Inter-Carrier Interference in MIMO OFDM

Anastasios Stamoulis, Suhas N. Diggavi, and Naofal Al-Dhahir
AT&T Shannon Laboratory
Florham Park, NJ 07932

Abstract

In this paper, we examine multicarrier transmission over time-varying channels. We first develop a model for such a transmission scheme and focus particularly on MIMO OFDM. Using this method, we analyze the impact of time-variation within a transmission block (time variation could arise both from Doppler spread of the channel and from synchronization errors). To mitigate the effects of such time-variations, we propose a time-domain approach. We design ICI-mitigating block linear filters, and we examine how they are modified in the context of space-time block-coded transmissions. Our approach reduces to the familiar single-tap frequency-domain equalizer when the channel is block time-invariant. Channel estimation in rapidly time-varying scenarios becomes critical, and we propose a scheme for estimating channel parameters varying within a transmission block. Along with the channel estimation scheme, we also examine the issue of pilot tone placement and show that in time-varying channels it may be better to group pilot tones together into clumps equispaced onto the FFT grid; this placement technique is in contrast to the common wisdom for time-invariant channels. Finally, we provide numerical results illustrating the performance of these schemes, both for uncoded and space-time block-coded systems.

Parts of this paper will appear in WCNC'02 and ICC'02.

I. INTRODUCTION

The explosive growth of wireless communications is creating the demand for high-speed, reliable, and spectrally-efficient communication over the wireless medium. There are several challenges in attempts to provide high-quality service in this dynamic environment. These pertain to channel time-variation and the limited spectral bandwidth available for transmission. In this paper we are concerned with the role of time-variation in transmission of multicarrier modulated signals and the design of schemes to handle such variations.

Multicarrier transmission for wireline channels has been well studied [1]. The main advantage of orthogonal frequency division multiplexing (OFDM[1]) transmission stems from the fact that the Fourier basis forms an eigenbasis for time-invariant(DMT) channels. This simplifies the receiver and leads to inexpensive hardware implementations, as the equalizer is just a single-tap filter in the frequency domain (under the assumption that the channel is time-invariant within a transmission block). Furthermore, combined with multiple antennas [2], [3], OFDM modems are attractive for high data rate wireless networks (Wireless LANs and Home Networking). However, the block time-invariance assumption may not be valid at high mobile speeds or when there are impairments such as synchronization errors (e.g., frequency offsets). In such situations, the Fourier basis need not be the eigenbasis, and the loss of orthogonality (of the carriers) at the receiver results in inter-carrier interference (ICI). Depending on the Doppler spread in the channel and the block length chosen for transmission, ICI can potentially cause a severe deterioration of quality of service (QoS) in OFDM systems. Essentially, ICI arises from "time-selectivity" of the channel and therefore is the frequency-domain dual of the inter-symbol interference (ISI) seen in single-carrier transmission over frequency-selective channels.

Given the direct relationship between achievable data rate and the Signal-To-Inter-carrier-Interference-Plus-Noise Ratio (SINR) at every subcarrier (see, e.g., [1]), ICI has a negative impact on the data throughput, as it effectively decreases the SINR. Starting from an analysis of the ICI impact on system performance, we develop herein an

[1] For wireline transmissions, OFDM is commonly referred to as Discrete Multitone [1].

ICI mitigation scheme for both single-input single-output (SISO) and multiple-input multiple-output (MIMO) OFDM systems. Founded upon optimum block linear filters and practically realizable channel estimation techniques, our proposed scheme significantly improves the SINR (which serves as a figure of merit for the system performance) over a wide range of channel and system configurations.

In SISO OFDM, carrier frequency errors induce a structured ICI pattern which can be eliminated using either training-based or blind techniques (see, e.g., [4] and references therein). The frequency offset manifests itself as a multiplicative factor in the frequency channel response (at each subcarrier of the FFT grid), and once the frequency offset has been estimated, it can be removed rather easily. Without estimating the carrier offset, transmit precoding can be employed to suppress its impact on ICI (see, e.g., [5] and references therein). On the other hand, Doppler-induced ICI appears to be more challenging. In SISO OFDM, previous studies have quantified the effects of ICI on the system performance: [6] uses central limit theorem arguments to model ICI as a Gaussian random process and quantify its impact on bit error rate (BER), whereas [7] presents results for a number of different Doppler spectra. Reference [8] focuses on the wide-sense stationary uncorrelated scattering (WSSUS) [9] channel to show that if the OFDM block duration is greater than 8% of the channel coherence time, then the SINR (due to ICI) is less than 20dB (see also [10] and references therein).

Similar to SISO OFDM, MIMO OFDM is sensitive to Doppler and carrier frequency errors which destroy the subcarrier orthogonality and give rise to ICI. MIMO OFDM may be equipped with transmit and/or receive diversity (see, e.g., [11], [2]), and may build upon beamforming, space-time coding, and interference cancellation (to name just a few techniques) to increase the achievable data rates over the wireless medium. Our ICI-mitigating designs encompass generic MIMO OFDM systems. In particular, for space-time-coded transmissions, we illustrate through simulations the significant SINR gains by our approach.

Compensation for Doppler-induced ICI hinges upon the estimation of the time-varying channel. When the OFDM block duration is less than 10% of the channel coherence time, [12] argues that the channel can be assumed to vary in a linear fashion (and, as a result, it can be obtained by linear interpolation between two channel estimates acquired from training blocks). When the OFDM block duration is much smaller than the channel coherence time (relatively mild Doppler), the channel can be assumed approximately constant (over an OFDM block), and its estimation has been thoroughly studied (see [10] and references therein). For rapidly time-varying environments, and for the cases where the channel obeys a parsimonious model, techniques in [13], [14] can be used for channel estimation and tracking. In this work, we look at *severe* Doppler cases, where the channel can no longer be assumed constant within a transmission block: for such rapidly time-varying channels, we develop a novel estimation method.

Our contributions in this paper can be summarized as follows

- We develop a model for inter-carrier interference applicable to MIMO channels.
- We present an ICI analysis for MIMO-OFDM.
- We propose a time-domain filtering scheme for mitigating ICI.
- We propose a channel estimation and tracking scheme for estimating the time-varying (within transmission block) channel parameters.
- We propose a pilot tone placement scheme (which works in conjunction with the channel estimation scheme).

The following is an overview of the paper organization. In Section II we present the model and the notation used in the paper. We present an ICI analysis for MIMO OFDM in Section III, and, in Section IV we develop optimum block linear filters which suppress ICI and yield significant SINR improvement. In Section V, we address the problem of estimating a rapidly time-varying channel, and in Section VI we present illustrative simulations. Finally, in Section VII we conclude with some discussion.

II. Data Model

The input data $\{x(k)\}$ is passed through the filter $g(t)$ to produce the transmitted signal $s(t)$. The received signal can be written as

$$y_c(t) = \int h_c(t;\tau)s(t-\tau)d\tau + z(t) \ , \qquad (1)$$

where $h_c(t;\tau)$ is the impulse response of the time-varying channel and $z(t)$ is the additive Gaussian noise. We collect sufficient statistics through Nyquist sampling. The basic idea is that if we sample at a rate larger than $2(W_I + W_s)$, where W_I is the input bandwidth and W_s is the bandwidth of the channel time-variation [15], then we get Nyquist sampling. A careful argument about the sampling rate required for time-varying channels can be found in [16]. In this paper, we assume that this criterion is met and therefore we have the following discrete-time model

$$y(k) = y_c(kT_s) = \sum_{l=0}^{\nu-1} h(k;l)x(k-l) + z(k)\ , \tag{2}$$

where $h(k;l)$ represents the sampled time-varying channel impulse response (which combines the transmit filter $g(t)$ with the physical channel $h_c(t;\tau)$). The approximation of having a finite impulse response in (2) can be made as good as we need by choosing ν [16]. In this paper, we focus on the discrete-time model given in (2). For the case of multiple transmitter and receiver diversity channel with M_r receive and M_t transmit antennas, we can easily generalize this model to

$$\mathbf{y}(k) = \sum_{l=0}^{\nu-1} \mathbf{H}(k;l)\mathbf{x}(k-l) + \mathbf{z}(k)\ , \tag{3}$$

where $\mathbf{H}(k;l) \in \mathbb{C}^{M_r \times M_t}$ is the l^{th} tap of the matrix response with $\mathbf{x}(k) \in \mathbb{C}^{M_t}$ as the input, $\mathbf{y}(k) \in \mathbb{C}^{M_r}$ is the output, and $\mathbf{z}(k) \in \mathbb{C}^{M_r}$ is the complex additive temporally-white circularly-symmetric Gaussian noise with $\mathbf{z}(k) \sim \mathbb{C}\mathcal{N}(0, \mathbf{R}_{zz})$, i.e., a complex Gaussian vector with mean $\mathbf{0}$ and covariance $\mathbf{R}_{zz}$. The noise is assumed independent of the input. Throughout this paper, we impose an average power constraint on the input which is assumed to be zero mean, i.e., $\mathbb{E}[||\mathbf{x}(k)||^2] \leq P$. The specific structure of $\{\mathbf{H}(k;l)\}$ could be constructed by assigning a special structure to $\mathbf{H}_c(t;\tau)$ (for example a discrete multipath channel).

Over a time block of N symbol durations, Equation (3) can be expressed in matrix notation as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}\ , \tag{4}$$

where $\mathbf{y}, \mathbf{z} \in \mathbb{C}^{N.M_r}$, $\mathbf{x} \in \mathbb{C}^{N.M_t}$, and $\mathbf{H} \in \mathbb{C}^{N.M_r \times N.M_t}$. We assume that a cyclic prefix of length equal to channel memory is inserted in each input block to eliminate inter–

block interference (IBI). This prefix serves just as a guard interval between blocks when we have channel time-variation within a transmission block. However, when the channel impulse response (CIR) is time–invariant over the block, the cyclic prefix serves the role played in OFDM, i.e., $\mathbf{H}$ becomes a *block-circulant* matrix.

Now, let us consider OFDM transmission where each of the M_t time–domain N–dimensional input vectors in x is generated by taking IDFT of an information–bearing vector, i.e.

$$\mathbf{x} = \mathbf{Q}^{(\mathrm{Tx})H}\mathbf{X} \ , \tag{5}$$

where[2] $\mathbf{Q}^{(\mathrm{Tx})} = \tilde{\mathbf{Q}} \otimes \mathbf{I}_{M_t}$ ($\otimes$ denotes the Kronecker product [17]), $\mathbf{I}_{M_t}$ is the identity matrix of dimension M_t, and $\tilde{\mathbf{Q}}(l,k) = \frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}lk}$ for $0 \le l,k \le N-1$ is the standard N-dimensional DFT matrix. We will first develop more detailed notation for the case $M_r = M_t = 1$, and then the generalization would not be difficult.

We can write the output of the DFT at the receiver for a time-block $[-(\nu-1),\ldots,N-1]$ (where N is the number of carriers) as

$$Y(p) = G(p,p)X(p) + \sum_{\substack{q=0 \\ p\neq q}}^{N-1} G(p,q)X(q) + Z(p) \tag{6}$$

for $p = 0,\ldots,N-1$, where $Y(p)$, $X(p)$ and $Z(p)$ are the DFTs of $\{y(k)\}$, $\{x(k)\}$ and $\{z(k)\}$ respectively. We have also defined $G(p,q)$ as the $(p,q)^{th}$ element of $\mathbf{G} = \tilde{\mathbf{Q}}\bar{\mathbf{H}}^{(N)}\tilde{\mathbf{Q}}^H/N$. Here $\tilde{\mathbf{Q}}$ is the DFT matrix defined as $\tilde{\mathbf{Q}}^H = [\mathbf{q}_0,\ldots\mathbf{q}_{N-1}]$ and $\mathbf{q}_s = \frac{1}{\sqrt{N}}[1,\ldots,\exp(j2\pi s(N-1)/N)]^T$, and $\bar{\mathbf{H}}^{(N)}$ is the equivalent channel matrix including the effects of the cyclic prefix (used in OFDM) defined as

$$\bar{\mathbf{H}}^{(N)} = \begin{bmatrix} h(0;0) & 0 & \cdots & h(0;2) & h(0;1) \\ h(1;1) & h(1;0) & \cdots & h(1;3) & h(1;2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h(\nu-1;\nu-1) & h(\nu-1;\nu-2) & \vdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & h(N-2;0) & 0 \\ 0 & 0 & \cdots & h(N-1;1) & h(N-1;0) \end{bmatrix} \tag{7}$$

[2] We denote complex matrix transpose with superscript H, complex conjugation by superscript $*$ and matrix transpose by superscript T.

We can easily evaluate $G(m,s)$ as

$$G(m,s) = \frac{1}{N}\sum_{r=0}^{N-1}\sum_{l=0}^{\nu-1} h(r;l)e^{j2\pi r(s-m)/N}e^{-j2\pi sl/N} . \tag{8}$$

Note that the form of the model in (6) is applicable to more general transmission scenarios than using OFDM. We can replace $\tilde{\mathbf{Q}}$ by any arbitrary matrix $\mathbf{B}$ and for a given structure of the guard interval (prefix), we can find the specific structure of $\mathbf{G}$ (as in (8) for OFDM) for that case.

In the case when we have transmit and receive diversities we can modify (6) to

$$\mathbf{Y}(p) = \mathbf{G}(p,p)\mathbf{X}(p) + \sum_{\substack{q=0\\q\neq p}}^{N-1}\mathbf{G}(p,q)\mathbf{X}(q) + \mathbf{Z}(p) , \tag{9}$$

where $[\mathbf{G}(p,q)]_{i,j} = [\tilde{\mathbf{Q}}\mathbf{H}^{(i,j)}\tilde{\mathbf{Q}}^H/N]_{p,q}$, $1 \leq i \leq M_r,\ 1 \leq j \leq M_t,\ 0 \leq p \leq N-1,\ 0 \leq q \leq N-1$, and $\mathbf{H}^{(i,j)}$ is the equivalent channel matrix (as in (7)) for the time-varying ISI channel between the i^{th} receiver and the j^{th} transmitter, denoted by $\{h^{(i,j)}(k;l)\}$. Also the received vector for the p^{th} frequency bin is $\mathbf{Y}(p) = [Y^{(1)}(p),\ldots,Y^{(M_r)}(p)]^T$, $\mathbf{X}(p) = [X^{(1)}(p),\ldots,X^{(M_t)}(p)]^T$ is the transmitted vector for the p^{th} frequency bin, and $\mathbf{Z}(p) = [Z^{(1)}(p),\ldots,Z^{(M_r)}(p)]^T$ is the noise vector. We can rewrite (9) in compact form as

$$\mathbf{Y} = \mathbf{G}\mathbf{X} + \mathbf{Z} , \tag{10}$$

where the $(p,q)^{th}$ block of $\mathbf{G}$ is given by $\mathbf{G}(p,q)$, $\mathbf{Y} = [\mathbf{Y}(0),\ldots,\mathbf{Y}(N-1)]^T \in \mathbb{C}^{N.M_r}$, $\mathbf{X} = [\mathbf{X}(0),\ldots,\mathbf{X}(N-1)]^T \in \mathbb{C}^{N.M_t}$, and $\mathbf{Z} = [\mathbf{Z}(0),\ldots,\mathbf{Z}(N-1)]^T \in \mathbb{C}^{N.M_r}$. We can also write the matrix $\mathbf{G}$ in more compact notation as

$$\mathbf{G} = \mathbf{Q}^{(\mathrm{Rx})}\mathbf{H}\mathbf{Q}^{(\mathrm{Tx})H}, \tag{11}$$

where $\mathbf{Q}^{(\mathrm{Rx})} = \tilde{\mathbf{Q}} \otimes \mathbf{I}_{M_r},\ \ \mathbf{Q}^{(\mathrm{Tx})} = \tilde{\mathbf{Q}} \otimes \mathbf{I}_{M_t}$ and $\mathbf{H}$ is obtained from (4).

III. ICI ANALYSIS

We start with the data model given in (9) which we rewrite below

$$\mathbf{Y}(m) = \mathbf{G}(m,m)\mathbf{X}(m) + \sum_{\substack{n=0\\n\neq m}}^{N-1}\mathbf{G}(m,n)\mathbf{X}(n) + \mathbf{Z}(m) . \tag{12}$$

We are interested in finding the covariance matrix of the ICI (with noise) $\tilde{\mathbf{Z}}(m) = \mathbf{Z}_{ICI}(m) + \mathbf{Z}(m)$, where $\mathbf{Z}_{ICI}(n) = \sum_{\substack{n=0 \\ n \neq m}}^{N-1} \mathbf{G}(m,n)\mathbf{X}(n)$. In order to find the covariance of $\mathbf{Z}_{ICI}(n)$ we need to make a few simplifying assumptions

• Each channel $\{h^{(i,j)}(k;l)\}$ is modeled as a WSSUS channel, i.e., $\mathbb{E}[h^{(i,j)}(k;l)h^{(i,j)*}(k;n)] = \mathbb{E}[|h^{(i,j)}(k;l)|^2]\delta_{l-n}$. Moreover, the channel is modeled as a complex Gaussian stochastic process.

• The channel between the i^{th} receiver and the j^{th} transmitter is uncorrelated with the channel between the k^{th} receiver and the l^{th} transmitter when $i \neq k$ or $j \neq l$. That is, elements of $\mathbf{H}^{(i,j)}$ are uncorrelated with the elements of $\mathbf{H}^{(k,l)}$. Moreover, we assume that the channels $\{h^{(i,j)}(k;l)\}$ are identically distributed over $\{(i,j)\}$.

These statistical assumptions on the channel are *not* crucial and the analysis can be accomplished without these assumptions. However, they do simplify the algebra and expose the important issues that we need to point out. Given these assumptions we can make the following observation.

Fact III.1: The elements of the matrix $\mathbf{G}(m,n)$ are independent and identically distributed (i.i.d.) Gaussian random variables.

Proof: As $[\mathbf{G}(m,n)]_{i,j} = [\tilde{\mathbf{Q}}\mathbf{H}^{(i,j)}\tilde{\mathbf{Q}}^H/N]_{m,n}$, the elements of $\mathbf{G}(m,n)$ are i.i.d. complex Gaussian because the channels $\mathbf{H}^{(i,j)}$ are modeled as complex Gaussian and independent across the transmit and receive antennas. □

Given these statistical assumptions we can now proceed to evaluate the covariance of $\mathbf{Z}_{ICI}(n)$

$$\begin{aligned} \mathbb{E}[\mathbf{Z}_{ICI}(n)\mathbf{Z}_{ICI}^H(n)] &= \mathbb{E}[\sum_{\substack{p=0 \\ n \neq m}}^{N-1}\sum_{\substack{n=0 \\ n \neq m}}^{N-1} \mathbf{G}(m,p)\mathbf{X}(p)\mathbf{X}^H(n)\mathbf{G}^H(m,n)] \\ &= \sum_{\substack{p=0 \\ n \neq m}}^{N-1}\sum_{\substack{n=0 \\ n \neq m}}^{N-1} \mathbb{E}[\mathbf{G}(m,p)\mathbf{R}_{XX}(p,n)\mathbf{G}^H(m,n)], \end{aligned} \tag{13}$$

where we have defined $\mathbf{R}_{XX}(p,n) = \mathbb{E}[\mathbf{X}(p)\mathbf{X}^H(n)]$. Equation (13) indicates that the basic component of the calculation is $\mathbb{E}[\mathbf{G}(m,p)\mathbf{R}_{XX}(p,n)\mathbf{G}^H(m,n)]$; this is what we focus on next.

We start from the observation that we can do an eigendecomposition of the signal covariance matrix as $\mathbf{R}_{XX}(p,n) = \mathbf{U}(p,n)\mathbf{\Lambda}_X(p,n)\mathbf{U}^H(p,n)$ where $\mathbf{U}(p,n)$ is a unitary matrix. From Fact (III.1), the matrix $\mathbf{G}(m,p)$ has i.i.d. Gaussian elements (i.e., $\mathbf{G}(m,p)$ is isotropically distributed), which implies that the matrix $\mathbf{G}(m,p)\mathbf{U}(p,n)$ is distributionally equivalent to $\mathbf{G}(m,p)$ (note that if we apply a unitary transformation on an i.i.d. Gaussian matrix the resultant matrix is also an i.i.d. Gaussian matrix with the same distribution). Hence, we can rewrite $\mathbb{E}[\mathbf{G}(m,p)\mathbf{R}_{XX}(p,n)\mathbf{G}^H(m,n)]$ in the following way

$$\mathbb{E}[\mathbf{G}(m,p)\mathbf{R}_{XX}(p,n)\mathbf{G}^H(m,n)] = \mathbb{E}[\mathbf{G}(m,p)\mathbf{\Lambda}_X(p,n)\mathbf{G}^H(m,n)], \tag{14}$$

where $\mathbf{\Lambda}_X(p,n) = diag(\lambda_1(p,n),\ldots,\lambda_{M_t}(p,n))$ is a diagonal matrix of the eigenvalues of $\mathbf{R}_{XX}(p,n)$.

Now, let us examine the elements of the matrix given in (14). The $(r,s)^{th}$ element is

$$\begin{aligned}
&\left\{\mathbb{E}[\mathbf{G}(m,p)\mathbf{\Lambda}_X(p,n)\mathbf{G}^H(m,n)]\right\}_{r,s} = \\
&= \sum_{k=1}^{M_t} \lambda_k(p,n)\mathbb{E}[\{\mathbf{G}(m,p)\}_{r,k}\left\{\mathbf{G}^H(m,n)\right\}_{k,s}] \\
&\stackrel{(a)}{=} \sum_{k=1}^{M_t} \lambda_k(p,n)\mathbb{E}[\{\mathbf{G}(m,p)\}_{r,k}\left\{\mathbf{G}^H(m,n)\right\}_{k,r}]\delta_{r-s}\ ,
\end{aligned} \tag{15}$$

where (a) is due to the assumption that the channels on different links are uncorrelated[3]. Hence the component matrices $\mathbb{E}[\mathbf{G}(m,p)\mathbf{R}_{XX}(p,n)\mathbf{G}^H(m,n)]$ are diagonal (a direct result of the fact that the channels are modeled as being spatially white).

Using the WSSUS modeling assumption, and using the Fourier basis, we can write

$$\begin{aligned}
&\mathbb{E}[\{\mathbf{G}(m,p)\}_{r,k}\left\{\mathbf{G}^H(m,n)\right\}_{k,r}] = \\
&\frac{1}{N^2}\left(\sum_{l=0}^{\nu-1} e^{-j2\pi l(p-n)/N}\right)\sum_{r_1=0}^{N-1}\sum_{r_2=0}^{N-1} r_{hh}(r_1-r_2)e^{j2\pi r_1(p-m)/N}e^{-j2\pi r_2(n-m)/N}\ ,
\end{aligned} \tag{16}$$

where we have defined $r_{hh}(r_1-r_2) = \mathbb{E}[h^{(k,r)}(r_1;l)h^{(k,r)*}(r_2;l)]$.

[3]This can be seen from the fact that $[\mathbf{G}(m,p)]_{r,k} = [\tilde{\mathbf{Q}}\mathbf{H}^{(r,k)}\tilde{\mathbf{Q}}^H/N]_{m,p}$ and $[\mathbf{G}^H(m,n)]_{k,s} = [\tilde{\mathbf{Q}}\mathbf{H}^{(s,k)H}\tilde{\mathbf{Q}}^H/N]_{m,n}$, which shows that if $r \neq s$ then $[\mathbf{G}(m,p)]_{r,k}$ and $[\mathbf{G}^H(m,n)]_{k,s}$ are uncorrelated.

Consequently, as the channels are spatially independent we can write the covariance matrix of the ICI as

$$\mathbb{E}[\mathbf{Z}_{ICI}(n)\mathbf{Z}_{ICI}^{H}(n)] = \sum_{\substack{p=0\\ n\neq m}}^{N-1}\sum_{\substack{n=0\\ n\neq m}}^{N-1}\frac{\text{tr}(\mathbf{\Lambda}_X(p,n))}{N^2}\left(\sum_{l=0}^{\nu-1}e^{-j2\pi l(p-n)/N}\right)\left(\sum_{r_1=0}^{N-1}\sum_{r_2=0}^{N-1}r_{hh}(r_1-r_2)e^{j2\pi r_1\frac{p-m}{N}}e^{-j2\pi r_2\frac{n-m}{N}}\right)\mathbf{I}\ ,$$

where $tr(.)$ denotes the trace of a matrix.

This analysis points to a few issues worth noting

- The ICI is accentuated with the presence of multiple transmit antennas (i.e., $M_t > 1$).
- The covariance of the ICI is spatially white in (17) due to the assumption that the channel matrices $\mathbf{H}^{(i,j)}$ were assumed to be so. We can analyze spatial correlation with similar techniques, but the expressions turn out to be a little more complicated without adding much more insight.
- Multiple receive antennas could be used to mitigate the ICI. For example, the ICI term in (12) can be seen as interference, and hence using multiple receive antennas we can do either zero-forcing or MMSE interference suppression purely in the spatial domain [18]. However, this might require many receive antennas depending on the amount of dominant ICI. In order to also utilize the additional available degrees of freedom (dimensions) in the time-domain, we examine next a combined time and spatial domain approach.

IV. Time-Domain ICI Mitigation Technique

Henceforth, we focus on the case $M_t = M_r = M$ as it brings out the essential ideas without unduly complicating the algebra. Extending the analysis to the general case is straightforward. Therefore, as $M_r = M_t = M$, for notational convenience using (11) we write $\mathbf{Q} = \mathbf{Q}^{(\mathrm{Rx})} = \mathbf{Q}^{(\mathrm{Tx})}$.

To combat ICI, the length–NM received vector $\mathbf{y}$ is multiplied by an $NM \times NM$ matrix (corresponding to an NM–tap time–varying filter) $\mathbf{W}$ resulting in

$$\tilde{\mathbf{y}} = \mathbf{W}\mathbf{y} = \mathbf{W}(\mathbf{H}\mathbf{x}+\mathbf{z}) = \mathbf{W}\mathbf{H}\mathbf{x}+\tilde{\mathbf{z}}\ , \tag{17}$$

where the colored noise vector $\tilde{\mathbf{z}}$ has auto–correlation matrix $\mathbf{R}_{\tilde{z}\tilde{z}} = \sigma_z^2 \mathbf{W}\mathbf{W}^H$. Taking the DFT of both sides of (17), we get

$$\begin{aligned} \tilde{\mathbf{Y}} &\stackrel{def}{=} \mathbf{Q}\tilde{\mathbf{y}} = \mathbf{QWHQ}^H\mathbf{X} + \mathbf{Q}\tilde{\mathbf{z}} \\ &\stackrel{def}{=} \bar{\mathbf{G}}\mathbf{X} + \tilde{\mathbf{Z}} \; . \end{aligned} \tag{18}$$

Note here that $\bar{\mathbf{G}}$ captures the effect of filtering with $\mathbf{W}$. In contrast, in Equation (10), $\mathbf{G}$ captures the effects of ICI before the filtering operation. Next, we show how to design linear ICI–cancellation filters for OFDM. We start with the SISO case $(M = 1)$ and then extend the design to the MIMO case.

A. Optimum Linear ICI-Cancellation Filter for SISO OFDM

Using (18), we can express the SINR at the m^{th} frequency bin $(1 \leq m \leq N)$ as follows [19]

$$SINR_m = \frac{E_{x,m}|\bar{\mathbf{G}}(m,m)|^2}{\frac{1}{N}\mathrm{tr}(\mathbf{R}_{\tilde{z}\tilde{z}}) + \sum_{n \neq m} E_{x,n}|\bar{\mathbf{G}}(m,n)|^2} \; , \tag{19}$$

where $E_{x,n}$ is the input energy allocated to the n^{th} frequency bin. We assume that the channel is unknown at the transmitter. Therefore, the total input energy $E_{x,tot}$ is divided equally across all frequency bins, i.e., $E_{x,n} = \frac{E_{x,tot}}{N} \stackrel{def}{=} E_x$. Under this assumption and defining $\mathbf{e}_i$ to be the i^{th} unit vector, Equation (19) becomes

$$\begin{aligned} SINR_m &= \frac{E_x|\mathbf{e}_m^H\bar{\mathbf{G}}\mathbf{e}_m|^2}{\frac{\sigma_{\tilde{z}}^2}{N}\mathrm{tr}(\mathbf{W}\mathbf{W}^H) + E_x\sum_{n \neq m}|\mathbf{e}_m^H\bar{\mathbf{G}}\mathbf{e}_n|^2} \\ &= \frac{\mathbf{e}_m^H\mathbf{QWHQ}^H\mathbf{e}_m\mathbf{e}_m^H\mathbf{QH}^H\mathbf{W}^H\mathbf{Q}^H\mathbf{e}_m}{\frac{1}{SNR} + \mathbf{e}_m^H\mathbf{QWHQ}^H(\sum_{n \neq m}\mathbf{e}_n\mathbf{e}_n^H)\mathbf{QH}^H\mathbf{W}^H\mathbf{Q}^H\mathbf{e}_m} \\ &= \frac{\mathbf{q}_m^H\mathbf{W}\mathbf{h}_m\mathbf{h}_m^H\mathbf{W}^H\mathbf{q}_m}{\frac{1}{SNR} + \mathbf{q}_m^H\mathbf{WH}(\mathbf{I}_N - \mathbf{q}_m\mathbf{q}_m^H)\mathbf{H}^H\mathbf{W}^H\mathbf{q}_m} \\ &= \frac{\mathbf{w}_m^H\mathbf{h}_m\mathbf{h}_m^H\mathbf{w}_m}{\mathbf{w}_m^H(\frac{1}{SNR}\mathbf{I}_N + \mathbf{R}_m)\mathbf{w}_m} \; , \end{aligned} \tag{20}$$

where we have defined $SNR = \frac{E_x}{\sigma_{\tilde{z}}^2}$, $\mathbf{q}_m \stackrel{def}{=} \mathbf{Q}^H\mathbf{e}_m$, $\mathbf{h}_m = \mathbf{H}\mathbf{q}_m$, $\mathbf{w}_m = \mathbf{W}^H\mathbf{q}_m$, and $\mathbf{R}_m = \mathbf{HQ}^H(\mathbf{I}_N - \mathbf{e}_m\mathbf{e}_m^H)\mathbf{QH}^H = \mathbf{HH}^H - \mathbf{h}_m\mathbf{h}_m^H$. Furthermore, we assumed that $\mathbf{w}_m^H\mathbf{w}_m = 1$ for $1 \leq m \leq N$, which implies that $\mathrm{tr}(\mathbf{W}\mathbf{W}^H) = \sum_{m=1}^N \mathbf{w}_m^H\mathbf{w}_m = N$.

The achievable bit rate in an OFDM system is (approximately) given by [1]

$$Rate = \sum_{m=1}^{N} \log_2(1 + \frac{SINR_m}{\Gamma}) \ , \tag{21}$$

where $\Gamma \approx \frac{\gamma_m}{3\gamma_c}(Q^{-1}(\frac{P_e}{N_e}))^2$ is the *SNR gap* (assumed the same for all frequency bins), the Q-function is defined by $Q(x) = \int_x^\infty e^{-x^2/2}\, dx$, and γ_m, γ_c denote the desired performance margin and overall coding gain, respectively. Furthermore, P_e and N_e are the target probability of error and the number of nearest neighbors in the used signal constellation, respectively. It is clear from (21) that optimizing $\mathbf{w}_m$ (and hence $\mathbf{W}$) in (20) to maximize $SINR_m$ also maximizes the bit rate.

We have the following optimization problem

$$max_{\mathbf{w}_m} \mathbf{w}_m^H \mathbf{h}_m \mathbf{h}_m^H \mathbf{w}_m \ \text{ subject to } \ \mathbf{w}_m^H(\frac{1}{SNR}\mathbf{I} + \mathbf{R}_m)\mathbf{w}_m = 1 \ \text{ and } \ \mathbf{w}_m^H \mathbf{w}_m = 1 \ .$$

This is a standard *generalized eigenvalue problem* whose solution is determined as follows

1. Define the Cholesky factorization $\tilde{\mathbf{R}}_m \stackrel{def}{=} \frac{1}{SNR}\mathbf{I} + \mathbf{R}_m \stackrel{def}{=} \mathbf{L}\mathbf{L}^H$.

2. Compute the *dominant* eigenvector, call it $\mathbf{v}_1$, of the matrix $\mathbf{L}^{-1}\mathbf{h}_m\mathbf{h}_m^H\mathbf{L}^{-H}$. Now, since the matrix $\mathbf{h}_m\mathbf{h}_m^H$ is rank–1, it can be shown that $\mathbf{v}_1 = \alpha\mathbf{L}^{-1}\mathbf{h}_m$ where α is a constant that makes $\mathbf{v}_1$ a unit–norm vector.

3. Compute $\mathbf{w}_{m,opt} = \mathbf{L}^{-H}\mathbf{v}_1 = \beta\tilde{\mathbf{R}}_m^{-1}\mathbf{h}_m$ where the constant β is chosen to make $\mathbf{w}_{m,opt}$ a unit–norm vector.

4. The corresponding optimum SINR for the m^{th} frequency bin is

$$SINR_{m,opt} = \mathbf{h}_m^H \tilde{\mathbf{R}}_m^{-1} \mathbf{h}_m \ . \tag{22}$$

B. An Alternative Form

We can avoid the need to compute $\tilde{\mathbf{R}}_m^{-1}$ for each m as follows. Using the matrix inversion lemma, we can write

$$\begin{aligned} \tilde{\mathbf{R}}_m &= \frac{1}{SNR}\mathbf{I} + \mathbf{R}_m = \frac{1}{SNR}\mathbf{I} + \mathbf{H}\mathbf{H}^H - \mathbf{h}_m\mathbf{h}_m^H \\ &\stackrel{def}{=} \mathbf{R}_{yy} - \mathbf{h}_m\mathbf{h}_m^H \\ \Rightarrow \tilde{\mathbf{R}}_m^{-1} &= (\frac{1}{SNR}\mathbf{I} + \mathbf{R}_m)^{-1} = \mathbf{R}_{yy}^{-1}(\mathbf{I} + \frac{\mathbf{h}_m\mathbf{h}_m^H\mathbf{R}_{yy}^{-1}}{(1 - \mathbf{h}_m^H\mathbf{R}_{yy}^{-1}\mathbf{h}_m)}) \ . \end{aligned}$$

Therefore, only one inverse, namely that of the output auto–correlation matrix $\mathbf{R}_{yy}$, needs to be computed. In summary, the optimum filters $\mathbf{w}_{m,opt}$ are computed as follows

1. Compute $\mathbf{R}_{yy}^{-1}$ (using channel estimates obtained with pilot tones and interpolation or using an ensemble average from several received vectors $\mathbf{y}$).
2. For each frequency bin $1 \leq m \leq N$, compute
   (a) $\tilde{\mathbf{w}}_m = \mathbf{R}_{yy}^{-1}\mathbf{h}_m$
   (b) $\mathbf{w}_{m,opt} = \frac{\tilde{\mathbf{w}}_m}{|\mathbf{w}_m|}$
   (c) The corresponding optimum SINR for m^{th} frequency bin is

$$SINR_{m,opt} = \frac{\mathbf{h}_m^H\mathbf{R}_{yy}^{-1}\mathbf{h}_m}{(1 - \mathbf{h}_m^H\mathbf{R}_{yy}^{-1}\mathbf{h}_m)} \; . \tag{23}$$

Finally, we note that computational complexity can be further reduced using a low–rank approximation of $\mathbf{R}_{yy}$ (this technique has been used in [20] as part of the channel estimation process).

C. *ICI Cancellation Filter for MIMO OFDM*

We can express the SINR for the j^{th} transmit antenna and the k^{th} frequency bin, denoted by $SINR_k^j$, in MIMO OFDM as follows [4]

$$\begin{aligned} SINR_k^j &= \frac{E_x \sum_{r=1}^{M} |\bar{\mathbf{G}}(kM+r, kM+j)|^2}{\frac{1}{NM}\mathrm{tr}(\mathbf{R}_{\tilde{z}\tilde{z}}) + E_x \sum_{r=1}^{M}\sum_{i=1, i\neq kM+j}^{MN} |\bar{\mathbf{G}}(kM+r, i)|^2} \\ &= \frac{\sum_{r=1}^{M} |\mathbf{e}_{kM+r}^H \bar{\mathbf{G}} \mathbf{e}_{kM+j}|^2}{(\frac{1}{SNR})(\frac{1}{NM})\mathrm{tr}(\mathbf{W}\mathbf{W}^H) + \sum_{r=1}^{M}\sum_{i=1, i\neq kM+j}^{MN} |\mathbf{e}_{kM+r}^H \bar{\mathbf{G}} \mathbf{e}_i|^2} \end{aligned} \tag{24}$$

[4] Note that in the MIMO case we define inter–carrier interference to include interference from all other frequency bins for the same antenna and interference from all frequency bins (including k^{th}) for other antennas.

Assuming that $\mathbf{w}_l^H \mathbf{w}_l = 1$ for $1 \leq l \leq MN$, we have $\sum_{r=1}^{M} \mathbf{w}_{kM+r}^H \mathbf{w}_{kM+r} = M$ and hence $\mathrm{tr}(\mathbf{W}\mathbf{W}^H) = NM$. Therefore, Equation (24) becomes (using (18))

$$\begin{aligned}
SINR_k^j &= \frac{\sum_{r=1}^{M} \mathbf{e}_{kM+r}^H \mathbf{Q}\mathbf{W}\mathbf{H}\mathbf{Q}^H \mathbf{e}_{kM+j}\mathbf{e}_{kM+j}^H \mathbf{Q}\mathbf{H}^H \mathbf{W}^H \mathbf{Q}^H \mathbf{e}_{kM+r}}{\frac{1}{SNR} + \sum_{r=1}^{M} \mathbf{e}_{kM+r}^H \mathbf{Q}\mathbf{W}\mathbf{H}\mathbf{Q}^H (\sum_{i=1, i\neq kM+j}^{MN} \mathbf{e}_i \mathbf{e}_i^H) \mathbf{Q}\mathbf{H}^H \mathbf{W}^H \mathbf{Q}^H \mathbf{e}_{kM+r}} \\
&= \frac{\sum_{r=1}^{M} \mathbf{q}_{kM+r}^H \mathbf{W}\mathbf{h}_{kM+j}\mathbf{h}_{kM+j}^H \mathbf{W}^H \mathbf{q}_{kM+r}}{\frac{1}{SNR} + \sum_{r=1}^{M} \mathbf{q}_{kM+r}^H \mathbf{W}\mathbf{H}(\mathbf{I}_N - \mathbf{q}_{kM+j}\mathbf{q}_{kM+j}^H)\mathbf{H}^H \mathbf{W}^H \mathbf{q}_{kM+r}} \\
&= \frac{\sum_{r=1}^{M} \mathbf{w}_{kM+r}^H \mathbf{h}_{kM+j}\mathbf{h}_{kM+j}^H \mathbf{w}_{kM+r}}{\frac{1}{SNR} + \sum_{r=1}^{M} \mathbf{w}_{kM+r}^H (\mathbf{H}\mathbf{H}^H - \mathbf{h}_{kM+j}\mathbf{h}_{kM+j}^H)\mathbf{w}_{kM+r}} \\
&= \frac{\sum_{r=1}^{M} \mathbf{w}_{kM+r}^H \mathbf{h}_{kM+j}\mathbf{h}_{kM+j}^H \mathbf{w}_{kM+r}}{\sum_{r=1}^{M} \mathbf{w}_{kM+r}^H (\frac{1}{M.SNR}\mathbf{I}_{NM} + \mathbf{H}\mathbf{H}^H - \mathbf{h}_{kM+j}\mathbf{h}_{kM+j}^H)\mathbf{w}_{kM+r}},
\end{aligned}$$

where we have defined $\mathbf{q}_i = \mathbf{Q}^H \mathbf{e}_i$, $\mathbf{h}_i = \mathbf{H}\mathbf{q}_i$, and $\mathbf{w}_i = \mathbf{W}^H \mathbf{q}_i$. If we further define $\tilde{\mathbf{R}}_{kj} = \frac{1}{M.SNR}\mathbf{I}_{NM} + \mathbf{H}\mathbf{H}^H - \mathbf{h}_{kM+j}\mathbf{h}_{kM+j}^H$ and $\bar{\mathbf{R}}_{kj} = \mathbf{h}_{kM+j}\mathbf{h}_{kM+j}^H$, we get

$$\begin{aligned}
SINR_k^j &= \frac{\sum_{r=1}^{M} \mathbf{w}_{kM+r}^H \bar{\mathbf{R}}_{kj} \mathbf{w}_{kM+r}}{\sum_{r=1}^{M} \mathbf{w}_{kM+r}^H \tilde{\mathbf{R}}_{kj} \mathbf{w}_k} \\
&= \frac{\tilde{\mathbf{w}}_k^H \bar{\bar{\mathbf{R}}}_{kj} \tilde{\mathbf{w}}_k}{\tilde{\mathbf{w}}_k^H \tilde{\tilde{\mathbf{R}}}_{kj} \tilde{\mathbf{w}}_k},
\end{aligned}$$

where we have made the definitions $\bar{\bar{\mathbf{R}}} = diag(\underbrace{\bar{\mathbf{R}}_{kj}, \cdots, \bar{\mathbf{R}}_{kj}}_{M})$, $\tilde{\tilde{\mathbf{R}}} = diag(\underbrace{\tilde{\mathbf{R}}_{kj}, \cdots, \tilde{\mathbf{R}}_{kj}}_{M})$, and $\tilde{\mathbf{w}}_k^H = \underbrace{\begin{bmatrix} \mathbf{w}_{kM+1}^H & \mathbf{w}_{kM+2}^H & \cdots & \mathbf{w}_{(k+1)M}^H \end{bmatrix}}_{M}$. It can be shown that $SINR_k^j$ is maximized by the filters

$$\mathbf{w}_{kM+r,opt} = \alpha \mathbf{R}_{yy}^{-1} \mathbf{h}_{kM+r} : \text{ for } 0 \leq k \leq N-1 \; ; \; 1 \leq r \leq M \; , \tag{25}$$

where α is a normalizing constant and $\mathbf{R}_{yy} = \frac{1}{M.SNR}\mathbf{I}_{NM} + \mathbf{H}\mathbf{H}^H$. Finally, the optimum filter $\mathbf{W}_{opt}$ is given by

$$\mathbf{W}_{opt} = \mathbf{Q}^H \begin{bmatrix} \mathbf{w}_{1,opt}^H \\ \mathbf{w}_{2,opt}^H \\ \vdots \\ \mathbf{w}_{NM,opt}^H \end{bmatrix}.$$

It is worth emphasizing that the filter $\mathbf{W}_{opt}$ calculated above performs joint inter-carrier and inter-antenna interference suppression. With zero Doppler, the matrix

$\mathbf{H}$ will be *block circulant*, hence the matrix $\mathbf{QHQ}^H$ will be *block diagonal* (i.e., inter–antenna interference is still present). However, the matrix $\mathbf{QW}_{opt}\mathbf{HQ}^H$ will be (approximately) *diagonal* indicating that the matrix filter $\mathbf{W}_{opt}$ in this case acts also as a linear "multi–user" detector that suppresses inter–antenna interference.

D. Space-Time-Coded OFDM

In this section, we will explore a special case of coded MIMO OFDM, namely a space-time block-coded OFDM with two transmit and one receive antennas, i.e., $M_t = 2$ and $M_r = 1$ (the analysis can be extended to the case $M_t > 2$ and $M_r > 1$). We study a space-time block-coded OFDM system based on the the Alamouti code [21] (as first proposed in [22]): though details can be found in [22], the basic mechanism is as follows. We consider two source symbols $c_1(m)$, $c_2(m)$ which, in a conventional OFDM system, would be transmitted over two consecutive OFDM blocks on the same subcarrier m. Similar to the Alamouti code, the two source symbols are mapped as

$$\mathbf{X}_{(1)}(m) = [c_1(m), c_2(m)]^T, \;\; \mathbf{X}_{(2)}(m) = [-c_2^*(m), c_1^*(m)]^T \,, \tag{26}$$

where $\mathbf{X}_{(1)}$ represents the information-bearing vector for the first OFDM block and $\mathbf{X}_{(2)}$ corresponds to the second OFDM block[5].

Our goal is to understand the effect of channel time-variation over such a transmission scheme, and the performance of optimal MMSE receiver structures. Using (12) we can write the output of two consecutive OFDM blocks as

$$\mathbf{Y}_{(1)}(m) = \mathbf{G}_{(1)}(m,m)\mathbf{X}_{(1)}(m) + \sum_{\substack{n=0 \\ n\neq m}}^{N-1} \mathbf{G}_{(1)}(m,n)\mathbf{X}_{(1)}(n) + \mathbf{Z}_{(1)}(m) \tag{27a}$$

$$\mathbf{Y}_{(2)}(m) = \mathbf{G}_{(2)}(m,m)\mathbf{X}_{(2)}(m) + \sum_{\substack{n=0 \\ n\neq m}}^{N-1} \mathbf{G}_{(2)}(m,n)\mathbf{X}_{(2)}(n) + \mathbf{Z}_{(2)}(m) \,, \tag{27b}$$

[5] Intuitively, each OFDM subcarrier can be thought of as a flat-fading channel and the Alamouti code is applied to each of the OFDM subcarriers. As a result, the Alamouti code yields diversity gains at every subcarrier.

and, with the aid of (26), obtain

$$\begin{bmatrix} \mathbf{Y}_{(1)}(m) \\ \mathbf{Y}_{(2)}{}^*(m) \end{bmatrix} = \begin{bmatrix} [\mathbf{G}_{(1)}(m,m)]_{1,1} & [\mathbf{G}_{(1)}(m,m)]_{1,2} \\ [\mathbf{G}_{(2)}(m,m)]_{1,2}^* & -[\mathbf{G}_{(2)}(m,m)]_{1,1}^* \end{bmatrix} \begin{bmatrix} c_1(m) \\ c_2(m) \end{bmatrix}$$
$$+ \sum_{\substack{n=0 \\ n\neq m}}^{N-1} \begin{bmatrix} [\mathbf{G}_{(1)}(m,n)]_{1,1} & [\mathbf{G}_{(1)}(m,n)]_{1,2} \\ [\mathbf{G}_{(2)}(m,n)]_{1,2}^* & -[\mathbf{G}_{(2)}(m,n)]_{1,1}^* \end{bmatrix} \begin{bmatrix} c_1(n) \\ c_2(n) \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{(1)}(m) \\ \mathbf{Z}_{(2)}{}^*(m) \end{bmatrix}, \quad (28)$$

which can be written as

$$\mathbf{Y}(m) \stackrel{def}{=} \mathbf{G}(m,m)\mathbf{c}(m) + \sum_{n=0, n\neq m}^{N-1} \mathbf{G}(m,n)\mathbf{c}(n) + \mathbf{Z}(m) .$$

Now it can be seen that the mathematical structure of the Alamouti-coded OFDM is identical to that of the MIMO OFDM (with ICI) structure defined in (12) and, therefore, the MMSE analysis done in the previous section carries through with the appropriate interpretations of notation. Furthermore, the derivation of the optimum linear MIMO ICI cancellation filter $\mathbf{W}$ is straightforward (by applying the results of Section IV-C). In this case, $\mathbf{W}$ is a $2N \times 2N$ matrix as it processes $2N$ tones over $2N$ symbol durations.

Computational complexity can be reduced by defining a per-tone 2×2 filter $\mathbf{W}_m$ which suboptimally attempts to minimize ICI by processing only $\mathbf{Y}_{(1)}(m)$ and $\mathbf{Y}_{(2)}(m)$. With time-invariant channels, and in the absence of noise, it can be seen that $\mathbf{W}_m$ reduces to

$$\mathbf{W}_m = \begin{bmatrix} [\mathbf{G}_{(1)}(m,m)]_{1,1} & [\mathbf{G}_{(1)}(m,m)]_{1,2} \\ [\mathbf{G}_{(2)}(m,m)]_{1,2}^* & -[\mathbf{G}_{(2)}(m,m)]_{1,1}^* \end{bmatrix}^* ,$$

which coincides with the Alamouti-decoding matrix. Note that with time-invariant channels, in Equation (28) all "off-diagonal" terms become zero, $[\mathbf{G}_{(1)}(m,m)]_{1,1} = [\mathbf{G}_{(2)}(m,m)]_{1,1}$, and $[\mathbf{G}_{(1)}(m,m)]_{1,2} = [\mathbf{G}_{(2)}(m,m)]_{1,2}$.

V. Channel Estimation and Tracking

In this section, we address the challenging problem of channel estimation in an OFDM system with severe Doppler and we propose a promising technique based on channel interpolation and judiciously-placed pilot tones. Throughout this section, we assume $M_r = M_t = 1$.

In a time-varying environment, estimation of $\mathbf{H}$ amounts to estimating N channels $\mathbf{h}_n := [h(n;0), \ldots, h(n;\nu-1)]^T$, $0 \leq n \leq N-1$, that comprise the rows of $\mathbf{H}$; in other words, we need to estimate $N\nu$ parameters. This number of parameters can be significantly reduced if a priori knowledge about the channel dynamics is available (see, e.g., [23]), or if it is assumed that the channel is quasi-static (i.e., constant within an OFDM block, but varying from block to block). Although it has been shown that the aforementioned methods have provided promising results, both of them rely on assumptions which do not necessarily hold true in the OFDM environment considered in this work. On the one hand, we have made no assumptions about the underlying channel structure; on the other hand, our work addresses scenarios where there is significant ICI and, as a result, the channel cannot be assumed quasi-static.

Given the minimal set of assumptions on the structure of the underlying channel, the estimation of $\mathbf{H}$ appears, at first sight, to be a daunting task. Even when a full training block is used (in other words, all OFDM subcarriers are pilot tones), we have available only N values for the estimation of $N\nu$ parameters. To reduce the number of parameters needed for channel estimation from $N\nu$ to less than N, we make the reasonable assumption that some of the channels $\mathbf{h}_n$ can be obtained by linear interpolation. Such an assumption holds true if there is not significant variation between channels $\mathbf{h}_n$ and $\mathbf{h}_{n+1}$, $0 \leq n \leq (N-2)$, and as we illustrate later on, channel interpolation does not have a significant negative impact on the achievable performance of our methods. Furthermore, channel interpolation transforms a seemingly very difficult problem to a tractable one, as we will shortly see.

The basic idea of using interpolation to reduce the number of parameters is as follows. We parameterize the matrix $\mathbf{H}$ by using a small number of its rows. We then express the entire matrix as a function of these rows and therefore reducing the number of parameters to be estimated. Physically, this puts "markers" in time where the channel is estimated, and the estimates at other times are interpolated using these estimates. Let us consider M channels $\mathbf{h}_{m(1)}, \ldots, \mathbf{h}_{m(M)}$ which form M rows of the channel matrix $\mathbf{H}$; let us denote by $\mathcal{M} := \{m(1), \ldots, m(M)\}$ the set of these rows. Our method is based on the assumption that each channel $\mathbf{h}_n$, $n \notin \mathcal{M}$, can be expressed as a linear combination of the M channels $\mathbf{h}_{m(1)}, \ldots, \mathbf{h}_{m(M)}$, i.e.,

$$h(n;l) = \mathbf{a}_n^T \left[h(m(1);l), \ldots, h(m(M);l)\right]^T , \quad 0 \le l \le \nu - 1 . \tag{29}$$

For example, when the channel process $\{\mathbf{h}_n\}$ is Gaussian, then such a linear interpolation is optimal in the MMSE sense. Note that in (29) we have used the same $M \times 1$ weight vector $\mathbf{a}_n$ for all of the ν channel taps of $\mathbf{h}_n$: using the same weight vector is a direct consequence of the assumption that $h(n;0), \ldots, h(n;\nu-1)$ are i.i.d. processes for every n (but there is correlation between $h(n;l)$ and $h(m;l)$ for $n \neq m$).

In Section V-C, we present guidelines on the choice of the weight vectors $\mathbf{a}_n$'s. In the following section, we study how channel estimation can be accomplished using pilot tones.

A. *Channel Estimation Using Pilot Tones*

Let us denote by $H_C(\mathbf{h}_m)$ the $N \times N$ circulant matrix which we would have in the OFDM system if the underlying channel was fixed and equal to $\mathbf{h}_m$, i.e.,

$$H_C(\mathbf{h}_m) := \begin{bmatrix} h(m;0) & 0 & \cdots & h(m;2) & h(m;1) \\ h(m;1) & h(m;0) & \cdots & h(m;3) & h(m;2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h(m;\nu-1) & h(m;\nu-2) & \vdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & h(m;0) & 0 \\ 0 & 0 & \cdots & h(m;1) & h(m;0) \end{bmatrix} . \tag{30}$$

Using (29), we can see that the channel matrix $\tilde{\mathbf{H}}$ obtained by the interpolation of the channels $\mathbf{h}_m(1), \ldots, \mathbf{h}_m(M)$ is

$$\tilde{\mathbf{H}} = \sum_{1 \le i \le M} \mathbf{A}_{m(i)} H_C(\mathbf{h}_{m(i)}) , \tag{31}$$

with $\mathbf{A}_{m(i)}$ an $N \times N$ diagonal matrix with entries

$$[\mathbf{A}_{m(i)}]_{n,n} = \begin{cases} 1 , & n = m(i) \\ 0 , & n \in \mathcal{M} \text{ and } n \neq m(i) \\ a_n(m(i)) , & \text{otherwise} \end{cases} . \tag{32}$$

Suppose $N = 3$, $\nu = 2$, $\mathcal{M} = \{0, 2\}$, $\mathbf{a}_2 = (0.3, 0.7)^T$. Then, $\tilde{\mathbf{H}}$ is given by

$$\begin{bmatrix} h(0;0) & 0 & h(0;1) \\ .3h(0;1) + .7h(2;1) & .3h(0;0) + .7h(2;0) & 0 \\ 0 & h(2;1) & h(2;0) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h(0;0) & 0 & h(0;1) \\ h(0;1) & h(0;0) & 0 \\ 0 & h(0;1) & h(0;0) \end{bmatrix}$$
$$+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & .7 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h(2;0) & 0 & h(2;1) \\ h(2;1) & h(2;0) & 0 \\ 0 & h(2;1) & h(2;0) \end{bmatrix}$$

and $\mathbf{H}$, $\tilde{\mathbf{H}}$ differ only in the second "interpolated" row.

Fig. 1. Example of interpolated matrix

An example of this formulation is given in Fig. 1.

Given the structured form of $\tilde{\mathbf{H}}$, its estimation amounts to estimating $M\nu$ parameters grouped in the $M\nu \times 1$ vector

$$\tilde{\mathbf{h}} := [\mathbf{h}_{m(1)}^T, \ldots, \mathbf{h}_{m(M)}^T]^T \,. \tag{33}$$

To perform channel estimation using pilot tones, we need to express $\tilde{G}(m, s) := [\mathbf{Q}\tilde{\mathbf{H}}\mathbf{Q}^H]_{m,s}$ as a function of $\tilde{\mathbf{h}}$. From (31), we can express $\tilde{G}(m, s)$ as

$$\tilde{G}(m, s) = \sum_{1 \le i \le M} [\mathbf{Q}\mathbf{A}_{m(i)} H_C(\mathbf{h}_{m(i)})\mathbf{Q}^H]_{m,s} \,. \tag{34}$$

We can easily evaluate

$$[\mathbf{Q}\mathbf{A}_{m(i)} H_C(\mathbf{h}_{m(i)})\mathbf{Q}^H]_{m,s} = \frac{1}{N} \sum_{r=0}^{N-1} \sum_{l=0}^{\nu-1} [\mathbf{A}_{m(i)}]_{r,r} \, h(m(i); l) \, e^{j\frac{2\pi r(s-m)}{N}} \, e^{-j\frac{2\pi s l}{N}} \,, \tag{35}$$

which implies that $h(m(i); l)$ is scaled by

$$b_{m(i)}^{m,s}(l) := e^{-j\frac{2\pi s l}{N}} \, \frac{1}{N} \sum_{r=0}^{N-1} [\mathbf{A}_{m(i)}]_{r,r} \, e^{j\frac{2\pi r(s-m)}{N}} \,. \tag{36}$$

Hence, by defining the $1 \times \nu$ row vector

$$\mathbf{b}_{m(i)}^{m,s} := \left[b_{m(i)}^{m,s}(0), \ldots, b_{m(i)}^{m,s}(\nu - 1) \right] \,, \tag{37}$$

we can write

$$[\mathbf{Q}\mathbf{A}_{m(i)} H_C(\mathbf{h}_{m(i)})\mathbf{Q}^H]_{m,s} = \mathbf{b}_{m(i)}^{m,s} \mathbf{h}_{m(i)} \,. \tag{38}$$

By defining the $1 \times M\nu$ row vector $\mathbf{b}^{m,s} := (\mathbf{b}^{m,s}_{m(1)}, \ldots, \mathbf{b}^{m,s}_{m(M)})$, it follows from (34), (38) that

$$\tilde{G}(m,s) = \mathbf{b}^{m,s}\tilde{\mathbf{h}} \,. \tag{39}$$

If the channel matrix was $\tilde{\mathbf{H}}$ (instead of $\mathbf{H}$), then the received tone $Y(m)$ would be

$$Y(m) = \tilde{G}(m,m)X(m) + \underbrace{\sum_{n \neq m} \tilde{G}(m,n)X(n)}_{\text{ICI}} + \text{noise} \,. \tag{40}$$

Inspired by how pilot-based channel estimation is done in OFDM systems, one could proceed as follows

1. Ignore the ICI term and assume that $Y(p) = \tilde{G}(p,p)X(p) + \text{noise} \overset{(39)}{=} X(p)\mathbf{b}^{p,p}\tilde{\mathbf{h}} + \text{noise}$.
2. Assuming that there are P pilot tones, which are placed at subcarriers $p(1), \ldots, p(P)$, form the $P \times M\nu$ system of linear equations

$$\begin{bmatrix} Y(p(1)) \\ \vdots \\ Y(p(P)) \end{bmatrix} = \begin{bmatrix} X(p(1))\mathbf{b}^{p(1),p(1)} \\ \vdots \\ X(p(P))\mathbf{b}^{p(P),p(P)} \end{bmatrix} \tilde{\mathbf{h}} + \text{noise} \,,$$

or $\mathbf{Y}_{(P)} = \mathbf{B}_{(P)}\tilde{\mathbf{h}} + \text{noise}$.

3. With P chosen such that $P \geq M\nu$, obtain $\tilde{\mathbf{h}}$ as the least–squares solution of the aforementioned system of linear equations, i.e., $\hat{\tilde{h}} = \mathbf{B}^{\dagger}_{(P)}\mathbf{Y}_{(P)}$, with $\mathbf{B}^{\dagger}_{(P)}$ the pseudoinverse of $\mathbf{B}_{(P)}$.

Unfortunately, this approach does not produce reliable channel estimates, because $\mathbf{B}_{(P)}$ is not full column rank. As we prove in the appendix, $\mathbf{B}_{(P)}$ has rank ν irrespective of the choice of interpolation coefficients $\{\mathbf{a}_n\}_{n \notin \mathcal{M}}$. Consequently, channel estimation is not possible if we look only at "isolated" pilot tones. Intuitively, this is not a surprising result as we study a system where ICI is not negligible – hence, we should account for the ICI in the channel estimation process. The important *practical implication* of this fact is that in OFDM systems with severe ICI, pilot tones should not be placed far apart from each other, but rather they should be grouped together. Before we explore the reasons for this grouping of pilot tones, let us note the contrast with OFDM in time-invariant or slowly time-varying environments, where it has been shown that the pilot tones should be equi-spaced [24], [25].

Revisiting (40) for a pilot tone p, we obtain

$$Y(p) = \tilde{G}(p,p)X(p) + \sum_{\substack{q \text{ pilot} \\ q \neq p}} \tilde{G}(m,q)X(q) + \sum_{n \text{ not pilot}} \tilde{G}(m,n)X(n) + \text{noise}\,, \tag{41}$$

which yields

$$Y(p) = \tilde{\mathbf{b}}^p \tilde{\mathbf{h}} + \sum_{n \text{ not pilot}} \tilde{G}(m,n)X(n) + \text{noise}\,, \quad \tilde{\mathbf{b}}^p := \left(\sum_{q \text{ pilot}} X(q)\mathbf{b}^{p,q} \right) . \tag{42}$$

Equation (42) is the basis of the following channel estimation method

1. Assume that for every pilot tone p

$$Y(p) = \tilde{\mathbf{b}}^p \tilde{\mathbf{h}} + e(p)\,, \quad e(p) := \left(\sum_{q \text{ not pilot}} X(q)\mathbf{b}^{p,q} \right) \tilde{\mathbf{h}} + Z(p)\,. \tag{43}$$

2. Form the $P \times M\nu$ system of linear equations

$$\begin{bmatrix} Y(p(1)) \\ \vdots \\ Y(p(P)) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{b}}^{p(1),p(1)} \\ \vdots \\ \tilde{\mathbf{b}}^{p(P),p(P)} \end{bmatrix} \tilde{\mathbf{h}} + \begin{bmatrix} e(p(1)) \\ \vdots \\ e(p(P)) \end{bmatrix} ,$$

or $\mathbf{Y}_{(P)} = \tilde{\mathbf{B}}_{(P)} \tilde{\mathbf{h}} + \mathbf{e}_{(P)}$.

3. Obtain $\tilde{\mathbf{h}}$ as the least squares solution of the aforementioned system of linear equations, $\hat{\tilde{h}} = \tilde{\mathbf{B}}^{\dagger}_{(P)} \mathbf{Y}_{(P)}$.

Extension of this technique to MIMO OFDM is straightforward.

B. Placement of Pilot Tones

Unlike time-invariant or slowly-time varying OFDM systems where pilot tones should be equi-spaced [24], [25], our proposed channel estimation method implies that pilot tones should be grouped together (but in more than one group). In this section, we quantify this observation and propose a pilot tone placement scheme. For analytical tractability, we start from the case $\nu = 1$ which corresponds to channels which are time-varying but not frequency-selective. Work on PSAM (Pilot Symbol Assisted Modulation) [26] has suggested that, for Rayleigh flat-fading channels, pilot symbols should be placed periodically in the time domain to produce channel estimates; the coherent detection of the transmitted symbols is based on the interpolation of these channel estimates. The periodic transmission of pilot symbols in the time domain suggests a grouping of pilot tones in the frequency domain – an intuitive idea explored by the ensuing discussion.

We adopt as figure of merit the approximation error $e(p)$ in (43) and we seek the pilot tone placement scheme which minimizes $\mathbf{E}\left[||e(p)||^2\right]$ for a specific pilot tone p. Assuming that the transmitted symbols are i.i.d. with $\mathbf{E}\left[||X(q)||^2\right] = 1$, we obtain

$$\mathbf{E}\left\{||e(p)||^2\right\} = \left(\sum_{q \text{ not pilot}} \mathbf{b}^{p,q} \mathbf{R}_{\tilde{h}\tilde{h}} (\mathbf{b}^{p,q})^H \right) + \sigma^2 \,, \tag{44}$$

where $\mathbf{R}_{\tilde{h}\tilde{h}} := \mathbf{E}\left[\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H\right]$ and $\sigma^2 = \mathbf{E}\left[||Z(p)||^2\right]$ is the variance of the AWGN (which is the same across all subcarriers). In the appendix we show that $\mathbf{b}^{p,q}\mathbf{R}_{\tilde{h}\tilde{h}}(\mathbf{b}^{p,q})^H$ is a decreasing function of $((p-q) \mod N)$. Consequently, $\mathbf{E}\left[||e(p)||^2\right]$ is minimized if we place all pilot tones $q \neq p$ as close to p as possible. It is implied that for time-selective channels, pilot tones should be grouped together. On the other hand, in frequency-selective time-invariant channels, placing the pilot tones equispaced on the FFT grid is the optimal scheme [24], [25]. Intuitively, for frequency-selective time-varying channels, the pilot tones should be grouped together, *but* in more than one group. As we discuss in the simulations section, we have found that our channel estimation method indeed produces the best results when the *pilot tones are partitioned into equispaced groups on the FFT grid.*

C. Selection of the Interpolation Weight Vectors $\mathbf{a}_n$

Without imposing any assumptions on the underlying channel variations, linear interpolation appears to be the simplest method for choosing the weight vectors $\mathbf{a}_n$. On the other hand, if a priori knowledge about the underlying channel model is available, more sophisticated channel interpolation schemes can be devised. Let us suppose that the channels follow the Jakes model [27]: $\mathbf{E}\left[(h(m;l)h^H(n;l))\right] = J_o(2\pi f_d(m-n)T)$, with f_d denoting the Doppler frequency, and T denoting the symbol period. Then, the calculation of the interpolation weights is straightforward. For example, if we fix rows $\mathbf{h}_0$, $\mathbf{h}_{N/2-1}$, $\mathbf{h}_{N-1}$, the set of weights $\mathbf{a}_n = [a_n(1), a_n(N/2-1), a_n(N-1)]^T$ which minimizes the error $\mathbf{E}[|h(n;l) - \mathbf{a}_n^H \tilde{\mathbf{h}}(l)|^2]$ with $\tilde{\mathbf{h}}(l) := [h(1;l), h(N/2-1;l), h(N-1;l)]^T$ can be obtained using the the orthogonality principle

$$\mathbf{a}_n^H = \mathbf{R}_{h_n\tilde{h}} \mathbf{R}_{\tilde{h}\tilde{h}}^{-1} \,,$$

where $\mathbf{R}_{h_n\tilde{h}} = \mathbf{E}[h(n;l)\tilde{\mathbf{h}}^H(l)]$ and $\mathbf{R}_{\tilde{h}\tilde{h}} = \mathbf{E}[\tilde{\mathbf{h}}(l)\tilde{\mathbf{h}}^H(l)]$. With $J_o[n] := J_o(2\pi f_d nT)$, it can be seen that

$$\begin{aligned} \mathbf{R}_{h_n\tilde{h}} &= [J_o[n], J_o[|N/2-1-n|], J_o[N-1-n]] \\ \mathbf{R}_{\tilde{h}\tilde{h}} &= \begin{bmatrix} 1 & J_o[N/2-1] & J_o[N-1] \\ J_o[N/2-1] & 1 & J_o[N/2] \\ J_o[N-1] & J_o[N/2] & 1 \end{bmatrix} . \end{aligned} \tag{45}$$

As we demonstrate in the simulations section, for Doppler values of practical importance (i.e., less than 200Hz), there is little to be gained by adopting the Jakes-based estimator in place of the linear interpolator. Hence, from an implementation point of view, the linear estimator appears to be an attractive solution as it dispenses with the estimation of the Doppler frequency without sacrificing performance.

D. Channel Tracking

Under relatively mild Doppler, the quality of the channel estimates (and, consequently, the SINR gains of our ICI mitigating methods) can be improved by channel tracking. A simple tracking scheme is the following: an initial channel estimate $\hat{\mathbf{H}}_{(0)}$ can be obtained by transmitting a full training block; subsequent blocks contain pilot tones which are used to acquire new estimates $\mathbf{H}'_{(n)}$. The channel estimate for the n^{th} OFDM block is obtained using a forgetting factor α as follows $\hat{\mathbf{H}}_{(n)} = \alpha\hat{\mathbf{H}}_{(n-1)} + (1-\alpha)\mathbf{H}'_{(n)}$ (see, e.g., [24] for a similar method in SISO OFDM). Frequent re-training can further improve the quality of the channel estimates at the expense of the overhead of the training symbols. As we illustrate in the simulations section, the aforementioned tracking technique results in higher SINR gains (as channel estimates become more accurate).

VI. Numerical Results

In this section, we present numerical examples which illustrate the SINR improvement due to the $\mathbf{W}$ filter in both SISO and MIMO OFDM systems. As the design of $\mathbf{W}$ depends on channel knowledge, we also study the accuracy of our channel estimation method and the penalty it incurs on the SINR. Throughout our simulation experiments, we investigate an OFDM system with $N = 64$ subcarriers, a total bandwidth of 200KHz, carrier frequency $f_c = 2.4$GHz, and the velocity of the mobile is set to 60MPH (the corresponding Doppler frequency is 214Hz). In all simulations, we have chosen as figure of merit the average $SINR$ gain, which is defined as

$$SINR = \left(\frac{\prod_{m=0}^{N-1} SINR_m^W}{\prod_{m=0}^{N-1} SINR_m} \right)^{1/N} ,$$

with $SINR_m^W$ ($SINR_m$) being the SINR at subcarrier m with (without) the filter $\mathbf{W}$. The $SINR$ metric expresses the SINR improvement over all subcarriers and since it depends on the channel realization, in all figures we depict the average SINR over 1000 random channel matrices $\mathbf{H}$. The entries of $\mathbf{H}$ were generated using the Jakes model, with $\nu = 4$, and each channel tap modeled as a complex Gaussian random variable $\mathcal{N}(0, 1)$.

SINR improvements with known channel: First, we look at the SINR improvement due to the $\mathbf{W}$ filter in SISO OFDM with perfect channel knowledge. Fig. 2 depicts, for various values of SNR, the SINR gain achieved with perfect channel knowledge for the carrier frequencies of $f_c = 2.4$GHz and $f_c = 60$GHz. For smaller values of the SNR, the SINR improvement is smaller, as $\mathbf{W}$ attempts to suppress both ICI and additive noise. As the SNR increases, the SINR improvement increases, because $\mathbf{W}$ attempts to suppress (essentially) only the ICI. It is also evident that at $f_c = 60$GHz, the ICI is much more severe and, consequently, $\mathbf{W}$ results in higher SINR improvements over a wider range of SNR. In both cases, the substantial SINR gains illustrate the merits of our ICI mitigation technique.

SINR improvements with estimated channel: Fig. 3 illustrates the performance penalty due to channel estimation ($f_c = 2.4$GHz). The channel estimates are obtained using the method described in Section V with $M = 2$ (the top and the bottom row of $\mathbf{H}$

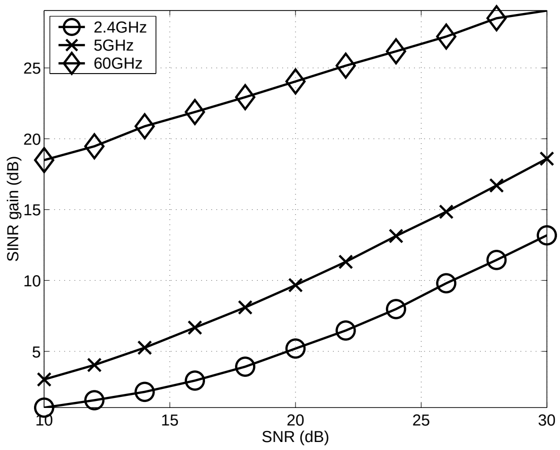

Fig. 2. Gains with known channel

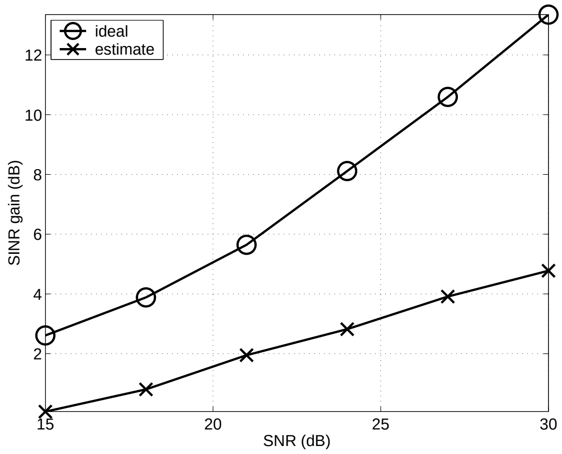

Fig. 3. Gains with estimated channel

are estimated, and the rest of the rows are obtained with linear interpolation). We have used 4ν pilot tones grouped into ν groups which are equispaced onto the FFT grid (i.e., subcarriers $\{6, 7, 8, 9, 22, 23, 24, 25, 38, 39, 40, 41, 54, 55, 56, 57\}$ are pilot tones). From Fig. 3 we observe that even with imperfect channel estimation, significant SINR gains can be achieved – this observation serves as testament to the effectiveness of our channel estimation technique. Additionally, under relatively mild Doppler (90Hz), 2-3dB of SNR can be gained by using frequent re-training and channel tracking, as Fig. 4 illustrates.

Pilot Tone Placement Scheme: Fig. 5 depicts the average SINR gains when $\mathbf{W}$ is based on channel estimation and a total of 4ν pilot tones are placed on the FFT grid according to a number of different schemes

• In scheme "A", the pilot tones $\{1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45, 49, 53, 57, 61\}$ are placed equispaced on the FFT grid.

• In scheme "B", the pilot tones $\{24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39\}$ are placed into a big group of subcarriers centered on the FFT grid.

• In scheme "C", the pilot tones $\{13, 14, 15, 16, 17, 18, 19, 20, 45, 46, 47, 48, 49, 50, 51, 52\}$ are placed into $M = 2$ groups of 2ν tones.

• In scheme "D", the pilot tones $\{6, 7, 8, 9, 22, 23, 24, 25, 38, 39, 40, 41, 54, 55, 56, 57\}$ are placed into ν groups of $2M$ tones.

It is clear from Fig. 5 that the placement of pilot tones has a significant impact on the

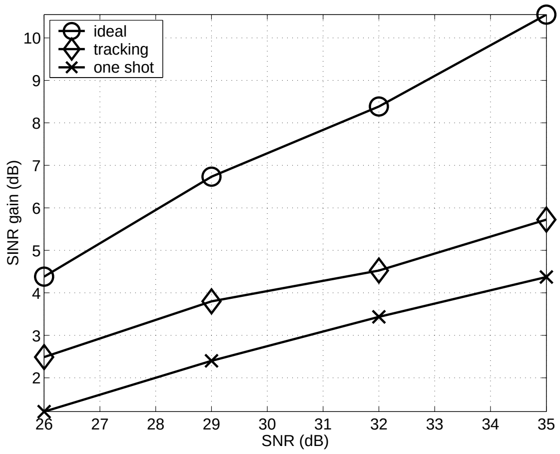

Fig. 4. Channel Tracking

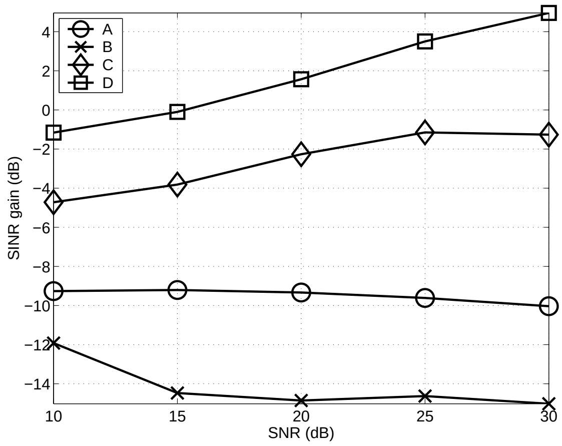

Fig. 5. Pilot Tone Placement

quality of the channel estimates and the resulting SINR gains. It is worth noting that placing the pilot tones equispaced on the FFT grid results in very poor performance unlike the time-invariant case [25], [24].

Interpolation Technique: Though pilot placement does have a significant impact on the achievable SINR gains, this does not seem to be necessarily the case with the choice of interpolation weights. To illustrate this observation, let us look at the case $\mathcal{M} = \{0, N/2, N-1\}$ (i.e., the top, middle, and bottom rows are to be estimated and the rest of the rows are obtained using interpolation). We compare the SINR gains obtained with linear interpolation weights and the performance obtained using the weights of Section V-C. To isolate the effects of the choice of the interpolation weights, we let all subcarriers be pilot tones (which amounts to having a full training OFDM block utilized for channel estimation), and Fig. 6 depicts the performance of the two interpolation methods. We observe that linear interpolation achieves the same performance as the "Jakes-based" interpolation which, intuitively, is a not a surprising result: for Doppler frequency values of practical interest, the variation of the rows of $\mathbf{H}$ is not very large and, as a result, linear interpolation performs adequately. On the other hand, for significant variation, the Jakes-based estimator outperforms the linear interpolator. For illustration purposes only, we look at the case of 858Hz Doppler. Fig. 6 depicts the achieved performance at this extremely high Doppler and it is evident that the Jakes based-estimator outperforms the linear interpolator. Also, it can

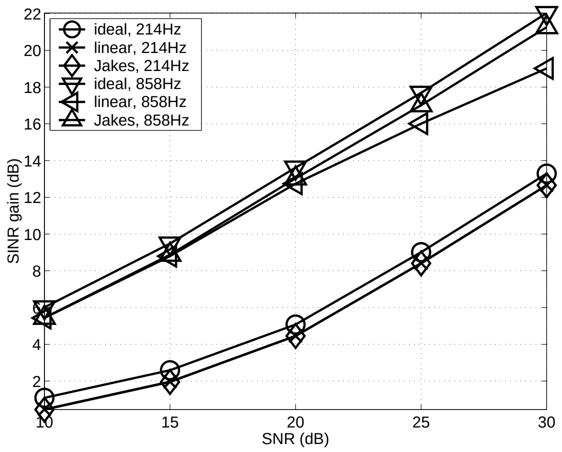

Fig. 6. Interpolation Scheme

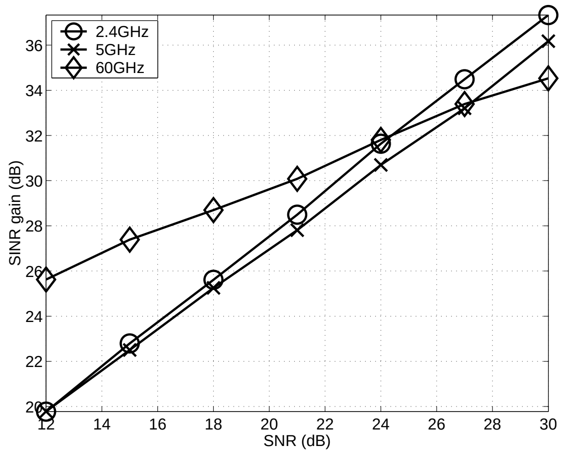

Fig. 7. SINR gains in Space-Time-coded OFDM

be deduced from Fig. 6 that even at high SNR's where high quality of the channel estimate is expected (with a full training block), there is significant difference between the performance achieved with perfect channel knowledge and channel estimates. This is due to the large variation across the channel matrix rows which invalidates the assumption that some channel rows can be expressed as a function of as few as three channel rows.

SINR improvements in MIMO OFDM: Finally, Fig. 7 illustrates the SINR gains achieved in space-time block-coded OFDM transmissions where the channel is perfectly known at the receiver. As described in Section IV-D, the filter $\mathbf{W}$ of Section IV-C operates on two consecutive transmission blocks and we define the $SINR$ as the geometric average of the gains achieved over the tones of the two blocks, i.e.,

$$SINR = \sqrt{SINR_{(1)} SINR_{(2)}}\,.$$

Similar to Fig. 2, we can clearly see the SNR improvements that the MIMO filter $\mathbf{W}$ yields.

VII. Conclusions

In this paper we have analyzed the effect of inter-carrier interference on MIMO-OFDM. The cause of such an impairment is time-variation within a transmission block and occurs in practice due to both Doppler spread of the transmission channel and synchronization errors. We proposed time-domain filtering-based ICI mitigation techniques which cascade the time-varying channel with a receive filter so that the overall channel is approximately (in the appropriate mean-square-error metric) time-invariant. The associated problem of estimating such a rapidly-varying channel was addressed and a novel estimation scheme was proposed. Coupled to the proposed channel estimation scheme is the issue of pilot tone placement. Examination of this issue showed that non-uniform placement of pilot tones and in particular grouping of tones into clumps equispaced on the FFT grid is effective for these situations. Numerical results illustrating these techniques were also presented.

APPENDIX

I. $\mathbf{B}_{(P)}$ HAS RANK ν

From (36), we obtain

$$\begin{aligned} b^{p,p}_{m(i)}(l) &= e^{-j\frac{2\pi pl}{N}} \frac{1}{N} \sum_{r=0}^{N-1} [A_{m(i)}]_{r,r} \\ &= \frac{\frac{1}{N}\sum_{r=0}^{N-1}[A_{m(i)}]_{r,r}}{\frac{1}{N}\sum_{r=0}^{N-1}[A_{m(1)}]_{r,r}}\, b^{p,p}_{m(1)} \\ &= c_{m(i)}\, b^{p,p}_{m(1)}(l)\ , \quad c_{m(i)} := \frac{\frac{1}{N}\sum_{r=0}^{N-1}[A_{m(i)}]_{r,r}}{\frac{1}{N}\sum_{r=0}^{N-1}[A_{m(1)}]_{r,r}}\ . \end{aligned}$$

Hence, with the aid of (37), we obtain $\mathbf{b}^{p,p}_{m(i)} = c_{m(i)} \mathbf{b}^{p,p}_{m(1)}$, which readily yields that

$$\mathbf{B}_{(P)} = \begin{pmatrix} \mathbf{b}^{p(1),p(1)}_{m(1)} & \dots & \mathbf{b}^{p(1),p(1)}_{m(i)} & \dots & \mathbf{b}^{p(1),p(1)}_{m(M)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{b}^{p(P),p(P)}_{m(1)} & \dots & \mathbf{b}^{p(P),p(P)}_{m(i)} & \dots & \mathbf{b}^{p(P),p(P)}_{m(M)} \end{pmatrix}$$

has rank ν. □

II. $\mathbf{b}^{p,q}\mathbf{R}_{\tilde{h}\tilde{h}}(\mathbf{b}^{p,q})^*$ AS A FUNCTION OF $(p-q)$

For analytical tractability, let us focus on the case $\nu = 1$ and $\mathcal{M} = \{0, N-1\}$. From (33) we have $\tilde{\mathbf{h}} := (h_0(0), h_{N-1}(0))^T$ and $\mathbf{R}_{\tilde{h}\tilde{h}}$ can be evaluated as

$$\mathbf{R}_{\tilde{h}\tilde{h}} = \begin{bmatrix} 1 & \gamma \\ \gamma & 1 \end{bmatrix}\ , \quad 0 < \gamma < 1\ .$$

Then, the term $\beta^{p,q}$:=$\mathbf{b}^{p,q}\mathbf{R}_{\tilde{h}\tilde{h}}(\mathbf{b}^{p,q})^*$ is written as (c.f. (37))

$$\begin{aligned}\beta^{p,q} &= \left(\begin{array}{cc} b_0^{p,q}(0) & b_{N-1}^{p,q}(0) \end{array} \right) \left[\begin{array}{cc} 1 & \gamma \\ \gamma & 1 \end{array} \right] \left[\begin{array}{c} (b_0^{p,q}(0))^* \\ (b_{N-1}^{p,q}(0))^* \end{array} \right] \\ &= |b_0^{p,q}(0)|^2 + |b_{N-1}^{p,q}(0)|^2 + \gamma\big((b_0^{p,q}(0))^* b_{N-1}^{p,q}(0) + (b_0^{p,q}(0))(b_{N-1}^{p,q}(0))^*\big) \, . \end{aligned} \tag{46}$$

Without any assumptions on the underlying physical channels, we adopt the linear interpolation scheme for the selection of the weights $\mathbf{a}_n$ of (29); hence, we set

$$a_n = \left(\begin{array}{cc} \frac{N-1-n}{N-1} & \frac{n}{N-1} \end{array} \right)^T \, , \quad 1 \leq n \leq N-2 \, . \tag{47}$$

With the aid of (32), (36), we can write $b_0^{p,q}(0)$, $b_{N-1}^{p,q}(0)$ in closed–form as follows

$$\begin{aligned} b_0^{p,q}(0) &= \frac{1}{N}\sum_{r=0}^{N-1}[\mathbf{A}_0]_{r,r}\, e^{j\frac{2\pi r(p-q)}{N}} \\ &\overset{(32),(47)}{=} \frac{1}{N}\sum_{r=0}^{N-1}(1-\frac{r}{N-1})e^{j\frac{2\pi r(p-q)}{N}} \\ &= \frac{(-1)}{N(N-1)}\sum_{r=0}^{N-1} r e^{j\frac{2\pi r(p-q)}{N}} \, , \text{ because } \sum_{r=0}^{N-1} e^{j\frac{2\pi r(p-q)}{N}} = 0 \text{ for } p \neq q \\ &= \frac{(-1)}{N(N-1)}\sum_{r=0}^{N-1} r\omega^r \, , \text{ with } \omega{:=}e^{j\frac{2\pi(p-q)}{N}} \\ &= \frac{(-1)}{N(N-1)}\frac{\omega^N(N(\omega-1)-\omega)+\omega}{(\omega-1)^2} \\ &= \frac{(-1)}{(N-1)(\omega-1)} \qquad \text{since } \omega^N = 1 \, , \end{aligned}$$

which finally yields

$$b_0^{p,q}(0) = \frac{(-1)}{(N-1)(e^{j\frac{2\pi(p-q)}{N}}-1)} \, , \text{ for } p \neq q \, . \tag{48}$$

Similarly, we can verify that $b_{N-1}^{p,q}(0) = -b_0^{p,q}(0)$, for $p \neq q$. Hence we obtain $\beta^{p,q} = 2(1-\gamma)|b_0^{p,q}(0)|^2$ which, with the aid of (48), finally yields

$$\beta^{p,q} = \frac{2(1-\gamma)}{(N-1)^2(1-cos(2\pi(p-q)/N))} \, , \text{ for } p \neq q \, . \tag{49}$$

Equation (49) constitutes an intuitive result: the further a tone q is from p, the less its interference on tone p is.

REFERENCES

[1] J.M.Cioffi, "A Multicarrier Primer," In *ANSI T1E1.4 Committee Contribution, No. 91-157*, Boca Raton, FL. Available online at http://www-isl.stanford.edu/ people/cioffi/papers.html, November 1991.

[2] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Transanctions on Communications*, vol. 46, no. 3, pp. 357–366, March 1998.

[3] L J. Cimini, B. Daneshrad, and N. R. Sollenberger, "Clustered OFDM with transmitter diversity and coding," in *GLOBECOM*, 1996, pp. 703–707.

[4] U. Tureli, D. Kivanc, and H. Liu, "Experimental and analytical studies on a high-resolution OFDM carrier frequency offset estimator," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 2, pp. 629–643, March 2001.

[5] J. Armstrong, "Analysis of new and existing methods of reducing intercarrier interference due to carrier frequency offset in OFDM," *IEEE Transactions on Communications*, pp. 365–369, March 1999.

[6] M. Russell and G. Stuber, "Interchannel interference analysis of OFDM in a mobile environment," in *VTC*, 1995, pp. 820–824.

[7] P. Robertson and S. Kaiser, "Analysis of the loss of orthogonality through doppler spread in OFDM systems," in *Globecom*, 1999, pp. 701–706.

[8] J. Li and M. Kavehrad, "Effects of time selective multipath fading on OFDM systems for broadband mobile applications," *IEEE Communications Letters*, pp. 332–334, December 1999.

[9] William C. Jakes, *Microwave Mobile Communications*, John Wiley and Sons, New York, 1974.

[10] Y. Li and L.J. Cimini, "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Transactions on Communications*, vol. 49, no. 3, pp. 401–404, March 2001.

[11] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *VTC '98. 48th IEEE Vehicular Technology Conference*, 1998, vol. 3, pp. 2232–6.

[12] W. Jeon, K. Chang, and Y.Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Transactions on Communications*, pp. 27–32, January 1999.

[13] M.K. Tsatsanis and Z. Xu, "Pilot symbol assisted modulation in frequency selective fading wireless channels," *IEEE Transactions on Signal Processing*, vol. 48, no. 8, pp. 2353–2365, August 2000.

[14] C. Tepedelenlioglu and G. B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Transactions on Signal Processing*, vol. 48, no. 7, pp. 2029–2043, July 2000.

[15] T. Kailath, "Channel characterization: time-variant dispersive channels," in *Lectures on Communication System Theory*, E. J. Baghdady, Ed., pp. 95–123. McGraw-Hill Book Co., New York, 1961.

[16] M. Medard, *Capacity of time-varying multiple user channels in wireless communications*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 1995.

[17] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, United Kingdom, 1990.

[18] J H. Winters, J. Salz, and R. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Transactions on Communications*, vol. 42, no. 2-4, pp. 1740–1751, Feb/Mar/Apr 1994.

[19] S. Diggavi, "Analysis of multicarrier transmission in time-varying channels," in *ICC*, 1997, pp. 1191–1195.

[20] Y.-S. Choi, P.J. Voltz, and F.A. Cassara, "On channel estimation and detection for multicarrier signals in fast

and selective Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 49, no. 8, pp. 1375–1387, August 2001.

[21] S.M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[22] Z. Liu, G.B. Giannakis, A. Scaglione, and S. Barbarossa, "Decoding and equalization of unknown multipath channels based on block precoding and transmit-antenna diversity," in *Asilomar Conf on Signals, Systems, and Computers*, 1999, pp. 1557–1561.

[23] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, pp. 1969–1986, October 1998.

[24] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 3, pp. 1122–8, August 1998.

[25] S. Ohno and G. B. Giannakis, "Capacity maximizing pilots and precoders for wireless OFDM over rapidly fading channels," *IEEE Transactions on Information Theory*, May 2001 (submitted).

[26] J.K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels (mobile radio)," *IEEE Transactions on Vehicular Technology*, vol. 40, pp. 686–693, November 1991.

[27] W.C. Jakes, *Microwave mobile communications*, John Wiley & Sons, Inc., 1994.