Realizable minimum mean-squared error channel shorteners

Roberto López-Valcarce*
Departamento de Teoría de la Señal y las Comunicaciones
Universidad de Vigo
36200 Vigo (Pontevedra), Spain
Ph. +34 986 812677, Fax +34 986 812116
E-mail: valcarce@gts.tsc.uvigo.es .

Abstract

We present an analysis of realizable (i.e., causal, stable, and of finite degree) minimum mean-squared error (MMSE) channel shorteners for multiple-input multiple-output (MIMO) systems, driven by spatially and temporally white signals, and subject to a constant output power constraint. This is of interest since this design has recently been shown to result in near optimal rate performance in multitone transceivers, and the performance of conventional finite impulse response (FIR) shorteners is upper bounded by that of realizable schemes. The MMSE shortener is shown to consist of a prewhitening filter followed by an FIR postfilter of order equal to the sum of the overall delay and the target shortening length. It is shown that this design results in output decorrelation, and that the shortener output enjoying the smallest MMSE sees a target impulse response without zeros inside the unit disk. The asymptotic behavior of the shortened system is explored, and performance bounds are provided in terms of the channel frequency response and the noise power spectral density.

I. Introduction

In FFT-based multicarrier systems, intersymbol and intercarrier interference can be avoided with the insertion of a cyclic prefix of length ν between consecutive symbols, for both single-input (e.g. discrete multitone (DMT) modulation [5]) and multiple-input (e.g. discrete matrix multitone (DMMT) [19]) frequency selective channels. If the channel impulse response spans no more than $\nu+1$ samples, then the effect of the channel appears as a circular convolution, enabling one-tap equalization on each subcarrier.

* This work was supported by the Ramón y Cajal program of the Spanish Ministry of Science and Technology.

The cyclic prefix reduces the data rate of the system by a factor of $N/(N+\nu)$, where N is the length of the multicarrier symbol, so ν has to be kept as small as possible. Usually the channel memory is much longer than the cyclic prefix length, so that a channel shortener (also known as a time-domain equalizer) is placed in the receiver with the purpose of appropriately shortening the overall impulse response. Channel shorteners also find application as 'preconditioners' for maximum-likelihood sequence estimation in order to reduce the computational complexity of the Viterbi algorithm, which increases exponentially with the channel memory [3].

The goal of this paper is the study of the structure and properties of channel shorteners under the only constraint of realizability, i.e., the filter may have an infinite impulse response (IIR), but it must be causal, stable, and of finite degree. This is in contrast to standard approaches which either assume shorteners with a doubly infinite number of taps [9], [14] or consider finite impulse response (FIR) filters of fixed order (thus, potentially suboptimal) [3]. We adopt the minimum mean-squared error (MMSE) criterion, by which the shortener minimizes the MSE between its output and that of a virtual channel with a short target impulse response (TIR), subject to a constant output power constraint. While not attempting to directly maximize the achievable bit rate (which constitutes the ultimate performance measure in DMT systems), our design bears a direct relation to that from [6], which was shown to be near to rate optimal.

Assuming a temporally and spatially white transmit signal, the general multiple-input multiple-output (MIMO) framework is adopted throughout. Multicarrier MIMO techniques are proving very valuable in wireless applications, in which two antenna arrays are deployed at the transmitter and receiver sides (see e.g. [18] and the references therein), and also in certain wireline DSL systems in which a whole bundle of twisted pairs can be accessed at both ends [10]. Multiple subchannels may be available as well by oversampling the received signals. The MIMO channel shortening problem has been considered in [2], [20], [22], under the assumption of a fixed order shortener. On the other hand, an analysis of single-input realizable equalizers (rather than shorteners) has been presented in [12]; this can be seen as a special case of the channel shortening problem in which $\nu = 0$, so that the target impulse response reduces to a single spike.

The paper is organized as follows. Section II introduces the input-output model and the mathematical description of the channel shortening problem. Section III presents the optimal TIR; the optimal shortener is given in Section IV, along with a discussion of its properties. Section V analyzes the asymptotic performance of the realizable design for large shortening delay. Numerical results are presented in Section VI, and the paper is concluded in Section VII.

In our notation, vectors and matrices are denoted in lowercase and uppercase respectively. Superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ denote respectively the conjugate, the transpose and the conjugate transpose. $\lambda_{\min}\{A\}$, $\lambda_{\max}\{A\}$ and $\operatorname{tr} A$ denote the smallest and largest eigenvalues and the trace of matrix A respectively, while '$\otimes$' denotes the Kronecker product. ψ_k denotes a vector of all zeros except for a 1 in the k-th position, and for a transfer function $P(z)$, the 'tilde' notation denotes paraconjugation, i.e. $\tilde{P}(z) = [P(1/z^*)]^H$.

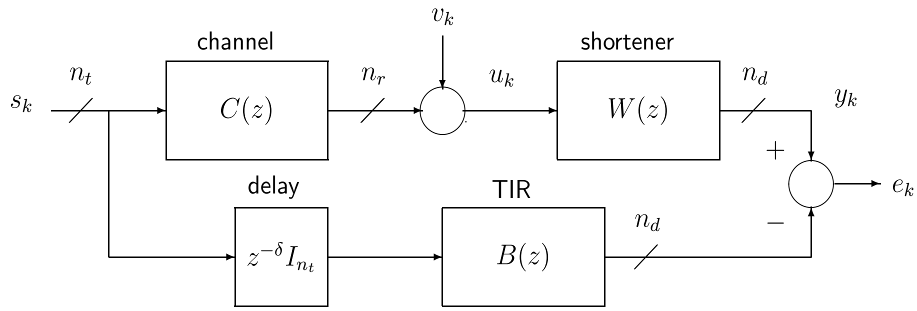

Fig. 1. Illustration of the shortening error generation.

II. Problem formulation

We consider a MIMO channel with transfer function $C(z) = \sum_{n=0}^{\infty} C_n^H z^{-n}$, where each term C_n is a $n_t \times n_r$ matrix. The $n_r \times 1$ received signal vector is then

$$u_k = C(z)s_k + v_k, \tag{1}$$

where $\{s_k\}$ is the $n_t \times 1$ vector sequence of transmitted symbols, assumed zero mean white with unit covariance, i.e. $E[s_k s_l^H] = I_{n_t}$ if $k = l$ and zero otherwise; $\{v_k\}$ is the $n_r \times 1$ vector noise process, assumed independent of $\{s_k\}$ and zero mean with power spectral density $S_v(z)$. It is also assumed that $\{v_k\}$ is a full rank process [4], meaning that its power spectral density (PSD) $S_v(z)$ is not singular for all z.

Fig. 1 illustrates the problem. A virtual TIR with $\nu + 1$ taps and $n_d \times n_t$ transfer function $B(z) = \sum_{n=0}^{\nu} B_n^H z^{-n}$ is introduced, and the shortening error e_k is defined as the difference between the shortener output $y_k = W(z)u_k$ and the TIR output, including a delay δ. This delay is a design parameter which can considerably impact the final performance. The goal is to determine $B(z)$ and the $n_d \times n_r$ transfer function $W(z)$ in order to minimize the MSE

$$J_\delta^\nu \triangleq \operatorname{tr} E[e_k e_k^H], \tag{2}$$

when the shortener $W(z)$ is constrained to be realizable: it must be causal and stable and its implementation must require a finite number of delay elements. n_d represents the number of effective antennas after shortening, and may or may not be equal to n_t or n_r, depending on the application.

Clearly, in order to avoid the trivial solution $W(z) = 0$, $B(z) = 0$, some additional constraint must be imposed. Traditionally, designs from the literature have considered identity tap or orthonormality constraints on either the shortener or the TIR [3], [2]. More recently, an intuitively more appealing constant output power constraint has been proposed in [23] for the $n_t = n_r = n_d = 1$ case. Several ways of generalizing this strategy to the MIMO case are possible. Perhaps the most natural one is to constrain the total output power: $\operatorname{tr} E[y_k y_k^H] = 1$. However, with this approach there is no guarantee that some of the shortener output signals are not nulled out. To avoid this problem, we shall use instead a uniform power constraint over the output vector, i.e. $E[|y_{k,l}|^2] = 1/n_d$, $1 \le l \le n_d$, where $y_{k,l}$ is the l-th element of y_k.

III. OPTIMAL TIR

To find the optimum $B(z)$, let us introduce the overall transfer function

$$Q(z) = W(z)C(z) = \sum_{n=0}^{\infty} Q_n^H z^{-n}. \tag{3}$$

Then the error covariance can be written as

$$E[e_k e_k^H] = \sum_{n=0}^{\delta-1} Q_n^H Q_n + \sum_{n=\delta}^{\delta+\nu} (Q_n^H - B_{n-\delta}^H)(Q_n - B_{n-\delta}) + \sum_{n>\delta+\nu} Q_n^H Q_n + E[\bar{v}_k \bar{v}_k^H], \tag{4}$$

where $\bar{v}_k = W(z)v_k$ is the filtered noise. Since the TIR coefficients do not affect the output power $E[|y_{k,l}|^2]$, minimization of the trace of (4) w.r.t. $B(z)$ can be done directly, leading to

$$B_n = Q_{n+\delta}, \qquad 0 \leq n \leq \nu. \tag{5}$$

With this, the 'reduced MSE cost' becomes

$$\begin{aligned} J_\delta^\nu &= \operatorname{tr}\left(\sum_{n=0}^{\infty} Q_n^H Q_n - \sum_{n=\delta}^{\delta+\nu} Q_n^H Q_n + E[\bar{v}_k \bar{v}_k^H]\right) && (6)\\ &= \operatorname{tr} E[y_k y_k^H] - \operatorname{tr}\left(\sum_{n=\delta}^{\delta+\nu} Q_n^H Q_n\right), && (7) \end{aligned}$$

which gives, in view of (5) and the output power constraints $E[|y_{k,l}|^2] = 1/n_d$,

$$J_\delta^\nu = 1 - ||B(z)||^2 = 1 - \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{tr}[B(e^{j\omega})\tilde{B}(e^{j\omega})]d\omega. \tag{8}$$

Hence, minimizing J_δ^ν s.t. $E[|y_{k,l}|^2] = 1/n_d$, $1 \leq l \leq n_d$, is equivalent to maximizing $\operatorname{tr}\sum_{n=\delta}^{\delta+\nu} Q_n^H Q_n$ subject to the same constraint.

Let us introduce the signal-to-interference-plus-noise ratio (SINR) as the ratio of the signal power inside the window $\delta \leq n \leq \delta+\nu$ to that outside (the 'interference') plus the noise power, all measured at the shortener output. Then we can write

$$\text{SINR}_\delta^\nu = \frac{\operatorname{tr}\sum_{n=\delta}^{\delta+\nu} Q_n^H Q_n}{\operatorname{tr}(E[y_k y_k^H] - \sum_{n=\delta}^{\delta+\nu} Q_n^H Q_n)} = \frac{\operatorname{tr} E[y_k y_k^H]}{J_\delta^\nu} - 1. \tag{9}$$

From (9), it is seen that maximizing SINR_δ^ν s.t. $E[|y_{k,l}|^2] = 1/n_d$, $1 \leq l \leq n_d$ is equivalent to minimizing J_δ^ν subject to the same constraints. Therefore, under a constant output power constraint on the shortener outputs, the MMSE and maximum SINR designs are equivalent.

Recently, Daly *et al.* [6] have proposed a design where they showed to achieve near-optimal bit rate performance in (single-input single-output) DMT transceivers. It is based on minimization of the sum of the interference plus noise powers at the shortener output, subject to the constraint of keeping constant the output useful (that is, inside the window) signal power. Since interference plus noise power equals the total power minus the useful signal power, this amounts to minimizing the total output power while keeping constant the useful signal power. The corresponding solution is the same (up to a scaling constant) as that in the MMSE or maximum SINR design with constant output power constraints; therefore the MMSE design can be expected to achieve good performance in DMT systems in terms of bit rate as well.

IV. OPTIMAL SHORTENER

In this section we derive the structure of the MMSE shortener as well as several of its properties. Following a Lagrange multiplier approach, we show that it consists of a prewhitener applied to the received signal followed by an FIR filter with $\delta+\nu+1$ (matrix valued) taps. Then it will be seen that the optimal shortener results in a drastic reduction of the diversity present in the output vector y_k. We get around this problem by suitably selecting a suboptimal shortener that preserves the degree of diversity available. Next it is found that the resulting TIR corresponding to the shortener output with largest SINR is devoid of zeros inside the unit circle. Certain relations in the achieved performance in terms of ν and δ are then given.

A. Shortener structure

Let $S_u(z) = C(z)\tilde{C}(z) + S_v(z)$ be the PSD of the received signal $\{u_k\}$. It can be factored as

$$S_u(z) = F(z)\Omega\tilde{F}(z), \tag{10}$$

where $\Omega = \Omega^H$ is positive definite because $\{v_k\}$ (and hence $\{u_k\}$) is assumed to be a full rank process, and $F(z)$ is $n_r \times n_r$ causal, monic ($F(z) = I$ for $z^{-1} = 0$) and minimum phase (all poles lie in $|z| < 1$, and $F(z)$ has constant rank in $|z| \geq 1$). We shall further assume that $F(z)$ has finite McMillan degree [21], so it is realizable. This will be the case provided that $C(z)$ has finite degree (though possibly IIR), and that $\{v_k\}$ can be accurately modeled as an ARMA process of finite order. In that case, we can assume without loss of generality that the shortener includes a prewhitening filter as front end,

$$W(z) = G(z)F^{-1}(z), \tag{11}$$

(note that $F^{-1}(z)$ is realizable), and then perform the optimization in terms of $G(z) = \sum_{n=0}^{\infty} G_n^H z^{-n}$. To do so, let us introduce the prewhitened channel

$$H(z) = F^{-1}(z)C(z) = \sum_{n=0}^{\infty} H_n^H z^{-n}. \tag{12}$$

With this, the matrix $Q = [\begin{array}{cccc} Q_\delta^H & Q_{\delta+1}^H & \cdots & Q_{\delta+\nu}^H \end{array}]^H$ of the coefficients in the window of interest of the overall transfer function (3), given by $Q(z) = G(z)H(z)$, can be written as $Q = \mathcal{H}_\delta G$, where $\mathcal{H}_\delta$ is a $(\nu+1)n_t \times (\delta+\nu+1)n_r$ block Toeplitz matrix given by

$$\mathcal{H}_\delta = \begin{bmatrix} H_\delta & H_{\delta-1} & \cdots & H_0 & 0 & \cdots & 0 \\ H_{\delta+1} & H_\delta & \cdots & H_1 & H_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \\ H_{\delta+\nu} & H_{\delta+\nu-1} & \cdots & H_\nu & H_{\nu-1} & \cdots & H_0 \end{bmatrix}, \tag{13}$$

and $G = [\begin{array}{cccc} G_0^H & G_1^H & \cdots & G_{\delta+\nu}^H \end{array}]^H$. Therefore, the quantity to maximize can be written as

$$\sum_{n=\delta}^{\delta+\nu} Q_n^H Q_n = Q^H Q = G^H \mathcal{H}_\delta^H \mathcal{H}_\delta G. \tag{14}$$

That is, the signal power inside the window $[\delta, \delta+\nu]$ depends only on the first $\delta+\nu+1$ coefficients of the impulse response of $G(z)$.

Now note that the output of the front end prefilter $F^{-1}(z)$ is a temporally white process with covariance equal to Ω, so that the shortener output covariance reads

$$E[y_k y_k^H] = \sum_{n=0}^{\infty} G_n^H \Omega G_n = G^H \Omega_\delta G + \sum_{n>\delta+\nu} G_n^H \Omega G_n, \tag{15}$$

where $\Omega_\delta = I_{\delta+\nu+1} \otimes \Omega$. Then, from (14) and (15), the problem is seen to become

$$\text{maximize} \quad \operatorname{tr} G^H \mathcal{H}_\delta^H \mathcal{H}_\delta G \quad \text{subject to} \quad \psi_l^H \left(G^H \Omega_\delta G + \sum_{n>\delta+\nu} G_n^H \Omega G_n \right) \psi_l = \frac{1}{n_d}, \quad 1 \le l \le n_d. \tag{16}$$

The corresponding Lagrangian for this problem is

$$\mathcal{L} = \operatorname{tr} G^H (\mathcal{H}_\delta^H \mathcal{H}_\delta) G + \sum_{l=1}^{n_d} \lambda_l \left(\frac{1}{n_d} - \psi_l^H G^H \Omega_\delta G \psi_l - \sum_{n>\delta+\nu} \psi_l^H G_n^H \Omega G_n \psi_l \right),$$

where λ_l are the Lagrange multipliers. Let $\Lambda \triangleq \operatorname{diag}(\lambda_1, \cdots, \lambda_{n_d})$. Equating the partial derivatives of $\mathcal{L}$ w.r.t. G and G_i with $i > \delta+\nu$ to zero, we obtain the following equations:

$$\Omega G_i \Lambda = 0, \qquad i > \delta + \nu \tag{17}$$

$$\mathcal{H}_\delta^H \mathcal{H}_\delta G = \Omega_\delta G \Lambda, \tag{18}$$

plus the original constraint (16). At any stationary point satisfying these equations, for each $l = 1, \ldots, n_d$, either (i) $\lambda_l = 0$ and $\mathcal{H}_\delta^H \mathcal{H}_\delta G \psi_l = 0$, or (ii) $(\lambda_l, G\psi_l)$ is a generalized eigenpair of the matrix pencil $(\mathcal{H}_\delta^H \mathcal{H}_\delta, \Omega_\delta)$, and $G_i \psi_l = 0$ for all $i > \delta + \nu$. The value of $\operatorname{tr} Q^H Q$ at these candidate points becomes

$$\begin{aligned}
\operatorname{tr} Q^H Q &= \operatorname{tr} G^H \Omega_\delta G \Lambda \\
&= \sum_{l=1}^{n_d} \lambda_l \cdot \psi_l^H G^H \Omega_\delta G \psi_l \\
&= \sum_{l=1}^{n_d} \lambda_l \left(\frac{1}{n_d} - \psi_l^H \sum_{n>\delta+\nu} G_n^H \Omega G_n \psi_l \right) \\
&= \frac{1}{n_d} \sum_{l=1}^{n_d} \lambda_l - \sum_{n>\delta+\nu} \operatorname{tr}(G_n^H \underbrace{\Omega G_n \Lambda}_{=0}) \\
&= \frac{1}{n_d} \sum_{l=1}^{n_d} \lambda_l.
\end{aligned} \tag{19}$$

Observe that this value is unaffected by the value of G_i, $i > \delta + \nu$. Hence, in order to satisfy (17), we can choose $G_i = 0$, $i > \delta + \nu$ without any loss of optimality, concluding that the optimum postfilter $G(z)$ is FIR of order $\delta + \nu$.

The operation of the optimal shortener is most easily visualized in the high SNR, SISO case. With $n_t = n_r = n_d = 1$ and neglecting the noise, the effect of the prewhitening filter $f^{-1}(z)$ is seen to be (i) cancel all of the channel poles, (ii) cancel all channel zeros inside the unit disk, and (iii) place poles at the reciprocal of the channel zeros outside the unit disk. The resulting effective channel $h(z)$ is an allpass system. If the order of $h(z)$ (which is the number of original channel zeros outside the unit disk) is no larger than ν, then $\nu+1$ degrees of freedom in the postfilter $g(z)$ are spent in canceling the poles of $h(z)$, while the remaining δ degrees introduce the required delay (recall that $g(z)$ has $\delta+\nu+1$ taps). When the noise power is not negligible, the prewhitening filter has to account for noise presence (and possibly coloring) as well, and the action of the FIR postfilter is not as easily perceived.

B. SINR vs. diversity

The global maximum of (19) is obtained when all λ_l's are equal to the largest generalized eigenvalue of $(\mathcal{H}_\delta^H \mathcal{H}_\delta, \Omega_\delta)$. However, this solution may not be acceptable in terms of channel capacity, as explained next.

Suppose that this largest generalized eigenvalue had multiplicity one. In such case, the columns of the optimal (in the MSE sense) matrix G are all equal (up to a scaling), and hence the rows of the transfer function $G(z)$ will be all equal as well, resulting in $y_{k,1} = y_{k,2} = \cdots y_{k,n_d}$ for all k. This means that the rank of the overall shortened channel $W(z)C(z)$ is at most one for all z, even if that of the original channel $C(z)$ was larger. This in turn implies a drop in the capacity of the overall $n_d \times n_t$ MIMO system (which grows with the channel rank, see [19]), a clearly undesirable effect.

A more sensible solution is to allow some excess MSE in order to avoid this potential capacity reduction of the global MMSE shortener:

choose $\lambda_1, \ldots, \lambda_{n_d}$ to be the n_d largest generalized eigenvalues (counting their multiplicities) of $(\mathcal{H}_\delta^H \mathcal{H}_\delta, \Omega_\delta)$, with the columns of G the corresponding generalized eigenvectors.

In that case the columns of G (the corresponding generalized eigenvectors) can be chosen to be linearly independent, thus avoiding the rank drop problem. With this design, the covariance matrix of the shortener output becomes diagonal:

$$E[y_k y_k^H] = G^H \Omega_\delta G = n_d^{-1} I_{n_d}, \tag{20}$$

due to the orthogonality property of the generalized eigenvectors.

Using (19), one finds that $\psi_l^H E[e_k e_k^H] \psi_l = (1-\lambda_l)/n_d$. Therefore, at each of the shortener outputs $y_{k,l}$, $1 \le l \le n_d$, the individual MMSE will follow the same ordering as the eigenvalue assignment. Note also that if $n_d > r \triangleq \operatorname{rank} \mathcal{H}_\delta^H \mathcal{H}_\delta$, then one would end up with $\lambda_{r+1} = \cdots = \lambda_{n_d} = 0$ and therefore $E[e_{k,l} e_{k,l}^*] = E[y_{k,l} y_{k,l}^*] = 1/n_d$, for $r+1 \le l \le n_d$. This means that for these values of l, the l-th shortener output $y_{k,l}$ comprises noise and interference only, but no useful signal. In view of this, it is of no advantage to increase n_d beyond r. Further, we have from (13) that

$$r \le n_t \cdot (\nu+1), \tag{21}$$

with equality holding if at least one of the taps $H_0, \ldots, H_\delta$ has full column rank. Thus the number of effective antennas cannot exceed the product of the number of transmit antennas times the number of taps in the target impulse response (a manifestation of transmit and multipath diversity).

C. Maximum phase property

Note from (18) that the matrix $Q = \mathcal{H}_\delta \, G$ satisfies

$$\mathcal{H}_\delta^H \, Q = \mathcal{H}_\delta^H \mathcal{H}_\delta \, G = \Omega_\delta G \Lambda \tag{22}$$

and therefore

$$\mathcal{H}_\delta \, \Omega_\delta^{-1} \, \mathcal{H}_\delta^H \, Q = \mathcal{H}_\delta \, G\Lambda = Q\Lambda. \tag{23}$$

That is, the columns of Q are the eigenvectors of the $n_t(\nu+1)$-square matrix $R_\delta^\nu \triangleq \mathcal{H}_\delta \, \Omega_\delta^{-1} \, \mathcal{H}_\delta^H$ associated to its largest eigenvalues $\lambda_1 \geq \ldots \geq \lambda_{n_d}$. This in turn implies the following fact about the row of the attained TIR $B(z) = \sum_{n=0}^{\nu} Q_{n+\delta}^H z^{-n}$ corresponding to the output with largest SINR.

Property 1: The $1 \times n_t$ transfer function $b_1(z) \triangleq \psi_1^H B(z)$ has no zeros inside the unit disk, provided that the largest eigenvalue λ_1 of R_δ^ν is simple.

The proof is given in the Appendix. The following remarks are in order:

1) Observe that if $n_t > 1$, then $b_1(z)$ need not have any zeros at all, i.e. the entries of $b_1(z)$ may be coprime. Property 1 shows that the greatest common divisor of these entries is a maximum phase polynomial. (If $n_t = 1$, then $b_1(z)$ itself is a scalar maximum phase polynomial).
2) When the largest eigenvalue of R_δ^ν has multiplicity $m > 1$, then multiple choices are possible for $b_1(z)$. While not all of these are necessarily devoid of zeros inside the unit disk in that case, at least one of them will have this property, as can be deduced following [13].
3) It must be noted that this result applies to realizable shorteners only. When the channel shortener is constrained to be FIR with a prespecified number of taps, as is the case with more traditional designs, the resulting MMSE TIR need not satisfy this maximum phase property.
4) It is known that in $n_t = n_d = 1$ systems, the performance of suboptimum trellis-based equalizers such as delayed decision-feedback sequence estimation (DDFSE) [7] or reduced-state sequence estimation (RSSE) [8] can be severely degraded if the overall discrete-time channel impulse response is not minimum phase. If the channel is shortened using the realizable design with the ultimate goal of using one of these sequence estimators, one possibility is to operate in reversed time (i.e. the receiver waits until reception of a data burst is complete and then the stored received samples are time-reversed and fed to the sequence estimator). This method effectively converts the maximum-phase TIR into a minimum-phase one.

D. Performance properties

Let us denote by $\lambda_\delta^\nu[1] \geq \lambda_\delta^\nu[2] \geq \cdots \geq \lambda_\delta^\nu[n_d]$ the n_d largest eigenvalues of the matrix R_δ^ν. The realizable design in section IV-B results in an MMSE

$$J_\delta^\nu = 1 - \frac{1}{n_d} \sum_{l=1}^{n_d} \lambda_\delta^\nu[l]. \tag{24}$$

It is readily checked that

$$R_{\delta+1}^{\nu} = R_{\delta}^{\nu} + \begin{bmatrix} H_{\delta+1} \\ H_{\delta+2} \\ \vdots \\ H_{\delta+\nu+1} \end{bmatrix} \Omega^{-1} \begin{bmatrix} H_{\delta+1}^H & H_{\delta+2}^H & \cdots & H_{\delta+\nu+1}^H \end{bmatrix}, \tag{25}$$

which implies $\lambda_{\delta+1}^{\nu}[l] \geq \lambda_{\delta}^{\nu}[l]$ for each l, and consequently

$$J_{\delta+1}^{\nu} \leq J_{\delta}^{\nu}, \tag{26}$$

meaning that for a given target length ν, increasing the delay improves performance (always measured in terms of SINR or MSE). On the other hand, it is also seen that the $n_t(\nu+1) \times n_t(\nu+1)$ southeast submatrix of $R_{\delta-1}^{\nu+1}$ is R_{δ}^{ν}, so that $\lambda_{\delta-1}^{\nu+1}[l] \geq \lambda_{\delta}^{\nu}[l]$ in view of Cauchy's interlacing theorem [17]. Therefore

$$J_{\delta-1}^{\nu+1} \leq J_{\delta}^{\nu}, \tag{27}$$

i.e. performance also improves if the target length is increased by one tap while the lag is decreased by the same amount. (By doing this, one is changing the 'don't care' window from $[\delta, \delta+\nu]$ to $[\delta-1, \delta+\nu]$, which contains the primitive window). Iterating this argument, the following relation is obtained:

$$J_{0}^{\delta+\nu} \leq \cdots \leq J_{\delta-1}^{\nu+1} \leq J_{\delta}^{\nu} \leq J_{\delta+1}^{\nu-1} \leq \cdots \leq J_{\delta+\nu}^{0}. \tag{28}$$

Finally, (26)-(27) together imply that $J_{\delta}^{\nu+1} \leq J_{\delta}^{\nu}$. That is, for a given delay, increasing the target length cannot result in worse performance, which is not surprising: it comes at the expense of lower bandwidth efficiency in a multicarrier system, or higher complexity for a trellis-based sequence estimator.

V. Asymptotic results

We now explore the properties of the solution as the delay δ is allowed to increase. The general MIMO case is examined first; then, several properties applicable to the single-input channel are presented.

A. The general case

We begin by introducing the $n_t(\nu+1)$-square matrix

$$R^{\nu} \triangleq \lim_{\delta\to\infty} R_{\delta}^{\nu}. \tag{29}$$

It can be easily checked that the (i,j) block of size $n_t \times n_t$ of R^{ν} is given by

$$(R^{\nu})_{i,j} = \sum_{n=0}^{\infty} H_n \Omega^{-1} H_{n+j-i}^H \triangleq K_{j-i}, \qquad 0 \leq i \leq j \leq \nu. \tag{30}$$

Thus R^{ν} is a block Toeplitz Hermitian matrix. As a consequence, if (λ, q) is an eigenpair of R^{ν} and λ is simple, then q must be either 'block symmetric' or 'block skew-symmetric', i.e. it must satisfy $q = cXq^*$

where c is a scalar with $|c|^2 = 1$ and X is the $n_t(\nu+1)$-square block reversal matrix

$$X = \begin{bmatrix} & & I_{n_t} \\ & \cdot^{\cdot^{\cdot}} & \\ I_{n_t} & & \end{bmatrix}.$$

Recall that the transfer function $b_1(z) = \psi_1^H B(z)$ that featured in Property 1 was constructed from an eigenvector of R_δ^ν associated to the largest eigenvalue. Thus, for $\delta \to \infty$, it follows that if the largest eigenvalue of R^ν is simple, then $b_1(z) = cz^{-\nu}[\tilde{b}_1(z)]^T$ for some c with $|c|^2 = 1$. Hence, if $b_1(z_0) = 0$, one must have $b_1(1/z_0^*) = 0$ as well. In addition, from Property 1, $b_1(z)$ has no zeros inside the unit disk, so it is concluded that all of the zeros of the asymptotic transfer function must be located on the unit circle. This fact was first pointed out in [1] for the $n_t = n_d = 1$ case; see also [14].

From (30), the power spectral density associated to R^ν is readily found to be

$$\begin{aligned} K(z) = \sum_{n=-\infty}^{\infty} K_n z^{-n} &= \tilde{H}(z)\Omega^{-1}H(z) \\ &= \tilde{C}(z)S_u^{-1}(z)C(z). \end{aligned} \tag{31}$$

Let us introduce the spectral SNR matrix (which has size $n_t \times n_t$)

$$\Gamma(e^{j\omega}) \triangleq \tilde{C}(e^{j\omega})S_v^{-1}(e^{j\omega})C(e^{j\omega}). \tag{32}$$

Then, as shown in the Appendix, one has

$$K(z) = \Gamma(z)[I_{n_t} + \Gamma(z)]^{-1} = I_{n_t} - T(z) \qquad \text{with} \qquad T(z) \triangleq [I_{n_t} + \Gamma(z)]^{-1}. \tag{33}$$

In view of (24) and (33), it is seen that the asymptotic MMSE $J^\nu \triangleq \lim_{\delta\to\infty} J_\delta^\nu$ is just the average of the n_d smallest eigenvalues of the $n_t(\nu+1) \times n_t(\nu+1)$ covariance matrix $I - R^\nu$, which is constructed from $T(z)$. From [16, Theorem 3.1], it follows that the eigenvalues of this matrix lie in the interval $[\min_\omega \lambda_{\min}\{T(e^{j\omega})\}, \max_\omega \lambda_{\max}\{T(e^{j\omega})\}]$. This means that the asymptotic MMSE can be bounded in the same way for all ν:

$$\min_\omega \lambda_{\min}\{T(e^{j\omega})\} \leq J^\nu \leq \max_\omega \lambda_{\max}\{T(e^{j\omega})\}. \tag{34}$$

A better upper bound is obtained by noticing that $J^\nu \leq J^0$ for all ν, and that

$$J^0 \leq \frac{1}{n_t}\operatorname{tr}(I_{n_t} - R^0) = \frac{1}{n_t}\frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{tr} T(e^{j\omega})d\omega, \tag{35}$$

with equality holding in (35) if $n_d = n_t$.

Note that the SINR is related to the MMSE via $\text{SINR}_\delta^\nu = (1/J_\delta^\nu) - 1$. From this and (33), the bounds in (34) can be rephrased in terms of the SINR to obtain

$$\min_\omega \lambda_{\min}\{\Gamma(e^{j\omega})\} \leq \text{SINR}^\nu \leq \max_\omega \lambda_{\max}\{\Gamma(e^{j\omega})\}, \tag{36}$$

where $\mathrm{SINR}^\nu \triangleq \lim_{\delta\to\infty}\mathrm{SINR}^\nu_\delta$. In addition, as $\nu\to\infty$ the n_d smallest eigenvalues of $I-R^\nu$ converge to $\min_\omega \lambda_{\min}\{T(e^{j\omega})\}$ [16, Corollary 3.5], so that the lower (resp. upper) bound in (34) (resp. (36)) is attained asymptotically:

$$\lim_{\nu\to\infty} J^\nu = \min_\omega \lambda_{\min}\{T(e^{j\omega})\}, \qquad \lim_{\nu\to\infty}\mathrm{SINR}^\nu = \max_\omega \lambda_{\max}\{\Gamma(e^{j\omega})\}. \tag{37}$$

Thus, the asymptyotic SINR is given by the peak value of the largest eigenvalue of the spectral SNR.

B. *The single-input case*

When $n_t = n_d = 1$, the matrix R^ν becomes Toeplitz and J^ν reduces to the smallest eigenvalue of $I-R^\nu$. The spectral SNR (32) becomes a scalar $\gamma(e^{j\omega})$, and so does $t(z)=1/[1+\gamma(z)]$. Then the MMSE bounds (34)-(35) read as

$$\min_\omega t(e^{j\omega}) \le J^\nu \le \frac{1}{2\pi}\int_{-\pi}^{\pi} t(e^{j\omega})d\omega, \tag{38}$$

and they are attained for $\nu\to\infty$ and $\nu=0$ respectively.

Observe now that, from the Cauchy-Schwarz inequality,

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} t(e^{j\omega})d\omega\right)\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{t(e^{j\omega})}d\omega\right)\ge 1,$$

and therefore the asymptotic MMSE value for $\nu=0$ is upper bounded as

$$J^0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} t(e^{j\omega})d\omega \ge \frac{1}{1+\frac{1}{2\pi}\int_{-\pi}^{\pi}\gamma(e^{j\omega})d\omega}. \tag{39}$$

Hence, from (36) and (39) it follows that the arithmetic mean of the spectral SNR lies between the asymptotic values of SINR for $\nu=0$ and $\nu\to\infty$:

$$\mathrm{SINR}^0 \le \underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi}\gamma(e^{j\omega})d\omega}_{\triangleq\,\gamma_{\mathrm{am}}} \le \lim_{\nu\to\infty}\mathrm{SINR}^\nu. \tag{40}$$

It is noteworthy that, in general, the arithmetic mean γ_{am} does not equal the average per subchannel SNR as usually defined:

$$\beta \triangleq \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi}||c(e^{j\omega})||^2 d\omega}{\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{tr}\, S_v(e^{j\omega})d\omega}. \tag{41}$$

Using the following factorization of the noise PSD,

$$S_v(e^{j\omega}) = \tilde{U}(e^{j\omega})\Sigma(e^{j\omega})U(e^{j\omega}) \tag{42}$$

where $\Sigma(e^{j\omega})>0$ is diagonal and $U(z)$ is paraunitary, i.e. $\tilde{U}(e^{j\omega})U(e^{j\omega}) = I_{n_r}$, it is shown in the Appendix that, in the special case of $\Sigma(e^{j\omega})=\Sigma$ constant, then

$$\beta \le \gamma_{\mathrm{am}}. \tag{43}$$

If $\Sigma=\sigma^2 I_{n_r}$, corresponding to temporally and spatially white noise, then $\gamma_{\mathrm{am}} = n_r\cdot\beta$, which shows the diversity advantage of having n_r subchannels.

Note that in the SISO channel case ($n_t=n_r=1$), one has $\gamma(e^{j\omega}) = |c(e^{j\omega})|^2/S_v(e^{j\omega})$ which is unaffected if a channel zero is reflected with respect to the unit circle, i.e. if $c(z)=(1-z_0z^{-1})\bar{c}(z)$ is changed to $(z_0^*-z^{-1})\bar{c}(z)$, and therefore the asymptotic performance of the realizable shorteners does not change. However, the rate of convergence in terms of δ will be different in general.

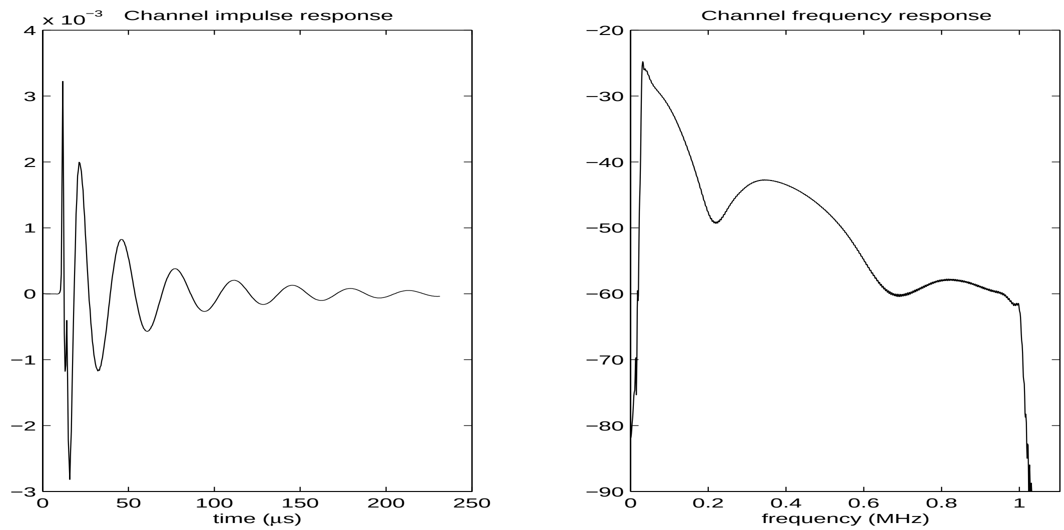

Fig. 2. Channel impulse and frequency responses of CSA loop 2 with splitter and front end filter.

VI. Numerical examples

In our first example we consider a single-input single-output channel as given by the standard DSL test loop CSA 2, combined with a POTS splitter and a twelfth-order Chebyshev bandpass filter for the 30-1000 kHz frequency band (see Fig. 2). The sampling frequency is 2.208 MHz. The resulting discrete-time impulse response is well approximated by an ARMA(4,34) system. Fig. 3(a) shows the asymptotic SINR for $\nu \leq 30$ (13.59 μs) considering a noise PSD of the form $S_v(e^{j\omega}) = \alpha^2(1+a^2-2a\cos\omega)/(1+a^2+2a\cos\omega)$ with $a = 0.5$, 0 and -0.5, corresponding to highpass, white and lowpass noise respectively. The factor α^2 is adjusted to yield a received SNR of 10 dB in all cases. The less harmful effect of highpass noise agrees with the fact that, in view of the channel transfer function, the maximum value of the spectral SNR $\gamma(e^{j\omega}) = |c(e^{j\omega})|^2/S_v(e^{j\omega})$ is larger in that case.

In practice, inclusion of a prewhitening filter in the design need not be resource optimal. The realizable (δ,ν) MMSE shortener has a total of $N_\delta^\nu = N_F + \delta + \nu + 1$ coefficients (in the SISO case), where N_F is the number of coefficients of the prewhitener. In this DSL example, the ARMA(4,34) channel model assuming white noise yields $N_F = 38$. Fig. 3(b) shows the SINR attained by the realizable design, as well as that of an N-tap FIR MMSE design optimized in terms of the overall delay, both as functions of the total number of shortener coefficients N and assuming $\nu = 15$. Observe that the smallest possible number of taps in the realizable design is $N_0^{15} = 54$. However, the performance of this approach can be attained by the FIR design with a much smaller number of taps, by optimizing over the delay range. This does not contradict the optimality of the realizable MMSE design, since it is possible for an FIR shortener with N_δ^ν taps to outperform the realizable (δ,ν) MMSE shortener for some delay $\delta' > \delta$.

For the second example we present the results obtained with a single-input two-output L-tap channel under Rayleigh fading with a constant power delay profile and white noise ($S_v(e^{j\omega}) = \sigma_v^2 I_2$). The channel taps are generated as complex zero-mean independent Gaussian random variables with variance σ_c^2, and

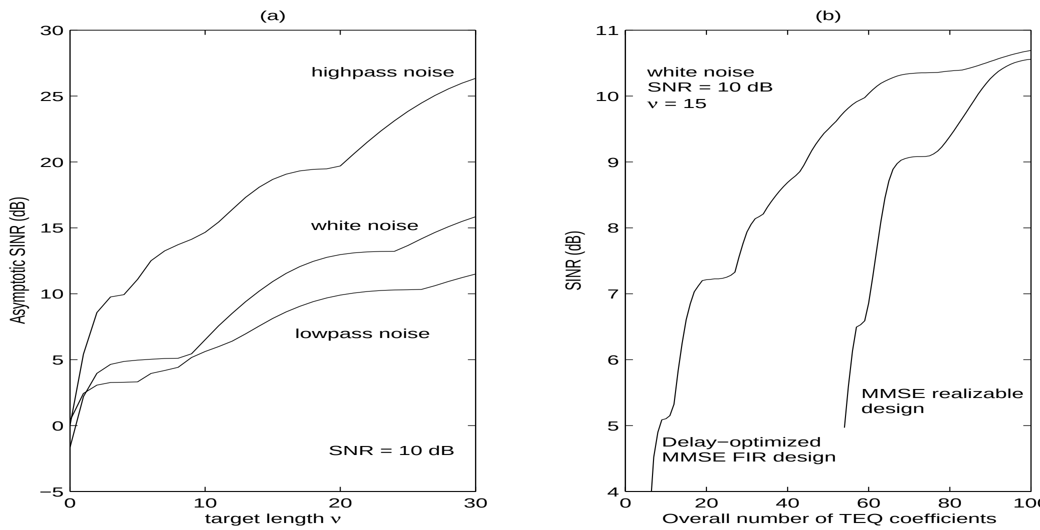

Fig. 3. (a) Asymptotic SINR for different noise spectral densities. (b) Comparison of MMSE realizable and MMSE FIR designs in terms of total number of taps.

performance is averaged over 1000 channel realizations. The target length is $\nu = 6$. Fig. 4(a) shows the attained SINR as a function of the number of channel taps, for an average SNR $L\sigma_c^2/\sigma_v^2 = 18$ dB and for several values of the delay. It is evident that the longer the channel delay spread, the larger the value of δ required for a given performance level. Also shown is the maximum of the spectral SNR (asymptotic SINR for an infinite target length) which, interestingly, increases with the channel length. Next we fixed $L = 16$ taps to estimate the performance of the realizable design for an SNR range of 10-30 dB; the results are shown in Fig. 4(b). The SINR is seen to increase linearly with the received SNR. For this value of L the SNR gap between the curves for $\delta = 0$ and $\delta = \infty$ is of 5 dB.

VII. Concluding remarks

Given a channel and a noise power spectral density, we have seen that the realizable MMSE (or maximum SINR) channel shortener for a target length ν and a given delay δ consists of a whitening prefilter followed by a $(\delta+\nu+1)$-tap FIR filter. For multiple-input channels, a tradeoff between shortening SINR and channel diversity is reached by assigning linearly independent generalized eigenvectors to different shortener outputs. The resulting shortened impulse response for the largest SINR output will be maximum phase, and the covariance matrix of the shortener output vector will be diagonal. Performance becomes a monotonic function of δ, and its asymptotic value (as the delay is increased) has been related to the spectral SNR available at the receiver.

These schemes provide an upper bound to the performance of any practical shortener for a given (δ, ν) pair. However, it is possible for a constrained shortener with fewer taps to yield higher SINR for some other delay $\delta' > \delta$ (which in practice must be optimized by exhaustive search). Analytically relating the performance of constrained shorteners to delay remains an interesting open problem. Extension of the

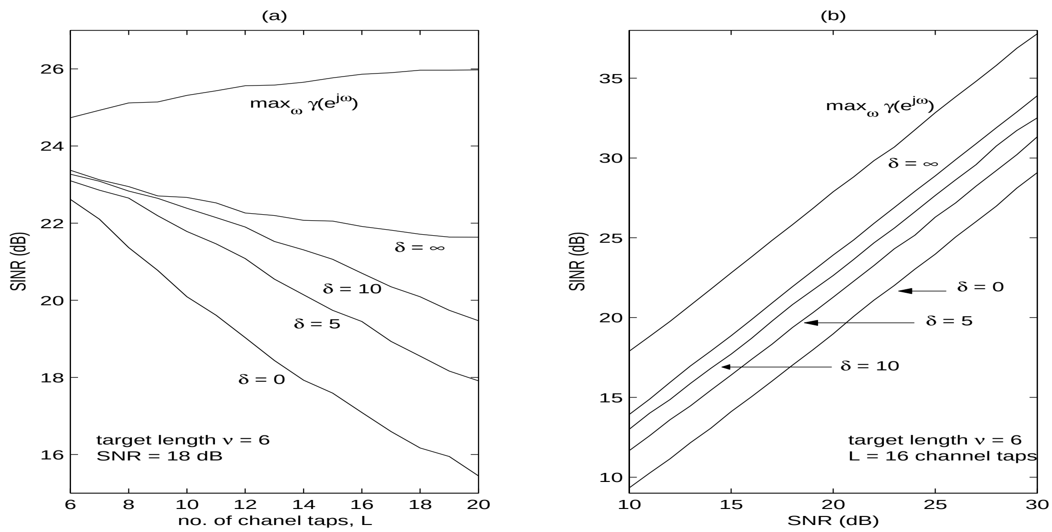

Fig. 4. Performance of the realizable design in 1-input 2-output Rayleigh channel with white noise. (a) SINR vs. number of channel taps, SNR = 18 dB . (b) SINR vs. SNR, 16 channel taps.

realizable design to allow for temporally and spatially colored transmit signals constitutes another open problem.

Appendix

Proof of Property 1: Note that the coefficients of $b_1(z)$ are given by the first column of Q, which from (23) is seen to be an eigenvector of R_δ^ν associated to its largest eigenvalue λ_1. Let us partition R_δ^ν into blocks as

$$R_\delta^\nu = [R_{ij}]_{0\le i,j\le \nu}$$

where each R_{ij} has size $n_t \times n_t$, and then introduce the block displacement matrix

$$\Delta \triangleq [R_{ij}]_{1\le i,j\le\nu} - [R_{ij}]_{0\le i,j\le \nu-1}.$$

It can be readily checked that Δ satisfies

$$\Delta = \begin{bmatrix} H_{\delta+1} \\ H_{\delta+2} \\ \vdots \\ H_{\delta+\nu} \end{bmatrix} \Omega^{-1} \begin{bmatrix} H_{\delta+1}^H & H_{\delta+2}^H & \cdots & H_{\delta+\nu}^H \end{bmatrix} \ge 0. \tag{44}$$

Using (44), a straightforward modification of Lemma 2 from [13] shows that if the eigenvalue is simple and $b_1(z_0) = 0$, then $|z_0| \ge 1$ must follow. ∎

Proof of (33): Starting with the expression (31) for $K(z)$, and ommitting the argument z for clarity, one has

$$\begin{aligned} K(z) &= \tilde{C} S_u^{-1} C \\ &= \tilde{C}(C\tilde{C} + S_v)^{-1} C \\ &= \tilde{C}[S_v^{-1} - S_v^{-1} C (I_{n_t} + \tilde{C} S_v^{-1} C)^{-1} \tilde{C} S_v^{-1}] C \\ &= \Gamma - \Gamma (I_{n_t} + \Gamma)^{-1} \Gamma \\ &= \Gamma - \Gamma (I_{n_t} + \Gamma)^{-1} [(I_{n_t} + \Gamma) - I_{n_t}] \\ &= \Gamma - \Gamma [I_{n_t} - (I_{n_t} + \Gamma)^{-1}] \\ &= \Gamma (I_{n_t} + \Gamma)^{-1}. \end{aligned}$$

In the third line we have used the matrix inversion lemma (a.k.a. Woodbury formula [11, p. 51]), and in the fourth, the definition of the spectral SNR (32). Now,

$$\begin{aligned} I_{n_t} - K(z) &= I_{n_t} - \Gamma (I_{n_t} + \Gamma)^{-1} \\ &= I_{n_t} - [(I_{n_t} + \Gamma) - I_{n_t}](I_{n_t} + \Gamma)^{-1} \\ &= I_{n_t} - [I_{n_t} - (I_{n_t} + \Gamma)^{-1}] \\ &= (I_{n_t} + \Gamma)^{-1}, \end{aligned}$$

which proves the last part of (33). ∎

Proof of (43): Write $\Sigma = \mathrm{diag}(\ \sigma_1^2 \quad \cdots \quad \sigma_{n_r}^2\ )$, and let $d(e^{j\omega}) = U(e^{j\omega}) c(e^{j\omega})$. Note that $||d(e^{j\omega})||^2 = ||c(e^{j\omega})||^2$. Then

$$\gamma(e^{j\omega}) = \tilde{d}(e^{j\omega}) \Sigma^{-1} d(e^{j\omega}) = \sum_{k=1}^{n_r} \frac{|d_k(e^{j\omega})|^2}{\sigma_k^2}.$$

Now note that, for arbitrary $\{x_k\}_{k=1}^p$ and nonzero $\{y_k\}_{k=1}^p$, the following holds:

$$\left(\sum_{k=1}^{p} \frac{x_k^2}{y_k^2} \right) \left(\sum_{n=1}^{p} y_n^2 \right) = \sum_{k=1}^{p} \left(x_k^2 + x_k^2 \sum_{n \neq k} \frac{y_n^2}{y_k^2} \right) \geq \sum_{k=1}^{p} x_k^2. \tag{45}$$

Therefore,

$$\begin{aligned} \gamma_{\mathrm{am}} &= \sum_{k=1}^{n_r} \frac{1}{\sigma_k^2} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |d_k(e^{j\omega})|^2 d\omega \right) \\ &\geq \frac{\sum_{k=1}^{n_r} \frac{1}{2\pi} \int_{-\pi}^{\pi} |d_k(e^{j\omega})|^2 d\omega}{\sum_{k=1}^{n_r} \sigma_k^2} \\ &= \frac{\frac{1}{2\pi} \int_{-\pi}^{\pi} ||c(e^{j\omega})||^2 d\omega}{\mathrm{tr}\, S_v(e^{j\omega})} = \beta, \end{aligned}$$

since $\mathrm{tr}\, S_v(e^{j\omega}) = \mathrm{tr}\, \Sigma$ is a constant in this case.

Interestingly, if the noise eigenvalues σ_k^2 are allowed to vary with ω, it is possible to have $\gamma_{\rm am} < \beta$. To see this, consider the $n_t = 1$ case in which $S_v(e^{j\omega})$ is not constant; assume that S_v is normalized so that the noise power is 1. If the channel transfer function is proportional to the noise PSD, i.e.

$$c(e^{j\omega}) = \alpha S_v(e^{j\omega}),$$

then, on the one hand,

$$\gamma_{\rm am} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\alpha|^2 S_v^2(e^{j\omega})}{S_v(e^{j\omega})} d\omega = |\alpha|^2 \frac{1}{2\pi}\int_{-\pi}^{\pi} S_v(e^{j\omega}) d\omega = |\alpha|^2,$$

while on the other hand

$$\beta = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} |\alpha|^2 S_v^2(e^{j\omega}) d\omega}{\frac{1}{2\pi}\int_{-\pi}^{\pi} S_v(e^{j\omega}) d\omega} = |\alpha|^2 \frac{1}{2\pi}\int_{-\pi}^{\pi} S_v^2(e^{j\omega}) d\omega > |\alpha|^2,$$

since the Cauchy-Schwarz inequality shows that $\frac{1}{2\pi}\int_{-\pi}^{\pi} S_v^2(e^{j\omega}) d\omega \geq 1$, with equality holding iff $S_v(e^{j\omega})$ is constant. ■

ACKNOWLEDGEMENT

The author thanks R. K. Martin at the Air Force Institute of Technology for his useful comments on an early draft of this paper.

REFERENCES

[1] N. Al-Dhahir, 'Optimized-transmitter reduced-complexity MMSE-DFE under finite-length constraints', Ph.D. thesis, Stanford University, 1994.

[2] N. Al-Dhahir, 'FIR channel shortening equalizers for MIMO ISI channels', *IEEE Trans. Communications*, vol. 49, pp. 213-218, Feb. 2001.

[3] N. Al-Dhahir and J. M. Cioffi, 'Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: a unified approach', *IEEE Trans. Info. Theory*, vol. 42, pp. 903-915, May 1996.

[4] B. D. O. Anderson and J. B. Moore, *Optimal Filtering.* Englewood Cliffs, NJ: Prentice-Hall, 1979.

[5] J. S. Chow, J. C. Tu and J. M. Cioffi, 'A discrete multitone transceiver system for HDSL applications', *IEEE J. Select. Areas Commun.*, vol. 9, pp. 895-908, Aug. 1991.

[6] D. Daly, C. Heneghan and A. D. Fagan, 'Minimum mean-squared error impulse response shortening for discrete multitone transceivers', *IEEE Trans. Signal Processing*, vol. 52, pp. 301-306, Jan. 2004.

[7] A. Duel-Hallen and C. Heegard, 'Delayed decision-feedback sequence estimation', *IEEE Trans. Communications*, vol. 37, pp. 428-436, May 1989.

[8] M. V. Eyuboğlu and S. U. Qureshi, 'Reduced-state sequence estimation with set partitioning and decision feedback', *IEEE Trans. Communications*, vol. 36, pp. 13-20, Jan. 1988.

[9] D. Falconer and F. Magee, 'Adaptive channel memory truncation for maximum likelihood sequence estimation', *Bell Syst. Tech. J.*, vol. 52 no. 9, pp. 1541-1562, Nov. 1973.

[10] G. Ginis and J. M. Cioffi, 'Vectored transmission for digital subscriber line systems', *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1085-1104, Jun. 2002.

[11] G. H. Golub and C. F. Van Loan, *Matrix Computations, 3rd ed.* Baltimore, MD: Johns Hopkins, 1996.

[12] R. López-Valcarce, 'Realizable linear and decision feedback equalizers: properties and connections', *IEEE Trans. Signal Processing*, vol. 52, pp. 757-773, Mar. 2004.

[13] R. López-Valcarce and S. Dasgupta, 'A new proof for the stability of equation-error models', *IEEE Signal Processing Letters*, vol. 6, pp. 148-150, Jun. 1999.

[14] R. K. Martin, M. Ding, B. L. Evans and C. R. Johnson Jr., 'Infinite length results and design implications for time-domain equalizers', *IEEE Trans. Signal Processing*, vol. 52, pp. 297-301, Jan. 2004.

[15] P. Melsa, R. Younce and C. Rohrs, 'Impulse response shortening for discrete multitone transceivers', *IEEE Trans. Communications*, vol. 44, pp. 1662-1672, Dec. 1996.

[16] M. Miranda and P. Tilli, 'Asymptotic spectra of Hermitian block Toeplitz matrices and preconditioning results', *SIAM J. Matrix Anal. Applic.*, vol. 21, no. 3, pp. 867-881, 2000.

[17] B. N. Parlett, *The symmetric eigenvalue problem*, Prentice Hall, 1980.

[18] A. J. Paulraj, D. A. Gore, R. U. Nabar and H. Bölcskei, 'An overview of MIMO communications – A key to gigabit wireless', *Proc. IEEE*, vol. 92, pp. 198-218, Feb. 2004.

[19] G. G. Raleigh and J. M. Cioffi, 'Spatio-temporal coding for wireless communication', *IEEE Trans. Communications*, vol. 46, pp. 357-366, Mar. 1998.

[20] A. Tkacenko and P. P. Vaidyanathan, 'Eigenfilter design of MIMO equalizers for channel shortening', *Proc. IEEE ICASSP '02*, vol. 3, pp. 2361-2364, 2002.

[21] P. P. Vaidyanathan, *Multirate systems and filterbanks*, Prentice Hall, 1993.

[22] D. Wübben and K.-D. Kammeyer, 'Impulse shortening and equalization of frequency-selective MIMO channels with respect to layered space-time architectures', *Signal Processing*, vol. 83, pp. 1643-659, Aug.. 2003.

[23] G. Ysebaert, K. Van Acker, M. Moonen and B. De Moor, 'Constraints in channel shortening equalizer design for DMT-based systems', *Signal Processing*, vol. 83, pp. 641-648, Mar. 2003.