

AFFINE PRECODERS FOR RELIABLE COMMUNICATIONS

Jonathan H. Manton, Iven Y. Mareels and Yingbo Hua

Department of Electrical and Electronic Engineering
The University of Melbourne, Parkville, Victoria 3052, Australia.

ABSTRACT

It is known that precoding a signal prior to its transmission through an unknown finite impulse response channel facilitates the equalisation of the channel. Moreover, since linear precoders spread the spectrum, the equalisation process mitigates the effects of channel spectral nulls caused by frequency selective fading. This paper argues that certain affine precoders are more efficient than linear precoders. Here, the efficiency of a precoder is a measure of its ability to enable the receiver to recover the source signal relative to the amount of redundancy introduced by the precoder. An efficient affine precoder is heuristically derived and simulations used to demonstrate that it is superior to both the traditional training sequence method and the more recent filter bank precoding scheme. This precoding and equalisation scheme naturally extends to time varying channels.

1. INTRODUCTION

This paper propounds an efficient precoding scheme for the reliable transmission of a signal through an unknown finite impulse response (FIR) channel. It is assumed that an upper bound on the channel length is known. This fundamental problem has been studied extensively in the literature with many channel equalisation schemes proposed. Recently it was shown that linear precoders not only enable the equalisation of the unknown channel [1, 5], they also spread the spectrum [4], making it possible to mitigate the effects of spectral nulls in the channel. This paper takes the idea one step further by introducing *affine precoders* and demonstrating their superiority over linear precoders as well as over traditional training sequence based methods.

Only precoding and equalisation schemes which accurately recover the source signal from a very short output sequence are considered here. This requirement is necessary if the channel parameters vary with time. It has been shown in [5] that certain linear precoders (and hence affine precoders) enable the receiver to identify the channel based solely on the algebraic structure of the precoded signal, that is, no statistical structure of the source signal is required. The philosophy adopted here is that statistical and finite alphabet properties of the source signal should be exploited only at a later stage, and only to enhance the overall performance. This eliminates the ill effects of short data lengths and statistical anomalies associated with receivers which rely on statistical structure.

Optimal linear precoders and equalisers, under the simplifying assumption that the channel is known, are derived in [2, 6] and references therein. When the channel is unknown though, the design of efficient precoders is an unresolved problem. This paper shows that certain affine precoders are more efficient than linear

precoders. It also argues that affine precoders are efficient even when compared to non-linear precoders.

The motivation for considering the class of affine precoders is the observation that neither a training sequence nor a linear precoder on its own is entirely adequate. A training sequence efficiently identifies the channel but cannot guard against channel spectral nulls. A linear precoder guards against channel spectral nulls but simulations show it is not efficient¹ at identifying the channel [3]. The explanation for this is that each channel parameter is a rational function (a ratio of two polynomial functions) of the channel output if a linear precoder is used [5]. Poor performance results when the denominator of this rational function is close to zero. This phenomenon shows up as heavy tails on the error distribution [3]; the performance of a linear precoder is sensitive to the actual sequence transmitted.

This paper is based on two key ideas. The first is that if both a training sequence and a linear precoder are used then each method's strength compensates for the other's weakness. The second key idea, justified in Section 3, is that the *constellation* of the precoder should be a subset of the surface of a sphere. Both these ideas are incorporated into a candidate affine precoder in Section 3. An interesting question to ask is how much power should be allocated to the training sequence, and how much to the linear precoder? Simulations in Section 4 study this issue.

The remainder of this paper is organised as follows. Section 2 states the channel equalisation problem and defines the class of affine precoders. Section 3 provides a geometrical interpretation of precoding followed by a series of arguments which lead to a candidate affine precoder. Also derived is an algorithm to determine the MLE of the channel. Simulation results in Section 4 reveal that the candidate affine precoder is more efficient than either a training sequence based method or a linear precoder on its own. Conclusions and further work are presented in Section 5.

2. PROBLEM FORMULATION

This paper studies the problem of transmitting a source vector $\mathbf{s} \in \mathbb{R}^p$ through an FIR channel of length L with unknown parameters $\mathbf{h} = [h_0, \dots, h_{L-1}]^T$. The source \mathbf{s} is mapped to the vector $\mathbf{x} \in \mathbb{R}^{n+L-1}$. The relationship between the transmitted vector $\mathbf{x} = [x_{2-L}, \dots, x_0, x_1, \dots, x_n]^T$ and the received vector $\mathbf{z} \in \mathbb{R}^n$ is

$$\mathbf{z} = \mathbf{y} + \mathbf{n}, \quad \mathbf{y} = \mathbf{h} \star \mathbf{x}, \quad \mathbf{n} \sim N(0, \sigma^2 \mathbf{I}) \quad (1)$$

where σ^2 is the unknown noise variance and \star denotes convolution, that is, $y_k = \sum_{i=0}^{L-1} h_i x_{k-i}$ for $k = 1, \dots, n$. The noise

¹Recall it is assumed that the source statistics are unknown. Linear precoders perform well if longer data lengths and knowledge of the source statistics are available.

This work was supported by the Australian Research Council.

vector \mathbf{n} is referred to as AWGN (additive white Gaussian noise). A fundamental problem is to choose the mapping from \mathbf{s} to \mathbf{x} in such a way as to minimise the sensitivity of the receiver's estimate of the source to different channel vectors \mathbf{h} and noise power σ^2 .

The idea of using an **affine precoder** is introduced in this paper. That is, the relationship between the transmitted vector \mathbf{x} and the source vector \mathbf{s} can be written as

$$\mathbf{x} = A\mathbf{s} + \mathbf{b} \quad (2)$$

for some matrix $A \in \mathbb{R}^{(n+L-1) \times p}$ and vector $\mathbf{b} \in \mathbb{R}^{n+L-1}$. It is assumed that A has full rank and that \mathbf{b} is orthogonal to the range space of A , that is, $\mathbf{b}^T A\mathbf{s} = 0$ for all \mathbf{s} .

For any precoder (2), define its **constellation** to be

$$C = \{\mathbf{x} : \mathbf{x} = A\mathbf{s} + \mathbf{b}, \mathbf{s} \in \Omega\}, \quad \Omega \subset \mathbb{R}^p \quad (3)$$

where Ω is the set of all possible source vectors \mathbf{s} . It is shown in Section 3 that the intrinsic performance of a precoder depends only on its constellation. Also used in Section 3 is the fact that certain affine precoders can embed a low dimensional unit sphere into a high dimensional one. To state this precisely, define the sphere

$$S^{p-1} = \{\mathbf{s} \in \mathbb{R}^p : \|\mathbf{s}\| = 1\}$$

where $\|\cdot\|$ denotes the Euclidean norm. Under the assumption that $\mathbf{b}^T A\mathbf{s} = 0$ for all \mathbf{s} it is clear that $\|\mathbf{x}\|^2 = \mathbf{s}^T A^T A\mathbf{s} + \mathbf{b}^T \mathbf{b}$. Thus, if all the eigenvalues of $A^T A$ are equal to λ^2 , $\|\mathbf{x}\|^2 = \lambda^2 \|\mathbf{s}\|^2 + \|\mathbf{b}\|^2$. In particular, there exist invertible affine precoders which map S^{p-1} into S^{n+L-2} .

3. PRECODER DESIGN

This section derives a candidate affine precoder which is robust to unknown channels and additive noise, as well as an algorithm to compute the corresponding MLE of the channel.

3.1. Geometric Interpretation of Precoding

The following argument shows that the intrinsic performance of a precoder depends only on its constellation. Let f_1 and f_2 denote two precoders with the same constellation, that is, $f_1(\Omega) = f_2(\Omega)$. (Note that any precoder is implicitly assumed to be invertible.) Then there exists an invertible function g such that $f_1 = f_2 \circ g$ where \circ denotes composition; the performance of f_1 is identical to the performance of f_2 under a suitable transformation of the source vector.

The performance depends on the actual channel \mathbf{h} through both its magnitude $\|\mathbf{h}\|$ and its direction $\mathbf{h}/\|\mathbf{h}\|$. Changing the magnitude $\|\mathbf{h}\|$ simply scales the noise-free output \mathbf{y} . The effect of changing $\mathbf{h}/\|\mathbf{h}\|$ is more interesting and motivates defining the set

$$\tilde{Y}_{\mathbf{x}} = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} = \mathbf{h} * \mathbf{x}, \mathbf{h} \in S^{L-1}\}$$

where S^{L-1} denotes the $L-1$ dimensional sphere. Then for each transmitted vector \mathbf{x} , the set $\tilde{Y}_{\mathbf{x}}$ is an ellipsoid. Moreover, for different \mathbf{x} , the shape of the ellipsoid $\tilde{Y}_{\mathbf{x}}$ is in general different. Two extreme cases are $\tilde{Y}_{\mathbf{x}}$ being a sphere, and $\tilde{Y}_{\mathbf{x}}$ being a line segment. An example of the former is $\mathbf{x} = [0, \dots, 0, 1, 0, \dots, 0]$ where there are $L-1$ zeros both before and after the one. In this case, the channel output power $\|\mathbf{y}\|^2$ does not depend on $\mathbf{h}/\|\mathbf{h}\|$.

This is therefore the most robust way of identifying an unknown channel. The other extreme occurs in the infinite (and complex-valued) data case when a single (complex exponential) sinusoid is transmitted; the channel output is identical to the channel input except for a constant scaling factor.

The above demonstrates the difficulty in designing an efficient precoder; different transmitted vectors \mathbf{x} perform differently depending on the channel \mathbf{h} . For simplicity, it is initially assumed that both the transmitter and the receiver know the channel vector \mathbf{h} . The problem thus reduces to that of overcoming the AWGN. Applying Shannon's ideas [7] suggests that the noise is best overcome in the finite alphabet case by choosing the set of possible output vectors \mathbf{y} to contain equally spaced points inside the solid ball $\{\mathbf{y} : \|\mathbf{y}\| \leq 1\}$. Moreover, Shannon argues that as the dimension of \mathbf{y} increases, these points will tend to cluster near the surface of the solid ball.

If the transmitter does not know the channel it cannot ensure that \mathbf{y} is equally spaced on the surface of the unit sphere. If however it is known that the channel is close to being the identity then it is sensible to constrain \mathbf{x} to the surface of the unit sphere. Moreover, simulations show that the expected frequency response of a randomly chosen channel approaches that of the identity channel. This motivates requiring the constellation to be a subset of the unit sphere. If Ω , the set of possible source vectors, is an infinite set then the requirement that different \mathbf{y} are "equally spaced" can be made precise in the following way. Assume that the "error" involved in choosing \mathbf{s}_1 over \mathbf{s}_2 is a function of $\|\mathbf{s}_1 - \mathbf{s}_2\|$. Then the precoder function $G : \Omega \rightarrow \mathbb{R}^{n+L-1}$ should ensure the ratio $\|G(\mathbf{s}_1) - G(\mathbf{s}_2)\|/\|\mathbf{s}_1 - \mathbf{s}_2\|$ is as large as possible. In summary, the precoder G should satisfy

$$G \in \arg \max_G \left\{ \inf_{\substack{\mathbf{s}_1, \mathbf{s}_2 \in \Omega \\ \mathbf{s}_1 \neq \mathbf{s}_2}} \frac{\|G(\mathbf{s}_1) - G(\mathbf{s}_2)\|}{\|\mathbf{s}_1 - \mathbf{s}_2\|} \right\} \quad (4)$$

where the maximisation is over some suitable function space, such as the class of all smooth functions G which satisfy $G(\Omega) \subset S^{n+L-2}$.

Henceforth it is assumed for convenience that the possible source vectors are points on the unit sphere, that is, $\Omega = S^{p-1}$. The objective is to find a precoder G which satisfies (4). Section 2 shows that certain affine precoders embed S^{p-1} in S^{n+L-2} . It is now shown that these affine precoders also satisfy (4). Since the largest distance between any two points on a unit sphere is 2, it follows that $\|G(\mathbf{s}_1) - G(\mathbf{s}_2)\|/\|\mathbf{s}_1 - \mathbf{s}_2\| \leq 1$ if $\mathbf{s}_1 = -\mathbf{s}_2 \in S^{p-1}$ and the precoder G satisfies $G(S^{p-1}) \subset S^{n+L-2}$. Let G be an affine precoder $G(\mathbf{s}) = A\mathbf{s} + \mathbf{b}$ for which $\mathbf{b}^T A\mathbf{s} = 0$ for all \mathbf{s} and all the eigenvalues of $A^T A$ are equal to λ^2 . Then, as shown in Section 2, $\|G(\mathbf{s})\|^2 = \lambda^2 \|\mathbf{s}\|^2 + \|\mathbf{b}\|^2$. Thus (4) is satisfied if $\lambda^2 = 1$ and $\|\mathbf{b}\| = 0$. Such a G is a linear precoder and, as eluded to in Section 1, suffers from the choice $\|\mathbf{b}\| = 0$. If $\|\mathbf{b}\| > 0$ then there is only $1 - \|\mathbf{b}\|^2$ units of power left for encoding \mathbf{s} . Moreover, the affine precoder G fully utilises this power when $\lambda^2 = 1 - \|\mathbf{b}\|^2$.

In summary, there is a compromise between choosing a precoder G which satisfies (4) and sending a fixed training sequence \mathbf{b} ; the greater the power $\|\mathbf{b}\|$ the less spread out the possible transmitted vectors can be. Nevertheless, subject to the constraint of sending a fixed training sequence \mathbf{b} , the class of affine precoders $G(\mathbf{s}) = A\mathbf{s} + \mathbf{b}$ satisfy (4) for any A such that the eigenvalues of $A^T A$ are all equal to $1 - \|\mathbf{b}\|^2$ and $\mathbf{b}^T A\mathbf{s} = 0$ for all \mathbf{s} .

3.2. Candidate Affine Precoder

This section imposes further requirements on an affine precoder. The first requirement is that a *consecutive* training sequence be used. If the training sequence is not consecutive then the channel parameters \mathbf{h} will be rational, rather than polynomial, functions of the channel output \mathbf{y} . The presence of a polynomial term in the denominator makes the channel estimates sensitive to the actual sequence transmitted. If L denotes the length of the channel then the most efficient training sequence is $[0, \dots, 0, 1, 0, \dots, 0]$ where there are $L - 1$ zeros both before and after the one, a total of $2L - 1$ symbols. Not only is the power output of the channel constant regardless of the channel parameters and thus best suited to resisting the effects of AWGN, this training sequence also introduces the least amount of redundancy. Indeed, it is shown in [5] that almost all linear precoders require the equivalent of $2L$ known symbols to identify a channel with L unknown parameters.

The second requirement is that a linear precoder must have a zero prefix. A zero prefix is a sequence of $L - 1$ zeros which reset the initial state of the channel to zero. Linear precoders which do not have a zero prefix perform poorly [3] because estimating the initial state of the channel is an ill-conditioned problem. Both requirements are expediently met by placing the training sequence in front of the linear precoder. The trailing $L - 1$ zeros of the training sequence double as a zero prefix for the linear precoder.

All that remains is the choice of a linear precoder which satisfies the eigenvalue constraint $\lambda^2 = 1 - \|\mathbf{b}\|^2$ made in Section 3.1. Perhaps the simplest such linear precoder is one which introduces evenly spaced zeros. This leads to the following family of candidate affine precoders:

Break the source \mathbf{s} into blocks so that $\mathbf{s}^T = [\mathbf{s}_1, \dots, \mathbf{s}_m]$. For instance, if \mathbf{s} has $p = 15$ elements then it can be grouped into $m = 3$ blocks $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$ of five elements each. Choose a λ which lies between zero and one. Then the transmitted vector \mathbf{x} is

$$\mathbf{x} = \left[0, \dots, 0, \sqrt{1 - \lambda^2}, 0, \dots, 0, \lambda \mathbf{s}_1, 0, \lambda \mathbf{s}_2, 0, \dots, \lambda \mathbf{s}_m \right]^T \quad (5)$$

where there are $L - 1$ zeros both before and after the $\sqrt{1 - \lambda^2}$ term.

Remark: If h_0 is close to zero then the estimate of the last element \mathbf{s}_p of the source \mathbf{s} will be inaccurate. To overcome this, simply append one or more trailing zeros to the above affine precoder.

3.3. Maximum-Likelihood Estimate of the Channel

Computing the MLE under the assumption $\|\mathbf{s}\| = 1$ is a constrained optimisation problem and can be solved by a number of standard routines. A suitable cost function is now derived.

Define for $k = 0, \dots, L - 1$ the $(p + m - 1) \times (p + m - 1)$ matrix J^k whose ij th element is $(J^k)_{ij} = \delta(i - j - k)$ where δ is Dirac's delta function. That is, J^0 is the identity matrix, J^1 the shift matrix, and in general $J^k = (J^1)^k$. Define the $(p + m - 1) \times (p + m - 1)$ Toeplitz channel matrix $H = h_0 J^0 + \dots + h_{L-1} J^{L-1}$ and the $(p + m - 1) \times p$ precoder matrix P such that $[\lambda \mathbf{s}_1, 0, \lambda \mathbf{s}_2, 0, \dots, \lambda \mathbf{s}_m]^T = P\mathbf{s}$. Thus, if the affine precoder (5) is used, the channel output \mathbf{y} is

$$\mathbf{y}^T = \left[\mathbf{y}_1^T, \mathbf{y}_2^T \right], \quad \mathbf{y}_1 = \sqrt{1 - \lambda^2} \mathbf{h}, \quad \mathbf{y}_2 = \lambda H P \mathbf{s}.$$

Partition \mathbf{z} analogously so that $\mathbf{z}^T = [\mathbf{z}_1^T, \mathbf{z}_2^T]$. The MLE of \mathbf{h} and \mathbf{s} is obtained by minimising the cost function $\phi(\mathbf{s}, \mathbf{h}) = \phi_1(\mathbf{h}) + \phi_2(\mathbf{s}, \mathbf{h})$ where

$$\phi_1 = \frac{1}{2} \left\| \sqrt{1 - \lambda^2} \mathbf{h} - \mathbf{z}_1 \right\|^2, \quad \phi_2 = \frac{1}{2} \left\| \lambda H P \mathbf{s} - \mathbf{z}_2 \right\|^2.$$

This is a separable least squares problem; for fixed \mathbf{h} the MLE of \mathbf{s} is

$$\hat{\mathbf{s}} = \frac{1}{\lambda} (HP)^+ \mathbf{z}_2, \quad (HP)^+ = \left(P^T H^T H P \right)^{-1} P^T H^T.$$

Define the separated cost function $\Psi(\mathbf{h}) = \phi(\hat{\mathbf{s}}, \mathbf{h})$. It can be shown [3] that its gradient vector is

$$\dot{\Psi}(\mathbf{h}) = (1 - \lambda^2) \mathbf{h} - \sqrt{1 - \lambda^2} \mathbf{z}_1 + \lambda Z^T (\lambda H P \hat{\mathbf{s}} - \mathbf{z}_2)$$

where $Z = [J^0 P \hat{\mathbf{s}}, \dots, J^{L-1} P \hat{\mathbf{s}}]$. Thus the MLE is calculated by minimising $\Psi(\mathbf{h})$ subject to the constraint $g(\mathbf{h}) = 0$ where $g(\mathbf{h}) = \frac{1}{2} (\|\hat{\mathbf{s}}\|^2 - 1)$. The gradient vector of $g(\mathbf{h})$ is determined by differentiating the pseudo-inverse; the element $\frac{\partial g}{\partial h_i}$ equals

$$-\hat{\mathbf{s}}^T (HP)^+ \left[J^i P \hat{\mathbf{s}} + \left(J^i P (HP)^+ \right)^T \left(HP \hat{\mathbf{s}} - \frac{1}{\lambda} \mathbf{z}_2 \right) \right]$$

for $i = 0, \dots, L - 1$.

4. SIMULATION STUDY

Two precoders were simulated. The MLE of \mathbf{s} and \mathbf{h} was found by using the Matlab function `constr` to minimise the cost function $\Psi(\mathbf{h})$ subject to the constraint $g(\mathbf{h}) = 0$. In addition, two other estimators were used. One estimated \mathbf{h} from the training sequence only, that is, $\hat{\mathbf{h}} = \frac{\mathbf{z}_1}{\sqrt{1 - \lambda^2}}$. The other estimated \mathbf{h} from the training sequence but under the constraint that $\|\hat{\mathbf{s}}\| = 1$, that is, by minimising $\phi_1(\mathbf{h})$ subject to the constraint $g(\mathbf{h}) = 0$.

Each precoder was simulated 200 times for various values of λ and σ^2 . For each iteration, $p = 12$ source symbols were uniformly randomly generated on the surface of the unit sphere $\|\mathbf{s}\| = 1$. The FIR channel \mathbf{h} of length $L = 3$ was also uniformly randomly generated on the unit sphere $\|\mathbf{h}\| = 1$ subject to the constraint that $|h_0| \geq 0.7$. This constraint ensures that at least half the power is transmitted without delay. As remarked in Section 3.2, if it is believed that h_0 is close to zero then trailing zeros should be appended to the precoder.

Precoder One is the affine precoder (5) where $m = 4$ and each of the four blocks $\mathbf{s}_1, \dots, \mathbf{s}_4$ contain three elements. For comparison, Precoder Two merely adds a training sequence, that is, it is the affine precoder (5) with $m = 1$ and $\mathbf{s}_1 = \mathbf{s}$. Because the encoded vector \mathbf{x} lies on the unit sphere, the signal to noise ratio is $\text{SNR} = -10 \log_{10}(n\sigma^2)$. The following two quantities were estimated by simulation.

$$\text{MSE} = 10 \log_{10} \left(\mathbf{E}[\|\mathbf{s} - \hat{\mathbf{s}}\|^2] \right)$$

$$\text{SD} = \text{standard deviation of } 10 \log_{10} \|\mathbf{s} - \hat{\mathbf{s}}\|^2.$$

The MSE is the mean-square error between the estimated and actual source vector, while SD is a measure of how sensitive the MSE is to different source and channel vectors. These values are represented in Tables 1 to 5 as $\text{MSE} \pm \text{SD}$.

λ	SNR		
	20dB	40dB	60dB
0.1	0.3 ± 2.4	-19.5 ± 2.6	-39.7 ± 2.7
0.5	-12.9 ± 2.7	-33.0 ± 2.6	-53.0 ± 2.7
0.9	-15.9 ± 2.9	-36.4 ± 2.6	-56.2 ± 2.8

Table 1: MSE \pm SD for Precoder One using the training sequence only to estimate the channel.

λ	SNR		
	20dB	40dB	60dB
0.1	-2.4 ± 2.2	-20.0 ± 2.6	-40.1 ± 2.8
0.5	-13.7 ± 2.6	-33.7 ± 2.6	-53.7 ± 2.7
0.9	-17.2 ± 2.8	-37.4 ± 2.7	-58.4 ± 2.9

Table 2: MSE \pm SD for Precoder One using the training sequence only to estimate the channel but subject to the constraint $\|\hat{s}\| = 1$.

λ	SNR		
	20dB	40dB	60dB
0.1	-2.5 ± 2.2	-20.0 ± 2.6	-40.1 ± 2.8
0.5	-13.7 ± 2.6	-33.7 ± 2.6	-53.6 ± 2.7
0.9	-17.6 ± 2.7	-38.0 ± 2.5	-58.8 ± 2.6
0.95	-17.9 ± 2.8	-37.6 ± 2.9	-58.7 ± 2.8
0.99999	-13.3 ± 4.7	-34.4 ± 4.4	-59.9 ± 2.7

Table 3: MSE \pm SD for Precoder One using the full MLE to estimate the channel subject to the constraint $\|\hat{s}\| = 1$.

λ	SNR		
	20dB	40dB	60dB
0.1	5.8 ± 4.4	-14.4 ± 4.4	-34.4 ± 4.3
0.5	-6.7 ± 4.5	-28.1 ± 4.5	-48.9 ± 4.2
0.9	-8.9 ± 5.0	-28.0 ± 5.5	-49.7 ± 4.9

Table 4: MSE \pm SD for Precoder Two using the training sequence only to estimate the channel.

λ	SNR		
	20dB	40dB	60dB
0.1	-2.2 ± 2.1	-17.4 ± 3.5	-35.7 ± 3.9
0.5	-11.6 ± 3.5	-31.0 ± 3.8	-50.4 ± 3.7
0.9	-14.3 ± 4.1	-30.7 ± 5.0	-52.3 ± 4.5

Table 5: MSE \pm SD for Precoder Two using the training sequence only to estimate the channel but subject to the constraint $\|\hat{s}\| = 1$.

Discussion of Results: The simulations verify that a training sequence used in conjunction with a linear precoder, such as Precoder One, outperforms either a training sequence on its own, such as Precoder Two, or a linear precoder on its own. That Precoder One outperforms Precoder Two can be seen by comparing Table 3 to Table 5; not only does Precoder One improve the MSE by on average 3–6dB, it is more robust to different source and channel vectors (as seen by the significantly lower SD values). That Precoder One outperforms a linear precoder can be seen from Table 3 which shows that as $\lambda \rightarrow 1$ the MSE and SD start to increase. This is consistent with the argument that linear precoders without a training sequence are sensitive to different source and channel vectors.

Tables 2 and 3 refer to Precoder One and show that using only the training sequence to identify the channel gives comparable results with the full MLE. This supports the intuition that the training sequence is best used for identifying the channel while the linear precoder is best used for spreading the spectrum to avoid channel spectral nulls.

Comparing Table 1 with Table 2, and Table 4 with Table 5, shows that the constraint $\|\hat{s}\| = 1$ plays an important role in reducing the MSE. The constraint regularises the inversion of the channel, which is especially important when the SNR is low since the channel inversion problem becomes ill-conditioned.

5. CONCLUSION

This paper studied the problem of designing a precoder to facilitate the reliable transmission of a short message through an unknown FIR channel. It was argued that it is efficient to use a training sequence in conjunction with a linear precoder. Simulation results support the claims made throughout this paper.

6. REFERENCES

- [1] G. B. Giannakis. Filterbanks for blind channel identification and equalization. *IEEE Signal Processing Letters*, 4(6):184–187, June 1997.
- [2] T. Li and Z. Ding. Joint transmitter-receiver optimization for partial response channels based on nonmaximally decimated filterbank precoding technique. *IEEE Transactions on Signal Processing*, 47(9):2407–2418, 1999.
- [3] J. H. Manton and Y. Hua. Deterministic and semi-blind channel identification and deconvolution. *Signal Processing*, 1999. Submitted.
- [4] J. H. Manton and Y. Hua. A frequency domain approach to deterministic channel identification. *IEEE Signal Processing Letters*, November 1999. To appear.
- [5] J. H. Manton, W. D. Neumann, P. T. Norbury, and Y. Hua. Algebraic analysis of linear precoding for channel identification. 1999. Preprint.
- [6] A. Scaglione, S. Barbarossa, and G. B. Giannakis. Filterbank transceivers optimizing information rate in block transmissions over dispersive channels. *IEEE Transactions on Information Theory*, 45(3):1019–1032, April 1999.
- [7] C. E. Shannon. Communication in the presence of noise. *Proc. IRE*, 37:10–21, January 1949.