A Convex Optimization Approach to Blind Channel Shortening in Multicarrier Modulations

Huy-Dung Han and Zhi Ding
Department of Electrical and Computer Engineering
University of California, Davis, CA, 95616
Email: {hdhan, zding}@ucdavis.edu

Abstract—We study the problem of channel shortening in multicarrier modulation systems without training. We reformulate two existing methods, the sum-squared and the sum-absolute autocorrelation minimization algorithms (SAM and SAAM), into semidefinite programming to overcome their shortcoming of local convergence. We present the original SAM and SAAM cost functions into a batch optimization problem before relaxing the original problem into globally convergent semidefinite programming algorithms. Our batch processor is superior to the original problem into globally convergent semidefinite programming algorithms. Our batch processor is superior to the original stochastic gradient algorithms in terms of achievable bit rate and signal to interference and noise ratio (SINR).

Index Terms—Blind, convex optimization, relaxation, channel shortening, DMT, OFDM, equalization

I. INTRODUCTION

Multicarrier (MC) communication systems have become popular in both wireless and wireline communications. They can deliver high speed broadband connection and provide channel adaptive data rate to maximize the utilization of available link capacity. In wireless environment, MC systems based on orthogonal frequency division multiplexing (OFDM) have been adopted in IEEE802.11a/g/n, HIPERLAN2, and IEEE802.16e standards. Wireline systems such as digital subscriber loop/lines (DSL) also utilize discrete multitone (DMT) modulation.

Multicarrier modulations enjoy many advantages such as simpler equalization and robustness to channel dispersions. To combat multipath fading effect, cyclic prefix (CP) need to be included in the modulated signal before transmission. One requirement is that the CP length must be at least as long as the channel delay spread to avoid inter-carrier interference. However, because the delay spread of many (wireless) channels can vary significantly, it can be wasteful to select the CP length to account for the longest possible channel delay spread because longer CP lowers bandwidth efficiency. In many practical applications such as DSL or Ultra Wideband (UWB), the channel length can sometimes be much longer than nominal CP length. Consequently, inter-symbol interference (ISI) and inter-carrier interference (ICI) can degrade the multicarrier receiver performance.

To overcome this problem, channel shortening filters also known as time domain equalizers (TEQ) are often utilized in practical receivers. TEQs are designed such that the combined response of the channel plus the TEQ has an effective length shorter than the CP length, thereby removing much of the ICI. Over the past decade, a number of TEQ approaches have been explored under different design criteria. Melsa et al. [1] proposed the design for maximum shortening signal to noise ratio (MSSNR) in an attempt to minimize the energy outside the desired window of delay spread while fixing the energy inside the window. Other designs have been proposed aimed at maximizing the bit rate [2], [3], [4]. In these TEQ algorithms, the transmission of training sequences is still necessary, which can lead to significant training overhead and lower bandwidth efficiency.

Blind algorithms for TEQ are preferred in systems where TEQ training is undesirable or costly. It reduces training overhead and improves spectral efficiency. In [5], Martin et al. proposed a Multicarrier Equalization by Restoration of Redundancy (MERRY) method in which the TEQ is designed to restore the CP structure of OFDM data symbols. The algorithm can be implemented in both batch and adaptive forms. The adaptation is carried out once per OFDM symbol. The “sum-squared autocorrelation minimization (SAM)” algorithm [6] is another example of blind adaptive TEQ approach. Its objective function forces the combined channel impulse response to be effectively zero outside a response window length. One of its weaknesses lies in its multimodal cost function whose global convergence cannot be guaranteed.

Adaptive blind algorithms for channel shortening are suitable for connection-based links with slowly time-varying channels because they can adapt and track channel changes. However, in bursty wireless systems where channel state changes rapidly between OFDM symbols and when frequent training symbols are not available, we prefer batch algorithms. In [7], Kameyama et al. considered batch algorithms based on second-order statistics for TEQ optimization. The TEQ coefficient vector is chosen from a null space that depends on the effective channel length. In the present work, we focus on alternative batch algorithms for TEQ.

In this manuscript, we formulate the channel shortening problem into a new convex optimization issue. We modify SAM and SAAM algorithms into simpler and globally convergent convex optimization methods. Through relaxation, our new convex optimization problems that can be solved by well-known and efficient numerical methods. Our solutions can

1This material is based upon work supported by the National Science Foundation under Grant 0515058 and, in part, by the Vietnam Education Foundation.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2009 proceedings.
directly be applied as TEQ parameters for channel shortening or can be further refined via local search algorithms.

II. SYSTEM MODEL

The basic block diagram of a multicarrier system with receiver TEQ is illustrated in Figure 1. Let $X [k]$ be a complex signal vector of length $N$ for transmission by the $N$ active subcarriers. Each element in $X[k]$ is an $M$-QAM symbol carried by a subcarrier. The time domain OFDM signal $x[k]$ is generated from an $N$-point IFFT plus a CP of length $L_{CP}$. The resulting serial signal $x[k]$ is transmitted through a distortive channel of impulse response $\{h_k\}$. The channel response can be denoted as $h[k] = [h_0 h_1 \ldots h_{L_h}]^T$ with delay spread $L_h + 1$. To eliminate inter-carrier interference, it is necessary that $L_{CP} \geq L_h$.

For linear channels with additive Gaussian noise $\eta[k]$ that is i.i.d. with zero-mean and variance $\sigma^2_\eta$, the received signal samples are related to the channel input via

$$ r[k] = \sum_{l=0}^{L_h} h[l] x[k-l] + \eta[k]. \quad (1) $$

By sending $r[k]$ through the receiver TEQ with parameters $w = [w_0 \ w_1 \ \ldots \ w_{L_w}]^T$, the TEQ output signal $y[k]$ is ready for processing by the conventional OFDM receiver for equalization and data recovery. To simplify notations, we let $c = [c_0 \ c_1 \ \ldots \ c_{L_c}]^T$ to denote the combined system response of the physical channel and the TEQ. This combined impulse response has total length of $L_c + 1 = L_h + L_w + 1$.

In a standard multicarrier system, a bank of 1-tap frequency domain equalizers would follow the FFT to recover the modulated data symbols. Longer cyclic prefix allows the system to tolerate longer multipath delay spread, at the cost of bandwidth efficiency. Thus, it is more efficient to utilize a CP of nominal length that can tackle a majority of multipath channels. For less common channels whose delay spread is longer than $L_{CP}$, we rely on the time domain equalizer (TEQ) to reduce the effective length of the combined response of channel plus TEQ such that it becomes effectively shorter than the cyclic prefix length. In short, the TEQ parameters $w_i$ should be such that $c_i \approx 0, \ l < 0, \ l > L_{CP}$.

III. BLIND TEQ ALGORITHMS BASED ON OUTPUT AUTO-CORRELATION

Blind algorithms do not rely on training for TEQ parameter selection. We consider several existing algorithms here. First, we write the TEQ output signal as

$$ y[k] = \sum_{l=0}^{L_w} w_l r[k-l] $$

$$ = \sum_{l=0}^{L_c} c_l r[k-l] + \sum_{l=0}^{L_w} w_l \eta[k-l] \quad (2) $$

For simplicity, without loss of generalization, we let the transmitted time-domain signal $x[k]$ be real with zero mean and unit variance. We now briefly introduce two existing blind TEQ algorithms that are based on TEQ output autocorrelation.

The objective of TEQ is to minimize the effective channel response $c$ outside a window of length $L_{CP} + 1$. Consequently, this requires that the autocorrelation of TEQ output $y[k]$ be zero outside the window of size $2L_{CP} + 1$ center at zero, i.e.,

$$ R_c(l) = \sum_{i=0}^{L_c} c_i c_{i-l} = 0 \quad \text{when} \ |l| > L_{CP}. \quad (3) $$

By exploiting this property, the TEQ coefficients can be determined, as shown in [6], by minimizing the cost function

$$ J_{SAM} = \sum_{l=L_{CP}+1}^{L_c} |R_c(l)|^2 \quad (4) $$

To avoid the trivial solution of $w = 0$, we impose the constraint of that the equalizer output $y[k]$ to have unit power. Consider two assumptions:

1) The transmitted sequence is white, zero-mean, wide-sense stationary.

2) The effective channel $c$ has length less than half FFT size, i.e. $L_c \leq N/2$.

Under these conditions, the autocorrelation function $R_g(l)$ of the TEQ output signal $y_k$ is related to $R_c(l)$ and the autocorrelation function of the TEQ parameters $R_w(l)$ via [6]

$$ R_g(l) = \sum_{l=0}^{L_c} |R_c(l)|^2 + 2\sigma^2_\eta R_w(l) \quad (5) $$

where we denote $R_w(l) = \sum w_l w_{l-\ell}$. Under additive channel noise, the cost function of (4) can be written as

$$ J_{SAM} = \sum_{l=L_{CP}+1}^{L_c} |R_g(l)|^2 $$

$$ = \sum_{l=L_{CP}+1}^{L_c} |R_c(l)|^2 + 2\sigma^2_\eta \sum_{l=L_{CP}+1}^{L_c} R_c(l) R_w(l) $$

$$ + \sigma^4_\eta \sum_{l=L_{CP}+1}^{L_c} |R_w(l)|^2 \quad (6) $$

Note that, if $L_w \leq L_{CP}$, then the last two sums in the preceding equation would vanish such that $J_{SAM} = J_{SAM}$. For this reason, the SAM algorithm was presented in [6] to find TEQ parameters for minimizing $J_{SAM}$.

As an alternative to minimizing the cost function defined in (4), the authors of [8] proposed a different cost for minimization. It takes the form as the sum of absolute values of auto-correlations (SAAM) [8]

$$ J_{SAAM} = \sum_{l=L_{CP}+1}^{L_c} |R_c(l)| \quad (7) $$

From the TEQ output, the SAAM cost function is approxi-
mated by by

\[
\hat{J}_{\text{SAAM}} = \sum_{l=L_{CP}+1}^{L_c} |R_y(l)|
\]

\[
= \sum_{l=L_{CP}+1}^{L_c} \left| R_c(l) + \sigma_n^2 R_w(l) \right| \tag{8}
\]

Similar to the SAM cost function, so long as \( L_w \leq L_{CP} \) is satisfied, we can find \( \hat{w} \) by minimizing \( \hat{J}_{\text{SAAM}} = \hat{J}_{\text{SAM}} \) as the noise correlation \( \sigma_n^2 R_w(l) \) vanishes.

In fact, we can further generalize the cost functions into

\[
\hat{J}_p = \sum_{l=L_{CP}+1}^{L_c} |R_y(l)|^p \tag{9}
\]

where \( p \geq 1 \). The selection of \( p = 1, 2 \) corresponds to the SAAM and SAM costs, respectively. When \( p = \infty \), the cost becomes

\[
\hat{J}_\infty = \max_{L_{CP}+1 \leq l \leq L_c} |R_y(l)|
\]

Given the cost function defined in terms of TEQ output autocorrelations, steepest descent search can be directly applied to determine the optimum TEQ parameter vector \( \hat{w} \) that minimizes \( \hat{J}_p \). Specifically, we can apply steepest descent by processing a block of received signal samples in a batch algorithm. Alternatively, adaptive algorithms can also be utilized to implement stochastic gradient descent searches. Details of these algorithms can be found in [8] and [6] for \( p = 1, 2 \), respectively.

It is important to note, however, that the costs \( \hat{J}_p \) are non-convex and multimodal functions of \( w_i \). In particular, good initial values are essential in practical implementations in order to find the optimum TEQ setting as the global minimum. The absence of global convergence guarantee motivates our investigation to search for more consistent blind TEQ solutions that can either be directly applied as TEQ parameters or as good initial points for further search of the global minimum of SAM or SAAM. We present our formulation of convex optimization algorithms with assured global convergence based on the principle of positive semidefinite relaxation in the next section.

IV. MINIMIZATION BY CONVEX OPTIMIZATION

A. Convex Optimization and Relaxation

For notational convenience, first define \( R_r(l) \) as the autocorrelation of the received signal and \( R_c(l) \) as a Toeplitz

matrix of autocorrelations of the channel output signal

\[
R_r(l) = \begin{bmatrix}
R_c(l) & \cdots & R_c(l-L_w) \\
\vdots & \ddots & \vdots \\
R_c(l+L_w) & \cdots & R_c(l)
\end{bmatrix}
\tag{10}
\]

We can now rewrite the autocorrelation of the TEQ output as

\[
R_y(l) = E \{ y[k] y[k-l] \} = \sum_{i=0}^{L_w} w_i r[k-i] \sum_{j=0}^{L_w} w_j r[k-l-j] = \sum_{j=0}^{L_w} w_j \sum_{i=0}^{L_w} r_k r_{k+l-i} = w^T R_r(l) w. \tag{11}
\]

We note that the nontrivial constraint is equivalent to requiring \( R_y(0) = w^T R_r(0) w = 1 \). This set of new notations allows us to write the minimization problem of the cost in (9) into

\[
\text{Minimize } \sum_{l=L_{CP}+1}^{L_c} |w^T R_r(l) w|^p \tag{12}
\]

subject to \( w^T R_r(0) w = 1 \).

Because \( R_r(0) \) is positive definite, we can decompose \( R_r(0) = Q^T Q \).

This decomposition allows us to define \( \hat{w} = Q\hat{w} \) to simplify the quadratic parameter constraint into \( \|\hat{w}\|^2 = \hat{w}^T Q^T Q \hat{w} = 1 \).

Moreover, we have

\[
w^T R_r(l) w = \hat{w}^T Q^{-T} R_r(l) Q^{-1} \hat{w} = \hat{w}^T \tilde{R}_r(l) \hat{w} \tag{13}
\]

by defining \( \tilde{R}_r(l) = Q^{-T} R_r(l) Q^{-1} \).

We can now write our generic optimization problem as

\[
\text{Minimize } \sum_{l=L_{CP}+1}^{L_c} |\hat{w}^T \tilde{R}_r(l) \hat{w}|^p \tag{14}
\]

subject to \( \|\hat{w}\|^2 = 1 \).
We introduce a new set of variables $\tau_l$ and change the problem to [9]

$$\text{Minimize } \sum_{l=L_{CP}+1}^{L_c} \tau_l^p$$
subject to $-\tau_l < \tilde{w}^T \tilde{R}_r (l) \tilde{w} < \tau_l$
and $\|\tilde{w}\|^2 = 1$

The problem remains nonlinear and non-convex. We can, however, resort to the technique of semidefinite relaxation and convert the problem into convex optimization. Define $W = \tilde{w} \tilde{w}^T$, we can apply semidefinite relaxation and solve the relaxed optimization problem

$$\text{Minimize } \sum_{l=L_{CP}+1}^{L_c} \tau_l^p$$
subject to $-\tau_l < \text{Tr} \left( W \tilde{R}_r (l) \right) < \tau_l$
and $W \succeq 0$
\[ \text{Tr}(W) = 1 \]

where $W \succeq 0$ denotes (symmetric) positive semidefinite.

Through relaxation, the problem of (16) is now convex. Its global optimal solution can be found using modern semidefinite programming (SDP) solvers such as SDPT-3 [10] by applying the interior point method. For the convex optimization, the choice of $p$ is essential and we can solve with the SDP for $p = 1, 2, \infty$.

### B. Post processing for the semidefinite relaxation

Once the optimum positive semidefinite solution $W$ is determined, it has to be mapped back into the solution space $w$. Typically, some post-processing techniques are needed to convert the SDP relaxation solution into an approximate solution of the original optimization problem. One approach often applied is the Gaussian randomized approximation [11]. More specifically, let $W_{\text{opt}}$ be the result of the optimizing solution to (16). We then randomly generate vectors $\tilde{w}$ based on zero mean Gaussian distribution $\mathcal{N} (0, W_{\text{opt}})$. Among the ensemble of random vectors, we select the one that yields the smallest cost in the original problem (14). Under certain conditions, the randomization can achieve a close to the global minima in polynomial time [12], [13].

Alternatively, we can find the optimum vector $\tilde{w}$ that can provide the best rank-1 decomposition of the relaxed matrix solution $W_{\text{opt}}$. In other words, we can use as the final TEQ parameter vector $\tilde{w}$ that minimizes

$$\|\tilde{w} \tilde{w}^T - W_{\text{opt}}\|^2$$

Thus, the solution to the minimization of (17) is the eigenvector that corresponds to the maximum eigenvalue of $W_{\text{opt}}$. In the following section we will refer to SDP optimization with the eigenvector technique as SDP.

### C. Computational Complexity

In addition to the estimation of TEQ input autocorrelation functions based on actual data samples, the major computational load takes place in the SDP. The computational complexity for the SDP is polynomial-time [14], with the worst case scenario of $O(L_{w}^{6.5})$. Additionally, the complexity for the post-processing by finding the optimum parameter vector is in the order of $O(L_{w}^{3.5})$. Typically, the channel length $L_h$ is much longer than the TEQ length. Therefore, for systems that adopt short length TEQ, the SDP plus the post-processing has complexity comparable to gradient-based algorithms in the original SAM and SAAM. In comparison, the number of floating point operations (FLOPs) of an adaptive method in one update step may be less but the total number is comparable to or more than the proposed batch convex optimization algorithm.

### V. Simulation Results

A simulation example is described for SAM in [15] which provides the details of the simulation setup. To compare various methods through a common benchmark, we adopt the same setup in which a set of standard Asymmetric DSL (ADSL) downstream parameters was chosen. In particular, the FFT size in the OFDM equals 512, whereas the cyclic prefix length is 32. The TEQ has 16 taps. We take, as simulation channel, the CSA test loop 1 from [16]. The noise power is set such that $\sigma_x^2 \|h\|^2 / \sigma_n^2 = 40$ dB. Both the SAM and SAAM algorithms are tested by collecting 75 OFDM frames (of 544 samples each). They are both initialized with a single unit spike in the middle of the tap vector. For each iteration, the cost functions and the TEQ updates are estimated by averaging over a moving window of length 32. For fair comparisons, we estimate the correlation matrices in our SDP formulation in a similar manner. Furthermore, to demonstrate the loss of performance due to the lack of training, we provide, as performance upper bound, training-based simulation results from the MSSNR method [1].

We use the achievable bit rate and the signal to interference and noise ratio (SINR) as the performance metrics in the simulations. More specifically, for multicarrier systems, the achievable bit rate is determined according to the theoretical analysis as

$$R = \sum_{i=0}^{N-1} \log_2 \left( 1 + \frac{\text{SNR}_i}{\Gamma} \right)$$

where $\text{SNR}_i$ is the matched filter bound of SNR for the $i$th sub-channel and $\Gamma$ is the gap for achieving the Shannon capacity. Details of this setup are available in [4]. The output SINR of the TEQ is calculated [17] according to

$$\text{SINR} = \frac{\sigma_x^2 \|c_d\|^2}{\sigma_x^2 \left( \|c_d\|^2 - \|c_{d^*}\|^2 \right) + \sigma_n^2 \|w\|^2}$$

where $c_d = [c_0 \cdots c_{d + L_{CP}}]^T$ and $d^* = \arg \max_d \|c_d\|^2$. In short, $c_{d^*}$ represents the segment of consecutive $L_{CP} + 1$
samples in the combined channel-TEQ response that has the largest energy.

Figure 2. Original channel and the shortened channel

Figure 3. TEQ parameters from SDP ($p = 2$)

Figure 4. The convergence of adaptive algorithms

To further improve the SDP results, we can use the TEQ parameter vector from SDP to initialize the two adaptive SAM and SAAM algorithms. Figure 4 shows the results from SDP initialization in comparison with those from the original SAM and SAAM algorithms. We can see that, given the SDP initialization, the costs in (6) and (8) are minimized very rapidly after several iterations whereas the regular adaptive SAM and SAAM algorithms require several hundreds iterations. We can observe from these results that the optimized parameter vectors from SDP are very close to minima of the SAM and the SAAM costs.

To further improve the SDP results, we can use the TEQ parameter vector from SDP to initialize the two adaptive SAM and SAAM algorithms. Figure 4 shows the results from SDP initialization in comparison with those from the original SAM and SAAM algorithms. We can see that, given the SDP initialization, the costs in (6) and (8) are minimized very rapidly after several iterations whereas the regular adaptive SAM and SAAM algorithms require several hundreds iterations. We can observe from these results that the optimized parameter vectors from SDP are very close to minima of the SAM and the SAAM costs.

Table I

<table>
<thead>
<tr>
<th>TEQ length</th>
<th>16</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSSNR</td>
<td>3.892</td>
<td>3.8271</td>
</tr>
<tr>
<td>SAM</td>
<td>2.6446</td>
<td>2.9653</td>
</tr>
<tr>
<td>SAAM</td>
<td>2.5836</td>
<td>2.818</td>
</tr>
<tr>
<td>SDP ($p = 2$)</td>
<td>3.1453</td>
<td>3.1209</td>
</tr>
<tr>
<td>SDP initialized SAM</td>
<td>2.6308</td>
<td>2.321</td>
</tr>
<tr>
<td>SDP initialized SAAM</td>
<td>2.5981</td>
<td>2.058</td>
</tr>
</tbody>
</table>

In Table I we compare the achievable bit rates for different schemes for TEQ taps of 16 and 5, respectively. The achievable bit rates of our newly proposed SDP algorithm are markedly better than those of other blind methods.

It is interesting to observe that, after using the result of SDP to initialize the SAM and the SAAM, the achievable bit rates become lower. This is because lower SAM/SAAM costs do not necessarily correspond to higher bit rates. Indeed, the optimized SDP solution is closer to the TEQ that maximizes the achievable bit rate.

For better illustration, Figure 5 compares the bit rate under different SNR from the three algorithms (SDP, SAM, and SAAM). In the low SNR regime, the performance of all three algorithms are comparable. However, under high SNR, the proposed SDP algorithm outperforms the adaptive methods substantially. It is interesting to note that the choice of $p$ does not appear to significantly affect the achievable bit rates. Therefore, we should consider using the best $p$ that leads to the lowest computational complexity and the greatest implementation convenience.

We note that, in some practical application, the SINR may be a more important performance metric. Figure 6 compares the resulting SINR for the algorithms under consideration. In general, all three TEQ algorithms can improve the SINR when compared to systems without the TEQ. The SINR of the resulting TEQ output tends to increase almost linearly as SNR grows before eventual saturation due to the non-zero residuals beyond the effective length of the channel-TEQ combination. For SNR values beyond certain level, the noise becomes insignificant in comparison with the residual total.
response which the TEQ failed to fully annihilate. This result illustrates the true capability and the limitation of each TEQ algorithm.

![Graph showing SINR vs. SNR for different algorithms](image)

Figure 6. SINR versus SNR for different algorithms

We also note that, the underlying ADSL channel response has quite a special feature, i.e. all the channel coefficients is positive. The example used to illustrate Theorem 1 in [6] also has the same feature. In [6], the authors showed that two of the four minima match the two global minima of the inverse of the shortening signal to noise ratio (SSNR) cost function defined in [6]. The SSNR is defined as the power ratio of the effective channel response inside and outside the window of interest (of width $L_{CP}$ + 1). Note that the SSNR is actually the SINR when the effect of noise is not considered. However, this phenomenon is not generally true for every channel. To illustrate, Figure 7 shows the contour of the 1/SSNR cost and the positions of the SAM minima for a channel with response $h = [0.1 1 -0.8 0.2]$ and $L_{CP} = 1$. Although the global minima of the SAM cost closely match some local minima of the 1/SSNR cost, and the global minima of the 1/SSNR cost do not necessarily manifest as minima of the SAM cost. This is because when the effective channel is shortened within the CP length the SAM cost (4) is minimized. However, the converse is not true.

![Contour of the 1/SSNR cost function](image)

Figure 7. Contour of the 1/SSNR cost function

VI. CONCLUSION

In this work, we formulate a convex optimization approach to the blind channel shortening problem. Unlike the existing gradient based search algorithms that are more susceptible to local convergence, the algorithm can provide better performance than the SAM and SAAM adaptive algorithms in terms of achievable bit rate and SINR. The newly proposed convex optimization is general and can be applied to different performance norms. It is also easy to impose additional parametric and signal constraints for performance improvement and/or for practical implementations.

REFERENCES