<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Joint robust weighted LMMSE transceiver design for dual-hop AF multiple-antenna relay systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Xing, C; Ma, S; Fei, Z; Wu, YC; Kuang, J</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>The 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011), Beijing, China, 5-9 December 2011. In Globecom. IEEE Conference and Exhibition, 2011, p. 1-5</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2011</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/165230">http://hdl.handle.net/10722/165230</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>Globecom. IEEE Conference and Exhibition. Copyright © IEEE.</td>
</tr>
</tbody>
</table>
Joint Robust Weighted LMMSE Transceiver Design for Dual-Hop AF Multiple-Antenna Relay Systems

Chengwen Xing*, Shaodan Ma†, Zesong Fei*, Yik-Chung Wu* and Jingming Kuang*

*School of Information and Electronics, Beijing Institute of Technology, Beijing, China
Email: {xingchengwen}@gmail.com {feizesong, jmkuang}@bit.edu.cn
†Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong
Email: {sdma, ycwu}@eee.hku.hk

Abstract—In this paper, joint transceiver design for dual-hop amplify-and-forward (AF) MIMO relay systems with Gaussian distributed channel estimation errors in both two hops is investigated. Due to the fact that various linear transceiver designs can be transformed to a weighted linear minimum mean-square-error (LMMSE) transceiver design with specific weighting matrices, weighted mean square error (MSE) is chosen as the performance metric. Precoder matrix at source, forwarding matrix at relay and equalizer matrix at destination are jointly designed with channel estimation errors taken care of by Bayesian philosophy. Several existing algorithms are found to be special cases of the proposed solution. The performance advantage of the proposed robust design is demonstrated by the simulation results.

I. INTRODUCTION

By deploying relays, cooperative communication has great potential to improve system performance [1]. Among various relaying strategies, amplify-and-forward (AF) scheme is attractive for practical implementation due to its low complexity. On other hand, it is well-established that employing multiple antennas is beneficial to improve wireless system performance due to spatial diversity and multiplexing gains. Consequently, in order to obtain the virtues of these two techniques, combination of AF transmission and multi-input multi-output (MIMO) systems attracts more and more researchers’ interest. Linear transceiver design for AF MIMO relay systems has been widely researched in [2]–[6]. Unfortunately, most of the existing works assume channel state information (CSI) is perfectly known. However channel estimation errors are inevitable in practical systems. Robust design, which takes channel estimation errors into account, is therefore desirable in practice. Joint robust design of relay forwarding matrix and destination equalizer under channel estimation errors has been considered in [7]. However in [7], source precoder design, which provides further degree of freedom for improving system performance, is not considered. In this paper, we take a step further to jointly design source precoder matrix, relay forwarding matrix and destination equalizer matrix under channel estimation errors.

There are two main kinds of criteria for transceiver design: capacity maximization and mean-square-error (MSE) minimization. In fact, both kinds of criteria can be formulated as weighted MSE minimization problems [8]. Therefore, weighted MSE is chosen as the objective function in this paper.

In this paper, the channel estimation errors are assumed to be Gaussian distributed, which is generally true when linear channel estimation algorithms are adopted [9]. The structures of the optimal solutions for the robust joint transceiver design with minimum weighted MSE as the objective are first derived. Based on these optimal structures, the transceiver design is then significantly simplified and iterative water-filling is adopted to solve the problem. Finally, the performance gain of the proposed robust design is verified by simulation.

The following notations are used throughout this paper. Boldface lowercase letters denote vectors, while boldface uppercase letters denote matrices. The notations $Z^T$, $Z^*$ and $Z^H$ denote the transpose, conjugate and conjugate transpose of the matrix $Z$, respectively and $\text{Tr}(Z)$ is the trace of the matrix $Z$. The symbol $I_M$ denotes an $M \times M$ identity matrix, while $0_{M,N}$ denotes an $M \times N$ all zero matrix. The notation $Z^{1/2}$ is the Hermitian square root of the positive semi-definite matrix $Z$, such that $Z^{1/2}Z^{1/2} = Z$ and $Z^{1/2}$ is also a Hermitian matrix. The symbol $E$ denotes statistical expectation.

II. SYSTEM MODEL

A. Transmitted and Received Signals

In this paper, a three node dual-hop amplify-and-forward (AF) cooperative communication system is considered. In the considered system, there is one source with $N_S$ antennas, one relay with $M_R$ receive antennas and $N_R$ transmit antennas, and one destination with $M_D$ antennas. Due to long distance and possibly deep fading, the direct link between the source and destination is not considered in this paper. At the first hop, the source transmits data to the relay. The received signal, $x$, at the relay is

$$x = H_{sr} P s + n_1$$  \hspace{1cm} (1)

where $H_{sr}$ is the MIMO channel matrix between the source and the relay, and $P$ is the precoder matrix at the source. The vector $s$ is the $N \times 1$ data vector transmitted by the source with the covariance matrix $R_s = E \{ s s^H \} = I_N$. Furthermore, $n_1$ is the additive Gaussian noise vector with correlation matrix $R_{n_1} = \sigma^2 I_{M_R}$.

At the relay, the received signal $x$ is multiplied by a forwarding matrix $F$. Then the resultant signal is transmitted
to the destination. The received signal $y$ at the destination can be written as

$$y = H_{rd} F_{sr} s + H_{rd} n_2,$$

(2)

where $H_{rd}$ is the MIMO channel matrix between the relay and the destination, and $n_2$ is the additive Gaussian noise vector at the second hop with correlation matrix $R_{n_2} = \sigma_n^2 I_{M_D}$. In order to guarantee the transmitted data $s$ can be recovered at the destination, it is assumed that $N_S, M_R, N_R,$ and $M_D$ are greater than or equal to $N$ [2].

B. Channel Estimation Error Model

Channel state information is usually estimated via using training sequences. In a dual-hop system, there are two channels to be estimated. For implementation simplicity, training sequences are transmitted from both source and destination to the relay. At the relay, the two channels are separately estimated and then based on the channel estimates, transceivers are jointly designed. After that, the designed transceiver matrices are forward to the source and destination, respectively. In general, two hop channels can be written as

$$H_{sr} = \hat{H}_{sr} + \Sigma_{sr}^{1/2} W_{sr} \Psi_{sr}^{1/2},$$

$$H_{rd} = \hat{H}_{rd} + \Sigma_{rd}^{1/2} W_{rd} \Psi_{rd}^{1/2},$$

(3)

where $\hat{H}_{sr}$ and $\hat{H}_{rd}$ are the estimated channels, and $\Sigma_{sr}$ and $\Sigma_{rd}$ are the corresponding channel estimation errors whose elements are zero mean Gaussian random variables. Furthermore, the $M_R \times N_S$ matrix $\Delta H_{sr}$ can be decomposed as $\Delta H_{sr} = \Sigma_{sr}^{1/2} W_{sr} \Psi_{sr}^{1/2}$, where the elements of the $M_R \times N_S$ matrix $\Sigma_{sr}$, are independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and unit variance. When MMSE channel estimator is used, the row and column covariance matrices of $\Delta H_{sr}$ can be derived as [9]

$$\Sigma_{sr} = I_{M_R},$$

(4)

$$\Psi_{sr} = (\sigma_{sr}^2 I_{N_S} + \frac{1}{\sigma_n^2} D_1 D_1^T)^{-1},$$

where $\sigma_{sr}^2$ is channel variance in the first hop and $D_1$ is the training sequence used in the first hop. Similarly, the row and column covariance matrices of $\Delta H_{rd}$, can also be derived to be

$$\Sigma_{rd} = (\sigma_{rd}^2 I_{M_D} + \frac{1}{\sigma_n^2} D_2 D_2^T)^{-1},$$

$$\Psi_{rd} = I_{N_R},$$

(5)

where $\sigma_{rd}^2$ denotes the channel variance of the second hop and $D_2$ represents the training sequence used in the second hop. Notice that the directions of training sequences transmitted in the two hops are opposite. Therefore, the results given by (4) and (5) are different.

III. Problem Formulation

At the destination, a linear equalizer $G$ is adopted to detect the data vector $s$. The mean-square-error (MSE) matrix is $\mathbb{E}\{ (Gy-s)^H (Gy-s) \}$, where $s$ is the expectation taken with respect to random data, channel estimation errors, and noise. In [7], it is shown that

$$\mathbb{E}\{ (Gy-s)^H (Gy-s) \} = \mathbb{E}\{(\hat{H}_{sr} F_{sr} F_{rd} + K_2^2)G^H + I_N - (P^H \hat{H}_{sr} F^H \hat{H}_{rd}^H)G^H \} - (\hat{H}_{sr} F_{sr} P),$$

(6)

where matrices $R_x$ and $K_2$ are defined as

$$R_x = \mathbb{E}\{xx^H\} = F_{sr} P F_{rd}^H \hat{H}_{sr}^H + K_1,$$

$$K_1 = Tr\{PP^H \Psi_{sr} \Sigma_{sr} + R_{n_2} \} + \Sigma_{rd} + R_{n_2},$$

$$K_2 = Tr\{FR_{sr}^H \Psi_{rd} \Sigma_{rd} + R_{n_2} \} + \Sigma_{rd} + R_{n_2}.$$
IV. THE STRUCTURE OF OPTIMAL $\tilde{F}$

With
\[
\begin{align*}
\mathbf{A} & \triangleq (\mathbf{I}_P^{-1/2} \mathbf{K}_1^{-1/2} \mathbf{H}_s \mathbf{P} \mathbf{W}^{1/2}) \\
\mathbf{M} & \triangleq \tilde{F}^H \bar{H}_d (\bar{H}_d \tilde{F}^H \bar{H}_d + \mathbf{K}_2)^{-1} \bar{H}_d \tilde{F},
\end{align*}
\]
the weighted MSE can be further reformulated as
\[
\text{MSE}_W(\tilde{F}, \mathbf{P}) = \text{Tr}(\mathbf{W}) - \text{Tr}(\mathbf{A}^H \mathbf{M} \mathbf{A}).
\]

Using the matrix inversion lemma, $\mathbf{M}$ can be reformulated as
\[
\text{MSE}_W(\tilde{F}, \mathbf{P}) = \text{Tr}(\mathbf{W}) - \text{Tr}(\mathbf{A}^H \mathbf{M} \mathbf{A}).
\]

Using the matrix inversion lemma, $\mathbf{M}$ can be reformulated as
\[
\text{MSE}_W(\tilde{F}, \mathbf{P}) = \text{Tr}(\mathbf{W}) - \text{Tr}(\mathbf{A}^H \mathbf{M} \mathbf{A}).
\]

It should be highlighted that the two constraints in (15) are independent and for any given $\mathbf{P}$ the optimization problem for $\tilde{F}$ is equivalent to
\[
\begin{align*}
\min_{\tilde{F}} \quad & \text{Tr}[\mathbf{A}^H (\tilde{F}^H \bar{H}_d \mathbf{K}_2^{-1} \bar{H}_d \tilde{F} + \mathbf{M})^{-1} \mathbf{A}] \\
\text{s.t.} \quad & \text{Tr}(\tilde{F}^H \mathbf{W} \tilde{F}) = P_r,
\end{align*}
\]
where the constraint parts independent of $\tilde{F}$ have been neglected. With the fact that $\mathbf{W} \mathbf{P} = \mathbf{I}$ from (5), we have
\[
\begin{align*}
\mathbf{K}_2 = \text{Tr}(\tilde{F}^H \mathbf{W} \tilde{F}) \mathbf{I} + s \mathbf{I}_M + \mathbf{M} &= P_r \mathbf{I} + s \mathbf{I}_M + \mathbf{M}.
\end{align*}
\]

which is a constant matrix. Furthermore, based on singular value decomposition
\[
\mathbf{A} = \mathbf{U}_A \mathbf{L}_A \mathbf{V}_A^H, \quad \mathbf{K}_2^{-1/2} \bar{H}_d = \mathbf{U}_A \mathbf{L}_A \bar{H}_d \mathbf{V}_A^H,
\]

where the diagonal elements of $\mathbf{L}_A$ and $\mathbf{L}_A$ are in decreasing order. Using Majorization theory, the objective function of (19) can be transformed into a Schur-concave function of $\mathbf{K}_1$, which is composed of $\mathbf{U}_A \mathbf{L}_A \mathbf{V}_A ^H$ of $\mathbf{A}$ in (16) and using the matrix inversion lemma again, the weighted MSE (25) can be further rewritten as
\[
\text{MSE}_W(\tilde{F}, \mathbf{P}) = \text{Tr}(\mathbf{W}) - \text{Tr}(\mathbf{A}^H \mathbf{M} \mathbf{A}).
\]

It should be highlighted that the two constraints in (15) are independent and for any given $\mathbf{P}$ the optimization problem for $\tilde{F}$ is equivalent to
\[
\begin{align*}
\min_{\tilde{F}} \quad & \text{Tr}[\mathbf{A}^H (\tilde{F}^H \bar{H}_d \mathbf{K}_2^{-1} \bar{H}_d \tilde{F} + \mathbf{M})^{-1} \mathbf{A}] \\
\text{s.t.} \quad & \text{Tr}(\tilde{F}^H \mathbf{W} \tilde{F}) = P_r,
\end{align*}
\]
where $\mathbf{M}$ is a $N \times N$ principal submatrix of $\mathbf{L}_A$ and $\mathbf{K}_2^{-1/2} \bar{H}_d$ are in decreasing order. The diagonal matrix $\mathbf{L}_A$ is the $N \times N$ principal submatrix of $\mathbf{L}_A$. The matrices $\mathbf{V}_{r,d,N}$ and $\mathbf{U}_{A,N}$ are the first $N$ columns of $\mathbf{V}_{r,d}$ and $\mathbf{U}_A$, respectively. As $\mathbf{A}$ is a function of $\mathbf{P}$, the value of $\mathbf{U}_{A,N}$ will be given in the next section, after giving the structure of optimal $\mathbf{P}$.

V. THE STRUCTURE OF OPTIMAL $\mathbf{P}$

Substituting the structure of optimal $\tilde{F}$ (22) into the definition of $\mathbf{M}$ in (16), and together with the fact that the diagonal elements of $\mathbf{A}_\mathbf{F} \mathbf{K}_2^{-1/2} \bar{H}_d \mathbf{A}_\mathbf{F}$ are in decreasing order, we have
\[
\begin{align*}
\mathbf{M} & = \mathbf{U}_{A,N} [\mathbf{I}_N - (\mathbf{A}_\mathbf{F} \mathbf{K}_2^{-1/2} \bar{H}_d + \mathbf{I}_N)^{-1}] \mathbf{U}_{A,N}^H, \\
\tilde{A}_\mathbf{M} & = \mathbf{A}_\mathbf{M} \mathbf{V}^H = \mathbf{A} \tilde{A}_\mathbf{M}^H \mathbf{A}^H. \quad (23)
\end{align*}
\]

where $\tilde{A}_\mathbf{M}$ is a $N \times N$ diagonal matrix with diagonal elements in decreasing order. It is also straightforward that (23) is the eigen-decomposition of $\mathbf{M}$. Based on (23), and the singular value decomposition of $\mathbf{A}$ given in (21), after a straightforward substitution we have the following identity
\[
\begin{align*}
\mathbf{A}^H \mathbf{M} \mathbf{A} & = \mathbf{V} \tilde{A}_\mathbf{M} \tilde{A}_\mathbf{M} \mathbf{V}^H = \mathbf{V} \tilde{A}_\mathbf{M} \mathbf{V}^H \mathbf{A}^H \mathbf{A}.
\end{align*}
\]

where $\tilde{A}_\mathbf{M}$ is the $N \times N$ principal submatrix of $\mathbf{L}_A$ and $\mathbf{K}_2^{-1/2} \bar{H}_d \mathbf{A}_\mathbf{F}$ are in decreasing order. The diagonal matrix $\tilde{A}_\mathbf{M}$ is the $N \times N$ principal submatrix of $\mathbf{L}_A$. The matrices $\mathbf{V}_{r,d,N}$ and $\mathbf{U}_{A,N}$ are the first $N$ columns of $\mathbf{V}_{r,d}$ and $\mathbf{U}_A$, respectively. As $\mathbf{A}$ is a function of $\mathbf{P}$, the value of $\mathbf{U}_{A,N}$ will be given in the next section, after giving the structure of optimal $\mathbf{P}$.

with
\[
\begin{align*}
\mathbf{M} & = \mathbf{U}_{A,N} [\mathbf{I}_N - (\mathbf{A}_\mathbf{F} \mathbf{K}_2^{-1/2} \bar{H}_d + \mathbf{I}_N)^{-1}] \mathbf{U}_{A,N}^H.
\end{align*}
\]
$$\min_{\hat{P}} \operatorname{Tr}[U_W A_{\hat{f}}^{1/2} \tilde{A}_{U} A_{U}^{1/2} U_W^H (\tilde{H}^H (P_s \Psi_{sr} + \sigma_{n1}^2 I_{N_s})^{-1/2} \tilde{H}^H (P_s \Psi_{sr} + \sigma_{n1}^2 I_{N_s})^{-1/2} \tilde{P} + I_N)^{-1}]$$

s.t. $\operatorname{Tr}(\tilde{P}^H) = P_s$ \hspace{1cm} (34)

as the optimization problem (19) for $\tilde{F}$. Following the same argument for $\tilde{F}$ and defining unitary matrices $U_{sr}$ and $V_{sr}$ based on the following singular value decomposition

$$\tilde{H}_{sr}(P_s \Psi_{sr} + \sigma_{n1}^2 I_{N_s})^{-1/2} = U_{sr} A_{sr} V_{sr}^H,$$ \hspace{1cm} (35)

with the diagonal elements of the diagonal matrix $A_{sr}$ in decreasing order, the optimal $\tilde{P}$ has the following structure

$$\tilde{P} = V_{sr,N} A_{\tilde{P}} U_{\tilde{P}}^H,$$ \hspace{1cm} (36)

where $A_{\tilde{P}}$ is a $N \times N$ diagonal matrix such that the diagonal elements of $A_{\tilde{P}}$, $A_{sr}$ are in decreasing order. The diagonal matrix $A_{sr}$ is the $N \times N$ principal submatrix of $A_{sr}$. Furthermore, based on the definition of $\tilde{P}$ given by (33), the optimal $\tilde{F}$ has the following structure

$$\tilde{F} = V_{sr,N} A_{\tilde{F}} U_{\tilde{F}}^H,$$ \hspace{1cm} (37)

Substituting (37) into the definition of $A$ in (16), we have $U_{A,N} = U_{sr,N}$. Therefore, the optimal $\tilde{F}$ has the following structure

$$\tilde{F} = V_{sr,N} A_{\tilde{F}} U_{\tilde{F}}^H.$$ \hspace{1cm} (38)

**Remark:** Given (37) and (38), the remaining problem is how to determine two diagonal matrices $A_{\tilde{F}}$ and $A_{\tilde{P}}$.

Notice that after $A_{\tilde{P}}$ is computed, the remaining unknown parameter in (37) is only $\eta_{\tilde{P}}$. In order to solve $\eta_{\tilde{P}}$, substitute the formulation of $P$ (37) into the definition of $\eta_{\tilde{P}}$ in (29), and then we get

$$\eta_{\tilde{P}} = \operatorname{Tr}(P \Psi_{sr} + \sigma_{n1}^2 I_{N_s})^{-1/2} \Psi_{sr}(P \Psi_{sr} + \sigma_{n1}^2 I_{N_s})^{-1/2} \times \Lambda_{\tilde{P}}^{H} + \sigma_{n1}^2.$$ \hspace{1cm} (39)

This is a simple linear function of $\eta_{\tilde{P}}$, and $\eta_{\tilde{P}}$ can be easily solved to be

$$\eta_{\tilde{P}} = \sigma_{n1}^2 / \{1 - \operatorname{Tr}[V_{sr,N}^H (P \Psi_{sr} + \sigma_{n1}^2 I_{N_s})^{-1/2} \Psi_{sr} (P \Psi_{sr} + \sigma_{n1}^2 I_{N_s})^{-1/2} \times \Lambda_{\tilde{P}}^{H} + \sigma_{n1}^2]\}. \hspace{1cm} (40)$$

**VI. PROPOSED SOLUTION FOR $A_{\tilde{F}}$ AND $A_{\tilde{P}}$**

In this section, the optimal $A_{\tilde{F}}$ and $A_{\tilde{P}}$ will be derived. Notice that $A_{sr}$ and $A_{rd}$ are the $N \times N$ principal submatrices of $A_{sr}$ and $A_{rd}$, respectively. Based on (31), substituting the optimal structures (37) and (38) into the original optimization problem (15), and denoting

$$\tilde{A}_{sr} = \text{diag}\{\lambda_{sr,i}\}, \quad \tilde{A}_{rd} = \text{diag}\{\lambda_{rd,i}\}, \quad A_{\tilde{F}} = \text{diag}\{f_i\}, \quad A_{\tilde{P}} = \text{diag}\{p_i\}, \quad A_{\tilde{W}} = \text{diag}\{w_i\}$$ \hspace{1cm} (41)

the optimization problem (15) becomes

$$\min_{f_i, p_i} \sum_{i=1}^{N} w_i (f_i^2 \lambda_{sr, i}^2 + \sigma_{n1}^2 \lambda_{sr, i}^2 + 1) / (p_i^2 \lambda_{sr, i}^2 + 1) (f_i^2 \lambda_{rd, i}^2 + 1)$$

s.t. $\sum_{i=1}^{N} f_i^2 = P_f$, \hspace{1cm} $\sum_{i=1}^{N} p_i^2 = P_s.$ \hspace{1cm} (42)
based on which, the weighted MSE (18) can be rewritten as
\[
\text{MSE}_W(\tilde{F}, P) = \text{Tr}(W) - \text{Tr}(WN) \\
\geq \text{Tr}(W) - \sum \lambda_i(W) \lambda_i(N),
\]  
(48)
where $\lambda_i(Z)$ denotes the $i$th largest eigenvalue of $Z$. Using Neumann inequality [12], for the minimum weighted MSE $\tilde{N}$ and $W$ have the same eigen-vectors. In other words, given the eigen-decomposition of $W$ as
\[
W = U_W \Lambda_W U_W^H
\]  
(49)
where the diagonal elements of the diagonal matrix $\Lambda_W$ are in decreasing order, for the minimum MSE $\tilde{N}$ could be eigen-decomposed as
\[
\tilde{N} = U_W \Lambda_N U_W^H.
\]  
(50)
where $\Lambda_N$ is a diagonal matrix whose diagonal elements are in decreasing order. Based on (49) and (50), for the minimum MSE it holds that
\[
A^H M A = W^{1/2} N W^{1/2} = U_W \Lambda_1 \Lambda_2 U_W^H
\]  
(51)
As the diagonal elements of the diagonal matrices $\Lambda_W$ and $\Lambda_N$ are positive and both in decreasing order, the diagonal elements of $\Lambda_2$ are also in decreasing order. Clearly, the second equation of (51) also denotes an eigen-decomposition of $A^H M A$. Comparing (45) and (51), it can be concluded that for the minimum MSE there exists an eigen-decomposition of $A^H M A$ with eigenvalues in decreasing order such that $V_A = U_W$.

REFERENCES