

Impact of Polarization on the Achievable Rate of MIMO Systems with Linear Receivers

Adrian Ispas*, Cuiling Sun*, Christian Schneider†, Xitao Gong*, Gerd Ascheid*, and Reiner Thomä†

*Chair for Integrated Signal Processing Systems, RWTH Aachen University, Germany

{ispas, gong, ascheid}@iss.rwth-aachen.de, cuiling.sun@rwth-aachen.de

†Institute of Information Technology, Ilmenau University of Technology, Germany

{christian.schneider, reiner.thomae}@tu-ilmenau.de

Abstract—The extension of multiple-input multiple-output (MIMO) systems to the polarization domain is known to have potential benefits in terms of the spectral efficiency. It is thus of importance to understand how the channel influences the spectral efficiency, e.g., when a dual-polarized (DP) antenna setup should be favored over a single-polarized (SP) one. In this work, we study the achievable rate, i.e., the sum mutual information (MI) of all transmitted streams, over DP MIMO channels with linear receivers. We derive an approximation of the sum MI with a linear minimum mean squared error (LMMSE) or a zero forcing receiver which is an explicit function of statistical channel parameters. The evaluations of selected 4×4 SP and DP MIMO setups are based on channel measurements performed at 2.53 GHz in an urban macrocell scenario. The results reveal that DP antenna setups can yield substantial gains in the sum MI on links with high K -factors and high signal-to-noise ratios (SNRs). Moreover, the degradation in the sum MI with two streams due to the use of an LMMSE receiver is found to be small, especially for the DP antenna setup. The approximate evaluation of the sum MI is able to reproduce the SNR threshold region in which a second transmitted stream should be activated or one should switch from an SP to a DP antenna setup.

I. INTRODUCTION

The use of multiple antennas at the transmitter (TX) and the receiver (RX) allows for substantial enhancements of the spectral efficiency. It is widely known that the efficiency of multiple-input multiple-output (MIMO) transmission is highly dependent on the signal-to-noise ratio (SNR) as well as the channel conditions. It is thus of interest to understand how the channel influences the spectral efficiency of a MIMO system. Moreover, MIMO systems can be extended to make use of the polarization domain and offer, e.g., dual-polarized (DP) transmission [1], [2]. DP MIMO systems can be advantageous over single-polarized (SP) MIMO systems in terms of the spectral efficiency, e.g., in the presence of line-of-sight (LOS) [3]–[6]. As a strong decorrelation occurs over orthogonal polarizations, DP MIMO systems can be realized with co-located antennas and thus allow for compact antenna array designs. A summary of experimental characteristics of DP propagation can be found in [3]. There are various papers that report measurement-based results on the spectral efficiency of DP MIMO systems with an optimal RX, see, e.g., [4]–[7].

Practically, MIMO transmission faces several challenges besides a low SNR or unfavorable channel conditions. Exam-

ples are the acquirement of accurate channel state information (CSI) [8] and equalization at the RX to separate the transmitted streams. While the RX is able to obtain an accurate estimate of the instantaneous channel, the TX must rely on CSI due to feedback from the RX or on channel reciprocity. Both approaches are challenging to realize in practice. It is thus more realistic to assume the TX to only have statistical CSI which changes on a much slower scale than instantaneous CSI [9]–[11]. Based on statistical CSI, a low-complexity precoding can be used such as a semi-unitary precoding with equal power allocation [12]. At the RX side, a practical solution to the equalization problem is to use a linear filter, e.g., the linear minimum mean squared error (LMMSE) or the zero forcing (ZF) RX, with independently decoded streams.

Contributions: We study the achievable rate over DP MIMO channels with linear RXs when the TX has statistical CSI and the RX has instantaneous CSI. The evaluations are based on measured channels at 2.53 GHz with selected 4×4 SP and DP antenna setups. Specifically, our contributions are as follows:

- We derive an approximate evaluation of the achievable rate, i.e., the sum mutual information (MI) of all transmitted streams, for LMMSE and ZF RXs. The approximation is an explicit function of statistical channel parameters.
- We evaluate the achievable rate and its approximation for LMMSE and ZF RXs on links with varying K -factors. We assess when there is a gain in the achievable rate due to DP instead of SP transmission or the use of two instead of a single transmitted stream.
- We demonstrate that the loss in sum MI with two streams, due to the restriction to linear RXs, is small.

Notation: For a matrix \mathbf{A} , the (element-wise) complex conjugate, the transpose, and the conjugate transpose are denoted by \mathbf{A}^* , \mathbf{A}^T , and \mathbf{A}^H , respectively. The trace of the matrix \mathbf{A} is denoted by $\text{tr}\{\mathbf{A}\}$. For two matrices \mathbf{A} and \mathbf{B} , $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product. The column-wise stacking of the matrix \mathbf{A} is denoted by $\text{vec}\{\mathbf{A}\}$. The $N \times N$ identity matrix is represented by \mathbf{I}_N . The $MN \times MN$ commutation matrix $\mathbf{K}_{M,N}$ fulfills $\mathbf{K}_{M,N} \text{vec}\{\mathbf{A}\} = \text{vec}\{\mathbf{A}^T\}$ for an $M \times N$ matrix \mathbf{A} . We use $[\mathbf{A}]_{k,l}$ to denote the element in the k th row and l th column of \mathbf{A} . The expectation of a random variable x is denoted by $\mathbb{E}\{x\}$. $\log(\cdot)$ is the logarithm to the base 2, $\ln(\cdot)$ is the natural logarithm, and j denotes the imaginary unit.

This work was supported by the Ultra high-speed Mobile Information and Communication (UMIC) research centre.

II. SYSTEM MODEL

We consider a time-varying and frequency-flat fading MIMO channel with N_{TX} antennas at the TX and N_{RX} antennas at the RX. The input-output relation between the TX and the RX is given by the length- N_{RX} received column vector

$$\mathbf{y}[m] = \mathbf{H}[m]\mathbf{x}[m] + \mathbf{n}[m] \quad (1)$$

for time slots $m \in \mathbb{Z}$. The random channel matrices $\mathbf{H}[m]$ of size $N_{\text{RX}} \times N_{\text{TX}}$ are jointly proper. The zero-mean column vectors $\mathbf{x}[m]$ of length N_{TX} denote the jointly proper Gaussian transmitted vectors that are uncorrelated in time with $\mathbb{E}\{\mathbf{x}^H[m]\mathbf{x}[m]\} = P_x$. The column vectors $\mathbf{n}[m]$ of length N_{RX} are the white jointly proper Gaussian noise vectors in time with covariance matrix $\mathbb{E}\{\mathbf{n}[m]\mathbf{n}^H[m]\} = \sigma_n^2 \mathbf{I}_{N_{\text{RX}}}$. For ease of exposition, we define the (nominal) SNR $\rho = P_x/\sigma_n^2$.

We assume the RX to have instantaneous CSI, i.e., the RX has knowledge of the current channel realization $\mathbf{H}[m]$. The TX, on the other hand, has only statistical CSI of the channel. We use a semi-unitary precoding with equal power allocation along N_{st} transmitted streams [12]. Thus, we have $\mathbf{x}[m] = \mathbf{F}[m]\mathbf{s}[m]$ with the $N_{\text{TX}} \times N_{\text{st}}$ semi-unitary precoding matrix $\mathbf{F}[m]$ and the length N_{st} zero-mean jointly proper Gaussian column vector $\mathbf{s}[m]$ with covariance matrix $\mathbb{E}\{\mathbf{s}[m]\mathbf{s}^H[m]\} = P_x/N_{\text{st}}\mathbf{I}_{N_{\text{st}}}$. At the RX side, we use a linear equalization scheme, i.e., either an LMMSE or a ZF RX, with independently decoded streams.

III. PERFORMANCE EVALUATION

We assess the performance by means of the (time-dependent) MI between the input and the output combined with CSI at the RX for each stream. Considering the system model in Section II, the sum MI across the N_{st} transmitted streams can be stated for the LMMSE and the ZF RX in bit/channel use (bit/c.u.) as [13], [14]

$$I_{\text{LMMSE}}[m] = \sum_{k=1}^{N_{\text{st}}} -\mathbb{E} \left\{ \log \left(\left[(\mathbf{I}_{N_{\text{st}}} + \rho \mathbf{X}[m])^{-1} \right]_{k,k} \right) \right\} \quad (2)$$

$$I_{\text{ZF}}[m] = \sum_{k=1}^{N_{\text{st}}} \mathbb{E} \left\{ \log \left(1 + \frac{\rho}{[\mathbf{X}^{-1}[m]]_{k,k}} \right) \right\} \quad (3)$$

respectively, with $\mathbf{X}[m] = \mathbf{F}^H[m]\mathbf{H}^H[m]\mathbf{H}[m]\mathbf{F}[m]/N_{\text{st}}$.

In order to understand how statistical channel parameters influence the sum MI, we derive approximations of (2) and (3) based on a second-order multivariate Taylor series expansion [15] of $\mathbf{H}^H[m]\mathbf{H}[m]$ at its expected value $\mathbf{R}_{\text{TX}}^*[m]$. Here, $\mathbf{R}_{\text{TX}}[m] = \mathbb{E}\{\mathbf{H}^T[m]\mathbf{H}^*[m]\}$ is the transmit correlation matrix of the channel. It can be shown that the following expressions are then obtained:

$$I_{\text{LMMSE}}[m] \approx \sum_{k=1}^{N_{\text{st}}} \left[-\log \left(\left[(\mathbf{I}_{N_{\text{st}}} + \rho \mathbb{E}\{\mathbf{X}[m]\})^{-1} \right]_{k,k} \right) + \frac{\log(e)\rho^2}{2N_{\text{st}}^2} \text{tr} \left\{ (\mathbf{F}^T[m] \otimes \mathbf{F}^H[m]) \times \mathbf{Z}[m] (\mathbf{F}[m] \otimes \mathbf{F}^*[m]) \mathbf{Y}_{\text{LMMSE},k}[m] \right\} \right] \quad (4)$$

TABLE I
SPECIFICATION AND PROPERTIES OF THE REFERENCE LINKS

Link	BS	Track	MT orient.	MT pos. [m]	K -Factors
1	1	41a-42	back	0 – 34.9	low
2	3	9a-9b	left	0 – 38.9	medium
3	2	10b-9a	front	9.8 – 56.8	high

$$I_{\text{ZF}}[m] \approx \sum_{k=1}^{N_{\text{st}}} \left[\log \left(1 + \rho / \left[(\mathbb{E}\{\mathbf{X}[m]\})^{-1} \right]_{k,k} \right) + \frac{\log(e)}{2N_{\text{st}}^2} \text{tr} \left\{ (\mathbf{F}^T[m] \otimes \mathbf{F}^H[m]) \times \mathbf{Z}[m] (\mathbf{F}[m] \otimes \mathbf{F}^*[m]) \mathbf{Y}_{\text{ZF},k}[m] \right\} \right] \quad (5)$$

with $\mathbb{E}\{\mathbf{X}[m]\} = \mathbf{F}^H[m]\mathbf{R}_{\text{TX}}^*[m]\mathbf{F}[m]/N_{\text{st}}$ which is assumed to be invertible. The $N_{\text{TX}}^2 \times N_{\text{TX}}^2$ Hessian matrices $\mathbf{Y}_{\text{LMMSE},k}[m]$ and $\mathbf{Y}_{\text{ZF},k}[m]$, derived and specified for $N_{\text{st}} \geq 2$ in Appendix A and Appendix B, respectively, are functions of $\mathbf{R}_{\text{TX}}^*[m]$ only. The derivation for $N_{\text{st}} = 1$ follows in a straightforward manner. We further have the $N_{\text{TX}}^2 \times N_{\text{TX}}^2$ fourth-order moment matrix of the channel

$$\mathbf{Z}[m] = \mathbb{E} \left\{ \text{vec} \left\{ \mathbf{H}^H[m]\mathbf{H}[m] - \mathbf{R}_{\text{TX}}^*[m] \right\} \times \left(\text{vec} \left\{ \mathbf{H}^H[m]\mathbf{H}[m] - \mathbf{R}_{\text{TX}}^*[m] \right\} \right)^T \right\}. \quad (6)$$

To the best of our knowledge, the approximations of the sum MIs (4) and (5) are new. Their entire derivation is left out due to space constraints.

IV. CHANNEL MEASUREMENTS AND ANTENNA SETUPS

The urban macrocell channel measurements used in this work cover 2 bands of 45 MHz at 2.53 GHz in Ilmenau, Germany. Overall, DP MIMO links from 3 base station (BS) positions with different heights to a mobile terminal (MT) moving with a maximal velocity of about 10 km/h were measured on several tracks in a sequential fashion. Here, we extract a 20 MHz band at 2.505 GHz, and we use the BS positions at a height of 25 m with the 3 MT reference tracks [16]. Regarding the antenna setups, we use a uniform linear array at the BS and two uniform circular arrays (UCAs) (lying on top of each other) at the MT. This allows us to study four orientations at the MT, i.e., the front (direction of motion), the back, and the two sides of the MT. We classify the measurements into links with low, medium, and high (co-polarized) K -factors [17], [18], see Table I. The low K -factor links have K -factor values in $[0, 2]$, and the medium and high K -factor links have several peaks above values of 5 and 10, respectively. For more details about the scenario classification, we refer to [18]. Since we aim to study the effects of propagation along different polarizations, we differentiate between three antenna setups: two SP antenna setups, i.e., a vertical-polarized (VP) and a horizontal-polarized (HP) setup, and one DP antenna setup with an equal number of VP and HP antennas. All setups have four antennas at the TX as well as at the RX. For the SP setups, we have antenna separations of λ_c at the TX and $0.5\lambda_c$ (different UCAs) or $0.327\lambda_c$ (same UCA) at the RX. For the DP setup, we have co-located DP antennas at the TX

and the RX separated by $3\lambda_c$ and $0.5\lambda_c$ (across the UCAs), respectively. Thus, the array length at the TX remains constant.

We post-process the acquired measurement data with a denoising step in the time-delay domain. Then, we estimate the statistical quantities by replacing the ensemble averaging with an averaging in time and frequency over $N_t = 16$ and $N_f = 128$ samples, respectively. Overall, we thus average over 2048 (≈ 500 non-coherent) realizations [16]. We assemble the N_{co} co-polarized sub-links into the vector $\mathbf{h}_{\text{co}}[m]$. In order to account for the power loss in the cross-polarized sub-links and thus allowing for a fair comparison between the SP and the DP setups, we normalize the channel matrices $\mathbf{H}[m]$ such that $E\{\|\mathbf{h}_{\text{co}}[m]\|_F^2\} = N_{\text{co}}$ holds inside each region of size N_t in time and N_f in frequency.

V. RESULTS

We present results on the achievable rate, i.e., the sum MI, with LMMSE and ZF RXs for the three considered setups and the three links specified in Section IV. The columns of the precoding matrices $\mathbf{F}[m]$ are chosen as the N_{st} strongest eigenvectors of $\mathbf{R}_{\text{TX}}^*[m]$ [12]. We note that in this case (4) and (5) can be simplified since $E\{\mathbf{X}[m]\} = \mathbf{F}^H[m]\mathbf{R}_{\text{TX}}^*[m]\mathbf{F}[m]/N_{\text{st}}$ is a diagonal matrix. Moreover, we study single- and double stream transmission, i.e., $N_{\text{st}} \leq 2$, as we observed that, in the considered SNR range, a third stream can only increase the sum MI of the LMMSE and the ZF RX on the low- K -factor link 1. The results on the sum MI (averaged over each link) are given in Fig. 1, Fig. 2, and Fig. 3, respectively. We first focus on the SP setups. We observe that the SNR threshold at which one should activate a second stream increases with increasing K -factor. On the low- K -factor link, the SNR threshold is close to 0 dB, whereas, for the medium- and high- K -factor links, it is around or above 10 dB. In the case of the DP setup, the SNR threshold is close to or below 0 dB for all three links. Furthermore, the DP setup is only advantageous in terms of the sum MI for the medium- and high- K -factor links starting at medium SNRs. On the low- K -factor link, SP setups are always preferred. In Fig. 4, we compare the VP to the DP setup with the LMMSE receiver; for the medium- and high- K -factor links, the DP setup yields higher sum MI values above an SNR of 7.5 dB and 6 dB, respectively. We observe that the approximate evaluation of the sum MI is able to reproduce the SNR threshold regions at which a second stream should be activated or at which a DP is preferred over an SP setup.

In Fig. 5, we present the sum MI relative to the MI of an optimal RX on link 2. Obviously, we observe a high degradation of the sum MI with a ZF RX at low SNRs. The LMMSE receiver only suffers from small degradations; in the SP cases the lowest relative sum MI is around 0.9. The DP antenna setup is able to reduce the degradation further due to the use of two orthogonal polarizations since the power loss in the cross-polarized links can suppress the interference.

VI. CONCLUSION

In this paper, we studied the achievable rate over DP MIMO channels with linear receivers. For the evaluations, we used selected 4×4 SP and DP MIMO setups and realistic channel data from measurements in an urban macrocell scenario at

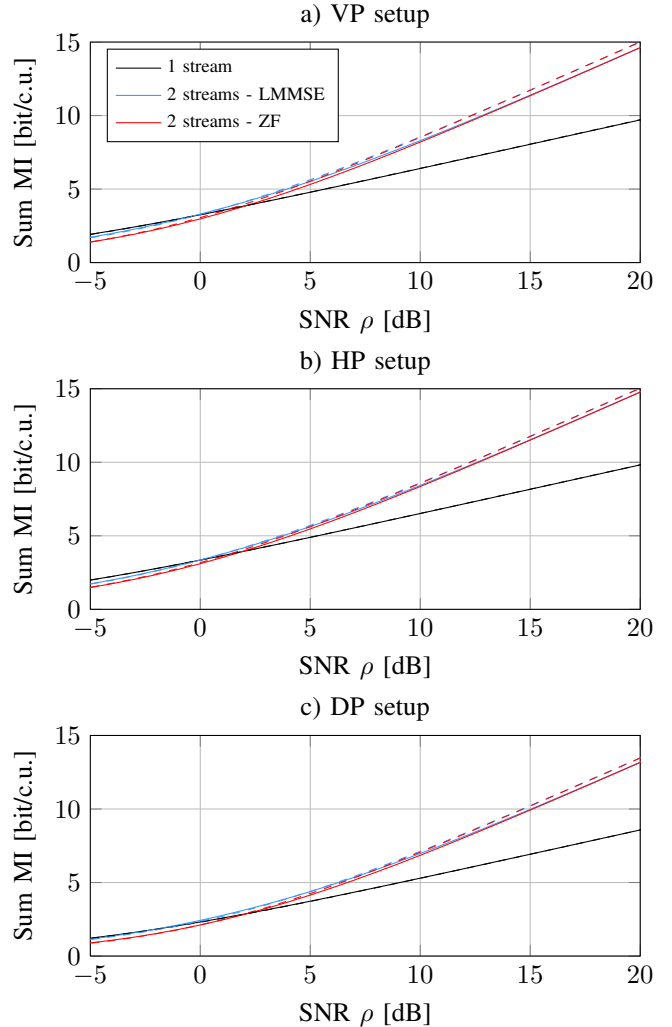


Fig. 1. Sum MI vs. SNR on the low- K -factor link 1 (dashed lines denote the corresponding approximation).

2.53 GHz. The results reveal that the activation of a second transmitted stream is only beneficial when the SNR crosses a threshold; for the SP setups, this threshold increases with the K -factors, whereas, for the DP setup, it remains at low SNR. We showed that DP antenna setups are only valuable when a Ricean component is present and a certain SNR is attained. Furthermore, we observed that the degradation of the sum MI with two streams due to the use of an LMMSE RX is rather small, especially for the DP setup.

APPENDIX A

HESSIAN MATRIX FOR THE LMMSE RECEIVER

We derive the Hessian matrix required for the second-order multivariate Taylor series expansion of the sum MI with the LMMSE RX. We consider $N_{\text{st}} \geq 2$ since the case $N_{\text{st}} = 1$ is trivial. In this appendix, we use results from [19], and we drop the time argument to simplify notation. First, we consider the function

$$f_k(\mathbf{A}) = \ln \left(\frac{1}{[\mathbf{A}^{-1}]_{k,k}} \right) \stackrel{(a)}{=} \ln \left(\frac{\det \mathbf{A}}{\det \mathbf{A}^{kk}} \right) \quad (7)$$

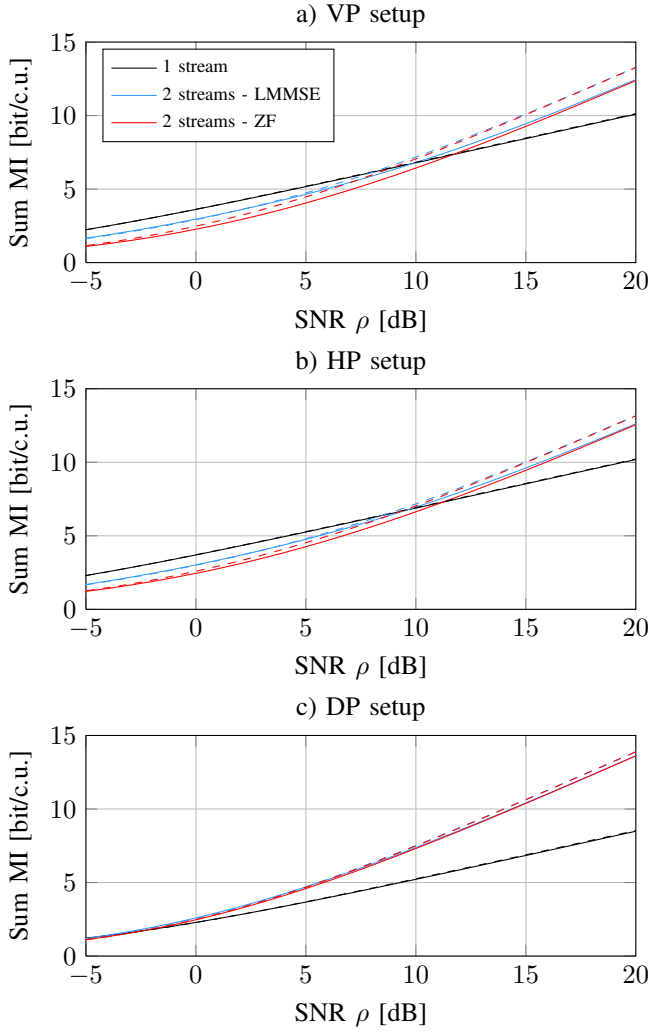


Fig. 2. Sum MI vs. SNR on the medium- K -factor link 2 (dashed lines denote the corresponding approximation).

with $\mathbf{A} = \mathbf{I}_{N_{\text{st}}} + \rho \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} / N_{\text{st}}$ and \mathbf{A}^{kk} being equal to \mathbf{A} with the k th row and k th column removed. In (a), we used the relation [20, Section 0.8.2]

$$[\mathbf{A}^{-1}]_{k,k} = \frac{\det \mathbf{A}^{kk}}{\det \mathbf{A}}. \quad (8)$$

The differential of $f_k(\mathbf{A})$ can be written as

$$\begin{aligned} df_k(\mathbf{A}) &= \text{tr} \{ \mathbf{A}^{-1} d\mathbf{A} \} - \text{tr} \{ \mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T d\mathbf{A} \} \\ &= \text{tr} \{ (\mathbf{A}^{-1} - \mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T) d\mathbf{A} \} \end{aligned} \quad (9)$$

where we denote by \mathbf{M}_k the identity matrix $\mathbf{I}_{N_{\text{st}}}$ with the k th column removed, thus $\mathbf{A}^{kk} = \mathbf{M}_k^T \mathbf{A} \mathbf{M}_k$ follows. The derivative of $f_k(\mathbf{A})$ with respect to \mathbf{A} follows as

$$\frac{\partial f_k(\mathbf{A})}{\partial \mathbf{A}} = \mathbf{A}^{-T} - (\mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T)^T. \quad (10)$$

and

$$\mathcal{D}_{\mathbf{A}} f_k(\mathbf{A}) = \left(\text{vec} \left\{ \frac{\partial f_k(\mathbf{A})}{\partial \mathbf{A}} \right\} \right)^T \quad (11)$$

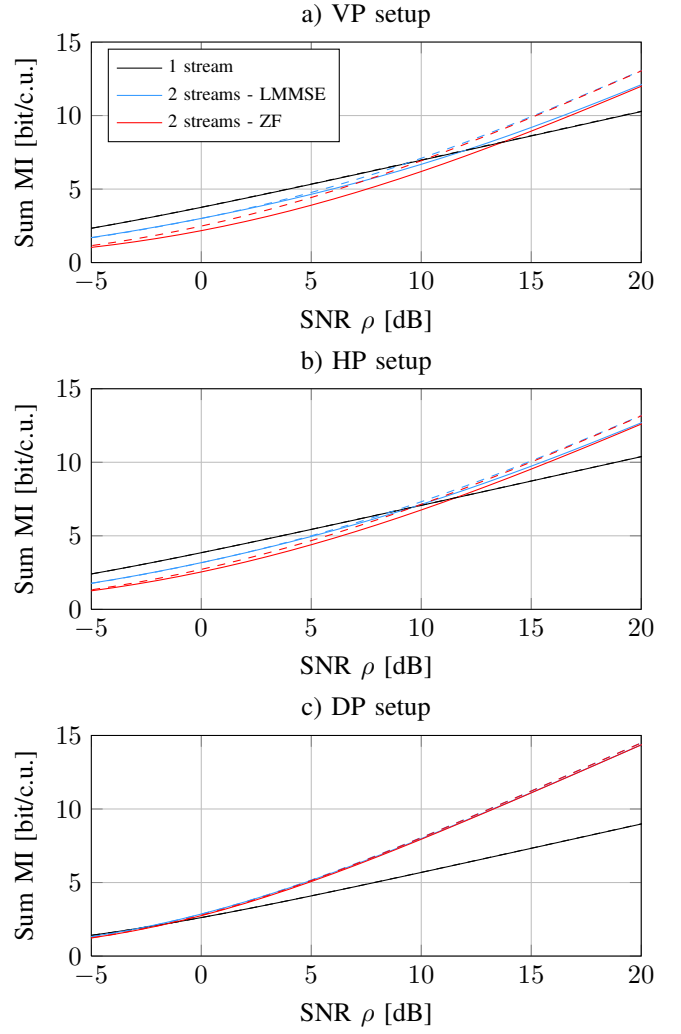


Fig. 3. Sum MI vs. SNR on the high- K -factor link 3 (dashed lines denote the corresponding approximation).

where we defined

$$\mathcal{D}_{\mathbf{A}} \mathbf{G} = \frac{\partial \text{vec} \{ \mathbf{G} \}}{\partial (\text{vec} \{ \mathbf{A} \})^T}. \quad (12)$$

Using results from [19], the derivative of the first term in (10) can be obtained as

$$\mathcal{D}_{\mathbf{A}} (\mathbf{A}^{-T}) = -\mathbf{K}_{N_{\text{st}}, N_{\text{st}}} (\mathbf{A}^{-T} \otimes \mathbf{A}^{-1}). \quad (13)$$

The derivative of the second term in (10) can be obtained by applying the chain rule [19]. First, we define $\mathbf{J} = \mathbf{M}_k \mathbf{B}$ with $\mathbf{B} = (\mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1})^T$. Then, using [21, Lemma 4.3.1]

$$\text{vec} \{ \mathbf{CDE} \} = (\mathbf{E}^T \otimes \mathbf{C}) \text{vec} \{ \mathbf{D} \} \quad (14)$$

for matrices \mathbf{C} , \mathbf{D} , and \mathbf{E} of appropriate sizes, the derivative of \mathbf{J} with respect to \mathbf{B} is obtained as $\mathcal{D}_{\mathbf{B}} \mathbf{J} = \mathbf{I}_{N_{\text{st}}} \otimes \mathbf{M}_k$. The rest of the analysis is performed in the same manner by

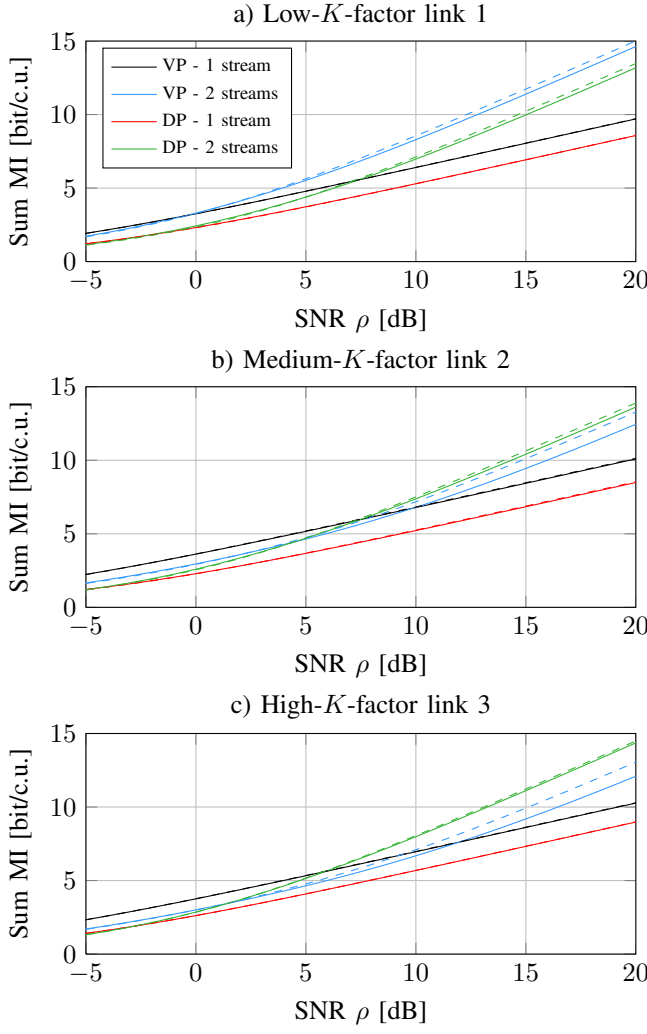


Fig. 4. Sum MI vs. SNR with the LMMSE receiver (dashed lines denote the corresponding approximation).

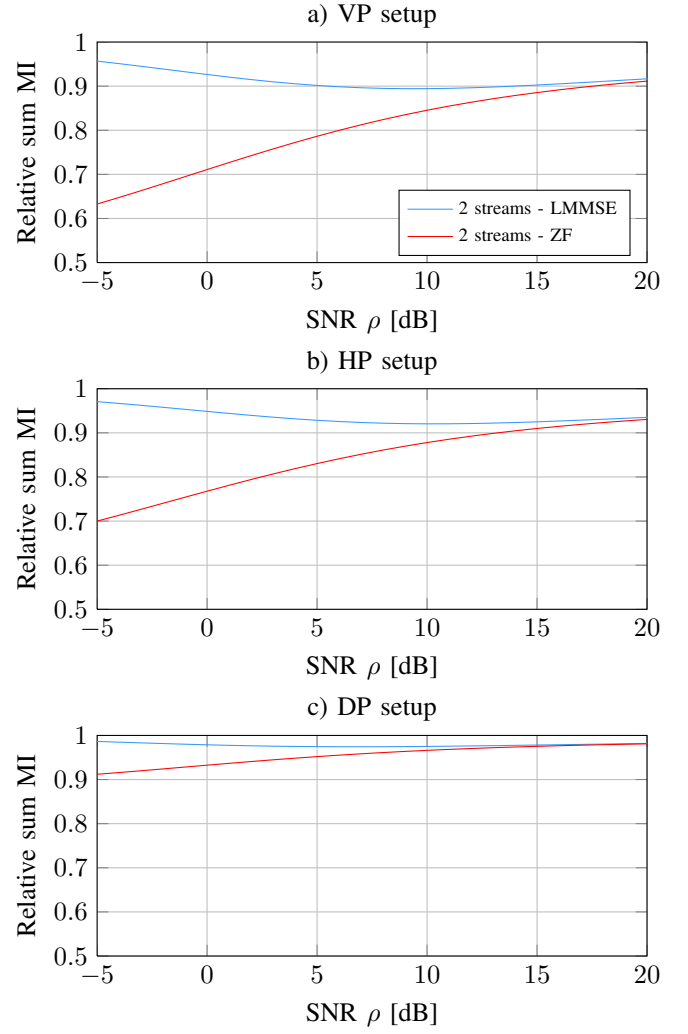


Fig. 5. Relative sum MI with respect to an optimal RX vs. SNR on the medium- K -factor link 2.

exploiting the chain rule recursively. This yields

$$\begin{aligned} \mathcal{D}_{\mathbf{A}}(\mathcal{D}_{\mathbf{A}}f_k(\mathbf{A}))^T &= -\mathbf{K}_{N_{\text{st}}, N_{\text{st}}}(\mathbf{A}^{-T} \otimes \mathbf{A}^{-1}) \\ &\quad + (\mathbf{I}_{N_{\text{st}}} \otimes \mathbf{M}_k)\mathbf{K}_{N_{\text{st}}, N_{\text{st}}-1}(\mathbf{I}_{N_{\text{st}}-1} \otimes \mathbf{M}_k) \\ &\quad \times ((\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-T} \otimes (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1}) \\ &\quad \times (\mathbf{M}_k^T \otimes \mathbf{M}_k^T). \end{aligned} \quad (15)$$

Finally, the Hessian matrix at $\mathbf{A} = \mathbb{E}\{\mathbf{A}\}$ is

$$\mathbf{Y}_{\text{LMMSE}, k} = \mathcal{D}_{\mathbf{A}}(\mathcal{D}_{\mathbf{A}}f_k(\mathbf{A}))^T \Big|_{\mathbf{A}=\mathbb{E}\{\mathbf{A}\}} \quad (16)$$

with $\mathbb{E}\{\mathbf{A}\} = \mathbf{I}_{N_{\text{st}}} + \rho \mathbf{F}^H \mathbf{R}_{\text{TX}}^* \mathbf{F} / N_{\text{st}}$.

APPENDIX B HESSIAN MATRIX FOR THE ZF RECEIVER

We derive the Hessian matrix required for the second-order multivariate Taylor series expansion of the sum MI with the ZF RX. We consider the non-trivial case $N_{\text{st}} \geq 2$, and we use results from [19]. The time argument is dropped for the

remainder of this appendix. First, we define the function

$$f_k(\mathbf{A}) = \ln \left(1 + \rho \frac{1}{[\mathbf{A}^{-1}]_{k,k}} \right) \stackrel{(a)}{=} \ln(1 + \rho w) \quad (17)$$

with $\mathbf{A} = \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} / N_{\text{st}}$, $w = \det \mathbf{A} / \det \mathbf{A}^{kk}$, and $\mathbf{A}^{kk} = \mathbf{M}_k^T \mathbf{A} \mathbf{M}_k$. In (a), we used (8). Then, we calculate the differential of $f_k(\mathbf{A})$:

$$\begin{aligned} df_k(\mathbf{A}) &= \frac{\partial}{\partial w} (\ln(1 + \rho w)) dw \\ &= \frac{\rho}{1 + \rho w} \left(\frac{\det \mathbf{A} \operatorname{tr}\{\mathbf{A}^{-1} d\mathbf{A}\} \det \mathbf{A}^{kk}}{(\det \mathbf{A}^{kk})^2} \right. \\ &\quad \left. - \frac{\det \mathbf{A} \det \mathbf{A}^{kk} \operatorname{tr}\{(\mathbf{A}^{kk})^{-1} d\mathbf{A}^{kk}\}}{(\det \mathbf{A}^{kk})^2} \right) \\ &= \frac{\rho \det \mathbf{A}}{1 + \rho \frac{\det \mathbf{A}}{\det \mathbf{A}^{kk}}} \left(\frac{\operatorname{tr}\{\mathbf{A}^{-1} d\mathbf{A}\}}{\det \mathbf{A}^{kk}} \right. \\ &\quad \left. - \frac{\operatorname{tr}\{\mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T d\mathbf{A}\}}{\det \mathbf{A}^{kk}} \right) \end{aligned}$$

$$= \frac{\rho \det \mathbf{A}}{\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A}} \times \text{tr} \left\{ (\mathbf{A}^{-1} - \mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T) d\mathbf{A} \right\}. \quad (18)$$

We can write the derivative of $f_k(\mathbf{A})$ with respect to \mathbf{A} as

$$\frac{\partial f_k(\mathbf{A})}{\partial \mathbf{A}} = w_{0,k} \mathbf{W}_{1,k} \quad (19)$$

with

$$w_{0,k} = \frac{\rho \det \mathbf{A}}{\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A}} \quad (20)$$

$$\mathbf{W}_{1,k} = (\mathbf{A}^{-1} - \mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T)^T. \quad (21)$$

We now calculate the derivative of $\frac{\partial f_k(\mathbf{A})}{\partial \mathbf{A}}$ with respect to \mathbf{A} . Consider first the differential

$$\begin{aligned} d \text{vec} \left\{ \frac{\partial f_k(\mathbf{A})}{\partial \mathbf{A}} \right\} &= w_{0,k} d \text{vec} \{ \mathbf{W}_{1,k} \} + \text{vec} \{ \mathbf{W}_{1,k} \} dw_{0,k} \\ &= w_{0,k} (\mathcal{D}_{\mathbf{A}} \mathbf{W}_{1,k}) d \text{vec} \{ \mathbf{A} \} \\ &\quad + \text{vec} \{ \mathbf{W}_{1,k} \} (\mathcal{D}_{\mathbf{A}} w_{0,k}) d \text{vec} \{ \mathbf{A} \}. \quad (22) \end{aligned}$$

Then, the derivative of (22) with respect to \mathbf{A} is

$$\mathcal{D}_{\mathbf{A}} (\mathcal{D}_{\mathbf{A}} f_k(\mathbf{A}))^T = w_{0,k} \mathcal{D}_{\mathbf{A}} \mathbf{W}_{1,k} + \text{vec} \{ \mathbf{W}_{1,k} \} \mathcal{D}_{\mathbf{A}} w_{0,k}. \quad (23)$$

The derivative of $\mathbf{W}_{1,k}$ with respect to \mathbf{A} can be calculated as in the case of the LMMSE RX:

$$\begin{aligned} \mathcal{D}_{\mathbf{A}} \mathbf{W}_{1,k} &= -\mathbf{K}_{N_{\text{st}}, N_{\text{st}}} (\mathbf{A}^{-T} \otimes \mathbf{A}^{-1}) \\ &\quad + (\mathbf{I}_{N_{\text{st}}} \otimes \mathbf{M}_k) \mathbf{K}_{N_{\text{st}}, N_{\text{st}}-1} (\mathbf{I}_{N_{\text{st}}-1} \otimes \mathbf{M}_k) \\ &\quad \times ((\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-T} \otimes (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1}) \\ &\quad \times (\mathbf{M}_k^T \otimes \mathbf{M}_k^T). \quad (24) \end{aligned}$$

The differential of $w_{0,k}$, can be calculated as

$$\begin{aligned} dw_{0,k} &= \frac{\rho \det \mathbf{A} \text{tr} \{ \mathbf{A}^{-1} d\mathbf{A} \} (\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A})}{(\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A})^2} \\ &\quad - \frac{\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) \text{tr} \{ \mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T d\mathbf{A} \}}{(\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A})^2} \\ &\quad \times \rho \det \mathbf{A} - \frac{(\rho \det \mathbf{A})^2 \text{tr} \{ \mathbf{A}^{-1} d\mathbf{A} \}}{(\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A})^2} \\ &= \frac{\rho \det \mathbf{A} \det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)}{(\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A})^2} \\ &\quad \times \text{tr} \left\{ (\mathbf{A}^{-1} - \mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T) d\mathbf{A} \right\}. \quad (25) \end{aligned}$$

The derivative of $w_{0,k}$ with respect to \mathbf{A} follows as

$$\begin{aligned} \frac{\partial w_{0,k}}{\partial \mathbf{A}} &= \frac{\rho \det \mathbf{A} \det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)}{(\det(\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k) + \rho \det \mathbf{A})^2} \\ &\quad \times (\mathbf{A}^{-1} - \mathbf{M}_k (\mathbf{M}_k^T \mathbf{A} \mathbf{M}_k)^{-1} \mathbf{M}_k^T)^T \quad (26) \end{aligned}$$

and we note that

$$\mathcal{D}_{\mathbf{A}} w_{0,k} = \left(\text{vec} \left\{ \frac{\partial w_{0,k}}{\partial \mathbf{A}} \right\} \right)^T. \quad (27)$$

Using (23), the Hessian matrix at $\mathbf{A} = \mathbf{E} \{ \mathbf{A} \}$, i.e.,

$$\mathbf{Y}_{\text{ZF},k} = \mathcal{D}_{\mathbf{A}} (\mathcal{D}_{\mathbf{A}} f_k(\mathbf{A}))^T \Big|_{\mathbf{A}=\mathbf{E}\{\mathbf{A}\}} \quad (28)$$

with $\mathbf{E} \{ \mathbf{A} \} = \mathbf{F}^H \mathbf{R}_{\text{TX}}^* \mathbf{F} / N_{\text{st}}$ is finally obtained.

REFERENCES

- [1] M. R. Andrews, P. P. Mitra, and R. deCarvalho, "Tripling the capacity of wireless communications using electromagnetic polarization," *Nature*, vol. 409, pp. 316–318, Jan. 2001.
- [2] A. S. Y. Poon and D. N. C. Tse, "Degree-of-freedom gain from using polarimetric antenna elements," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5695–5709, Sep. 2011.
- [3] C. Oestges, B. Clerckx, M. Guillaud, and M. Debbah, "Dual-polarized wireless communications: From propagation models to system performance evaluation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 4019–4031, Oct. 2008.
- [4] M. Coldrey, "Modeling and capacity of polarized MIMO channels," in *Proc. 67th IEEE Veh. Technol. Conf. (VTC)*, Singapore, May 2008, pp. 440–444.
- [5] V. Erceg, H. Sampath, and S. Catreux-Erceg, "Dual-polarization versus single-polarization MIMO channel measurement results and modeling," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 28–33, Jan. 2006.
- [6] P. Kyritsi, D. C. Cox, R. A. Valenzuela, and P. W. Wolniansky, "Effect of antenna polarization on the capacity of a multiple element system in an indoor environment," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1227–1239, Aug. 2002.
- [7] V. R. Anreddy and M. A. Ingram, "Capacity of measured Ricean and Rayleigh indoor MIMO channels at 2.4 GHz with polarization and spatial diversity," in *Proc. IEEE Conf. Wireless Commun. and Netw. (WCNC)*, Las Vegas, NV, USA, Apr. 2006, pp. 946–951.
- [8] V. Raghavan, A. M. Sayeed, and V. V. Veeravalli, "Semiunitary precoding for spatially correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1284–1298, Mar. 2011.
- [9] M. Herdin, N. Czink, H. Özcelik, and E. Bonek, "Correlation matrix distance, a meaningful measure for evaluation of non-stationary MIMO channels," in *Proc. 61st IEEE Veh. Technol. Conf. (VTC)*, Stockholm, Sweden, Jun. 2005, pp. 136–140.
- [10] M. Herdin, "Non-stationary indoor MIMO radio channels," Ph.D. dissertation, Vienna Univ. of Technology, Vienna, Austria, 2004.
- [11] A. Ispas, C. Schneider, G. Ascheid, and R. Thomä, "On non-stationary urban macrocell channels in a cooperative downlink beamforming scenario," in *Proc. 72nd IEEE Veh. Technol. Conf. (VTC)*, Ottawa, Canada, Sep. 2010.
- [12] V. Raghavan and A. M. Sayeed, "Impact of spatial correlation on statistical precoding in MIMO channels with linear receivers," in *Proc. 44th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2006, pp. 244–253.
- [13] D. A. Gore, R. W. Heath, Jr., and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov. 2002.
- [14] M. R. McKay, I. B. Collings, and A. M. Tulino, "Achievable sum rate of MIMO MMSE receivers: A general analytic framework," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 396–410, Jan. 2010.
- [15] A. Hjørungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge, UK: Cambridge Univ. Press, 2011.
- [16] A. Ispas, C. Schneider, G. Ascheid, and R. Thomä, "Analysis of local quasi-stationarity regions in an urban macrocell scenario," in *Proc. 71st IEEE Veh. Technol. Conf. (VTC)*, Taipei, Taiwan, May 2010.
- [17] V. Erceg, P. Soma, D. S. Baum, and A. J. Paulraj, "Capacity obtained from multiple-input multiple-output channel measurements in fixed wireless environments at 2.5 GHz," in *Proc. IEEE Int. Conf. Commun. (ICC)*, New York, NY, USA, Apr. 2002, pp. 396–400.
- [18] A. Ispas, J. Hölscher, X. Gong, C. Schneider, G. Ascheid, and R. Thomä, "Modeling and performance evaluation for mobile Ricean MIMO channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Ottawa, Canada, Jun. 2012.
- [19] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.
- [20] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1990.
- [21] —, *Topics in Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1994.