A Low-Complexity Receiver for OFDM in Doubly-Selective Channels

Philip Schniter
Dept. of Electrical Engineering, The Ohio State University
2015 Neil Ave., Columbus OH 43210
e-mail: schniter@ee.eng.ohio-state.edu

***Abstract*— Orthogonal frequency division multiplexing (OFDM) systems may experience significant inter-carrier interference (ICI) when used in time- and frequency-selective, or doubly-selective, channels. In such cases, the classical symbol estimation schemes, e.g., minimum mean-squared error (MMSE) and zero-forcing (ZF) estimation, require matrix inversion that is prohibitively complex for large symbol lengths. An analysis of the ICI generation mechanism leads us to propose a novel two-stage equalizer whose complexity (apart from the FFT) is linear in the OFDM symbol length. The first stage applies optimal linear pre-processing to restrict ICI support and the second stage uses iterative MMSE estimation to estimate finite-alphabet frequency-domain symbols. Simulation results indicate that our equalizer has significant performance and complexity advantages over the classical linear MMSE estimator in doubly-selective channels.**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) [1] has emerged as one of the most practical techniques for data communication over frequency-selective fading channels. In OFDM, the computationally-efficient fast Fourier transform (FFT) is used to transmit data in parallel over a large number of orthogonal subcarriers. When an adequate number of subcarriers are used in conjunction with a cyclic prefix of adequate length, subcarrier orthogonality is maintained even in the presence of frequency-selective fading. Orthogonality implies a lack of subcarrier interference and permits simple, high-performance data detection.

In time- and frequency-selective—or *doubly-selective*—fading, however, the orthogonality of OFDM is lost, leading to subcarrier interference which greatly complicates optimal data detection. Historically, OFDM has been applied to scenarios in which time-selectivity can be effectively ignored. But future wireless applications are expected to operate at high transmit-frequencies, at high levels of mobility, and at high capacities, resulting in fading that is doubly-selective. Thus, the primary advantage of classical OFDM—interference-free operation—will not carry over to important future applications.

While the vast majority of OFDM literature ignores intra-symbol channel variation, a few OFDM modifications have been proposed to cope with the resulting inter-carrier interference (ICI) [2]–[6]. These schemes require $\geq \mathcal{O}(N^2)$ complexity, however, where N denotes the OFDM symbol length, making them impractical for large N. Note that symbol lengths of 4096 and 8192 have been adopted by the 802.16 and DVB-T standards, respectively.

In this paper, we propose $\mathcal{O}(N)$-complexity symbol estimation strategies for OFDM systems in the presence of doubly-selective fading. Rather than simply ignoring small-valued ICI coefficients (as in [2], [5]) we use signal-to-interference-plus-noise ratio (SINR)-optimal low-complexity linear pre-processing to squeeze ICI into a few coefficients. Then we propose low-complexity iterative symbol estimation schemes that leverage the ICI-shortened channel representation. Full channel knowledge is assumed throughout; low-complexity doubly-selective channel estimation is treated elsewhere (see, e.g., [2], [4], [6]–[8]).

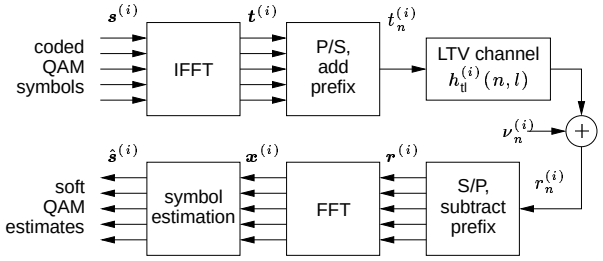

Fig. 1. OFDM system model.

Notation: We use $(\cdot)^t$ to denote transpose, $(\cdot)^*$ conjugate, and $(\cdot)^H$ conjugate transpose. $\mathcal{C}(\boldsymbol{b})$ denotes the circulant matrix with first column $\boldsymbol{b}$, $\mathcal{D}(\boldsymbol{b})$ the diagonal matrix created from the vector $\boldsymbol{b}$, $\mathrm{diag}(\boldsymbol{B})$ the diagonal matrix with the same diagonal terms as matrix $\boldsymbol{B}$, and $\boldsymbol{I}$ the identity matrix. $\|\cdot\|_F$ denotes the Frobenius norm and $\odot$ element-wise multiplication. Expectation is denoted by $\mathrm{E}\{\cdot\}$ and covariance by $\mathrm{Cov}\{\boldsymbol{b},\boldsymbol{c}\} := \mathrm{E}\{\boldsymbol{b}\boldsymbol{c}^H\} - \mathrm{E}\{\boldsymbol{b}\}\,\mathrm{E}\{\boldsymbol{c}^H\}$. Finally, $\delta(\cdot)$ denotes the Kronecker delta, $\langle\cdot\rangle_N$ the modulo-N operation, $*$ convolution, $\mathbb{R}$ the field of reals, and $\mathbb{Z}$ the set of integers.

II. System Model

The OFDM system model is illustrated in Fig. 1. At each index $i \in \mathbb{Z}$, a set of N coded QAM "frequency-domain" symbols $\{s_k^{(i)}\}$ is collected to form an OFDM symbol $\boldsymbol{s}^{(i)} = [s_0^{(i)}, \ldots, s_{N-1}^{(i)}]^t$. The OFDM symbol is converted into the time-domain samples $\{t_n^{(i)}\}$ according the (norm-preserving) N-point inverse DFT operation

$$t_n^{(i)} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k^{(i)} e^{j\frac{2\pi}{N}kn}, \quad -N_p \leq n < N, \qquad (1)$$

which are then serially transmitted over a noisy multipath channel. Note that $\{t_n^{(i)}\}$ incorporates a cyclic prefix of length $N_p \leq N$. The multipath channel is modeled by the time-variant discrete impulse response $h_{\mathrm{tl}}(n,l)$, defined as the time-n response to an impulse applied at time $n-l$. The channel response during the i^{th} OFDM symbol interval is $h_{\mathrm{tl}}^{(i)}(n,l) := h_{\mathrm{tl}}(iN + iN_p + n, l)$ for $-N_p \leq n < N$. Assuming a causal channel with maximum delay spread $N_h \leq N_p$, the received

samples collected during the i^{th} OFDM symbol interval are

$$r_n^{(i)} \;=\; \sum_{l=0}^{N_h-1} h_{\mathsf{tl}}^{(i)}(n,l)\,t_{n-l}^{(i)} + \nu_n^{(i)}, \quad 0 \le n < N, \qquad (2)$$

where $\nu_n^{(i)}$ are samples of white Gaussian noise (AWGN) with variance σ^2. Note that $r_n^{(i)}$ contains contributions from only the i^{th} transmitted symbol; this is a consequence of assuming that the multipath-corrupted cyclic prefix is discarded by the receiver. The receiver then computes an N-point DFT of $\{r_n^{(i)}\}$

$$x_d^{(i)} \;=\; \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} r_n^{(i)} e^{-j\frac{2\pi}{N}dn}. \qquad (3)$$

Using $\boldsymbol{F}$ to denote the N-point unitary DFT matrix, $\mathcal{H}_{\mathsf{tl}}^{(i)}$ to denote a (time-variant, circular) convolution matrix such that $[\mathcal{H}_{\mathsf{tl}}^{(i)}]_{n,l} \;:=\; h_{\mathsf{tl}}^{(i)}\big(n, \langle n-l\rangle_N\big)$, and defining $\boldsymbol{r}^{(i)} := [r_0^{(i)}, \ldots, r_{N-1}^{(i)}]^t$ and $\boldsymbol{\nu}^{(i)} := [\nu_0^{(i)}, \ldots, \nu_{N-1}^{(i)}]^t$, equation (2) can be written in vector form as follows.

$$\boldsymbol{r}^{(i)} \;=\; \mathcal{H}_{\mathsf{tl}}^{(i)}\boldsymbol{t}^{(i)} + \boldsymbol{\nu}^{(i)} \;=\; \mathcal{H}_{\mathsf{tl}}^{(i)}\boldsymbol{F}^H\boldsymbol{s}^{(i)} + \boldsymbol{\nu}^{(i)} \qquad (4)$$

Defining $\boldsymbol{x}^{(i)} := [x_0^{(i)}, \ldots, x_{N-1}^{(i)}]^t$ and the *subcarrier coupling matrix* $\mathcal{H}_{\mathsf{df}}^{(i)} := \boldsymbol{F}\mathcal{H}_{\mathsf{tl}}^{(i)}\boldsymbol{F}^H$, equation (3) can be written

$$\boldsymbol{x}^{(i)} \;=\; \boldsymbol{F}\boldsymbol{r}^{(i)} \;=\; \mathcal{H}_{\mathsf{df}}^{(i)}\boldsymbol{s}^{(i)} + \boldsymbol{w}^{(i)} \qquad (5)$$

where $\boldsymbol{w}^{(i)} \;=\; \boldsymbol{F}\boldsymbol{\nu}^{(i)}$. It is straightforward to show that $[\mathcal{H}_{\mathsf{df}}^{(i)}]_{d,k} = h_{\mathsf{df}}^{(i)}(d-k,k)$, where

$$h_{\mathsf{df}}^{(i)}(d,k) \;:=\; \frac{1}{N}\sum_{n=0}^{N-1}\sum_{l=0}^{N-1} h_{\mathsf{tl}}^{(i)}(n,l)e^{-j\frac{2\pi}{N}(lk+dn)}. \qquad (6)$$

Note that $\{h_{\mathsf{df}}^{(i)}(0,:)\}$ appear on the main diagonal of $\mathcal{H}_{\mathsf{df}}^{(i)}$, $\{h_{\mathsf{df}}^{(i)}(1,:)\}$ on the first sub-diagonal, $\{h_{\mathsf{df}}^{(i)}(-1,:)\}$ on the first super-diagonal, and so on. This implies that $h_{\mathsf{df}}^{(i)}(d,k)$ can be interpreted as the response, at subcarrier $k+d$, to a frequency-domain impulse centered at subcarrier k.

We assume the typical wide-sense stationary uncorrelated scattering (WSSUS) model [9] such that

$$\mathrm{E}\{h_{\mathsf{tl}}(n,l)h_{\mathsf{tl}}^*(n-q,l-m)\} \;=\; r_{\mathsf{t}}(q)\sigma_l^2\delta(m). \qquad (7)$$

In (7), $r_{\mathsf{t}}(q)$ denotes the normalized tap autocorrelation (where $r_{\mathsf{t}}(0)=1$) and σ_l^2 denotes the variance of the l^{th} tap.

III. ICI-Generating Mechanism

A non-diagonal subcarrier coupling matrix introduces inter-carrier interference (ICI), complicating the symbol estimation task. The variance of the subcarrier coupling coefficients $\{h_{\mathsf{df}}^{(i)}(d,k)\}$ can be written [10]

$$\mathrm{E}\{|h_{\mathsf{df}}(d,\cdot)|^2\} \;=\; \big(S(\phi) * V(\phi)\big)\big|_{\phi=\frac{2\pi}{N}d} \cdot \sum_l \sigma_l^2 \qquad (8)$$

where $S(\phi) := \sum_q r_{\mathsf{t}}(q)e^{-j\phi q}$ denotes the Doppler spectrum for $\phi \in \mathbb{R}$ and $V(\phi) := \left(\frac{\sin(\phi N/2)}{N\sin(\phi/2)}\right)^2$ the Dirichlet sinc.

Equation (8) interprets the ICI-generating mechanism: the Doppler spectrum $S(\phi)$ is convolved with the Dirichlet sinc $V(\phi)$ and then sampled on the regular grid $\{\phi : \phi = \frac{2\pi}{N}d, d \in \mathbb{Z}\}$. With a linear time-invariant (LTI) channel, i.e., zero Doppler spread, the nulls of $S(\phi) * V(\phi)$ fall on the grid, implying $\mathrm{E}\{|h_{\mathsf{df}}(d,\cdot)|^2\} = \delta(\langle d\rangle_N)\sum_l \sigma_l^2$. With a linear time-variant (LTV) channel, i.e., non-zero Doppler spread, the nulls of $S(\phi) * V(\phi)$ no longer fall on the grid, implying ICI.

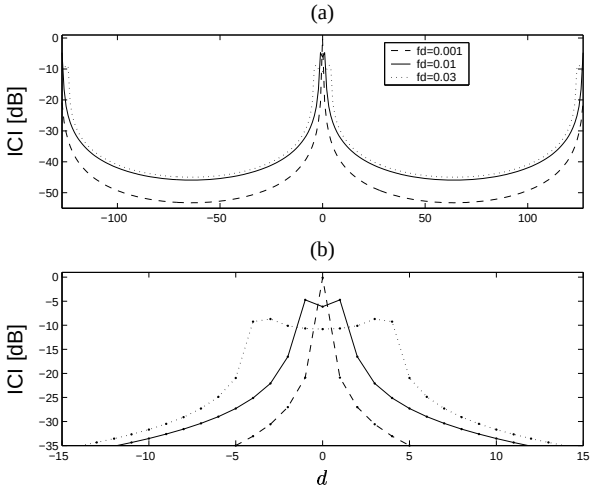

Fig. 2. ICI variance $\mathrm{E}\{|h_{\mathsf{df}}(d,\cdot)|^2\}$ versus d for $N = 128$, WSSUS Rayleigh fading, and various f_{d}. Subplot (b) zooms the view.

In the case of Rayleigh fading [9], $r_{\mathsf{t}}(q) = J_0(2\pi f_{\mathsf{d}} q)$ and

$$S(\phi) \;=\; \begin{cases} \frac{1}{\sqrt{(2\pi f_{\mathsf{d}})^2 - \phi^2}} & |\phi| < 2\pi f_{\mathsf{d}} \\ 0 & |\phi| > 2\pi f_{\mathsf{d}} \end{cases}$$

where $J_0(\cdot)$ denotes the zeroth-order Bessel function of the first kind and f_{d} denotes the maximum Doppler frequency normalized to the *signaling* rate (rather than the OFDM symbol rate). Figure 2 plots $\mathrm{E}\{|h_{\mathsf{df}}(d,\cdot)|^2\}$ as a function of d assuming $N = 128$, Rayleigh fading, and various f_{d}. Here we see that even a Doppler frequency equal to approximately one DFT bin width (i.e., $f_{\mathsf{d}} = 0.01 \approx 1/N$) induces widespread ICI. We will see that ignoring this residual ICI can have serious consequences. Thus the finite-duration nature of the observation plays a critical role; the *time-frequency uncertainty principle* strikes again.

IV. Symbol Estimation

From the observation $\boldsymbol{x}^{(i)}$ in (5), the receiver attempts to detect the true symbol $\boldsymbol{s}^{(i)}$. We assume a detection procedure which consists of an estimation (i.e., equalization) stage following by a decoding stage. Decoding is outside the scope of this paper and so we focus on symbol estimation. Since the decoding performance is expected to be proportional to the *subcarrier-averaged* SINR [11], we employ this criterion in the design of our estimators.

A. Classical Methods

The linear MMSE estimate [4] is given in (9) assuming $\mathrm{E}\{\boldsymbol{s}\} = \boldsymbol{0} = \mathrm{E}\{\boldsymbol{w}\}$, $\mathrm{E}\{\boldsymbol{s}^{(i)}\boldsymbol{s}^{(i)H}\} = \boldsymbol{I}$, $\mathrm{E}\{\boldsymbol{s}^{(i)}\boldsymbol{w}^{(i)H}\} = \boldsymbol{0}$, $\mathrm{E}\{\boldsymbol{w}^{(i)}\boldsymbol{w}^{(i)H}\} = \sigma^2\boldsymbol{I}$, and knowledge of the channel.

$$\hat{\boldsymbol{s}}_{\mathsf{mmse}}^{(i)} \;=\; \big(\mathcal{H}_{\mathsf{df}}^{(i)H}\mathcal{H}_{\mathsf{df}}^{(i)} + \sigma^2\boldsymbol{I}\big)^{-1}\mathcal{H}_{\mathsf{df}}^{(i)H}\boldsymbol{x}^{(i)} \qquad (9)$$

With an LTI channel, $\mathcal{H}_{\mathsf{df}}$ is diagonal and MMSE estimation can be implemented in $\mathcal{O}(N)$ operations; this is the principle motivation for OFDM. With an LTV channel, (9) requires

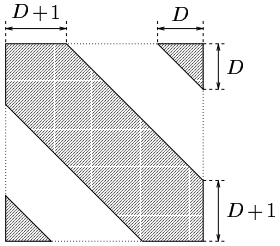

Fig. 3. Desired structure of windowed subcarrier coupling matrix $\mathcal{C}(\boldsymbol{\beta})\mathcal{H}_{\mathsf{df}}^{(i)}$.

non-trivial matrix inversion, making linear MMSE estimation impractical for large N.

B. Linear Pre-processing

In place of $\mathcal{O}(N^2)$ matrix multiplication, we propose low-complexity $\mathcal{O}(N)$ linear pre-processing that renders the ICI response *sparse*, thereby simplifying subsequent symbol estimation. The ICI-generating mechanism described in Sec. III suggests pre-processing that "squeezes" the significant coefficients of $\mathcal{H}_{\mathsf{df}}$ into the shaded regions in Fig. 3. This might be interpreted as the frequency-domain dual of inter-symbol interference (ISI)-response shortening [12]. The parameter $D \in \{0,\ldots,\frac{N}{2}-1\}$ controls the target ICI-response length: larger D corresponds to a longer ICI span and thus increased estimation complexity. In Sec. V we confirm that $D = \lceil f_{\mathsf{d}} N \rceil + 1$ is appropriate for Rayleigh fading.

1) Time-Domain Windowing: While single-carrier systems typically achieve ISI-shortening via convolutive linear filtering, we leverage the receiver's FFT operation to achieve ICI-shortening via fast convolution. Using $\boldsymbol{\beta}$ as the shortening filter's impulse response, the shortened observation becomes

$$\breve{\boldsymbol{x}}^{(i)} \;=\; \mathcal{C}(\boldsymbol{\beta})\boldsymbol{x}^{(i)} \;=\; \mathcal{C}(\boldsymbol{\beta})\mathcal{H}_{\mathsf{df}}^{(i)}\boldsymbol{s}^{(i)} + \mathcal{C}(\boldsymbol{\beta})\boldsymbol{w}^{(i)}, \qquad (10)$$

where we desire that $\mathcal{C}(\boldsymbol{\beta})\mathcal{H}_{\mathsf{df}}^{(i)}$ has the structure illustrated in Fig. 3. Defining $\boldsymbol{b} := \sqrt{N}\boldsymbol{F}^H\boldsymbol{\beta}$, the property $\mathcal{C}(\boldsymbol{g}) = \boldsymbol{F}\,\mathcal{D}\big(\sqrt{N}\boldsymbol{F}^H\boldsymbol{g}\big)\boldsymbol{F}^H$ implies $\mathcal{C}(\boldsymbol{\beta}) = \boldsymbol{F}\,\mathcal{D}(\boldsymbol{b})\boldsymbol{F}^H$ and so

$$\breve{\boldsymbol{x}}^{(i)} \;=\; \boldsymbol{F}\,\mathcal{D}\big(\boldsymbol{b}\big)\boldsymbol{r}^{(i)}. \qquad (11)$$

Thus ICI-shortening can be accomplished through application of the N-point window $\boldsymbol{b}$ to the time-domain observation $\boldsymbol{r}^{(i)}$.

2) Max-SINR Window Design: The coefficients $\boldsymbol{b}$ are designed to maximize the subcarrier-averaged SINR. Using the assumptions on $\boldsymbol{s}^{(i)}$ and $\boldsymbol{w}^{(i)}$ from Sec. IV-A, signal energy $\mathcal{E}_s^{(i)}$ and noise-plus-interference energy $\mathcal{E}_{ni}^{(i)}$ become

$$\mathcal{E}_s^{(i)} \;=\; \big\|\mathcal{M}_{\mathrm{D}}\big(\mathcal{C}(\boldsymbol{\beta})\mathcal{H}_{\mathsf{df}}^{(i)}\big)\big\|_F^2 \qquad (12)$$

$$\mathcal{E}_{ni}^{(i)} \;=\; \big\|\overline{\mathcal{M}}_{\mathrm{D}}\big(\mathcal{C}(\boldsymbol{\beta})\mathcal{H}_{\mathsf{df}}^{(i)}\big)\big\|_F^2 + \sigma^2\,\|\mathcal{C}(\boldsymbol{\beta})\|_F^2 \qquad (13)$$

where $\mathcal{M}_{\mathrm{D}}(\cdot)$ is a mask-operator that preserves the shaded region in Fig. 3 and $\overline{\mathcal{M}}_{\mathrm{D}}(\cdot)$ is its complement. The window coefficients $\boldsymbol{b}_\star^{(i)}$ which maximize $\mathrm{SINR}^{(i)} := \mathcal{E}_s^{(i)}/\mathcal{E}_{ni}^{(i)}$ are derived in [10]. There it is found that

$$\boldsymbol{b}_\star^{(i)} \;=\; \boldsymbol{v}_\star\big(\boldsymbol{A}\odot\hat{\boldsymbol{R}}^{(i)*},\,\mathrm{diag}(\hat{\boldsymbol{R}}^{(i)}+\sigma^2\boldsymbol{I}) - \boldsymbol{A}\odot\hat{\boldsymbol{R}}^{(i)*}\big)$$

where $\boldsymbol{v}_\star(\boldsymbol{B},\boldsymbol{C})$ denotes the principle generalized eigenvalue of the matrix pencil $(\boldsymbol{B},\boldsymbol{C})$, $[\boldsymbol{H}_{\mathsf{tl}}^{(i)}]_{n,l} := h_{\mathsf{tl}}^{(i)}(n,l)$, $\hat{\boldsymbol{R}}^{(i)} := \boldsymbol{H}_{\mathsf{tl}}^{(i)}\boldsymbol{H}_{\mathsf{tl}}^{(i)H}$, and $[\boldsymbol{A}]_{m,n} := \frac{\sin\left(\frac{\pi}{N}(2D+1)(n-m)\right)}{N\sin\left(\frac{\pi}{N}(n-m)\right)}$.

As a function of the channel realization, design of $\boldsymbol{b}_\star^{(i)}$ may be impractical. Maximization of $\overline{\mathrm{SINR}} := \mathrm{E}\{\mathcal{E}_s^{(i)}\}/\mathrm{E}\{\mathcal{E}_{ni}^{(i)}\}$ instead leads to realization-independent coefficients $\bar{\boldsymbol{b}}_\star$ given by (14) with $\hat{\boldsymbol{R}}^{(i)}$ replaced by $\boldsymbol{R} := r_{\mathrm{t}}(m-n)\sum_l \sigma_l^2$. Design of $\bar{\boldsymbol{b}}_\star$ requires knowledge of only $r_{\mathrm{t}}(\cdot)$ and $\sigma^{-2}\sum_l\sigma_l^2$.

C. Iterative MMSE Symbol Estimation

Here we propose two high-performance low-complexity estimators of $\boldsymbol{s}^{(i)}$ from $\breve{\boldsymbol{x}}^{(i)}$ that leverage the ICI-shortened structure of $\mathcal{C}(\boldsymbol{\beta})\mathcal{H}_{\mathsf{df}}^{(i)}$. Both schemes are iterative in that previously-estimated interference (and/or pilots) are used as priors for the current estimate. In the sequel, we use $\breve{\mathcal{H}}^{(i)} := \mathcal{C}(\boldsymbol{\beta})\mathcal{H}_{\mathsf{df}}^{(i)}$ and omit the symbol-index superscript $^{(i)}$, turning (10) into $\breve{\boldsymbol{x}} = \breve{\mathcal{H}}\boldsymbol{s} + \mathcal{C}(\boldsymbol{\beta})\boldsymbol{w}$. All indexing is taken modulo-N.

1) MMSE Estimation Using Priors: The structure of $\breve{\mathcal{H}}$ (recall Fig. 3) implies that s_k contributes primarily to the observation elements $\breve{\boldsymbol{x}}_k := \big[\breve{x}_{k-D}\;\cdots\;\breve{x}_{k+D}\big]^t$. Then defining $\breve{h}_{n,m} := [\breve{\mathcal{H}}]_{n,m}$, $\boldsymbol{s}_k := \big[s_{k-2D}\;\cdots\;s_{k+2D}\big]^t$, and

$$\breve{\mathcal{H}}_k \;:=\; \begin{bmatrix} \breve{h}_{k-D,k-2D} & \cdots & \breve{h}_{k-D,k} & & \\ & \ddots & \vdots & \ddots & \\ & & \breve{h}_{k+D,k} & \cdots & \breve{h}_{k+D,k+2D} \end{bmatrix}$$

$$\boldsymbol{C}_k \;:=\; \begin{bmatrix} \beta_{k-D} & \beta_{k-D-1} & \cdots & \beta_{k-D-N+1} \\ \beta_{k-D+1} & \beta_{k-D} & \cdots & \beta_{k-D-N+2} \\ \ddots & \ddots & \ddots & \ddots \\ \beta_{k+D} & \beta_{k+D-1} & \cdots & \beta_{k+D-N+1} \end{bmatrix},$$

(14) will be exact in the case of perfect-ICI shortening:

$$\breve{\boldsymbol{x}}_k \;\approx\; \breve{\mathcal{H}}_k\boldsymbol{s}_k + \boldsymbol{C}_k\boldsymbol{w} \qquad (14)$$

Note that, as a consequence of modulo-N indexing, the elements of $\breve{\mathcal{H}}$ from the top-right and bottom-left shaded triangles in Fig. 3 are included in $\breve{\mathcal{H}}_k$.

The MMSE linear estimate of s_k given $\breve{\boldsymbol{x}}_k$ is

$$\hat{s}_k \;=\; \mathrm{E}\{s_k\} + \mathrm{Cov}(\boldsymbol{s}_k,\breve{\boldsymbol{x}}_k)\,\mathrm{Cov}(\breve{\boldsymbol{x}}_k,\breve{\boldsymbol{x}}_k)^{-1}\big(\breve{\boldsymbol{x}}_k - \mathrm{E}\{\breve{\boldsymbol{x}}_k\}\big).$$

If we assume $\mathrm{E}\{\boldsymbol{w}\} = \boldsymbol{0}$, $\mathrm{E}\{\boldsymbol{w}\boldsymbol{w}^H\} = \sigma^2\boldsymbol{I}$, $\mathrm{E}\{\boldsymbol{s}\boldsymbol{w}^H\} = \boldsymbol{0}$, and independence among $\{s_k\}$, and if we define $\bar{s}_k := \mathrm{E}\{s_k\}$, $v_k := \mathrm{Cov}(s_k,s_k)$, $\bar{\boldsymbol{s}}_k := [\bar{s}_{k-2D},\ldots,\bar{s}_{k+2D}]^t$, $\boldsymbol{v}_k := [v_{k-2D},\ldots,v_{k+2D}]^t$, and $\breve{\boldsymbol{h}}_k := \big[\breve{h}_{k-D,k}\;\cdots\;\breve{h}_{k+D,k}\big]^t$ then it is straightforward to show that the estimate can be re-written

$$\boldsymbol{f}_k \;=\; \big(\sigma^2\boldsymbol{C}_k\boldsymbol{C}_k^H + \breve{\mathcal{H}}_k\,\mathcal{D}(\boldsymbol{v}_k)\breve{\mathcal{H}}_k^H\big)^{-1}\breve{\boldsymbol{h}}_k v_k \qquad (15)$$

$$\hat{s}_k \;=\; \bar{s}_k + \boldsymbol{f}_k^H(\breve{\boldsymbol{x}}_k - \breve{\mathcal{H}}_k\bar{\boldsymbol{s}}_k). \qquad (16)$$

We choose to use only *extrinsic* information, i.e., only the priors from $\{s_d, d\neq k\}$ when estimating s_k. Henceforth, then, we always set $\bar{s}_k = 0$ and $v_k = 1$ when using (15)-(16).

2) Updating the Priors: The symbol estimate $\hat{s}_k$ can be used to update $\bar{s}_k$ and v_k. For simplicity, we consider only i.i.d. BPSK symbols $s_k \in \mathcal{B} := \{-1,+1\}$; QAM extensions are straightforward. Defining $\mu_k(s) := \mathrm{E}\{\hat{s}_k | s_k = s\}$ and $\sigma_k^2(s) := \mathrm{Cov}(\hat{s}_k,\hat{s}_k|s_k=s)$, and assuming a conditionally Gaussian model for the estimates:

$$p(\hat{s}_k|s_k=s) \;\approx\; \phi\big((\hat{s}_k-\mu_k(s))/\sigma_k(s)\big)/\sigma_k(s),$$

where $\phi(w) := e^{-w^2}/\sqrt{\pi}$ is the proper complex Gaussian density, it can be shown that $\mu_k(s) = \boldsymbol{f}_k^H \breve{\boldsymbol{h}}_k s$ and $\sigma_k^2(s) = \boldsymbol{f}_k^H \breve{\boldsymbol{h}}_k (1 - \breve{\boldsymbol{h}}_k^H \boldsymbol{f}_k)$. If we define the prior and posterior log-likelihood ratios (LLR) as $L(s_k) := \ln \frac{P(s_k=+1)}{P(s_k=-1)}$ and $L(s_k|\hat{s}_k) := \ln \frac{P(s_k=+1|\hat{s}_k)}{P(s_k=-1|\hat{s}_k)}$, respectively, their difference can be expressed as

$$\begin{aligned} \Delta L(\hat{s}_k) &:= L(s_k|\hat{s}_k) - L(s_k) \\ &= \ln \frac{p(\hat{s}_k|s_k=+1)}{p(\hat{s}_k|s_k=-1)} = \frac{4\,\mathrm{Re}(\hat{s}_k)}{1 - \breve{\boldsymbol{h}}_k^H \boldsymbol{f}_k}. \end{aligned} \quad (17)$$

The posterior LLR leads to an update of the priors:

$$\begin{aligned} \bar{s}_{k,\mathsf{new}} &= (+1)\cdot P(s_k=+1|\hat{s}_k) + (-1)\cdot P(s_k=-1|\hat{s}_k) \\ &= \frac{e^{L_k}}{e^{L_k}+1} - \frac{1}{e^{L_k}+1} = \tanh(L_k/2) \quad (18) \\ v_{k,\mathsf{new}} &= \sum_{s\in\mathcal{B}} \big(s - \mathrm{E}\{s_k|\hat{s}_k\}\big)^2 P(s_k=s|\hat{s}_k) = 1 - \bar{s}_{k,\mathsf{new}}^2 \quad (19) \\ L_{\mathsf{new}}(s_k) &= L(s_k) + \Delta L(\hat{s}_k) \quad (20) \end{aligned}$$

which, in turn, can be used to estimate $\{s_{d\neq k}\}$ via (15)-(16).

3) Iterative Joint Estimators: To initialize the first stage of the iteration, we set $\bar{s}_k = 0$ and $v_k = 1$ for indices k that do not correspond to pilots. For k corresponding to pilots, $\bar{s}_k$ are assigned the pilot amplitudes and $v_k = 0$. Two methods of iterative OFDM-symbol estimation are proposed below.

In *block iterative estimation* (BIE), we calculate the entire batch of estimates $\{\hat{s}_k, k = 0,\ldots,N-1\}$ via (15)-(16) before updating the priors via (17)-(19). Using updated priors, a new batch of estimates can be computed via (15)-(16), and so on. The algorithm terminates when the LLRs surpass a threshold or a specified number of iterations elapse.

In *sequential iterative estimation* (SIE), we calculate $\hat{s}_0$ via (15)-(16) and then immediately update the priors $\bar{s}_0$ and v_0 via (17)-(19). Next, we calculate $\hat{s}_1$ and then immediately update $\bar{s}_1$ and v_1. This continues until $\hat{s}_{N-1}$, $\bar{s}_{N-1,\mathsf{new}}$, and $v_{N-1,\mathsf{new}}$ have been computed, then repeats again, starting with $\hat{s}_0$. The algorithm terminates when the LLRs surpass a threshold or a specified number of iterations elapse.

4) Computational Complexity: The implementation complexity of the BIE and SIE algorithms is dominated by the $(2D+1)\times(2D+1)$ Hermitian matrix inversion in (15). As this requires only $\mathcal{O}(D^2)$ operations, a total of $\mathcal{O}(D^2N)$ operations is needed per iteration. It should be noted that $\boldsymbol{C}_k\boldsymbol{C}_k^H$ is fixed for all k and equal to a sub-block of $\mathcal{C}\big(\boldsymbol{F}(\boldsymbol{b}\odot\boldsymbol{b})/\sqrt{N}\big)$, the latter of which can be pre-computed for realization-independent $\boldsymbol{b}$.

5) Relation to Other Known Schemes: The iterative algorithms proposed in Sec. IV-C are related to, yet different from, a number of existing algorithms. SIE is perhaps closest to the estimation stage in the "turbo equalization" scheme of Tüchler et al. [13]. Unlike our approach, however, [13] assumes a LTI channel in white noise and inserts a decoding iteration after each equalization iteration. BIE bears some similarity to the "probabilistic data association" (PDA). method for synchronous CDMA detection proposed by Luo et al. [14]. The latter scheme, however, requires $\mathcal{O}(N^3)$ operations per iteration and a special symbol pre-ordering since their channel and noise are structured differently than ours.

V. Numerical Results and Discussion

Here we discuss the proposed algorithms and study the results of numerical simulations. All experiments employed $N = 128$ i.i.d., unit-variance BPSK symbols per OFDM symbol, SNR^{-1}-variance circular AWGN noise, and an energy-preserving WSSUS Rayleigh-fading channel with $\sigma_l^2 = N_h^{-1}$ (for $0 \le l < N_h$) and $N_h = N/4$. Channel knowledge was assumed and so no pilots were employed.

As a benchmark, consider symbol estimation given perfect knowledge of interfering symbols. This generates the so-called *matched filter bound* (MFB). The MFB does not use assume perfect ICI-shortening; it makes use of all coefficients in $\breve{\mathcal{H}}$. We also consider the *approximate MFB* (AMFB) in which the estimates are calculated assuming perfect knowledge of interference and perfect-ICI-shortening (although the data is generated as usual). The AMFB lower bounds the MSE performance of our algorithms since they were also designed assuming perfectly-shortened ICI [as a result of (14)].

Figures 4 and 5 investigate the subcarrier-averaged SINR performance of various windows averaged over 1000 channel realizations. Figure 4 supports the rule $D \ge \lceil f_{\mathsf{d}} N \rceil + 1$ and verifies that, while complexity increase discourages larger D, performance does not. Figure 5 shows that, with max-SINR windowing and proper selection of D, SINR $\approx$ SNR over the expected operating region. This suggests that the interference is dominated by channel noise and not residual ICI, i.e., that max-SINR windowing does indeed suppress the undesired ICI.

Both Fig. 4 and Fig. 5 show that there is little difference between the performance of the max-SINR window $\boldsymbol{b}_\star^{(i)}$ and its channel-realization-independent approximation $\bar{\boldsymbol{b}}_\star$. In fact, for typical SNRs and a conservative choice of D, there is little difference between $\boldsymbol{b}_\star^{(i)}$ and the Hamming window. For high SNR or $D < \lceil f_{\mathsf{d}} N \rceil + 1$, however, the Hamming window is suboptimal. The rectangular window is clearly suboptimal for all but the lowest SNR; receivers designed around "basis expansion models" that ignore small ICI-coefficients (as in [2], [5]) can suffer serious losses for all but the lowest SNRs.

Figure 6 compares the subcarrier-averaged MSE performance of the SIE and BIE iterative symbol estimation algorithms—after three iterations—to the MFB, the AMFB, and the classical linear MMSE estimation (9). For each trace, the estimations were performed on $\bar{\boldsymbol{b}}_\star$-windowed data and averaged over 5000 channel/data realizations.

We see in Fig. 6 that the SIE and BIE algorithms have, for all practical purposes, reached the AMFB. Recall that the AMFB lower bounds the MSE of our iterative algorithms since they were designed around a sparse approximation (14) of the windowed subcarrier coupling matrix. The difference between the AMFB and the true MFB can be attributed to this low-complexity-enabling sparse approximation. It is interesting to note that, since equalization alone reaches the AMFB, incorporating symbol reliability information from a decoder (i.e., turbo equalization) would not improve our estimates.

Linear MMSE equalization using the full (i.e., non-sparse) subcarrier coupling matrix performs equivalently to our iterative estimation algorithms (and the AMFB) only in the case of very low Doppler and SNR; at higher Dopplers and SNRs,

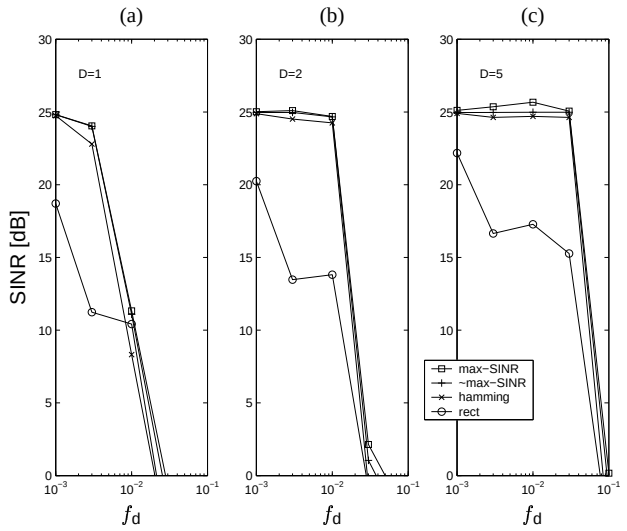

Fig. 4. SINR versus f_d for $N = 128$, SNR=25 dB, WSSUS Rayleigh fading, and various D.

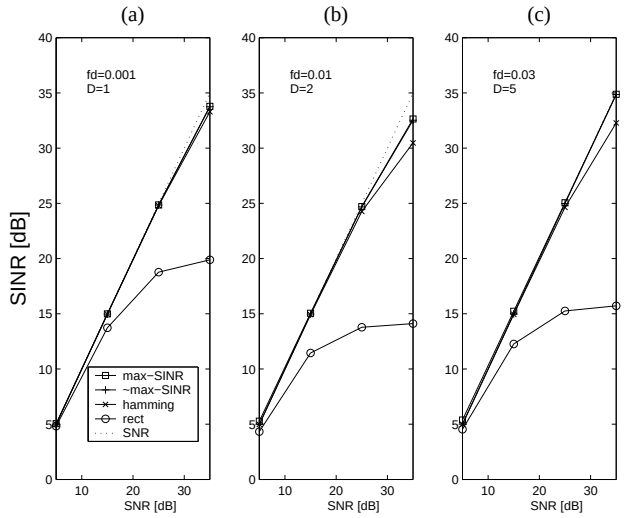

Fig. 5. SINR versus SNR for $N = 128$, WSSUS Rayleigh fading, and various $\{f_\mathrm{d}, D\}$.

iterative estimation significantly outperforms linear MMSE equalization. This is especially meaningful given the fact that the complexity of our iterative estimators is far less than that of the linear MMSE estimator (9).

VI. Conclusions

Equalization of OFDM in doubly-selective channels is complicated by the existence of ICI: the classical frequency-domain equalizer—a simple scaling of each sub-carrier—is no longer sufficient. Previously proposed doubly-selective OFDM equalizers either approximate the linear MMSE estimator with an $\mathcal{O}(N)$ scheme [2], resulting in relatively poor performance, or require at least $\mathcal{O}(N^2)$ operations per OFDM symbol [3]–[6], making them infeasible for large symbol length N. In response, we proposed a low-complexity two-stage equalizer whose performance far surpasses the linear MMSE estimator. The first stage, requiring $\mathcal{O}(N)$ operations, applies SINR-optimal windowing to squeeze ICI into a range of $2D + 1$ subcarrier intervals. The second stage, requiring $\mathcal{O}(D^2 N)$ operations, uses iterative soft ICI-cancellation to estimate the frequency-domain symbols. Simulations indicate that our equalizer performs close to the MFB after only two iterations.

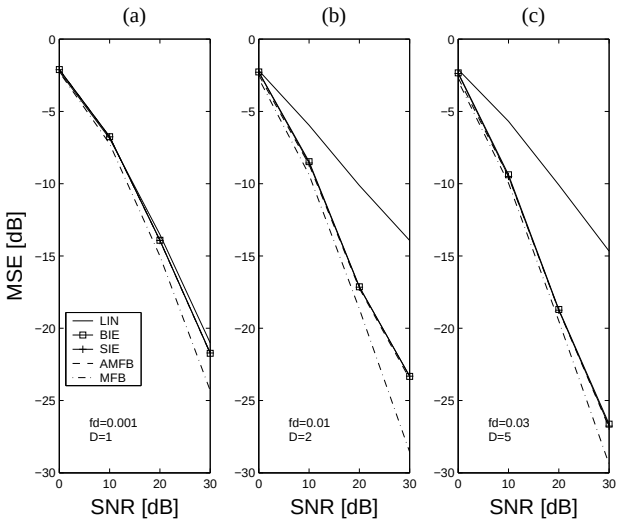

Fig. 6. MSE versus SNR after 3 iterations for $N = 128$, WSSUS Rayleigh fading, and approximately-max-SINR windowing.

References

[1] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, vol. 19, pp. 628–634, Oct. 1971.

[2] 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, Jan. 1999.

[3] J.-P. M. G. Linnartz and A. Gorokhov, "New equalization approach for OFDM over dispersive and rapidly time varying channel," in *Proc. IEEE Int. Symposium Personal Indoor Mobile Radio Commun.*, vol. 2, pp. 1375–1379, 2000.

[4] 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 Trans. Commun.*, vol. 49, pp. 1375–1387, Aug. 2001.

[5] X. Cai and G. B. Giannakis, "Low-complexity ICI suppression for OFDM over time- and frequency-selective Rayleigh fading channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Nov. 2002.

[6] A. Stamoulis, S. N. Diggavi, and N. Al-Dhahir, "Intercarrier interference in MIMO OFDM," *IEEE Trans. Signal Processing*, vol. 50, pp. 2451–2464, Oct. 2002.

[7] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective wireless fading channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1351–1366, May 2003.

[8] P. Schniter, "Low-complexity estimation of doubly-selective channels," in *Proc. IEEE Workshop Signal Processing Advances in Wireless Commun.*, 2003. (to appear).

[9] W. C. Jakes, *Microwave Mobile Communications*. Wiley, 1974.

[10] P. Schniter and S. H. D'Silva, "Low-complexity detection of OFDM in doubly-dispersive channels," in *Proc. Asilomar Conf.*, Nov. 2002.

[11] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, pp. 100–109, Feb. 1995.

[12] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell System Tech. J.*, vol. 52, pp. 1541–1562, Nov. 1973.

[13] M. Tüchler, A. Singer, and R. Koetter, "Minimum mean square error equalization using *a priori* information," *IEEE Trans. Signal Processing*, vol. 50, pp. 673–683, Mar. 2002.

[14] J. Luo, K. R. Pattipati, P. K. Willett, and F. Hasegawa, "Near-optimal multiuser detection in synchronous CDMA using probabilistic data association," *IEEE Commun. Letters*, vol. 5, pp. 361–363, Sep. 2001.