Linear Unbiased Channel Estimation and Data Detection in Superimposed OFDM Systems

Malihe Ahmadi and Majid Ghanbarinejad
Department of Electrical and Computer Engineering
University of Alberta
Edmonton, Canada
Email: {ahmadina,madjid}@ece.ualberta.ca

Aryan Saadat Mehr
Department of Electrical Engineering
University of Saskatchewan
Saskatoon, Canada
Email: aryan.saadat@usask.ca

***Abstract*—Reliable channel estimation is necessary for orthogonal frequency-division multiplexing (OFDM) systems employing coherent detection to achieve high data rates. Pilot-aided detection algorithms benefit from rapid convergence, but suffer from non-efficient bandwidth usage. The idea of superimposed data transmission can be enabled for channel estimation without sacrificing the data rate. In this paper, we first derive the best linear unbiased estimator (BLUE) for channel impulse response and then, we propose an iterative joint channel estimator and data detector to improve the performance of both the estimator and the detector. Furthermore, we derive the variance of the proposed estimator and show that the equispace and equipower conditions hold for the superimposed pilots to attain the minimum variance of the estimator. Simulations show that the performance with perfect knowledge of the channel impulse response can be achieved closely by the proposed iterative data detector.**

I. Introduction

Orthogonal frequency-division multiplexing (OFDM) is used for high data rate wireless communication and is considered for the fourth-generation (4G) of mobile wireless systems. Coherent data detection in OFDM systems requires knowledge of the channel impulse response (CIR). Pilot-based channel estimation [1], [2] allocates pilot and data symbols to separate subcarriers, i.e., it dedicates part of the bandwidth explicitly to pilot symbols. Desta *et al.* [3] provide a study of selected pilot-based and subspace-based blind channel estimation schemes and compare their performance and limitations.

The time-varying channel in wireless communications requires frequent insertion of pilot symbols, which introduces significant bandwidth loss especially for fast-time-varying channels. The idea of superimposed pilot tones was first described in [4] for analog communications, and was later extended to digital communication systems in [5]. The main advantage of the superimposed-pilot scheme, in the OFDM context, is that the information symbols can be transmitted over all time-frequency slots, hence saving the bandwidth compared to time-multiplexed-pilot scheme. In other words, none of the subcarriers need to be dedicated completely to the pilots.

In [6], the potential of the superimposed-pilot scheme for channel estimation and tracking was compared to the inserted-pilot scheme. The peak-to-average-power ratio (PAPR) for superimposed OFDM was studied in [7]. The superimposed pilots can be used for channel estimation without sacrificing the data rate [8]. In [9], a two-dimensional Wiener filter performs an initial channel estimation. The estimated channel is, then, used iteratively for data detection and improvement of the channel estimation. Maximum likelihood (ML) and minimum mean square error (MMSE) iterative channel estimators were derived in [10]. The paper proposes two data detectors by eliminating the CIR variable from the likelihood function.

In this paper, we derive the best linear unbiased estimator (BLUE) for channel estimation. Then, we use this estimation as an initial CIR to detect data. With the knowledge of the detected data (obtained through a decision-directed algorithm), we perform iterations to improve the channel estimation and the data detection. We also derive the estimation variance for the proposed channel estimator and optimize the pilot placement to minimize the variance.

Notations: $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$ denote transpose and conjugate transpose, respectively. A signal constellation is denoted by $\mathcal{Q}$. $\mathsf{E}[\cdot]$ denotes mean and $\mathsf{Var}(\cdot)$ denotes variance. $\mathsf{tr}\{\cdot\}$ denotes the trace of a matrix. Bold letters are used for vector and matrices. $\mathbf{I}_N$ is the $N \times N$ identity matrix. The $N \times N$ discrete Fourier transform (DFT) matrix is

$$\mathbf{F} = \frac{1}{\sqrt{N}}[e^{-j\frac{2\pi}{N}kl}], \quad k,l \in 0,1,\cdots,N-1, \quad j=\sqrt{-1} \ .$$

$\mathbf{A}_D$ denotes a diagonal matrix whose diagonal entries are the components of vector $\mathbf{A}$. $\mathcal{CN}(\mu,\sigma^2)$ identifies a complex random variable with mean μ and variance σ^2. A vector whose elements are the diagonal elements of matrix $\mathbf{A}$ is denoted by $\mathsf{diag}(\mathbf{A})$. $\mathbb{C}^L$ denotes the set of L-dimensional complex vectors. $\|\cdot\|$ denotes the Euclidean norm.

The rest of the paper is as follows. Section II describes the system model of the superimposed OFDM system. An iterative channel estimation and data detection is proposed in Section III. Pilot optimization is, then, studied in Section IV. Section V provides simulation results to demonstrate the performance of the proposed scheme. Finally, Section VI concludes the paper.

II. Superimposed OFDM System Model

A baseband model of a discrete-time OFDM system over a frequency-selective channel is considered. Data are drawn

978-1-4673-1881-5/12/$31.00 ©2012 IEEE

from a constellation $\mathcal{Q}$. We follow the same notation as in [10]: a transmitted symbol X_k at the kth subcarrier is a linear combination of a pilot symbol and a data symbol

$$X_k = \sqrt{\varphi_k} S_k + \sqrt{\phi_k} P_k, \quad k = 0, 1, \cdots, N-1, \qquad (1)$$

where $P_k \in \mathcal{Q}$ is a known pilot, $S_k \in \mathcal{Q}$ is a zero-mean randomly-distributed data symbol, and both P_k and S_k have unit average power [10]. The coefficients ϕ_k and φ_k are the power of the pilot and data symbols, respectively, and N is the total number of subcarriers. In the OFDM system, a block of N transmit symbols are modulated by the inverse DFT (IDFT). Thus, the resulting time-domain signal samples after the IDFT can be expressed by

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

A guard interval is then inserted to combat inter-symbol interference.

The transmit and receive pulse shaping as well as the physical channel can be considered as a composite CIR

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

where $h_l \sim \mathcal{CN}(0, \sigma_l^2)$ is the amplitude and τ_l is the delay of the lth tap. It is assumed that $\tau_l = lT_s$ so that the channel has a finite impulse response with an effective length L. Assuming perfect synchronization, the samples of the received signal can be represented as

$$y_n = \sum_{l=0}^{L-1} h_l x_{n-l} + w_n,$$

where $w_n \sim \mathcal{CN}(0, \sigma_w^2)$ is the additive white Gaussian noise. After removing the guard interval from the samples of the received signal and performing the DFT demodulation, we have

$$\begin{aligned} Y_k &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} y_n e^{-j(2\pi kn/N)} \\ &= H_k X_k + W_k, \quad 0 \le k \le N-1, \end{aligned} \qquad (2)$$

with

$$H_k = \sum_{l=0}^{L-1} h_l e^{-j2\pi kl/N}, \quad 0 \le k \le N-1, \qquad (3)$$

$$W_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} w_n e^{-j2\pi kn/N}, \quad 0 \le k \le N-1, \qquad (4)$$

where H_k is the channel frequency response on the kth subcarrier and W_k is the noise on the kth subcarrier. Since $\mathbf{F}$ is a unitary matrix, we have $W_k \sim \mathcal{CN}(0, \sigma_w^2)$.

Defining $\mathbf{H} = [H_0, H_1, \cdots, H_{N-1}]^\mathsf{T} = \mathbf{F}_L \mathbf{h}$, where $\mathbf{h} = [h_0, h_1, \cdots, h_{L-1}]^\mathsf{T} \in \mathbb{C}^L$ is the CIR and $\mathbf{F}_L$ is the $N \times L$ submatrix of $\mathbf{F}$, (2) can be vectorized as

$$\mathbf{Y} = \mathbf{X}_D \mathbf{F}_L \mathbf{h} + \mathbf{W} = (\mathbf{\Phi} \mathbf{P}_D + \mathbf{\Psi} \mathbf{S}_D) \mathbf{F}_L \mathbf{h} + \mathbf{W}, \qquad (5)$$

where

$$\begin{aligned} \mathbf{X} &= [X_0, X_1, \cdots, X_{N-1}]^\mathsf{T}, \\ \mathbf{Y} &= [Y_0, Y_1, \cdots, Y_{N-1}]^\mathsf{T}, \\ \mathbf{P} &= [P_0, P_1, \cdots, P_{N-1}]^\mathsf{T}, \\ \mathbf{S} &= [S_0, S_1, \cdots, S_{N-1}]^\mathsf{T}, \\ \mathbf{W} &= [W_0, W_1, \cdots, W_{N-1}]^\mathsf{T}, \\ \mathbf{\Phi} &= \mathsf{diag}\left([\sqrt{\phi_0}, \sqrt{\phi_1}, \cdots, \sqrt{\phi_{N-1}}]\right), \\ \mathbf{\Psi} &= \mathsf{diag}\left([\sqrt{\varphi_0}, \sqrt{\varphi_1}, \cdots, \sqrt{\varphi_{N-1}}]\right) \ . \end{aligned}$$

III. Channel Estimator and Iterative Data Detector

A. Channel Estimator

Let $S_k, \ k = 0, \cdots, N-1$ be zero-mean randomly-distributed data symbols, and let the noise $\mathbf{W}$ and the CIR $\mathbf{h}$ be vectors of zero-mean complex Gaussian random variables (CGRV). The Gauss-Markov Theorem [11, Theorem 6.1] states that, since $\mathbf{Y}$ is of the general linear form

$$\mathbf{Y} = \mathbf{\Phi} \mathbf{P}_D \mathbf{F}_L \mathbf{h} + \mathbf{\Psi} \mathbf{S}_D \mathbf{F}_L \mathbf{h} + \mathbf{W} = \mathbf{\Phi} \mathbf{P}_D \mathbf{F}_L \mathbf{h} + \mathbf{V}, \qquad (6)$$

where $\mathbf{V}$ is a zero-mean complex vector with covariance matrix $\mathbf{C}$, the BLUE (or identically, the weighted least squares estimator) of $\mathbf{h}$ is given by

$$\hat{\mathbf{h}} = ((\mathbf{\Phi} \mathbf{P}_D \mathbf{F}_L)^\mathsf{H} \mathbf{C}^{-1} \mathbf{\Phi} \mathbf{P}_D \mathbf{F}_L)^{-1} (\mathbf{\Phi} \mathbf{P}_D \mathbf{F}_L)^\mathsf{H} \mathbf{C}^{-1} \mathbf{Y} \ . \qquad (7)$$

Defining $\alpha = \sum_{l=0}^{L-1} \sigma_l^2 / N$, we have

$$\begin{aligned} \mathbf{C} &= \mathsf{E}\left[(\mathbf{\Psi} \mathbf{S}_D \mathbf{F}_L \mathbf{h} + \mathbf{W})(\mathbf{\Psi} \mathbf{S}_D \mathbf{F}_L \mathbf{h} + \mathbf{W})^\mathsf{H}\right] \\ &= \alpha \mathbf{\Psi}^2 + \sigma_w^2 \mathbf{I}_N, \end{aligned} \qquad (8)$$

where expectation is over both the CIR and the information symbols. In (8), it is assumed that information symbols on different subcarriers are independent, unit-average random variables. The minimum variance of $\hat{\mathbf{h}}_i$ is

$$\mathsf{Var}\left(\hat{\mathbf{h}}_i\right) = [((\mathbf{\Phi} \mathbf{P}_D \mathbf{F}_L)^\mathsf{H} \mathbf{C}^{-1} \mathbf{\Phi} \mathbf{P}_D \mathbf{F}_L)^{-1}]_{ii} \ . \qquad (9)$$

B. Iterative Data Detector

Now, we use the estimated channel to form a decision-directed joint channel estimator and data detector (JCEDD). The probability distribution function of the received signal $\mathbf{Y}$ conditioned on $\mathbf{h}$ and $\mathbf{S}$ is obtained by

$$f(\mathbf{Y}|\mathbf{h}, \mathbf{S}) = \frac{1}{(\pi \sigma_w)^N} \exp\{-\frac{1}{\sigma_w^2} \|\mathbf{Y} - \mathbf{X}_D \mathbf{F}_L \mathbf{h}\|^2\} \ .$$

Thus, the ML JCEDD is described by

$$\{\hat{\mathbf{S}}, \hat{\mathbf{h}}\} = \underset{\mathbf{S} \in \mathcal{Q}^N, \mathbf{h} \in \mathbb{C}^L}{\arg\min} \|\mathbf{Y} - (\mathbf{\Phi} \mathbf{P}_D + \mathbf{\Psi} \mathbf{S}_D) \mathbf{F}_L \mathbf{h}\|^2 \ . \qquad (10)$$

The output of the channel estimator proposed in the previous subsection can be used in the first iteration of an iterative decision-directed data detector. Starting from (7) as $\hat{\mathbf{h}}^0$, the data detection problem can be solved by one of the following approaches:

- In the ith iteration, the vector $\mathbf{S}$ of data symbols can be estimated as [10]

$$\hat{\mathbf{S}}^i = \mathcal{M}_{\mathcal{Q}}(\mathbf{\Psi}^{-1}[(\mathbf{H}_D^i)^{-1}\mathbf{Y} - \mathbf{\Phi}\mathbf{P}]), \tag{11}$$

where $\mathbf{H}_D^i = \mathsf{diag}\left(\mathbf{F}_L\hat{\mathbf{h}}^i\right)$ and $\mathcal{M}_{\mathcal{Q}}(.)$ maps its argument to the nearest element in $\mathcal{Q}$.
- The data detection problem is the so-called integer least squares (LS) problem

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{Q}^N}{\arg\min}\ \|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2, \tag{12}$$

where $\mathbf{r} = \mathbf{Y} - \mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L\hat{\mathbf{h}}^i$, $\mathbf{H} = \mathbf{H}_D^i\mathbf{\Psi}$ and $\mathbf{x} = \mathbf{S}$. The VBLAST detection algorithm [12], which is a sequence of nulling and interference cancelation in order to achieve the maximum post detection SNR, can be performed for data detection.
- Another important algorithm for solving (12) is sphere decoding (SD) [13], which only tests the lattice point lying inside a hypersphere of radius d, i.e.,

$$\|\mathbf{r} - \mathbf{H}\mathbf{x}\|^2 \leq d^2\ . \tag{13}$$

SD can offer ML decoding at lower complexity, as opposed to the exponential complexity incurred by the exhaustive search. The expected complexity of SD is shown to be polynomial for a wide range of SNR and lattice dimensions [14].

Given $\hat{\mathbf{S}}^i$, the LS solution to estimating CIR in (10) is

$$\hat{\mathbf{h}}^i = ((\mathbf{X}_D^i\mathbf{F}_L)^{\mathsf{H}}\mathbf{X}_D^i\mathbf{F}_L)^{-1}(\mathbf{X}_D^i\mathbf{F}_L)^{\mathsf{H}}\mathbf{Y}, \tag{14}$$

where $\mathbf{X}_D^i = \mathbf{\Phi}\mathbf{P}_D + \mathbf{\Psi}\hat{\mathbf{S}}_D^i$.

IV. Pilot Distribution Optimization

The place and power distribution of pilot and data symbols can be optimized to minimize the variance of the CIR estimator given in (15) subject to the power constraint $\sum_{k=0}^{N-1}\varphi_k = D$ and $\sum_{k\in I_p}\phi_k = P$, where I_p is the index of N_p subcarriers with superimposed pilots, and P and D are the total power on pilots and data symbols, respectively. Starting from (9), we obtain

$$\sigma^2_{\hat{\mathbf{h}}^0} = \mathsf{tr}\left\{\left((\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L)^{\mathsf{H}}(\alpha\mathbf{\Psi}^2 + \sigma_w^2\mathbf{I}_N)^{-1}\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L\right)^{-1}\right\}\ . \tag{15}$$

From [15], for an $L\times L$ positive definite matrix $\mathbf{A}$, we know $\mathsf{tr}\left\{\mathbf{A}\right\}\mathsf{tr}\left\{\mathbf{A}^{-1}\right\} \geq L^2$, where the equality holds if and only if $\mathbf{A} = a\mathbf{I}_L,\ a\neq 0$. Since

$$\mathbf{A} = (\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L)^{\mathsf{H}}(\alpha\mathbf{\Psi}^2 + \sigma_w^2\mathbf{I}_N)^{-1}\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L$$

is positive definite (assuming $N_p \geq L$), we have

$$\sigma^2_{\hat{\mathbf{h}}^0} \geq \frac{L^2}{\mathsf{tr}\left\{(\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L)^{\mathsf{H}}(\alpha\mathbf{\Psi}^2 + \sigma_w^2\mathbf{I}_N)^{-1}\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L\right\}}, \tag{16}$$

where the equality holds if and only if

$$(\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L)^{\mathsf{H}}(\alpha\mathbf{\Psi}^2 + \sigma_w^2\mathbf{I}_N)^{-1}\mathbf{\Phi}\mathbf{P}_D\mathbf{F}_L = a\mathbf{I}_L\ .$$

The (r,s)th $(0\leq r,s\leq L-1)$ entry of $\mathbf{A}$ can be written as

$$[\mathbf{A}]_{r,s} = \sum_{k\in I_p} e^{j\frac{2\pi(r-s)k}{N}}\frac{\phi_k}{\alpha\varphi_k + \sigma_w^2}\ .$$

To satisfy $\mathbf{A} = a\mathbf{I}_L$, it is required that

$$\sum_{k\in I_p} e^{j\frac{2\pi(r-s)k}{N}}\frac{\phi_k}{\alpha\varphi_k + \sigma_w^2} = 0, \quad r\neq s\ . \tag{17}$$

Equation (17) is satisfied if the following conditions are satisfied:

$$I_p = \{z_0 + k'z,\ k' = 0,1,\cdots,N_p-1\}, \quad z = N/N_p \in \mathbb{Z},$$
$$z_0 \in \{0,1,\cdots,z-1\}, \quad N_p \geq L, \tag{18}$$

$$\frac{\phi_k}{\alpha\varphi_k + \sigma_w^2} = \mathsf{cons.}, \quad k\in I_p\ . \tag{19}$$

Condition (19) holds when the pilots are equipowered ($\phi_k = \phi = P/N_p,\ k\in I_p$) and the power of the data symbols on the subcarriers with superimposed pilots are identical as well ($\varphi_k = \varphi,\ k\in I_p$). Condition (18) means that at least L superimposed pilots are required and they need to be equispaced. When conditions (18) and (19) are satisfied, we can write $\mathbf{A} = P/(\alpha\varphi + \sigma_w^2)\mathbf{I}_L$, and therefore $\sigma^2_{\hat{\mathbf{h}}^0} \geq L(\alpha\varphi + \sigma_w^2)/P$. One can pick the values of P and φ based on this lower bound on the variance of the initial channel estimation. Note that if $\varphi = 0$, the initial channel estimation with the minimum variance of $L\sigma_w^2/P$ is achievable, which means the proposed initial channel estimator achieves the lowest variance for non-superimposed OFDM (i.e., when data and pilot symbols are separated).

The Cramér-Rao bound (CRB) of the LS channel estimator in (14) is given by [10]

$$\mathsf{CRB}_{\mathbf{h}} = \left(\frac{1}{\sigma_w^2}\mathbf{F}_L^{\mathsf{H}}(\mathbf{\Psi}^2 + \mathbf{\Phi}^2)\mathbf{F}_L + \mathbf{R}_{\mathbf{h}}^{-1}\right)^{-1}, \tag{20}$$

where $\mathbf{R}_{\mathbf{h}} = \mathsf{E}\left[\mathbf{h}\mathbf{h}^{\mathsf{H}}\right]$.

V. Simulation Results

In our simulations, a superimposed OFDM system with 64 subcarriers is developed. The channel model considered for the simulations is a Rayleigh fading 5-tap channel with power profile $\sigma_l^2 = \sigma_0^2 e^{-l/5}$. Each path is an independently generated CGRV which is constant over one block of OFDM symbols, but can vary from one block to another. Both information and pilot symbols are drawn from a BPSK constellation. The implemented OFDM system has 8 equispaced superimposed pilots.

In Fig. 1, we compare the performance of the channel estimator (represented by BLUE) with the ML estimator given in [10], where

$$\phi_k = 0.8,\ \varphi_k = 0.2,\ k\in I_p,\ \varphi_k = 1,\ k\in \bar{I}_p\ .$$

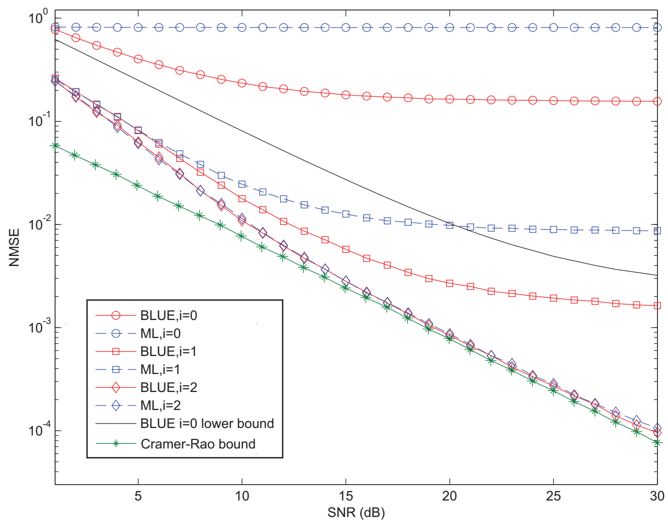

Fig. 1. NMSE of the BLUE CIR estimator vs. that of the ML CIR estimator.

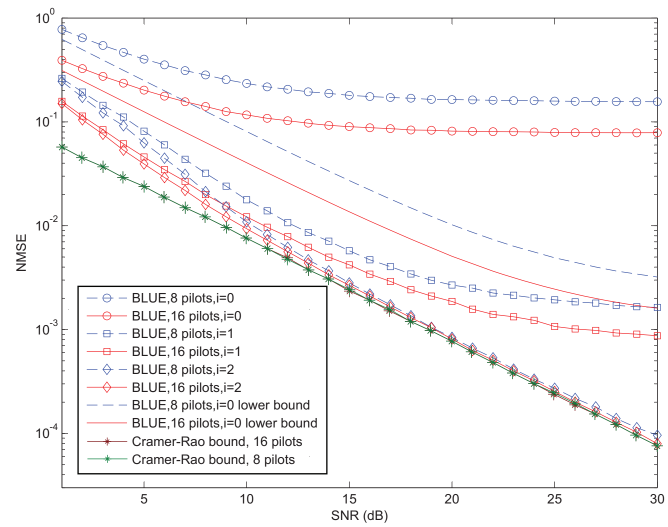

Fig. 3. NMSE of the BLUE CIR estimator, 8 vs. 16 superimposed pilots.

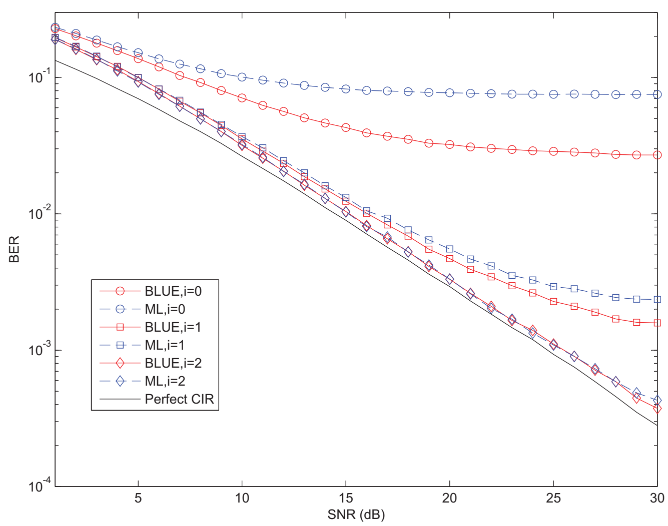

Fig. 2. BER of the iterative data detector in III-B vs. that of the data detector in [10, Section III-A].

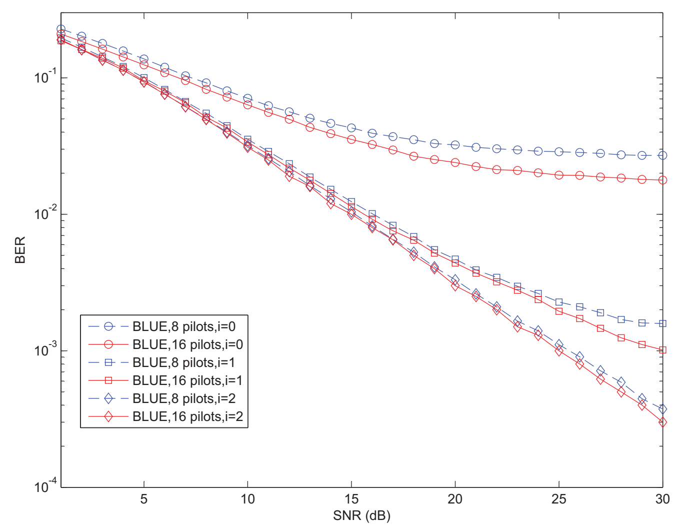

Fig. 4. BER of the iterative data detector in III-B, 8 vs. 16 superimposed pilots.

i denotes the number of iterations. This figure shows the normalized mean square error (NMSE) of the channel estimation, which is defined as

$$\frac{\sum_{i=1}^{M} \|\hat{\mathbf{h}}_{\mathbf{i}} - \mathbf{h}\|^2}{\sum_{i=1}^{M} \|\mathbf{h}\|^2},$$

where M is the number of runs of the Monte Carlo simulation. It also includes the normalized lower bound of the mean square error (MSE) of the initial estimation (16) as well as the CRB of the LS estimator given by (20). As can be seen, in high SNR, the first iteration improves the NMSE of the estimator significantly (almost 20dB), while the second iteration provides about 10dB improvement. The proposed estimator performs better than the ML estimator at the cost of higher complexity.

Fig. 2 shows the bit error rate (BER) performance of the first data detector explained in Section III-B for different numbers of iterations. It also compares the performance of the proposed JCEDD with that of the ML JCEDD presented in [10] with

$$\phi_k = 0.8,\ \varphi_k = 0.2,\ k \in I_p,\ \varphi_k = 1,\ k \in \bar{I}_p\ .$$

The performance of the data detector with perfect CIR knowledge provides a benchmark. As the figure shows, in high SNR, BER of the data detector with zero and one iteration reaches an error floor while the detector with perfect CIR knowledge performs only about 1dB better than our data detector with two (or more) iterations.

Figs. 3 and 4 show the effect of the number of pilots (comparing 8 versus 16 pilots) on the performance of the JCEDD, where

$$\phi_k = 0.8,\ \varphi_k = 0.2,\ k \in I_p,\ \varphi_k = 1,\ k \in \bar{I}_p\ .$$

As can be seen, in high SNR and with more number of

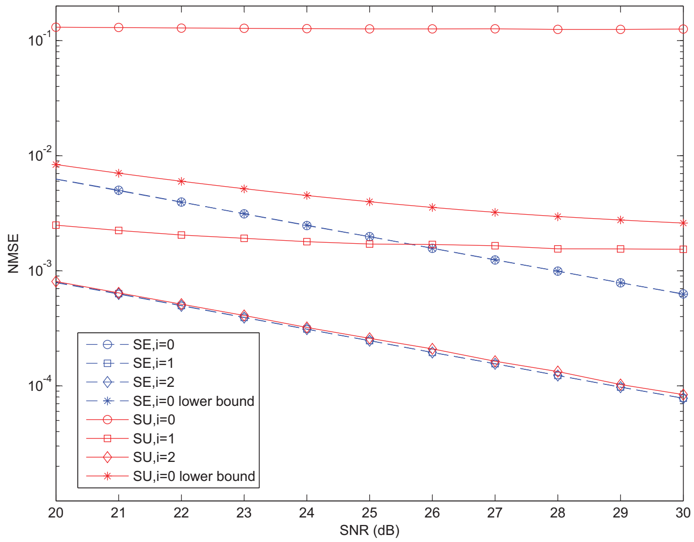

Fig. 5. NMSE of the BLUE CIR estimator, separated (SE) vs. superimposed (SU) pilots.

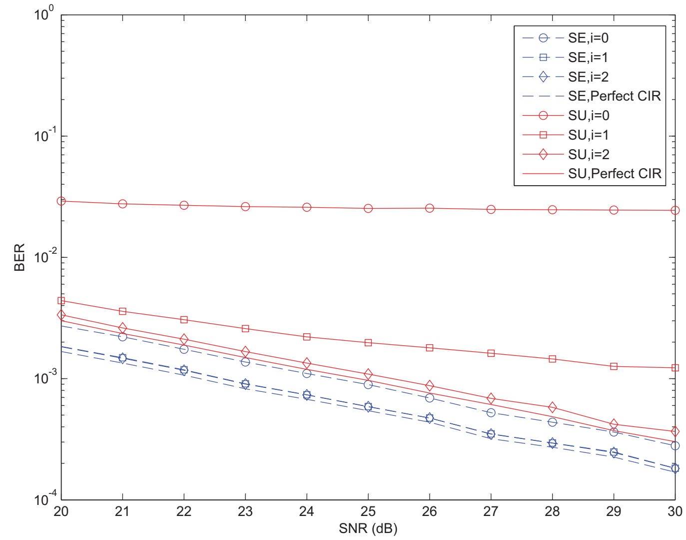

Fig. 6. BER of the iterative data detector, separated (SE) vs. superimposed (SU) pilots.

iterations, the improvement made by including more pilots is insignificant.

Figs. 5 and 6 compare the performance of 8 superimposed pilots with that of 8 separated pilots. They are respectively denoted as SU and SE, and the total power of pilots is the same in both ($\phi_k = 1,\ k \in I_p,\ \varphi_k = 1,\ k \in \bar{I}_p$; $\varphi_k = 0.2,\ k \in I_p$ for the superimposed case). As Fig. 5 shows, with at least two iterations, the difference in the NMSE is insignificant while SU has higher bandwidth efficiency. Also, as mentioned in Section IV, the initial channel estimator for SE reaches the MSE lower bound. Fig. 6 demonstrates that, although SU with no iteration has an error floor, with two iterations and in high SNR, SE performs only about 2dB better than SU.

VI. Conclusion

We proposed a linear, unbiased channel estimator and iterative data detector for superimposed OFDM systems. The channel estimator requires only the value of the SNR, pilots, the power distribution of data symbols on superimposed subcarriers and the CIR total power to be known at the receiver. To improve the channel estimation, we used a decision-directed algorithm. We derived the variance of the channel estimator and optimized the allocation and the power distribution of the superimposed pilots to minimize the MSE of the estimator. In our simulations, we investigated the effect of the number of iterations on the performance of the the proposed JCEDD. We also showed that the proposed JCEDD performs better than the ML JCEDD described in [10].

References

[1] 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.

[2] M. Dong and L. Tong, “Optimal design and placement of pilot symbols for channel estimation,” *IEEE Trans. Signal Process.*, vol. 50, no. 12, pp. 3055–3069, Dec. 2002.

[3] Y. T. Desta, J. Tao, and W. Zhang, “Review on selected channel estimation algorithms for orthogonal frequency division multiplexing system,” *Info. Technology Jnl.*, vol. 10, no. 5, pp. 914–926, 2011.

[4] C. E. Kastenholz and W. P. Birkemeier, “A simultaneous information transfer and channel sounding modulation technique for wide-band channels,” *IEEE Trans. Comm. Technol.*, vol. 13, no. 2, pp. 162–165, 1965.

[5] B. Farhang-Boroujeny, “Pilot-based channel identification: proposal for semi-blind identification of communication channels,” *IEEE Elect. Letters*, vol. 31, no. 13, pp. 1044–1046, 1995.

[6] S. Balasubramanian, B. Farhang-Boroujeny, and V. J. Mathews, “Pilot embedding for channel estimation and tracking in ofdm systems,” in *Proc. of IEEE Globecom*, vol. 2, Dallas, TX, Nov. 2004, pp. 1244–1248.

[7] N. Chen and G. T. Zhou, “What is the price paid for superimposed training in ofdm?” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, Montreal, CA, May 2004, pp. iv–421–iv–424.

[8] ——, “A superimposed periodic pilot scheme for semi-blind channel estimation of ofdm systems,” in *Proc. of IEEE 10th Digital Signal Processing Workshop*, Pine Mountain, GA, Oct. 2002, pp. 362–365.

[9] C. K. Ho, B. Farhang-Boroujeny, and F. Chin, “Added pilot semi-blind channel estimation scheme for ofdm in fading channels,” in *Proc. of IEEE Globecom*, vol. 5, San Antonio, TX, Nov. 2001, pp. 3075–3079.

[10] T. Cui and C. Tellambura, “Superimposed pilot symbols for channel estimation in ofdm systems,” in *Proc. of IEEE Globecom*, vol. 4, St. Louis, MO, Nov. 2005, pp. 2229–2233.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[12] G. Golden, C. Foschini, R. Valenzuela, and P. Wolniansky, “Detection algorithm and initial laboratory results using v-blast space-time communication architecture,” *Electronics Letters*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[13] E. Viterbo and J. Boutros, “A universal lattice code decoder for fading channels,” *IEEE Trans. Info. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

[14] B. Hassibi and H. Vikalo, “On the sphere-decoding algorithm i. expected complexity,” *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[15] S. Ohno and G. B. Giannakis, “Optimal training and redundant precoding for block transmissions with application to wireless ofdm,” *IEEE Trans. Comm.*, vol. 50, no. 12, pp. 2113–2123, Dec. 2002.