On the Sensitivity of the MIMO Tomlinson–Harashima Precoder With Respect to Channel Uncertainties

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

***Abstract*—The multiple-input multiple-output Tomlinson–Harashima (MIMO-TH) precoder is a well-known structure that mitigates interstream interference in flat fading MIMO systems. The MIMO-TH filters are designed by assuming perfect channel state information (CSI) at both the transmitter and the receiver. However, in practice, channel estimates are available instead of the true channels. In this work, we assess the MIMO-TH performance degradation in the cases where the channel estimates are used as if they were the true channels. More specifically, we develop second-order and high-SNR approximations to the excess mean-square error (EMSE) induced by channel uncertainties, uncovering the factors that determine the MIMO-TH performance degradation in practice. Numerical experiments are in agreement with our theoretical developments.**

***Index Terms*—Channel estimation errors, channel time variations, MIMO systems, Tomlinson–Harashima precoding.**

I. Introduction

INTERSTREAM interference is a problem commonly encountered in multiple-input multiple-output (MIMO) communication systems. Many receiver structures mitigating interstream interference have been proposed in the literature, achieving various levels of performance with varying complexity. Prominent among them is the MIMO decision feedback equalizer (DFE). This nonlinear receiver works efficiently but may suffer from error propagation. This disadvantage can be overcome by moving the feedback loop of the DFE to the transmitter, resulting in the so-called Tomlinson–Harashima (TH) precoder. In this work, we consider the TH precoder proposed in Appendix E of [1]. The design of the TH precoder assumes perfect channel state information (CSI) at both the transmitter and the receiver; see, for example, [1]–[5]. However, since CSI uncertainties *always* exist in real-world systems, due to, e.g., channel estimation errors, this assumption is *not* realistic. One way to proceed is to use the channel estimate as if it were the true channel; this is sometimes called the *mismatched* or *naive* approach. Another way is to exploit the statistical description of the channel uncertainties and develop *robust* designs; see, for example, [6]–[8]. However, in all cases, the design of the MIMO-TH filters is based on *inexact* channel estimates and thus performance degradation is inevitable.

In this work, we consider a packet-based communication scenario where the channel may change (slowly) between successive packets. During each packet, the receiver estimates the channel and feeds its estimate back to the transmitter. This estimate is used for the design of the TH precoding filter that will be applied to the next packet. Thus, the TH precoding filter suffers from channel estimation errors (that occur at the receiver) and usually also suffers from mismatch due to channel time-variations, because the next packet may pass through a (slightly) different channel. Upon arrival of the packet, the receiver estimates the current channel (which is fed back to the transmitter) and proceeds to equalization and detection. Thus, the processing of each packet suffers from errors at both the transmitter and the receiver. Obviously, these errors degrade the MIMO-TH performance. We quantify this degradation by assessing the associated excess mean-square error (EMSE). We show that the EMSE consists of two components that can be studied separately. The first component is due to the mismatch between the previous channel estimate and the current channel, while the second is due to the mismatch between the current channel and its estimate. We develop a second-order approximation to the EMSE which, in our experiments, is very accurate for SNR higher than 5 dB. However, this approximation is quite complicated and thus difficult to interpret. We focus on the high-SNR regime and derive a simple, informative, and tight (for sufficiently high SNR) EMSE upper bound, which uncovers the basic factors that determine the MIMO-TH performance degradation.

A. Notation and Matrix Results

Superscripts ^{T}, ^{H}, and * denote transpose, conjugate transpose and elementwise conjugation, respectively. $\mathrm{tr}(\cdot)$, $\mathrm{vec}(\cdot)$, and $\mathrm{vech}(\cdot)$ denote the trace, the vectorization and the half-vectorization operator, respectively. $\otimes$ denotes the Kronecker product and $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number. $\mathbf{I}_M$ and $\mathbf{O}_M$ denote the $M \times M$ identity and zero matrix, respectively. Finally, a_{ij} denotes the (i,j)th element of matrix $\mathbf{A}$.

We remind that for matrices with compatible dimensions [9, pp. 17-19]

$$\mathrm{tr}(\mathbf{ABCD}) = \mathrm{vec}^T(\mathbf{D}^T)(\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}) \quad (1)$$

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}) \quad (2)$$

$$\mathbf{AB} \otimes \mathbf{CD} = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) \quad (3)$$

$$\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B}) \quad (4)$$

Manuscript received December 16, 2008; accepted September 02, 2009. First published October 13, 2009; current version published March 10, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mounir Ghogho.

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).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2034302

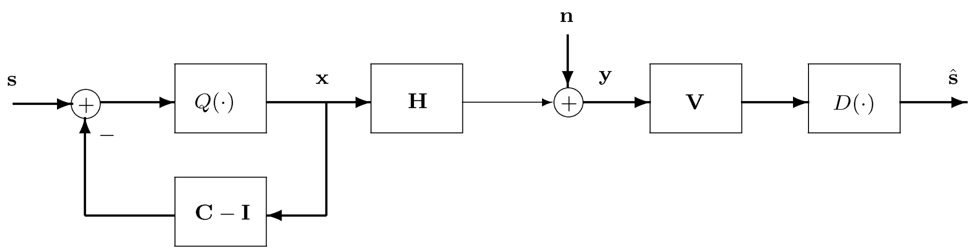

Fig. 1. System model.

and [9, p. 117]

$$\mathbf{K}(\mathbf{A}\otimes\mathbf{B})\mathbf{K}^{\mathbf{H}} = \mathbf{B}\otimes\mathbf{A} \tag{5}$$

where $\mathbf{K}$ denotes the commutation matrix. If $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite, then [9, p. 44]

$$\mathrm{tr}(\mathbf{AB}) \leq \mathrm{tr}(\mathbf{A})\mathrm{tr}(\mathbf{B}). \tag{6}$$

For any matrix $\mathbf{A}$ [9, p. 97]

$$\mathrm{vec}(\mathbf{A}) = \mathbf{K}\mathrm{vec}(\mathbf{A}^T). \tag{7}$$

If $\mathbf{A}$ is lower triangular, then [9, p. 99]

$$\mathrm{vec}(\mathbf{A}) = \mathbf{L}^T\mathrm{vech}(\mathbf{A}) \tag{8}$$

and

$$\mathrm{vec}\left(\mathrm{diag}\left(\mathrm{diag}(\mathbf{A})\right)\right) = \mathbf{L}^T\mathbf{LKL}^T\mathbf{L}\mathrm{vec}(\mathbf{A}) \tag{9}$$

where $\mathbf{L}$ is the elimination matrix [9, ch. 9] and $\mathrm{diag}(\mathrm{diag}(\mathbf{A}))$ is the diagonal matrix whose elements are the diagonal elements of $\mathbf{A}$. Finally, we remind that [14, p. 130]

$$(\mathbf{A}+\Delta\mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\Delta\mathbf{A}\mathbf{A}^{-1} + O\left(\|\Delta\mathbf{A}\|^2\right). \tag{10}$$

During our study, we shall develop first- and second-order approximations, with respect to channel uncertainties, as well as high-SNR approximations. In order to distinguish among these cases, we shall use the symbols $\simeq$, $\approx$, and $\approxeq$, respectively.

The rest of the paper is structured as follows. In Section II, we present the MIMO-TH structure assuming that the receiver and the transmitter have perfect CSI, we describe the channel uncertainties, compute the filters that result from the naive approach and define the EMSE. In Section III, we develop a second-order approximation to the EMSE, while in Section IV we derive a simple high-SNR EMSE upper bound. In Section V, we support our theoretical developments with numerical experiments. Some conclusions appear in Section VI.

II. The MIMO-TH Precoder

A. The System Model

We consider the baseband-equivalent discrete-time frequency-flat MIMO system depicted in Fig. 1, with n_t transmit and n_r receive antennas (with $n_r \geq n_t$). The input–output relation of the channel is

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n} \tag{11}$$

where $\mathbf{x}$ is the $n_t \times 1$ channel input vector, $\mathbf{H}$ is the $n_r \times n_t$ channel matrix, and $\mathbf{n}$ is the $n_r \times 1$ additive channel noise. The channel input symbols x_k, $k = 1,\ldots,n_t$, are successively generated from the data symbols s_k, $k = 1,\ldots,n_t$, as shown in Fig. 1, where the feedback loop consists of the feedback matrix $\mathbf{C}$ and the modulo operator $Q_M(\cdot)$. If $\mathbf{s}$ is a vector with independent identically distributed (i.i.d.) elements s_k (drawn from an M-QAM constellation), then it can be shown that $\mathbf{x}$ consists of uncorrelated random variables, uniformly distributed in $(-M, M]$ and has covariance matrix $\mathbf{R_x} = \sigma_x^2\mathbf{I}_{n_t}$, where $\sigma_x^2 = 2M^2/12$ [1, p. 462]. The noise vector $\mathbf{n}$ is assumed to be complex-valued circular Gaussian with covariance matrix $\mathbf{R_n} = \sigma_n^2\mathbf{I}_{n_r}$.

B. Optimal MMSE MIMO-TH

In this subsection, we briefly present the computation of the MMSE MIMO-TH filters, following the approach of the Appendix E of [1]. The error signal before the receiver's modulo operator (see Fig. 1) is

$$\mathbf{e} = \mathbf{Vy} - \mathbf{Cx} \tag{12}$$

and the mean-squared error is defined as

$$\mathrm{mse}(\mathbf{C},\mathbf{V}) := \mathcal{E}\left[\|\mathbf{e}\|_2^2\right]. \tag{13}$$

The function $\mathrm{mse}(\cdot)$ can be expressed as

$$\mathrm{mse}(\mathbf{C},\mathbf{V}) = \mathrm{tr}\left(\mathbf{V}\left(\sigma_x^2\mathbf{HH}^H + \sigma_n^2\mathbf{I}_{n_r}\right)\mathbf{V}^H\right) - 2\sigma_x^2\mathrm{Re}\left\{\mathrm{tr}(\mathbf{VHC}^H)\right\} + \sigma_x^2\mathrm{tr}(\mathbf{CC}^H). \tag{14}$$

For any $\mathbf{C}$, minimization of $\mathrm{mse}(\mathbf{C},\mathbf{V})$ with respect to $\mathbf{V}$, yields

$$\mathbf{V} = \mathbf{CH}^H\left(\mathbf{HH}^H + \zeta\mathbf{I}_{n_r}\right)^{-1}$$

where $\zeta := \sigma_n^2/\sigma_x^2$. By substituting this value to $\mathrm{mse}(\mathbf{C},\mathbf{V})$, we obtain

$$\mathrm{MSE}(\mathbf{C}) := \sigma_n^2\mathrm{tr}\left(\mathbf{C}\left(\mathbf{H}^H\mathbf{H} + \zeta\mathbf{I}_{n_t}\right)^{-1}\mathbf{C}^H\right). \tag{15}$$

Minimization of $\mathrm{MSE}(\mathbf{C})$ with respect to $\mathbf{C}$, subject to the constraint that $\mathbf{C}$ be a lower triangular matrix with diagonal elements equal to 1, gives [1], [10]

$$\mathbf{C}_o = \mathbf{GR} \tag{16}$$

where $\mathbf{R}$ is the lower triangular matrix satisfying the modified Cholesky factorization

$$\mathbf{H}^H\mathbf{H} + \zeta\mathbf{I}_{n_t} = \mathbf{R}^H\mathbf{R} \tag{17}$$

and

$$\mathbf{G} = \text{diag}\left(r_{11}^{-1}, \ldots, r_{n_tn_t}^{-1}\right). \tag{18}$$

Using $\mathbf{C}_o$, we compute the optimal $\mathbf{V}$ as

$$\mathbf{V}_o = \mathbf{C}_o\mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H + \zeta\mathbf{I}_{n_r}\right)^{-1}. \tag{19}$$

Substituting $\mathbf{C}_o$ and $\mathbf{V}_o$ in (14), we obtain

$$\begin{aligned}\text{MMSE} &:= \text{mse}(\mathbf{C}_o, \mathbf{V}_o) = \text{MSE}(\mathbf{C}_o)\\ &= \sigma_n^2 \text{tr}\left(\mathbf{C}_o\left(\mathbf{H}^H\mathbf{H} + \zeta\mathbf{I}_{n_t}\right)^{-1}\mathbf{C}_o^H\right).\end{aligned} \tag{20}$$

Using (16), we derive the alternative expression for the MMSE

$$\begin{aligned}\text{MMSE} &= \sigma_n^2 \text{tr}(\mathbf{G}\mathbf{G}^H)\\ &= \sigma_n^2 \text{tr}\left(\text{diag}\left(|r_{11}|^{-2}, \ldots, |r_{n_tn_t}|^{-2}\right)\right).\end{aligned} \tag{21}$$

C. Channel Uncertainties

After the description of the ideal case, where we assumed that the channel $\mathbf{H}$ is perfectly known at the receiver and the transmitter, we proceed to a realistic scenario where both the transmitter and the receiver possess channel estimates. More specifically, we consider a frequency division duplex system and focus on the transmission of packet i. During the transmission of packet $(i-1)$, the receiver estimates the true channel, $\mathbf{H}_{i-1}$, as $\bar{\mathbf{H}}_{i-1}$. This estimate is communicated to the transmitter through a feedback channel and is used for precoding packet i. The true channel during the transmission of packet i, $\mathbf{H}_i$, is estimated at the receiver as $\bar{\mathbf{H}}_i$. Thus, in general, the channel estimate used at the transmitter for precoding packet i, $\bar{\mathbf{H}}_{i-1}$, suffers from *both* estimation errors and errors due to channel time-variations (other potential error sources are quantization errors and feedback channel errors—the following analysis can easily incorporate quantization errors, while the same does not happen for the feedback channel errors). On the other hand, the channel estimate at the receiver for packet i, $\bar{\mathbf{H}}_i$, suffers *only* from estimation errors.

In order to assess the associated performance degradation, we adopt the following statistical models for the channel inaccuracies.

1) *Channel estimation errors*: During each packet, we use training and estimate the channel using the maximum-likelihood (ML) method, i.e., we assume that the channel is *constant* but *unknown*. The $n_t \times N_{\text{tr}}$ training block for packet i, $\mathbf{S}_i$, is multiplexed with the precoded information vectors (for example, it may be at the start of the packet) but is *not* precoded (we note that $\mathbf{S}_i$ may be the same for all i). If $\mathbf{Y}_i$ denotes the channel output corresponding to $\mathbf{S}_i$, then the ML estimate of $\mathbf{H}_i$ is [11, p. 174]

$$\bar{\mathbf{H}}_i = \mathbf{Y}_i\mathbf{S}_i^H\left(\mathbf{S}_i\mathbf{S}_i^H\right)^{-1}. \tag{22}$$

The channel estimation error is defined as

$$\Delta\mathbf{H}_{\text{est},i} := \bar{\mathbf{H}}_i - \mathbf{H}_i. \tag{23}$$

Optimal channel estimates are obtained for semi-unitary training matrices, i.e., $\mathbf{S}_i\mathbf{S}_i^H = \sigma_x^2 N_{\text{tr}}\mathbf{I}_{n_t}$, and the optimal channel estimation error covariance matrix is [11, p. 175]

$$\boldsymbol{\Sigma}_{\text{est}} := \mathcal{E}\left[\text{vec}(\Delta\mathbf{H}_{\text{est},i})\text{vec}^H(\Delta\mathbf{H}_{\text{est},i})\right] = \frac{\sigma_n^2}{\sigma_x^2 N_{\text{tr}}}\mathbf{I}_{n_tn_r}. \tag{24}$$

We note that channel estimation errors associated with different packets are independent due to the assumed noise independence.

2) *Channel time-variations*: We adopt a commonly used statistical model describing the time evolution of the channel (the model is used only for analysis purposes and is not exploited during channel estimation). We denote with τ the time difference between two successive packets. We assume that $\{\mathbf{H}_i\}$ is a stationary matrix random process where, for all i, the elements of $\mathbf{H}_i$ are unit variance i.i.d. circular Gaussian random variables, i.e.,

$$\text{vec}(\mathbf{H}_{i-1}), \text{vec}(\mathbf{H}_i) \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{I}_{n_tn_r}\right).$$

We assume that the channel coefficients are time-varying according to Jakes' model, with common maximum Doppler frequency f_d. Thus, $\mathbf{H}_{i-1}$ and $\mathbf{H}_i$ can be modeled as jointly Gaussian with cross-correlation [12, p. 93]

$$\mathcal{E}\left[\text{vec}(\mathbf{H}_i)\text{vec}^H(\mathbf{H}_{i-1})\right] = \rho_\tau\mathbf{I}_{n_tn_r} \tag{25}$$

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. If we define the channel error due to time-variations as

$$\Delta\mathbf{H}_{\text{tv},i} := \mathbf{H}_i - \mathbf{H}_{i-1} \tag{26}$$

then the associated error covariance matrix is independent of i and is given by

$$\boldsymbol{\Sigma}_{\text{tv}} := \mathcal{E}\left[\text{vec}(\Delta\mathbf{H}_{\text{tv},i})\text{vec}^H(\Delta\mathbf{H}_{\text{tv},i})\right] = 2(1-\rho_\tau)\mathbf{I}_{n_tn_r}. \tag{27}$$

Finally, we note that it is natural to assume that the errors due to channel time-variations are independent of the channel estimation errors because they are originating from independent phenomena, i.e., the first from the random channel evolution in time and the second from the additive channel noise.

In the sequel, for notational convenience, we neglect index i. We denote with $\mathbf{H}$ the true channel, with $\hat{\mathbf{H}}$ the channel estimate at the transmitter, and with $\tilde{\mathbf{H}}$ the channel estimate at the receiver. We define the mismatch at the transmitter and the receiver as

$$\Delta\mathbf{H}_{\text{Tx}} := \hat{\mathbf{H}} - \mathbf{H}, \quad \Delta\mathbf{H}_{\text{Rx}} := \tilde{\mathbf{H}} - \mathbf{H}. \tag{28}$$

It can be easily shown that $\Delta\mathbf{H}_{\mathrm{Tx}}$ and $\Delta\mathbf{H}_{\mathrm{Rx}}$ are zero mean with covariance matrices

$$\mathcal{E}\left[\mathrm{vec}(\Delta\mathbf{H}_{\mathrm{Tx}})\mathrm{vec}^H(\Delta\mathbf{H}_{\mathrm{Tx}})\right] = \mathbf{\Sigma}_{\mathrm{est}} + \mathbf{\Sigma}_{\mathrm{tv}} \tag{29}$$

and

$$\mathcal{E}\left[\mathrm{vec}(\Delta\mathbf{H}_{\mathrm{Rx}})\mathrm{vec}^H(\Delta\mathbf{H}_{\mathrm{Rx}})\right] = \mathbf{\Sigma}_{\mathrm{est}} \tag{30}$$

respectively. Furthermore, $\Delta\mathbf{H}_{\mathrm{Rx}}$ and $\Delta\mathbf{H}_{\mathrm{Tx}}$ are independent.

We close this subsection by mentioning that $\zeta = \sigma_n^2/\sigma_x^2$ is also required for the computation of the filters at both the transmitter and the receiver. We assume that $\sigma_x^2 = 2M^2/12$ is known at both sides and σ_n^2 is estimated at the receiver (for more details we refer to [11, Sec. 9.4]); then, the estimate is sent to the transmitter through a feedback channel. It turns out that the variance of the noise variance estimation error is $O(\sigma_n^4)$ and thus, for sufficiently high SNR, the error in ζ is negligible compared with the channel estimation error. Thus, we assume that ζ is perfectly known.

D. MIMO-TH: The Mismatched Approach

In this subsection, we follow the mismatched approach and compute the MIMO-TH filters using the channel estimates $\hat{\mathbf{H}}$ and $\tilde{\mathbf{H}}$ as if they were the true channel $\mathbf{H}$. The transmitter, based on (16), computes and uses

$$\hat{\mathbf{C}} = \hat{\mathbf{G}}\hat{\mathbf{R}} \tag{31}$$

where $\hat{\mathbf{R}}^H\hat{\mathbf{R}} = \hat{\mathbf{H}}^H\hat{\mathbf{H}} + \zeta\mathbf{I}_{n_t}$ and $\hat{\mathbf{G}} = \mathrm{diag}(\hat{r}_{11}^{-1}, \ldots, \hat{r}_{n_t n_t}^{-1})$. We note that since the receiver knows $\hat{\mathbf{H}}$, it can compute and use $\hat{\mathbf{C}}$.

Given that the transmitter uses $\hat{\mathbf{C}}$, the input estimation error becomes

$$\hat{\mathbf{e}} = \mathbf{V}\hat{\mathbf{y}} - \hat{\mathbf{C}}\hat{\mathbf{x}} \tag{32}$$

where $\hat{\mathbf{x}}$ is the channel input produced by the feedback filter $\hat{\mathbf{C}}$ and $\hat{\mathbf{y}} = \mathbf{H}\hat{\mathbf{x}} + \mathbf{n}$. In order to compute the "optimal" filter at the receiver, we follow steps analogous to those of Section II-B. Then, it can be shown that the filter that minimizes $\mathcal{E}[\|\hat{\mathbf{e}}\|_2^2]$ is

$$\hat{\mathbf{V}} = \hat{\mathbf{C}}\mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H + \zeta\mathbf{I}_{n_r}\right)^{-1}. \tag{33}$$

The best the receiver can do is to use its current channel estimate $\tilde{\mathbf{H}}$ as if it were $\mathbf{H}$ and compute[1]

$$\hat{\tilde{\mathbf{V}}} = \hat{\mathbf{C}}\tilde{\mathbf{H}}^H\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \zeta\mathbf{I}_{n_r}\right)^{-1}. \tag{34}$$

Using (31) and (34) in (14), we obtain that the MSE achieved by the mismatched approach is

$$\begin{aligned}\mathrm{mse}(\hat{\mathbf{C}}, \hat{\tilde{\mathbf{V}}}) = \sigma_x^2\mathrm{tr}\left(\hat{\tilde{\mathbf{V}}}\left(\mathbf{H}\mathbf{H}^H + \zeta\mathbf{I}_{n_r}\right)\hat{\tilde{\mathbf{V}}}^H\right) \\ -2\sigma_x^2\mathrm{Re}\left\{\mathrm{tr}\left(\hat{\tilde{\mathbf{V}}}\mathbf{H}\hat{\mathbf{C}}^H\right)\right\} + \sigma_x^2\mathrm{tr}(\hat{\mathbf{C}}\hat{\mathbf{C}}^H).\end{aligned} \tag{35}$$

[1]It can be proven that if the receiver uses $\hat{\mathbf{H}}$ instead of $\tilde{\mathbf{H}}$, then the performance degrades dramatically. The proof can be made available by the authors upon request.

The EMSE is defined as

$$\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\tilde{\mathbf{V}}}) := \mathcal{E}\left[\mathrm{mse}(\hat{\mathbf{C}}, \hat{\tilde{\mathbf{V}}}) - \mathrm{mse}(\mathbf{C}_o, \mathbf{V}_o)\right] \tag{36}$$

where the expectation is with respect to the channel uncertainties. Our main task in the sequel is to quantify $\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\tilde{\mathbf{V}}})$.

III. EMSE—Second-Order Analysis

In this section, we develop a second-order approximation to $\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\tilde{\mathbf{V}}})$, with respect to channel uncertainties.

We start by considering two unrealistic and, thus, seemingly, useless cases. Their usefulness will become evident shortly.

1) *Channel uncertainties* only *at the transmitter*: We assume that the transmitter possesses the channel estimate $\hat{\mathbf{H}}$ while the receiver has perfect CSI. Thus, the transmitter and the receiver use filters $\hat{\mathbf{C}}$ and $\hat{\mathbf{V}}$, defined in (31) and (33), respectively. Substituting these values into (14), we obtain that the associated MSE is

$$\begin{aligned}\mathrm{MSE}(\hat{\mathbf{C}}) &= \mathrm{mse}(\hat{\mathbf{C}}, \hat{\mathbf{V}}) \\ &= \sigma_n^2\mathrm{tr}\left(\hat{\mathbf{C}}\left(\mathbf{H}^H\mathbf{H} + \zeta\mathbf{I}_{n_t}\right)^{-1}\hat{\mathbf{C}}^H\right).\end{aligned} \tag{37}$$

The corresponding EMSE is defined as

$$\mathrm{EMSE}(\hat{\mathbf{C}}) := \mathcal{E}\left[\mathrm{MSE}(\hat{\mathbf{C}}) - \mathrm{MSE}(\mathbf{C}_o)\right]. \tag{38}$$

2) *Channel uncertainties* only *at the receiver*: We assume that the transmitter has perfect CSI and the receiver possesses the channel estimate $\tilde{\mathbf{H}}$. Thus, the transmitter uses $\mathbf{C}_o$ defined in (16), while the receive filter, denoted as $\tilde{\mathbf{V}}$, is computed using the optimal transmit filter $\mathbf{C}_o$ and the channel estimate $\tilde{\mathbf{H}}$, as

$$\tilde{\mathbf{V}} = \mathbf{C}_o\tilde{\mathbf{H}}^H\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \zeta\mathbf{I}_{n_r}\right)^{-1}. \tag{39}$$

Substituting (16) and (39) into (14), we obtain that the associated MSE is

$$\begin{aligned}\mathrm{MSE}_o(\tilde{\mathbf{V}}) :=& \mathrm{mse}(\mathbf{C}_o, \tilde{\mathbf{V}}) \\ =& \sigma_x^2\mathrm{tr}\left(\tilde{\mathbf{V}}\left(\mathbf{H}\mathbf{H}^H + \zeta\mathbf{I}_{n_r}\right)\tilde{\mathbf{V}}^H\right) \\ &- 2\sigma_x^2\mathrm{Re}\left\{\mathrm{tr}\left(\tilde{\mathbf{V}}\mathbf{H}\mathbf{C}_o^H\right)\right\} \\ &+ \sigma_x^2\mathrm{tr}\left(\mathbf{C}_o\mathbf{C}_o^H\right).\end{aligned} \tag{40}$$

The corresponding EMSE is defined as

$$\mathrm{EMSE}(\tilde{\mathbf{V}}) := \mathcal{E}\left[\mathrm{MSE}_o(\tilde{\mathbf{V}}) - \mathrm{MSE}_o(\mathbf{V}_o)\right]. \tag{41}$$

The next result shows that $\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\tilde{\mathbf{V}}})$ can be decomposed into two terms that correspond to these unrealistic cases.

Proposition 1: The EMSE induced by channel inaccuracies at both the transmitter and the receiver can be approximated as

$$\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\tilde{\mathbf{V}}}) \approx \mathrm{EMSE}(\hat{\mathbf{C}}) + \mathrm{EMSE}(\tilde{\mathbf{V}}). \tag{42}$$

Proof: The proof is provided in Appendix I and is based on the fact that the channel errors $\Delta\mathbf{H}_{\rm Tx}$ and $\Delta\mathbf{H}_{\rm Rx}$ are zero-mean and independent. □

In the sequel, we develop second-order approximations to $\mathrm{EMSE}(\hat{\mathbf{C}})$ and $\mathrm{EMSE}(\tilde{\mathbf{V}})$.

A. Channel Uncertainties Only at the Transmitter

Using a Taylor expansion of the function $\mathrm{MSE}(\mathbf{C})$ in (37) around $\mathbf{C}_o$, we obtain

$$\mathrm{MSE}(\hat{\mathbf{C}}) = \mathrm{MSE}(\mathbf{C}_o) + \mathrm{tr}\left(\Delta\mathbf{C}\mathrm{MSE}''(\mathbf{C}_o)\Delta\mathbf{C}^H\right) \quad (43)$$

where $\Delta\mathbf{C} := \hat{\mathbf{C}} - \mathbf{C}_o$ and $\mathrm{MSE}''(\mathbf{C}_o)$ is the second derivative of $\mathrm{MSE}(\mathbf{C})$ evaluated at $\mathbf{C}_o$.[2] It can be shown that [13]

$$\mathrm{MSE}''(\mathbf{C}_o) = \sigma_n^2\left(\mathbf{H}^H\mathbf{H} + \zeta\mathbf{I}_{n_t}\right)^{-1}. \quad (44)$$

Using (38), (43), and (44), we obtain

$$\begin{aligned}\mathrm{EMSE}(\hat{\mathbf{C}}) &= \mathcal{E}\left[\mathrm{tr}\left(\Delta\mathbf{C}\mathrm{MSE}''(\mathbf{C}_o)\Delta\mathbf{C}^H\right)\right]\\ &= \sigma_n^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{C}\boldsymbol{\mathcal{A}}^{-1}\Delta\mathbf{C}^H)\right] \quad (45)\end{aligned}$$

where

$$\boldsymbol{\mathcal{A}} := \mathbf{H}^H\mathbf{H} + \zeta\mathbf{I}_{n_t}. \quad (46)$$

The following lemma gives a second-order approximation to $\mathrm{EMSE}(\hat{\mathbf{C}})$.

Lemma 1: A second-order approximation to $\mathrm{EMSE}(\hat{\mathbf{C}})$ is given by

$$\mathrm{EMSE}(\hat{\mathbf{C}}) \approx \sum_{i=1}^{3}\mathbf{B}_i \quad (47)$$

where terms $\mathbf{B}_i$ are defined in (48)–(50)

$$\begin{aligned}\mathbf{B}_1 :=& \left(\alpha + \frac{\sigma_n^2}{\sigma_x^2 N_{\rm tr}}\right)\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\mathbf{L}^T\mathbf{L}\right.\\ &\left.\times(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{R}^{-H}\mathbf{H}^H\mathbf{H}\mathbf{R}^{-1})\mathbf{L}^T\mathbf{L}\right) \quad (48)\\ \mathbf{B}_2 :=& -2\left(\alpha + \frac{\sigma_n^2}{\sigma_x^2 N_{\rm tr}}\right)\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\right.\\ &\left.\times\,\mathbf{P}(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{R}^{-H}\mathbf{H}^H\mathbf{H}\mathbf{R}^{-1})\mathbf{P}\right) \quad (49)\\ \mathbf{B}_3 :=& \left(\alpha + \frac{\sigma_n^2}{\sigma_x^2 N_{\rm tr}}\right)\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\,\mathbf{L}^T\mathbf{L}\right.\\ &\left.\times(\mathbf{R}^{-T}\mathbf{H}^T\mathbf{H}^*\mathbf{R}^{-*}\otimes\mathbf{R}^{-H}\mathbf{R}^{-1})\mathbf{L}^T\mathbf{L}\right). \quad (50)\end{aligned}$$

In these expressions, $\mathbf{L}$ is the elimination matrix and $\mathbf{P} := \mathbf{L}^T\mathbf{L}\mathbf{K}\mathbf{L}^T\mathbf{L}$, where $\mathbf{K}$ is the commutation matrix. The scalar α is defined as $\alpha := 2(1-\rho_\tau)$ (see (27)).

Proof: The proof is provided in Appendix II. □

[2]The first derivative of $\mathrm{MSE}(\mathbf{C})$ at $\mathbf{C}_o$ vanishes because $\mathbf{C}_o$ is the minimizer of $\mathrm{MSE}(\mathbf{C})$.

B. Channel Uncertainties Only at the Receiver

Using a Taylor expansion of the function $\mathrm{MSE}_o(\mathbf{V})$ in (40) around $\mathbf{V}_o$, we obtain

$$\mathrm{MSE}_o(\tilde{\mathbf{V}}) = \mathrm{MSE}_o(\mathbf{V}_o) + \mathrm{tr}\left(\Delta\mathbf{V}\mathrm{MSE}_o''(\mathbf{V}_o)\Delta\mathbf{V}^H\right) \quad (51)$$

where $\Delta\mathbf{V} := \tilde{\mathbf{V}} - \mathbf{V}_o$, and $\mathrm{MSE}_o''(\mathbf{V}_o)$ is the second derivative of $\mathrm{MSE}_o(\mathbf{V})$ evaluated at $\mathbf{V}_o$.[3] It can be shown that [13]

$$\mathrm{MSE}_o''(\mathbf{V}_o) = \sigma_x^2\left(\mathbf{H}\mathbf{H}^H + \zeta\mathbf{I}_{n_r}\right). \quad (52)$$

Using (41), (40), and (52), we obtain

$$\begin{aligned}\mathrm{EMSE}(\tilde{\mathbf{V}}) &= \mathcal{E}\left[\mathrm{tr}\left(\Delta\mathbf{V}\mathrm{MSE}_o''(\mathbf{V}_o)\Delta\mathbf{V}^H\right)\right]\\ &= \sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{V}\boldsymbol{\mathcal{B}}\Delta\mathbf{V}^H)\right] \quad (53)\end{aligned}$$

where

$$\boldsymbol{\mathcal{B}} := \mathbf{H}\mathbf{H}^H + \zeta\mathbf{I}_{n_r}. \quad (54)$$

The following lemma gives a second-order approximation to $\mathrm{EMSE}(\tilde{\mathbf{V}})$.

Lemma 2: A second-order approximation to the $\mathrm{EMSE}(\tilde{\mathbf{V}})$ is given by

$$\mathrm{EMSE}(\tilde{\mathbf{V}}) \approx \mathbf{T}_1 + \mathbf{T}_2 \quad (55)$$

where

$$\mathbf{T}_1 := \frac{\sigma_n^2}{N_{\rm tr}}\mathrm{tr}(\mathbf{H}^T\boldsymbol{\mathcal{B}}^{-T}\mathbf{H}^*)\mathrm{tr}\left(\mathbf{V}_o\mathbf{V}_o^H\right) \quad (56)$$

and

$$\mathbf{T}_2 := \frac{\sigma_n^2}{N_{\rm tr}}\mathrm{tr}\left((\mathbf{C}_o - \mathbf{V}_o\mathbf{H})^H(\mathbf{C}_o - \mathbf{V}_o\mathbf{H})\right)\mathrm{tr}(\boldsymbol{\mathcal{B}}^{-T}). \quad (57)$$

Proof: The proof is provided in Appendix III. □

Substituting (47) and (55) into (42), we obtain a second-order approximation to the EMSE induced by channel uncertainties at both the transmitter and the receiver. Admittedly, this approximation is complicated and difficult to interpret. In the sequel, we shall develop simple and insightful high-SNR expressions.

IV. EMSE—High-SNR Approximations

In this section, we focus on the high-SNR regime and we derive a simple upper bound to $\mathrm{EMSE}(\hat{\mathbf{C}})$ and a simple approximation to $\mathrm{EMSE}(\tilde{\mathbf{V}})$. Putting these expressions together, we obtain a simple high-SNR EMSE upper bound for the mismatched MIMO-TH precoder. Finally, we average over the channel statistics and obtain a simple high-SNR upper bound for the expected value of the EMSE to MMSE ratio.

High-SNR regime means "small" σ_n^2. Our results will be derived either by ignoring $O(\sigma_n^2)$ terms compared with $O(1)$ terms or by ignoring $O(\sigma_n^4)$ terms compared with $O(\sigma_n^2)$ terms. We

[3]The first derivative of $\mathrm{MSE}_o(\mathbf{V})$ at $\mathbf{V}_o$ is zero because $\mathbf{V}_o$ minimizes the function $\mathrm{MSE}_o(\mathbf{V})$.

proceed by presenting some high-SNR approximations that will be useful in the sequel.

Using the definition of matrix $\mathbf{R}$ in (17), it can be shown that for high SNR

$$\mathbf{R}^{-H}\mathbf{H}^{H}\mathbf{H}\mathbf{R}^{-1} \approxeq \mathbf{I}_{n_t} \tag{58}$$

and

$$\mathrm{tr}(\mathbf{R}^{-T}\mathbf{R}^{-*}) \approxeq \mathrm{tr}\left((\mathbf{H}^{H}\mathbf{H})^{-1}\right). \tag{59}$$

Furthermore (the proof is provided in Appendix V)

$$\mathrm{tr}\left(\mathbf{V}_o\mathbf{V}_o^{H}\right) \approxeq \frac{1}{\sigma_n^2}\mathrm{MMSE}. \tag{60}$$

Using the matrix inversion lemma [16], it can be shown that

$$\mathbf{H}^{H}\left(\mathbf{H}\mathbf{H}^{H}+\zeta\mathbf{I}_{n_r}\right)^{-1} = \left(\mathbf{H}^{H}\mathbf{H}+\zeta\mathbf{I}_{n_t}\right)^{-1}\mathbf{H}^{H}. \tag{61}$$

Then, using (54), (61), and the high-SNR assumption, we get

$$\begin{aligned}\mathrm{tr}(\mathbf{H}^{T}\boldsymbol{\mathcal{B}}^{-T}\mathbf{H}^{*}) &= \mathrm{tr}\left(\mathbf{H}^{T}\left(\mathbf{H}^{*}\mathbf{H}^{T}+\zeta\mathbf{I}_{n_r}\right)^{-1}\mathbf{H}^{*}\right)\\ &= \mathrm{tr}\left(\left(\mathbf{H}^{T}\mathbf{H}^{*}+\zeta\mathbf{I}_{n_t}\right)^{-1}\mathbf{H}^{T}\mathbf{H}^{*}\right)\\ &\approxeq \mathrm{tr}\left(\mathbf{I}_{n_t}\right).\end{aligned} \tag{62}$$

Finally, using (19) and (61), we can write matrix $\mathbf{C}_o-\mathbf{V}_o\mathbf{H}$ as

$$\begin{aligned}\mathbf{C}_o-\mathbf{V}_o\mathbf{H} &= \mathbf{C}_o\left(\mathbf{I}_{n_t}-\mathbf{H}^{H}\left(\mathbf{H}\mathbf{H}^{H}+\zeta\mathbf{I}_{n_r}\right)^{-1}\mathbf{H}\right)\\ &= \mathbf{C}_o\left(\mathbf{I}_{n_t}-\left(\mathbf{H}^{H}\mathbf{H}+\zeta\mathbf{I}_{n_t}\right)^{-1}\mathbf{H}^{H}\mathbf{H}\right)\end{aligned} \tag{63}$$

and for high SNR

$$\mathbf{C}_o-\mathbf{V}_o\mathbf{H} \approxeq \mathbf{O}_{n_t}. \tag{64}$$

A. High SNR—Channel Uncertainties Only at the Transmitter

Lemma 3: In the high-SNR regime, the following approximate inequality holds

$$\mathrm{EMSE}(\hat{\mathbf{C}}) \lessapprox 2(1-\rho_\tau)(n_t-1)\mathrm{tr}\left((\mathbf{H}^{H}\mathbf{H})^{-1}\right)\mathrm{MMSE}. \tag{65}$$

Proof: Using (58) in (48) and ignoring the term that involves σ_n^4, we obtain

$$\begin{aligned}\mathbf{B}_1 &\approxeq \alpha\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\mathbf{L}^T\mathbf{L}\left(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{I}_{n_t}\right)\mathbf{L}^T\mathbf{L}\right)\\ &\overset{(a)}{\leq} \alpha\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\left(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{I}_{n_t}\right)\right)\\ &\overset{(3)}{=} \alpha\sigma_n^2\mathrm{tr}(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{G}^2)\\ &\overset{(4)}{=} \alpha\sigma_n^2\mathrm{tr}(\mathbf{R}^{-T}\mathbf{R}^{-*})\mathrm{tr}(\mathbf{G}^2)\end{aligned} \tag{66}$$

where at point (a) we used the structure of the elimination and the commutation matrices and the fact that matrices $\mathbf{G}^2$ and $\mathbf{R}^{-T}\mathbf{R}^{-*}$ have positive diagonal elements.

Using (58) in (49) and ignoring the term that involves σ_n^4, we obtain

$$\begin{aligned}\mathbf{B}_2 &\approxeq -2\alpha\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\mathbf{P}\left(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{I}_{n_t}\right)\mathbf{P}\right)\\ &= -2\alpha\sigma_n^2\mathrm{tr}(\mathbf{G}^2\mathbf{R}^{-T}\mathbf{R}^{-*}).\end{aligned} \tag{67}$$

The proof of the last equality is provided in Appendix IV for the $n_r\times 2$ case (the generalization is easy).

Finally, using (58) in (50) and ignoring the term involving σ_n^4, we obtain

$$\begin{aligned}\mathbf{B}_3 &\approxeq \alpha\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\mathbf{L}^T\mathbf{L}\left(\mathbf{I}_{n_t}\otimes\mathbf{R}^{-H}\mathbf{R}^{-1}\right)\mathbf{L}^T\mathbf{L}\right)\\ &\overset{(b)}{\leq} \alpha\sigma_n^2\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\left(\mathbf{I}_{n_t}\otimes\mathbf{R}^{-H}\mathbf{R}^{-1}\right)\right)\\ &\overset{(3)}{=} \alpha\sigma_n^2\mathrm{tr}\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\mathbf{R}^{-H}\mathbf{R}^{-1}\right)\\ &= n_t\alpha\sigma_n^2\mathrm{tr}(\mathbf{G}^2\mathbf{R}^{-H}\mathbf{R}^{-1})\\ &= n_t\alpha\sigma_n^2\mathrm{tr}(\mathbf{G}^2\mathbf{R}^{-T}\mathbf{R}^{-*})\end{aligned} \tag{68}$$

where at point (b) we used the structure of the elimination and the commutation matrices and the fact that matrices $\mathbf{G}^2$ and $\mathbf{R}^{-H}\mathbf{R}^{-1}$ have positive diagonal elements.

Combining expressions (47) and (66)–(68), we obtain

$$\begin{aligned}\mathrm{EMSE}(\hat{\mathbf{C}}) &\lessapprox \alpha\sigma_n^2\mathrm{tr}(\mathbf{R}^{-T}\mathbf{R}^{-*})\mathrm{tr}(\mathbf{G}^2)\\ &\quad+(n_t-2)\alpha\sigma_n^2\mathrm{tr}(\mathbf{G}^2\mathbf{R}^{-T}\mathbf{R}^{-*})\\ &\overset{(6)}{\leq} \alpha\sigma_n^2(n_t-1)\mathrm{tr}(\mathbf{R}^{-T}\mathbf{R}^{-*})\mathrm{tr}(\mathbf{G}^2).\end{aligned} \tag{69}$$

Using (69), (21) and (59), we conclude with the following bound:

$$\mathrm{EMSE}(\hat{\mathbf{C}}) \lessapprox \alpha(n_t-1)\mathrm{tr}\left((\mathbf{H}^{H}\mathbf{H})^{-1}\right)\mathrm{MMSE}. \tag{70}$$

Finally, recalling the definition of α as $\alpha := 2(1-\rho_\tau)$, we obtain (65) to prove Lemma 3. □

B. High SNR—Channel Uncertainties Only at the Receiver

Lemma 4: In the high-SNR regime, the following approximation holds

$$\mathrm{EMSE}(\tilde{\mathbf{V}}) \approxeq \frac{n_t}{N_{\mathrm{tr}}}\mathrm{MMSE}. \tag{71}$$

Proof: Starting with $\mathbf{T}_1$ in (56) and using (60) and (62), we obtain

$$\mathbf{T}_1 \approxeq \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\mathbf{V}_o\mathbf{V}_o^{H}\right)\mathrm{tr}\left(\mathbf{I}_{n_t}\right) \approxeq \frac{n_t}{N_{\mathrm{tr}}}\mathrm{MMSE}. \tag{72}$$

Using (64) in (57), we get

$$\mathbf{T}_2 \approxeq 0. \tag{73}$$

We conclude that, for sufficiently high SNR, term $\mathbf{T}_2$ is negligible compared with $\mathbf{T}_1$. Combining expressions (55), (72), and (73), we obtain (71) to prove Lemma 4. □

TABLE I
ELEMENTS OF CHANNEL MATRIX $\mathbf{H}$

0.2877 + 0.2097*j	1.5537 - 1.0653*j	0.3450 - 0.5177*j	-0.8714 - 0.2760*j
-0.2433 - 0.5179*j	-0.8435 - 0.2245*j	0.1448 + 1.0146*j	-0.2163 - 0.5226*j
-1.0170 - 0.5243*j	0.5028 - 0.4757*j	-1.1749 + 0.5322*j	1.0848 + 0.5731*j
-0.7636 - 0.6970*j	-0.6756 + 0.3384*j	0.4666 - 0.0437*j	0.2691 + 1.0737*j
0.4901 + 0.0910*j	0.0804 + 0.0937*j	-0.5181 - 0.7709*j	-0.4012 - 0.0189*j
-0.5236 + 0.9038*j	0.2247 + 0.2181*j	0.6343 - 0.8332*j	-0.5841 + 0.1868*j

C. High SNR—Channel Uncertainties at Both the Transmitter and the Receiver

Proposition 2: The high-SNR MIMO-TH EMSE induced by channel uncertainties at both the transmitter and the receiver is upper bounded as

$$\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\hat{\mathbf{V}}}) \lessapprox 2(1-\rho_\tau)(n_t-1)\mathrm{tr}\left((\mathbf{H}^H\mathbf{H})^{-1}\right)\mathrm{MMSE} + \frac{n_t}{N_{\mathrm{tr}}}\mathrm{MMSE}. \tag{74}$$

Proof: The proof requires only the substitution of (65) and (71) into (42). □

We observe that the EMSE is upper bounded by an expression proportional to the MMSE. The proportionality factor is determined by the system parameters n_t and N_{tr}, the channel correlation coefficient ρ_τ and the conditioning of the channel matrix through $\mathrm{tr}((\mathbf{H}^H\mathbf{H})^{-1})$.

In the simulations section, we shall observe that this high-SNR bound is in many cases tight because the EMSE due to the channel inaccuracies only at the receiver dominates the EMSE due to channel inaccuracies only at the transmitter.

D. High SNR—Averaging Over the Channels

In this subsection, we compute the average, over the channels, of the EMSE to MMSE ratio.

Proposition 3: Taking expectation with respect to the channels in (74), we obtain the following bound for the average EMSE to MMSE ratio, for $n_r > n_t$,

$$\mathcal{E}\left[\frac{\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\hat{\mathbf{V}}})}{\mathrm{MMSE}}\right] \lessapprox 2(1-\rho_\tau)\frac{n_t(n_t-1)}{n_r-n_t} + \frac{n_t}{N_{\mathrm{tr}}}. \tag{75}$$

Proof: Bound (74) can be written as

$$\frac{\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\hat{\mathbf{V}}})}{\mathrm{MMSE}} \lessapprox 2(1-\rho_\tau)(n_t-1)\mathrm{tr}\left((\mathbf{H}^H\mathbf{H})^{-1}\right) + \frac{n_t}{N_{\mathrm{tr}}}. \tag{76}$$

If we take expectation with respect to the channel, we get

$$\mathcal{E}\left[\frac{\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\hat{\mathbf{V}}})}{\mathrm{MMSE}}\right] \lessapprox 2(1-\rho_\tau)(n_t-1)\mathcal{E}\left[\mathrm{tr}\left((\mathbf{H}^H\mathbf{H})^{-1}\right)\right] + \frac{n_t}{N_{\mathrm{tr}}}. \tag{77}$$

It can be shown that if the elements of $\mathbf{H}$ are zero-mean, unit variance i.i.d. circular complex Gaussian random variables and $n_r > n_t$, then [15]

$$\mathcal{E}\left[\mathrm{tr}\left((\mathbf{H}^H\mathbf{H})^{-1}\right)\right] = \frac{n_t}{n_r-n_t}. \tag{78}$$

Substituting (78) in (77), we prove (75). □

We observe that the average EMSE to MMSE ratio is upper bounded by an expression which depends on the system parameters n_t, n_r, and N_{tr}, and the channel correlation coefficient ρ_τ.

V. SIMULATION RESULTS

In the first part of our experiments, we illustrate Propositions 1 and 2 using a *specific* channel realization, by taking averages over the channel uncertainties. More specifically, we consider a system with $n_t = 4$ transmit antennas and $n_r = 6$ receive antennas and channel matrix $\mathbf{H}$ with elements given in Table I.[4] The noise is spatially and temporally white, circularly symmetric complex Gaussian with variance σ_n^2. The input symbols are i.i.d., drawn from a 4-QAM constellation. We assume that the training block consists of $N_{\mathrm{tr}} = 10$ columns. We set the channel correlation coefficient equal to $\rho_\tau = 0.99$. We define the SNR as the ratio of the total receive power to the total noise power

$$\mathrm{SNR} := \frac{\mathcal{E}_{\mathbf{x}}\left[\mathrm{tr}(\mathbf{H}\mathbf{x}\mathbf{x}^H\mathbf{H}^H)\right]}{\mathcal{E}_{\mathbf{n}}\left[\mathrm{tr}(\mathbf{n}\mathbf{n}^H)\right]} = \frac{\sigma_x^2\|\mathbf{H}\|_F^2}{n_r\sigma_n^2}. \tag{79}$$

In Fig. 2, we plot the MMSE (20), the average of the MSEs for the case of channel inaccuracies only at the transmitter (the average is over $\Delta\mathbf{H}_{\mathrm{Tx}}$), the average of the MSEs for the case of channel inaccuracies only at the receiver (the average is over $\Delta\mathbf{H}_{\mathrm{Rx}}$), and finally the average of the MSEs for inaccuracies at both the transmitter and the receiver. We observe that the EMSE component due to $\Delta\mathbf{H}_{\mathrm{Rx}}$ is significantly larger than that due to $\Delta\mathbf{H}_{\mathrm{Tx}}$. This observation is in agreement with our theoretical results because the high-SNR approximations (65) and (71) indicate that both EMSEs are proportional to the MMSE, with the proportionality factor in (71) being larger than the one in (65), as long as the channel matrix $\mathbf{H}$ is well conditioned and the channel correlation coefficient ρ_τ is relatively large. An explanation of this phenomenon might be the fact that in the first case the receiver is optimized by taking into account the channel uncertainties at the transmitter while something analogous does not happen in the latter case.

In Fig. 3, we present the experimentally computed EMSE, the theoretical second-order approximation as the sum of (47) and (55), and the EMSE bound in (74). We observe that the experimental and theoretical EMSE values practically coincide for SNR higher than 5 dB. Also, the EMSE bound is very close to the true EMSE for SNR higher than 15 dB.

[4]Analogous results have been obtained in extended simulations with other channel realizations.

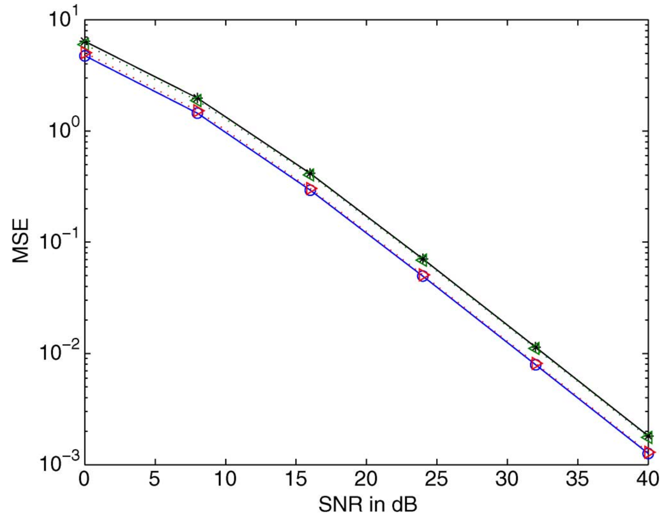

Fig. 2. MMSE using the true channel ("—o"), expectation of the MSEs for channel inaccuracies only at transmitter ("▷"), expectation of the MSEs for channel inaccuracies only at the receiver ("◁") and expectation of the MSEs for channel inaccuracies at both the transmitter and the receiver ("—∗").

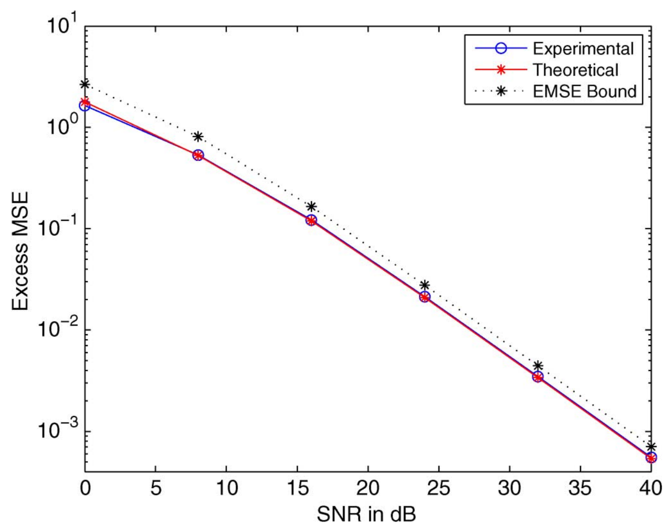

Fig. 3. Experimentally computed EMSE, theoretical second-order approximation (sum of (47) and (55)), and EMSE bound in (74).

In the second part of our experiments, we take averages over the channel matrices by assuming that the elements of $\mathbf{H}$ are i.i.d. $\mathcal{CN}(0,1)$. The SNR in this case is defined as

$$\mathrm{SNR} := \frac{\mathcal{E}_{\mathbf{x},\mathbf{H}}\left[\mathrm{tr}(\mathbf{H}\mathbf{x}\mathbf{x}^H\mathbf{H}^H)\right]}{\mathcal{E}_{\mathbf{n}}\left[\mathrm{tr}(\mathbf{n}\mathbf{n}^H)\right]} = \frac{\sigma_x^2 n_t}{\sigma_n^2}. \qquad (80)$$

In Fig. 4, we plot the experimentally computed EMSE and the theoretical second-order approximation, i.e., the sum of (47) and (55), averaged over different channel realizations, for the parameters defined above. We observe that the two curves coincide for SNR higher than 7 dB, meaning that our analysis holds for this case too, although it is difficult to give a simple expression for the theoretical second-order approximation.

In Fig. 5, we plot the experimental average ratio EMSE/MMSE and the simple bound in (75). We observe that the bound in (75) is very close to the true average EMSE to MMSE ratio, which attains a constant value for sufficiently high SNR.

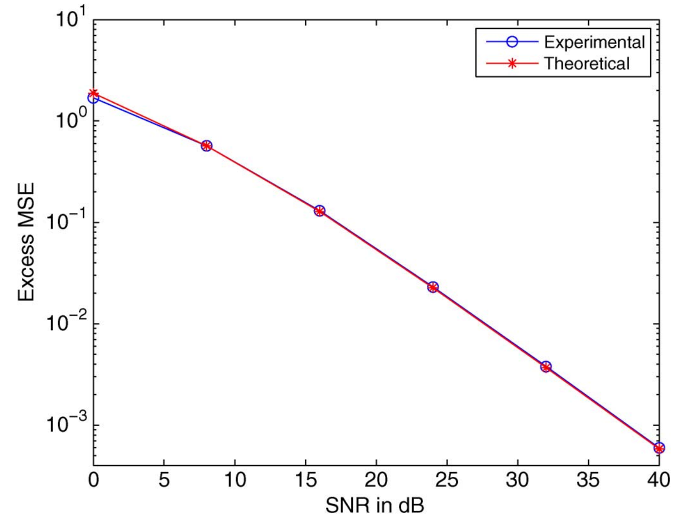

Fig. 4. Experimentally computed EMSE and theoretical second-order approximation (sum of (47) and (55)) averaged over different channel realizations.

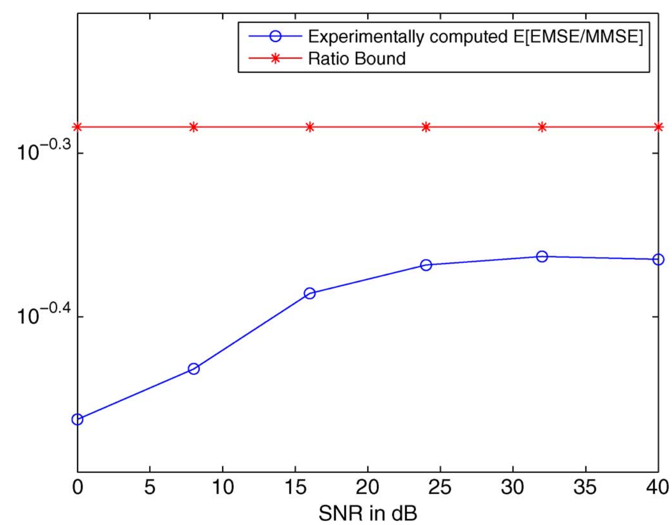

Fig. 5. Experimentally computed averaged ratio EMSE/MMSE and the corresponding bound in (75).

VI. Conclusion

We considered the sensitivity of the mismatched MIMO-TH with respect to channel estimation errors and channel time-variations. We developed a second-order EMSE approximation which, unfortunately, was difficult to interpret. We focused on the high-SNR regime and derived a simple and informative EMSE upper bound that uncovers the factors that determine the sensitivity of the MIMO-TH precoder with respect to channel uncertainties at both the transmitter and the receiver. Numerical experiments were in agreement with our theoretical analysis.

Appendix I

Proof of Proposition 1: The aim is to compute the EMSE assuming channel inaccuracies at both the transmitter and the receiver. The matrix filters used in this case are given by (31) and (34).

We have already defined $\Delta\mathbf{C}$ and $\Delta\mathbf{V}$, as $\Delta\mathbf{C} := \hat{\mathbf{C}} - \mathbf{C}_o$ and $\Delta\mathbf{V} := \tilde{\mathbf{V}} - \mathbf{V}_o$, respectively. We have also mentioned (and prove in Appendixes II and III) that $\Delta\mathbf{C}$ depends only on $\Delta\mathbf{H}_{\mathrm{Tx}}$, while $\Delta\mathbf{V}$ depends only on $\Delta\mathbf{H}_{\mathrm{Rx}}$. We recall that $\Delta\mathbf{H}_{\mathrm{Tx}}$ and $\Delta\mathbf{H}_{\mathrm{Rx}}$ are independent.

In order to compute the EMSE defined in (36), we define $\Delta\check{\mathbf{V}} := \hat{\check{\mathbf{V}}} - \mathbf{V}_o$ and use (34) and (39). Then

$$\begin{aligned}\Delta\check{\mathbf{V}} &:= \hat{\check{\mathbf{V}}} - \mathbf{V}_o\\ &= (\mathbf{C}_o + \Delta\mathbf{C})\tilde{\mathbf{H}}^H\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \zeta\mathbf{I}_{n_r}\right)^{-1} - \mathbf{V}_o\\ &= \tilde{\mathbf{V}} - \mathbf{V}_o + \Delta\mathbf{C}\tilde{\mathbf{H}}^H\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \zeta\mathbf{I}_{n_r}\right)^{-1}\\ &= \Delta\mathbf{V} + \Delta\mathbf{C}\underbrace{\tilde{\mathbf{H}}^H\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \zeta\mathbf{I}_{n_r}\right)^{-1}}_{\mathcal{X}}. \end{aligned}\tag{81}$$

We observe that $\Delta\check{\mathbf{V}}$ depends on both $\Delta\mathbf{H}_{\mathrm{Tx}}$ and $\Delta\mathbf{H}_{\mathrm{Rx}}$, through $\Delta\mathbf{C}$ and $\Delta\mathbf{V}$, respectively. If we write term $\mathcal{X}$ using (10) and then keep only the first-order terms, we get

$$\begin{aligned}\mathcal{X} &= \mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1} - \mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1}\left(\underbrace{\mathbf{H}\Delta\mathbf{H}_{\mathrm{Rx}}^H + \Delta\mathbf{H}_{\mathrm{Rx}}\mathbf{H}^H}_{\Delta\boldsymbol{\mathcal{B}}}\right)\boldsymbol{\mathcal{B}}^{-1}\\ &\quad + \Delta\mathbf{H}_{\mathrm{Rx}}^H\boldsymbol{\mathcal{B}}^{-1} + O\left(\|\Delta\mathbf{H}_{\mathrm{Rx}}\|^2\right)\\ &\simeq \mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1} - \mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1}\Delta\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^{-1} + \Delta\mathbf{H}_{\mathrm{Rx}}^H\boldsymbol{\mathcal{B}}^{-1}\end{aligned}\tag{82}$$

where matrix $\boldsymbol{\mathcal{B}}$ is defined in (54). Combining (81) and (82), we get

$$\Delta\check{\mathbf{V}} \simeq \Delta\mathbf{V} + \Delta\mathbf{C}\mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1}.\tag{83}$$

Next, we return to the EMSE definition in (36). We first substitute $\hat{\mathbf{C}}$ and $\hat{\check{\mathbf{V}}}$ with $\hat{\mathbf{C}} = \mathbf{C}_o + \Delta\mathbf{C}$ and $\hat{\check{\mathbf{V}}} = \Delta\check{\mathbf{V}} + \mathbf{V}_o$, respectively, in (35). Then, using the definition (36), we get

$$\begin{aligned}&\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\check{\mathbf{V}}})\\ &\quad= \underbrace{\sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\check{\mathbf{V}}\boldsymbol{\mathcal{B}}\Delta\check{\mathbf{V}}^H)\right]}_{\boldsymbol{\mathcal{Y}}_1} - \underbrace{\sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\check{\mathbf{V}}\mathbf{H}\Delta\mathbf{C}^H)\right]}_{\boldsymbol{\mathcal{Y}}_2}\\ &\qquad - \underbrace{\sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{C}\mathbf{H}^H\Delta\check{\mathbf{V}})\right]}_{\boldsymbol{\mathcal{Y}}_3} + \sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{C}\Delta\mathbf{C}^H)\right].\end{aligned}\tag{84}$$

Using (83), and the fact that $\Delta\mathbf{H}_{\mathrm{Tx}}$ and $\Delta\mathbf{H}_{\mathrm{Rx}}$ are zero mean and independent (which implies independence between $\Delta\mathbf{C}$ and $\Delta\mathbf{V}$), terms $\boldsymbol{\mathcal{Y}}_i$ become

$$\begin{aligned}\boldsymbol{\mathcal{Y}}_1 &\approx \sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{V}\boldsymbol{\mathcal{B}}\Delta\mathbf{V}^H)\right] + \sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{C}\mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1}\mathbf{H}\Delta\mathbf{C}^H)\right]\\ &\overset{(53)}{=} \mathrm{EMSE}(\tilde{\mathbf{V}}) + \sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{C}\mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1}\mathbf{H}\Delta\mathbf{C}^H)\right]\end{aligned}\tag{85}$$

and

$$\boldsymbol{\mathcal{Y}}_2 = \boldsymbol{\mathcal{Y}}_3 \approx \sigma_x^2\mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{C}\mathbf{H}^H\boldsymbol{\mathcal{B}}^{-1}\mathbf{H}\Delta\mathbf{C}^H)\right].\tag{86}$$

Finally, we combine (84)–(86) and use (61) and (45) to get

$$\mathrm{EMSE}(\hat{\mathbf{C}}, \hat{\check{\mathbf{V}}}) \approx \mathrm{EMSE}(\hat{\mathbf{C}}) + \mathrm{EMSE}(\tilde{\mathbf{V}}).$$

Appendix II

Proof of Lemma 1: The aim is to develop a second-order approximation to $\mathrm{EMSE}(\hat{\mathbf{C}})$. Towards this purpose, we must develop a first-order approximation to $\Delta\mathbf{C}$ with respect to $\Delta\mathbf{H}_{\mathrm{Tx}}$. Using (31) and defining $\Delta\mathbf{G} := \hat{\mathbf{G}} - \mathbf{G}$ and $\Delta\mathbf{R} := \hat{\mathbf{R}} - \mathbf{R}$, we obtain

$$\hat{\mathbf{C}} = (\mathbf{G} + \Delta\mathbf{G})(\mathbf{R} + \Delta\mathbf{R}) = \mathbf{C} + \mathbf{G}\Delta\mathbf{R} + \Delta\mathbf{G}\mathbf{R} + \Delta\mathbf{G}\Delta\mathbf{R}.$$

Thus, a first-order approximation to $\Delta\mathbf{C}$, with respect to $\Delta\mathbf{R}$ and $\Delta\mathbf{G}$, is

$$\Delta\mathbf{C} \simeq \mathbf{G}\Delta\mathbf{R} + \Delta\mathbf{G}\mathbf{R}.\tag{87}$$

Next, we derive first-order approximations to $\Delta\mathbf{R}$ and $\Delta\mathbf{G}$, with respect to $\Delta\mathbf{H}_{\mathrm{Tx}}$. We start with $\Delta\mathbf{R}$. We remind that $\mathbf{R}^H\mathbf{R} = \boldsymbol{\mathcal{A}}$ and

$$\hat{\mathbf{R}}^H\hat{\mathbf{R}} = \hat{\mathbf{H}}^H\hat{\mathbf{H}} + \zeta\mathbf{I}_{n_t} \simeq \boldsymbol{\mathcal{A}} + \underbrace{\left(\mathbf{H}^H\Delta\mathbf{H}_{\mathrm{Tx}} + \Delta\mathbf{H}_{\mathrm{Tx}}^H\mathbf{H}\right)}_{\Delta\boldsymbol{\mathcal{A}}}.$$

Using a result for the Cholesky factorization of a perturbed positive definite matrix [14], we obtain

$$\hat{\mathbf{R}} := \mathbf{R} + \Delta\mathbf{R} \simeq \mathbf{R} + \mathbf{F}_L\mathbf{R}$$

where $\mathbf{F}_L$ is the lower triangular part of matrix $\mathbf{F} := \mathbf{R}^{-H}\Delta\boldsymbol{\mathcal{A}}\mathbf{R}^{-1}$, with diagonal elements equal to half the diagonal elements of $\mathbf{F}$. Thus,

$$\Delta\mathbf{R} \simeq \mathbf{F}_L\mathbf{R}.\tag{88}$$

For term $\Delta\mathbf{G}$, we have

$$\begin{aligned}\Delta\mathbf{G} &:= \hat{\mathbf{G}} - \mathbf{G}\\ &= \mathrm{diag}\left(\frac{1}{\hat{r}_{11}}, \cdots, \frac{1}{\hat{r}_{n_tn_t}}\right) - \mathrm{diag}\left(\frac{1}{r_{11}}, \cdots, \frac{1}{r_{n_tn_t}}\right)\\ &\overset{(a)}{\simeq} \mathrm{diag}\left(\frac{1}{r_{11}} - \frac{\Delta r_{11}}{r_{11}^2}, \cdots, \frac{1}{r_{n_tn_t}} - \frac{\Delta r_{n_tn_t}}{r_{n_tn_t}^2}\right)\\ &\quad - \mathrm{diag}\left(\frac{1}{r_{11}}, \cdots, \frac{1}{r_{n_tn_t}}\right)\end{aligned}$$

where at point (a) we used the first-order approximation

$$\frac{1}{\hat{r}_{ii}} \simeq \frac{1}{r_{ii}} - \frac{\Delta r_{ii}}{r_{ii}^2}.$$

Thus,

$$\Delta\mathbf{G} \simeq -\mathrm{diag}\left(\frac{\Delta r_{11}}{r_{11}^2}, \ldots, \frac{\Delta r_{n_tn_t}}{r_{n_tn_t}^2}\right).\tag{89}$$

Finally, using (88), (89), and the definition of matrix $\mathbf{G}$ in (18), we get

$$\begin{aligned}\Delta\mathbf{G} &\simeq -\,\mathrm{diag}\left(\mathrm{diag}(\mathbf{F}_L)\right)\mathrm{diag}\left(\mathrm{diag}(\mathbf{R})\right)^{-1}\\ &= -\,\frac{1}{2}\mathrm{diag}\left(\mathrm{diag}(\mathbf{F})\right)\mathbf{G}.\end{aligned}\tag{90}$$

Up to this point, we have expressed terms $\Delta\mathbf{R}$ and $\Delta\mathbf{G}$ as functions of the matrix $\mathbf{F}$, which, in turn, is a linear function of $\Delta\mathbf{H}_{\rm Tx}$. Next, we return to (45) and using (1), we write the EMSE as

$$\mathrm{EMSE}(\hat{\mathbf{C}})=\sigma_n^2\mathrm{tr}\left(\left(\boldsymbol{\mathcal{A}}^{-T}\otimes\mathbf{I}_{n_t}\right)\mathcal{E}\left[\mathrm{vec}(\Delta\mathbf{C})\mathrm{vec}^H(\Delta\mathbf{C})\right]\right). \tag{91}$$

Using (87), (88), and (90), and defining $\mathbf{D}_{\mathbf{F}_L} := \mathrm{diag}(\mathrm{diag}(\mathbf{F}_L))$, we can express term $\mathrm{vec}(\Delta\mathbf{C})$ as

$$\begin{aligned}\mathrm{vec}(\Delta\mathbf{C}) &\simeq \mathrm{vec}(\mathbf{G}\Delta\mathbf{R})+\mathrm{vec}(\Delta\mathbf{G}\mathbf{R})\\ &\simeq \mathrm{vec}(\mathbf{G}\mathbf{F}_L\mathbf{R})+\mathrm{vec}\left(-\mathbf{D}_{\mathbf{F}_L}\mathbf{G}\mathbf{R}\right)\\ &\overset{(2)}{=}(\mathbf{R}^T\otimes\mathbf{G})\mathrm{vec}(\mathbf{F}_L)-\left(\mathbf{R}^T\mathbf{G}^T\otimes\mathbf{I}_{n_t}\right)\mathrm{vec}\left(\mathbf{D}_{\mathbf{F}_L}\right)\\ &\overset{(*)}{=}(\mathbf{R}^T\otimes\mathbf{G})\mathrm{vec}(\mathbf{F}_L)-\left(\mathbf{R}^T\mathbf{G}^T\otimes\mathbf{I}_{n_t}\right)\mathrm{vec}\left(\mathbf{D}_{\mathbf{F}_L}\right)\\ &\quad\pm(\mathbf{R}^T\otimes\mathbf{G})\mathrm{vec}\left(\mathbf{D}_{\mathbf{F}_L}\right)\\ &=(\mathbf{R}^T\otimes\mathbf{G})\left(\mathrm{vec}(\mathbf{F}_L)+\mathrm{vec}\left(\mathbf{D}_{\mathbf{F}_L}\right)\right)\\ &\quad-\left(\left(\mathbf{R}^T\mathbf{G}^T\otimes\mathbf{I}_{n_t}\right)+(\mathbf{R}^T\otimes\mathbf{G})\right)\mathrm{vec}\left(\mathbf{D}_{\mathbf{F}_L}\right)\\ &=(\mathbf{R}^T\otimes\mathbf{G})\mathrm{vec}(\mathbf{F}_l)\\ &\quad-\left(\left(\mathbf{R}^T\mathbf{G}^T\otimes\mathbf{I}_{n_t}\right)+(\mathbf{R}^T\otimes\mathbf{G})\right)\mathrm{vec}\left(\mathbf{D}_{\mathbf{F}_L}\right)\end{aligned} \tag{92}$$

where $\mathbf{F}_l$ is the lower triangular part of $\mathbf{F}$. At point $(*)$ we add and subtract the same term in order to simplify our calculations.

Next, we express $\mathrm{vec}(\mathbf{F}_l)$ and $\mathrm{vec}(\mathbf{D}_{\mathbf{F}_L})$ in terms of $\mathrm{vec}(\Delta\mathbf{H}_{\rm Tx})$. Using (2), (3) and (8), we can write

$$\begin{aligned}\mathrm{vec}(\mathbf{F}_l) &= \mathbf{L}^T\mathrm{vech}(\mathbf{F}_l)=\mathbf{L}^T\mathrm{vech}(\mathbf{F})\\ &=\mathbf{L}^T\mathrm{vech}(\mathbf{R}^{-H}\Delta\boldsymbol{\mathcal{A}}\mathbf{R}^{-1})\\ &\overset{(8)}{=}\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H})\mathrm{vec}(\Delta\boldsymbol{\mathcal{A}})\\ &=\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H})\mathrm{vec}\left(\mathbf{H}^H\Delta\mathbf{H}_{\rm Tx}+\Delta\mathbf{H}_{\rm Tx}^H\mathbf{H}\right)\\ &\overset{(2)}{=}\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H})\left(\mathbf{I}_{n_t}\otimes\mathbf{H}^H\right)\mathrm{vec}(\Delta\mathbf{H}_{\rm Tx})\\ &\quad+\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H})\left(\mathbf{H}^T\otimes\mathbf{I}_{n_t}\right)\mathrm{vec}\left(\Delta\mathbf{H}_{\rm Tx}^H\right)\\ &\overset{(3)}{=}\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H}\mathbf{H}^H)\mathrm{vec}(\Delta\mathbf{H}_{\rm Tx})\\ &\quad+\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\mathbf{H}^T\otimes\mathbf{R}^{-H})\mathrm{vec}\left(\Delta\mathbf{H}_{\rm Tx}^H\right).\end{aligned} \tag{93}$$

We continue with $\mathrm{vec}(\mathbf{D}_{\mathbf{F}_L})$. Using (2), (3), (9) and defining $\mathbf{P} := \mathbf{L}^T\mathbf{L}\mathbf{K}\mathbf{L}^T\mathbf{L}$, we obtain

$$\begin{aligned}\mathrm{vec}\left(\mathbf{D}_{\mathbf{F}_L}\right) &\overset{(9)}{=}\frac{1}{2}\mathbf{L}^T\mathbf{L}\mathbf{K}\mathbf{L}^T\mathbf{L}\mathrm{vec}(\mathbf{F})=\frac{1}{2}\mathbf{P}\mathrm{vec}(\mathbf{R}^{-H}\Delta\boldsymbol{\mathcal{A}}\mathbf{R}^{-1})\\ &\overset{(2)}{=}\frac{1}{2}\mathbf{P}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H})\mathrm{vec}\left(\mathbf{H}^H\Delta\mathbf{H}_{\rm Tx}+\Delta\mathbf{H}_{\rm Tx}^H\mathbf{H}\right)\\ &=\frac{1}{2}\mathbf{P}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H})\left(\mathbf{I}_{n_t}\otimes\mathbf{H}^H\right)\mathrm{vec}(\Delta\mathbf{H}_{\rm Tx})\\ &\quad+\frac{1}{2}\mathbf{P}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H})\left(\mathbf{H}^T\otimes\mathbf{I}_{n_t}\right)\mathrm{vec}\left(\Delta\mathbf{H}_{\rm Tx}^H\right)\\ &\overset{(3)}{=}\frac{1}{2}\mathbf{P}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H}\mathbf{H}^H)\mathrm{vec}(\Delta\mathbf{H}_{\rm Tx})\\ &\quad+\frac{1}{2}\mathbf{P}(\mathbf{R}^{-T}\mathbf{H}^T\otimes\mathbf{R}^{-H})\mathrm{vec}\left(\Delta\mathbf{H}_{\rm Tx}^H\right).\end{aligned} \tag{94}$$

We return to (92), and using (93), (94) and (7), after some calculations, we obtain

$$\begin{aligned}\mathrm{vec}(\Delta\mathbf{C}) &\simeq \mathbf{M}_1\mathrm{vec}(\Delta\mathbf{H}_{\rm Tx})+\mathbf{M}_2\mathrm{vec}\left(\Delta\mathbf{H}_{\rm Tx}^H\right)\\ &=\mathbf{M}_1\mathrm{vec}(\Delta\mathbf{H}_{\rm Tx})+\mathbf{M}_2\mathbf{K}\mathrm{vec}\left(\Delta\mathbf{H}_{\rm Tx}^*\right)\end{aligned} \tag{95}$$

where

$$\begin{aligned}\mathbf{M}_1 &:= (\mathbf{R}^T\otimes\mathbf{G})\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H}\mathbf{H}^H)\\ &\quad-\frac{1}{2}\left(\left(\mathbf{R}^T\mathbf{G}^T\otimes\mathbf{I}_{n_t}\right)+(\mathbf{R}^T\otimes\mathbf{G})\right)\mathbf{P}(\mathbf{R}^{-T}\otimes\mathbf{R}^{-H}\mathbf{H}^H)\end{aligned} \tag{96}$$

and

$$\begin{aligned}\mathbf{M}_2 &:= (\mathbf{R}^T\otimes\mathbf{G})\mathbf{L}^T\mathbf{L}(\mathbf{R}^{-T}\mathbf{H}^T\otimes\mathbf{R}^{-H})\\ &\quad-\frac{1}{2}\left(\left(\mathbf{R}^T\mathbf{G}^T\otimes\mathbf{I}_{n_t}\right)+(\mathbf{R}^T\otimes\mathbf{G})\right)\mathbf{P}(\mathbf{R}^{-T}\mathbf{H}^T\otimes\mathbf{R}^{-H}).\end{aligned} \tag{97}$$

Using the circular symmetry of $\Delta\mathbf{H}_{\rm Tx}$, (29), (27), (24) and (91), we obtain

$$\begin{aligned}&\mathrm{EMSE}(\hat{\mathbf{C}})\\ &\approx\sigma_n^2\mathrm{tr}\left(\left(\boldsymbol{\mathcal{A}}^{-T}\otimes\mathbf{I}_{n_t}\right)\mathbf{M}_1(\boldsymbol{\Sigma}_{\rm est}+\boldsymbol{\Sigma}_{\rm tv})\mathbf{M}_1^H\right)\\ &\quad+\sigma_n^2\mathrm{tr}\left(\left(\boldsymbol{\mathcal{A}}^{-T}\otimes\mathbf{I}_{n_t}\right)\mathbf{M}_2\mathbf{K}(\boldsymbol{\Sigma}_{\rm est}+\boldsymbol{\Sigma}_{\rm tv})^*\mathbf{K}^H\mathbf{M}_2^H\right)\\ &=\left(\alpha+\frac{\sigma_n^2}{\sigma_x^2N_{\rm tr}}\right)\sigma_n^2\mathrm{tr}\left(\left(\boldsymbol{\mathcal{A}}^{-T}\otimes\mathbf{I}_{n_t}\right)\mathbf{M}_1\mathbf{M}_1^H\right)\\ &\quad+\left(\alpha+\frac{\sigma_n^2}{\sigma_x^2N_{\rm tr}}\right)\sigma_n^2\mathrm{tr}\left(\left(\boldsymbol{\mathcal{A}}^{-T}\otimes\mathbf{I}_{n_t}\right)\mathbf{M}_2\mathbf{M}_2^H\right)\end{aligned} \tag{98}$$

where we also defined the scalar α, as $\alpha := 2(1-\rho_\tau)$, and used that $\mathbf{K}\mathbf{K}^H = \mathbf{I}$. Using (96) and (97) and after some calculations, it can be shown that the second-order EMSE approximation (98) can be expressed as

$$\mathrm{EMSE}(\hat{\mathbf{C}})\approx\sum_{i=1}^{3}\mathbf{B}_i.$$

where terms $\mathbf{B}_i$ are given in (48)–(50). During the calculations, we also used that, from the definition of matrices $\mathbf{R}$ and $\boldsymbol{\mathcal{A}}$ in (17) and (46), respectively, we get

$$\mathbf{R}^*\boldsymbol{\mathcal{A}}^{-T}\mathbf{R}^T=\mathbf{I}_{n_t}. \tag{99}$$

Appendix III

Proof of Lemma 2: The aim is to develop a second-order approximation to $\mathrm{EMSE}(\tilde{\mathbf{V}})$. In order to compute the EMSE defined in (53), we must develop a first-order approximation to $\Delta\mathbf{V}$ with respect to $\Delta\mathbf{H}_{\rm Rx}$, which is defined as $\Delta\mathbf{H}_{\rm Rx} := \tilde{\mathbf{H}}-\mathbf{H}$. We can write $\tilde{\mathbf{V}}$ from (39) as

$$\begin{aligned}\tilde{\mathbf{V}} &= \mathbf{C}_o\left(\mathbf{H}^H+\Delta\mathbf{H}_{\rm Rx}^H\right)\\ &\quad\times\Bigg(\mathbf{H}\mathbf{H}^H+\zeta\mathbf{I}_{n_r}+\underbrace{\mathbf{H}\Delta\mathbf{H}_{\rm Rx}^H+\Delta\mathbf{H}_{\rm Rx}\mathbf{H}^H}_{\Delta\boldsymbol{\mathcal{B}}}+O\left(\|\Delta\mathbf{H}_{\rm Rx}\|^2\right)\Bigg)^{-1}.\end{aligned} \tag{100}$$

Using (10) and the definition of $\mathbf{V}_o$ in (19), we obtain

$$\tilde{\mathbf{V}} \simeq \mathbf{V}_o - \left(\mathbf{V}_o \Delta\boldsymbol{\mathcal{B}} - \mathbf{C}_o \Delta\mathbf{H}_{\mathrm{Rx}}^H\right)\boldsymbol{\mathcal{B}}^{-1}.$$

Thus, a first-order approximation to $\Delta\mathbf{V}$ is

$$\Delta\mathbf{V} \simeq \underbrace{-\left(\mathbf{V}_o \Delta\boldsymbol{\mathcal{B}} - \mathbf{C}_o \Delta\mathbf{H}_{\mathrm{Rx}}^H\right)}_{\mathbf{K}_\Delta}\boldsymbol{\mathcal{B}}^{-1} \tag{101}$$

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

$$\begin{aligned}\mathrm{EMSE}(\tilde{\mathbf{V}}) &= \sigma_x^2 \mathcal{E}\left[\mathrm{tr}(\Delta\mathbf{V}\boldsymbol{\mathcal{B}}\Delta\mathbf{V}^H)\right] \\ &\overset{(101)}{\approx} \sigma_x^2 \mathcal{E}\left[\mathrm{tr}\left(\mathbf{K}_\Delta \boldsymbol{\mathcal{B}}^{-1}\mathbf{K}_\Delta^H\right)\right] \\ &= \sigma_x^2 \mathcal{E}\left[\mathrm{tr}\left(\mathbf{I}_{n_t}\mathbf{K}_\Delta \boldsymbol{\mathcal{B}}^{-1}\mathbf{K}_\Delta^H\right)\right] \\ &\overset{(1)}{=} \sigma_x^2 \mathcal{E}\left[\mathrm{vec}^H(\mathbf{K}_\Delta)\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\mathrm{vec}(\mathbf{K}_\Delta)\right] \\ &= \sigma_x^2 \mathrm{tr}\left(\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\mathcal{E}\left[\mathrm{vec}(\mathbf{K}_\Delta)\mathrm{vec}^H(\mathbf{K}_\Delta)\right]\right).\end{aligned} \tag{102}$$

From the definitions of $\mathbf{K}_\Delta$ in (101), $\Delta\boldsymbol{\mathcal{B}}$ in (100), and (2), we obtain

$$\begin{aligned}\mathrm{vec}(\mathbf{K}_\Delta) &= -\,\mathrm{vec}(\mathbf{V}_o \Delta\mathbf{H}_{\mathrm{Rx}}\mathbf{H}^H) \\ &\quad + \mathrm{vec}\left((\mathbf{C}_o - \mathbf{V}_o\mathbf{H})\Delta\mathbf{H}_{\mathrm{Rx}}^H\right) \\ &= \underbrace{-(\mathbf{H}^* \otimes \mathbf{V}_o)}_{\boldsymbol{\mathcal{T}}_1}\mathrm{vec}(\Delta\mathbf{H}_{\mathrm{Rx}}) \\ &\quad + \underbrace{(\mathbf{I}_{n_r}\otimes(\mathbf{C}_o - \mathbf{V}_o\mathbf{H}))}_{\boldsymbol{\mathcal{T}}_2}\mathrm{vec}\left(\Delta\mathbf{H}_{\mathrm{Rx}}^H\right).\end{aligned} \tag{103}$$

Using (7), we get

$$\mathrm{vec}(\mathbf{K}_\Delta) = \boldsymbol{\mathcal{T}}_1 \mathrm{vec}(\Delta\mathbf{H}_{\mathrm{Rx}}) + \boldsymbol{\mathcal{T}}_2\mathbf{K}\mathrm{vec}\left(\Delta\mathbf{H}_{\mathrm{Rx}}^*\right).$$

Using the circular symmetry of $\Delta\mathbf{H}_{\mathrm{Rx}}$ and (30), we obtain

$$\begin{aligned}\mathrm{EMSE}(\tilde{\mathbf{V}}) \approx\ & \sigma_x^2 \mathrm{tr}\left(\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\boldsymbol{\mathcal{T}}_1 \boldsymbol{\Sigma}_{\mathrm{est}}\boldsymbol{\mathcal{T}}_1^H\right) \\ & + \sigma_x^2 \mathrm{tr}\left(\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\boldsymbol{\mathcal{T}}_2\mathbf{K}\boldsymbol{\Sigma}_{\mathrm{est}}^*\mathbf{K}^H\boldsymbol{\mathcal{T}}_2^H\right).\end{aligned}$$

Finally, using (24) we obtain the expression

$$\mathrm{EMSE}(\tilde{V}) \approx \mathbf{T}_1 + \mathbf{T}_2$$

where

$$\begin{aligned}\mathbf{T}_1 &:= \sigma_x^2 \mathrm{tr}\left(\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\boldsymbol{\mathcal{T}}_1 \boldsymbol{\Sigma}_{\mathrm{est}}\boldsymbol{\mathcal{T}}_1^H\right) \\ &= \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\boldsymbol{\mathcal{T}}_1^H\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\boldsymbol{\mathcal{T}}_1\right) \\ &\overset{(103)}{=} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\mathbf{H}^T\boldsymbol{\mathcal{B}}^{-T}\mathbf{H}^* \otimes \mathbf{V}_o\mathbf{V}_o^H\right) \\ &\overset{(4)}{=} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\mathbf{H}^T\boldsymbol{\mathcal{B}}^{-T}\mathbf{H}^*\right)\mathrm{tr}\left(\mathbf{V}_o\mathbf{V}_o^H\right)\end{aligned}$$

and

$$\begin{aligned}\mathbf{T}_2 &:= \sigma_x^2 \mathrm{tr}\left(\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\boldsymbol{\mathcal{T}}_2\mathbf{K}\boldsymbol{\Sigma}_{\mathrm{est}}^*\mathbf{K}^H\boldsymbol{\mathcal{T}}_2^H\right) \\ &= \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\mathbf{K}^H\boldsymbol{\mathcal{T}}_2^H\left(\boldsymbol{\mathcal{B}}^{-T}\otimes \mathbf{I}_{n_t}\right)\boldsymbol{\mathcal{T}}_2\mathbf{K}\right) \\ &\overset{(103)}{=} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\mathbf{K}^H\left(\boldsymbol{\mathcal{B}}^{-T}\otimes(\mathbf{C}_o - \mathbf{V}_o\mathbf{H})^H(\mathbf{C}_o - \mathbf{V}_o\mathbf{H})\right)\mathbf{K}\right) \\ &\overset{(5)}{=} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left(\left((\mathbf{C}_o - \mathbf{V}_o\mathbf{H})^H(\mathbf{C}_o - \mathbf{V}_o\mathbf{H})\otimes\boldsymbol{\mathcal{B}}^{-T}\right)\right) \\ &\overset{(4)}{=} \frac{\sigma_n^2}{N_{\mathrm{tr}}}\mathrm{tr}\left((\mathbf{C}_o - \mathbf{V}_o\mathbf{H})^H(\mathbf{C}_o - \mathbf{V}_o\mathbf{H})\right)\mathrm{tr}(\boldsymbol{\mathcal{B}}^{-T}).\end{aligned}$$

APPENDIX IV

In this Appendix, we prove the second equality in (67) for the $n_r \times 2$ case (i.e., $n_t = 2$). The aim is to simplify the trace term of the first line of (67)

$$\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\mathbf{P}\left(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{I}_{n_t}\right)\mathbf{P}\right). \tag{104}$$

For notational simplicity, we define matrices $\mathbf{Q}$ and $\mathbf{Z}$, as $\mathbf{Q} := \mathbf{R}^{-T}\mathbf{R}^{-*}$ and $\mathbf{Z} := \mathbf{G}^2$.

We first write the matrices inside the trace operator of (104). For the $n_r \times 2$ case,

$$\mathbf{P} = \begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix}$$

and

$$\mathbf{I}_2\otimes\mathbf{Z} = \begin{bmatrix}\mathbf{Z}&\mathbf{O}_2\\\mathbf{O}_2&\mathbf{Z}\end{bmatrix}.$$

For the other Kronecker product, we get

$$\mathbf{Q}\otimes\mathbf{I}_2 = \begin{bmatrix}q_{11}&0&q_{12}&0\\0&q_{11}&0&q_{12}\\q_{21}&0&q_{22}&0\\0&q_{21}&0&q_{22}\end{bmatrix}.$$

Then, the product of the matrices inside the trace operator is

$$(\mathbf{I}_{n_t}\otimes\mathbf{G}^2)\mathbf{P}\left(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{I}_{n_t}\right)\mathbf{P} = \begin{bmatrix}z_{11}q_{11}&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&z_{22}q_{22}\end{bmatrix}$$

and is obvious that

$$\begin{aligned}\mathrm{tr}\left(\left(\mathbf{I}_{n_t}\otimes\mathbf{G}^2\right)\mathbf{P}\left(\mathbf{R}^{-T}\mathbf{R}^{-*}\otimes\mathbf{I}_{n_t}\right)\mathbf{P}\right) &= \mathrm{tr}(\mathbf{Z}\mathbf{Q}) \\ &= \mathrm{tr}(\mathbf{G}^2\mathbf{R}^{-T}\mathbf{R}^{-*}).\end{aligned} \tag{105}$$

Using an analogous procedure, it can be shown that result (105) holds for the general $n_r \times n_t$ case.

APPENDIX V

In this Appendix, we simplify term $\mathrm{tr}(\mathbf{V}_o\mathbf{V}_o^H)$ in the high-SNR regime (i.e., $\zeta \rightarrow 0$). Using (16), (17), (19), and (61), we write matrix $\mathbf{V}_o$ as

$$\mathbf{V}_o = \mathbf{G}\mathbf{R}^{-H}\mathbf{H}^H.$$

Then, using (58), we get

$$\begin{aligned}\mathrm{tr}\left(\mathbf{V}_o\mathbf{V}_o^H\right) &= \mathrm{tr}(\mathbf{G}\mathbf{R}^{-H}\mathbf{H}^H\mathbf{H}\mathbf{R}^{-1}\mathbf{G}^H)\\ &\approxeq \mathrm{tr}(\mathbf{G}\mathbf{G}^H).\end{aligned}$$

Finally, using (21), we get

$$\mathrm{tr}\left(\mathbf{V}_o\mathbf{V}_o^H\right) \approxeq \frac{1}{\sigma_n^2}\mathrm{MMSE}.$$

REFERENCES

[1] R. F. H. Fischer, *Precoding and Signal Shaping for Digital Communications*. New York: Wiley, 2002.

[2] A. A. D' Amico and M. Morelli, "Joint TX-RX MMSE design for MIMO multicarrier systems with Tomlinson–Harashima pre-coding," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3118–3127, Aug. 2008.

[3] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1305–1316, Jul. 2004.

[4] Y. Zhu and K. B. Letaief, "Frequency domain equalization with Tomlinson–Harashima precoding for single carrier broadband MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 12, pp. 4420–4431, Dec. 2007.

[5] M. Joham, D. Schmidt, J. Brehmer, and W. Utschick, "Finite-length MMSE Tomlinson–Harashima precoding for frequency selective vector channels," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 3073–3088, Jun. 2007.

[6] M. B. Shenouda and T. N. Davidson, "Tomlinson–Harashima precoding for broadcast channels with uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1380–1389, Sep. 2007.

[7] F. A. Dietrich, P. Breun, and W. Utschick, "Robust Tomlinson–Harashima precoding for the wireless broadcast channel," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 631–644, Feb. 2007.

[8] M. Payaro, A. Pascual-Iserte, A. I. Perez-Neira, and M. A. Lagunas, "Robust design of spatial Tomlinson–Harashima Precoding in the presence of errors in the CSI," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2396–2401, Jul. 2007.

[9] H. Luetkepohl, *Handbook of Matrices*. New York: Wiley, 1996.

[10] N. Al-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Process.*, vol. 48, no. 10, pp. 2921–2936, Oct. 2000.

[11] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[12] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[13] 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.

[14] G. W. Stewart, "On the perturbation of LU, Cholesky, and QR factorizations," *SIAM J. Matrix Anal. Appl.*, vol. 14, no. 4, pp. 1141–1145, Oct. 1993.

[15] A. Lozano, A. M. Tulino, and S. Verdu, "Multiple-antenna capacity in the low-power regime," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2527–2544, Oct. 2003.

[16] G. Golub and C. Van Loan, *Matrix Computations*. Baltimore, MD: The John Hopkins Univ. Press, 1996.

Despoina Tsipouridou was born in Drama, Greece, in 1980. She received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece, and the M.Sc. degree in electronic and computer engineering from the Technical University of Crete, Greece, in 2004 and 2006, respectively. She is currently working toward the Ph.D. degree in the Telecommunications Division, Department of Electronic and Computer Engineering, Technical University of Crete, Chania, Greece.

Her research interests are in the area of Signal Processing for Communications.

Athanasios P. Liavas (M'89) received the Diploma and Ph.D. degrees from the University of Patras, Greece, in 1989 and 1993, respectively.

From 1993 to 1995, he served in the Greek Army. From 1996 to 1998, he was a Marie Curie postdoctoral Fellow at the Institut National des Telecommunications, Evry, France. From 1999 to 2001, he was visiting Assistant Professor at the Department of Informatics, University of Ioannina. In 2001, he became Assistant Professor at the Department of Mathematics, University of the Aegean. In 2004, he joined the Department of Electronic and Computer Engineering, Technical University of Crete, as Associate Professor. In September 2009, he became Professor and Department Chair.

Dr. Liavas served as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2005 to 2009. He is an elected member of the IEEE SPCOM Technical Committee (first election 2006, re-elected 2009).