Optimal Channel Equalization
for MIMO Time-Varying Channels

Ehab F. Badran   and   Guoxiang Gu

Abstract— We consider optimal channel equalization for MIMO (multi-input/multi-output) time-varying channels in the sense of MMSE (minimum mean-squared-error), where the observation noise can be non-stationary. We show that all ZF (zero-forcing) receivers can be parameterized in an affine form which eliminates completely the ISI (inter-symbol-interference), and optimal channel equalizers can be designed through minimization of the MSE (mean-squared-error) between the detected signals and the transmitted signals, among all ZF receivers. We demonstrate that the optimal channel equalizer is a modified Kalman filter, and show that under the AWGN (additive white Gaussian noise) assumption, the proposed optimal channel equalizer minimizes the BER (bit error rate) among all possible ZF receivers. Our results are applicable to optimal channel equalization for DWMT (discrete wavelet multitone), multirate trans-multiplexers, OFDM (orthogonal frequency division multiplexing), and DS (direct sequence) CDMA (code division multiple access) wireless data communication systems. A design algorithm for optimal channel equalization is developed, and several simulation examples are worked out to illustrate the proposed design algorithm.

Keywords— Channel equalization, Kalman filtering, Time-varying systems.

I. INTRODUCTION

Inter-symbol interference (ISI) has been one of the major obstacles for performance improvement in wireless data communications. Channel equalization is a common and effective technique to suppress ISI. For MIMO channels, often precoders need be employed to provide the required redundancies in order for causal and stable equalizers to exist [6], [10]. Data communication systems such as digital subscriber loops, DWMT (discrete wavelet multitone), FDMA (frequency division multiple access), and DS-CDMA systems admit MIMO channels, and have a common structure in that precoders are designed to provide redundancies, which enable not only error control coding, but also channel equalization such that ISI can be eliminated completely. Many researchers [6], [9], [10] have contributed to this important problem area, focusing on the design of FIR receivers to achieve MMSE. It should be clear that MMSE receivers in the existing literature are actually suboptimal by the fact that optimization is carried out over all FIR filters of a fixed order. Their optimality over all possible linear, or nonlinear filters of possibly infinite order can not be claimed thus far.

In this paper we consider channel equalization for multi-input/multi-output time-varying channels in the presence of possibly non-stationary noises. Wireless channels often experience selective fadings, implying that actual radio channels are likely to be time-varying. It should be clear that channel estimation algorithms which track fast fading channels exist [4], [5], and most other estimation algorithms [8], [12] can also be adapted to estimate time-varying channels by introducing weighting factors with less weights given to old measurement data. On the other hand the noise variance is dependent on the location of the cellular user, which can be pre-determined via the experimental method. Thus we will assume that the channel model and noise covariance are given at each sampling time t. We will study design of optimal receivers that achieve not only PR or zero-forcing, but also MMSE among all possible linear and time-varying filters of arbitrary orders. We will show that for redundant MIMO channels, the zero-forcing receivers are equivalent to causal and stable left inverses of some "tall" matrix of linear dynamic systems, consisting of the channel and precoder. Such left inverses are not unique. Any zero-forcing receiver filterbank accomplishes the goal of channel equalization, and eliminates completely ISI. We will first parameterize all zero-forcing channel equalizers in an affine form, and then seek one of them to minimize the MSE caused by possible non-stationary noises at the receivers, thereby converting the constrained optimization into unconstrained optimization for receivers design. It will be shown that the design of optimal channel equalizers is equivalent to the design of optimal state estimators for some augmented system subject to the same noise processes. Hence the celebrated Kalman filtering can be used successfully to design the optimal detectors among all channel equalizers, which are truly optimal over all linear and time-varying receiver filters of arbitrary order. The effectiveness of our proposed optimal channel equalization method is illustrated by illustrative examples.

II. MIMO TIME-VARYING CHANNELS AND ZERO FORCING CONDITION

We consider the discrete-time MIMO time-varying composite channel whose input-output relation is described by

$$y(t) = \phi(t, \tau) \star \mathbf{s}(t) + \eta(t), \tag{1}$$

with $\phi(t, \tau) = h(t, \tau) \star f(t, \tau) = \sum_{k=\tau}^{t} h(t, t - k) f(k, \tau)$, where the received signal $y(t)$, and the observation noise $\eta(t)$ have dimension $P \times 1$, and the transmitted data symbol $\mathbf{s}(t)$ has dimension $M \times 1$. The impulse response at time t of the composite channel $\phi(t, \tau)$ has size $P \times M$, which is the

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:2, No:11, 2008

convolution, denoted by $\star$, of the channel impulse response $\underline{h}(t,\tau)$, and precoder or transmitter impulse response $\underline{f}(t,\tau)$. It is assumed that $\underline{h}(t,\tau)$ has size $P \times P$, and $\underline{f}(t,\tau)$ has size $P \times M$, where $P > M$. Thus the precoder $\underline{f}(t,\tau)$ provides redundancies, which enable the zero-forcing condition, and allow equalization. Our objective is to synthesize a detector, or decoder $\underline{g}(t,\tau)$ such that it minimizes

$$
\begin{aligned}
J(\underline{g})(t) &= E[\underline{e}^*(t)\underline{e}(t)] = \text{trace}\{E[\underline{e}(t)\underline{e}^*(t)]\} \quad (2) \\
\underline{e}(t) &= \underline{g}(t,\tau) \star \underline{\eta}(t),
\end{aligned}
$$

with $\underline{\eta}$ possibly non-stationary, subject to the zero-forcing or PR condition

$$
\underline{g}(t,\tau) \star \underline{\phi}(t,\tau) = \delta(t-\tau)I_M = \begin{cases} I_M, & t=\tau, \\ 0, & t \neq \tau, \end{cases} \quad (3)
$$

where $(\cdot)^*$ denotes conjugate transpose.

Dynamic MIMO discrete-time systems such as the multi-channel model (multi-transmitter/multi-receiver model), multirate filterbank model, versatile multirate filterbank transceiver model, and multirate transmultiplexers have a common mathematical input-output relation equivalent to (1).

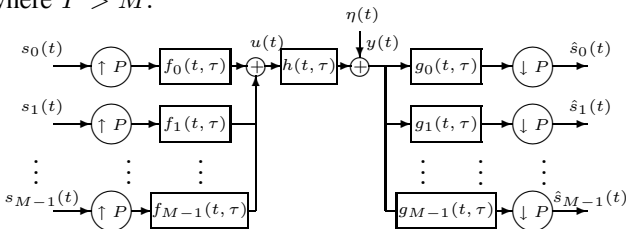Consider the multirate transmultiplexers as in Figure 1 where $P > M$.

Fig. 1. multirate transmultiplexer model.

The channel is represented by $h(t,\tau)$ which is a scalar impulse response, and M transmitter filters $\{f_i(t,\tau)\}_{i=0}^{M-1}$ are employed to precode the transmitted data signals, and M receiver filters $\{g_i(t,\tau)\}_{i=0}^{M-1}$ to decode, or detect the transmitted data signals. It is indicated in [13] that the CDMA wireless data communication system for the forward channel is equivalent to the model in Figure 1, where $f_i(t,\tau)$ consists of the PN codes of the ith user for $0 \leq i \leq M-1$. The fact that $P > M$ provides the redundancy. It can be shown that with $\mathbf{s}(t) = \begin{bmatrix} s_0(t) & s_1(t) & \cdots & s_{M-1}(t) \end{bmatrix}^T$, and $\underline{y}(t) = \begin{bmatrix} y(tP) & y(tP+1) & \cdots & y(tP+P-1) \end{bmatrix}^T$, the model (1) holds with $\underline{\phi}(t,\tau) = \underline{h}(t,\tau) \star \underline{f}(t,\tau)$ for some $\underline{h}(t,\tau)$, and $\underline{f}(t,\tau)$, which can be obtained based on Figure 1.

For the design of the optimal channel equalizer in the MMSE sense, we first need the existence condition for the receiver $\underline{g}(t,\tau)$ such that the zero forcing condition (3) holds. For this purpose, we consider the noise-free case for the rest of the section. It is convenient to introduce the state-space realization for the composite channel model $\underline{\phi}(t,\tau)$:

$$
\begin{aligned}
x(t+1) &= A(t)x(t) + B(t)\mathbf{s}(t), \quad x(0) = 0, \quad (4) \\
\underline{y}(t) &= C(t)x(t) + D(t)\mathbf{s}(t),
\end{aligned}
$$

where $x(t)$ is the n-dimensional state vector at time t. Thus $A(t)$ has size $n \times n$, $B(t)$ has size $n \times M$, $C(t)$ has size $P \times n$, and $D(t)$ has size $P \times M$ for all integer valued time t.

Because of the requirement on stability, the following stability notion is needed.

Definition 2.1: The time-varying system described by the state space model (4) is exponentially stable, or simply stable, if there exists an $N_0 > 0$ such that

$$
\rho\left(\prod_{t=t_0}^{N+t_0} A(t)\right) \leq \alpha\beta^N, \quad \alpha > 0, \quad 0 < \beta < 1, \quad (5)
$$

for all $N \geq N_0 > t_0 \geq 0$, where $\rho(\cdot)$ denotes the spectral radius, and α, β are independent of N.

Thus, the exponential stability can be ensured only by having $\rho(A(t)) < 1 \; \forall \; t$, which if holds true, then there exist $\alpha, \; \beta, \; N_0$ such that (5) is true.

Throughout the paper, exponential stability is used in place of stability. For time-invariant systems, exponential stability reduces to $\rho(A) < 1$. Because stability of the state-space system (4) depends only on $A(t)$, we will say that $A(t)$ is exponentially stable, if the condition in Definition 2.1 is true.

Remark 2.2: We will assume exponential stability for the channel model (4), and

$$
\text{rank}[\ D(t)\ ] = M \quad \forall\ t. \quad (6)
$$

If the condition (6) is violated for some t, then we assume the existence of the factorization:

$$
\underline{\phi}(t,\tau) = \text{diag}(q^{-\ell_0}, q^{-\ell_1}, \cdots, q^{-\ell_{P-1}}) \star \underline{\phi}_0(t,\tau),
$$

for some positive integers $\{\ell_i\}_{i=0}^{P-1}$, such that the state-space realization of $\underline{\phi}_0(t,\tau)$ has its $D(t)$-term full column rank as in (6). In this case, the receiver can be assumed of the form

$$
\underline{g}(t,\tau) = \underline{g}_0(t,\tau)\text{diag}(q^{-(\ell_{\max}-\ell_0)}, \cdots, q^{-(\ell_{\max}-\ell_{P-1})})
$$

with $\ell_{\max} = \max_{0 \leq k < P} \ell_k$. It follows that $\underline{g}(t,\tau) \star \underline{\phi}(t,\tau) = \underline{g}_0(t,\tau) \star \underline{\phi}_0(t - \ell_{\max}, \tau)$, and $J(\underline{g}) = J(\underline{g}_0)$. Hence there is no loss of generality to assume (6). ∎

Since $P > M$, $D(t)$ having full column rank implies the existence of $D_\perp(t)$ such that $D_a(t) = \begin{bmatrix} D(t) & D_\perp(t) \end{bmatrix}$ is square and nonsingular. A particular $D^+(t)$, left inverse of $D(t)$, and $D_\perp(t)$ can be chosen as $D^+(t)$ and $D_\perp(t)$ exist).

$$
\begin{aligned}
D^+(t) &= (D^*(t)D(t))^{-1}D^*(t), \quad (7) \\
D_\perp(t)D_\perp^*(t) &= I_P - D(t)(D^*(t)D(t))^{-1}D^*(t)
\end{aligned}
$$

where $D_\perp(t)$ is the minimum rank Cholesky factor.

It follows that

$$
\begin{bmatrix} D^+(t) \\ D_\perp^+(t) \end{bmatrix} \begin{bmatrix} D(t) & D_\perp(t) \end{bmatrix} = \begin{bmatrix} I_M & 0 \\ 0 & I_{P-M} \end{bmatrix}, \quad (8)
$$

$$
D_\perp^+(t) = D_\perp^*(t).
$$

There are many choices for $D^+(t)$, and $D_\perp(t)$ such that the above holds. Those in (7) are the simplest, and most commonly used.

Lemma 2.3: Let the state-space realization for $\underline{\phi}(t,\tau)$ be given as in (4) where $D(t)$ satisfies the condition (6). Then a causal and stable $\underline{g}(t,\tau)$ exists and satisfies the zero-forcing condition (3), if and only if there exists a state estimation gain $L(t)$ such that

$$
A_L(t) = A(t) - B(t)D^+(t)C(t) + L(t)D_\perp^*(t)C(t) \quad (9)
$$

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:2, No:11, 2008

is exponentially stable. Such a state estimator gain $L(t)$ will be called stabilizing.

Proof: If $A_L(t)$ is exponentially stable for some $L(t)$, then it is claimed that $\underline{g}(t;\tau) = \underline{\phi}^+(t;\tau)$, described by the state-space model:

$$\begin{aligned}
\tilde{x}(t+1) &= A_L(t)\tilde{x}(t) + B_L(t)\underline{y}(t), \; \tilde{x}(0) = 0, \quad (10)\\
B_L(t) &= B(t)D^+(t) - L(t)D_\perp^*(t),\\
\hat{\mathbf{s}}(t) &= -D^+(t)C(t)\tilde{x}(t) + D^+(t)\underline{y}(t).
\end{aligned}$$

is a causal and stable left inverse of $\underline{\phi}(t;\tau)$, and thus the ZF condition (3) holds. Indeed taking the difference between the state-space models as in (4) and (10) gives the error system

$$\tilde{x}_s(t+1) = A_L\tilde{x}_s(t), \quad \hat{\mathbf{s}}(t) = -D^+(t)C(t)\tilde{x}_s(t) + \mathbf{s}(t)$$

where $\tilde{x}_s(t) = \tilde{x}(t) - x(t)$, in light of (8), and expression of $A_L(t)$, $B_L(t)$, and $\underline{y}(t)$. Hence by the initial condition of $\tilde{x}(0) = x(0) = 0$, we have that $\tilde{x}_s(0) = 0$ and $\tilde{x}_s(t) = 0 \; \forall\, t$. It follows that $\hat{\mathbf{s}}(t) = \mathbf{s}(t)$. Conversely, assume that a stable and causal $\underline{g}(t;\tau)$ exists such that the ZF condition (3) is true. Since the state-space model (4) for $\underline{\phi}(t;\tau)$ has an equivalent form

$$\begin{aligned}
x(t+1) &= A(t)x(t) + B_L(t)\mathbf{d}(t), \quad x(0) = 0,\\
\underline{y}(t) &= C(t)x(t) + \mathbf{d}(t), \quad (11)
\end{aligned}$$

for any $L(t)$ due to (8), where $\mathbf{d}(t) = D(t)\mathbf{s}(t)$, and $B_L(t)$ is as in (10),

$$\underline{y}(t) = \underline{\phi}(t;\tau) \star \mathbf{s}(t) = \underline{\phi}_1(t;\tau) \star \mathbf{d}(t)$$

for some $\underline{\phi}_1(t;\tau)$. The existence of stable and causal $\underline{g}(t;\tau)$ such that (3) holds then implies that

$$\begin{aligned}
\hat{\mathbf{s}}(t) &= \underline{g}(t;\tau) \star \left[\underline{\phi}(t;\tau) \star \mathbf{s}(t)\right]\\
&= \left[\underline{g}(t;\tau) \star \underline{\phi}_1(t;\tau)\right] \star \mathbf{d}(t) = \mathbf{s}(t).
\end{aligned}$$

Because $\mathbf{d}(t) = D(t)\mathbf{s}(t)$, we have that $\underline{g}(t;\tau) \star \underline{\phi}_1(t;\tau) = D^+(t)$, a left inverse of $D(t)$, which in turn implies that $\underline{\phi}_1(t;\tau)$ as described in (11) has a stable and causal left inverse. Noticing that $\underline{\phi}_1(t;\tau)$, described by the state-space model as in (11), has an equal number of inputs and outputs, and the direct transmission from $\mathbf{d}(t)$ to $\underline{y}(t)$ is identity, we conclude that it has a unique inverse, given by

$$\tilde{x}(t+1) = A_L(t)\tilde{x}(t) + B_L(t)\underline{y}(t), \mathbf{d}(t) = -C(t)\tilde{x}(t) + \underline{y}(t).$$

Causality and stability of $\underline{g}(t;\tau)$ then concludes that

$$\begin{aligned}
A_L(t) &= A(t) - B_L(t)C(t)\\
&= A(t) - B(t)D^+(t)C(t) + L(t)D_\perp^*(t)C(t)
\end{aligned}$$

is exponentially stable. ∎

For the case of time-invariant models, the equivalent zero-forcing condition in Lemma 2.3 reduces to the strictly minimum phase condition [2], [3]:

$$\text{rank}\left\{\begin{bmatrix} zI_n - A & B \\ C & D \end{bmatrix}\right\} = n + M \;\; \forall\, |z| \geq 1. \quad (12)$$

In the remainder of this section we establish a parallel result to (8) for the composite channel model $\underline{\phi}(t;\tau)$. Let $\mathbf{v}(t)$ be of

length $(P-M) \; \forall\, t$. We define a time-varying system $\underline{\phi}_\perp(t,\tau)$ by the state-space model as

$$\begin{aligned}
x_\perp(t+1) &= A(t)x_\perp(t) + B_L(t)D_\perp(t)\mathbf{v}(t), \quad (13)\\
\underline{w}(t) &= C(t)x_\perp(t) + D_\perp(t)\mathbf{v}(t), \quad x_\perp(0) = 0.
\end{aligned}$$

It is noted that $B(t) = B_L(t)D(t)$, by the relation in (8). Thus the state-space model for $\underline{\phi}_\perp(t,\tau)$ is the same as for $\underline{\phi}(t,\tau)$ except that $D_\perp(t)$ replaces $D(t)$. Hence by Lemma 2.3, a specific left inverse $\underline{\phi}_\perp^+(t,\tau)$ to $\underline{\phi}_\perp(t,\tau)$ is described by the state-space model:

$$\begin{aligned}
\tilde{x}_\perp(t+1) &= A_L(t)\tilde{x}_\perp(t) + B_L(t)\underline{w}(t), \quad (14)\\
\mathbf{v}(t) &= -D_\perp^*(t)C(t)\tilde{x}_\perp(t) + D_\perp^*(t)\underline{w}(t), \quad \tilde{x}(0) = 0.
\end{aligned}$$

Let $x_a(t) = x(t) + x_\perp(t)$ and $\underline{y}_a(t) = \underline{y}(t) + \underline{w}(t)$ for the state space models in (4) and (13). Then it can be verified that the augmented system $\underline{\phi}_a(t,\tau) = \begin{bmatrix} \underline{\phi}(t,\tau) & \underline{\phi}_\perp(t,\tau) \end{bmatrix}$ is square, and has the state-space model

$$x_a(t+1) = A(t)x_a(t) + B_L(t)D_a(t)\begin{bmatrix} \mathbf{s}(t) \\ \mathbf{v}(t) \end{bmatrix}, \quad (15)$$

$$y_a(t) = C(t)x_a(t) + D_a(t)\begin{bmatrix} \mathbf{s}(t) \\ \mathbf{v}(t) \end{bmatrix}, \; x_a(0) = 0.$$

Lemma 2.4: Consider the state-space model as in (4), which satisfies the condition (6), and that there exists a state estimation gain $L(t)$ such that $A_L(t)$ as in (9) is exponentially stable. Then for $\underline{\phi}_a(t,\tau)$ as in (15), there exists a unique $\underline{\phi}_a^{-1}(t,\tau) = \begin{bmatrix} \underline{\phi}^+(t,\tau) \\ \underline{\phi}_\perp^+(t,\tau) \end{bmatrix}$, which is a causal and stable inverse for the square augmented system $\underline{\phi}_a(t,\tau)$, and described by the state-space model:

$$\begin{aligned}
\tilde{x}_a(t+1) &= A_L(t)\tilde{x}_a(t) + B_L(t)y_a(t), \; \tilde{x}_a(0) = 0,\\
\begin{bmatrix} \hat{\mathbf{s}}(t) \\ \hat{\mathbf{v}}(t) \end{bmatrix} &= \begin{bmatrix} D^+(t) \\ D_\perp^*(t) \end{bmatrix}[C(t)\tilde{x}_a(t) + y_a(t)]. \quad (16)
\end{aligned}$$

That is, for each $D^+(t)$, $D_\perp^*(t)$ satisfying (8), and $L(t)$ stabilizing, there exists a unique causal and stable inverse $\underline{\phi}_a^{-1}(t,\tau)$ for $\underline{\phi}_a(t,\tau)$ such that

$$\begin{bmatrix} \underline{\phi}^+(t,\tau) \\ \underline{\phi}_\perp^+(t,\tau) \end{bmatrix} \star \underline{\phi}_a(t,\tau) = \delta(t-\tau)\begin{bmatrix} I_M & 0 \\ 0 & I_{P-M} \end{bmatrix}. \quad (17)$$

Proof: by using the state-space models as in (15) and (16) to construct the composite system from $\begin{bmatrix} \mathbf{s}(t) \\ \mathbf{v}(t) \end{bmatrix}$ to $\begin{bmatrix} \hat{\mathbf{s}}(t) \\ \hat{\mathbf{v}}(t) \end{bmatrix}$ (which is the output of $\begin{bmatrix} \underline{\phi}^+(t,\tau) \\ \underline{\phi}_\perp^+(t,\tau) \end{bmatrix}$), and applying the similarity transformation $\begin{bmatrix} x_{as}(t) \\ \tilde{x}_{as}(t) \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -I_n & I_n \end{bmatrix}\begin{bmatrix} x_a(t) \\ \tilde{x}_a(t) \end{bmatrix}$ on the composite system. It follows that $\begin{bmatrix} \hat{\mathbf{s}}(t) \\ \hat{\mathbf{v}}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{s}(t) \\ \mathbf{v}(t) \end{bmatrix}$ with the initial condition of $\tilde{x}_a(0) = x_a(0) = 0$. ∎

III. MMSE CHANNEL EQUALIZATION

In this section, we study the design of the decoder, or detector $g(t,\tau)$ such that it not only achieves the zero-forcing condition (3), but also minimizes the MSE $J(\underline{g})$ as defined in (2). Our strategy is to parameterize all the channel

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:2, No:11, 2008

equalizers $\underline{g}(t,\tau)$, and then search for the optimal one over the parameterized set. The results in the previous section are useful.

Theorem 3.1: Under the same hypotheses as in Lemma 2.4, the set of all causal and stable left inverses of $\underline{\phi}(t,\tau)$ is given by

$$\mathbf{\Phi}^+ := \{\underline{\phi}^+(t,\tau) + \gamma(t,\tau) \star \underline{\phi}_\perp^+(t,\tau) \qquad (18)$$
$$\gamma(t,\tau) \text{ is causal and stable}\},$$

where $\underline{\phi}^+(t,\tau)$, and $\underline{\phi}_\perp^+(t,\tau)$ are described by the state-space models (10), and (14), respectively.

Proof: It is noted that any $\underline{\phi}^{\text{inv}}(t,\tau) \in \mathbf{\Phi}^+$ is causal and stable, by causality and stability of $\underline{\phi}^+(t,\tau)$, $\underline{\phi}_\perp^+(t,\tau)$, and $\gamma(t,\tau)$. Employing the identity (17) yields $\underline{\phi}^{\text{inv}}(t,\tau) \star \underline{\phi}(t,\tau) = \delta(t-k)I_M$. Thus any $\underline{\phi}^{\text{inv}}(t,\tau) \in \mathbf{\Phi}^+$ is indeed a causal and stable left inverse of $\underline{\phi}(t,\tau)$. Conversely consider any causal and stable left inverse $\underline{\tilde{\phi}}^+(t,\tau)$. Then with $\underline{\phi}^{\text{inv}}(t,\tau) \in \mathbf{\Phi}^+$,

$$\underline{\phi}^{\text{inv}}(t,\tau) \star \begin{bmatrix} \underline{\phi}(t,\tau) & \underline{\phi}_\perp(t,\tau) \end{bmatrix}$$
$$= \begin{bmatrix} \underline{\phi}^+(t,\tau) + \gamma(t,\tau) \star \underline{\phi}_\perp^+(t,\tau) \end{bmatrix} \star \begin{bmatrix} \underline{\phi}(t,\tau) & \underline{\phi}_\perp(t,\tau) \end{bmatrix}$$
$$= \begin{bmatrix} \delta(t-\tau)I_M & \gamma(t,\tau) \end{bmatrix},$$

and

$$\underline{\tilde{\phi}}^+(t,\tau) \quad \star \quad \begin{bmatrix} \underline{\phi}(t,\tau) & \underline{\phi}_\perp(t,\tau) \end{bmatrix}$$
$$= \begin{bmatrix} \delta(t-\tau)I_M & \underline{\phi}^+(t,\tau) \star \underline{\phi}_\perp(t,\tau) \end{bmatrix}.$$

Thus with $\gamma(t,\tau) = \underline{\phi}^+(t,\tau) \star \underline{\phi}_\perp(t,\tau)$ which is causal and stable, $\underline{\tilde{\phi}}^+(t,\tau)$ is indeed in $\mathbf{\Phi}^+$, by the fact that $\underline{\phi}_a(t,\tau) = \begin{bmatrix} \underline{\phi}(t,\tau) & \underline{\phi}_\perp(t,\tau) \end{bmatrix}$ is square, and has a unique causal and stable inverse. ∎

Theorem 3.1 is the time-varying version for the parameterization of all stable left inverses in [2], [3] associated with the time-invariant systems. It is instrumental to obtaining the optimal channel equalizers. With the presence of the observation noise at the output of the channel, the composite system $\underline{\phi}(t,\tau)$ as in (4) is modified into

$$x(t+1) = A(t)x(t) + B(t)\mathbf{s}(t), \quad x(0) = 0, \qquad (19)$$
$$\underline{y}(t) = C(t)x(t) + D(t)\mathbf{s}(t) + \underline{\eta}(t).$$

Our objective is to seek $\underline{g}(t,\tau)$, among all the zero-forcing receiver filterbanks parameterized in Theorem 3.1, which minimizes the MSE $J(\underline{g})$. We assume that the noise process has zero mean, with known covariance for all t:

$$E[\eta(t)] = 0, \quad E\left[\underline{\eta}(t)\underline{\eta}^*(k)\right] = R_{\underline{\eta}}(t)\delta(t-k). \qquad (20)$$

Under the zero-forcing condition, $\underline{g}(t,\tau) \in \mathbf{\Phi}^+$, by Theorem 3.1. Consequently the symbol detection error at the output of the receiving filterbank is

$$\underline{e}(t) = \hat{\mathbf{s}}(t) - \mathbf{s}(t) = \underline{g}(t,\tau) \star \underline{\eta}(t), \quad \underline{g}(t,\tau) \in \mathbf{\Phi}^+, \quad (21)$$

in light of (3), and Theorem 3.1. Thus the MSE, or the variance of the symbol detection error to be minimized is

$$J \quad (\underline{g} \in \mathbf{\Phi}^+)(t)$$
$$= \text{trace}\left\{ E\left[\left(\underline{g}(t,\tau) \star \underline{\eta}(t)\right) \left(\underline{g}(t,\tau) \star \underline{\eta}(t)\right)^* \right] \right\}$$
$$= \text{trace}\left\{ \sum_{\tau=-\infty}^{t} \underline{g}(t,\tau)R_{\underline{\eta}}(\tau)\underline{g}^*(t,\tau) \right\}. \qquad (22)$$

Theorem 3.2: Suppose that the hypotheses in Lemma 2.4 hold. Let $\underline{g}(t,\tau) = \underline{\phi}_{\text{opt}}^+(t,\tau) \in \mathbf{\Phi}^+$ be the optimal receiver filterbank to be designed with $\mathbf{\Phi}^+$ parameterized in Theorem 3.1. That is,

$$J(\underline{g} = \underline{\phi}_{\text{opt}}^+) = \inf_{\underline{g} \in \mathbf{\Phi}^+} \text{trace}\{ \sum_{\tau=-\infty}^{t} \underline{g}(t,\tau)R_{\underline{\eta}}(\tau)\underline{g}^*(t,\tau)\}$$

Then the optimal receiver $\underline{g}(t,\tau) = \underline{\phi}_{\text{opt}}^+(t,\tau) \in \mathbf{\Phi}^+$, which minimizes the MSE, is equivalent to the optimal state estimator for the process

$$\begin{bmatrix} x_K(t+1) \\ \hline y_K(t) \end{bmatrix} = \begin{bmatrix} \underline{A}(t) & \underline{B}(t) \\ \hline \underline{C}(t) & \underline{D}(t) \end{bmatrix} \begin{bmatrix} x_K(t) \\ \hline \underline{\eta}(t) \end{bmatrix} \qquad (23)$$

where $\underline{\eta}(t)$ is the same as in (19) and

$$\underline{A}(t) = \begin{bmatrix} A_0(t) & 0 \\ -D^+(t)C(t) & 0 \end{bmatrix}, \quad \underline{B}(t) = \begin{bmatrix} B(t)D^+(t) \\ D^+(t) \end{bmatrix}$$

$$\underline{C}(t) = \begin{bmatrix} D_\perp^* C(t) & 0 \end{bmatrix}, \quad \underline{D}(t) = -D_\perp^*(t), \qquad (24)$$

$A_0(t) = A(t) - B(t)D^+(t)C(t)$ and a state estimator gain $\underline{L}(t,\tau) = \begin{bmatrix} L(t,\tau) \\ -\gamma(t,\tau) \end{bmatrix}$. The state estimator is schematically illustrated in Figure 2.
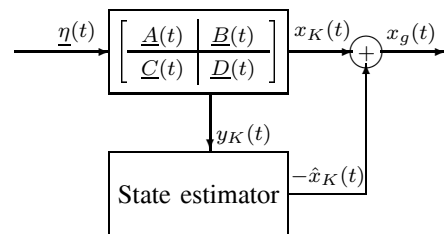

Fig. 2. State estimator.

Proof: Let $\underline{g}(t,\tau) \in \mathbf{\Phi}^+$. Then by (18), and $\underline{\phi}_a^{-1}(t,\tau) = \begin{bmatrix} \underline{\phi}^+(t,\tau) \\ \underline{\phi}_\perp^+(t,\tau) \end{bmatrix}$,

$$\underline{g}(t,\tau) = \underline{\phi}^+(t,\tau) + \gamma(t,\tau) \star \underline{\phi}_\perp^+(t,\tau)$$
$$= \begin{bmatrix} \delta(t-\tau)I_M & \gamma(t,\tau) \end{bmatrix} \star \phi_a^{-1}(t,\tau).$$

By the MSE to be minimized as in (21), and In light of the state-space model for $\underline{\phi}_a^{-1}(t,\tau)$ as in (16), and the MSE to be minimized as in (21), we have that the above $\underline{g}(t,\tau)$ can be described by the following state-space model (recall that $A_L(t)$ is exponentially stable, by the hypotheses):

$$x_g(t+1) = A_L(t)x_g(t) + B_L(t)\underline{\eta}(t), \qquad (25)$$
$$B_L(t) = B(t)D^+(t) - L(t)D_\perp^*(t),$$
$$\underline{e}(t) = -D^{(\text{inv})}(t) \star \left[C(t)x_g(t) + D_\perp^*(t)\underline{\eta}(t) \right],$$
$$D^{(\text{inv})}(t) = D^+(t) + \gamma(t,\tau) \star D_\perp^+(t)$$

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:2, No:11, 2008

by an abuse of notations. It follows that the state-space model of $\underline{g}(t,\tau) \in \mathbf{\Phi}^+$ as in (25) has the same form as the state-space model for $\underline{\phi}^+(t,\tau)$ as in (10), with the difference in inputs, and $D^+(t)$ replaced by $D^{(\mathrm{inv})}(t)$, which is again a left inverse of $D(t)$. Since $J(g)$ as in (22) does not change with $\underline{g}(t,\tau)$ replaced by $\underline{\tilde{g}}(t,\tau) = q^{-1} \star \underline{g}(t,\tau)$, where q^{-1} is the unit delay operator, we have an equivalent minimization problem for $J(\tilde{g})$ where $\underline{\tilde{g}}(t,\tau) = q^{-1} \star \underline{g}(t,\tau)$ is described by the state-space model

$$\tilde{x}_g(t+1) := \begin{bmatrix} x_g(t+1) \\ \hat{\mathbf{s}}(t+1) \end{bmatrix} \tag{26}$$

$$= \begin{bmatrix} A_L(t) & 0 \\ -D^{(\mathrm{inv})}(t) \star C(t) & 0 \end{bmatrix} \tilde{x}_g(t) + \begin{bmatrix} B_L(t) \\ D^{(\mathrm{inv})}(t) \end{bmatrix} \star \underline{\eta}(t),$$

$$\tilde{\mathbf{s}}(t) = \hat{\mathbf{s}}(t-1) = \underline{C}_1 \tilde{x}_g(t), \quad \underline{C}_1 = \begin{bmatrix} 0 & I_M \end{bmatrix},$$

by the state-space model of $\underline{g}(t,\tau)$ as in (25), and

$$\begin{aligned} A_L(t) &= A_0(t) + L(t)C_0(t), \\ B_L(t) &= B(t)D^+(t) - L(t)D_\perp^*(t). \end{aligned}$$

Therefore with $\underline{A}(t), \underline{B}(t), \underline{C}(t), \underline{D}(t)$ as defined in (24), and

$$\underline{L}(t,\tau) = \begin{bmatrix} L(t,\tau) \\ -\gamma(t,\tau) \end{bmatrix}, \tag{27}$$

(26) is equivalent to

$$\begin{aligned} \tilde{x}_g(t+1) &= [\underline{A}(t) + \underline{L}(t,\tau) \star \underline{C}(t)] \tilde{x}_g(t) \\ &\quad + [\underline{B}(t) - \underline{L}(t,\tau) \star D_\perp^*(t)] \underline{\eta}(t), \quad (28) \\ \tilde{\mathbf{s}}(t) &= C_1 \tilde{x}_g(t). \end{aligned}$$

Now consider the time-varying system as in (23). In light of the Kalman filtering theory [1], the optimal state estimate for $x_K(t)$ based on output measurements $y_K(\cdot)$ up to time $(t-1)$ is the conditional mean $\hat{x}_K(t|t-1)$, satisfying

$$\begin{aligned} &E\left[\{x_K(t) - \hat{x}(t|t-1)\}\{x_K(t) - \hat{x}(t|t-1)\}^*\right] \geq \\ &E\left[\{x_K(t) - \hat{x}_K(t|t-1)\}\{x_K(t) - \hat{x}_K(t|t-1)\}^*\right], \quad (29) \end{aligned}$$

for any other (linear) estimate $\hat{x}(t|t-1)$. Moreover the optimal state estimator has the form

$$\begin{aligned} \hat{x}_K(t+1|t) &= \underline{A}(t)\hat{x}_K(t|t-1) \\ &\quad + \underline{K}(t,\tau) \star [y_K(t) - \underline{C}\hat{x}_K(t|t-1)], \quad (30) \end{aligned}$$

with $\underline{K}(t,\tau) = \underline{K}_{\mathrm{opt}}(t)$ the optimal state estimation gain which is non-dynamic, but time-varying. The above yields the error system

$$\begin{aligned} x_e(t+1) &= [\underline{A}(t) + \underline{K}(t,\tau) \star \underline{C}(t)] x_e(t) \\ &\quad + [\underline{B}(t) - \underline{K}(t,\tau) \star D_\perp^*(t)] \underline{\eta}(t), \quad (31) \end{aligned}$$

with $x_e(t) = x_K(t) - \hat{x}_K(t|t-1)$ which is identical to (28), if $x_e(t) = \tilde{x}_g(t)$, and $\underline{K}(t,\tau) = \underline{L}(t,\tau)$. Thus the augmented system $\underline{\tilde{g}}(t,\tau) = q^{-1} \star \underline{g}(t,\tau)$ has the form of state estimator. Hence minimization of $J(g)$ over all possible $\underline{g}(t,\tau) \in \mathbf{\Phi}^+$ is equivalent to optimal state estimator design over all possible state estimator gain $\underline{K}(t,\tau) = \underline{L}(t,\tau)$, which can be dynamical. ∎

In light of the celebrated Kalman filtering theory, the optimality of the state estimator is achieved by the static time-varying gain, and therefore our MMSE design for optimal channel equalization needs consider only the state estimation gain $\underline{L}(t)$, without searching over the dynamical gain $\underline{L}(t,\tau)$. However direct use of the Kalman filtering on the augmented system increases the order of the system by M, which can be large. The following demonstrate that we can obtain an nth order receiver filterbank to achieve optimal channel equalization.

Theorem 3.3: Suppose that the set of all zero-forcing receiving filters $\mathbf{\Phi}^+$ is nonempty. Let the time-varying system $\underline{\phi}(t,\tau)$ be as in (4). Then the blocked time-varying receiver filterbank $\underline{g}(t,\tau) \in \mathbf{\Phi}^+$ achieving optimal channel equalization is described by the state-space model

$$\begin{aligned} \tilde{x}(t+1) &= \left[A_0(t) + L^{\mathrm{opt}}(t)C_0(t)\right] \tilde{x}(t) \quad (32) \\ &\quad + \left[B(t)D^+(t) - L^{\mathrm{opt}}(t)D_\perp^*(t)\right] \underline{y}(t), \\ \hat{\mathbf{s}}(t) &= -D^{(\mathrm{inv})}(t) \left[C(t)\tilde{x}(t) - \underline{y}(t)\right], \end{aligned}$$

where $D^{(\mathrm{inv})}(t) = D^+(t) + \Gamma^{(\mathrm{opt})}(t)D_\perp^*(t)$, and

$$A_0(t) = A(t) - B(t)D^+(t)C(t), \quad C_0(t) = D_\perp^*(t)C(t). \tag{33}$$

The optimal state estimator gain $L(t) = L^{\mathrm{opt}}(t)$, and $\gamma(t,\tau) = \Gamma^{\mathrm{opt}}(t)$, are given by

$$-L^{\mathrm{opt}}(t) = [A_0(t)X(t|t-1)C_0^*(t) + B(t)S_0(t)]\, \tilde{R}^{-1}(t)$$

$$\Gamma^{\mathrm{opt}}(t) = \left[S_0(t) - D^+(t)C(t)X(t|t-1)C_0^*(t)\right] \tilde{R}^{-1}(t). \tag{34}$$

That is, $\gamma(t,\tau)$ can be chosen as non-dynamic. The matrix $X(t|t-1)$ is the covariance of $\tilde{x}(t)$, calculated from the following recursive difference Riccati equation (DRE)

$$\begin{aligned} X(t+1|t) &= A_0(t)X(t|t-1)A_0^*(t) \quad (35) \\ &\quad - \mathcal{X}(t)\tilde{R}^{-1}(t)\mathcal{X}^*(t) + B(t)R_0(t)B^*(t) \end{aligned}$$

with $\mathcal{X}(t) = [A_0(t)X(t|t-1)C_0^*(t) + B(t)S_0(t)]$, and the covariance for the initial value of the state vector $\tilde{x}(0)$ is $X(0|-1) = X_0$, and

$$\tilde{R}(t) = D_\perp^*(t)R_\eta(t)D_\perp(t) + C_0(t)X(t)C_0^*(t) \tag{36}$$

$$S_0(t) = -D^+(t)R_\eta(t)D_\perp(t), \tag{37}$$

$$R_0(t) = D^+(t)R_\eta(t)(D^+(t))^*.$$

Proof: By the proof of Theorem 3.2, and [1], $\hat{x}_K(t|t-1) = E[x_K(t)|Y_{t-1}]$ is the optimal state estimate for the process in (23), based on measurements $Y_{t-1} = \{y_K(\tau)\}_{\tau=0}^{t-1}$, which satisfies (29) for any other state estimate $\hat{x}(t|t-1)$. That is $\hat{\Sigma}(t|t-1) \geq \Sigma(t|t-1)$ with $\hat{\Sigma}(t|t-1)$ the covariance for $\hat{x}(t|t-1)$, and $\Sigma(t|t-1)$ for $\hat{x}_K(t|t-1)$, as proven in [1]. By the equivalence established in Theorem 3.2, and the state-space model (28), we have

$$\begin{aligned} E\left[\{x_K(t) - \hat{x}_K(t|t-1)\}\{x_K(t) - \hat{x}_K(t|t-1)\}^*\right] = \\ E\left[\tilde{x}_g(t)\tilde{x}_g^*(t)\right], \end{aligned}$$

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:2, No:11, 2008

with $\tilde{x}_g(t)$ as in (28). Denote $\Sigma_{t|t-1} = \Sigma(t|t-1)$. Then the error covariance for $\tilde{x}_g(t)$ is given by

$$
\begin{aligned}
\Sigma_{t|t-1} &= E\left[\tilde{x}_g(t)\tilde{x}_g^*(t)\right] \qquad (38)\\
&= E\left\{ \begin{bmatrix} x_g(t) \\ \hat{\mathbf{s}}(t) \end{bmatrix} \begin{bmatrix} x_g^*(t) & \hat{\mathbf{s}}(t)^* \end{bmatrix} \right\}\\
&= E\begin{bmatrix} x_g x_g^*(t) & x_g(t)\hat{\mathbf{s}}(t)^* \\ \hat{\mathbf{s}}(t)x_g^*(t) & \hat{\mathbf{s}}(t)\hat{\mathbf{s}}(t)^* \end{bmatrix}\\
&= \begin{bmatrix} \Sigma_{t|t-1}^{(1,1)} & \Sigma_{t|t-1}^{(12)} \\ \Sigma_{t|t-1}^{(21)} & \Sigma_{t|t-1}^{(22)} \end{bmatrix}
\end{aligned}
$$

Applying the Kalman filtering results [1], we have that $\Sigma_{t|t-1}$ satisfies the following DRE,

$$
\begin{aligned}
\Sigma_{t+1|t} =\ &\underline{A}(t)\Sigma_{t|t-1}\underline{A}^*(t) + \underline{B}(t)R_\eta(t)\underline{B}^*(t)\\
&- \left[\underline{A}(t)\Sigma_{t|t-1}\underline{C}^*(t) + \underline{B}(t)R_\eta(t)D_\perp\right]\\
&\times \tilde{\underline{R}}^{-1}(t)\left[\underline{A}(t)\Sigma_{t|t-1}\underline{C}^*(t) + \underline{B}(t)R_\eta(t)D_\perp\right]^*,
\end{aligned}
$$

where $\tilde{\underline{R}}(t)$ as in (36), and $R_\eta(t)$ as in (20). By the expressions in (24), the (1,1) position of the above DRE is the same as

$$
\begin{aligned}
\Sigma_{t|t-1}^{(11)} =\ &A_0(t)\Sigma_{t|t-1}^{(11)}A_0^*(t) + B(t)R_0(t)B^*(t)\\
&- \left[A_0(t)\Sigma_{t|t-1}^{(11)}C_0^*(t) + B(t)S_0(t)\right]\\
&\times \tilde{R}^{-1}(t)\left[A_0(t)\Sigma_{t|t-1}^{(11)}C_0^*(t)B(t)S_0(t)\right]^*
\end{aligned}
$$

which is identical to (35), with $X(t|t-1) = \Sigma_{11}(t|t-1) = \Sigma_{t|t-1}^{(11)}$. Because

$$
\underline{A}(t)\Sigma(t|t-1)\underline{C}^*(t) = \begin{bmatrix} A_0(t) \\ -D^+C(t) \end{bmatrix} X(t|t-1)C^*(t)D_\perp,
$$
$$
\underline{C}(t)\Sigma(t|t-1)\underline{C}^*(t) = D_\perp^*C(t)X(t|t-1)C^*(t)D_\perp,
$$

the optimal state estimation gain formula as in [1] yields the expressions in (34). ∎

In light of the various properties of the Kalman filter, the following can be easily deduced.

Corollary 3.4: The optimal channel equalizer given in Theorem 3.3 is stable in the sense that $A_0(t) + L^{\text{opt}}(t)C_0(t)$ is exponentially stable. In particular, if $\underline{\phi}(t,\tau)$ converges to a time-invariant system, and the state space model matrices $A(t)$, $B(t)$, $C(t)$, and $D(t)$ converge to A, B, C, and D respectively for which the PR condition holds and D has full column rank, then the DRE as in (35) converges to the following ARE

$$
\begin{aligned}
X\ -\ &A_0XA_0^* - BR_0B^* \qquad (39)\\
&+ \left[A_0XC_0^* + BS_0\right]\tilde{R}^{-1}\left[A_0XC_0^* + BS_0\right]^* = 0,
\end{aligned}
$$

with X the stabilizing solution, and the optimal state estimator gain and Γ converge to

$$
\begin{aligned}
L = L^{\text{opt}} &:= -\left[A_0XC_0^* + BS_0\right]\tilde{R}^{-1},\\
\Gamma = \Gamma^{\text{opt}} &:= \left[S_0 - D^+CXC_0^*\right]\tilde{R}^{-1},
\end{aligned}
$$

where $A_L = A_0 + LC_0$ is exponentially stable.

Remark 3.5: In light of (29), our proposed channel equalizer is optimal over all possible linear equalizers with respect to the MSE. If in addition the noise is AWGN, then it is also optimal over all nonlinear equalizers. In fact by [1] the Kalman filter in presence of AWGN is the maximum a posteriori (MAP) estimator, implying that our proposed channel equalizer minimizes the BER as well, which will be discussed again in the next section.

We summarize this section with the following design algorithm:

Design Algorithm for Optimal Channel Equalizers:

Step 1: Find state-space realizations for the blocked time-varying systems: $\underline{h}(t,\tau)$, and $\underline{f}(t,\tau)$ $\forall$ t. Then find a state space realization for $\underline{\phi}(t,\tau)$ in (1), which satisfies (6), and which are exponentially stabilizable.

Step 2: Set the state-space model for the optimal channel equalizer as in (32), with the *a priori* initial condition $\tilde{x}(0) = \tilde{x}_0$, which has the covariance $X_0 \geq 0$.

Step 3: For $t = 0, 1, \cdots$, do the following:
- Compute $\hat{s}(t)$ according to (32).
- Compute DRE (35). For $t = 0$, use $X(0|-1) = X_0$. Set $L(t) = L^{\text{opt}}(t)$ and $\Gamma(t) = \Gamma^{\text{opt}}(t)$.
- Compute $\tilde{x}(t+1)$ according to (32). For $t = 0$, use $\tilde{x}(0) = \tilde{x}_0$.

End.

It is noted that the initial condition $\tilde{x}(0) = \tilde{x}_0$, with covariance matrix X_0, is assumed *a priori*. Because of the optimality of the Kalman filter, $\tilde{x}(t)$ converges rather quickly to its steady-state value, and thus only the first a few estimates may have large errors, which is shown for the first simulation example in the next section.

IV. BER AND ILLUSTRATIVE EXAMPLES

Bit Error Rate Probability

A commonly used performance measurement is BER, rather than the MSE. To find an expression for the BER at the receiver, we denote $\mathcal{J}(t) = E\left[\underline{e}(t)\underline{e}^*(t)\right]$. Thus there holds $J(\underline{g} = \phi_{\text{opt}}^+)(t) = \text{trace}\{\mathcal{J}(t)\}|_{\underline{g}=\phi_{\text{opt}}^+}$. The MSE value $J_i(t)$, the ith diagonal element of $\mathcal{J}(t)$, is given by $J_i(t) := E[e_i(t)e_i^*(t)]$, $e_i(t) = \underline{g}_i(t,\tau) \star \underline{\eta}(t)$, where $\underline{g}_i(t,\tau)$ is the ith row of the blocked receivers filters $\underline{g}(t,\tau)$. Thus, the MSE for the ith received data stream $\mathbf{s}_i(t)$ is $J_i(t)$, and the average error probability for the M symbol streams $\mathbf{s}(t)$ is defined as $P_e := \frac{1}{M}\sum_{i=0}^{M-1}P_e^{(i)}$, where $P_e^{(i)}$ denotes the error probability of the ith symbol stream. For BPSK constellations, the error probability of the ith symbol in the case of additive Gaussian noise (AGN) is given by $P_e^{(i)} = \frac{1}{2}\text{erfc}\left(\frac{1}{\sqrt{J_i(t)}}\right)$

Illustrative Examples

We consider examples which illustrate our proposed optimal channel equalization, where both deterministic time-varying channels, and fading channels are examined.

Example 1: We consider the fading channel given by its impulse response

$$
\mathbf{h}(t) = E[\mathbf{h}(t)] + \mathbf{dh}(t), \qquad (40)
$$

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
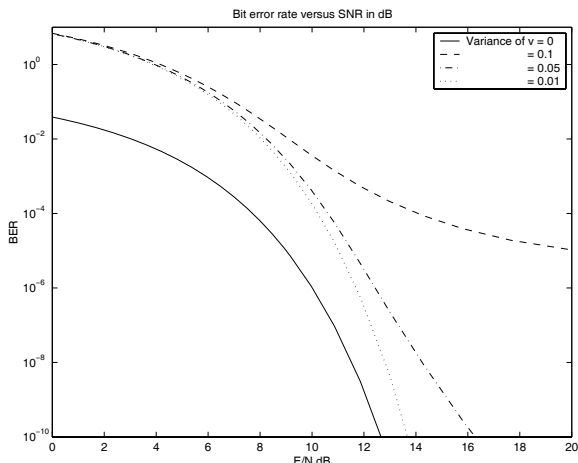Vol:2, No:11, 2008

Fig. 3.   Bit error rate sketches for the fading channel (Example 1).

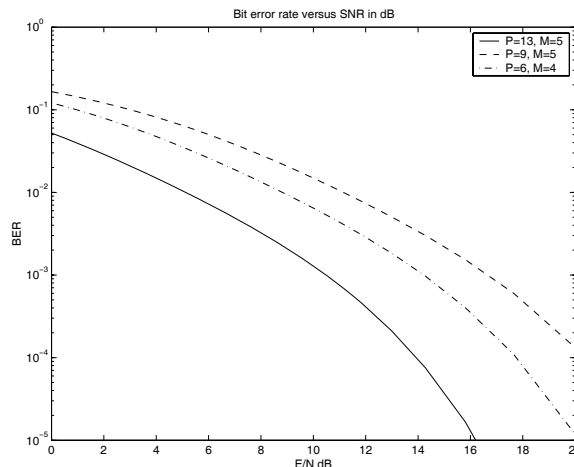Fig. 4.   Bit error rate sketches for fading channel (Example 2).

with $\mathbf{dh}(t+1) = A\mathbf{dh}(t) + B\mathbf{v}(t)$, where A, and B are both diagonal matrices, satisfying $a_i^2 + b_i^2 = 1$, and a_i, b_i are the ith diagonal elements of A and B respectively [5]. The random process $v(\cdot)$ has zero mean with a fixed variance σ_v^2. We assume that $E[\mathbf{h}(t)]$ is known to the receiver (estimated via some estimation algorithms), and equal $\mathbf{h}(t) = \begin{bmatrix} h_0(t) & h_1(t) & \cdots & h_L(t) \end{bmatrix} = \mathbf{h}_{ss} + \alpha^{-t}, \quad \sigma(t) = \sigma_{ss} + \alpha^{-t} \quad \forall \ t$. The values $\alpha = 1.2$, $\sigma_{ss} = 0.1$ are chosen, and $\mathbf{h}_{ss}$ has length $L + 1 = 8$ given by $\mathbf{h}_{ss} = \begin{bmatrix} 1 & -0.3 & 0.5 & -0.4 & 0.1 & -0.02 & 0.3 & -0.1 \end{bmatrix}$.
$E[\mathbf{h}(t)]$ is used to compute the optimal equalizer. Then we compute the BER curve for the true channel as in (40) generated for three different values of $\sigma_v^2 = 0.1, 0.05, 0.01$. Figure 2 shows the BER as a function of E_b/N_0, using the Monte Carlo method, and averaging. As we can see that the BER improves as the variance σ_v^2 as in decreases. Clearly for large variance channel process noise, knowing the mean channel is not adequate to achieve good BER performance. Hence channel estimation has to be performed to obtain better information of the channel, in order to improve the BER performance.

Example 2:   We consider a time-varying fading channel model; the variation of each tap $h(t, \tau)$ can be simulated by the following sum [11]:

$$h(t, \tau) = \sum_{q=0}^{Q} c_q(\tau) e^{j\theta_q} \exp\left[j(2\pi v\lambda)\cos(2\pi q/Q)t\right] \quad (41)$$

where the parameters are c_q amplitude of the qth path, θ_q uniformly distributed random variable in $[0, 2\pi]$, λ wavelength corresponding to the carrier frequency, and v speed of the mobile.
The variance for the non-stationary noise $\eta(t)$ is given by $\sigma^2(t) = (\sigma_1 + \frac{t}{N}(\sigma_2 - \sigma_1))^2$ with $\sigma_1 = 0.1$, $\sigma_2 = 1$. For the case $Q = 4$, which is a suitable value for the number of paths as in [11], $\lambda = 0.333\ m$ which is corresponding to the carrier frequency 900MHz, and $v = 22.22\ m/s$, we generate the time-varying fading channel with order $L = 10$. For different values for P and M, we set $\underline{f}(t, \tau) = \delta(t - \tau)\begin{bmatrix} I_M & 0 \end{bmatrix}^T$, which is the precoder commonly used in the literature. We

obtain a simple state-space model of $\underline{\phi}(t, \tau)$.

Following the steps of the design algorithm in the previous section. The optimal receiver filters $\{g_p(t, \tau)\}_{p=0}^{P-1}$ can then be implemented according to Theorem 3.3. For different values of P and M, we compute the BER Figure 3 shows the BER's for each pair of P and M as a function of $E_b/N_0(t)$. As we can see from Figure 3, the variation of P and M changes the BER, indicating the trade-off between the data rate efficiency, and BER performance. ■

Example 3   This example is from [11], where the channel is given by:

$$h(t, \tau) = c(\tau)\frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \exp[j(2\pi f_{D_{\max}} \cos(\alpha_{k,\tau} t) + \theta_{k,\tau})], \tag{42}$$

where the parameters are $c(\tau) = 1$ controls the power of the τth tap chosen according to the power delay profile, $f_{D_{\max}}$ the desired maximum Doppler frequency and has the value of 66Hz, K the number of sinusoids and equal 100, and $\alpha_{k,\tau}$, $\theta_{k,\tau}$ mutually independent, uniformly distributed random variables.
We used the variance as in the previous example. The random channel was generated with order $L = 10$. For different values for P and M, the optimal receiver filters $\{g_p(t, \tau)\}_{p=0}^{P-1}$ are computed according to Theorem 3.3, assuming that the channel information is available at the receiver. Figure 6 shows the BER curves for each pair of P and M as a function of $E_b/N_0(t)$, where $E_b/N_0(0) = 1$ and $E_b/N_0(\infty) = 100$.

The BER for the random channel is expected to be higher than that of the previous example, because two independent random variables are used to generate the fading channel, versus only one in Example 2. As we can see from Figure 3, the variation of P and M changes the BER. The higher the P value and the larger the ratio of P to M, the better the BER. ■

V. Conclusion

In this paper we investigated the optimal channel equalization problem for time-varying channels in the presence

World Academy of Science, Engineering and Technology
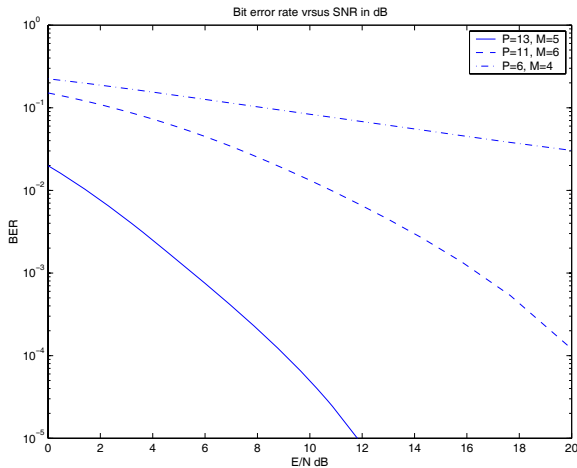International Journal of Electronics and Communication Engineering
Vol:2, No:11, 2008

Fig. 5. Bit error rate sketches for random channel (Example 3).

of possible non-stationary noises. A state-space approach is adopted to obtain the optimal channel equalizer, which is an optimal state estimator for some modified system. A necessary and sufficient condition on zero-forcing is established for the noise-free case in Section 2. If the zero-forcing condition holds for the given time-varying channel and the transmitter filterbank, and the noise is present, optimal time-varying receiver filterbank is obtained in Section 3, which is a modified Kalman filter. The optimal channel equalization results in this paper are especially useful for wireless communication, as emphasized in [9].

REFERENCES

[1] B.D. Anderson and J.B. Moore, *Optimal Filtering*, Prentice Hall, New Jersey, 1979.
[2] E. F. Badran, *Optimal Channel Equalization for Filterbank Transceivers in Presence of White Noise*, Ph.D Dissertation, Electrical and Computer Eng. Dept., Louisiana State Univ., Baton Rouge, USA, May 2002.
[3] G. Gu and E.F. Badran, "Optimal design for channel equalization via filterbank approach," *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 536-545 Feb. 2004.
[4] S. Haykin, A.H. Sayed, J.R. Zeidler, P. Yee, and P.C. Wei, "Adaptive tracking of linear time-variant systems by extended RLS algorithms," *IEEE Trans. Signal Processing*, vol. 45, pp. 1118-1128, 1997.
[5] C. Komninakis, C. Fragouli, A.H. Sayed, "MIMO fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Processing*, vol. 50, pp. 1065-1076, 2002.
[6] S.-Y. Kung, Y. Wu, and X. Zhang, "Bezout space-time precoders and equalizers for MIMO channels," *IEEE Trans. Signal Processing*, vol. 50, 2499-2514, oct. 2002.
[7] T.J. Lim, L.K. Rasmussen, and H. Sugimoto, "An asynchronous multiuser CDMA detector based on the Kalman filter," *IEEE J. Select. Areas Commun.*, vol. 16, no. 9, pp. 1711-1722, December 1998.
[8] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516-525, Feb. 1995.
[9] A. Scaglione, G.B. Giannakis and S. Barbarossa "Redundant filterbank precoders and equalizers, part I: Unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1983-2006, July 1999.
[10] A. Scaglione, P. Stoica, S. Barbarossa, G.B. Giannakis, and H. Sampath, "Optimal design for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051-1064, May 2002.
[11] C. Tepedelenlioglu and G.B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Trans. on Signal Processing*, vol. 48, No. 7, pp. 2029-2043, July 2000.
[12] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proceedings of IEEE*, vol. 86, 1951-1968, Oct. 1998.
[13] M.K. Tsatsanis and G.B. Giannakis, "Optimal decorrelating receivers for DS-CDMA systems: A signal processing framework," *IEEE Trans. on Signal Processing*, vol. 44, 3044-3055, 1996.

Ehab F. Badran received the B.S. and M.S. degrees in Electrical Engineering from Assiut University, Assiut, Egypt, in 1995 and 1998, respectively, and the M.S. and Ph.D. degrees in Electrical Engineering from Louisiana State University, Baton Rouge, Louisiana, USA, in 2001 and 2002, respectively. From 1995 to 1998, he was assigned to work as an instructor in the Department of Electrical Engineering, Assiut University, Egypt. In May 1998, he was promoted to work as an assistant lecturer in the same department. From 2000 to 2002, he worked as a teaching assistant and research assistant in the Electrical and Computer Engineering, Louisiana State University, Baton Rouge, Louisiana, USA, during his graduate study at LSU. From September 2002 to September 2003, he worked as an Assistant Professor in the Department of Electrical Engineering, Assiut University, Assiut, Egypt. Since September 2003, he is an Assistant Professor in the Department of Electronics and Communications Engineering, Arab Academy for Science and Technology & Maritime Transportation, Alexandria, Egypt. His research and teaching interests are in communications, signal processing, and speech signal processing.
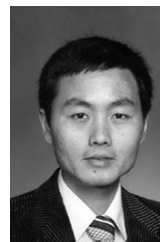
Guoxiang Gu (S'87-M'90-SM'02) received the BS degree from the East China University in 1981, the MS and Ph.D. degrees from University of Minnesota in 1985 and 1988, respectively, all in electrical engineering. Since 2000, he has been professor at Louisiana State University, Baton Rouge. He is now F. Hugh Coughlin/CLECO Distinguished Professor of Electrical Engineering at LSU. His previous academic experience includes: visiting assistant professor (1988-1990) at the Wright State University, Dayton, Ohio, research faculty at the Wright Laboratory (summers of 1994 and 1995) of US Air Force. He was an associate editor for IEEE Transactions on Automatic Control from 1988 to 2000. He co-authored one book, numerous book chapters, and over 50 journal papers and 60 conference papers. His research has focused on system identification, robust control, and communication-oriented digital signal processing.