SUPERIMPOSED TRAINING FOR CHANNEL SHORTENING EQUALIZATION IN OFDM

Xiaoli Ma[1], *Robert J. Baxley*[1], *John Kleider*[2], and *G. Tong Zhou*[1]

[1]School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332
[2]General Dynamics, C4 Systems, Scottsdale, AZ 85257

ABSTRACT

In orthogonal frequency division multiplexing (OFDM) systems, the cyclic prefix (CP) is required to be greater than the length of the channel impulse response to avoid inter-block interference. However, a long CP decreases power and bandwidth efficiencies. Recently channel shortening equalizers (CSEs) have been introduced to enable the use of a shorter CP, but they either require perfect channel knowledge, or demand high complexity and long decoding delay when the channel is unknown. In this paper, we propose a low complexity CSE for OFDM with unknown channels by using superimposed training. Our unique design achieves bandwidth efficient channel estimation, and low complexity channel shortening equalization as well as symbol decoding. Simulation results demonstrate the effectiveness of the proposed scheme.

I. INTRODUCTION

Modern wireless communications ask for high data rates with wideband signals. When the symbol period is comparable or shorter than the delay spread, the underlying wireless channel becomes frequency-selective. Orthogonal frequency division multiplexing (OFDM) has been widely adopted for wireless systems [3], because of its low complexity equalization and decoding over frequency-selective channels. For OFDM and other multi-carrier systems, a cyclic prefix (CP) no shorter than the channel order has to be added in front of each OFDM block to avoid the inter-block interference (IBI). When the channel impulse response is long, a lengthy CP results in reduced bandwidth and power efficiencies. When the CP is shorter than the channel order, IBI degrades performance dramatically. In this case, the performance degradation may be alleviated by parallel interference cancellation (PIC) or serial interference cancellation (SIC) (see e.g., [11]) when the channel is known. However, PIC and SIC sacrifice the low complexity of OFDM decoding. Furthermore, either blind or data-aided channel estimation costs extra computational complexity or bandwidth. This motivates us to develop a low complexity method which avoids insertion of a long CP when the channel is unknown.

Recently, channel shortening equalizers (CSEs) have been introduced to reduce the order of the equivalent channel [1, 8–10, 12, 14, 17], and thus a shorter CP than the original channel order may be inserted in OFDM symbol blocks. However, most CSE designs either require perfect channel knowledge at the receiver [9, 17], or demand high computational complexity and long decoding delay when the channel is unknown, e.g., blind or non-data-aided CSE designs in [1, 8, 10, 12, 14].

In this paper, we develop a low complexity CSE for OFDM systems with unknown channels by using superimposed training. The CSE is a finite impulse response (FIR) filter that is designed based on the channel estimated with a superimposed training sequence. Our unique scheme achieves bandwidth efficient channel estimation, and low complexity channel shortening equalization as well as symbol decoding. Superimposed training has been proposed for frequency-selective channel estimation [15, 18], and it has also been shown to have the potential of reducing the peak-to-average power ratio (PAR) of the transmitted signal [2, 4, 5]. The major advantage of superimposed training relative to preamble based training

This material is based on work supported by the U.S. Army Research Laboratory and the U.S. Army Research Office under grant number W911NF-06-1-0090 and through collaborative participation in the Collaborative Technology Alliance for Communications & Networks sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAD19-01-2-0011. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. This work was also supported in part by the US National Science Foundation Graduate Research Fellowship Program and Grant No. 0218778.

or optimal training in [13] is that it does not sacrifice the transmission rate, and thus is bandwidth efficient. Furthermore, compared to blind methods, superimposed training enables lower complexity channel estimation and achieves better performance.

The rest of the paper is organized as follows. In Section II, we introduce the system model of OFDM with superimposed training. Section III describes the CSE design by using superimposed training. Section IV presents the low complexity symbol decoding. Simulation results are shown in Section V, and Section VI concludes the paper.

Notation: Upper (lower) bold face letters indicate matrices (column vectors). Superscript $(\cdot)^{\mathcal{H}}$ denotes Hermitian, $(\cdot)^T$ transpose, and $(\cdot)^*$ conjugate. $E[\cdot]$ stands for expectation; diag$[\boldsymbol{x}]$ for a diagonal matrix with $\boldsymbol{x}$ on its main diagonal; $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix; $\|\cdot\|$ denotes the Frobenius norm; and $\boldsymbol{F}_N$ is the normalized $N \times N$ fast Fourier transform (FFT) matrix with its (m, n)th element being $N^{-1/2}\exp(-j2\pi mn/N)$.

II. SYSTEM MODEL

Consider single-user OFDM transmissions with N_c subcarriers over an L_hth-order frequency-selective fading channel ($N_c > L_h$), and assume that $L_h + 1$ is much greater than the desired CP length L_{cp}. The discrete-time equivalent impulse response vector of the channel is $\boldsymbol{h} = [h_0, ..., h_{L_h}]^T$. Let the transmitted signal be $x(n)$. Then, the received signal $y(n)$ can be written as

$$y(n) = \sum_{\ell=0}^{L_h} h_\ell x(n-\ell) + \eta(n), \tag{1}$$

where $\eta(n)$ is additive white Gaussian noise (AWGN) with zero mean and variance N_0. Suppose that we collect $P = N_c + L_{cp}$ symbols as one block at the receiver, then the matrix-vector counterpart of (1) becomes

$$\boldsymbol{y}_k = \boldsymbol{H}_0 \boldsymbol{x}_k + \boldsymbol{H}_1 \boldsymbol{x}_{k-1} + \boldsymbol{\eta}_k, \tag{2}$$

where $\boldsymbol{x}_k = [x(kP), \ldots, x((k+1)P-1)]^T$, $\boldsymbol{y}_k$ and $\boldsymbol{\eta}_k$ are defined similarly, $\boldsymbol{H}_0$ is a $P \times P$ Toeplitz matrix with first column $[\boldsymbol{h}^T\ \boldsymbol{0}^T_{P-L_h-1}]^T$, and $\boldsymbol{H}_1$ is a $P \times P$ Toeplitz matrix with the first row $[\boldsymbol{0}^T_{P-L_h}\ h_{L_h}, \ldots, h_1]$ (see e.g., [16] for details).

The transmitted signal $x(n)$ consists of the information part which is unknown at the receiver and the superimposed training part which is known at both transmitter and receiver sides; i.e.,

$$x(n) = u(n) + c(n), \tag{3}$$

where $u(n)$ is the information symbol and $c(n)$ is the superimposed training symbol. It is straightforward to show that the kth transmitted block can be represented as

$$\boldsymbol{x}_k = \boldsymbol{u}_k + \boldsymbol{c}_k. \tag{4}$$

The information-containing block $\boldsymbol{u}_k$ is generated from information block $\boldsymbol{s}_k$ with each entry drawn from a finite alphabet by performing inverse fast Fourier transform (IFFT) and CP insertion. Mathematically, $\boldsymbol{u}_k = \boldsymbol{T}_{cp}\boldsymbol{F}^{\mathcal{H}}\boldsymbol{s}_k$, where $\boldsymbol{T}_{cp} = [\boldsymbol{I}^T_{cp}, \boldsymbol{I}_{N_c}]^T$ is the CP insertion matrix with $\boldsymbol{I}_{cp}$ denoting the last L_{cp} rows of $\boldsymbol{I}_{N_c}$, and $\boldsymbol{F}^{\mathcal{H}}$ is IFFT matrix. The baseband equivalent system block diagram is given in Fig. 1.

Here our goal is to estimate the information symbols $\boldsymbol{s}_k$ with low complexity from the observations $\boldsymbol{y}_k$'s using the knowledge of the superimposed training sequence $\boldsymbol{c}_k$ and without knowing the channel taps.

III. THE CSE DESIGN

To introduce the CSE design, we need the following assumptions:

as1) The symbols are drawn independently from a finite alphabet with zero mean; i.e., $E[\boldsymbol{s}_k] = \boldsymbol{0}$;

as2) Noise is independent from information symbols and also has zero mean; i.e., $E[\boldsymbol{\eta}_k] = \boldsymbol{0}$.

From (3), we observe that unlike preamble training or optimal training in [13], superimposed training symbols are coupled with information symbols. To filter out this known information, we need to use statistical methods. Here we adopt the first-order method for channel estimation which has also been used in [15, 18] and then design a low-complexity CSE.

A. Channel Estimation

Based on *as1)* and *as2)*, the mean of the received block is given as [c.f. (2)]

$$\bar{\boldsymbol{y}} = E[\boldsymbol{y}_k] = \boldsymbol{H}_0 E[\boldsymbol{c}_k] + \boldsymbol{H}_1 E[\boldsymbol{c}_{k-1}]. \tag{5}$$

We further assume that the training sequence satisfies the following condition:

as3) If the training sequence $\{c(n)\}$ is random, then it is wide sense stationary (WSS) with nonzero

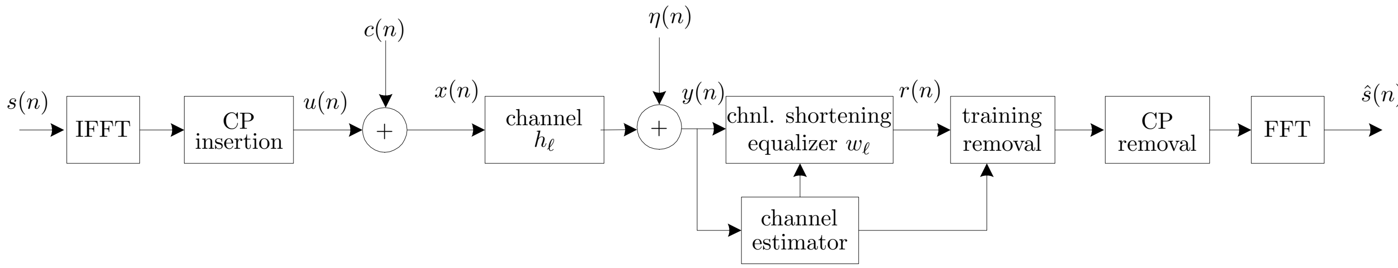

Fig. 1. Discrete-time baseband equivalent system model

mean; if the training sequence $\{c(n)\}$ is deterministic, then it has a constant mean for a known period.

With *as3)*, $E[\boldsymbol{c}_k] = E[\boldsymbol{c}_{k-1}] = \bar{\boldsymbol{c}}$, or $\frac{1}{N_b}\sum_{k=1}^{N_b} \boldsymbol{c}_k = \bar{\boldsymbol{c}}$. In this case, the mean of the received block in (5) can be represented as

$$\bar{\boldsymbol{y}} = \tilde{\boldsymbol{H}}\bar{\boldsymbol{c}}, \tag{6}$$

where $\tilde{\boldsymbol{H}} = \boldsymbol{H}_0 + \boldsymbol{H}_1$ is a circulant matrix generated by $[\boldsymbol{h}^T, \boldsymbol{0}^T_{P-L_h-1}]^T$. Given the model in (6), using the property that $\tilde{\boldsymbol{H}}\bar{\boldsymbol{c}} = \bar{\boldsymbol{C}}\boldsymbol{h}$ where $\bar{\boldsymbol{C}}$ is the first $L_h + 1$ columns of the circulant matrix $\tilde{\boldsymbol{C}}$ generated by $\bar{\boldsymbol{c}}$, the least-squares (LS) channel estimator is given as

$$\hat{\boldsymbol{h}} = \bar{\boldsymbol{C}}^\dagger \bar{\boldsymbol{y}}, \tag{7}$$

where † denotes the pseudo-inverse. Note that here we need to design the superimposed training sequence properly so that $\bar{\boldsymbol{C}}$ is invertible.

B. CSE Design

The CSE is designed as an FIR filter with coefficients $\{w_\ell\}_{\ell=0}^{L_w}$, where the length of the FIR filter satisfies

$$L_w + L_h + 1 < N_c. \tag{8}$$

As shown in Fig. 1, the output of the CSE can be written as

$$r(n) = \sum_{\ell=0}^{L_w} w_\ell y(n-\ell). \tag{9}$$

Similar to (2), the $P \times 1$ vector $\boldsymbol{r}_k$ can be represented as

$$\boldsymbol{r}_k = \boldsymbol{W}_0 \boldsymbol{y}_k + \boldsymbol{W}_1 \boldsymbol{y}_{k-1}, \tag{10}$$

where $\boldsymbol{W}_0$ is a $P \times P$ Toeplitz matrix generated by first column $[w_0 \ \cdots \ w_{L_w} \ \boldsymbol{0}^T_{P-L_w-1}]^T$, and $\boldsymbol{W}_1$ is a $P \times P$ Toeplitz matrix generated by first row $[\boldsymbol{0}^T_{P-L_w} \ w_{L_w}, \ldots, w_1]$. Based on (6) and (10), the first-order statistic of $\boldsymbol{r}_k$ is given as:

$$\bar{\boldsymbol{r}} = E[\boldsymbol{r}_k] = \tilde{\boldsymbol{W}}\bar{\boldsymbol{y}} = \tilde{\boldsymbol{W}}\tilde{\boldsymbol{H}}\bar{\boldsymbol{c}} = \tilde{\boldsymbol{C}}\boldsymbol{H}\boldsymbol{w}, \tag{11}$$

where $\tilde{\boldsymbol{W}} = \boldsymbol{W}_0 + \boldsymbol{W}_1$ is a circulant matrix, $\boldsymbol{w}$ is a vector containing $L_w + 1$ CSE coefficients $\{w_\ell\}$, and $\boldsymbol{H}$ denotes the first $L_w + 1$ columns of $\tilde{\boldsymbol{H}}$. The fourth equality in (11) is based on the community property of the product of circulant matrices.

Given (11) and CSE coefficients $\{w_\ell\}$, the shortened channel can be represented as

$$\hat{\boldsymbol{h}}_{sh} = \tilde{\boldsymbol{C}}^{-1}\bar{\boldsymbol{r}}. \tag{12}$$

To estimate $\bar{\boldsymbol{y}}$ and $\bar{\boldsymbol{r}}$, we use their sample means based on N_b blocks as

$$\bar{\boldsymbol{y}} \approx \frac{1}{N_b}\sum_{k=1}^{N_b} \boldsymbol{y}_k, \quad \bar{\boldsymbol{r}} \approx \frac{1}{N_b}\sum_{k=1}^{N_b} \boldsymbol{r}_k. \tag{13}$$

Note that $\hat{\boldsymbol{h}}_{sh}$ is the circular convolution of zero-padded $\{w_\ell\}$ and zero-padded $\{h_\ell\}$ with length P. Define $\hat{\boldsymbol{h}}_{sh} = [\hat{\boldsymbol{h}}^T_{main}, \hat{\boldsymbol{h}}^T_{res}]^T$, where $\hat{\boldsymbol{h}}_{main}$ contains the first $(L_{cp}+1)$ entries of $\hat{\boldsymbol{h}}_{sh}$ and $\hat{\boldsymbol{h}}_{res}$ contains the rest. To design a CSE, our objective is to fix $\|\hat{\boldsymbol{h}}_{res}\|$ while maximizing $\|\hat{\boldsymbol{h}}_{main}\|$.

Based on (12) and (11), we find that

$$\begin{bmatrix} \hat{\boldsymbol{h}}_{main} \\ \hat{\boldsymbol{h}}_{res} \end{bmatrix} = \boldsymbol{H}\boldsymbol{w} = \begin{bmatrix} \boldsymbol{H}_{main} \\ \boldsymbol{H}_{res} \end{bmatrix} \boldsymbol{w}, \tag{14}$$

where $\boldsymbol{H}_{main}$ with size $(L_{cp}+1) \times (L_w + 1)$ denotes the first $(L_{cp}+1)$ rows of $\boldsymbol{H}$ and $\boldsymbol{H}_{res}$ consists of the rest. It can be verified that $\|\hat{\boldsymbol{h}}_{res}\| = \|\boldsymbol{H}_{res}\boldsymbol{w}\|$ and $\|\hat{\boldsymbol{h}}_{main}\| = \|\boldsymbol{H}_{main}\boldsymbol{w}\|$. Since the true channel is unknown, we use the estimated channel in (7) to obtain $\hat{\boldsymbol{H}}_{main}$ and $\hat{\boldsymbol{H}}_{res}$. The objective of the CSE design is to find $\boldsymbol{w}$ such as

$$\max \|\hat{\boldsymbol{H}}_{main}\boldsymbol{w}\|^2, \ \text{ subject to } \|\hat{\boldsymbol{H}}_{res}\boldsymbol{w}\|^2 = 1. \tag{15}$$

Similar CSE designs have been shown in [7, 9]. Note that the size of $\hat{\boldsymbol{H}}_{res}$ is $(N_c-1)\times(L_w+1)$ and thus the size of $\hat{\boldsymbol{H}}_{res}^{\mathcal{H}}\hat{\boldsymbol{H}}_{res}$ is $(L_w+1)\times(L_w+1)$. Because $N_c > L_h + L_w + 1$, $\hat{\boldsymbol{H}}_{res}^{\mathcal{H}}\hat{\boldsymbol{H}}_{res}$ is a symmetric and positive definite matrix with probability one. It can be decomposed by the Cholesky decomposition as $\hat{\boldsymbol{H}}_{res}^{\mathcal{H}}\hat{\boldsymbol{H}}_{res} = \boldsymbol{Q}^{\mathcal{H}}\boldsymbol{Q}$. Therefore, the solution of (15) is given by

$$\hat{\boldsymbol{w}} = \boldsymbol{Q}^{-1}\boldsymbol{u}_{\max}, \tag{16}$$

where $\boldsymbol{u}_{\max}$ is the eigenvector corresponding to the maximum eigenvalue of $\boldsymbol{Q}^{-\mathcal{H}}\hat{\boldsymbol{H}}_{main}^{\mathcal{H}}\hat{\boldsymbol{H}}_{main}\boldsymbol{Q}^{-1}$, and $\|\boldsymbol{u}_{\max}\| = 1$.

Note that here our CSE design does not depend on the exact knowledge of the channel order L_h. It only requires an upper bound of the channel order so that assumption *as3)* can be satisfied and $\hat{\boldsymbol{h}}$ can be obtained.

C. *Training Sequence Design*

Apparently, the design of the training sequence affects the performance of the channel estimation and channel shortening equalization. Here we give several training sequences, all of which satisfy the condition in *as3)*.

Periodic delta sequence: In this case, the training sequence is deterministic and periodic with period P; i.e. $\boldsymbol{c}_k = \boldsymbol{c}_{k-1}$. Within each period, the structure of the sequence is given as follows:

$$\boldsymbol{c}_k = [\sigma_c, \boldsymbol{0}_{P-1}^T]^T. \tag{17}$$

With this sequence, $\tilde{\boldsymbol{C}}$ becomes a diagonal matrix. Therefore, the channel estimator in (7) and the effective channel estimator in (12) have low complexity since $\tilde{\boldsymbol{C}}^{\dagger} = \frac{1}{\sigma_c}\boldsymbol{I}_P$. Based on (7) and (13), the channel mean square error (MSE) becomes

$$E[\|\hat{\boldsymbol{h}} - \boldsymbol{h}\|^2] = \frac{L_h+1}{\sigma_c^2 N_b}\left(\sigma_u^2\sum_{\ell=0}^{L_h}|h_\ell|^2 + N_0\right), \tag{18}$$

where σ_u^2 is the variance of $u(n)$. From (18), we observe that as the number of blocks N_b and/or the power of training part σ_c increase, the channel MSE decreases. This is consistent with most statistical channel estimators. Increasing signal power σ_u^2 with all other variables fixed, the channel MSE increases because the power from the residual average of the information symbols increases.

The quality of the channel estimation affects the CSE design and also the decoding performance. It is desirable to increase the power on superimposed training σ_c^2. However, the signal-to-noise ratio (SNR) σ_u^2/N_0 also affects the decoding performance. Therefore, if the total power $\mathcal{P} = \sigma_c^2 + P\sigma_u^2$ is fixed, there may exist an optimal power allocation factor to distribute the power on training and information symbols so that the bit-error-rate is minimized. This issue will be further examined by simulations.

Periodic training sequence: The training sequence can also be picked as a periodic m-sequence. For example, in [15], a periodic training sequence is given as $[1, -1, -1, 1, 1, 1, -1]$ with period 7 and the total power of training sequence in one block is still σ_c^2. The performance analysis is provided in [15].

Non-zero mean random sequence: Transmitting periodic non-zero mean symbols is another way to perform synchronization and channel estimation [6]. In this design, the information sequence $s(n)$ satisfying

$$E[s(kN_c + n)] = \mu_0\delta(n) \text{ with } \mu_0 \neq 0, \tag{19}$$

where $\delta(n)$ is Kronecker delta function. A separate design for superimposed training sequence $c(n)$ is not necessary in this case, since it has already been taken into account in (19). The non-zero mean symbol is drawn from non-symmetric constellation. For example, the non-symmetric BPSK is $\{\theta, -1\}$ where $\theta > 1$. For this type of sequence, after averaging, the input-output relationship is the same as the one in (2). Thus, the performance of channel estimation and CSE design is the similar to the one with the deterministic periodic delta sequence.

IV. SYMBOL DETECTION

At the receiver, after channel shortening equalization, we perform symbol decoding. As shown in Fig. 1, for each received block $\boldsymbol{r}_k$, we first remove the effect of the superimposed training, and then perform OFDM demodulation which includes CP removal and FFT operation. After removing the effect of training sequence from $\boldsymbol{r}_k$, the signal prior to CP removal and FFT operation can be written as

$$\begin{aligned}\boldsymbol{v}_k = \boldsymbol{r}_k &- \hat{\boldsymbol{W}}_0(\hat{\boldsymbol{H}}_0\boldsymbol{c}_k + \hat{\boldsymbol{H}}_1\boldsymbol{c}_{k-1}) \\ &- \hat{\boldsymbol{W}}_1(\hat{\boldsymbol{H}}_0\boldsymbol{c}_{k-1} + \hat{\boldsymbol{H}}_1\boldsymbol{c}_{k-2}),\end{aligned} \tag{20}$$

where $\hat{\boldsymbol{W}}_0$ and $\hat{\boldsymbol{W}}_1$ are generated by $\hat{\boldsymbol{w}}$ in (16) and have the same structure as those in (10).

We perform CP removal by left multiplying $\boldsymbol{v}_k$ a CP removal matrix $\boldsymbol{R}_{cp} = [\boldsymbol{0}_{L_{cp}\times N_c}, \boldsymbol{I}_{N_c}]$, and then perform FFT operation by an FFT matrix $\boldsymbol{F}$. The output is an $N_c \times 1$ vector

$$\hat{\boldsymbol{z}}_k = \boldsymbol{F}\boldsymbol{R}_{cp}\boldsymbol{v}_k. \tag{21}$$

Taking into account the OFDM operation at the transmitter and CSE effects, if the channel is perfectly known, we can rewrite (21) as

$$\begin{aligned}\boldsymbol{z}_k = \boldsymbol{D}_h\boldsymbol{s}_k + \boldsymbol{F}\boldsymbol{R}_{cp}\,((\boldsymbol{W}_0\boldsymbol{H}_1 + \boldsymbol{W}_1\boldsymbol{H}_0)\boldsymbol{u}_{k-1} \\ + \boldsymbol{W}_1\boldsymbol{H}_1\boldsymbol{u}_{k-2}) + \boldsymbol{\xi}_k,\end{aligned} \tag{22}$$

where $\boldsymbol{\xi}_k = \boldsymbol{F}\boldsymbol{R}_{cp}(\boldsymbol{W}_0\boldsymbol{\eta}_k + \boldsymbol{W}_1\boldsymbol{\eta}_{k-1})$, and $\boldsymbol{D}_h = \mathrm{diag}[H(0), \ldots, H(N_c - 1)]$ with $H(n) = \sum_{\ell=0}^{L_w+L_h} h_{sh}(\ell)\; e^{-j2\pi\ell/N_c}$. It is not difficult to verify that $\boldsymbol{W}_1\boldsymbol{H}_1 = \boldsymbol{0}$ and $\boldsymbol{W}_0\boldsymbol{H}_1 + \boldsymbol{W}_1\boldsymbol{H}_0$ is a banded matrix with non-zero entries in the last $\max\{L_h, L_w\}$ columns and $L_h + L_w$ rows. By multiplying $\boldsymbol{R}_{cp}$, the IBI term is reduced. Since the CSE is not ideal, there exists residual IBI which causes error-floor as shown by simulations later.

If the channel is unknown, we use the estimated channel to replace $\boldsymbol{D}_h$ as $\hat{\boldsymbol{D}}_h$. Thus, based on $\hat{\boldsymbol{z}}_k$ and $\hat{\boldsymbol{D}}_h$, we can estimate the symbol as

$$\hat{s}(n) = \arg\min_{s\in\mathcal{A}} |\hat{z}_k(n) - \hat{H}(n)s|^2, \tag{23}$$

where $\hat{z}_k(n)$ is the nth entry of $\hat{\boldsymbol{z}}_k$ and $\hat{H}(n)$ is the (n, n)th element of $\hat{\boldsymbol{D}}_h$. In this way, we keep the low-complexity decoder of OFDM.

Therefore, by combining CSE with channel estimation based on superimposed training, we have designed a low-complexity scheme for decoding OFDM symbols in unknown frequency-selective channels, which is bandwidth efficient because it allows the CP length to be shorter than the channel order and training sequences do not need extra bandwidth.

V. SIMULATION RESULTS

In this section, we present three simulated examples to illustrate our design and test the effects of the different parameters on the performance.

Example 1: Fig. 2 is a plot of the QPSK symbol error rate (SER) versus SNR for three cases: i) unknown channel, $N_b = 150$; ii) unknown channel, $N_b = 300$; and iii) perfect channel state information (PCSI) is known at the receiver. The plot was generated using 1000 monte carlo realizations of a channel $h(l)$, where the channel taps are zero mean complex gaussian with variance $\sigma_l^2 = \lambda e^{-0.2l}$ and λ is chosen so that $\sum_{l=0}^{L_h}\sigma_l^2 = 1$. The other parameters for the simulation are the number of subcarriers $N_c = 64$, the channel order $L_h = 16$, the CP length $L_{cp} = 8$, the number of coefficients for CSE $L_w = 41$, the training sequence $\boldsymbol{c}_k = [\sigma_c, \boldsymbol{0}^T]^T$ is chosen according to (17) and the power allocation factor $\rho = \sigma_c^2/\sigma_x^2 = 0.25$, where $\sigma_x^2 = P\sigma_u^2 + \sigma_c^2$. In addition to the SER for the proposed CSE, Fig. 2 also includes the SER achieved when CSE is not implemented. The plot shows that, when the channel is perfectly known at the receiver and the proposed CSE is implemented, an order of magnitude improvement in SER is possible in the high-SNR regime. Even when the first-order channel estimate in (7) is used with the proposed CSE, more than a two-fold improvement in SER is possible in the high-SNR regime.

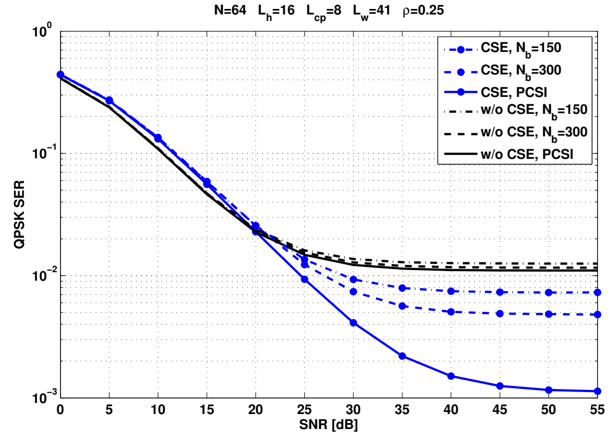

Fig. 2. SER vs. SNR with and without CSE for $N_b = 150$, $N_b = 300$ and PCSI.

Example 2: Fig. 3 is a plot of the SER versus the power allocation factor $\rho = \sigma_c^2/\sigma_x^2$. The monte carlo channel realizations were generated just as they were in the first example and the simulation parameters are $N = 64$, $L_h = 16$, $L_{cp} = 8$, $L_w = 41$, SNR= 35dB and $\boldsymbol{c}_k = [\sigma_c, \boldsymbol{0}^T]^T$. The plot illustrates that it is desirable to choose ρ so that the SER is minimized. When ρ is small, channel MSE is large, and thus the SER is high. Notice that in this case, the SER with CSE is even worse than the case without CSE. This is because the unreliably estimated channel misleads the

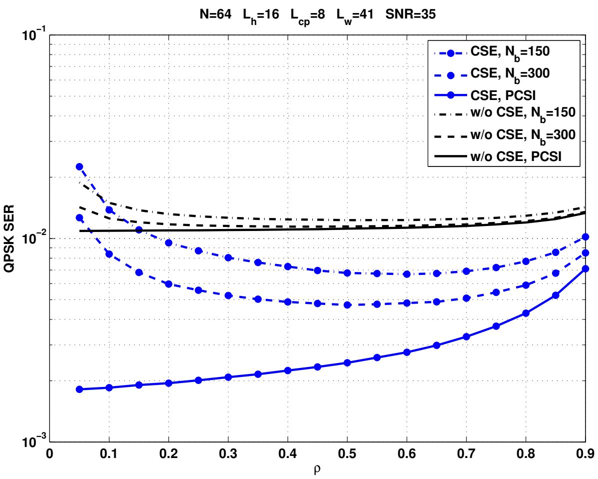

Fig. 3. SER vs. ρ with and without CSE for $N_b = 150$, $N_b = 300$ and PCSI.

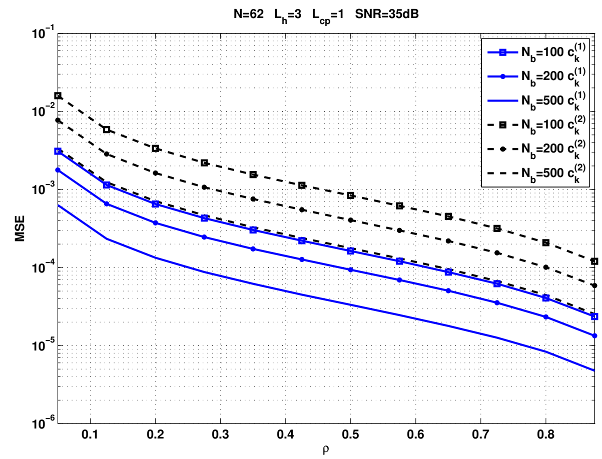

Fig. 4. MSE vs. ρ with CSE for two different superimposed sequences $\boldsymbol{c}_k^{(1)}$ and $\boldsymbol{c}_k^{(2)}$.

CSE so that the performance is worse. When ρ is too high, the power for information symbols σ_u^2 is low. Therefore, the decoding is unreliable. For this set of simulation parameters, the plot shows that $\rho = 0.5$ is the best choice if the proposed CSE is used when $N_b = 300$ and $\rho = 0.6$ if $N_b = 150$.

Example 3: Fig. 4 is a plot of the MSE versus ρ for two different superimposed sequences, $\boldsymbol{c}_k^{(1)} = [\sigma_c, \mathbf{0}^T]^T$ and $\boldsymbol{c}_k^{(2)} = [\boldsymbol{v}, \boldsymbol{v}, ..., \boldsymbol{v}]^T$, where $\boldsymbol{v} = [1, -1, -1, 1, 1, 1, -1]$. The monte carlo channel realizations were generated just as they were in the first example and the simulation parameters are $N = 62$, $L_h = 3$, $L_{cp} = 1$ and SNR$=$ 35dB. The plot shows that using $\boldsymbol{c}_k^{(1)}$ as the superimposed training sequence leads to a large MSE reduction over using $\boldsymbol{c}_k^{(2)}$ as the superimposed training sequence.

VI. **CONCLUSIONS**

In this paper, using superimposed training, we design a first-order channel estimator and then construct a channel shortening equalizer (CSE) that is a finite impulse response (FIR) filter to shorten the channel when the affordable CP length is shorter than the channel order. With this CSE, the residual IBI is reduced but the symbol decoding complexity is still low. Simulation results show that our technique is effective[1].

[1]The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government.

REFERENCES

[1] J. Balakrishnan, R. K. Martin, and C. R. Johnson Jr., "Blind, adaptive channel shortening by sum-squared auto-correlation minimization (SAM)," *IEEE Trans. Signal Processing*, vol. 51, pp. 3086-3093, Dec. 2003.

[2] R. J. Baxley and J. Kleider, "Embedded synchronization/pilot sequence creation using POCS," *Proc. ICASSP*, vol. IV, pp. 321-324, May 2006.

[3] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5-14, May 1990.

[4] N. Chen and G. T. Zhou, "What is the price paid for superimposed training in OFDM?," *Proc. ICASSP*, vol. VI, pp. 421-424, May 2004.

[5] N. Chen and G. T. Zhou, "Superimposed training for OFDM: a peak-to-average power ratio analysis," *IEEE Trans. Signal Processing*, vol. 54, no. 6, June 2006.

[6] X. Luo and G. B. Giannakis, "Low-complexity blind synchronization and demodulation for (ultra-) wideband multi-user ad hoc access," *IEEE Trans. Wireless Communications*, vol. 5, no. 5, May 2006.

[7] X. Ma, L. Song, and J. Kleider, "Channel shortening equalization for differential OFDM systems," *Proc. IEEE Workshop on Signal Processing Advances in Wireless Commun.*, pp. 860-864, Jun. 2005.

[8] R. K. Martin, J. Balakrishnan, W. A. Sethares, and C. R. Johnson Jr., "A blind adaptive TEQ for multicarrier systems," *IEEE Signal Processing Letters*, vol. 9, pp. 341-343, Nov. 2002.

[9] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone

transceivers," *IEEE Trans. Communications*, vol. 44, pp. 1662-1672, Dec, 1996.

[10] T. Miyajima and Z. Ding, "Second-order statistical approaches to channel shortening in multicarrier systems," *IEEE Trans. Signal Processing*, vol. 52, no. 11, pp. 3253-3264, Nov. 2004.

[11] M. Munster and L. Hanzo, "Parallel-interference-cancellation-assisted decision-directed channel estimation for OFDM systems using multiple transmit antennas," *IEEE Trans. Wireless Communications*, vol. 4, no. 5, pp. 2148-2162, Sep. 2005.

[12] R. Nawaz and J. A. Chambers, "A novel single lag autocorrelation minimization (SLAM) algorithm for blind adaptive channel shortening," *Proc. ICASSP*, vol. III, pp. 885-888, Mar. 2005.

[13] S. Ohno and G. B. Giannakis, "Optimal training and redundant precoding for block transmissions with application to wireless OFDM," *IEEE Trans. Communications*, vol. 50, no. 12, pp. 2113-2123, Dec. 2002.

[14] F. Romano and S. Barbarossa, "Non-data aided adaptive channel shortening for efficient multi-carrier systems," *Proc. ICASSP*, vol. 4, pp. 233-236, Apr. 2003.

[15] J. Tugnait and X. Meng, "On superimposed training for channel estimation: performance analysis, training power allocation, and frame synchronization," *IEEE Trans. Signal Processing*, vol. 54, no. 2, pp. 752-765, Feb. 2006.

[16] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: where Fourier meets Shannon" *IEEE Signal Processing Magazine*, Vol. 17, No. 3, pp. 29-48, May 2000.

[17] J. Zhang, W. Ser and J. Zhu, "Effective optimisation method for channel shortening in OFDM systems," *IEE Proc. Communications*, vol. 150, no. 2, pp. 85–90, Apr. 2003.

[18] G. T. Zhou, M. Viberg, and T. McKelvey, "A first-order statistical method for channel estimation," *IEEE Signal Processing Letters*, vol. 10, no. 3, pp. 57-60, Mar. 2003.