BLIND CHANNEL SHORTENERS

C. R. Johnson, Jr., R. K. Martin, J. M. Walsh, A. G. Klein, C. E. Orlicki, and T. Lin [1]

School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853

Abstract: Although blind, adaptive algorithms for equalization are widely studied, hitherto there has been little academic attention given to blind, adaptive algorithms for channel shortening. Channel shortening is needed to preserve subcarrier orthogonality in multicarrier modulation, and it can be used to dramatically reduce the complexity of maximum likelihood sequence estimation and multiuser detection. This paper reviews the channel shortening problem from a tutorial perspective, and shows how it is an extension of traditional equalization. It is shown that traditional methods of devising blind, adaptive equalization algorithms cannot be easily applied to the channel shortening problem. The paper concludes with a discussion of several new property restoral algorithms that enable blind, adaptive channel shortening.

Keywords: Blind, Adaptive, Channel Shortening, Equalization, Inverse Control

1. INTRODUCTION

The ordinary objective of communication channel equalization is to reproduce at the equalizer output a delayed version of the channel input sequence, with the channel output the equalizer input. When using a linear, baud-spaced equalizer, this translates, in the channel-noise-free case, to a channel and equalizer combination with a transfer function of $z^{-\delta}$ with δ a positive integer and z^{-1} a unit sample delay.

In various applications, such as multicarrier communication systems for wired (e.g. DMT for xDSL) and wireless (e.g. OFDM for DVB and WLAN) scenarios, the impulse response of the channel equalizer combination need not be just a single nonzero term. Instead a channel-equalizer combination with an impulse response of suitably limited duration, e.g. $\sum_{i=0}^{P} h_i z^{-(i+\delta)}$, can be adequate for desired system performance. The specific values of the h_i are less important (as long as some are nonzero) than the fact that outside this $P+1$ sample window the impulse response coefficients are all (nearly) zero. Generally, channel-shortening can be achieved with a shorter equalizer and with less noise gain than equalization to a single spike. An extreme example is a finite impulse response (FIR) channel containing a zero on the unit circle. A zero-forcing linear FIR equalizer would have infinite gain at the frequency of the channel null, thereby catastrophically amplifying any channel noise. However, the single zero on the unit circle could be retained in a shortened channel.

While adaptive channel shorteners relying on training have been developed and implemented, prior to last year (2002) no blind method for direct adaptive channel shortening existed. This paper describes the channel-shortening problem formulation as a metamorphosis from a traditional channel equalization problem first into a model-

[1] This work was supported in part by Applied Signal Technology.

following inverse controller, then into a channel-shortener. Trained adaptive algorithms for the linear combiner format that emerges in each case are readily proposed. The inappropriateness of establishing blind channel shortening schemes based on decision-direction or dispersion minimization, which were successful in producing blind channel equalizers, will become apparent.

This paper discusses several blind, adaptive channel shorteners proposed in 2002. These algorithms arose from exploitation of transmitted signal properties found in practical communication systems; in one case whiteness (Balakrishnan *et al.*, 2002) and in the other an inclusion of replicated segments (Martin *et al.*, 2002*a*). Using a classical stochastic gradient descent approach, these algorithms are based on cost functions that penalize deviation from some property of the signal to be recovered, and require only the equalizer input and output to form their parameter updates (and are therefore considered blind).

Other researchers have proposed blind, adaptive algorithms for systems that require channel shortening, although these efforts have focused on equalization, rather than channel shortening. However, since channel shortening is the goal rather than equalization, the use of an equalizer is expected to yield suboptimal performance compared to a channel shortener. In (Jones, 2003), a time domain approach similar to (Martin *et al.*, 2002*a*) was proposed, but the algorithm leads to full equalization rather than just channel shortening. (de Courville *et al.*, 1996) makes use of the common practice with OFDM of transmission of zeros on some carriers as a substitute for training data on those carriers, again leading to a single-spike equalization algorithm. (Romano and Barbarossa, 2003) extended this method through the use of frequency-hopping in the transmitter to allow for blind, adaptive channel shortening.

2. ADAPTIVE PARAMETER ESTIMATION PROBLEM MUTATION

We keep the description basic in drawing a thread from traditional channel equalization to channel shortening via model-following. The intent is to draw attention to the issue of adaptive channel shortening and the need for new approaches to blind solutions.

2.1 Traditional Channel Equalization

In traditional channel equalization the goal is to process the channel output sequence to reproduce the channel input sequence, as depicted in Figure 1. The source s, a delayed version of which is to

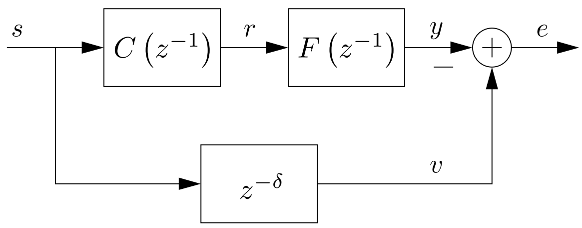


Fig. 1. Traditional Channel Equalization

be recovered so that $\mathrm{E}[e^2]$ is minimized, takes on values from a discrete set, e.g. $\{\pm 1\}$ or $\{\pm 1, \pm 3\}$. The linear channel is presumed to have a causal, FIR model

$$r(k) = \sum_{j=0}^{M-1} c_j s(k-j) \tag{1}$$

with transfer function $C(z^{-1})$. The output of the channel is the received signal in the noise-free idealization of Figure 1. The equalizer also has a causal FIR model

$$y(k) = \sum_{i=0}^{N-1} f_i r(k-i) \tag{2}$$

with transfer function $F(z^{-1})$ that filters the received signal $r(k)$ and produces the output $y(k)$ we would like to have match $s(k-\delta)$. The delayed source recovery error $e(k)$ for delay δ and f_i values at time k is

$$e(k) = s(k-\delta) - y(k) = s(k-\delta) - \sum_{i=0}^{N-1} f_i r(k-i) \tag{3}$$

The recovery error description of (3) fits the format of a linear combiner's prediction error

$$e(k) = d(k) - X^T(k)\theta(k) \tag{4}$$

with its difference between the desired combiner output d and its actual output formed by the inner product of the regressor vector X and the adapted parameter vector θ. The definitions that match (3) to (4) are $d(k) = s(k-\delta)$,

$$X(k) = [r(k) \;\; r(k-1) \;\; \ldots \;\; r(k-N+1)]^T, \tag{5}$$

and

$$\theta(k) = [f_0(k) \;\; f_1(k) \;\; \ldots \;\; f_{N-1}(k)]^T. \tag{6}$$

For the linear combiner error of (4), the LMS algorithm (Widrow and Stearns, 1985)

$$\theta(k+1) = \theta(k) + \mu e(k) X(k) \tag{7}$$

is a classic adaptive approach to minimization of $\mathrm{E}[e^2]$.

Implementing (7) requires knowledge of $s(k-\delta)$ at the recevier. This is accomplished by the (periodic) transmission of a prearranged training sequence. We can use the property that the source signal takes on only certain values, e.g. ± 1, to produce blind alternatives that do not require explicit knowledge at the receiver of $s(k-\delta)$.

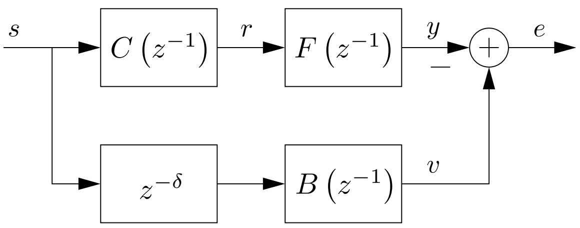


Fig. 2. Model-Following

One strategy presumes that the (initial) equalizer setting, while not perfectly recovering $s(k-\delta)$, generates a signal $y(k)$ that is close to $s(k-\delta)$. In fact, we assume that it is close enough so a quantizer produces error-free delayed source recovery, i.e. $Q[y(k)] = s(k-\delta)$. For $s \in \{\pm 1\}$, a sign operator serves as the quantizer Q, and $\text{sign}[y(k)]$ replaces $s(k-\delta)$ in e in (3) which is substituted in (7) to produce the blind decision-directed LMS (DDLMS) algorithm

$$\theta(k+1) = \theta(k) + \mu\left(Q[y(k)] - y(k)\right) X(k). \quad (8)$$

Another blind channel equalization algorithm can be created by differently exploiting the discrete alphabet nature of the source to be recovered. For example, with s either 1 or -1, s^2 is always 1. This suggests the cost function $\text{E}[(1-y^2(k))^2]$. A stochastic gradient descent minimizing this dispersion cost yields the constant modulus algorithm (CMA) (Treichler *et al.*, 2001, Ch. 6)

$$\theta(k+1) = \theta(k) + \mu(1-y^2(k))y(k)X(k) \quad (9)$$

where $y(k) = X^T(k)\theta(k)$. CMA in (9) is termed blind because it does not need knowledge of the specific source sequence $\{s\}$ that $\{y\}$ is to match. Dispersion minimization also works for multilevel discrete alphabet sources.

2.2 Model-Following Inverse Control

Now we will mutate the problem into a version reminiscent of inverse control (Widrow and Stearns, 1985, Ch. 11). Rather than cause the transfer function of the channel-equalizer combination $C(z^{-1})F(z^{-1})$ to equal $z^{-\delta}$, we will choose to have it match a prespecified model transfer function $B(z^{-1})$, as in Figure 2. Writing the error as

$$e(k) = v(k) - \sum_{i=0}^{N-1} f_i r(k-i) \quad (10)$$

reveals a linear combiner format. With this e, but the same definitions of θ and X as in the traditional channel equalization problem, LMS in (7) provides a trained adaptive solution.

The blind algorithms DDLMS and CMA do not survive the mutation because the signal to be matched by y, i.e. v in Figure 2, is not drawn from the same alphabet as the source.

2.3 Channel Shortening

We again mutate the problem. Though the model-following schematic of Figure 2 still applies, we no longer preselect $B(z^{-1})$. Instead the objective is only to have the channel-equalizer combination FIR response have nonzero values only in a window P samples wide. This task only presents a challenge if P is less than the sample width M of the channel. Hence, the labeling of this task as channel shortening.

If we judge our success by minimizing the mean of the square of the model-following error e in Figure 2, we must not admit the trivial solution with the coefficients of both $F(z^{-1})$ and $B(z^{-1})$ set to zero. A simple way to enforce this constraint is to fix $b_\zeta = 1$ for some ζ. Thus,

$$\begin{aligned} e(k) &= \sum_{j=0}^{P-1} b_j(k)s(k-j-\delta) - \sum_{i=0}^{N-1} f_i r(k-i) \\ &= s(k-\zeta-\delta) + \sum_{j=0, j\neq\zeta}^{P-1} b_j(k)s(k-j-\delta) \\ &\quad - \sum_{i=0}^{N-1} f_i r(k-i) \\ &= s(k-\zeta-\delta) - X^T(k)\theta(k) \end{aligned} \quad (11)$$

where

$$\begin{aligned} X(k) = [&-s(k-\delta), \ \ldots, \ -s(k-\delta-\zeta+1), \\ &-s(k-\delta-\zeta-1), \ \ldots, \ -s(k-\delta-P+1), \\ &r(k), \ \ldots, \ r(k-N+1)]^T \end{aligned} \quad (12)$$

and

$$\theta(k) = [b_0, ..., b_{\zeta-1}, b_{\zeta+1}, ..., b_{P-1}, f_0, ..., f_{N-1}]^T \quad (13)$$

Under the assumption that the preselection of δ and ζ leads to an acceptable solution, these definitions of e, X, and θ and (7) provide a trained adaptive channel shortener similar to that in (Falconer and Magee, 1973). But, just as noted in the case of model-following, the desired output y is no longer drawn from the source alphabet and loses the finite-alphabet property exploited by decision direction and dispersion minimization. Can other signal properties can be used to establish a blind channel shortener?

3. THE NEED FOR BLIND CHANNEL SHORTENING

A primary application of channel shortening occurs in multicarrier communication systems (Pollet *et al.*, 2000) present in wired digital subscriber loop (DSL) and wireless local area network (LAN) standards. These standards use a cyclic prefix in the transmitted sequence. Each data block of a prespecified length has its last several samples prepended to the front of the block prior to transmission. Intersymbol (and intercarrier) interference is removed as long as the length of this cyclic prefix segment is greater than the delay spread of the channel-equalizer combination. Adaptation will be needed if the channel is time-varying. Blind adaptation will be desirable if the downtime for (re)training represents a large portion of system operating activity.

Currently, in the DSL application the cyclic prefix set in transmission standards is acknowledged to be much shorter than the channel delay spread encountered in practice. This substantiates the use of a channel shortener. However, the channel is presumed time invariant (aside from glacially-paced, temperature-induced, performance-degrading channel parameter drifts which occur in practice), which obviates the need for periodic retraining and the primary motivation for a blind channel shortener.

In the wireless LAN application, there is general agreement that the channel model is time-varying. However, so far, the belief is that such schemes will not be deployed in circumstances in which the channel delay spread exceeds the standards-mandated cyclic prefix. Thus, while adaptation would be needed in a channel shortener, and a blind scheme would be welcome, the need for a channel shortener is typically denied.

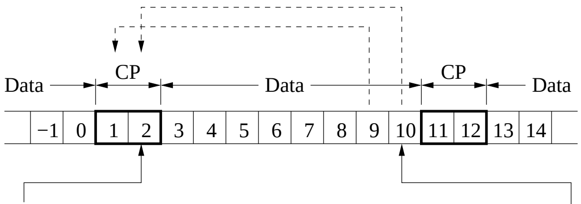
The compelling need for a blind channel shortener remains in the future. Based on a similar experience in the adoption of blind implementation of traditional channel equalization, our contention is that such a need will arise. The previous lack of blind channel shortening schemes and our belief in their eventual utility motivated our search for candidate algorithms for blind channel shortening.

4. TWO PROPERTY RESTORAL ALGORITHMS

Two signal properties that can be exploited include the traditonally assumed whiteness of the transmitted source sequence (especially common with coded and scrambled signals) and the cyclic prefix mentioned that is common with multicarrier system application of a channel shortener. Restoring each of these two properties generates a candidate blind channel shortener.

4.1 The MERRY and FRODO Algorithms

The MERRY algorithm (Multicarrier Equalization by Restoration of Redundancy) makes use of the redundancy of the data introduced by the



$$y(2) = h_0 s(2) + h_1 s(1) + h_2 s(0) + h_3 s(-1) + h_4 s(-2)$$
$$= h_0 s(10) + h_1 s(9) + [h_2 s(0) + h_3 s(-1) + h_4 s(-2)]$$
$$y(10) = h_0 s(10) + h_1 s(9) + [h_2 s(8) + h_3 s(7) + h_4 s(6)]$$

Fig. 3. Illustration of the difference in the ISI at the received CP and at the end of the received symbol.

cyclic prefix (CP). When the CP is added, the last P samples in the length N block are prepended to the start of the block, which makes the transmitted data appear periodic over those $N + P$ samples (called a “symbol”). This periodicity is necessary for maintaining the orthogonality of the subcarriers (Pollet *et al.*, 2000). After the CP is added, the last P samples are identical to the first P samples in the symbol, i.e.

$$s\,(Mk + i) = s\,(Mk + i + N)\,, \quad i \in \{1, \ldots, P\}\,, \tag{14}$$

where $M = N + P$ is the total symbol duration and k is the symbol index. Figure 3 shows an example of this, with $N = 8$, $P = 2$, and $M = N + P = 10$. The symbol pictured is for $k = 0$. The received data $\mathbf{r}$ is obtained from $\mathbf{s}$ by

$$r(Mk+i) = \sum_{l=0}^{L_c} c_l \cdot s(Mk+i-l) + n(Mk+i), \tag{15}$$

and the equalized data $\mathbf{y}$ is obtained from $\mathbf{r}$ by

$$y(Mk + i) = \sum_{j=0}^{L_f} f_j \cdot r(Mk + i - j). \tag{16}$$

The effective channel is $\mathbf{h} = \mathbf{c} \star \mathbf{f}$.

The channel destroys the relationship analogous to (14) in the received data, because the ISI that affects the CP is different from the ISI that affects the last P samples in the symbol. Consider the example in Figure 3. The transmitted samples 2 and 10 are identical. However, at the receiver, the interfering samples before sample 2 are not all equal to their counterparts before sample 10. If h_2, h_3, and h_4 were zero, then $y(2) = y(10)$. Trying to force $y(2) = y(10)$ should force $h_2 = h_3 = h_4 = 0$, thus forcing the effective channel to be as short as the CP. The location of the window of P non-zero taps can be varied by comparing $y(3)$ to $y(11)$, or $y(4)$ to $y(12)$, etc.

The MERRY cost function is

$$J = \mathrm{E}\left|y(Mk + P + \delta) - y(Mk + P + N + \delta)\right|^2, \quad \delta \in \{0, \ldots, M - 1\}, \tag{17}$$

where δ is the desired delay. A stochastic gradient descent of (17) leads to the blind, adaptive MERRY algorithm:

$$\boxed{\begin{aligned} &\text{For symbol } k = 0, 1, 2, \ldots, \\ &\quad \tilde{\mathbf{r}}(k) = \mathbf{r}(Mk + P + \delta) \\ &\qquad\qquad - \mathbf{r}(Mk + P + N + \delta) \\ &\quad e(k) = \mathbf{f}^T(k)\ \tilde{\mathbf{r}}(k) \\ &\quad \hat{\mathbf{f}}(k+1) = \mathbf{f}(k) - \mu\ e(k)\ \tilde{\mathbf{r}}^*(k) \\ &\quad \mathbf{f}(k+1) = \frac{\hat{\mathbf{f}}(k+1)}{\|\hat{\mathbf{f}}(k+1)\|} \end{aligned}} \tag{18}$$

where $\mathbf{r}(i) = [r(i), r(i-1), \ldots, r(i - L_f)]^T$, and * denotes complex conjugation. Note that a constraint (e.g. $\|\mathbf{f}\| = 1$) must be enforced in order to prevent the trivial solution $\mathbf{f} = \mathbf{0}$.

The MERRY algorithm in (18) finds the minimum eigenvector of the matrix $\mathbf{A} = \mathrm{E}\left[\tilde{\mathbf{r}}\tilde{\mathbf{r}}^H\right]$. Let $\mathbf{C}$ be the channel convolution matrix (so that $\mathbf{h} = \mathbf{Cf}$) and let $\mathbf{C}_{wall}$ be obtained from $\mathbf{C}$ by removing rows δ through $\delta + P - 1$. If the input $s(k)$ is white, then $\mathbf{A} = \mathbf{C}_{wall}^T \mathbf{C}_{wall}$, and

$$J_\delta = 2\ \sigma_s^2 \left(\sum_{j=0}^{\delta-1} |h_j|^2 + \sum_{j=\delta+P}^{L_h} |h_j|^2 \right) + 2\ \mathbf{f}^T \mathbf{R}_n \mathbf{f}^*, \tag{19}$$

the energy of the effective channel outside the window plus the noise gain (Martin *et al.*, 2002*b*).

If a blind, non-adaptive channel shortener is required, then the matrix $\mathbf{A} = \mathrm{E}\left[\tilde{\mathbf{r}}\tilde{\mathbf{r}}^H\right]$ can be estimated from the data, and its eigenvector corresponding to its minimum eigenvalue can be computed. This may also provide an initialization technique that avoids slow modes of convergence.

MERRY can be extended to compare multiple samples in the CP to multiple samples at the end of the symbol. As shown in Figure 4, each difference term that is added to the cost function produces a different window with a different delay, and the (somewhat smaller) overall window is the union of the individual windows. This allows the option of using more data, increasing the convergence rate at the expense of over-shortening the channel. This cousin to MERRY is called Forced Redundancy with Optional Data Omission (FRODO). Figure 4 shows an example in which $P = 4$. In this case, comparing three of the four points in the CP to their brethren at the end of the symbol yields a union of three "don't care" windows, in which the impulse reponse doesn't matter so long as it is non-zero. Thus, three times as much data can be used per update, but the shortened channel will be much smaller than necessary. Since the best solution can be found by constraining the filter as little as possible, the "over-shortened" FRODO solution is sub-optimal.

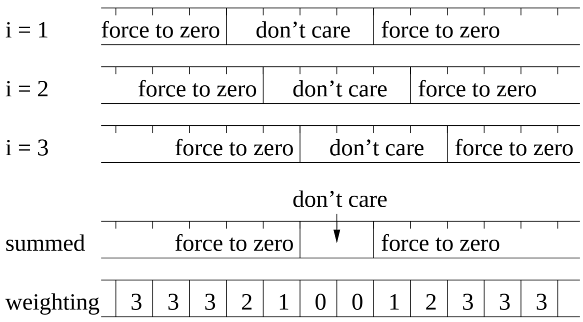


Fig. 4. The relation of the "don't care" windows in the different terms of the FRODO cost function, for $P = 4$. The line "summed" indicates the effect of considering three terms at once, and the line "weighting" indicates how much emphasis the total cost function places on forcing each tap to zero.

4.2 The SAM Algorithm

The SAM algorithm (Sum-squared Auto-correlation Minimization) relies on fourth-order statistics of the received data rather than on properties of multicarrier modulation. The idea is that if the effective channel ($\mathbf{h} = \mathbf{c} \star \mathbf{w}$) is short, its autocorrelation should be short:

$$R_h(l) = \sum_{k=0}^{L_h} h_k h_{k-l} \cong 0, \quad |l| > P. \tag{20}$$

This suggests the cost function

$$\hat{J} = \sum_{l=P+1}^{L_h} |R_h(l)|^2, \tag{21}$$

again with a constraint such as $\|\mathbf{f}\| = 1$ to prevent $\mathbf{f} = \mathbf{0}$. If the source $s(k)$ and channel noise $w(k)$ are white, and if $L_f \le P$, then

$$R_y(l) = \mathrm{E}[y(n)y(n-l)] = R_h(l), \quad |l| > P, \tag{22}$$

allowing use of the cost function

$$J_{sam} = \sum_{l=P+1}^{L_h} |R_y(l)|^2. \tag{23}$$

If $L_f > P$, (22) is still approximately true so long as the noise is small.

A gradient descent of (23) leads to

$$\boxed{\begin{aligned} \mathbf{f}(n+1) = \mathbf{f}(n) - \mu \sum_{l=P+1}^{L_h} &\mathrm{E}\left[y(n)y(n-l)\right] \\ &\cdot \mathrm{E}\left[y(n)\mathbf{r}_{n-l} + y(n-l)\mathbf{r}_n\right] \end{aligned}} \tag{24}$$

For implementation, the expectations can be replaced with instantaneous, moving average, or auto-regressive estimates. The latter yields the fastest convergence rate for the lowest computational cost (Balakrishnan *et al.*, 2002).

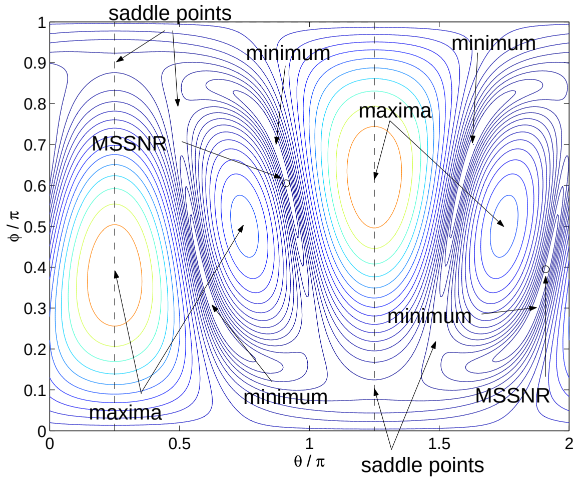


Fig. 5. Contours of the SAM cost function. The two circles are the global maxima of the shortening SNR.

Since (23) involves fourth-order statistics of the data, it is multimodal and difficult to analyze. The auto-correlation is invariant to flipping the filter's zero locations over the unit circle, so there are as many as 2^{L_f} minima that all have the same value of the SAM cost. However, they have very different values of whetever the true cost function is, e.g. MSE, bit rate, or bit error rate. When the filter length is reasonably long [2], the convergence of SAM does not appear to be troubled by the multimodality. One odd effect is that the filter $\mathbf{f}$ is generally symmetric, perhaps because time-reversing $\mathbf{f}$ is equivalent to flipping its zeros over the unit circle (which does not change the cost).

A plot of the SAM cost function is shown in Figure 5. The channel is $\mathbf{c} = [1, 0.3, 0.2]$, $P = 1$ (so a 2-tap channel is desired), there is no noise, and the 3-tap filter $\mathbf{f}$ satisfies $\|\mathbf{f}\| = 1$. $\mathbf{f}$ can be represented in spherical coordinates as $f_0 \triangleq f_x = \cos(\theta)\sin(\phi)$, $f_1 \triangleq f_z = \cos(\phi)$, $f_2 \triangleq f_y = \sin(\theta)\sin(\phi)$. Then time-reversing $\mathbf{f}$ is equivalent to reflecting θ over $\frac{\pi}{4}$ or $\frac{5\pi}{4}$, and $\mathbf{f} \to -\mathbf{f}$ is equivalent to the combination of reflecting ϕ over $\frac{\pi}{2}$ and adding π to θ (mod 2π). The four minima all have equivalent values of the SAM cost, due to the equivalencies of $\mathbf{f} \Leftrightarrow -\mathbf{f}$ and of time-reversing $\mathbf{f}$.

The circles in Figure 5 correspond to the Maximum Shortening SNR (MSSNR) design (Melsa *et al.*, 1996), which maximizes the ratio of the energy inside the window to the energy outside the window. Two of the global minima of the SAM cost nearly match the global maxima of the SSNR. The other two minima can be avoided by reversing the order of taps in the final settings for $\mathbf{f}$.

[2] "Reasonably long" here means possibly shorter than the channel, but comfortably longer than the minimum length needed to achieve a "good" solution.

5. CONCLUSION

Although blind, adaptive equalization algorithms are widely studied, hitherto there has been little academic attention given to blind, adaptive channel shortening algorithms. Traditional approaches to making adaptive equalization algorithms blind cannot be applied to channel shortening, since a channel shortener's output has different signal properties than an equalizer's output. This paper has reviewed some of the salient points of several new blind channel shortening algorithms.

REFERENCES

Balakrishnan, J., R. K. Martin and C. R. Johnson, Jr. (2002). Blind, Adaptive Channel Shortening by Sum-squared Auto-correlation Minimization (SAM). In: *Proc. 36th Asilomar Conf. on Signals, Systems, and Computers.*

de Courville, M., P. Duhamel, P. Madec and J. Palicot (1996). Blind equalization of OFDM systems based on the minimization of a quadratic criterion. In: *Proc. Int. Conf. on Comm.* Dallas, TX. pp. 1318–1321.

Falconer, D. D. and F. R. Magee (1973). Adaptive Channel Memory Truncation for Maximum Likelihood Sequence Estimation. *Bell Sys. Tech. Journal* **52**(9), 1541–1562.

Jones, D. (2003). Property-Restoral Algorithms for Blind Equalization of OFDM. Submitted to *Proc. 37th Asilomar Conf. on Signals, Systems, and Comp.* Pacific Grove, CA.

Martin, R. K., J. Balakrishnan, W. A. Sethares and C. R. Johnson, Jr. (2002*a*). A Blind, Adaptive TEQ for Multicarrier Systems. *IEEE Signal Proc. Letters* **9**(11), 341–343.

Martin, R. K., J. Balakrishnan, W. A. Sethares and C. R. Johnson, Jr. (2002*b*). Blind, Adaptive Channel Shortening for Multicarrier Systems. In: *Proc. 36th Asilomar Conf. on Signals, Systems, and Comp.* Pacific Grove, CA.

Melsa, P. J. W., R. C. Younce and C. E. Rohrs (1996). Impulse Response Shortening for Discrete Multitone Transceivers. *IEEE Trans. on Comm.* **44**(12), 1662–1672.

Pollet, T., M. Peeters, M. Moonen and L. Vandendorpe (2000). Equalization for DMT-Based Broadband Modems. *IEEE Communications Magazine* **38**(5), 106–113.

Romano, F. and S. Barbarossa (2003). Non-data Aided Adaptive Channel Shortening for Efficient Multi-carrier Systems. In: *IEEE International Conf. on Accoustics, Speech, and Signal Processing.* Hong Kong.

Treichler, J. R., C. R. Johnson, Jr. and M. G. Larimore (2001). *Theory and Design of Adaptive Filters.* Prentice Hall. New York.

Widrow, B. and S. D. Stearns (1985). *Adaptive Signal Processing.* Prentice-Hall. Englewood Cliffs, N.J.