MMSE Decision Feedback Equalization of Orthogonal Multipulse Modulated Signals

A.G. Klein†, D.R. Brown III§, and C.R. Johnson, Jr.†

†School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853
email: agk5@cornell.edu

§Department of Electrical and Computer Engineering
Worcester Polytechnic Institute, Worcester, MA 01609
email: drb@ece.wpi.edu

***Abstract*— We propose a finite impulse response (FIR) minimum mean squared error (MMSE) decision feedback equalizer (DFE) for orthogonal multipulse modulated signals received through a multipath channel. Our system model accounts for both the loss of orthogonality caused by the multipath channel as well as intersymbol interference (ISI). First, we review previous work on the subject which used the zero-forcing criterion under strict assumptions about the channel and equalizer lengths. Then, we derive a computationally efficient MMSE equalizer which removes these restrictions, and is suitable for use with stochastic gradient descent algorithms. Finally, we demonstrate the performance of the proposed equalizer with simulations.**

I. Introduction

Orthogonal multipulse modulation is a modulation scheme that has been studied for many applications, and its many variants include frequency shift keying (FSK) and pulse position modulation (PPM). The problem of equalizing M-ary orthogonal signals has been known to be difficult [1]. In fact, traditional uses of orthogonal modulation have been in situations with little or no ISI, and thus there has been little motivation to explore equalization of such signals to the extent that equalization has been explored for linearly modulated signals. Orthogonal modulation is a power efficient scheme, but is bandwidth inefficient, and thus has attracted attention for use in ultra wideband (UWB) communication systems where ISI *will* be an issue.

The equalization of orthogonal multipulse signals has been proposed in [2] and in [3] for the specific case of PPM. In [2] a zero-forcing (ZF) decision feedback equalizer (DFE) is proposed that employs an infinite impulse response (IIR) feedforward filter. In [3] the ZF decision feedback equalizer is derived under the following assumptions: the channel is monic and minimum phase, the additive noise is ignored (i.e. since it is a ZF equalizer), and the feedback portion of the equalizer is as long as the channel (and possibly infinite).

In this paper, we propose a minimum mean-squared error (MMSE) DFE for orthogonally modulated signals, and we show the design equations and performance of the proposed structure. Furthermore, we remove several assumptions inherent in the previous work [2][3] – that is, we permit non-monic and non-minimum phase channels, we accommodate noise from any stationary random process, and we permit the lengths of the FIR feedforward and feedback portions of the equalizer to be design parameters. We then make several modifications to the equalization structure, thereby permitting a computational savings and the use of stochastic gradient decent techniques for determining the MMSE equalizer tap values. Finally, we include simulation results which demonstrate the performance of the proposed equalizer.

In this paper, we use $^\top$ to denote matrix transpose, $\otimes$ to denote matrix direct product, $\boldsymbol{I}_m$ to denote the $m \times m$ identity matrix, $\mathbf{1}_{m\times n}$ to denote the $m \times n$ matrix of all ones, and $\mathbf{0}_{m\times n}$ to denote the $m \times n$ matrix of all zeros.

II. System Model

M-ary orthogonal multipulse modulation is accomplished by transmitting one of M orthogonal waveforms, whose time-reversed versions are denoted $\boldsymbol{s}_i$. Each waveform is assumed to consist of P chips, where $P \geq M$. Thus, orthogonal multipulse modulation can be thought of as a block coding scheme where information is conveyed by transmitting one of M codes. Let $\boldsymbol{S} \in \mathbb{R}^{P\times M}$ be the matrix whose columns are the M waveforms $\boldsymbol{s}_i$. We assume the waveforms are mutually orthogonal and have unit energy, so that $\boldsymbol{S}^\top \boldsymbol{S} = \boldsymbol{I}_M$. In this paper, we consider a discrete-time model where each chip is sampled once.

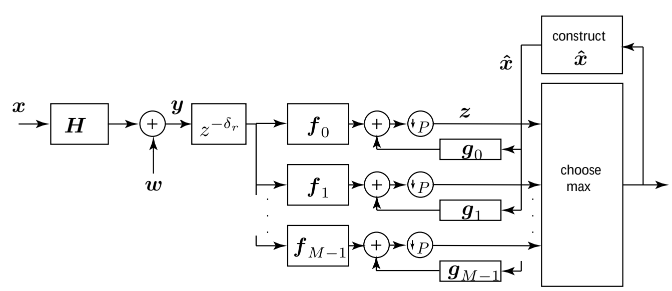

Fig. 1. MIMO DFE block diagram.

The system model is shown in Fig. 1. We assume i.i.d. M-ary orthogonal symbols are transmitted through a linear time-invariant FIR channel of length N_h with impulse response $\boldsymbol{h} = [h[0] \ \dots \ h[N_h-1]]^\top$. We assume additive noise $w[n]$ arises from a stationary random process with autocorrelation $\boldsymbol{R}_{ww} \triangleq E[\boldsymbol{w}[n]\boldsymbol{w}[n]^\top]$, assumed to be uncorrelated with the data.

A vector model for the length N_f received vector at time n is then

$$\boldsymbol{y}[n] = \boldsymbol{H}\boldsymbol{x}[n] + \boldsymbol{w}[n] \tag{1}$$

where $\boldsymbol{y}[n] \in \mathbb{R}^{N_f}$ is the received vector, $\boldsymbol{H} \in \mathbb{R}^{N_f \times (N_f+N_h-1)}$ is the channel convolution matrix, $\boldsymbol{x}[n] \in \mathbb{R}^{N_f+N_h-1}$ is the vector of transmitted chips whose symbols belong to columns of $\boldsymbol{S}$, and $\boldsymbol{w}[n] \in \mathbb{R}^{N_f}$ is the noise vector. More precisely, we have

$$\begin{aligned}
\boldsymbol{x}[n] &= \begin{bmatrix} x[n] & x[n-1] & \cdots & x[n-N_f-N_h+2] \end{bmatrix}^\top \\
\boldsymbol{H} &= \begin{bmatrix} h[0] & \cdots & h[N_f-1] & 0 & \cdots & & 0 \\ 0 & h[0] & \cdots & h[N_f-1] & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ & & h[0] & \cdots & & h[N_f-1] & 0 \\ 0 & \cdots & 0 & h[0] & \cdots & & h[N_f-1] \end{bmatrix}
\end{aligned}$$

Note that a complete symbol is generated at times that are multiples of P, so $\{x[-P+1], \ldots, x[-1], x[0]\}$ would be a complete symbol. In general, $x[n]$ is not a stationary process at the chip level, but is cyclostationary. Thus, the correlation statistics of chips at even sampling times $2n$ are different from those at odd sampling times $2n+1$ when $P = 2$, for example.

At the receiver we employ a bank of M decision-feedback equalizers. Each equalizer is operating only once every P chips; hence, the number of computations for the bank of M equalizers operating at the symbol rate is equal to M/P times the number of computations required for a single equalizer operating at the chip rate. The feedforward equalizers have impulse response $\boldsymbol{f}_i = [f_i[0] \ \ldots \ f_i[N_f-1]]^\top$ and length $N_f \geq P$. The feedback equalizers $\boldsymbol{g}_i = [g_i[0] \ \ldots \ g_i[N_g-1]]^\top$ each have length N_g (assumed to be a multiple of M), and operate on the signal $\hat{\boldsymbol{x}}[n] \in \mathbb{R}^{N_g}$. Note that the receiver front-end has a bulk delay δ_r to align the feedforward equalizer outputs to the desired phase before downsampling. If we let $\boldsymbol{F} = [\boldsymbol{f}_0 \cdots \boldsymbol{f}_{M-1}] \in \mathbb{R}^{N_f \times M}$ and $\boldsymbol{G} = [\boldsymbol{g}_0 \cdots \boldsymbol{g}_{M-1}] \in \mathbb{R}^{N_g \times M}$, then the output of the M equalizers before downsampling becomes $\boldsymbol{F}^\top \boldsymbol{y}[n-\delta_r] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[n]$. However, we are only interested in the equalizer outputs at times that are multiples of P, and after downsampling we keep the δ_r-th polyphase component. Thus, the output of the M equalizers is

$$\boldsymbol{z}[n] = \boldsymbol{F}^\top \boldsymbol{y}[Pn-\delta_r] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn]. \tag{2}$$

Due to the downsampling operation, as well as the operation of the decision device (discussed below), $\boldsymbol{z}[n] \in \mathbb{R}^M$, $\boldsymbol{x}[Pn-\delta_r]$, and $\hat{\boldsymbol{x}}[Mn]$ do not live on the same time-scale. That is, $\boldsymbol{z}[n]$ is a symbol-rate signal, $\boldsymbol{y}[Pn-\delta_r]$ is a chip-rate signal, and $\hat{\boldsymbol{x}}[Mn]$ operates at M/P times the chip rate.

The optimal (i.e. maximum likelihood) decision device for the AWGN channel is the minimum distance detector, which amounts to choosing the maximum output of M correlators $\max\{\boldsymbol{S}^\top \boldsymbol{y}\}$ as shown in Fig. 2. While such a memoryless decision device is by no means optimal in the presence of ISI, this is the decision device that is assumed in this paper due to its simple implementation and low latency. The input to the decision device is a block of M values at the symbol rate, and the output is a serial stream of M samples where

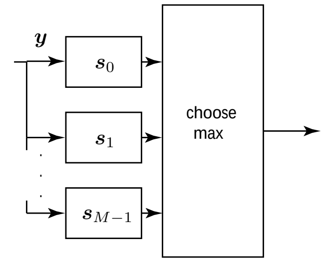

Fig. 2. ML detector for M-ary orthogonal signaling in AWGN channel.

only one chip per symbol is set to 1 and all others are 0. The output operates at a rate M/P times the chip rate, and a 1 in the ith position represents a decision that the ith column of $\boldsymbol{S}$ was the transmitted signal (i.e. the output is essentially a stream of M-ary PPM symbols). The detector output is then accumulated in the vector $\hat{\boldsymbol{x}}[Mn]$, which is input to the feedback equalizer. While the proposed receiver structure in Fig. 1 does not explicitly show the correlation with each of the signal waveforms $\boldsymbol{s}_i$ that appear in Fig. 2, these correlations will be accomplished by the equalizers.

A single DFE typically has a delay of 1 in the feedback path. However, since the decision device for orthogonal multipulse modulation requires a block of data before making a decision, the feedback signal estimates need to be delayed by one whole symbol, or M detector output samples.

III. Previous Work: Zero-Forcing Equalizers

In [2], ZF equalization of multipulse modulated signals was proposed. A feedforward IIR filter is employed, and as such this scheme does not fit within the framework of our proposed structure. However, it is a ZF equalizer, and thus will suffer from noise enhancement.

One of the schemes proposed in [3] for equalization of PPM, referred to as the block DFE, is a special case of our proposed structure with the following parameter choices: $\delta_r = 0$, $N_f = M$, $\boldsymbol{F} = ([\boldsymbol{I}_M \ \boldsymbol{0}_{M\times(N_h-1)}]\boldsymbol{H}^\top)^{-1}$, $N_g = N_h - 1$, and $\boldsymbol{G} = -[\boldsymbol{0}_{(N_h-1)\times M} \ \boldsymbol{I}_{N_h-1}]\boldsymbol{H}^\top \boldsymbol{F}$. While [3] concerns the equalization of PPM, this will be the (non-unique) zero-forcing decision feedback equalizer for any M-ary orthogonal signaling so long as $P = M$ (though the equalizer output would need to be multiplied by $\boldsymbol{S}^\top$). As mentioned above, several additional structures are considered in [3] which involve feedback of tentative chip decisions, but these structures are not considered in our paper.

Making the standard assumption of feedback of correct decisions, we have $\hat{x}[n] = x[n-M]$ and then $\hat{\boldsymbol{x}}[Mn] = [\boldsymbol{0}_{(N_h-1)\times M} \boldsymbol{I}_{N_h-1}]\boldsymbol{x}[Mn]$. Then, partition the channel matrix as $\boldsymbol{H} = [\boldsymbol{H}_0 \ \boldsymbol{H}_1]$ where

$$\begin{aligned}
\boldsymbol{H}_0 &= \boldsymbol{H}[\boldsymbol{I}_M \ \boldsymbol{0}_{M\times(N_h-1)}]^\top \\
\boldsymbol{H}_1 &= \boldsymbol{H}[\boldsymbol{0}_{(N_h-1)\times M} \ \boldsymbol{I}_{N_h-1}]^\top.
\end{aligned}$$

Note that $\boldsymbol{F} = (\boldsymbol{H}_0^\top)^{-1}$ and $\boldsymbol{G} = -\boldsymbol{H}_1^\top \boldsymbol{F} = -(\boldsymbol{H}_0^{-1}\boldsymbol{H}_1)^\top$. From (2) and (1), the output of the zero-forcing equalizer is then

$$\begin{aligned} \boldsymbol{z}[n] &= \boldsymbol{F}^\top \boldsymbol{y}[Mn] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn] \\ &= \underbrace{[\boldsymbol{I}_M \ \mathbf{0}_{M\times(N_h-1)}]\boldsymbol{x}[Mn]}_{\text{ISI-free symbol}} + \underbrace{\boldsymbol{H}_0^{-1}\boldsymbol{w}[Mn]}_{\text{noise term}} \end{aligned} \quad (3)$$

and so the output of the zero-forcing equalizer proposed in [3] is the ISI-free symbols plus filtered noise.

IV. The Proposed MMSE Equalizer

The previously proposed zero-forcing schemes have several shortcomings, as mentioned in the introduction: the channel needs to be monic and minimum phase so that $\boldsymbol{H}_0^{-1}$ exists, the additive noise is ignored since it is a zero-forcing equalizer, and the feedback portion of the equalizer needs to be as long as the channel. Here, we propose a constrained-length MMSE block DFE which can handle a larger class of channels, does not suffer from noise enhancement, and is suitable for use with adaptive gradient descent algorithms (e.g. least mean squares).

The schemes proposed in [2][3] implicitly assume the delay through the channel and equalizer is zero, which may be reasonable in monic and minimum phase channels. For the MMSE equalizer, we wish to allow arbitrary target delays through the combined response of the channel and equalizer, thereby accommodating a larger class of channels. Consequently, we need to introduce several delay parameters. Let δ be the desired delay through the channel and equalizer chain. However, δ may not be a multiple of P; thus, we will adjust for this with the bulk delay at the front of the receiver. Let $\delta_t = \lceil \delta/P \rceil P$ and $\delta_r = \delta_t - \delta$. So, δ_r is the residual delay to align the chips to the correct sampling phase, and δ_t is the new total target delay through the entire channel and equalizer chain. Since δ_r and δ_t are functions of δ, the only design parameter is δ.

Let $\boldsymbol{E}_\delta \in \mathbb{R}^{(N_f+N_h-1)\times M}$ be the matrix of the desired combined responses through the channel and each of the M equalizers. We will leave $\boldsymbol{E}_\delta$ unspecified for now, but certainly we will want the desired equalizer response to resemble $\boldsymbol{S}^\top$ so that the equalizer performs the correlation shown in the decision device of Fig. 2. The output error is then

$$\boldsymbol{\epsilon} = \boldsymbol{F}^\top \boldsymbol{y}[Pn-\delta_r] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn] - \boldsymbol{E}_\delta^\top \boldsymbol{x}[Pn-\delta_r]. \quad (4)$$

As is common in MIMO problems, we can use the trace or determinant of the autocorrelation of the output error as a measure of the mean-squared error. As was pointed out in [4], the same set of equalizer coefficients minimizes both measures. Here, we choose the trace:

$$\begin{aligned} J(\boldsymbol{F},\boldsymbol{G}) &= \operatorname{tr}(E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top]) \\ &= \operatorname{tr}(\boldsymbol{F}^\top \boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^\top \boldsymbol{F} + 2\boldsymbol{F}^\top \boldsymbol{H}\boldsymbol{R}_{x\hat{x}}\boldsymbol{G} \\ &\quad -2\boldsymbol{F}^\top \boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{E}_\delta + \boldsymbol{F}^\top \boldsymbol{R}_{ww}\boldsymbol{F} + \boldsymbol{G}^\top \boldsymbol{R}_{\hat{x}\hat{x}}\boldsymbol{G} \\ &\quad -2\boldsymbol{G}^\top \boldsymbol{R}_{x\hat{x}}^\top \boldsymbol{E}_\delta + \boldsymbol{E}_\delta^\top \boldsymbol{R}_{xx}\boldsymbol{E}_\delta) \end{aligned} \quad (5)$$

where we have used the fact that the data and noise are uncorrelated, and we define

$$\boldsymbol{R}_{xx} \triangleq E[\boldsymbol{x}[Pn-\delta_r]\boldsymbol{x}[Pn-\delta_r]^\top] \quad (6)$$

$$\boldsymbol{R}_{x\hat{x}} \triangleq E[\boldsymbol{x}[Pn-\delta_r]\hat{\boldsymbol{x}}[Mn]^\top] \quad (7)$$

$$\boldsymbol{R}_{\hat{x}\hat{x}} \triangleq E[\hat{\boldsymbol{x}}[Mn]\hat{\boldsymbol{x}}[Mn]^\top]. \quad (8)$$

To find the MMSE equalizer settings, we begin by setting the derivative of (5) with respect to $\boldsymbol{G}$ to zero

$$\begin{aligned} \frac{1}{2}\frac{\partial}{\partial \boldsymbol{G}} J(\boldsymbol{F},\boldsymbol{G}) &= \boldsymbol{R}_{x\hat{x}}^\top \boldsymbol{H}^\top \boldsymbol{F} + \boldsymbol{R}_{\hat{x}\hat{x}}\boldsymbol{G} - \boldsymbol{R}_{x\hat{x}}^\top \boldsymbol{E}_\delta \\ &\triangleq \mathbf{0}_{M\times 1} \\ \Longrightarrow \boldsymbol{G}^* &= \underbrace{\boldsymbol{R}_{\hat{x}\hat{x}}^{-1}\boldsymbol{R}_{x\hat{x}}^\top}_{\triangleq \boldsymbol{B}}(\boldsymbol{E}_\delta - \boldsymbol{H}^\top \boldsymbol{F}). \end{aligned} \quad (9)$$

And setting the derivative with respect to $\boldsymbol{F}$ to zero gives

$$\begin{aligned} \frac{1}{2}\frac{\partial}{\partial \boldsymbol{F}} J(\boldsymbol{F},\boldsymbol{G}^*) &= \boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^\top \boldsymbol{F} + \boldsymbol{H}\boldsymbol{R}_{x\hat{x}}\boldsymbol{G}^* \\ &\quad -\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{E}_\delta + \boldsymbol{R}_{ww}\boldsymbol{F} \\ &= \boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^\top \boldsymbol{F} \\ &\quad +\boldsymbol{H}\boldsymbol{R}_{x\hat{x}}(\boldsymbol{R}_{\hat{x}\hat{x}}^{-1}\boldsymbol{R}_{x\hat{x}}^\top(\boldsymbol{E}_\delta - \boldsymbol{H}^\top \boldsymbol{F})) \\ &\quad -\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{E}_\delta + \boldsymbol{R}_{ww}\boldsymbol{F} \\ &\triangleq \mathbf{0}_{M\times 1} \end{aligned}$$

After solving this equation for $\boldsymbol{F}$, we have an equation for the MMSE feedforward equalizer coefficients, shown on the next page in (10).

A. The choice of $\boldsymbol{E}_\delta$

We have left the desired response matrix $\boldsymbol{E}_\delta$ undefined up to now. A reasonable choice might be to force each of the M responses to the M signal waveforms (with appropriate delay), so $\boldsymbol{E}_\delta = [\mathbf{0}_{M\times\delta}\ \boldsymbol{S}^\top\ \mathbf{0}_{M\times(N_f+N_h-1-P-\delta)}]^\top$. However, there are other choices for the desired response that do not effect the operation of the decision device. The "choose max" decision device is invariant to an added constant, so long as the same constant k is added to each of the M inputs.

Now we will motivate a specific choice of the matrix $\boldsymbol{E}_\delta$, and make a modification to our equalizer structure. First, consider a binary case where $M = 2$. The "choose max" decision device will then perform the operation $\max\{z_0[n], z_1[n]\}$. However, this decision rule is equivalent to deciding $z_0[n] - z_1[n] \gtrless 0$. This implies that for the binary case, we can make the decision on a single statistic. Note that $z_0[n] - z_1[n] = (\boldsymbol{f}_0 - \boldsymbol{f}_1)^\top \boldsymbol{y}[Pn-\delta_r] + (\boldsymbol{g}_0 - \boldsymbol{g}_1)^\top \hat{\boldsymbol{x}}[Mn]$ which means that we can reduce the bank of two equalizers to a single equalizer, and then make our decision. We will now show that for the M-ary case, so long as the decision rule is "choose max", we can reduce the bank of M equalizers to a bank of $M-1$ equalizers.

First, choose

$$\boldsymbol{E}_\delta = [\mathbf{0}_{M\times\delta}\ (\boldsymbol{I}_M - 1/M\mathbf{1}_{M\times M})\boldsymbol{S}^\top\ \mathbf{0}_{M\times(N_f+N_h-1-P-\delta)}]^\top$$

which corresponds to a response where the same constant $k = 1/M$ has been subtracted from each of the M decision device inputs, and thus will not change the operation of the decision

$$F^* = \underbrace{\left[H(R_{xx} - R_{x\hat{x}}R_{\hat{x}\hat{x}}^{-1}R_{x\hat{x}}^\top)H^\top + R_{ww}\right]^{-1} H(R_{xx} - R_{x\hat{x}}R_{\hat{x}\hat{x}}^{-1}R_{x\hat{x}}^\top)}_{\triangleq A} E_\delta. \tag{10}$$

device. Note that the matrix $I_M - 1/M \mathbf{1}_{M\times M}$ has a zero eigenvalue, which is a fact that we will exploit. Let $U^\top U = I_M - 1/M\mathbf{1}_{M\times M}$ be the Cholesky factorization. Since the matrix is only *semi*-positive definite, the last row of the upper triangular matrix U will be all zero. Next, define

$$E'_\delta = [\mathbf{0}_{M\times\delta} \;\; US^\top \;\; \mathbf{0}_{M\times(N_f+N_h-1-P-\delta)}]^\top$$

so that $E_\delta = E'_\delta U$. From (2), (9), and (10) the MMSE equalizer output is given by

$$z[n] = E_\delta^\top A^\top y[Pn - \delta_r] + (E_\delta - H^\top F)^\top B^\top \hat{x}[Mn]. \tag{11}$$

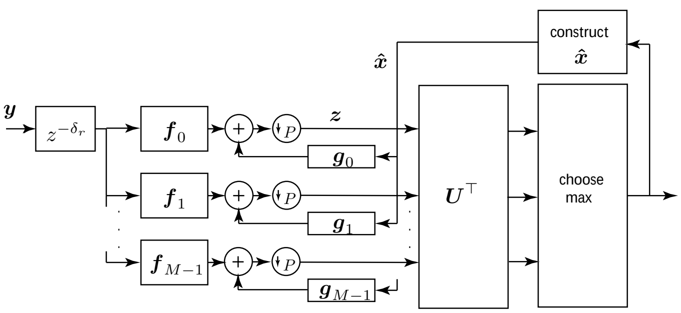

Fig. 3. Modified MIMO DFE block diagram.

Now, we will modify the equalizer structure by multipling the M outputs of the equalizer by the $M \times M$ matrix $U^\top$, as shown in Figure 3. Furthermore, instead of designing the equalizers to have response E_δ, we will design them to have response E'_δ. The output of the modified structure will be

$$\begin{aligned} U^\top z'[n] &= U^\top \Bigg[\underbrace{(E'_\delta)^\top A^\top}_{\triangleq (F')^\top} y[Pn-\delta_r] \\ &\quad + \underbrace{(E'_\delta - H^\top F')^\top B^\top}_{\triangleq (G')^\top} \hat{x}[Mn] \Bigg] \\ &= E_\delta^\top A^\top y[Pn-\delta_r] \\ &\quad + (E_\delta - H^\top F)^\top B^\top \hat{x}[Mn] \end{aligned} \tag{12}$$

where we have used $E_\delta = E'_\delta U$. Thus, the output of the modified structure in (12) is equivalent to the output of the original structure in (11). Note that the last row of U is all zero, so the last column of E'_δ is all zero, and therefore the last column of both F' and G' is all zero. We have effectively reduced the bank of M equalizers to a bank of $M-1$ equalizers, at the expense of multiplication by an $M \times M$ lower triangular matrix $U^\top$. We are trading $N_f + N_g$ multiply operations for $2(M-1)$ multiplies since each column of $U^\top$ has only 2 unique elements. In an adaptive setting, the computational savings is even greater since we need to update $N_f + N_g$ fewer parameters.

Let us return to the $M = 2$ example to show that our previous intuition applies. For $M = 2$, we have

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \implies U^\top U = I_2 - \frac{1}{2}\mathbf{1}_{2\times 2}$$

and we see that we are effectively designing a single equalizer to be what would have been the difference of 2 equalizers, as expected.

We can think of this from a geometric perspective. Closer inspection of the entries of the desired response U reveals that the equalizer is effectively mapping the M orthogonal waveforms to M equidistant points on an $M-1$ dimensional hypersphere (i.e. to the vertices of the regular $M-1$ dimensional simplex). This is the so-called simplex signal set [5], whose signals have dimensionality $M-1$ but maintain the same Euclidean distance as that between M-ary orthogonal signals. By translating and rotating the original signal set, we incur a reduction in the signal dimensionality (and also the number of equalizers), but we maintain the same performance.

B. The Feedback Signal

One point that we have overlooked thus far is the definition of $\hat{x}[Mn]$. We assume, as is common in DFE literature, that the feedback signal consists of the correct symbol decisions represented by unit vectors, so that $\hat{x}[Mn]$ is a windowed and cross correlated version of the original signal $x[Pn - \delta_r]$, as

$$\hat{x}[Mn] = \underbrace{\begin{bmatrix} \mathbf{0}_{\delta+P\times N_g} \\ I_{N_g/M} \otimes S \\ \mathbf{0}_{(N_f+N_h-1-\delta-P-N_gP/M)\times N_g} \end{bmatrix}^\top}_{\triangleq W} x[Mn - \delta_r] \tag{13}$$

where for convenience we assume N_g is a multiple of M, and $W \in \mathbb{R}^{N_g\times(N_f+N_h-1)}$. Due to the way have constructed W, we observe that $N_f + N_h - 1 - \delta - P - N_gP/M$ must be non-negative in order for W to make sense. We can always artificially append zeros to the channel impulse response, however, thereby increasing N_h to meet this condition.

Note from (10) that $R_{\hat{x}\hat{x}}$ needs to be inverted to calculate the equalizer taps. While the data autocorrelation matrices will be discussed in more detail in section IV-C, this matrix will not be invertible for $N_g \geq 2M$. For example when $N_g = 2M$, we will have

$$R_{\hat{x}\hat{x}} = 1/M^2 \begin{bmatrix} MI_M & \mathbf{1}_{M\times M} \\ \mathbf{1}_{M\times M} & MI_M \end{bmatrix} \tag{14}$$

which is not invertible.

The non-invertibility of this matrix implies that an additional constraint is necessary. Here, we motivate a modification to the feedback signal that effectively causes the autocorrelation matrix to be full rank while providing a computational savings. Let us consider a specific example where $M = 2$ and $N_g = 4$, so our decision feedback equalizer spans $N_g/M = 2$ symbols. Consider 2 possible responses for the feedback filter, $\boldsymbol{g}$ and $\boldsymbol{g}'$ where

$$\begin{aligned} \boldsymbol{g} &= \begin{bmatrix} g[0] & g[1] & g[2] & g[3] \end{bmatrix}^\top \\ \boldsymbol{g}' &= \boldsymbol{g} + \begin{bmatrix} +a & +a & -a & -a \end{bmatrix}^\top . \end{aligned}$$

Note that both of these feedback equalizers give the same output, regardless of the value of a since $\hat{\boldsymbol{x}} \in \{[1\,0\,1\,0], [1\,0\,0\,1], [0\,1\,1\,0], [0\,1\,0\,1]\}$, and so $\boldsymbol{g}^\top \hat{\boldsymbol{x}} = (\boldsymbol{g}')^\top \hat{\boldsymbol{x}} \in \{g[0] + g[2],\ g[0] + g[3],\ g[1] + g[2],\ g[1] + g[3]\}$. It is no coincidence that the vector $[+a\,+a\,-a\,-a]^\top$ is a basis vector for the nullspace of $\boldsymbol{R}_{\hat{x}\hat{x}}$ in (14). Thus, by choosing the free parameter a to be $a = g[3]$, for example, we can obtain the same output from a feedback equalizer with 3 non-zero taps since the last tap of $\boldsymbol{g}'$ is equal to zero.

In the M-ary case with arbitrary $N_g \geq M$, any set of feedback equalizer taps $\boldsymbol{g}$ has a corresponding set of taps $\boldsymbol{g}'$ that gives an identical output, but has at least $N_g/M - 1$ of its taps equal to zero. In particular, we can find a $\boldsymbol{g}'$ where $g'[kM-1] = 0$ for all integers k such that $2 \leq k \leq N_g/M$. We omit the proof due to space constraints.

This equalizer with selected taps constrained to zero can be represented more conveniently in matrix notation as an equalizer of length $N_g - N_g/M + 1$ that operates on a decimated version of $\hat{\boldsymbol{x}}$. The decimated entries correspond to the zeroed equalizer taps, and the decimated signal can be represented as

$$\underbrace{\begin{bmatrix} \boldsymbol{I}_M & \boldsymbol{0}_{M\times N_g - M} \\ \boldsymbol{0}_{N_g - N_g/M+1-M\times M} & \boldsymbol{I}_{N_g/M-1} \otimes [\boldsymbol{I}_{M-1}\ \boldsymbol{0}_{M-1\times 1}] \end{bmatrix}}_{\triangleq \boldsymbol{V}} \hat{\boldsymbol{x}}[Mn]. \quad (15)$$

Note that $\boldsymbol{V} \in \mathbb{R}^{(N_g - N_g/M+1)\times N_g}$ and our new shortened equalizer $\boldsymbol{G}'' \in \mathbb{R}^{(N_g - N_g/M+1)\times M}$ has output $(\boldsymbol{G}'')^\top \boldsymbol{V}\hat{\boldsymbol{x}}[Mn]$ which we claim can made to have identical output to $\boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn]$ for arbitrary $\boldsymbol{G}$. It turns out that $\boldsymbol{R}_{\hat{x}\hat{x}} \in \mathbb{R}^{N_g \times N_g}$ will have rank $N_g - N_g/M + 1$, and by throwing out $N_g/M - 1$ carefully selected samples (i.e. by forming $\boldsymbol{V}\hat{\boldsymbol{x}}[Mn]$), the autocorrelation of the decimated signal becomes $\boldsymbol{V}\boldsymbol{R}_{\hat{x}\hat{x}}\boldsymbol{V}^\top$ which will be full rank.

C. Summary of MMSE Equalizer

It is now useful to summarize the proposed MMSE equalizer and to include the modifications made in sections IV-A and IV-B. The structure shown in Figure 3 still applies. Now, we will drop the primes (e.g. $\boldsymbol{F}' \rightarrow \boldsymbol{F}$) that were adopted in previous sections. The design parameters are δ, N_f, and N_g; the channel and noise characteristics are assumed to be known.

First, find the Cholesky factorization $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_M - 1/M\boldsymbol{1}_{M\times M}$, which is a fixed matrix that depends only on M. Next, form the desired response as $\boldsymbol{E}_\delta = [\boldsymbol{0}_{M\times\delta}\,\boldsymbol{U}\boldsymbol{S}^\top\,\boldsymbol{0}_{M\times(N_f+N_h-1-P-\delta)}]^\top$. Then, we can determine $\boldsymbol{F} \in \mathbb{R}^{N_g\times M}$ and $\boldsymbol{G} \in \mathbb{R}^{(N_g - N_g/M+1)\times M}$ using (9) and (10). Note that the last column of $\boldsymbol{F}$ and $\boldsymbol{G}$ will both be zero, so effectively we have $M-1$ equalizers. Also note that, under the assumption of correct decisions, the feedback equalizers are operating on a decimated version of the decisions given by

$$\hat{\boldsymbol{x}}[Mn] = \boldsymbol{V}\boldsymbol{W}\boldsymbol{x}[Pn - \delta_r] \quad (16)$$

where $\hat{\boldsymbol{x}}[Mn] \in \mathbb{R}^{N_g - N_g/M+1}$, $\boldsymbol{W}$ windows and cross correlates the appropriate samples of $\boldsymbol{x}$ with $\boldsymbol{S}$ as defined in (13), and $\boldsymbol{V}$ performs the decimation as defined in (15).

The equalizer output is then given by

$$\boldsymbol{z}[n] = \boldsymbol{F}^\top \boldsymbol{y}[Pn - \delta_r] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn].$$

The equalizer output is then multiplied by $\boldsymbol{U}^\top$, and the result is passed to the "choose max" decision device.

The data autocorrelation matrices needed in the equalizer design equations (9) and (10) remain to be defined. Consider a length NM vector of data $\boldsymbol{x}'[Mn]$ which contains a stream of PPM chips from the signal set consisting of the columns of the identity matrix. Define $\mathcal{S} \triangleq (\boldsymbol{I}_N \otimes \boldsymbol{S})$. Then, any data stream $\boldsymbol{x}[Pn]$ consisting of chips from an arbitrary orthogonal signal set $\boldsymbol{S}$ can be represented as $\mathcal{S}\boldsymbol{x}'[Mn]$. Note that $\mathcal{S}^\top \mathcal{S} = \boldsymbol{I}_{NM}$.

The autocorrelation of a length NM stream of i.i.d. PPM symbols aligned to the symbol boundary is given by the block Tœplitz matrix

$$\begin{aligned} \boldsymbol{R} &= E[(\boldsymbol{x}'[Mn])(\boldsymbol{x}'[Mn])^\top] \\ &= \frac{1}{M^2}(\boldsymbol{1}_{NM} + (\boldsymbol{I}_N \otimes (M\boldsymbol{I}_M - \boldsymbol{1}_M))) \\ &= \frac{1}{M^2}\begin{bmatrix} M\boldsymbol{I}_M & \boldsymbol{1}_M & \boldsymbol{1}_M & \cdots \\ \boldsymbol{1}_M & M\boldsymbol{I}_M & \boldsymbol{1}_M & \\ \boldsymbol{1}_M & \boldsymbol{1}_M & \ddots & \\ \vdots & & & \end{bmatrix}. \end{aligned} \quad (17)$$

Thus, the autocorrelation of a stream of chips from an arbitrary signal set is given by $\mathcal{S}\boldsymbol{R}\mathcal{S}^\top$. We use the notation $(\mathcal{S}\boldsymbol{R}\mathcal{S}^\top)_{i:j,\,k:l}$ to denote the extraction of rows i through j and columns k through l of this matrix, and we assume N is as large as necessary to permit this extraction.

From (6)-(8), (16), and (17) we have

$$\begin{aligned} \boldsymbol{R}_{xx} &\triangleq E[\boldsymbol{x}[Pn-\delta_r]\boldsymbol{x}[Pn-\delta_r]^\top] \\ &= (\mathcal{S}\boldsymbol{R}\mathcal{S}^\top)_{\delta_r:\delta_r+N_f+N_h-2,\delta_r:\delta_r+N_f+N_h-2} \\ \boldsymbol{R}_{x\hat{x}} &\triangleq E[\boldsymbol{x}[Pn-\delta_r]\hat{\boldsymbol{x}}[Mn]^\top] \\ &= \boldsymbol{R}_{xx}\boldsymbol{W}^\top\boldsymbol{V}^\top \\ \boldsymbol{R}_{\hat{x}\hat{x}} &\triangleq E[\hat{\boldsymbol{x}}[Mn]\hat{\boldsymbol{x}}[Mn]^\top] \\ &= \boldsymbol{V}\boldsymbol{W}\boldsymbol{R}_{xx}\boldsymbol{W}^\top\boldsymbol{V}^\top. \end{aligned}$$

Observe that for the special case of an AWGN channel with no ISI, we have $\boldsymbol{H} = \boldsymbol{I}$ and $\boldsymbol{R}_{ww} = \sigma_w^2\boldsymbol{I}$. In this case, the detector reduces to deciding $\max\{\boldsymbol{U}^\top\boldsymbol{F}^\top\boldsymbol{y}\} = \max\{(\boldsymbol{I}_M - 1/M * \boldsymbol{1}_{M\times M})\boldsymbol{S}^\top/(1+\sigma_w^2)\boldsymbol{y}\} = \max\{\boldsymbol{S}^\top\boldsymbol{y}\}$ which is the ML detector.

While we have assumed in this paper that the channel $\boldsymbol{H}$ and noise statistics $\boldsymbol{R}_{ww}$ were known, this is not likely to be the case in practice, and thus an adaptive scheme is desirable. Furthermore, the direct computation of the equalizer coefficients from equations (9) and (10) may not be feasible. Fortunately, the cost functions are quadratic, indicating we could use the LMS algorithm to calculate $\boldsymbol{F}$ and $\boldsymbol{G}$ via the update equations

$$\begin{aligned} \boldsymbol{\epsilon}[n] &= \boldsymbol{F}[n]^\top \boldsymbol{y}[Pn-\delta_r] + \boldsymbol{G}[n]^\top \hat{\boldsymbol{x}}[Mn] \\ &\quad -\boldsymbol{E}_\delta^\top \boldsymbol{x}[Pn-\delta_r] \\ \boldsymbol{F}[n+1] &= \boldsymbol{F}[n] - \mu_1 \boldsymbol{y}[Pn-\delta_r]\boldsymbol{\epsilon}[n]^\top \\ \boldsymbol{G}[n+1] &= \boldsymbol{G}[n] - \mu_2 \hat{\boldsymbol{x}}[Mn]\boldsymbol{\epsilon}[n]^\top \end{aligned}$$

where μ_1 and μ_2 are small positive constants. Note that the above equations implicitly assume training data is available since $\boldsymbol{x}[Pn-\delta_r]$ appears in the error calculation. Alternatively, the system could be operated in a decision-directed mode.

Lastly, while our block structure may seem complicated at first glance, it requires significantly fewer operations than a single scalar equalizer operating at the chip rate. A scalar equalizer operating at the chip rate would require $P(N_f + N_g)$ multiply-accumulate (MAC) operations to equalize one symbol. On the other hand, our structure only requires $(M-1)(N_f+N_g-N_g/M+1)$ MACs, plus an additional $2(M-1)$ multiplies for the matrix $\boldsymbol{U}^\top$.

V. Simulations

Here, we consider a simulation setup with the following properties. We employ 4-ary Walsh codes for the signal waveforms, and we include 2 samples of guard time. Thus,

$$\boldsymbol{S}^\top = \frac{1}{2}\begin{bmatrix} +1 & +1 & +1 & +1 & 0 & 0 \\ -1 & +1 & -1 & +1 & 0 & 0 \\ +1 & -1 & -1 & +1 & 0 & 0 \\ -1 & -1 & +1 & +1 & 0 & 0 \end{bmatrix}$$

and so $M=4$ and $P=6$. As the channel impulse response, we choose $\boldsymbol{h} = [(2/3)(-8/15)(1/5)(2/5)(-4/15)]^\top$, and we assume the noise is AWGN so $\boldsymbol{R}_{ww} = \sigma^2 \boldsymbol{I}_{N_f}$. The equalizer design parameters are chosen to be $N_f = 2P$, $N_g = M$, and $\delta = 4$.

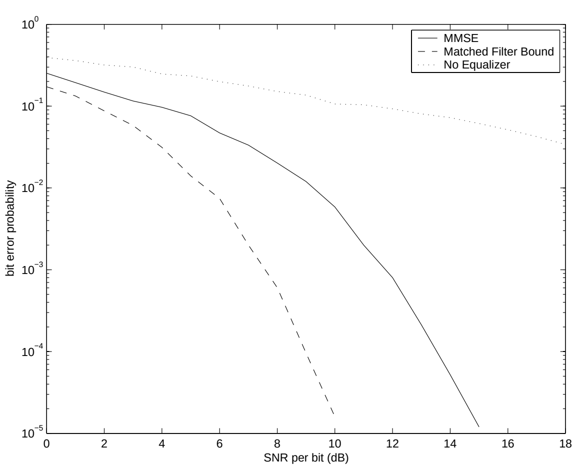

Fig. 4. Simulated performance of proposed structure

The performance is shown in Fig. 4, and is compared with the case of no equalization, and the matched filter bound (which may not be attainable, even for an ML receiver with exponential complexity). Further simulations of the proposed equalizer appeared in [6] for the PPM case, where we obtained performance 3 dB better than that of [3].

Another simulation was conducted to demonstrate the convergence of an adaptive equalizer when the LMS is employed. For the same setup above, with $SNR = 10dB$, we have plotted the time history of the MSE in Fig. 5. The equalizers were initialized to all zeroes, and converged to the MMSE solution after approximately 1000 symbols.

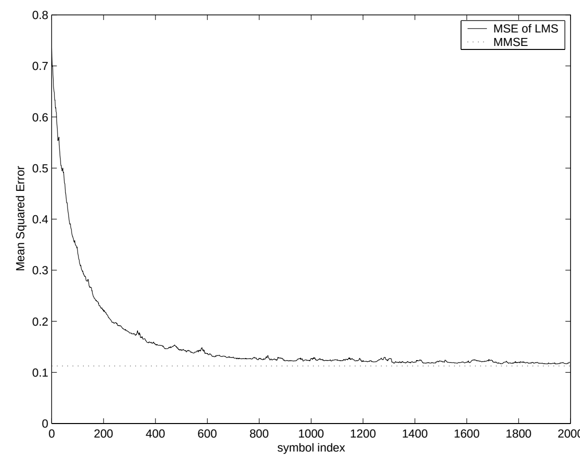

Fig. 5. MSE for LMS algorithm

VI. Conclusion

In this paper we proposed a constrained-length FIR MMSE DFE for orthogonal multipulse modulated signals, and we showed the design equations and performance of the proposed structure. This structure permitted several benefits over previously proposed unconstrained-length ZF schemes – namely, less noise enhancement, lower computational load, generalization to a larger class of channels, and the ability to easily apply gradient decent algorithms.

Further work could investigate properties of (possibly blind) adaptive implementations of this structure and the effect of DFE error propagation.

References

[1] E.A. Lee and D.G. Messerschmitt, *Digital Communications*, Kluwer Academic Publishers, 1994.

[2] M.K. Varanasi, "Equalization for Multipulse Modulation," *1997 IEEE International Conference on Personal Wireless Communications*, pp. 48-51, December 1997.

[3] J.R. Barry, "Sequence Detection and Equalization for Pulse-Position Modulation," *IEEE International Conference on Communications (ICC 94) Conference Record*, vol. 3, pp. 1561-1656, May 1994.

[4] N. Al-Dhahir and A.H. Sayed, "The Finite-Length Multi-Input Multi-Output MMSE-DFE," *IEEE Transactions on Signal Processing*, vol. 48, pp. 2921-2936, October 2000.

[5] J.G. Proakis, *Digital Communications*, McGraw-Hill, 3rd ed., 1995.

[6] A.G. Klein and C.R. Johnson, Jr., "MMSE Decision Feedback Equalization of Pulse Position Modulated Signals," *2004 IEEE International Conference on Communications (ICC 04) Conference Record*, June 2004.