Per-Survivor Processing for Channel Acquisition, Data Detection and Modulation Classification

Norman E. Lay

Communications Sciences Institute
University of Southern California
Los Angeles, CA 90089-2565
e-mail: lay@amtsun.jpl.nasa.gov

Andreas Polydoros

Communications Sciences Institute
University of Southern California
Los Angeles, CA 90089-2565
e-mail: andreas@solar.usc.edu

Abstract

Per-survivor processing (PSP) is a technique for jointly estimating the data sequence and unknown parameters of a communications signal which exhibits memory. PSP utilizes a trellis to retain and update distinct data sequence and channel estimates. In this paper, we apply PSP to signals affected by ISI due to a FIR channel. The unknown channel acquisition performance of PSP is compared to that obtained from two blind equalization techniques. PSP may also be used to perform modulation classification of signals in ISI environments. Independent trellises matched to the modulations are created and metrics from each trellis are compared to produce a classification decision.

1 Introduction

Per-Survivor Processing is a joint channel and data estimation technique [1] which has demonstrated significant performance gains in various tracking applications in which a received signal has been corrupted with memory, when compared to conventional adaptive sequence estimation. Unlike blind equalization, which seeks to recover an undistorted (memoryless) data signal through the use of an adaptive inverse channel filter [2] [3], PSP exploits the channel memory within its receiver structure. Multiple candidate data sequences are retained according to a selection rule (e.g. the Viterbi algorithm) and individual channel estimates are produced based upon these sequences. A true maximum likelihood implementation of PSP would require the retention of *all* possible data sequences and a least squares channel estimate corresponding to each of those sequences [1] [4]. Even Viterbi based full-state implementations of PSP require an exponential increase in processing as a function of the channel memory. Alternatively, blind equalization approaches typically demand a much lower computational burden which is independent of the channel length. The objective of this paper is to provide a direct performance comparison between these techniques when considering the acquisition of a static unknown FIR channel. The application of PSP to modulation classification is also introduced to demonstrate the versatility of PSP.

2 Channel Model

The receiver observations, r_k, for a discrete time finite impulse response (FIR) channel model are given by the following equation.

$$r_k = \sum_{l=0}^{L-1} h_l \cdot a_{k-l} + n_k \tag{1}$$

The channel, of length L, has coefficients specified by the h_l's. The transmitted symbols are the a_k's, and the n_k's represent a complex AWGN sequence. Table 1 details the impulse responses of the channels examined in this paper. Channel B exhibits a spectral null.

Channel Name	*Impulse Response*
Channel A [4]	[-0.205 -0.513 0.718 0.369 0.205]
Channel B [4]	[0.407 0.815 0.407]
Channel C [3]	[0.348 0.870 0.348]

Table 1: Channel Impulse Responses

1058-6393/95 $4.00 © 1995 IEEE

3 Known Channel Results

Prior to comparing the performance of PSP to adaptive linear equalization for unknown FIR channels, we first compare the two techniques on the basis of their known channel counterparts. For known channel equalization, we derive the optimal mean square error coefficients by using the Wiener Equations. These equalizer taps are calculated as follows.

$$\mathbf{W}_{opt} = \mathbf{R}_{\mathbf{ar}} \cdot \mathbf{R}^{-1} \tag{2}$$

$$= E\left[\mathbf{a} \cdot \mathbf{r}^H\right] \cdot \left(\mathbf{H} \cdot E\left[\mathbf{a} \cdot \mathbf{a}^H\right] \cdot \mathbf{H}^H + \sigma_n^2\, \mathbf{I}\right)^{-1}$$

The rows of $\mathbf{H}$ consist of delayed versions of the FIR channel sequence, h_k, $\mathbf{r}$ represents the receiver observations, $[r_1\ r_2 \ldots r_M]^T$, and $\mathbf{a}$ represents the data sequence, $[a_k\ a_{k-1} \ldots a_{k-L-M+1}]^T$, where M is the length of the equalizer. For the optimal weighting matrix, $\mathbf{W}_{opt}$, the rows represent different realizations of the MMSE weights for estimating the transmitted data with different delays. The row corresponding to the minimum residual ISI is selected from this matrix as the "best" set of MMSE weights. A plot of the optimum achievable residual ISI for channel A is shown in the next figure for SNR's ranging from 0 to 30 dB, and for equalizer lengths up to 80 taps. On the basis

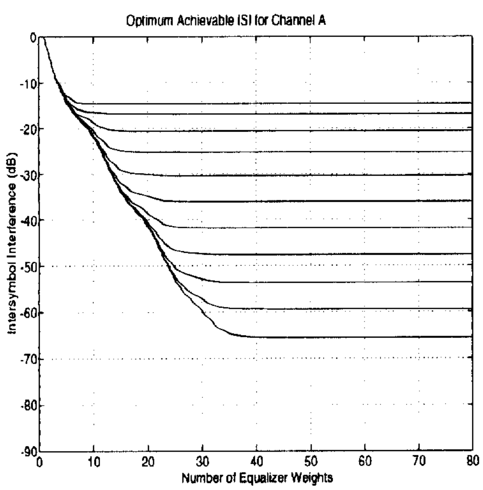

Figure 1: Equalizer Length vs. ISI, Channel A

of this plot an equalizer length of 31 taps is selected as sufficient for achieving the best reduction in ISI over a wide range of SNR's. Comparable selections made for channels B and C resulted in equalizer lengths 15 taps for each.

The known channel performance of the equalizers is then simulated by using the Wiener coefficients to invert the channel followed by a standard AWGN correlation detector at the equalizer output to obtain the symbol error rate. The known channel trellis based performance is obtained via maximum likelihood sequence estimation (MLSE) by using the Viterbi algorithm [5]. A comparison of the QPSK error rate performances for the three channels is shown in Figure 2. For the easily invertible channel A, the two techniques yield comparable results. However, when a spectral null channel is considered, the equalizer degrades significantly due to large residual ISI.

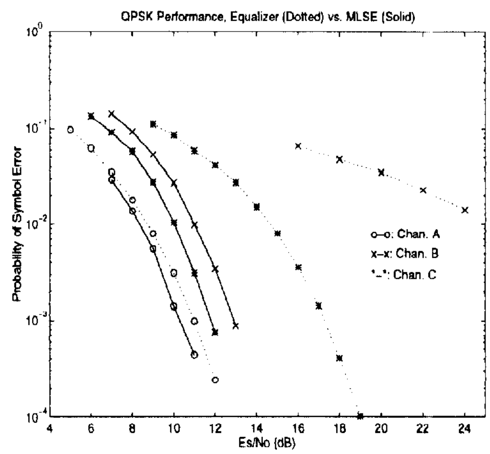

Figure 2: Known Channel QPSK Detection

4 Acquisition Performance

We select two blind equalization techniques for comparison to PSP. All three (the two blind equalizers plus PSP) are tested for speed of acquisition and overall symbol error rate performance. The first of these two is an orthogonalized form [6] of the constant modulus algorithm [7] wherein the weight update is given in (3) and will be referred to as the OCMA algorithm. This algorithm premultiplies the received data sequence by an estimate of the covariance matrix inverse to speed convergence time. An LMS update, with step size μ, is used to adapt equalizer coefficients.

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \mu \hat{\mathbf{R}}_k^{-1} \mathbf{r}_k \epsilon_{k-1}^* \tag{3}$$

$$where\ \ \epsilon_n^* = \frac{z_n \overline{||a_i||^p}}{||z_n||} - z_n, \quad \mathbf{R}_n = E[\mathbf{r}_n \cdot \mathbf{r}_n^H],$$

$$and\ \ z_n = \mathbf{w}_n^H \cdot \mathbf{r}_n$$

The second algorithm is the "unconstrained" version taken from [8] and its weight update equation is shown in (4). The term, $K(a)$, represents the kurtosis of the transmitted symbols, and γ_1, γ_2, and α are additional constants determined by the statistics of the source constellation. The algorithm adaptation step sizes are given by δ, δ_ϵ, and δ_p. This algorithm is subsequently referred to as the SW blind equalizer.

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \delta \cdot sgn\ K(a) \cdot [(|z_k|^2 \qquad (4)$$
$$+\gamma_1 \langle |z|^2 \rangle_k + \gamma_2) z_k - \langle z^2 \rangle_k z_k^*] \cdot \mathbf{r}_k^H$$
$$where\ \ \langle z^2 \rangle_k = (1-\delta_\epsilon)\langle z^2 \rangle_{k-1} + \delta_\epsilon z_k^2,$$
$$and\ \ \langle |z|^2 \rangle_k = (1-\delta_p)\langle |z|^2 \rangle_{k-1} + \delta_p |z_k|^2$$

PSP differs from blind equalization in that it does not attempt to produce a channel inverse to return the received signal to its undistorted form. Rather, it directly estimates the coefficients of the FIR channel on a per surviving data sequence basis. For a specific information sequence, $\mathbf{a}_{k+1}$, of length equal to the channel's, the channel estimate is updated in an LMS adaptive fashion (5) with the adaptive step size set by the constant μ [1] [4]. The surviving data sequences are updated with the Viterbi algorithm using the per path channel estimates.

$$\hat{\mathbf{h}}_{k+1} = \hat{\mathbf{h}}_k + \mu \cdot \epsilon_{k+1} \cdot \mathbf{a}_{k+1}^* \qquad (5)$$
$$where\ \ \epsilon_{k+1} = r_k - \hat{\mathbf{h}}_k^T \cdot \mathbf{a}_{k+1}$$

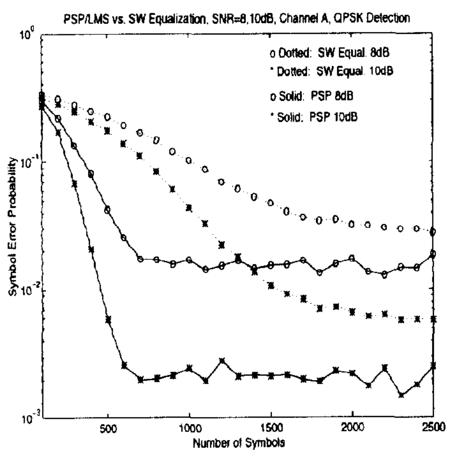

Figure 3: Unknown Channel A, PSP/MLSE vs. SW Blind Equalizer, QPSK Detection Performance

Figure 3 is a representative examples of the QPSK acquisition performance measured in symbol error rate as a function of observed symbols. Error rates are measured in 100 symbol block averages which are ensemble averaged over multiple simulation runs. An effort was made to optimize adaptive step size for both PSP and blind equalization for speed of convergence over the range of tested SNR's.

Channel	**OCMA**	**SW**	**PSP**
Channel A	1000	1700	700
Channel B	N/A	N/A	500
Channel C	1000-2000	1500-2500	400

Table 2: QPSK Modulation, Channel Acquisition Times (Symbols)

The acquisition performance of all three algorithms is summarized in Table 2 which indicates the approximate number of symbols required by each technique before the detected symbol error rate stabilizes to a constant value for a given E_s/N_0. For the spectral null channel, B, both blind equalizers are unable to adaptively equalize the received signal within the constraints of the tested parameters. The wide range in acquisition performance of the two equalizers for channel C is due to different acquisition times at different SNR's.

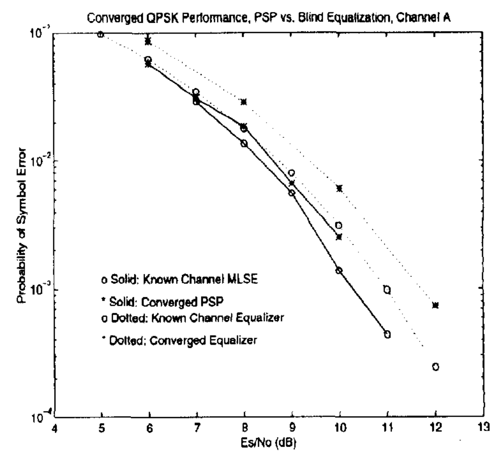

Figure 4: Converged & Known Channel QPSK Detection Performance, Channel A, PSP vs. SW

Figures 4-6 detail comparisons of the converged error rate performance between the blind equalizers and PSP. For both channels A and C, the converged performance of each technique is very close to the known channel values. PSP performance is virtually coincident with those curves, while equalizer detection is within 1 dB. The advantage of known channel MLSE over equalization is maintained in the blind adaptive case.

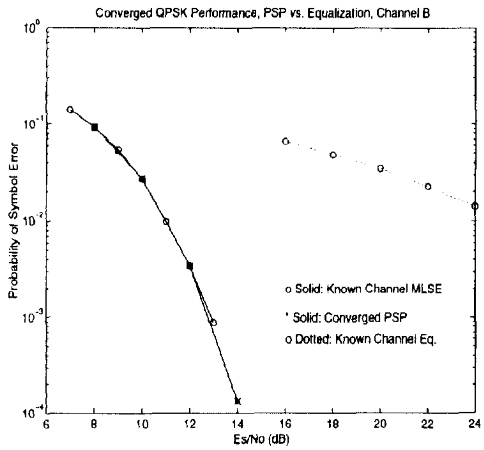

Figure 5: Converged & Known Channel QPSK Detection Performance, Channel B, PSP only

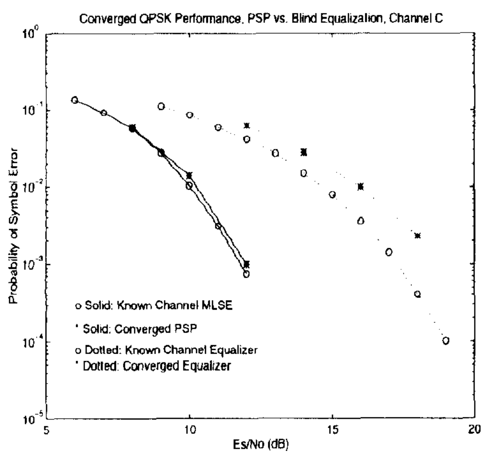

Figure 6: Converged & Known Channel QPSK Detection Performance, Channel C, PSP vs. OCMA

5 Modulation Classification

By virtue of its channel acquisiton properties, PSP is also examined for its suitability in performing modulation classification in ISI environments. Again, prior to treating the adaptive case, we first discuss the construction of known channel likelihood based classifcation tests. For a received sequence, $\mathbf{r}$, and two possible candidate modulations, C_0 and C_1, the composite hypothesis [9] likelihood ratio test is given by (6) and (7). Since the actual transmitted data, $\mathbf{a}$, are unknown, an average over all possible data sequences is required for the average likelihood ratio test (ALRT).

$$\Lambda(\mathbf{r}) = \frac{pr(\mathbf{r}|C_0)}{pr(\mathbf{r}|C_1)} \underset{H_1}{\overset{H_0}{\gtrless}} \eta \equiv \frac{P(C_1)}{P(C_0)} \quad (6)$$

$$pr(\mathbf{r}|C_i) = \int_{\mathbf{a}} pr(\mathbf{r}|\mathbf{a}, C_i) p(\mathbf{a}|C_i) d\mathbf{a} \quad (7)$$
$$= \langle pr(\mathbf{r}|\mathbf{a}, C_i) p(\mathbf{a}|C_i) \rangle_{\mathbf{a} \in C_i}$$

A different decision statistic based upon the generalized likelihood ratio test (GLRT) is obtained by replacing unknown parameters in the likelihood ratio by their maximumlikelihood estimates [9]. In the classification application, the data are initially unknown but can be obtained using the Viterbi algorithm to produce ML data estimates.

$$\Lambda_g(\mathbf{r}) = \frac{pr(\mathbf{r}|\hat{\mathbf{a}}_{ML(C_0)}, C_0)}{pr(\mathbf{r}|\hat{\mathbf{a}}_{ML(C_1)}, C_1)} \underset{C_1}{\overset{C_0}{\gtrless}} \eta \quad (8)$$

$$where,\ pr(\mathbf{r}|\hat{\mathbf{a}}_{ML(C_i)}, C_i) = \underset{\mathbf{a} \in C_i}{max}\ pr(\mathbf{r}|\mathbf{a}, C_i)$$

$$\hat{\mathbf{a}}_{ML(C_i)} \equiv MLSE\ of\ Constellation\ C_i$$

An efficient means of calculating the average likelihood function of signals affected with memory has been applied to the detection of CPM signals [10] and to the estimation of parameters of multiple direct sequence spread spectrum signals [11]. We apply this technique to memory induced by an FIR channel for our ALRT classification test. The known FIR channel consists of two identical taps, $h_0 = h_1 = .707$. The ALRT decision threshold is set by the *a priori* probabilities of the two modulations. Alternatively, the GLRT decision statistic is obtained by comparing the accumulated state metrics corresponding to the "best" state for two different constellation trellises implementing the Viterbi algorithm. An appropriate threshold is obtained by comparing histograms of the decision statistics conditioned upon the transmitted modulations.

The classification performance of the ALRT and the GLRT for two 16-ary modulations is shown in Figure 7. A performance improvement of roughly one dB is observed in the ALRT over the GLRT, for 100, 200 and 500 observed symbols. Knowledge of the noise variance and signal power is required for the ALRT but not for the GLRT.

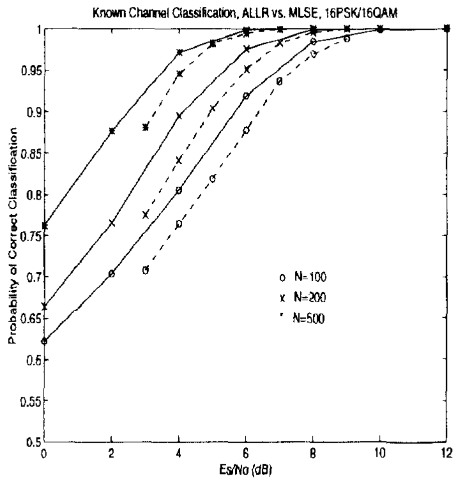

Figure 7: Known Channel ALRT vs. GLRT Classification Tests, 16PSK/16QAM, 2 Tap Channel

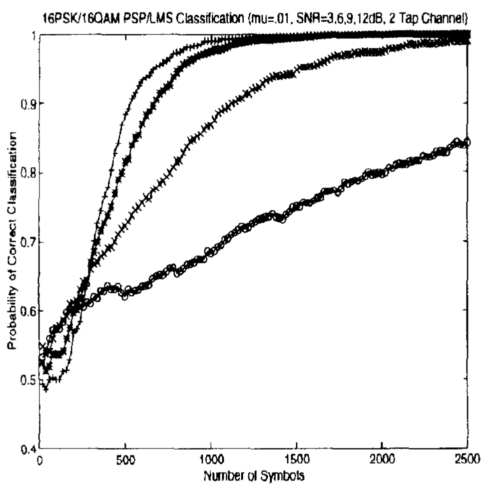

Figure 8: Unknown Channel PSP/GLRT Classification, 16PSK/16QAM, 2 Tap Channel

Figure 8 plots the time domain classification performance of the GLRT for an unknown channel. An adaptive step size of $\mu = .01$ is used for PSP channel estimation. The delayed rise in the probability of correct classification reflects the time required to adaptively acquire the channel prior to reliable classification decisions.

References

[1] R. Raheli, A. Polydoros, C. Tzou, "The Principle of Per-Survivor Processing: A General Approach to Approximate and Adaptive MLSE", accepted for publication in *IEEE Trans. on Comm.*; see also: "Per-Survivor Processing", *Digital Signal Processing - A Review Journal*, Academic Press, Vol. 3 No. 3, pp. 175-187, July 1993.

[2] S. Haykin ed., *Blind Deconvolution*, Prentice-Hall, NJ, 1994.

[3] E. Satorius and J. Mulligan, "An Alternative Methodology for Blind Equalization", *Digital Signal Processing*, Vol. 3 No. 3, pp. 199-209, July 1993.

[4] N. Seshadri, "Joint Data and Channel Estimation Using Blind Trellis Search Techniques", *Blind Deconvolution*, S. Haykin ed., Prentice-Hall, NJ, 1994.

[5] G. Forney, "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference", *IEEE Trans. on Info.Theory*, Vol. IT-18 No.3, pp. 363-378, May 1972.

[6] R. Gooch and J. Lundell, "The CM Array: An Adaptive Beamformer for Constant Modulus Signals", *ICASSP Tokyo 1986*, pp. 2523- 2526.

[7] D. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems", *IEEE Trans. on Comm.*, Vol. COM-28 No.11, pp. 1867-1875, November 1980.

[8] O. Shalvi and E. Weinstein, "New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)", *IEEE Trans. on Info. Theory*, Vol. 36 No. 2, pp. 312-321, March 1990.

[9] H. Van Trees, *Detection, Estimation, and Modulation Theory*, New York, NY: John Wiley & Sons, 1968.

[10] C. Hwang, "Advanced Methods for Digital Modulation Classification", Ph. D. Dissertation, University of Southern California, 1991.

[11] U. Cheng, "Parameter Estimation of Multiple Direct-Sequence Signals", unpublished memo.