in Fig. 6. It can be seen that the performance of this system varies greatly with different SR link quality. Furthermore, the performance of the system approaches the CF bounds. The gap varies within 1.4 and 1.8 dB with different SR links.

V. Conclusion

In this paper, we have proposed a DMIPQ for the half-duplex ID-MARC when the relay cannot recover the initial bits correctly. The proposed DMIPQ can remove the noise of the received signal while preserving the useful information. To find the quantization intervals of the proposed scheme, an iterative suboptimal algorithm was provided. To further reduce the computational complexity, an analytical approximation of the DMIPQ was also proposed. It was shown that the approximation is equivalent to the MMSE Lloyd-max quantizer in the ID-MARC. Simulation results showed that the proposed DMIPQ has a great improvement on the BER performance. Due to the significant performance improvement, the proposed DMIPQ may be an attractive choice for the future cellular network architecture, such as the cloud-radio access network [19], [20], etc. Note that the proposed DMIPQ is a scalar quantizer, and the vector quantizer usually performs better. Therefore, the design of vector DMIPQ can be an interesting future work.

Acknowledgment

The authors would like to thank Prof. Y. L. Guan and Prof. C. Yuen for fruitful discussions.

References

[1] G. Kramer and A. J. van Wijngaarden, "On the white Gaussian multiple-access relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, p. 40.

[2] C. Hausl and P. Dupraz, "Joint network-channel coding for the multiple-access relay channel," in *Proc. IEEE Conf. Sensor Ad Hoc Commun. Netw.*, Sep. 2006, pp. 817–822.

[3] Y. Li, G. Song, and L. Wang, "Design of joint network-low density parity check codes based on the EXIT charts," *IEEE Commun. Lett.*, vol. 13, no. 8, pp. 600–602, Aug. 2009.

[4] J. Li, J. Yuan, R. Malaney, M. H. Azmi, and M. Xiao, "Network Coded LDPC Code Design for a Multi-Source Relaying System," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1538–1551, May 2011.

[5] Y. Zhang and Z. Zhang, "Joint network-channel coding with rateless code over multiple access relay system," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 320–332, Jan. 2013.

[6] G. Zeitler, G. Bauch, and J. Widmer, "Quantize-and-forward schemes for the orthogonal multiple-access relay channel," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1148–1158, Apr. 2012.

[7] A. Winkelbauer, N. Goertz, and G. Matz, "Compress-and-forward in the multiple-access relay channel: With or without network coding?" in *Proc. 7th ISTC*, Gothenburg, Sweden, Aug. 2012, pp. 817–822.

[8] L. Ping, L. Liu, K. Y. Wu, and W. K. Leung, "Interleave-division multiple-access (IDMA) communications," in *Proc. Int. Symp. Turbo Codes Related Topics*, Brest, France, Sep. 2003, pp. 173–180.

[9] L. Ping, L. Liu, K. Y. Wu, and W. K. Leung, "Approaching the capacity of multiple access channels using interleaved low-rate codes," *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 4–6, Jan. 2004.

[10] W. Rave, "Quantization of log-likelihood ratios to maximize mutual information," *IEEE Signal Process. Lett.*, vol. 16, no. 4, pp. 283–286, Apr. 2009.

[11] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 192–137, Mar. 1982.

[12] A. Chakrabarti, A. de Baynast, A. Sabharwal, and B. Aazhang, "Half-duplex estimate-and-forward relaying: Bounds and code design," in *Proc. ISIT*, Jul. 2006, pp. 1239–1243.

[13] A. Chakrabarti, A. Sabharwal, and B. Aazhang, "Practical quantizer design for half-duplex estimate-and-forward relaying," *IEEE Trans. Commun.*, vol. 59, no. 1, pp. 74–83, Jan. 2011.

[14] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 2020–2040, Jun. 2005.

[15] D. G. Brennan, "Linear diversity combining techniques," *Proc. IRE*, vol. 47, no. 6, pp. 1075–1102, Jun. 1959.

[16] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York, NY, USA: McGraw-Hill, 1984.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. 2nd ed. New York, NY, USA: Wiley, 2006.

[18] L. Ping, W. K. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3213–3324, Dec. 2003.

[19] X. Wang, "C-RAN: The road toward green RAN," *China Commun. J.*, Jun. 2010.

[20] "C-RAN: The road towards green RAN," China Mobile Res. Inst., Beijing, China, Ver. 2.5, White Paper, Oct. 2012.

Achieving Optimality in Robust Joint Optimization of Linear Transceiver Design

Hongying Tang, Wen Chen, *Senior Member, IEEE*, and Jun Li, *Member, IEEE*

***Abstract*—This paper presents new results on linear transceiver designs in a multiple-input–multiple-output (MIMO) link. By considering the minimal total mean-square error (MSE) criterion, we prove that the robust optimal linear transceiver design has a channel-diagonalizing structure, which verifies the conjecture in the previous work. Based on this property, the original design problem can be transformed into a scalar problem, whose global optimal solution is first obtained in this work. Simulation results show the performance advantages of our solution over the existing schemes.**

***Index Terms*—Convex optimization, linear transceiver design, mean-square error (MSE), robust design.**

I. Introduction

The multiple-input–multiple-output (MIMO) technique has attracted considerable interest from both academia and industry in recent years. By exploiting the multiplexing and diversity property, it can significantly improve the spectral efficiency and link the reliability of the system [2]. In the literature, transceiver designs in MIMO systems have been extensively studied in [1]–[13]. One approach of the designs is to allow a nonlinear process at the transmitter or the receiver, such

Manuscript received May 7, 2014; revised December 16, 2014; accepted March 5, 2015. Date of publication March 11, 2015; date of current version March 10, 2016. This work was supported by the National 973 Project under Grant 2012CB316106, by the National Science Foundation of China under Grant 61328101, by the Science and Technology Commission of Shanghai Municipality (STCSM) Science and Technology Innovation Program under Grant 13510711200, and by the Southeast University (SEU) National Key Laboratory on Mobile Communications under Grant 2013D11. The review of this paper was coordinated by Prof. S. Muhaidat.

H. Tang and W. Chen are with the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200241, China, and also with the School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China (e-mail: lojordan@sjtu.edu.cn; wenchen@sjtu.edu.cn).

J. Li is with the School of Electrical and Information Engineering, University of Sydney, Sydney, N.S.W. 2006, Australia (e-mail: jun.li1@sydney.edu.au).

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/TVT.2015.2411855

as the successive interference cancelation receiver design discussed in [2], or the maximum-likelihood detector investigated in [3] and [4].

As an alternative approach, the linear transceiver design, which only allows linear matrix multiplication of the signal, is more preferable in a practical system, due to low implementation complexity, and is the focus of this paper. In [5], the joint optimal linear transceiver design problem was addressed, and a closed-form solution was derived. Their result was generalized into the multicarrier MIMO system in [6] by developing a unified optimization framework. The aforementioned works in [5] and [6] enjoy a common favorable feature that the transceiver processing matrix parallelized the original channel and allocated power to each data stream. In light of the optimality of this *channel-diagonalizing* structure in the perfect channel-state information (CSI) case, one may wonder whether the same property holds for the robust design in the imperfect CSI case.

Robust design, which aims to reduce the sensitivity of the imperfect CSI to the system performance, has attracted much attention [1], [7]–[9]. Generally, there are two widely used CSI uncertainty models in the literature: the stochastic model and the deterministic model. For the statistical CSI uncertainty model, where the distribution of CSI uncertainties is assumed to be known, this channel-diagonalizing structure has been well established in MIMO channels [7], [8]. However, for the deterministic CSI uncertainty model, which assumes that the instantaneous value of CSI error is norm bounded, this problem remains unsolved, and only some restricted results were obtained in [1] and [9]. The authors in [9] proposed a semirobust scheme, by optimizing only the transmit processing matrix with some fixed equalizer. Obviously, this scheme cannot fully exploit the performance gain by the equalizer, since the fixed equalizer may not be optimal. Later in [1], the authors considered joint linear transceiver design and showed a superior performance over [9]. By imposing certain structural constraints on the processing matrix at the transmitter or receiver side, they observed the favorable channel-diagonalizing structure. Then, they transformed the original problem into the issues of power loading among each data stream, which were further solved by the alternation optimization method. However, two problems in [1] were left unsolved:

$\mathcal{Q}1$) *Joint optimal structure*: Without any additional structural restriction, is this channel-diagonalizing structure joint optimal?

$\mathcal{Q}2$) *Global optimal solution*: If it is, does the alternating-optimization-based method converge to the global optimal solution?

In this paper, we will answer the aforementioned two questions raised in [1]. Without assuming any specific structure for the linear transmitter-equalizer matrix, we show that the optimal design actually admits a channel-diagonalizing structure. Based on this property, the original problem reduces to a scalar convex problem, whose optimal solution can thus be efficiently obtained. Simulation results in Section V show the superior performance of our solution over that in [1].

Notations: $[\cdot]^H$ denotes the conjugate transpose of a matrix or a vector. $\mathbf{I}$ and $\mathbf{0}$ denote the identity and zero matrix, respectively. $\mathbb{R}^N$ and $\mathbb{C}^N$ denote the N-dimensional real field and complex field, respectively. $||\cdot||_2$ and $\|\cdot\|_F$ denote the Frobenius norm of a vector and a matrix, respectively. We will use boldface lowercase letters to denote column vectors and boldface uppercase letters to denote matrices. The positive semidefinite matrix $\mathbf{X}$ is denoted by $\mathbf{X}\succeq 0$. $\text{diag}\{\mathbf{x}_1,\ldots,\mathbf{x}_R\}$ denotes the diagonal concatenation of block matrices $\mathbf{x}_1,\ldots,\mathbf{x}_R$. The $\text{tr}(\cdot)$ is the trace of a matrix. $\text{vec}(\mathbf{X})$ stacks the columns of matrix $\mathbf{X}$ into a vector. $\otimes$ denotes the Kronecker product. $\Re\{\cdot\}$ denotes the real part of a complex number. $\lambda_{\max}(\cdot)$ is the maximum eigenvalue of a matrix.

II. Problem Statement

We consider a MIMO communication system equipped with N transmit antennas at the source and M receive antennas at the destination. The symbol vector $\mathbf{s}\in\mathbb{C}^L$ is linearly precoded by a source precoding matrix $\mathbf{F}\in\mathbb{C}^{N\times L}$, through the MIMO channel $\mathbf{H}\in\mathbb{C}^{M\times N}$, and then received by the destination. We assume that $E\{\mathbf{s}\mathbf{s}^H\}=\mathbf{I}$ without loss of generality. Generally, the transmitter imposes a power constraint on the precoding matrix $\mathbf{F}$ as $\text{tr}(\mathbf{F}\mathbf{F}^H)\leq P$. A linear equalizer $\mathbf{G}\in\mathbb{C}^{L\times M}$ is usually applied on the received signal to obtain the estimated symbol vector $\hat{\mathbf{s}}$ as

$$\hat{\mathbf{s}}=\mathbf{GHFs}+\mathbf{Gn}$$

where $\mathbf{n}\in\mathbb{C}^M$ is the additive white Gaussian noise observed at the destination with variance $\sigma_n^2\mathbf{I}$. Then, the MSE between $\hat{\mathbf{s}}$ and $\mathbf{s}$ is given by

$$\text{MSE}\triangleq\mathcal{E}\left\{\|\hat{\mathbf{s}}-\mathbf{s}\|^2\right\}=\|\mathbf{GHF}-\mathbf{I}\|_F^2+\sigma_n^2\|\mathbf{G}\|_F^2.$$

We assume that $L\leq\text{rank}(\mathbf{H})$, since the number of degrees of freedom is upper bounded by $L\leq\text{rank}(\mathbf{H})=\min\{M,N\}$.

In a practical wireless communication scenario, perfect CSI is usually difficult to obtain. With only imperfect CSI, the system performance will be deteriorated. This motivates us to investigate the robust design taking the CSI errors into account. To characterize the mismatched CSI, we adopt a common deterministic imperfect CSI model [1], [9] and write the channel matrix as

$$\mathbf{H}=\tilde{\mathbf{H}}+\mathbf{E}\tag{1}$$

where $\tilde{\mathbf{H}}$ is the estimated channel matrix, and $\mathbf{E}$ is the corresponding CSI error matrix satisfying $\|\mathbf{E}\|_F\leq\varepsilon$ for some $\varepsilon\geq 0$. As in [1] and [9], we assume that only $\tilde{\mathbf{H}}$ and ε are available at both ends.

By taking the imperfect CSI model (1) into account, the robust transmitter-equalizer design is given by the solution of the following min–max problem:

$$\begin{aligned}\min_{\mathbf{G},\mathbf{F}}\max_{\|\mathbf{E}\|_F\leq\varepsilon}\quad&\|\mathbf{G}(\tilde{\mathbf{H}}+\mathbf{E})\mathbf{F}-\mathbf{I}\|_F^2+\sigma_n^2\|\mathbf{G}\|_F^2\\ \text{s.t.}\quad&\text{tr}(\mathbf{F}\mathbf{F}^H)\leq P.\end{aligned}\tag{2}$$

III. Robust Joint Optimal Structure of $\mathbf{F}$ and $\mathbf{G}$

Here, we will determine the joint optimal structure of $\mathbf{F}$ and $\mathbf{G}$ in problem (2), showing that they diagonalize the MIMO channel into eigensubchannels. We also figure out that the worst-case CSI uncertainty $\mathbf{E}$ has the similar singular value decomposition (SVD) structure as the nominal channel $\tilde{\mathbf{H}}$, which simplifies problem (2) into a scalar problem, as we will shown in Section IV.

Denote the SVD structure of $\mathbf{F}$ and $\mathbf{G}$ by $\mathbf{F}=\mathbf{U}_f\mathbf{\Sigma}_f\mathbf{V}_f^H$ and $\mathbf{G}=\mathbf{U}_g\mathbf{\Sigma}_g\mathbf{V}_g^H$, respectively, where $\mathbf{U}_f$, $\mathbf{V}_f$, $\mathbf{U}_g$, and $\mathbf{V}_g$ are unitary matrices. The matrices $\mathbf{\Sigma}_f$ and $\mathbf{\Sigma}_g$ can be written as

$$\mathbf{\Sigma}_f=[\hat{\mathbf{\Sigma}}_f,\mathbf{0}]^T,\quad\Sigma_g=[\hat{\mathbf{\Sigma}}_g,\mathbf{0}]$$

where $\hat{\mathbf{\Sigma}}_f\triangleq\text{diag}\{f_1,\ldots,f_L\}$ and $\hat{\mathbf{\Sigma}}_g\triangleq\text{diag}\{g_1,\ldots,g_L\}$ are real diagonal matrices. Denote the nominal channel $\tilde{\mathbf{H}}$ by $\tilde{\mathbf{H}}=\mathbf{U}_h\mathbf{\Sigma}_h\mathbf{V}_h^H$, and let $\hat{\mathbf{\Sigma}}_h$ be the $L\times L$ diagonal matrix containing the largest L singular values $\gamma_1\geq\cdots\geq\gamma_L$. Then, the following theorem determines the optimal structure of $\mathbf{F}$ and $\mathbf{G}$.

Theorem 1: The robust optimal $\mathbf{F}$ and $\mathbf{G}$ problem (2) can be expressed in the following structure:

$$\mathbf{F}=\mathbf{V}_h\mathbf{\Sigma}_f\tag{3}$$

$$\mathbf{G}=\mathbf{\Sigma}_g\mathbf{U}_h^H.\tag{4}$$

Meanwhile, the corresponding worst-case channel uncertainty is given by $\mathbf{E} = \mathbf{U}_h \boldsymbol{\Delta}_D \mathbf{V}_h^H$, with $\boldsymbol{\Delta}_D = \text{diag}\{\hat{\boldsymbol{\Delta}}_D \mathbf{0}\}$, and $\hat{\boldsymbol{\Delta}}_D \in \mathbb{R}^{L\times L}$ being diagonal.

Proof: We write the first additive term of the objective function of (2) as

$$\left\|\mathbf{G}(\hat{\mathbf{H}} + \mathbf{E})\mathbf{F} - \mathbf{I}\right\|_F^2 \overset{(a)}{=} \|\mathbf{G}\mathbf{U}_h(\boldsymbol{\Sigma}_h + \boldsymbol{\Delta})\mathbf{V}_h^H\mathbf{F} - \mathbf{I}\|_F^2 \overset{(b)}{=} \|\mathbf{G}'(\boldsymbol{\Sigma}_h + \boldsymbol{\Delta})\mathbf{F}' - \mathbf{I}\|_F^2$$

where in (a) we have defined $\boldsymbol{\Delta} \triangleq \mathbf{U}_h^H \mathbf{E}\mathbf{V}_h$. By the unitary-invariant property of $\|\cdot\|_F$, we can see that $\boldsymbol{\Delta}$ still satisfies $\|\boldsymbol{\Delta}\|_F \le \varepsilon$. In (b), we have defined $\mathbf{G}' \triangleq \mathbf{G}\mathbf{U}_h$ and $\mathbf{F}' \triangleq \mathbf{V}_h^H\mathbf{F}$. Now, problem (2) can be rewritten as

$$\begin{aligned}\min_{\mathbf{G}',\mathbf{F}'} \max_{\|\boldsymbol{\Delta}\|_F\le\varepsilon} \quad & \text{MSE}(\mathbf{F}', \mathbf{G}', \boldsymbol{\Delta}) \\ & \triangleq \|\mathbf{G}'(\boldsymbol{\Sigma}_h + \boldsymbol{\Delta})\mathbf{F}' - \mathbf{I}\|_F^2 + \sigma_d^2\|\mathbf{G}'\|_F^2 \\ \text{s.t.} \quad & \text{tr}(\mathbf{F}'\mathbf{F}'^H) \le P \end{aligned} \tag{5}$$

which is an optimization problem with respect to $\mathbf{F}'$ and $\mathbf{G}'$. To proceed, we first discuss a particular case when $(\mathbf{F}'\mathbf{G}') = ([\hat{\boldsymbol{\Sigma}}_f, \mathbf{0}]^T, [\hat{\boldsymbol{\Sigma}}_g, \mathbf{0}])$. Then, problem (5) becomes

$$\begin{aligned}\min_{\hat{\boldsymbol{\Sigma}}_g,\hat{\boldsymbol{\Sigma}}_f} \max_{\|\hat{\boldsymbol{\Delta}}\|_F\le\varepsilon} \quad & \|\hat{\boldsymbol{\Sigma}}_g(\hat{\boldsymbol{\Sigma}}_h + \hat{\boldsymbol{\Delta}})\hat{\boldsymbol{\Sigma}}_f - \mathbf{I}\|_F^2 + \sigma_d^2\|\hat{\boldsymbol{\Sigma}}_g\|_F^2 \\ \text{s.t.} \quad & \text{tr}(\hat{\boldsymbol{\Sigma}}_f\hat{\boldsymbol{\Sigma}}_f^H) \le P \end{aligned} \tag{6}$$

where $\hat{\boldsymbol{\Delta}}$ is the upper left $L \times L$ submatrix of $\boldsymbol{\Delta}$. We will then show that there exists an optimal $\hat{\boldsymbol{\Delta}}$ in (6) that is diagonal.

After some matrix manipulations and noticing the fact that the maximization of a convex function is achieved on the boundary [9], the inner maximization of problem (6) can be transformed into the following problem:

$$\min_{\|\boldsymbol{\delta}\|=\varepsilon} \boldsymbol{\delta}^H(-\mathbf{B}^T \otimes \mathbf{C})\boldsymbol{\delta} - 2\Re\{\mathbf{d}^H\boldsymbol{\delta}\} \tag{7}$$

where $\boldsymbol{\delta} \triangleq \text{vec}(\hat{\boldsymbol{\Delta}})$, $\mathbf{C} \triangleq \hat{\boldsymbol{\Sigma}}_g\hat{\boldsymbol{\Sigma}}_g^H$, $\mathbf{B} \triangleq \hat{\boldsymbol{\Sigma}}_f\hat{\boldsymbol{\Sigma}}_f^H$, and $\mathbf{d} \triangleq \text{vec}(\hat{\boldsymbol{\Sigma}}_g^H(\hat{\boldsymbol{\Sigma}}_g\hat{\boldsymbol{\Sigma}}_h\hat{\boldsymbol{\Sigma}}_f - \mathbf{I})\hat{\boldsymbol{\Sigma}}_f^H)$.

By the result in [9], $\boldsymbol{\delta}$ is a global minimizer of (7) if and only if there exists an ω such that

$$(-\mathbf{B}^T \otimes \mathbf{C} + \omega\mathbf{I})\boldsymbol{\delta} = \mathbf{d}, \quad -\mathbf{B}^T \otimes \mathbf{C} + \omega\mathbf{I} \succeq \mathbf{0}, \quad \|\boldsymbol{\delta}\| = \varepsilon$$

which is equivalent to

$$\omega\hat{\boldsymbol{\Delta}} - \mathbf{C}\hat{\boldsymbol{\Delta}}\mathbf{B} = \hat{\boldsymbol{\Sigma}}_g^H(\hat{\boldsymbol{\Sigma}}_g\hat{\boldsymbol{\Sigma}}_h\hat{\boldsymbol{\Sigma}}_f - \mathbf{I})\hat{\boldsymbol{\Sigma}}_f^H, \tag{8}$$

$$\text{tr}(\hat{\boldsymbol{\Delta}}\hat{\boldsymbol{\Delta}}^H) = \varepsilon^2, \tag{9}$$

$$\omega \ge \lambda_{\max}(\mathbf{B}^T \otimes \mathbf{C}). \tag{10}$$

Since both $\mathbf{C}$ and $\mathbf{B}$ are diagonal, (8)–(10) tell us that, for any given $\hat{\boldsymbol{\Sigma}}_f$ and $\hat{\boldsymbol{\Sigma}}_g$, there exists an optimal $\hat{\boldsymbol{\Delta}}$ that is diagonal. Denote the optimal solution of (6) by $(\hat{\boldsymbol{\Sigma}}_f^\sharp \hat{\boldsymbol{\Sigma}}_g^\sharp, \hat{\boldsymbol{\Delta}}_D^\sharp)$, with $\hat{\boldsymbol{\Delta}}_D^\sharp$ being diagonal. To facilitate the analysis, we further define $\mathbf{F}'^\sharp \triangleq [\hat{\boldsymbol{\Sigma}}_f^\sharp, \mathbf{0}]^T$, $\mathbf{G}'^\sharp \triangleq [\hat{\boldsymbol{\Sigma}}_g^\sharp, \mathbf{0}]$, and $\boldsymbol{\Delta}_D^\sharp \triangleq \text{diag}\{\hat{\boldsymbol{\Delta}}_D^\sharp, \mathbf{0}\}$. Then, by the aforementioned definitions and discussions, we have

$$\max_{\|\hat{\boldsymbol{\Delta}}\|_F\le\varepsilon} \text{MSE}(\mathbf{F}'^\sharp, \mathbf{G}'^\sharp, \boldsymbol{\Delta}) = \text{MSE}(\mathbf{F}'^\sharp, \mathbf{G}'^\sharp, \boldsymbol{\Delta}_D^\sharp). \tag{11}$$

Now, we discuss another particular situation when $\boldsymbol{\Delta} = \boldsymbol{\Delta}_D^\sharp$. Then, problem (5) becomes

$$\begin{aligned}\min_{\mathbf{F}',\mathbf{G}'} \quad & \|\mathbf{G}'(\boldsymbol{\Sigma}_h + \boldsymbol{\Delta}_D^\sharp)\mathbf{F}' - \mathbf{I}\|_F^2 + \sigma_d^2\|\mathbf{G}'\|_F^2 \\ \text{s.t.} \quad & \text{tr}(\mathbf{F}'\mathbf{F}'^H) \le P \end{aligned} \tag{12}$$

which has been discussed in [5], [7] and [10], and the optimal solution is given by

$$(\mathbf{F}', \mathbf{G}') = \left([\hat{\boldsymbol{\Sigma}}_f, \mathbf{0}]^T, [\hat{\boldsymbol{\Sigma}}_g, \mathbf{0}]\right). \tag{13}$$

Substituting (13) into problem (12), we have

$$\begin{aligned}\min_{\hat{\boldsymbol{\Sigma}}_f,\hat{\boldsymbol{\Sigma}}_g} \quad & \left\|\hat{\boldsymbol{\Sigma}}_g(\hat{\boldsymbol{\Sigma}}_h + \hat{\boldsymbol{\Delta}}_D^\sharp)\hat{\boldsymbol{\Sigma}}_f - \mathbf{I}\right\|_F^2 + \sigma_d^2\|\hat{\boldsymbol{\Sigma}}_g\|_F^2 \\ \text{s.t.} \quad & \text{tr}\left(\hat{\boldsymbol{\Sigma}}_f\hat{\boldsymbol{\Sigma}}_f^H\right) \le P. \end{aligned} \tag{14}$$

Remember that the optimal solution of (6) is denoted by $(\hat{\boldsymbol{\Sigma}}_f^\sharp, \hat{\boldsymbol{\Sigma}}_g^\sharp, \hat{\boldsymbol{\Delta}}_D^\sharp)$. Then, it is easy to know that the optimal solution of (14) is given by $(\hat{\boldsymbol{\Sigma}}_f^\sharp, \hat{\boldsymbol{\Sigma}}_g^\sharp)$, which means that, when the channel uncertainty is $\boldsymbol{\Delta}_D^\sharp = \text{diag}\{\hat{\boldsymbol{\Delta}}_D^\sharp, \mathbf{0}\}$, the optimal $(\mathbf{F}', \mathbf{G}')$ of problem (12) is given by $([\hat{\boldsymbol{\Sigma}}_f^\sharp, \mathbf{0}]^T, [\hat{\boldsymbol{\Sigma}}_g^\sharp, \mathbf{0}])$, or we have

$$\text{MSE}(\mathbf{F}', \mathbf{G}', \boldsymbol{\Delta}_D^\sharp) \ge \text{MSE}(\mathbf{F}'^\sharp, \mathbf{G}'^\sharp, \boldsymbol{\Delta}_D^\sharp). \tag{15}$$

We will next show that, from these two special cases given in (11) and (15), the joint optimal structure of $\mathbf{F}'$ and $\mathbf{G}'$ can be obtained. This technique has also been used in [11]–[13] and is detailed as follows:

$$\begin{aligned}\max_{\|\boldsymbol{\Delta}\|_F\le\varepsilon} \text{MSE}(\mathbf{F}', \mathbf{G}', \boldsymbol{\Delta}) & \overset{(a)}{\ge} \text{MSE}(\mathbf{F}', \mathbf{G}', \boldsymbol{\Delta}_D^\sharp) \\ & \overset{(b)}{\ge} \text{MSE}(\mathbf{F}'^\sharp, \mathbf{G}'^\sharp, \boldsymbol{\Delta}_D^\sharp) \overset{(c)}{=} \max_{\|\boldsymbol{\Delta}\|_F\le\varepsilon} \text{MSE}(\mathbf{F}'^\sharp, \mathbf{G}'^\sharp, \boldsymbol{\Delta}) \end{aligned} \tag{16}$$

where (a) is due to the fact that $\boldsymbol{\Delta}_D^\sharp$ is only a particular channel, (b) is due to (15), and (c) is due to (11). Inequality (16) shows that the optimal $\mathbf{F}'$ and $\mathbf{G}'$ must be given by $(\mathbf{F}'^\sharp, \mathbf{G}'^\sharp)$. The proof is completed. ■

Remark 1: Theorem 1 provides some interesting insights into the robust optimal transceiver design, showing that its optimal structure is given by $\mathbf{V}_f = \mathbf{U}_g$, $\mathbf{U}_f = \mathbf{V}_h$, and $\mathbf{V}_g = \mathbf{U}_h$, which diagonalizes the MIMO channel into eigensubchannels, and is consistent with the results under perfect and stochastic CSI assumptions. Therefore, the answer to question $\mathcal{Q}1$ is yes. Note that problem (2) was also considered in [1], where only partial results of Theorem 1 was obtained. That is, by assuming $\mathbf{V}_g = \mathbf{U}_h$, then $\mathbf{V}_f = \mathbf{U}_g$ and $\mathbf{U}_f = \mathbf{V}_h$ were proved to be optimal; on the other hand, given $\mathbf{U}_f = \mathbf{V}_h$, then $\mathbf{V}_f = \mathbf{U}_g$ and $\mathbf{V}_g = \mathbf{U}_h$ proved to be optimal.

IV. Robust Global Optimal Design Based on Scalar Optimization

Based on Theorem 1, we know that the optimal solution of problem (2) is determined by problem (6), where $\hat{\boldsymbol{\Delta}}$ is diagonal. Denote

$\hat{\boldsymbol{\Delta}} \triangleq \text{diag}\{x_1, \ldots, x_L\}$, $\mathbf{f} \triangleq [f_1, \ldots, f_L]^T$, and $\mathbf{g} \triangleq [g_1, \ldots, g_L]^T$; then, problem (6) is equivalent to

$$\begin{aligned} \min_{\mathbf{f},\mathbf{g}} \max_{\sum_{i=1}^L x_i^2 \le \varepsilon^2} \quad & \sum_{i=1}^L (f_i g_i(\gamma_i + x_i) - 1)^2 + \sigma_n^2 \sum_{i=1}^L g_i^2 \\ \text{s.t.} \quad & \sum_{i=1}^L f_i^2 \le P. \end{aligned} \tag{17}$$

Introducing a slack variable t, problem (17) can be converted to

$$\min_{\mathbf{f},\mathbf{g},t} \quad t + \sigma_n^2 \sum_{i=1}^L g_i^2 \tag{18a}$$
$$\text{s.t.} \quad \sum_{i=1}^L (f_i g_i(\gamma_i + x_i) - 1)^2 \le t, \sum_{i=1}^L x_i^2 \le \varepsilon^2 \tag{18b}$$
$$\sum_{i=1}^L f_i^2 \le P. \tag{18c}$$

Generally speaking, it is difficult to derive a closed-form solution of (18). Thus, we will solve it in the numerical results. Let $\boldsymbol{\eta} \triangleq [f_1 g_1 \gamma_1 - 1, \ldots, f_L g_L \gamma_L - 1]^T$, $\boldsymbol{\Gamma} \triangleq \text{diag}\{f_1 g_1, \ldots, f_L g_L\}$, and $\mathbf{x} \triangleq [x_1, \ldots, x_L]^T$. Constraint (18b) can be rewritten as

$$\|\boldsymbol{\eta} + \boldsymbol{\Gamma}\mathbf{x}\|_2^2 \le t, \quad \|\mathbf{x}\|_2 \le \varepsilon. \tag{19}$$

Following the similar lines in [1] and [9], one can transform (19) into

$$\begin{bmatrix} t-\mu & \boldsymbol{\eta}^H & \mathbf{0} \\ \boldsymbol{\eta} & \mathbf{I} & \varepsilon\boldsymbol{\Gamma} \\ \mathbf{0} & \varepsilon\boldsymbol{\Gamma}^H & \mu\mathbf{I} \end{bmatrix} \succeq \mathbf{0}, \quad \exists \mu \ge 0. \tag{20}$$

By applying Schur's complement [14], (20) is equivalent to the following constraint:

$$\begin{aligned} &\begin{bmatrix} t-\mu & \boldsymbol{\eta}^H \\ \boldsymbol{\eta} & \mathbf{I} \end{bmatrix} - \frac{1}{\mu} \begin{bmatrix} 0 \\ \varepsilon\boldsymbol{\Gamma} \end{bmatrix} [0 \quad \varepsilon\boldsymbol{\Gamma}^H] \\ &= \begin{bmatrix} t-\mu & \boldsymbol{\eta}^H \\ \boldsymbol{\eta} & \mathbf{I} - \frac{1}{\mu}\varepsilon^2 \boldsymbol{\Gamma}\boldsymbol{\Gamma}^H \end{bmatrix} \succeq \mathbf{0}. \end{aligned} \tag{21}$$

Using Schur's complement again, (21) can be written as

$$t - \mu - \boldsymbol{\eta}^H \left(\mathbf{I} - \frac{1}{\mu}\varepsilon^2 \boldsymbol{\Gamma}\boldsymbol{\Gamma}^H \right)^{-1} \boldsymbol{\eta} \ge \mathbf{0}$$

or equivalently

$$\mu + \sum_{i=1}^L \frac{(f_i g_i \gamma_i - 1)^2}{1 - \varepsilon^2 f_i^2 g_i^2 / \mu} \le t. \tag{22}$$

Combining (18) and (22), we get the following problem:

$$\begin{aligned} \min_{\mathbf{f},\mathbf{g},\mu} \quad & \sum_{i=1}^L \frac{(\gamma_i f_i g_i - 1)^2}{1 - \varepsilon^2 f_i^2 g_i^2 / \mu} + \mu + \sigma_n^2 \sum_{i=1}^L g_i^2 \\ \text{s.t.} \quad & \sum_{i=1}^L f_i^2 \le P, \quad \mu \ge \varepsilon^2 f_i^2 g_i^2 \end{aligned} \tag{23}$$

where the constraint $\mu \ge \varepsilon^2 f_i^2 g_i^2$ is implicitly included in the constraint (21).

Problem (23) is still difficult to deal with. However, we will next show that, by some variable transformations, the global optimal solution of problem (23) can be obtained. Define $\mathbf{m} \triangleq [m_1, \ldots, m_L]^T$ and $\mathbf{n} \triangleq [n_1, \ldots, n_L]^T$, where $m_i = f_i g_i$ and $n_i = g_i^2$, for $i = 1, \ldots, L$. Then, problem (23) becomes

$$\begin{aligned} \min_{\mathbf{m},\mathbf{n},\mu} \quad & \phi_1(\mathbf{m}, \mathbf{n}, \mu) \triangleq \sum_{i=1}^L \frac{(\gamma_i m_i - 1)^2}{1 - \varepsilon^2 m_i^2 / \mu} + \mu + \sigma_n^2 \sum_{i=1}^L n_i \\ \text{s.t.} \quad & \sum_{i=1}^L m_i^2 / n_i \le P, \quad \mu \ge \varepsilon^2 m_i^2. \end{aligned} \tag{24}$$

Let $\mathbf{s} \triangleq [s_1, \ldots, s_L]^T$. We claim that (24) is equivalent to the following problem:

$$\begin{aligned} \min_{\mathbf{m},\mathbf{n},\mu,\mathbf{s}} \quad & \phi_2(\mathbf{m}, \mathbf{n}, \mu, \mathbf{s}) \triangleq \sum_{i=1}^L \frac{(\gamma_i m_i - 1)^2}{1 - s_i} + \mu + \sigma_n^2 \sum_{i=1}^L n_i \\ \text{s.t.} \quad & \sum_{i=1}^L m_i^2 / n_i \le P, \ \frac{\varepsilon^2 m_i^2}{\mu} \le s_i < 1. \end{aligned} \tag{25}$$

This can be explained as follows. First, suppose that $(\mathbf{m}^\sharp, \mathbf{n}^\sharp, \mu^\sharp, \mathbf{s}^\sharp)$ is the optimal solution of (25). In view of (24), it follows that

$$\begin{aligned} &\min_{\mathbf{m},\mathbf{n},\mu} \phi_1(\mathbf{m}, \mathbf{n}, \mu) \le \phi_1(\mathbf{m}^\sharp, \mathbf{n}^\sharp, \mu^\sharp) \\ &\qquad \overset{(a)}{\le} \phi_2(\mathbf{m}^\sharp, \mathbf{n}^\sharp, \mu^\sharp, \mathbf{s}^\sharp) = \min \phi_2(\mathbf{m}, \mathbf{n}, \mu, \mathbf{s}) \end{aligned}$$

where in (a), we have used the constraint $\varepsilon^2 m_i^{\sharp 2} / \mu^\sharp \le s_i^\sharp$. Then, we know that $\min \phi_1(\mathbf{m}, \mathbf{n}, \mu) \le \min \phi_2(\mathbf{m}, \mathbf{n}, \mu, \mathbf{s})$. On the other hand, for any feasible $(\mathbf{m}, \mathbf{n}, \mu)$ of (24), we can always find some $\mathbf{s}$, which makes the equality hold in (25). This means that it is also feasible to (25). Then, we must have $\min \phi_1(\mathbf{m}, \mathbf{n}, \mu) \ge \min \phi_2(\mathbf{m}, \mathbf{n}, \mu, \mathbf{s})$. Thereby, we must have $\min \phi_1(\mathbf{m}, \mathbf{n}, \mu) = \min \phi_2(\mathbf{m}, \mathbf{n}, \mu, \mathbf{s})\}$.

Introducing the slack variable $\mathbf{z} \triangleq [z_1, \ldots, z_L]^T$, problem (25) can be written as

$$\begin{aligned} \min_{\mathbf{m},\mathbf{n},\mu,\mathbf{s},\mathbf{z}} \quad & \sum_{i=1}^L z_i + \mu + \sigma_n^2 \sum_{i=1}^L n_i \\ \text{s.t.} \quad & \begin{bmatrix} z_i & \gamma_i m_i - 1 \\ \gamma_i m_i - 1 & 1 - s_i \end{bmatrix} \succeq \mathbf{0}, \quad i = 1, \ldots, L \\ & \begin{bmatrix} \mu & \varepsilon m_i \\ \varepsilon m_i & s_i \end{bmatrix} \succeq \mathbf{0}, \quad s_i < 1, \ i = 1, \ldots, L \\ & \sum_{i=1}^L \frac{m_i^2}{n_i} \le P. \end{aligned} \tag{26}$$

Problem (26) is a semidefinite programming (SDP) problem, which can be efficiently solved by the MATLAB package tools such as CVX [15]. Then, the optimal f_i and g_i are obtained by $f_i = m_i / \sqrt{n_i}$ and $g_i = \sqrt{n_i}$.

Remark 2: By fixing $\mathbf{f}$ or $\mathbf{g}$, problem (23) becomes problem (6) or (7) in [1], where they were proved to be convex and can be optimally solved, respectively. This process is repeated until convergence. However, the solution of this alternating-optimization-based method depends on the initial point $\mathbf{f}^{(0)}$ and may not be optimal if problem (23) has local minimal. Thus, it is natural to ask question $\mathcal{Q}2$, does this method converge to the global optimal solution? Unfortunately, this is not guaranteed. As will be seen in Section V, different initial point $\mathbf{f}^{(0)}$ will lead to different results and may also incur some performance loss. Therefore, our answer to question $\mathcal{Q}2$ is *not always, and it depends on the initial point*.

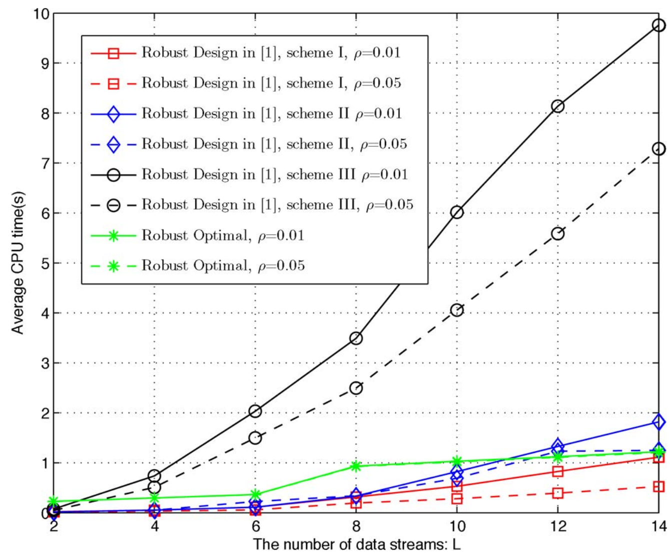

Fig. 1. Average CPU time comparison versus different L for $M = N = L$ and $P = 20$ dBW.

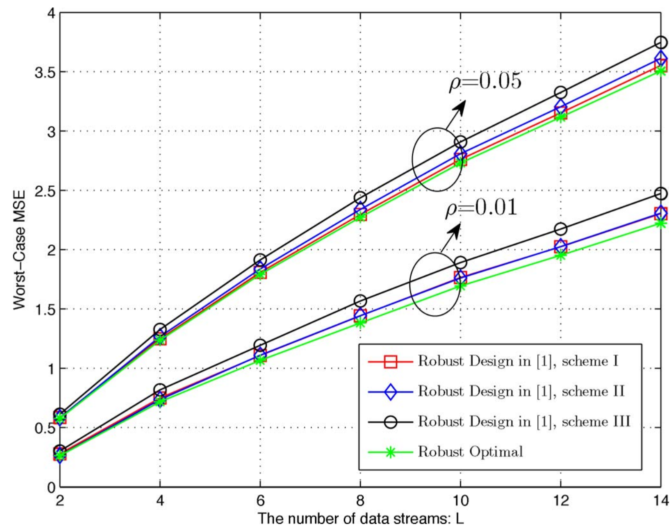

Fig. 2. Worst-case MSE versus different L for $M = N = L$ and $P = 20$ dBW.

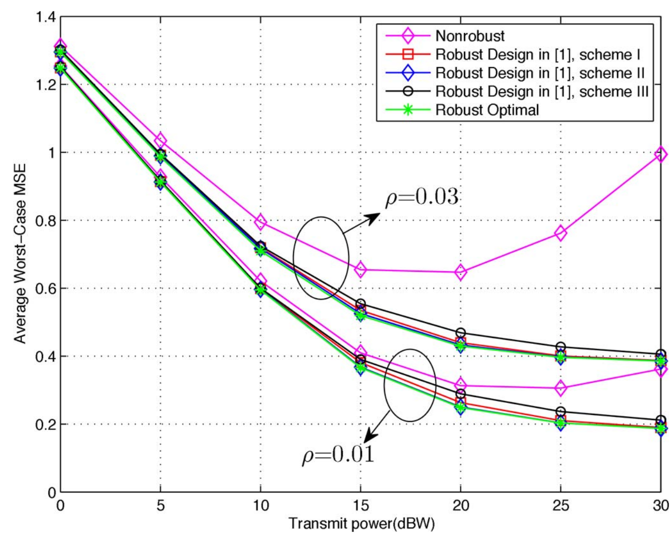

Fig. 3. Worst-case MSE versus different transmit power for $M = N = L = 2$.

V. Discussions and Simulations

Here, we first provide the complexity comparison between our joint optimal design and the robust design in [1] and then give numerical results to compare the two robust designs.

The channel fading is modeled as Rayleigh fading, and each channel entry satisfies the complex normal distribution $\mathcal{CN}(0,1)$. The noise is assumed to be zero-mean unit-variance complex Gaussian random variables. In our simulations, we set $M = N = L$ and vary ε through the normalized parameter $\rho \in [0,1)$, i.e., $\varepsilon^2 = \rho\|\tilde{\mathbf{H}}\|_F^2$. Then, the larger the ρ is, the poorer the CSI quality will be. All results are averaged over 1000 channel realizations.

A. Complexity Comparison

Since we have derived the optimal structure of transceiver design in Theorem 1, the major computing step in our work remains in solving problem (26). The complexity for solving the SDP problem (26) is $\mathcal{O}(L^2)$ per iteration, and the number of iterations typically lies between 5 and 50, for an SDP problem [16]. On the other hand, the complexity analysis of the method in [1] is a little complicated. In their work, when fixing $\mathbf{g}$, the optimal $\mathbf{f}$ was obtained by the three-level primal–primal decomposition method, where a close-form solution was given in the lowest level, while the bisection method and the gradient method were applied at the middle and third level, respectively. Similarly, the problem for determining optimal $\mathbf{g}$ under fixed $\mathbf{f}$ was also solved in two levels, where a close-form solution was derived at the first level, while the bisection method was used at the second level. It can be seen that it is hard to determine the complexity of each iteration, as well as the exact (or even approximate) iteration number in this method. Upon this observation, we resort to the CPU time comparison required by the two methods.

Fig. 1 shows the average CPU time comparison between our robust optimal method and the robust method in [1]. We set $P = 20$ dBW and choose different initial points for the method in [1]: a) Scheme I: Set the initial point $\mathbf{f}^{(0)}$ with equal elements; b) Scheme II: Set $\mathbf{f}^{(0)}$ as the nonrobust solution that takes the nominal channel $\tilde{\mathbf{H}}$ as the actual channel [5]; c) Scheme III: Set $\mathbf{f}^{(0)}$ as a random variable that satisfies $\|\mathbf{f}^{(0)}\|_2^2 = P$. It can be observed in Fig. 1 that Scheme III is the most time-consuming scheme among all the schemes. Although the time cost by Schemes I and II is similar to that in our global optimal solution, it increases more rapidly as the number of data streams L becomes large. In Fig. 1, we know that the required CPU time for the method in [1] tends to be more random in nature and heavily depends on the initial point. On the other hand, our method is not only efficient but also has a more stable runtime performance.

B. Numerical Results

We now study the system MSE performance in different scenarios. In Fig. 2, we set the same network configuration as that in Fig. 1 and investigate the average worst-case MSE performance versus L of our method and three different schemes in [1]. As shown in the plot, Scheme I and Scheme II suffer some marginal performance loss, while scheme III incurs some apparent loss, which grows even larger when L increases. Hence, the method in [1] does not always lead to the optimal solution.

As another example, Fig. 3 depicts the average worst case MSE performance versus different transmit power with $\rho = 0.01$ and $\rho = 0.03$. We consider the case when $M = N = L = 2$. Fig. 3 verifies the superior performance of robust schemes over the nonrobust scheme in [5]. Moreover, it can be observed that Scheme II has an almost optimal performance, while Scheme I approximately approaches to the optimal solution. Therefore, when the advanced software package (such as CVX) is not available, Scheme I (or Scheme II) can be viewed as a simple implementation of the global optimal method.

VI. Conclusion and Future Work

In this paper, we have investigated the global optimal transceiver design in the MIMO link under a deterministic CSI uncertainty model. We first prove that the optimal design of the transmitter equalizer has a favorable channel-diagonalizing structure. Then, we simplify the original problem into a scalar optimization problem and obtain the global optimal solution via an SDP problem. Simulation results show that our method outperforms the existing schemes.

We only considered the point-to-point MIMO system in this work. However, as pointed out in [4], the transmission can also be affected by the multiple-access interference, if the transmit–receive nodes are active over a communication network, which employs nonorthogonal multiplexing. In this case, the received signal at the destination is contaminated by the spatially colored Gaussian noise, and the robust optimal transceiver design must be reconsidered. Hence, it would be interesting to address this issue in our future research.

Acknowledgment

The authors would like to thank Prof. Y. L. Guan and Prof. C. Yuen for fruitful discussions.

References

[1] J. Wang and M. Bengtsson, "Joint optimization of the worst-case robust MMSE MIMO transceiver," *IEEE Signal Process. Lett.*, vol. 18, no. 5, pp. 295–298, May 2011.

[2] Z. Wang and W. Chen, "Relay beamforming design with SIC detection for MIMO multi-relay networks with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 62, no. 8, pp. 3774–3785, Oct. 2013.

[3] E. Baccarelli, M. Biagi, C. Pelizzoni, and N. Cordeschi, "Optimal MIMO UWB-IR transceiver for Nakagami-fading and Poisson-arrivals," *J. Commun.*, vol. 3, no. 1, pp. 27–40, Jan. 2008.

[4] E. Baccarelli, M. Biagi, C. Pelizzoni, and N. Cordeschi, "Optimized power allocation for multiantenna systems impaired by multiple access interference and imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 3089–3105, Sep. 2007.

[5] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[6] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[7] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1141–1150, Mar. 2009.

[8] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.

[9] J. Wang and D. P. Palomar, "Robust MMSE precoding in MIMO channels with pre-fixed receivers," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5802–5818, Nov. 2010.

[10] N. Khaled, S. Thoen, and L. Deneire, "Optimizing the joint transmit and receive MMSE design using mode selection," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 730–737, Apr. 2005.

[11] H. Shen, J. Wang, B. C. Levy, and C. Zhao, "Robust optimization for amplify-and-forward MIMO relaying from a worst-case perspective," *IEEE Trans. Signal Process.*, vol. 61, no. 21 pp. 5458–5471, Nov. 2013.

[12] H. Shen, J. Wang, W. Xu, Y. Rong, and C. Zhao, "A worst-case robust MMSE transceiver design for nonregenerative MIMO relaying," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 695–709, Feb. 2014.

[13] H. Tang, W. Chen, J. Li, and H. Wan, "Achieving global optimality for joint source and relay beamforming design in two-hop relay channels," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4422–4435, Nov. 2014.

[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[15] M. Grant and S. Boyd, CVX' Users' Guide, 2009. [Online]. Available: http://cvxr.com/cvx/doc/index.html

[16] L. Vandenberghe and S. Boyd, "Semidefinte programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, Mar. 1996.

A Low-Complexity Power Allocation Algorithm for Multiple-Input–Multiple-Output Spatial Modulation Systems

Ping Yang, *Member, IEEE*, Yue Xiao, *Member, IEEE*, Shaoqian Li, *Senior Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—In this paper, a low-complexity Euclidean distance-based power allocation (PA) algorithm is proposed for spatial modulation (SM)-aided multiple-input–multiple-output (MIMO) systems. The proposed algorithm only exploits the Euclidean distances of a few dominant error vectors for optimizing the PA matrix; hence, its computational complexity is considerably reduced compared with the exhaustive-search-based algorithm. Our simulation results show that the proposed algorithm provides beneficial bit error ratio (BER) performance improvements compared with both the conventional SM- and PA-based spatial multiplexing arrangements.**

***Index Terms*—Euclidean distance, link adaptation (LA), power allocation (PA), spatial modulation (SM).**

I. Introduction

Link adaptation (LA) has an important role in wireless communication [1]. Recently, LA techniques have also been extended to the family of spatial modulation (SM) techniques [2]–[5], which constitute a novel class of low-complexity yet energy-efficient [6], [7] multiple-input–multiple-output (MIMO) transmission techniques. More specifically, in [8]–[11], several transmit antenna selection (TAS) algorithms relying on different optimization criteria, such as the norm-based TAS of [8] and the antenna-correlation-based TAS of [10], were proposed for striking an attractive multiplexing–diversity gain tradeoff. Moreover, by exploiting the MIMO scheme's degrees of freedom, both adaptive modulation (AM) arrangements [12], [13] and meritorious constellation designs [14]–[16] have been investigated in the literature of SM-MIMO systems.

Manuscript received August 19, 2014; revised January 6, 2015; accepted February 27, 2015. Date of publication March 4, 2015; date of current version March 10, 2016. This work was supported in part by the National Basic Research Program of China under Grant 2013CB329001, by the National Science Foundation of China under Grant 61471090, by the European Research Council under an Advanced Fellow Grant, and by the National Natural Science Foundation of China under Grant 61101101. The review of this paper was coordinated by Prof. H.-H. Chen.

P. Yang, Y. Xiao, and S. Li are with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Sichuan 611731, China (e-mail: yplxw@163.com; xiaoyue@uestc.edu.cn; lsq@uestc.edu.cn).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

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/TVT.2015.2410252