https://doi.org/10.1109/KES.1998.725840

Optimal HMM filtering and Decision Feedback Equalisation for Differential Encoded Transmission Systems.

Jason Ford & John Moore
Department of Systems Engineering,
Research School of Information Sciences and Engineering,
Australian National University, Canberra ACT Australia 0200.
Ph +61(2)62798683 and email Jason.Ford@anu.edu.au.

Keywords: Hidden Markov Model, Differential transmission, Decision Feedback Equalisation.

Abstract

In this paper conditional hidden Markov model (HMM) filters and conditional Kalman filters (KF) are coupled together to improve demodulation of differential encoded signals in noisy fading channels.

We present an indicator matrix representation for differential encoded signals and the optimal HMM filter for demodulation. The filter requires $O(N^3)$ calculations per time iteration, where N is the number of message symbols. Decision feedback equalisation is investigated via coupling the optimal HMM filter for estimating the message, conditioned on estimates of the channel parameters, and a KF for estimating the channel states, conditioned on soft infomation message estimates. Here the soft message estimates are conditional mean estimates. The conditional KF is an adaptive channel estimation scheme based on modelling the phase and amplitude variations as a stochastic linear system.

The key to our coupled HMM-KF filter approach is that the HMM filter provides immediate soft information message estimates and our Kalman Filter exploits the idempotent nature of Markov chains.

The particular differential encoding scheme examined in this paper is differential phase shift keying (DPSK). However, the techniques developed can be extended to other forms of differential modulation. The channel model we use allows for multiplicative channel distortions and additive white Gaussian noise. Simulation studies are also presented.

1 Introduction

A frequently occurring problem encountered in wireless digital communication systems is signal fading which results from multiple propagation paths [8]. Two complementary techniques to enhance detection of a message sequence transmitted over unknown channels are decision feedback equalisation (DFE) and differential encoding schemes.

In message estimation, channel knowledge can be used to counter the distortions introduced by the channel, yet the transmission channel is frequently unknown, see [14, 15, 10, 2]. In 1965, Lucky [13] proposed an adaptive method for estimating the channel distortions now known as decision feedback equalisation. Decision feedback equalisation uses 'best' estimates of the transmitted message to estimate the transmission channel and hence improve message detection.

In slowly varying channel environments, message estimates from a Viterbi algorithm, proposed in 1967 by Viterbi[9], invariably delayed to give improved estimation, can be used in a decision feedback structure [8, Page 651]. However, in wireless transmission the fading environment typically varies relatively rapidly [12] and channel estimates are required immediately.

In differentially encoded signals, the message is encoded in the difference between symbols, see Proakis[8, Page 187] for further explanation. This means that the signals can be detected even when the transmission channel is unknown as long as the channel variation between successive symbols is small. Demodulation of differential encoded signals is therefore not contingent on explicit knowledge of the transmission channel. An important property of differential modulation is that there is rapid recovery after channel nulls such as are typical in Rayleigh fading transmission channels. However, the performance of difference modulation can be poor when there is rapid channel variation. Standard demodulation of differential signals should be viewed as a form of digital demodulation (or decision feedback equalisation) where the channel estimate is derived only from the previous symbol [8].

Decision feedback equalisation based on immediate message estimates is one technique for overcoming channel distortion introduced into differential transmission. Here we propose a receiver that uses more than the last received symbol to estimate the transmission channel.

For clarity, in this paper we only consider the simplest encoder structure, that is, differential modulation without trellis coding or bit interleaving, and consider a simply decision feedback equalisation structure. A similar approach to here was taken by Collings and Moore in [10], the key difference here is that we propose the optimal HMM filter and exploit the idempotent property of Markov chains in the formulation of our Kalman filter, parallelling the work in [3, 4]. The techniques developed is this paper can be easily extended for more complicated transmitter structures.

We begin this paper by introducing a indicator vector state space formulation of differential signalling. Using this state space formulation and by exploiting the idempotent nature of Markov states, an $O(N^3)$ optimal HMM filter is presented in an informative way which highlights the structure of the problem.

We next incorporate our HMM filter into a decision feedback equalisation structure. We propose a receiver structure where a conditional HMM filter is coupled to a conditional Kalman filter to incorporate decision feedback equalisation to aid demodulation. In order to investigate the proposed decision feedback equalisation we first present a standard differential demodulation scheme and highlight the effect of rapidly changing channel conditions on its performance. Following this we investigate the use of decision feedback on differentially coded transmission systems.

This paper is organised as follows: In Section 2, we formulate the HMM, signal model and channel model for a differential encoded system. In Section 3, we introduce our optimal HMM filter. In Section 4, we present a standard differential receiver and highlight the effect of rapid channel variation on receiver performance. In Section 5, a decision feedback equaliser is proposed based on a Kalman filter. In Section 6, simulation studies are presented. Finally, some conclusions are presented in Section 7.

2 State Space Formulation

In this section we present a differential phase modulation signal model, a channel model and then reformulated these as an HMM signal model in state space form and an associated state space stochastic channel model respectfully.

To simplify the discussion in this paper, we assume that digital phase modulation is used to transmit the signal. This type of digital phase modulation is usually called phase-shift keying (PSK). Other forms of modulation not considered here that could be handled by this approach include: Pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM) and others.

We assume that the relevant match or correlation filter demodulators have been implemented and symbol synchronisation and timing issues have be resolved. We proceed now with a signal space analysis.

2.1 Signal Model

In MDPSK transmission schemes the carrier signal is transmitted as phase information over the channel. Let f_k be our message signal, being a real values discrete-time signal, where

$$f_k \in Z_f = \{Z_f^{(1)}, \dots, Z_f^{(N)}\} , \quad Z_f^{(i)} = (i/N)2\pi \in \mathbb{R} \tag{2.1}$$

and we denote the vector z_f as follows

$$z_f = (Z_f^{(1)}, \dots, Z_f^{(N)}). \tag{2.2}$$

In differential transmission systems the carrier symbol is the modulo sum of the message sequence. If we let θ_k denote the carrier symbol, then

$$\theta_k = (\theta_{k-1} + f_k)_{2\pi} \tag{2.3}$$

where $(.)_{2\pi}$ denotes a modulo 2π operation.

The transmitted symbol at time k, represented in the customary complex baseband notation,

$$m_k = \exp(j\theta_k) \quad \in \mathbb{C} \tag{2.4}$$

where imaginary and real components are transmitted using the quadrature and in-phase components of a carrier waveform.

2.2 Channel Model

The baseband signal m_k is transmitted via a channel which can cause both amplitude attenuation and phase shift. The channel can be represented as a multiplicative disturbance, g_k.

$$g_k = \kappa_k \exp(j\phi_k) = g_k^R + jg_k^I \quad \in \mathbb{C} \tag{2.5}$$

where the superscripts R and I refer to the real and imaginary parts. This disturbance introduces time-varying gain and phase changes to the signal and is assumed to vary slowly.

The baseband observation process y_k is thus assumed to have the form

$$y_k = g_k m_k + w_k \quad \in \mathbb{C} \tag{2.6}$$

We define $Y_k \triangleq (y_0, ..., y_k)$ and $w_k \sim N(0, R_k)$. We assume w_k is complex with real and imaginary parts that are i.i.d., with zero mean and Gaussian density, ie. $w_k^R \sim N(0, \sigma_R^2)$ and $w_k^I \sim N(0, \sigma_I^2)$, where w_k^R and w_k^I are the real part and imaginary parts of w_k respectively. Let $\mathcal{Y}_k$ denote the complete filtration generated by $y_\ell,\ \ell \leq k$. As a consequence,

$$E[w_{k+1}|\mathcal{Y}_k] = 0. \tag{2.7}$$

This model of the channel is simplistic in that it allows no inter-symbol interference (ISI) and assumes Gaussian noise, but it is realistic in narrow band communication. This channel model can represent fading channels through the variation in ϕ_k and κ_k.

2.3 State Space Signal Model

A discrete-time state space model for the signal model in the previous section is now presented. Consider the following assumption on the message sequence, f_k

Assumption on the message source

$$f_k \text{ is a first order Markov chain} \tag{2.8}$$

For linear modulation without memory (such as PSK and QAM) this assumption appears inappropriate because the symbols from the message source are usually assumed to be mutually independent. There would seem no advantage in viewing the message sequence as a Markov chain. However, for various other modulation techniques such as NRZI and Miller coding this Markov assumption appears more natural. This assumption also appears appropriate for the case when message symbols are convolutional coded. The HMM filter we present below could be used to generate preliminary state estimates for the sole purpose of estimating the channel in systems with trellis or turbo coded signals with interleaving. The Viterbi algorithm then later used in parallel to produce the final message estimates.

Let us define an indicator vector $X_k^f \in \{e_1, ..., e_n\}$ associated with message symbol, $f_k \in Z_f$, where $e_i = (0, ..., 0, 1, 0, ...0)'$ with 1 in the ith position. That is, to each possible message symbol, $Z_f^{(i)}$, we associate an indicator vector, e_i. We can now write f_k in terms of X_k^f as

$$f_k = z_f' X_k^f. \tag{2.9}$$

Hence, under assumption (2.8) the transition probability matrix of the message process is

$$\mathbf{A} = (a_{ij}) \quad 1 \leq i, j \leq N \tag{2.10}$$

where

$$a_{ij} = P(X_{k+1}^f = e_i | X_k^f = e_j) \tag{2.11}$$

so that

$$E[X_{k+1}^f | X_k^f] = AX_k^f \tag{2.12}$$

where $E[.]$ denotes the expectation operator. We also denote $\{\mathcal{F}_\ell, \ell \in \mathcal{Z}^+\}$ the complete filtration generated by X_k^f, that is, for any $k \in \mathcal{Z}^+$, $\mathcal{F}_k$ is the complete filtration generated by $X_\ell^f,\ \ell \leq k$.

Lemma 2.1 *Under the assumption that f_k is a first order Markov chain the dynamics of X_k^f are given by the state equation*

$$X_{k+1}^f = AX_k^f + M_{k+1} \tag{2.13}$$

where M_{k+1} is a $(A, \mathcal{F}_k)$ martingale increment, in that $E[M_{k+1}|\mathcal{F}_k] = 0$.

Proof See [1]

Likewise, let us define an indicator vector, $X_k^\theta \in \{e_1, ..., e_n\}$, associated with carrier symbol, θ_k, such that

$$\theta_k = z_f' X_k^\theta \tag{2.14}$$

Lemma 2.2 *The indicator functions X_k^θ and X_k^f are related as follows*

$$X_{k+1}^\theta = D(X_{k+1}^f) X_k^\theta \quad or \tag{2.15}$$

$$X_{k+1}^\theta = D(X_k^\theta) X_{k+1}^f \tag{2.16}$$

where $D(e_i) = S^{\underline{n}' e_i}$ the shift operator. Note, S is defined as

$$S = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \tag{2.17}$$

and $\underline{n} = (1, \ldots, N)'$.

Proof Equation (2.15) comes from equation (2.3). Equation (2.16) follows from noting that X^f_{k+1} and X^θ_k are indicator vectors and as such non-linear functions can be written as linear functions, see [1]. □

The observation process can now be expressed in terms of the indicator vectors as follows.

$$\begin{aligned} y_k &= \begin{pmatrix} y^R_k \\ y^I_k \end{pmatrix} \\ &= \begin{pmatrix} m^R_k & -m^I_k \\ m^I_k & m^R_k \end{pmatrix} \begin{pmatrix} g^R_k \\ g^I_k \end{pmatrix} + \begin{pmatrix} w^R_k \\ w^I_k \end{pmatrix} \\ &= \begin{pmatrix} (z^R_f)' X^\theta_k & -(z^I_f)' X^\theta_k \\ (z^I_f)' X^\theta_k & (z^R_f)' X^\theta_k \end{pmatrix} \begin{pmatrix} g^R_k \\ g^I_k \end{pmatrix} \\ &\quad + \begin{pmatrix} w^R_k \\ w^I_k \end{pmatrix} \\ &= H(z_f, X^\theta_k) \begin{pmatrix} g^R_k \\ g^I_k \end{pmatrix} + w_k \end{aligned}$$

with appropriate definition of $H(z_f, X^\theta_k)$ and where the superscripts R and I refer to the real and imaginary parts.

We shall define two vectors of parametrised probability densities as $\mathbf{B}_k = diag(b_k(e_1), \ldots, b_k(e_N))$, and $b_k(e_i) = P[Y_k | X^\theta_k = e_i]$ and $\mathbf{B}_{k|i} = diag(b_{k|i}(e_1), \ldots, b_{k|i}(e_N))$, and $b_{k|i}(e_j) = P[Y_k | X^f_k = e_j, X^\theta_{k-1} = e_i]$.

In the special case, as here, when w_k is complex and its components have Gaussian densities, we can write

$$\begin{aligned} b_k &= \frac{1}{2\pi\sigma_R\sigma_I} \exp\left(-\frac{(y_R - H(z_f, e_i)^R g_k)^2}{2\sigma_R^2} \right. \\ &\quad \left. - \frac{(y_I - H(z_f, e_i)^I g_k)^2}{2\sigma_I^2} \right); \\ &\quad b_{k|i}(e_j) = a_{ji} b_{k-1}(e_i) \end{aligned}$$

where $H(z_f)^R$ and $H(z_f)^I$ are the real and imaginary parts of $H(z_f)$ respectively.

3 Optimal HMM filter

Standard demodulation of DPSK is performed by phase comparison. Optimal path demodulation of convolution coded signals is performed by the Viterbi algorithm. Here we present the optimal HMM filter for message demodulation in two ways.

The key advantage of the HMM filter over standard demodulation techniques is that it provides instantaneous soft decision information. The standard techniques being the maximum *a posterior* estimates of standard demodulation techniques such as matched filter-phase comparison demodulation or the delayed maximum likelihood estimates of the Viterbi algorithm.

The hidden Markov signal model offers the flexibility of being able to model communication systems with various degrees of complexity. For example, the HMM filter can be based on knowledge of only the digital modulation layer of a communication system, as we consider in this paper, or include knowledge of the channel encoder layer as in the case when the information source has been convolutional coded. Due to the soft information the HMM filter provides it also appears well suited to communication systems in which bit interleaving is performed. The HMM filter structure presented in this paper could be used to generate preliminary soft message estimation for the purpose of channel estimation and in parallel the Viterbi algorithm conditioned on these channel estimates could be used to produce the hard decision on message estimates.

3.1 State Space Representation

Let $\mathcal{X}^0_k$ denote the space of the modulation scheme. This space is also represented by the indicator vectors X^f_k and X^θ_k. The approach taken in the previous formulation of HMM filters is obtain $\mathcal{X}^0_k$ from the Kronecker product of these indicator vectors. That is,

$$\mathcal{X}^0_k := X^f_k \otimes X^\theta_{k-1} \tag{3.1}$$

where $\otimes$ is the Kronecker product. $\mathcal{X}^0_k$ is known to be a Markov process and standard HMM filtering theory can be applied. However, $\mathcal{X}^0_k$ is $(N^2 \times 1)$ and hence the filter calculations are of order N^4, including zero operations.

If instead we define an indicator matrix $\mathcal{X}_k$ as follows

$$\mathcal{X}_k := X^f_k X^{\theta'}_{k-1} \tag{3.2}$$

then we note that

$$\begin{aligned} X^f_k &= \mathcal{X}_k \underline{1}_N \quad \text{and} \\ X^\theta_{k-1} &= \mathcal{X}'_k \underline{1}_N \end{aligned} \tag{3.3}$$

where $\underline{1}_N = (1, \ldots, 1)'$, an N-vector of ones.

Lemma 3.1 *The dynamics of* $\mathcal{X}_k$ *are given by the state equation*

$$\mathcal{X}_{k+1} = \sum_{i=1}^{N} (e_i(\mathcal{A}_i(A)(e'_i \mathcal{X}_k)')' + \mathcal{M}_{k+1} \tag{3.4}$$

where $e_i \mathcal{X}_k$ gives the ith row of matrix $\mathcal{X}_k$ written as a column vector, $\mathcal{A}_i(A)$ is a transition matrix, and $\mathcal{M}_{k+1} = M_{k+1} e'_{(i+j \ mod \ N)}$ and is a $(A, \mathcal{F}_k)$ martingale increment, in that $E[\mathcal{M}_{k+1}|\mathcal{F}_k] = 0$

Proof

$$\begin{aligned}\mathcal{X}_{k+1} &= X^f_{k+1} \, X^{\theta\,\prime}_k \\ &= (AX^f_k + M_{k+1})(D(X^f_k)X^\theta_{k-1})' \\ &= AX^f_k X^{\theta\prime}_{k-1} D(X^f_k)' + M_{k+1}(D(X^f_k)X^\theta_{k-1})'\end{aligned}$$

Now by noting that $\mathcal{X}_k$ is a function of X^f_k and X^θ_{k-1}, that is, $\mathcal{X}_k(X^f_k, X^\theta_{k-1}) = X^f_k X^{\theta\prime}_{k-1}$ and denoting $\mathcal{M}_{k+1} = M_{k+1}(D(X^f_k)X^\theta_{k-1})'$ we obtain

$$\begin{aligned}\mathcal{X}_{k+1} &= A\mathcal{X}_k(X^f_k, X^\theta_{k-1})D(X^f_k)' + \mathcal{M}_{k+1} \\ &= \left(\sum_{i=1}^N \sum_{j=1}^N A\mathcal{X}_k \langle e_j, e_i \rangle D(X^f_k)' \right) + \mathcal{M}_{k+1}\end{aligned}$$

Using the idempotent property of indicator vectors we can write

$$\mathcal{X}_{k+1} = \left(\sum_{i=1}^N \sum_{j=1}^N A e_j e'_i D(e_j) \mathcal{X}^{(j,i)}_k \right) + \mathcal{M}_{k+1}$$

where the jith element of $\mathcal{X}_k$ as $\mathcal{X}^{(j,i)}_k$. Performing the inner summation and writing as a matrix product we obtain

$$\mathcal{X}_{k+1} = \left(\sum_{i=1}^N \mathcal{A}_i (e'_i \mathcal{X}_k) \right)' + \mathcal{M}_{k+1}$$

where

$$\mathcal{A}_i = \begin{bmatrix} Ae_1 e'_i D(e_1) \\ Ae_2 e'_i D(e_2) \\ \cdots \\ Ae_N e'_i D(e_N) \end{bmatrix}$$

It follows from Lemma 2.2 that $\mathcal{M}^{(i,j)}_{k+1} = M_{k+1} e_{(i+j \ mod \ N)}$. From Lemma 2.1, M_{k+1} is a $(A, \mathcal{F}_k)$ martingale increment and hence $\mathcal{M}_{k+1}$ is a $(A, \mathcal{F}_k)$ martingale increment.

□

From Lemma 3.1 it can be seen that the dynamics of $\mathcal{X}_k$ can be viewed as N parallel independent HMMs, that is if we denote the ith row of $\mathcal{X}_{k+1}$ by $\mathcal{X}^{(i)}_{k+1}$

$$\mathcal{X}^{(i)\prime}_{k+1} = \mathcal{A}_i \mathcal{X}^{(i)\prime}_{k+1}$$

We use standard HMM filter techniques on each of the rows of $\mathcal{X}$ to obtain an estimate of $\widehat{\mathcal{X}}_k = E[\mathcal{X}_k|Y_k]$. That is for each row ,

$$\widehat{\mathcal{X}}^{(i)\prime}_{k+1} = N_k \, \mathbf{B}_{k+1|i} \, \mathcal{A}_i \widehat{\mathcal{X}}^{(i)\prime}_k \qquad (3.5)$$

where $N_k = \langle \mathbf{B}_{k+1|i} \, \mathcal{A}_i \widehat{\mathcal{X}}^{(i)\prime}_k, \underline{1} \rangle^{-1}$ where $\underline{1}$ is the vector of ones.

The estimate $\widehat{X}^f_k$ can be obtained from $\widehat{\mathcal{X}}_{k+1}$ using property (3.3).

Remarks

1. Lemma 3.1 states that the rows of $\mathcal{X}_k$ can be considered as Markov states that evolve to form rows of $\mathcal{X}_{k+1}$. This structure is hidden in the $\mathcal{X}^0_k$ formulation.

2. This filter requires order N^3 calculations per time instant.

3.2 Conditional Filters Formulation

In this subsection we present a more convenient formulation of our optimal HMM filter using coupled conditional HMM filters. In this formulation we introduce conditional HMM filters for X^f_{k+1} and X^θ_{k+1} which exploiting the interdependence between the signals.

Let $\widehat{X}^f_{k+1}$ and $\widehat{X}^\theta_{k+1}$ denote the conditional filtered normalised state estimates of X^f_{k+1} and X^θ_{k+1} respectively. That is,

$$\begin{aligned}\widehat{X}^f_{k+1} &:= E[X^f_{k+1}|Y_k] \\ \widehat{X}^\theta_{k+1} &:= E[X^\theta_{k+1}|Y_k]\end{aligned} \qquad (3.6)$$

First consider the intermediate conditional state estimate, $\widehat{X}^f_{k+1|i}$, given by

$$\widehat{X}^f_{k+1|i} := E[X^f_{k+1}|Y_k, X^\theta_k = e_i] \qquad (3.7)$$

From Bayes rule it is clear that

$$\widehat{X}^f_{k+1} = \sum_{i=1}^N \widehat{X}^f_{k+1|i} \widehat{X}^\theta_k(i) \qquad (3.8)$$

where $\widehat{X}^\theta_k(i)$ is the ith element if $\widehat{X}^\theta_k$. Note that $\widehat{X}^\theta_k(i) = P(X^\theta_k = e_i|Y_k)$.

Lemma 3.2 *The following forward recursion exists to estimate $\widehat{X}^f_{k+1|i}$*

$$\widehat{X}^f_{k+1|i} = N^{(1)}_k \mathbf{B}_{k+1|i} A \widehat{X}^f_k \qquad (3.9)$$

where $N^{(1)}_k = \langle \mathbf{B}_{k+1|i} A \widehat{X}^f_k, \underline{1} \rangle^{-1}$ is a normalising factor.

Proof Follows from assumption (2.8), Lemma 2.1 and standard HMM theory. □

Lemma 3.3 *The conditional filtered normalised state estimates* $\widehat{X}^{\theta}_{k+1}$ *is given by*

$$\widehat{X}^{\theta}_{k+1} = N^{(2)}_k \mathbf{B}_{k+1} \sum_{i=1}^{N} D(e_i) \widehat{X}^{f}_{k+1|i} \quad (3.11)$$

where $N^{(2)}_k = \langle \mathbf{B}_{k+1} \sum_{i=1}^{N} D(e_i) \widehat{X}^{f}_{k+1|i}, \underline{1} \rangle^{-1}$ *is a normalising factor.*

Proof Follows from (2.16) and Bayes rule. □

Lemma 3.4 *The conditional filtered normalised state estimates* $\widehat{X}^{f}_{k+1}$ *is given by*

$$\widehat{X}^{f}_{k+1} = N^{(3)}_k \sum_{i=1}^{N} \widehat{X}^{f}_{k+1|i} e_i' \widehat{X}^{\theta}_{k+1} \quad (3.12)$$

where $N^{(3)}_k = \langle \sum_{i=1}^{N} \widehat{X}^{f}_{k+1|i} e_i' \widehat{X}^{\theta}_{k+1}, \underline{1} \rangle^{-1}$ *is a normalising factor.*

Proof Follows from definition of $\widehat{X}^{f}_{k+1}$ and (3.8) □

Application of these last three Lemmas gives a recursive filter for estimating $\widehat{X}^{f}_{k+1}$, and hence f_{k+1}, at each time instant.

Remarks

1. The primary difference between the suboptimal approach in [5] and here appears in Lemma 3.3. The suboptimal approach would update $\widehat{X}^{\theta}_{k+1}$ as follows

$$\widehat{X}^{\theta}_{k+1} = \mathbf{B}_{k+1} \widehat{X}^{\theta}_{k}$$

rather than (3.11).

2. These filters take $O(N^3)$ calculations per time step to implement.

4 Differential Receiver

In this section we introduce the standard differential receiver and highlight the effect of channel variation on receiver performance.

The transmitted symbol is recovered from the quadrature and in-phase components of the carrier waveform. Then the message symbol is recovered from the difference between successive received carrier symbols, for the case of DPSK the message is stored in the carrier phase. Let $\bar{\theta}_k$ denote the received phase at time k. The received phase $\bar{\theta}_k$ is related to the transmitted phase θ_k as follows,

$$\bar{\theta}_k = (\theta_k + \theta^c_k + \phi(w_k))_{2\pi} \quad (4.1)$$

where θ^c_k is the channel phase and time k and $\phi(w_k)$ is the phase of w_k. The difference between successive received phase signals is

$$(\bar{\theta}_k - \bar{\theta}_{k-1})_{2\pi} = (f_k + \Delta^c_k + \phi(w_k - w_{k-1}))_{2\pi} \quad (4.2)$$

where $\Delta^c_k = \theta^c_k - \theta^c_{k-1}$.

When no information is available about the channel phase then the standard estimate of the message symbol, see Figure 1, is

$$\widehat{f}_k = (\bar{\theta}_k - \bar{\theta}_{k-1})_{2\pi} \quad (4.3)$$

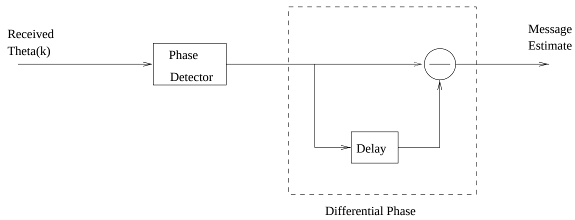

Figure 1: Block diagram of M-ary differential coherent receiver

It is clear from (4.2) that the quality of the estimate (4.3) will degrade as the rate of change of the channel, $\sim \Delta^c_k$, increases.

However, if estimates of the channel phase, $\widehat{\theta}^c_k$, are available then these can be used to improve the performance of the receiver by including these estimates into the received structure. Consider the message estimate, $\bar{f}_k$ which includes channel estimates, see also Figure 2.

$$\bar{f}_k = (\bar{\theta}_k - \bar{\theta}_{k-1} - \widehat{\Delta}^c_k)_{2\pi} \quad (4.4)$$

where $\widehat{\Delta}^c_k = \widehat{\theta}^c_k - \widehat{\theta}^c_{k-1}$,

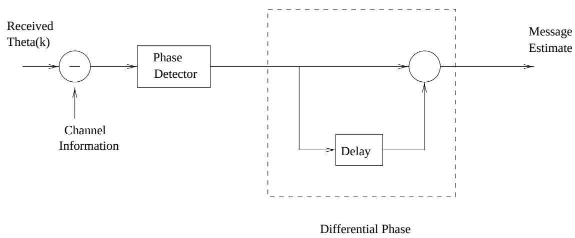

Figure 2: Block diagram of differential receiver using channel estimates

In this receiver the message estimate are related to the message as follows:

$$\bar{f}_k = (f_k + \Delta^c_k - \widehat{\Delta}^c_k + \phi(w_k - w_{k-1}))_{2\pi}. \quad (4.5)$$

comparing with (4.2) it is clear that the performance of the receiver that uses the channel estimates will be better when $|\Delta_k^c - \widehat{\Delta}_k^c| < |\Delta_k^c|$

This is the motivation for the following section.

5 Decision Feedback Channel Estimation

To produce channel estimates we propose the use of a coupled conditional hidden Markov model (HMM) filter and conditional Kalman filter (KF). The HMM filter provides soft decision information about the received message symbol. In this paper we consider a decision feedback structure in which the Kalman filter estimates are based on the maximum *a posteriori* (MAP) estimate of the message symbol from the HMM filter. Similarly, the HMM filter is conditioned on this channel estimate.

More complicated decision feedback equalisation structures are possible using banks of Kalman filters conditioned according to the soft estimate information (as suggested in [14]), or Kalman filters with more complicated structures are possible but not considered here.

5.1 Conditional HMM filter

Let $\widehat{\mathcal{X}}_{k|\mathcal{C}_k}$ denote the conditional filtered state estimate of $\mathcal{X}_k$, where $\mathcal{C}_k = \{\bar{g}_o, \dots, \bar{g}_k\}$. By definition,

$$\widehat{\mathcal{X}}_{k|\mathcal{C}_k} = E[\mathcal{X}_k|\mathcal{Y}_k, \mathcal{C}_k] \tag{5.1}$$

Lemma 5.1 *The estimates $\widehat{\mathcal{X}}_{k|\mathcal{C}_k}$ can be founding using the forward recursion*

$$\widehat{\mathcal{X}}^{(i)'}_{k+1|\mathcal{C}_k} = N_k \, \mathbf{B}_{k+1|i,\mathcal{C}_k} \, \mathcal{A}_i \widehat{\mathcal{X}}^{(i)'}_{k|\mathcal{C}_k} \tag{5.2}$$

where $N_k = \langle \mathbf{B}_{k+1|i,\mathcal{C}_k} \quad \mathcal{A}_i \widehat{\mathcal{X}}^{(i)'}_{k|\mathcal{C}_k}, \underline{1}\rangle^{-1}$ *is a normalising factor for each row and* $\mathbf{B}_{k|i,\mathcal{C}_k} = diag(b_{k|i,\bar{g}_k}(e_1), \dots, b_{k|i,\bar{g}_k}(e_N))$, *and* $b_{k|i,\bar{g}_k}(e_j) = P[Y_k | X^f_{k+1} = e_j, X^\theta_k = e_i, g_k = \bar{g}_k]$,

$$\mathcal{A}_i = \begin{bmatrix} Ae_1 e_i' D(e_1) \\ Ae_2 e_i' D(e_2) \\ \dots \\ Ae_N e_i' D(e_N) \end{bmatrix}$$

and $\mathcal{M}_{k+1} = M_{k+1} e_{(i+j \bmod N)}$

Proof See [11]

5.2 Conditional KF Channel Estimate

The observation process (2.6) is bi-linear in the channel parameter, g_k, and the message symbol, m_k. Note that $\theta_k^c = \phi(g_k)$. In this subsection a conditional KF is proposed for estimation of the channel parameter given the message symbol.

If we assume the channel dynamics are given by the following linear time invariant stochastic system

$$\begin{aligned} g_{k+1} &= F g_k + v_k \qquad in \mathbb{C} \\ y_k &= H(z_f, X_k^\theta) g_k + w_k \in \mathbb{C} \end{aligned} \tag{5.3}$$

for some known F and where $v_k = N[0, Q_k]$. Then the KF equation for estimating g_k given the message estimate X_k^θ is,

$$\begin{aligned} \widehat{g}_k &= F\widehat{g}_{k-1} + K_k[y_k - \widehat{g}_{k-1} m_k] \\ K_k &= \Sigma_{k|k-1} \bar{H}_k (\bar{H}_k' \Sigma_{k|k-1} \bar{H}_k + R_k)^{-1} \\ \Sigma_{k|k} &= (I - K_k \bar{H}_k') \Sigma_{k|k-1} \\ \Sigma_{k+1|k} &= F \Sigma_{k|k} F' + Q_k \end{aligned} \tag{5.4}$$

where we have $\bar{H}_k := H(z_f, X_k^\theta)$. We note that

$$\begin{aligned} \Sigma_{k|k}^{-1} &= \Sigma_{k|k-1}^{-1} + H(z_f, X_k^\theta) R_k^{-1} H(z_f, X_k^\theta)' \\ &= \Sigma_{k|k-1}^{-1} + \begin{bmatrix} \sigma_w^R L_k(X_k^\theta) & 0 \\ 0 & \sigma_w^I L_k(X_k^\theta) \end{bmatrix} \end{aligned} \tag{5.5}$$

where

$$\begin{aligned} L_k(X_k^\theta) &= (z_f^R)' X_k^\theta X_k^{\theta\prime} z_f^R + (z_f^I)' X_k^\theta X_k^{\theta\prime} z_f^I \\ &= (z_f^R)' diag(X_k^\theta) z_f^R \\ &\quad + (z_f^I)' diag(X_k^\theta) z_f^I, \end{aligned} \tag{5.6}$$

and where $diag(X_k^\theta)$ is the diagonal matrix with X_k^θ on its diagonal.

In equation (5.6) we have exploited the idempotent nature of Markov chains by replacing the product $X_k^\theta X_k^{\theta\prime}$ by $diag(X_k^\theta)$. This is a key contribution of this paper. This forces the Kalman filter to incorporate more *a priori* knowledge. In the following section we will replace X_k^θ by the conditional mean estimate $\widehat{X}_k^\theta$. It is important to including this extra structure here, that $X_k^\theta X_k^{\theta\prime}$ equals $diag(X_k^\theta)$ before replacing X_k^θ by $\widehat{X}_k^\theta$ because $\widehat{X}_k^\theta \widehat{X}_k^{\theta\prime} \neq diag(\widehat{X}_k^\theta)$. Other formulations of the Kalman filter for this problem, for example [10] and [8, Page 656], do not exploit this structure.

To reduce computational effort we note that if $\Sigma_{0|0}^{-1}$ is diagonal, F is diagonal and Q_k is diagonal then $\Sigma_{k|k}^{-1}$ and $\Sigma_{k+1|k}^{-1}$ will be diagonal and the Kalman filter equations can be simplified. In particular, the recursion for K_k can be rewritten in terms of $\Sigma^{-1}_{k+1|k}$. The simplified Kalman filter equations are as follows

$$\begin{aligned}
\widehat{g}_k &= F\widehat{g}_{k-1} + K_k[y_k - \widehat{g}_{k-1}m_k] \\
K_k &= F\Sigma^{-1}_{k+1|k}\bar{H}_k R_k^{-1} \\
\Sigma^{-1}_{k|k} &= \Sigma^{-1}_{k-1|k-1} + \begin{bmatrix} \sigma_w^R L_k(X_k^\theta) & 0 \\ 0 & \sigma_w^I L_k(X_k^\theta) \end{bmatrix} \\
\Sigma_{k+1|k} &= F\Sigma_{k|k}F' + Q_k \qquad (5.8)
\end{aligned}$$

Estimates for the channel phase, $\widehat{\theta}^c_k$, are given by $\widehat{\theta}^c_k = \phi(\widehat{g}_k)$.

Remarks

- Even when the channel dynamics are not given exactly in the form (5.3) they can often be approximated by (5.3).
- F is typically approximated to be of the form fI for some $f < 1$ which allows the use of the (5.8) equations and reduces computational effort. A forgetting factor can be introduced into line 4 of recursions (5.8) to allow for model variations and modelling errors.
- The Kalman filter requires $O(N^2)$ calculations per time instant. The simplified Kalman filter requires $O(N)$ calculations per time instant.

5.3 The Complete Algorithm

To allow simultaneous estimation of the channel, $\widehat{g}_k$, and message, $\widehat{m}_k$ or $\widehat{X}_k$, the conditional filters (5.2) and (5.4) are coupled together, see Figure 3.

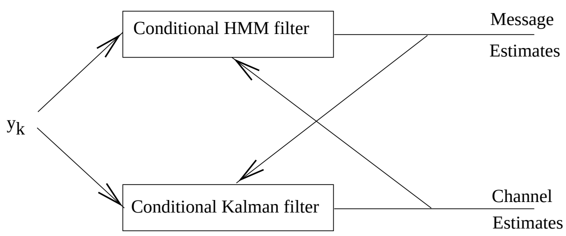

Figure 3: Block diagram of coupled filter structure

Remarks

- The coupled algorithm requires $O(N^3)$ calculations per iteration.
- Need to boost noise in each model to "model" estimation errors as is suggested in [3].
- The decision feedback structure does not suffer from the usual error propagation effects. While message estimation errors may indeed introduce constant phase errors in the channel phase estimate, these absolute phase errors do not degrade the performance of the difference receivers, as is well known [7]. In the usual way, a message error will be followed by at least one more message error.

6 Simulations

6.1 Differential decoding with and without Decision Feedback

In our simulations we investigate the gain in performing decision feedback on a un-coded transmission system. To evaluate the performance of the receiver in a likely environment the transmitted message sequence is actually generated by convolutional coded a i.i.d. message sequence using a rate 2/3 convolutional code. However, the HMM-KF filter does not use this information to produce estimates of the i.i.d. message sequence and only estimates the encoded sequences. This environment represents the case when the HMM-KF filter is used to produce channel estimates that are passed to a Viterbi algorithm to perform optimal decoding of the i.i.d. message sequence.

The encoded signal is transmitted over a transmission channel that is time varying and unknown. The variations are deterministic, with the phase and amplitude varying sinusoidally from 50% to 150% of a nominal value. The optimal HMM decoder is used to demodulate the encoded symbols, but the original i.i.d. message sequence is not decoded from the demodulated symbols. In a realistic application the Viterbi algorithm would be used in parallel for optimal decoding.

Figure 4 shows an improvement in bit error performance in terms of the encoded symbols due to decision feedback. Likewise, the channel estimates are seen to improve the bit error rate in terms of the i.i.d. message sequence This curve demonstrate that there is over half a dB gain to achieve $P_e = 10^{-3}$.

7 Conclusion

In this paper we have investigated optimal hidden Markov model (HMM) filtering and decision feedback equalisation of differentially encoded transmission signals. We have presented the optimal (HMM) filter for demodulation of differentially encoded signals. We then proposed a decision feedback structure that coupled together a conditional HMM filter to a conditional Kalman filter. A key point being that the Kalman filter exploited the idempotent nature of Markov chains include more *a priori* structural information. We also present simplified Kalman filters for a particular channel assumptions. The HMM filter requires $O(N^3)$ calculations per time instant, where N is the number of message symbols. The Kalman filter requires $O(N^2)$ calculations per time instant which is reduced to $O(N)$ under the channel assumptions.

Simulation studies demonstrated a half dB gain to achieved a bit error rate of 1×10^{-3}.

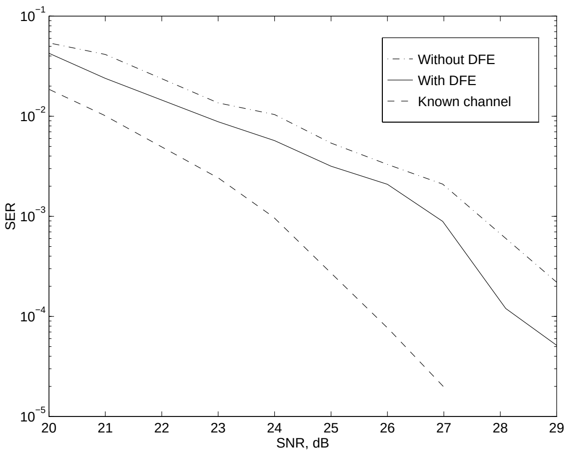

Figure 4: Improvement in BER performance due to Decision Feedback Equalisation.

References

[1] R.J. Elliott, L. Aggoun and J.B. Moore, *Hidden Markov Models, Estimation and Control*, Springer-Verlag, 1995.

[2] I.B. Collings and J.B. Moore, "Adaptive HMM Filters for Signals in Noisy Fading Channels", Proc. of Int. Conf. on Acoustics, Speech & Signal Processing ICASSP94, Adelaide, Australia, Vol. 3, pp 305-308, 1994.

[3] J.J. Ford and J.B. Moore, "On Adaptive HMM state estimation", to appear IEEE Trans. on Signal Processing.

[4] Jason Ford and John B. Moore, "Adaptive Estimation of HMM Transition Probabilities", to appear IEEE Transactions on Signal Processing.

[5] I.B. Collings, *Hidden Markov Model Signal Processing and Control*, PhD thesis, Australian National University, January 1995.

[6] L.B. White, "Adaptive Tracking of Frequency Modulated Signals using Hidden Markov Models", Presented at the CRASys/IEEE Workshop on Hidden Markov Models and Tracking Wirrinae, South Australia, February 1992.

[7] R.E. Ziemer and R.L. Peterson, *Introduction to Digital Communication*, Maxwell Macmillan, 1992.

[8] John G. Proakis, *Digital Communications*, 3rd ed., McGraw-Hill, New York, 1995.

[9] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm", IEEE Trans. on Information Theory, Vol. IT-13, No. 2,, pp 260-269, 1967.

[10] I.B. Collings and J.B. Moore, "An Adaptive Hidden Markov Model Approach to FM and M-ary DPSK Demodulation in Noisy Fading Channels", Signal Processing, Vol 47, No. 1, pp71-84, November 1995.

[11] J.J. Ford, E. Lane and J.B. Moore, "Optimal HMM processing of Differential Phase Modulated Communication Systems", Proc. ISSPA, Gold Coast, Australia, August 1996.

[12] R. Steele, *Mobile Radio Communications*, Pentech Press, London, 1992.

[13] R.W. Lucky, "Automatic Equalization of Digital Communications" Bell Syst. Tech. J., vol 44. pp. 547-588, April 1965.

[14] K. Giridhar, J.J. Shynk and R.A. Iltis, "A Modified Bayesian Algorithm with Decision Feedback for Blind Adaptive Equalization", IFAC, adaptive Systems in Control and Signal Processing, Grenoble, France, pp 511-516, 1992.

[15] Y. Liu and S.D. Blostein, "Identification of frequency Non-selective Fading Channels Using Decision Feedback and Adaptive Linear Prediction", IEEE Transactions on Communications, April, 1995.