The proposed algorithms produce simultaneously stable tracking and size estimates converging to the true parameters. The DA procedure provides the most accurate results, since it processes the cumulative (in a window) measurement information which increases the computational time. The relative computational time IMM-DA:MKF:IMM-PF corresponds approximately to the proportions: 18:6:1. It should be noted that the PF involves an additional "artificial" noise, necessary for prediction. The proper choice of noise parameters can lead to a good result. However, the PF aspect ratio RMSE slowly increases over time [Fig. 3(d)]. This is observed over various scenarios and different sample sizes, as shown in Fig. 4. A similar tendency is indicated also in [9]. Taking this fact into consideration, we may conclude, that the MKFm provides a reasonable compromise between accuracy and computational time. The model validation scheme, incorporated within the MKFm, gives an additional size type information: if we are interested in the size type, which is not the true one, the Kolmogorov–Smirnov test certainly rejects this hypothesis. For example, if we want to check the hypothesis $\boldsymbol{\theta}_3 = \boldsymbol{\theta}_{\rm true}$, the estimated test statistic $ktest2 = 8$ definitely exceeds a 5% critical value of 1.36, since $\boldsymbol{\theta}_{\rm true} \equiv \boldsymbol{\theta}_2$.

VIII. Conclusion

A suboptimal solution to the problem of extended object tracking is proposed in this correspondence. MC algorithms (DA, MKFm and PF) are developed for the object extent parameter estimation, based on positional and along-range object extent measurements. The kinematic states are estimated with an IMM filter and with a MKFm, respectively. The approach of separation of states from parameters is implemented in the IMM-DA and IMM-PF. The overall state vector has a decreased dimension compared with the joint state-parameter estimation, the type of maneuver can be identified relatively quickly, and the kinematic states are estimated with small peak dynamic errors. The developed techniques offer a reasonable trade-off between accuracy and computational time and successfully deal with the complex target-observer geometry.

Acknowledgment

The authors would like to thank the anonymous reviewers and the Associate Editor for their constructive comments and suggestions.

References

[1] R. Chen and J. Liu, "Mixture Kalman filters," *J. Roy. Statist. Soc. B*, vol. 62, pp. 493–508, 2000.
[2] K. Gilholm and D. Salmond, "Spatial distribution model for tracking extended objects," *Proc. Inst. Elect. Eng.—Radar, Sonar Navig.*, vol. 152, no. 5, pp. 364–371, 2005.
[3] C. Andrieu, A. Doucet, and V. Tadić, "On-line parameter estimation in general state space models," in *Proc. 44th IEEE Conf. Decision Control*, 2005, pp. 332–337, Paper MoA10.4.
[4] C. Andrieu, A. Doucet, S. Singh, and V. Tadić, "Particle methods and change detection, system identification, and control," *Proc. IEEE*, vol. 92, no. 3, pp. 423–438, Mar. 2004.
[5] J. Diebolt and C. Robert, "Estimation of finite mixture distributions through Bayesian sampling," *J. Roy. Stat. Soc. B*, vol. 56, no. 4, pp. 363–375, 1994.
[6] J. Vermaak, N. Ikoma, and S. Godsill, "Sequential Monte Carlo framework for extended object tracking," *Proc. Inst. Elect. Eng.—Radar, Sonar Navig.*, vol. 152, no. 5, pp. 353–363, 2005.
[7] D. Salmond and M. Parr, "Track maintenance using measurements of target extent," *Proc. Inst. Elect. Eng.—Radar Sonar Navig.*, vol. 150, no. 6, pp. 389–395, 2003.
[8] Y. Boers and J. N. Driessen, "Track before detect approach for extended objects," in *Proc. 9th Int. Conf. Inf. Fusion (ISIF)*, Florence, Italy, Jul. 2006, [CD-ROM].
[9] B. Ristic and D. Salmond, "A study of a nonlinear filtering problem for tracking an extended target," in *Proc. 7th Int. Conf. Inf. Fusion*, 2004, pp. 503–509.
[10] D. Angelova and L. Mihaylova, "A Monte Carlo algorithm for state and parameter estimation of extended targets," in *LNCS Proceedings from the Sixth International Conference*, V. N. Alexandrov, G. Dick van Albada, P. M. A. Sloot, and J. Dongarra, Eds. Berlin, Germany: Springer-Verlag, 2006, vol. 3993, pt. III, pp. 624–631.
[11] V. Jilkov, X. R. Li, and D. Angelova, "Estimation of Markovian jump systems with unknown transition probabilities through Bayesian sampling," in *LNCS Proceedings from the Fifth International Conference on Numerical Methods and Applications*, I. Dimov, I. Lirkov, S. Margenov, and Z. Zlatev, Eds. Berlin, Germany: Springer-Verlag, 2003, vol. 2542, pp. 307–315.
[12] J. Dezert, "Tracking manoeuvring and bending extended target in cluttered environment," in *Proc. SPIE*, 1998, vol. 3373, pp. 283–294.
[13] A. Doucet, N. Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, 2001.
[14] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation: Theory, Algorithms, and Software*. New York: Wiley, 2001.
[15] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Norwood, MA: Artech House, 2004.
[16] J. Vermaak, C. Andrieu, A. Doucet, and S. Godsill, "Particle methods for Bayesian modeling and enhancement of speech signals," *IEEE Trans. Speech Audio Process.*, vol. 10, no. 3, pp. 173–185, 2002.

On the Sensitivity of the Transmit MIMO Wiener Filter With Respect to Channel and Noise Second-Order Statistics Uncertainties

Despoina Tsipouridou and Athanasios P. Liavas, *Member, IEEE*

***Abstract*—We consider the sensitivity of the transmit multiple-input multiple-output (MIMO) Wiener filter with respect to channel and noise second-order statistics (SOS) uncertainties. Using results from matrix perturbation theory, we derive second-order approximations to the excess mean-square error (EMSE) induced by using the channel or noise SOS estimates as if they were the true quantities. Assuming optimal training and sufficiently high signal-to-noise ratio (SNR), we develop simple and informative approximations to the EMSE, which indicate that the channel estimation errors are much more significant than the noise SOS estimation errors. Uncertainties due to channel time variations induce EMSE that increases with increasing SNR and asymptotically tends to a constant value.**

***Index Terms*—Multiple-input multiple-output (MIMO) systems, pre-equalization, Wiener filtering.**

I. Introduction

Joint optimization of transmit and receive filters for combatting frequency selectivity and/or interstream interference in multiple-input multiple-output (MIMO) or multiuser systems has been extensively studied (see, for example, [1] and the references therein). In order to keep the mobile units as simple as possible, 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 pre-equalization (or precoding) structures that combat frequency selectivity and/or interstream interference and keep the receivers simple, because the only processing required at the receiver is a scalar scaling [1], [2].

The TxWF, which outperforms the two other structures in terms of mean-square error (MSE) and bit-error rate (BER) [1], can be computed if the channel and the input and noise second-order statistics (SOS) are perfectly known at the transmitter. This may happen, for

Manuscript received February 26, 2007; revised July 6, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Philippe Loubaton.

The authors are with the Department of Electronic and Computer Engineering, Technical University of Crete, 73100 Kounoupidiana, Chania, Greece (e-mail: despoina@telecom.tuc.gr; liavas@telecom.tuc.gr).

Digital Object Identifier 10.1109/TSP.2007.907840

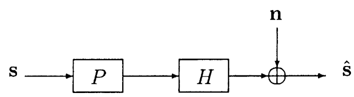

Fig. 1. System model.

example, in time-division duplex (TDD) systems or systems with a feedback information channel. 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 then use the estimates as if they were the true quantities. Estimation errors and/or time variations introduce uncertainties in the estimated quantities and induce excess MSE (EMSE) leading to TxWF performance degradation.

In this correspondence, we consider the sensitivity of the TxWF with respect to channel and noise SOS uncertainties and we develop second-order approximations to the associated EMSEs. While the general expressions are complicated and difficult to interpret, we are able to derive simple and informative EMSE approximations for the high signal-to-noise ratio (SNR) cases. It turns out that the EMSE due to channel estimation errors is proportional to the minimum MSE (MMSE), while the EMSE due to noise SOS estimation errors is proportional to the squared noise variance. On the other hand, the EMSE due to channel time variations increases for increasing SNR and asymptotically reaches a constant value.

The rest of the correspondence is structured as follows. In Section II, we compute the TxWF assuming that the channel and the noise SOS are known at the transmitter and we give an expression for the MMSE [1]. Also, we derive a general expression for the EMSE due to channel or noise SOS uncertainties. In Sections III and IV, we develop second-order approximations to the EMSE assuming channel and noise SOS uncertainties, respectively. In Section V, we present simulations that support our theoretical results and we conclude the correspondence in Section VI.

Notation: Superscripts ^{T}, ^{H}, and * denote transpose, conjugate transpose and elementwise conjugation, respectively. $\mathrm{Re}\{\cdot\}$ extracts the real part of a complex number, $\otimes$ denotes the Kronecker product and $\mathrm{vec}(\cdot)$ denotes the vectorization operator. The eigenvalues of matrix A are denoted as $\lambda_i(A)$. If A is an $n \times n$ semi-positive-definite matrix, then its eigenvalues are ordered such that $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$.

II. The Transmit Wiener Filter

A. The System Model

We consider the pre-equalized, baseband-equivalent, discrete-time frequency-flat MIMO system, with n_t transmit and n_r receive antennas (with $n_r \leq n_t$), depicted in Fig. 1. This system is 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 $n_t \times n_r$ pre-equalization matrix, H is the $n_r \times n_t$ channel matrix, and $\mathbf{n}$ is the $n_r \times 1$ additive channel noise. The input and noise vectors, $\mathbf{s}$ and $\mathbf{n}$, are assumed to be complex-valued, independent, circular, with covariance matrices $R_s = I_{n_r}$ and $R_n = \sigma_n^2 I_{n_r}$, respectively; furthermore, the noise is assumed to be Gaussian. This model is particularly suitable for the broadcast scenario, where the users cannot cooperate in order to combat interstream interference and, thus, the need for pre-equalization is imperative. In this case, the ith element of $\mathbf{s}$ is the symbol intended for the ith user.

B. Computation of the TxWF

Our aim is to compute the TxWF P and the scalar β that minimize the cost function [2]

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

subject to the transmit power constraint

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

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

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

The optimal values for this constrained optimization problem are [2]

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

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

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

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 the MMSE is

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

C. Channel and Noise SOS Uncertainties

In the above development, we assumed that the channel matrix H and the noise covariance matrix R_n are perfectly known at the transmitter. Perhaps, the easiest way to obtain estimates of these quantities is through training. In frequency division duplex (FDD) systems, the estimates can be computed at the receiver and communicated to the transmitter via a feedback channel, while in TDD systems they can be computed at the transmitter. In this work, we consider channel estimation at the receiver, i.e., FDD systems. Of course, analogous results hold for TDD systems.

The channel estimate $\hat{H}$ at the transmitter may suffer from two basic sources of uncertainty, namely, estimation errors and time variations. We define the channel error or mismatch as

$$\Delta H := \hat{H} - H \tag{9}$$

and we assume that $\mathrm{vec}(\Delta H)$ is complex valued, circular with

$$R_{\mathrm{vec}(\Delta H)} := \mathcal{E}\left[\mathrm{vec}(\Delta H)\mathrm{vec}^H(\Delta H)\right] = \Sigma. \tag{10}$$

Next, we compute Σ for the two cases of interest.

1) *Channel estimation errors:* In this case, we assume that the channel is time invariant and we estimate it using training. If we denote the $n_t \times N_{\mathrm{tr}}$ training block as S_{tr} and the corresponding channel output as Y_{tr}, then the maximum-likelihood (ML) channel estimate is [6, p. 174]

$$\hat{H} = Y_{\mathrm{tr}} S_{\mathrm{tr}}^H \left(S_{\mathrm{tr}} S_{\mathrm{tr}}^H \right)^{-1}. \tag{11}$$

Optimal channel estimates are obtained for semi-unitary training matrix S_{tr}, i.e., $S_{\mathrm{tr}} S_{\mathrm{tr}}^H \propto I_{n_t}$. The corresponding channel estimation error covariance matrix is given by [6, p. 175]

$$\Sigma = \frac{\sigma_n^2}{N_{\mathrm{tr}}} I_{n_t n_r}. \tag{12}$$

2) *Channel time variations:* In this case, we assume that uncertainties due to channel time variations dominate those due to channel estimation errors (i.e., we assume that the channel estimate is perfect and we focus on channel time variations).[1] We denote with $H = H_t$ the true channel at time instant t and with $\hat{H} = H_{t-\tau}$ the outdated channel version at the transmitter, where τ is the time needed for the feedback loop. We assume that $\{H_t\}$ is a stationary matrix random process and, at each time instant t, the elements of H_t are zero-mean, unit variance independent and identically distributed (i.i.d.) Gaussian random variables, yielding $\mathrm{vec}(H), \mathrm{vec}(\hat{H}) \sim \mathcal{CN}(\mathbf{0}, I_{n_t n_r})$. The channel coefficients are time varying according to Jakes' model, with common maximum Doppler frequency f_d. Thus, $\hat{H}$ and H can be modeled as jointly Gaussian with cross-correlation [7, p. 93]

$$\mathcal{E}[\mathrm{vec}(H)\mathrm{vec}^H(\hat{H})] = \rho_\tau I_{n_t n_r}$$

where ρ_τ is the normalized correlation coefficient specified by the Jakes' model, i.e., $\rho_\tau = J_0(2\pi f_d \tau)$, with $J_0(\cdot)$ the zeroth-order Bessel function of the first kind. In this case, it can be easily proved that

$$\Sigma = 2(1 - \rho_\tau) I_{n_t n_r}. \tag{13}$$

We continue with the noise SOS uncertainties. Since we assume that $R_n = \sigma_n^2 I_{n_r}$, we define the SOS estimation error as

$$\Delta R_n := \left(\hat{\sigma}_n^2 - \sigma_n^2 \right) I_{n_r}. \tag{14}$$

Using training data S_{tr}, it can be shown that an unbiased noise variance estimate is [9, p. 697]

$$\hat{\sigma}_n^2 = \frac{1}{n_r (N_{\mathrm{tr}} - n_t)} \mathrm{tr}\left(Y_{\mathrm{tr}} \Pi_{S_{\mathrm{tr}}^H}^{\perp} Y_{\mathrm{tr}}^H \right) \tag{15}$$

where $\Pi_{S_{\mathrm{tr}}^H}^{\perp}$ is the orthogonal projector onto the orthogonal complement of the column space of S_{tr}^H. For more details, the reader is referred to, for example [6, Sec. 9.4]. Using optimal training, it can be shown that the noise variance estimate (15) has variance

$$\mathcal{E}\left[\left(\hat{\sigma}_n^2 - \sigma_n^2 \right)^2 \right] = \frac{\sigma_n^4}{n_r (N_{\mathrm{tr}} - n_t)}. \tag{16}$$

D. *EMSE of the TxWF With Uncertainties*

In this subsection, we develop a second-order approximation to the EMSE induced by channel or noise SOS uncertainties. We denote with $\hat{\tilde{P}}$ and $\hat{\beta}$ the scaled TxWF and the Wiener scalar computed by using the channel or the noise SOS estimates as if they were the true quantities. The corresponding TxWF is $\hat{P} := \hat{\beta}\hat{\tilde{P}}$. The MSE associated with $\hat{P}$ and $\hat{\beta}$ is

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

Using a Taylor expansion of function MSE$(\cdot)$ around point $\tilde{P}_o$, we obtain

$$\mathrm{MSE}(\hat{\tilde{P}}) = \mathrm{MSE}(\tilde{P}_o) + \mathrm{tr}(\Delta\tilde{P}^H \mathrm{MSE}''(\tilde{P}_o)\Delta\tilde{P}) \tag{17}$$

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), we obtain that [8]

$$\mathrm{MSE}''(\tilde{P}_o) = H^H H + \alpha I_{n_t}. \tag{18}$$

We define the EMSE as

$$\begin{aligned} \mathrm{EMSE}(\hat{\tilde{P}}) &:= \mathcal{E}[\mathrm{MSE}(\hat{\tilde{P}}) - \mathrm{MSE}(\tilde{P}_o)] \\ &= \mathcal{E}[\mathrm{tr}(\Delta\tilde{P}^H \mathrm{MSE}''(\tilde{P}_o)\Delta\tilde{P})] \\ &= \mathcal{E}\left[\mathrm{tr}\left(\Delta\tilde{P}^H \left(H^H H + \alpha I_{n_t} \right) \Delta\tilde{P} \right) \right]. \end{aligned} \tag{19}$$

III. EMSE Due to Channel Uncertainties

In this section, we assume that the transmitter perfectly knows the noise SOS and has obtained a channel estimate $\hat{H}$, which is used for the computation of the TxWF. In order to compute the EMSE in (19), we must develop a first-order approximation to $\Delta\tilde{P}$ with respect to ΔH. This is our task in the sequel. If we use in (6) the estimate $\hat{H}$ as if it were the true channel H, then we compute the scaled pre-equalization matrix

$$\hat{\tilde{P}} = \left(\hat{H}^H \hat{H} + \alpha I_{n_t} \right)^{-1} \hat{H}^H \tag{20}$$

which can be written as

$$\hat{\tilde{P}} = \Big(H^H H + \alpha I_{n_t} + \underbrace{H^H \Delta H + \Delta H^H H}_{K_\Delta} + O(\|\Delta H\|^2) \Big)^{-1} \times (H^H + \Delta H^H). \tag{21}$$

Using the first-order approximation [5, p. 131]

$$(A + \Delta A)^{-1} = A^{-1} - A^{-1}\Delta A A^{-1} \tag{22}$$

and definition (6), we obtain

$$\hat{\tilde{P}} = \tilde{P}_o - \left(H^H H + \alpha I_{n_t} \right)^{-1} (K_\Delta \tilde{P}_o - \Delta H^H) + O(\|\Delta H\|^2).$$

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

$$\Delta\tilde{P} = -\underbrace{\left(H^H H + \alpha I_{n_t} \right)^{-1}}_{\mathcal{A}} \underbrace{(K_\Delta \tilde{P}_o - \Delta H^H)}_{\Delta} \tag{23}$$

and a second-order approximation of the EMSE is given by

$$\begin{aligned} \mathrm{EMSE}(\hat{\tilde{P}}) &= \mathcal{E}\left[\mathrm{tr}\left(\Delta\tilde{P}^H (H^H H + \alpha I_{n_t}) \Delta\tilde{P} \right) \right] \\ &\overset{(23)}{=} \mathcal{E}[\mathrm{tr}(\Delta^H \mathcal{A} \Delta)] \\ &= \mathcal{E}\left[\mathrm{tr}\left(\mathcal{A}\Delta I_{n_r} \Delta^H \right) \right] \\ &\overset{(a)}{=} \mathcal{E}\left[\mathrm{vec}^H(\Delta)\,(I_{n_r} \otimes \mathcal{A})\,\mathrm{vec}(\Delta) \right] \\ &= \mathrm{tr}\left((I_{n_r} \otimes \mathcal{A})\; \mathcal{E}[\mathrm{vec}(\Delta)\mathrm{vec}^H(\Delta)] \right) \end{aligned} \tag{24}$$

[1]We introduce the statistical model for the channel time variations just for *analysis purposes*. Robust precoders exploiting this knowledge (see, for example, [3]) are beyond the scope of this correspondence.

where at point (a) we used expression [4, p. 42]

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

From the definitions of Δ in (23) and K_Δ in (21), we obtain

$$\begin{aligned}\mathrm{vec}(\Delta) &= \mathrm{vec}(H^H \Delta H \tilde{P}_o) + \mathrm{vec}\left(\Delta H^H \left(H\tilde{P}_o - I_{n_r}\right)\right)\\ &= \underbrace{\left(\tilde{P}_o^T \otimes H^H\right)}_{\mathcal{T}_1}\mathrm{vec}(\Delta H)\\ &\quad + \underbrace{\left(\left(\tilde{P}_o^T H^T - I_{n_r}\right)\otimes I_{n_t}\right)}_{\mathcal{T}_2}\mathrm{vec}(\Delta H^H)\end{aligned} \quad (25)$$

where we made use of [4, p. 17]

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

Using the commutation matrix $K_{n_t n_r}$ [4, p. 9], we obtain

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

yielding

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

where, for notational simplicity, the commutation matrix is denoted as K. Using the circular symmetry of ΔH and (10), we obtain

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

Finally, we obtain the expression

$$\mathrm{EMSE}(\hat{\tilde{P}}) = \mathbf{T}_1 + \mathbf{T}_2 \quad (26)$$

where

$$\begin{aligned}\mathbf{T}_1 &:= \mathrm{tr}\left((I_{n_r}\otimes\mathcal{A})\,\mathcal{T}_1\Sigma\mathcal{T}_1^H\right)\\ &= \mathrm{tr}\left(\mathcal{T}_1^H (I_{n_r}\otimes\mathcal{A})\,\mathcal{T}_1\Sigma\right)\\ &\overset{(25)}{=} \mathrm{tr}\left(\left(\tilde{P}_o^*\tilde{P}_o^T \otimes H\mathcal{A}H^H\right)\Sigma\right)\end{aligned} \quad (27)$$

and

$$\begin{aligned}\mathbf{T}_2 &:= \mathrm{tr}\left((I_{n_r}\otimes\mathcal{A})\,\mathcal{T}_2 K\Sigma^* K^H\mathcal{T}_2^H\right)\\ &= \mathrm{tr}\left(K^H\mathcal{T}_2^H (I_{n_r}\otimes\mathcal{A})\,\mathcal{T}_2 K\Sigma^*\right)\\ &\overset{(25)}{=} \mathrm{tr}\left(K^H\left(\left(H^*\tilde{P}_o^* - I_{n_r}\right)\right.\right.\\ &\qquad \left.\left.\times\left(\tilde{P}_o^T H^T - I_{n_r}\right)\otimes\mathcal{A}\right)K\Sigma^*\right)\\ &= \mathrm{tr}\left(\left(\mathcal{A}\otimes\left(H^*\tilde{P}_o^* - I_{n_r}\right)\right.\right.\\ &\qquad \left.\left.\times\left(\tilde{P}_o^T H^T - I_{n_r}\right)\right)\Sigma^*\right)\end{aligned} \quad (28)$$

where we made use of [4, p. 16]

$$AB\otimes CD = (A\otimes C)(B\otimes D)$$

and [4, p. 117]

$$K(A\otimes B)K^H = (B\otimes A)$$

for matrices with compatible dimensions.

Until now, we have expressed the EMSE in terms of Σ. Expressions (26)–(28) are admittedly complicated and do not provide significant insight. In the sequel, we assume sufficiently high SNR, and we derive simple and informative approximations of the EMSE.

A. Channel Estimation Errors and High SNR

In this subsection, we assume that the channel uncertainties are due to estimation errors, implying that $\Sigma = (\sigma_n^2/N_{\mathrm{tr}})\,I_{n_t n_r}$. Using the SVD of H, it can be shown that, for $i = 1,\ldots,n_r$ (recall the definition of $\mathcal{A}$ in (23))

$$\lambda_i(H\mathcal{A}H^H) = \frac{\lambda_i\left(H^H H\right)}{\lambda_i(H^H H)+\alpha}. \quad (29)$$

For $\alpha \ll \lambda_{n_r}(H^H H)$, that is, α much smaller than the smallest nonzero eigenvalue of $H^H H$, implying sufficiently high SNR, we obtain

$$\mathrm{tr}(H\mathcal{A}H^H) = \sum_{i=1}^{n_r}\frac{\lambda_i(H^H H)}{\lambda_i(H^H H)+\alpha} \approx \mathrm{tr}\,(I_{n_r}). \quad (30)$$

Another high-SNR approximation that will prove useful in the sequel is (the proof is provided in the Appendix)

$$\mathrm{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\right) \approx \frac{1}{\alpha}\,\mathrm{MMSE}. \quad (31)$$

Starting with $\mathbf{T}_1$ in (27) and using expression

$$\mathrm{tr}(A\otimes B) = \mathrm{tr}(A)\,\mathrm{tr}(B)$$

we obtain

$$\begin{aligned}\mathbf{T}_1 &\overset{(12)}{=} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\otimes H\mathcal{A}H^H\right)\\ &\overset{(30)}{\approx} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\right)\mathrm{tr}\,(I_{n_r})\\ &= \frac{n_r\,\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\right)\\ &\overset{(31)}{\approx} \frac{n_r\sigma_n^2}{N_{\mathrm{tr}}}\frac{1}{\alpha}\mathrm{MMSE}.\end{aligned} \quad (32)$$

In order to compute an approximation of $\mathbf{T}_2$, we use an expression analogous to (29), for $i = 1,\ldots,n_r$

$$\lambda_i\left(\left(H^*\tilde{P}_o^* - I_{n_r}\right)\left(\tilde{P}_o^T H^T - I_{n_r}\right)\right) = \frac{\alpha^2}{(\lambda_i(H^H H)+\alpha)^2}. \quad (33)$$

For high SNR, the right-hand side of (33) goes to zero, yielding

$$\mathrm{tr}\left(\left(H^*\tilde{P}_o^* - I_{n_r}\right)\left(\tilde{P}_o^T H^T - I_{n_r}\right)\right) \approx 0. \quad (34)$$

Thus

$$\mathbf{T}_2 \overset{(12)}{=} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}(\mathcal{A})\mathrm{tr}\left(\left(H^*\tilde{P}_o^* - I_{n_r}\right)\left(\tilde{P}_o^T H^T - I_{n_r}\right)\right) \overset{(34)}{\approx} 0. \quad (35)$$

We conclude that, for sufficiently high SNR, term $\mathbf{T}_2$ is negligible compared with $\mathbf{T}_1$; this statement is in agreement with simulations in Section V. Combining expressions (26), (32) and (35), we obtain

$$\mathrm{EMSE}(\hat{\tilde{P}}) \approx \frac{n_r\sigma_n^2}{N_{\mathrm{tr}}}\frac{1}{\alpha}\mathrm{MMSE} \overset{(7)}{=} \frac{\mathrm{E}}{N_{\mathrm{tr}}}\mathrm{MMSE}.$$

Thus, for optimal training and sufficiently high SNR

$$\mathrm{EMSE}(\hat{\tilde{P}}) \approx \frac{\mathrm{E}}{N_{\mathrm{tr}}}\,\mathrm{MMSE}. \quad (36)$$

We observe that the EMSE is approximately proportional to the MMSE, with the proportionality factor being the ratio of the transmit power, E, to the length of the training block used for channel estimation, N_{tr}. Expression (36) can be used as a criterion for the choice of the length of the training block N_{tr} and/or the total transmit power E.

B. Time-Varying Channels and High SNR

In this subsection, we assume that the uncertainties due to time variations dominate those due to estimation errors, yielding $\Sigma = 2(1-\rho_\tau)I_{n_t n_r}$. The only term of the previous analysis that is affected is (27), which becomes

$$\begin{aligned}\mathbf{T}_1 &\overset{(13)}{=} 2(1-\rho_\tau)\,\mathrm{tr}\left(\tilde{P}_o^* \tilde{P}_o^T \otimes H\mathcal{A}H^H\right)\\ &\overset{(30)}{\approx} 2(1-\rho_\tau)\mathrm{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\right)\mathrm{tr}\left(I_{n_r}\right).\end{aligned} \tag{37}$$

giving that

$$\mathrm{EMSE}(\hat{\tilde{P}}) \approx 2n_r(1-\rho_\tau)\,\mathrm{tr}(\tilde{P}_o^*\tilde{P}_o^T). \tag{38}$$

It is easy to see that

$$\begin{aligned}\mathrm{tr}\left(\tilde{P}_o^*\tilde{P}_o^T\right) &= \mathrm{tr}\left(\tilde{P}_o\tilde{P}_o^H\right) = \mathrm{tr}\left(H\mathcal{A}^2H^H\right)\\ &= \sum_{i=1}^{n_r}\frac{\lambda_i(HH^H)}{(\lambda_i(HH^H+\alpha)^2}\end{aligned}$$

is an increasing function of SNR and tends to $\mathrm{tr}((HH^H)^{-1})$ for SNR tending to infinity. Thus, the EMSE increases for increasing SNR and asymptotically attains the value

$$\mathrm{EMSE}(\hat{\tilde{P}}) \approx 2n_r(1-\rho_\tau)\,\mathrm{tr}((HH^H)^{-1}). \tag{39}$$

Of course, the above approximations are accurate for slow time variations because fast time variations introduce large channel uncertainties rendering our asympotic analysis inaccurate.

IV. EMSE DUE TO NOISE SOS UNCERTAINTIES

In this section, we assume that the channel is perfectly known at the transmitter and the noise SOS estimate $\hat{R}_n$ is used as if it were the true R_n. Then, the scaled precoding matrix becomes

$$\hat{\tilde{P}} = \left(H^HH + \frac{\mathrm{tr}(R_n+\Delta R_n)}{\mathrm{E}}I_{n_t}\right)^{-1}H^H. \tag{40}$$

Using (22), a first-order approximation to $\Delta\tilde{P}$, with respect to ΔR_n, is given by

$$\begin{aligned}\Delta\tilde{P} &= -\frac{\mathrm{tr}(\Delta R_n)}{\mathrm{E}}\left(H^HH+\alpha I_{n_t}\right)^{-1}\tilde{P}_o\\ &\overset{(23)}{=} -\frac{\mathrm{tr}(\Delta R_n)}{\mathrm{E}}\mathcal{A}\tilde{P}_o.\end{aligned} \tag{41}$$

Substituting the above expression in (19), we obtain the second-order approximation

$$\begin{aligned}\mathrm{EMSE}(\hat{\tilde{P}}) &= \mathcal{E}\left[\mathrm{tr}\left(\Delta\tilde{P}^H(H^HH+\alpha I_{n_t})\Delta\tilde{P}\right)\right]\\ &= \frac{\mathcal{E}[\mathrm{tr}^2(\Delta R_n)]}{\mathrm{E}^2}\mathrm{tr}\left(\tilde{P}_o^H\mathcal{A}\tilde{P}_o\right).\end{aligned} \tag{42}$$

A. Optimal Training and High SNR

Considering optimal training and high SNR, recalling that we consider the spatially and temporally white Gaussian noise case and using (16), we get

$$\begin{aligned}\mathcal{E}[\mathrm{tr}^2(\Delta R_n)] &= \mathcal{E}\left[n_r^2\left(\hat{\sigma}_n^2-\sigma_n^2\right)^2\right]\\ &= n_r^2\mathcal{E}\left[\left(\hat{\sigma}_n^2-\sigma_n^2\right)^2\right].\end{aligned} \tag{43}$$

Using (16) and (43), (42) becomes

$$\begin{aligned}\mathrm{EMSE}(\hat{\tilde{P}}) &= \frac{n_r^2\mathcal{E}\left[\left(\hat{\sigma}_n^2-\sigma_n^2\right)^2\right]}{\mathrm{E}^2}\mathrm{tr}\left(\tilde{P}_o^H\mathcal{A}\tilde{P}_o\right)\\ &= \frac{n_r\,\sigma_n^4}{\mathrm{E}^2(N_{\mathrm{tr}}-n_t)}\mathrm{tr}\left(\tilde{P}_o^H\mathcal{A}\tilde{P}_o\right).\end{aligned} \tag{44}$$

Using the definitions of $\tilde{P}_o$ and $\mathcal{A}$ in (6) and (23), respectively, we write

$$\begin{aligned}\mathrm{tr}\left(\tilde{P}_o^H\mathcal{A}\tilde{P}_o\right) &= \mathrm{tr}\left(H\left(H^HH+\alpha I_{n_t}\right)^{-3}H^H\right)\\ &= \mathrm{tr}(H\mathcal{A}^3H^H).\end{aligned} \tag{45}$$

An expression analogous to (29), for $i = 1,\ldots,n_r$, is

$$\lambda_i(H\mathcal{A}^3H^H) = \frac{\lambda_i(H^HH)}{(\lambda_i(H^HH)+\alpha)^3}$$

which, for high SNR, gives

$$\lambda_i(H\mathcal{A}^3H^H) \approx \frac{1}{\lambda_i^2(H^HH)} = \frac{1}{\lambda_i^2(HH^H)}.$$

Thus

$$\mathrm{tr}\left(\tilde{P}_o^H\mathcal{A}\tilde{P}_o\right) \approx \sum_{i=1}^{n_r}\frac{1}{\lambda_i^2(HH^H)} = \|\left(HH^H\right)^{-1}\|_F^2 \tag{46}$$

and finally, combining expressions (44) and (46), we obtain

$$\mathrm{EMSE}(\hat{\tilde{P}}) \approx \frac{n_r\sigma_n^4}{\mathrm{E}^2(N_{\mathrm{tr}}-n_t)}\|\left(HH^H\right)^{-1}\|_F^2. \tag{47}$$

This approximation states that the EMSE is proportional to the squared noise variance, σ_n^4, which decreases very fast for increasing SNR. The proportionality factor is determined by the transmit power, E, the length of the training block, N_{tr}, the number of the transmit and receive antennas n_t and n_r, and the conditioning of the matrix channel H, through the Frobenius norm $\|(HH^H)^{-1}\|_F^2$. In the simulations section, we will see that this bound is a good approximation to the EMSE, especially at high SNR.

V. SIMULATIONS

We consider a system with $n_t = 3$ transmit antennas and $n_r = 2$ receive antennas. We consider the channel matrix H with elements given in Table I.[2] The noise is spatially and temporally white, circularly symmetric complex Gaussian with variance σ_n^2. We set the transmit

[2]Our results hold for *any* channel matrix. Since in our theoretical developments we did not average over the channels, but only over the channel uncertainties, in the simulations we use only one channel realization. We have made analogous observations in extensive simulation studies.

TABLE I
ELEMENTS OF CHANNEL MATRIX H

-0.2646 + 0.1212*j	-0.0456 - 0.2588*j	-0.0081 - 0.7268*j
0.0664 + 0.0179*j	-0.1597 + 0.4986*j	0.0656 - 0.1866*j

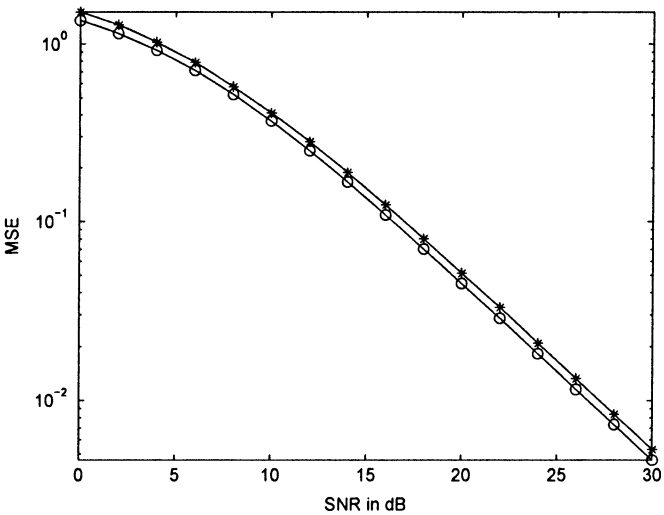

Fig. 2. MMSE using the true channel ("-o-") and expectation of the MSEs using the channel estimate ("-∗-").

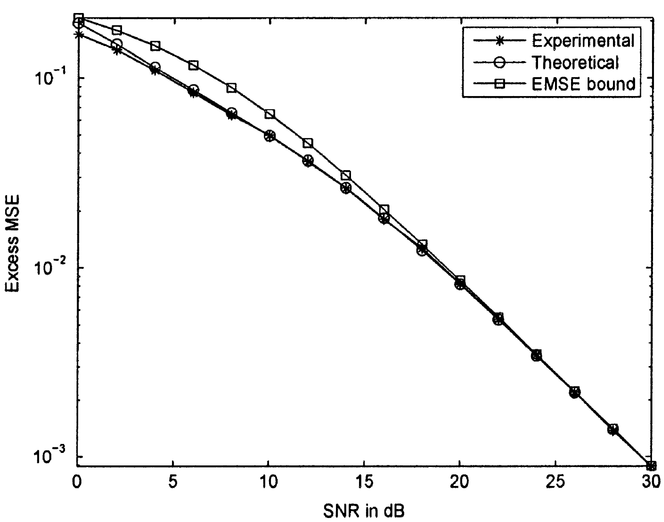

Fig. 3. Experimentally computed EMSE, theoretical second-order approximation (26), and high-SNR approximation (36) for the case of channel estimation errors.

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

Simulation 1. Channel Estimation Errors: In Fig. 2, we plot the MMSE (8) and the mean of the MSEs computed using the channel estimate (the average is over different realizations of the channel estimation error ΔH). We observe that the distance of these two quantities is approximately constant and does not depend on the SNR, verifying expression (36).

In Fig. 3, we present the experimentally computed EMSE, the theoretical second-order approximation (26) and approximation (36). We observe that the experimental and theoretical EMSE values practically coincide for SNR higher than 5 dB, while approximation (36) is very close to the EMSE, especially at high SNR.

In Fig. 4, we plot terms $\mathbf{T}_1$ and $\mathbf{T}_2$ of the theoretical EMSE of (26). We observe that, for SNR higher than 7 dB, the contribution of term

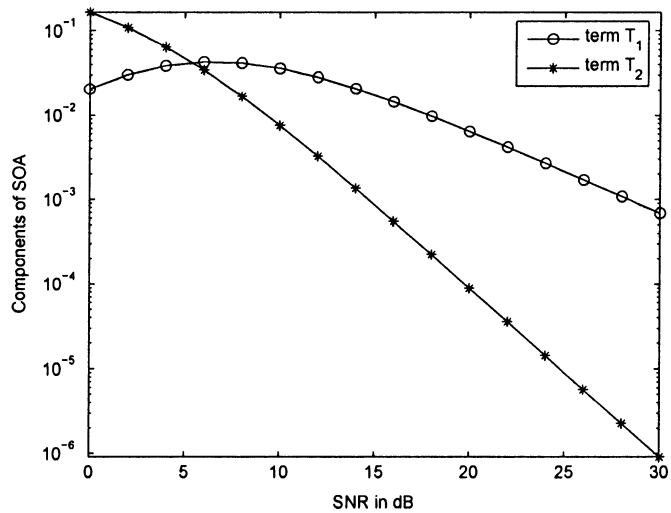

Fig. 4. Terms $\mathbf{T}_1$ and $\mathbf{T}_2$ of the EMSE second-order approximation (26) for the case of channel estimation errors.

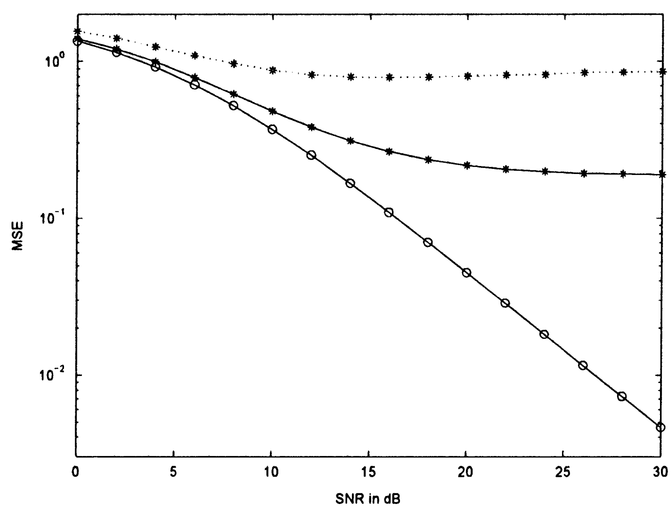

Fig. 5. MMSE using the true channel ("-o-") and expectation of the MSEs using the channel estimate for $\rho = 0.99$ ("-∗-") and for $\rho = 0.9$ (dotted line).

$\mathbf{T}_2$ to the EMSE is much smaller than the contribution of term $\mathbf{T}_1$, supporting our claim that the EMSE is approximately equal to term $\mathbf{T}_1$ for the high-SNR cases.

Simulation 2. Channel Time Variations: In Fig. 5, we plot the MMSE (8) and the mean of the MSEs computed using the outdated channel versions for channel correlation coefficients $\rho_\tau = 0.99, 0.9$ (the average is over different realizations of the channel uncertainties due to channel time variations ΔH). We observe that the distance of the two curves from the MMSE increases for increasing SNR, and the mean of the MSEs reaches a floor. This happens because the EMSE induced by the channel time variations increases for increasing SNR and asymptotically attains a limit value.

In Fig. 6, we present the theoretical second-order approximation (26), the corresponding experimentally computed EMSE, the high-SNR approximation (38) and the asymptotic value (39) for channel correlation coefficient equal to $\rho = 0.99$, implying very accurate channel information at the transmitter. We observe that the second-order approximation is very accurate, while (38) is a good approximation to the EMSE for SNR higher than 20 dB.

Simulation 3. Noise Estimation Errors: In Fig. 7, we present the theoretical second-order approximation (44), the corresponding experimentally computed EMSE and the high-SNR approximation (47). We

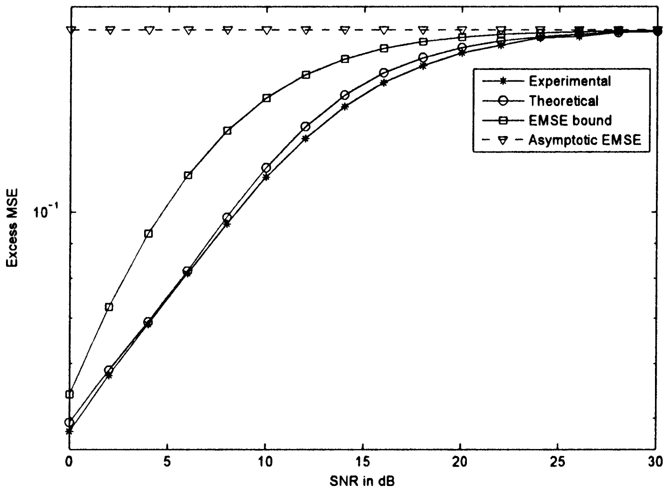

Fig. 6. Experimentally computed EMSE, theoretical second-order approximation (26), high-SNR approximation (38) and asymptotic EMSE value (39) for channel time variations ($\rho = 0.99$).

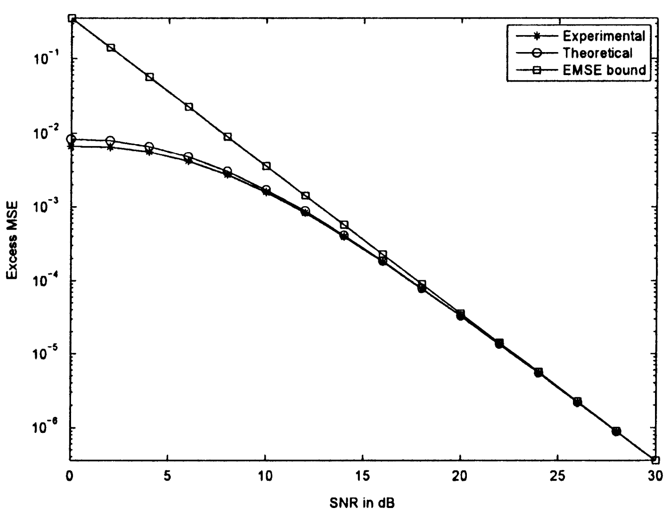

Fig. 7. Experimentally computed EMSE, theoretical second-order approximation (44), and high-SNR approximation (47) for the case of noise SOS estimation errors.

observe that the first two quantities practically coincide and approximation (47) is very close to the true EMSE for SNR higher than 15 dB.

Comparing the EMSEs for the cases of estimation errors only (see Figs. 3 and 7), we observe that the error induced by the channel estimation errors is much more significant than that induced by the noise SOS estimation errors.

VI. Conclusion

We considered the behavior of the TxWF under channel and noise SOS uncertainties by developing second-order EMSE approximations. We derived simple EMSE approximations in the high-SNR cases. Considering the channel estimation errors, we concluded that the EMSE is proportional to the MMSE, with the proportionality factor determined by the transmit power E and the length of the training block $N_{\rm tr}$. Considering the channel time variations, we found that the EMSE increases and, for high SNR, it reaches an asymptotic value. For the case of noise SOS estimation errors, we showed that the EMSE is proportional to the squared noise variance, σ_n^4. A comparison of the EMSEs for the cases of estimation errors only, shows that the error induced by the channel estimate is much more significant than that induced by the noise SOS estimate.

Appendix
A Useful Approximation

In order to simplify term $\mathrm{tr}(\tilde{P}_o^* \tilde{P}_o^T)$ in the high-SNR cases, i.e., $\alpha \ll \lambda_{n_r}(H^H H)$, we notice that

$$\mathrm{tr}\left(\tilde{P}_o^* \tilde{P}_o^T\right) = \mathrm{tr}\left(\tilde{P}_o \tilde{P}_o^H\right).$$

Using the definitions of MMSE and $\tilde{P}_o$ in (8) and (6), respectively, we get

$$\begin{aligned}\mathrm{MMSE} &= \mathrm{tr}(I_{n_r}) - 2\mathrm{tr}(\tilde{P}_o H) + \mathrm{tr}\left(\tilde{P}_o \tilde{P}_o^H H^H H\right) \\ &\quad + \alpha \mathrm{tr}\left(\tilde{P}_o \tilde{P}_o^H\right) \\ &= \mathrm{tr}(I_{n_r}) - 2\mathrm{tr}\left(\mathcal{A} H^H H\right) + \mathrm{tr}\left(\mathcal{A} H^H H \mathcal{A} H^H H\right) \\ &\quad + \alpha \mathrm{tr}\left(\tilde{P}_o \tilde{P}_o^H\right).\end{aligned}$$

Using approximations analogous to (30), we can write for the high-SNR cases

$$\begin{aligned}\mathrm{MMSE} &\approx \mathrm{tr}\left(I_{n_r}\right) - 2\mathrm{tr}\left(I_{n_r}\right) + \mathrm{tr}\left(I_{n_r}\right) + \alpha \mathrm{tr}\left(\tilde{P}_o \tilde{P}_o^H\right) \\ &= \alpha \mathrm{tr}\left(\tilde{P}_o \tilde{P}_o^H\right).\end{aligned}$$

Finally, we get

$$\mathrm{tr}\left(\tilde{P}_o \tilde{P}_o^H\right) \approx \frac{1}{\alpha}\mathrm{MMSE}. \tag{48}$$

References

[1] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.
[2] M. Joham, K. Kusume, M. H. Gzara, W. Utschick, and J. A. Nossek, "Transmit Wiener filter," Munich Univ. of Technology, Munich, Germany, Tech. Rep. TUM-LNS-TR-02-1, Jan. 2002.
[3] F. A. Dietrich, R. Hunger, M. Joham, and W. Utschick, "Robust transmit Wiener filter for time division duplex systems," in *Proc. 3rd IEEE Int. Symp. Signal Processing Information Technology*, Darmstadt, Dec. 2003, pp. 415–418.
[4] H. Luetkepohl, *Handbook of Matrices*. New York: Wiley, 1996.
[5] G. Stewart and J.-G. Sun, *Matrix Perturbation Theory*. New York: Academic, 1990.
[6] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[7] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[8] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Commun.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.
[9] A. Sayed, *Fundamentals of Adaptive Filtering*. New York: IEEE-Wiley, 2003.