

# Semi-Blind Equalization for OFDM Systems over Fast Fading Channels

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

**Abstract**—We consider semi-blind data detectors for OFDM systems over fast fading channels. Three maximum-likelihood (ML) data detectors (MLD) are derived assuming exact channel correlation. The first two exploit the correlation among pre-DFT and post-DFT received signals and the third one average the ML metric over channel impulse response (CIR). We show that the first two MLDs for input data symbols are given by an integer least-squares (LS) minimization problem, which can be efficiently solved by V-BLAST detection (suboptimal) or by a sphere decoder (SD) (optimal). We contribute a low-complexity near-ML SD and an iterative detector for the third MLD. Despite its low complexity, MLD1 performs robustly for normalized Doppler rates less than 3%, which meets the requirement of many practical systems. With iterative detection, both MLD1 and MLD3 can exploit time-diversity.

## I. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) divides the transmission bandwidth into many narrow sub-channels which are transmitted in parallel. If the channel is time-invariant, channel equalization can be easily accomplished by a single-tap frequency equalizer. OFDM is sensitive to Doppler spread and carrier frequency offset (CFO), which destroys the subcarrier orthogonality and gives rise to Inter-Carrier-Interference (ICI). In OFDM, CFO can be estimated by either training-based or blind techniques and can be compensated. Without estimating the carrier offset, transmit precoding can be employed to suppress its impact on ICI. On the other hand, Doppler-induced ICI appears to be more challenging. The effects of ICI are analyzed in [1], where central limit theorem arguments are used to model ICI as a Gaussian random process and to quantify its impact on the bit error rate (BER). Bounds on signal-to-interference-and-noise ratio (SINR) (due to ICI) are given in [2].

Many channel estimation algorithms have been proposed for compensating time-varying channels and the Doppler-induced ICI. 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 its estimation has been thoroughly studied in [3], [4] (and references therein). When the OFDM symbol duration is less than 10% of the channel coherence time, the channel varies linearly [5] and can be estimated by linear interpolation

between two channel estimates acquired by training blocks. In [6], an MMSE channel estimator and a successive interference cancellation (SIC) scheme with optimal ordering are proposed. Since the channel is time-varying, it is difficult or complicated to estimate the channel impulse response (CIR) directly. The principal motivation of this paper is therefore to derive semi-blind data detectors. That is, input data symbols are detected without first estimating the CIR.

In this paper, we derive three MLDs by using the maximum-likelihood principle. The term "semi-blind" is used since we assume the channel correlation and noise variance are known at the receiver. The first two MLDs exploit the correlation among pre-DFT and post-DFT received samples. The third one averages the maximum likelihood (ML) function over the CIR. Exhaustive search of the resulting solution space yields the ML solution but has exponential complexity in the number of subcarriers and is computationally prohibitive. When  $f_d T < 0.03$ , the cost function of the first two MLDs can be written in quadratic form. The vertical Bell Labs layered space time (V-BLAST) [7] algorithm can be used to suboptimally solve the quadratic form. However, near ML performance is given by the sphere decoder (SD) [8]. When  $f_d T < 0.03$  in MLD1 or MLD3, we give an iterative detection algorithm, which is a variant of both exhaustive search and greedy algorithms.

*Notation:*  $(\cdot)^T$  and  $(\cdot)^H$  denote transpose and conjugate transpose. The set of complex  $K \times 1$  vectors is denoted by  $\mathcal{C}^K$ . A  $M$ -ary signal constellation is denoted by  $\mathcal{Q}$ . A complex Gaussian variable with mean  $\mu$  and variance  $\sigma^2$  is denoted by  $z \sim \mathcal{CN}(\mu, \sigma^2)$ . The discrete Fourier transform (DFT) matrix of size  $N \times N$  is given by  $\mathbf{F} = 1/\sqrt{N}[e^{j\frac{2\pi}{N}kl}]$ ,  $k, l \in 0, 1, \dots, N-1$ ,  $j = \sqrt{-1}$ . The diagonal matrix formed by a vector  $\mathbf{A}$  is  $\mathbf{A}_D$ .

## II. SYSTEM MODEL

### A. OFDM baseband model

The binary source data are grouped and/or mapped into the multi-phase symbols from a finite constellation  $\mathcal{Q}$ , which are modulated by inverse DFT (IDFT) on  $N$  parallel subcarriers.

The resulting time domain samples are

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

where  $X_k \in \mathcal{Q}$ . Note that  $X_k, k = 0, 1, \dots, N-1$ , are called modulation symbols and the term ‘‘OFDM symbol’’ is used to denote the entire sequence  $\{x(0), x(1), \dots, x(N-1)\}$ . The input symbol duration is  $T_s$  and the OFDM symbol duration  $T = NT_s$ . These samples are appropriately pulse shaped to construct the time domain signal  $x(t)$  for transmission.

A guard interval, inserted to prevent inter-frame interference, includes a cyclic prefix of  $\{x(N_g - 1), \dots, x(N-1)\}$  where  $N_g$  is the number of samples in the guard interval ( $N_g$  is assumed to be larger than the delay spread of the channel). The composite response which includes transmit and receive pulse shaping and the physical channel response between the transmitter and receiver may be modelled as [3]

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

where  $h_l(t) \sim \mathcal{CN}(0, \sigma_l^2)$  and  $\tau_l$  is the delay of the  $l$ -th tap. Typically, it is assumed that  $\tau_l = lT_s$  and this results in a finite impulse response filter with an effective length  $L$ . The received signal after sampling can be represented as

$$\begin{aligned} y(n) &= \sum_{l=0}^{L-1} h_l(n) x(n-l) + w(n) \\ &= \frac{1}{\sqrt{N}} \sum_{l=0}^{L-1} h_l(n) \sum_{k=0}^{N-1} X_k e^{j(2\pi k(n-l)/N)} + w(n) \end{aligned} \quad (2)$$

where  $w(n)$  is Additive White Gaussian Noise (AWGN). At the receiver, the guard interval is removed and DFT demodulation is performed, resulting

$$\begin{aligned} Y_k &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y(n) e^{-j(2\pi kn/N)}, \quad k = 0, 1, 2, \dots, N-1. \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} h_l(n) x(n-l) e^{-j(2\pi kn/N)} + W_k \\ &= X_k H_k + \alpha_k + W_k \end{aligned} \quad (3)$$

where  $H_k = 1/N \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} h_l(n) e^{-j2\pi kl/N}$ ,  $W_k = 1/\sqrt{N} \sum_{n=0}^{N-1} w(n) e^{-j2\pi nk/N}$  and

$$\alpha_k = \frac{1}{N} \sum_{\substack{m=0, \\ m \neq k}}^{N-1} X_m \sum_{n=0}^{N-1} \left( \sum_{l=0}^{L-1} h_l(n) e^{-j2\pi ml/N} \right) e^{j2\pi n(m-k)/N}.$$

The  $\alpha_k$ 's represent ICI caused by the channel time-variation.

## B. Channel model

A wide-sense stationary uncorrelated scattering (WSSUS) channel (1) is characterized by its power delay profile (PDP) and scattering function. We assume that  $h_l(t)$  has the same normalized correlation function  $r_t(\Delta t)$ . Hence

$$r_{h_l}(\Delta t) = E\{h_l(t + \Delta t) h_l^*(t)\} = \sigma_l^2 r_t(\Delta t). \quad (4)$$

The autocorrelation function of the channel is

$$E\{h_{l_1}(t + \Delta t) h_{l_2}^*(t)\} = \sigma_{l_1}^2 r_t(\Delta t) \delta(l_1 - l_2). \quad (5)$$

From Jakes' model [9],  $r_t(\Delta t) = J_0(2\pi f_d \Delta t)$ .  $J_0(\cdot)$  denotes the zeroth-order Bessel function of the first kind, and  $f_d$  is Doppler frequency in hertz. We define frequency domain correlation as

$$r_f(\Delta f) = \sum_{l=0}^{L-1} \sigma_l^2 e^{-j2\pi \Delta f \tau_l}. \quad (6)$$

Clearly  $r_t(\Delta t)$  is dependent on the Doppler frequency, while  $r_f(\Delta f)$  depends on the PDP.

## III. MAXIMUM LIKELIHOOD DATA DETECTION IN A FAST FADING CHANNEL

### A. MLD 1

This estimator exploits the correlation among pre-DFT received samples. The correlation between  $y(n_1)$  and  $y(n_2)$  in (2) can be written as

$$\begin{aligned} \mathbf{R}_{yy}(n_1, n_2) &= E\{y(n_1) y^*(n_2)\} \\ &= \frac{1}{N} \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} X_{k_1} X_{k_2}^* e^{j2\pi k_1 n_1/N} e^{-j2\pi k_2 n_2/N} \\ &\quad \times r_t[(n_1 - n_2)T_s] \sum_{l=0}^{L-1} \sigma_l^2 e^{-j2\pi(k_1 - k_2)l/N} + \sigma_n^2 \delta(n_1 - n_2) \\ &= \frac{1}{N} \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-1} X_{k_1} X_{k_2}^* e^{j2\pi k_1 n_1/N} e^{-j2\pi k_2 n_2/N} \\ &\quad \times r_t[(n_1 - n_2)T_s] r_f\left(\frac{k_1 - k_2}{NT_s}\right) + \sigma_n^2 \delta(n_1 - n_2) \\ &= r_t[(n_1 - n_2)T_s] \mathbf{f}_{n_1}^H \mathbf{X}_D \mathbf{R}_f \mathbf{X}_D^H \mathbf{f}_{n_2} + \sigma_n^2 \delta(n_1 - n_2) \end{aligned} \quad (7)$$

where  $\mathbf{X}_D = \text{diag}\{X_0, X_1, \dots, X_{N-1}\}$ , the frequency domain correlation matrix  $[\mathbf{R}_f]_{i,j} = r_f(i - j)/(NT_s)$  for  $i = 0, 1, \dots, N-1, j = 0, 1, \dots, N-1$ , and  $\mathbf{f}_i$  is the  $i$ -th column of DFT matrix  $\mathbf{F}$ . The second equality follows from (5) while the third equality follows from (6).

Unfortunately  $\mathbf{R}_{yy}$  cannot be written in compact matrix form. Instead, we use the Taylor-series expansion of  $r_t(\Delta t) = r_0 + r_1 \Delta t + r_2 (\Delta t)^2 + \dots$ . Since  $r_t(\Delta t)$  is an even function,  $r_{2k+1} = 0$  for  $k = 0, 1, \dots$ . For Jakes' model, the zeroth-order Bessel function of the first kind can be expanded as

$J_0(2\pi x) \approx 1 - (\pi x)^2$ . Hence,

$$\begin{aligned} r_t [(n_1 - n_2)T_s] &= J_0(2\pi f_d T_s (n_1 - n_2)) \\ &\simeq 1 - (\pi f_d T (n_1 - n_2)/N)^2 \\ &= 1 + r_2 (n_1 - n_2)^2 \end{aligned} \quad (8)$$

where  $r_2 = -(\pi f_d T/N)^2$  and  $f_d T$  is the normalized Doppler frequency. When  $\pi f_d T < 0.1$  or  $f_d T < 0.03$ , the second term can be neglected, i.e.,  $r_t [(n_1 - n_2)T_s] = 1$ . Therefore, the autocorrelation matrix  $\mathbf{R}_{yy}$  of the received signal  $\mathbf{y}$  ( $\mathbf{y} = [y(0), \dots, y(N-1)]$ ) can be written as

$$\mathbf{R}_{yy} = \mathbf{F}^H \mathbf{X}_D \mathbf{R}_f \mathbf{X}_D^H \mathbf{F} + \sigma_n^2 \mathbf{I}_N. \quad (9)$$

The probability density function of  $\mathbf{y}$  conditional on the transmitted data is therefore

$$p(\mathbf{y}|\mathbf{X}_D) = (\pi^N \det(\mathbf{R}_{yy}))^{-1} \exp(-\mathbf{y}^H \mathbf{R}_{yy}^{-1} \mathbf{y}). \quad (10)$$

Provided  $|X_k| = 1$ , the determinant of  $\mathbf{R}_{yy}$  can be expressed as

$$\begin{aligned} \det(\mathbf{R}_{yy}) &= \det(\mathbf{X}_D) \det(\mathbf{R}_f + \sigma_n^2 \mathbf{I}_N) \det(\mathbf{X}_D^H) \\ &= \det(\mathbf{R}_f + \sigma_n^2 \mathbf{I}_N). \end{aligned} \quad (11)$$

The determinant of  $\mathbf{R}_{yy}$  is thus independent of  $\mathbf{X}_D$ , when  $X_k$ 's are from constant modulus constellation. Ignoring terms that are independent of  $\mathbf{X}_D$ , we find that maximizing the log likelihood function is equivalent to minimizing the following cost function

$$\begin{aligned} \hat{\mathbf{X}}_D &= \arg \min_{\mathbf{X}_D} \mathbf{y}^H \mathbf{R}_{yy}^{-1} \mathbf{y} \\ &= \arg \min_{\mathbf{X}_D} \mathbf{y}^H \mathbf{F}^H \mathbf{X}_D (\mathbf{F}_L \mathbf{R}_f \mathbf{F}_L^H + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{X}_D^H \mathbf{F} \mathbf{y}. \end{aligned} \quad (12)$$

Solving (12) is equivalent to solving

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathcal{Q}^N} \mathbf{x}^T \mathbf{Y}_D^H (\mathbf{R}_f + \sigma_n^2 \mathbf{I}_N)^{-1} \mathbf{Y}_D \mathbf{x}^* \quad (13)$$

where  $\mathbf{x} = [X_0, X_1, \dots, X_{N-1}]$ . Eq. (13) gives the MLD1. For  $f_d T < 0.03$ , the cost function (13) remains the same, which means (13) is robust for normalized Doppler frequencies less than 0.03.

When  $f_d T > 0.03$ , the first order Taylor-series expansion is not accurate. In Section IV, we contribute an iterative detector to solve this problem.

### B. MLD 2

This estimator exploits the correlation among post-DFT received samples. The correlation between  $Y_{k_1}$  and  $Y_{k_2}$  in (3) can be written as

$$\begin{aligned} &E\{Y_{k_1} Y_{k_2}^*\} \\ &= E\{X_{k_1} H_{k_1} H_{k_2}^* X_{k_2}^* + X_{k_1} H_{k_1} \alpha_{k_2}^* + \alpha_{k_1} H_{k_2}^* X_{k_2}^* \\ &\quad + \alpha_{k_1} \alpha_{k_2}^* + W_{k_1} W_{k_2}^*\} \end{aligned} \quad (14)$$

From (3), (5) and (6), we have

$$\begin{aligned} &E\{H_{k_1} H_{k_2}^*\} \\ &= \frac{1}{N^2} \sum_{n_1, n_2=0}^{N-1} \sum_{l_1, l_2=0}^{L-1} E\{h_{l_1}(n_1) h_{l_2}^*(n_2)\} e^{-j2\pi k_1 l_1 / N} e^{j2\pi k_2 l_2 / N} \\ &= \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} r_t [(n_1 - n_2)T_s] \sum_{l=0}^{L-1} \sigma_l^2 e^{-j2\pi(k_1 - k_2)l / N} \\ &= \frac{1}{N^2} r_f \left( \frac{k_1 - k_2}{NT_s} \right) \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} r_t [(n_1 - n_2)T_s] \\ &= \sigma_{\text{ICI}}^2 r_f \left( \frac{k_1 - k_2}{NT_s} \right) \end{aligned} \quad (15)$$

where  $\sigma_{\text{ICI}}^2 = 1/N^2 \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} r_t [(n_1 - n_2)T_s]$ . We assume that the information symbols are statistical independent or equivalently  $E\{X_{k_1} X_{k_2}^*\} = \delta(k_1 - k_2)$ , which yields

$$\begin{aligned} E\{\alpha_{k_1} H_{k_2}^* X_{k_2}^*\} &= \frac{1 - \delta(k_1 - k_2)}{N^2} \left( \sum_{l=0}^{L-1} \sigma_l^2 \right) \\ &\quad \times \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} r_t [(n_1 - n_2)T_s] e^{-j2\pi n_1 (k_1 - k_2) / N}. \end{aligned} \quad (16)$$

$E\{X_{k_1} H_{k_1} \alpha_{k_2}^*\} = E\{\alpha_{k_2} H_{k_1}^* X_{k_1}^*\}^*$  can be obtained similarly. The correlation of the ICI terms can be obtained as

$$\begin{aligned} &E\{\alpha_{k_1} \alpha_{k_2}^*\} \\ &= \left( \sum_{l=0}^{L-1} \sigma_l^2 \right) \left\{ \delta(k_1 - k_2) - \frac{1}{N^2} \sum_{n_1=0}^{N-1} \sum_{n_2=0}^{N-1} r_t [(n_1 - n_2)T_s] \right. \\ &\quad \times [e^{-j2\pi n_1 (k_1 - k_2) / N} + (1 - \delta(k_1 - k_2))] \\ &\quad \left. \times e^{-j2\pi n_2 (k_1 - k_2) / N} \right\} \end{aligned} \quad (17)$$

Therefore, the autocorrelation matrix of  $\mathbf{Y}$  can be written as

$$\mathbf{R}_{YY} = \sigma_{\text{ICI}}^2 (\mathbf{X}_D \mathbf{R}_f \mathbf{X}_D^H + \sigma_{en}^2 \mathbf{I}_N) \quad (18)$$

$$\sigma_{en}^2 = \frac{\left( \left( \sum_{l=0}^{L-1} \sigma_l^2 \right) (1 - \sigma_{\text{ICI}}^2) + \sigma_n^2 \right)}{\sigma_{\text{ICI}}^2}. \quad (19)$$

$\sigma_{en}^2$  can be considered as the equivalent noise variance, which incorporates the effect of ICI. For sufficiently large  $N$ ,  $\mathbf{Y}$  can be modelled as Gaussian via the central limit theorem. Hence, similar to (13), MLD2 for  $\mathbf{x}$  is given by

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathcal{Q}^N} \mathbf{x}^T \mathbf{Y}_D^H (\sigma_{\text{ICI}}^2 \mathbf{R}_f + \sigma_{en}^2 \mathbf{I}_N)^{-1} \mathbf{Y}_D \mathbf{x}^*. \quad (20)$$

### C. MLD 3

Eq. (2) can be written in matrix form as

$$\mathbf{y} = \mathcal{X} \mathbf{h} + \mathbf{w} \quad (21)$$

where  $\mathbf{y} = [y(0), y(1), \dots, y(N-1)]^T$ ,  $\mathbf{h} = [\mathbf{h}(0)^T, \mathbf{h}(1)^T, \dots, \mathbf{h}(N-1)^T]^T$  where  $\mathbf{h}(n) = [h_0(n), h_1(n), \dots, h_{L-1}(n)]^T$ ,  $\mathbf{w} = [w(0), w(1), \dots, w(N-$

1)]<sup>T</sup> and  $\mathcal{X}$  is given at the top of the fifth page. The received symbol vector  $\mathbf{y}$  is Gaussian with mean  $\mathcal{X}\mathbf{h}$  and covariance matrix  $\sigma_n^2\mathbf{I}_N$ . The likelihood function for the unknown parameters  $\mathcal{X}$  and  $\mathbf{h}$  is given by

$$\Lambda(\mathbf{y}|\mathcal{X}, \mathbf{h}) = \exp \left\{ -\frac{1}{\sigma_n^2} \|\mathbf{y} - \mathcal{X}\mathbf{h}\|^2 \right\}. \quad (22)$$

We evaluate the marginal likelihood function  $\Lambda(\mathbf{y}|\mathcal{X})$ , which is the average of  $\Lambda(\mathbf{y}|\mathcal{X}, \mathbf{h})$  with respect to  $\mathbf{h}$  and can be expressed as

$$\Lambda(\mathbf{y}|\mathcal{X}, \mathbf{h}) = E_{\mathbf{h}}\{\Lambda(\mathbf{y}|\mathcal{X}, \mathbf{h})\}. \quad (24)$$

Using [10, p. 595, eq. (B-3-20)] and dropping irrelevant factors, (24) becomes

$$\Lambda(\mathbf{y}|\mathbf{x}, \epsilon) = \exp \{ \mathbf{y}^H \mathcal{X} \mathbf{G} \mathcal{X}^H \mathbf{y} \} \quad (25)$$

where  $\mathbf{G} = (\mathcal{X}\mathcal{X}^H + \sigma_n^2\mathbf{R}_h^{-1})^{-1}/\sigma_n^2$  and  $\mathbf{R}_h = E\{\mathbf{h}\mathbf{h}^H\}$  is the  $N \times N$  autocorrelation matrix of  $\mathbf{h}$  and can be computed using (4)-(6). The ML estimate of  $\mathcal{X}$  can be obtained as

$$\hat{\mathcal{X}} = \arg \max_{\mathcal{X}} \mathbf{y}^H \mathcal{X} \mathbf{G} \mathcal{X}^H \mathbf{y}. \quad (26)$$

This gives the MLD3. Since  $\mathbf{G}$  depends on  $\mathcal{X}$  and  $\mathcal{X}$  cannot be separated from  $\mathbf{G}$ , it is difficult to write (26) in quadratic form. The iterative algorithm in Section IV is used to solve (26).

*Remarks:*

- The MLDs (13) (20) and (26) need the knowledge of  $\mathbf{R}_f$ ,  $r_t(\cdot)$  and  $\sigma_n^2$ . Since the channel covariance matrix remains unchanged over a long period, the MLDs need not estimate it every data block. The BER by the channel covariance mismatch is concealed in noise in low SNR, while in high SNR such mismatch dominates the BER. Hence  $\sigma_n^2$  is chosen corresponding to a high SNR value.
- The MLD1 does not contain  $f_d$ , so it is robust to  $r_t(\cdot)$  mismatch when  $f_d T < 0.03$ .

#### IV. LOWER COMPLEXITY DETECTION ALGORITHMS

##### A. V-BLAST detection

The MLDs (13) and (20) can be solved by using V-BLAST detection algorithm [7]. They have the general form as

$$\begin{aligned} \hat{\mathbf{x}} &= \arg \min_{\mathbf{x} \in \mathcal{Q}^N} \mathbf{x}^T \mathbf{G} \mathbf{x}^* \\ &= \arg \min_{\mathbf{x} \in \mathcal{Q}^N} \mathbf{x}^T \mathbf{M}^H \mathbf{M} \mathbf{x}^* \end{aligned} \quad (27)$$

where the positive definite matrix  $\mathbf{G}$  can be Cholesky factored as  $\mathbf{G} = \mathbf{M}^H \mathbf{M}$ . The problem (27) can be reformulated as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathcal{Q}^N} \|\mathbf{M} \mathbf{x}^*\|^2 \quad (28)$$

The cost function in (28) is similar to that in BLAST MIMO systems. Hence, V-BLAST can be adapted to solve (28). We omit the details for brevity. As V-BLAST suffers from error propagation, more accurate detection algorithm needs to be derived. We use SD algorithm [8] and Schnorr-Euchner strategy [11] with our data detectors. Details are omitted for

brevity. When  $f_d T > 0.03$ , (13) is not valid. Neither MLD1 nor MLD3 can be written in quadratic form. Note that even  $f_d T > 0.03$ , the solution quality of (13) and (20) is usually good. We thus choose these solutions as the initial estimate denoted by  $\mathbf{x}^{(0)}$ , where the superscript <sup>(i)</sup> denotes the  $i$ -th iteration. The vector  $\mathbf{x}$  is partitioned into several groups. Let the size of group be denoted by  $S$  and the number of groups is  $M$ . In the  $i$ -th iteration, when it proceeds to the  $m$ -th group  $\mathbf{g}_m = \{X_{(m-1)S+1}, \dots, X_{mS}\}$ , all the other  $M-1$  groups are fixed and  $\mathbf{g}_m$  is chosen as

$$\hat{\mathbf{g}}_m = \arg \max_{\mathbf{g}_m \in \mathcal{Q}^S} f(\mathbf{g}_m | \mathbf{g}_1, \dots, \mathbf{g}_{m-1}, \mathbf{g}_{m+1}, \dots, \mathbf{g}_M) \quad (29)$$

where  $f(\cdot)$  is the corresponding cost function in (26). Then it proceeds to the  $m+1$ -th group and so on. The same process continues until no element in  $\mathbf{x}^{(i)}$  changes in the  $(i+1)$ -th iteration. It is evident when  $M=N$ , this detector reduces to exhaustive search. Since the initial vector is usually good and when  $M=1$  it reduces to the greedy algorithm. We choose  $M$  to be 1 or 2 to ensure good performance.

#### V. SIMULATION RESULTS

We now test the proposed MLDs via computer simulations. We assume the following system specifications:

- Both the data and pilot symbols are chosen from the binary phase-shift keying (BPSK).
- The carrier frequency of the OFDM system is 5GHz and the data rate is 3MHz.
- 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]  $\mu s$  is considered [12]. Each path is a complex Gaussian random process independently generated with the classical Doppler spectrum based on the inverse discrete fourier transform method in [13]. All the paths are rounded to integer and we do not consider leakage. The channel correlation matrix is assumed perfectly known at the receiver.
- The number of subcarriers  $N=32$ . 4 pilots are inserted to solve the phase ambiguity.

We set  $M=1$  in the iterative detector and compare the proposed detectors with an ideal, reference least-squares (LS) detector, which is assumed to have perfect CIR.

Figs. 1-3 compare the BER performance of different MLDs with perfect CIR for Doppler frequency  $f_d T = 0.005, 0.01, 0.1$ . The LS data detector with perfect CIR is used as a benchmark<sup>1</sup>. The MLD ID denotes MLD with iterative detection. When  $f_d T = 0.005$  (Fig. 1), all the proposed MLDs perform well and close to each other. The gap between the MLD1 with SD and the ideal LS detector is less than 0.5 dB at BER= $10^{-3}$  and no error floor is observed. When  $f_d T = 0.01$  (Fig. 2), all our proposed MLDs with SD are comparable with the benchmark in high SNR, which shows that they exploit

<sup>1</sup>The LS data detector is not the optimal detector for fast fading channels, which can be exploited the time diversity.

