} \tag{3.7}$$

where $\varepsilon_x \equiv E\left[|x_{k-\Delta}|^2\right]$ and $\mathbf{r}_\Delta = \mathbf{R}_{yx}\mathbf{1}_\Delta$. By defining the Cholesky factorization [73] of $\mathbf{R}_{yy}$ as $\mathbf{R}_{yy} = \mathbf{L}\mathbf{L}^H$ where $\mathbf{L}$ is an $lN_f \times lN_f$ lower-triangular matrix, we can rewrite (3.7) as follows

$$\xi = \varepsilon_x - \mathbf{W}^H \mathbf{L}\mathbf{L}^{-1} \mathbf{r}_\Delta - \mathbf{r}_\Delta^H \mathbf{L}^{-H}\mathbf{L}^H \mathbf{W} + \mathbf{W}^H \mathbf{L}\mathbf{L}^H \mathbf{W} \tag{3.8}$$

where $(.)^{-H} \equiv ((.)^H)^{-1}$. Completing the squares in (3.8) yields

$$\xi = \underbrace{\varepsilon_x - \mathbf{r}_\Delta^H \mathbf{L}^{-H}\mathbf{L}^{-1}\mathbf{r}_\Delta}_{\equiv \xi_{\min}} + \underbrace{\|\ \mathbf{L}^H\mathbf{W} - \mathbf{L}^{-1}\mathbf{r}_\Delta\ \|_2^2}_{\equiv \xi_{\text{excess}}} \tag{3.9}$$

Note that $\mathbf{W}$ controls $\xi$ via the term $\xi_{\text{excess}}$ only because $\xi_{\min}$ does not depend on $\mathbf{W}$. Since $\xi_{\text{excess}} \geq 0$, $\xi$ is minimized by choosing $\mathbf{W}$ such that $\xi_{\text{excess}} = 0$ and $\xi = \xi_{\min}$. Consequently, the optimum choice for $\mathbf{W}$ in the MMSE sense is

$$\mathbf{W}_{\text{opt}} = \mathbf{L}^{-H}\mathbf{L}^{-1}\mathbf{r}_\Delta = \mathbf{R}_{yy}^{-1}\mathbf{r}_\Delta \tag{3.10}$$

In general, $\mathbf{W}_{\text{opt}}$ is not sparse and, hence, the complexity of computing and implementing $\mathbf{W}_{\text{opt}}$ will increase proportional to $(lN_f)^2$ which can be prohibitively large. Any choice for $\mathbf{W}$ different from $\mathbf{W}_{\text{opt}}$ increases $\xi_{\text{excess}}$ which translates into performance degradation. A practical performance-complexity trade-off can be achieved if we design a sparse $\mathbf{W}$ such that $\xi_{\text{excess}} \leq \epsilon$ where $\epsilon$ controls the tolerable performance loss in terms of MSE increase. According to Section 3.2, this is achieved either by solving the convex optimization program in (3.11) or by calling the OMP algorithm in (3.12); i.e.,

$$\min_{\mathbf{W}_{\text{s}} \in \mathbb{C}^{lN_f}} \|\ \mathbf{W}_{\text{s}}\ \|_1 \quad \text{subject to} \quad \|\ \mathbf{L}^H\mathbf{W}_{\text{s}} - \mathbf{L}^{-1}\mathbf{r}_\Delta\ \|_2^2 \ \leq\ \epsilon \tag{3.11}$$

$$\mathbf{W}_{\mathrm{s}} = \mathrm{OMP}\left(\mathbf{L}^{-1}\mathbf{r}_{\Delta},\ \mathbf{L}^{H},\ \parallel \mathbf{L}^{H}\mathbf{W}_{\mathrm{s}} - \mathbf{L}^{-1}\mathbf{r}_{\Delta} \parallel_2^2 \ \leq \ \epsilon\right) \tag{3.12}$$

Since $\mathbf{L}$ is a lower-triangular matrix, the vector $\mathbf{L}^{-1}\mathbf{r}_{\Delta}$ is easily computed using the forward-substitution method [73]. Observe that the matrix $\mathbf{L}^{H}$ is a square matrix, so this problem is different from the CS setup as pointed out in Section 3.2. Also, note that $\mathbf{W}_{\mathrm{s}}$ is computed each time the channel estimate is updated. We define the decision-point signal-to-noise ratio as $\widehat{\mathrm{SNR}} \equiv \frac{\varepsilon_x}{\xi}$ and write $\widehat{\mathrm{SNR}}$ corresponding to $\mathbf{W}_{\mathrm{s}}$ as

$$\widehat{\mathrm{SNR}}(\mathbf{W}_{\mathrm{s}}) = \frac{\varepsilon_x}{\xi_{\min} + \xi_{\mathrm{excess}}(\mathbf{W}_{\mathrm{s}})} \geq \frac{\widehat{\mathrm{SNR}}(\mathbf{W}_{\mathrm{opt}})}{1 + \frac{\epsilon}{\xi_{\min}}} \tag{3.13}$$

where $\widehat{\mathrm{SNR}}(\mathbf{W}_{\mathrm{opt}}) \triangleq \frac{\varepsilon_x}{\xi_{\min}}$ is the $\widehat{\mathrm{SNR}}$ corresponding to $\mathbf{W}_{\mathrm{opt}}$. The performance loss is quantified by $\gamma$ where

$$\gamma \equiv \ell\left(\frac{\widehat{\mathrm{SNR}}(\mathbf{W}_{\mathrm{opt}})}{\widehat{\mathrm{SNR}}(\mathbf{W}_{\mathrm{s}})}\right) \leq \ell\left(1 + \frac{\epsilon}{\xi_{\min}}\right) \triangleq \gamma_{\max} \tag{3.14}$$

where $\ell(.) \equiv 10\,\mathrm{Log}_{10}(.)$ and $\mathrm{Log}_{10}$ is the base-10 logarithm. Furthermore, $\xi_{\min} = \varepsilon_x - \parallel \mathbf{L}^{-1}\mathbf{r}_{\Delta} \parallel_2^2$. To summarize, we compute $\epsilon$ based on the acceptable $\gamma_{\max}$ and compute the sparse solution $\mathbf{W}_{\mathrm{s}}$ through (3.11) or (3.12).

### 3.3.3 Sparse FIR SISO-DFE

For the SISO-DFE, we denote the spans of the FFF, $\mathbf{W}$, and the FBF, $\mathbf{b}$, by $lN_f$ and $N_b$ taps, respectively. The $k^{\mathrm{th}}$ error sample is defined as [74]

$$e_k = x_{k-\Delta} - \left(\mathbf{W}^{H}\mathbf{y}_{k:k-N_f+1} - \mathbf{b}^{H}\hat{\mathbf{x}}_{k-\Delta-1:k-\Delta-N_b}\right) \tag{3.15}$$

where $\hat{\mathbf{x}}_{k-\Delta-1:k-\Delta-N_b}$ is the slicer output vector representing the hard decisions for $\mathbf{x}_{k-\Delta-1:k-\Delta-N_b}$. Assuming correct past decisions, $\hat{\mathbf{x}}_{k-\Delta-1:k-\Delta-N_b}$ is replaced by

$\mathbf{x}_{k-\Delta-1:k-\Delta-N_b}$ and (3.15) becomes [75]

$$e_k = x_{k-\Delta} - \underbrace{\begin{bmatrix}\mathbf{W}^H & -\mathbf{b}^H\end{bmatrix}}_{\equiv\widetilde{\mathbf{W}}^H} \underbrace{\begin{bmatrix}\mathbf{y}_{k:k-N_f+1} \\ \mathbf{x}_{k-\Delta-1:k-\Delta-N_b}\end{bmatrix}}_{\equiv\tilde{\mathbf{y}}} \tag{3.16}$$

where $\widetilde{\mathbf{W}}$ is a length-$(lN_f + N_b)$ vector combining the FFF and the FBF taps. Hence, the MSE $\xi$ is given by

$$\xi = \varepsilon_x - \widetilde{\mathbf{W}}^H\tilde{\mathbf{r}}_\Delta - \tilde{\mathbf{r}}_\Delta^H\widetilde{\mathbf{W}} + \widetilde{\mathbf{W}}^H\mathbf{R}_{\tilde{y}\tilde{y}}\widetilde{\mathbf{W}} \tag{3.17}$$

where $\mathbf{R}_{\tilde{y}\tilde{y}} \equiv E\left[\tilde{\mathbf{y}}\tilde{\mathbf{y}}^H\right]$ and $\tilde{\mathbf{r}}_\Delta = \mathbf{R}_{\tilde{y}x}\mathbf{1}_\Delta$ where $\mathbf{R}_{\tilde{y}x} \equiv E\left[\tilde{\mathbf{y}}\,\mathbf{x}_{k:k-N_f-\nu+1}^H\right]$. Assuming $\mathbf{R}_{xx} = \varepsilon_x\mathbf{I}_{N_f+\nu}$, it can be shown that [75]

$$\mathbf{R}_{\tilde{y}\tilde{y}} = \begin{bmatrix}\mathbf{R}_{yy} & \varepsilon_x\mathbf{H}\mathbf{J}_\Delta \\ \varepsilon_x\mathbf{J}_\Delta^H\mathbf{H}^H & \varepsilon_x\mathbf{I}_{N_b}\end{bmatrix} \qquad \text{and} \qquad \mathbf{R}_{\tilde{y}x} = \begin{bmatrix}\varepsilon_x\mathbf{H} \\ \varepsilon_x\mathbf{J}_\Delta^H\end{bmatrix} \tag{3.18}$$

where $\mathbf{J}_\Delta$ is an $(N_f + \nu) \times N_b$ matrix whose structure for $\Delta \leq s$ is given by $\mathbf{J}_\Delta = \begin{bmatrix}\mathbf{0}_{N_b\times(\Delta+1)} & \mathbf{I}_{N_b} & \mathbf{0}_{N_b\times(s-\Delta)}\end{bmatrix}^H$ where $s \equiv N_f + \nu - N_b - 1$. As in the LE case, we define the Cholesky factorization of $\mathbf{R}_{\tilde{y}\tilde{y}}$ as $\mathbf{R}_{\tilde{y}\tilde{y}} = \widetilde{\mathbf{L}}\widetilde{\mathbf{L}}^H$ and write the MSE as follows

$$\xi = \underbrace{\varepsilon_x - \tilde{\mathbf{r}}_\Delta^H\widetilde{\mathbf{L}}^{-H}\widetilde{\mathbf{L}}^{-1}\tilde{\mathbf{r}}_\Delta}_{\equiv\xi_{\min}} + \underbrace{\parallel \widetilde{\mathbf{L}}^H\widetilde{\mathbf{W}} - \widetilde{\mathbf{L}}^{-1}\tilde{\mathbf{r}}_\Delta \parallel_2^2}_{\equiv\xi_{\text{excess}}}. \tag{3.19}$$

where $\widetilde{\mathbf{L}}$ depends on $\Delta$. Next, either the convex program in (3.20) or the OMP algorithm in (3.21) is used to compute the sparse solution $\widetilde{\mathbf{W}}_{\text{s}}$ such that $\xi_{\text{excess}}$ does not exceed $\epsilon$; i.e.,

$$\min_{\widetilde{\mathbf{W}}_{\text{s}}\in\mathbb{C}^{lN_f+N_b}} \parallel \widetilde{\mathbf{W}}_{\text{s}} \parallel_1 \quad \text{subject to} \quad \parallel \widetilde{\mathbf{L}}^H\widetilde{\mathbf{W}}_{\text{s}} - \widetilde{\mathbf{L}}^{-1}\tilde{\mathbf{r}}_\Delta \parallel_2^2 \;\leq\; \epsilon \tag{3.20}$$

$$\widetilde{\mathbf{W}}_{\text{s}} = \text{OMP}\left(\widetilde{\mathbf{L}}^{-1}\tilde{\mathbf{r}}_\Delta,\ \widetilde{\mathbf{L}}^H,\ \parallel \widetilde{\mathbf{L}}^H\widetilde{\mathbf{W}}_{\text{s}} - \widetilde{\mathbf{L}}^{-1}\tilde{\mathbf{r}}_\Delta \parallel_2^2 \;\leq\; \epsilon\right) \tag{3.21}$$

After computing $\widetilde{\mathbf{W}}_{\text{s}}$, its top $lN_f$ elements are assigned to the FFF and the remaining $N_b$ elements are assigned to the FBF. The quantities $\gamma$ and $\gamma_{\max}$ can also be defined for the DFE case as in the LE case.

### 3.3.4 Sparse FIR MIMO-DFE

For the FIR MIMO-DFE [16], the length-($lN_f n_o$) FFF $\mathbf{W}_i$ and the length-($N_b n_i$) FBF $\mathbf{b}_i$ are designed to recover the $i^{\text{th}}$ input stream, i.e., the transmitted stream from the $i^{\text{th}}$ input antenna. We assume that only previous decisions on other streams are available at the present time and that these decisions are correct. Then, the $k^{\text{th}}$ error sample for the $i^{\text{th}}$ input stream ($1 \leq i \leq n_i$) is given by [16]

$$e_{k,i} = \mathbf{x}_{k-\Delta}(i) - \underbrace{\begin{bmatrix}\mathbf{W}_i^H & -\mathbf{b}_i^H\end{bmatrix}}_{\equiv \widetilde{\mathbf{W}}_i^H} \underbrace{\begin{bmatrix}\mathbf{y}_{k:k-N_f+1} \\ \mathbf{x}_{k-\Delta-1:k-\Delta-N_b}\end{bmatrix}}_{\equiv \tilde{\mathbf{y}}} \tag{3.22}$$

where $\mathbf{x}_{k-\Delta}(i)$ is the transmitted symbol by the $i^{\text{th}}$ input antenna at time $(k-\Delta)$. Note that $\mathbf{W}_i$ and $\mathbf{b}_i$ are applied, respectively, to the received samples and the decision samples to remove inter-antenna interference. Similar to the SISO-DFE case, the MSE for the $i^{\text{th}}$ input stream is

$$\xi_i = \underbrace{\varepsilon_x - \tilde{\mathbf{r}}_{\Delta,i}^H \widetilde{\mathbf{L}}^{-H} \widetilde{\mathbf{L}}^{-1} \tilde{\mathbf{r}}_{\Delta,i}}_{\equiv \xi_{\min,i}} + \underbrace{\| \widetilde{\mathbf{L}}^H \widetilde{\mathbf{W}}_i - \widetilde{\mathbf{L}}^{-1} \tilde{\mathbf{r}}_{\Delta,i} \|_2^2}_{\equiv \xi_{\text{excess},i}} \tag{3.23}$$

where $\widetilde{\mathbf{L}}$ is defined as in the SISO-DFE case and $\tilde{\mathbf{r}}_{\Delta,i} = \mathbf{R}_{\tilde{y}x} \mathbf{1}_{\Delta,i}$ where

$$\mathbf{1}_{\Delta,i}^H = [\mathbf{0}_{1\times n_i\Delta} \quad \underbrace{0 \; .. \; 0}_{i-1} \; 1 \; \underbrace{0 \; .. \; 0}_{n_i - i} \quad \mathbf{0}_{1\times n_i(N_f+\nu-\Delta-1)}].$$

To compute a sparse solution $\widetilde{\mathbf{W}}_i^H$ such that $\xi_{\text{excess},i} \leq \epsilon_i$, we solve the $l_1$-norm minimization program in (3.24) or call the OMP algorithm in (3.25)

$$\min_{\widetilde{\mathbf{W}}_{\text{s},i} \in \mathbb{C}^{lN_f n_o + N_b n_i}} \| \widetilde{\mathbf{W}}_{\text{s},i} \|_1 \quad \text{subject to } \| \widetilde{\mathbf{L}}^H \widetilde{\mathbf{W}}_{\text{s},i} - \widetilde{\mathbf{L}}^{-1} \tilde{\mathbf{r}}_{\Delta,i} \|_2^2 \leq \epsilon_i \tag{3.24}$$

$$\widetilde{\mathbf{W}}_{\text{s},i} = \text{OMP}\left(\widetilde{\mathbf{L}}^{-1} \tilde{\mathbf{r}}_{\Delta,i},\ \widetilde{\mathbf{L}}^H,\ \| \widetilde{\mathbf{L}}^H \widetilde{\mathbf{W}}_{\text{s},i} - \widetilde{\mathbf{L}}^{-1} \tilde{\mathbf{r}}_{\Delta,i} \|_2^2 \leq \epsilon_i\right) \tag{3.25}$$

where $\epsilon_i$ controls the performance loss for the $i^{\text{th}}$ input antenna. If the input antennas represent different users with different required quality of service (QoS) levels, the system

designer can assign small (large) values of $\epsilon_i$ for users demanding high (low) QoS levels, respectively.

## 3.4 Sparse FIR Channel Shortening

For simplicity, we use the signal model in Section 3.3.1 with $n_i = n_o = 1$. However, our algorithm can be extended to MIMO systems as well based on the results in [22].

### 3.4.1 Algorithm Design

In FIR channel shortening [21], the goal is to design the length-$(lN_f)$ fractionally-spaced CSE, $\mathbf{W}$, and the length-$(N_f + \nu)$ TIR, $\mathbf{b}$, which minimize the mean square of the error signal, $e_k$. In other words, the CSE is to be designed such that the overall impulse response of the channel, $\mathbf{h}$, and the CSE, $\mathbf{W}$, best approximates a TIR with few, namely $(N_b + 1)$, taps. Then, the ML or MAP detectors are designed based on the new short TIR. Although the length of the TIR vector $\mathbf{b}$ is $(N_f+\nu)$, only $(N_b+1)$ of its taps are designed to be nonzero. The choice of $N_b$ represents a performance-complexity tradeoff. In [21], the nonzero $(N_b+1)$ taps were chosen to be *contiguous* and their location (delay) within the $(N_f + \nu)$-span of $\mathbf{b}$ was optimized. In this chapter, we relax the contiguousness constraint and allow the nonzero $(N_b + 1)$ taps to be anywhere within the $(N_f + \nu)$-span of the TIR to achieve better performance without increasing complexity. The $k^{\text{th}}$ sample of the channel shortening error sequence, $e_k$, is given by [21]

$$e_k = \mathbf{W}^H \mathbf{y}_{k:k-N_f+1} - \mathbf{b}^H \mathbf{x}_{k:k-N_f-\nu+1} \tag{3.26}$$

Hence, the channel shortening MSE is given by

$$\text{MSE} \triangleq \xi = E\left[|e_k|^2\right] = \mathbf{W}^H\mathbf{R}_{yy}\mathbf{W} - \mathbf{W}^H\mathbf{R}_{yx}\mathbf{b} - \mathbf{b}^H\mathbf{R}_{yx}^H\mathbf{W} + \mathbf{b}^H\mathbf{R}_{xx}\mathbf{b} \tag{3.27}$$

Minimizing the MSE over $\mathbf{W}$ by differentiating it with respect to (w.r.t) $\mathbf{W}$ and equating the result to zero, we get

$$\mathbf{W}_{\text{opt}} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{b} \tag{3.28}$$

Substituting $\mathbf{W}_{\text{opt}}$ for $\mathbf{W}$ in (3.27), we get

$$\xi = \mathbf{b}^H \underbrace{\left(\mathbf{R}_{xx} - \mathbf{R}_{yx}^H\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\right)}_{\triangleq \mathbf{R}_{x/y}^{\perp}} \mathbf{b} \tag{3.29}$$

Minimizing $\xi$ over $\mathbf{b}$ yields the trivial solution $\mathbf{b} = 0$. To avoid this, we perform the minimization subject to the UTC where one of the $(N_b + 1)$ taps is constrained to be unity. The index of the unit tap is denoted by $i_t$ $(0 \leq i_t \leq N_f + \nu - 1)$. Our goal is to design $\mathbf{b}$ which minimizes $\xi$ such that its nonzero taps are not constrained to be contiguous unlike [21]. Towards this end, we define the Cholesky factorization [73] of $\mathbf{R}_{x/y}^{\perp}$ as $\mathbf{R}_{x/y}^{\perp} \triangleq \mathbf{U}^H\mathbf{U}$ where $\mathbf{U}$ is an upper-triangular matrix and rewrite (3.29) as follows

$$\xi = \mathbf{b}^H\mathbf{U}^H\mathbf{U}\mathbf{b} = \| \mathbf{U}\mathbf{b} \|_2^2 = \| \bar{\mathbf{U}}\bar{\mathbf{b}} + \mathbf{u}_{i_t} \|_2^2 \tag{3.30}$$

where $\bar{\mathbf{U}}$ is formed by all the columns of $\mathbf{U}$ except for the $i_t^{\text{th}}$ column, $\mathbf{u}_{i_t}$ is the $i_t^{\text{th}}$ column of $\mathbf{U}$, and $\bar{\mathbf{b}}$ is formed by all the elements of $\mathbf{b}$ except for the $i_t^{\text{th}}$ unity element. Observe that the length-$(N_f + \nu - 1)$ $\bar{\mathbf{b}}$ contains only $N_b$ nonzero taps whose locations and values need to be determined. To achieve this goal, we formulate the following convex optimization problem

$$\min_{\bar{\mathbf{b}}} \| \bar{\mathbf{b}} \|_1 \quad \text{subject to} \quad \| \bar{\mathbf{U}}\bar{\mathbf{b}} + \mathbf{u}_{i_t} \|_2^2 \leq \epsilon_{\text{ch}} \tag{3.31}$$

where $\epsilon_{\text{ch}} > 0$ is a design parameter that controls the performance. The matrix $\bar{\mathbf{U}}$ is a tall matrix, so we have more measurements than unknowns unlike the CS theory setup. Using (3.31), the designer has no direct control over the resulting number of nonzero taps $N_b$. Hence, in situations where a specific $N_b$ is desired (e.g. due to complexity constraints), the OLS algorithm[2] is used instead of the above convex program as follows

$$\bar{\mathbf{b}} = \text{OLS}\left(-\mathbf{u}_{i_t},\ \bar{\mathbf{U}},\ \parallel \bar{\mathbf{b}} \parallel_0 =\ N_b\right) \tag{3.32}$$

where $\parallel \bar{\mathbf{b}} \parallel_0$ is equal to the number of iterations in the OLS algorithm. The constraint in (3.32) only specifies the number of OLS iterations. After computing $\bar{\mathbf{b}}$, we construct the sparse TIR, denoted by $\mathbf{b}_{\text{s}}$, by simply inserting the unit tap in the $i_t^{\text{th}}$ location. Finally, the optimum (in the MMSE sense) CSE taps and the resulting MMSE are, respectively, found to be

$$\mathbf{W}_{\text{opt}} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{b}_{\text{s}} \qquad \text{and} \qquad \xi_{\min} = \mathbf{b}_{\text{s}}^H\mathbf{R}_{xx}\mathbf{b}_{\text{s}} - \mathbf{b}_{\text{s}}^H\mathbf{R}_{yx}^H\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{b}_{\text{s}}. \tag{3.33}$$

Note that the unit tap index, $i_t$, needs to be optimized such that the resulting MSE is minimized. However, unlike [21], we do not need to additionally optimize the location (i.e., starting index which is commonly known as the decision delay parameter) of the contiguous taps because we do not constrain the taps to be contiguous. Instead, we use the OLS program in (3.32) to compute their locations and values.

### 3.4.2 Performance Analysis

Denoting the indices of the nonzero taps by $I$ and the vector containing the values of the nonzero TIR taps by $\mathbf{b}_I$, the MMSE tap values and the resulting MMSE subject to the UTC

[2]We found that OLS yields a significantly smaller residual error than OMP especially for small $N_f$

are given by [21]

$$\mathbf{b}_{I,\text{opt}} = \frac{\mathbf{R}_I^{-1}(:, i_t')}{\mathbf{R}_I^{-1}(i_t', i_t')} \qquad \text{and} \qquad \text{MMSE} = \frac{1}{\mathbf{R}_I^{-1}(i_t', i_t')} \tag{3.34}$$

where $i_t'$ is the unit tap index within $\mathbf{b}_I$ and $\mathbf{R}_I$ is a submatrix of $\mathbf{R}_{x/y}^{\perp}$ formed by the intersections of the rows and columns whose indices are those in $I$. We observe from (3.34) that the MMSE is controlled by the indices in $I$. If the indices in $I$ are not constrained to be contiguous, then the rows and columns forming $\mathbf{R}_I$ are also not constrained to be contiguous. Due to this extra degree of freedom, the noncontiguous solution enjoys a larger search space and, hence, the resulting MMSE will be less than (or at least the same as) the MMSE of the contiguous solution. The search operation is implemented by calling the OLS algorithm in (3.32).

### 3.4.3 Sparse FIR CSE

In general, the MMSE CSE in (3.33) is non-sparse, i.e., all the tap weights have nonzero values. Furthermore, long CSEs are needed to achieve good performance especially for highly-dispersive channels. However, the complexity of computing and implementing these long non-sparse FIR CSEs is high. Motivated by these considerations, we propose a sparse implementation for the FIR CSE, $\mathbf{W}$, as follows. After computing the sparse TIR $\mathbf{b}_\text{s}$, we write the MSE as a function of $\mathbf{W}$ as follows

$$\begin{aligned}\xi(\mathbf{W}) &= \mathbf{W}^H\mathbf{R}_{yy}\mathbf{W} - \mathbf{W}^H\underbrace{\mathbf{R}_{yx}\mathbf{b}_\text{s}}_{:=\mathbf{q}} - \mathbf{b}_\text{s}^H\mathbf{R}_{yx}^H\mathbf{W} + \mathbf{b}_\text{s}^H\mathbf{R}_{xx}\mathbf{b}_\text{s} \\ &= \underbrace{\mathbf{b}_\text{s}^H\mathbf{R}_{xx}\mathbf{b}_\text{s} - \mathbf{q}^H\mathbf{R}_{yy}^{-1}\mathbf{q}}_{=\xi_{\min}\text{ in (3.33)}} + \underbrace{\| \mathbf{L}^H\mathbf{W} - \mathbf{L}^{-1}\mathbf{q} \|_2^2}_{\equiv\xi_{\text{excess}}(\mathbf{W})}\end{aligned} \tag{3.35}$$

Setting $\xi_{\text{excess}}(\mathbf{W}) = 0$ yields the MMSE non-sparse solution for $\mathbf{W}$ in (3.33). However, we design a sparse $\mathbf{W}$ such that $\xi_{\text{excess}}(\mathbf{W}) \leq \epsilon_{\text{cse}}$ where $\epsilon_{\text{cse}} > 0$ controls the acceptable performance loss. This is achieved by either solving the convex optimization program in (3.36) or by calling the OMP algorithm in (3.37); i.e.,

$$\min_{\mathbf{W}_{\text{s}}} \| \mathbf{W}_{\text{s}} \|_1 \quad \text{subject to} \quad \| \mathbf{L}^H \mathbf{W}_{\text{s}} - \mathbf{L}^{-1}\mathbf{q} \|_2^2 \leq \epsilon_{\text{cse}} \tag{3.36}$$

$$\mathbf{W}_{\text{s}} = \text{OMP}\left(\mathbf{L}^{-1}\mathbf{q},\ \mathbf{L}^H,\ \| \mathbf{L}^H \mathbf{W}_{\text{s}} - \mathbf{L}^{-1}\mathbf{q} \|_2^2 \leq \epsilon_{\text{cse}}\right) \tag{3.37}$$

where $\mathbf{L}^{-1}\mathbf{q}$ is easily computed using the forward substitution method [73]. Note that the TIR and the CSE are computed each time the CIR estimate is updated. We define the resulting output (or decision-point) SNR when using $\mathbf{W}$ as $\text{SNR}_{\text{o}}(\mathbf{W}) \triangleq \frac{\varepsilon_x}{\xi(\mathbf{W})}$. The performance loss is quantified, as in (3.14), by

$$\gamma_{\text{cse}} \equiv \ell\left(\frac{\text{SNR}_{\text{o}}(\mathbf{W}_{\text{opt}})}{\text{SNR}_{\text{o}}(\mathbf{W}_{\text{s}})}\right) \leq \ell\left(1 + \frac{\epsilon_{\text{cse}}}{\xi_{\min}}\right) \triangleq \gamma_{\text{cse,max}} \tag{3.38}$$

Furthermore, $\xi_{\min} = \mathbf{b}_{\text{s}}^H \mathbf{R}_{xx} \mathbf{b}_{\text{s}} - \| \mathbf{L}^{-1}\mathbf{q} \|_2^2$. To summarize, $\epsilon_{\text{cse}}$ is computed based on the acceptable $\gamma_{\text{cse,max}}$ and the sparse solution $\mathbf{W}_{\text{s}}$ is computed through (3.36) or (3.37).

## 3.5 Partial Self FEXT Cancellation for Cellular Backhaul in Vectored DSL

### 3.5.1 Signal Model

We consider vectored VDSL [25] transmission with discrete multitone (DMT) modulation for the cellular backhaul application [76]. We assume that $L$ equal-length twisted-pair lines are used to transport traffic from a cellular base station to the backhaul network. The electromagnetic coupling between the vectored twisted-pair lines causes their signals to interfere with each other at the receiver side. This kind of interference is called self far-end

crosstalk (FEXT) which is the dominant impairment source in VDSL systems operating at high frequencies and using short loops. We assume all transmissions are synchronized and employ DMT transmission with a cyclic-prefix whose length equals or exceeds the maximum memory of the direct and crosstalk channels. Grouping the $L$ received signals over the $k^{\text{th}}$ tone, we cast the input-output relation as follows [25, 26]

$$\mathbf{Y}^k = \mathbf{H}^k \mathbf{X}^k + \mathbf{Z}^k \tag{3.39}$$

where $\mathbf{Y}^k$ and $\mathbf{X}^k$ group the $L$ received and transmitted symbols, respectively, on the $k^{\text{th}}$ tone. The $m^{\text{th}}$ elements of the vectors $\mathbf{Y}^k$ and $\mathbf{X}^k$, denoted by $Y_m^k$ and $X_m^k$, represent the received and transmitted symbols, respectively, on the $m^{\text{th}}$ line at the $k^{\text{th}}$ tone. Thermal noise, alien crosstalk, and other impairments experienced by the $L$ lines over the $k^{\text{th}}$ tone are lumped and grouped into $\mathbf{Z}^k$. Furthermore, $\mathbf{H}^k$ is the $L \times L$ MIMO channel matrix whose $(m, n)$ element, denoted by $H_{(m,n)}^k$, represents the frequency-domain response of the crosstalk channel from the $n^{\text{th}}$ line to the $m^{\text{th}}$ line at the $k^{\text{th}}$ tone. Also, $H_{(m,m)}^k$ represents the direct channel of the $m^{\text{th}}$ line at the $k^{\text{th}}$ tone.

### 3.5.2 Partial Self FEXT Cancellation

Self FEXT experienced by the $m^{\text{th}}$ line on the $k^{\text{th}}$ tone is linearly cancelled by applying the filter $\mathbf{W}_m^k$ such that the estimate of $X_m^k$ is given by $\widetilde{X}_m^k = \left(\mathbf{W}_m^k\right)^H \mathbf{Y}^k$. Dropping the tone index $k$, for simplicity, and rewriting $\widetilde{X}_m$ in terms of the residual self FEXT interference plus noise, denoted by $r_m$, we get

$$\widetilde{X}_m = X_m + \underbrace{\left( \left(\mathbf{W}_m^H \mathbf{h}_m - 1\right) X_m + \sum_{n \neq m} \mathbf{W}_m^H \mathbf{h}_n X_n \right) + \mathbf{W}_m^H \mathbf{Z}}_{\triangleq r_m} \tag{3.40}$$

where $\mathbf{h}_n$ is the $n^{\text{th}}$ column of $\mathbf{H}$. The residual self FEXT variance is given by

$$\zeta_m \triangleq E\left[|r_m|^2\right] = S_m \left|\mathbf{W}_m^H \mathbf{h}_m - 1\right|^2 + \sum_{n \neq m} S_n \left|\mathbf{W}_m^H \mathbf{h}_n\right|^2 + \mathbf{W}_m^H \mathbf{S}_{\mathbf{Z}} \mathbf{W}_m \tag{3.41}$$

where $\mathbf{S}_{\mathbf{Z}} \triangleq E\left[\mathbf{Z}\mathbf{Z}^H\right]$, and the lumped noise in $\mathbf{Z}$ is assumed to be uncorrelated with the data. We also assume that $E\left[\mathbf{X}\mathbf{X}^H\right] \triangleq \mathbf{S}_{\mathbf{X}} = \text{diag}\left(S_1, S_2, \ldots, S_L\right)$ where $\{S_m, 1 \leq m \leq L\}$ are real nonnegative numbers representing the input power spectral density of the $m^{\text{th}}$ line. It is not difficult to show that

$$\zeta_m = \mathbf{W}_m^H \underbrace{\left(\mathbf{H}\mathbf{S}_{\mathbf{X}}\mathbf{H}^H + \mathbf{S}_{\mathbf{Z}}\right)}_{\triangleq \mathbf{S}_{\mathbf{Y}}} \mathbf{W}_m - S_m \mathbf{W}_m^H \mathbf{h}_m - S_m \mathbf{h}_m^H \mathbf{W}_m + S_m \tag{3.42}$$

Defining $\mathbf{S}_{\mathbf{Y}} \triangleq \mathbf{L}_{\text{xt}} \mathbf{L}_{\text{xt}}^H$ where $\mathbf{L}_{\text{xt}}$ is an $L \times L$ lower-triangular matrix, it can be shown that

$$\zeta_m = \underbrace{S_m - \bar{\mathbf{h}}_m^H \mathbf{S}_{\mathbf{Y}}^{-1} \bar{\mathbf{h}}_m}_{\triangleq \zeta_{m,\min}} + \underbrace{\parallel \mathbf{L}_{\text{xt}}^H \mathbf{W}_m - \mathbf{L}_{\text{xt}}^{-1} \bar{\mathbf{h}}_m \parallel_2^2}_{\triangleq \zeta_{m,\text{excess}}(\mathbf{W}_m)} \tag{3.43}$$

where $\bar{\mathbf{h}}_m = S_m \mathbf{h}_m$. To reduce run-time complexity, we design a sparse self FEXT canceler $\widehat{\mathbf{W}}_m$ such that $\zeta_{m,\text{excess}}(\mathbf{W}_m) \leq \epsilon_{\text{xt}}$ where $\epsilon_{\text{xt}} > 0$ controls the acceptable data rate loss as will be shown shortly. Again, this is done by either solving the convex program in (3.44) or calling the OMP algorithm in (3.45); i.e.,

$$\min_{\widehat{\mathbf{W}}_m} \parallel \widehat{\mathbf{W}}_m \parallel_1 \quad \text{subject to } \parallel \mathbf{L}_{\text{xt}}^H \widehat{\mathbf{W}}_m - \mathbf{L}_{\text{xt}}^{-1} \bar{\mathbf{h}}_m \parallel_2^2 \ \leq \ \epsilon_{\text{xt}} \tag{3.44}$$

$$\widehat{\mathbf{W}}_m = \text{OMP}\left(\mathbf{L}_{\text{xt}}^{-1} \bar{\mathbf{h}}_m,\ \mathbf{L}_{\text{xt}}^H,\ \parallel \mathbf{L}_{\text{xt}}^H \widehat{\mathbf{W}}_m - \mathbf{L}_{\text{xt}}^{-1} \bar{\mathbf{h}}_m \parallel_2^2 \ \leq \ \epsilon_{\text{xt}}\right) \tag{3.45}$$

Instead, the constraint in (3.45) can be placed on the sparsity level (number of active taps) of the canceler if the system is constrained by a certain complexity level. Next, we derive an expression for the data rate loss and show how $\epsilon_{\text{xt}}$ can be used to control this loss. Denoting the tone width by $B$ Hz, the resulting data rate of the $m^{\text{th}}$ line at the $k^{\text{th}}$ tone using the

designed sparse self FEXT canceler is given from (3.40) by

$$R_m(\widehat{\mathbf{W}}_m) = B\mathrm{Log}_2\left(1+\frac{S_m/\Gamma}{\zeta_{m,\min}+\zeta_{m,\mathrm{excess}}(\widehat{\mathbf{W}}_m)}\right)$$
$$\simeq \underbrace{B\mathrm{Log}_2\left(\frac{S_m/\Gamma}{\zeta_{m,\min}}\right)}_{\triangleq R_{m,\max}} - \underbrace{B\mathrm{Log}_2\left(1+\frac{\zeta_{m,\mathrm{excess}}(\widehat{\mathbf{W}}_m)}{\zeta_{m,\min}}\right)}_{\triangleq R_{m,\mathrm{loss}}} \tag{3.46}$$

where the above approximation is valid since DSL systems are high-SNR environments and $R_{m,\mathrm{loss}}$ represents the data rate loss incurred by the sparse self FEXT canceler design. Furthermore, $\Gamma$ is the SNR gap defined as $\Gamma = \frac{\gamma_m}{3\gamma_c}\left(Q^{-1}(Pe/4)\right)^2$ where $\gamma_c$, $\gamma_m$, and $P_e$ denote the coding gain, noise margin, and the target bit error-rate, respectively [77]. The function $Q^{-1}(.)$ is the well-known inverse $Q$-function. Since $\zeta_{m,\mathrm{excess}}(\widehat{\mathbf{W}}_m) \leq \epsilon_{\mathrm{xt}}$ (as ensured by the design in (3.44)), the data rate loss is bounded as $R_{m,\mathrm{loss}} \leq B\mathrm{Log}_2\left(1+\epsilon_{\mathrm{xt}}/\zeta_{m,\min}\right)$. Hence, $\epsilon_{\mathrm{xt}}$ is chosen to control the acceptable data rate loss. Furthermore, it has been empirically shown that most of the self FEXT emanates from few neighboring pairs. Hence, the matrix $\bar{\mathbf{H}}$ is sparse and, thus, our proposed technique can effectively cancel most of the self FEXT using substantially fewer taps per tone than those of the ZF design which requires $L$ multiplications per line per tone per DMT symbol.

## 3.6 Simulation Results

In all the simulations, we use the greedy algorithms[3] described in Section 3.2 to obtain the sparse solutions because they, unlike the $l_1$-norm minimization approach, provide control over the number of nonzero filter coefficients.

[3]We use OMP for all problems except for finding the TIR taps where we use OLS

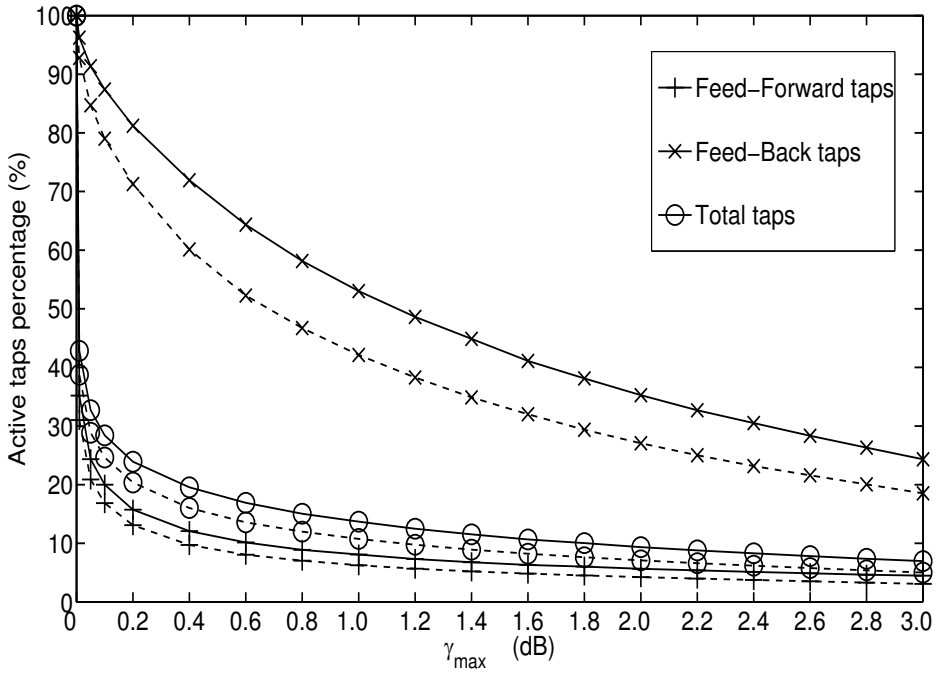


Figure 3.1. Active taps percentage versus $\gamma_{\max}$ for SISO-DFEs (dashed lines) and MIMO-DFEs (solid lines) both with $\mathrm{SNR_I} = 10$ dB for the ITU Vehicular-A channel model.

### 3.6.1 Sparse Equalization Results

We simulate the performance of our proposed sparse equalizer designs for various channel models as follows. For a given $\gamma_{\max}$, we compute the corresponding $\epsilon$ via (3.14), compute the sparse tap weights vector using the OMP algorithm, and calculate the active taps (having nonzero weights) percentage of the total filter span. For example, in SISO-DFEs, the FFF span is $lN_f$ and the FBF span is $N_b$. When the optimum MMSE solution is used, the number of the active filter taps equals the filter span. For LEs, a good choice for $\Delta$ is known to be $\Delta \approx \frac{N_f+\nu}{2}$ [75], while for DFEs, a good choice is $\Delta = N_f - 1$ when $N_f$ is sufficiently large [74]. These are the values used for $\Delta$ in this chapter. The noise is modeled as additive white Gaussian noise (AWGN) with one-sided power spectral density level denoted by $N_o$. We define the input SNR as $\mathrm{SNR_I} \equiv 10\mathrm{Log}_{10}\left(\frac{\varepsilon_x}{N_o}\right)$ and use $N_f = 35$ and $N_b = \nu$.

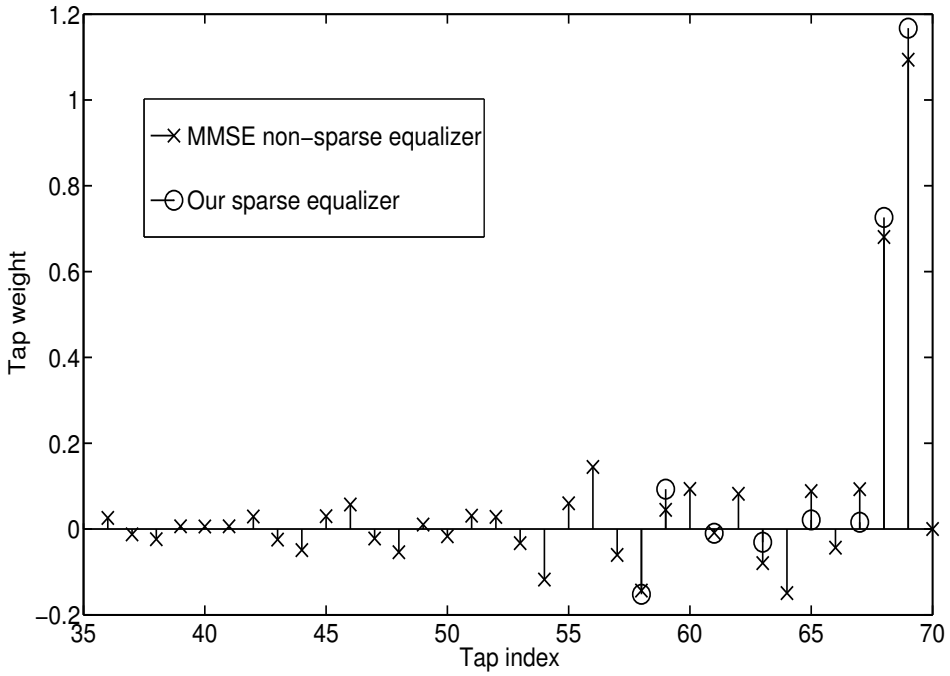


Figure 3.2. Single realizations of MMSE and sparse SISO-DFEs FFF impulse responses with $\mathrm{SNR_I} = 25$ dB and $\gamma_{\max} = 0.3$ dB for the ITU Vehicular-A channel. The equalizer taps for the indices below 35 are too insignificant to show.

In Figure 3.1, we show the performances of the sparse SISO-DFE and MIMO-DFE for the ITU Vehicular A channel [78] which has six nonzero paths and spans about eleven symbol durations. For the sparse SISO-DFE, we plot the active FFF taps percentage of $lN_f$, the active FBF percentage of $N_b$, and the total active taps percentage of $(lN_f + N_b)$ versus $\gamma_{\max}$ with $l = 2$ samples per symbol. The percentage savings in the FFF tap weights are more than those in the FBF tap weights which is a desirable feature for systems using binary or quadrature phase shift keying (BPSK or QPSK) modulation where the real and imaginary parts of the FBF input are either +1 or -1 thanks to the slicer preceding the FBF. Hence, the FBF operations are simple additions and subtractions, i.e., no multiplications, unlike the FFF whose input is continuous-valued. Since $lN_f > N_b$, the total active taps percentage is closer to the active FFF taps percentage than to the FBF taps percentage as shown in Figure

3.1. The performance of the sparse MIMO-DFE is also shown in Figure 3.1 where the active FFF, FBF, and total taps percentages are plotted versus $\gamma_{\max}$. We observe that allowing a maximum $\widehat{\text{SNR}}$ loss of 0.2dB reduces the total number of active taps by more than 70% for both SISO-DFEs and MIMO-DFEs! Our sparse FIR equalization approach becomes even more advantageous in the MIMO case where the total number of taps, for the non-sparse solution, increases with the products $n_i n_o$ and $n_i^2$. Comparing the active taps percentages of the SISO-DFE and the MIMO-DFE at the same $\gamma_{\max}$ reveals that the latter needs more taps to mitigate both ISI and inter-antenna interference. Figure 3.2 shows a single realization of both the MMSE solution and the sparse solution for SISO-DFEs. The performances of the sparse SISO-DFE and MIMO-DFE are shown in Figure 3.3 for non-sparse channel PDPs; namely, a 12-paths uniform PDP and an 8-paths exponentially decaying PDP with a decaying rate of 1 dB per path. On the same figure, we show the performance for the ITU Vehicular A channel for comparison. Inspecting Figure 3.3, we observe that as the number of significant paths in the channel PDP increases, more active taps are needed to equalize the channel.

Next, we investigate the bit-error-rate (BER) performance of our sparse FIR equalizers with uncoded BPSK modulation. Both SISO-DFE and MIMO-DFE cases are investigated in Figure 3.4 where we plot the BER versus $\text{SNR}_\text{I}$ for different values of $\gamma_{\max}$. The value $\gamma_{\max} = 0$dB represents the optimum MMSE non-sparse solution where all the equalizer taps are active. For the SISO-DFE case, our sparse design with $\gamma_{\max} = 0.2$ dB and 0.4 dB results in a maximum $\text{SNR}_\text{I}$ loss of 0.25dB and 0.5dB; however, the number of equalizer taps is reduced by 78.3% and 82.9%, respectively. Furthermore, as expected, the diversity gains of the MIMO-DFE over the SISO-DFE are observed at high values of $\text{SNR}_\text{I}$. In Figure 3.5, we

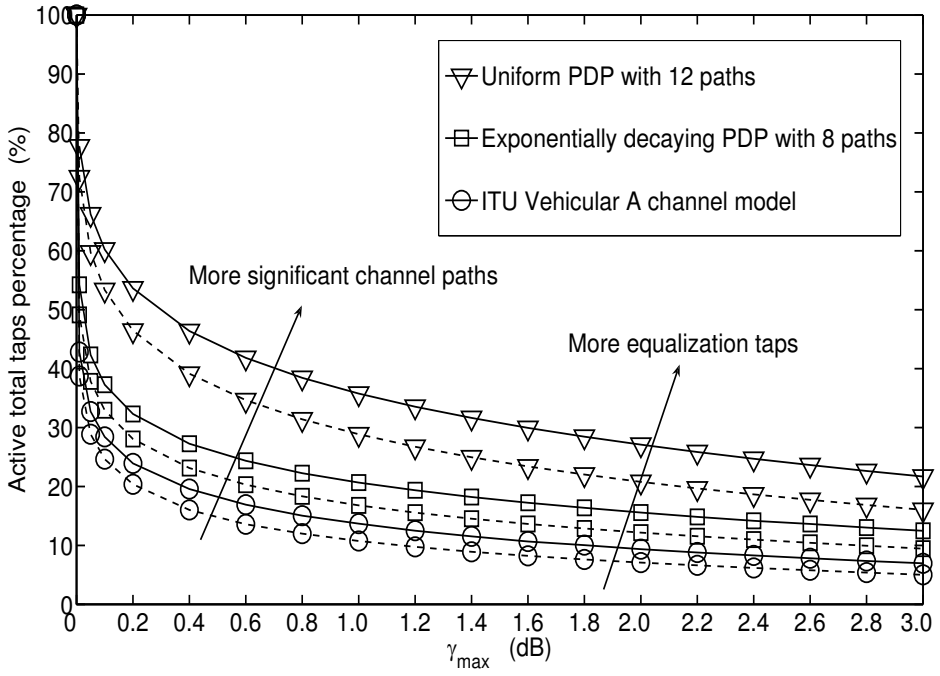


Figure 3.3. Active total taps percentage versus $\gamma_{\max}$ for SISO-DFEs (dashed lines) and MIMO-DFEs (solid lines) for sparse and non-sparse channels with $\mathrm{SNR_I} = 10$ dB.

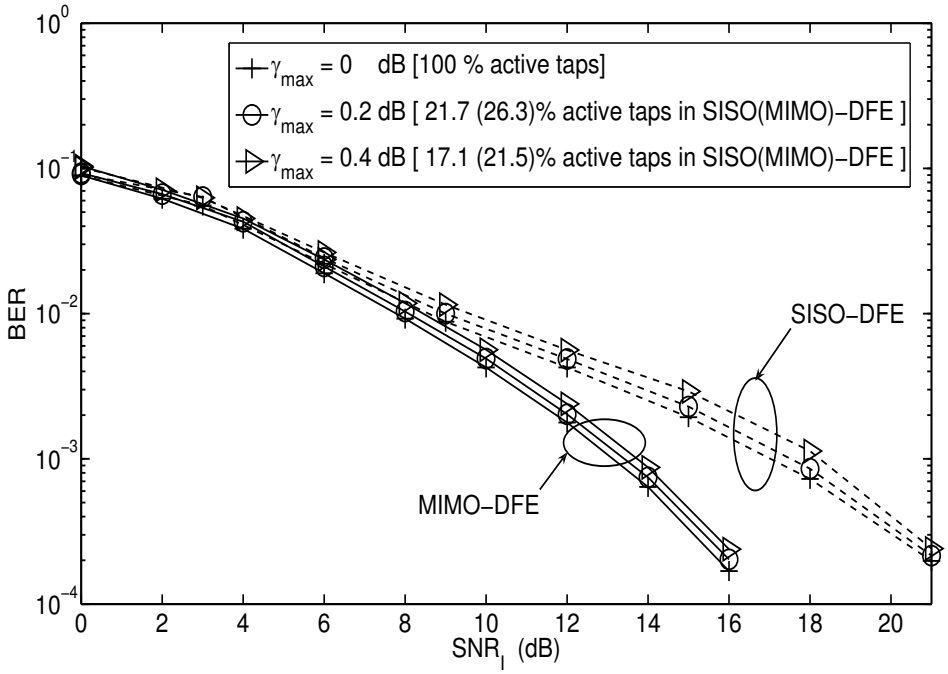


Figure 3.4. BER versus $\mathrm{SNR_I}$ for SISO-DFEs (dashed lines) with $l = 2$ and for MIMO-DFEs (solid lines) with $n_i = 2$, $n_o = 2$, and $l = 2$. BPSK modulation is used.

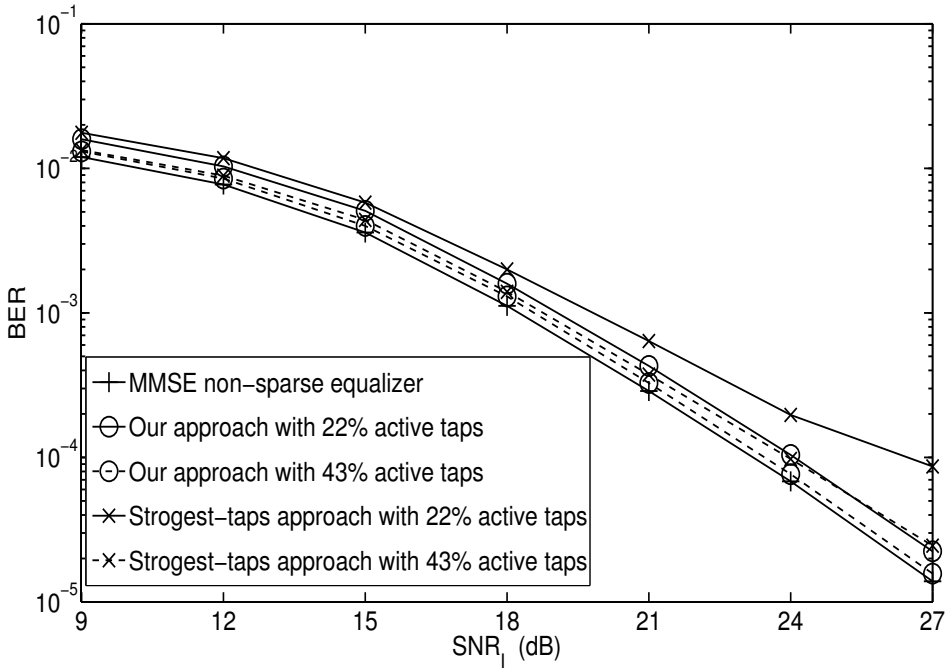


Figure 3.5. BER comparison between our sparse equalization approach and the strongest-taps approach for LEs with different sparsity levels (active taps percentages), $l = 2$, and BPSK modulation.

compare our sparse equalization approach with that in [13] where the whole MMSE equalizer tap vector is computed but only the $F$ strongest taps are selected. We refer to the approach in [13] as the "strongest-taps" approach. In addition to the performance superiority of our approach compared to the strongest-taps approach, the computational complexity of the latter is less than that of the former where the inversion of an $lN_f \times lN_f$ matrix is required to compute the whole MMSE solution as shown in (3.10). However, our OMP-based approach requires the inversion of an $F \times F$ matrix where $F << lN_f$ for all practical purposes.

### 3.6.2 Channel Shortening Results

We simulate the performance of our proposed sparse design techniques for sparse and non-sparse CIRs and denote the number of nonzero CIR taps by $L_{\text{ch}}$. The sparse CIR follows

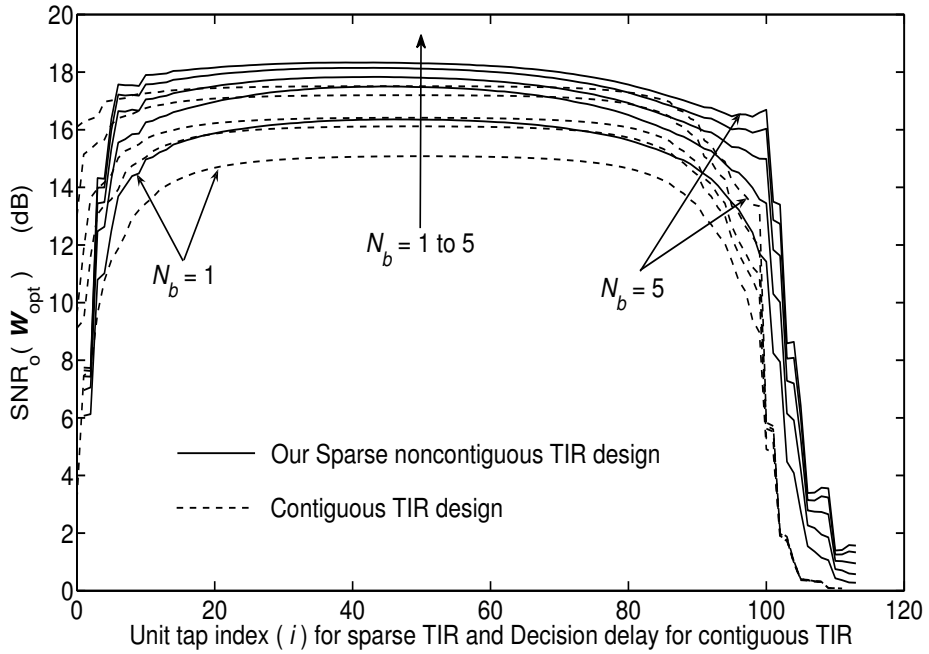


Figure 3.6. $\mathrm{SNR_o}(\mathbf{W}_{\mathrm{opt}})$ versus unit tap index and decision delay for sparse (solid lines) and contiguous (dashed lines) TIRs, respectively, for Vehicular A CIR with $N_f = 100$ and $\mathrm{SNR_I} = 20$ dB.

the ITU Vehicular A channel model and the non-sparse one has a uniform PDP whose taps are contiguous and have the same average power. We compare our sparse TIR design with the TIR design in [23] which we denote by the "strongest-paths" approach. Furthermore, we follow the noise model and the input SNR definition in Section 3.6.1.

Figure 3.6 depicts the variations of the output SNR with the unit tap index, $i_t$, and the decision delay for our sparse noncontiguous and contiguous TIR designs, respectively. It is clear that the unit tap index and the decision delay should be chosen carefully since suboptimum choices can degrade the performance significantly. However, we observe that a near optimum performance is achieved as long as these parameters are chosen in the vicinity of $(N_f + \nu)/2$. In the rest of the simulations, we optimize these parameters to compare the contiguous, strongest-paths, and our sparse noncontiguous TIR designs.

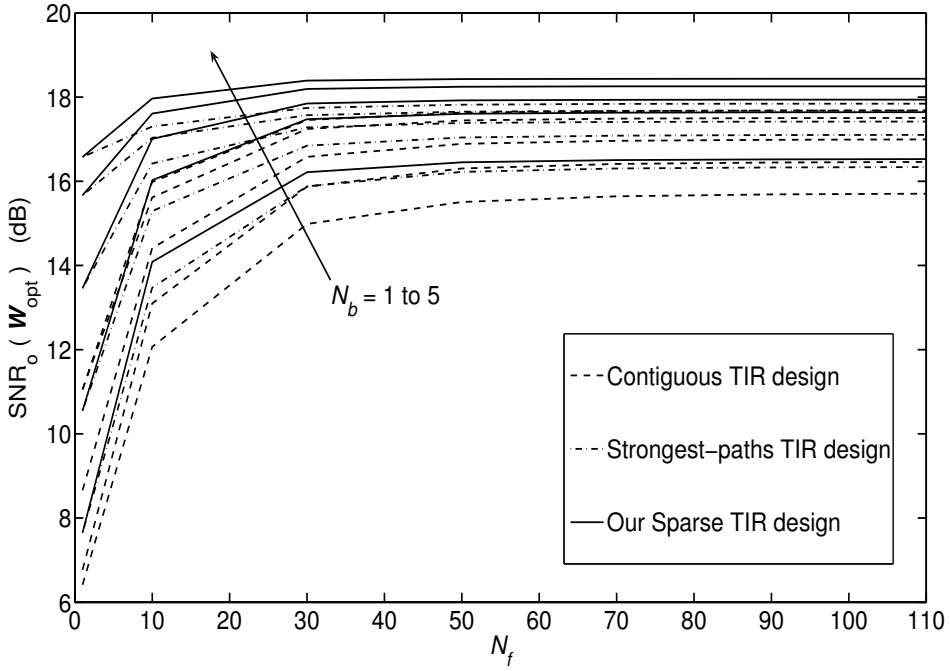


Figure 3.7. $\mathrm{SNR_o}(\mathbf{W}_{\mathrm{opt}})$ versus $N_f$ for three TIR designs for Vehicular A CIR with $\mathrm{SNR_I} = 20$ dB.

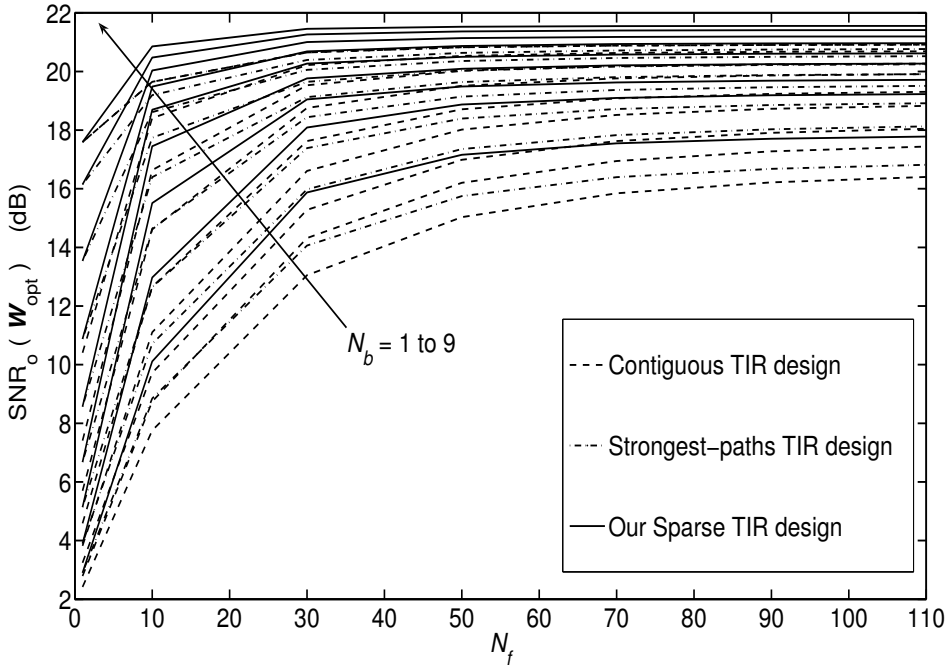


Figure 3.8. $\mathrm{SNR_o}(\mathbf{W}_{\mathrm{opt}})$ versus $N_f$ for three TIR designs for the 10-paths uniform PDP CIR with $\mathrm{SNR_I} = 20$ dB.

In Figures 3.7 and 3.8, we plot $\mathrm{SNR_o}(\mathbf{W}_{\mathrm{opt}})$ versus $N_f$ for the ITU Vehicular A channel and the 10-paths uniform PDP channel, respectively, and for $N_b$ ranging from 1 (lower curves) to $L_{\mathrm{ch}} - 1$ (upper curves). Recall that $(N_b + 1)$ represents the number of TIR taps. As $N_b$ increases, the output SNR increases for all TIR designs as expected which means that the TIR becomes more accurate in approximating the actual CIR. However, we observe that our sparse noncontiguous design of the TIR performs better than the other two designs for all choices of $N_f$ and $N_b$. From Figure 3.7, we observe that our sparse design with $N_b = 3$ outperforms both the contiguous MMSE and the strongest-paths designs with $N_b = 5$ over the whole range of $N_f$ and for $N_f \geq 30$, respectively. Therefore, allowing the TIR taps to be noncontiguous results in a better performance at the same complexity (measured by $N_b$) or in a lower complexity at the same performance. The complexity is measured by $N_b$ because the number of TIR trellis states (directly impacts the ML/MAP detectors complexity) increases exponentially with $N_b$. The number of states for the contiguous design is $2^{(N_b+1)}$. Although the number of trellis states for noncontiguous TIR designs is larger than that for the contiguous one due to the spread of the TIR taps, the algorithm in [79] can be employed to reduce the number of computations. Alternatively, the belief propagation (BP) algorithm can be used since its complexity is exponential only in $N_b$ regardless of the nonzero taps locations [23]. However, the BP algorithm suffers from short cycles which impact the accuracy of the messages independence assumption. In our simulations, we use the ML(BP) algorithm if the span of nonzero TIR taps is less(greater) than the maximum capacity of the simulation tool, respectively.[4] The superiority of our sparse TIR design over the contiguous and strongest-taps TIR designs is evident in Figure 3.9 where the former outperforms the

[4]BP is not used by itself to reduce the impact of short cycles.

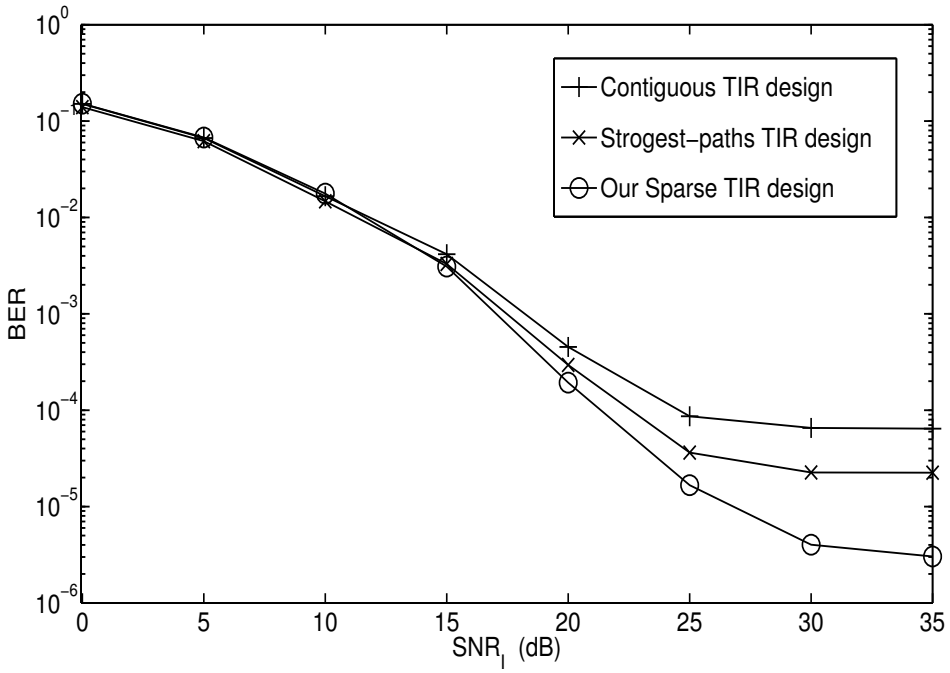


Figure 3.9. BER versus $\mathrm{SNR_I}$ for various TIR designs with MLSE detection for Vehicular A CIR with $N_f = 40$, $N_b = 3$, and BPSK modulation.

contiguous and strongest-taps designs by about 3.3 dB and 1.3 dB, respectively, at BER $\approx 10^{-4}$. Furthermore, both contiguous and strongest-taps designs reach their error floors 5 dB earlier than our sparse TIR design where the error floors in Figure 3.9 are due to the residual channel shortening MSEs. We also observe that our sparse TIR design reduces the error floor by about one order of magnitude compared to the strongest-paths TIR designs and by more than one order of magnitude compared to the contiguous TIR design.

Next, we study the effect of our sparse CSE and TIR designs on the performance in Figure 3.10 where we plot the active CSE taps (i.e those with nonzero weights) percentage of the total CSE span $lN_f$ versus $\gamma_{\max}$ which is the maximum loss in the output SNR. We observe that allowing a higher loss in the output SNR yields a bigger reduction in the number of CSE taps and that the active CSE taps percentage increases as $N_b$ decreases because the

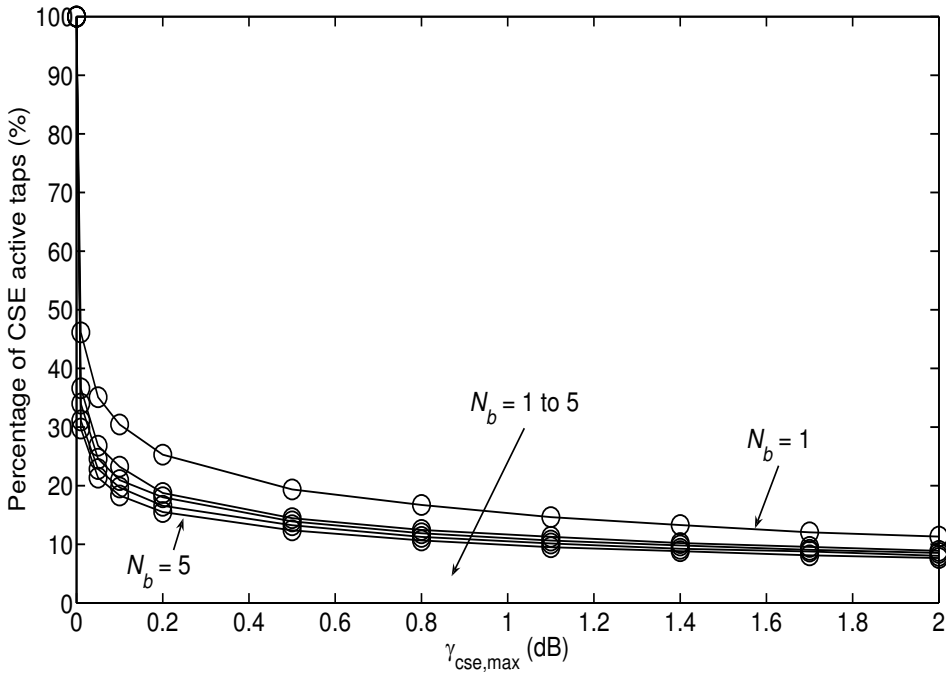


Figure 3.10. Active CSE taps percentage versus $\gamma_{\max}$ for our sparse CSE and TIR designs for Vehicular A CIR with $\mathrm{SNR_I} = 20$ dB and $N_f = 100$.

equalizer has to work harder (i.e., needs more taps) to shorten the CIR to a shorter TIR. We also observe that allowing a maximum of 0.2 dB loss in $\mathrm{SNR_o}$ with $N_b = 3$ results in about 82% reduction in the number of CSE active taps, i.e., the equalizer can shorten the channel using only 18 out of 100 taps.

### 3.6.3 Sparse Self FEXT Cancellation Results

The parameters of the simulated vectored VDSL system are listed in Table 3.1. We use the analytical model $\left|H^k_{(m,m)}\right|^2 = \exp\left(-1.158\frac{l_m\sqrt{f_k}}{l_o}\right)$ with linear phase for the direct transfer functions [80] where exp(.) is the exponential function, $f_k$ is the frequency of the $k^{\text{th}}$ tone in KHz, $l_m$ is the $m^{\text{th}}$ loop length in meters, and $l_o = 548.8$ meters is a reference length. For self FEXT, we use the model in [81] where $H^k_{(m,n)} = \left|H^k_{(n,n)}\right| \exp\left(j\phi(f_k)\right) 10^{-0.05\,\rho(f_k)}$ where

Table 3.1. Simulation Parameters

| | |
|---|---|
| Number of vectored lines ($L$) | 25 loops |
| Number of active DS tones | 1568 tones |
| Number of active US tones | 1209 tones |
| Subcarrier width ($B$) | 4.3125 KHz |
| Coding gain ($\gamma_c$) | 3 dB |
| Noise margin ($\gamma_m$) | 6 dB |
| Target bit error rate ($P_e$) | $10^{-7}$ |
| SNR gap ($\Gamma$) | 13 dB |
| Transmit power spectral density (PSD) | Flat at -60 dBm/Hz |
| Noise PSD | Flat at -140 dBm/Hz |
| Twisted pair length ($l_m$) | 1000 meters, $\forall m$ |
| Twisted pair type | 24-Gauge |

$j \equiv \sqrt{-1}$ and $\phi(f_k)$ is a random variable uniformly distributed in the interval $[0, 2\pi]$ radians which models the channel phase dispersion over the $k^{\text{th}}$ tone.

Furthermore, $\rho\,(f_k)$ is a Gaussian random variable expressed in dB with standard deviation $\sigma$ and mean $\mu = 2.33\sigma$ which models the spatial magnitude dispersion of the FEXT coupling channel over the $k^{\text{th}}$ tone. The value of $\sigma$ in dB specifies the spatial coupling dispersion, i.e., as $\sigma$ increases, the number of dominant crosstalkers decreases and $\bar{\mathbf{H}}$ becomes more sparse. Moreover, if $\sigma = \infty$, $\bar{\mathbf{H}}$ becomes diagonal, i.e., there is no self FEXT. No alien crosstalk is assumed since vectoring is performed over all $L$ pairs in the binder. Figure 3.11 shows the coupling dispersion variation with $\sigma$ where the y-axis represents the total self FEXT power induced by the $n_s$ strongest crosstalkers as a percentage of the total FEXT power and the x-axis represents $n_s$ where $1 \leq n_s \leq L-1$ and $L-1$ is the number of

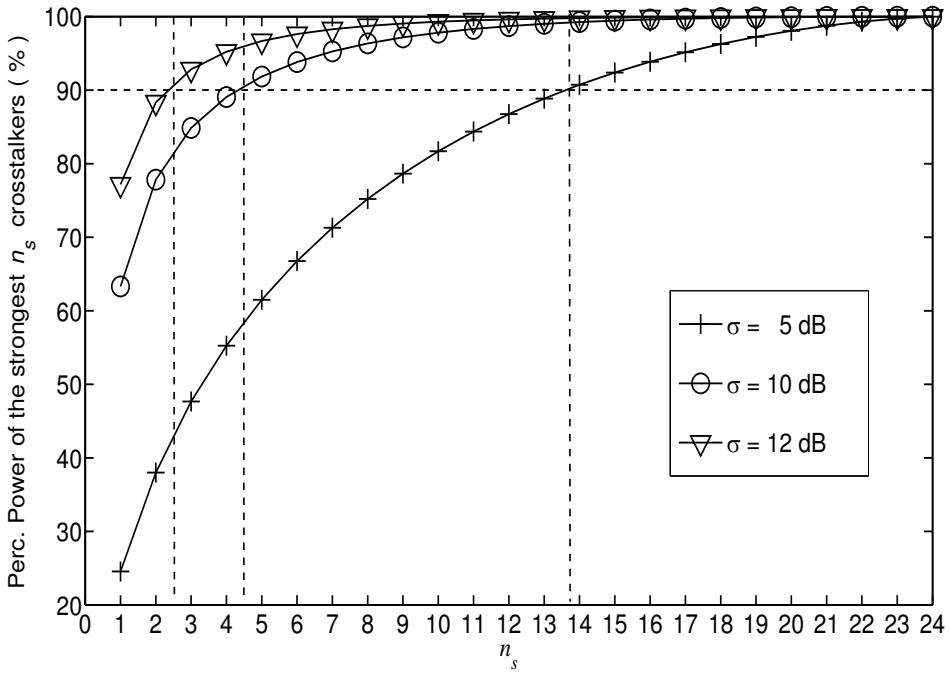


Figure 3.11. Total FEXT power from the strongest $n_s$ crosstalkers as a percentage of the total FEXT power versus $n_s$.

crosstalkers. When $\sigma = 10$ dB, about 90% of the FEXT power originates from just 4 or 5 lines which is typical in practice. Figure 3.12 shows the achievable downstream (DS) and upstream (US) data rates per line in Mbps versus the percentage active taps in $\widehat{\mathbf{W}}_m$ for various values of $\sigma$. For simulations, we use the frequency plan and the tone assignment in [82]. For practical values of $\sigma$ such as $\sigma = 10$ dB, we achieve 50% reduction in run-time complexity at a DS data rate loss of only 0.59 Mbps per line. The achievable data rates of all lines will be the same because they are all assumed to have the same length.

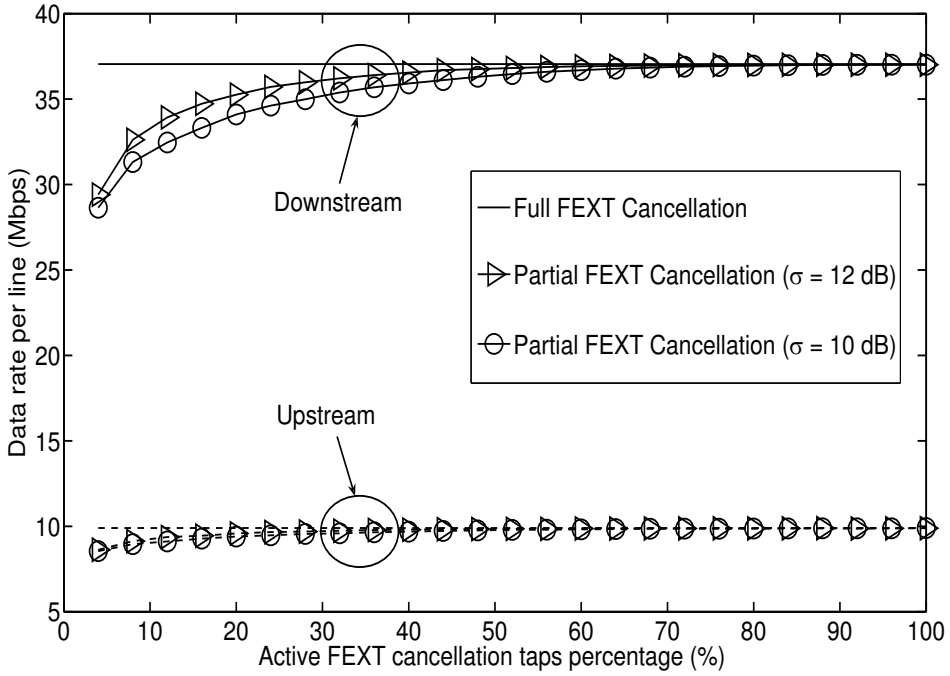


Figure 3.12. The achievable downstream (solid lines) and upstream (dashed lines) data rates per line versus the percentage active FEXT cancellation taps for different values of $\sigma$.

# CHAPTER 4

# PHASE NOISE IN SC-FDMA SYSTEMS: ANALYSIS AND COMPENSATION

## 4.1 Introduction

Carrier frequency offset (CFO) and phase noise (PN) are major oscillator impairments in direct-conversion transceivers, and single-carrier frequency-division-multiple access (SC-FDMA) is the uplink transmission scheme in the long-term evolution (LTE) standard. In this chapter, we derive a new analytical expression for the error vector magnitude (EVM) in asynchronous SC-FDMA systems under CFO and joint transmit-receive PN. The derived EVM expression reveals an interesting cross-layer relationship between the subcarrier mapping scheme at the medium-access control (MAC) layer and the immunity to CFO and PN at the physical layer. Furthermore, we propose an iterative reduced-complexity joint decoding and PN compensation scheme which does not require any pilots in PN tracking and exploits the low-pass nature of the PN process without assuming a specific PN model. Simulation results show the effectiveness of our proposed digital baseband compensation scheme in PN mitigation with only three iterations.

The rest of this chapter is organized as follows. The system model is described in Section 4.2 and the EVM expression is presented in Section 4.3 with the detailed derivation relegated to the Appendix. In Section 4.4, we simplify the derived EVM expression and analytically compare the performances of localized and distributed subcarrier mapping

schemes under CFO and PN. Our proposed iterative joint data decoding and PN compensation scheme is described in Section 4.5. Simulation results are presented in Section 4.6. The contents of this chapter are adapted and reprinted from [83][1] and [84][2].

*Notations*: Lower and upper-case bold letters denote vectors and matrices, respectively, and $\mathbf{A}(m,n)$ denotes the element in the $m$-th row and $n$-th column of $\mathbf{A}$. The $j$-th element of $\mathbf{a}$ is denoted by $a(j)$, and $\mathbf{a}(j:k)$ denotes the portion of $\mathbf{a}$ starting at index $j$ and ending at index $k \geq j$. The matrices $\mathbf{I}$ and $\mathbf{F}$ denote, respectively, the identity and Fast Fourier Transform (FFT) matrices, and their subscripts denote theirs sizes. The matrix $\mathbf{0}$ is the all-zero matrix. For matrices, the notation $\bar{\mathbf{A}}$ denotes $\mathbf{F}\mathbf{A}\mathbf{F}^H$. Also, $(\ )^H$, $(\ )^*$, and $(\ )^T$ denote the matrix Hermitian, complex conjugate, and transpose operations, respectively, while $(\ )^{-1}$ denotes both the scalar reciprocal and the matrix inverse operations. The operators $\mathbb{E}[\ ]$, $|\ |$, $\lceil . \rceil$, and $\lfloor . \rfloor$ denote, respectively, the statistical expectation, absolute, ceiling and floor functions. Also, $\mathrm{diag}(x_0, \ldots, x_{N-1})$ denotes a diagonal matrix whose diagonal entries are given by the argument; while $\mathrm{diag}(\mathbf{A})$ denotes a column vector containing the diagonal entries of $\mathbf{A}$.

## 4.2 System Model

### 4.2.1 Signal Model

We consider the uplink SC-FDMA transmission scenario with one receive antenna at the Base Station (BST) and $N_u$ single-antenna users. The total number of subcarriers is $N$

[1]Copyright [2011] IEEE. Reprinted, with permission, from A. Gomaa and N. Al-Dhahir, SC-FDMA Performance in Presence of Oscillator Impairments: EVM and Subcarrier Mapping Impact, IEEE GLOBECOM, 2011

[2]Copyright [2012] IEEE. Reprinted, with permission, from A. Gomaa and N. Al-Dhahir, Blind Phase Noise Compensation for SCFDMA with Application to LTE-Uplink, IEEE ICC, 2012

where each user is exclusively assigned $M < N$ subcarriers such that each subcarrier is not assigned to more than one user in the same time slot. The subcarrier mapping of the $l$-th user is defined by the $N \times M$ matrix $\mathbf{S}_l$ whose elements are zeros except for a single '1' in each column. The indices of the rows containing these ones are denoted by the set $J_l$ and correspond to the locations of the assigned subcarriers, and the cardinality of $J_l$ is $M$, $\forall\, l$. The $t$-th SC-FDMA time-domain (TD) symbol generated by the $l$-th user is given by

$$\mathbf{g}_l^t = \mathbf{F}_N^H \mathbf{S}_l \mathbf{F}_M \mathbf{x}_l^t \tag{4.1}$$

where $\mathbf{x}_l^t$ is the $M \times 1$ data vector generated by the $l$-th user in the $t$-th SC-FDMA symbol with $\mathbb{E}\left[\mathbf{x}_l \mathbf{x}_l^H\right] = \eta_l \mathbf{I}_M$ and $\eta_l$ is the received power from the $l$-th user. Then, a cyclic-prefix

Table 4.1. Key Parameters/Variables

| Parameters | Meaning | Dimens. |
|---|---|---|
| $\mathbf{x}_l^t$ | Data vector generated by $l$-th user at $t$-th SC-FDMA symbol | $M \times 1$ |
| $\mathbf{y}^t$ | Received time-domain vector at $t$-th SC-FDMA symbol | $N \times 1$ |
| $\mathbf{S}_l$ | Subcarrier mapping matrix of $l^{\text{th}}$ user. $\mathbf{S}_k^H \mathbf{S}_k = \mathbf{I}_M$ and $\mathbf{S}_k \mathbf{S}_k^H \triangleq \mathsf{S}_k$ | $N \times M$ |
| $\mathbf{H}_l$ ($\bar{\mathbf{H}}_l$) | Circulant (Diagonal) channel matrix seen by $l$-th user in TD (FD) | $N \times N$ |
| $\mathbf{Q}_l$ ($\bar{\mathbf{Q}}_l$) | Diagonal (circulant) CFO matrix in time (frequency) domain | $N \times N$ |
| $\mathbf{P}_l$ ($\bar{\mathbf{P}}_l$) | Diagonal (circulant) transmit PN matrix of $l$-th user in TD (FD) | $N \times N$ |
| $\mathbf{P}_r$ ($\bar{\mathbf{P}}_r$) | Diagonal (circulant) receive PN matrix in TD (FD) | $N \times N$ |
| $\mathbf{R}_l$ ($\bar{\mathbf{R}}_l$) | Auto-correlation matrix of the CFR of $l$-th user | $N \times N$ |
| $\mathbf{D}_k$ | Diagonal FD equalization matrix of $k$-th user. $\mathbb{E}\left[\mathbf{D}_k \mathbf{D}_k^H\right] \triangleq \mathsf{D}_k$ | $N \times N$ |
| $\gamma_l$ | SNR of the $l$-th user with $\left(\gamma_k \mathbf{R}_k(n,n)\right)^{-1} \triangleq p_k^n$ | $1 \times 1$ |
| $\alpha_l$ | Normalized CFO between $l$-th user and BST | $1 \times 1$ |
| $\beta_l$ ($\beta_r$) | Two-sided 3-dB linewidths of PSD of $l$-th user (BST) PN | $1 \times 1$ |

(CP) of length $L_p$ greater than or equal to the channel memory is appended to the vector $\mathbf{g}_l$ before transmission. The parameters used throughout the chapter are listed in Table 4.1.

### 4.2.2 PN and CFO Models

CFO results from inaccuracies of the crystal oscillators used at the transmitter and receiver to generate the carrier signals. These inaccuracies cause the carrier frequency to deviate from its nominal value. Furthermore, practical oscillators do not generate pure single-frequency signals (tones) and, hence, the power spectrum density (PSD) of the generated signal is not a Dirac-Delta function. Instead, the PSD typically follows the Lorentzian shape [85] causing energy leakage from each subcarrier into its neighbors resulting in inter-carrier interference (ICI). This impairment causes the oscillator phase to be noisy and, hence, the name *phase noise* where this noise follows the Wiener random process. In the presence of CFO and joint transmit-receive PN, the received TD signal is given by

$$\mathbf{y}^t = \mathbf{P}_r^t \sum_{l=1}^{N_u} \mathbf{Q}_l^t \mathbf{H}_l^t \mathbf{P}_l^t \mathbf{g}_l^t + \mathbf{z}^t \tag{4.2}$$

where $\mathbf{H}_l^t$ is the $N \times N$ circulant channel matrix whose first column contains the zero-padded channel impulse response (CIR) between the $l$-th user and the BST, and $\mathbf{z}^t$ is the complex additive white Gaussian noise (AWGN) vector with a single-sided PSD of $N_o$ Watt/Hz. The CFO between the $l$-th user and the BST is modeled by the $N \times N$ diagonal matrix

$$\mathbf{Q}_l^t \triangleq \text{diag}\left(\left\{\exp\left(\frac{j2\pi m\alpha_l}{N}\right)\right\}_{m=t(N+L_p)}^{t(N+L_p)+N-1}\right) \tag{4.3}$$

with $\alpha_l \triangleq \frac{\Delta f_l}{f_{\text{sub}}}$ where $\Delta f_l$ is the difference in Hz between the carrier frequencies of the $l$-th user and the BST and $f_{\text{sub}}$ denotes the subcarrier frequency spacing in Hz. The transmit

and receive PN processes are modeled [85] by the multiplication of the TD signals by the diagonal matrices $\mathbf{P}_l^t$ and $\mathbf{P}_r^t$, respectively, given by

$$\begin{aligned}\mathbf{P}_l^t &\triangleq \text{diag}\left(\exp\left(j\phi_0^{l,t}\right), \exp\left(j\phi_1^{l,t}\right), \ldots, \exp\left(j\phi_{N-1}^{l,t}\right)\right)\\ \mathbf{P}_r^t &\triangleq \text{diag}\left(\exp\left(j\phi_0^{r,t}\right), \exp\left(j\phi_1^{r,t}\right), \ldots, \exp\left(j\phi_{N-1}^{r,t}\right)\right)\end{aligned} \tag{4.4}$$

where $\phi_n^{l,t}$ and $\phi_n^{r,t}$ are the PN samples perturbing the transmitted and received signals, respectively, at the $n$-th sample of the $t$-th SC-FDMA symbol. The discrete-time PN model of free-running oscillators [86] follows the following first-order auto-regressive AR(1) processes[3]

$$\phi_n^{l,t} = \phi_{n-1}^{l,t} + \epsilon_n^{l,t}, \qquad \phi_n^{r,t} = \phi_{n-1}^{r,t} + \epsilon_n^{r,t} \tag{4.5}$$

where $\epsilon_n^{l,t} \sim \mathcal{N}(0, \sigma_l^2 \triangleq \frac{2\pi\beta_l}{Nf_{\text{sub}}})$ and $\epsilon_n^{r,t} \sim \mathcal{N}(0, \sigma_r^2 \triangleq \frac{2\pi\beta_r}{Nf_{\text{sub}}})$ are independent Gaussian-distributed random variables. Without loss of generality, we assume that $\phi_0^{l,0} = \phi_0^{r,0} = 0,\ \forall l$. The parameters $\beta_l$ and $\beta_r$ denote the two-sided 3-dB linewidths of the Lorentzian-shaped PSD of the oscillators feeding the $l$-th user and the BST, respectively. For simplicity, we drop the SC-FDMA symbol index, $t$, in Sections 4.3 and 4.4.2 where we analyze the performance under CFO and PN.

### 4.3 EVM Expression

Denoting by $k$ the index of the user of interest, we compute its decision statistic in four steps. First, we apply the $N$-point FFT to $\mathbf{y}^t$. Second, we select its assigned $M$ subcarriers by applying the selection matrix $\mathbf{S}_k^H$. Third, we apply frequency-domain (FD) equalization to mitigate the channel frequency-selectivity effects. Fourth, we apply the $M$-point inverse FFT to go back to TD for data detection. We do not use joint-subcarrier processing at the

[3]We assume this model to simplify the EVM expression derivation. However, we do not assume it for PN compensation in Section 4.5.

third step, i.e., the equalization matrix is diagonal; therefore, the second and third steps can be interchanged. Following this procedure, we write the decision statistic of the $k$-th user as follows

$$\hat{\mathbf{x}}_k = \mathbf{U}_{kk}\mathbf{x}_k + \sum_{\substack{l=1 \\ l\neq k}}^{N_u} \mathbf{U}_{kl}\mathbf{x}_l + \breve{\mathbf{z}} \tag{4.6}$$

with the three terms in (4.6) given by

$$\mathbf{U}_{kk} = \mathbf{F}_M^H \mathbf{S}_k^H \mathbf{D}_k \bar{\mathbf{P}}_r \bar{\mathbf{Q}}_k \bar{\mathbf{H}}_k \bar{\mathbf{P}}_k \mathbf{S}_k \mathbf{F}_M$$

$$\mathbf{U}_{kl} = \mathbf{F}_M^H \mathbf{S}_k^H \mathbf{D}_k \bar{\mathbf{P}}_r \bar{\mathbf{Q}}_l \bar{\mathbf{H}}_l \bar{\mathbf{P}}_l \mathbf{S}_l \mathbf{F}_M$$

$$\breve{\mathbf{z}} = \mathbf{F}_M^H \mathbf{S}_k^H \mathbf{D}_k \mathbf{F}_N \mathbf{z}$$

where $\mathbf{D}_k \equiv \text{diag}\left( \left\{ \frac{\bar{\mathbf{H}}_k^H(n,n)}{\left|\bar{\mathbf{H}}_k(n,n)\right|^2 + \gamma_k^{-1}} \right\}_{n=0}^{N-1} \right)$ is the diagonal minimum-mean squared-error (MMSE) FD equalization matrix of the $k$-th user with $\gamma_k \triangleq \frac{\eta_k}{N_o}$ being the signal-to-noise ratio (SNR) of the $k$-th user. As defined in Section 4.1, $\bar{\mathbf{P}}_r \equiv \mathbf{F}_N \mathbf{P}_r \mathbf{F}_N^H$, $\bar{\mathbf{Q}}_l \equiv \mathbf{F}_N \mathbf{Q}_l \mathbf{F}_N^H$, and $\bar{\mathbf{Q}}_k \equiv \mathbf{F}_N \mathbf{Q}_k \mathbf{F}_N^H$. Also, $\bar{\mathbf{H}}_k \equiv \mathbf{F}_N \mathbf{H}_k \mathbf{F}_N^H$ is the $N \times N$ diagonal FD channel matrix. Note that the diagonal and off-diagonal entries of the matrix $\mathbf{U}_{kk}$ correspond to the desired and ICI terms, respectively. The matrix $\mathbf{U}_{kl}$ corresponds to the inter-user interference (IUI) from the $l$-th user into the $k$-th user. The EVM of the $k$-th user is defined as follows

$$\text{EVM}_k \triangleq \sqrt{\frac{\mathbb{E}\left[\mathbf{e}^H\mathbf{e}\right]}{M\eta_k}} = \sqrt{\frac{\mathbb{E}\left[\text{Tr}\left(\mathbf{e}\mathbf{e}^H\right)\right]}{M\eta_k}} \times 100\ (\%) \tag{4.7}$$

where the error vector $\mathbf{e} \triangleq \hat{\mathbf{x}}_k - \mathbf{x}_k$ and Tr(.) denotes the matrix trace. Assuming that the data vectors of different users are independent with zero mean, that the user's data symbols are independent, and that the data and noise vectors are independent, we write

$$\mathbb{E}\left[\text{Tr}\left(\mathbf{e}\mathbf{e}^H\right)\right] = \eta_k \underbrace{\mathbb{E}\left[\text{Tr}\left(\mathbf{U}'_{kk}\left(\mathbf{U}'_{kk}\right)^H\right)\right]}_{\triangleq t_1} + \sum_{\substack{l=1 \\ l\neq k}}^{N_u} \eta_l \underbrace{\mathbb{E}\left[\text{Tr}\left(\mathbf{U}_{kl}\mathbf{U}_{kl}^H\right)\right]}_{\triangleq t_2} + \underbrace{\mathbb{E}\left[\text{Tr}\left(\breve{\mathbf{z}}\breve{\mathbf{z}}^H\right)\right]}_{\triangleq t_3} \tag{4.8}$$

where $\mathbf{U}'_{kk} \triangleq \mathbf{U}_{kk} - \mathbf{I}_M$. The first term, denoted by $t_1$, in (4.8) is evaluated as follows

$$t_1 = \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}'_{kk}\left(\mathbf{U}'_{kk}\right)^H\right)\right] = M - 2\mathrm{Re}(\underbrace{\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\right)\right]}_{\triangleq t_{1A}}) + \underbrace{\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\mathbf{U}_{kk}^H\right)\right]}_{\triangleq t_{1B}} \tag{4.9}$$

where Re(.) denotes the real part of its argument. To simplify the presentation, the derivations of the quantities $t_{1A}$, $t_{1B}$, $t_2$, and $t_3$ in (4.8) and (4.9) are relegated to the Appendix while their final expressions are given below. Starting with $t_{1A}$ in (4.9), we have

$$t_{1A} \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\right)\right] = \mathrm{Tr}\left(\mathbb{E}\left[\bar{\mathbf{P}}_r\right]\bar{\mathbf{Q}}_k\left(\mathbf{T}_k \circ \mathbb{E}\left[\bar{\mathbf{P}}_k\right]\right)\mathsf{S}_k\right) \tag{4.10}$$

where $\mathsf{S}_k \triangleq \mathbf{S}_k\mathbf{S}_k^H$ is a diagonal matrix, the notation $\circ$ denotes the element-wise Hadamard product, and

$$\mathbb{E}\left[\bar{\mathbf{P}}_{r/k}\right] = \mathbf{F}_N \mathrm{diag}\left(\left\{\exp\left(\frac{-m\sigma_{r/k}^2}{2}\right)\right\}_{m=0}^{N-1}\right)\mathbf{F}_N^H \tag{4.11}$$

where we used the fact that $\mathbb{E}\left[e^{jy}\right] = e^{-\sigma^2/2}$ [87] for $y \sim N(0,\sigma^2)$. The diagonal and off-diagonal entries of the matrix $\mathbf{T}_k$ are, respectively, given by

$$\mathbf{T}_k(m,n) = \frac{\mathbf{R}_k(m,n)}{\mathbf{R}_k(n,n)}\left(1 + p_k^n \exp\left(p_k^n\right)\mathrm{Ei}\left(-p_k^n\right)\right),\ \forall\, m,n \tag{4.12}$$

where $p_k^n \triangleq \left(\gamma_k \mathbf{R}_k(n,n)\right)^{-1}$, $\mathbf{R}_k(m,n) \triangleq \mathbb{E}\left[\bar{\mathbf{H}}_k(m,m)\bar{\mathbf{H}}_k^*(n,n)\right]$ is the FD channel auto-correlation matrix, and $\mathrm{Ei}(-p_k^n) \triangleq \int_{-\infty}^{-p_k^n} \frac{\exp(x)}{x}dx$ is the exponential integral function [88].

Next, the quantity $t_{1B}$ in Equation (4.9) is given by

$$t_{1B} \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\mathbf{U}_{kk}^H\right)\right] \simeq \frac{1}{N}\sum_{n\in J_k}\sum_{m=0}^{N-1}\mathbf{f}_{((m-n))_N}^H \mathbf{V}_r \mathbf{f}_{((m-n))_N}\mathbf{L}_k(m,n) \tag{4.13}$$

where $\mathbf{f}_{((m-n))_N}$ is the $((m-n) \bmod N)$-th column of $\mathbf{F}_N$ and

$$\mathbf{V}_r(m,n) = \exp\left(\frac{-|m-n|\,\sigma_r^2}{2} + \frac{j2\pi(m-n)\alpha_k}{N}\right),\ \forall\, m,n = 0,1,\ldots,N-1 \tag{4.14}$$

$$\mathbf{L}_k = \mathbf{F}_N\left(\left(\mathbf{F}_N^H \mathsf{S}_l \mathbf{F}_N\right) \circ \mathbf{V}_l\right)\mathbf{F}_N^H \mathbf{Y}_k \tag{4.15}$$

$$\mathbf{V}_l(m,n) = \exp\left(\frac{-|m-n|\,\sigma_l^2}{2}\right),\ \forall\, m,n = 0,1,\ldots,N-1 \tag{4.16}$$

$$\begin{aligned}\mathbf{Y}_k(m,n) = {} & \frac{|\mathbf{R}_k(m,n)|^2}{|\mathbf{R}_k(n,n)|^2}\left(1 + p_k^n + p_k^n(p_k^n+2)\exp(p_k^n)\mathrm{Ei}(-p_k^n)\right) \\ & - \frac{\mathbf{R}_k(m,m)\mathbf{R}_k(n,n) - |\mathbf{R}_k(m,n)|^2}{|\mathbf{R}_k(n,n)|^2}\left(\left(1 + (1+p_k^n)\exp\left(p_k^n\right)\mathrm{Ei}\left(-p_k^n\right)\right)\right),\ \forall\, m,n\end{aligned} \tag{4.17}$$

Next, the quantity $t_2$ in Equation (4.8) is given by

$$t_2 \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kl}\mathbf{U}_{kl}^H\right)\right] = \sum_{n\in J_k} \frac{-\mathbf{W}_r(n,n)}{\mathbf{R}_k(n,n)}\left(1 + (1+p_k^n)\exp\left(p_k^n\right)\mathrm{Ei}\left(-p_k^n\right)\right) \tag{4.18}$$

where the diagonal matrices $\mathbf{W}_r$ and $\mathbf{W}_l$ are, respectively, given by $\mathbf{W}_r = \mathbf{F}_N\left(\left(\mathbf{F}_N^H\left(\mathbf{W}_l \circ \mathbf{R}_l\right)\mathbf{F}_N\right) \circ \mathbf{V}_r\right)\mathbf{F}_N^H$ and $\mathbf{W}_l = \mathbf{F}_N\left(\left(\mathbf{F}_N^H \mathsf{S}_l \mathbf{F}_N\right) \circ \mathbf{V}_l\right)\mathbf{F}_N^H$. Finally, the quantity $t_3$ is given by

$$t_3 \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\breve{\mathbf{z}}\breve{\mathbf{z}}^H\right)\right] = \sum_{n\in J_k} \frac{-N_o}{\mathbf{R}_k(n,n)}\left(1 + (1+p_k^n)\exp\left(p_k^n\right)\mathrm{Ei}\left(-p_k^n\right)\right) \tag{4.19}$$

## 4.4 EVM Expression Simplification and Subcarrier Mapping Impact

### 4.4.1 EVM Expression Simplification

To gain insight into the EVM expression given in Section 4.3, we simplify it by considering the *high-SNR scenario with only transmit PN*[4], i.e., $\bar{\mathbf{P}}_r = \bar{\mathbf{Q}}_k = \mathbf{I}_N,\ \forall\, k$, and $\mathbf{D}_k\bar{\mathbf{H}}_k \simeq \mathbf{I}_N$. In uplink transmissions, receive PN is significantly smaller than transmit PN since the

[4]Receive PN effects can be modeled as transmit PN [41] by lumping their effects and summing their effective 3-dB linewidths.

BST oscillator quality is typically better than that of the user terminal. With the above assumptions, we re-write the EVM expression terms as follows

$$\begin{aligned} t_{1A} &= \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{H}}_k\bar{\mathbf{P}}_k\mathsf{S}_k\mathbf{D}_k\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{P}}_k\mathsf{S}_k\mathbf{D}_k\bar{\mathbf{H}}_k\right)\right] \simeq \mathrm{Tr}\left(\mathbb{E}\left[\bar{\mathbf{P}}_k\right]\mathsf{S}_k\right) \\ &= \frac{M}{N}\sum_{m=0}^{N-1}\exp\left(\frac{-m\sigma_k^2}{2}\right) \simeq M \end{aligned} \tag{4.20}$$

where the second equality follows from the identity $\mathrm{Tr}\left(\mathbf{AB}\right) = \mathrm{Tr}\left(\mathbf{BA}\right)$ and the approximation $\frac{1}{N}\sum_{m=0}^{N-1}\exp\left(\frac{-m\sigma_k^2}{2}\right) \simeq 1$ holds under practical values of $\sigma_k^2$ ranging from $10^{-6}$ to $10^{-3}$. Next, we re-write $t_{1B}$ as follows

$$\begin{aligned} t_{1B} &= \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{H}}_k\mathbf{W}_k\bar{\mathbf{H}}_k^H\mathbf{D}_k\mathbf{D}_k^H\mathsf{S}_k\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{W}_k\bar{\mathbf{H}}_k\bar{\mathbf{H}}_k^H\mathbf{D}_k\mathbf{D}_k^H\mathsf{S}_k\right)\right] \\ &\simeq \mathrm{Tr}\left(\mathbf{W}_k\mathsf{S}_k\right) = \sum_{n\in J_k}\mathbf{W}_k(n,n) \end{aligned} \tag{4.21}$$

where $\mathbf{W}_k$ is defined in (A.10) and the commutative property of diagonal matrices is used. In the Appendix, *we show that* $\mathbf{W}_k$ *can be approximated by a diagonal matrix.* Inspecting (A.10) and (4.16), we observe that the diagonal of $\mathbf{W}_k$ is the circular convolution between the diagonal of $\mathsf{S}_k$ and the FFT of the sequence $\left\{\exp\left(\frac{-|m|\sigma_k^2}{2}\right)\right\}_{m=-\lfloor N/2\rfloor}^{\lceil N/2\rceil-1}$ which is nothing but the Lorentzian spectrum. Furthermore, the diagonal of $\mathsf{S}_k$ is nonzero only at the subcarrier indices of the $k$-th user. We use the FFT property that element-wise multiplication in TD corresponds to circular convolution in FD. Recalling from Section 4.2.2 that the PN PSD is Lorentzian-shaped, we conclude that *the sequence* $\{\mathbf{W}_k(n,n)\}$ *in* (4.21) *is the circular convolution between the user's spectrum and the PN spectrum.* This circular convolution

quantifies the PN-induced intra-user interference. Similarly, we re-write $t_2$ as follows

$$\begin{aligned} t_2 &= \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{H}}_l \mathbf{W}_l \bar{\mathbf{H}}_l^H \mathsf{D}_k \mathsf{S}_k\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{W}_l \bar{\mathbf{H}}_l^H \bar{\mathbf{H}}_l \mathsf{D}_k \mathsf{S}_k\right)\right] \\ &\simeq \sum_{n\in J_k} \mathbf{W}_l(n,n)\mathsf{D}_k(n,n)\mathbb{E}\left[\left|\bar{\mathbf{H}}_l(n,n)\right|^2\right] \\ &= \sum_{n\in J_k} \mathbf{W}_l(n,n)\frac{-\mathbf{R}_l(n,n)}{\mathbf{R}_k(n,n)}\left(1+(1+p_k^n)\exp\left(p_k^n\right)\mathrm{Ei}\left(-p_k^n\right)\right) \\ &\simeq \sum_{n\in J_k} \mathbf{W}_l(n,n)\frac{-\mathbf{R}_l(n,n)}{\mathbf{R}_k(n,n)}\mathrm{Ei}\left(-p_k^n\right) \end{aligned} \tag{4.22}$$

where $\mathsf{D}_k(n,n)$ is given in (A.17) and the last approximation follows from the high-SNR assumption where $p_k^n \to 0^+$ and $\mathrm{Ei}\left(-p_k^n\right) \to -\infty$. If *the channels experienced by different users are identically distributed*, i.e., $\mathbf{R}_l = \mathbf{R}_k$, then

$$t_2 \simeq \sum_{n\in J_k} -\mathbf{W}_l(n,n)\mathrm{Ei}\left(-p_k^n\right) \tag{4.23}$$

Similar to $\mathbf{W}_k$, the sequence $\{\mathbf{W}_l(n,n)\}$ represents the circular convolution between the PN spectrum and the $l$-th user spectrum. Due to the non-impulsive nature of the PN spectrum, spectrum energy leakage of the $l$-th user causes inter-user interference. The quantity $t_2$ quantifies the $l$-th user's energy leakage at the $k$-th user's subcarriers. Applying the high-SNR approximation to $t_3$, we write the simplified EVM expression as follows

$$\mathrm{EVM}_k \simeq \sqrt{\underbrace{\frac{1}{M}\left(\sum_{n\in J_k}\mathbf{W}_k(n,n)\right)-1}_{\text{Intra-user Interference}} + \underbrace{\frac{1}{M}\sum_{l\neq k}\sum_{n\in J_k} w_l \mathbf{W}_l(n,n)}_{\text{Inter-user interference}} + \underbrace{\frac{1}{M}\sum_{n\in J_k} -p_k^n \mathrm{Ei}\left(-p_k^n\right)}_{\text{Noise contribution}}}, \tag{4.24}$$

where $w_l = -\frac{\eta_l \mathbf{R}_l(n,n)}{\eta_k \mathbf{R}_k(n,n)}\mathrm{Ei}\left(-p_k^n\right)$. Furthermore, $\mathrm{Ei}\left(-p_k^n\right)$ can be approximated by $\mathrm{Ei}\left(-p_k^n\right) \simeq 0.577 + \mathrm{Log}\left(p_k^n\right) - p_k^n$ for $p_k^n \leq 0.3$ or, equivalently, for $\gamma_k \geq 5$ dB where Log(.) is the natural logarithm function.

### 4.4.2 Localized versus Distributed Subcarrier Mapping

In light of the simplified EVM expression in (4.24), we investigate the effect of the user's subcarrier mapping at the MAC layer on its immunity to CFO and PN at the physical layer. We compare the localized and distributed subcarrier mapping schemes shown in Figure 4.1. We assume that only the subcarrier $n_k$ is assigned to the $k$-th user. Then, we consider

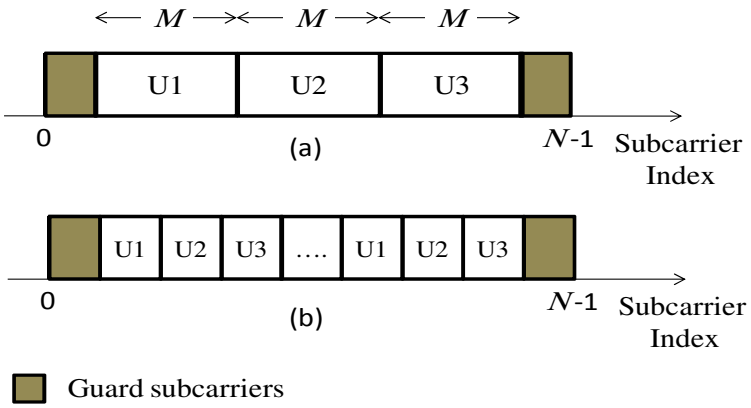


Figure 4.1. Subcarriers Mapping Schemes (a) Localized (b) Distributed

two scenarios where the neighboring subcarrier, $n_k + 1$, is assigned to the same user in one scenario and to the $l$-th user (where $l \neq k$) in the second scenario. Clearly, the first and second scenarios represent, respectively, the localized and distributed subcarrier mapping schemes [47]. In the presence of PN and CFO, the energy of the $(n_k + 1)$-th subcarrier leaks into its neighbor, the $(n_k)$-th subcarrier, causing intra and inter-user interference in the first (localized) and second (distributed) scenarios, respectively. The quantities $t_{1B}$ and $t_2$ were shown in Section 4.4.1 to quantify the contributions of the intra and inter-user interference, respectively, to the EVM expression. To compare the localized and distributed schemes, we compare the intra and inter user interference on the $k$-th user and re-write the quantities

$t_{1B}$ and $t_2$ in (4.21) and (4.23), respectively, with $J_k$ containing only $n_k$ as follows

$$t_{1B} = \mathbf{W}_k(n_k, n_k), \qquad t_2 = -\mathbf{W}_k(n_k, n_k)\mathrm{Ei}\left(-p_k^{n_k}\right) \tag{4.25}$$

where we assume that *both users have the same channel auto-correlation function and the same PN parameters*, i.e., $\mathbf{R}_l = \mathbf{R}_k$ and $\mathbf{W}_l = \mathbf{W}_k$. Inspecting (4.9), we observe that the diagonal entries of $\mathbf{U}_{kk}\mathbf{U}_{kk}^H$ are represented in $t_{1B}$ in addition to the off-diagonal entries accounting for the intra-user interference. For $t_{1B}$ to account only for the intra-user interference, we set the $n_k$-th diagonal entry in $\mathsf{S}_k$ to zero yielding $\mathsf{S}_k = \mathsf{S}_l$ where only the $(n_k+1)$-th diagonal entry is nonzero and, hence, $\mathbf{W}_l = \mathbf{W}_k$. Furthermore, the high-SNR assumption implies that $p_k^{n_k} \rightarrow 0^+$, $-\mathrm{Ei}\left(-p_k^{n_k}\right) >> 1$ and, hence, $t_2 >> t_{1B}$. Based on the above discussion and assuming that the received power levels from all users are the same, i.e., no near-far effects, we conclude that, for the EVM performance metric, inter-user interference is more harmful than intra-user interference. Therefore, **the localized subcarrier mapping is more immune to PN and CFO than the distributed subcarrier mapping**. In Section 4.6, we illustrate this result numerically.

## 4.5 Iterative Reduced-Complexity Joint Channel Decoding and PN Compensation

The CFO-induced phase deviation is linearly structured as shown in Equation (4.3); hence, CFO compensation is easy once the slope (which is proportional to the CFO value) is estimated. Several efficient techniques have been proposed in the literature to estimate and compensate CFO using training symbols. However, the PN-induced phase deviation varies randomly from sample to sample and is not parameterized by a single parameter as in the CFO case making PN compensation more challenging than CFO compensation. Hence, in

this section, we assume no CFO and focus on PN compensation. Since no pilots are inserted in the data SC-FDMA symbols to keep the PAPR low, we use the decoded symbols as pilots for PN estimation and compensation. Signal detection and channel equalization are first described in Section 4.5.1. Next, our proposed algorithm for iterative joint decoding and PN compensation is presented in Section 4.5.2 and its initialization is investigated in Section 4.5.3.

### 4.5.1 Reduced-Complexity Signal Detection

The *PN low-pass nature* makes commuting the matrices $\mathbf{H}_l^t$ and $\mathbf{P}_r^t$ an accurate approximation [41, 89] assuming that the channel coherence bandwidth is greater than $\beta_r$. Hence, we re-write Equation (4.2) without CFO as follows

$$\mathbf{y}^t \simeq \sum_{l=1}^{N_u} \mathbf{H}_l^t \mathbf{P}_r^t \mathbf{P}_l^t \mathbf{g}_l^t + \mathbf{z}^t \triangleq \sum_{l=1}^{N_u} \mathbf{H}_l^t \mathbf{P}_{\mathrm{C},l}^t \mathbf{g}_l^t + \mathbf{z}^t \tag{4.26}$$

where $\mathbf{P}_{\mathrm{C},l}^t \triangleq \mathbf{P}_r^t \mathbf{P}_l^t$. The transmit and receive PN effects are lumped together and treated as transmit PN only whose effective 3-dB linewidth is $\beta_{\mathrm{C},l} = \beta_l + \beta_r$. Denoting by $k$ the index of the user of interest, we obtain its decision statistic by applying the $N$-point FFT to $\mathbf{y}^t$ and selecting its assigned $M$ subcarriers by applying the matrix $\mathbf{S}_k^H$. Hence, we write the decision statistic of the $k$-th user as follows

$$\mathbf{Y}_k^t = \mathbf{S}_k^H \mathbf{F}_N \mathbf{y}^t = \mathbf{S}_k^H \bar{\mathbf{H}}_k^t \bar{\mathbf{P}}_{\mathrm{C},k}^t \mathbf{S}_k \mathbf{F}_M \mathbf{x}_k^t + \mathbf{Z}_k^t \tag{4.27}$$

where $\mathbf{Z}_k^t = \mathbf{S}_k^H \mathbf{F}_N \mathbf{z}^t + \mathbf{S}_k^H \mathbf{F}_N \sum_{l \neq k} \mathbf{H}_l^t \mathbf{P}_{\mathrm{C},l}^t \mathbf{g}_l^t$ represents the noise plus the inter-user interference (IUI) caused by the PN-induced leakage between users' subcarriers. Furthermore, $\bar{\mathbf{H}}_k^t = \mathbf{F}_N \mathbf{H}_k^t \mathbf{F}_N^H$ is the $N \times N$ diagonal FD channel matrix and $\bar{\mathbf{P}}_{\mathrm{C},k}^t = \mathbf{F}_N \mathbf{P}_{\mathrm{C},k}^t \mathbf{F}_N^H$ is the $N \times N$ circulant PN matrix. It is worth mentioning that the ICI is induced by the matrix

$\bar{\mathbf{P}}^t_{\mathrm{C},k}$ due to its non-diagonal (circulant) structure. However, for practical PN levels, $\bar{\mathbf{P}}^t_{\mathrm{C},k}$ is a *banded* matrix with few significant diagonals resulting in ICI only from few neighboring subcarriers. Since $\mathbf{S}_k^H\mathbf{S}_k = \mathbf{I}_M$, we write

$$\mathbf{Y}^t_k = \mathbf{S}_k^H\mathbf{S}_k\mathbf{S}_k^H\bar{\mathbf{H}}^t_k\bar{\mathbf{P}}^t_{\mathrm{C},k}\mathbf{S}_k\mathbf{F}_M\mathbf{x}^t_k + \mathbf{Z}^t_k \tag{4.28}$$

where $\mathbf{S}_k\mathbf{S}_k^H$ is a diagonal matrix whose diagonal entries are zeros except for those whose indices lie in $J_k$. Since both $\bar{\mathbf{H}}^t_k$ and $\mathbf{S}_k\mathbf{S}_k^H$ are diagonal matrices, we can interchange them as follows

$$\mathbf{Y}^t_k = \underbrace{\mathbf{S}_k^H\bar{\mathbf{H}}^t_k\mathbf{S}_k}_{\triangleq\bar{\mathsf{H}}^t_k}\underbrace{\mathbf{S}_k^H\bar{\mathbf{P}}^t_{\mathrm{C},k}\mathbf{S}_k}_{\triangleq\bar{\mathsf{P}}^t_{\mathrm{C},k}}\mathbf{F}_M\mathbf{x}^t_k + \mathbf{Z}^t_k \equiv \bar{\mathsf{H}}^t_k\bar{\mathsf{P}}^t_{\mathrm{C},k}\mathbf{F}_M\mathbf{x}^t_k + \mathbf{Z}^t_k \tag{4.29}$$

where $\bar{\mathsf{H}}^t_k$ and $\bar{\mathsf{P}}^t_{\mathrm{C},k}$ are sub-matrices of $\bar{\mathbf{H}}^t_k$ and $\bar{\mathbf{P}}^t_{\mathrm{C},k}$, respectively, whose row and column indices lie in $J_k$. Note that $\bar{\mathsf{H}}^t_k$ is an $M \times M$ *diagonal* matrix while $\bar{\mathsf{P}}^t_{\mathrm{C},k}$ is an $M \times M$ *Toeplitz* [61] matrix. The next detection step is to apply FD equalization to equalize the channel effects in the matrix $\bar{\mathsf{H}}^t_k$. Applying the MMSE equalizer, we get

$$\mathbf{Y}^t_{k,eq} = \mathbf{E}^t_k\mathbf{Y}^t_k \tag{4.30}$$

where $\mathbf{E}^t_k = \left(\left(\bar{\mathsf{H}}^t_k\right)^H\bar{\mathsf{H}}^t_k + \frac{1}{\gamma_k}\right)^{-1}\left(\bar{\mathsf{H}}^t_k\right)^H \equiv \mathbf{S}_k^H\mathbf{D}_k\mathbf{S}_k$. Note that $\bar{\mathsf{H}}^t_k$ is a *diagonal* matrix and so is $\mathbf{E}^t_k$. Next, we apply the $M$-point inverse FFT to go back to TD where the $k$-th user's data exists. Meanwhile, we remove the bias factor, $\alpha^t_k$ , induced by the MMSE equalization to get

$$\mathbf{Y}^t_{k,TD} = \frac{1}{\alpha^t_k}\mathbf{F}_M^H\mathbf{Y}^t_{k,eq} = \mathbf{F}_M^H\bar{\mathsf{P}}^t_{\mathrm{C},k}\mathbf{F}_M\mathbf{x}^t_k + \mathbf{Z}^t_{k,TD} \equiv \mathsf{P}^t_{\mathrm{C},k}\mathbf{x}^t_k + \mathbf{Z}^t_{k,TD} \tag{4.31}$$

where $\mathsf{P}^t_{\mathrm{C},k} \triangleq \mathbf{F}_M^H\bar{\mathsf{P}}^t_{\mathrm{C},k}\mathbf{F}_M$. The vector $\mathbf{Z}^t_{k,TD}$ contains the noise and residual ICI, and $\alpha^t_k$ is the average of the diagonal entries of the diagonal matrix $\mathbf{E}^t_k\bar{\mathsf{H}}^t_k$. Exploiting the fact that *Toeplitz*

*matrices of large sizes can be approximated as circulant matrices* [61], we approximate the Toeplitz matrix $\bar{\mathsf{P}}^t_{\mathrm{C},k}$ by a circulant one. Hence, we approximate the matrix $\mathsf{P}^t_{\mathrm{C},k}$ by a diagonal matrix since circulant matrices are diagonalized by the FFT matrix. Note that $\bar{\mathsf{P}}^t_{\mathrm{C},k}$ is a sub-matrix of $\bar{\mathbf{P}}^t_{\mathrm{C},k} = \mathbf{F}_N \mathbf{P}^t_{\mathrm{C},k} \mathbf{F}_N^H$ where $\mathbf{P}^t_{\mathrm{C},k}$ is a diagonal matrix whose diagonal entries are complex exponentials of unit magnitude as in (4.4). Hence, the magnitudes of the diagonal entries of $\mathsf{P}^t_{\mathrm{C},k}$ are also approximately unity. Intuitively, the matrix $\mathsf{P}^t_{\mathrm{C},k}$ can be regarded as a PN matrix directly multiplying the complex data in $\mathbf{x}^t_k$ as in (4.31). In SC-FDMA, the complex data in $\mathbf{x}^t_k$ are in TD; hence, adding the PN samples directly to their phases is intuitively appealing. Finally, $\mathbf{Y}^t_{k,\,TD}$ in (4.31) can be re-written in an equivalent form as follows

$$\mathbf{Y}^t_{k,\,TD} = \mathbf{X}^t_k \mathbf{p}^t_k + \mathbf{Z}^t_{k,\,TD} \tag{4.32}$$

where $\mathbf{X}^t_k$ is an $M \times M$ diagonal matrix whose diagonal entries are given by the vector $\mathbf{x}^t_k$ while $\mathbf{p}^t_k$ is an $M \times 1$ vector containing the diagonal entries of the matrix $\mathsf{P}^t_{\mathrm{C},k}$.

### 4.5.2 Iterative Processing

In this section, we describe our proposed iterative channel decoding and PN compensation approach depicted in Figure 4.2. For the $t$-th SC-FDMA symbol, proceed as follows

**Initialization**: Initialize the estimate of the diagonal PN matrix $\mathsf{P}^t_{\mathrm{C},k}$ with the $M \times M$ diagonal matrix $\hat{\mathsf{P}}^{t,0}_{\mathrm{C},k}$, and initialize the iteration counter with $i = 0$.

**The $i^{\mathrm{th}}$ iteration**:

1. Using $\hat{\mathsf{P}}^{t,i}_{\mathrm{C},k}$ and based on Equation (4.31), we compensate for PN in $\mathbf{Y}^t_{k,\,TD}$ and obtain

$$\mathbf{Y}^{t,i}_{k,cmp} = \left(\hat{\mathsf{P}}^{t,i}_{\mathrm{C},k}\right)^{-1} \mathbf{Y}^t_{k,\,TD} \simeq \mathbf{x}^t_k + \left(\hat{\mathsf{P}}^{t,i}_{\mathrm{C},k}\right)^{-1} \mathbf{Z}^t_{k,\,TD} \triangleq \mathbf{x}^t_k + \mathbf{Z}^t_{k,cmp} \tag{4.33}$$

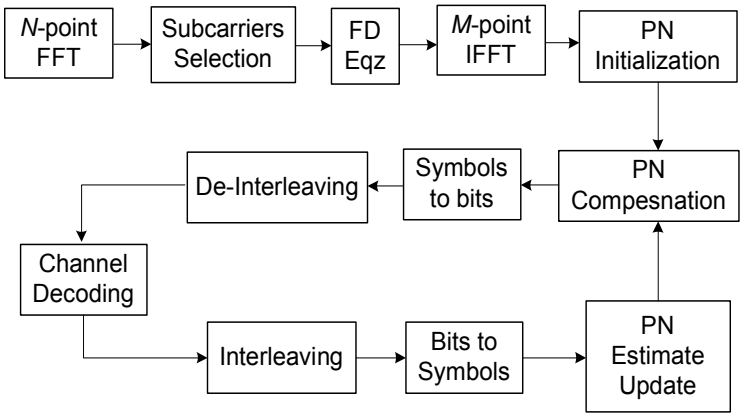


Figure 4.2. Iterative Channel Decoding and PN Compensation

where $\left(\hat{\mathsf{P}}_{\mathrm{C},k}^{t,i}\right)^{-1}$ is easily calculated since $\hat{\mathsf{P}}_{\mathrm{C},k}^{t,i}$ is an $M \times M$ diagonal matrix.

2. Compute the Log-likelihood-ratios (LLRs) of the code bits corresponding to the $m$-th entry of $\mathbf{Y}_{k,cmp}^{t,i}$, $0 \leq m \leq M-1$, as follows

$$L_m^{q,i} = \mathrm{Log}\left(\frac{\sum_{\theta \in S_1} \exp\left(\frac{-1}{\sigma_m^2}\left|\mathbf{Y}_{k,cmp}^{t,i}(m) - \theta\right|^2\right)}{\sum_{\theta \in S_0} \exp\left(\frac{-1}{\sigma_m^2}\left|\mathbf{Y}_{k,cmp}^{t,i}(m) - \theta\right|^2\right)}\right) \tag{4.34}$$

where $0 \leq q \leq Q-1$, $Q$ is the number of code bits represented by any constellation symbol, $2^Q$ is the constellation size. Furthermore, $\sigma_m^2 = \mathbb{E}\left[\left|\mathbf{Z}_{k,cmp}^{t}(m)\right|^2\right]$ and $S_1$ and $S_0$ are the sets of signal constellation symbols where the $q$-th bit is equal to 1 and 0, respectively. The LLR in (4.34) can be approximated by

$$L_m^{q,i} \simeq \frac{-1}{\sigma_m^2}\left(\min_{\theta \in S_1}\left|\mathbf{Y}_{k,cmp}^{t,i}(m) - \theta\right|^2 - \min_{\theta \in S_0}\left|\mathbf{Y}_{k,cmp}^{t,i}(m) - \theta\right|^2\right) \tag{4.35}$$

3. De-interleave the LLRs obtained in Step 2 and pass them to the channel decoder which updates them using the channel code structure. Then, interleave the updated LLRs and denote them by $\left\{\bar{L}_m^{q,i}\right\}$.

4. Using $\left\{\bar{L}_m^{q,i}\right\}$, compute soft estimates for the signal constellation symbols in $\mathbf{x}_k^t$ as follows

$$\hat{\mathbf{x}}_k^{t,i}(m) = \mathbb{E}\left[\mathbf{x}_k^t(m)\right] = \sum_{b=1}^{2^Q} \theta_b P\left(\mathbf{x}_k^t(m) = \theta_b\right) \tag{4.36}$$

where $\theta_b$ is a summation variable that takes on all values of the signal constellation points and $P\left(\mathbf{x}_k^t(m) = \theta_b\right)$ is the probability that $\mathbf{x}_k^t(m)$ equals $\theta_b$ which is calculated as follows

$$P\left(\mathbf{x}_k^t(m) = \theta_b\right) = \prod_{q=0}^{Q-1} P\left(c_m^q = c_b^q\right) = \prod_{q=0}^{Q-1} \frac{1}{2}\left(1 + (2c_b^q - 1)\tanh\left(\frac{\bar{L}_m^{q,i}}{2}\right)\right) \tag{4.37}$$

where tanh(.) is the hyperbolic tangent function and $c_m^q$, $c_b^q \in \{0, 1\}$ represent the $q$-th code bit corresponding to the signal constellation symbols $\mathbf{x}_k^t(m)$ and $\theta_b$, respectively. The second equality in (4.37) is obtained by observing that $\bar{L}_m^{q,i}$ can written as $\text{Log}\frac{P(c_m^q=1)}{P(c_m^q=0)}$.

5. An updated estimate of the PN vector $\mathbf{p}_k^t$ can be obtained by a simple point-wise division of the vector $\mathbf{Y}_{k,TD}^t$ by the vector $\hat{\mathbf{x}}_k^{t,i}$ obtained in (4.36). However, this approach does not exploit the low-pass nature of the PN process rendering the PN samples in $\mathbf{p}_k^t$ correlated. Two possible approaches can be followed to *exploit the correlation between the PN samples*. We call the first approach "**FD-processing**" where the PN vector $\mathbf{p}_k^t$ is written in terms of its FD transform $\bar{\mathbf{p}}_k^t$ as follows $\mathbf{p}_k^t = \mathbf{F}_M^H \bar{\mathbf{p}}_k^t$. Furthermore, most of the entries of $\bar{\mathbf{p}}_k^t$ are negligible due to the low-pass nature of the PN process. Hence, (4.32) can be written as follows

$$\mathbf{Y}_{k,TD}^t = \mathbf{X}_k^t \mathbf{F}_{M,LP}^H \bar{\mathbf{p}}_{k,LP}^t + \mathbf{Z}_{k,TD}^t \tag{4.38}$$

where $\bar{\mathbf{p}}_{k,LP}^t$ contains $D << M$ significant entries of $\bar{\mathbf{p}}_k^t$ and $\mathbf{F}_{M,LP}^H$ is an $M \times D$ tall matrix containing the columns of $\mathbf{F}_M^H$ corresponding to the significant entries of $\bar{\mathbf{p}}_k^t$.

Next, the linear least-squares estimate (LLSE) of $\bar{\mathbf{p}}_{k,LP}^{t}$ is obtained as follows

$$\hat{\bar{\mathbf{p}}}_{k,LP}^{t} = \left(\mathbf{A}^{H}\mathbf{A}\right)^{-1}\mathbf{A}^{H}\mathbf{Y}_{k,TD}^{t} \tag{4.39}$$

where $\mathbf{A} = \hat{\mathbf{X}}_{k}^{t,i}\mathbf{F}_{M,LP}^{H}$ and $\hat{\mathbf{X}}_{k}^{t,i}$ is a *diagonal* matrix whose diagonal entries are given by the estimate symbols $\left\{\hat{\mathbf{x}}_{k}^{t,i}(m)\right\}$ obtained in (4.36). The size of the matrix $\mathbf{A}^{H}\mathbf{A}$ is $D \times D$; hence, its inverse can be obtained easily since $D << M$ (typically, $D = 3$). Then, an updated estimate of the PN vector $\mathbf{p}_{k}^{t}$ is obtained as follows

$$\hat{\mathbf{p}}_{k}^{t,i} = \mathbf{F}_{M,LP}^{H}\hat{\bar{\mathbf{p}}}_{k,LP}^{t} \tag{4.40}$$

The second approach is called "**sub-block processing**" where the TD vector $\mathbf{Y}_{k,TD}^{t}$ is divided into C sub-blocks each of length $G = \frac{M}{\mathrm{C}}$. Thanks to the low-pass nature of the PN process, the PN samples are assumed constant over each sub-block. Hence, we write the $c$-th sub-block vector as follows

$$\mathbf{Y}_{k,c}^{t} \triangleq \mathbf{Y}_{k,TD}^{t}(cG : (c+1)G - 1) = p_{k}^{t}(c)\mathbf{x}_{k,c}^{t} + \mathbf{Z}_{k,c}^{t} \tag{4.41}$$

where $\mathbf{x}_{k,c}^{t} = \mathbf{x}_{k}^{t}(cG : (c+1)G - 1)$, $\mathbf{Z}_{k,c}^{t} = \mathbf{Z}_{k,TD}^{t}(cG : (c+1)G - 1)$, and the scalar $p_{k}^{t}(c)$ represents the phase noise sample perturbing the $c$-th block. Next, we compute the LLSE of $p_{k}^{t}(c)$ as follows

$$\hat{p}_{k}^{t,i}(c) = \frac{\left(\hat{\mathbf{x}}_{k,c}^{t,i}\right)^{H}\mathbf{Y}_{k,c}^{t}}{\parallel \hat{\mathbf{x}}_{k,c}^{t,i} \parallel_{2}^{2}} \tag{4.42}$$

where $\parallel . \parallel_{2}$ denotes the $l_2$-norm and $\hat{\mathbf{x}}_{k,c}^{t,i}$ is constructed from the estimates obtained in (4.36) for the $c$-th sub-block. Then, the estimate $\hat{p}_{k}^{t,i}(c)$ is assigned to the middle of the corresponding sub-block, and the PN vector $\mathbf{p}_{k}^{t}$ is estimated by inter/extrapolating $\left\{\hat{p}_{k}^{t,i}(c)\right\}_{c=0}^{\mathrm{C}-1}$ over the $M$ samples of the $t$-th SC-FDMA symbol. Note that

the low-pass property of the PN process is exploited in both the "FD-processing" and "sub-block processing" approaches to reduce the number of unknowns to be estimated. Hence, more averaging is done in (4.40) and (4.42) which helps in reducing the error propagation effect due to the inaccuracies of the symbols estimates in (4.36).

6. Construct the $M \times M$ diagonal matrix $\hat{\mathsf{P}}_{\mathrm{C},k}^{t,i+1}$ whose diagonal entries are given by the PN vector estimate $\hat{\mathbf{p}}_k^{t,i}$ calculated in Step 5 using either the "FD-processing" approach or the "sub-block processing" approach.

7. Set $i = i + 1$ and check the stopping criterion. If met, exit the algorithm, else go to Step 1. Several stopping criteria can be used, e.g. the maximum number of iterations or the cyclic-redundancy check.

*We emphasize that modeling both transmit and receive PN as solely transmit PN moved the equalization step outside the iterations which significantly reduced the overall complexity.*

### 4.5.3 Initialization

In this section, we investigate the PN initialization needed for the proposed iterative approach in Section 4.5.2. Specifically, we show how to choose the diagonal matrix $\hat{\mathsf{P}}_{\mathrm{C},k}^{t,0}$. If pilots were inserted in SC-FDMA symbols, they would be used to estimate the CPE for the initialization process as done in [41–43,90] for OFDM systems. The CPE is the average of the PN samples perturbing the SC-FDMA symbol. However, the model adopted in this work **assumes no pilots within the data symbols** in order to keep the low-PAPR property of the SC-FDMA modulation technique. To overcome this problem, we exploit the correlation between PN

samples perturbing consecutive SC-FDMA symbols. Hence, we choose the PN initialization of the $t$-th SC-FDMA symbol to be a function of the estimated PN in the $(t-1)$-th SC-FDMA symbol. Two choices are proposed in this work. First, we choose the *average* of the PN samples estimated in the $(t-1)$-th SC-FDMA symbol to initialize the PN of the $t$-th SC-FDMA symbol as follows

$$\hat{\mathsf{P}}_{\mathrm{C},k}^{t,0} = \left(\frac{1}{M}\mathbf{1}_M^T \hat{\mathbf{p}}_k^{t-1}\right)\mathbf{I}_M \tag{4.43}$$

where $\mathbf{1}_M$ is the $M \times 1$ all-ones vector and $\hat{\mathbf{p}}_k^{t-1}$ is the PN vector estimate obtained at the last iteration of the $(t-1)$-th SC-FDMA symbol detection. The second choice is the *last sample* of the estimated PN vector as follows

$$\hat{\mathsf{P}}_{\mathrm{C},k}^{t,0} = \hat{\mathbf{p}}_k^{t-1}(M)\mathbf{I}_M \tag{4.44}$$

The PN of the first SC-FDMA symbol is initialized with $\hat{\mathsf{P}}_{\mathrm{C},k}^{t,0} = \mathbf{I}_M, t = 0$. This is reasonable since a training symbol usually precedes the data symbols in all frames, as in the LTE standard [46], for timing and phase synchronization.

## 4.6 Simulation Results

In this section, we present numerical results for our derived EVM expression and our proposed PN compensation approach. We model SC-FDMA systems with $N = 512$, $L_p = 32$, $f_{\mathrm{sub}} = 15KHz$, $M = 96$, and the total number of guard subcarriers on both sides is 212 as specified in the LTE uplink standard [46] for 5 MHz bandwidth. We use 16-QAM modulation for all simulations except for Figure 4.7 where we use QPSK. We also use rate-1/2 convolutional coding with MAP decoding.

### 4.6.1 EVM Expressions Verification

In this subsection, we numerically verify the derived EVM expression by Monte-Carlo simulations. For all users, we use the exponentially-decaying multi-path channel model with 6 uncorrelated paths and 1 dB decaying factor per path. We use the localized subcarrier mapping scheme shown in Figure 4.1a. We assume that all users have the same received SNR level and transmit PN and CFO parameters, i.e., $\gamma_l = \gamma$, $\beta_l = \beta$, and $\Delta f_l = \Delta f, \forall l$. In Figure 4.3, we compare the analytical EVM expression versus the simulated one for user $u_1$ under various levels of joint transmit-receive PN and CFO. The accuracy of the EVM expression

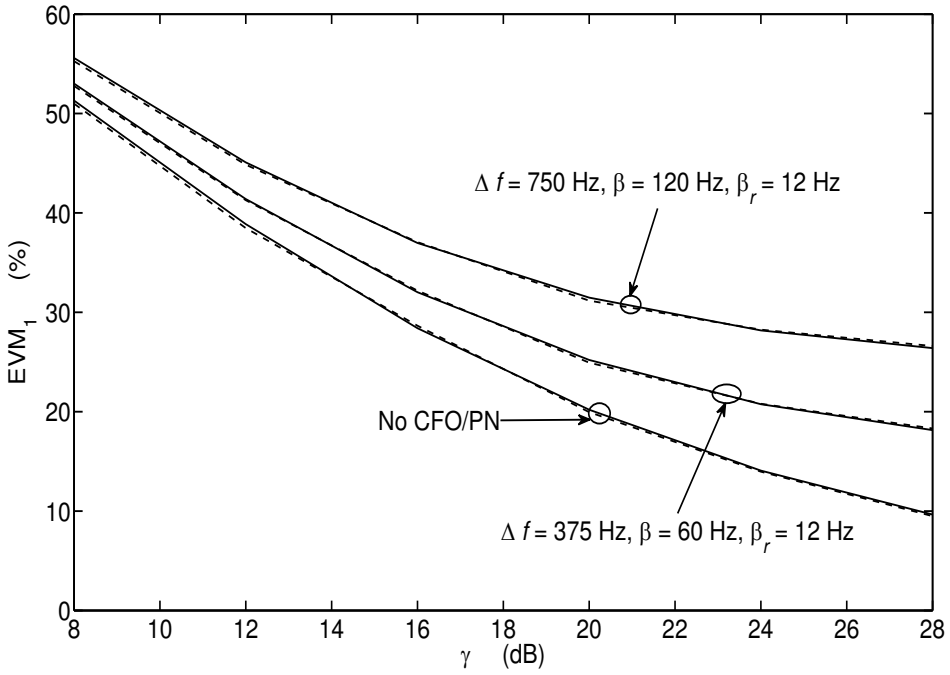


Figure 4.3. Comparison between analytical (solid lines) and simulated (dashed lines) EVM of $u_1$ with joint transmit-receive PN and CFO

derived in Section 4.3 is evident even for SNR levels as low as 8 dB. Furthermore, in Figure 4.4, we verify the observation made in Section 4.4.2 about the effect of subcarrier mapping on the EVM. We compare the analytical and simulated EVM of $u_1$ with both distributed

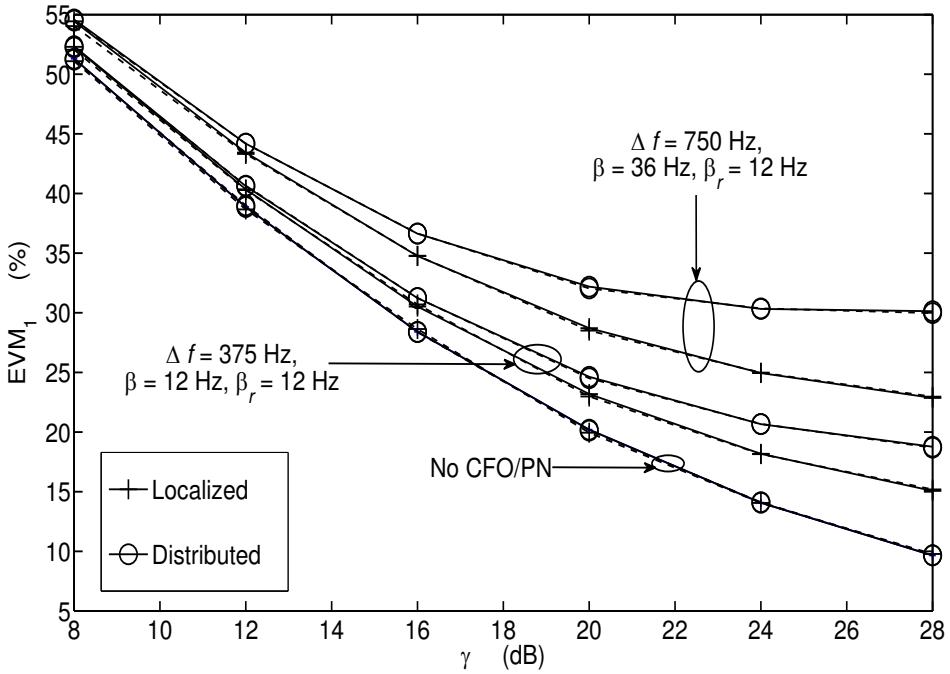


Figure 4.4. Comparison between analytical (solid lines) and simulated (dashed lines) EVM of $u_1$ with localized and distributed mapping schemes under various levels of CFO and PN

and localized subcarrier mapping schemes under various CFO and PN levels. In addition to the accuracy of the analytical EVM expression, we notice that the localized subcarrier mapping scheme is more immune to CFO and PN than the distributed subcarrier mapping scheme as discussed in Section 4.4.2. Note that the EVMs of both mapping schemes are identical in the case of no CFO/PN because there is neither IUI nor ICI. We also observe that the EVM gap between both schemes increases as the CFO/PN level increases. The derived EVM expression can be used to investigate the near-far effect by allowing different receive SNR levels ($\gamma_l$) for different users.

### 4.6.2 Performance Evaluation of the Proposed PN Compensation Algorithm

In this subsection, we simulate the performance of our proposed iterative compensation algorithm. The data frame consists of 6 data SC-FDMA symbols over which the channel is quasi-static as in the LTE uplink subframe [46]. The transmit and receive PN levels are set to $\beta_{t,l} = 110$ Hz $\forall\, l$, $\beta_r = 12$ Hz which correspond to oscillator PSDs of $-47.6$ dBc/Hz and $-57.2$ dBc/Hz, respectively, at 1 kHz frequency offset from the oscillator carrier frequency. We choose $\beta_{t,l}$ to be greater than $\beta_r = 12$ since users' oscillators are less accurate than the BST's oscillator. Perfect channel knowledge is assumed at the BST only. We use the localized subcarrier mapping scheme and assume enough guard subcarriers between users; hence, we ignore the PN-induced leakage between users' edge subcarriers. In Figure 4.5, we show the frame-error-rate (FER) performance of five iterations of our "sub-block

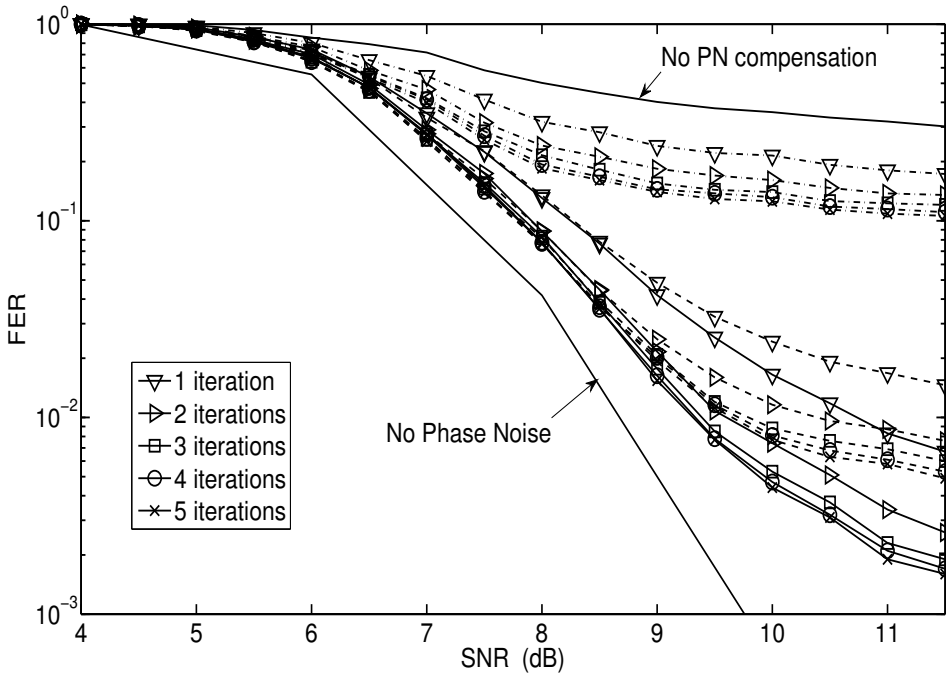


Figure 4.5. Performance of "sub-block processing" using no initialization (dash-dotted), initialization in (4.43) (dashed), and initialization in (4.44) (solid) for the AWGN channel

processing-based" approach without initialization, i.e., $\hat{\mathsf{P}}_{\mathrm{C},k}^{t,0} = \mathbf{I}_M$, $\forall\, t$, and with the two initializations proposed in (4.43) and (4.44). Clearly, both proposed initializations show significant performance improvement over the case without initialization and converge after only three iterations. Furthermore, the initialization in (4.44) outperforms that in (4.43) because the last PN sample of the $(t-1)$-th SC-FDMA symbol is the closest one to the PN samples of the $t$-th SC-FDMA symbol. *Linear interpolation is used and the number of sub-blocks $C$ is varied over iterations as follows.* In the early iterations, the SC-FDMA symbol is divided into few sub-blocks since the symbols estimates are still inaccurate and, hence, more averaging is needed. In later iterations with improved data symbols estimates, the SC-FDMA symbol is divided into more sub-blocks to improve the accuracy of the PN interpolation. In the 1st and 2nd iterations, $C$ is chosen to be 1 and 2, respectively. Then, we

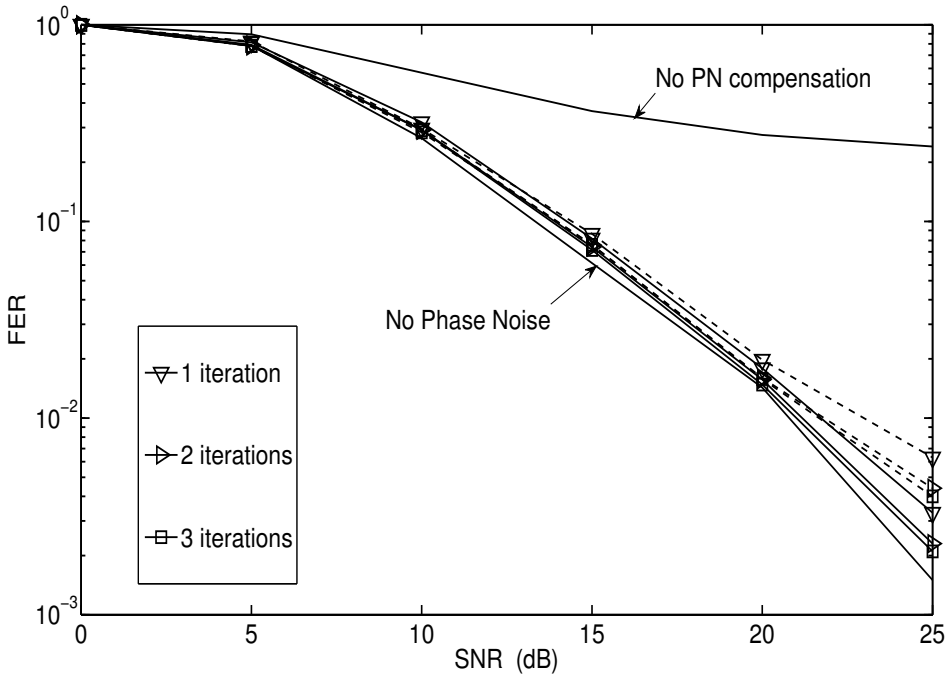


Figure 4.6. Performance of "FD-processing" (dashed) and "sub-block processing" (solid) with the initialization in (4.44) for the SCM macro suburban channel

increment $C$ by 2 for each subsequent iteration. Similarly, for the "FD-processing" approach, the parameter $D$ is chosen to increase over iterations where $D = 1$ in the $1^{\text{st}}$ iteration and then incremented by 2 for each subsequent iteration. In Figures 4.6 and 4.7, we compare the performances of our "FD-processing-based" and "sub-block processing-based" iterative approaches for the SCM [91] suburban and urban macro channel models, respectively, using the initialization in (4.44). The latter is shown to outperform the former which is intuitively appealing since "FD-processing" approximates the non-periodic PN process by a periodic one by keeping only few significant FFT coefficients. This leads to high estimation errors especially at both edges of the PN block [43, 92]. Furthermore, the performance of the "sub-block processing-based" iterative approach converges after three iterations to that with no PN for FER values of practical interest as shown in Figure 4.6. The impact of channel esti-

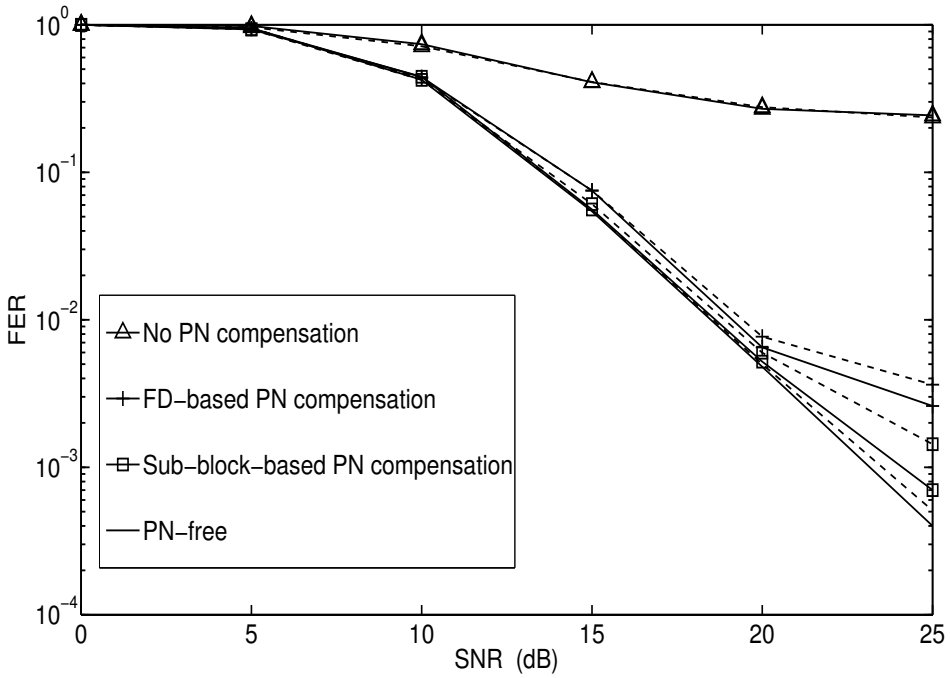


Figure 4.7. Performance of various algorithms using known channel (solid) and estimated channel (dashed) for the SCM macro urban channel. The initialization in (4.44) is used for our PN compensation algorithms with 3 iterations

mation errors is shown also in Figure 4.7 where the pilots preceding each data frame are used to compute the standard LLSE of the channel frequency response (CFR) at each subcarrier. The pilots are generated using the Zadoff-Chu sequences as in the LTE standard [46].

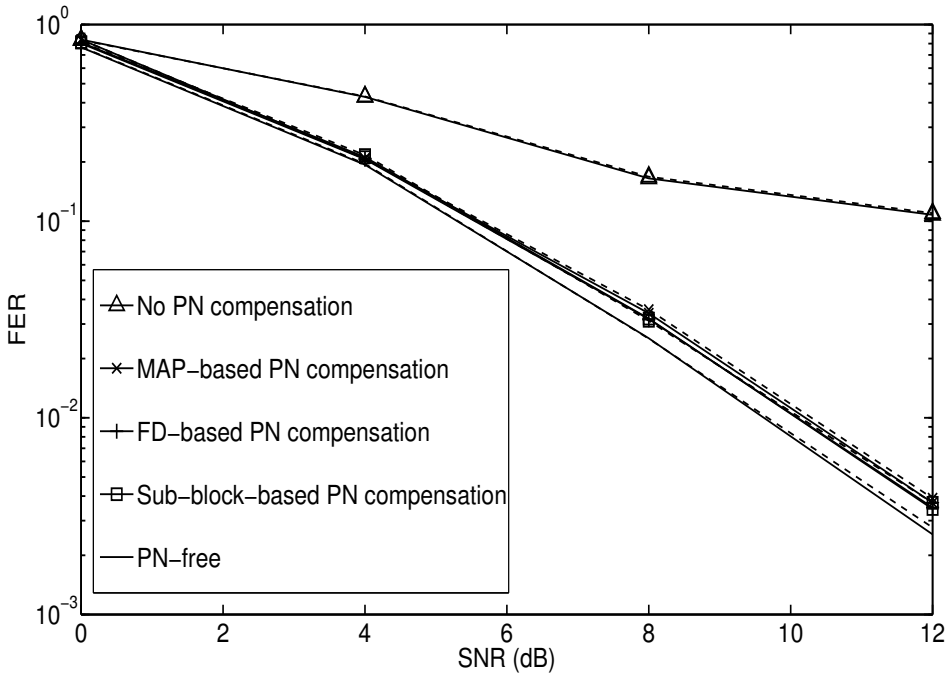


Figure 4.8. Performance comparison with MAP-based PN compensation in [1] for QPSK using known (solid) and estimated (dashed) channel for the SCM macro urban channel. The initialization in (4.44) is used for our PN compensation algorithms with 3 iterations

Our proposed algorithms experience insignificant SNR loss of 0.3 dB at FER=$10^{-2}$ due to channel estimation errors; hence, they are robust under channel estimation errors. In Figure 4.8, we compare the performances of our proposed PN compensation algorithms with the MAP-based one proposed in [1] for QPSK[5] with known and estimated channels. We observe that our proposed algorithms and the MAP-based one almost achieve the same performance with known and estimated channels. However, our algorithms have the advantage that they are applicable to any modulation and do not require the optimization of any parameters

[5]Algorithm in [1] is originally designed for constant-amplitude modulation

unlike the MAP algorithm where careful choices of some parameters are required to satisfy the stability conditions. Furthermore, our sub-block-based algorithm requires less complexity than the MAP algorithm where the latter requires $5M$ less real multiplications than the former per iteration/SC-FDMA/iteration. Also, our algorithm does not involve the evaluation of several hyperbolic functions as in [1]. The performances of our proposed algorithms

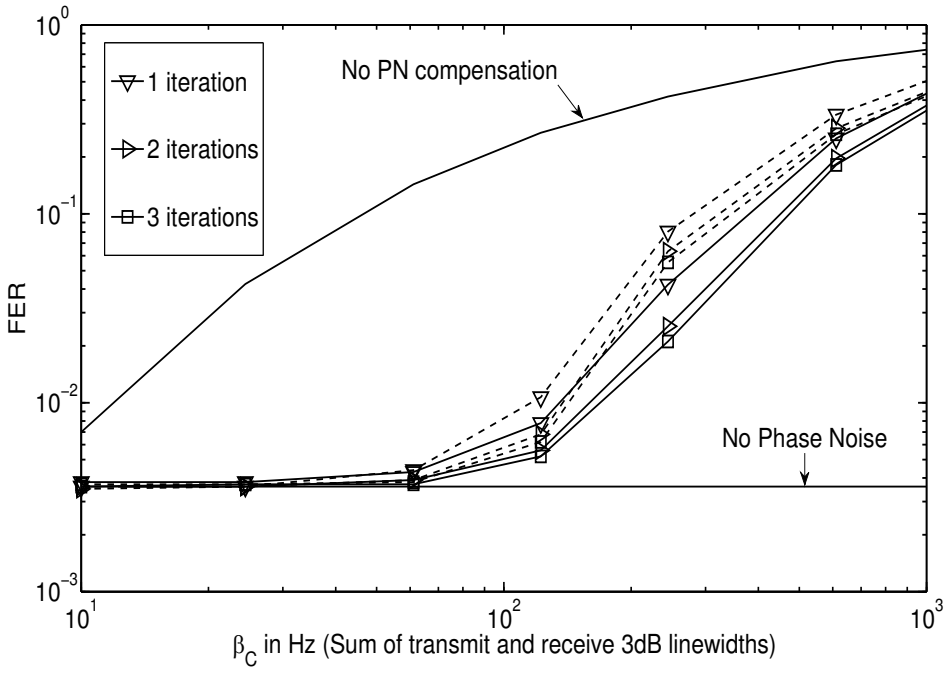


Figure 4.9. FER versus PN level for "FD-processing" (dashed) and "sub-block processing" (solid) with the initialization in (4.44) for the SCM macro urban channel with SNR = 20 dB

versus the PN level are shown in Figure 4.9. We conclude by evaluating the computational complexity of the sub-block-based PN estimation approach in Table 4.2. Assuming that each real multiplication or division operation is implemented in one instruction and neglecting

Table 4.2. Complexity analysis of sub-block processing-based PN estimation per SC-FDMA symbol per user per iteration

| Real multiplications | Real additions | Real divisions |
|---|---|---|
| $\mathrm{O}(8M)$ | $\mathrm{O}(9M + 2C - 3)$ | $\mathrm{O}(2C)$ |

the complexity of real additions, the sub-block processing algorithm with $C = 4$ requires approximately 9.4 millions instructions per second (MIPS) per user per iteration which is well within the processing power available at LTE base stations.

# CHAPTER 5

# CONCLUSIONS

## 5.1 Narrow-band interference mitigation

We presented a novel approach to exploit the sparsity inherent in the frequency-domain representation of NBI signals to estimate and mitigate their effects on SISO and MIMO-OFDM systems. The approach can be extended to single-carrier systems [93] as well. In addition, a novel CS-based technique is proposed for channel estimation in the presence of NBI where NBI is first estimated and cancelled before channel estimation. Simulations showed that the performance loss due to channel estimation errors is insignificant. Several extensions were studied including multiple and mobile NBI whose channels are time-varying and frequency-selective. Furthermore, we proposed a reduced-complexity implementation for our CS-based approach where the inversion of a channel-dependent matrix is avoided. Finally, we showed that our CS-based approach for NBI estimation is applicable not only to ZP-OFDM but also to CP-OFDM systems as well. A direct application of NBI cancellation in MIMO systems is the WLANs standard IEEE 802.11n where multiple antennas are adopted in the presence of NBI from other devices operating in the same unlicensed (2.4 and/or 5 GHz) bands.

## 5.2 Sparse FIR MIMO equalizers

We proposed a novel approach for sparse FIR filter design and applied it to three broadband communication scenarios. First, we formulated convex programs and greedy algorithms to

design sparse fractionally-spaced FIR LEs, DFEs, and MIMO-DFEs where dramatic complexity reductions are achieved at a small performance loss compared to the conventional MMSE non-sparse FIR equalizer design. Second, we proposed a novel approach to design sparse FIR channel shortening equalizers. The key idea behind this new approach is to allow the TIR taps to be noncontiguous to have a larger search space for the TIR taps. The new sparse TIR design was shown to yield a lower channel shortening MSE compared to the conventional MMSE contiguous TIR design. Third, we designed a novel per-tone partial self-FEXT canceler for vectored VDSL systems where only a subset of the cancellation taps is activated resulting in substantial complexity reductions. Our design uses the OMP algorithm to determine the locations and weights of the active cancellation taps. We demonstrate a significant reduction in crosstalk cancellation complexity at small data rate losses compared to theoretical limits.

### 5.3 Phase noise in SC-FDMA

We derived a new EVM expression for SC-FDMA uplink transmissions under CFO and joint transmit-receive PN taking into account MMSE channel equalization effects. The resulting expression is a function of the CFO and PN levels, the active users' received SNR levels and subcarrier assignments, and the users' channel delay profiles. Based on the derived EVM expression, we compared the localized and distributed subcarrier mapping schemes from the viewpoint of their immunity to CFO and PN. In addition, we proposed an iterative fast-converging reduced-complexity algorithm for joint decoding and PN compensation in SC-FDMA systems. As in the LTE uplink, no pilots are multiplexed with data subcarriers to keep the low-PAPR feature of SC-FDMA systems. Hence, our proposed PN compensation

algorithm does not use pilots in PN tracking. Instead, we use the decoded data iteratively for PN estimation and compensation. Furthermore, the PN low-pass nature is exploited without assuming a-priori knowledge about the exact PN model to avoid model mismatch problems. Simulation results showed that our PN compensation algorithm converges after few iterations and achieves significant performance gains.

## 5.4 Future work

### 5.4.1 Sparsity-based estimation

Our sparsity-aware narrow-band estimation approach can be extended to single carrier systems, e.g. single-carrier frequency-domain equalization [93]. Furthermore, it can be extended to other multiple-input multiple-output flavors, e.g. spatial multiplexing and beamforming.

### 5.4.2 Sparse equalization

Our sparse channel shortening approach can be further extended to multiple-input multiple-output systems. Also, our sparse channel shortening approach can be extended to multiple-access schemes, e.g. single-carrier frequency-division multiple-access systems. Furthermore, our sparse framework design can be applied to different applications and various finite impulse response filters.

### 5.4.3 RF impairments

Further extension to our RF impairments work is the studying of single-carrier frequency-division multiple-access systems systems under inphase/quadrature imbalance impairing the

RF front-end. The error-vector magnitude can be derived and simplified under I/Q imbalance. Furthermore, our compensation approach can be extended to compensate the I/Q imbalance.

# APPENDIX

Using the identity $\mathrm{Tr}\left(\mathbf{AB}\right) = \mathrm{Tr}\left(\mathbf{BA}\right)$ for any two matrices $\mathbf{A}$ and $\mathbf{B}$, we write

$$t_{1A} \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{P}}_r \bar{\mathbf{Q}}_k \bar{\mathbf{H}}_k \bar{\mathbf{P}}_k \mathsf{S}_k \mathbf{D}_k\right)\right] \tag{A.1}$$

Using the fact that diagonal matrices are commutative under multiplication and assuming that the channel and PN processes are statistically independent, we write

$$\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\mathbb{E}\left[\bar{\mathbf{P}}_r\right] \bar{\mathbf{Q}}_k \bar{\mathbf{H}}_k \mathbb{E}\left[\bar{\mathbf{P}}_k\right] \mathbf{D}_k \mathsf{S}_k\right)\right] \tag{A.2}$$

Since both $\bar{\mathbf{H}}_k$ and $\mathbf{D}_k$ are diagonal matrices, we rewrite (A.2) as follows

$$\begin{aligned}\mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\right)\right] &= \mathrm{Tr}\left(\mathbb{E}\left[\bar{\mathbf{P}}_r\right] \bar{\mathbf{Q}}_k \left(\mathbb{E}\left[\mathrm{diag}(\bar{\mathbf{H}}_k)\left(\mathrm{diag}(\mathbf{D}_k)\right)^T\right] \circ \mathbb{E}\left[\bar{\mathbf{P}}_k\right]\right) \mathsf{S}_k\right) \\ &\triangleq \mathrm{Tr}\left(\mathbb{E}\left[\bar{\mathbf{P}}_r\right] \bar{\mathbf{Q}}_k \left(\mathbf{T}_k \circ \mathbb{E}\left[\bar{\mathbf{P}}_k\right]\right) \mathsf{S}_k\right)\end{aligned}$$

where the matrix $\mathbf{T}_k \triangleq \mathbb{E}\left[\mathrm{diag}(\bar{\mathbf{H}}_k)\left(\mathrm{diag}(\mathbf{D}_k)\right)^T\right]$. We assume that the CIR follows the zero-mean circular symmetric complex Gaussian distribution and so does $\bar{\mathbf{H}}_k(n,n)$.

**Lemma 1.** Let $f(x)$ and $g(y)$ denote two functions of the jointly Gaussian random variables $x$ and $y$, respectively, with the operator $|$ denoting the conditional expectation. Then,

$$\mathbb{E}_{x,y}\left[f(x)g(y)\right] = \mathbb{E}_x\mathbb{E}_{y|x}\left[f(x)g(y)\right] = \mathbb{E}_x\left[f(x)\mathbb{E}_{y|x}\left[g(y)\right]\right] = \int_x f(x)\int_y g(y)p(y|x)dy\,dx \tag{A.3}$$

where $p(y|x)$ is the conditional probability density function of $y$ given $x$. Since $x$ and $y$ are

jointly Gaussian, $p(y|x)$ is also Gaussian whose conditional mean and variance are [66]

$$\mathbb{E}\left[y|x\right] = \mathbb{E}\left[y\right] + \frac{\mathrm{cov}\left[y,x\right]}{\mathrm{var}\left[x\right]}\left(x - \mathbb{E}\left[x\right]\right), \qquad \mathrm{var}\left[y|x\right] = \mathrm{var}\left[y\right] - \frac{\left|\mathrm{cov}\left[y,x\right]\right|^2}{\mathrm{var}\left[x\right]} \tag{A.4}$$

where $\mathrm{var}\left[x\right] \triangleq \mathbb{E}[|x - \mathbb{E}[x]|^2]$ and $\mathrm{cov}\left[y,x\right] \triangleq \mathbb{E}[(y - E[y])(x - E[x])^*]$. Denoting $\bar{\mathbf{H}}_k(m,m)$ and $\bar{\mathbf{H}}_k(n,n)$ by $y$ and $x$, respectively, we compute the entries of $\mathbf{T}_k$ as follows

$$\mathbf{T}_k(m,n) = \mathbb{E}_{x,y}\left[\frac{yx^*}{|x|^2 + \gamma_k^{-1}}\right] = \mathbb{E}_x\ \mathbb{E}_{y|x}\left[\frac{yx^*}{|x|^2 + \gamma_k^{-1}}\right] = \mathbb{E}_x\left[\frac{x^*\mathbb{E}_{y|x}\left[y\right]}{|x|^2 + \gamma_k^{-1}}\right], \quad \forall\, m,n \tag{A.5}$$

Since $\bar{\mathbf{H}}_k(m,m)$ and $\bar{\mathbf{H}}_k(n,n)$ are jointly Gaussian, we use Lemma 1 and write

$$\mathbb{E}_{y|x}\left[y\right] = \mathbb{E}\left[y\right] + \frac{\mathbb{E}\left[yx^*\right]}{\mathbb{E}\left[|x|^2\right]}\left(x - \mathbb{E}\left[x\right]\right) = \frac{\mathbf{R}_k(m,n)}{\mathbf{R}_k(n,n)}\ x \tag{A.6}$$

$$\mathbf{T}_k(m,n) = \frac{\mathbf{R}_k(m,n)}{\mathbf{R}_k(n,n)}\mathbb{E}_x\left[\frac{|x|^2}{|x|^2 + \gamma_k^{-1}}\right] = \frac{\mathbf{R}_k(m,n)}{\mathbf{R}_k(n,n)}\int_{z=0}^{\infty}\frac{z}{z+\gamma_k^{-1}}\frac{\exp\left(-z/\mathbb{E}[z]\right)}{\mathbb{E}[z]}dz \tag{A.7}$$

$$= \frac{\mathbf{R}_k(m,n)}{\mathbf{R}_k(n,n)}\left(1 + p_k^n \exp\left(p_k^n\right)\mathrm{Ei}\left(-p_k^n\right)\right), \quad \forall\, m,n \tag{A.8}$$

where $z \triangleq |x|^2$ is exponentially distributed and $\mathbb{E}[z] = \mathbb{E}\left[\left|\bar{\mathbf{H}}_k(n,n)\right|^2\right] = \mathbf{R}_k(n,n)$.

Next, we evaluate the quantity $t_{1B}$ as follows

$$t_{1B} \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kk}\mathbf{U}_{kk}^H\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{P}}_r\bar{\mathbf{Q}}_k\bar{\mathbf{H}}_k\mathbf{W}_k\bar{\mathbf{H}}_k^H\bar{\mathbf{Q}}_k^H\bar{\mathbf{P}}_r^H\mathbf{D}_k\mathbf{D}_k^H\mathsf{S}_k\right)\right] \tag{A.9}$$

where

$$\mathbf{W}_k \triangleq \mathbb{E}\left[\bar{\mathbf{P}}_k\mathsf{S}_k\bar{\mathbf{P}}_k^H\right] = \mathbf{F}_N\mathbb{E}\left[\mathbf{P}_k\left(\mathbf{F}_N^H\mathsf{S}_k\mathbf{F}_N\right)\mathbf{P}_k^H\right]\mathbf{F}_N^H = \mathbf{F}_N\left(\left(\mathbf{F}_N^H\mathsf{S}_k\mathbf{F}_N\right)\circ\mathbf{V}_k\right)\mathbf{F}_N^H \tag{A.10}$$

and $\mathbf{V}_k \triangleq \mathbb{E}\left[\mathrm{diag}(\mathbf{P}_k)\left(\mathrm{diag}(\mathbf{P}_k)\right)^H\right]$ whose entries are given in (4.16). Inspecting the structure of $\mathbf{W}_k$ in (A.10), we find that $\mathbf{F}_N^H\mathbf{W}_k\mathbf{F}_N$ is a Toeplitz matrix [61] because $\mathbf{V}_k$ is Toeplitz and $\mathbf{F}_N^H\mathsf{S}_k\mathbf{F}_N$ is circulant. Furthermore, Toeplitz matrices with large sizes can be approximated as circulant matrices [61]; hence, for large FFT sizes, the matrix $\mathbf{W}_k$ can be approximated by its main diagonal while setting the off-diagonal entries to zero. Using this

approximation, we interchange $\mathbf{W}_k$ and the diagonal matrix $\bar{\mathbf{H}}_k$ and proceed as follows

$$t_{1B} \simeq \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{P}}_r\bar{\mathbf{Q}}_k\mathbf{W}_k\bar{\mathbf{H}}_k\bar{\mathbf{H}}_k^H\bar{\mathbf{Q}}_k^H\bar{\mathbf{P}}_r^H\mathbf{D}_k\mathbf{D}_k^H\mathsf{S}_k\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{P}}_r\bar{\mathbf{Q}}_k\mathbf{W}_k\left(\left(\bar{\mathbf{Q}}_k^H\bar{\mathbf{P}}_r^H\right)\circ\mathbf{Y}_k\right)\mathsf{S}_k\right)\right] \tag{A.11}$$

where

$$\mathbf{Y}_k \triangleq \mathbb{E}\left[\mathrm{diag}(\bar{\mathbf{H}}_k\bar{\mathbf{H}}_k^H)\left(\mathrm{diag}(\mathbf{D}_k\mathbf{D}_k^H)\right)^T\right]$$

whose entries, $\mathbf{Y}_k(m,n) = \mathbb{E}\left[\frac{\left|\bar{\mathbf{H}}_k(m,m)\right|^2\left|\bar{\mathbf{H}}_k(n,n)\right|^2}{\left(\left|\bar{\mathbf{H}}_k(n,n)\right|^2+\gamma_k^{-1}\right)^2}\right]$, are computed in (4.17) using Lemma 1 and the relation $\mathbb{E}[|y|^2|x] = \mathrm{var}[y|x] + |E[y|x]|^2$ yielding

$$\mathbb{E}\left[\left|\bar{\mathbf{H}}_k(m,m)\right|^2 \mid \bar{\mathbf{H}}_k(n,n)\right] = \frac{\left|\mathbf{R}_k(m,n)\right|^2}{\left|\mathbf{R}_k(n,n)\right|^2}\left|\bar{\mathbf{H}}_k(n,n)\right|^2 + \frac{\mathbf{R}_k(m,m)\mathbf{R}_k(n,n) - \left|\mathbf{R}_k(m,n)\right|^2}{\mathbf{R}_k(n,n)} \tag{A.12}$$

Since $\mathbf{W}_k$ is approximated by a diagonal matrix, it can be moved and multiplied by $\mathbf{Y}_k$ as follows

$$\begin{aligned} t_{1B} &\simeq \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{P}}_r\bar{\mathbf{Q}}_k\left(\left(\bar{\mathbf{Q}}_k^H\bar{\mathbf{P}}_r^H\right)\circ(\mathbf{W}_k\mathbf{Y}_k)\right)\mathsf{S}_k\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\left(\mathbf{C}_k\circ\mathbf{C}_k^*\right)\mathbf{L}_k\mathsf{S}_k\right)\right] \\ &= \sum_{n\in J_k}\sum_{m=0}^{N-1}\mathbb{E}\left[\left|\mathbf{C}_k(1,((m-n))_N\right|^2\right]\mathbf{L}_k(m,n) \end{aligned} \tag{A.13}$$

where $\mathbf{C}_k \triangleq \bar{\mathbf{P}}_r\bar{\mathbf{Q}}_k$, $\mathbf{L}_k \triangleq \mathbf{W}_k\mathbf{Y}_k$, and$((m-n))_N \triangleq (m-n) \bmod N$ is the module-$N$ operation which originates from the circulant structure of the matrix $\mathbf{C}_k$, defined in (A.13). Furthermore,

$$\mathbf{C}_k(1,((m-n))_N) = \frac{1}{\sqrt{N}}\mathbf{f}_{((m-n))_N}^H\mathrm{diag}\left(\mathbf{P}_r\mathbf{Q}_k\right) \tag{A.14}$$

$$\mathbb{E}\left[\left|\mathbf{C}_k(1,((m-n))_N)\right|^2\right] = \frac{1}{N}\mathbf{f}_{((m-n))_N}^H\underbrace{\mathbb{E}\left[\mathrm{diag}\left(\mathbf{P}_r\mathbf{Q}_k\right)\left(\mathrm{diag}\left(\mathbf{P}_r\mathbf{Q}_k\right)\right)^H\right]}_{\triangleq\mathbf{V}_r}\mathbf{f}_{((m-n))_N} \tag{A.15}$$

Then, it can be easily shown that the entries of $\mathbf{V}_r$ are those given in (4.14). Substituting (A.15) back into (A.13), we get the final expression in (4.13).

Next, we derive the quantity $t_2$ assuming that the channels experienced by different users are independent.

$$t_2 \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\mathbf{U}_{kl}\mathbf{U}_{kl}^H\right)\right] = \mathbb{E}\left[\mathrm{Tr}\left(\bar{\mathbf{P}}_r\bar{\mathbf{Q}}_l\bar{\mathbf{H}}_l\mathbf{W}_l\bar{\mathbf{H}}_l^H\bar{\mathbf{Q}}_l^H\bar{\mathbf{P}}_r^H\mathsf{D}_k\mathsf{S}_k\right)\right] \tag{A.16}$$

where $\mathbf{W}_l \triangleq \mathbb{E}\left[\bar{\mathbf{P}}_l\mathsf{S}_l\bar{\mathbf{P}}_l^H\right]$ is given in (A.10), and $\mathsf{D}_k \triangleq \mathbb{E}\left[\mathbf{D}_k\mathbf{D}_k^H\right]$ is a *diagonal* matrix with the diagonal entries

$$\mathsf{D}_k(n,n) = \mathbb{E}\left[\frac{\left|\bar{\mathbf{H}}_k(n,n)\right|^2}{\left(\left|\bar{\mathbf{H}}_k(n,n)\right|^2+\gamma_k^{-1}\right)^2}\right] = \frac{-1}{\mathbf{R}_k(n,n)}\times\left(1+(1+p_k^n)\exp\left(p_k^n\right)\mathrm{Ei}\left(-p_k^n\right)\right) \tag{A.17}$$

Exploiting the *diagonal* structure of $\bar{\mathbf{H}}_l$, we re-write (A.16) as follows

$$t_2 = \mathrm{Tr}\left(\mathbb{E}\left[\bar{\mathbf{P}}_r\bar{\mathbf{Q}}_l\left(\mathbf{W}_l\circ\mathbb{E}\left[\mathrm{diag}\left(\bar{\mathbf{H}}_l\right)\left(\mathrm{diag}\left(\bar{\mathbf{H}}_l\right)\right)^H\right]\right)\bar{\mathbf{Q}}_l^H\bar{\mathbf{P}}_r^H\mathsf{D}_k\mathsf{S}_k\right]\right) \tag{A.18}$$

Observing that $\mathbf{R}_l = \mathbb{E}\left[\mathrm{diag}\left(\bar{\mathbf{H}}_l\right)\left(\mathrm{diag}\left(\bar{\mathbf{H}}_l\right)\right)^H\right]$ and defining $\mathbf{G}_l \triangleq \mathbf{W}_l\circ\mathbf{R}_l$, we write

$$\begin{aligned} t_2 &= \mathrm{Tr}\left(\mathbb{E}\left[\bar{\mathbf{P}}_r\bar{\mathbf{Q}}_l\mathbf{G}_l\bar{\mathbf{Q}}_l^H\bar{\mathbf{P}}_r^H\mathsf{D}_k\mathsf{S}_k\right]\right) = \mathrm{Tr}\left(\mathbb{E}\left[\mathbf{F}_N\mathbf{P}_r\mathbf{Q}_l\mathbf{F}_N^H\mathbf{G}_l\mathbf{F}_N\mathbf{Q}_l^H\mathbf{P}_r^H\mathbf{F}_N^H\mathsf{D}_k\mathsf{S}_k\right]\right) \\ &= \mathrm{Tr}\left(\mathbf{F}_N\left(\mathbf{F}_N^H\mathbf{G}_l\mathbf{F}_N\circ\mathbb{E}\left[\mathrm{diag}\left(\mathbf{P}_r\mathbf{Q}_l\right)\left(\mathrm{diag}\left(\mathbf{P}_r\mathbf{Q}_l\right)\right)^H\right]\right)\mathbf{F}_N^H\mathsf{D}_k\mathsf{S}_k\right) \end{aligned} \tag{A.19}$$

where the last equality holds thanks to the diagonal structure of the matrix $\mathbf{P}_r\mathbf{Q}_l$. Inspecting the definition of $\mathbf{V}_r$ in (A.15), we write

$$t_2 = \mathrm{Tr}\left(\mathbf{F}_N\left(\mathbf{F}_N^H\mathbf{G}_l\mathbf{F}_N\circ\mathbf{V}_r\right)\mathbf{F}_N^H\mathsf{D}_k\mathsf{S}_k\right) \triangleq \mathrm{Tr}\left(\mathbf{W}_r\mathsf{D}_k\mathsf{S}_k\right) = \sum_{n\in J_k}\mathbf{W}_r(n,n)\mathsf{D}_k(n,n) \tag{A.20}$$

Finally, we evaluate $t_3$ (which is the noise contribution to the EVM) as follows

$$t_3 \triangleq \mathbb{E}\left[\mathrm{Tr}\left(\breve{\mathbf{z}}\breve{\mathbf{z}}^H\right)\right] = N_o\mathrm{Tr}\left(\mathsf{D}_k\mathsf{S}_k\right) = N_o\sum_{n\in J_k}\mathsf{D}_k(n,n) \tag{A.21}$$

Substituting the entries of $\mathsf{D}_k$ from (A.17) into (A.20) and (A.21), we get the expressions in (4.18) and (4.19), respectively.

## REFERENCES


[1] M. Sabbaghian and D. Falconer, "Joint turbo frequency domain equalization and carrier synchronization," *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 204 –212, 2008.

[2] IEEE Standard 802.11n, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," 2009.

[3] J. Park, D. Kim, C. Kang, and D. Hong, "Effect of bluetooth interference on OFDM-based WLAN," in *Proc. IEEE VTC-Fall*, 2003.

[4] A. Batra and J. Zeidler, "Narrowband interference mitigation in OFDM systems," in *Proc. IEEE MILCOM*, 2008.

[5] ——, "Narrowband interference mitigation in BICM OFDM systems," in *Proc. IEEE ICASSP*, 2009.

[6] D. Darsena and F. Verde, "Successive NBI cancellation using soft decisions for OFDM systems," *IEEE Signal Processing Letters*, vol. 15, pp. 873–876, 2007.

[7] M. Morelli and M. Moretti, "Channel estimation in OFDM systems with unknown interference," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 5338–5347, 2009.

[8] M. Han *et. al.*, "OFDM channel estimation with jammed pilot detector under narrowband jamming," *IEEE Trans. on Veh. Tech.*, vol. 57, pp. 1934–1939, 2008.

[9] R. Nilsson, F. Sjoberg, and J. LeBlanc, "A rank-reduced LMMSE canceller for narrowband interference suppression in OFDM-based systems," *IEEE Trans. on Commun.*, vol. 51, no. 12, pp. 2126–2140, 2003.

[10] D. Darsena, G. Gelli, L. Paura, and F. Verde, "A constrained maximum-SINR NBI-resistant receiver for OFDM systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 3032–3047, 2007.

[11] D. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289 – 1306, April 2006.

[12] S. Raghavan, J. Wolf, L. Milstein, and L. Barbosa, "Nonuniformly Spaced Tapped-Delay-Line Equalizers," *IEEE Transactions on Communications*, vol. 41, pp. 1290–1295, 1993.

[13] M. Melvasalo, P. Janis, and V. Koivunen, "Sparse equalization in high data rate WCDMA systems," in *IEEE SPAWC*, June 2007.

[14] S. Ariyavisitakul, N. Sollenberger, and L. Greenstein, "Tap-Selectable Decision Feedback Equalization," *IEEE Transactions on Communications*, vol. 45, pp. 1497–1500, 1997.

[15] N. Al-Dhahir and C. Fragouli, "How to choose the number of taps in a DFE," in *Proc. CISS*, March 2002.

[16] N. Al-Dhahir and A. Sayed, "The Finite-Length Multi-Input Multi-Output MMSE-DFE," *IEEE Trans. on Signal process.*, vol. 48, pp. 2921–2936, 2000.

[17] K. Berberidis and A. Rontogiannis, "Efficient Decision-Feedback Equalizer for Sparse Multipath Channels," in *IEEE ICASSP*, 2000.

[18] G. Kutz and D. Raphaeli, "Determination of tap positions for sparse equalizers," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1712–1724, 2007.

[19] H. Sui, E. Masry, and B. Rao, "Chip-level DS-CDMA downlink interference suppression with optimized finger placement," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3908–3921, 2006.

[20] T. Baran, D. Wei, and A. Oppenheim, "Linear programming algorithms for sparse filter design," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1605 –1617, 2010.

[21] N. Al-Dhahir and J. M. Cioffi, "Efficiently Computed Reduced-Parameter Input-Aided MMSE Equalizers for ML Detection: A Unified Approach," *IEEE Trans. on Inf. Theory*, vol. 42, pp. 903–915, 1996.

[22] N. Al-Dhahir, "FIR Channel-Shortening Equalizers for MIMO ISI Channels," *IEEE Trans. on Commun.*, vol. 49, pp. 213–218, 2001.

[23] S. Roy, T. Duman, and V. McDonald, "Error rate improvement in underwater MIMO communications using sparse partial response equalization," *IEEE J. of Ocean. Eng.*, vol. 34, no. 2, pp. 181–201, 2009.

[24] J. Cioffi, V. Oksman, J. Werner, T. Pollet, P. Spruyt, J. Chow, and K. Jacobsen, "Very-High-Speed Digital Subscriber Lines," *IEEE Communications Magazine*, vol. 37, pp. 72–79, 1999.

[25] G. Ginis and J. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1085–1104, 2002.

[26] R. Cendrillon, G. Ginis, E. Bogaert, and M. Moonen, "A Near-Optimal Linear Crosstalk Canceler for Upstream VDSL," *IEEE Transactions on Signal Processing*, vol. 54, pp. 3136–3146, 2006.

[27] R. Cendrillon, M. Moonen, G. Ginis, K. V. Acker, T. Bostoen, and P. Vandaele, "Partial Crosstalk Cancellation for Upstream VDSL," *EURASIP J. on Applied Signal Process.*, vol. 10, pp. 1520–1535, 2004.

[28] B. Razavi, *RF Microelectronics*. NJ, USA: Prentice Hall, 1998.

[29] F. Horlin and A. Bourdoux, *Digital Compensation for Analog Front-Ends*. England: Wiley, 2008.

[30] A. Tarighat and A. H. Sayed, "Joint compensation of transmitter and receiver impairments in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 240–247, 2007.

[31] M. Valkama, A. Springer, and G. Hueber, "Digital signal processing for reducing the effects of RF imperfections in radio devices - An overview," in *Proceedings of IEEE ISCAS*, 2010.

[32] B. Narasimhan, D. Wang, S. Narayanan, H. Minn, and N. Al-Dhahir, "Digital compensation of frequency-dependent joint Tx/Rx I/Q imbalance in OFDM systems under high mobility," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 3, pp. 405–417, 2009.

[33] D. Tandur and M. Moonen, "Joint adaptive compensation of transmitter and receiver IQ imbalance under carrier frequency offset in OFDM-based systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5246–5252, 2007.

[34] L. Anttila, P. Handel, and M. Valkama, "Joint mitigation of power amplifier and I/Q modulator impairments in broadband direct-conversion transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 4, pp. 730–739, 2010.

[35] A. Yamaoka, K. Yamaguchi, T. Kato, and Y. Tanabe, "Experimental performance evaluation of IQ imbalance and DC offset estimation and compensation technique for 3GPP LTE base station," in *IEEE MTT-S International Microwave Symposium Digest*, 2010.

[36] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "MIMO-OFDM channel estimation in the presence of I/Q imbalance and phase noise for IEEE 802.11n," in *Proceedings of IEEE ICASSP*, 2010.

[37] A. Gomaa, Y. Chi, N. Al-Dhahir, and R. Calderbank, "On training signal design for multi-user MIMO-OFDM: Performance analysis and tradeoffs," in *Proceedings of IEEE VTC-Fall*, 2011.

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

[39] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "A non-iterative technique for phase noise ICI mitigation in packet-based OFDM systems," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5945–5950, 2010.

[40] ——, "Reduced-complexity joint baseband compensation of phase noise and I/Q imbalance for MIMO-OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3450–3460, 2010.

[41] V. Syrjala and M. Valkama, "Receiver DSP for OFDM systems impaired by transmitter and receiver phase noise," in *IEEE International Conference on Communications (ICC)*, 2011.

[42] M. K. Lee, K. Yang, and K. Cheun, "Iterative receivers based on subblock processing for phase noise compensation in OFDM systems," *IEEE Trans. on Communications*, vol. 59, no. 3, pp. 792 –802, 2011.

[43] S. Bittner, A. Frotzscher, G. Fettweis, and E. Deng, "Oscillator phase noise compensation using Kalman tracking," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2009.

[44] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Trans. on Communications*, vol. 55, no. 8, pp. 1607 –1616, 2007.

[45] Y. Chi, A. Gomaa, N. Al-Dhahir, and A. R. Calderbank, "Training signal design and tradeoffs for spectrally-efficient multi-user MIMO-OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 7, pp. 2234–2245, 2011.

[46] I. T. S. Sesia and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*. Wiley, 2009.

[47] H. Myung and D. Goodman, *Single Carrier FDMA: A New Air Interface for Long Term Evolution*. Wiley, 2008.

[48] A. Georgiadis, "Gain, phase imbalance, and phase noise effects on error vector magnitude," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 2, pp. 443–449, 2004.

[49] "Physical channels and modulation," *3GPP TS 36.211*, March 2009.

[50] B. Muquet, Z. Wang, G. Giannakis, M. Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. on Comm.*, vol. 50, pp. 2136–2148, 2002.

[51] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744 –765, 1998.

[52] A. Gomaa and N. Al-Dhahir, "A sparsity-aware approach for NBI estimation in MIMO-OFDM," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1854–1862, 2011.

[53] S. Zhou and G. Giannakis, "Space-time coding with maximum diversity gains over frequency-selective fading channels," *IEEE Signal Processing Letters*, vol. 8, pp. 269–272, 2001.

[54] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Info. Theory*, vol. 52, no. 2, pp. 489 – 509, Feb. 2006.

[55] ——, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. on Pure and Appl. Math*, vol. 59, no. 9, pp. 1207–1223, 2006.

[56] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: cambridge Univ. Press, 2004.

[57] A. Batra, J. Balakrishnan, G. Aiello, J. Foerster, and A. Dabak, "Design of a multiband OFDM system for realistic UWB channel environments," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 9, pp. 2123 – 2138, 2004.

[58] D. Chi and P. Das, "Effects of jammer and nonlinear amplifiers in MIMO-OFDM with application to 802.11n WLAN," in *Proc. IEEE MILCOM 2008*.

[59] W. Bai, C. He, L. Jiang, and X. Li, "Robust channel estimation in MIMO-OFDM systems," *Electronics Letters*, vol. 39, no. 2, pp. 242–244, 2003.

[60] S. Kalyani, V. Raj, and K. Giridhar, "Narrowband interference mitigation in turbo-coded OFDM systems," in *Proc. IEEE ICC 2007*.

[61] R. Gray, *Toeplitz and Circulant Matrices: A review*. Hanover, MA: Now Publishers Inc., 2006.

[62] G. H. Golub and C. F. V. Loan, *Matrix Computations-Third Edition*. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[63] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. on Info. Theory*, vol. 45, pp. 1456 – 1467, 1999.

[64] T. K. Paul and T. Ogunfunmi, "Wireless LAN Comes of Age: Understanding the IEEE 802.11n Amendment," *IEEE Circuits and Systems Magazine*, vol. 8, pp. 28–54, 2008.

[65] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Comm.*, vol. 16, pp. 1451–1458, 1998.

[66] S. M. Kay, *Fundamentals Of Statistical Signal Processing: Estimation Theory*. New Jersey: Prentice Hall, 1993.

[67] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. on Comm.*, vol. 47, pp. 27–32, 1999.

[68] K. Shi, Y. Zhou, B. Kelleci, T. Fischer, E. Serpedin, and A. Karsilayan, "Impacts of narrowband interference on OFDM-UWB receivers: Analysis and mitigation," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 1118–1128, 2007.

[69] A. Gomaa and N. Al-Dhahir, "A new design framework for sparse FIR MIMO equalizers," *IEEE Transactions on Communications*, vol. 59, no. 8, pp. 2132–2140, 2011.

[70] S. Mallat, G. Davis, and Z. Zhang, "Adaptive time-frequency decompositions," *SPIE J. of Optical Eng.*, vol. 33, pp. 2183–2191, 1994.

[71] S. Chen, S. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *International Journal of Control*, vol. 50, no. 5, pp. 1873–1896, 1989.

[72] D. Needell and J. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*.

[73] R. Horn and C. Johnson, *Matrix Analysis*. New York, USA: Cambridge Univ. Press, 2009.

[74] N. Al-Dhahir and J. Cioffi, "MMSE Decision-Feedback Equalizer: Finite-Length Results," *IEEE Trans. Inf. Theory*, vol. 41, pp. 961–975, 1995.

[75] J. Cioffi, *EE379A notes, Chapter 3*. Stanford University, USA.

[76] I. Kalet and S. Shamai, "The Next Challenge for Cellular Networks: Backhaul," *IEEE Microwave Magazine*, vol. 10, pp. 54–66, 2009.

[77] J. Cioffi, "A Multicarrier Primer," ANSI T1E1.4 Committee. Contribution, no. 91-157, Boca Raton, FL, Nov. 1991.

[78] "Guidelines for The Evaluation of Radio Transmission Technologies for IMT-2000," *Recommendation ITU-R M.1225*, 1997.

[79] N. Benvenuto and R. Marchesani, "The Viterbi Algorithm for Sparse Channels," *IEEE Trans. on Commun.*, vol. 44, pp. 287–289, 1996.

[80] I. Kalet and S. Shamai, "On The Capacity of a Twisted-Wire Pair: Gaussian Model," *IEEE Trans. on Commun.*, vol. 38, pp. 379–383, 1990.

[81] M. Sorbara, P. Duvaut, F. Shmulyian, S. Singh, and A. Mahadevan, "Construction of a DSL-MIMO Channel Model for Evaluation of FEXT Cancellation Systems in VDSL2," in *IEEE Sarnoff Symposium*, 2007.

[82] "Very high speed digital subscriber line transceivers 2 (VDSL2)," *ITU-T Recommendation G.993.2*, Feb. 2006.

[83] A. Gomaa and N. Al-Dhahir, "SC-FDMA performance in presence of oscillator impairments: EVM and subcarrier mapping impact," in *IEEE GLOBECOM*, 2011.

[84] ——, "Blind phase noise compensation for SCFDMA with application to LTE-uplink," in *IEEE ICC*, 2012.

[85] T. Schenk, *RF Imperfections in High-rate Wireless Systems: Impact and Digital Compensation*. The Netherlands: Springer, 2008.

[86] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and wiener phase noise," *IEEE Transcations on Communications*, vol. 43, 1995.

[87] J. G. Proakis, *Digital Communication*, 4th ed. New York, NY: McGraw-Hill, 2001.

[88] M. Abramovitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. NY, USA: Dover Publications Inc., 1964.

[89] P. Liu, Y. Bar-Ness, and J. Zhu, "Effects of phase noise at both transmitter and receiver on the performance of OFDM systems," in *Annual Conference on Information Sciences and Systems*, 2006.

[90] D. Petrovic, W. Rave, and G. Fettweis, "Common phase error due to phase noise in OFDM-estimation and suppression," in *IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2004.

[91] "Spatial channel model for mimo simulations," *3GPP TR 25.996*, Sep. 2003.

[92] S. Bittner, E. Zimmermann, and G. Fettweis, "Exploiting phase noise properties in the design of MIMO-OFDM receivers," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2008.

[93] F. Pancaldi, G. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-Carrier Frequency Domain Equalization," *IEEE Signal Processing Magazine*, vol. 25, pp. 37–56, 2008.

## VITA

Ahmad Abdulrahman Gomaa Mohammed received the B.Sc. with honor degree in Electronics and Communications Engineering from Cairo University, Egypt, in 2005 and the M.S. degree in Communications Engineering from Cairo University, Egypt, in 2008. He is currently pursuing the Ph.D. degree at The University of Texas at Dallas, Richardson, Texas, USA.

He worked as a teacher-assistant for Cairo University, Egypt, from 2005 to 2008. Since 2009, he has been with the Broadband Information Transmission and Signaling (BITS) Lab at The University of Texas at Dallas, Richardson, Texas, USA. His research interests include sparse FIR filters design, RF impairments compensation at the baseband, interference cancellation and compressive sensing applications to digital communications.

**Journal Publications:**

1. A. Gomaa and N. Al-Dhahir, "A new Design Framework for Sparse FIR MIMO Equalizers," *IEEE Transactions on Communications*, vol. 59, no. 8, pp.2132-2140, August 2011.
2. A. Gomaa and N. Al-Dhahir, "A Sparsity-Aware Approach for NBI Estimation in MIMO-OFDM," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1854-1862, June 2011.
3. Y. Chi, A. Gomaa, N. Al-Dhahir and R. Calderbank, "Training Signal Design and Tradeoffs for Spectrally-Efficient Multi-User MIMO-OFDM Systems" *IEEE Transactions on Wireless Communications*, vol. 10, no. 7, pp. 2234-2245, July 2011.
4. A. Gomaa and N. Al-Dhahir, "Phase Noise in Asynchronous SC-FDMA Systems: EVM Analysis and Data-Aided Compensation," *Submitted*.
5. A. Gomaa and A. Elezabi, "Iterative MMSE estimation of SINR for turbo coded FH-SS systems in partial-band interference and fading channels" *European Transactions on Telecommunications*, vol. 22, issue. 2, pp. 93-97, March 2011.

6. A. Gomaa and A. Elezabi, "Blind and Semi-Blind Decoder Metrics for Turbo-Coded Frequency-Hopped Spread-Spectrum in Partial-Band Interference and Fading Channels" *International Journal on Wireless Personal Communications*, vol. 52, no. 4, pp. 735-751, March 2010.

7. A. Elezabi and A. Gomaa, "Diversity Combining and SNR Estimation for Turbo-Coded Frequency-Hopped Spread-Spectrum in Partial-Band Interference and Fading Channels" *International Journal on Wireless Personal Communications*, vol. 50, no. 3, pp. 351-365, August 2009.

**Conference Publications:**

1. A. Gomaa, M. Mokhtar and N. Al-Dhahir, "Amplify-and-Forward Relaying under I/Q Imbalance," *IEEE Global Telecommunications Conference (GLOBECOM), 2012*, (*accepted*).

2. A. ElSamadouny, A. Gomaa and N. Al-Dhahir, "Likelihood-Based Spectrum Sensing of OFDM Signals in the Presence of Tx/Rx IQ Imbalance," *IEEE Global Telecommunications Conference (GLOBECOM), 2012*, (*accepted*).

3. F. Rusek, N. Al-Dhahir and A. Gomaa, "A Rate-Maximizing Channel Shortening Detector with Side Information Soft Feedback," *IEEE Global Telecommunications Conference (GLOBECOM), 2012*, (*accepted*).

4. A. Gomaa and N. Al-Dhahir, "Blind Phase Noise Compensation for SCFDMA with Application to LTE-Uplink," *IEEE International Conference on Communications (ICC), 2012.*

5. A. Gomaa and N. Al-Dhahir, "SC-FDMA Performance in Presence of Oscillator Impairments: EVM and Subcarrier Mapping Impact," *IEEE Global Telecommunications Conference (GLOBECOM), 2011.*

6. A. Gomaa and N. Al-Dhahir, "Multi-User SC-FDMA Systems Under IQ Imbalance: EVM and Subcarrier Mapping Impact," *IEEE Global Telecommunications Conference (GLOBECOM), 2011.*

7. A. Gomaa and N. Al-Dhahir, "Compressive-Sensing-Based Approach for NBI Cancellation in MIMO-OFDM," *IEEE Global Telecommunications Conference (GLOBECOM), 2011.*

8. A. Gomaa and N. Al-Dhahir, "Sparse FIR Equalization: A New Design Framework," *IEEE Vehicular Technology Conference (VTC-fall), 2011.*

9. A. Gomaa, Y. Chi, N. Al-Dhahir and R. Calderbank, "On Training Signal Design for Multi-User MIMO-OFDM: Performance Analysis and Tradeoffs," *IEEE Vehicular Technology Conference (VTC-fall), 2011.*

10. Y. Chi, A. Gomaa, N. Al-Dhahir and R. Calderbank, "MMSE-Optimal Training Sequences For Spectrally-Efficient Multi-User MIMO-OFDM Systems," *IEEE European Signal Processing Conference (EUSIPCO), 2011.*

11. A. Gomaa and N. Al-Dhahir, "A Compressive Sensing Approach to NBI Cancellation in Mobile OFDM Systems," *IEEE Global Telecommunications Conference (GLOBECOM), 2010.*

12. A. Gomaa and N. Al-Dhahir, "Low-Complexity Sparse FIR Channel Shortening," *IEEE Global Telecommunications Conference (GLOBECOM), 2010.*

13. A. Gomaa, K. M. Z.Islam and N. Al-Dhahir, "Two novel compressed-sensing algorithms for NBI detection in OFDM systems," *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), 2010.*

14. A. Gomaa, M. Nafie and M. Abdallah, "Novel Reliability-Based Hybrid ARQ Technique," *IEEE Global Telecommunications Conference (GLOBECOM), 2009.*

15. A. Elezabi and A. Gomaa, "Diversity combining and SNR estimation for turbo-coded Frequency-Hopped Spread-Spectrum in partial-band interference and fading channels," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2008.*