# On Doubly Dispersive Channel Estimation for Pilot-Aided Pulse-Shaped Multi-Carrier Modulation

Philip Schniter
Dept. ECE, The Ohio State University, Columbus OH 43210
schniter@ece.osu.edu

***Abstract*— In this paper, we propose several methods for the pilot-aided estimation of significant ICI coefficients resulting from pulse-shaped multicarrier modulation (PS-MCM) over DD channels. Specifically, we outline Wiener and reduced-rank (RR) Wiener estimation schemes that leverage statistical channel structure, as well as deterministic least-squares (LS) schemes based on basis expansion modeling (BEM). We then report the results of a numerical study which suggests that RR Wiener estimation outperforms LS estimation based on polynomial and oversampled complex exponential BEM, even under significant statistical mismatch. In addition, the RR Wiener estimator is computationally cheaper than the LS-BEM techniques. These findings have implications on the practical design of PS-MCM channel estimation schemes.**[1]

## I. Introduction

One of the most channeling aspects of multicarrier communication (MCM) over doubly dispersive (DD) channels is joint mitigation of inter-symbol interference (ISI) and inter-carrier interference (ICI). The ICI and ISI profiles are a function of the channel's dispersion characteristics as well as the pulse shapes used in modulation and demodulation. For example, cyclic-prefixed orthogonal frequency division multiplexing (CP-OFDM) is known for excellent ISI suppression but poor ICI suppression in DD channels (e.g., [1]). Generalizations of CP-OFDM based on smooth, rather than rectangular, pulses allow better joint suppression of ICI and ISI [2]–[7]. While it is impossible to completely suppress both ICI and ISI in a spectrally efficient multicarrier system, it is possible to design pulses which make the ISI negligible and reduce the ICI span so that each subcarrier sees significant interference from only $\pm D$ adjacent sub-carriers. In an $N$-sub-carrier system, then, equalization would require knowledge of only $(2D+1)N$ *significant* ICI coefficients [8]–[13], where typically $D \ll N$. This reduction in unknown parameters is key to practical implementation.

Still, given only $N$ observations per multicarrier symbol, it is impossible to accurately estimate $(2D+1)N$ ICI coefficients without assuming and exploiting *structure* in the channel response [14]–[19]. This channel structure could be statistical, via an assumed correlation structure, or deterministic, via an assumed basis expansion model (BEM). In either case, however, poor estimation performance might result if the structural assumptions do not match the true channel properties. For this reason, the *robustness* of these channel estimation schemes is of particular importance.

In this paper, we propose Wiener, rank-reduced (RR) Wiener, and least-squares (LS) methods for pilot-aided estimation of significant ICI coefficients arising from general pulse-shaped (PS) MCM over DD channels. We focus on pilot-aided methods, rather than decision-directed methods, for reasons of complexity: the decision-directed methods typically require the inversion of large data-dependent matrices. Within the class of *general PS-MCM* we include both classical schemes like CP-OFDM as well as modern schemes (e.g., [2]–[7]) which use smooth overlapping pulses. While the estimation of time-domain DD channel coefficients (e.g., [9], [15], [20]–[22]) or frequency-domain DD channel coefficients (e.g., [23], [24]) is well studied, we are not aware of much work on the estimation of pulse-shaped ICI coefficients whose structure depends both on channel and pulse properties (e.g., [25] allows smooth demodulation pulses but assumes a CP-OFDM transmitter).

For each proposed estimator, we derive an expression for mean-squared estimator error which is then examined in a detailed numerical study. We pay special attention to the performance of *mismatched* Wiener estimators, i.e., Wiener estimators designed under incorrect statistical assumptions. Our numerical performance study suggests that Wiener estimates compare favorably to the LS-BEM estimates, even under significant statistical mismatch. In addition, our study shows that the rank-reduced Wiener estimator can be implemented with a fraction of the complexity required for LS-BEM. These findings have implications on the practical design of PS-MCM channel estimation schemes.

## II. System Model

At each symbol index $i \in \mathbb{Z}$, $N$ QAM data points $\{s_k(i)\}_{k=0}^{N-1}$ are collected to form a (multicarrier) symbol $\boldsymbol{s}(i) = [s_0(i), \ldots, s_{N-1}(i)]^T$. These symbols are used to modulate pulsed subcarriers as follows:

$$t_n = \sum_{i=-\infty}^{\infty} a_{n-iN_s} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k(i) e^{j\frac{2\pi}{N}(n-iN_s-N_o)k}. \quad (1)$$

In (1), $\{a_n\}$ is the modulation pulse, $N_s$ is the symbol interval, and $N_o \in \{0, \ldots, N-1\}$ delays the subcarrier origin relative to the pulse origin. The multipath channel is described by its time-variant discrete impulse response $h_{n,l}$, defined as the time-$n$ response to an impulse applied at time $n-l$. We assume

[1]This work was supported by the National Science Foundation CAREER grant CCR-0237037.

a causal impulse response of length of $N_h$. The signal observed by the receiver is then

$$r_n \;=\; \nu_n + \sum_{l=0}^{N_h-1} h_{n,l} t_{n-l}, \tag{2}$$

where $\{\nu_n\}$ is a circular white Gaussian noise (CWGN) process with variance $\sigma^2$. Defining $r_n(i) := r_{iN_s+n}$, $\nu_n(i) := \nu_{iN_s+n}$, and $h_{n,l}(i) := h_{iN_s+n,l}$, equations (1) and (2) imply

$$\begin{aligned} r_n(i) \;&=\; \nu_n(i) + \sum_{l=0}^{N_h-1} h_{n,l}(i) \sum_{q=-\infty}^{\infty} a_{qN_s+n-l} \\ &\quad\times \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} s_k(i-q) e^{j\frac{2\pi}{N}(n-l+qN_s-N_o)k}. \end{aligned} \tag{3}$$

The receiver employs the modulation pulse $\{b_n\}$ to calculate $\{x_d(i)\}_{d=0}^{N-1}$, where

$$x_d(i) \;=\; \frac{1}{\sqrt{N}} \sum_{n=-\infty}^{\infty} r_n(i) b_n e^{-j\frac{2\pi}{N}d(n-N_o)}. \tag{4}$$

Plugging (3) into (4), we find

$$x_d(i) \;=\; w_d(i) + \sum_{q=-\infty}^{\infty} \sum_{k=0}^{N-1} H_{d-k,k}(i,q)\, s_k(i-q) \tag{5}$$

where

$$w_d(i) \;=\; \frac{1}{\sqrt{N}} \sum_{n=-\infty}^{\infty} b_n \nu_n(i) e^{-j\frac{2\pi}{N}d(n-N_o)} \tag{6}$$

$$\begin{aligned} H_{d,k}(i,q) \;&=\; \frac{1}{N} \sum_{n=-\infty}^{\infty} \sum_{l=0}^{N_h-1} h_{n,l}(i) b_n a_{qN_s+n-l} \\ &\quad\times e^{-j\frac{2\pi}{N}d(n-N_o)} e^{-j\frac{2\pi}{N}k(l-qN_s)} \end{aligned} \tag{7}$$

Equation (5) indicates that $H_{d,k}(i,q)$ can be interpreted as the response, at time $i$ and subcarrier $k+d$, to a frequency-domain impulse applied at time $i-q$ and subcarrier $k$. Note that $H_{d,k}(i,q)$ depends on the pulses $\{a_n\}$ and $\{b_n\}$.

In the sequel, we assume wide-sense stationary uncorrelated scattering (WSSUS) [26] so that $\mathrm{E}\{h_{n,l} h^*_{n-m,l-\ell}\} = \rho_m \sigma_l^2 \delta_\ell$. Here, $\rho_m$ denotes the normalized autocorrelation at lag $m$ (i.e., $\rho_0 = 1$) and $\sigma_l^2$ denotes the variance of the $l^{th}$ tap. In the case of Rayleigh fading, we have $\rho_m = J_0(2\pi f_{\mathsf{d}} T_c m)$, where $f_{\mathsf{d}} T_c$ denotes the normalized single-sided Doppler spread and $J_0(\cdot)$ denotes a $0^{th}$-order Bessel function of the first kind.

In practice we implement finite-duration causal pulses $\{a_n\}$ and $\{b_n\}$ of length $N_a$ and $N_b$, respectively, implying that only a finite number of terms in the set $\{H_{d,k}(i,q)\}_{q\in\mathbb{Z}}$ will be non-zero. Specifically, (7) implies that non-zero terms result from indices $q$ which satisfy $0 \le qN_s + n - l \le N_a - 1$ for some $n \in \{0, \ldots, N_b - 1\}$ and some $l \in \{0, \ldots, N_h - 1\}$. It is straightforward to show that $H_{d,k}(i,q)$ may be non-zero for $q \in \{-L_{\mathsf{pre}}, \ldots, L_{\mathsf{pst}}\}$, where $L_{\mathsf{pre}} = \lfloor \frac{N_b-1}{N_s} \rfloor$ and $L_{\mathsf{pst}} = \lfloor \frac{N_a+N_h-2}{N_s} \rfloor$.

With the definitions $\boldsymbol{x}(i) := [x_0(i), \ldots, x_{N-1}(i)]^T$, $\boldsymbol{w}(i) := [w_0(i), \ldots, w_{N-1}(i)]^T$, and $[\boldsymbol{H}(i,q)]_{d,k} := H_{d-k,k}(i,q)$, (5) implies the block formulation

$$\boldsymbol{x}(i) \;=\; \boldsymbol{w}(i) + \sum_{q=-L_{\mathsf{pre}}}^{L_{\mathsf{pst}}} \boldsymbol{H}(i,q) \boldsymbol{s}(i-q). \tag{8}$$

It will be convenient to write

$$\boldsymbol{w}(i) \;=\; \boldsymbol{B}\boldsymbol{\nu}(i) \tag{9}$$

$$\boldsymbol{B} \;=\; \boldsymbol{F}\boldsymbol{J}\,\mathcal{D}(\boldsymbol{b}) \tag{10}$$

$$\boldsymbol{J} \;=\; \left[ \begin{array}{c|ccc|c} \boldsymbol{0}_{N-N_o\times N_o} & & & & \boldsymbol{I}_{\bar{N}_o} \\ \boldsymbol{I}_{N_o} & \boldsymbol{I}_N & \cdots & \boldsymbol{I}_N & \boldsymbol{0}_{N-\bar{N}_o\times\bar{N}_o} \end{array} \right], \tag{11}$$

where $\boldsymbol{\nu}(i) := [\nu_0(i), \ldots, \nu_{N_b-1}(i)]^T$, $\boldsymbol{F}$ denotes the unitary $N$-DFT matrix, $\bar{N}_o = \langle N_b - N_o \rangle_N$, and the number of $\boldsymbol{I}_N$ matrices in $\boldsymbol{J}$ is $\lfloor \frac{N_b - N_o}{N} \rfloor$.

In this paper, we will assume that the pulses $\{a_n\}$ and $\{b_n\}$ are designed so that inter-symbol interference (ISI) becomes negligible relative to $\boldsymbol{w}(i)$, in which case (8) reduces to

$$\boldsymbol{x}(i) \;=\; \boldsymbol{w}(i) + \boldsymbol{H}(i,0)\boldsymbol{s}(i). \tag{12}$$

In addition, we will assume that we are interested in estimating only $N(2D+1)$ coefficients within $\boldsymbol{H}(i,0)$, namely, those within the shaded region of Fig. 1. For convenience, we collect them in $\boldsymbol{g}_D(i) \in \mathbb{C}^{(2D+1)N}$:

$$\boldsymbol{g}_D(i) \;:=\; \left[\mathrm{diag}_{-D}(\boldsymbol{H}(i,0))^T, \ldots, \mathrm{diag}_D(\boldsymbol{H}(i,0))^T\right]^T, \tag{13}$$

where $\mathrm{diag}_k(\cdot)$ extracts the $k^{th}$ sub-diagonal of its matrix argument, i.e., $\mathrm{diag}_k(\boldsymbol{H}) := \left[[\boldsymbol{H}]_{k,0}, [\boldsymbol{H}]_{k+1,1}, \ldots, [\boldsymbol{H}]_{k+N-1,N-1}\right]^T$ with modulo-$N$ indexing assumed.

## III. Channel Estimation

Below we propose Wiener, rank-reduced Wiener, and LS-BEM schemes for pilot-aided estimation of $\boldsymbol{g}(i)$. Before discussing the estimation schemes, we describe the pilot pattern.

### A. Choice of Pilot Pattern

We choose a pilot pattern where one out of every $P \ge 2$ multicarrier symbols is used as a pilot. These pilot symbols are then used to estimate the channel coefficients of the $P-1$ multicarrier data symbols in-between. Pilot patterns of this form are relatively common, having been used in several other works (e.g., [9], [18]). Since design of optimal pilot symbols appears to be a challenging problem, we used values obtained from a semi-exhaustive search.

We choose this pattern over one where each multicarrier symbol contains a mixture of pilot and data sub-carriers for the following reason. Assuming a significant ICI radius equal to $D$, the pilot and data sub-carriers would interfere unless a frequency-domain guard with radius $2D$ was placed around each pilot tone.[2] Since Nyquist sampling considerations imply

[2] This pilot strategy corresponds to the MMSE-optimal pilot pattern from [27] for a DD channel satisfying a $(2D+1)$-coefficient CE-BEM.

the need for $N_h$ pilot tones, prevention of pilot/data interference would require that at least $(4D+1)N_h$ sub-carriers are spared from data transmission. For many applications of interest (e.g., the setup in Section IV), however, $(4D+1)N_h > N$, making this scheme impractical.

We now define some quantities that follow from our pilot pattern. Say that, for all indices $i$ corresponding to pilot symbols, we have $\boldsymbol{s}(i) = \boldsymbol{p}$. For these $i$ we can write

$$\boldsymbol{x}(i) = \boldsymbol{P}\boldsymbol{g}(i) + \boldsymbol{w}(i) \quad (14)$$

$$\boldsymbol{g}(i) := \left[\mathrm{diag}_0(\boldsymbol{H}(i,0))^T, \ldots, \mathrm{diag}_{N-1}(\boldsymbol{H}(i,0))^T\right]^T \quad (15)$$

$$\boldsymbol{P} = \left[\boldsymbol{\Theta}^0 \mathcal{D}(\boldsymbol{p}) \ \cdots \ \boldsymbol{\Theta}^{N-1} \mathcal{D}(\boldsymbol{p})\right] \quad (16)$$

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{0}_{N-1}^T & 1 \\ \boldsymbol{I}_{N-1} & \boldsymbol{0}_{N-1} \end{bmatrix}, \quad (17)$$

where $\mathcal{D}(\cdot)$ transforms a vector argument into a diagonal matrix. From (7), we can write

$$\boldsymbol{g}(i) = \boldsymbol{C}\boldsymbol{h}(i), \quad (18)$$

$$\boldsymbol{g}_D(i) = \boldsymbol{C}_D\boldsymbol{h}(i), \quad (19)$$

where $\boldsymbol{C} \in \mathbb{C}^{N^2 \times N_b N_h}$, $\boldsymbol{C}_D \in \mathbb{C}^{(2D+1)N \times N_b N_h}$, and $\boldsymbol{h}(i) \in \mathbb{C}^{N_b N_h}$ are defined element-wise as

$$[\boldsymbol{h}(i)]_m = h_{\langle m \rangle_{N_b}, \left\lfloor \frac{m}{N_b} \right\rfloor}(i) \quad (20)$$

$$[\boldsymbol{C}]_{n,m} = \frac{1}{N} b_{\langle m \rangle_{N_b}} a_{\langle m \rangle_{N_b} - \left\lfloor \frac{m}{N_b} \right\rfloor} e^{-j\frac{2\pi}{N}\left\lfloor \frac{m}{N_b} \right\rfloor n} \times e^{-j\frac{2\pi}{N}\left\lfloor \frac{n}{N} \right\rfloor \left(\langle m \rangle_{N_b} - N_o\right)}. \quad (21)$$

$$[\boldsymbol{C}_D]_{n,m} = \frac{1}{N} b_{\langle m \rangle_{N_b}} a_{\langle m \rangle_{N_b} - \left\lfloor \frac{m}{N_b} \right\rfloor} e^{-j\frac{2\pi}{N}\left\lfloor \frac{m}{N_b} \right\rfloor n} \times e^{-j\frac{2\pi}{N}\left(\left\lfloor \frac{n}{N} \right\rfloor - D\right) \left(\langle m \rangle_{N_b} - N_o\right)}. \quad (22)$$

Note that $\boldsymbol{h}(i)$ contains all time-domain impulse response coefficients affecting $\boldsymbol{H}(i,0)$, and that its statistics are easily written in terms of $\{\sigma_l^2\}_{l=0}^{N_h-1}$ and $\{\rho_m\}_{m=0}^{N_b-1}$.

Our goal is to estimate $\underline{\boldsymbol{g}}_D := [\boldsymbol{g}_D(i+1)^T, \ldots, \boldsymbol{g}_D(i+P-1)^T]^T$, the channel coefficients required for coherent data detection [via (4)], from $\underline{\boldsymbol{x}} := [\boldsymbol{x}(i)^T, \boldsymbol{x}(i+P)^T]^T$, the pilot observations.

### B. Wiener Channel Estimation

We now derive a pilot-aided Wiener channel estimation procedure based on the pilot structure in Section III-A. The linear MMSE estimate of $\underline{\boldsymbol{g}}_D$ from $\underline{\boldsymbol{x}}$ is [28]

$$\hat{\underline{\boldsymbol{g}}}_{D,\mathrm{w}} = \boldsymbol{R}_{gx}\boldsymbol{R}_{xx}^{-1}\underline{\boldsymbol{x}}, \quad (23)$$

where $\boldsymbol{R}_{gx} := \mathrm{E}\{\underline{\boldsymbol{g}}_D \underline{\boldsymbol{x}}^H\}$ and $\boldsymbol{R}_{xx} := \mathrm{E}\left\{\underline{\boldsymbol{x}}\underline{\boldsymbol{x}}^H\right\}$. From (9), (14), and (18),

$$\boldsymbol{R}_{gx} = \begin{bmatrix} \boldsymbol{R}_{gx}^{(1)} & \boldsymbol{R}_{gx}^{(1-P)} \\ \boldsymbol{R}_{gx}^{(2)} & \boldsymbol{R}_{gx}^{(2-P)} \\ \vdots & \vdots \\ \boldsymbol{R}_{gx}^{(P-1)} & \boldsymbol{R}_{gx}^{(-1)} \end{bmatrix} \quad (24)$$

$$\boldsymbol{R}_{xx} = \begin{bmatrix} \boldsymbol{R}_{xx}^{(0)} & \boldsymbol{R}_{xx}^{(-P)} \\ \boldsymbol{R}_{xx}^{(P)} & \boldsymbol{R}_{xx}^{(0)} \end{bmatrix} \quad (25)$$

where

$$\boldsymbol{R}_{gx}^{(q)} := \boldsymbol{C}_D\boldsymbol{R}_{hh}^{(q)}\boldsymbol{C}^H\boldsymbol{P}^H \quad (26)$$

$$\boldsymbol{R}_{xx}^{(q)} := \boldsymbol{P}\boldsymbol{C}\boldsymbol{R}_{hh}^{(q)}\boldsymbol{C}^H\boldsymbol{P}^H + \delta_q\sigma^2\boldsymbol{B}\boldsymbol{B}^H \quad (27)$$

$$\boldsymbol{R}_{hh}^{(q)} := \mathrm{E}\{\boldsymbol{h}(i)\boldsymbol{h}(i-q)^H\}. \quad (28)$$

In (27) we assumed $PN_s > N_b$, so that $\mathrm{E}\{\boldsymbol{w}(i)\boldsymbol{w}(i+P)^H\} = \boldsymbol{0}$. The WSSUS assumption implies that

$$\boldsymbol{R}_{hh}^{(q)} = \mathcal{D}([\sigma_0^2, \ldots, \sigma_{N_h-1}^2]^T) \otimes \boldsymbol{R}_\rho^{(q)} \quad (29)$$

$$[\boldsymbol{R}_\rho^{(q)}]_{m,n} = \rho_{m-n+qN_s}, \quad m,n \in \{0, \ldots, N_b-1\}. \quad (30)$$

It is well known that the Wiener estimation error $\tilde{\underline{\boldsymbol{g}}}_{D,\mathrm{w}} = \hat{\underline{\boldsymbol{g}}}_{D,\mathrm{w}} - \underline{\boldsymbol{g}}_D$ has covariance [28]

$$\mathrm{E}\{\tilde{\underline{\boldsymbol{g}}}_{D,\mathrm{w}}\tilde{\underline{\boldsymbol{g}}}_{D,\mathrm{w}}^H\} = \boldsymbol{R}_{gg} - \boldsymbol{R}_{gx}\boldsymbol{R}_{xx}^{-1}\boldsymbol{R}_{gx}^H, \quad (31)$$

where $\boldsymbol{R}_{gg} := \mathrm{E}\{\underline{\boldsymbol{g}}_D\underline{\boldsymbol{g}}_D^H\}$ is given by

$$\boldsymbol{R}_{gg} = \begin{bmatrix} \boldsymbol{R}_{gg}^{(0)} & \boldsymbol{R}_{gg}^{(-1)} & \cdots & \boldsymbol{R}_{gg}^{(1-P)} \\ \boldsymbol{R}_{gg}^{(1)} & \boldsymbol{R}_{gg}^{(0)} & \cdots & \boldsymbol{R}_{gg}^{(2-P)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{R}_{gg}^{(P-1)} & \boldsymbol{R}_{gg}^{(P-2)} & \cdots & \boldsymbol{R}_{gg}^{(0)} \end{bmatrix} \quad (32)$$

$$\boldsymbol{R}_{gg}^{(q)} := \boldsymbol{C}_D\boldsymbol{R}_{hh}^{(q)}\boldsymbol{C}_D^H. \quad (33)$$

### C. BEM-Constrained Least-Squares Estimation

When it is difficult to obtain accurate knowledge of statistical quantities like $\{\rho_m\}$, $\{\sigma_l^2\}$, and $\sigma^2$, Wiener channel estimation becomes infeasible. As an alternative, one could assume that the channel obeys a basis expansion model (BEM) and estimate the BEM coefficients via least-squares (LS) fit. A generic LS-BEM channel estimation procedure is outlined below for the pilot structure specified in Section III-A.

The BEM models the (estimated) time-domain channel coefficients over the pilot/data/pilot interval $N_f = N_b + PN_s$, using the same basis expansion at each delay:

$$[\hat{\underline{\boldsymbol{h}}}]_m = \hat{h}_{\langle m \rangle_{N_f}, \left\lfloor \frac{m}{N_f} \right\rfloor}(i), \quad \hat{\underline{\boldsymbol{h}}} \in \mathbb{C}^{N_f N_h} \quad (34)$$

$$\hat{\underline{\boldsymbol{h}}} = (\boldsymbol{I}_{N_h} \otimes \boldsymbol{Q})\hat{\underline{\boldsymbol{\eta}}}. \quad (35)$$

In (35), $\boldsymbol{Q} \in \mathbb{C}^{N_f \times K}$ contains the basis vectors and $\hat{\underline{\boldsymbol{\eta}}} \in \mathbb{C}^{N_h K}$ contains the (estimated) BEM coefficients. We can relate $\hat{\underline{\boldsymbol{h}}}$ to $\hat{\boldsymbol{g}}(i+q)$ and $\hat{\boldsymbol{g}}_D(i+q)$ via

$$\hat{\boldsymbol{g}}(i+q) = \boldsymbol{C}^{(q)}\hat{\underline{\boldsymbol{h}}} \quad (36)$$

$$\hat{\boldsymbol{g}}_D(i+q) = \boldsymbol{C}_D^{(q)}\hat{\underline{\boldsymbol{h}}} \quad (37)$$

$$[\boldsymbol{C}^{(q)}]_{n,m} = \frac{1}{N} b_{\langle m \rangle_{N_f} - qN_s} a_{\langle m \rangle_{N_f} - qN_s - \left\lfloor \frac{m}{N_f} \right\rfloor} e^{-j\frac{2\pi}{N}\left\lfloor \frac{m}{N_f} \right\rfloor n} \times e^{-j\frac{2\pi}{N}\left\lfloor \frac{n}{N} \right\rfloor \left(\langle m \rangle_{N_f} - qN_s - N_o\right)} \quad (38)$$

$$[\boldsymbol{C}_D^{(q)}]_{n,m} = \frac{1}{N} b_{\langle m \rangle_{N_f} - qN_s} a_{\langle m \rangle_{N_f} - qN_s - \left\lfloor \frac{m}{N_f} \right\rfloor} e^{-j\frac{2\pi}{N}\left\lfloor \frac{m}{N_f} \right\rfloor n} \times e^{-j\frac{2\pi}{N}\left(\left\lfloor \frac{n}{N} \right\rfloor - D\right) \left(\langle m \rangle_{N_f} - qN_s - N_o\right)} \quad (39)$$

so that we get

$$\underline{\hat{\boldsymbol{g}}}_D \;=\; \underline{\boldsymbol{C}}_D(\boldsymbol{I}_{N_h}\otimes\boldsymbol{Q})\underline{\hat{\boldsymbol{\eta}}} \tag{40}$$

$$\underline{\boldsymbol{C}}_D \;=\; \begin{bmatrix}\boldsymbol{C}_D^{(1)T} & \cdots & \boldsymbol{C}_D^{(P-1)T}\end{bmatrix}^T. \tag{41}$$

For channel estimation, we choose BEM coefficients $\underline{\boldsymbol{\eta}}$ to LS-fit the pilot observations $\underline{\boldsymbol{x}}$. This yields, via (14) and (37),

$$\underline{\hat{\boldsymbol{\eta}}}_{\text{ls}} \;=\; \arg\min_{\underline{\boldsymbol{\eta}}}\left\|\underline{\boldsymbol{x}}-\begin{bmatrix}\boldsymbol{P}\boldsymbol{C}^{(0)}\\ \boldsymbol{P}\boldsymbol{C}^{(P)}\end{bmatrix}(\boldsymbol{I}_{N_h}\otimes\boldsymbol{Q})\underline{\boldsymbol{\eta}}\right\|^2. \tag{42}$$

We then plug $\underline{\hat{\boldsymbol{\eta}}}_{\text{ls}}$ into (40) to obtain the estimated ICI coefficients $\underline{\hat{\boldsymbol{g}}}_{D,\text{ls}}$:

$$\underline{\hat{\boldsymbol{g}}}_{D,\text{ls}} \;=\; \boldsymbol{F}_{\text{ls}}\,\underline{\boldsymbol{x}} \tag{43}$$

$$\boldsymbol{F}_{\text{ls}} \;=\; \underline{\boldsymbol{C}}_D(\boldsymbol{I}_{N_h}\otimes\boldsymbol{Q})\left(\begin{bmatrix}\boldsymbol{P}\boldsymbol{C}^{(0)}\\ \boldsymbol{P}\boldsymbol{C}^{(P)}\end{bmatrix}(\boldsymbol{I}_{N_h}\otimes\boldsymbol{Q})\right)^+, \tag{44}$$

where $(\cdot)^+$ denotes the pseudo-inverse. The covariance of the LS-BEM estimation error $\underline{\tilde{\boldsymbol{g}}}_{D,\text{ls}}=\underline{\hat{\boldsymbol{g}}}_{D,\text{ls}}-\underline{\boldsymbol{g}}_D$ is then

$$\mathrm{E}\{\underline{\tilde{\boldsymbol{g}}}_{D,\text{ls}}\underline{\tilde{\boldsymbol{g}}}_{D,\text{ls}}^H\}=\boldsymbol{F}_{\text{ls}}\boldsymbol{R}_{xx}\boldsymbol{F}_{\text{ls}}^H-\boldsymbol{R}_{gx}\boldsymbol{F}_{\text{ls}}^H-\boldsymbol{F}_{\text{ls}}\boldsymbol{R}_{gx}^H+\boldsymbol{R}_{gg}, \tag{45}$$

for $\boldsymbol{R}_{xx}$, $\boldsymbol{R}_{gx}$, and $\boldsymbol{R}_{gg}$ defined in Section III-B.

Examples of BEMs which do not require statistical channel knowledge include the polynomial BEM [15], [17]:

$$[\boldsymbol{Q}]_{m,k} \;=\; (\sqrt{N_f})^{-1}\left(m-\tfrac{N_f-1}{2}\right)^k, \tag{46}$$

and oversampled complex exponential (OCE) BEM with oversampling factor $M$ [16], [21]:

$$[\boldsymbol{Q}]_{m,k} \;=\; (\sqrt{N_f})^{-1}e^{j\frac{2\pi}{MN_f}(k-\frac{K-1}{2})m}. \tag{47}$$

BEMs which require statistical knowledge include the Slepian BEM [18] and the Karhunen-Loeve BEM [19].

### D. Rank-Reduced Wiener Estimation

We now derive a rank-reduced (RR) version of the Wiener channel estimation procedure outlined in Section III-B and give a BEM interpretation. The intuition is that each of the $N_h$ channel taps changes slowly over the $N_f$-duration pilot/data/pilot interval and thus contributes only about $K=1+\lceil 2f_{\mathsf{d}}T_cN_f\rceil$ non-negligible singular values to $\boldsymbol{R}_{gx}\boldsymbol{R}_{xx}^{-1}$. Thus, optimal rank reduction [28] can be used to significantly reduce the complexity of channel estimation with little performance degradation [9].

The optimal rank-$N_hK$ estimator of $\underline{\boldsymbol{g}}(i)$ is constructed as follows [28]. From the SVD $\boldsymbol{R}_{gx}\boldsymbol{R}_{xx}^{-1}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H$, we build $\boldsymbol{U}_K$ and $\boldsymbol{V}_K$ from the first $N_hK$ columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively, and we build $\boldsymbol{\Sigma}_K$ from the first $N_hK$ rows and columns of $\boldsymbol{\Sigma}$. We find that $\boldsymbol{R}_{gx}\boldsymbol{R}_{xx}^{-1}\approx\boldsymbol{U}_K\boldsymbol{W}_K^H$ for $\boldsymbol{U}_K\in\mathbb{C}^{(P-1)(2D+1)N\times N_hK}$ and $\boldsymbol{W}_K:=\boldsymbol{V}_K\boldsymbol{\Sigma}_K\in\mathbb{C}^{2N\times N_hK}$. Note that $\boldsymbol{U}_K$ can be interpreted as the MMSE-optimal order-$N_hK$ BEM for $\underline{\boldsymbol{g}}_D$ and $\boldsymbol{W}_K$ can be interpreted as the linear MMSE estimator of the corresponding BEM coefficients $\underline{\boldsymbol{\lambda}}$. The resulting rank-reduced estimation procedure

$$\underline{\hat{\boldsymbol{\lambda}}} \;=\; \boldsymbol{W}_K^H\underline{\boldsymbol{x}} \tag{48}$$

$$\underline{\hat{\boldsymbol{g}}}_{D,\text{rr}} \;=\; \boldsymbol{U}_K\underline{\hat{\boldsymbol{\lambda}}} \tag{49}$$

requires only $N_hK[2N+(P-1)(2D+1)N]$ complex MACs per $P-1$ MCM data symbols.

Using $\boldsymbol{F}_{\text{rr}}=\boldsymbol{U}_K\boldsymbol{W}_K^H$, the covariance of the RR Wiener estimation error $\underline{\tilde{\boldsymbol{g}}}_{D,\text{rr}}=\underline{\hat{\boldsymbol{g}}}_{D,\text{rr}}-\underline{\boldsymbol{g}}_D$ can be expressed as

$$\mathrm{E}\{\underline{\tilde{\boldsymbol{g}}}_{D,\text{rr}}\underline{\tilde{\boldsymbol{g}}}_{D,\text{rr}}^H\} \;=\; \boldsymbol{F}_{\text{rr}}\boldsymbol{R}_{xx}\boldsymbol{F}_{\text{rr}}^H-\boldsymbol{R}_{gx}\boldsymbol{F}_{\text{rr}}^H-\boldsymbol{F}_{\text{rr}}\boldsymbol{R}_{gx}^H+\boldsymbol{R}_{gg}. \tag{50}$$

## IV. Numerical Results

The coefficient-averaged MSEs from (31), (45), and (50), i.e., $\mathcal{E}_{\text{w}}=\frac{1}{(2D+1)N_h}\operatorname{tr}(\mathrm{E}\{\underline{\tilde{\boldsymbol{g}}}_{D,\text{w}}\underline{\tilde{\boldsymbol{g}}}_{D,\text{w}}^H\})$, $\mathcal{E}_{\text{ls}}=\frac{1}{(2D+1)N_h}\operatorname{tr}(\mathrm{E}\{\underline{\tilde{\boldsymbol{g}}}_{D,\text{ls}}\underline{\tilde{\boldsymbol{g}}}_{D,\text{ls}}^H\})$, and $\mathcal{E}_{\text{rr}}=\frac{1}{(2D+1)N_h}\operatorname{tr}(\mathrm{E}\{\underline{\tilde{\boldsymbol{g}}}_{D,\text{rr}}\underline{\tilde{\boldsymbol{g}}}_{D,\text{rr}}^H\})$, respectively, are now analyzed under various parameter settings. Both the OCE-BEM and polynomial-BEM were tested. We employed the pulse-shaped MCM system from [7],[3] which chooses the modulator/demodulator pulses to maximize SINR, where "signal energy" is defined as that received through the channel constructed using the diagonal elements of $\boldsymbol{H}(i,0)$, and "interference energy" is defined as that received through ISI as well as the ICI coefficients outside the shaded region in Fig. 1. The system under consideration used $N=16$ sub-carriers, significant ICI radius $D=2$, multicarrier symbol interval $N_s=N$ (i.e., operation at 1 symbol/second/Hz), pulse lengths $N_a=24$ and $N_b=26$, OCE-BEM oversampling factor $M=3$, and (unless otherwise noted) pilot spacing $P=2$. We used Jakes model to generate realizations of a Rayleigh fading WSSUS channel with maximum normalized delay spread $N_h=4$ and exponential power decay $\sigma_l^2=2^{-l/N_3}$ for $l\in\{0,\dots,N_h-1\}$. Unless otherwise noted, the half-power length was $N_3=4$, the normalized Doppler frequency was $f_{\mathsf{d}}T_c=0.01$, and SNR=15dB. Note that, for large $N_3$, the power profile becomes uniform, while, for small $N_3$, the channel becomes frequency-flat.

### A. Effect of Rank $K$

Figure 2 shows the MSE of RR-Wiener and LS-BEM methods versus the rank parameter $K$ under the nominal conditions described earlier. For comparison, Fig. 2 also shows the minimum MSE (i.e., that attained by full-rank Wiener estimation). First, we see that the Wiener estimator is extremely robust to rank reduction. Next, we see that, while Wiener error decreases with rank, LS-BEM error does not. In fact, LS-BEM faces an inherent compromise between imposing too much structure (i.e., $K$ too low) or not enough (i.e., $K$ too high). For the remainder of our experiments, we use $K=3$ in an attempt to get near-optimal performance out of all algorithms, keeping in mind that the Wiener estimator could be operated at rank $K=1$ or $K=2$ without much performance loss.

### B. Effect of $f_{\mathsf{d}}T_c$, $f_{\mathsf{d}}T_c$-Mismatch, and Pilot Spacing $P$

Figures 3-4 show MSE versus $f_{\mathsf{d}}T_c$ for LS-BEM, RR-Wiener under $f_{\mathsf{d}}T_c$ mismatch, and Wiener under both mismatched and perfect knowledge of $f_{\mathsf{d}}T_c$. Figure 3 uses pilot

[3]Similar results were observed for other MCM systems, though the results are not reported here.

spacing $P = 2$ while Fig. 4 uses $P = 5$. There we see that the OCE and polynomial BEMs perform similarly, with a relatively constant MSE at low $f_{\mathsf{d}}T_c$ and increasing MSE at high $f_{\mathsf{d}}T_c$. Wiener estimation performs substantially better than LS-BEM, and *mismatched* RR-Wiener estimation performs about the same as LS-BEM over the entire range of mismatch. Comparing Fig. 3 to Fig. 4, we see that MSE decreases as the pilot density increases, as would be expected.

### C. Effect of SNR and SNR-Mismatch

Figure 5 shows MSE versus SNR for LS-BEM, RR-Wiener under SNR mismatch, and Wiener under both mismatched and perfect knowledge of SNR. Here again, the OCE and polynomial LS-BEMs perform similarly, the Wiener estimator outperforms both LS-BEMs (significantly so at low SNR), and the mismatched reduced-rank Wiener outperforms the LS-BEMs over a wide range of mismatch.

### D. Effect of Decay Parameter $N_3$ and Its Mismatch

Figure 5 shows MSE versus decay parameter $N_3$ for the LS-BEMs, RR-Wiener under $N_3$ mismatch, and Wiener under both mismatched and perfect knowledge of $N_3$. It turns out that estimation performance is almost completely invariant to $N_3$, so that the Wieners scheme significantly outperform the LS-BEMs, regardless of mismatch and rank-reduction.

## V. Conclusion

In this paper, we proposed several methods for the pilot-aided estimation of significant ICI coefficients resulting from pulse-shaped multicarrier transmissions over DD channels. The key to accurate estimation of these coefficients is the exploitation of structure within the channel response. We outlined Wiener and RR-Wiener estimation schemes that leverage statistical structure, as well as deterministic LS schemes that leverage BEM structure. We then reported the results of a numerical study which suggested that RR-Wiener estimation outperforms LS estimation based on polynomial and OCE BEMs, even under significant statistical mismatch. In addition, it suggests that the rank-reduced Wiener estimator can be made computationally cheaper than the LS-BEM techniques without much loss in performance. These findings have implications on the practical design of PS-MCM channel estimation schemes.

## References


[1] M. Russell and G. L. Stüber, "Interchannel interference analysis of OFDM in a mobile environment," in *Proc. IEEE Vehicular Technology Conference*, vol. 2, pp. 820–824, 1995.

[2] B. Le Floch, M. Alard, and C. Berrou, "Coded orthogonal frequency division multiplex," *Proceedings of the IEEE*, vol. 83, pp. 982–996, June 1995.

[3] R. Haas and J.-C. Belfiore, "A time-frequency well-localized pulse for multiple carrier transmission," *Wireless Personal Communications*, vol. 5, pp. 1–18, July 1997.

[4] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," *IEEE Journal on Selected Areas In Communications*, vol. 16, pp. 1579–1589, Oct. 1998.

[5] D. Schafhuber, G. Matz, and F. Hlawatsch, "Pulse-shaping OFDM/BFDM systems for time-varying channels: ISI/ICI analysis, optimal pulse design, and efficient implementation," in *Proc. IEEE Symp. PIMRC*, pp. 1012–1016, Sep. 2002.

[6] T. Strohmer and S. Beaver, "Optimal OFDM design for time-frequency dispersive channels," *IEEE Trans. on Communications*, vol. 51, pp. 1111–1122, July 2003.

[7] P. Schniter, "A new approach to multicarrier pulse design for doubly-dispersive channels," in *Proc. Allerton Conf. on Communication, Control, and Computing*, Oct. 2003.

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

[9] 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. on Communications*, vol. 49, pp. 1375–1387, Aug. 2001.

[10] X. Cai and G. B. Giannakis, "Bounding performance and suppressing inter-carrier interference in wireless mobile OFDM," *IEEE Trans. on Communications*, vol. 51, pp. 2047–2056, Dec. 2003.

[11] I. Barhumi, G. Leus, and M. Moonen, "Time-domain and frequency-domain per-tone equalization for OFDM in doubly-selective channels," *Signal Processing*, vol. 84, pp. 2055–2066, Nov. 2004.

[12] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Trans. on Signal Processing*, vol. 52, pp. 1002–1011, Apr. 2004.

[13] L. Rugini, P. Banelli, and G. Leus, "Simple equalization of time-varying channels for OFDM," *IEEE Communications Letters*, vol. 9, pp. 619–621, July 2005.

[14] M. K. Tsatsanis and G. B. Giannakis, "Modeling and equalization of rapidly fading channels," *Internat. Journal of Adaptive Control & Signal Processing*, vol. 10, pp. 159–176, Mar. 1996.

[15] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. on Communications*, vol. 47, pp. 862–873, June 1999.

[16] T. A. Thomas and F. W. Vook, "Multi-user frequency-domain channel identification, interference suppression, and equalization for time-varying broadband wireless communications," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, (Boston, MA), pp. 444–448, Mar. 2000.

[17] A. Gorokhov and J.-P. Linnartz, "Robust OFDM receivers for dispersive time-varying channels: Equalization and channel acquisition," *IEEE Trans. on Communications*, vol. 52, pp. 572–583, Apr. 2004.

[18] T. Zemen and C. F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheriodal sequences," *IEEE Trans. on Signal Processing*, vol. 53, pp. 3597–3607, Sept. 2005.

[19] K. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *Proc. IEEE Global Telecommunications Conf.*, (St. Louis, MO), Nov. 2005.

[20] G. Leus and M. Moonen, "Deterministic subspace based blind channel estimation for doubly-selective channels," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communication*, pp. 210–214, June 2003.

[21] G. Leus, "On the estimation of rapidly time-varying channels," in *Proc. European Signal Processing Conf.*, Sept. 2004.

[22] I. Barhumi, G. Leus, and M. Moonen, "MMSE estimation of basis expansion models for rapidly time-varying channels," in *Proc. European Signal Processing Conf.*, Sept. 2005.

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

[24] Y. Mostafi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. on Wireless Communications*, vol. 4, pp. 765–774, Mar. 2005.

[25] Z. Tang, G. Leus, R. C. Cannizzaro, and P. Banelli, "Pilot-assisted time-varying OFDM channel estimation," in *Proc. IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing*, 2006. submitted.

[26] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 4th ed., 2001.

[27] A. P. Kannu and P. Schniter, "MSE-optimal training for linear time-varying channels," in *Proc. IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing*, 2005.

[28] L. L. Scharf, *Statistical Signal Processing*. Reading, MA: Addison-Wesley, 1991.

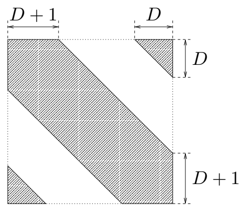


Fig. 1. Quasi-banded channel matrix.

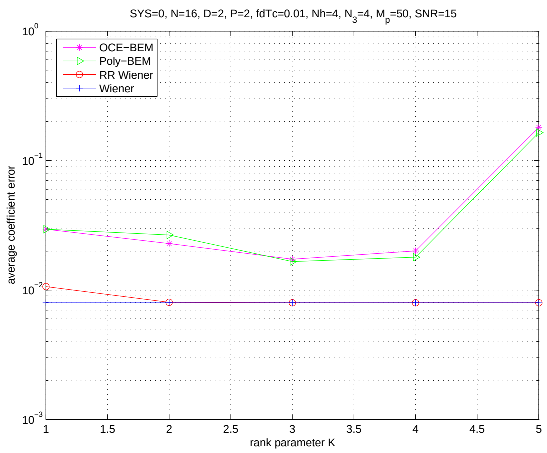


Fig. 2. Channel estimation MSE versus rank parameter $K$ for LS-BEM, Wiener, and rank-reduced Wiener schemes.

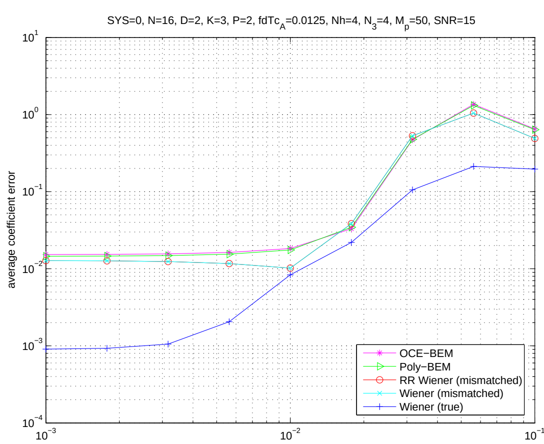


Fig. 3. Channel estimation MSE versus $f_{\mathsf{d}}T_c$ for LS-BEM, Wiener, mismatched Wiener, and mismatched rank-reduced Wiener schemes when $P = 2$. The mismatched schemes assumed $f_{\mathsf{d}}T_c = 0.0125$.

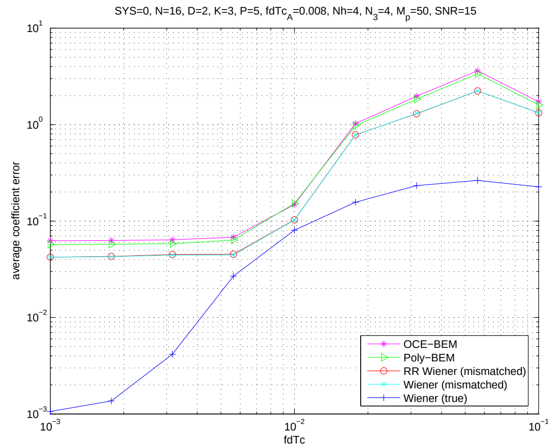


Fig. 4. Channel estimation MSE versus $f_{\mathsf{d}}T_c$ for LS-BEM, Wiener, mismatched Wiener, and mismatched rank-reduced Wiener schemes when $P = 5$. The mismatched schemes assumed $f_{\mathsf{d}}T_c = 0.008$.

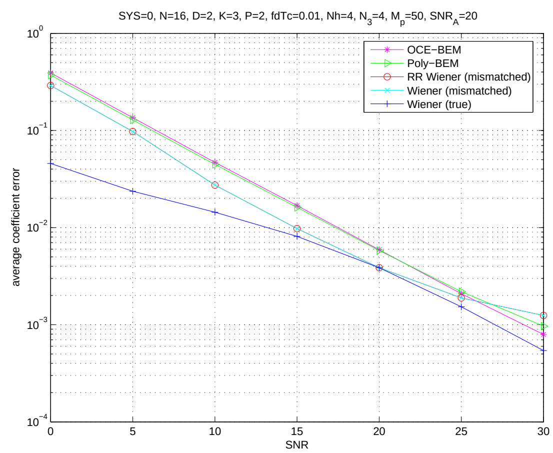


Fig. 5. Channel estimation MSE versus SNR for LS-BEM, Wiener, mismatched Wiener, and mismatched rank-reduced Wiener schemes. The mismatched schemes assumed SNR$=$ 20dB.

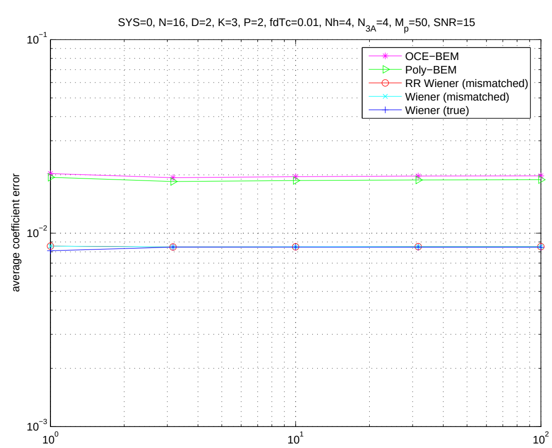


Fig. 6. Channel estimation MSE versus exponential decay parameter $N_3$ for LS-BEM, Wiener, mismatched Wiener, and mismatched rank-reduced Wiener schemes. The mismatched schemes assumed $N_3 = 4$.