

# Low-Complexity Pilot-Aided Channel Estimation for OFDM Systems Over Doubly-Selective Channels

Tao Cui, Chintha Tellambura and Yue Wu  
Department of Electrical and Computer Engineering  
University of Alberta  
Edmonton, AB, Canada T6G 2V4  
Email: {taocui, chintha,yuewu}@ece.ualberta.ca

**Abstract**—In this paper, we investigate channel estimation (CE) and data detection for OFDM systems over doubly-selective channels. We derive an oversampling basis expansion model (BEM) for doubly-selective channels and its statistical properties. The time diversity in the Doppler-induced inter-carrier-interference (ICI) and its relationship to the carrier frequency offset (CFO) induced ICI are illustrated using the BEM. We derive two low complexity linear minimum mean-square-error (LMMSE) channel estimators using the BEM. The sphere decoder (SD) is modified to equalize the ICI channel. A low-complexity iterative equalizer without matrix inversion is also proposed. Our proposed channel estimators have low complexity and achieve good performance. Furthermore, the low-complexity iterative equalizer performs close to SD.

## I. INTRODUCTION

Broadband wireless communications often require complicated channel estimation (CE) and equalization to compensate for severe frequency-selective fading. Orthogonal frequency division multiplexing (OFDM) partitions the entire bandwidth into parallel subchannels so that each narrow subchannel only suffers from frequency-flat fading and only needs a one-tap equalizer to compensate for the corresponding multiplicative channel distortion. However, CE and equalization for OFDM in doubly-selective (time and frequency) channels is challenging. The Doppler-induced channel variation introduces ICI which destroys the subchannel orthogonality and hence degrades the system performance.

When the OFDM block duration is much smaller than the channel coherence time (relatively mild Doppler), the channel remains approximately constant over an OFDM symbol, and the CE for this case has been thoroughly studied. When the OFDM symbol duration is less than 10% of the channel coherence time, the channel varies linearly [1] and it can be estimated by linear interpolation between two channel estimates acquired by training symbols. In [2], a linear minimum mean-square error (LMMSE) and a successive interference cancellation (SIC) scheme with the optimal ordering are proposed, which can exploit the time diversity provided by the time-selective channel. This estimator has high complexity since it needs to estimate the  $N$  different channel impulse responses (CIR) in each OFDM block.

In this paper, we use the basis expansion model (BEM) [3] for estimating doubly-selective channels. Using the BEM, we analyze the ICI in OFDM and show the diversity introduced by time-selectivity. The time-varying channel within each OFDM block can be represented by a few BEM coefficients. Therefore, two LMMSE estimators are proposed. The first one assumes that the BEM coefficients remain constant for several OFDM symbols, while the second one assumes that they vary from symbol to symbol and Wiener filtering is then used to estimate the coefficients between pilot symbols. For data detection, in addition to using the well-known vertical Bell Labs layered space-time (V-BLAST) algorithm [4], we develop a fast sphere decoder (SD) [5] for OFDM systems. Observing that most of the ICI on each subcarrier comes from several neighboring subcarriers [6], we develop a low complexity iterative detector which avoids matrix inversion.

*Notation:* Bold symbols denote matrices or vectors.  $(\cdot)^T$ ,  $(\cdot)^H$  and  $(\cdot)^\dagger$  denote transpose, conjugate transpose and Moore-Penrose pseudo-inverse respectively. The set of all complex  $K \times 1$  vectors is denoted by  $\mathcal{C}^K$ . The discrete Fourier transform (DFT) matrix of size  $N \times N$  is given by  $\mathbf{F} = 1/\sqrt{N}[e^{j\frac{2\pi}{N}kl}]$ ,  $k, l \in 0, 1, \dots, N-1$ .  $\mathbf{I}_N$  denotes the  $N \times N$  identity matrix. The Circulant Column-Sylvester Matrix of a vector  $\mathbf{a} = [a_1, a_2, \dots, a_n]^T$  ( $a_k$  can be scalar, vector or matrix) with  $\xi$  block columns is defined as

$$\mathbf{S}^{(\xi)}[\mathbf{a}] = \begin{bmatrix} a_1 & a_2 & \cdots & a_n & & & \\ & & \ddots & & & \ddots & \\ & & & a_1 & \cdots & a_{n-1} & a_n \\ a_n & & & a_1 & \cdots & a_{n-1} & \\ & & & & \ddots & & \vdots \\ a_2 & \cdots & a_n & & & & a_1 \end{bmatrix} \quad (1)$$

## II. SYSTEM AND CHANNEL MODEL

### A. System model

In OFDM systems, source data are grouped and mapped into  $X_k$ , which is selected from a complex signal constellation  $\mathcal{Q}$  with unitary energy. Complex data are modulated by inverse discrete Fourier transform (IDFT) on  $N$  parallel subcarriers.

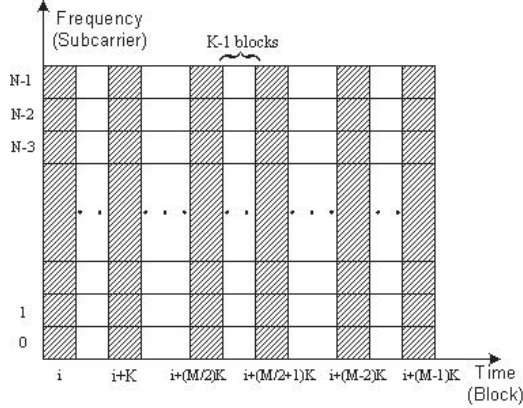


Fig. 1. The pilot pattern used for channel estimation.

The symbol interval and block interval are denoted by  $T_s$  and  $T_f = NT_s$ . The resulting OFDM symbol during the  $m$ th block interval that comprises  $N$  samples is given by

$$x_n^{(m)} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k^{(m)} e^{j(2\pi kn/N)}, \quad -N_g \leq n \leq N-1 \quad (2)$$

where  $N$  is the IFFT size,  $N_g$  is the length of the guard interval, which is larger than the maximum channel delay.

For channel estimation, pilot symbols are necessary. While in the literature, several pilot patterns are possible [7], there are two patterns frequently used in a practical system. The first one dedicates certain subcarriers to pilot symbols in each OFDM block, while the second one dedicates all subcarriers in certain OFDM blocks to pilot symbols. In this paper, we assume the second pattern (Fig. 1), in which the  $i, i+K, \dots, i+(M-1)K$ -th blocks are pilot blocks.

### B. Channel Model

The doubly-selective fading channel in mobile wireless communications is often modelled as wide-sense-stationary uncorrelated-scattering (WSSUS) channel with the complex baseband representation described by

$$h(t, \tau) = \sum_{l=0}^{L-1} h(t, \tau_l) \delta(\tau - \tau_l) \quad (3)$$

where  $\tau_l$  is the delay,  $\delta(\cdot)$  denotes Dirac's delta function and  $L$  is the number of paths. The maximum delay is denoted by  $T_m$ . The autocorrelation function of the WSSUS channel is given by  $E\{h(t, \tau)h^*(t + \Delta t, \tau + \Delta \tau)\} = \phi_h(\Delta t, \tau)\delta(\tau)$ . The autocorrelation function can be separated into the multiplication of time correlation  $\phi_t(\Delta t)$  and power delay profile (PDP)  $\phi_\tau(\tau)$ . In a rich-scattering environment,  $\phi_h(\Delta t, \tau) = \phi_t(\Delta t)\phi_\tau(\tau)$ . The Fourier transform of the  $\phi_t(\Delta t)$  gives the Doppler power spectrum  $\phi_D(f)$ . In the classic Jakes' model,  $\phi_t(\Delta t) = J_0(2\pi f_d \Delta t)$  and  $\phi_D(f) = \frac{1}{\pi f_d \sqrt{1-(f/f_d)^2}} |f| < f_d$ , where  $J_0(x)$  is the zeroth-order Bessel function of the first kind and  $f_d$  is the maximum Doppler frequency. In a Rayleigh fading channel,  $h(t, \tau_l)$ 's are complex Gaussian processes with zero mean and variance  $\sigma_l^2 = \phi_\tau(\tau_l)$ , which

are independent for different paths. The channel is normalized so that  $\sum_{l=0}^{L-1} \sigma_l^2 = 1$ .

The Fourier transform of  $h(t, \tau_l)$  over time  $t$  is the delay Doppler-spread function  $S(f, \tau_l)$ , which is given by

$$S(f, \tau_l) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} h(t, \tau_l) e^{-j2\pi ft} dt. \quad (4)$$

This function is band limited in Doppler frequency, i.e.  $S(f, \tau_l) = 0$  for  $|f| > f_d$ . The inverse Fourier transform gives

$$h(t, \tau_l) = \int_{-f_d}^{+f_d} S(f, \tau_l) e^{j2\pi ft} df. \quad (5)$$

In the time interval  $NT_s$ , the Riemann integral in (5) can be approximated by the Riemann summation after sampling. The discrete-time baseband equivalent channel model at the  $i$ -th block can be described as

$$h^{(i)}(n, d_l) \doteq \sum_{q=-Q}^Q h_q^{(i)}(d_l) e^{j\frac{2\pi q n}{N_s}} \quad (6)$$

where  $h^{(i)}(n, d_l) = h[(i-1)(N+N_g)T_s + nT_s, d_l]$ ,  $d_l = \lceil \tau_l/T_s \rceil$  is the discrete delay after sampling,  $N_s = GN$  is the number of samples in Doppler frequency domain and  $Q = \lceil f_d N_s T_s \rceil = \lceil f_d G N T_s \rceil$ . If  $G = 1$ , (6) is equivalent to the BEM in [3]. Therefore (6) can be viewed as the oversampling BEM. Without oversampling, the reconstructed channel differs from  $h(n, d_l)$  at the borders of the observation interval. The time varying channel  $h(t, \tau_l)$  within  $NT_s$  can now be represented by  $2Q+1$  BEM coefficients. From simulation, we find that (6) is valid for  $0 \leq n \leq N_c = \lfloor GN/4 \rfloor$ . Eq. (6) can be written in matrix form as

$$\mathbf{h}^{(i)}(d_l) = \mathbf{\Phi} \mathbf{h}_b^{(i)}(d_l) \quad (7)$$

where  $\mathbf{h}^{(i)}(d_l) = [h^{(i)}(0, d_l), \dots, h^{(i)}(N_c, d_l)]^T$ ,  $\mathbf{h}_b(d_l) = [h_{-Q}^{(i)}(d_l), \dots, h_Q^{(i)}(d_l)]^T$  and  $\mathbf{\Phi}(n, \cdot) = [e^{-j\frac{2\pi Q}{N_s}n}, \dots, e^{j\frac{2\pi Q}{N_s}n}]$ . Here we use the Matlab notation and subscript  $b$  denotes the use of BEM. If  $N_c \geq 2Q+1$ ,  $\mathbf{\Phi}$  is invertible. From (7), we can obtain

$$E\{\mathbf{h}_b^{(i)}(d_l)\} = E\{\mathbf{\Phi}^{-1} \mathbf{h}^{(i)}(d_l)\} = \mathbf{0} \quad (8)$$

and

$$\mathbf{R}_b(d_l) = E\{\mathbf{h}_b^{(i)}(d_l) [\mathbf{h}_b^{(i)}(d_l)]^H\} = \sigma_l^2 \mathbf{\Phi}^{-1} \mathbf{R}_h (\mathbf{\Phi}^{-1})^H \quad (9)$$

where  $\mathbf{R}_h(k_1, k_2) = \phi_t((k_2 - k_1)T_s)$ .  $\mathbf{R}_b(d_l)$  is independent of  $i$ . If  $G = 1$  and  $N_c = N$ , it can be readily verified that  $\mathbf{R}_b$  becomes a diagonal matrix, which means  $h_q^{(i)}(d_l)$ 's are independent and verifies the Assumption 3 in [3]. If  $G > 1$ ,  $h_q^{(i)}(d_l)$ 's are correlated but  $h_q^{(i)}(d_l)$ 's for different paths are independent. From (9),  $\mathbf{R}_b$  depends on  $N_c$ . We find when  $N_c > N$ , the change of  $\mathbf{R}_b$  with the increase of  $N_c$  is negligible.

### III. ICI AND DIVERSITY ANALYSIS

The received signal  $y^{(i)}(n)$  in the  $i$ -th OFDM block can be expressed as

$$y_n^{(i)} = \sum_{l=0}^{L-1} h^{(i)}(n, d_l) x_{n-d_l}^{(i)} + w_n^{(i)} \quad (10)$$

where  $w_n^{(i)}$  is additive white Gaussian noise with zero-mean and variance  $\sigma_n^2$ . Substituting the BEM (6) into (10), we have

$$y_n^{(i)} = \sum_{l=0}^{L-1} \sum_{q=-Q}^Q h_q^{(i)}(d_l) e^{j\frac{2\pi q}{N_s} n} x_{n-d_l}^{(i)} + w_n^{(i)}. \quad (11)$$

If  $G = 1$ ,  $h_q^{(i)}(d_l)$ 's are independent. The representation (11) can be interpreted as  $L(2Q+1)$  independent, flat-fading channels. Therefore, the maximum diversity order is proportional to the product  $T_m/T_s(2f_dNT_s + 1)$ . The additional diversity due to the time-selectivity is proportional to  $NT_m f_d$ . With appropriate linearly precoding, the maximum diversity can be achieved [3]. Eq. (11) can be written in matrix form as

$$\mathbf{y}^{(i)} = \sum_{q=-Q}^Q \mathbf{\Gamma}_q \mathcal{S}^{(N)} [\mathbf{h}_q^{(i)}] \mathbf{x}^{(i)} + \mathbf{w}^{(i)} \quad (12)$$

where  $\mathbf{\Gamma}_q = \text{diag}\{1, e^{j\frac{2\pi q}{N_s}}, \dots, e^{j\frac{2\pi q}{N_s}(N-1)}\}$ ,  $\mathbf{h}_q^{(i)} = [h_q^{(i)}(d_0), 0, \dots, h_q^{(i)}(d_{L-1})]^T$ ,  $\mathbf{x}^{(i)} = [x_0^{(i)}, x_1^{(i)}, \dots, x_{N-1}^{(i)}]^T$  and  $\mathbf{w}^{(i)} = [w_0^{(i)}, w_1^{(i)}, \dots, w_{N-1}^{(i)}]^T$ . After DFT, (12) becomes

$$\mathbf{Y}^{(i)} = \sum_{q=-Q}^Q \mathbf{\Xi}_q \mathbf{H}_q^{(i)} \mathbf{X}^{(i)} + \mathbf{W}^{(i)} \quad (13)$$

where  $\mathbf{X}^{(i)} = \mathbf{F}^H \mathbf{x}^{(i)}$ ,  $\mathbf{H}_q^{(i)} = \text{diag}\{H_q^{(i)}(0), \dots, H_q^{(i)}(N-1)\}$ ,  $H_q^{(i)}(k) = \sum_{l=0}^{L-1} h_q^{(i)}(d_l) e^{-j2\pi k d_l / N}$  and  $\mathbf{W}^{(i)} = \mathbf{F}^H \mathbf{w}^{(i)}$ .  $\mathbf{\Xi}_q$  is the CFO induced ICI matrix with frequency offset  $\epsilon = q/G$ . Therefore, the effect of ICI in (13) is the summation of CFO with  $-Q/G, \dots, Q/G$ . When  $G \rightarrow +\infty$ , the ICI induced by Doppler frequency  $f_d$  is the accumulation of the ICI induced by frequency offset within  $[-f_d, f_d]$ , which establishes the relationship between Doppler induced ICI and CFO induced ICI. Different from the CFO induced ICI, the received vector  $\mathbf{Y}^{(i)}$  (13) has  $2Q+1$  replicas of  $\mathbf{X}^{(i)}$ , which offers time-diversity. When only frequency offset exists, no diversity can be achieved. The compensation of Doppler induced ICI is more challenging since it requires many more parameters to estimate than for the case of CFO. Next we will show that with our channel estimators and near-optimal equalizers, time-diversity can be achieved with low complexity.

#### IV. CHANNEL ESTIMATION

In a doubly selective channel, it is impossible to estimate all the  $LN$  parameters with only one pilot block. In [2],  $M$  pilot blocks are collected to perform MMSE channel estimation. However, the complexity of this estimator is high. In this section, we present two computationally efficient LMMSE channel estimators using the BEM (6). The pilot pattern is as Fig. 1.

##### A. Channel estimator 1

As indicated in Section II B., the BEM (6) is valid for  $0 \leq n \leq N_c = \lfloor GN/4 \rfloor$ . If we assume that the BEM coefficients are constant within  $(M-1)K+1$  OFDM blocks,

$G$  must satisfy  $\lfloor GN/4 \rfloor > ((M-1)K+1)N$ . The input-output relationship (11) can be written in matrix form as

$$\mathbf{y}^{(i)} = \sum_{q=-Q}^Q \mathbf{\Gamma}_q \mathcal{S}^{(L)} [\mathbf{x}^{(i)}] \mathbf{h}_q + \mathbf{w}^{(i)} \quad (14)$$

$$= \mathbf{D}^{(i)} \mathbf{h}_b + \mathbf{w}^{(i)},$$

where  $\mathbf{\Gamma}_q$  is defined in (12),  $\mathbf{h}_q = [h_q(d_0), \dots, h_q(d_{L-1})]^T$ ,  $\mathbf{h}_b = [\mathbf{h}_{-Q}^T, \dots, \mathbf{h}_Q^T]^T$  and  $\mathbf{D}^{(i)} = [\mathbf{\Gamma}_{-Q} \mathcal{S}^{(L)} [\mathbf{x}^{(i)}], \dots, \mathbf{\Gamma}_Q \mathcal{S}^{(L)} [\mathbf{x}^{(i)}]]$ . Since  $\mathbf{h}_q$  is constant within  $(M-1)K+1$  OFDM blocks, we omit the superscript  $(i)$  in  $\mathbf{h}_b$ . When collecting all the received vector in  $M$  pilot blocks, e.g.  $i, i+K, \dots, i+(M-1)K$  blocks, the input-output relationship becomes

$$\mathbf{y} = \mathbf{D} \mathbf{h}_b + \mathbf{w} \quad (15)$$

where  $\mathbf{y} = [(\mathbf{y}^{(i)})^T, \dots, (\mathbf{y}^{(i+(M-1)K)})^T]^T$ ,  $\mathbf{D} = [(\mathbf{D}^{(i)})^T, \dots, (\mathbf{D}^{(i+(M-1)K)})^T]^T$  and  $\mathbf{w} = [(\mathbf{w}^{(i)})^T, \dots, (\mathbf{w}^{(i+(M-1)K)})^T]^T$ . Since we have  $(2Q+1)L$  unknown BEM coefficients, to ensure uniqueness in estimating the channel, the total number of pilot symbols must satisfy  $MN \geq (2Q+1)L$ . The LMMSE channel estimator (CE1) is given by

$$\hat{\mathbf{h}}_b = (\sigma_n^2 \mathbf{R}_b^{-1} + \mathbf{D}^H \mathbf{D})^{-1} \mathbf{D}^H \mathbf{y} \quad (16)$$

where  $\mathbf{R}_b = E\{\mathbf{h}_b \mathbf{h}_b^H\}$  is the covariance matrix of  $\mathbf{h}_b$  with  $(\mathbf{R}_b)_{\text{I.L:L}(2Q+1), \text{I.L:L}(2Q+1)} = \mathbf{R}_b(d_l)$  and  $\mathbf{R}_b(d_l)$  can be obtained from (9). The covariance of the error  $\tilde{\mathbf{h}}_b = \mathbf{h}_b - \hat{\mathbf{h}}_b$  is given by

$$\mathbf{R}_{\tilde{\mathbf{h}}_b} = E\{\tilde{\mathbf{h}}_b \tilde{\mathbf{h}}_b^H\} = \sigma_n^2 (\sigma_n^2 \mathbf{R}_b^{-1} + \mathbf{D}^H \mathbf{D})^{-1} \quad (17)$$

and the mean square error (MSE) of  $\hat{\mathbf{h}}_b$  is thus

$$\sigma_{\tilde{\mathbf{h}}_b}^2 = \text{tr}(\mathbf{R}_{\tilde{\mathbf{h}}_b}) = \sigma_n^2 \text{tr} \left( (\sigma_n^2 \mathbf{R}_b^{-1} + \mathbf{D}^H \mathbf{D})^{-1} \right). \quad (18)$$

With oversampling  $\mathbf{R}_b$  is non-diagonal, the pilots design approach in [3] fails in this case. The optimal pilots can be found by minimizing the MSE (18) via exhaustive search.

Although the BEM (6) is valid for  $0 \leq n \leq N_c = \lfloor GN/4 \rfloor$ , (6) is still a good approximation when  $n$  is relatively small. To improve the accuracy of CE1, we use the  $i, \dots, i+(M-1)K$ -th pilot blocks to estimate the BEM coefficients in the  $i+1, \dots, i+K-1$ -th blocks. The coefficients in the  $i+(m-1)K+1, \dots, i+mK-1$ -th blocks are estimated using the  $i+(m-1)K, i+mK, \dots, i+(m+M-2)K$ -th pilot blocks.

##### B. Channel estimator 2

We now consider a more general case where the BEM coefficients are allowed to vary. The BEM coefficients are calculated independently for different pilot blocks and Wiener filtering gets the BEM coefficients at other blocks. The estimation accuracy can be improved using this estimator.

In the  $i$ th block, the input-output relationship (11) can be written in matrix form as

$$\mathbf{y}^{(i)} = \mathbf{D}^{(i)} \mathbf{h}_b^{(i)} + \mathbf{w}^{(i)}. \quad (19)$$

where  $\Gamma_q$  is defined in (12),  $\mathbf{h}_q^{(i)} = [h_q^{(i)}(d_0), \dots, h_q^{(i)}(d_{L-1})]^T$ ,  $\mathbf{h}_b^{(i)} = \left[ \left( \mathbf{h}_{-Q}^{(i)} \right)^T, \left( \mathbf{h}_{-Q+1}^{(i)} \right)^T, \dots, \left( \mathbf{h}_Q^{(i)} \right)^T \right]^T$  and  $\mathbf{D}^{(i)} = [\Gamma_{-Q} \mathcal{S}^{(L)}[\mathbf{x}^{(i)}], \dots, \Gamma_Q \mathcal{S}^{(L)}[\mathbf{x}^{(i)}]]$ . The LMMSE channel estimate for  $\mathbf{h}_b^{(i)}$  is given by

$$\hat{\mathbf{h}}_b^{(i)} = \left( \sigma_n^2 \mathbf{R}_b^{-1} + \left( \mathbf{D}^{(i)} \right)^H \mathbf{D}^{(i)} \right)^{-1} \left( \mathbf{D}^{(i)} \right)^H \mathbf{y}^{(i)} \quad (20)$$

$$= \Psi^{(i)} \mathbf{y}^{(i)}$$

where  $\Psi^{(i)} = \left( \sigma_n^2 \mathbf{R}_b^{-1} + \left( \mathbf{D}^{(i)} \right)^H \mathbf{D}^{(i)} \right)^{-1} \left( \mathbf{D}^{(i)} \right)^H$  and  $\mathbf{R}_b$  is defined in (16).  $\mathbf{h}_b^{(i+K)}, \dots, \mathbf{h}_b^{(i+(M-1)K)}$  can be estimated in the same way. All the  $M$  vectors in the pilot blocks are then combined to estimate the BEM coefficients in the data blocks. Define the estimated BEM coefficients vector  $\mathbf{h}_b = \left[ \left( \hat{\mathbf{h}}_b^{(i)} \right)^T, \left( \hat{\mathbf{h}}_b^{(i+K)} \right)^T, \dots, \left( \hat{\mathbf{h}}_b^{(i+(M-1)K)} \right)^T \right]^T$ . The BEM coefficients in the  $k$ th ( $i < k < i + K$ ) block can be obtained via LMMSE as (CE2)

$$\hat{\mathbf{h}}_b^{(k)} = \mathbf{R}_{\mathbf{h}_b^{(k)} \mathbf{h}_b} \mathbf{R}_{\mathbf{h}_b \mathbf{h}_b}^{-1} \mathbf{h}_b \quad (21)$$

where  $\mathbf{R}_{\mathbf{h}_b^{(k)} \mathbf{h}_b} = E\{\mathbf{h}_b^{(k)} (\mathbf{h}_b)^H\}$  and  $\mathbf{R}_{\mathbf{h}_b \mathbf{h}_b} = E\{\mathbf{h}_b (\mathbf{h}_b)^H\}$ . From (7)

$$E\left\{ \mathbf{h}_b^{(k)} \left( \hat{\mathbf{h}}_b^{(i)} \right)^H \right\} = \left( \Phi \right)^{-1} \mathbf{R}_{\mathbf{h}^{(k)} \mathbf{y}^{(i)}} \left( \Psi^{(i)} \right)^H \quad (22)$$

where  $\mathbf{h}^{(k)} = \left[ \left( \mathbf{h}^{(k)}(d_0) \right)^T, \dots, \left( \mathbf{h}^{(k)}(d_{L-1}) \right)^T \right]^T$  and  $\mathbf{R}_{\mathbf{h}^{(k)} \mathbf{y}^{(i)}} = E\{\mathbf{h}^{(k)} (\mathbf{y}^{(i)})^H\}$ . From (20)

$$E\left\{ \hat{\mathbf{h}}_b^{(i)} \left( \hat{\mathbf{h}}_b^{(j)} \right)^H \right\} = \Psi^{(i)} \mathbf{R}_{\mathbf{y}^{(i)} \mathbf{y}^{(j)}} \left( \Psi^{(j)} \right)^H \quad (23)$$

where  $\mathbf{R}_{\mathbf{y}^{(i)} \mathbf{y}^{(j)}} = E\{\mathbf{y}^{(i)} (\mathbf{y}^{(j)})^H\}$ . Both  $\mathbf{R}_{\mathbf{h}^{(k)} \mathbf{y}^{(i)}}$  and  $\mathbf{R}_{\mathbf{y}^{(i)} \mathbf{y}^{(j)}}$  can be obtained from (10). The MSE of CE2 is given by

$$\sigma_k^2 = \text{tr} \left( \mathbf{R}_b - \mathbf{R}_{\mathbf{h}_b^{(k)} \mathbf{h}_b} \mathbf{R}_{\mathbf{h}_b \mathbf{h}_b}^{-1} \mathbf{R}_{\mathbf{h}_b \mathbf{h}_b^{(k)}} \right) \quad (24)$$

where  $\mathbf{R}_b$  is defined in (16). From (21), the MSE depends on the block index  $k$ . The optimal pilots for CE2 can be found by minimizing the summation of all the MSE's at data blocks.

### C. Complexity analysis

In CE1, the size of  $\mathbf{y}$  in (16) is  $MN$ . If the estimated BEM coefficients are used for the blocks  $i+1, \dots, i+(M-1)K$ , the average number of complex multiplications in each block is  $\frac{(2Q+1)LNM}{(K-1)(M-1)}$ . If they are only used for the blocks  $i+1, \dots, i+K-1$ , the average number of complex multiplications in each block is  $\frac{(2Q+1)LNM}{K-1}$ . In CE2, estimating  $\hat{\mathbf{h}}_b^{(i+(m-1)K)}$  for  $m = 1, 2, \dots, M$  using (20) needs  $NML(2Q+1)$  complex multiplications and estimating each  $\hat{\mathbf{h}}_b^{(k)}$  using (21) needs  $ML^2(2Q+1)^2$ . Therefore the average number of complex multiplications is  $\frac{NML(2Q+1)}{(K-1)(M-1)} + ML^2(2Q+1)^2$ . Compared with the channel estimator in [2], whose complexity is  $LMN^2$ , our proposed two channel estimators have much lower complexity.

## V. LOW COMPLEXITY EQUALIZATION

If the BEM vector  $\mathbf{h}_q^{(i)}$  is estimated, (13) becomes

$$\mathbf{Y}^{(i)} = \sum_{q=-Q}^Q \Xi_q \hat{\mathbf{H}}_q^{(i)} \mathbf{X}^{(i)} + \mathbf{W}^{(i)} \quad (25)$$

$$= \mathbf{H}^{(i)} \mathbf{X}^{(i)} + \mathbf{W}^{(i)}$$

where  $\mathbf{H}^{(i)} = \sum_{q=-Q}^Q \Xi_q \hat{\mathbf{H}}_q^{(i)}$ . For simplicity, we ignore the superscript  $^{(i)}$ . The data detection problem is thus

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X} \in \mathcal{Q}^N} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|^2 \quad (26)$$

where  $\mathcal{Q}$  is the finite constellation set. Eq. (26) can be solved by both V-BLAST and SD. Details are omitted for brevity.

The complexity of V-BLAST is  $O(N^3)$  and the complexity of SD is larger than  $O(N^3)$ . When  $N$  is very large, the complexity is still high. To further reduce the complexity, we propose an iterative and low complexity equalizer (ILE) without performing matrix inversion.

As indicated in [6], most of the ICI on each subcarrier comes from several neighboring subcarriers. The matrix  $\mathbf{H}$  in (26) can thus be approximated by a band matrix where the half of the band width is  $P$ , meaning  $H_{i,j} = 0$  for  $|i-j| > P$ . Since the matrix becomes sparse for  $P \leq N$ , it is not efficient to calculate the matrix inverse of  $\mathbf{H}$  in V-BLAST and SD. We start by using one-tap equalization without considering the ICI term. The initial estimate of  $X_k^{(0)}$  is chosen from  $\mathcal{Q}$  close to  $Y_k/H_{k,k}$ , where the superscript  $^{(i)}$  denotes the  $i$ -th iteration.  $\mathbf{X}$  is partitioned into  $K$  groups and each group has  $S$  elements.  $\mathbf{e} = \mathbf{Y} - \mathbf{H}\mathbf{X}^{(0)}$  is first computed. In the  $i$ -th iteration, the  $k$ -th group  $\mathbf{x}_k^{(i)} = \{X_{(k-1)S+1}^{(i)}, \dots, X_{kS}^{(i)}\}$  is determined by

$$\mathbf{x}_k^{(i)} = \arg \min_{\mathbf{x}_k^{(i)} \in \mathcal{Q}^S} \|\mathbf{e}_{(k-1)S+1:P:kS+P} + \mathbf{H}_{(k-1)S+1:P:kS+P, (k-1)S+1:kS} \mathbf{x}_k^{(i-1)} - \mathbf{x}_k^{(i)}\|^2. \quad (27)$$

The same iteration continues. In each iteration, the complexity is proportional to  $|\mathcal{Q}|^S(2P+1)K$ . From simulation, we find  $S = 1$  and only 2-3 iterations can ensure convergence. Therefore the complexity of ILE is  $O((2P+1)N)$ . Compared with the V-BLAST  $O(N^3)$  and the equalizer in [6]  $O(N^2(2P+1))$ , the complexity of ILE is low and offers good performance.

## VI. SIMULATION RESULTS

We now present numerical results to illustrate the effectiveness of our proposed channel estimators and equalizers for a practical OFDM system. We assume the following system specifications:

- Both the data and pilot symbols are chosen from binary phase-shift keying (BPSK).
- The carrier frequency of the OFDM system is 5 GHz and the data rate is 6 MHz.
- The 6-ray COST 207 TU model with the PDP [0.189, 0.379, 0.239, 0.095, 0.061, 0.037] and delay profile [0.0, 0.2, 0.5, 1.6, 2.3, 5.0]  $\mu\text{s}$  is considered. A normalized Doppler frequency  $Nf_d T_s = 0.1$  is considered. Each path is a complex Gaussian random process independently generated with the classical Doppler spectrum using the inverse discrete Fourier transform method. All

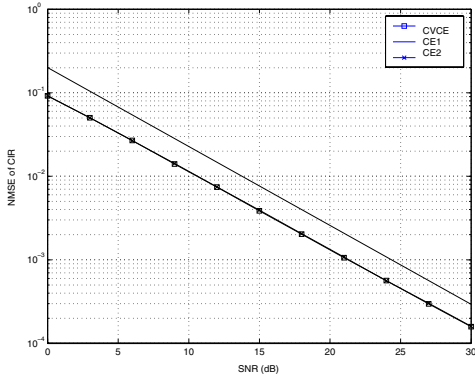


Fig. 2. NMSE of CIR versus SNR in a BPSK OFDM system.

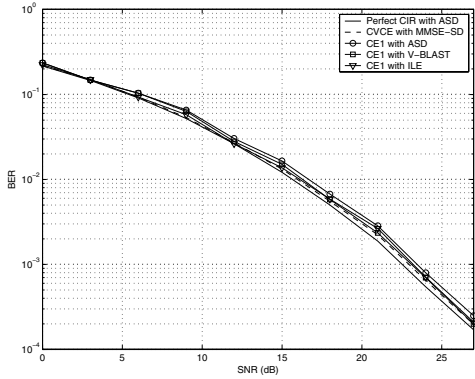


Fig. 3. BER versus SNR with in a BPSK OFDM system with CE1, CVCE and perfect CIR.

the paths are rounded to integer and we do not consider leakage. The PDP is assumed perfectly known at the receiver.

- The number of subcarriers  $N = 64$ . The guard interval  $N_d = 16$ . We set  $M = 6$  and  $K = 2$  in Fig. 1.

To make CE1 valid for the selected pilot pattern,  $[G_1 N/4] > ((M - 1)K + 1)N$ ,  $G_1 = 60$  and  $Q_1 = 6$ . For CE2,  $G_2 = 20$  and  $Q_2 = 2$ . For  $N f_d T_s = 0.1$ , more than 98% of the  $X_k$ 's energy is distributed on the  $k$ -th and its two neighboring subcarriers [6]. Therefore we choose  $P = 4$  and  $S = 1$  in the ILE.

Fig. 2 compares the average MSE of our proposed estimators with that of the estimator in [2] (CVCE). The CE2 performs close to CVCE. The CE1 has a 2.2-dB gap with CE2 and CVCE, which may be due to the assumption that the BEM coefficients are constant within  $(M - 1)K + 1$  blocks. The gap can be reduced with the increase of oversampling factor  $G$  but will increase complexity. The average number of complex multiplications is 5990.4 for CE1, 7704 for CE2 and 147456 for CVCE. Our proposed two channel estimators hence reduce complexity about 200 times while still performing well. The parameter  $Q$  offers a tradeoff between complexity and performance.

Figs. 3, 4 compare the BER of different equalizers. The ASD with perfect CIR is used as a benchmark. The CE1 with SD has a 0.4-dB gain over that with V-BLAST but it is 0.1 dB worse than the CVCE with MMSE-SD and 0.55 dB worse than

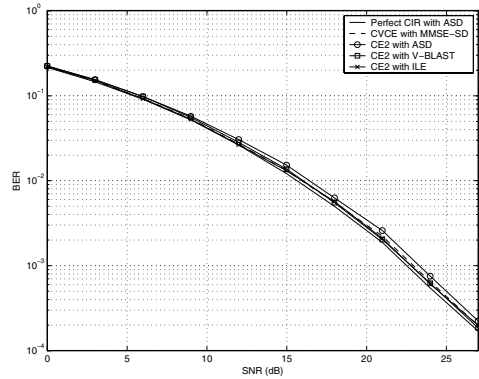


Fig. 4. BER versus SNR with in a BPSK OFDM system with CE2, CVCE and perfect CIR.

the perfect CIR with ASD at  $\text{BER}=10^{-3}$  (Fig. 3). This is due to the 2.2-dB loss in channel estimation with CE1. The CE2 with ASD has a 0.5-dB gain over that with V-BLAST and a 0.15-dB gain over the CVCE with MMSE-SD but it is 0.3 dB worse than the perfect CIR with ASD at  $\text{BER}=10^{-3}$  (Fig. 4). The performance loss of CE2 with ASD is negligible compared with the benchmark. The ILE performs close to ASD and has low complexity.

## VII. CONCLUSION

This paper considered the estimation of channel and data for OFDM systems over doubly-selective channels. We derived an oversampling BEM for doubly-selective channels. Its statistical properties are derived and used for the LMMSE channel estimation. Our proposed two channel estimators have low complexity but perform similar to that of [2]. To exploit the time-diversity introduced by time-selectivity, we contribute an iterative low-complexity equalizer, which performs nearly ML. The complexity of Channel Estimator 2 can be further reduced using a reduced-rank approximation. Our proposed estimators can also be extended to coded OFDM systems.

## 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, no. 1, pp. 27 – 32, Jan. 1999.
- [2] Y.-S. Choi, P. Voltz, and F. 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.
- [3] X. Ma and G. B. Giannakis, "Maximum-diversity transmissions over doubly selective wireless channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1832 – 1840, July 2003.
- [4] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14–15, Jan. 1999.
- [5] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Computation*, vol. 44, pp. 463–471, Apr. 1985.
- [6] X. Cai and G. B. Giannakis, "Bounding performance and suppressing intercarrier interference in wireless mobile OFDM," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2047 – 2056, Dec. 2003.
- [7] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.