# LOW-COMPLEXITY ESTIMATION OF DOUBLY-SELECTIVE CHANNELS

*Philip Schniter*

The Ohio State University, Dept. of Electrical Engineering, Columbus OH 43210
schniter.1@osu.edu

**ABSTRACT**

Doubly-selective wireless communication channels, i.e., those with significant selectivity in both time and frequency domains, arise in applications operating over wide bandwidths, at high carrier frequencies, and under high mobility conditions. Many systems designed to communicate over such channels are OFDM-based and employ large block lengths to combat capacity loss resulting from the use of a redundant prefix. The use of large block length renders channel estimation schemes whose complexities scale in proportion with the square or cube of the block length impractical. In this paper, we present two channel estimation schemes which leverage FFTs and a careful choice of pilots to achieve low estimation error with low complexity.

## 1. INTRODUCTION

The study of doubly-selective wireless communication channels, i.e., channels with significant selectivity in both time and frequency domains, has garnered recent interest. Such channels are expected to arise in applications operating over wide bandwidths, at high carrier frequencies, and under high mobility conditions.

Several multi-carrier schemes have been recently proposed for communication over doubly-selective channels [1–5]. Most are based on orthogonal frequency division multiplexing (OFDM) systems [6] that are modified to handle Doppler- as well as delay-spread. When a cyclic prefix (CP) is used, OFDM systems gain computational advantages at the cost of capacity reduction. Because CP duration is proportional to channel delay spread, capacity can be recovered through the use of a large OFDM block length. Consequently, block lengths of 4096 [7] and 8192 [8] are common.

The use of large block lengths is feasible with OFDM systems that leverage FFTs to achieve $\mathcal{O}(N \log N)$ per-block processing complexity, where $N$ denotes block length. The recently proposed doubly-selective OFDM modifications [1–4] however, exhibit $\mathcal{O}(N^2)$ or $\mathcal{O}(N^3)$ per-block complexities, making them impractical when block length is large.

In [5], the authors proposed an $\mathcal{O}(N \log N)$ *symbol detection* scheme for the doubly-selective case. In this paper, we present some preliminary work on doubly-selective *channel estimation*. As in [5], our goals include *low computational complexity* in addition to high performance. While we acknowledge the existence of more sophisticated methods of time-varying channel estimation (e.g., [9]), we stress the importance of low computational complexity to our target applications. Such low complexity might be achieved through, e.g., the use of FFTs in place of large matrix multiplies or large matrix inverses.

Perhaps the simplest and most well-known means of estimating a doubly-selective impulse response is to embed Kronecker-delta pilot sequences, yielding sampled estimates of the time-varying impulse response, and then to linearly interpolate between these samples (e.g., [1, 10]). One such Kronecker-delta (KD) scheme will be compared to the two estimation schemes proposed in this paper via numerical simulation. The specifics of the KD scheme used for comparison will be given in a later section.

The structure of this paper is as follows. In Sec. 2, we describe a compact channel representation which forms the basis for our estimation strategies and in Sec. 3 we describe the received signal model. In Secs. 4 and 5 we derive two low-complexity channel estimation schemes and in Sec. 6 we present numerical evaluations of these schemes. Section 7 concludes the paper.

## 2. A COMPACT CHANNEL MODEL

Consider first the time/lag representation $h_{\mathsf{tl}}(n,\delta)$, defined as the response at time $n$ to an impulse applied at time $n-\delta$. Assuming wide-sense stationary uncorrelated scattering (WSSUS) [11],

$$\mathrm{E}\big\{h_{\mathsf{tl}}(n,d)h^*_{\mathsf{tl}}(n-p,d-m)\big\} \;=\; r_{\mathsf{tl}}(p,d)\delta(m). \qquad (1)$$

where $\delta(\cdot)$ is the Kronecker delta. With finite delay-spread $N_h$,

$$r_{\mathsf{tl}}(p,d) = 0 \;\text{ for }\; d \notin \{0 \dots N_h-1\}. \qquad (2)$$

Adding the "Rayleigh fading" assumption [11],

$$\begin{aligned} r_{\mathsf{tl}}(p,d) \;&=\; \sigma_d^2 J_0(2\pi f_{\mathsf{d}} p) \\ S_{\mathsf{tl}}(\phi,d) \;&:=\; \sum_{p=-\infty}^{\infty} r_{\mathsf{tl}}(p,d)e^{-j\phi p} \\ &=\; \begin{cases} \frac{\sigma_d^2}{\sqrt{(2\pi f_{\mathsf{d}})^2-\phi^2}} & |\phi| \le 2\pi f_{\mathsf{d}} \\ 0 & \text{else} \end{cases}, \end{aligned} \qquad (3)$$

where $J_0$ denotes the $0^{th}$-order Bessel function of the first kind. Note that the "Doppler spectrum" $S_{\mathsf{tl}}(\phi,\cdot)$ is strictly bandlimited to the normalized Doppler frequency $f_{\mathsf{d}}$.

Next consider the sampled-Doppler/lag representation $h^{(i)}_{\mathsf{dl}}(k,d)$, defined over the block with time indices $i,\dots,i{+}N{-}1$:

$$h^{(i)}_{\mathsf{dl}}(k,d) \;:=\; \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} h_{\mathsf{tl}}(n+i,d)e^{-j\frac{2\pi}{N}k(n+i)}.$$

Note that

$$\begin{aligned} \mathrm{E}\{|h^{(i)}_{\mathsf{dl}}(k,d)|^2\} \;&=\; \frac{1}{N}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1} r_{\mathsf{tl}}(n-m,d)e^{-j\frac{2\pi}{N}k(n-m)} \\ &=\; \sum_{p=-N+1}^{N-1} \tfrac{N-|p|}{N}\, r_{\mathsf{tl}}(p,d)e^{-j\frac{2\pi}{N}kp} \\ &=\; \Big[S_{\mathsf{tl}}(\phi,d) * N\big(\tfrac{\sin(\phi N/2)}{N\sin(\phi/2)}\big)^2\Big]\Big|_{\phi=\frac{2\pi k}{N}}, \end{aligned} \qquad (4)$$

where $*$ denotes convolution. Eqn. (4) suggests that, for large $N$,

$$\mathrm{E}\{|h_{\mathsf{dl}}^{(i)}(k,d)|^2\} \;\approx\; S_{\mathsf{tl}}(\tfrac{2\pi k}{N},d), \tag{5}$$

in which case (2), (3), and (5) imply that the coefficients $h_{\mathsf{dl}}^{(i)}(k,d)$ will be small for either

$$\begin{aligned} &k \;:\; \langle |k| \rangle_N > f_{\mathsf{d}} N \quad \text{or} \\ &d \;:\; d \notin \{0 \ldots N_h - 1\}. \end{aligned}$$

Thus, $h_{\mathsf{dl}}^{(i)}$ gives a compact representation of the channel. Specifically, we expect $h_{\mathsf{dl}}^{(i)}$ to have only $2 f_{\mathsf{d}} N N_h$ "big" coefficients.

## 3. THE RECEIVED SIGNAL MODEL

Using $\{t_n\}$ to denote the transmitted signal and $\{w_n\}$ additive channel noise, the receiver observes the sequence $\{x_n\}$, where

$$x_n \;=\; w_n + \sum_{d=0}^{N_h-1} h_{\mathsf{tl}}(n,d) t_{n-d}.$$

Since $h_{\mathsf{tl}}(n,d) \;=\; \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} h_{\mathsf{dl}}^{(i)}(k,d) e^{j\frac{2\pi}{N}kn}$ for $n \in \mathcal{N}_i := \{i,\ldots,i+N-1\}$, we can write

$$\begin{aligned} x_n \;&=\; w_n + \frac{1}{\sqrt{N}} \sum_{d=0}^{N_h} \sum_{k=0}^{N-1} h_{\mathsf{dl}}^{(i)}(k,d) e^{j\frac{2\pi}{N}kn} t_{n-d} \\ &=\; w_n + \mathbf{f}_n^H \mathbf{H}_{\mathsf{dl}}^{(i)} \mathbf{t}_n, \end{aligned}$$

where

$$\begin{aligned} \mathbf{f}_n^H \;&:=\; \frac{1}{\sqrt{N}} \left[ e^{j\frac{2\pi}{N}n\cdot 0} \;\ldots\; e^{j\frac{2\pi}{N}n(N-1)} \right] \\ \mathbf{H}_{\mathsf{dl}}^{(i)} \;&:=\; \begin{bmatrix} h_{\mathsf{dl}}^{(i)}(0,0) & \ldots & h_{\mathsf{dl}}^{(i)}(0,N_h-1) \\ \vdots & & \vdots \\ h_{\mathsf{dl}}^{(i)}(N-1,0) & \ldots & h_{\mathsf{dl}}^{(i)}(N-1,N_h-1) \end{bmatrix} \\ \mathbf{t}_n \;&:=\; \left[ t_n \;\ldots\; t_{n-N_h+1} \right]^t. \end{aligned}$$

Note $\mathbf{f}_n$ is the $n^{th}$ column of an $N$-point unitary DFT matrix $\mathbf{F}_N$.

We now rearrange the rows of $\mathbf{H}_{\mathsf{dl}}^{(i)}$ to place the $2D$ "big" rows on top. From Sec. 2, we expect $D \approx \lceil f_{\mathsf{d}} N \rceil$ for large $N$. Thus, for $n \in \mathcal{N}_i$,

$$\begin{aligned} x_n \;&=\; w_n + \mathbf{f}_{\mathsf{b},n}^H \mathbf{H}_{\mathsf{dl,b}}^{(i)} \mathbf{t}_n + \mathbf{f}_{\mathsf{s},n}^H \mathbf{H}_{\mathsf{dl,s}}^{(i)} \mathbf{t}_n \\ &=\; w_n + (\mathbf{t}_n^t \otimes \mathbf{f}_{\mathsf{b},n}^H) \mathbf{h}_{\mathsf{dl,b}}^{(i)} + (\mathbf{t}_n^t \otimes \mathbf{f}_{\mathsf{s},n}^H) \mathbf{h}_{\mathsf{dl,s}}^{(i)}, \end{aligned}$$

where $\mathbf{h}_{\mathsf{dl,b}}^{(i)} = \mathrm{vec}(\mathbf{H}_{\mathsf{dl,b}}^{(i)})$, $\mathbf{h}_{\mathsf{dl,s}}^{(i)} = \mathrm{vec}(\mathbf{H}_{\mathsf{dl,s}}^{(i)})$, and

$$\begin{aligned} \mathbf{f}_{\mathsf{b},n}^H \;&=\; \frac{1}{\sqrt{N}} \left[ e^{j\frac{2\pi}{N}n(-D)} \;\ldots\; e^{j\frac{2\pi}{N}n(D-1)} \right] \\ \mathbf{f}_{\mathsf{s},n}^H \;&=\; \frac{1}{\sqrt{N}} \left[ e^{j\frac{2\pi}{N}nD} \;\ldots\; e^{j\frac{2\pi}{N}n(N-D-1)} \right] \\ \mathbf{H}_{\mathsf{dl,b}}^{(i)} \;&:=\; \begin{bmatrix} h_{\mathsf{dl}}^{(i)}(-D,0) & \ldots & h_{\mathsf{dl}}^{(i)}(-D,N_h-1) \\ \vdots & & \vdots \\ h_{\mathsf{dl}}^{(i)}(D-1,0) & \ldots & h_{\mathsf{dl}}^{(i)}(D-1,N_h-1) \end{bmatrix} \\ \mathbf{H}_{\mathsf{dl,s}}^{(i)} \;&:=\; \begin{bmatrix} h_{\mathsf{dl}}^{(i)}(D,0) & \ldots & h_{\mathsf{dl}}^{(i)}(D,N_h-1) \\ \vdots & & \vdots \\ h_{\mathsf{dl}}^{(i)}(N-D-1,0) & \ldots & h_{\mathsf{dl}}^{(i)}(N-D-1,N_h-1) \end{bmatrix}. \end{aligned}$$

## 4. ZERO-FORCING CHANNEL ESTIMATION

Collecting an $L$-element subset of $\{x_n : n \in \mathcal{N}_i\}$ into the vector $\underline{\mathbf{x}}^{(i)}$, we have

$$\underline{\mathbf{x}}^{(i)} \;=\; \mathbf{T}_{\mathsf{b}}^{(i)} \mathbf{h}_{\mathsf{dl,b}}^{(i)} + \mathbf{T}_{\mathsf{s}}^{(i)} \mathbf{h}_{\mathsf{dl,s}}^{(i)} + \underline{\mathbf{w}}^{(i)}, \tag{6}$$

where $\underline{\mathbf{w}}^{(i)}$ is the corresponding vector of noise samples and

$$\mathbf{T}_{\mathsf{b}}^{(i)} := \begin{bmatrix} \vdots \\ \mathbf{t}_n^t \otimes \mathbf{f}_{\mathsf{b},n}^H \\ \vdots \end{bmatrix}, \qquad \mathbf{T}_{\mathsf{s}}^{(i)} := \begin{bmatrix} \vdots \\ \mathbf{t}_n^t \otimes \mathbf{f}_{\mathsf{s},n}^H \\ \vdots \end{bmatrix}.$$

If the measurement $\underline{\mathbf{x}}^{(i)}$ is constructed with samples spaced by at least $N_h$ time indices, different symbols will appear in every row of $\mathbf{T}_{\mathsf{b}}^{(i)}$, giving freedom to pilot design. Say, for example, that the measurement vector is constructed with regular spacing $K \geq N_h$:

$$\underline{\mathbf{x}}^{(iK)} \;=\; \left[ x_{iK} \;\; x_{(i+1)K} \;\; x_{(i+2)K} \;\; \cdots \;\; x_{(i+L-1)K} \right]^t, \tag{7}$$

so that $[\underline{\mathbf{x}}^{(iK)}]_l = x_{(i+l)K}$. If we choose the pilots according to

$$\mathbf{t}_{iK} \;=\; e^{j\frac{2\pi}{N}DKi} \left[ e^{j\frac{2\pi}{N}2DKi\cdot 0} \;\ldots\; e^{j\frac{2\pi}{N}2DKi\cdot(N_h-1)} \right]^t, \tag{8}$$

then

$$\begin{aligned} &\mathbf{t}_{(i+l)K}^t \otimes \mathbf{f}_{\mathsf{b},(i+l)K}^H \\ &\quad=\; \tfrac{1}{\sqrt{N}} \left[ e^{j\frac{2\pi}{N}K(i+l)\cdot 0} \;\ldots\; e^{j\frac{2\pi}{N}K(i+l)(2DN_h-1)} \right]. \end{aligned}$$

For the particular choices $L = 2DN_h$ and $K = \frac{N}{2DN_h}$, we see that $\mathbf{t}_{(i+l)K}^t \otimes \mathbf{f}_{\mathsf{b},(i+l)K}^H$ becomes the $(i+l)^{th}$ row of an $L$-point IDFT matrix, in which case

$$\begin{aligned} \mathbf{T}_{\mathsf{b}}^{(iK)} \;&=\; \begin{bmatrix} \mathbf{t}_{iK}^t \otimes \mathbf{f}_{\mathsf{b},iK}^H \\ \mathbf{t}_{(i+1)K}^t \otimes \mathbf{f}_{\mathsf{b},(i+1)K}^H \\ \vdots \\ \mathbf{t}_{(i+L-1)K}^t \otimes \mathbf{f}_{\mathsf{b},(i+L-1)K}^H \end{bmatrix} \\ &=\; \frac{1}{\sqrt{K}} \mathbf{F}_L^H \,\mathrm{diag}\left( \left[ e^{j\frac{2\pi}{L}i\cdot 0} \;\cdots\; e^{j\frac{2\pi}{L}i(L-1)} \right] \right). \end{aligned}$$

In summary, these choices for $\{t_n\}$, $L$, and $K$ yield a system whose unknown parameters $\mathbf{h}_{\mathsf{dl,b}}^{(iK)}$ are related to the measurement $\underline{\mathbf{x}}^{(iK)}$ via point-wise multiplication and an IFFT. This implies that a "zero-forcing" (ZF) estimation of $\mathbf{h}_{\mathsf{dl,b}}^{(iK)}$ could be efficiently computed using $\mathcal{O}(U \log U)$ samples, where $U = 2DN_h$ denotes the number of unknowns. Specifically, the ZF estimate is

$$\begin{aligned} \widehat{\mathbf{h}_{\mathsf{dl,b}}^{(iK)}} \;&=\; \left( \mathbf{T}_{\mathsf{b}}^{(iK)} \right)^{-1} \underline{\mathbf{x}}^{(iK)} \\ &=\; \sqrt{K}\,\mathrm{diag}\left( [e^{-j\frac{2\pi}{L}i\cdot 0} \;\cdots\; e^{-j\frac{2\pi}{L}i(L-1)}] \right) \mathbf{F}_L \underline{\mathbf{x}}^{(iK)}. \end{aligned}$$

Though low-complexity, the ZF estimator has two potential drawbacks. First, it does not satisfy any well-established criterion of optimality. Second, there are limitations on the allowable Doppler frequency $f_{\mathsf{d}}$ and delay spread $N_h$. Using the large-$N$ approximation $D \approx f_{\mathsf{d}} N$ from Sec. 2, the constraints $K = \frac{N}{2DN_h}$ and $K \geq N_h$ imply a maximum tolerable Doppler frequency of

$$f_{\mathsf{d}}\big|_{\max} \;\approx\; \frac{D}{N} \;=\; \frac{1}{2KN_h} \;\leq\; \frac{1}{2N_h^2}, \tag{9}$$

where equality in (9) corresponds to the case $K = N_h$, i.e., a persistent stream of pilot symbols.

The performance of the ZF estimator will be evaluated numerically in Sec. 6.

## 5. MMSE CHANNEL ESTIMATION

The existence of a low-complexity ZF channel estimation scheme motivates us to investigate the possibility of a low-complexity MMSE approach.

The linear MMSE estimate of $\mathbf{h}_{\mathsf{dl,b}}^{(iK)}$ from $\underline{\mathbf{x}}^{(iK)}$ takes the well known form [12]

$$\begin{aligned}\widehat{\mathbf{h}_{\mathsf{dl,b}}^{(iK)}} &= \mathbf{C}^{(iK)H}\underline{\mathbf{x}}^{(iK)}\\ \mathbf{C}^{(iK)} &= \mathrm{E}\big\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\big\}^{-1}\,\mathrm{E}\big\{\underline{\mathbf{x}}^{(iK)}\mathbf{h}_{\mathsf{dl,b}}^{(iK)H}\big\}\end{aligned} \tag{10}$$

and achieves the performance

$$\begin{aligned}\mathcal{E}_{\mathsf{ms,b}} &:= \mathrm{E}\big\{\|\widehat{\mathbf{h}_{\mathsf{dl,b}}^{(iK)}} - \mathbf{h}_{\mathsf{dl,b}}^{(iK)}\|_2^2\big\} \qquad (11)\\ &= \mathrm{E}\big\{\|\mathbf{h}_{\mathsf{dl,b}}^{(iK)}\|_2^2\big\} - \mathrm{tr}\left(\mathbf{C}^{(iK)H}\,\mathrm{E}\big\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\big\}\mathbf{C}^{(iK)}\right).\end{aligned}$$

Below we derive the MMSE linear estimator $\mathbf{C}^{(iK)}$ assuming the WSSUS Rayleigh fading channel from Sec. 2, the measurement construction (7), the pilots (8), and uncorrelated zero-mean noise samples $\{w_n\}$ with variance $\sigma_w^2$.

First we analyze $\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\}$. The entries of this matrix take the form

$$\begin{aligned}\mathrm{E}\{x_{(i+p)K}x_{iK}^*\} &= \mathrm{E}\Big\{\Big(w_{iK}^* + \sum_{d=0}^{N_h-1} h_{\mathsf{tl}}^*(iK,d)\,t_{iK-d}^*\Big)\\ &\quad\cdot\Big(w_{(i+p)K} + \sum_{c=0}^{N_h-1} h_{\mathsf{tl}}(iK+pK,c)t_{iK+pK-c}\Big)\Big\}\\ &= \sum_{c=0}^{N_h-1}\sum_{d=0}^{N_h-1}\mathrm{E}\big\{h_{\mathsf{tl}}(iK+pK,c)h_{\mathsf{tl}}^*(iK,d)\big\}\\ &\quad\cdot t_{(i+p)K-c}\,t_{iK-d}^* + \sigma_w^2\delta(p)\\ &= \sum_{c=0}^{N_h-1}\sum_{d=0}^{N_h-1} r_{\mathsf{tl}}(pK,d)\delta(c-d)t_{(i+p)K-c}\,t_{iK-d}^* + \sigma_w^2\delta(p)\\ &= \sum_{d=0}^{N_h-1} r_{\mathsf{tl}}(pK,d)t_{(i+p)K-d}\,t_{iK-d}^* + \sigma_w^2\delta(p).\end{aligned}$$

Since $t_{(i+p)K-d}\,t_{iK-d}^* = e^{j\frac{2\pi}{N}KDp(2d+1)}$, we find

$$\begin{aligned}&\mathrm{E}\{x_{(i+p)K}x_{iK}^*\}\\ &\quad= \sum_{d=0}^{N_h-1} r_{\mathsf{tl}}(pK,d)e^{j\frac{2\pi}{N}KDp(2d+1)} + \sigma_w^2\delta(p)\\ &\quad= e^{j\frac{2\pi}{N}KDp}\sum_{d=0}^{N_h-1} r_{\mathsf{tl}}(pK,d)e^{j\frac{4\pi}{N}KDpd} + \sigma_w^2\delta(p).\end{aligned}$$

Assuming a WSSUS Rayleigh-fading channel,

$$\begin{aligned}\mathrm{E}\{x_{(i+p)K}x_{iK}^*\} &= e^{j\frac{2\pi}{N}KDp}J_0(2\pi f_{\mathsf{d}}pK)\\ &\quad\cdot\sum_{d=0}^{N_h-1}\sigma_d^2 e^{j\frac{4\pi}{N}KDpd} + \sigma_w^2\delta(p),\end{aligned}$$

and under the additional assumption of a *uniform power decay profile*, i.e., $\sigma_d^2 = \frac{1}{N_h}\ \forall d \in \{0,\dots,N_h-1\}$,

$$\begin{aligned}&\mathrm{E}\{x_{(i+p)K}x_{iK}^*\}\\ &\quad= e^{j\frac{2\pi}{N}KDp}J_0(2\pi f_{\mathsf{d}}pK)\,\delta\big(\langle p\rangle_{\frac{N}{2KD}}\big) + \sigma_w^2\delta(p).\end{aligned}$$

The previous expression implies that the $L\times L$ matrix $\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\}$ will be Toeplitz with active sub-diagonal indices $\{0,\pm\frac{N}{2KD},\pm2\frac{N}{2KD},\dots\}$, where 0 denotes the index of the main diagonal. In fact, we can write

$$\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\} = \mathbf{R}_x\otimes\mathbf{I}_{\frac{N}{2KD}}, \tag{12}$$

where $\mathbf{R}_x$ is a $\frac{2KDL}{N}\times\frac{2KDL}{N}$ symmetric Toeplitz matrix:

$$[\mathbf{R}_x]_{i+q,i} = \sigma_w^2\delta(q) + (-1)^q J_0\left(\pi f_{\mathsf{d}}ND^{-1}q\right).$$

This is convenient because

$$\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\}^{-1} = \mathbf{R}_x^{-1}\otimes\mathbf{I}_{\frac{N}{2KD}},$$

which permits matrix multiplication using only $\frac{2KDL^2}{N}$ multiplies. More specifically,

$$\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\}^{-1}\underline{\mathbf{x}}^{(iK)} = \mathrm{vec}\big(\mathrm{mat}(\underline{\mathbf{x}}^{(iK)})\mathbf{R}_x^{-1}\big),$$

where $\mathrm{mat}(\underline{\mathbf{x}}^{(iK)})$ is a $\frac{N}{2KD}\times\frac{2KDL}{N}$ matrix composed column-wise from $\underline{\mathbf{x}}^{(iK)}$.

Next we analyze $\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\mathbf{h}_{\mathsf{dl,b}}^{(iK)H}\}$. The entries of this matrix take the form

$$\begin{aligned}&\mathrm{E}\{x_{(i+p)K}h_{\mathsf{dl}}^{(iK)*}(l,d)\}\\ &= \mathrm{E}\Big\{\Big(w_{(i+p)K} + \sum_{c=0}^{N_h-1} h_{\mathsf{tl}}(iK+pK,c)t_{(i+p)K-c}\Big)\\ &\quad\cdot\Big(\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} h_{\mathsf{tl}}^*(n+iK,d)e^{j\frac{2\pi}{N}l(n+iK)}\Big)\Big\}\\ &= \tfrac{1}{\sqrt{N}}\sum_{c=0}^{N_h-1}\sum_{n=0}^{N-1}\mathrm{E}\big\{h_{\mathsf{tl}}(iK+pK,c)h_{\mathsf{tl}}^*(n+iK,d)\big\}\\ &\quad\cdot t_{(i+p)K-c}\,e^{j\frac{2\pi}{N}l(n+iK)}\\ &= \tfrac{1}{\sqrt{N}}\sum_{c=0}^{N_h-1}\sum_{n=0}^{N-1} r_{\mathsf{tl}}(pK-n,d)\delta(c-d)t_{(i+p)K-c}\,e^{j\frac{2\pi}{N}l(n+iK)}\\ &= \tfrac{1}{\sqrt{N}}\sum_{n=0}^{N-1} r_{\mathsf{tl}}(pK-n,d)t_{(i+p)K-d}\,e^{j\frac{2\pi}{N}l(n+iK)}\\ &= \tfrac{1}{\sqrt{N}}e^{j\frac{2\pi}{N}K[D(i+p)(2d+1)+li]}\sum_{n=0}^{N-1} r_{\mathsf{tl}}(pK-n,d)e^{j\frac{2\pi}{N}ln}.\end{aligned}$$

The matrix $\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\mathbf{h}_{\mathsf{dl,b}}^{(iK)H}\}$ is structured so that

$$\begin{aligned}&\big[\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\mathbf{h}_{\mathsf{dl,b}}^{(iK)H}\}\big]_{p,q}\\ &\quad= \mathrm{E}\{x_{(i+p)K}h_{\mathsf{dl}}^{(iK)*}(l,d)\}\quad\text{for}\quad\begin{cases}d = \lfloor\frac{q}{2D}\rfloor\\ l = \langle q\rangle_{2D} - D\end{cases}\\ &\quad= \tfrac{1}{\sqrt{N}}e^{j\frac{2\pi}{N}K\left[D(i+p)(2\lfloor\frac{q}{2D}\rfloor+1)+(\langle q\rangle_{2D}-D)i\right]}\\ &\qquad\cdot\sum_{n=0}^{N-1} r_{\mathsf{tl}}(pK-n,\lfloor\tfrac{q}{2D}\rfloor)e^{j\frac{2\pi}{N}(\langle q\rangle_{2D}-D)n}\\ &\quad= \tfrac{1}{\sqrt{N}}e^{j\frac{2\pi}{N}K\left[i(2D\lfloor\frac{q}{2D}\rfloor+\langle q\rangle_{2D})+p(2D\lfloor\frac{q}{2D}\rfloor+D)\right]}\\ &\qquad\cdot\sum_{n=0}^{N-1} r_{\mathsf{tl}}(pK-n,\lfloor\tfrac{q}{2D}\rfloor)e^{j\frac{2\pi}{N}(\langle q\rangle_{2D}-D)n}.\end{aligned}$$

Leveraging the fact that $q = 2D\lfloor \frac{q}{2D} \rfloor + \langle q \rangle_{2D}$, we find

$$\begin{aligned}
\big[\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\mathbf{h}_{\mathsf{dl,b}}^{(iK)H}\}\big]_{p,q} &= \tfrac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}K\left[iq+p(q-\langle q\rangle_{2D}+D)\right]} \\
&\quad \cdot \sum_{n=0}^{N-1} r_{\mathsf{tl}}(pK-n, \lfloor \tfrac{q}{2D} \rfloor) e^{j\frac{2\pi}{N}(\langle q\rangle_{2D}-D)n} \\
&= \tfrac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}K[iq+pq]} \\
&\quad \cdot \sum_{n=0}^{N-1} r_{\mathsf{tl}}(pK-n, \lfloor \tfrac{q}{2D} \rfloor) e^{-j\frac{2\pi}{N}(\langle q\rangle_{2D}-D)(pK-n)} \\
&= \tfrac{1}{\sqrt{K}} e^{j\frac{2\pi}{N}Kiq} \cdot \sqrt{\tfrac{K}{N}} e^{j\frac{2\pi}{N}Kpq} \\
&\quad \underbrace{\cdot \sum_{m=pK-N+1}^{pK} r_{\mathsf{tl}}(m, \lfloor \tfrac{q}{2D} \rfloor) e^{-j\frac{2\pi}{N}(\langle q\rangle_{2D}-D)m}}_{:= [\mathbf{A}]_{p,q}}.
\end{aligned}$$

$[\mathbf{A}]_{p,q}$ can be recognized as a rectangular-windowed version of $S_{\mathsf{tl}}\big(\frac{2\pi}{N}(\langle q\rangle_{2D} - D), \lfloor \frac{q}{2D} \rfloor\big)$. A uniform power decay profile implies

$$[\mathbf{A}]_{p,q} = \frac{1}{N_h} \sum_{m=pK-N+1}^{pK} J_0(2\pi f_{\mathsf{d}} m) e^{-j\frac{2\pi}{N}(\langle q\rangle_{2D}-D)m},$$

where we note that the first $2D$ columns of $\mathbf{A}$ repeat $N_h$ times. If we choose $K = \frac{N}{2DN_h}$ and $L = P \cdot 2DN_h$ for some $P \in \mathbb{Z}^+$ (where typically $P = 1$),

$$\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\mathbf{h}_{\mathsf{dl,b}}^{(iK)H}\} = \frac{1}{\sqrt{K}} \Big( (\mathbf{1}_P \otimes \mathbf{F}_{2DN_h}^H) \odot \mathbf{A} \Big) \mathbf{D}^{(iK)}, \quad (13)$$

where $\odot$ denotes the Hadamard (i.e., element-wise) product, $\mathbf{1}_P$ denotes a $P \times 1$ vector of ones, and

$$\mathbf{D}^{(iK)} = \mathrm{diag}\left( \left[ e^{j\frac{2\pi}{N}iK\cdot 0} \ \cdots \ e^{j\frac{2\pi}{N}iK(2DN_h-1)} \right] \right).$$

Multiplication by the matrix $\mathrm{E}\{\underline{\mathbf{x}}^{(iK)}\mathbf{h}_{\mathsf{dl,b}}^{(iK)H}\}$ can be reduced to the computation of $2DP$ FFTs of size $2DN_h$. The factor $2D$ pertains to the case where we window the FFT input vector with each of the $2D$ unique columns of $\mathbf{A}$, calculate an FFT for each window, and keep only one out of every $2D$ FFT outputs. More efficient implementations which do not "throw away" DFT outputs are, of course, possible.

It follows from (10), (12), and (13) that we can write the MMSE linear estimator $\mathbf{C}^{(iK)H}$ that assumes $L$–dimensional measurements ($L = 2PDN_h$) constructed according to (7), $K$–spaced pilot sequences ($K = \frac{N}{2DN_h}$) chosen as in (8), and a Rayleigh-fading WSSUS channel with uniform power decay profile as

$$\begin{aligned}
\mathbf{C}^{(iK)H} &= \tfrac{1}{\sqrt{K}} \mathbf{D}^{(iK)H} \big( \mathbf{A}^H \odot (\mathbf{1}_P^t \otimes \mathbf{F}_{2DN_h}) \big) \\
&\quad \cdot \big( \mathbf{R}_x^{-1} \otimes \mathbf{I}_{N_h} \big).
\end{aligned} \quad (14)$$

Equation (14) shows a cascade of three matrix operations: the leftmost (diagonal) matrix multiply requires $2DN_h$ multiplies, the middle (FFT-based) one requires no more than $\mathcal{O}(2PD \cdot 2DN_h \log 2DN_h)$ multiplies, and the rightmost one requires $\frac{2KDL^2}{N} = 4P^2D^2N_h$ multiplies. Note that multiplication by an unstructured $\mathbf{C}^{(iK)H}$ would require $4PD^2N_h^2$ multiplies.

As with the ZF technique, the MMSE technique returns reliable estimates over a limited range of $f_d$ and $N_h$. Our choices for $K$ and $L$ imply that

$$f_{\mathsf{d}}\big|_{\max} \approx \frac{D}{N} = \frac{1}{2KN_h} \le \frac{1}{2N_h^2}, \quad (15)$$

where equality in (15) corresponds to the case $K = N_h$, i.e., a persistent stream of pilot symbols.

## 6. NUMERICAL RESULTS

We now compare the Kronecker-delta (KD) scheme described in Sec. 1 with the ZF and MMSE schemes derived in Secs. 4 and 5 using block-normalized mean-squared estimation error:

$$\mathcal{E} := \frac{1}{N}\mathrm{E}\{\|\mathbf{H}_{\mathsf{tl}}^{(iK)} - \widehat{\mathbf{H}_{\mathsf{tl}}^{(iK)}}\|_F^2\} = \frac{1}{N}\mathrm{E}\{\|\mathbf{H}_{\mathsf{dl}}^{(iK)} - \widehat{\mathbf{H}_{\mathsf{dl}}^{(iK)}}\|_F^2\}.$$

With the ZF and MMSE schemes we estimate only the "big" $\mathbf{H}_{\mathsf{dl,b}}^{(iK)}$ coefficients, and so in these cases $\mathcal{E}$ includes errors caused by the "compact channel representation" of Sec. 2:

$$\mathcal{E} = \frac{1}{N}\mathrm{E}\{\|\mathbf{H}_{\mathsf{dl,b}}^{(iK)} - \widehat{\mathbf{H}_{\mathsf{dl,b}}^{(iK)}}\|_F^2\} + \frac{1}{N}\mathrm{E}\{\|\mathbf{H}_{\mathsf{dl,s}}^{(iK)}\|_F^2\}.$$

For the MMSE scheme, $\mathrm{E}\{\|\mathbf{H}_{\mathsf{dl}}^{(iK)}\|_F^2\} = N$ and (11) imply

$$\begin{aligned}
\mathcal{E}_{\mathsf{ms}} &= \frac{1}{N}\mathcal{E}_{\mathsf{ms,b}} + \frac{1}{N}\mathrm{E}\{\|\mathbf{H}_{\mathsf{dl,s}}^{(iK)}\|_F^2\} \\
&= 1 - \frac{1}{N}\mathrm{tr}\left( \mathbf{C}^{(iK)H}\mathrm{E}\big\{\underline{\mathbf{x}}^{(iK)}\underline{\mathbf{x}}^{(iK)H}\big\}\mathbf{C}^{(iK)} \right).
\end{aligned}$$

The KD scheme we adopted for the simulations was constructed to give a fair comparison with the ZF and MMSE schemes. In our KD, length-$2N_h$ pilot sequences of the form $t_{iK-N_h-d} = \delta(d - N_h)$, for $0 \le d < 2N_h$ and $i \in \mathbb{Z}$, were spaced $K \ge 2N_h$ samples apart. These pilots gave estimates of the time-varying impulse response at $K$-spaced intervals:

$$\begin{aligned}
\begin{bmatrix} \hat{h}_{\mathsf{tl}}(iK,0) \\ \hat{h}_{\mathsf{tl}}(iK,1) \\ \vdots \\ \hat{h}_{\mathsf{tl}}(iK,N_h-1) \end{bmatrix} &:= \begin{bmatrix} x_{iK} \\ x_{iK+1} \\ \vdots \\ x_{iK+N_h-1} \end{bmatrix} \\
&= \begin{bmatrix} h_{\mathsf{tl}}(iK,0) \\ h_{\mathsf{tl}}(iK+1,1) \\ \vdots \\ h_{\mathsf{tl}}(iK+N_h-1,N_h-1) \end{bmatrix} + \begin{bmatrix} w_{iK} \\ w_{iK+1} \\ \vdots \\ w_{iK+N_h-1} \end{bmatrix}.
\end{aligned}$$

Linear interpolation was then used to construct estimates of the time-varying impulse response at points in-between. For $N = K$, this corresponds to a linear fit, and for $N > K$ where $\frac{N}{K} \in \mathbb{Z}$, a piecewise linear fit. In all schemes (i.e., KD, ZF, and MMSE), we chose a unit-modulus pilot sequence in accordance with a peak-power constraint that is standard in practical implementations.

Figures 1 and 2 show the results of numerical simulations with channels generated using Jakes' method [11]. In both figures we assume $N = 256$, $K = 16$, $N_h = 4$, $D = 2$, and $P = 1$. Figure 1 shows the estimation error $\mathcal{E}$ versus normalized Doppler frequency $f_{\mathsf{d}}$ for two different SNRs, while Figure 2 shows $\mathcal{E}$ versus SNR for two different Dopplers. Here we observe that the KD scheme is predominantly *noise-limited*; it is relatively insensitive to $f_{\mathsf{d}}$ (for

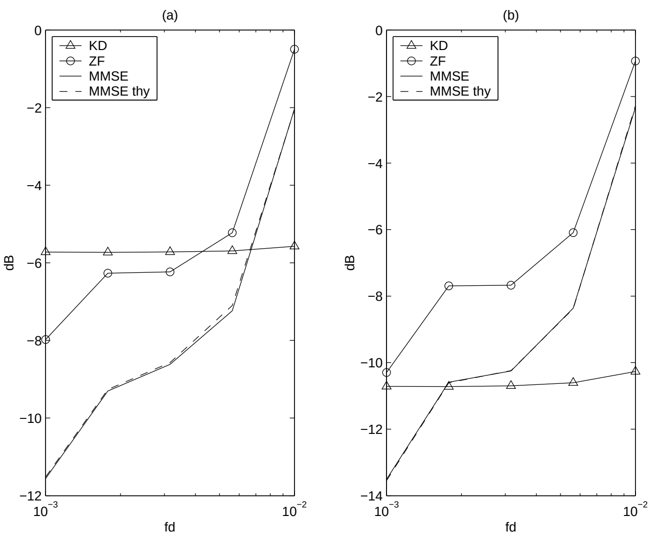


**Fig. 1**. MSE vs. $f_d$ for (a) SNR = 10dB, (b) SNR = 15dB.

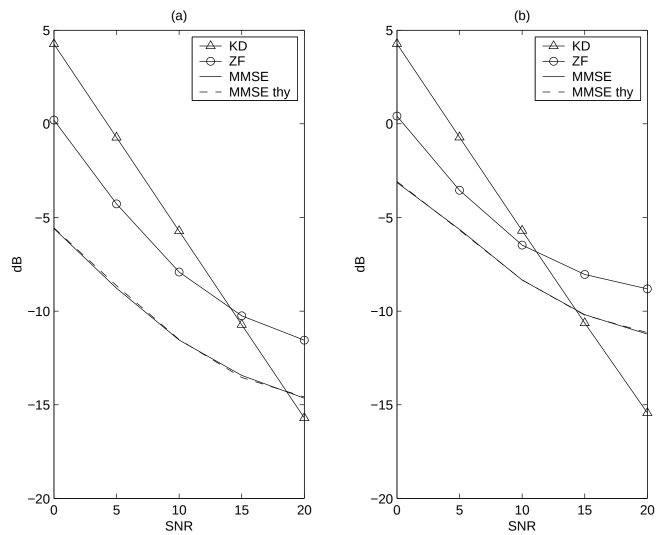


**Fig. 2**. MSE vs. SNR for (a) $f_{\mathrm{d}}$ =0.001, (b) $f_{\mathrm{d}}$ =0.005.

the parameters chosen in our simulations). The ZF and MMSE schemes, while more sensitive to increases in $f_{\mathrm{d}}$, are more tolerant of noise. Hence, *the ZF and MMSE approaches outperform the KD approach in the low-SNR and low-$f_{\mathrm{d}}$ operating region*. Finally, we observe that the MMSE approach outperforms the ZF approach at the cost of a greater implementation complexity.

## 7. CONCLUSIONS

We have derived two low-complexity schemes for the estimation of doubly-selective channel responses, one based on a ZF criterion and one based on a MMSE criterion. Both estimate a reduced set of channel parameters in the Doppler/lag domain by leveraging the usual assumptions of finite delay-spread and finite Doppler-spread. Pilot sequences were chosen so that the implementation of the ZF estimation scheme reduces to a $U$-point FFT at the receiver, where $U = 2DN_h$ represents the number of unknown parameters. With the same pilots, the MMSE estimation scheme can be implemented using $2D$ of these $U$-point FFTs. In fact, it was recognized that more efficient MMSE implementations are possible, though the details are beyond the scope of this paper.

The zero-forcing and MMSE schemes were compared to a well-known technique whereby Kronecker-delta pilot sequences are embedded to provide samples of the time-varying impulse response and piecewise linear interpolation is used to construct the response between these samples. Numerical simulations suggest that the ZF and MMSE schemes proposed herein outperform the Kronecker-delta scheme in the low-SNR and low-Doppler region.

## 8. REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] 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.

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

[6] 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.

[7] "Air interface for fixed broadband wireless access systems. Part A: Systems between 2-11 GHz." IEEE Standard 802.16, 01/01r1, July 2001.

[8] "Digital video broadcasting (DVB): Framing structure, channel coding and modulation for digital terrestrial television." ETSI Standard, ETS 300 744, July 1999.

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

[10] J. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Tech.*, vol. 40, pp. 686–693, Nov. 1991.

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

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