BLIND FRACTIONALLY-SPACED EQUALIZATION OF NOISY FIR CHANNELS: ADAPTIVE AND OPTIMAL SOLUTIONS

Georgios B. Giannakis *Steven D. Halford*

Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903-2442, USA

ABSTRACT

Blind Fractionally Spaced (FS) equalizers only require output samples taken at rates higher than the symbol rate to estimate the channel or the equalizer. Methods for finding FIR zero forcing equalizers directly from the observations are described and adaptive versions are developed. In contrast, most current methods require channel estimation as a first step to estimating the equalizer. For the noisy channel, the FIR equalizer is shown to be minimum mean-square error. FS equalizers are not unique, thereby allowing optimum zero-forcing parametric FIR or nonparametric IIR equalizers to be derived such that in addition to being zero-forcing, they also minimize the noise power at the output. These optimum equalizers do not depend on the input distribution and are also valid for deterministic inputs. Finally, if the additive noise is white, they do not depend on the SNR.

1. INTRODUCTION

Blind Equalizers remove the distortion caused by intersymbol interference (ISI) without the need for training sequences. In order to capture the complete phase characteristics of the ISI, Higher- (than second-) Order Statistics (HOS) must be used to find the equalizer taps [7]. However, [12] showed that most channels can be identified from the second-order cyclostationary statistics generated by fractionally-spaced (FS) equalizers. In addition, [11] pointed out that for FIR channels, unlike symbol rate equalizers, exact zero-forcing FIR equalizers can be found in the FS case. Hence, cyclostationarity offers the advantages of reliable estimates from smaller data records (relative to HOS methods) as well as the possibility of exact zero-forcing of non-minimum phase FIR channels with FIR equalizers.

These properties have motivated considerable research interest in fractionally-spaced channel identification and equalization. The discrete-time equivalent model for the "overall" fractionally-spaced system is seen in Figure 1, where the oversampling rate P implies there are P samples

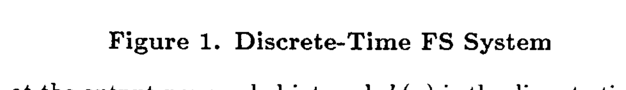

Figure 1. Discrete-Time FS System

at the output per symbol interval, $h(n)$ is the discrete-time equivalent FIR channel of order QP, and the noisy observations are $y(n)$. The symbols $w(n)$ are estimated from $y(n)$ by subsampling the output of $g(n)$. Mathematically, the input-output relations are:

$$y(n) = \sum_m w(m)h(n-mP) + v(n) = x(n) + v(n), \quad (1)$$

and the FS equalized symbols

$$\hat{w}(n) = \sum_m g(m)y(nP-m). \quad (2)$$

The observations $y(n)$ are cyclostationary – that is, the correlation $c_{2y}(n;m) := \mathrm{E}\{y(n)y^*(n+m)\}$ is periodically time-varying with period P. Assuming, as usual, the symbol stream $w(n)$ is i.i.d. (with zero mean and variance $c_{2w}(0)$) and independent of $v(n)$, it follows from (1):

$$c_{2y}(n;m) = c_{2w}(0)\sum_\ell h(n-\ell P)h^*(n+m-\ell P) + c_{2v}(n;m). \quad (3)$$

Impulse response $h(n)$ will be identifiable from the $c_{2y}(n;m)$ provided its transfer function $H(z)$ satisfies the following identifiability condition [12] (see also [3]).

ID condition: *There are no zeros of $H(z)$ equispaced on a circle with angle $\frac{2\pi}{P}$ separating one zero from the next.*

Most current methods treat the scalar cyclostationary FS system of Figure 1 as a stationary multichannel system. This leads to vectorization of the output and ESPRIT [12], MUSIC [10], multivariate linear prediction methods [11], and other eigen-based solutions [9] have been developed to identify $h(n)$. As an alternative to vectorizing $y(n)$, the periodic (in n) sequence $c_{2y}(n;m)$ can be expanded in Fourier Series. The k^{th} Fourier Series coefficient, known as the cyclic correlation at cycle k, is given by (c.f. (3)):

$$C_{2y}(k;m) = \frac{1}{P}\sum_{n=0}^{P-1} c_{2y}(n;m)e^{-j\frac{2\pi}{P}kn} \quad (4)$$
$$= \frac{c_{2w}(0)}{P}\sum_\ell h(\ell)h^*(\ell+m)e^{-j\frac{2\pi}{P}k\ell} + c_{2v}(m)\delta(k)$$

where the Kronecker delta $\delta(k)$ is present because the noise $v(n)$ is assumed stationary. Eq. (4) leads to simple linear and eigen-based channel identification algorithms [4]. In addition, performance analysis is possible with the scalar cyclostationary approach and optimal linear [4] and nonlinear [5] channel identification algorithms have been developed.

Once $h(n)$ is estimated, equalization can be performed by adopting an appropriate performance measure. For example, the maximum likelihood approach would use $h(n)$ in a Viterbi algorithm to estimate $w(n)$. If on the other hand, a

0-7803-2431-5/95 $4.00 © 1995 IEEE

linear equalizer $g(n)$ is desired, a least-squares symbol rate solution is possible. An alternative performance criterion is to require the equalizer to be zero-forcing (ZF) or, in other words, to require the transfer function from $w(n)$ to $\hat{w}(n)$ to be the identity system. If the ZF criterion is met, the ISI is completely removed. For the system in Figure 1, the **ZF condition** is:

$$\sum_m g(m)h(nP-m) = \delta(n), \quad (5)$$

or in the frequency domain:

$$\frac{1}{P}\sum_{k=0}^{P-1} G\left(\frac{\omega-2k\pi}{P}\right)H\left(\frac{\omega-2k\pi}{P}\right) = 1. \quad (6)$$

If the ID condition is satisfied for $h(n)$, then the matrix of linear equations resulting from (5) for different n's can be shown to be full rank and hence solutions for $g(n)$ exist. Equivalent conditions for the multichannel case and a procedure for estimating $g(n)$ from the observations are given in [11] (see also [8]).

One drawback of estimating the channel and using (5), or, the direct procedure of [11] is the inability to track time-variations in the channel. For this common practical situation, feasible adaptive implementations have not been studied thoroughly. In addition, in both the time-varying and the time-invariant case, the same overdeterminancy which allows the phase to be determined from second-order statistics, will allow a multitude of solutions for (5).

In this paper, we present a recursive procedure for finding the ZF equalizer directly from the observed samples. This will lead to the adaptive algorithms described in Section 2. More importantly, we will show that it is possible to find equalizers that are both exactly ZF and at the same time minimize the variance of the noise at the output. The derivation of these optimum equalizers is given in Section 3. They can be parametric FIR or nonparametric IIR. In addition, these optimum equalizers do not require knowledge of the input. They are solved independent of the input distribution if the input is random and are also valid when the input is considered as deterministic.

2. BLIND ZF EQUALIZERS

For a recursive method to directly find the $\{g(n)\}_{n=0}^{L_g}$, we need to express (5) in terms of the data correlations. Consider the following noiseless (i.e., $v(n) = 0 \ \forall \ n$ and $x(n) = y(n)$) time-varying correlation from (3):

$$c_{2x}(-\ell;\ell+m) = c_{2w}(0)\sum_i h(-\ell-iP)h^*(m-iP). \quad (7)$$

Multiplying both sides by $g(\ell)$ and summing over ℓ yields:

$$\sum_\ell g(\ell)c_{2x}(-\ell;\ell+m) = c_{2w}(0)\sum_i h^*(m-iP)\sum_\ell g(\ell)h(-\ell-iP). \quad (8)$$

The last sum in (8) equals $\delta(-i)$ due to (5); hence,

$$\sum_\ell g(\ell)c_{2x}(-\ell;\ell+m) = c_{2w}(0)h^*(m). \quad (9)$$

Eq. (9) can also be expressed in the frequency domain as:

$$\sum_{k=0}^{P-1} G^*\left(\omega-\frac{2\pi k}{P}\right)S_{2x}(k;\omega) = \frac{c_{2w}(0)}{Q}H(\omega), \quad (10)$$

where $G(\omega)$ and $H(\omega)$ are the Fourier transforms of $g(n)$ and $h(n)$ respectively, and $S_{2x}(k;\omega)$ is the cyclic spectrum (i.e., the Fourier transform of $C_{2x}(k;m)$ w.r.t. m).

Let us assume (w.l.o.g.) that $h(n)$ is causal and $h(0)$ is known. Concatenating equations from (9) with $m \le 0$ we obtain a matrix equation $\mathbf{C}\mathbf{g} = \mathbf{e}_h$. If $h(n)$ satisfies the ID condition, $\mathbf{C}$ is full rank and thus solving for $\mathbf{g}$ the aforementioned matrix equation implies $\{g(n)\}_{n=0}^{L_g}$ can be found uniquely using a blind algorithm without estimating $\{h(n)\}_{n=0}^{QP}$. In practice, we estimate $\{g(n)\}_{n=0}^{L_g}$ in a batch method by replacing the ensemble correlations with their sample estimates, $\hat{c}_{2x}(n;m)$, computed from the observations as we show next.

2.1. Recursive and adaptive methods

To compute the time-varying correlations recursively, we adopt the following (normalized) sample estimator:

$$\begin{aligned}\hat{c}_T(n;m) &:= \sum_{k=0}^{T-1} x(n+kP)x^*(n+m+kP) \\ &\quad + x(n+TP)x^*(n+m+TP) \\ &= \hat{c}_{T-1}(n;m) \\ &\quad + x(n+TP)x^*(n+m+TP) \end{aligned} \quad (11)$$

where $\hat{c}_T(n;m)$ and $\hat{c}_{T-1}(n;m)$ are the sample estimates obtained at time TP and $(T-1)P$ respectively. In order to track time-variations, a standard "forgetting" factor λ ($0 < \lambda \le 1$) is also included:

$$\hat{c}_T(n;m) = \lambda\hat{c}_{T-1}(n;m) + x(n+TP)x^*(n+m+TP). \quad (12)$$

The zero-forcing equalizer at time TP is in vector form:

$$\hat{\mathbf{g}}_T = \hat{\mathbf{C}}_T^{-1}\mathbf{e}_h, \quad (13)$$

where $\hat{\mathbf{C}}_T$ is the appropriate $L_g \times L_g$ matrix of sample TV correlations calculated from (12), $\hat{\mathbf{g}}_T$ is the $L_g \times 1$ vector of unknown taps, and $\mathbf{e}_h$ is a vector of constants. Similarly, vectorizing (12) and substituting into (13) yields

$$\hat{\mathbf{g}}_T = (\lambda\hat{\mathbf{C}}_{T-1} + \mathbf{x}_n\mathbf{x}_{n+m}^{'})^{-1}\mathbf{e}_h, \quad (14)$$

where the prime indicates transpose. As in standard recursive least-squares (RLS), the well-known matrix inversion lemma is used to simplify (14) into:

$$\hat{\mathbf{g}}_T = \left[\lambda\hat{\mathbf{C}}_{T-1}^{-1} - \frac{1}{\lambda^2}\frac{\hat{\mathbf{C}}_{T-1}^{-1}\mathbf{x}_n\mathbf{x}_{n+m}^{'}\hat{\mathbf{C}}_{T-1}^{-1}}{\mathbf{x}_{n+m}^{'}\hat{\mathbf{C}}_{T-1}^{-1}\mathbf{x}_n + 1}\right]\mathbf{e}_h. \quad (15)$$

where the term inside the brackets denotes $\mathbf{C}_T^{-1}$.

Since (15) requires only a scalar division, it is computationally feasible to use (15) to equalize a TV channel. Or, if $\lambda = 1$ is chosen, (15) provides a method to recursively compute the equalizer taps. Using the ergodicity results of [2], it can also be shown that the $\hat{\mathbf{g}}_T$ estimator is mean-square sense consistent as $T \to \infty$, at least when the true channel is time-invariant. In addition to the RLS solution, simpler LMS alternatives appear to be feasible and will be reported in the future.

Both (9) and (14) (provided $\lambda = 1$) yield an equalizer which is FIR and is exactly ZF. However, this is only true when there is no noise (i.e., when $v(n) = 0$). In the presence of noise, $c_{2x}(n;m)$ is replaced with $c_{2y}(n;m)$ and the equalizer found from (9) will not satisfy the ZF condition of (5). A natural question which then arises is, if the equalizer is not ZF in the presence of noise, is it optimal in any sense? And, are there other solutions which are exactly ZF in the presence of noise or have other optimality? As will be seen, the FIR Wiener filter (i.e., the optimal linear predictor of $w(n)$ based on $y(n)$) will yield the same set of equations as (9).

2.2. FS Wiener equalizers

Our goal is to find the $\{g(n)\}_{n=0}^{L_g}$ such that $E\{|\hat{w}(n) - w(n)|^2\}$ is minimized. As usual, we substitute $\hat{w}(n)$ from (2), take the complex derivative of $E\{|\hat{w}(n) - w(n)|^2\}$ w.r.t. the equalizer coefficients and set these equations to zero:

$$\frac{\partial}{\partial g^*(m)} E\{|\sum_{\ell} g(\ell)y(nP-\ell) - w(n)|^2\} = 0. \qquad (16)$$

After manipulation and upon using the facts that $w(n)$ is i.i.d. and independent of $v(n)$, (16) yields:

$$\sum_{\ell}\sum_{i} E\{y(nP-\ell)y^*(nP-i)\}g(\ell)\delta(i-m)$$
$$-\sum_{i} E\{y^*(nP-i)w(n)\}\delta(i-m) = 0.$$

Since the cross-correlation of $y(\cdot)$ is periodic with period P, and using (1), we infer:

$$\sum_{\ell} g(\ell)c_{2y}(-\ell;\ell+m) = c_{2w}(0)h^*(m), \qquad (17)$$

which is equal to (9) when $y(\cdot)$ is used in place of $x(\cdot)$. Since in practice, $y(\cdot)$ must be used, this blind equalizer will not be zero ZF, but rather minimum mean square error (MMSE). Since MMSE may not be the best criterion for equalization, we are motivated to examine the properties of "other" solutions satisfying (5). As alluded to before, the ZF equalizer which satisfies (5) is not unique. As an example, for a given $H(\omega)$, if $G(\omega)$ satisfies (6), it follows that

$$\tilde{G}(\omega) = G(\omega) - F(\omega)\prod_{\ell=1}^{P-1} H(\omega - \frac{2\pi\ell}{P}), \qquad (18)$$

also satisfies (6), provided that

$$\sum_{k=0}^{P-1} F(\omega - \frac{2\pi k}{P}) = 0. \qquad (19)$$

Using the substitution $\omega \rightarrow \omega P$, we find that (19) is satisfied by any FIR or IIR filter $\tilde{f}(n)$ if we downsample its impulse response by P; i.e., $f(n) = \tilde{f}(nP)$ satisfies (19). For example, an IIR filter which satisfies (6) is:

$$G(\omega) = \frac{P\ H^*(\omega)}{\sum_{k=0}^{P-1} |H(\omega - \frac{2\pi k}{P})|^2}\ . \qquad (20)$$

Motivated by the multitude of ZF solutions, it is natural to look for the one(s) which not only satisfy ZF (5), but also optimally suppress the noise.

3. OPTIMAL ZERO FORCING EQUALIZERS

First, we consider FIR-FSE's. Define $\mathbf{1} := [1\ 0 \ldots 0]'$ and let $\mathcal{T}(\mathbf{h}_i)$ denote the Toeplitz matrix whose first column is formed by the vector $\mathbf{h}_i$. If $\mathbf{h}_i$ denotes the i^{th} subchannel of $h(n)$, $\mathbf{h}_i := [h(i)h(P+i)\ldots h(QP-P+i)]$, and $\mathbf{g} = [\mathbf{g}_0'\ \mathbf{g}_1' \cdots \mathbf{g}_{P-1}']'$, the ZF condition (5) can be written as:

$$\mathcal{T}(\mathbf{h}_{0:P-1})\ \mathbf{g} = \mathbf{1}, \qquad (21)$$

where $\mathcal{T}(\mathbf{h}_{0:P-1})$ denotes a block Toeplitz formed from the $Q \times 1$ vectors $\mathbf{h}_0, \mathbf{h}_1, \ldots\ \mathbf{h}_{P-1}$ as $\mathcal{T}(\mathbf{h}_{0:P-1}) := [\mathcal{T}(\mathbf{h}_0)|\mathcal{T}(\mathbf{h}_1)|\ldots|\mathcal{T}(\mathbf{h}_{P-1})]$. Among the FSEs which satisfy (21), we will select the one which minimizes the noise contribution at the output of the downsampler which follows the equalizer (see Figure 1). The noise contribution, $\epsilon(n)$, is:

$$\epsilon(n) = \sum_{m} g(m)v(nP-m), \qquad (22)$$

and its variance is given by:

$$\sigma_\epsilon^2 = \mathbf{g}^{*'}\mathbf{R}_{vv}\mathbf{g}, \qquad (23)$$

where $\mathbf{R}_{vv}$ is the noise covariance matrix and $*'$ denotes conjugate transpose. Hence, the optimal FIR ZF equalizer is:

$$\mathbf{g}_o = \arg\min\{\mathbf{g}^{*'}\mathbf{R}_{vv}\mathbf{g}\}$$

subject to $\mathcal{T}(\mathbf{h}_{0:P-1})\mathbf{g} = \mathbf{1}$. This constrained minimization problem can be solved using the Lagrange multipliers for complex vector unknowns (see e.g., [6, Appendix C]) and yields:

$$\mathbf{g}_o = \mathbf{R}_{vv}^{-1}\mathcal{T}^{*'}(\mathbf{h}_{0:p-1})[\mathcal{T}(\mathbf{h}_{0:p-1})\mathbf{R}_{vv}^{-1}\mathcal{T}(\mathbf{h}_{0:p-1})]^{-1}\mathbf{1}. \quad (24)$$

In the white noise case, (24) simplifies to:

$$\mathbf{g}_o = \mathcal{T}^{\dagger}(\mathbf{h}_{0:p-1})\mathbf{1}, \qquad (25)$$

where $\dagger$ denotes pseudoinverse. Note that knowledge of σ_v^2 is not required in (25). In comparison with the Wiener filter seen previously, the $\mathbf{g}_o$ of (24) guarantees ZF yet minimizes σ_ϵ^2. This is the same as minimizing $E\{|\hat{w}(n) - w(n)|^2\}$ if $w(n)$ is treated as deterministic. This is evident since $\mathbf{g}_o$ does not, unlike the Wiener filter, depend on $c_{2w}(0)$.

Depending on the noise color, if we do not limit our search to FIR g's, we may obtain better noise suppression. Restated, we want to search for the ZF equalizer without regard to length, and find the one with the optimum noise suppression. If we cast (5) and (23) into the frequency domain, then again this search becomes a constrained minimization problem for every frequency ω. The solution in the frequency domain can be shown to be:

$$G_o(\omega) = \frac{PH^*(\omega)/S_{vv}(\omega)}{\sum_{k=0}^{P-1} |H(\omega - \frac{2\pi k}{P}|^2/S_{vv}(\omega - \frac{2\pi k}{P})}, \qquad (26)$$

where $S_{vv}(\omega)$ is the power spectral density of the noise.

From IIR $G_o(\omega)$ we can see how fast $g_o(n)$ decays to zero, and then use this to design the FIR $\mathbf{g}_o$ of appropriate length. It should also be emphasized that neither (24) nor (26) require information about the input. In the case of colored input, the optimum equalizers do not need to know the coloring. Also, both (24) and (26) are valid for deterministic input. This provides a nice complementary result to the deterministic channel estimator of [9] and the associated equalizer described in [8].

For multidimensional processes of continuous support a result similar to our nonparametric IIR equalizer has been reported independently in [1] for characterization of imaging sensors. The link can be established after casting [1] in discrete-time and taking into account the univariate-cyclostationary to multivariate-stationary equivalence.

Another interesting link appears to exist between our optimal FIR equalizer and Capon's minimum variance distortionless response beamformer (e.g., [6, pp. 550-558]). This research direction will be explored in the future.

4. SIMULATIONS

In this section we present simulation results for the zero-forcing condition (12) and its adaptive version (14). Figure 2A shows a nonminimum phase channel of order 4 ($Q = 2$, $P = 2$) and the solid line in Figure 2B shows the magnitude of the channel's frequency response. The dotted line in Figure 2B shows the frequency response after equalization. The equalizer taps were computed from (12) using $N = 200$ data to compute the statistics. The SNR was 30 dB and the symbols were chosen i.i.d. from a BPSK alphabet. Figure 3A shows the received symbols before equalization while Figure 3B shows the same symbols after equalization. To evaluate the adaptive procedure, the channel in Figure 2A was modified by allowing $h(4)$ to change linearly in time. The resulting "ideal" ZF coefficients are calculated from (5) and are shown as a function of time by Figure 3. The dashed lines represent the mean of the estimated ZF equalizer coefficients using (14) and the dashed-dotted lines represent the standard deviation. The mean and standard deviation was computed over 100 Monte Carlo runs and the SNR was again 30 dB.

Acknowledgments: The authors wish to thank Dr. M. K. Tsatsanis for exchanging ideas on this and related topics. Also we thank Dr. M. Unser from NIH for bringing (during ICIP'95) to our attention the work of [1]. This work was supported by ONR grant N0014-93-1-0485.

REFERENCES

[1] C. A. Berenstein and E. V. Patrick, "Exact deconvolution for multiple convolution operators - An overview plus performance characterization for imaging sensors," *Proc. of the IEEE*, pp. 723-734, April 1990.

[2] A. V. Dandawaté and G. B. Giannakis, "Asymptotic theory of mixed time averages and k^{th}-order cyclic-moment and cumulant statistics," *IEEE Trans. on IT*, Jan. 1995.

[3] Z. Ding and Y. Li, "On channel identification based on second-order cyclic spectra," *IEEE Trans. on Signal Processing*, pp. 1260-1264, May 1994.

[4] G. B. Giannakis, "Linear cyclic correlation approaches for blind identification of FIR channels," *Proc. of 28th Annual Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, October 31 - November 2, 1994.

[5] S. Halford and G. B. Giannakis, "Asymptotically optimal blind equalizers based on cyclostationary statistics," *Proc. IEEE Military Com. Conf.*, pp. 306-310, Fort Monmouth, NJ, October 2-5, 1994.

[6] S. Haykin, *Adaptive Filter Theory*, 2nd Edition, Prentice Hall, Englewood Cliffs, NJ, 1991.

[7] D. Hatzinakos, and C. L. Nikias, "Blind Equalization Using a Tricepstrum Based Algorithm," *IEEE Trans. on Communications*, vol. 38, pp.669-682, May 1991.

[8] H. Liu and G. Xu, "A deterministic approach to blind symbol estimation," *Proc. of 28th Annual Asilomar Conf. on Signals, Systems, and Computers*, 1994.

[9] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind identification of multichannel FIR systems," *Proc. of ICASSP*, vol. IV, pp. 581-584, Adelaide, Australia 1994.

[10] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *Proc. of Intl. Conf. on ASSP*, vol. IV, pp. 573-576, Adelaide, Australia, 1994.

[11] D. T. M. Slock, "Blind fractionally-spaced equalization based on cyclostationarity and second-order statistics", *Proc. of ATHOS Workshop on System Identification and Higher Order Statistics*, France, September 1993.

[12] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. on IT*, pp. 340-349, 1994.

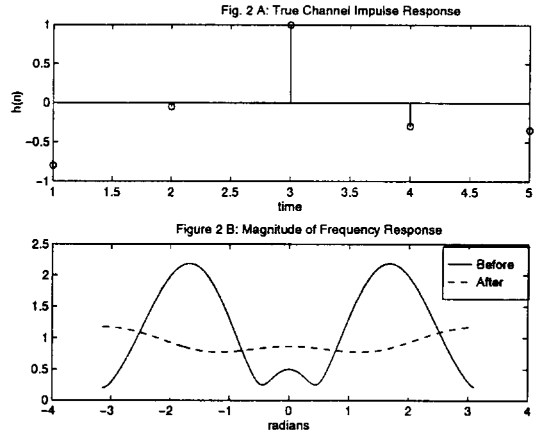

Figure 2. Time & Freq. Domain for SNR = 30 dB

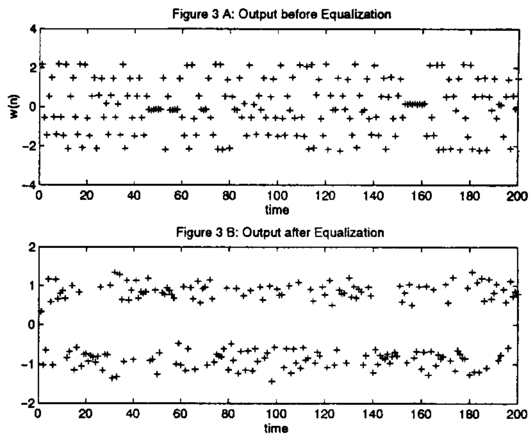

Figure 3. Rec'd & Equalized Symbs.:SNR=30 dB

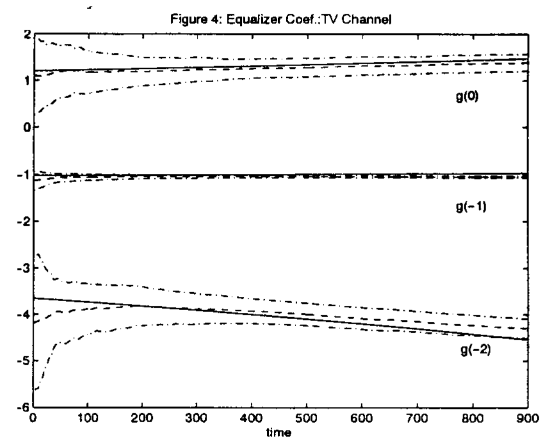

Figure 4. TV Equal. Taps for SNR = 30 dB