

A Sparsity-Aware Approach for NBI Estimation in MIMO-OFDM

Ahmad Gomaa, *Student Member, IEEE*, and Naofal Al-Dhahir, *Fellow, IEEE*

Abstract—In this paper, we present a novel approach based on compressive sensing theory 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 and approaching the interference-free performance limit over practical ranges of NBI power levels, spectral widths, and mobility levels.

Index Terms—Narrow-band interference (NBI), OFDM, compressive sensing, sparsity, MIMO.

I. INTRODUCTION

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 [1], which employs multiple-input multiple-output (MIMO) and OFDM technologies, NBI is generated by Bluetooth devices operating in the same band [2]. The same problem also arises in other WLAN 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 paper. Among the NBI mitigation techniques, the one in [3] 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 [4] as an erasure

insertion mechanism which localizes the erasures to the subcarriers surrounding the NBI without affecting the remaining subcarriers. In [3] and [4], only a single-tone NBI is assumed. However, in this paper, we consider NBI affecting several OFDM subcarriers. In [5], 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 [6] 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 [6]. In [7], 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 [8], 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 [9] 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 [10] showed how to reconstruct a sparse vector from insufficient noisy measurements. In this paper, we exploit the inherent sparsity of the NBI signals in the frequency-domain (FD) and use CS theory convex optimization techniques to estimate them. In the next section, we provide a review of CS theory principles and describe the system model. We apply our proposed NBI estimation approach to single-input single-output (SISO) zero-padded (ZP) OFDM systems in Section III. The extension to MIMO systems is presented in Section IV. In the same section, we address the problem of channel estimation in the presence of NBI. In Section V, 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

Manuscript received June 23, 2010; revised December 13, 2010 and February 19, 2011; accepted March 8, 2011. The associate editor coordinating the review of this paper and approving it for publication was H. Nguyen.

The authors are with The University of Texas at Dallas, 2601 N. Floyd Road, Richardson, Texas 75080, USA (e-mail: aag083000@utdallas.edu).

This work was supported by RIM Inc. and Cisco Inc. and was presented in part at ICASSP 2010 and Globecom 2010.

Digital Object Identifier 10.1109/TWC.2011.101118.

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 and conclusions are presented in Sections VI and VII, respectively.

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 their sizes. The matrix $\mathbf{0}_{m \times n}$ denotes the all-zero matrix of size $m \times n$ and the notation $\text{diag}(x_1, \dots, x_N)$ denotes an $N \times N$ diagonal matrix whose diagonal elements are $\{x_1, \dots, x_N\}$. Also, $(\cdot)^H$, $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^{-1}$ 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 [11] where the k^{th} element in the vector $\mathbf{P}_P^n \mathbf{a}$ is $a((P - k + n) \bmod P)$.

II. CS THEORY BACKGROUND AND SYSTEM MODEL

A. Compressive Sensing Theory Background

CS theory [12] asserts that we can efficiently recover a sparse vector $\mathbf{x} \in \mathbb{C}^N$ from a noisy measurement vector $\mathbf{y} \in \mathbb{C}^M$ where $M \ll N$. In other words, the solution of the noisy under-determined system of equations $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{z}$ can be computed under some conditions on \mathbf{A} where \mathbf{A} denotes the $M \times N$ measurement matrix and $\mathbf{z} \in \mathbb{C}^M$ is a bounded noise vector. 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

$$\min_{\tilde{\mathbf{x}} \in \mathbb{C}^N} \|\tilde{\mathbf{x}}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{A}\tilde{\mathbf{x}}\|_2^2 \leq \epsilon \quad (1)$$

where $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the l_1 -norm and the l_2 -norm, respectively, and the parameter ϵ is chosen such that it bounds the amount of noise in the measurements. In fact, the convex optimization problem in (1) is a second-order cone program and can be solved efficiently [13].

B. 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

$$\mathbf{y} = \mathbf{H} \underbrace{[\mathbf{F}_N \quad \mathbf{0}_{N \times \nu}]^H}_{:= \mathbf{F}_{zp}} \mathbf{X} + \mathbf{\Lambda}_{f_0} \mathbf{H}_J \mathbf{j} + \mathbf{z} \quad (2)$$

where \mathbf{X} , \mathbf{j} and \mathbf{z} denote, respectively, the 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}_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 ν denoting the guard sequence length. Instead of cyclic prefixing, we use zero-padding [14] (implied by the precoding matrix \mathbf{F}_{zp}) to

eliminate inter-block interference (IBI) in the guard sequence. ZP-OFDM is adopted in multi-band OFDM ultra-wide-band (UWB) [15] 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}_{zp} , we can replace it with a circulant matrix without changing the received OFDM signal [14]. The first column of this circulant matrix is the zero-padded channel impulse response (CIR) coefficients vector denoted by $[h_0 \ h_1 \ \dots \ h_{L-1} \ \mathbf{0}_{1 \times (P-L)}]^T$ where h_k is the k^{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}_J is a $P \times P$ Toeplitz matrix whose first column is the zero-padded NBI CIR vector. Furthermore, $\mathbf{\Lambda}_{f_0} = \text{diag}\left(1, \exp\left(i\frac{2\pi\alpha}{P}\right), \dots, \exp\left(i\frac{2\pi\alpha(P-1)}{P}\right)\right)$ where $i \triangleq \sqrt{-1}$, $\exp(\cdot)$ is the exponential function, and α is a uniformly-distributed random variable [4] over the interval $[-\frac{1}{2}, \frac{1}{2}]$. The variable α is used, as in [4] and [16], 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 [17] and [18] 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^{th} element, $J(k)$, represents the NBI signal at the k^{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.

C. 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}_J and the frequency offset [4] 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

$$\mathbf{Y} = \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{y} = \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{H} \mathbf{F}_{zp} \mathbf{X} + \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{\Lambda}_{f_0} \mathbf{H}_J \mathbf{F}_P^H \mathbf{J} + \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{z} \quad (3)$$

where $\mathbf{\Lambda}_W = \text{diag}(w(0), w(1), \dots, w(P-1))$ and $w(n)$ is the n^{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 FFT of the first column of \mathbf{H} . For large values of P , \mathbf{H}_J can be approximated as a circulant¹ matrix [19] and, hence, expressed as $\mathbf{H}_J = \mathbf{F}_P^H \mathbf{\Lambda}_J \mathbf{F}_P$ where $\mathbf{\Lambda}_J$ is a diagonal matrix whose diagonal is the P -point FFT of the first column of \mathbf{H}_J . Therefore, we re-write (3) as follows

$$\mathbf{Y} = \underbrace{\mathbf{H}_W \mathbf{\Lambda} \mathbf{V} \mathbf{X}}_{\triangleq \tilde{\mathbf{A}}} + \underbrace{\mathbf{H}_{\text{eqv}} \mathbf{\Lambda}_J \mathbf{J}}_{\triangleq \mathbf{J}_{\text{eqv}}} + \underbrace{\mathbf{F}_P \mathbf{\Lambda}_W \mathbf{Z}}_{\mathbf{Z}} = \tilde{\mathbf{A}}_{P \times N} \mathbf{X} + \mathbf{J}_{\text{eqv}} + \mathbf{Z} \quad (4)$$

where $\mathbf{H}_W \triangleq \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{F}_P^H$ and $\mathbf{H}_{\text{eqv}} \triangleq \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{\Lambda}_{f_0} \mathbf{F}_P^H$ are circulant matrices, and $\mathbf{V} \triangleq \mathbf{F}_P \mathbf{F}_{zp}$. Since $\mathbf{\Lambda}_J$ is a diagonal matrix, it does not affect the sparsity of \mathbf{J} ; hence, $\mathbf{\Lambda}_J \mathbf{J}$ 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}_{eqv} such that the equivalent NBI vector, \mathbf{J}_{eqv} , is spectrally contained, i.e. sparse.

III. 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 (4), the sparse FD NBI vector \mathbf{J}_{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 (4) first. This technique is known as signal-blocking in the context of array processing. Assuming knowledge of $\tilde{\mathbf{A}}$ and, hence, $\tilde{\mathbf{A}}$ at the receiver, we filter \mathbf{Y} by \mathbf{W} such that $\mathbf{W} \tilde{\mathbf{A}} = \mathbf{0}$. To this end, \mathbf{W} is designed to be the projection matrix on the left null-subspace of $\tilde{\mathbf{A}}$ as follows²

$$\mathbf{W} = \mathbf{I}_P - \tilde{\mathbf{A}} \tilde{\mathbf{A}}^\dagger \quad (5)$$

where $\tilde{\mathbf{A}}^\dagger \triangleq (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H$ is the Moore–Penrose pseudo-inverse [20]. Since $\tilde{\mathbf{A}}$ 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 $\tilde{\mathbf{A}}$ a tall matrix. Multiplying \mathbf{Y} by \mathbf{W} , we get

$$\tilde{\mathbf{Y}} = \mathbf{W} \mathbf{Y} = \mathbf{W} \mathbf{J}_{\text{eqv}} + \mathbf{W} \mathbf{Z} \triangleq \mathbf{W} \mathbf{J}_{\text{eqv}} + \tilde{\mathbf{Z}} \quad (6)$$

Note that \mathbf{W} is a rank- ν matrix; hence, the linear system in (6) is under-determined and falls under CS theory framework. Therefore, we formulate and solve the following convex program to estimate \mathbf{J}_{eqv}

$$\min_{\tilde{\mathbf{J}}_{\text{eqv}} \in \mathbb{C}^P} \|\tilde{\mathbf{J}}_{\text{eqv}}\|_1 \quad \text{subject to} \quad \|\tilde{\mathbf{Y}} - \mathbf{W} \tilde{\mathbf{J}}_{\text{eqv}}\|_2^2 \leq \epsilon \quad (7)$$

where ϵ is a design parameter chosen high enough to bound $\|\tilde{\mathbf{Z}}\|_2^2$ with high probability. Next, we cancel $\tilde{\mathbf{J}}_{\text{eqv}}$ from \mathbf{Y} , remove the window, and proceed to the data detection and decoding stage.

¹However, in simulations, we generate \mathbf{H}_J as a Toeplitz matrix.

²Note 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.

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 $\tilde{\mathbf{J}} = \mathbf{\Lambda}^{-1} \mathbf{\Lambda}_J \mathbf{J}$ 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 ν equations with r unknowns because \mathbf{W} is a rank- ν 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 ν (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, e.g. by $[\mathbf{I}_N \ \mathbf{0}_{N \times \nu}]^H$, then this assumption is violated as the coefficients matrix would be $[\mathbf{0}_{\nu \times N} \ \mathbf{I}_\nu]$ that has N columns with all-zero entries.

IV. 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 [21] to larger numbers of transmit antennas; for example four transmit antennas as in the IEEE 802.11n standard [22]. We start by developing a space time block coding (STBC) scheme for ZP-OFDM systems.

A. STBC for ZP-OFDM

Inspired by the Alamouti STBC scheme [23] and the STBC scheme of ZP single-carrier transmission [11], we propose the following transmit diversity scheme for ZP-OFDM with 2 transmit antennas. The coded symbols are divided into $N \times 1$ blocks $\{\mathbf{X}(k), k = 0, 1, \dots\}$. Denoting the transmitted OFDM block at time k from the l^{th} antenna by $\mathbf{x}_l(k)$, we write the outputs of the 2 transmit antennas at times 0 and 1 as follows

$$\begin{aligned} \mathbf{x}_1(0) &= \mathbf{F}_{zp} \mathbf{X}(0), \quad \mathbf{x}_2(0) = \mathbf{F}_{zp} \mathbf{X}(1), \\ \mathbf{x}_1(1) &= -\mathbf{F}_{zp} \mathbf{P}_{tx} \mathbf{X}^*(1), \quad \mathbf{x}_2(1) = \mathbf{F}_{zp} \mathbf{P}_{tx} \mathbf{X}^*(0) \end{aligned} \quad (8)$$

where \mathbf{P}_{tx} is a precoding matrix introduced to make the equivalent MIMO channel matrix in the form of the Quaternion matrix³ 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

³A Quaternion matrix has the form $\begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}$

NBI-free received OFDM blocks at times 0 and 1 at the q^{th} receive antenna as follows

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

where $1 \leq q \leq N_r$ with N_r denoting the number of receive antennas, $\mathbf{z}^q(k)$ is the complex AWGN vector of the q^{th} receive antenna at time k , and \mathbf{H}_l^q is the $P \times P$ circulant channel matrix between the l^{th} transmit antenna and the q^{th} receive antenna. We remove the window after estimating the NBI, so we do not consider it while designing the STBC scheme. It was shown in [11] 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 I. Negating, conjugating, and multiplying $\mathbf{y}^q(1)$ by $\mathbf{P}_P^{(n)}$, we get

$$\begin{aligned} \hat{\mathbf{y}}^q(1) &\triangleq -\mathbf{P}_P^{(n)} (\mathbf{y}^q(1))^* = \mathbf{P}_P^{(n)} (\mathbf{H}_1^q)^* \mathbf{F}_{zp}^* \mathbf{P}_{tx}^* \mathbf{X}(1) \\ &\quad - \mathbf{P}_P^{(n)} (\mathbf{H}_2^q)^* \mathbf{F}_{zp}^* \mathbf{P}_{tx}^* \mathbf{X}(0) - \mathbf{P}_P^{(n)} (\mathbf{z}^q(1))^* \end{aligned} \quad (9)$$

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

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

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 FFTs of $\mathbf{y}^q(0)$ and $\hat{\mathbf{y}}^q(1)$ in a single vector, we get

$$\mathbf{Y}^q \triangleq \begin{bmatrix} \mathbf{F}_P \mathbf{y}^q(0) \\ \mathbf{F}_P \hat{\mathbf{y}}^q(1) \end{bmatrix} = \underbrace{\begin{bmatrix} \Lambda_1^q & \Lambda_2^q \\ -(\Lambda_2^q)^H & (\Lambda_1^q)^H \end{bmatrix}}_{\triangleq \Lambda^q} \begin{bmatrix} \mathbf{V} \mathbf{X}(0) \\ \mathbf{V} \mathbf{X}(1) \end{bmatrix} + \begin{bmatrix} \mathbf{Z}^q(0) \\ \mathbf{Z}^q(1) \end{bmatrix} \quad (10)$$

where $\Lambda_l^q \triangleq \mathbf{F}_P \mathbf{H}_l^q \mathbf{F}_P^H$ is a diagonal matrix whose diagonal is the P -point FFT 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 MIMO channel matrix, Λ^q , is a Quaternion as desired.

B. 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 FFT 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

$$\begin{aligned} \mathbf{Y}(0) &= \mathbf{F}_P \Lambda_W \mathbf{y}(0) = \mathbf{H}_W \Lambda_1 \mathbf{V} \mathbf{X}(0) + \mathbf{H}_W \Lambda_2 \mathbf{V} \mathbf{X}(1) \\ &\quad + \mathbf{J}_{\text{eqv}} + \mathbf{Z}(0) \\ \mathbf{Y}(1) &= -\mathbf{F}_P \mathbf{P}_P^{(N)} \Lambda_W \mathbf{y}^*(1) = -\mathbf{H}'_W \Lambda_2^H \mathbf{V} \mathbf{X}(0) \\ &\quad + \mathbf{H}'_W \Lambda_1^H \mathbf{V} \mathbf{X}(1) + \tilde{\mathbf{L}} \mathbf{J}_{\text{eqv}}^* + \mathbf{Z}(1) \end{aligned} \quad (11)$$

where $\tilde{\mathbf{L}} = -\mathbf{F}_P \mathbf{P}_P^{(N)} \mathbf{F}_P^T$, $\mathbf{H}'_W = \mathbf{F}_P \mathbf{P}_P^{(N)} \Lambda_W \mathbf{P}_P^{(N)} \mathbf{F}_P^H$, $\mathbf{Z}(0) = \mathbf{F}_P \Lambda_W \mathbf{z}(0)$, and $\mathbf{Z}(1) = -\mathbf{F}_P \mathbf{P}_P^{(N)} \Lambda_W \mathbf{z}^*(1)$. Thanks

to $\mathbf{P}_P^{(N)}$, the matrix $\tilde{\mathbf{L}}$ is diagonal and \mathbf{H}'_W is circulant. Stacking $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$ into a single vector, we write

$$\begin{aligned} \mathbf{Y} \triangleq \begin{bmatrix} \mathbf{Y}(0) \\ \mathbf{Y}(1) \end{bmatrix} &= \underbrace{\begin{bmatrix} \mathbf{H}_W \Lambda_1 \mathbf{V} & \mathbf{H}_W \Lambda_2 \mathbf{V} \\ -\mathbf{H}'_W \Lambda_2^H \mathbf{V} & \mathbf{H}'_W \Lambda_1^H \mathbf{V} \end{bmatrix}}_{\triangleq \tilde{\mathbf{A}}} \underbrace{\begin{bmatrix} \mathbf{X}(0) \\ \mathbf{X}(1) \end{bmatrix}}_{\triangleq \mathbf{X}} \\ &\quad + \underbrace{\begin{bmatrix} \mathbf{I}_P & \mathbf{0}_{P \times P} \\ \mathbf{0}_{P \times P} & \tilde{\mathbf{L}} \end{bmatrix}}_{\triangleq \tilde{\mathbf{L}}} \underbrace{\begin{bmatrix} \mathbf{J}_{\text{eqv}} \\ \mathbf{J}_{\text{eqv}}^* \end{bmatrix}}_{\triangleq \mathbf{J}} + \underbrace{\begin{bmatrix} \mathbf{Z}(0) \\ \mathbf{Z}(1) \end{bmatrix}}_{\triangleq \mathbf{Z}} \end{aligned} \quad (12)$$

For a better CS-based estimate of \mathbf{J} , we must cancel the unknown data term \mathbf{X} in (12) as we did in Section III. Assuming knowledge⁴ of $\tilde{\mathbf{A}}$ at the receiver, we filter \mathbf{Y} by \mathbf{W} such that $\mathbf{W} \tilde{\mathbf{A}} = \mathbf{0}$. To this end, \mathbf{W} is chosen to be the projection matrix on the left null-subspace of $\tilde{\mathbf{A}}$ given by

$$\mathbf{W} = \mathbf{I}_{2P} - \tilde{\mathbf{A}} \tilde{\mathbf{A}}^\dagger \quad (13)$$

Note that $\tilde{\mathbf{A}}$ 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 $\tilde{\mathbf{A}}$ a tall matrix. Multiplying \mathbf{Y} by \mathbf{W} , we get

$$\begin{aligned} \tilde{\mathbf{Y}} \triangleq \mathbf{W} \mathbf{Y} &= \underbrace{\mathbf{W} \tilde{\mathbf{A}}}_{\triangleq \tilde{\mathbf{W}}} \underbrace{\begin{bmatrix} \mathbf{I}_P & \mathbf{I}_P \end{bmatrix}^H}_{\triangleq \mathbf{I}_1} \mathbf{J}_{\text{eqv},R} \\ &\quad + i \underbrace{\mathbf{W} \tilde{\mathbf{L}}}_{\triangleq \tilde{\mathbf{L}}_2} \underbrace{\begin{bmatrix} \mathbf{I}_P & -\mathbf{I}_P \end{bmatrix}^H}_{\triangleq \mathbf{I}_2} \mathbf{J}_{\text{eqv},I} + \underbrace{\mathbf{W} \mathbf{Z}}_{\triangleq \tilde{\mathbf{Z}}} \end{aligned} \quad (14)$$

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

$$\begin{aligned} \tilde{\mathbf{Y}}_R &= \begin{bmatrix} \tilde{\mathbf{W}}_R \mathbf{I}_1 & -\tilde{\mathbf{W}}_I \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{J}_{\text{eqv},R}^T & \mathbf{J}_{\text{eqv},I}^T \end{bmatrix}^T + \tilde{\mathbf{Z}}_R, \\ \tilde{\mathbf{Y}}_I &= \begin{bmatrix} \tilde{\mathbf{W}}_I \mathbf{I}_1 & \tilde{\mathbf{W}}_R \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{J}_{\text{eqv},R}^T & \mathbf{J}_{\text{eqv},I}^T \end{bmatrix}^T + \tilde{\mathbf{Z}}_I \end{aligned} \quad (15)$$

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

$$\tilde{\mathbf{Y}}_T = \underbrace{\begin{bmatrix} \tilde{\mathbf{W}}_R \mathbf{I}_1 & -\tilde{\mathbf{W}}_I \mathbf{I}_2 \\ \tilde{\mathbf{W}}_I \mathbf{I}_1 & \tilde{\mathbf{W}}_R \mathbf{I}_2 \end{bmatrix}}_{\triangleq \mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{J}_{\text{eqv},R} \\ \mathbf{J}_{\text{eqv},I} \end{bmatrix}}_{\triangleq \mathbf{J}_T} + \underbrace{\begin{bmatrix} \tilde{\mathbf{Z}}_R \\ \tilde{\mathbf{Z}}_I \end{bmatrix}}_{\triangleq \tilde{\mathbf{Z}}_T} \quad (16)$$

Since \mathbf{W} is a rank-deficient matrix (its rank is 2ν), the linear system described by (16) represents an under-determined linear system of equations. Hence, a convex optimization program similar to those in (1) and (7) 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.

C. 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 (5) and (13). 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 2×1 systems; however, our approach can also be extended to other MIMO configurations. The idea is to

⁴Channel estimation is investigated in Section IV-C.

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

$$\mathbf{x}_1(0) = \mathbf{x}_2(0) = \mathbf{F}_{zp}\mathbf{m}_0, \quad \mathbf{x}_1(1) = -\mathbf{x}_2(1) = \mathbf{F}_{zp}\mathbf{m}_1 \quad (17)$$

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 (8) 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

$$\begin{aligned} \mathbf{y}(0) &= \mathbf{H}_1\mathbf{F}_{zp}\mathbf{m}_0 + \mathbf{H}_2\mathbf{F}_{zp}\mathbf{m}_0 + \mathbf{\Lambda}_{f0}\mathbf{H}_J\mathbf{j} + \mathbf{z}(0), \\ \mathbf{y}(1) &= \mathbf{H}_1\mathbf{F}_{zp}\mathbf{m}_1 - \mathbf{H}_2\mathbf{F}_{zp}\mathbf{m}_1 + \mathbf{\Lambda}_{f0}\mathbf{H}_J\mathbf{j} + \mathbf{z}(1) \end{aligned} \quad (18)$$

Since $\{\mathbf{H}_k, k = 1, 2\}$ are circulant matrices, we only need to estimate their first columns. However, the channel length is unlikely to exceed the guard sequence length ν . Then, we only need to estimate \mathbf{h}_k which denotes the first ν 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, (18) is re-written as follows

$$\begin{aligned} \mathbf{y}(0) &= \mathbf{M}_0(\mathbf{h}_1 + \mathbf{h}_2) + \mathbf{\Lambda}_{f0}\mathbf{H}_J\mathbf{j} + \mathbf{z}(0), \\ \mathbf{y}(1) &= \mathbf{M}_1(\mathbf{h}_1 - \mathbf{h}_2) + \mathbf{\Lambda}_{f0}\mathbf{H}_J\mathbf{j} + \mathbf{z}(1) \end{aligned} \quad (19)$$

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

$$\begin{aligned} \mathbf{Y}(0) &\triangleq \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{y}(0) = \tilde{\mathbf{M}}_0(\mathbf{h}_1 + \mathbf{h}_2) + \mathbf{J}_{\text{eqv}} + \mathbf{Z}(0), \\ \mathbf{Y}(1) &\triangleq \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{y}(1) = \tilde{\mathbf{M}}_1(\mathbf{h}_1 - \mathbf{h}_2) + \mathbf{J}_{\text{eqv}} + \mathbf{Z}(1) \end{aligned} \quad (20)$$

where $\tilde{\mathbf{M}}_0 = \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{M}_0, \tilde{\mathbf{M}}_1 = \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{M}_1, \mathbf{Z}(0) = \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{z}(0), \mathbf{Z}(1) = \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{z}(1)$, and $\mathbf{J}_{\text{eqv}} = \mathbf{F}_P \mathbf{\Lambda}_W \mathbf{\Lambda}_{f0} \mathbf{H}_J \mathbf{j}$. Since $\tilde{\mathbf{M}}_0$ and $\tilde{\mathbf{M}}_1$ are tall matrices and known at the receiver, 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 terms containing the unknown CIR vectors as follows

$$\tilde{\mathbf{Y}}(k) = \mathbf{A}_k \mathbf{Y}(k) = \mathbf{A}_k \mathbf{J}_{\text{eqv}} + \tilde{\mathbf{Z}}(k), \quad k = 0, 1 \quad (21)$$

where $\mathbf{A}_k = (\mathbf{I}_P - \tilde{\mathbf{M}}_k \tilde{\mathbf{M}}_k^\dagger)$ and $\tilde{\mathbf{Z}}(k) = \mathbf{A}_k \mathbf{Z}(k)$. Observing that (21) represents two under-determined linear systems of equations, we recover \mathbf{J}_{eqv} from either one by formulating and solving a convex l_1 -norm minimization program as before. Note that the transmission scheme in (17) makes both \mathbf{h}_1 and \mathbf{h}_2 multiplied by the same matrix in (19) 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 fixed and known at the receiver. Next, we cancel the NBI estimate from $\mathbf{Y}(0)$ and $\mathbf{Y}(1)$, convert back to TD, and remove the window to get

$$\begin{aligned} \mathbf{y}_c(0) &= \mathbf{M}_0(\mathbf{h}_1 + \mathbf{h}_2) + \tilde{\mathbf{z}}(0) \triangleq \mathbf{M}_0 \mathbf{h}_a + \tilde{\mathbf{z}}(0), \\ \mathbf{y}_c(1) &= \mathbf{M}_1(\mathbf{h}_1 - \mathbf{h}_2) + \tilde{\mathbf{z}}(1) \triangleq \mathbf{M}_1 \mathbf{h}_s + \tilde{\mathbf{z}}(1) \end{aligned} \quad (22)$$

where $\tilde{\mathbf{z}}(k) = \mathbf{z}(k) + \mathbf{j}_{\text{res}}$, and \mathbf{j}_{res} is the residual NBI after cancellation. The linear least-squares (LLS) estimators [24]

of \mathbf{h}_a and \mathbf{h}_s have the following closed forms

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

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

$$\hat{\mathbf{h}}_1 = (\hat{\mathbf{h}}_a + \hat{\mathbf{h}}_s)/2, \quad \hat{\mathbf{h}}_2 = (\hat{\mathbf{h}}_a - \hat{\mathbf{h}}_s)/2 \quad (23)$$

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.

V. EXTENSIONS

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} matrix used to cancel out the unknown data term. Furthermore, we extend our approach to CP-OFDM.

A. Mobile NBI

In some 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}_J is no longer Toeplitz and cannot be approximated as circulant. Therefore, $\mathbf{\Lambda}_J = \mathbf{F}_P \mathbf{H}_J \mathbf{F}_P^H$ is no longer diagonal; however, as shown in [25], 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} , $\mathbf{\Lambda}_J$ causes the FD NBI signal to spill over the D adjacent subcarriers from each side. However, the window, that we apply to the received signal, reduces this spectral leakage such that the overall NBI vector, \mathbf{J}_{eqv} , is still sparse. Nevertheless, \mathbf{J}_{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 IV-B since the NBI channel actually varies within each OFDM symbol. For mobile NBI, (12) becomes

$$\mathbf{Y} = \tilde{\mathbf{A}} \mathbf{X} + \mathbf{L} [\mathbf{J}_{\text{eqv},0}^T \quad \mathbf{J}_{\text{eqv},1}^H]^T + \mathbf{Z} \triangleq \tilde{\mathbf{A}} \mathbf{X} + \mathbf{L} \mathbf{J}_T + \mathbf{Z} \quad (24)$$

where \mathbf{L} was defined in (12) 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 (13) and applying it to \mathbf{Y} , we get

$$\tilde{\mathbf{Y}} = \mathbf{W} \mathbf{Y} = \mathbf{W} \mathbf{L} \mathbf{J}_T + \mathbf{W} \mathbf{Z} \triangleq \mathbf{A} \mathbf{J}_T + \tilde{\mathbf{Z}} \quad (25)$$

CS-based techniques can be used to estimate \mathbf{J}_T from (25).

B. 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, (12) becomes

$$\mathbf{Y} = \tilde{\mathbf{A}}\mathbf{X} + \mathbf{Z} + \mathbf{L} \underbrace{\left[\sum_{l=1}^{N_J} \left(\mathbf{J}_{\text{eqv},0}^l \right)^T \quad \sum_{l=1}^{N_J} \left(\mathbf{J}_{\text{eqv},1}^l \right)^H \right]^T}_{\triangleq \mathbf{J}_T} \quad (26)$$

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^{th} source over OFDM symbols 0 and 1, respectively. From (26), 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.

C. 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 (5) and (13) require the computation of the pseudo-inverse matrix which is computationally complex. The idea is to replace $\tilde{\mathbf{A}}^\dagger$ in (5) and (13) by a product of easy-to-compute matrices such that this product is equal to the identity matrix when multiplied by $\tilde{\mathbf{A}}$ which is the same role of the matrix $\tilde{\mathbf{A}}^\dagger$. Investigating the structure of $\tilde{\mathbf{A}}$ defined in (12), we propose the following alternative \mathbf{W} design for the MIMO⁵ case

$$\mathbf{W} = \mathbf{I}_{2P} - \tilde{\mathbf{A}} \begin{bmatrix} \mathbf{V}^H & \mathbf{0}_{N \times P} \\ \mathbf{0}_{N \times P} & \mathbf{V}^H \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_t^{-1} & \mathbf{0}_{P \times P} \\ \mathbf{0}_{P \times P} & \mathbf{\Lambda}_t^{-1} \end{bmatrix} \quad (27)$$

$$\times \begin{bmatrix} \mathbf{\Lambda}_1^H & -\mathbf{\Lambda}_2 \\ \mathbf{\Lambda}_2^H & \mathbf{\Lambda}_1 \end{bmatrix} \begin{bmatrix} \mathbf{H}_W^{-1} & \mathbf{0}_{P \times P} \\ \mathbf{0}_{P \times P} & (\mathbf{H}_W')^{-1} \end{bmatrix}$$

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, $\mathbf{H}_W^{-1} = \mathbf{F}_P \mathbf{\Lambda}_W^{-1} \mathbf{F}_P^H$ and $(\mathbf{H}_W')^{-1} = \mathbf{F}_P \mathbf{P}_P^{(N)} \mathbf{\Lambda}_W^{-1} \mathbf{P}_P^{(N)} \mathbf{F}_P^H$ where $\mathbf{\Lambda}_W^{-1}$ is easily computed since $\mathbf{\Lambda}_W$ 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^{th} column of \mathbf{W} is computed by filtering the k^{th} column of \mathbf{I}_{2P} by \mathbf{W} . Assuming that P is a power of 2, the numbers of real multiplications and additions needed to filter a length- $2P$ vector by the reduced-complexity design of \mathbf{W} are $O(40P + 24P \log_2 P)$ and $O(28P + 24P \log_2 P)$, respectively.

D. 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 [1], for example, \mathbf{X} contains 12 (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 IV-B, the extension to MIMO systems becomes straightforward with the only difference being that the STBC scheme would be the conventional Alamouti [23] scheme. The received CP-OFDM symbol after CP removal can be written as follows

$$\mathbf{y} = \mathbf{H}\mathbf{F}_N^H \mathbf{Q} \mathbf{X} + \mathbf{\Lambda}_{\text{fo}} \mathbf{H}_J \mathbf{j} + \mathbf{z} \quad (28)$$

where \mathbf{H} and \mathbf{H}_J are the $N \times N$ channel matrices of the desired and NBI signals, respectively, \mathbf{Q} is a unitary precoding matrix, $\mathbf{\Lambda}_{\text{fo}} = \text{diag} \left(1, \exp \left(i \frac{2\pi\alpha}{N} \right), \dots, \exp \left(i \frac{2\pi\alpha(N-1)}{N} \right) \right)$, 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

$$\mathbf{y}_1 = \mathbf{y} - \mathbf{H}\mathbf{F}_N^H \mathbf{Q}_K \mathbf{X}_K = \mathbf{H}\mathbf{F}_N^H \mathbf{Q}_U \mathbf{X}_U + \mathbf{\Lambda}_{\text{fo}} \mathbf{H}_J \mathbf{j} + \mathbf{z} \quad (29)$$

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 FFT, we get

$$\mathbf{Y}_1 \triangleq \mathbf{F}_N \mathbf{\Lambda}_W \mathbf{y}_1 = \mathbf{F}_N \mathbf{\Lambda}_W \mathbf{H}\mathbf{F}_N^H \mathbf{Q}_U \mathbf{X}_U + \mathbf{F}_N \mathbf{\Lambda}_W \mathbf{\Lambda}_{\text{fo}} \mathbf{H}_J \mathbf{j} + \mathbf{F}_N \mathbf{\Lambda}_W \mathbf{z} \triangleq \tilde{\mathbf{A}} \mathbf{X}_U + \mathbf{J}_{\text{eqv}} + \mathbf{Z} \quad (30)$$

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

$$\tilde{\mathbf{Y}}_1 = \mathbf{WY}_1 = \mathbf{WJ}_{\text{eqv}} + \tilde{\mathbf{Z}} \quad (31)$$

where $\tilde{\mathbf{Z}} = \mathbf{WZ}$. Since the rank of \mathbf{W} is $u < N$, the linear system in (31) is under-determined. Hence, we formulate and solve an l_1 -norm minimization program similar to those in (1) and (7). Then, we estimate \mathbf{J}_{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 $\mathbf{W} \triangleq \mathbf{I}_N - \tilde{\mathbf{A}} \mathbf{Q}_U^H \mathbf{F}_N \mathbf{H}^{-1} \mathbf{\Lambda}_W^{-1} \mathbf{F}_N^H$ 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 IV-B to increase the number of measurements and improve the performance.

VI. SIMULATION RESULTS

We simulate the performance of a 2×2 MIMO-OFDM⁶ system with $N = 128$ and $\nu = 32$. Selection receive diversity

⁵The design of \mathbf{W} for SISO systems is even simpler.

⁶Except for Fig. 3 where we simulate a SISO system.

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 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, α is randomly generated in each OFDM symbol from the interval $[-\frac{1}{2}, \frac{1}{2}]$.

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 $\frac{E_b}{J_{NBI}}$. The SIR is set to -10 dB unless otherwise is stated. Figure 1 shows the bit error rate (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 [1] 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 [16] 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 II-C. 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) [26] scheme where the energies of all the frequency components are computed and compared with a preset threshold. Then, the frequency subcarriers whose energies 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 Fig. 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, Fig. 1 shows that single-tone NBI signals cause more degradation than multiple-tone NBI as reported in [16]. 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. In the same figure, 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. Obviously, our CS-based approach retains its robustness to multiple NBI signals.

The case of NBI experiencing a time-varying FS channel is investigated in Fig. 2 with $D = 2$, i.e. the NBI FD

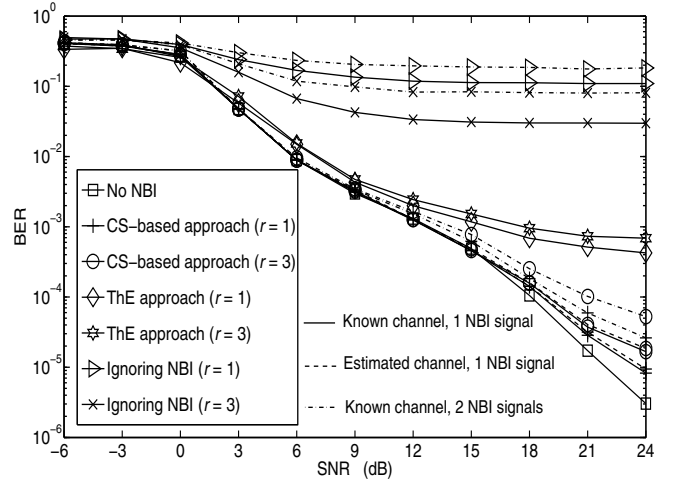


Fig. 1. BER for ZP-OFDM with known (solid lines) and estimated (dashed lines) channel for one NBI signal. Dash-dotted curves represent the case of two NBI signals.

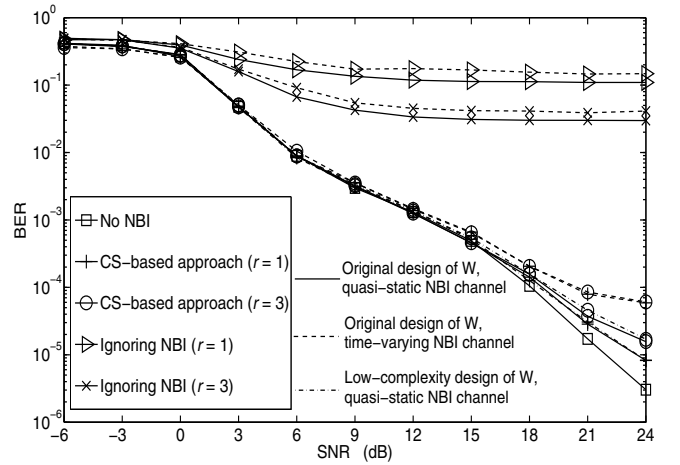


Fig. 2. BER for ZP-OFDM with the original design in (13) when NBI channel is quasi-static (solid lines) and time-varying with $D = 2$ (dashed lines). Dash-dotted curves represent the reduced-complexity design in (27).

channel matrix has 5 main diagonals. 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 Fig. 2. However, the performance gains over the case where NBI is ignored are still substantial. In the same figure, we compare the performance of the reduced-complexity design of the \mathbf{W} matrix to that of the original (more-complex) design when the NBI experiences a quasi-static channel. The reduced-complexity design virtually achieves the same performance of the more-complex design while avoiding the inversion of a channel-dependent matrix. The impact of increasing the NBI power on the performance is presented in [27]. The performance of our CS-based approach for CP-OFDM systems is shown in Fig. 3 where we follow the OFDM symbol structure of the IEEE802.11n HT 40 MHz mode [1], i.e. 20 known subcarriers out of 128 total subcarriers in each OFDM symbol. In the same figure, we show the performances of two other approaches from the literature; namely, the LMMSE and CMMOE approaches proposed in [8] and [9], respectively. We provided both approaches with the exact auto-correlation matrix of the NBI signal which

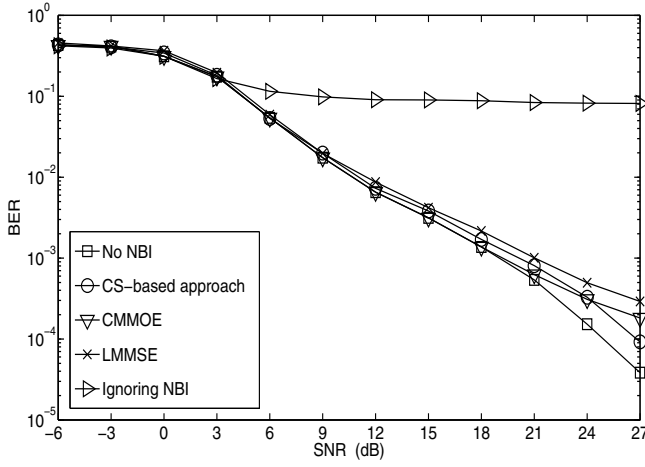


Fig. 3. Performance comparison between CS, LMMSE, and CMMOE approaches for SISO CP-OFDM systems with the HT 40MHz mode for $r = 3$.

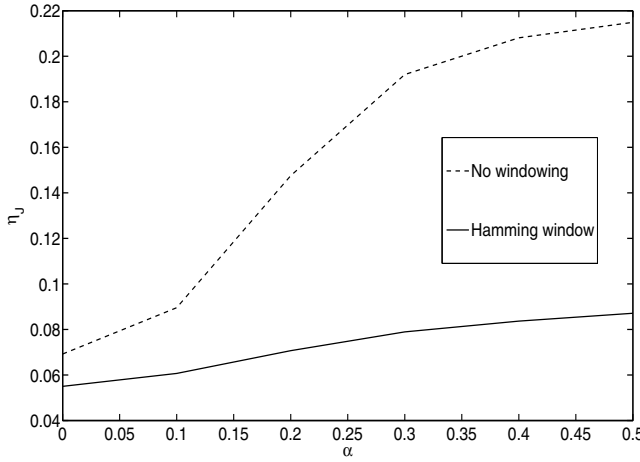


Fig. 4. Impact of frequency-offset between desired signal and NBI signal on the performance with SNR = 20 dB and $r = 3$.

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. 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 approach for ZP-OFDM systems through the metric $\eta_J \triangleq \frac{\sum_{m=1}^{M_s} \|\mathbf{J}_{eqv} - \hat{\mathbf{J}}_{eqv}\|_2^2}{\sum_{m=1}^{M_s} \|\mathbf{J}_{eqv}\|_2^2}$ where $M_s = 10^3$ is the number of realizations. The metric η_J is the normalized average error vector magnitude squared (AEVMS) of the NBI; therefore, small values of η_J indicate accurate NBI estimation, and vice versa. In Fig. 4, we investigate the effect of the FO on the performance with and without windowing. The NBI estimation accuracy degrades as α increases because the NBI signal becomes more spread in FD; hence, its sparsity

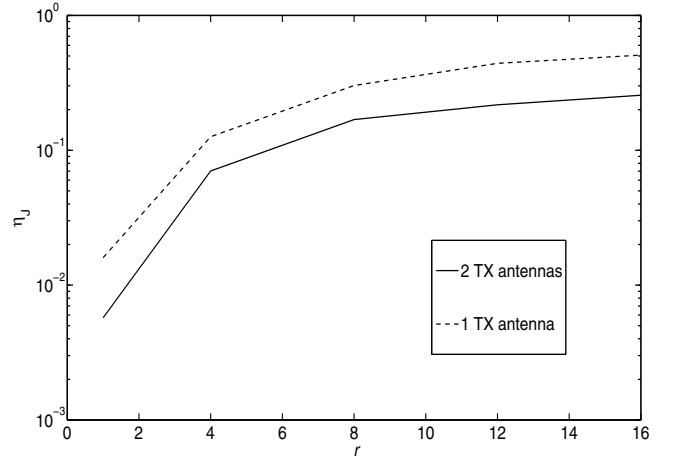


Fig. 5. Impact of transmitted NBI spectral width r on the performance with SNR = 25 dB, $\alpha = 0.1$, and 2 receive antennas with selection diversity.

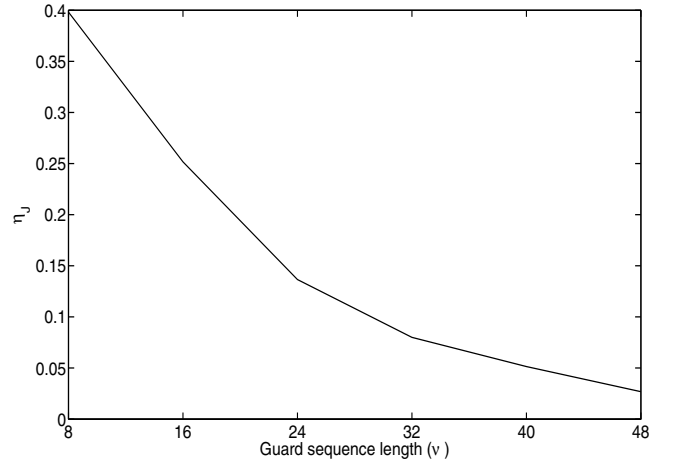


Fig. 6. Impact of guard sequence length ν on the performance with $N = 128$, SNR = 20 dB, $r = 3$, and $\alpha = 0.2$

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 Fig. 5 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 increases the rank⁷ of the measurement matrix \mathbf{W} and, hence, the number of equations, so the NBI estimation accuracy is improved. In Fig. 6, we show that increasing the guard sequence length ν 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 Fig. 7 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.

⁷According to Sections III and IV-B, rank of $\mathbf{W} = \nu$ and 2ν for 1 and 2 transmit antennas, respectively.

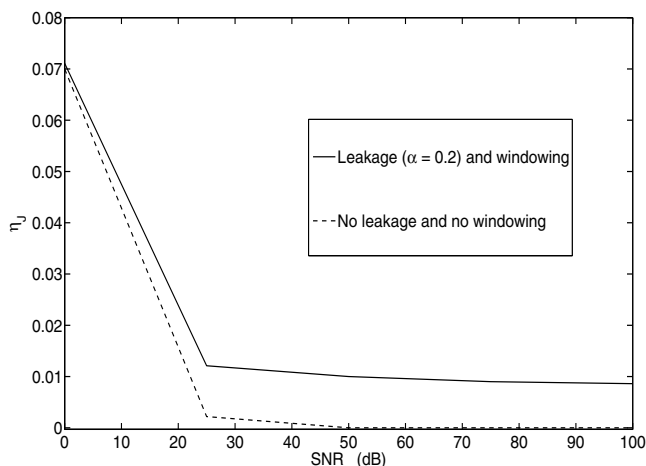


Fig. 7. Performance at high SNR values with $r = 1$.

VII. CONCLUSION

In this paper, 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 [28] 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.

REFERENCES

- [1] IEEE Standard 802.11n, "Wireless LAN medium access control (MAC) and physical layer (PHY) specifications," 2009.
- [2] J. Park, D. Kim, C. Kang, and D. Hong, "Effect of bluetooth interference on OFDM-based WLAN," in *Proc. IEEE VTC-Fall 2003*.
- [3] A. Batra and J. Zeidler, "Narrowband interference mitigation in OFDM systems," in *Proc. IEEE MILCOM 2008*.
- [4] —, "Narrowband interference mitigation in BICM OFDM systems," in *Proc. IEEE ICASSP 2009*.
- [5] D. Darsena and F. Verde, "Successive NBI cancellation using soft decisions for OFDM systems," *IEEE Signal Process. Lett.*, vol. 15, pp. 873-876, 2007.
- [6] M. Morelli and M. Moretti, "Channel estimation in OFDM systems with unknown interference," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 5338-5347, Oct. 2009.
- [7] M. Han et al., "OFDM channel estimation with jammed pilot detector under narrow-band jamming," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 1934-1939, May 2008.
- [8] R. Nilsson, F. Sjöberg, and J. LeBlanc, "A rank-reduced LMMSE canceller for narrowband interference suppression in OFDM-based systems," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2126-2140, 2003.
- [9] D. Darsena, G. Gelli, L. Paura, and F. Verde, "A constrained maximum-SINR NBI-resistant receiver for OFDM systems," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 3032-3047, 2007.
- [10] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

- [11] S. Zhou and G. Giannakis, "Space-time coding with maximum diversity gains over frequency-selective fading channels," *IEEE Signal Process. Lett.*, vol. 8, pp. 269-272, 2001.
- [12] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 9, pp. 1207-1223, 2006.
- [13] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY: Cambridge Univ. Press, 2004.
- [14] B. Muquet, Z. Wang, G. Giannakis, M. Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. Commun.*, vol. 50, pp. 2136-2148, Dec. 2002.
- [15] A. Batra, J. Balakrishnan, G. Aiello, J. Foerster, and A. Dabak, "Design of a multiband OFDM system for realistic UWB channel environments," *IEEE Trans. Microwave Theory Techniques*, vol. 52, no. 9, pp. 2123-2138, 2004.
- [16] 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*.
- [17] W. Bai, C. He, L. Jiang, and X. Li, "Robust channel estimation in MIMO-OFDM systems," *Electron. Lett.*, vol. 39, no. 2, pp. 242-244, 2003.
- [18] S. Kalyani, V. Raj, and K. Giridhar, "Narrowband interference mitigation in turbo-coded OFDM systems," in *Proc. IEEE ICC 2007*.
- [19] R. Gray, *Toeplitz and Circulant Matrices: A review*. Hanover, MA: Now Publishers Inc., 2006.
- [20] G. H. Golub and C. F. V. Loan, *Matrix Computations-Third Edition*. Baltimore, MD: Johns Hopkins Univ. Press, 1996.
- [21] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1456-1467, 1999.
- [22] T. K. Paul and T. Ogunfunmi, "Wireless LAN Comes of Age: Understanding the IEEE 802.11n Amendment," *IEEE Circuits Syst. Mag.*, vol. 8, pp. 28-54, 2008.
- [23] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451-1458, 1998.
- [24] S. M. Kay, *Fundamentals Of Statistical Signal Processing: Estimation Theory*. NJ: Prentice Hall, 1993.
- [25] 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. Commun.*, vol. 47, pp. 27-32, 1999.
- [26] 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 Trans. Signal Process.*, vol. 55, no. 3, pp. 1118-1128, 2007.
- [27] A. Gomaa and N. Al-Dhahir, "Compressive-sensing-based approach for NBI cancellation in MIMO-OFDM," *Submitted to IEEE GLOBECOM 2011*.
- [28] F. Pincaldi, G. Vitetta, R. Kalbasi, N. Al-Dhahir, M. Uysal, and H. Mheidat, "Single-carrier frequency domain equalization," *IEEE Signal Process. Mag.*, vol. 25, pp. 37-56, 2008.



Ahmad Gomaa (S'09) received his B.S and M.S degrees in electronics and communications engineering from Cairo University, Egypt, in 2005 and 2008, respectively. He is currently pursuing the Ph.D. degree at the University of Texas at Dallas, USA. His research interests include sparse FIR filters design, RF impairments compensation at the baseband, training sequence design for channel estimation, and compressive sensing applications to digital communications.

Naofal Al-Dhahir (F'08) is a Jonsson Distinguished Professor of Engineering at the University of Texas at Dallas and an IEEE Fellow. More information is available at <http://www.utdallas.edu/~aldhahir>.