MMSE Decision Feedback Equalization of Pulse Position Modulated Signals

A.G. Klein and C.R. Johnson, Jr.
School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853
email: agk5@cornell.edu

***Abstract*— We propose a minimum mean squared error (MMSE) decision feedback equalizer (DFE) for pulse position modulated (PPM) signals in the presence of intersymbol interference (ISI). While traditional uses of PPM may not have ISI, PPM is a candidate modulation scheme for ultra wideband (UWB), and may experience ISI in that application. 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

Pulse position modulation (PPM) is a modulation scheme that has been studied for many applications, including wireless optical communication and ultra wideband (UWB) communication [1]. Traditional uses of PPM have been in situations with little or no intersymbol interference (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. PPM is a power efficient scheme, but is bandwidth inefficient, and thus has attracted recent attention for use in UWB communication systems where ISI *will* be an issue. The unconstrained-length zero-forcing (ZF) decision feedback equalizer (DFE) for PPM was derived in [2] 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 constrained length minimum mean-squared error (MMSE) DFE for PPM signals, and we show the design equations and performance of the proposed structure. Furthermore, we completely remove the above assumptions – 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 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.

We will 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. As PPM is typically used as a baseband transmission scheme, all signals are assumed to be real-valued.

II. System Model

M-ary PPM is an orthogonal transmission scheme where a symbol consists of M "chips", only one of which is non-zero. PPM can be thought of as a block coding scheme where information is conveyed by the location of the non-zero sample within the block of M chips. Thus, the symbol alphabet comprises the M columns of the identity matrix $\boldsymbol{I}_M$. We assume that adjacent symbols are sent with no guard time between them, and we assume a discrete-time model where each chip is sampled once.

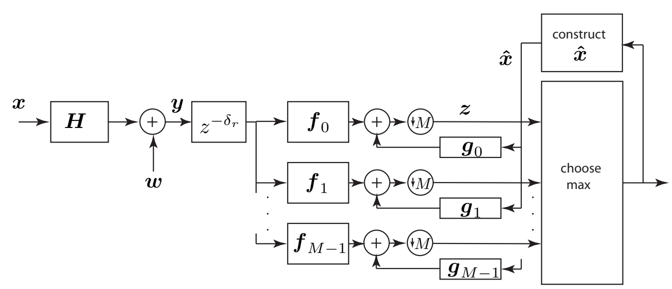

Fig. 1. MIMO DFE block diagram.

The system model is shown in Fig. 1. We have i.i.d. M-ary PPM symbols transmitted through a linear time-invariant FIR channel of length N_h with impulse response $\boldsymbol{h} = [h[0] \ \ldots \ h[N_h - 1]]^\top$, and 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 PPM chips, and $\boldsymbol{w}[n] \in \mathbb{R}^{N_f}$ is the noise vector. More precisely, we have

$$\boldsymbol{x}[n] = \begin{bmatrix} x[n] & x[n-1] & \cdots & x[n-N_f-N_h+2] \end{bmatrix}^\top$$

$$\boldsymbol{H} = \begin{bmatrix} h[0] & \cdots & h[N_f-1] & 0 & \cdots & & 0 \\ 0 & h[0] & \cdots & & h[N_f-1] & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ & & h[0] & & \cdots & & h[N_f-1] & 0 \\ 0 & \cdots & 0 & & h[0] & \cdots & & h[N_f-1] \end{bmatrix}.$$

A complete PPM symbol is generated at times that are multiples of M, so $\{x[-M+1], \dots, x[-1], x[0]\}$ would be a complete PPM symbol. The autocorrelation of a length MN stream of i.i.d. PPM symbols aligned to the symbol boundary is given by the block Tœplitz matrix

$$\begin{aligned} \boldsymbol{R} &= \frac{1}{M^2}(\mathbf{1}_{NM} + (\boldsymbol{I}_N \otimes (M\boldsymbol{I}_M - \mathbf{1}_M))) \\ &= \frac{1}{M^2} \begin{bmatrix} M\boldsymbol{I}_M & \mathbf{1}_M & \mathbf{1}_M & \cdots \\ \mathbf{1}_M & M\boldsymbol{I}_M & \mathbf{1}_M & \\ \mathbf{1}_M & \mathbf{1}_M & \ddots & \\ \vdots & & & \end{bmatrix}. \end{aligned} \quad (2)$$

In the sequel, we will refer to this matrix using the notation $\boldsymbol{R}_{i:j,\, k:l}$ to denote the extraction of rows i through j and columns k through l, and we assume N is as large as necessary to permit this extraction.

Note that $x[n]$ 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 $M = 2$, for example. This motivates the use of a different set of equalizer coefficients for each of the M chips of a symbol, and so at the receiver we employ a bank of M decision-feedback equalizers. Each equalizer is operating only once every M chips; hence, the number of computations for the bank of M equalizers operating at the symbol rate is equal to the number of computations required for a single equalizer operating at the chip rate. The feedforward equalizers have impulse response $\boldsymbol{f}_i = [f_i[0] \dots f_i[N_f-1]]^\top$ and length $N_f \geq M$. The feedback equalizers $\boldsymbol{g}_i = [g_i[0] \dots g_i[N_g-1]]^\top$ each have length N_g, assumed to be a multiple of M, and they operate on the signal $\hat{\boldsymbol{x}}[n] \in \mathbb{R}^{N_g}$. 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 M, and after downsampling we keep the δ_r-th polyphase component. Thus, the output of the M equalizers is

$$\boldsymbol{z}[n] = \boldsymbol{F}^\top \boldsymbol{y}[Mn - \delta_r] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn]. \quad (3)$$

Due to the downsampling operation, $\boldsymbol{z}[n] \in \mathbb{R}^M$ does not live on the same time-scale as the other signals. That is, $\boldsymbol{z}[n]$ is a symbol-rate signal, whereas the other signals are all 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 of the M received chips of the symbol. While such a memoryless decision device is by no means optimal in the presence of ISI, this is the decision device that is assumed here 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 chip estimates where only one chip per symbol is set to 1 and all others are 0. The detector output is then accumulated in the vector $\hat{\boldsymbol{x}}[Mn]$, which is input to the feedback equalizer. Note that $\hat{\boldsymbol{x}}$ is once again on the same time scale as all the other signals (i.e. the chip rate).

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

III. Previous Work: The Zero-Forcing Equalizer

One of the schemes proposed in [2], referred to as the block DFE, uses the above structure with the following parameter choices: $\delta_r = 0$, $N_f = M$, $\boldsymbol{F} = ([\boldsymbol{I}_M\, \mathbf{0}_{M\times(N_h-1)}]\boldsymbol{H}^\top)^{-1}$, $N_g = N_h - 1$, and $\boldsymbol{G} = -[\mathbf{0}_{(N_h-1)\times M}\, \boldsymbol{I}_{N_h-1}]\boldsymbol{H}^\top \boldsymbol{F}$. This is the (non-unique) ZF DFE. Several additional structures are considered in [2] 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 $\hat{\boldsymbol{x}}[Mn] = [\mathbf{0}_{(N_h-1)\times M}\, \boldsymbol{I}_{N_h-1}]\boldsymbol{x}[Mn]$. Then, we partition the channel matrix as $\boldsymbol{H} = [\boldsymbol{H}_0\, \boldsymbol{H}_1]$ where

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

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 (3) and (1), the output of the ZF equalizer is then

$$\begin{aligned} \boldsymbol{z}[n] &= \boldsymbol{F}^\top \boldsymbol{y}[Mn] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn] \\ &= \underbrace{[\boldsymbol{I}_M\, \mathbf{0}_{M\times(N_h-1)}]\boldsymbol{x}[Mn]}_{\text{ISI-free symbol}} + \underbrace{\boldsymbol{H}_0^{-1}\boldsymbol{w}[Mn]}_{\text{noise term}} \end{aligned} \quad (4)$$

and so the output of the ZF equalizer proposed in [2] is the ISI-free PPM symbols plus filtered noise.

IV. The Proposed MMSE Equalizer

The previously proposed ZF scheme has 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 ZF 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 and is suitable for use with adaptive gradient descent algorithms (e.g. least mean squares).

The scheme proposed in [2] implicitly assumes 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 M; thus, we will adjust for this with the bulk delay at the front of the receiver. Let $\delta_t = \lceil \delta/M \rceil M$ 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. The output error is then

$$\boldsymbol{\epsilon} = \boldsymbol{F}^\top \boldsymbol{y}[Mn-\delta_r] + \boldsymbol{G}^\top \hat{\boldsymbol{x}}[Mn] - \boldsymbol{E}_\delta^\top \boldsymbol{x}[Mn-\delta_r]. \quad (5)$$

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 [3], the same set of equalizer coefficients minimizes both measures. Here, we choose the trace:

$$\begin{aligned} J(\boldsymbol{F},\boldsymbol{G}) &= \operatorname{tr}(E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top]) \\ &= \operatorname{tr}[\boldsymbol{F}^\top(\boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{H}^\top + \boldsymbol{R}_{ww})\boldsymbol{F} + 2\boldsymbol{F}^\top \boldsymbol{H}\boldsymbol{R}_{x\hat{x}}\boldsymbol{G} \\ &\quad -2\boldsymbol{F}^\top \boldsymbol{H}\boldsymbol{R}_{xx}\boldsymbol{E}_\delta + \boldsymbol{G}^\top \boldsymbol{R}_{\hat{x}\hat{x}}\boldsymbol{G} \\ &\quad -2\boldsymbol{G}^\top \boldsymbol{R}_{x\hat{x}}^\top \boldsymbol{E}_\delta + \boldsymbol{E}_\delta^\top \boldsymbol{R}_{xx}\boldsymbol{E}_\delta] \end{aligned} \quad (6)$$

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

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

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

$$\boldsymbol{R}_{\hat{x}\hat{x}} \triangleq E[\hat{\boldsymbol{x}}[Mn]\hat{\boldsymbol{x}}[Mn]^\top]. \quad (9)$$

To find the MMSE equalizer settings, we begin by setting the derivative of (6) with respect to $\boldsymbol{G}$ to zero

$$\begin{aligned} \frac{1}{2}\frac{\partial}{\partial \boldsymbol{G}} J(\boldsymbol{F},\boldsymbol{G}) &= \boldsymbol{R}_{x\hat{x}}^\top \boldsymbol{H}^\top \boldsymbol{F} + \boldsymbol{R}_{\hat{x}\hat{x}}\boldsymbol{G} - \boldsymbol{R}_{x\hat{x}}^\top \boldsymbol{E}_\delta \\ &\triangleq \boldsymbol{0}_{M\times 1} \\ \Longrightarrow \boldsymbol{G}^* &= \underbrace{\boldsymbol{R}_{\hat{x}\hat{x}}^{-1}\boldsymbol{R}_{x\hat{x}}^\top}_{\triangleq \boldsymbol{B}}(\boldsymbol{E}_\delta - \boldsymbol{H}^\top \boldsymbol{F}). \end{aligned} \quad (10)$$

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

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

After solving this equation for $\boldsymbol{F}$, we have an equation for the MMSE feedforward equalizer coefficients, shown on the next page in (11).

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 response to a single spike with appropriate delay, so $\boldsymbol{E}_\delta = [\boldsymbol{0}_{M\times\delta}\ \boldsymbol{I}_M\ \boldsymbol{0}_{M\times(N_f+N_h-1-M-\delta)}]^\top$. However, the desired response does not need to be a single spike. The "choose max" decision device is invariant to an added constant, so long as the same constant is added to each of the M chips.

Now we will motivate a specific choice of the matrix $\boldsymbol{E}_\delta$, and make a modification to our equalizer structure. In the binary PPM case where $M = 2$, the "choose max" decision device will 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, thereby reducing the bank of two equalizers to a single equalizer. 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

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

which corresponds to a response where the constant $1/M$ has been subtracted from each of the M chips, and thus will not change the operation of the decision device. The matrix $\boldsymbol{I}_M - 1/M\boldsymbol{1}_{M\times M}$ has a zero eigenvalue, which is a fact that we will exploit. Let $\boldsymbol{U}^\top\boldsymbol{U} = \boldsymbol{I}_M - 1/M\boldsymbol{1}_{M\times M}$ be the Cholesky factorization. Since the matrix is only *semi*-positive definite, the last row of the upper triangular matrix $\boldsymbol{U}$ will be all zero. Next, define

$$\boldsymbol{E}'_\delta = [\boldsymbol{0}_{M\times\delta}\ \boldsymbol{U}\ \boldsymbol{0}_{M\times(N_f+N_h-1-M-\delta)}]^\top$$

so that $\boldsymbol{E}_\delta = \boldsymbol{E}'_\delta\boldsymbol{U}$. From (3), (10), and (11) the MMSE equalizer output is given by

$$\boldsymbol{z}[n] = \boldsymbol{E}_\delta^\top \boldsymbol{A}^\top \boldsymbol{y}[Mn-\delta_r] + (\boldsymbol{E}_\delta - \boldsymbol{H}^\top \boldsymbol{F})^\top \boldsymbol{B}^\top \hat{\boldsymbol{x}}[Mn]. \quad (12)$$

Now, we will modify the equalizer structure by multiplying the M outputs of the equalizer by the $M\times M$ matrix $\boldsymbol{U}^\top$, just prior to the decision device. Furthermore, instead of designing the equalizers to have response $\boldsymbol{E}_\delta$, we will design them to have response $\boldsymbol{E}'_\delta$. The output of the modified structure will be

$$\begin{aligned} \boldsymbol{U}^\top \boldsymbol{z}'[n] &= \boldsymbol{U}^\top \Bigg[\underbrace{(\boldsymbol{E}'_\delta)^\top \boldsymbol{A}^\top}_{\triangleq (\boldsymbol{F}')^\top} \boldsymbol{y}[Mn-\delta_r] \\ &\quad + \underbrace{(\boldsymbol{E'}_\delta - \boldsymbol{H}^\top \boldsymbol{F}')^\top \boldsymbol{B}^\top}_{\triangleq (\boldsymbol{G}')^\top} \hat{\boldsymbol{x}}[Mn] \Bigg] \\ &= \boldsymbol{E}_\delta^\top \boldsymbol{A}^\top \boldsymbol{y}[Mn-\delta_r] \\ &\quad + (\boldsymbol{E}_\delta - \boldsymbol{H}^\top \boldsymbol{F})^\top \boldsymbol{B}^\top \hat{\boldsymbol{x}}[Mn] \end{aligned} \quad (13)$$

where we have used $\boldsymbol{E}_\delta = \boldsymbol{E}'_\delta\boldsymbol{U}$. Thus, the output of the modified structure in (13) is equivalent to the output of the

$$F^* = \underbrace{\left[H(R_{xx} - R_{x\hat{x}}R_{\hat{x}\hat{x}}^{-1}R_{x\hat{x}}^\top)H^\top + R_{ww}\right]^{-1} H(R_{xx} - R_{x\hat{x}}R_{\hat{x}\hat{x}}^{-1}R_{x\hat{x}}^\top)}_{\triangleq A} E_\delta \quad (11)$$

original structure in (12). 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 $2(M-2)$ multiplies since each column of $U^\top$ has only 2 unique elements, one column is all zeros, and one column contains ± 1 after normalization. In an adaptive setting, the computational savings is even greater since we need to update $N_f + N_g$ fewer parameters.

Closer inspection of the entries of the desired response U reveals that the equalizer is effectively mapping the M PPM alphabet members 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 [4], whose signals have dimensionality $M-1$ but maintain the same Euclidean distance as that between PPM signals. By translating and rotating the original PPM 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, so that $\hat{x}[n] = x[n - M - \delta_t]$ or

$$\hat{x}[Mn] = \underbrace{\begin{bmatrix} 0_{\delta+M\times N_g} \\ I_{N_g} \\ 0_{(N_f+N_h-1-\delta-M-N_g)\times N_g} \end{bmatrix}^\top}_{\triangleq W} x[Mn - \delta_r] \quad (14)$$

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

Note from (11) 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

$$R_{\hat{x}\hat{x}} = 1/M^2 \begin{bmatrix} MI_M & 1_{M\times M} \\ 1_{M\times M} & MI_M \end{bmatrix} \quad (15)$$

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, g and g' where

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

Both of these feedback equalizers give the same output, regardless of the value of a since $\hat{x} \in \{[1\,0\,1\,0],\ [1\,0\,0\,1],\ [0\,1\,1\,0],\ [0\,1\,0\,1]\}$, and so $g^\top \hat{x} = (g')^\top \hat{x} \in \{g[0]+g[2],\ g[0]+g[3],\ g[1]+g[2],\ g[1]+g[3]\}$. It is no coincidence that the vector $[+a\ +a\ -a\ -a]^\top$ is a basis vector for the nullspace of $R_{\hat{x}\hat{x}}$ in (15). 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 g' is equal to zero.

In the M-ary case with arbitrary $N_g \geq M$, any set of feedback equalizer taps g has a corresponding set of taps g' that gives an identical output, but has at least $N_g/M - 1$ of its taps equal to zero. In particular, taps at locations $g'[Mn-1]$ for all integers n such that $2 \leq n \leq N_g/M$ will be zeroed. 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 reduced length $N_g - N_g/M + 1$ that operates on a decimated version of $\hat{x}$. The decimated entries correspond to the zeroed equalizer taps, and the decimated signal can be represented as

$$\underbrace{\begin{bmatrix} I_M & 0_{M\times N_g-M} \\ 0_{N_g-N_g/M+1-M\times M} & I_{N_g/M-1}\otimes[I_{M-1}\ 0_{M-1\times 1}] \end{bmatrix}}_{\triangleq V} \hat{x}[Mn]. \quad (16)$$

Note that $V \in \mathbb{R}^{(N_g-N_g/M+1)\times N_g}$ and our new shortened equalizer $G'' \in \mathbb{R}^{(N_g-N_g/M+1)\times M}$ has output $(G'')^\top V\hat{x}[Mn]$ which we claim can made to have identical output to $G^\top \hat{x}[Mn]$ for arbitrary G. It turns out that $R_{\hat{x}\hat{x}} \in \mathbb{R}^{N_g\times N_g}$ has rank $N_g - N_g/M + 1$, and by throwing out $N_g/M - 1$ carefully selected samples (i.e. by forming $V\hat{x}[Mn]$), the autocorrelation of the decimated signal becomes $VR_{\hat{x}\hat{x}}V^\top$ 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 Fig. 1 still applies, with the exception that z must be pre-multiplied by $U^\top$ before entering the decision device. Now, we will drop the primes (e.g. $F' \to F$) 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 $\boldsymbol{U}$ via the Cholesky factorization $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_M - 1/M \mathbf{1}_{M\times M}$, which is a fixed matrix that depends only on M. Next, form the desired response as $\boldsymbol{E}_\delta = [\mathbf{0}_{M\times\delta}\ \boldsymbol{U}\ \mathbf{0}_{M\times(N_f+N_h-1-M-\delta)}]^\top$. 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 (10) and (11). The last column of $\boldsymbol{F}$ and $\boldsymbol{G}$ will both be zero, so effectively we have $M-1$ equalizers. Under the assumption of correct decisions, the feedback equalizers are operating on a decimated version of $\boldsymbol{x}[Mn-\delta_r]$ given by

$$\hat{\boldsymbol{x}}[Mn] = \boldsymbol{V}\boldsymbol{W}\boldsymbol{x}[Mn-\delta_r] \tag{17}$$

where $\hat{\boldsymbol{x}}[Mn] \in \mathbb{R}^{N_g-N_g/M+1}$, $\boldsymbol{W}$ windows the appropriate samples of $\boldsymbol{x}$ as defined in (14), and $\boldsymbol{V}$ performs the decimation as defined in (16). The input to the "choose max" decision device is then

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

The correlation matrices needed in the equalizer design equations follow directly from (2), (7)-(9) and (17) giving

$$\begin{aligned}
\boldsymbol{R}_{xx} &= \boldsymbol{R}_{\delta_r:\delta_r+N_f+N_h-2,\delta_r:\delta_r+N_f+N_h-2} \\
\boldsymbol{R}_{x\hat{x}} &= \boldsymbol{R}_{xx}\boldsymbol{W}^\top \boldsymbol{V}^\top \\
\boldsymbol{R}_{\hat{x}\hat{x}} &= \boldsymbol{V}\boldsymbol{W}\boldsymbol{R}_{xx}\boldsymbol{W}^\top \boldsymbol{V}^\top.
\end{aligned}$$

While we have assumed 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 via (10) and (11) 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

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

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

V. Simulations

Here, we consider the same simulation setup as was performed in [2]. That is, $M = 2$, $\boldsymbol{h} = \begin{bmatrix}1 & -1 & 1\end{bmatrix}^\top/\sqrt{3}$, and the noise is AWGN. For the MMSE equalizer, we chose $N_f = 5$, $N_g = 2$, and $\delta = 3$. The results are shown in Fig. 2, and the proposed MMSE equalizer demonstrates approximately a 3 dB performance gain over the ZF scheme of [2], even at high SNR.

Typically ZF equalizers have similar performance to MMSE equalizers in an ISI-dominated regime where the SNR is high. However, this is not the case here because our proposed structure is fundamentally different in the way we perform the feedforward equalization. In the scheme of [2], the feedfoward equalizer $\boldsymbol{F}$ will always be a square lower triangular Tœplitz matrix, so the M-th chip of each symbol effectively has no feedforward equalization – only a gain. Since our proposed scheme does not impose this structure, it has a better ability to suppress ISI in addition to having less noise enhancement.

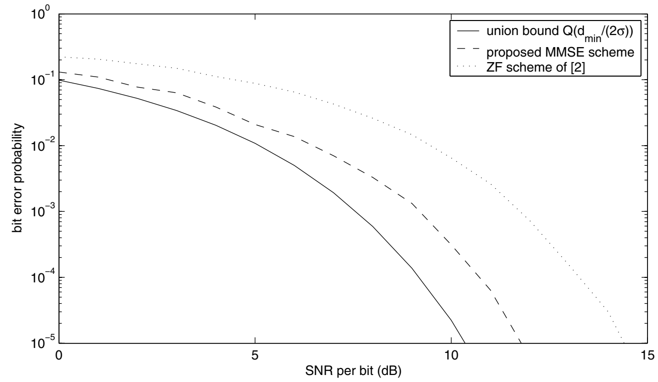

Fig. 2. Comparison of ZF and MMSE bit error rate.

The complexity of the proposed MMSE equalizer is less than that of a single scalar equalizer operating at the chip rate. To equalize one symbol, a single scalar equalizer at the chip rate requires $M(N_f+N_g)$ multiplies, while our proposed equalizer requires $(M-1)(N_f+N_g-N_g/M+1)+2(M-2)$ multiplies, and the scheme proposed in [2] requires $M(M+1)/2+(N_h-1)M$ multiplies. For the simulation in Fig. 2, the ZF and MMSE schemes have the same complexity.

VI. Conclusion

In this paper we proposed a MMSE DFE for PPM signals, and we showed the design equations and performance of the proposed structure. This structure permitted several benefits over the previously proposed ZF scheme – namely, better performance, 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 adaptive implementations of this structure, or extend the model to accommodate guard times between symbols, which are common in UWB PPM systems.

Acknowledgment

This work supported in part by National Science Foundation Grant CCR 0310023, Applied Signal Technology, and Texas Instruments.

References

[1] R. Scholtz, "Multiple Access with Time-Hopping Impulse Modulation," *Military Communications Conference (MILCOM) Conference Record*, vol. 2, pp. 447-450, October 1993.

[2] 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.

[3] 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.

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