

Research Article

Performance of Post-Demodulator Adaptive Filters for FSK Signals in a Multipath Environment

Shu-Ting Lee,¹ Sally L. Wood,¹ Michael J. Ready,² and John R. Treichler²

¹ Department Of Electrical Engineering, Santa Clara University, Santa Clara, CA 95053, USA

² Corporate Division, Applied Signal Technology, Sunnyvale, CA 94086, USA

Correspondence should be addressed to Shu-Ting Lee, slee1@scu.edu

Received 2 July 2010; Revised 17 October 2010; Accepted 11 December 2010

Academic Editor: Antonio Napolitano

Copyright © 2011 Shu-Ting Lee et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

FSK continues to be an important component of modern communication systems, and discoveries of lower impact methods for mitigating performance degradation due to multipath propagation are needed in many application areas. Previous work suggested the benefit of a simple post-demodulator LMS filter but the focus was narrow and analysis was hindered by the nonlinearity of the demodulation process and the nonstationary signal environment. This paper significantly extends understanding of the post-demodulator filter and demonstrates that the statistical assumptions that align BER performance with MSE performance do not apply in this context, and that simply decreasing the MSE might increase the BER rather than decrease it. An alternative post-demodulator adaptive filter with similar complexity to the LMS filter is proposed which shows improved BER performance. Analysis provides a simplified interpretation of the interaction of system and channel model parameters which may form the basis for more general use of a post-demodulator adaptive filter.

1. Introduction

Frequency shift keying (FSK) has been used since virtually the dawn of both wired and wireless communications to carry data signals over both short and long distances [1]. The reasons for its continued use in the modern context of wireless communication and optical fiber are the same as for its initial use. It is simple to generate, it is robust in the presence of noise, and it is resistant to degradations induced by nonlinear amplifiers and other circuit elements [2–5]. These attributes will motivate its use in future telecommunications systems as well. As transmission speeds grow, however, FSK-based radio systems must address another form of impairment as well, which is self-interference induced by multipath propagation.

Most modern receivers include adaptive equalizers to mitigate multipath-induced dispersion which is especially prevalent in urban areas. In the case of frequency-modulated signals, a constant modulus algorithm (CMA) adaptive filter is typically placed between the received signal and the FM demodulator to minimize envelope fluctuations caused by a multipath channel [6–8]. To be effective the CMA filter

length must be long enough to reasonably represent the inverse channel filter, which is typically about three times the delay spread of the channel [8, 9]. For a multipath channel with long-delay reflections, the combined effect of a long filter length, complex multiplications for coefficient updates, and slow adaptation time [10] may be too heavy a computational burden in applications with limited power or computing resources. The alternative of optimal detection using a maximum likelihood detector [11] or the Viterbi algorithm [12, 13] is also computationally intensive.

Ready and Harp [14] introduced the novel alternative of a post-Demodulator (post-D) architecture for FSK signals by placing a real-valued adaptive filter after the nonlinear demodulator block in the receiver. Although this adaptive filter, driven by the least mean square (LMS) adaptation algorithm of Widrow and Hoff [15], could not function as a traditional equalizer, empirical results showed significant performance and computational advantages compared to the commonly used pre-Demodulator (pre-D) architecture using the same amount of computation. Results were demonstrated for several multipath channels derived from physical measurements of signals received in multipath

environments. Ready and Harp's simulations with these channel models representing moderate and severe multipath indicated that relatively short post-D filters improved performance even when the filter length was much shorter than a channel inverse filter. However, they did not investigate performance for arbitrary multipath channels.

The post-D filter approach is counterintuitive in several ways. This application is not a typical linear channel equalization or system identification application because the nonlinear frequency demodulator between the channel and the adaptive filter destroys any linear relationship between system parameters and the statistical characteristics of the received signal. However, it is possible that although normal assumptions of linearity and stationarity do not apply, there may be signal correlations that could be used to advantage by an LMS filter.

Although results from some special cases have been reported [14, 16], a unified understanding of the behavior and performance of a post-D filter is still an open question. The shorter filter cannot directly compensate for long channels, and because it operates on a real-valued signal, it cannot match the spectrum of the complex envelope signal which does not have complex conjugate symmetry. Chung et al. [16] used a time varying model to bound the behavior of the post-D LMS adaptive filter. However that analysis only derived bounds on the time variation for channel models limited to a single reflection with a very small magnitude and delay. It did not provide any insight about the function of the post-D filter or the expected performance for longer delays or stronger reflections.

For the post-D architecture to become a useful component for system design, the performance benefit of using it must be understood as a function of design parameters across a wide range of typical application environments. In FSK applications, the impact of noise and interference from multipath addition of delayed signals before demodulation is difficult to model analytically because the nonlinearity of the demodulator block creates complex interrelationships between parameters of the channel model, the added noise, and the modulation. Because of the nonlinear demodulator block, the impact of individual parameters cannot be analyzed in isolation, and a strategy is needed to make useful generalizations in this nonstationary nonlinear context [17].

In this paper we have explored the general behavior of a post-D LMS filter and derived and demonstrated results that can be applied to longer channels with more severe multipath. Our work initially uncovered the lack of validity of the assumption in previously reported results that the minimum mean squared error (MSE) could be used as a performance measure which would correspond to a minimum bit error rate (BER) condition. While this assumption is valid in linear systems where Gaussian statistical models are reasonable, in the nonstationary nonlinear environment of the post-D filter a number of unexpected and discouraging characteristics of the relationship between MSE and BER were discovered. Investigation of the relationship between BER for multipath channels and typical measures of channel impact led to a better understanding of equivalent sets of parameters and

simplified further analysis for a wide range of mild and moderate channel models.

We show that simply inserting an LMS adaptive filter after the demodulator may not result in minimizing the BER despite uncomplicated convergence to the correct minimum MSE. There is a wide variation in performance depending on the channel model, and, in all but the mildest multipath models, there is danger that the post-D filter can increase the BER. Based on further analysis, two different modifications of the basic LMS filter are developed which are not computationally demanding. Both improve performance compared to the LMS filter, and one shows significant improvement in BER in both training and decision-directed operation.

2. Background

In general, an FSK signal shifts the carrier frequency by an amount $f_k = 1/2 \cdot \Delta f \cdot b[k]$, where Δf is the frequency deviation and the message symbols are $b[k] = \pm 1, \pm 3, \dots, \pm(M-1)$. When this type of FSK signal is implemented by switching among oscillators with desired frequencies, it requires a relatively large bandwidth for accurate transmission compared to the more efficient continuous-phase FSK (CPFSK) [18–20] obtained as a special case of continuous phase modulation (CPM) [21]. CPM signals exhibit the cyclostationarity property which poses research challenges for the design and analysis of communication systems, such as blind equalizers [22–25]. The CPM signal starts with a pulse amplitude modulation (PAM) signal as shown in (1) where T is the baud interval and $g(t)$ is the pulse shaping filter selected to limit the bandwidth of the signal:

$$x(t) = \sum_k b[k]g(t - kT). \quad (1)$$

A raised cosine (RC) pulse-shaping filter is often chosen for $g(t)$ to fit Nyquist's criterion for zero intersymbol interference (ISI) [26]. In most communication systems, the RC pulse shape is factored into two square-root-raised cosine (SRRC) pulse shapes as shown in (2) for bandwidth limitation at both the transmitter and the receiver:

$$g(t) = \frac{4\gamma}{\pi\sqrt{T}} \frac{\cos((1+\gamma)\pi t/T) + T \sin((1-\gamma)\pi t/T)/(4\gamma t)}{1 - (4\gamma t/T)^2}. \quad (2)$$

An infinite duration-matched SRRC filter pair minimizes the effects of ISI and channel noise [27]. However, in a practical implementation, the pulse shape is windowed, which results in nonzero side lobes in the frequency domain. The roll-off factor, γ , with $0 \leq \gamma \leq 1$ determines the excess bandwidth of the filter.

From (1) the integral of $x(t)$ scaled by the modulation index h is used in (3) to generate $\varphi(t)$, the phase of the complex envelope $s(t)$ in (4) which has a normalized

amplitude of 1. This complex envelope modulates a carrier frequency to produce the transmitted signal:

$$\varphi(t) = 2\pi h \int_{-\infty}^t x(\tau) d\tau, \quad (3)$$

$$s(t) = e^{j\varphi(t)}. \quad (4)$$

A baseband equivalent multipath channel $c(t)$ for N_d indirect paths can be modeled by the impulse response shown in (5) where the reflection coefficients, α_i , are in general complex and t_i represents the delay to the i th reflection. The channel model has been normalized for a unity gain for the direct path:

$$c(t) = \delta(t) + \sum_{i=1}^{N_d} \alpha_i \delta(t - t_i). \quad (5)$$

At the receiver it is assumed that the carrier frequency is accurately detected and used to recover $r(t)$ in (6), which is the sum of the complex envelope of the transmitted signal convolved with the baseband equivalent multipath channel model and an added noise, $\eta(t)$. It is also assumed that the additive noise is uncorrelated with $s(t)$. Here $\theta(t)$ is the instantaneous phase of the signal observed at the receiver. The received signal can be frequency demodulated using (7), where $r(t)$ has been written in terms of its real and imaginary parts as $r(t) = r_R(t) + jr_I(t)$:

$$r(t) = s(t) * c(t) + \eta(t) = a(t)e^{j\theta(t)}, \quad (6)$$

$$\begin{aligned} \hat{x}(t) &= \frac{1}{2\pi h} \cdot \frac{d}{dt}(\theta(t)) = \frac{1}{2\pi h} \cdot \frac{d}{dt} \tan^{-1} \left(\frac{r_I(t)}{r_R(t)} \right) \\ &= \frac{1}{2\pi h} \cdot \frac{1}{|r(t)|^2} (r_R(t)\dot{r}_I(t) - r_I(t)\dot{r}_R(t)). \end{aligned} \quad (7)$$

3. Post-D Filter

A discrete time receiver samples $r(t)$ with a sampling interval $T_s = T/L$ to produce $r[n]$ as shown in (8). The discrete time demodulator in (9) uses digital filters to estimate the derivatives to implement the demodulator of (7). The noise and multipath interference can cause variation in the amplitude of $r[n]$, and a limiter follows the demodulator to protect against very large output values due to an infrequent very small amplitude of $r[n]$. We assume for simplicity that the receiver knows the carrier frequency and there is no carrier drift, so only the channel and noise alter $r[n]$:

$$r[n] = s[n] + \sum_{i=1}^{N_d} \alpha_i s[n - n_i] + \eta[n], \quad (8)$$

$$\hat{x}[n] = \frac{L}{2\pi h} \cdot \frac{r_R[n] \cdot \dot{r}_I[n] - r_I[n] \cdot \dot{r}_R[n]}{|r[n]|^2}. \quad (9)$$

To perform simulations a corresponding discrete time model for $s[n]$ following the models of (1) to (4) is

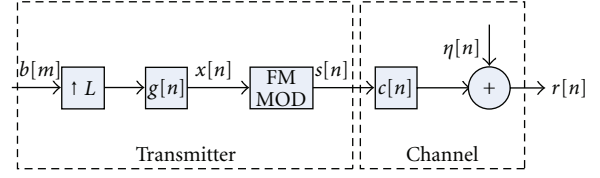


FIGURE 1: Discrete time transmitter and channel models.

needed. A block diagram of this model is shown in Figure 1. The message signal $b[m]$ is upsampled by factor of L for a sampling interval of T_s to create $b_u[n]$ which is then convolved with the sampled pulse-shaping filter $g[n]$ to generate $x[n]$ as shown in (10). The pulse-shaping filter has been windowed to $2K$ baud intervals. The trapezoidal discrete time approximation to integration is used in (11) to compute the sampled complex envelope signal $s[n]$ given by (12). Then, using a known carrier frequency, the complex envelope signal $s[n]$ creates the transmitted signal:

$$\begin{aligned} x[n] &= \sum_{m=0}^{N-1} b[m]g[n - mL] \\ &= \sum_{k=-KL}^{KL} g[k]b_u[n - k], \end{aligned} \quad (10)$$

$$\varphi[n] = 2\pi h \frac{T}{L} \left(0.5 \cdot x[n] + \sum_{k=-\infty}^{n-1} x[k] \right) \quad (11)$$

$$s[n] = e^{j\varphi[n]}. \quad (12)$$

The pulse-shaping filter for the upsampled data signal can be interpreted as L polyphase filters with length reduced by a factor of L which creates a cyclostationary output $x[n]$. Using $g_i[k] = g[kL + i]$, (10) can be rewritten as (13) where $K_1 = \text{floor}((KL - i)/L)$ and i indexes the L polyphase filters from 0 to $L - 1$:

$$x[mL + i] = \sum_{k=-K}^{K_1} g_i[k]b[m - k]. \quad (13)$$

The discrete channel $c[n]$ defined in (14) is obtained from (5):

$$c[n] = \delta[n] + \sum_{i=1}^{N_d} \alpha_i \delta[n - n_i]. \quad (14)$$

Performance of a system is measured by BER where errors may be due to the added noise, $\eta[n]$, and the frequency selective effects of the channel, $c[n]$. For 2-level FSK, let $\hat{b}[m]$ be the message signal obtained from $\hat{x}[n]$. The BER is given by

$$P_e = \frac{1}{N} \sum_{m=0}^{N-1} \frac{1}{2} |b[m] - \hat{b}[m]|. \quad (15)$$

A post-D adaptive filter shown in Figure 2 will use the LMS algorithm [17] to minimize the mean-squared value of $e[m]$. Let \mathbf{w}_n be the filter of L_f coefficients which produces output $y[n]$ using input $\hat{\mathbf{x}}_n$, the n th snapshot of the input

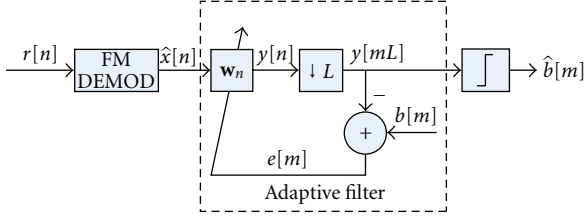


FIGURE 2: FM demodulator and post-D adaptive filter.

data vector given by (16). This snapshot holds ℓ_0 data samples following $\hat{x}[n]$ corresponding to a bulk delay to allow for cases in which the amplitude of a reflection with $n_i < \ell_0$ is larger than the direct path amplitude:

$$\hat{\mathbf{x}}_n = [\hat{x}[n + \ell_0], \dots, \hat{x}[n], \dots, \hat{x}[n - L_f + 1 + \ell_0]]^T. \quad (16)$$

The adaptive filter output $y[n]$, given by (17), is subtracted from the desired output $b[m]$ to create an error $e[m]$ given by (18):

$$y[n] = \mathbf{w}_n^T \hat{\mathbf{x}}_n, \quad (17)$$

$$e[m] = b[m] - y[mL]. \quad (18)$$

Here $y[mL]$ is the downsampled output of the post-D filter. At every L th sample, the weight vector is updated [17] using (19) to minimize the average squared error, $e_s = E[e^2[m]]$. The adaptation constant μ controls the rate of convergence and misadjustment error:

$$\mathbf{w}_{n+L} = \mathbf{w}_n + \mu \cdot e \left[\frac{n}{L} \right] \cdot \hat{\mathbf{x}}_n. \quad (19)$$

Let $\mathbf{R}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}$ be the autocorrelation matrix of the input data vector $\hat{\mathbf{x}}_n$ computed from only the data vectors used in (19). Due to the cyclostationary properties caused by the upsampled data signal and polyphase pulse shaping filter, $\mathbf{R}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}$ will not in general be Toeplitz. Following the typical analysis of an LMS adaptive filter, for convergence approaching the Wiener filter, \mathbf{w}_{LS} , as shown in (20), it is necessary that the adaptation constant μ be less than $2/\lambda_{\max}$ [17], where λ_{\max} is the maximum eigenvalue of $\mathbf{R}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}$. Here \mathbf{p} is the cross-correlation vector between the input $\hat{\mathbf{x}}_n$ and the desired output $b[n/L]$ as shown in (21) where $E[\cdot]$ denotes the stochastic expectation. For a decision-directed feedback system, the estimated message bit, $\hat{b}[m]$, made by a slicer can replace $b[m]$ in (18):

$$\mathbf{w}_{LS} = \mathbf{R}_{\hat{\mathbf{x}}\hat{\mathbf{x}}}^{-1} \cdot \mathbf{p}, \quad (20)$$

$$\mathbf{p} = E \left[\hat{\mathbf{x}}_n \cdot b \left[\frac{n}{L} \right] \right]. \quad (21)$$

For long channels, it is reasonable to assume that there should be no correlation between $s[n]$ and $s[n + n_1]$ for a finite duration pulse shaping filter and independently distributed message bits. From (14) for $N_d = 1$, the difference between the received and transmitted signal is the reflection and noise term, $\alpha_1 s[n - n_1] + \eta[n]$, which may cause a phase

and amplitude difference between $r[n]$ and $s[n]$. If $s[n - n_1]$ is not correlated with $s[n]$, then the expected value of the magnitude of the received signal error will be the magnitude of $\alpha_1 s[n - n_1] + \eta[n]$. Since $s[n - n_1]$ and $\eta[n]$ are assumed to be uncorrelated, the power of the interference term can be represented by (22) where σ^2 denotes the power of the added noise. Thus for long channels we expect the performance to strongly depend on $|\alpha_1|$ but have little dependence on n_1 :

$$E[|\alpha_1 s[n - n_1] + \eta[n]|^2] = |\alpha_1|^2 + \sigma^2. \quad (22)$$

4. Results

Analysis and simulation studies were used to address the questions of when and if a post-D adaptive filter could provide performance improvement in a multipath environment, how much improvement could be expected, and what generalizations would be possible in this nonlinear context. Because of the nonlinear characteristics of the FM demodulator between the channel and the adaptive filter, it cannot be assumed that the performance impact of channel parameters, SNR levels, modulation index, and adaptive filter structure could be studied independently. Simulation experiments were designed to determine the effects of the interaction of the parameters and to validate analytical predictions.

To systematically investigate the relationship of the many parameters to MSE and BER performance, the discrete time models in Figures 1 and 2 were used. A square-root-raised cosine filter of width $12T$ was used for $g[n]$. Although Chung et al. [16] implied that the choice of modulation index was not significant, results for two different modulation indices were compared for most of the questions addressed below. One modulation index, h_1 , was selected to be close the modulation index which produces Minimum-shift keying (MSK) [28] and the other, h_2 , expanded the signal bandwidth to approach the limit of the unaliased bandwidth at a sampling interval of $T_s = T/2$. MSK, a special case of CPFSK signaling with frequency deviation equal to 1/2 the bit rate, has been investigated and shown to have equivalent bandwidth and error probability to Offset Quadrature Phase Shift Keying (OQPSK) while shaping and filtering the data pulses [29].

The first question addressed was whether any generalizations could be made about the impact on BER of channel length or number of reflections. Previous analytical work [16] has focused on a model with a single reflection and a short delay of $2T_s$ or one symbol interval. However, because of the nonlinear demodulation block, it is not clear that conclusions drawn from the case of a single reflection in a short channel would apply to channels with multiple reflections and more realistic lengths. Without some way to draw general conclusions about length or number of reflections, it would be almost impossible to develop general design rules for use of a post-D filter.

To empirically explore the performance impact of the number of reflections in a multipath channel, the BER was measured for simulated random channels using the channel model of (14) with N_d set to 1, 2, 6, or 10 at several SNR

levels. For each channel three general measures of channel impact, the total power of the multipath reflections, P_M , the mean delay τ_d , and the rms delay spread, τ_{rms} , were computed using the definitions in (23) [30, 31]. Each of the N_d multipath reflections of the baseband equivalent channel $c[n]$ is defined by a complex-valued reflection coefficient α_k and a delay n_k . To simulate both mild and moderate multipath channels, the amplitude of the reflections was uniformly distributed between 0.1 and 0.5 with phase uniformly distributed over 0 to 2π . The delay intervals from the main signal to the multipath reflection were uniformly distributed over the interval $10T_s$ to $100T_s$:

$$P_M = \sum_{k=1}^{N_d} (|\alpha_k|^2),$$

$$\tau_d = \frac{1}{P_M} \sum_{k=1}^{N_d} |\alpha_k|^2 n_k T_s, \quad (23)$$

$$\tau_{rms} = \sqrt{\left(\frac{1}{P_M} \sum_{k=1}^{N_d} (|\alpha_k| n_k T_s)^2 \right) - \tau_d^2}.$$

Comparison of the observed average BERs for each random channel with the corresponding channel measures showed no clear relationship between the BER and either the mean delay τ_d or the rms delay spread τ_{rms} . However, in all cases the total noise and interference power, P_t , was a good predictor of BER, where $P_t = P_M + P_\eta$ is the sum of the noise power, P_η , and the total power of the multipath reflections, P_M . For an SNR of 15 dB, Figure 3 shows a scatter plot of the BER at the demodulator output of each random channel as a function of P_t . The symbols of the plotted points identify the number of reflections in the channel model. As the number of reflections increases, P_M tends to increase on average. The plot shows a smooth trajectory of overlapping regions dominated by a specific number of reflections indicating that the BER is clearly dependent on the magnitude of the reflection coefficients and the noise level but relatively independent of the number of reflections or the channel delay over the range of 5 to 50 symbol intervals. Similar plots for lower values of SNR overlap this plot but do not have observation points at the lowest values of P_t because of the higher values of P_η at lower SNR levels.

Based on these results, initial investigations focused on channel models with a single reflection in the confidence that the behavior should reasonably extend to multiple reflections at the same value of P_t . In addition, an SNR of 15 dB is used for most results presented in this paper because it allows a wide range of BER performance depending on P_M . However, a range of SNR values were investigated and included in discussion of the results presented.

There is no simple analytical representation of the effect of the channel parameters on the BER of the demodulator output, which is the input to the adaptive filter, and the adaptive filter minimizes $e^2[m]$, not the BER. The performance of the LMS filter will depend on the autocorrelation of the input signal, $\mathbf{R}_{\hat{x}\hat{x}}$, and the correlation of the input signal and the desired output, \mathbf{p} , as shown in (20) as well

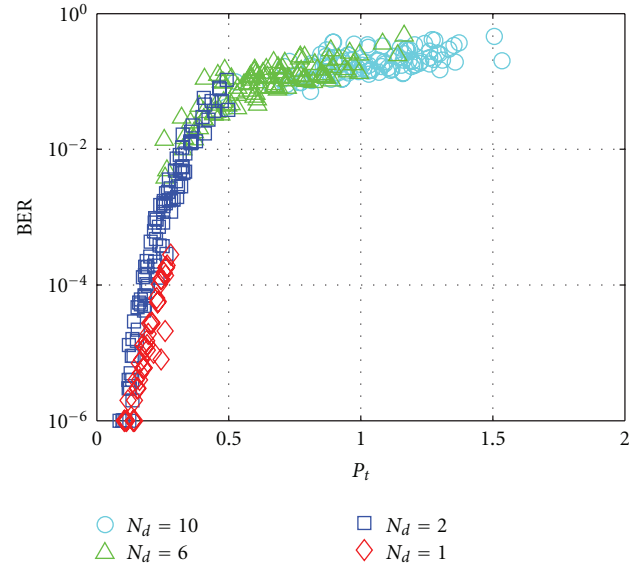


FIGURE 3: BER versus interference and noise power for an SNR of 15 dB.

as the statistical characteristics of the input signal. Although the nonlinearity of the demodulator block would make the channel equalization typical of a pre-D CMA filter impossible for a real-valued post-D filter, a post-D LMS filter could still in principle improve performance by taking advantage of any correlation of a desired signal within a time window of the received signal. The phase of the reflections and the length of the channel did not appear to have much effect on the BER compared to the effect of P_t in Figure 3, but it is reasonable to expect that they might affect the correlation functions and thus the performance of a post-D filter.

For sequences in which the message bits are i.i.d., a system with perfect modulation and demodulation and no multipath channel or noise should produce a correlation between the demodulator output and the desired message bit that is identical to the pulse shaping filter $g[n]$. However the impact of the channel and noise before the demodulator block is not easily predicted. Typical correlation between the desired message bit and the demodulator output for a signal degraded by a multipath channel with a single reflection is shown in Figure 4 for $n_1 = 20$. Results for $|\alpha_1| = 0.5$ with three different phase values are shown for both modulation index values in Figure 4(a). Results for $h = h_1$ are shown in solid lines while results for $h = h_2$ are shown in dashed lines. In all cases the correlation due to the multipath reflection is clearly visible at a lag of 20 although closer examination of the correlation in Figure 4(b) shows that the shape of the correlation function at a lag of 20 is dependent on the phase of the reflection coefficient and the modulation index.

Similar results for other values of n_1 show that the length of the channel affects the shape as well as the lag position of the secondary peaks. For example, results for $h = h_1$ when n_1 is set to 16 or 24 are very similar in shape to those shown by the solid lines in Figure 4 for $n_1 = 20$. However, results for $n_1 = 18$ would show that although the same shapes

occur, they occur at different phase values of the reflection coefficient. The results for $\alpha_1 = 0.5$ and $\alpha_1 = -0.5$, which are almost identical for $n_1 = 20$, are seen at $\alpha_1 = 0.5j$ for $n_1 = 18$. For $h = h_2$ the shapes for $\alpha_1 = 0.5$ and $\alpha_1 = -0.5$ are different at $n_1 = 20$ and correspond to the opposite values of α_1 at $n_1 = 18$. Thus, changing the phase of the reflection coefficient in the channel model can produce the same changes in the correlation function as changing the delay of the reflection or the modulation index.

For values of n_1 large enough to avoid overlap of the main and secondary peaks of the correlation function, the amplitudes and shapes of the correlation function are similar to those shown in Figure 4; so analysis of a small subset of channel models can be used to predict behavior of channels with much longer delays. It is noted that the analytical work of Chung focused on $n_1 = 2$, which is a special case because of the high degree of overlap and correlation of the added signal. Thus results for $n_1 = 2$ are not likely to be typical of results for longer channels delays, and later simulations demonstrate this. Regardless of the complex relationship between the shape of the correlation function and the channel parameters and modulation index, there is a significant level of correlation that could be used by an adaptive filter.

These observations are consistent with a mathematical representation of $\hat{x}(t)$. Consider the case of a single reflection with $\alpha_1 = |\alpha_1|e^{j\phi_1}$ from (6):

$$r(t) = e^{j\varphi(t)} + |\alpha_1|e^{j\varphi(t-t_1)+j\phi_1} + \eta(t), \quad (24)$$

$$\dot{r}(t) = j\dot{\varphi}(t)e^{j\varphi(t)} + j|\alpha_1|\dot{\varphi}(t-t_1)e^{j\varphi(t-t_1)+j\phi_1} + \dot{\eta}(t).$$

Using (7) in the noise-free case we can write $\hat{x}(t)$ in terms of $x(t)$ and $x(t-t_1)$ where from (3) $\dot{\varphi}(t) = 2\pi h x(t)$:

$$\begin{aligned} \hat{x}(t) = & \frac{x(t)(1 + |\alpha_1| \cos(\Delta\varphi(t, t_1, \phi_1)))}{1 + 2|\alpha_1| \cos(\Delta\varphi(t, t_1, \phi_1)) + |\alpha_1|^2} \\ & + \frac{x(t-t_1)(|\alpha_1|^2 + |\alpha_1| \cos(\Delta\varphi(t, t_1, \phi_1)))}{1 + 2|\alpha_1| \cos(\Delta\varphi(t, t_1, \phi_1)) + |\alpha_1|^2}. \end{aligned} \quad (25)$$

Here $\Delta\varphi(t, t_1, \phi_1) = \varphi(t) - \varphi(t-t_1) + \phi_1$. Although the factors multiplying $x(t)$ and $x(t-t_1)$ are time varying and dependent on the message sequence, the impact of the modulation index, the channel delay n_1 , the message sequence, and the phase of α_1 are always contained within $\Delta\varphi(t, t_1, \phi_1)$. These factors may have complex interactions, but the range of the impact is limited by the cosine function. However the magnitude of α_1 has impact on the relative contribution of $x(t)$ and $x(t-t_1)$ to $\hat{x}(t)$. This is consistent with observations included in Figures 3 and 4 and further justifies the validity of using a small subset of values of h , n_1 , and ϕ_1 to represent general behavior.

Adaptive filters were tested for a range of delay values for an SNR of 15 dB with $\alpha_1 = 0.5$ corresponding to moderate multipath in a reasonable operational range of performance. Simulations for reflection coefficients with

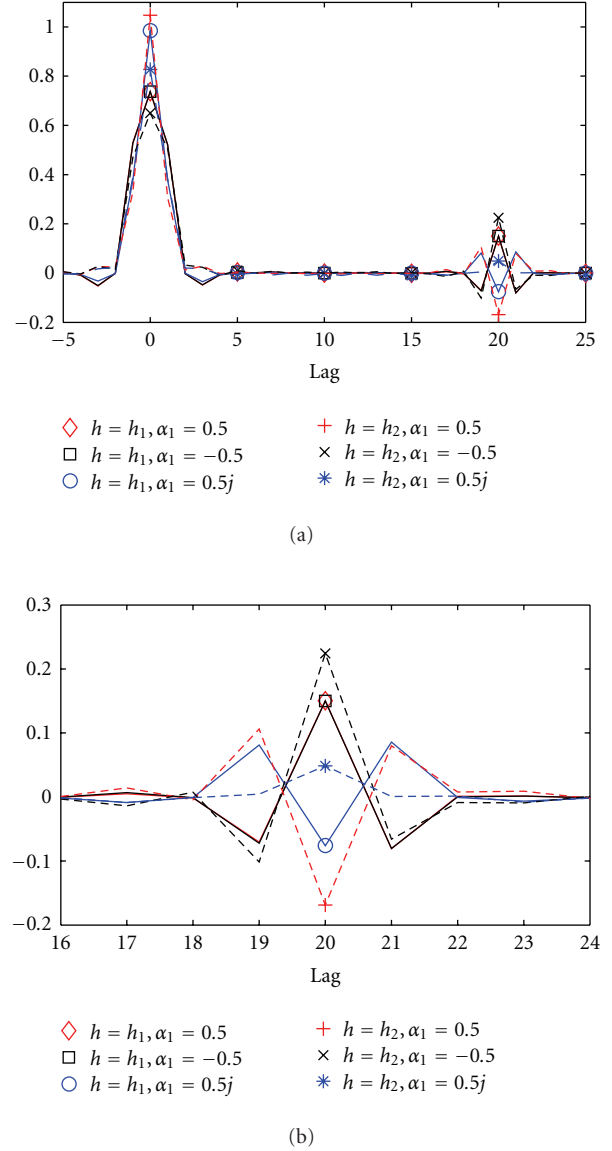


FIGURE 4: (a) Expected value of \hat{x}_n when $b[m] = +1$ for $n_1 = 20$ at SNR = 15 dB; (b) expanded scale view of (a) at a lag = n_1 .

smaller magnitudes of α_1 showed results similar to those reported here but with lower BERs. The behaviors observed as the delays vary are representative of the behaviors due to phase changes in the reflection coefficient. Adaptive filters of several lengths were placed between the demodulator output and the decision block as shown in Figure 2. The filters were initialized to have one tap set to 1.0 and all the rest set to 0.0 so that adaptation would start from the case of no post-D filtering and a fixed transport delay. To avoid numerical problems from an $r[n]$ value instantaneously close to zero, a limiter was placed at the demodulator output. The adaptation constant μ was set to 0.0005, which was the value used by Chung. However, simulations at other values of μ as well as computation of the stable range of μ showed that this was a somewhat conservative choice at the power levels of these simulations. Increasing μ by a factor of 5 resulted

in a corresponding decrease in convergence time without significant impact on convergence values or misadjustment.

Tests for four filter lengths are shown in the results that follow. A filter of length 3 with an initialization of 1.0 at the center tap was used to test the limit of the hypothesis that a very short LMS adaptive filter could produce benefits as proposed by Ready and Harp [14]. This filter spans only the large central peak of the correlation function shown in Figure 4. It is also short enough to easily allow the exact computation of the filter coefficients that produce the minimum BER. Filter lengths of 8, 16, and 32 were tested with initializations of 1.0 at the quarter length position to show the impact of filter length and to allow comparison with previously published results. Ready and Harp's work [14] indicated that a relatively short post-D adaptive filter of length 8 to 32 was found to be effective in reducing the MSE for physically observed channels with total delays much longer than the adaptive filter length. Chung et al. [16] used filter lengths of 10 for a reflection delay of $2T_s$. Although results for other filter lengths are not included in the following discussion, the results for a filter length of 8 are representative of other short filters with lengths from 5 to 10.

Multipath channel delays from 2 to 20 were selected for simulation. The lower values were used for comparison with other published work but were not expected to be representative of longer channels. For delays larger than 8 or 10, the results were expected to be similar to results for longer channels with appropriate selection of modulation index and phase of the reflection coefficient. In addition for the two longest filter lengths, this range of delay values would show the effect of the filter length compared to the extent of the correlation function.

The results of this set of simulations showed a number of unexpected results and provided some insight into the variability of previously reported results obtained when only a narrow set of examples were explored. Figure 5 shows both the BER and MSE for simulations for an SNR of 15 dB and a modulation index of $h = h_1$ using a channel with $\alpha_1 = 0.5$ and $n_1 = 10$. The MSE shown in Figure 5(b) converges for all filter lengths with the two longest filters producing the lowest values of MSE. However, the trajectories of the corresponding BERs did not follow the steadily decreasing trajectory of the MSE and instead attained a number of local maxima and minima. In Figure 5(a) the BER when no post-D filter is used is 0.00041, and after convergence of the LMS filter, all filters have an increased BER rather than a decreased BER. The length 32 filter has the lowest BER at convergence, but its BER of 0.00113 more than 2.5 times higher than the original BER. The BER for the shortest filter is more than ten times as large as the initial BER. Exact computation of the autocorrelation and crosscorrelation was used to compute the Weiner filter in (20) and verify the correct convergence in these simulations.

Figure 6 shows similar results for the case of $n_1 = 12$. Although the MSE behavior is similar to that of Figure 5, here the BER at convergence is higher for the longer filters than it is for the shorter filters. Even though it has the lowest MSE, the longest filter has the highest BER rate at convergence at 0.00092. This is just slightly lower than the initial BER of

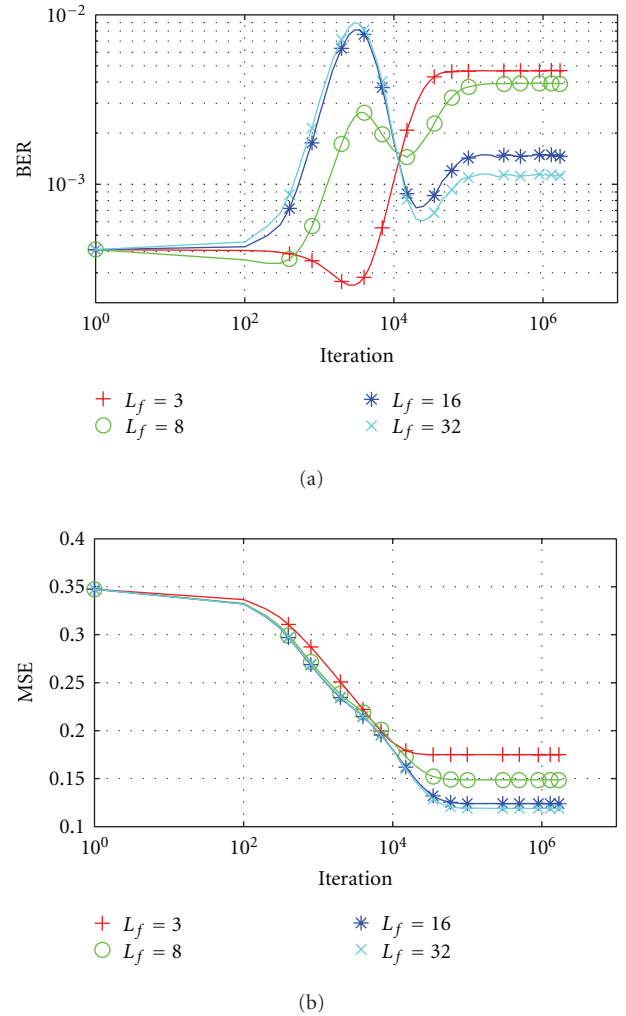


FIGURE 5: (a) BER and (b) Average MSE of four post-D LMS filters with SNR = 15 dB, modulation index $h = h_1$, and channel model $\alpha_1 = 0.5$ and $n_1 = 10$.

0.00113 and well above the minimum observed BER. The lowest converged BER is 0.00072 for the shortest filter.

These surprising results are important because previously reported analyses of the post-D filter implicitly assumed that minimizing $e^2[m]$ would minimize BER, and so only MSE or cluster variance was reported. The BER in many applications where Gaussian statistical models apply is expected to follow the MSE [32], but for a post-D filter an assumption of Gaussian statistical models is not reasonable, and previous work using MSE or cluster variance as a performance measure would miss this detrimental and unexpected behavior of the post-D filter with respect to the BER. Changing the channel delay n_1 by only two sample time intervals caused a large quantitative difference in converged BER as well as a large qualitative difference in that for one value of n_1 the longest filter was the best performer while at the other value of n_1 the shortest filter performed best. Considering these results, it is not clear how a post-D LMS filter could be used reliably in this context.

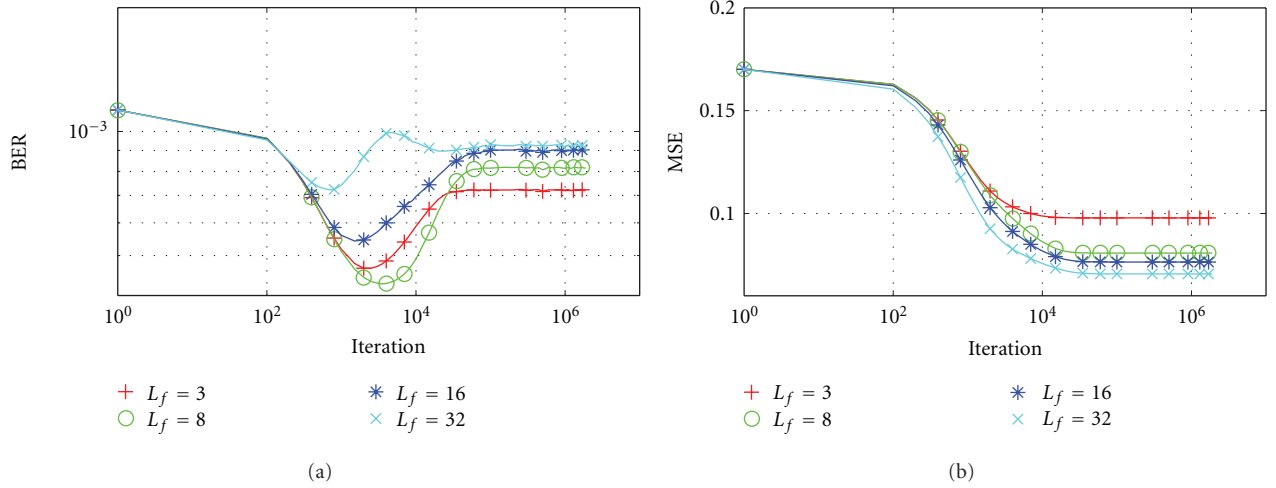


FIGURE 6: (a) BER and (b) Average MSE of four post-D LMS filters for the case in Figure 5 with $n_1 = 12$.

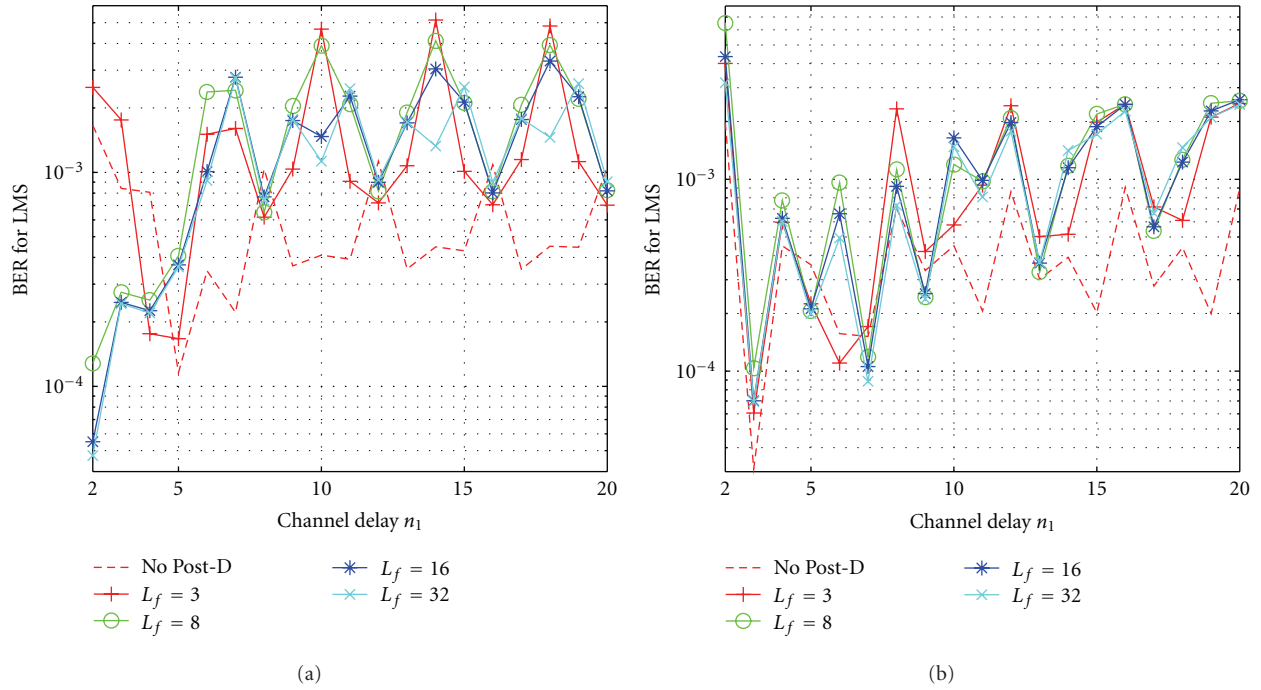


FIGURE 7: Converged LMS BER for $n_1 = 2$ to 20 for modulation indices (a) $h = h_1$ and (b) $h = h_2$.

Simulations to systematically determine the BER for converged LMS filters as a function of n_1 are shown in Figure 7 for both modulation indices at the four filter lengths shown in Figures 5 and 6. The BER for a system without a post-D filter is also included as an unmarked dashed line for reference. Figure 7(a), which uses $h = h_1$, shows a cyclic pattern of BER for values of n_1 larger than 8. This pattern is consistent with correlation results such as those shown in Figure 4. There is a wide range of BER performance as a function of n_1 , and the converged LMS filter does not always improve the BER. The performance of the filters of lengths 16 and 32 is similar until n_1 increases, above 14. As n_1 continues

to increase the performance of the length 16 filter is more similar to the length 8 filter because at those channel delay values neither filter is long enough to include the secondary peak of the correlation functions. Figure 7(b), which uses $h = h_2$, shows a different cyclic pattern of BER for values of n_1 of larger than 8. In this case there is very little difference in the performance of adaptive filters of different lengths. These cyclic patterns are consistent with (25).

These results are completely unexpected. Although the reduction in MSE and convergence is consistent with reported work [14, 16], the BER performance is poor and erratic and does not show a consistent relationship to filter

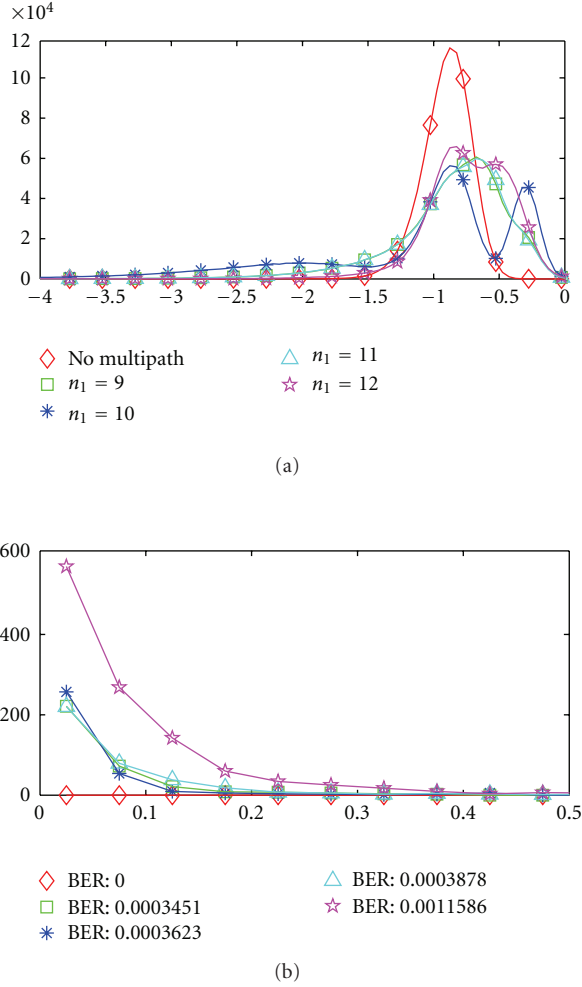


FIGURE 8: Histogram of $\hat{x}[m]$ when $b[m] = -1$ for four channel models with delays of $n_1 = 9$ to 12 , SNR = 15 dB, and modulation index $h = h_1$; the case of no multipath is shown for reference: (a) range of correct decision, and (b) range of bit errors at expanded scale.

length. It is not immediately obvious why the BER of the converged LMS filter is often higher than the initial BER before filtering. New insight into the behavior of the post-D filter can be gained by observing the histogram of the demodulator output. This data could be collected easily from an operating system and would be an estimate of the probability density function of $\hat{x}(t)$ in (25). Figure 8 shows the histogram conditioned on a message bit value of -1 for four channel models with n_1 set to 9 , 10 , 11 , and 12 . The same cyclical pattern shown in Figure 7 is seen in the histograms; so the behavior for longer channel delays is effectively included in this figure. Figure 8(a) shows the histogram for negative received values which would not cause a bit error, and Figure 8(b) shows the histogram at a magnified scale for positive demodulator outputs which would cause an error. The histogram for the case of no channel is included for reference. Results obtained conditioned on a message bit value of $+1$ are the same after reflection on the horizontal axis.

The interference from the multipath signal additions creates multimodal density functions characteristic of the lengths shown. Similar density functions are observed at other channel delay values and at other phases of the reflection coefficient. This is expected based on (25) since if $x(t - t_1) = x(t)$, then $\hat{x}(t) = x(t)$ and the value of $\Delta\varphi(t, t_1, \phi_1)$ does not affect $\hat{x}(t)$. For other values of $x(t - t_1)$ there will be a distribution of values away from $x(t)$ based on $\Delta\varphi(t, t_1, \phi_1)$. As $|\alpha_1|$ increases, this distribution becomes wider and causes more received values to fall on the wrong side of the decision boundary.

Figure 8(b) shows the histogram for the bits that caused errors, and the reason for the variability of the BERs as n_1 is changed can be seen from the difference in these plots. As the LMS filter adapts, the histogram of the outputs will migrate toward a more Gaussian shape, and as the number of values with a large deviation on both sides of the desired value is reduced, more outputs move across the error boundary.

Figure 9 shows the histograms for the post-D filter outputs for three representative delays of $n_1 = 10$, 11 and 12 at three significant times in the adaptation process: the initial time, the time of minimum BER, and convergence. In Figure 9(a) where $n_1 = 10$, the histogram has a much lower variance at convergence. However this lower variance histogram has a significant increase in positive values corresponding to an increased BER as shown in Figure 9(b). Similar effects are observed for $n_1 = 11$ and $n_1 = 12$ although they are not as dramatic as for $n_1 = 10$. For all values of n_1 the variance was reduced by the adaption process, skew was reduced although it remained negative, and kurtosis was reduced and moved closer to the Gaussian value. Thus in minimizing the MSE the adaptive filter generates an output with a probability density function that is closer to Gaussian than the filter input, but this often moves more values across the error threshold.

5. An Improved Post-D Filter

These simulations show that an LMS post-D filter will not reliably reduce BER even when successfully minimizing MSE. Although results were shown for $|\alpha_1| = 0.5$, increasing or decreasing the value of $|\alpha_1|$ shows similar behavior but at correspondingly higher or lower values of BER. As predicted by (25) the amount of the spread of the histograms compared to the case of no channel increases as $|\alpha_1|$ increases. It is clear from Figures 8 and 9 that part of the problem with the performance of the LMS filter is that the large values at the tails of the density functions which do not cause bit errors have a high cost in terms of squared error and will cause relatively larger changes in the LMS adaptation process. A new adaptive filter is proposed here which minimizes this impact by introducing a weighting function for the error. The proposed alternative adaptive filter would cause the filter to emphasize the updates when a bit error occurs compared to updates where no bit error occurs and would prevent adaptation on large values that do not cause bit errors. One

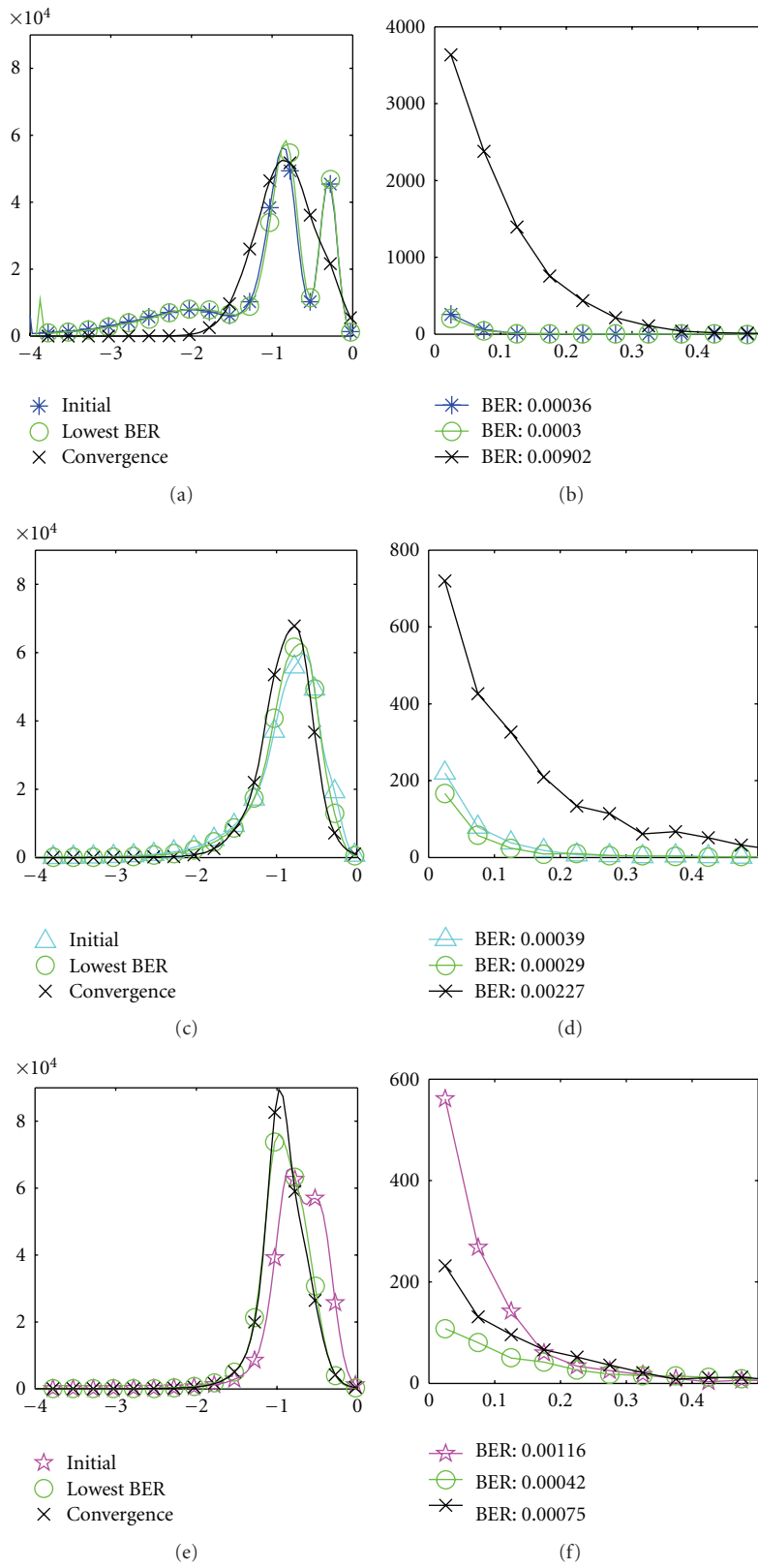


FIGURE 9: Histograms of the post-D filter outputs when $b[m] = -1$ for channels with delays of $n_1 = 10$ in (a) and (b), $n_1 = 11$ in (c) and (d), and $n_1 = 12$ in (e) and (f).

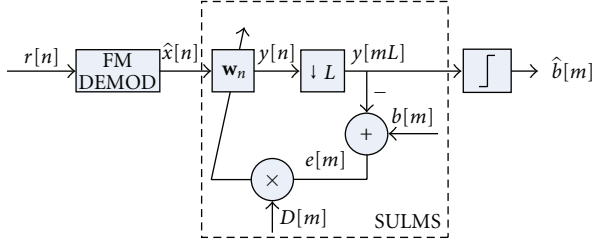


FIGURE 10: Block diagram for SULMS post-D filter.

possible method would modify the LMS cost function to create the selected update LMS (SULMS) filter defined by (26) and (27) as shown in Figure 10:

$$\mathbf{w}_{n+L} = \mathbf{w}_n + \mu \cdot D[m] \cdot e\left[\frac{n}{L}\right] \cdot \hat{\mathbf{x}}_n, \quad (26)$$

$$\rho[m] = \zeta \cdot (b[m] \cdot y[mL] - \beta), \quad (27)$$

$$D[m] = \frac{e^{-\rho[m]}}{e^{-\rho[m]} + e^{\rho[m]}}.$$

In (27), $D[m]$ is defined as a sigmoid curve controlled by the parameters ζ and β to adjust the contribution of the error $e[n/L]$. The steepness of the transition of the sigmoid curve is controlled by ζ and the position of the transition is controlled by β . The performance of this filter is relatively insensitive to the parameter ζ . If the parameter β is large compared to the width of the histogram of the demodulated values, the behavior of the SULMS is indistinguishable from the traditional LMS filter because $D[m]$ will have a value of 1 for almost all values of $e[n/L]$. As β decreases, fewer values of $e[n/L]$ contribute significantly to the adaptation which slows the convergence process. Although a wide range of values of β give good results, a value of $\beta = 0.25$ is used in results presented here. Since the histogram of demodulator output values is easily measured, the peak values observed operationally can be used to set reasonable values of parameters for the SULMS filter.

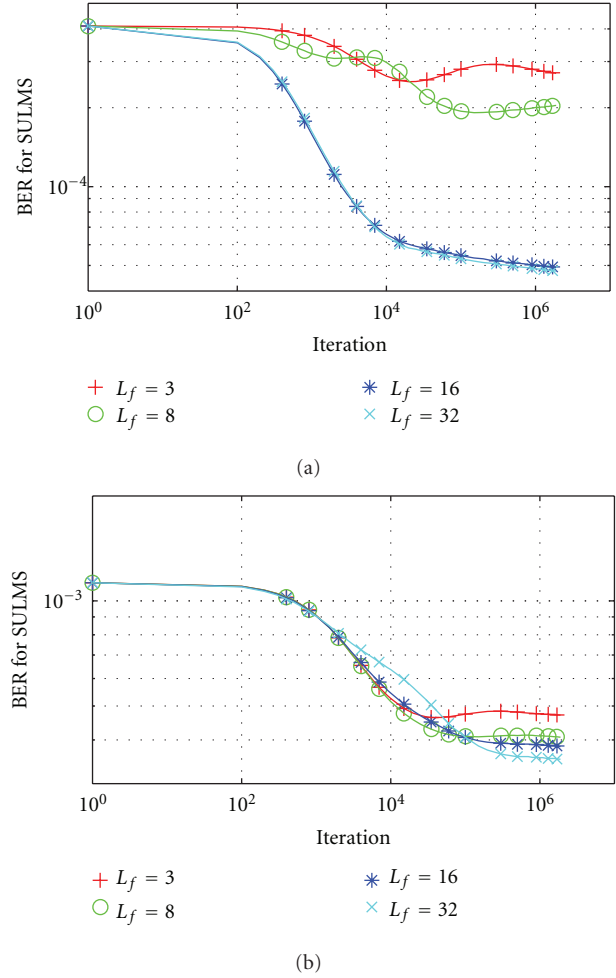
As defined in (27), $\rho[m]$ requires training data. An alternative that could be used in a decision-directed application is defined in (28):

$$\rho[m] = \zeta \cdot (|y[mL]| - \beta). \quad (28)$$

This would suppress adaptation on all large values of the demodulator output.

The trajectory of the BER as the SULMS filter adapts is shown in Figure 11, with the same channel models and noise level that were used in Figures 5 and 6 shown in Figures 11(a) and 11(b), respectively. In both cases there is significant reduction in the BER as the filter converges with the longest filter reducing the BER by almost a factor of 10. In both cases the longest filter has the lowest BER. At this SNR level results are the same for $\rho[m]$ defined by (27) or (28) and show a significant improvement compared to Figures 5 and 6.

The performance improvement of the SULMS post-D filter is plotted in Figure 12(a) as a function of delay value for the four filter lengths. The gain is defined as the ratio of

FIGURE 11: BER of four post-D SULMS filters with SNR = 15 dB, modulation index $h = h_1$, channel model $\alpha_1 = 0.5$, and (a) $n_1 = 10$, (b) $n_1 = 12$.

the BER with no post-D filter to the BER of the converged SULMS filter and is plotted on a logarithmic scale to allow comparison of gains over a wide dynamic range. For all delays the length 32 filter showed better performance than the two shortest filters. For $n_1 > 8$ the cyclic pattern of the BER for a filter of length 32 is clear. Depending on the channel delay, using the SULMS filter lowers the BER by a factor of 3 to 9. The BER for the filter of length 16 follows the value of the length 32 filter until n_1 is 12. Above that the delay is longer than the filter memory and the BER performance begins to follow a trajectory similar to the length 8 filter. For $n_1 > 8$ SULMS filters of lengths 3 and 8 also show a cyclic pattern with a maximum BER reduction of a factor of 3. These adaptive filters can take advantage of correlation introduced by the pulse shaping filter, but they are not long enough to take advantage of correlation introduced by the multipath reflection. For shorter channel delays the BER improvement is generally larger, but these cases are special because of the channel delay dependent overlap of the correlation peaks and it is not expected that the same level of improvement would be seen for longer channel delays.

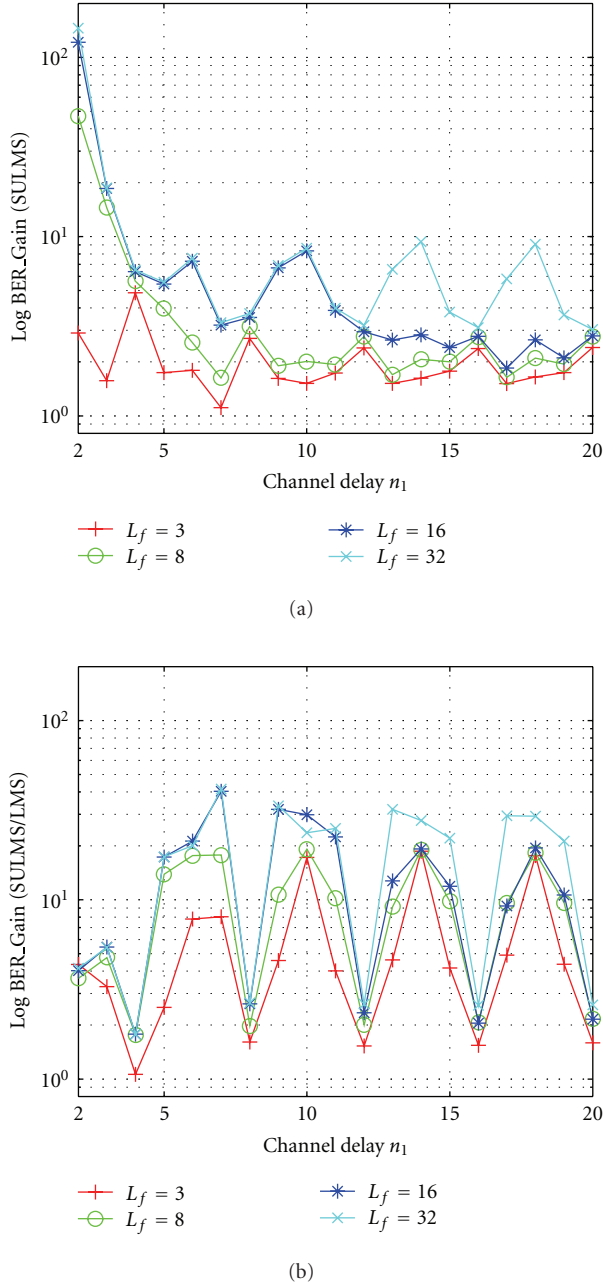


FIGURE 12: BER gain as a function of channel delay with SNR = 15 dB and modulation index $h = h_1$ for (a) converged SULMS filter compared to no post-D filter and (b) converged SULMS filter compared to LMS filter.

Importantly, there were no cases in which the SULMS filter caused the BER to increase.

The performance of the SULMS filter is compared to the conventional LMS filter in Figure 12(b). It shows BER improvements by as much as a factor of 40 because in some cases the LMS post-D filter increases the BER compared to the case where no filter is used. For all filter lengths and all delay values the SULMS filters lowered the BER of the unfiltered output, and in all cases SULMS outperformed the simple LMS filter. For tracking changes in a propagation

channel, the speed of convergence required of an adaptive filter depends on the ratio of the rate of change of the channel to the symbol rate of the signal [8]. This ratio varies widely depending on the application, but the convergence rate of the post-D SULMS method should be appropriate in the class of applications where other stochastic gradient descent methods such as LMS or CMA are suitable.

A second approach to improving the performance of the LMS filter was also tested. Instead of using $D[m]$, the limiter could be used to prevent the large outputs of the demodulation evident in Figures 8 and 9. Although the purpose of the limiter was to prevent very large demodulation outputs when the received signal had an uncharacteristically low amplitude, the limit value could be reduced to a lower level closer to the histogram peaks. Figure 13 shows that this simple approach is not as effective as the SULMS. Figure 13(a) shows the performance improvement of the reduced limit LMS post-D filter compared to no post-D filter and Figure 13(b) shows the performance improvement of the reduced limit LMS post-D filter compared to the unmodified LMS filter. The improvement shown in Figure 13(b) is less than the improvement shown by the SULMS filter in Figure 12(b), but it is still a substantial improvement over the unmodified LMS post-D filter. However, Figure 13(a) shows that use of this filter often increases the BER.

Using (28) for applications where training is not possible, the SULMS filter was tested in a decision-directed mode in which the decision $\hat{b}[m]$ was used instead of $b[m]$ to compute $e[m]$. There was no significant difference in either the convergence or the BER for the SNR = 15 dB cases in Figures 11(a) and 11(b). When the SNR was reduced to 10 dB both trained and decision-directed SULMS systems showed a BER reduction from 0.018 to 0.007 for channels with behavior similar to $n_1 = 10$. For channels with behavior similar to $n_1 = 12$, the unfiltered BER of 0.0037 was reduced to 0.0028 for the trained system or 0.0031 for the decision-directed system. Further reduction to SNR = 0 dB results in an unfiltered BER of approximately 0.3. At this noise level the trained systems show a reduction BER of 4% to 20% while the decision-directed systems show an increase in BER of 4% to 10%.

The decision-directed mode of operation has the additional advantage that for nonminimum phase channels with $|\alpha_1| > 1$ and $n_1 > \ell_0$, the decision block following the filter produces the message signal delayed by n_1 with the same BER level associated with $1/|\alpha_1|$ because the filter adapts to the stronger signal. In the training mode, when $n_1 < \ell_0$, the SULMS filter adapts to take advantage of the strong signal in the \hat{x}_n vector in (16) with the same $1/|\alpha_1|$ BER performance. However for $n_1 > \ell_0$ the direct path signal is so weak compared to the interference that the BER in a training mode is poor.

6. Conclusions

The analysis and results presented in this paper make a significant advance in understanding the behavior, benefits, limitations, and practical implementation of a post-Demodulator

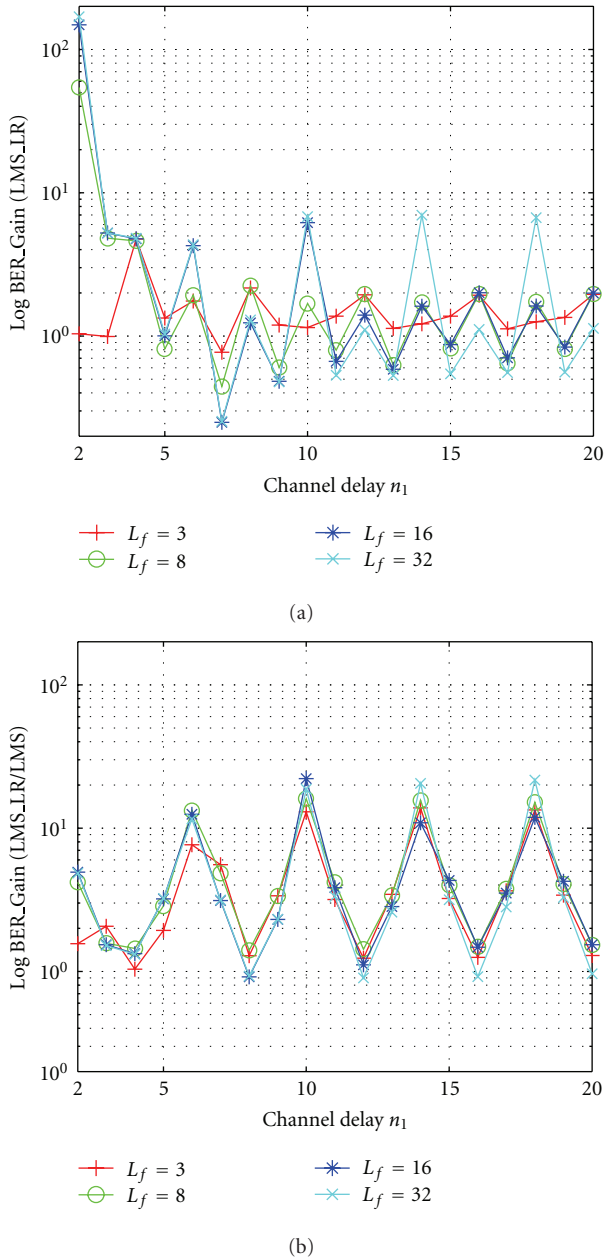


FIGURE 13: BER gain as a function of channel delay with SNR = 15 dB and modulation index $h = h_1$ for (a) reduced limit LMS filter compared to no post-D filter and (b) reduced limit LMS filter compared to LMS filter.

filter in a multipath environment. The post-D filter was proposed for an FSK communication application because it offered potential performance improvement at a lower computational cost than a traditional pre-D CMA equalizer or ML optimal decoding, and this could be of benefit in many communications applications with significant cost and power constraints. However, because of the nonlinearity of the frequency demodulator, general analysis of the impact of system and channel parameters on performance was not found in prior work.

Previously reported results used MSE or cluster variance to measure the performance benefit of the post-D adaptive filter, but our work shows that the statistical assumptions that link minimum BER to minimum MSE do not apply to the post-D filter, and that in fact simply inserting an LMS filter at the demodulator output may increase the BER even when the MSE is decreasing and the adaptation is stable. Probability density functions estimated from histograms and correlation functions were analyzed to understand the impact of the system and model parameters on the demodulator output. Analysis predicted and simulation showed that changes in reflection delay, reflection phase, and modulation index have similar impacts on the estimated probability density functions of the demodulator output. This allows a few selected cases to be representative of the range of these parameters despite the nonlinearity of the system. In addition, this approach shows why the LMS algorithm can increase BER while reducing MSE.

Based on this analysis, two modifications of the post-D filter were developed and shown to provide significant improvement in BER performance compared to the LMS algorithm. One, the SULMS filter applied a sigmoid function to the error used for adaptation and the other limited the range of values on the input to the adaptive filter. Use of the SULMS filter substantially reduced the BER compared to the unfiltered BER, and both trained and decision-directed implementations were effective. Although previous work suggested that very short filters perform as well as longer ones, these results suggest that additional performance improvement is possible if the filter is long enough to take advantage of the correlation between the demodulated output and the desired message bit. In addition, a short post-D SULMS filter could improve performance of a system with a pre-D CMA filter if the channel model included delays too long for the CMA equalizer. Preliminary post-D results for channels with multiple reflections show BER performance and general behavior consistent with the single reflection results for the same total noise and interference power. The results of this work may form the basis for understanding of how to effectively use a post-D filter in a multipath environment.

Acknowledgment

The authors thank Applied Signal Technology, Sunnyvale, California, for partial support of this work.

References

- [1] *A History of Science and Engineering in the Bell System—The Early Years (1875–1925)*, The Bell Telephone Laboratories, Inc, 1975.
- [2] J. Boykin Jr., "Spectrum economy for filtered and limited FSK signals," *IEEE Transactions on Communications*, vol. 19, no. 1, pp. 92–96, 1971.
- [3] D. L. Hedderly and L. Lundquist, "Computer simulation of a digital satellite communications link," *IEEE Transactions on Communications*, vol. 21, no. 4, pp. 321–325, 1973.

- [4] L. J. Greenstein and P. J. Fitzgerald, "Envelope fluctuation statistics of filtered PSK and other digital modulations," *IEEE Trans Commun*, vol. 27, no. 4, pp. 750–760, 1979.
- [5] I. Korn, "Error floors in the satellite and land mobile channels," *IEEE Transactions on Communications*, vol. 39, no. 6, pp. 833–837, 1991.
- [6] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 2, pp. 459–472, 1983.
- [7] J. R. Treichler and M. G. Larimore, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 420–431, 1985.
- [8] J. R. Treichler, M. G. Larimore, and J. C. Harp, "Practical blind demodulators for high-order QAM signals," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1907–1925, 1998.
- [9] J. R. Treichler, I. Fijalkow, and C. R. Johnson Jr., "Fractionally spaced equalizers: how long should they really be?" *IEEE Signal Processing Magazine*, vol. 13, no. 3, pp. 65–80, 1996.
- [10] S. L. Wood, M. G. Larimore, and J. R. Treichler, "CMA convergence for constant envelope, non-zero bandwidth signals," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, vol. 3, pp. 2525–2528, April 1997.
- [11] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 363–378, 1972.
- [12] G. D. Forney Jr., "The viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.
- [13] T. Aulin, "Symbol error probability bounds for coherently viterbi detected continuous phase modulated signals," *IEEE Transactions on Communications*, vol. 29, no. 11, pp. 1707–1715, 1981.
- [14] M. J. Ready and J. Harp, "Performance improvements achieved by equalizing intermediate rate FSK signals," in *Proceedings of the 32nd Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 379–383, November 1998.
- [15] B. Widrow and M. E. Hoff, "Adaptive switching circuits," in *IRE WESCON Convention Record*, pp. 96–104, August 1960.
- [16] W. Chung, C. R. Johnson Jr., and M. J. Ready, "Characterization of multipath distortion of FSK signals," *IEEE Signal Processing Letters*, vol. 9, no. 1, pp. 26–28, 2002.
- [17] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proceedings of the IEEE*, vol. 64, no. 8, pp. 1151–1162, 1976.
- [18] M. G. Pelchat, R. C. Davis, and M. B. Luntz, "Coherent demodulation of continuous phase binary FSK signals," in *Proceedings of International Telemetry Conference*, pp. 181–190, Washington, DC, USA, November 1971.
- [19] W. P. Osborne and M. B. Luntz, "Coherent and noncoherent detection of CPFSK," *IEEE Transactions on Communications*, vol. 22, no. 8, pp. 1023–1036, 1974.
- [20] T. A. Schonhoff, "Symbol error probabilities for M-ary CPFSK: coherent and noncoherent detection," *IEEE Transactions on Communications*, vol. 24, no. 6, pp. 644–652, 1976.
- [21] T. Aulin and C. E. W. Sundberg, "Continuous phase modulation—part I: full response signaling," *IEEE Transactions on Communications Systems*, vol. 29, no. 3, pp. 196–209, 1981.
- [22] A. Napolitano and C. M. Spooner, "Cyclic spectral analysis of continuous-phase modulated signals," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 30–44, 2001.
- [23] G. B. Giannakis and E. Serpedin, "Blind identification of ARMA channels with periodically modulated inputs," *IEEE Transactions on Signal Processing*, vol. 46, no. 11, pp. 3099–3104, 1998.
- [24] G. B. Giannakis and C. Tepedelenlioglu, "Direct blind equalizers of multiple FIR channels: a deterministic approach," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 62–74, 1999.
- [25] R. W. Heath and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Transactions on Signal Processing*, vol. 47, no. 3, pp. 848–856, 1999.
- [26] T. Aulin, N. Rydbeck, and C. E. W. Sundberg, "Continuous phase modulation—part II: partial response signaling," *IEEE Transactions on Communications Systems*, vol. 29, no. 3, pp. 210–225, 1981.
- [27] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 4th edition, 2001.
- [28] S. Pasupathy, "Minimum shift keying: a spectrally efficient modulation," *IEEE Communications Magazine*, vol. 17, no. 4, pp. 14–22, 1979.
- [29] D. H. Morais and K. Feher, "Bandwidth efficiency and probability of error performance of MSK and offset QPSK systems," *IEEE Transactions on Communications Systems*, vol. 27, no. 12, pp. 1794–1801, 1979.
- [30] P. Bello and B. Nelin, "The effect of frequency selective fading on the binary error probabilities of incoherent and differentially coherent matched filter receivers," *IEEE Transactions on Communications Systems*, vol. 11, no. 2, pp. 170–186, 1963.
- [31] D. C. Cox, "Delay Doppler characteristics of multipath propagation at 910 MHz in a suburban mobile radio environment," *IEEE Transactions on Antennas and Propagation*, vol. 20, no. 5, pp. 625–635, 1972.
- [32] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection theory*, Prentice-Hall PTR, Upper Saddle River, NJ, USA, 1998.