ON THE SENSITIVITY OF TRANSMIT WIENER FILTERING FOR BROADCAST CHANNELS WITH RESPECT TO CHANNEL ESTIMATION ERRORS

Despoina Tsipouridou and Athanasios P. Liavas

Department of Electronic and Computer Engineering
Technical University of Crete
73100, Kounoupidiana, Chania, Greece
E-mail: `(despoina, liavas)@telecom.tuc.gr`

ABSTRACT

We consider the behavior of the transmit Wiener filter for broadcast channels with non-cooperating users under channel uncertainties. We derive a second-order approximation to the excess mean-square error (EMSE) induced by using a channel estimate as if it were the true channel. In the high SNR cases, we develop a simple approximation to the EMSE. It turns out that the EMSE is proportional to the minimum mean square error with the proportionality factor determined by the total transmit power and the length of the training block.

***Index Terms*—** MIMO systems, broadcast channels, pre-equalization, Wiener filtering.

1. INTRODUCTION

Joint optimization of transmit and receive filters for combatting frequency selectivity and/or interstream interference in MIMO or multiuser systems has been extensively studied (see, for example, [1] and the references therein).

If we want to keep the mobile units as simple as possible, then we may consider separate transmit or receive processing. The transmit matched filter (TxMF), the transmit zero-forcing filter (TxZF) and the transmit Wiener filter (TxWF) are three linear precoding (or pre-equalization) structures that combat frequency selectivity and/or inter-stream interference and keep the receivers simple, because the only assumed receiver processing is a scalar scaling (see [2], [1] and the references therein). This is particularly appealing in the broadcast scenario, where we want to keep the receivers of non-cooperative users as simple as possible.

The TxWF outperforms the two other structures in terms of mean-square error (MSE) and bit-error rate (BER). If the channel matrix and the input and noise second-order statistics (SOS) are perfectly known at the transmitter (due to, for example, TDD or feedback information channel), then the TxWF can be computed. If the channel and/or the noise SOS are unknown at the transmitter, as it is usually the case, then a common approach towards the design of the TxWF is to estimate the unknown quantities and use the estimates as if they were the true quantities.

In this work, we consider the case where the channel estimate is used for the design of the TxWF as if it were the true channel and we develop a second-order approximation to the associated excess MSE (EMSE). Furthermore, in the high SNR cases, we derive a simple approximation to the EMSE which provides significant insight into the operation of the TxWF with a channel estimate. It turns out that, in the high SNR cases, the EMSE is approximately proportional to the minimum MSE (MMSE) with the proportionality factor determined by the total transmit power and the length of the training block.

Notation: Superscripts ^{T}, ^{H} and * denote transpose, conjugate transpose and elementwise conjugation, respectively. $\mathrm{Re}\{\cdot\}$ extracts the real part of a complex number, symbol $\otimes$ denotes the Kronecker product and $\mathrm{vec}(\cdot)$ denotes the vectorization operator.

The structure of the paper is as follows. In Section 2, we present the derivation of the TxWF, assuming that the CSI and the noise SOS are known at the transmitter and we give the expression of the MMSE in this case [1]. In Section 3, we develop a second-order approximation to the excess MSE, assuming CSI estimation errors. In Section IV, we present simulations that support our theoretical results.

2. THE TRANSMIT WIENER FILTER

We consider the pre-equalized, baseband-equivalent, discrete-time broadcast channel, with n_t transmit antennas and n_r non-cooperative receivers (with $n_r \leq n_t$), depicted in Fig. 1 and described by the expression

$$\hat{\mathbf{s}} = HP\mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{s}$ is the $n_r \times 1$ input signal, P is the $n_t \times n_r$ precoding matrix, H is the $n_r \times n_t$ channel matrix and $\mathbf{n}$ is the $n_r \times 1$ additive channel noise. The i-th element of vector $\mathbf{s}$ is the symbol intended for the i-th user. The input and noise

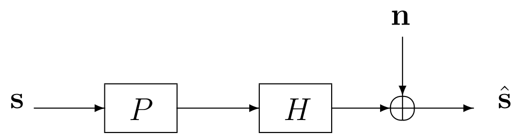

Fig. 1. System model

vectors, $\mathbf{s}$ and $\mathbf{n}$, are assumed to be independent, complex-valued, circular, zero-mean, with covariance matrices $R_s = I_{n_r}$ and $R_n = \sigma_n^2 I_{n_r}$, respectively.

Our aim is to find the TxWF P and the scalar β (with $\beta \neq 1$) that minimize the function

$$\mathrm{mse}(P, \beta) := \mathbf{E}\left[\|\mathbf{s} - \beta^{-1}\hat{\mathbf{s}}\|_2^2\right] \tag{2}$$

subject to the transmit power constraint

$$\mathbf{E}\left[\|P\mathbf{s}\|_2^2\right] = \mathrm{E}_{\mathrm{tr}}. \tag{3}$$

Function $\mathrm{mse}(\cdot)$ can be analytically expressed as

$$\begin{aligned}\mathrm{mse}(P, \beta) = \mathrm{tr}(I_{n_r}) - 2\beta^{-1}\mathrm{Re}\left\{\mathrm{tr}(HP)\right\} \\ + \beta^{-2}\mathrm{tr}(HPP^H H^H) + \beta^{-2}\mathrm{tr}(R_n).\end{aligned} \tag{4}$$

The solution of this constrained optimization problem is given by [2]

$$\beta_o = \sqrt{\frac{\mathrm{E}_{\mathrm{tr}}}{\mathrm{tr}(\tilde{P}\tilde{P}^H)}} \tag{5}$$

and $P_o = \beta_o \tilde{P}_o$, where

$$\tilde{P}_o := \left(H^H H + \alpha I_{n_t}\right)^{-1} H^H \tag{6}$$

and

$$\alpha := \frac{\mathrm{tr}(R_n)}{\mathrm{E}_{\mathrm{tr}}}. \tag{7}$$

In [1], quantity α has been defined as inverse SNR.

Using the optimal values P_o and β_o in (4), it can be shown that

$$\begin{aligned}\mathrm{mse}(P_o, \beta_o) &= \mathrm{tr}\big(I_{n_r}\big) - 2\,\mathrm{Re}\left\{\mathrm{tr}\big(\tilde{P}_o H\big)\right\} \\ &\quad + \mathrm{tr}\big(H\tilde{P}_o\tilde{P}_o^H H^H\big) + \alpha\,\mathrm{tr}\big(\tilde{P}_o\tilde{P}_o^H\big) \\ &=: \mathrm{MSE}(\tilde{P}_o).\end{aligned} \tag{8}$$

In the high SNR cases, an approximation that will be useful in the sequel is [7]

$$\mathrm{tr}\big(\tilde{P}_o\tilde{P}_o^H\big) \approx \frac{1}{\alpha}\,\mathrm{MMSE}. \tag{9}$$

3. COMPUTATION OF THE EXCESS MSE

In this section, we develop a second-order approximation to the excess MSE induced by channel estimation errors.

We denote the channel estimate $\hat{H}$ and we define the channel estimation error as

$$\Delta H := \hat{H} - H. \tag{10}$$

We assume that $\mathrm{vec}(\Delta H)$ is zero-mean, complex-valued, circular, with covariance matrix

$$R_{\mathrm{vec}(\Delta H)} := \mathbf{E}\left[\mathrm{vec}(\Delta H)\mathrm{vec}^H(\Delta H)\right] = \Sigma. \tag{11}$$

We denote with $\hat{\tilde{P}}$ and $\hat{\beta}$ the scaled TxWF and the Wiener scalar computed by using the channel estimate as if it were the true channel. The corresponding TxWF is $\hat{P} := \hat{\beta}\hat{\tilde{P}}$. The MSE associated with $\hat{P}$ and $\hat{\beta}$ is

$$\mathrm{mse}\big(\hat{P}, \hat{\beta}\big) = \mathrm{MSE}\big(\hat{\tilde{P}}\big).$$

Expansion of function MSE($\cdot$) around $\tilde{P}_o$, yields

$$\mathrm{MSE}(\hat{\tilde{P}}) = \mathrm{MSE}(\tilde{P}_o) + \mathrm{tr}(\Delta\tilde{P}^H \mathrm{MSE}''(\tilde{P}_o)\Delta\tilde{P}) \tag{12}$$

where $\Delta\tilde{P} := \hat{\tilde{P}} - \tilde{P}_o$ and $\mathrm{MSE}''(\tilde{P}_o)$ is the second derivative of the function MSE, evaluated at the point $\tilde{P}_o$. From (8), it becomes obvious that

$$\mathrm{MSE}''(\tilde{P}_o) = H^H H + \alpha I_{n_t}. \tag{13}$$

If we take expectation with respect to estimation errors in (12), we obtain

$$\begin{aligned}\mathrm{EMSE}(\hat{\tilde{P}}) &:= \mathbf{E}\left[\mathrm{MSE}\big(\hat{\tilde{P}}\big) - \mathrm{MSE}\big(\tilde{P}_o\big)\right] \\ &= \mathbf{E}\left[\mathrm{tr}\big(\Delta\tilde{P}^H\mathrm{MSE}''(\tilde{P}_o)\Delta\tilde{P}\big)\right] \\ &= \mathbf{E}\left[\mathrm{tr}\left(\Delta\tilde{P}^H(H^H H + \alpha I_{n_t})\Delta\tilde{P}\right)\right].\end{aligned} \tag{14}$$

In order to compute the EMSE, we must express $\Delta\tilde{P}$ in terms of the channel estimation error ΔH.

We start by providing a first-order approximation to $\Delta\tilde{P}$ with respect to channel estimation error ΔH. If we use in (6) estimate $\hat{H}$ as if it were the true channel, then we compute the scaled pre-coding matrix

$$\hat{\tilde{P}} = \left(\hat{H}^H\hat{H} + \alpha I_{n_t}\right)^{-1}\hat{H}^H. \tag{15}$$

If we ignore products of error terms inside the parenthesis, we obtain

$$\hat{\tilde{P}} = \Big(H^H H + \alpha I_{n_t} + \underbrace{H^H\Delta H + \Delta H^H H}_{K_\Delta}\Big)^{-1}(H^H + \Delta H^H). \tag{16}$$

Using the first-order approximation [4, p. 131]

$$(A + \Delta A)^{-1} = A^{-1} - A^{-1}\Delta A A^{-1} \tag{17}$$

definition (6) and ignoring products of error terms, we obtain

$$\hat{\tilde{P}} = \tilde{P}_o - (H^H H + \alpha I_{n_t})^{-1}(K_\Delta \tilde{P}_o - \Delta H^H).$$

Thus, a first-order approximation to $\Delta\tilde{P}$ is

$$\Delta\tilde{P} = -\underbrace{(H^H H + \alpha I_{n_t})^{-1}}_{\mathcal{A}}\underbrace{(K_\Delta \tilde{P}_o - \Delta H^H)}_{\Delta}. \tag{18}$$

Having computed a first-order approximation of $\Delta\tilde{P}$ as a function of the channel estimation error ΔH, we may proceed to a second-order approximation to the EMSE as follows:

$$\begin{aligned}
\text{EMSE}(\hat{\tilde{P}}) &= \mathbf{E}\left[\text{tr}\left(\Delta\tilde{P}^H(H^H H + \alpha I_{n_t})\Delta\tilde{P}\right)\right] \\
&\overset{(18)}{=} \mathbf{E}\left[\text{tr}\left(\Delta^H \mathcal{A}\,\Delta\right)\right] \\
&= \mathbf{E}\left[\text{tr}\left(\mathcal{A}\,\Delta\, I_{n_r}\,\Delta^H\right)\right] \\
&\overset{(a)}{=} \mathbf{E}\left[\text{vec}^H(\Delta)\big(I_{n_r}\otimes\mathcal{A}\big)\text{vec}(\Delta)\right] \\
&= \text{tr}\left(\left(I_{n_r}\otimes\mathcal{A}\right)\mathbf{E}\left[\text{vec}(\Delta)\text{vec}^H(\Delta)\right]\right)
\end{aligned} \tag{19}$$

where at point (a) we used expression [3]

$$\text{tr}(ABCD) = \text{vec}^T(D^T)\,(C\otimes A)\,\text{vec}(B).$$

¿From the definitions of Δ in (18) and K_Δ in (16), we obtain

$$\begin{aligned}
\text{vec}(\Delta) &= \text{vec}(H^H \Delta H \tilde{P}_o) + \text{vec}\left(\Delta H^H\left(H\tilde{P}_o - I_{n_r}\right)\right) \\
&= \underbrace{(\tilde{P}_o^T \otimes H^H)}_{\mathcal{T}_1}\text{vec}(\Delta H) + \\
&\qquad \underbrace{\left((\tilde{P}_o^T H^T - I_{n_r})\otimes I_{n_t}\right)}_{\mathcal{T}_2}\text{vec}(\Delta H^H)
\end{aligned} \tag{20}$$

where we made use of

$$\text{vec}(ABC) = (C^T\otimes A)\,\text{vec}(C).$$

Using the commutation matrix $K_{n_t n_r}$ [3, p. 9], we obtain

$$\text{vec}(\Delta H^H) = K_{n_t n_r}\text{vec}(\Delta H^*)$$

yielding

$$\text{vec}(\Delta) = \mathcal{T}_1\text{vec}(\Delta H) + \mathcal{T}_2 K\,\text{vec}(\Delta H^*)$$

where, for notational simplicity, the commutation matrix is denoted as K. Thus, using (11), we obtain

$$\text{EMSE}(\hat{\tilde{P}}) = \text{tr}\left((I_{n_r}\otimes\mathcal{A})\left(\mathcal{T}_1\,\Sigma\,\mathcal{T}_1^H + \mathcal{T}_2 K\Sigma^* K^H \mathcal{T}_2^H\right)\right).$$

After some algebra, using the definitions of the involved quantities and properties of the Kronecker product and the commutation matrix, the EMSE can be expressed as [7]

$$\text{EMSE}(\hat{\tilde{P}}) = \mathbf{T}_1 + \mathbf{T}_2 \tag{21}$$

where

$$\begin{aligned}
\mathbf{T}_1 &:= \text{tr}\left((I_{n_r}\otimes\mathcal{A})\,\mathcal{T}_1\Sigma\mathcal{T}_1^H\right) \\
&= \text{tr}\left(\left(\tilde{P}_o^*\tilde{P}_o^T\otimes H\mathcal{A}H^H\right)\Sigma\right)
\end{aligned} \tag{22}$$

and

$$\begin{aligned}
\mathbf{T}_2 &:= \text{tr}\left((I_{n_r}\otimes\mathcal{A})\,\mathcal{T}_2 K\Sigma^* K^H\mathcal{T}_2^H\right) \\
&= \text{tr}\left(\left(\mathcal{A}\otimes(H^*\tilde{P}_o^* - I_{n_r})(\tilde{P}_o^T H^T - I_{n_r})\right)\,\Sigma^*\right).
\end{aligned} \tag{23}$$

3.1. Simplifications in the high SNR cases

In this subsection, we assume that the SNR is sufficiently high and we derive a simple approximation for the EMSE.

Assuming that we have used optimal training, it can be shown that the channel estimation covariance matrix is given by [5, p.175]

$$\Sigma = \frac{\sigma_n^2}{N_{tr}} I_{n_t n_r} \tag{24}$$

where N_{tr} is the length of the training block.

For $\alpha \ll \lambda_{\min}(H^H H)$, an approximation that will prove useful in the sequel is ([4, p. 138])

$$\begin{aligned}
\text{tr}\left(H\mathcal{A}H^H\right) &= \sum_{i=1}^{n_r}\frac{\lambda_i\left(H^H H\right)}{\lambda_i\left(H^H H\right)+\alpha} \\
&\approx \text{tr}\left(I_{n_r}\right).
\end{aligned} \tag{25}$$

Starting with $\mathbf{T}_1$ in (22), we obtain

$$\begin{aligned}
\mathbf{T}_1 &\overset{(24)}{=} \frac{\sigma_n^2}{N_{tr}}\text{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\otimes H\mathcal{A}H^H\right) \\
&\overset{(25)}{\approx} \frac{\sigma_n^2}{N_{tr}}\text{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\right)\text{tr}\left(I_{n_r}\right) \\
&= \frac{n_r\,\sigma_n^2}{N_{tr}}\text{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\right) \\
&\overset{(9)}{\approx} \frac{n_r\,\sigma_n^2}{N_{tr}}\frac{1}{\alpha}\text{MMSE}.
\end{aligned} \tag{26}$$

In order to compute $\mathbf{T}_2$, we use an expression analogous to (25)

$$\lambda_i\left((H^*\tilde{P}_o^* - I_{n_r})(\tilde{P}_o^T H^T - I_{n_r})\right) = \frac{a^2}{\left(\lambda_i\left(H^H H\right)+\alpha\right)^2}.$$

For high SNR, the above eigenvalues go to zero, yielding

$$\text{tr}\left((H^*\tilde{P}_o^* - I_{n_r})(\tilde{P}_o^T H^T - I_{n_r})\right) \approx 0. \tag{27}$$

Thus,

$$\begin{aligned}
\mathbf{T}_2 &\overset{(24)}{=} \frac{\sigma_n^2}{N_{tr}}\text{tr}\left(\mathcal{A}\right)\text{tr}\left((H^*\tilde{P}_o^* - I_{n_r})(\tilde{P}_o^T H^T - I_{n_r})\right) \\
&\overset{(27)}{\approx} 0.
\end{aligned} \tag{28}$$

Table I
Elements of channel matrix H

-0.0648+0.0388*j	-0.0547+0.2974*j	0.2588-0.0954*j
0.4186+0.2072*j	-0.5157-0.3955*j	0.3897-0.1856*j

We conclude that term $\mathbf{T}_2$ is negligible compared to$\mathbf{T}_1$, for sufficiently high SNR. Combining expressions (21), (26) and (28), we obtain

$$\begin{aligned}\mathrm{EMSE}(\hat{\hat{P}}) &\approx \frac{n_r\,\sigma_n^2}{N_{tr}}\,\frac{1}{\alpha}\,\mathrm{MMSE}\\ &\overset{(7)}{=} \frac{\mathrm{E_{tr}}}{N_{tr}}\,\mathrm{MMSE}.\end{aligned}$$

Thus, in the high SNR cases

$$\boxed{\mathrm{EMSE}(\hat{\hat{P}}) \approx \frac{\mathrm{E_{tr}}}{N_{tr}}\,\mathrm{MMSE}.} \tag{29}$$

We observe that, in the high SNR cases, the EMSE is (approximately) proportional to the MMSE, with the proportionality factor being the ratio of the transmit power, $\mathrm{E_{tr}}$, to the length of the training block used for channel estimation, N_{tr}.

4. SIMULATIONS

In this section we support our theoretical results with simulations. We consider a broadcast system with $n_t = 3$ transmit antennas and $n_r = 2$ non-cooperative receivers.

The filtering matrix H is a realization of a 2×3 random matrix, with elements i.i.d. complex, circular, zero-mean Gaussian random variables, normalized so that $\|H\|_F^2 = 1$. Its elements are given in Table I.

We set the transmit power $\mathrm{E_{tr}} = n_t$. We assume that the training block is composed of $N_{\mathrm{tr}} = 20$ columns.

In Fig. 2, we plot the experimentally computed EMSE, the theoretical second-order approximation (21), and approximation (29). We observe that the experimental and theoretical EMSE values practically coincide for SNR higher than 5 dB, while expression (29) is a good approximation to the EMSE, especially at high SNR. Analogous results have been observed in extensive simulations.

In [7], we have also considered the sensitivity of the TxWF with respect to noise second-order statistics estimation errors.

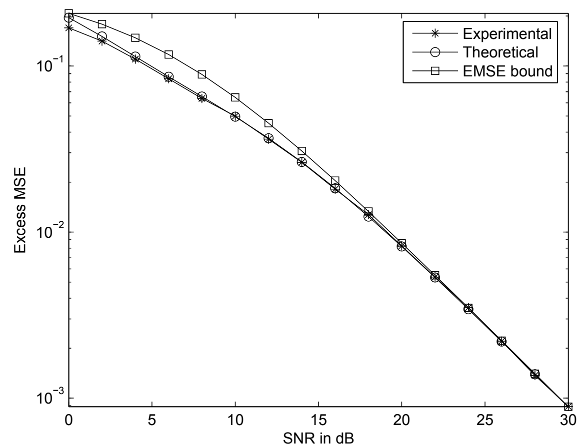

Fig. 2. Experimentally computed EMSE, theoretical second-order approximation (21), and high-SNR approximation (29).

5. REFERENCES

[1] M. Joham, W. Utschick, and J. A. Nossek, "Linear Transmit Processing in MIMO Communications Systems," *IEEE Trans. Signal Processing*, Vol. 53, no. 11, pp. 2700-2712, August 2005.

[2] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener Filter," Tech. Rep. TUM-LNS-TR-02-1, Munich University of Technology, January 2002.

[3] H. Luetkepohl. *Handbook of matrices*. John Wiley, 1996.

[4] G. Stewart and J.-G. Sun. *Matrix perturbation theory.* Academic Press, 1990.

[5] E. G. Larsson and P. Stoica. *Space-Time Block Coding for Wireless Communications.* Cambridge University Press, 2003.

[6] A. Sayed. *Fundamentals of adaptive filtering.* IEEE-Wiley, 2003.

[7] D. Tsipouridou and A. P. Liavas, "On the sensitivity of the transmit Wiener filter with respect to channel and noise second-order statistics estimation errors," in preparation.