Bit Error Rate Minimizing Channel Shortening Equalizers for Cyclic Prefixed Systems

Richard K. Martin, *Member, IEEE*, Geert Ysebaert, *Associate Member, IEEE*, and Koen Vanbleu, *Member, IEEE*

***Abstract*—Cyclic prefixed communications, such as multicarrier communications, first became widely used in the context of digital subscriber lines (DSL). In DSL, bit loading is allowed at the transmitter, and the performance metric is the bit rate that can be provided without exceeding a given bit error rate (BER). Wireless cyclic prefixed systems are now becoming increasingly popular, and in such systems the appropriate performance metric is the BER for a given bit loading at the transmitter. Cyclic prefixed systems perform well in the presence of multipath, provided that the channel delay spread is shorter than the guard interval between transmitted blocks. If this condition is not met, a channel shortening equalizer can be used to shorten the channel to the desired length. Previous work on channel shortening has largely been in the context of DSL, thus it has focused on maximizing the bit rate. In this paper, we propose a channel shortener that attempts to directly minimize the BER for a multiple-input multiple-output channel model. We simulate the performance of the resulting channel shortener and compare it to existing designs and the matched filter bound.**

***Index Terms*—Bit error rate, channel shortening, cyclic prefix, equalization, multicarrier.**

I. Introduction

THERE are currently two types of cyclic prefixed communication systems: multicarrier modulation and single-carrier cyclic prefixed (SCCP) modulation [also known as single-carrier frequency domain equalization (SC-FDE)]. In wireless systems, multicarrier modulation is called orthogonal-frequency-division multiplexing (OFDM), and in wireline systems, it is referred to as discrete multitone (DMT). Examples of wireless multicarrier systems include wireless local area networks (IEEE 802.11a/g, HIPERLAN/2, MMAC) [1], wireless metropolitan area networks (IEEE 802.16) [2], digital video and audio broadcasting in Europe [3], [4], satellite radio (e.g., Sirius and XM Radio) [5], and the proposed standard for multiband ultrawideband (IEEE 802.15.3a). Examples of wireline multicarrier systems include power line communications (e.g., HomePlug) [6] and digital subscriber lines (DSLs) [7]. SCCP modulation has not been implemented as widely by industry, but it is gaining support in the literature [8]–[10].

Cyclic prefixed systems are very robust to multipath, provided that the delay spread of the transmission channel is less than the length of the cyclic prefix (CP) inserted between transmitted blocks. If the channel is short, then equalization of the channel can be done tonewise in the frequency domain by a bank of complex scalars. This is called a frequency-domain equalizer (FEQ). However, if the channel is longer than the CP, additional equalization is required. Typically, this takes the form of a channel shortening equalizer (CSE) [also known as a time-domain equalizer (TEQ)], which is a filter at the front end of the receiver. Typically, the CSE is designed so that the effective channel is shorter than the CP (but not necessarily a single impulse); a survey of CSE design for DSL can be found in [11].

Channel shorteners have a wide variety of applications. The first application of channel shortening was to maximum-likelihood sequence estimation (MLSE) during the 1970s [12]–[14]. Since the mid-1990s, it has been used to shorten the long impulse responses of twisted pair telephone lines encountered by DSL [15]–[18]. Channel shortening can also be used to reduce the computational complexity of multiuser detection by suppressing the signals from a subset of users and detecting the remainder [19]. Channel shorteners can also be used for complexity reduction for ultra wideband systems, in which the number of correlators needed for detection of a pulse can be reduced by shortening the multipath channel [20]. Yet another application is in acoustics. Psychoacoustics defines the D50-measure for intelligibility of speech as the ratio of energy in a 50-ms window of the room impulse response to the total energy of the impulse response, and optimization of this measure can be performed by a channel shortener [21].

While early channel shortening algorithms were based on heuristic objective functions, more recent designs have focused on maximizing the bit rate for a given bit error rate (BER) [22]–[24], which is the appropriate measurement of performance in wireline multicarrier systems (such as DSL) that allow bit loading across the subcarriers. In contrast to the DSL application, wireless multicarrier and SCCP systems usually have a fixed bit loading, and receiver performance is measured in terms of bit error rate (BER) for a fixed bit rate. Moreover, even in DSL, once the initial bit loading has been done, the CSE can still be updated to minimize the BER for that bit loading. To date, no CSE design in the literature explicitly attempts to minimize the BER. Hence, the main goals of this paper are to investigate the BER cost surfaces for multicarrier and SCCP systems and to develop and assess CSE designs that attempt to minimize the BER at the output of the receiver.

Manuscript received July 7, 2006; revised September 15, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Petr Tichavsky. The work of R. K. Martin is funded in part by the Air Force Office of Scientific Research. The views expressed in this paper are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government. This document has been approved for public release; distribution unlimited.

R. K. Martin is with the Department of Electrical and Computer Engineering, The Air Force Institute of Technology/ENG, Wright-Patterson AFB, OH 45433 USA (e-mail: richard.martin@afit.edu).

G. Ysebaert is with Alcatel Bell, 2018 Antwerpen, Belgium.

K. Vanbleu is with Broadcom Corporation, 2800 Mechelen, Belgium (e-mail: koen.vanbleu@broadcom.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2007.893913

1053-587X/$25.00 © 2007 IEEE

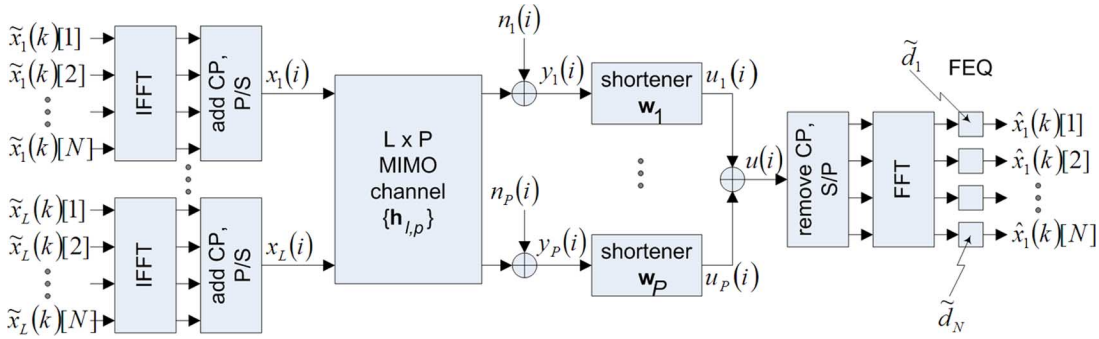


Fig. 1. Complex baseband multicarrier system model. No null tones are shown (i.e., $N_a = N$), but if null tones are present, they can be represented by replacing the corresponding inputs and FEQ coefficients by zeros.

The remainder of this paper is organized as follows. In Section II, we present the multicarrier and SCCP system models. In Section III, we derive the BER as a function of the CSE weights, for both multicarrier and SCCP systems. In Section IV, we present the minimum error rate CSE design algorithm for multicarrier systems and discuss a heuristic approach for minimizing the BER of an SCCP system. In Section V, we compare the BER of the proposed algorithms to the BER of existing popular channel shorteners, using a simulated wireless channel model. Section VI concludes the paper.

II. System Model

In this section, we first describe the multicarrier system model and then discuss the related SCCP system model. We assume a multiple-input multiple-output (MIMO) channel model with L transmit antennas and P receive antennas (or oversampling the received signal by a factor of P). Throughout, $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote complex conjugate, matrix transpose, and conjugate transpose, respectively, and $\mathcal{E}\{\cdot\}$ denotes statistical expectation. The notation was chosen to be as consistent with [25] as possible.

The MIMO multicarrier system model is shown in Fig. 1. There are N_a active tones out of a possible N tones; for example, in IEEE 802.11a and HIPERLAN/2 wireless LANs, $N = 64$, but only $N_a = 52$ tones are used to transmit data or pilot tones. The set of active tones is denoted $\mathcal{S}_a$. The complex finite-alphabet data symbols (usually multilevel QAM data) are blocked into groups of size N_a and then zero-padded to length N. The kth such block for transmitter l is denoted as the column vector $\tilde{\mathbf{x}}_l(k)$ whose nth element is $\tilde{\mathbf{x}}_l(k)[n]$, and it is referred to as the *frequency-domain transmitted data*. Then, an inverse discrete Fourier transform (IDFT) is taken to get N time-domain samples. In practical systems, the DFT is implemented by the fast Fourier transform (FFT), though, for purposes of analysis, we use $\mathcal{F}$ to denote the unitary $N \times N$ matrix that computes the DFT, with element (m, n) given by $(1/\sqrt{N})e^{-\jmath 2\pi mn/N}$. A cyclic prefix is inserted by copying the last ν samples of the block to the beginning of the block, and then the $S = N + \nu$ samples are transmitted serially. The ith transmitted data sample is denoted $x_l(i)$. Note that $l \in \{1, \ldots, L\}$ is the index of the transmit antenna, $p \in \{1, \ldots, P\}$ is the index of the receive antenna, k is the block index, n is the tone index, i is the sample index, and $\jmath = \sqrt{-1}$ is the unit imaginary number.

The redundancy in the transmitted signal due to the CP can be represented by

$$x_l(Sk + i) = x_l(Sk + i + N), \quad 1 \le l \le L, \quad 1 \le i \le \nu, \quad -\infty < k < \infty. \tag{1}$$

Let $\mathbf{h}_{l,p}$ be an FIR filter of length L_h, which models the channel from transmit antenna l to receive antenna p, and let $\mathbf{w}_p$ be an FIR filter of length L_w, which is the CSE to be designed for antenna p. Let $\mathbf{H}_{l,p}$ be the tall channel convolution matrix for $\mathbf{h}_{l,p}$, which is an $L_c \times L_w$ Toeplitz matrix, where $L_c = L_h + L_w - 1$ is the length of the effective channel.

Defining the transmitted signal, noise, and received signal vectors as

$$\mathbf{x}_l(i) = [x_l(i), \ldots, x_l(i - L_c + 1)]^T \tag{2}$$
$$\boldsymbol{\eta}_p(i) = [\eta_p(i), \ldots, \eta_p(i - L_w + 1)]^T \tag{3}$$
$$\mathbf{y}_p(i) = [y_p(i), \ldots, y_p(i - L_w + 1)]^T \tag{4}$$

and the "stacked" column vectors

$$\mathbf{x}(i) = \left[\mathbf{x}_1^T(i), \ldots, \mathbf{x}_L^T(i)\right]^T \tag{5}$$
$$\boldsymbol{\eta}(i) = \left[\boldsymbol{\eta}_1^T(i), \ldots, \boldsymbol{\eta}_P^T(i)\right]^T \tag{6}$$
$$\mathbf{y}(i) = \left[\mathbf{y}_1^T(i), \ldots, \mathbf{y}_P^T(i)\right]^T \tag{7}$$
$$\mathbf{w} = \left[\mathbf{w}_1^T, \ldots, \mathbf{w}_P^T\right]^T \tag{8}$$

we can compactly write the CSE input vector as

$$\mathbf{y}(i) = \underbrace{\begin{bmatrix} \mathbf{H}_{1,1}^T, & \cdots, & \mathbf{H}_{L,1}^T \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{1,P}^T, & \cdots, & \mathbf{H}_{L,P}^T \end{bmatrix}}_{\mathbf{H}^T} \mathbf{x}(i) + \boldsymbol{\eta}(i). \tag{9}$$

The input to the DFT at the receiver is then obtained by passing the signal through the bank of CSEs

$$u(i) = \mathbf{w}^T \mathbf{y}(i). \tag{10}$$

After channel shortening, the cyclic prefix is discarded from each block and a DFT is used to return the data to the frequency domain. This requires an estimate of the transmission delay Δ, since the length N DFT input vector for block k is

$$\mathbf{u}(k) = [u(Sk + \nu + 1 + \Delta), \ldots, u(S(k + 1) + \Delta)]^T. \tag{11}$$

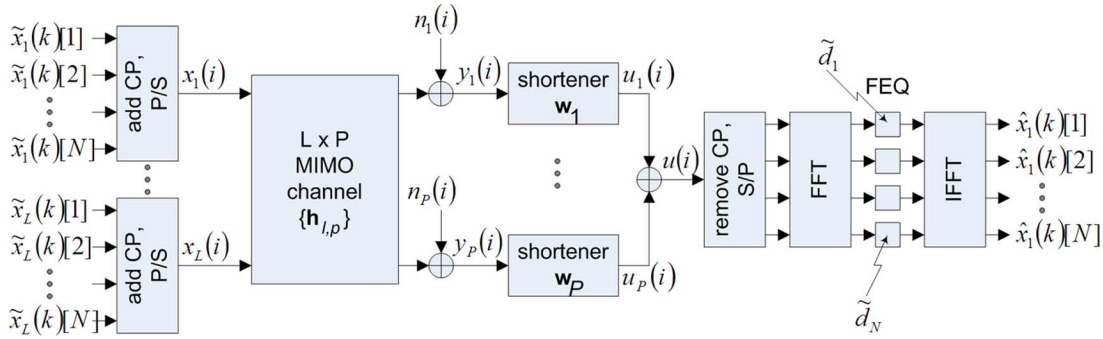


Fig. 2. Complex baseband SCCP system model.

The delay Δ is a design parameter whose choice affects the values of the optimal CSE as well as the performance that can be attained [26]. The FEQ input vector is obtained via a DFT

$$\tilde{\mathbf{u}}(k) = \mathcal{F}\mathbf{u}(k). \tag{12}$$

Finally, the FEQ is used to invert the channel in the frequency domain to get the estimate of the data. Without loss of generality, we assume that the receiver is attempting to recover the data from transmitter $l = 1$. Thus, the soft estimate (before the decision device) of the frequency-domain data is

$$\hat{\mathbf{x}}_1(k) = \tilde{\mathbf{d}}_o \odot \tilde{\mathbf{u}}(k) \tag{13}$$

where $\odot$ denotes element-by-element multiplication; and $\tilde{\mathbf{d}}_o$ contains the FEQ $\tilde{\mathbf{d}}$, a bank of N_a complex scalars, zero-padded to length N. The zero locations correspond to the zero locations in the transmitted vector $\tilde{\mathbf{x}}(k)$. In a multiuser scheme, the data for $l = 2, \ldots, L$ can be ignored, or a multiuser detection technique can be used to mitigate the interference. In a single-user scheme, the data $\tilde{x}_l$ may be the same on all transmitters l; or an Alamouti transmit diversity space–time code may be used [27].

The SCCP model is almost identical to the multicarrier system model. There are only two changes. The inverse fast Fourier transform (IFFT) is removed from the transmitter and placed at the receiver, after the FEQ but before the decision device. Also, since the transmitter operates entirely in the time domain, there are no null tones, so $N_a = N$. This SCCP system model is shown in Fig. 2.

In order to minimize repetition, the SCCP notation will be kept as similar as possible to the multicarrier modulation notation. The kth source data block of user l is the $N \times 1$ vector $\tilde{\mathbf{x}}_l(k)$. This data is converted to a parallel data stream with a cyclic prefix, and transmitted as $x_l(i)$ with time index i. The channel model and all other intermediate signals are represented in an identical fashion, up through the FEQ input vector $\tilde{\mathbf{u}}(k)$. The $N \times 1$ FEQ vector is denoted $\tilde{\mathbf{d}}$, which is not zero-padded. The FEQ output is denoted

$$\hat{\mathbf{u}}(k) = \tilde{\mathbf{d}} \odot \tilde{\mathbf{u}}(k) \tag{14}$$

and the final IFFT output vector is denoted

$$\hat{\mathbf{x}}_1(k) = \mathcal{F}^H \hat{\mathbf{u}}(k). \tag{15}$$

Note that, as before, we are assuming the receiver is attempting to recover the transmission from user 1.

Interleaving and forward-error correction (FEC) blocks can be inserted at the beginning and end of the system models of Figs. 1 and 2, although for conciseness, they are not depicted here.

III. BER Models

The goal of this section is to derive models for the BER of a multicarrier and an SCCP system, which we will attempt to optimize in Section IV. In Section III-A, we will model the BER for a multicarrier system in terms of the effective output signal-to-noise ratio (SNR) on the various subchannels, in Section III-B, we will review the subchannel SNR model for multicarrier systems in [25] and extend it to the MIMO case, in Section III-C, we will model the BER for an SCCP system in terms of the effective output SNR of the final N time-domain IFFT outputs, and in Section III-D, we will model the output SNR of the IFFT outputs of an SCCP system. Although a portion of Section III-B has appeared in [25], the remainder of Section III has not appeared in the literature.

A. Multicarrier BER Model

In computing the BER, we assume that the total residual interference and noise at the output of each tone has a Gaussian PDF and that M-level QAM signaling (consisting of two orthogonal $\sqrt{M}$-level PAM constellations) is used on each tone. For example, the IEEE 802.11a and HIPERLAN/2 wireless LAN standards support 4-QAM, 16-QAM, and 64-QAM. The probability of error on each of the PAM components is given by [28, pp. 225–226]

$$P_{\sqrt{M}}(n) = 2\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3}{M-1}\mathrm{SNR}_n}\right) \tag{16}$$

hence, the symbol error rate (SER) on tone n is

$$P_M(n) = 2P_{\sqrt{M}}(n) - (P_{\sqrt{M}}(n))^2 \tag{17}$$

where SNR_n is the effective signal-to-interference-and-noise ratio on tone n (usually called the subchannel SNR); and $Q(x)$ is the Q-function, which is the integral of a unit Gaussian PDF from x to infinity. For the $M = 4$ case, (16) is the BER for tone n, and it reduces to $Q(\sqrt{\mathrm{SNR}_n})$, which we use here for simplicity of notation. Averaging (16) and (17) over all of the

active tones, the BER and the SER for the output of an OFDM system with $M = 4$ are

$$\mathrm{BER}_{\mathrm{mcm}} = \frac{1}{N_a} \sum_{n \in \mathcal{S}_a} Q(\sqrt{\mathrm{SNR}_n}) \tag{18}$$

$$\mathrm{SER}_{\mathrm{mcm}} = \frac{1}{N_a} \sum_{n \in \mathcal{S}_a} (2Q(\sqrt{\mathrm{SNR}_n}) - Q^2(\sqrt{\mathrm{SNR}_n})). \tag{19}$$

Either can be used as the objective function to be minimized, although in the remainder of the paper we will focus on the BER. In the next section, we discuss the model of the subchannel SNR in terms of the CSE coefficients.

B. Multicarrier Subchannel SNR Model

The subchannel SNR can be modeled in various ways [16], [22]–[25], [29]. Most of the proposed models take the form

$$\mathrm{SNR}_n = \frac{\mathbf{w}^H \mathbf{B}_n \mathbf{w}}{\mathbf{w}^H \mathbf{A}_n \mathbf{w}} \tag{20}$$

where $\mathbf{A}_n$ and $\mathbf{B}_n$ are Hermitian positive semi-definite matrices. The distinction between the models lies in how the intercarrier interference (ICI), intersymbol interference (ISI), and DFT leakage are taken into account. In this section, we review the subchannel SNR model proposed in [24] and [25], although in the more general context of a MIMO channel model.

The DFT output can be written as

$$\mathcal{F}\underbrace{(\mathbf{Y}_k \mathbf{w})}_{\mathbf{u}(k)} = \underbrace{(\mathcal{F}\mathbf{Y}_k)}_{\tilde{\mathbf{Y}}_{k,N}} \mathbf{w} \tag{21}$$

where $\mathbf{Y}_k$ is a block Toeplitz matrix of size $N \times (PL_w)$, where each $N \times L_w$ sub-block contains the data that will be convolved with the pth CSE, $\mathbf{w}_p$, and successive rows are vectors $\mathbf{y}^T(i)$ for successive values of i. Then, $\tilde{\mathbf{Y}}_{k,N}$ is an $N \times (PL_w)$ matrix as well, with the nth row denoted by $\tilde{\mathbf{y}}_{k,n}$. For conciseness of notation, as in [25], define the correlation terms (still assuming that the user of interest is $l = 1$)

$$\sigma_n^2 \triangleq \mathcal{E}\{|\tilde{\mathbf{x}}_1^*(k)[n]|^2\} \tag{22}$$

$$\boldsymbol{\varphi}_n \triangleq \mathcal{E}\{\tilde{\mathbf{x}}_1^*(k)[n]\tilde{\mathbf{y}}_{k,n}\} \tag{23}$$

$$\mathbf{C}_n^2 \triangleq \mathcal{E}\left\{\tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{y}}_{k,n}\right\} \tag{24}$$

which have dimensions 1×1, $1 \times PL_w$, and $PL_w \times PL_w$, respectively.

The desired signal at the output of the FEQ (which is where we measure the subchannel SNR) is an undistorted copy of the transmitted signal on that tone at the input to the transmitter. Thus, the subchannel SNR on tone n is the ratio of the power of the desired signal $\tilde{\mathbf{x}}_1(k)[n]$ to the power of the total interference and noise (i.e., the error signal for block k and tone n), $\tilde{e}_{k,n} = \tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w} - \tilde{\mathbf{x}}_1(k)[n]$. Under this model, the subchannel SNR is

$$\mathrm{SNR}_n = \frac{\sigma_n^2}{\mathcal{E}\{|\tilde{e}_{k,n}|^2\}} = \frac{\sigma_n^2}{\mathcal{E}\{|\tilde{d}_n \tilde{\mathbf{y}}_{k,n} \mathbf{w} - \tilde{\mathbf{x}}_1(k)[n]|^2\}} \tag{25}$$

$$= \frac{\sigma_n^2}{|\tilde{d}_n|^2 \mathbf{w}^H \mathbf{C}_n^2 \mathbf{w} - \tilde{d}_n \boldsymbol{\varphi}_n \mathbf{w} - \tilde{d}_n^* \boldsymbol{\varphi}_n^* \mathbf{w}^* + \sigma_n^2}. \tag{26}$$

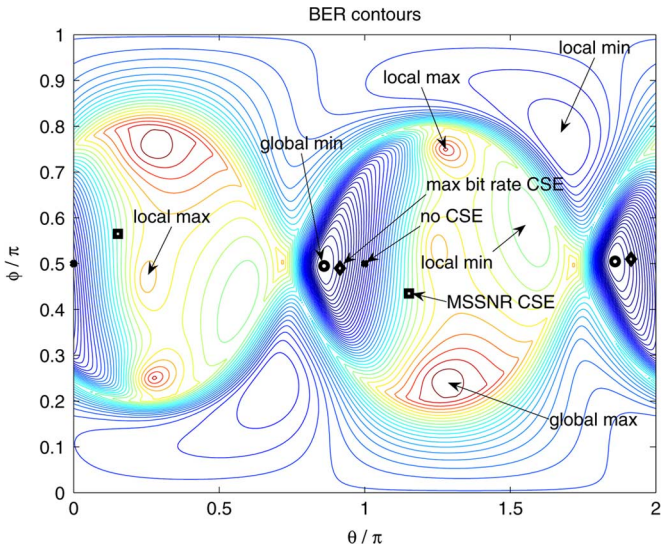


Fig. 3. Contours of the BER for a 3-tap CSE under a unit norm constraint for a multicarrier system. The CSE is parameterized in spherical coordinates by the angles θ and ϕ. The channel is $\mathbf{h} = [1, -0.3, 0.7]$, the FFT size is $N = 8$, the CP length is $\nu = 1$, and the SNR is 20 dB. Note that the cost function is symmetric with respect to a sign change in the CSE ($\mathbf{w} \rightarrow -\mathbf{w}$), which corresponds to $(\theta, \phi) \rightarrow (\theta + \pi, \pi - \phi)$ in spherical coordinates; hence, only half of the critical points are labeled.

The unbiased MMSE FEQ for tone n is found by setting the correlation of the input and output on tone n equal to the transmitted power on tone n [25] and is thus given by

$$\tilde{d}_n^{\mathrm{uMMSE}} = \frac{\sigma_n^2}{\boldsymbol{\varphi}_n \mathbf{w}}. \tag{27}$$

Substituting the FEQ (27) into the subchannel SNR model (26) yields

$$\mathrm{SNR}_n = \frac{1}{\sigma_n^2 \left(\frac{\mathbf{w}^H \mathbf{C}_n^2 \mathbf{w}}{\mathbf{w}^H \boldsymbol{\varphi}_n^H \boldsymbol{\varphi}_n \mathbf{w}} \right) - 1} \tag{28}$$

$$= \frac{\mathbf{w}^H \mathbf{B}_n \mathbf{w}}{\mathbf{w}^H \mathbf{A}_n \mathbf{w}} \tag{29}$$

with

$$\mathbf{A}_n = \sigma_n^2 \mathbf{C}_n^2 - \boldsymbol{\varphi}_n^H \boldsymbol{\varphi}_n, \tag{30}$$

$$\mathbf{B}_n = \boldsymbol{\varphi}_n^H \boldsymbol{\varphi}_n. \tag{31}$$

In order to compute the SNR, the auto- and cross-correlations can be empirically determined for each of the N_a active subchannels, assuming sufficient training data is available.

In order to visualize the model of the BER of (18) with this subchannel SNR model, consider the channel $\mathbf{h} = [1, -0.3, 0.7]$, with a CP length of $\nu = 1$, so that we wish to shorten the channel to length 2. A 3-tap CSE under a unit norm constraint (which does not affect the BER) can be parameterized by the two angles of spherical coordinates. Fig. 3 shows log-spaced contours of the BER, and Fig. 4 shows a surface plot of the log of the BER. The FFT size was $N = 8$, and the SNR was 20 dB. Observe that the BER is multimodal even for this low-dimensional example. If a gradient descent procedure was used to minimize this cost surface, a significant fraction of the possible initializations would lead to the local minimum rather than the global minimum. For example, initialization

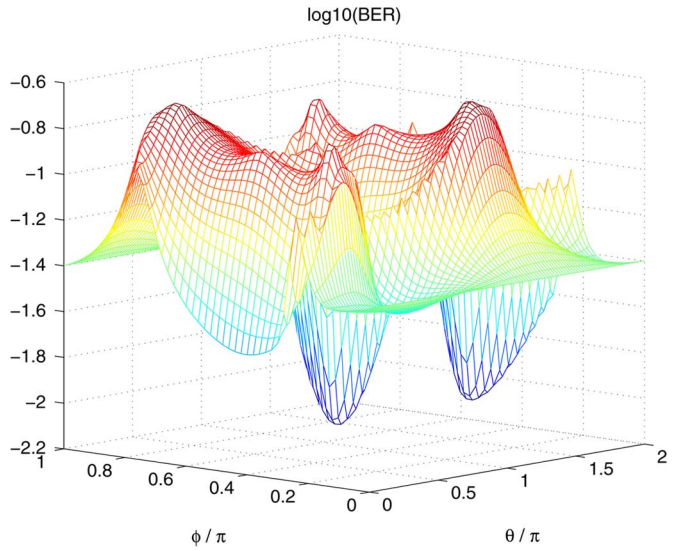


Fig. 4. Surface plot of $\log_{10}(\text{BER})$ for a 3-tap CSE under a unit norm constraint, for a multicarrier system. The CSE is parameterized in spherical coordinates by the angles θ and ϕ. The channel is $\mathbf{h} = [1, -0.3, 0.7]$, the FFT size is $N = 8$, the CP length is $\nu = 1$, and the SNR is 20 dB. Note that the cost function is symmetric with respect to a sign change in the CSE ($\mathbf{w} \rightarrow -\mathbf{w}$), which corresponds to $(\theta, \phi) \rightarrow (\theta + \pi, \pi - \phi)$ in spherical coordinates.

at the center of the figure, which is $[1, 0, 0]$ in Cartesian coordinates, will eventually lead to the global minimum, whereas initialization at the "north pole," which is $[0, 1, 0]$ in Cartesian coordinates, will lead to the local minimum in the top right.

The case of no CSE (or a CSE of $\mathbf{w} = [1, 0, 0]$) is shown as a dot, the popular maximum shortening SNR (MSSNR) CSE [18] is shown as a square, and the CSE that minimizes the BER is shown as a circle. Note that in this case, the MSSNR CSE actually degrades the BER, as evidenced by the fact that it lies on a higher BER contour than the case of no CSE. The authors have found that this relative performance is the same for other channels and other SNR values, although for very high SNR values ($\geq$25 dB) the MSSNR design sometimes performs better than not using a CSE at all.

This is motivation to seek CSE designs that attempt to directly minimize the BER rather than a proxy cost function (such as MSE [13], the shortening SNR [18], or the bit rate used as a metric in DSL systems [25]). Based on the subchannel SNR model discussed earlier in this section and the BER model of (18), the BER can be minimized using numerical methods. This is the subject of Section IV.

C. SCCP BER Model

The SCCP BER will be averaged over the N samples of the final IFFT output, $\hat{\mathbf{x}}_1(k)$. We again assume that the total residual interference and noise on each output sample is Gaussian, and that M-level QAM signalling is used on each tone. The probability of error on the PAM component of sample m is given by [28, pp. 225–226]

$$P_{\sqrt{M}}(m) = 2\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3}{M-1}\text{SNR}_n}\right) \tag{32}$$

hence the SER of sample m is

$$P_M(m) = 2P_{\sqrt{M}}(m) - (P_{\sqrt{M}}(m))^2 \tag{33}$$

where SNR_m is the effective signal-to-interference-and-noise ratio on sample m (which we will refer to as the output SNR); and $Q(x)$ is the Q-function. For the $M = 4$ case, (32) is the BER for sample m, and it reduces to $Q(\sqrt{\text{SNR}_m})$, which we use here for simplicity of notation. Averaging (32) and (33) over the N output samples, the BER and the SER for the output of an SCCP system with $M = 4$ are

$$\text{BER}_{\text{sccp}} = \frac{1}{N}\sum_{m=1}^{N} Q(\sqrt{\text{SNR}_m}) \tag{34}$$

$$\text{SER}_{\text{sccp}} = \frac{1}{N}\sum_{m=1}^{N} (2Q(\sqrt{\text{SNR}_m}) - Q^2(\sqrt{\text{SNR}_m})). \tag{35}$$

Again, either can be used as the objective function to be minimized for an SCCP system, although we focus on the BER. In the next section, we discuss the model of the output SNR in terms of the CSE coefficients.

D. SCCP Output SNR Model

As before, the DFT output can be written as

$$\mathcal{F}\underbrace{(\mathbf{Y}_k\mathbf{w})}_{\mathbf{u}(k)} = \underbrace{(\mathcal{F}\mathbf{Y}_k)}_{\tilde{\mathbf{Y}}_{k,N}}\mathbf{w} \tag{36}$$

where $\mathbf{Y}_k$ is a block Toeplitz matrix of size $N \times (PL_w)$, where each $N \times L_w$ sub-block contains the data that will be convolved with the pth CSE, $\mathbf{w}_p$, and successive rows are vectors $\mathbf{y}^T(i)$ for successive values of i. Then, $\tilde{\mathbf{Y}}_{k,N}$ is an $N \times (PL_w)$ matrix as well, with the nth row denoted by $\tilde{\mathbf{y}}_{k,n}$. Define $\mathbf{Q}_{m,n}$ to be element (m, n) of the (unitary) IDFT matrix, with $\mathbf{Q} = \mathcal{F}^H$. Passing (36) through the FEQ $\tilde{\mathbf{d}}$ and the IDFT and taking output sample m for user $l = 1$ yields

$$\hat{\mathbf{x}}_1(k)[m] = \sum_{n=1}^{N} \mathbf{Q}_{m,n}\tilde{d}_n\tilde{\mathbf{y}}_{k,n}\mathbf{w}. \tag{37}$$

Define the correlation terms

$$\sigma^2 \triangleq \mathcal{E}\{|\tilde{\mathbf{x}}_1^*(k)[m]|^2\} \tag{38}$$

$$\boldsymbol{\varphi}_{m,n} \triangleq \mathcal{E}\{\tilde{\mathbf{x}}_1^*(k)[m]\tilde{\mathbf{y}}_{k,n}\} \tag{39}$$

$$\mathbf{C}_{m,n}^2 \triangleq \mathcal{E}\left\{\tilde{\mathbf{y}}_{k,m}^H\tilde{\mathbf{y}}_{k,n}\right\} \tag{40}$$

which have dimensions 1×1, $1 \times PL_w$, and $PL_w \times PL_w$, respectively. For simplicity, we assume that σ^2 is independent of m, i.e., the N transmit samples have identical power, which is usually the case. The output SNR of sample m is the ratio of the power of the desired signal $\tilde{\mathbf{x}}_1(k)[m]$ to the power of the total interference and noise (i.e., the error signal for block k and

sample m), $\tilde{e}_{k,m} = \hat{\mathbf{x}}_1(k)[m] - \tilde{\mathbf{x}}_1(k)[m]$. Under this model, the output SNR is

$$\begin{aligned}\mathrm{SNR}_m &= \frac{\sigma^2}{\mathcal{E}\{|\tilde{e}_{k,m}|^2\}} \\ &= \frac{\sigma^2}{\mathcal{E}\left\{\left|\sum_{n=1}^{N} \mathbf{Q}_{m,n}\tilde{d}_n\tilde{\mathbf{y}}_{k,n}\mathbf{w} - \tilde{\mathbf{x}}_1(k)[m]\right|^2\right\}}. \end{aligned} \quad (41)$$

The denominator expands as

$$\begin{aligned}\mathcal{E}\{|\tilde{e}_{k,m}|^2\} = &\sum_{n_1,n_2} \mathbf{Q}^*_{m,n_1}\mathbf{Q}_{m,n_2}\tilde{d}^*_{n_1}\tilde{d}_{n_2}\mathbf{w}^H\mathbf{C}^2_{n_1,n_2}\mathbf{w} \\ &-\sum_n \mathbf{Q}_{m,n}\tilde{d}_n\boldsymbol{\varphi}_{m,n}\mathbf{w} \\ &-\sum_n \mathbf{Q}^*_{m,n}\tilde{d}^*_n\boldsymbol{\varphi}^*_{m,n}\mathbf{w}^* + \sigma^2, \quad m \in \{1,\ldots,N\}. \end{aligned} \quad (42)$$

The unbiased MMSE FEQ for sample m is found by setting the correlation of the input and output for sample m equal to the transmitted power for sample m, or equivalently

$$\begin{aligned}&\mathcal{E}\left\{\tilde{\mathbf{x}}^*_1(k)[m]\sum_{n=1}^{N}\mathbf{Q}_{m,n}\tilde{d}_n\tilde{\mathbf{y}}_{k,n}\mathbf{w}\right\} \\ &\quad = \mathcal{E}\{\tilde{\mathbf{x}}^*_1(k)[m]\tilde{\mathbf{x}}_1(k)[m]\} \quad (43) \\ &\sum_{n=1}^{N}\mathbf{Q}_{m,n}\tilde{d}_n\boldsymbol{\varphi}_{m,n}\mathbf{w} = \sigma^2. \quad (44)\end{aligned}$$

With this value of the FEQ, the output SNR becomes

$$\mathrm{SNR}_m = \frac{\sigma^2}{\sum_{n_1,n_2}\mathbf{Q}^*_{m,n_1}\mathbf{Q}_{m,n_2}\tilde{d}^*_{n_1}\tilde{d}_{n_2}\mathbf{w}^H\mathbf{C}^2_{n_1,n_2}\mathbf{w} - \sigma^2}. \quad (45)$$

Equation (44) can be rewritten in matrix form as

$$\tilde{\mathbf{d}}^T\mathbf{A}_m\mathbf{w} = \sigma^2, \quad m \in \{1,\ldots,N\} \quad (46)$$

where row $\mathbf{A}_m[n,:] = \mathbf{Q}_{m,n}\boldsymbol{\varphi}_{m,n}$. Collecting these N equations into a vector and solving for $\tilde{\mathbf{d}}$ yields

$$\tilde{\mathbf{d}} = \sigma^2\begin{bmatrix}\mathbf{w}^T\mathbf{A}_1^T \\ \vdots \\ \mathbf{w}^T\mathbf{A}_N^T\end{bmatrix}^{-1}\begin{bmatrix}1 \\ \vdots \\ 1\end{bmatrix}. \quad (47)$$

Unfortunately, this is as far as we can simplify the output SNR model. However, (34), (45), and (46) together allow us to evaluate the BER for a given CSE $\mathbf{w}$ and an unbiased MMSE FEQ $\tilde{\mathbf{d}}$ based on that CSE.

As in the multicarrier case, we can visualize the model of the BER of (34) with this output SNR model by using a 3-tap unit norm CSE parameterized by the two angles of spherical coordinates. As before, the channel is $\mathbf{h} = [1, -0.3, 0.7]$ with a CP length of $\nu = 1$ so that we wish to shorten the channel to length 2. For each value of the CSE $\mathbf{w}$, we will use (46) to solve for the unbiased MMSE FEQ $\tilde{\mathbf{d}}$. Fig. 5 shows log-spaced contours of the BER, and Fig. 6 shows a surface plot of the log of the BER. The FFT size was $N = 8$, and the SNR was 20 dB. Observe that the BER is extremely multimodal, and numerical optimization of this cost surface is an ambitious goal. Numerical minimization of the SCCP BER model is the subject of Section IV.

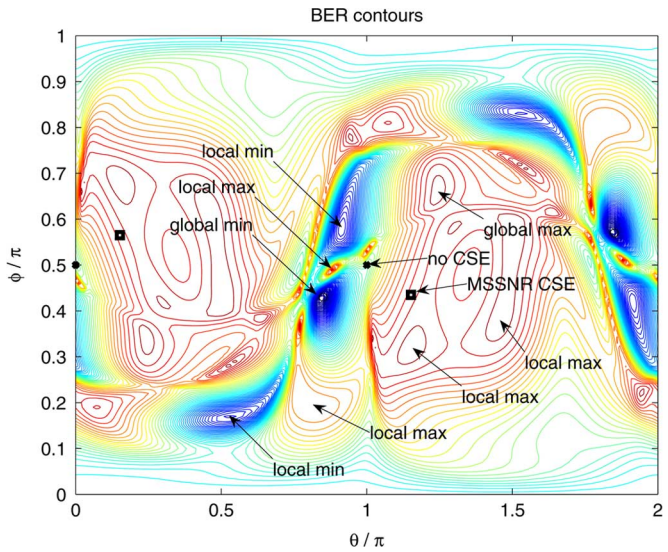


Fig. 5. Contours of the BER for a 3-tap CSE under a unit norm constraint, for an SCCP system. The CSE is parameterized in spherical coordinates by the angles θ and ϕ. The channel is $\mathbf{h} = [1, -0.3, 0.7]$, the FFT size is $N = 8$, the CP length is $\nu = 1$, and the SNR is 20 dB. Note that the cost function is symmetric with respect to a sign change in the CSE ($\mathbf{w} \rightarrow -\mathbf{w}$), which corresponds to $(\theta, \phi) \rightarrow (\theta + \pi, \pi - \phi)$ in spherical coordinates. Only a few of the critical points are labelled, since there are so many.

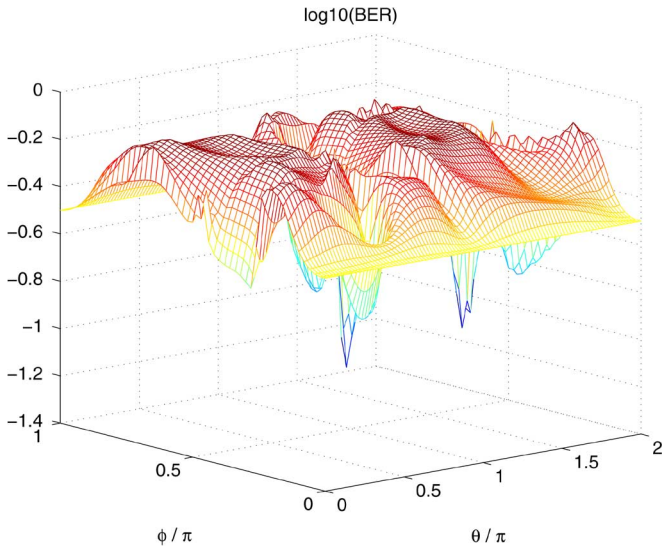


Fig. 6. Surface plot of $\log_{10}(\mathrm{BER})$ for a 3-tap CSE under a unit norm constraint, for an SCCP system. The CSE is parameterized in spherical coordinates by the angles θ and ϕ. The channel is $\mathbf{h} = [1, -0.3, 0.7]$, the FFT size is $N = 8$, the CP length is $\nu = 1$, and the SNR is 20 dB.

IV. BER Minimization Algorithms

In this section, we derive algorithms to minimize the multicarrier and SCCP BER models of Section III. First, we derive an iteratively reweighted (IR) minimum error rate (MER) algorithm for multicarrier systems, and then we derive a Gauss–Newton (GN) update rule for multicarrier systems that has better convergence properties (but is more expensive). The development parallels the development of the IR and GN designs in [25], which attempt to maximize the bit rate for systems that allow bit loading across the tones. Thus, for some of the intermediate mathematical jumps, the reader will be referred to [25]. In principle, since the cost function has been derived explicitly, there are other numerical optimization techniques that could be applied. The IR and GN techniques we use were chosen because the solutions to least-squares and Gauss–Newton updates can be written in closed form and because the simulation results show that they converge reasonable quickly. Finally, since gradient analysis of the SCCP BER model is analytically intractable, we discuss the use of a greedy search as a heuristic approach to minimize the BER.

A. Iteratively Reweighted Update Rule

We wish to minimize the objective function (18) with respect to the parameter vector $\boldsymbol{\theta} = [\mathbf{w}^T, \tilde{\mathbf{d}}^T]^T$, subject to the unbiased FEQ constraint of (27). Since minimization of a function is not affected by multiplying the function by a constant, we will ignore the leading factor of $1/N_a$, i.e., we will minimize $J = N_a \text{BER}_{\text{mcm}}$, to simplify the mathematics. A constrained minimization problem is typically transformed into an unconstrained minimization problem by adding a Lagrangian term to the cost function to be minimized,

$$J_L = \sum_{n \in \mathcal{S}_a} \left[Q(\sqrt{\text{SNR}_n}) - \mathcal{R}\left\{ \lambda_n^* \left(\tilde{d}_n^{\text{uMMSE}} \boldsymbol{\varphi}_n \mathbf{w} - \sigma_n^2 \right) \right\} \right] \tag{48}$$

where the N_a scalars λ_n are the Lagrange multipliers, and the real operator $\mathcal{R}\{\cdot\}$ ensures a real constraint term. The procedure we will use to derive the IR update rule is as follows.

1) For $1 \le n \le N_a$, set the gradient of (48) with respect to $\tilde{d}_n$ equal to zero, $\nabla_{\tilde{d}_n} J_L = 0$, and use the resulting equations along with the MMSE constraint on $\tilde{d}_n$ to solve for the Lagrange multipliers λ_n.
2) Compute the gradient of (48) with respect to the full parameter vector $\boldsymbol{\theta}$, then substitute in the values of the Lagrange multipliers from step 1).
3) Set the gradient resulting from step 2) equal to zero, which gives an equation that will be satisfied by the BER minimizing parameter settings.
4) Replace the expectation in the gradient with a time average, and replace the equality with a least-squares minimization.
5) Iteratively solve this minimization problem in the manner of Section II-D of [25].

For step 1), the element of the gradient from the nth FEQ parameter is

$$\nabla_{\tilde{d}_n} J_L = \underbrace{\frac{1}{\sqrt{2\pi}} \frac{e^{(-\text{SNR}_n)/2}}{\sqrt{\text{SNR}_n}} \cdot \frac{\text{SNR}_n^2}{\sigma_n^2}}_{\gamma_n} \times \mathcal{E}\{\tilde{\mathbf{y}}_{k,n}^* \mathbf{w}^* \tilde{e}_{k,n}\} - \lambda_n \mathcal{E}\{\tilde{\mathbf{y}}_{k,n}^* \mathbf{w}^* \tilde{\mathbf{x}}_1(k)[n]\} \tag{49}$$

which equals zero at the location of a minimum. Note that this derivative is very similar to the corresponding (66) in [25], with the only difference being how γ_n is defined. With this redefinition of γ_n, the Lagrange multipliers have the same form as in [25]

$$\lambda_n = \frac{\gamma_n}{\text{SNR}_n}. \tag{50}$$

For step 2), the remaining component of the gradient, from the $\mathbf{w}$ portion of $\boldsymbol{\theta}$, has the form

$$\nabla_{\mathbf{w}} J_L = \sum_{n \in \mathcal{S}_a} \gamma_n \tilde{d}_n^{\text{uMMSE}*} \mathcal{E}\left\{ \tilde{\mathbf{y}}_{k,n}^H \tilde{e}_{k,n} \right\} - \lambda_n \tilde{d}_n^{\text{uMMSE}*} \mathcal{E}\left\{ \tilde{\mathbf{y}}_{k,n}^H \tilde{\mathbf{x}}_1(k)[n] \right\}. \tag{51}$$

Stacking (51) and (49) for all $n \in \mathcal{S}_a$, the gradient of (18) is given by

$$\nabla_{\boldsymbol{\theta}} J_L = \mathcal{E}\left\{ \begin{bmatrix} \sum_n \tilde{\mathbf{y}}_{k,n}^H \tilde{d}_n^{\text{uMMSE}*} [\gamma_n \tilde{e}_{k,n} - \lambda_n \tilde{\mathbf{x}}_1(k)[n]] \\ \tilde{\mathbf{y}}_{k,1}^* \mathbf{w}^* [\gamma_1 \tilde{e}_{k,1} - \lambda_1 \tilde{\mathbf{x}}_1(k)[1]] \\ \vdots \\ \tilde{\mathbf{y}}_{k,N_a}^* \mathbf{w}^* [\gamma_{N_a} \tilde{e}_{k,N_a} - \lambda_{N_a} \tilde{\mathbf{x}}_1(k)[N_a]] \end{bmatrix} \right\}. \tag{52}$$

Substituting in the value of λ_n from (50) and simplifying

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} J_L &= \mathcal{E}\left\{ \begin{bmatrix} \check{\mathbf{Y}}_k^H \text{diag}\left(\tilde{d}^{\text{uMMSE}} \right)^* \\ \text{diag}\left(\tilde{\mathbf{Y}}_k^H \mathbf{w} \right)^* \end{bmatrix} \text{diag}(\boldsymbol{\gamma}_k) \right. \\ &\quad \left. \times \left[\text{diag}\left(\tilde{d}^{\text{uMMSE}} \right) \tilde{\mathbf{Y}}_k^H \mathbf{w} - \text{diag}(\boldsymbol{\rho}^{-2}) \tilde{\mathbf{x}}_1(k) \right] \right\} \\ &= \mathcal{E}\left\{ \check{\mathbf{Y}}_k^H \text{diag}\left(\check{\boldsymbol{\gamma}}_k \right) \tilde{\mathbf{e}}_k \right\} \end{aligned} \tag{53}$$

where $\boldsymbol{\rho}^{-2}$ is a vector with nth element

$$\rho_n^{-2} \triangleq \frac{\sigma_n^2 \mathbf{w}^H \mathbf{C}_n^2 \mathbf{w}}{\mathbf{w}^H \boldsymbol{\varphi}_n^H \boldsymbol{\varphi}_n \mathbf{w}}. \tag{54}$$

Using this formulation, the rest of the algorithm development parallels that in [25], hence we simply outline the resulting algorithm in Fig. 7.

Remarks: If one wishes to minimize the SER of (19) instead of the BER of (18), then the only change in the algorithm of Fig. 7 is that the weights γ_n have an additional factor of $(1 - Q(\sqrt{\text{SNR}_n}))$. This is a minor difference, especially in the vicinity of the optimum solution, where SNR_n should be large. If one wishes to minimize the BER for an M-level QAM constellation with $M > 4$, then the weights must be modified to be

$$\gamma_n^{M-\text{QAM}} = \alpha\sqrt{\beta} \frac{1}{\sqrt{2\pi}} \frac{e^{(-\beta \text{SNR}_n)/2}}{\sqrt{\text{SNR}_n}} \cdot \frac{\text{SNR}_n^2}{\sigma_n^2} \tag{55}$$

where $\alpha = 2(1 - 1/\sqrt{M})$ and $\beta = 3/(M-1)$.

After estimating the correlation terms σ_n^2, $\boldsymbol{\varphi}_n$, and $\mathbf{C}_n^2$ in (22), (23), and (24) from time averages, loop through the following three steps:

1. Estimate the weights $\check{\gamma}_n$ based on the previous value of the CSE, $\mathbf{w}_{prev}$:
   $\rho^{-2} = \frac{\sigma_n^2 \mathbf{w}^H \mathbf{C}_n^2 \mathbf{w}}{\mathbf{w}^H \boldsymbol{\varphi}_n^H \boldsymbol{\varphi}_n \mathbf{w}}$
   $\mathrm{SNR}_n = \frac{1}{\rho^{-2}-1}$
   $\check{\gamma}_n = \frac{(1+\mathrm{SNR}_n)^2}{\mathrm{SNR}_n^2}\gamma_n = \frac{e^{-(\mathrm{SNR}_n)/2}}{\sqrt{2\pi}\sqrt{\mathrm{SNR}_n}} \cdot \frac{(1+\mathrm{SNR}_n)^2}{\sigma_n^2}$
2. Compute the LS estimate of the biased MMSE FEQ based on the updated weights:
   $\widetilde{d}_n^{MMSE} = \frac{\mathbf{w}^H \boldsymbol{\varphi}_n^H}{\mathbf{w}^H \mathbf{C}_n^2 \mathbf{w}}$
3. Compute the LS estimate of the new CSE based on the new FEQ:
   $\mathbf{w} = \left\{\sum_{n\in\mathcal{S}_a} \check{\gamma}_n \left|\widetilde{d}_n\right|^2 \mathbf{C}_n^2\right\}^{-1} \left\{\sum_{n\in\mathcal{S}_a} \check{\gamma}_n \widetilde{d}_n^* \boldsymbol{\varphi}_n^H\right\}$

Fig. 7. Iteratively reweighted minimum error rate (IR-MER) CSE design algorithm.

B. Gauss–Newton Update Rule

As with the IR bit-rate-maximizing design in [25], the IR-MER CSE converges only linearly, and this can be improved by using an iterative GN algorithm. The GN algorithm will have a higher complexity per iteration than the IR algorithm, but as will be seen in Section V, the GN algorithm is more robust and has better performance than the IR algorithm. The GN algorithm is essentially a gradient descent algorithm, but the direction of the update vector is adjusted by a factor of the inverse of the Hessian matrix of the cost function, as follows:

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \mu_i \left(\nabla_{\boldsymbol{\theta}}^2 J\right)^{-1} \nabla_{\boldsymbol{\theta}} J\Big|_{\boldsymbol{\theta}_i}. \tag{56}$$

The gradient is given in (53); thus, the stochastic gradient and the approximate Hessian at the current iteration are given by

$$\nabla_{\boldsymbol{\theta}} J_L = \sum_k \check{\mathbf{Y}}_k^H \mathrm{diag}(\check{\boldsymbol{\gamma}}_k)\tilde{\mathbf{e}}_k \tag{57}$$

$$\nabla_{\boldsymbol{\theta}}^2 J_L = \sum_k \check{\mathbf{Y}}_k^H \mathrm{diag}(\check{\boldsymbol{\gamma}}_k)\check{\mathbf{Y}}_k. \tag{58}$$

The Hessian is approximate because we have neglected the dependence of $\check{\boldsymbol{\gamma}}_k$ on $\boldsymbol{\theta}$ (as was done in [25]), which greatly simplifies the analysis and computation of the Hessian. The difference between (58) and the similar-looking Hessian in [25] lies in the redefinition of the weights $\check{\boldsymbol{\gamma}}_k$. Thus, an efficient implementation of (56) can be found by paralleling the derivation (46) to (53) in [25]. The resulting algorithm is summarized in Fig. 8.

Remarks: As in the IR-MER algorithm, if one wishes to minimize the SER of (35) instead of the BER of (34), then the only change in the algorithm of Fig. 8 is that the weights γ_n have an additional factor of $(1 - Q(\sqrt{\mathrm{SNR}_n}))$. This is a minor difference, especially in the vicinity of the optimum solution, where SNR_n should be large. If one wishes to minimize the BER for an M-level QAM constellation with $M > 4$, then the weights must be modified as in (55).

After estimating the correlation terms σ_n^2, $\boldsymbol{\varphi}_n$, and $\mathbf{C}_n^2$ in (22), (23), and (24) from time averages, loop through the following steps:

1. Same as Fig. 7 (estimate $\check{\gamma}_n$)
2. Same as Fig. 7 (compute $\widetilde{\mathbf{d}}$)
3. Compute the Hessian, $\mathbf{Q}_{\boldsymbol{\theta}}$:
   $\mathbf{E}_{\widetilde{\mathbf{d}}} = \sum_{n\in\mathcal{S}_a} \check{\gamma}_n \left|\widetilde{d}_n\right|^2 \mathbf{C}_n^2$
   $\mathbf{F}_{\boldsymbol{\theta}}^H = \begin{bmatrix} \vdots \\ \left[\check{\gamma}_n \widetilde{d}_n \left(\mathbf{w}^H \mathbf{C}_n^2\right)\right] \\ \vdots \end{bmatrix}$
   $\mathbf{G}_{\mathbf{w},D} = \mathrm{diag}\left[\cdots \check{\gamma}_n \left(\mathbf{w}^H \mathbf{C}_n^2 \mathbf{w}\right) \cdots\right]$
   $\mathbf{Q}_{\boldsymbol{\theta}} = \mathbf{E}_{\widetilde{\mathbf{d}}} - \mathbf{F}_{\boldsymbol{\theta}} \mathbf{G}_{\mathbf{w},D}^{-1} \mathbf{F}_{\boldsymbol{\theta}}^H$
4. Compute the gradient:
   $\nabla_{\mathbf{w}} J = \sum_{n\in\mathcal{S}_a} \check{\gamma}_n \widetilde{d}_n^* \left(\widetilde{d}_n \mathbf{C}_n^2 \mathbf{w} - \boldsymbol{\varphi}_n^H\right)$
5. For a small step size μ, compute the new CSE:
   $\mathbf{w}_{new} = \mathbf{w}_{old} - \mu\, \mathbf{Q}_{\boldsymbol{\theta}}^{-1} (\nabla_{\mathbf{w}} J)$
6. Renormalize the CSE:
   $\mathbf{w}_{new} = \mathbf{w}_{new} / \|\mathbf{w}_{new}\|$

Fig. 8. Approximate Gauss–Newton minimum error rate (GN-MER) CSE design algorithm.

Estimate the correlation terms σ_n^2, $\boldsymbol{\varphi}_n$, and $\mathbf{C}_n^2$ in (38), (39), and (40) from time averages. Choose a neighborhood size σ_{step} and a unit norm initial guess $\mathbf{w}_{best}$, and loop through the following steps:

1. Generate $\mathbf{w}_{step}$ as a complex Gaussian random vector, i.i.d. with zero mean and variance $\frac{1}{2}\sigma_{step}^2$ on each of the real and imaginary components
2. $\mathbf{w}_{trial} = \frac{\mathbf{w}_{best} + \mathbf{w}_{step}}{\|\mathbf{w}_{best} + \mathbf{w}_{step}\|}$
3. Use (34), (45), and (46) to compute BER_{trial}
4. If $BER_{trial} < BER_{best}$, then $\mathbf{w}_{best} = \mathbf{w}_{trial}$ and $BER_{best} = BER_{trial}$

Stop at a given number of iterations or when a target BER is met.

Fig. 9. Greedy minimum error rate (G-MER) CSE design algorithm.

C. BER Minimization for SCCP Systems

Since the BER is intractable to direct minimization, a heuristic approach is needed. The approach we take here is a greedy search, in which at each iteration, we search in some neighborhood of the current best solution, and if the new solution has a lower BER, we accept it. The algorithm is summarized in Fig. 9. Note that if one wishes to minimize the BER for an M-level QAM constellation with $M > 4$, then in step 3), the use of (34) must be modified to include the additional constants inside and outside of the Q-function as shown in (32).

The BER model is invariant to scale factors in the CSE, hence the unit norm constraint is used for convenience. If the step size is $\sigma_{\mathrm{step}} = \alpha(L_w)^{-1/2}$, then the average step size becomes $\sqrt{\mathcal{E}\{\mathbf{w}_{\mathrm{step}}^H \mathbf{w}_{\mathrm{step}}\}} = \alpha$. For our simulations, we use $\alpha = 0.01$, i.e., each step has a magnitude of about 1% of the magnitude of the current CSE. The initialization for $\mathbf{w}_{\mathrm{best}}$ should be a

cheap-to-compute CSE that has the best performance of all such designs so that there is a reasonable chance of starting in the valley of the global minimum of the BER. In the simulations, we will use the minimum interblock interference (min-IBI) design [30] for initialization, although other choices are valid.

There are two drawbacks to the greedy search: 1) it requires computation of the BER model at each iteration, which is very expensive, and 2) as with a gradient descent method, the global minimum is only achieved if the initialization lies somewhere in the valley of the global minimum. In order to not become trapped in a local minimum, the greedy search can be generalized to simulated annealing [31]. In short, simulated annealing occasionally allows upwards steps, but the probability of allowing upwards steps decreases according to a user-defined "cooling schedule." It is known that under certain conditions (including infinite run time and an infinitesimally slow cooling schedule), simulated annealing will find the global minimum of the cost function. However, this further adds to the complexity, since a large number of iterations is required.

Computational complexity can be assessed in terms of the number of complex multiply-and-accumulate (MAC) operations that are required. If N_{corr} symbols are used to compute the correlation terms, then both the IR-MER and GN-MER algorithms require $(1/2)(PL_w)^2N_aN_{\text{corr}}$ complex MACs to compute the correlation terms. The IR-MER algorithm requires a further $(3/2)(PL_w)^2N_a+(1/3)(PL_w)^3$ complex MACs per iteration, whereas the GN-MER algorithm requires a further $(PL_w)^2N_a + 4(PL_w)N_a + (1/3)(PL_w)^3$ complex MACs per iteration. In sharp contrast, the greedy search requires $(1/2)(PL_w)^2N^2N_{\text{corr}}$ complex MACs to compute the correlation terms, and a further $(PL_w)^2N^3$ complex MACs per iteration. This computational complexity is the drawback of the proposed technique. Existing heuristic channel shorteners designed for DSL typically have much lower complexity. The MSSNR design, for example, has a computational complexity of approximately $(11/3)(PL_w)^3 + (L_h + L_w)(PL_w)^2$, when only a single value of the delay Δ is considered (as assumed in the rest of this paper). That means that for the representative parameters used in the simulations, the proposed techniques require $\mathcal{O}(10^7)$ MACs, whereas the MSSNR design only requires $\mathcal{O}(10^5)$ MACs to implement. See [32] for a detailed survey of the computational complexity of channel shorteners designed for DSL.

V. Simulations

In this section, we simulate the proposed IR-MER CSE of Fig. 7 and compare it to several designs in the literature. Performance will be assessed in terms of the BER. The algorithms compared throughout this section are the MSSNR design [18], the MMSE design [13], the minimum delay spread (MDS) design [33], the min-IBI design [30], the proposed IR-MER design and the proposed GN-MER design, and the GN implementation of the bit-rate-maximizing design (GN-BM) [25]. We also compare to the matched filter bound (MFB) and to the BER when no CSE is used. The plots include BER versus SNR, BER versus delay (averaged over many channels, and for a single channel), and BER versus algorithm iteration (for two different SNR values), all for a multicarrier system; and BER versus SNR for an SCCP system.

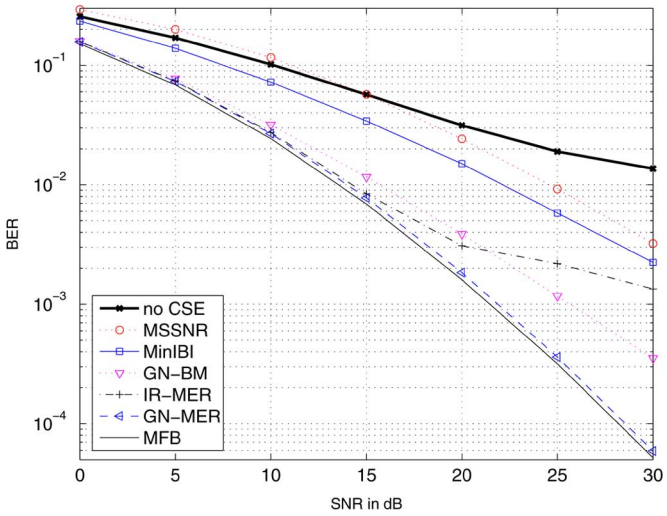


Fig. 10. Bit error rate versus SNR when no CSE is used, when the MSSNR CSE is used [18], when the proposed IR-MER CSE and the proposed GN-MER CSE are each used, and when the GN-BM design of [25] is used. There were two receive antennas, the channels were Rayleigh fading with 32 taps each, each CSE had length 16, there were 52 active tones and 12 null tones, and the CP length was 16. All of the iterative algorithms used 40 iterations.

Consider a wireless system with FFT size $N = 64$, CP length $\nu = 16$, and $N_a = 52$ active tones (hence, 12 null tones), as in the IEEE 802.11a, HIPERLAN/2, and MMAC wireless LAN standards. The active tones all use 4-QAM signal constellations. Let the channels be Rayleigh fading with approximately 32 significant taps, using the approximately exponential delay profile described in [34]. There are $L = 1$ transmit antenna and $P = 2$ receive antennas. The CSE has $L_w = 16$ taps per antenna. The correlation parameters σ_n^2, $\boldsymbol{\varphi}_n$, and $\mathbf{C}_n^2$ will be estimated using 2000 symbols of training, and the iterative algorithms will be run for 40 iterations to ensure convergence. The results will be averaged over 500 independently generated channels, input data sequences, and noise sequences. The BER is measured over 1000 symbols for each channel realization (generated independently of those used to estimate the correlation terms), except at 25-dB SNR we used 2000 and at 30-dB SNR we used 4000. In all cases, the MSSNR design will be used as the initial setting for the iterative algorithms, and the desired delay will be chosen heuristically (similar to the method of [34]) rather than performing a global search. No channel coding has been used, since the use of linear channel coding involves an additional matrix operation on the output, making the cost function for the multicarrier case as intractable as for the SCCP case.

The MFB is a bound on the SNR which assumes that the equalizer captures all of the signal energy, including ISI, and uses it for detection purposes. In traditional single-carrier systems, the MFB is the received SNR with the entire received signal power (including ISI) used as the numerator. For our purposes, the MFB was calculated by inserting the traditional single carrier MFB on the SNR for each subcarrier into (18).

Fig. 10 shows the BER versus SNR in decibels. Below 15-dB SNR, there is no benefit to using an MSSNR CSE compared with no CSE, but the proposed design improves the SER by a

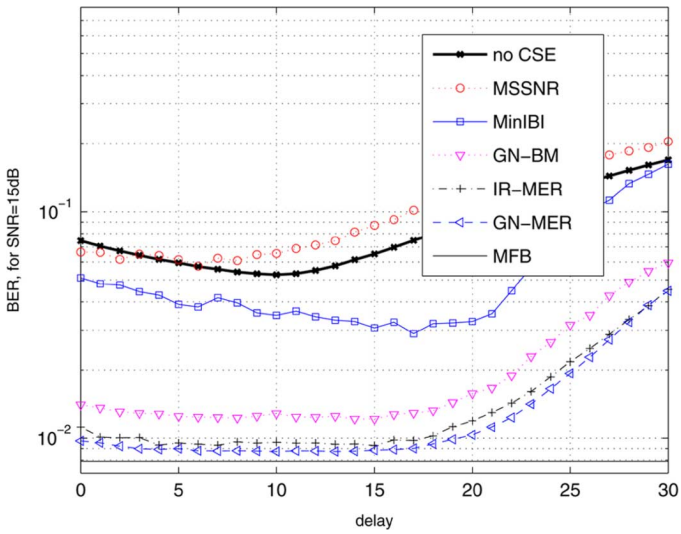


Fig. 11. BER versus delay, averaged over 50 realizations of the channel, input, and noise sequences. The SNR is 15 dB. The *average* performance of the various designs is not that sensitive to the delay choice.

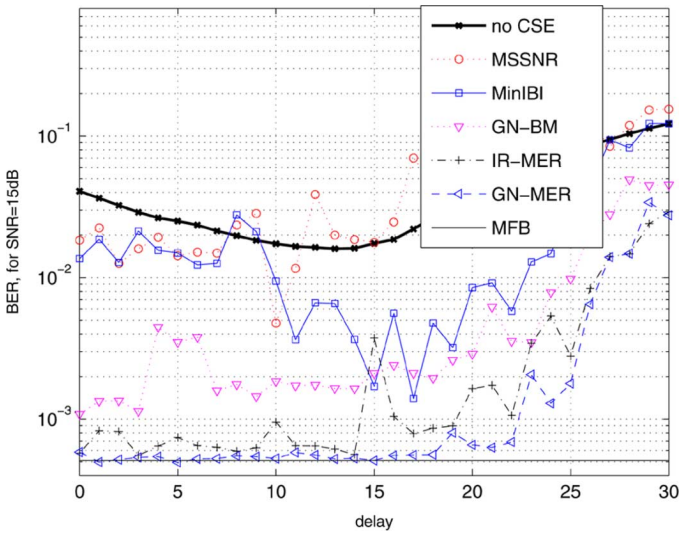


Fig. 12. BER versus delay, for a single realization of the channel, input, and noise sequences. The SNR is 15 dB. Note that for a given channel, the performance of the CSE may vary somewhat with the choice of delay, and the proposed GN-MER CSE shows the least sensitivity to slight changes in the delay choice.

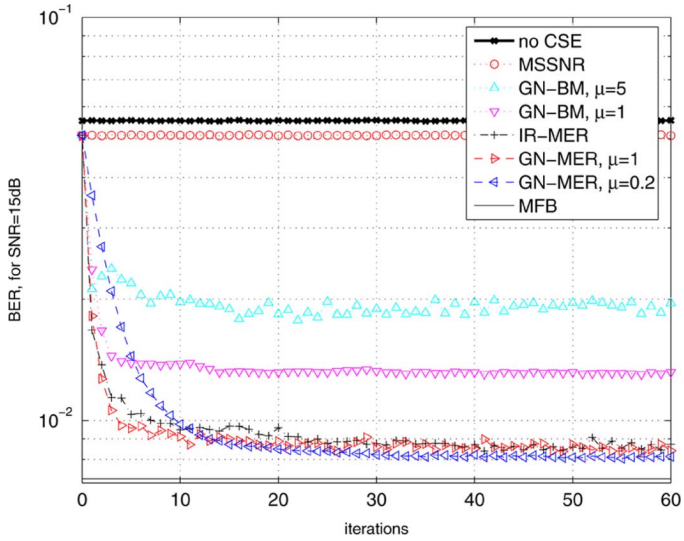


Fig. 13. BER versus iteration number, averaged over 50 realizations of the channel, input, and noise sequences. The SNR is 15 dB. Higher step sizes lead to faster convergence but worse asymptotic performance. Convergence takes place largely within the first 10 to 20 iterations.

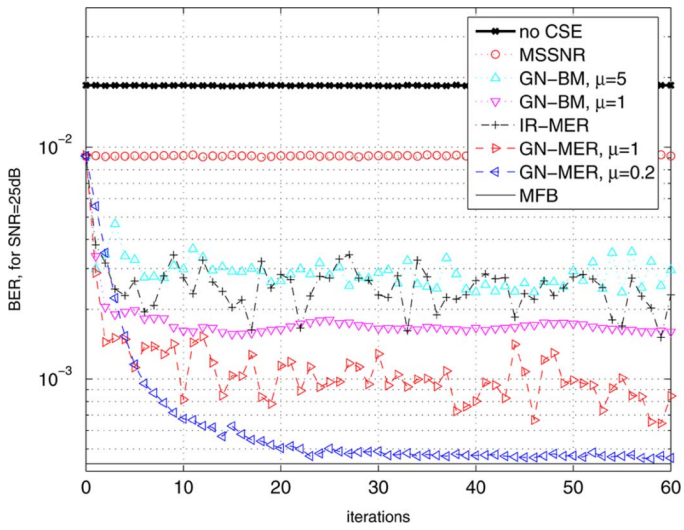


Fig. 14. BER versus iteration number, averaged over 50 realizations of the channel, input, and noise sequences. The SNR is 25 dB. The trends are largely the same as in Fig. 13.

factor of 2 to 200, for low (0 db) to high (30 dB) SNR values, respectively. Note that the BER of the IR-MER algorithm does decrease with increasing SNR, but it does not decrease as rapidly as the GN algorithms. Thus, the IR algorithm appears best suited to situations of low SNR (below 15 dB in this experiment), whereas the more computationally intensive GN-MER algorithm is better for higher SNR values. Also note that the min-IBI design and the GN-BM design exhibit intermediate performance, better than no CSE but not as good as the proposed design.

Figs. 11 and 12 show the BER versus desired delay for an SNR of 15 dB. In Fig. 11, the BER was averaged over all 50 realizations of the channel, input sequence, and noise sequence, whereas Fig. 12 shows the results for a single realization. The feature to note is that although the *average* performance of each algorithm does not vary much with small changes in delay, for an *individual* channel, some of the algorithms are very sensitive to the delay choice. However, the proposed GN-MER algorithm exhibits the least sensitivity, so the delay can be chosen via a rough estimate rather than a global search, without fear of a significant performance penalty.

Figs. 13 and 14 show learning curves of the BER versus iteration number for the iterative algorithms, using SNR values of 15 and 25 dB, respectively. The first salient feature to note is that the majority of the convergence occurs within 10 to 20 iterations in all cases. The second feature to note is that increasing the step size μ of the GN algorithms leads to faster convergence, but asymptotic performance suffers due to misadjustment of the filter taps. In principle, an optimal step size could be found at

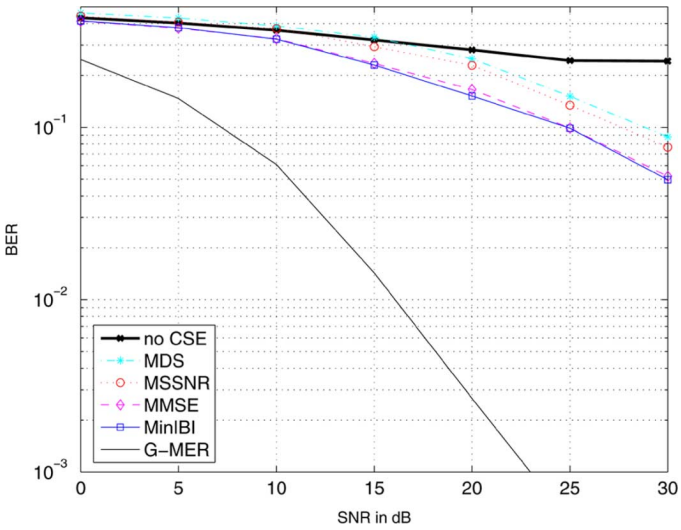


Fig. 15. Bit error rate versus SNR for an SCCP system, when no CSE is used, and when the MMSE [13], MSSNR [18], MDS [33], and min-IBI [30] designs are used. There were two receive antennas, the channels were Rayleigh fading with ten taps each, each CSE had length 16, the FFT size was 16, and the CP length was 4.

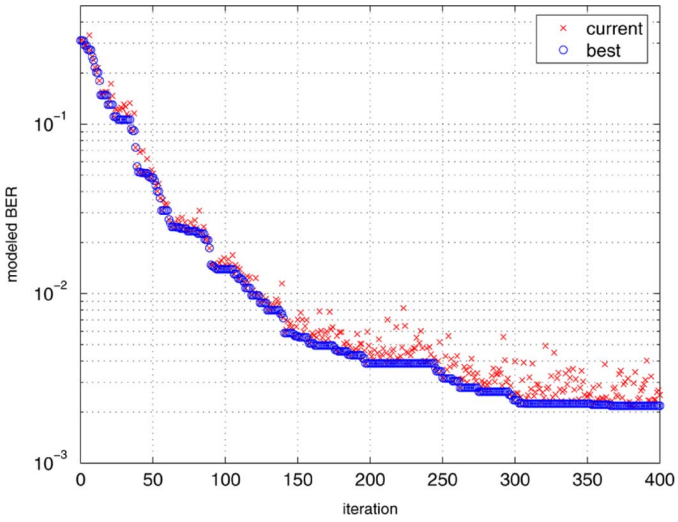


Fig. 16. History of the BER model for one run of the G-MER algorithm for an SCCP system. The SNR was 15 dB.

each iteration using a line search [35], although that would increase the complexity per update. The fact that the GN-BM and GN-MER algorithms have different step sizes is irrelevant, because their cost functions are markedly different, which makes the magnitudes of the update terms different.

Now consider an SCCP system with FFT size $N = 16$, CP length $\nu = 4$; these parameter sizes are chosen so small since the output SNR model of (45) is so expensive to compute. The data is drawn from a 4-QAM signal constellation. Let the channels be Rayleigh fading with approximately ten significant taps, using the approximately exponential delay profile described in [34]. There are $L = 1$ transmit antenna and $P = 2$ receive antennas. The CSE has $L_w = 16$ taps per antenna. For the greedy search, the correlation parameters σ_n^2, $\boldsymbol{\varphi}_n$, and $\mathbf{C}_n^2$ will be estimated using 100 blocks of training, and the min-IBI design will be used as the initialization. The BER values will be measured over 250 independently generated channels, input data sequences, and noise sequences, using 2000 blocks of data each. In all cases, the desired delay will be chosen heuristically (similar to the method of [34]) rather than performing a global search.

Fig. 15 shows the measured BER versus SNR in decibels, for this SCCP system; and Fig. 16 shows the calculated BER versus iteration number for the greedy search. Note that the BER model of (34), (45), and (46) is only used to perform the greedy search, and the actual BER assessment in Fig. 15 uses the actual measured BER, not the model. Aside from the greedy search, the channel shorteners considered (MDS, MSSNR, MMSE, and min-IBI) are the only ones that the authors are aware of that do not explicitly take into account the multicarrier signal structure, hence they are the only ones that can be directly applied to the SCCP system for comparison. The design that performs the best by far is the greedy search (G-MER). However, it is too computationally intensive for real-time implementation; of the remaining designs, the min-IBI design performs the best. Clearly, there is a need for a new design that is computationally cheaper than the greedy search but performs better than the min-IBI design.

VI. Conclusion

A channel shortening equalizer can be used to mitigate multipath effects in wireless cyclic prefixed communication systems. Previous channel shorteners were designed for heuristic objective functions and/or objective functions targeted at wireline systems. We have investigated the use of the BER as an objective function, examined it as a function of the channel shortener coefficients, and proposed several iterative algorithms for optimizing the BER for multicarrier and SCCP systems. Simulations show a significant improvement over existing designs.

References

[1] R. D. J. van Nee, G. A. Awater, M. Morikura, H. Takanashi, M. A. Webster, and K. W. Halford, "New high-rate wireless LAN standards," *IEEE Commun Mag*, vol. 37, no. 12, pp. 82–88, Dec. 1999.

[2] *Air Interface For Fixed Broadband Wireless Access Systems, MAC and Additional PHY Specifications For 2–11 GHz*, IEEE Std. 802.16a, 2003.

[3] *Digital Video Broadcast. (DVB); Framing Structure, Channel Coding and Modulation For Digital Terrestrial Telev.*, ETSI EN 300 744 V1.4.1, The European Telecomm. Standards Inst., 2001.

[4] *Radio Broadcast. System, Digital Audio Broadcast. (DAB) to Mobile, Portible, and Fixed Receivers*, ETSI 300 401, The European Telecomm. Standards Inst., 1995–1997.

[5] D. H. Layer, "Digital radio takes to the road," *IEEE Spectrum*, vol. 38, no. 7, pp. 40–46, Jul. 2001.

[6] S. Galli, A. Scaglione, and K. Dostert, "Broadband is power: Internet access through the power line network," *IEEE Commun. Mag.*, vol. 41, no. 5, pp. 82–83, May 2003.

[7] T. Starr, J. Cioffi, and P. Silverman, *Understanding Digital Subscriber Line Technology*. Upper Saddle River, NJ: Prentice-Hall, 1999.

[8] D. D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.

[9] H. Sari, G. Karam, and I. Jeanclaude, "Frequency-domain equalization of mobile radio and terrestrial broadcast channels," in *Proc. IEEE Global Communications Conf. (IEEE GLOBECOM)'94*, San Fransisco, CA, Nov. 1994, pp. 1–5.

[10] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, Feb. 1995.

[11] R. K. Martin, K. Vanbleu, M. Ding, G. Ysebaert, M. Milosevic, B. L. Evans, M. Moonen, and C. R. Johnson, Jr., "Unification and evaluation of equalization structures and design algorithms for discrete multitone modulation systems," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3880–3894, Oct. 2005.
[12] S. Qureshi and E. Newhall, "Adaptive receiver for data transmission over time-dispersive channels," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 448–457, Jul. 1973.
[13] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell Sys. Tech. J.*, vol. 52, pp. 1541–1562, Nov. 1973.
[14] W. Lee and F. Hill, "A maximum-likelihood sequence estimator with decision-feedback equalization," *IEEE Trans. Commun.*, vol. C-25, no. 9, pp. 971–979, Sep. 1977.
[15] J. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "Equalizer training algorithms for multicarrier modulation systems," in *Proc. IEEE Int. Conf. Communications*, Geneva, Switzerland, May 1993, pp. 761–765.
[16] N. Al-Dhahir and J. M. Cioffi, "Optimum finite-length equalization for multicarrier transceivers," *IEEE Trans. Commun.*, vol. 44, no. 1, pp. 56–64, Jan. 1996.
[17] N. Al-Dhahir and J. M. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 903–915, May 1996.
[18] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Commun.*, vol. 44, no. 12, pp. 1662–1672, Dec. 1996.
[19] I. Medvedev and V. Tarokh, "A channel-shortening multiuser detector for DS-CDMA systems," in *Proc. 53rd Vehicular Technology Conf.*, Rhodes, Greece, May 2001, vol. 3, pp. 1834–1838.
[20] S. I. Husain and J. Choi, "Single correlator based UWB receiver implemetation through channel shortening equalizer," in *Proc. 2005 Asia-Pacific Conf. Communications*, Perth, Western Australia, Oct. 2005, pp. 610–614.
[21] M. Kallinger and A. Mertins, "Room impulse response shortening by channel shortening concepts," in *Conf. Rec. 39th Asilomar Conf. Signals, Syst. Computers*, Pacific Grove, CA, CA, Nov. 2005, pp. 898–902.
[22] G. Arslan, B. L. Evans, and S. Kiaei, "Equalization for discrete multitone transceivers to maximize bit rate," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3123–3135, Dec. 2001.
[23] M. Milosevic, L. F. C. Pessoa, B. L. Evans, and R. Baldick, "DMT bit rate maximization with optimal time domain equalizer filter bank architecture," in *Proc. IEEE Asilomar Conf. Signals, Systems, Comp.*, Pacific Grove, CA, Nov. 2002, vol. 1, pp. 377–382.
[24] K. Vanbleu, G. Ysebaert, G. Cuypers, M. Moonen, and K. Van Acker, "Bitrate-maximizing time-domain equalizer design for dmt-based systems," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 871–876, Jun 2004.
[25] K. Vanbleu, G. Ysebaert, G. Cuypers, and M. Moonen, "On time-domain and frequency-domain mmse-based TEQ design for DMT transmission," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3311–3324, Aug. 2005.
[26] K. Van Acker, G. Leus, M. Moonen, O. van de Wiel, and T. Pollet, "Per tone equalization for DMT-based systems," *IEEE Trans. Commun.*, vol. 49, no. 1, pp. 109–119, Jan. 2001.
[27] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
[28] G. L. Stuber, *Principles of Mobile Communication*. Norwell, MA: Kluwer, 1996.
[29] W. Henkel and T. Kessler, "Maximizing the channel capacity of multicarrier transmission by suitable adaptation of the time-domain equalizer," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2000–2004, Dec. 2000.
[30] S. Celebi, "Interblock interference (IBI) minimizing time-domain equalizer (TEQ) for OFDM," *IEEE Signal Process. Lett.*, vol. 10, no. 8, pp. 232–234, Aug. 2003.
[31] C. R. Reeves, *Modern Heuristic Techniques For Combinatorial Problems*. London, U.K.: McGraw-Hill, 1995.
[32] R. K. Martin, K. Vanbleu, M. Ding, G. Ysebaert, M. Milosevic, B. L. Evans, M. Moonen, and C. R. Johnson, Jr., "Implementation complexity and communication performance tradeoffs in discrete multitone modulation equalizers," *IEEE Trans. Signal Process.*, vol. 54, no. 8, pp. 3216–3230, Aug. 2006.
[33] R. Schur and J. Speidel, "An efficient equalization method to minimize delay spread in OFDM/DMT systems," in *Proc. IEEE Int. Conf. Communications*, Helsinki, Finland, Jun. 2001, vol. 5, pp. 1481–1485.
[34] R. K. Martin, J. M. Walsh, and C. R. Johnson, Jr., "Low-complexity MIMO blind, adaptive channel shortening," *IEEE Trans. Signal Process.*, vol. 53, no. 4, pp. 1324–1334, Apr. 2005.
[35] A. Bjorck, *Least-Squares Methods*. New York, NY: Elsevier, 1987.



Richard K. Martin (M'04) received dual B.S. degrees (*summa cum laude*) in physics and electrical engineering from the University of Maryland, College Park, in 1999 and the M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 2001 and 2004, respectively.

Since August 2004, he has been an Assistant Professor at the Air Force Institute of Technology (AFIT), Dayton, OH, where he is the Signal Processing Curriculum Chair and the Graduate Electrical Engineering Program Chair. His research interests include equalization for multicarrier and single-carrier cyclic-prefixed systems; blind, adaptive filters; sparse adaptive filters; reduced complexity equalizer design; and automatic modulation recognition. He has authored 12 journal papers, 20 conference papers, and the book *Theory and Design of Adaptive Filters Answer Book* (Upper Saddle River, NJ: Prentice-Hall, 2002); and he has three patents.

Dr. Martin was twice elected "Instructor of the Quarter" for the Electrical and Computer Engineering Department by the AFIT Student Association.



Geert Ysebaert (S'97–A'01) was born in Leuven, Belgium, in 1976. He received the Master degree and Ph.D. degree in electrical engineering from the Katholieke Universiteit Leuven (KULeuven), Leuven, Belgium, in 1999 and 2004, respectively. From 1999 to 2003, he was supported by the Flemish Institute for Scientific and Technological Research in Industry (IWT). Since September 2004, he is with Alcatel, Research and Innovation in Antwerp, Belgium. His research interests are in the area of digital signal processing for DSL communications.



Koen Vanbleu (S'04–M'04) was born in Bonheiden, Belgium, in 1976. He received the Master's degree and the Ph.D. degree in electrical engineering from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1999 and 2004, respectively.

From 1999 to 2003, he was supported by the Fonds voor Wetenschappelijk Onderzoek (FWO) Vlaanderen. Since November 2004, he has been with the DSL Engineering division of Broadcom, Mechelen, Belgium. His research interests are in the area of digital signal processing for DSL communications.