Multiple H^∞ Filter based Deterministic Sequence Estimation in Non-Gaussian Channels

Harini Kulatunga, *Student Member, IEEE,* and Visakan Kadirkamanathan, *Member, IEEE*

Abstract

A novel and more robust implementation is proposed for sequence estimation in uncertain environments with additive non-Gaussian ambient noise and inter-symbol interference. This method is based on a deterministic performance index which minimizes the effect of worst case disturbances on the estimation error. The decoder has multiple H^∞ filters and is in the fashion of per-survivor processing with a Viterbi trellis for decoding. There is a substantial performance improvement over maximum likelihood sequence estimation as shown by simulation results obtained for joint channel estimation and symbol decoding in non-Gaussian channels.

EDICS Category: COM-ESTI

I. Introduction

In many physical channels, such as urban and indoor radio channels and underwater acoustic channels, the ambient noise is known to be non-Gaussian, due to man-made interference and natural noise as well, see [1] and references therein. Hence there is an increased interest in demodulation techniques for non-Gaussian Inter-Symbol Interference (ISI) channels [1], [2].

Equalization in ISI channels with additive Gaussian noise has been a well studied problem in the area of communication [3]. [4] introduced a receiver structure, consisting of a linear causal transversal filter (called the Whitened Matched Filter (WMF)), a symbol-rate sampler, and a recursive Viterbi algorithm, which is a maximum-likelihood sequence estimator (MLSE) of the entire transmitted sequence. The Viterbi algorithm requires the knowledge of the channel information and per-survivor processing (PSP) provides a zero-delay adaptive channel estimator [5], [6]. A joint channel estimator and symbol detector based on multiple model theory was developed for Gaussian channels in [7]. The Bit-Error-Rate (BER) performance of the Viterbi and PSP algorithms for joint channel and sequence estimation depends on the channel model, decoding delay and the noise distribution. Therefore an implementation based on the white Gaussian noise assumption will

H. Kulatunga* and V. Kadirkamanathan are with the Dept. of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK (Tel: +44 114 222 5250, Fax: +44 114 222 5661 email: {cop01thk*, visakan}@sheffield.ac.uk).

suffer performance loss when the additive noise is non-Gaussian. Hence, it is of interest to develop robust sequence estimation techniques for uncertain non-Gaussian environments, which is the subject of this paper.

Recently, [2] has used H^∞ theory for the equalization problem. The novel idea proposed in this paper is a Viterbi algorithm based on minimizing a worst-case deterministic cost with multiple H^∞ filters for channel estimation in the fashion of PSP. In this work a recursive deterministic performance index is derived from a norm bounded transfer operator which minimizes the estimation error with respect to worst-case measurement noise, driving noise and initialization error, effectively mitigating the impulsive noise based on a min-max strategy.

II. System Model

The received signal is passed through a WMF, sampled and input to the Viterbi decoder [4]. The discrete-time signal at the input to the decoder is expressed as ,

$$y_i = \sum_{n=0}^{K-1} x_n^{(i)} b_{i-n} + \nu_i \tag{1}$$

where $\{b_i\}$ is the information symbol sequence, T is the symbol period, $\{x_n^{(i)}\}_{n=0}^{K-1}$ are the time-varying coefficients of an equivalent discrete-time filter of $K-1$, S-state time-varying inter-symbol interference effects [3]. ν_i is the additive noise. The main assumptions made in this paper are,

Assumption 2.1 The channel impulse response $f(t)$ has finite duration of length K symbol intervals.

Assumption 2.2 The WMF is $W(z) \triangleq \frac{1}{X(z)} F(z)$ where $X(z)$ is a minimum-phase polynomial in z of degree K with coefficients x_n and $x_0 \neq 0$.

Assumption 2.3 $\{b_i\}$ is a complex, zero-mean and uncorrelated wide-sense stationary (WSS) sequence, i.e., $E(b_i b_j^*) = \delta(i-j)$ drawn from a signal constellation $\mathcal{B}$ of size S.

III. The Proposed H^∞-based Robust Estimation

The channel state information (CSI) is rarely known at the receiver. Denoting the channel coefficients at time instant i as $\mathbf{x}_i = [x_0^{(i)}, \dots, x_{K-1}^{(i)}]^*$, the coefficients are estimated based on the following state-space model,

$$\begin{aligned} \mathbf{x}_{i+1} &= \mathbf{x}_i + \mathbf{w}_i \\ y_i &= \mathbf{h}_i \mathbf{x}_i + \nu_i \\ s_i &= \mathbf{l}_i \mathbf{x}_i \end{aligned} \tag{2}$$

where w_i is a zero-mean random process, $\mathbf{w}_i = [w_0^{(i)}, \dots, w_{K-1}^{(i)}]^*$ and $\mathbf{h}_i = \mathbf{l}_i = [b_i, \dots, b_{i-K+1}]$. Thus we estimate an arbitrary linear combination, say $\check{s}_{i|i}$, of channel coefficients given the received signal $\{y_i\}$ of an ISI channel with non-Gaussian noise. Estimate of s_i indirectly gives an estimate of the state $\mathbf{x}_i$ similar to the innovation approach of the Kalman Filter [8]. The estimation error is,

$$e_i = \check{s}_{i|i} - \mathbf{l}_i \mathbf{x}_i \tag{3}$$

The worst-case performance measure is defined as,

$$\mathcal{J} \triangleq \frac{\sum_{i=1}^{N} e_i^* e_i}{\mathbf{x}_0^* \mathbf{\Pi_0}^{-1} \mathbf{x}_0 + \sum_{i=1}^{N} \mathbf{w}_i^* \mathbf{Q}_i^{-1} \mathbf{w}_i + \sum_{i=1}^{N} \nu_i^* \nu_i} \tag{4}$$

where the H^2 norm of a causal sequence $\{g_i\}$ is $\sum_{i=1}^{N} g_i^* g_i$ and transfer operator $\mathcal{J}$ maps unknown disturbances $\{\mathbf{\Pi_0}^{-1/2}\mathbf{x}_0, \{\mathbf{Q}_i^{-1/2}\mathbf{w}_i\}_{i=1}^{N}, \{\nu_i\}_{i=1}^{N}\}$ (where $\mathbf{\Pi_0}, \mathbf{Q}_i$ are positive definite weighting matrices, a design choice) to the estimated errors $\{e_i\}_{i=1}^{N}$. The optimal estimate $\{\hat{s}_i\}_{i=1}^{N}$, of all possible estimates $\{\check{s}_{i|i}\}_{i=1}^{N}$, should minimize the H^∞ norm of $\mathcal{J}$, i.e.,

$$\inf_{s_i} \|\mathcal{J}\|_\infty = \inf_{s_i} \sup_{\mathbf{x}_0, \mathbf{w} \in h^2, \nu \in h^2} \mathcal{J} < \gamma_f^2 \tag{5}$$

where $\gamma_f^2 > 0$ is a level of disturbance attenuation (a prescribed positive value) for $i = 1, \ldots, N$. The following state-space model can be introduced to describe the estimation problem in (2),(3),

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{w}_i \tag{6}$$

$$\begin{bmatrix} y_i \\ \check{s}_{i|i} \end{bmatrix} = \begin{bmatrix} \mathbf{h}_i \\ \mathbf{l}_i \end{bmatrix} \mathbf{x}_i + \begin{bmatrix} \nu_i \\ e_i \end{bmatrix} \Longrightarrow \mathbf{z}_i = \Theta_i \mathbf{x}_i + \mathbf{v}_i \tag{7}$$

where $\mathbf{v}_i = [\nu_i^*, e_i^*]^*$. The condition (5) can be related to a positivity of an indefinite-quadratic form J_N for the sequence length N obtained by substituting (4) in (5). Because of the underlying state-space model the resulting indefinite quadratic form can be written as,

$$J_N(\mathbf{x}_0, \omega, \mathbf{z}) = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{w} \\ \mathbf{z} \end{bmatrix}^* \left\{ \begin{bmatrix} \mathbf{I}_K & 0 & 0 \\ 0 & \mathbf{I}_{KN} & 0 \\ \mathbf{\Omega} & \mathbf{\Gamma} & \mathbf{I}_{2N} \end{bmatrix} \begin{bmatrix} \mathbf{\Pi_0} & 0 & 0 \\ 0 & \mathbf{Q} & 0 \\ 0 & 0 & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{I}_K & 0 & 0 \\ 0 & \mathbf{I}_{KN} & 0 \\ \mathbf{\Omega} & \mathbf{\Gamma} & \mathbf{I}_{2N} \end{bmatrix}^* \right\}^{-1} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{w} \\ \mathbf{z} \end{bmatrix} > 0 \tag{8}$$

where $\Omega \in \Re^{2N}$, $\Gamma \in \Re^{2N \times N}$ are the observability map and the impulse response matrix, respectively of the state-space model. $\mathbf{I}_x \in \Re^{x \times x}$ is an identity matrix. $\mathbf{z} = [\mathbf{z}_1^*, \ldots, \mathbf{z}_N^*]^*$, $\mathbf{w} = [w_1^*, \ldots, w_N^*]^*$, $\mathbf{v} = [\mathbf{v}_1^*, \ldots, \mathbf{v}_N^*]^*$. Stationary point τ_0 of quadratic form $J_N(\tau, \mathbf{z})$, where $\tau^* = [\mathbf{x}_0^*, \mathbf{w}^*]$, is easily obtained by partial differentiation w.r.t. τ, which gives,

$$J_N(\tau_0, \mathbf{z}) = \mathbf{z}^* \mathbf{R}_z^{-1} \mathbf{z} \tag{9}$$

where $\tau = \tau_0$ for sequence length N. $\mathbf{R}_z$ is a coefficient matrix. However, in order to satisfy condition (8), the minimum has to be positive for the estimates $\{\check{s}_{i|i}\}_{i=1}^{N}$, i.e. $J_N(\tau_0, \mathbf{z}) > 0$.

A. Recursive Calculation of the Performance Measure

A partially equivalent Krein space model (corresponding with (2)) can be defined (partial as Krein space variables are stochastic).

$$\begin{aligned} \overline{\mathbf{x}}_{i+1} &= \overline{\mathbf{x}}_i + \overline{\mathbf{w}}_i; \\ \overline{\mathbf{z}}_i &= \Theta_i \overline{\mathbf{x}}_i + \overline{\mathbf{v}}_i \end{aligned} \tag{10}$$

Krein space values are made distinct with an overline. [8] has detailed Krein space state-space representation and development of inner product matrices which are *equivalent* to the coefficient matrices in the deterministic quadratic form (derived from model (2) and described in (8)). A recursive formula, based on innovations in an *equivalent* Krein space, which can recursively compute the state estimates, subject to the positivity condition of $J_N(\tau_0, \mathbf{z})$, is presented in detail in [8]. These recursions are none other than the H^∞ filter given in [9],[10].

From innovations in Krein space, the column vector $\mathbf{z}$ can be expressed in terms of innovations as $\mathbf{z} = \mathbf{L}\alpha$ where $\mathbf{z}_i = \langle \overline{\mathbf{z}}_i, \overline{\alpha}_j \rangle \langle \overline{\alpha}_j, \overline{\alpha}_j \rangle^{-1} \alpha_j + \alpha_i$ and $\alpha = [\alpha_1^*, \ldots, \alpha_N^*]^*$, innovations of the deterministic quantities and $\langle \cdot, \cdot \rangle$ denotes an *gramian* matrix in stochastic Krein space. The minimum value of the quadratic form (8) expressed in terms of the innovations is $J_N(\tau_0, \mathbf{z}) = \alpha^* \mathbf{R}_\alpha^{-1} \alpha$, where $\mathbf{R}_\alpha = \mathrm{diag}(\mathbf{R}_{\alpha,1} \cdots \mathbf{R}_{\alpha,N})$ and coefficient matrix $\mathbf{R}_{\alpha,j} = \langle \overline{\alpha}_j, \overline{\alpha}_j \rangle$. Thus $\mathbf{R}_\alpha$ can be estimated recursively by Krein space projection of innovations leading to the recursive computation of $J_N(\tau_0, \mathbf{z})$ as,

$$J_i(\tau_{0|i}, \mathbf{z}) = J_{i-1}(\tau_{0|i-1}, \mathbf{z}) + \alpha_i^* \mathbf{R}_{\alpha,i}^{-1} \alpha_i \tag{11}$$

Therefore the condition for the existence of the H^∞ estimator is

$$J_i(\tau_{0|i}, \mathbf{z}) = \sum_{j=1}^{i} \begin{bmatrix} \alpha_{y,j} \\ \alpha_{s,j} \end{bmatrix}^* \mathbf{R}_{\alpha,j}^{-1} \begin{bmatrix} \alpha_{y,j} \\ \alpha_{s,j} \end{bmatrix} > 0 \tag{12}$$

where $\alpha_i = [\alpha_{y,i}^*, \alpha_{s,i}^*]^*$ and $\alpha_{y,i} = y_i - \hat{y}_{i|i-1}$ and $\alpha_{s,i} = \check{s}_{i|i} - \hat{s}_{i|i-1}$. From state-space model (10), $\mathbf{R}_{\alpha,j} =< \Theta_j \tilde{\overline{\mathbf{x}}}_j, \Theta_j \tilde{\overline{\mathbf{x}}}_j > + < \overline{v}_i, \overline{v}_i >$ where $\tilde{\overline{\mathbf{x}}}_j = \overline{\mathbf{x}}_j - \hat{\overline{\mathbf{x}}}_{j|j-1}$. Block upper-lower triangular factorization of $\mathbf{R}_{\alpha,j} = \mathcal{L}\mathbf{D}\mathcal{U}$ leads to a simplified form of (12),

$$\sum_{j=1}^{i} (y_j - \hat{y}_{j|j-1})^* (I + \mathbf{h}_j \mathbf{P}_j \mathbf{h}_j^*)^{-1} (y_j - \hat{y}_{j|j-1}) + \sum_{j=1}^{i} (\check{s}_{j|j} - \hat{s}_{j|j})^* (-\gamma_f^2 I + \mathbf{l}_j (\mathbf{P}_j^{-1} + \mathbf{h}_j^* \mathbf{h}_j)^{-1} \mathbf{l}_j^*)^{-1} (\check{s}_{j|j} - \hat{s}_{j|j}) > 0 \tag{13}$$

where $\hat{s}_{j|j} = \hat{s}_{j|j-1} + \mathbf{l}_j \mathbf{P}_j \mathbf{h}_j^* (I + \mathbf{h}_j \mathbf{P}_j \mathbf{h}_j^*)^{-1} (y_j - \hat{y}_{j|j-1})$. Therefore the simplest choice that satisfies condition (13) would be,

$$\check{s}_{j|j} = \hat{s}_{j|j} \tag{14}$$

giving one possible level-γ_f filter. The performance measure at time instant i is,

$$J_i(\tau_{0|i}, \mathbf{z}) = \sum_{j=1}^{i} (y_j - \hat{y}_{j|j-1})^* (I + \mathbf{h}_j \mathbf{P}_j \mathbf{h}_j^*)^{-1} (y_j - \hat{y}_{j|j-1}) \tag{15}$$

which is a unique minimum at stationary point $\tau_{0|i}$ for sequence length i and guarantees the existence condition (13) of a certain level-γ_f H^∞ filter. $\hat{y}_{j|j-1} = \mathbf{h}_j \hat{\mathbf{x}}_{j|j-1}$. $\hat{\mathbf{x}}_{j|j-1} = \hat{\mathbf{x}}_{j-1|j-1}$ and $\hat{\mathbf{x}}_{j-1|j-1}$ is calculated recursively using the Riccati H^∞ filter.

B. H^∞ Filtering

The discrete-time Riccati solution for the H^∞ filter is given in [10], [9], and $\mathbf{P}_j$ in (15) satisfies the Riccati recursion,

$$\mathbf{P}_{j+1} = \mathbf{P}_j + \mathbf{Q}_j - \mathbf{P}_j \begin{bmatrix} \mathbf{h}_j^* & \mathbf{l}_j^* \end{bmatrix} \mathbf{R}_{\alpha,j}^{-1} \begin{bmatrix} \mathbf{h}_j \\ \mathbf{l}_j \end{bmatrix} \mathbf{P}_j \tag{16}$$

$$\mathbf{R}_{\alpha,j} = \begin{bmatrix} 1 & 0 \\ 0 & -\gamma_f^{-2} \end{bmatrix} + \begin{bmatrix} \mathbf{h}_j \\ \mathbf{l}_j \end{bmatrix} \mathbf{P}_j \begin{bmatrix} \mathbf{h}_j^* & \mathbf{l}_j^* \end{bmatrix}; \tag{17}$$

The adaptive H^∞ filtering problem can have many solutions, each for a proper γ_f value. The $\gamma_f > 0$ value satisfies the condition,

$$\mathbf{P}_j^{-1} + \mathbf{h}_j^* \mathbf{h}_j - \gamma_f^{-2} \mathbf{l}_j^* \mathbf{l}_j > 0, \quad \forall j \tag{18}$$

If $\mathbf{P}_j$ exists then one possible level-γ_f filter will be,

$$\check{s}_{j|j} = \mathbf{l}_j \hat{\mathbf{x}}_{j|j},$$

$$\hat{\mathbf{x}}_{j+1|j+1} = \hat{\mathbf{x}}_{j+1} + \mathbf{K}_{j+1}(y_{j+1} - \mathbf{h}_{j+1}\hat{\mathbf{x}}_{j+1}) \tag{19}$$

$$\hat{\mathbf{x}}_{j+1} = \hat{\mathbf{x}}_{j|j},$$

$$\mathbf{K}_{j+1} = \mathbf{P}_{j+1}\mathbf{h}_{j+1}^*(I + \mathbf{h}_{j+1}\mathbf{P}_{j+1}\mathbf{h}_{j+1}^*)^{-1} \tag{20}$$

Effectively, the states $\hat{\mathbf{x}}_{j|j}$ are estimated recursively (19),(20) with filters initialized with $\hat{\mathbf{x}}_0$.

IV. Deterministic Sequence Estimation - PSP-H^∞ Algorithm

The discrete-time linear filter model for ISI channels has S^{K-1} states, the state at any time is given by the $K-1$ most recent inputs, $s_i \triangleq (b_{i-1}, \ldots, b_{i-K+1})$. The output at any time instant is determined by two successive states, $y_i = y(s_i, s_{i+1})$. Deterministic sequence estimation is defined as the choice of that $\{b_i\}_{i=1}^N$ for which the minimum worst-case quadratic cost $J_N(\cdot)$ w.r.t. measurement and driving noise and initialization errors, is minimized. To construct the recursive algorithm we use (15) from the previous section. The prediction $\hat{y}_{i|i-1}$ at time instant i is determined by transition from state s_i to s_{i+1}, i.e. $\hat{y}_{i|i-1} = \hat{y}(s_i, s_{i+1})$. Then min-max cost for a particular sequence breaks up into sum of independent increments:

$$J_N(\tau_{0|N}, \mathbf{z}) = \sum_i \Delta J_i(y_i - \hat{y}(s_i, s_{i+1})), \tag{21}$$

where $\Delta J_i(\cdot)$ is the cost increment at time instant i. The multiple mode representation [11], [7] is introduced which can be used to expound the joint sequence decoder and estimator, and provides an alternative view of the Viterbi trellis structure.

The state $s_i = (b_{i-1}, \ldots, b_{i-K+1})$ in the trellis diagram can be represented by a *mode sequence* denoted as $M[i]$. One possible mode sequence $M[i] = M_r$ is written in short as M_r^i and $r \in \{1, \ldots, S^{K-1}\}$. State s_i

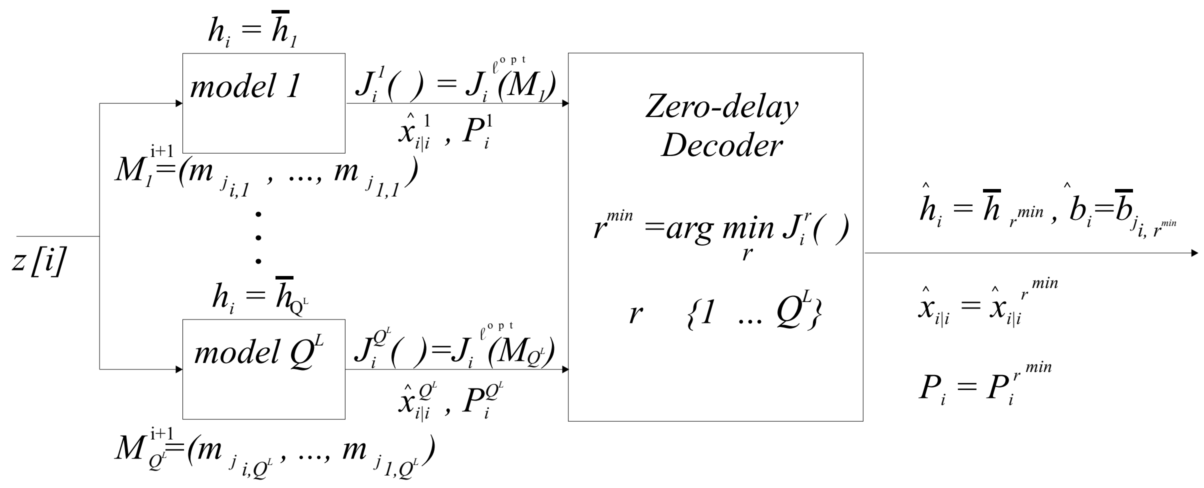


Fig. 1. Multiple Model Decoder Framework.

associated with mode sequence M_r^i can be expressed in terms of modes as,

$$M_r^i(s_i) = (m_{j_{i-1,r}}, \ldots, m_{j_{i-K+1,r}}), \quad \text{for } r \in \{1, \ldots, S^{K-1}\}. \tag{22}$$

where the indices $j_{i-n,r}$ can take values in the range $1 \leq j_{i-n,r} \leq S$. Each *symbol-mode* association $\{\bar{b}_{j_{i-n,r}}, m_{j_{i-n,r}}\}_{j_{i-n,r}=1}^{S}$ represents a possible value of the transmitted symbol $b_{i-n} \in \{\bar{b}_1, \ldots, \bar{b}_S\}$ and the associated mode. For $n = 1, \ldots, K-1$ the sequence $(m_{j_{i-1,r}}, \ldots, m_{j_{i-K+1,r}})$ forms the rth possible symbol sequence from time $i-K+1$ to $i-1$. Then $\{\bar{h}_r, M_r^i\}_{r=1}^{S^{K-1}}$ represents a possible value of the transmitted symbol sequence $(b_{i-1}, \ldots, b_{i-K+1}) \in \{\bar{h}_1, \ldots, \bar{h}_{S^{K-1}}\}$ and the associated system mode. A state sequence from state s_1 to s_i is then denoted based on (22) as a sequence of modes,

$$M_r^i(s_1, \ldots, s_i) = (m_{j_{i-1,r}}, \ldots, m_{j_{1,r}}), \quad \text{for } r \in \{1, \ldots, S^{i-1}\}. \tag{23}$$

B. Viterbi Decoder

Cost $\Delta J_i(\cdot)$ (21) is associated with the transition from state s_i to s_{i+1}. We denote the partial sum from state s_i to state s_j as $\Gamma[M_r^j(s_i, \ldots, s_j)] \triangleq \sum_{k=i}^{j-1} \Delta J_k(y_k - \hat{y}(s_k, s_{k+1}))$. Then the deterministic minimum worst-case cost (15) can be expressed in terms of partial costs as follows,

$$J_N(\tau_{0|N}, \mathbf{z}) = J_i(\tau_{0|i}, \mathbf{z}) + \Gamma[M_r^{N+1}(s_{i+1}, \ldots, s_{N+1})] \tag{24}$$

Let $M_{r,opt}^{i+1}(s_1, \ldots, s_{i+1})$ be the mode sequence from time 1 to i that gives the least minimum worst-case cost $J_i(\tau_{0|i}, \mathbf{z}) = \Gamma[M_r^{i+1}(s_1, \ldots, s_{i+1})]$ among all allowable mode sequences starting from state s_1 and ending with state s_{i+1} (i.e. $M_{r,opt}^{i+1}(\cdot)$ is a *survivor* sequence which is ending at state s_{i+1}). Then, from the premise of dynamic programming the initial segment $J_i(\tau_{0|i}, \mathbf{z}) = M_r^{i+1}(\cdot)$ of any other sequence, passing through s_{i+1}, can be replaced with the segment $M_{r,opt}^{i+1}(\cdot)$ and obtain a sequence which minimizes the sequence cost $J_i(\cdot)$ even further.

Now apply this argument to a finite state machine of length $K-1$. The state $s_{i+1} = (b_i, \ldots, b_{i-K+2})$ is not known but it is one of a finite number of states s_{i+1}^r, $1 \leq r \leq S^{K-1}$ of the trellis diagram representing the finite state machine. That is the s_{i+1}^rth state at time-instant i represents the $M_{r,opt}^{i+1}(s_1, \ldots, s_i)$th mode sequence at time-instant i. The latter segment of each of the mode sequences at each state has length $K-2$ and must be among the initial segment of S^{K-2} out of the S^{K-1} mode sequences at states $\{s_i^r\}$. Therefore there are always S paths converging on each state and the path with the least cost (path metric) is the survivor sequence $M_{r,opt}^{i+1}(\cdot)$. If $J_i(\tau_{0|i}, \mathbf{z}) = \Gamma[M_{r,opt}^{i+1}(s_1, \ldots, s_{i+1})]$ is the path metric at each state, then the path metric is updated at time $i+1$ as,

1) For each of the S continuations from each of the S^{K-1} survivors from time instant i compute

$$J_{i+1}^l(\cdot) = \Gamma[M_{r,opt}^{i+1}(s_1, \ldots, s_{i+1})] + \Delta J_{i+1}^l(y_{i+1} - \hat{y}(s_{i+1}, s_{i+2})), \quad \text{for } l = 1, \ldots, S. \tag{25}$$

$\{\Delta J_{i+1}^l(\cdot)\}$ are called the branch metrics.The channel state estimates from the H^∞ filter associated with the surviving path at state s_{i+1}^r is used to calculate the branch matrix, $\Delta J_{i+1}^l(y_{i+1} - \hat{y}(s_{i+1}, s_{i+2})) = \mathbf{h}_{i+1}\hat{\mathbf{x}}_{i|i}^r$, associated with each continuation from state s_{i+1}. $\mathbf{h}_{i+1}$ is not known and will take values according to the mode sequence associated with each continuation. The channel state and variance for each of these paths will be $\{\hat{\mathbf{x}}_{i+1|i}^l = \hat{\mathbf{x}}_{i|i}^r, \mathbf{P}_{i+1|i}^l = \mathbf{P}_{i|i}^r\}$.

2) For each of the states s_{i+2}^r, $1 \leq r \leq S^{K-1}$ representing the transitions at time $i+1$, compare the costs of the S continuations terminating in each state and select the smallest as the corresponding survivor,

$$l^{opt} = \min_l J_{i+1}^l(M_r) \tag{26}$$

Therefore $J_{i+1}^r(\tau_{0|i+1}, \mathbf{z}) = \Gamma[M_{r,opt}^{i+2}(s_1, \ldots, s_{i+2})] = J_{i+1}^{l^{opt}}(M_r)$. The corresponding channel estimate and Riccati recursion at each state will be (19),(16), initialized with $\{\hat{\mathbf{x}}_{i+1} = \hat{\mathbf{x}}_{i+1|i}^{l^{opt}}, \mathbf{P}_i = \mathbf{P}_{i+1|i}^{l^{opt}}\}$.

The PSP-H^∞ algorithm is given in Table I.

V. Simulations and Results

An illustration of the performance of the PSP-H^∞ algorithm for jointly estimating the symbol sequence and channel coefficients in the presence of severe inter-symbol interference (ISI) is obtained by Monte Carlo (MC) simulation. Consider the discrete time-invariant wireless channel characteristics [0.227 0.460 0.688 0.460 0.227] given in [3]. The additive noise ν_i in the received signal is impulsive, uncorrelated with the information sequence $\{b_i\}$, and modeled by a two-term Gaussian mixture model with probability density function (pdf) of,

$$p(\nu) = (1-\epsilon)\mathcal{N}(0, \rho_n^2) + \epsilon\mathcal{N}(0, \rho_I^2) \tag{27}$$

where $\mathcal{N}(0, \rho^2)$ represents the zero-mean Gaussian distribution function with variance ρ^2. The simulations show Bit-error rate (BER) performance versus Signal-to-noise ratio (SNR) where SNR is calculated by,

$$SNR = \frac{1}{N_0}\sum_n \|x_n\|^2$$
$$N_0 = (1-\epsilon)n_0 + \epsilon\sigma n_0 \tag{28}$$

TABLE I

PSP-H^{∞} FOR JOINT SYMBOL AND CHANNEL ESTIMATION IN ISI CHANNELS WITH NON-GAUSSIAN NOISE.

Initialized with $\mathbf{x}_0, \mathbf{P}_0$; For sequence length $i = 1, \ldots, N$,

FOR $r = 1 : S^{K-1}$ DO

Consider the paths from S number of previous states, $s_i = s_i^{r'}$ to state $s_{i+1} = s_{i+1}^{r}$, $1 \leq r' \leq S^{K-1}$ but $s_i^{r'}(1 : K-2) = s_{i+1}^{r}(2 : K-1)$ therefore only S number of previous states are considered.

FOR $l = 1 : S$ DO

$\mathbf{h}_i^l = \overline{\mathbf{h}}_r$

$\mathbf{x}_{i|i-1}^l = \hat{\mathbf{x}}_{i-1|i-1}^{r'}$, $\mathbf{P}_{i|i-1}^l = \mathbf{P}_{i-1}^{r'}$

$\hat{y}(s_i^{r'}, s_{i+1}^r) = \mathbf{h}_i^l \mathbf{x}_{i|i-1}^l$

Calculate $\Delta J_i^l(y_i - \hat{y}(s_i^{r'}, s_{i+1}^r))$

$J_i^l = J_{i-1}^{r'}(\tau_{0|i-1}, \mathbf{z}) + \Delta J_i^l(y_i - \hat{y}(s_i^{r'}, s_{i+1}^r)) = \Gamma[M_{r',opt}^i(s_1, \ldots, s_i)] + \Delta J_i^l(y_i - \hat{y}(s_i^{r'}, s_{i+1}^r))$

END

$l^{opt} = \min_l J_i^l \Rightarrow J_i^r(\tau_{0|i}, \mathbf{z}) = J_i^{l^{opt}} = \Gamma[M_{r,opt}^{i+1}(s_1, \ldots, s_{i+1})]$

$\mathbf{x}_{i|i}^r = \mathbf{x}_{i|i-1}^{l^{opt}} + \mathbf{K}_i(y_i - \mathbf{h}_i^{l^{opt}} \mathbf{x}_{i|i-1}^{l^{opt}})$; $\mathbf{P}_i^r = \left[\mathbf{I} - \begin{bmatrix} \mathbf{h}_i^{l^{opt}*} & \mathbf{l}_i^* \end{bmatrix} \mathbf{R}_{e,i}^{-1} \begin{bmatrix} \mathbf{h}_i^{l^{opt}*} \\ \mathbf{l}_i \end{bmatrix}\right] \mathbf{P}_{i|i-1}^{l^{opt}} + \mathbf{Q}_i$; $\mathbf{l}_i = \mathbf{h}_i^{l^{opt}*}$

END

IF $i >=$ decoding delay D,

$r^{min} = \min_r J_i^r(\tau_{0|i}, \mathbf{z})$

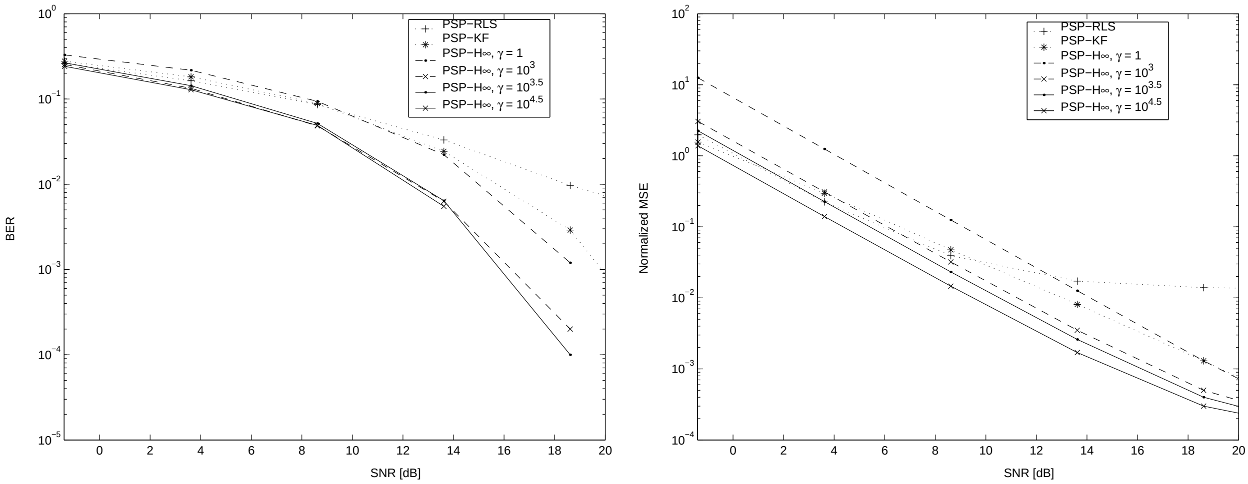
Work back along the mode sequence $M_{r^{min},opt}^{i+1}(s_1, \ldots, s_{i+1})$ associated with path ending at state $s_{i+1} = s_{i+1}^{r^{min}}$ and the 'symbol-mode' association will indicate the transmitted symbol at time $i - D$ as shown in Fig. 1 for zero-delay, $D = 0$.

END

where ϵ, σ are design parameters for the impulsive noise model and $\rho_n^2 = n_0$ and $\rho_I^2 = \sigma n_0$ are the variances of the two Gaussian distributions in (27). Considering an anti-podal (BPSK) transmission where $S = 2$, the performance of the robust algorithm is compared with Recursive Least Squares (RLS) and Kalman filter based PSP algorithms. The results in Figure 2 show that the robust version of the PSP algorithm has better BER performance over the whole SNR range. At very low SNR noise dominates the received signal and all filters exhibit similar BER, as the SNR improves the H^{∞} based PSP algorithm performs better than the other methods. This is attributed to the fact that H^{∞} estimator is robust to non-Gaussian noise whereas the convergence of the channel state estimates of the Kalman and RLS methods will be perturbed due to impulsive noise. In the case of time-invariant channel with AWGN, Figure 3, the H^{∞} filter that considers the worst-case scenario and the Kalman and RLS filters that expects noise to be Gaussian converge to a constant channel gain.

VI. Conclusion

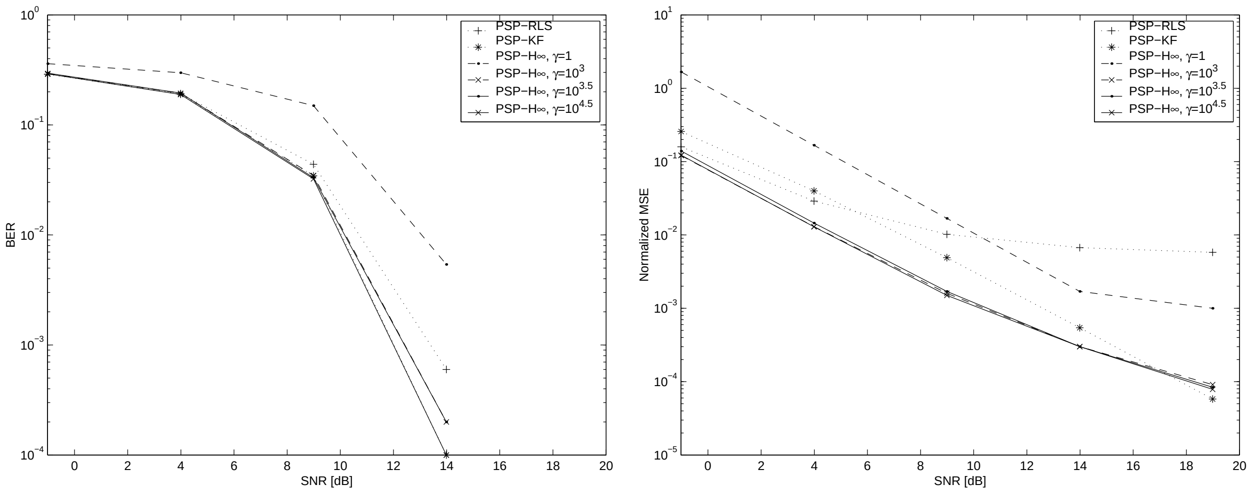
An adaptive H^{∞} filtering algorithm is developed which is presented from a multiple model viewpoint to equalize wireless ISI channels. Essentially, it is a parallel filtering algorithm with Viterbi decoding and the performance index minimizes the effect of worst-case disturbances by minimizing a deterministic quadratic



(a) BER vs. SNR.

(b) Averaged Normalized MSE vs. SNR.

Fig. 2. Equalizer length is 6, Training period = 60, $T = 10,000$, $\sigma = 100$, $\epsilon = 0.1$. Varying γ_f.



(a) BER vs. SNR.

(b) Averaged Normalized MSE vs. SNR.

Fig. 3. Equalizer length is 6, Training period = 150, $T = 10,000$, Gaussian noise. Varying γ_f.

function. The proposed robust sequence estimator outperforms, at low SNR, the maximum likelihood based sequence estimators in non-Gaussian noise.

References

[1] H. V. Poor and M. Tanda, "Multiuser detection in flat fading non-gaussian channels," *IEEE Trans. Commun.*, vol. 50, pp. 1769–1777, Nov. 2002.

[2] A. T. Erdogan, B. Hassibi, and T. Kailath, "FIR H^∞ equalization," *Signal Processing*, vol. 81, pp. 907–917, 2001.

[3] J. G. Proakis, *Digital Communications.* New York, NY: McGraw-Hill, Inc., 2001.

[4] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363–378, May 1972.

[5] R. Raheli, A. Polydoros, and C. Tzou, "Per-Survivor Processing: A general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 43, pp. 354–364, Feb./Mar./Apr. 1995.

[6] P. R. Chevillat and E. Eleftheriou, "Decoding of trellis-encoded signals in the presence of intersymbol interference and noise," *IEEE Trans. Commun.*, vol. 37, pp. 669–676, July 1989.

[7] H. Kulatunga and V. Kadirkamanathan, "Adaptive joint detection and estimation in MIMO systems: Hybrid systems approach," *IEEE Trans. Signal Processing*, Accepted for publication.

[8] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic Estimation and Control. A Unified approach to H^2 and H^∞ Theories.* Philadelphia, PA: Siam studies in Applied and Numerical Mathematics, 1999.

[9] P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an H^∞-setting," *IEEE Trans. Automat. Contr.*, vol. 36(2), pp. 151–166, Feb. 1991.

[10] U. Shaked and Y. Theodor, "H^∞ - Optimal estimation: A tutorial," in *Proc. IEEE Decision and Control'92*, Phoenix, AZ, 1992, pp. 2278–2286.

[11] Y. Bar-Shalom and X.-R. Li, *Estimation and tracking : principles, techniques and software.* Boston, NY: Artech House, 1993.