Optimal Training Design and Individual Channel Estimation for MIMO Two-Way Relay Systems in Colored Environment

Rui Wang*, Hani Mehrpouyan†, Meixia Tao**, and Yingbo Hua‡
*Department of Information and Communications, Tongji University, Shanghai, China.
†Department of Computer and Electrical Engineering and Computer Science, California State University, Bakersfield, CA, USA.
**Department of Electronic Engineering at Shanghai Jiao Tong University, Shanghai, China.
‡Department of Electrical Engineering at the University of California, Riverside, CA, USA.
Emails: liouxingrui@gmail.com, hani.mehr@ieee.org, mxtao@sjtu.edu.cn, yhua@ee.ucr.edu.

***Abstract*—In this paper, while considering the impact of antenna correlation and the interference from neighboring users, we study the problem of channel estimation and training sequence design in multi-input multi-output (MIMO) two-way relaying (TWR) systems. To this end, we propose to decompose the bidirectional transmission links into two phases, i.e., the multiple access (MAC) and the broadcasting (BC) phases. By deriving the optimal linear minimum mean-square-error estimators, the corresponding training design problems for the MAC and BC phases are formulated and solved. Subsequently, algorithms and, in some special cases, closed-form solutions for obtaining the optimal training sequences for channel estimation in TWR systems are derived. Moreover, to further reduce channel estimation overhead, the minimum required length of the training sequences are determined. Simulation results verify the effectiveness of the proposed training designs in improving channel estimation performance in TWR systems.**

I. Introduction

Two-way relaying (TWR) has received great attention recently due to its high spectrum efficiency, and in the mean time, maintaining the advantages of traditional relay assisted communications. The improvement in spectrum efficiency in TWR is achieved by applying self-interference cancelation at each source node and extracting the desired information from the received network-coded messages. In this case, the accuracy of the self-interference cancelation process significantly affects the performance of TWR systems. Moreover, when using the popular amplify-and-forward (AF) relaying strategy, the accuracy of self-interference cancelation process is highly dependent on the precision of the channel estimation process. Thus, obtaining highly accurate channel state information (CSI) becomes more important in TWR systems compared to traditional one-way relaying systems [1].

On the other hand, the multi-input multi-output (MIMO) technique can be introduced to TWR systems to further improve transmission reliability and bandwidth efficiency. One efficient way to realize such performance improvement is to exploit the estimated CSI for the application of source and relay precoding [2]. Therefore, in MIMO TWR systems, in addition to affecting the performance of self-interference cancelation, inaccurate channel estimation also imposes a negative effect on the precoder design.

This work is supported by the NSF of China under grant 61322102 and the National 973 project under grant 2012CB316100.

Some contributions have been reported for the matrix-form channel estimation of the MIMO two-way relay system. For example, in [3], a MIMO channel estimator is proposed that uses the self-interference as a training sequence to estimate the channel matrices corresponding to the broadcasting (BC) phase. In [4], an LS estimator is used to obtain the cascaded channel matrices corresponding to the BC and the multiple access (MAC) phases. Very recently, the authors in [5], [6] investigate the minimum mean-square-error (MMSE) channel estimation for TWR systems based on a correlated Gaussian MIMO channel model. In particular, in [5], the cascaded channel matrices for AF TWR systems are estimated and the training sequences at the two source nodes are optimized. Different from [5], the authors in [6] aim to estimate the individual channel matrices for each link.

In this paper, similar to [5], [6], we study channel estimation for correlated MIMO TWR systems by considering the Kronecker-structured channel model. However, *unlike* [5], [6], we take into account the interference from the nearby users. Thus, in this model, the disturbances at the source nodes and the relay node consist of both noise and interference. Note that the considered colored estimation environment may be more practical for applications in today's more densely deployed wireless networks. Although channel estimation of point-to-point MIMO systems in colored environments has been studied in [7], [8], to the best of our knowledge, this topic has not been addressed in the TWR scenario.

To enhance TWR performance, we seek to estimate the individual channel matrices corresponding to source-to-relay and relay-to-source links. To this end, we propose to decompose the bidirectional transmission of the TWR system into the MAC and BC phases, in which the channel estimations are performed at the relay node and two source nodes, respectively. The proposed estimation scheme is different from the ones in [5], [6], where the channel estimation is assumed to only be conducted at the user ends. As such, our proposed estimation scheme can more efficiently support precoding at the relay since it requires significantly less feedback overhead [2]. Based on the proposed estimation scheme, we derive the optimal linear MMSE (LMMSE) estimator for each phase. Next, the corresponding training design problems are formulated with the aim of minimizing the total MSE of channel estimation process for each phase. The training design problems considered here are different from that of

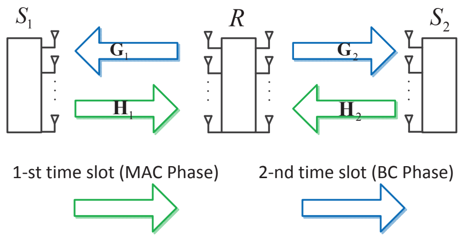


Fig. 1. An illustration of MIMO two-way relay system.

[5], [6], since we take into account the effect of colored disturbances caused by interference at the relay node and user ends. Moreover, the training design scenarios for point-to-point systems in [7], [8] are different from the scenario under consideration in this paper, since our proposed training sequence design is optimized to simultaneously enhance channel estimation accuracy over both links in the BC and MAC phases. Although, for the general scenario, it is difficult to derive the optimal training sequence structures as in [5]–[8], we propose two iterative design algorithms to solve the training design problems. These algorithms are verified to converge quickly to the near optimal solution and to not be sensitive to the initialization process. For some special cases, where the covariance matrices of the channels or disturbances have specific forms, the optimal structures of the training sequences are derived. The minimum required training lengths for channel estimation in both the MAC and BC phases are also derived.

Notations: $\otimes$ denotes the Kronecker operator. $\mathrm{vec}(\cdot)$ signifies the matrix vectorization operator. $\Re(z)$ denotes the real part of z. $\mathrm{Blkdiag}(\mathbf{A}_0, \mathbf{A}_1, \cdots, \mathbf{A}_{N-1})$ denotes a block diagonal matrix with $\mathbf{A}_0, \mathbf{A}_1, \cdots, \mathbf{A}_{N-1}$ as its diagonal matrices.

II. System Model

Consider a TWR system where two source nodes S_1 and S_2 equipped with N_1 and N_2 antennas, respectively, exchange information with each other through an M-antenna relay node R as shown in Fig. 1. The channel matrices from S_1 and S_2 to the relay are denoted by $\mathbf{H}_1$ and $\mathbf{H}_2$, respectively, and the channel matrices from the relay to S_1 and S_2 are denoted by $\mathbf{G}_1$ and $\mathbf{G}_2$, respectively. We assume that the individual channels are estimated within two phases, i.e., $\mathbf{H}_1$ and $\mathbf{H}_2$ are estimated in the MAC phase by using the training signals sent from the two sources, and $\mathbf{G}_1$ and $\mathbf{G}_2$ are estimated during the BC phase by utilizing the training signal transmitted from the relay node.

The received training signals in the MAC phase can be expressed as

$$\mathbf{Y}_R = \mathbf{H}_1\mathbf{S}_1 + \mathbf{H}_2\mathbf{S}_2 + \mathbf{N}_R, \tag{1}$$

where $\mathbf{Y}_R \in \mathbb{C}^{M\times L_S}$ with L_S being the length of the source training sequences, $\mathbf{S}_i \in \mathbb{C}^{N_i\times L_S}$ denotes the training sequence at the source S_i, and $\mathbf{N}_R \in \mathbb{C}^{M\times L_S}$ represents the correlated Gaussian disturbance at the relay node including the total background noise and interference from nearby communications, which is modeled by $\mathrm{vec}(\mathbf{N}_R) \sim \mathcal{CN}(\mathbf{0}, \mathbf{K}_R)$. The channel matrix $\mathbf{H}_i \in \mathbb{C}^{M\times N_i}$ is modeled via Rayleigh fading distribution with mean zero and covariance $\mathbf{Z}_{H_i} \in \mathbb{S}^{MN_i\times MN_i}$, i.e., $\mathrm{vec}(\mathbf{H}_i) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Z}_{H_i})$. Suppose that the source S_i has the maximum power of P_i during the training phase, i.e., $\mathrm{Tr}(\mathbf{S}_i\mathbf{S}_i^H) \le P_i$.

In the BC phase, the received training signals at the two source nodes are given by

$$\mathbf{Y}_i = \mathbf{G}_i\mathbf{S}_R + \mathbf{N}_i, \ i = 1, 2, \tag{2}$$

where $\mathbf{Y}_i \in \mathbb{C}^{N_i\times L_R}$, $\mathbf{S}_R \in \mathbb{C}^{N_i\times L_R}$ denotes the training signal at the relay node R and $\mathbf{N}_i \in \mathbb{C}^{M\times L_R}$ represents the correlated Gaussian disturbance at the source S_i, which is modeled by $\mathrm{vec}(\mathbf{N}_i) \sim \mathcal{CN}(\mathbf{0}, \mathbf{K}_i)$. It is assumed that the channel matrix $\mathbf{G}_i \in \mathbb{C}^{M\times N_i}$ follows the distribution $\mathrm{vec}(\mathbf{G}_i) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Z}_{G_i})$. Here, we assume that the length of the relay training sequence is L_R. To satisfy the relay power constraint, we have $\mathrm{Tr}(\mathbf{S}_R\mathbf{S}_R^H) \le P_R$, where P_R is the maximum power at the relay node during the training phase.

In this work, we assume that the covariances of the channels $\mathbf{Z}_{H_i}$, $\mathbf{Z}_{G_i}$ and the covariances of disturbances $\mathbf{K}_R$ and $\mathbf{K}_i$, for $i = 1, 2$, are structured and known. We assume that the channel matrices take the *kronecker-structured* model, where the covariance matrices are separated between the transmitter and receiver sides. Specifically, the covariance of the channel matrices can be decomposed as $\mathbf{Z}_{H_i} = \mathbf{Z}_{t,H_i} \otimes \mathbf{Z}_{r,H}$, $\mathbf{Z}_{G_i} = \mathbf{Z}_{t,G} \otimes \mathbf{Z}_{r,G_i}$, which allows us to model the channels as

$$\mathbf{H}_i = \mathbf{C}_{r,H}\mathbf{W}_{H_i}\mathbf{C}_{t,H_i}^T, \ \mathbf{G}_i = \mathbf{C}_{r,G_i}\mathbf{W}_{G_i}\mathbf{C}_{t,G}^T,$$

where $\mathbf{C}_{a,b}$ satisfies $\mathbf{Z}_{a,b} = \mathbf{C}_{a,b}\mathbf{C}_{a,b}^H$ with $a \in \{r, t\}, b \in \{H, H_1, H_2, G, G_1, G_2\}$. In (3), $\mathbf{W}_{H_i}$ and $\mathbf{W}_{G_i}$ are the unknown matrices and each element is a zero mean and unit variance Gaussian variable.

The structured disturbance covariances $\mathbf{K}_i$, for $i \in \{R, 1, 2\}$, are assumed to be modeled by [7]–[9]

$$\mathbf{K}_i = \mathbf{K}_{q,i} \otimes \mathbf{K}_{r,i}, \ i = R, 1, 2, \tag{3}$$

where $\mathbf{K}_{q,1}$, $\mathbf{K}_{q,2}$, $\mathbf{K}_{q,R}$ denote the temporal covariance matrix and $\mathbf{K}_{r,1}$, $\mathbf{K}_{r,2}$, and $\mathbf{K}_{r,R}$ denote the received spatial covariance matrix. Moreover, it is assumed that $\mathbf{K}_{r,1}$, $\mathbf{K}_{r,2}$, and $\mathbf{K}_{r,R}$ share the same eigenvectors with $\mathbf{Z}_{r,G_1}$, $\mathbf{Z}_{r,G_2}$ and $\mathbf{Z}_{r,H}$, respectively. This assumption is valid when the disturbances are either spatially uncorrelated or share the same spatial structure as the channels [8], [9]. The singular value decomposition (SVD) of $\mathbf{Z}_{t,H_i}$, $\mathbf{Z}_{r,H}$, $\mathbf{Z}_{t,G}$, and $\mathbf{Z}_{r,G_i}$ are given by

$$\begin{aligned} \mathbf{Z}_{a,b} =& \mathbf{U}_{a,b}\mathbf{\Sigma}_{a,b}\mathbf{U}_{a,b}^H, \\ & a \in \{r, t\}, b \in \{H, H_1, H_2, G, G_1, G_2\}, \end{aligned} \tag{4}$$

where $\mathbf{U}_{a,b}$ denotes the unitary eigenvector matrix and $\mathbf{\Sigma}_{a,b}$ is a diagonal matrix with $[\mathbf{\Sigma}_{a,b}]_{n,n} = \sigma_{a,b,n}$ being the n-th eigenvalue of $\mathbf{Z}_{a,b}$. Accordingly, the SVD decomposition of $\mathbf{C}_{a,b}$ is denoted by $\mathbf{C}_{a,b} = \mathbf{U}_{a,b}\mathbf{\Sigma}_{a,b}^{1/2}\tilde{\mathbf{U}}_{a,b}^H$ with $\tilde{\mathbf{U}}_{a,b}$ repesting a unitrary matrix. The SVD decomposition of $\mathbf{K}_{q,i}$ and $\mathbf{K}_{r,i}$ is denoted by

$$\mathbf{K}_{a,b} = \mathbf{V}_{a,b}\mathbf{\Delta}_{a,b}\mathbf{V}_{a,b}^H, \ a \in \{r, t\}, b \in \{1, 2, R\}, \tag{5}$$

where $\mathbf{V}_{a,b}$ denotes the unitary eigenvector matrix, $\mathbf{\Delta}_{a,b}$ is a diagonal matrix with $[\mathbf{\Delta}_{a,b}]_{n,n} = \delta_{a,b,n}$ being the n-th eigenvalue of $\mathbf{K}_{a,b}$.

III. Channel Estimation for Two-Way Relay Systems

For the estimation in the MAC phase, we rewrite (1) as

$$\begin{aligned}\mathbf{Y}_R &= \mathbf{C}_{r,H}\mathbf{W}_{H_1}\mathbf{C}_{t,H_1}^T\mathbf{S}_1 + \mathbf{C}_{r,H}\mathbf{W}_{H_2}\mathbf{C}_{t,H_2}^T\mathbf{S}_2 + \mathbf{N}_R \\ &= \mathbf{C}_{r,H}\mathbf{W}_H\mathbf{C}_{t,H}^T\mathbf{S} + \mathbf{N}_R,\end{aligned} \quad (6)$$

where $\mathbf{W}_H \triangleq [\mathbf{W}_{H_1}, \mathbf{W}_{H_2}]$, $\mathbf{C}_{t,H}^T \triangleq \text{Blkdiag}(\mathbf{C}_{t,H_1}^T, \mathbf{C}_{t,H_2}^T)$, and $\mathbf{S} \triangleq [\mathbf{S}_1^T, \mathbf{S}_2^T]^T$. Vectorizing $\mathbf{Y}_R$ and applying the identity $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$, we can rewrite (6) as

$$\mathbf{y}_R = \left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)\mathbf{w}_H + \mathbf{n}_R, \quad (7)$$

where $\mathbf{y}_R \triangleq \text{vec}(\mathbf{Y}_R)$, $\mathbf{w}_H \triangleq \text{vec}(\mathbf{W}_H)$ and $\mathbf{n}_R \triangleq \text{vec}(\mathbf{N}_R)$. The estimation of $\mathbf{w}_H$ based on the LMMSE criterion can be obtained as $\hat{\mathbf{w}}_H = \mathbf{T}_R\mathbf{y}_R$. The estimation matrix $\mathbf{T}_R$ has the following form

$$\mathbf{T}_R = \mathfrak{R}_{w_h y_R}\mathfrak{R}_{y_R y_R}^{-1}, \quad (8)$$

where $\mathfrak{R}_{w_h y_R} \triangleq \mathcal{E}(\mathbf{w}_H\mathbf{y}_R^H) = \mathbf{C}_{t,H}^H\mathbf{S}^* \otimes \mathbf{C}_{r,H}^H$, $\mathfrak{R}_{y_R y_R} \triangleq \mathcal{E}(\mathbf{y}_R\mathbf{y}_R^H) = \left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)\left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)^H + \mathbf{K}_R = \mathbf{S}^T\mathbf{C}_{t,H}\mathbf{C}_{t,H}^H\mathbf{S}^* \otimes \mathbf{C}_{r,H}\mathbf{C}_{r,H}^H + \mathbf{K}_R$. Let us define $\mathbf{h} \triangleq \text{vec}(\mathbf{H}) = (\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H})\mathbf{w}$ with $\mathbf{H} \triangleq [\mathbf{H}_1, \mathbf{H}_2]$, the resulting estimation error, or mean-square-error (MSE), e_R can be obtained as

$$\begin{aligned}e_R =& \mathcal{E}\left(||\mathbf{h} - \hat{\mathbf{h}}||_2^2\right) = \text{Tr}\left[\mathbf{C}_{0,H}(\mathbf{w}_H - \hat{\mathbf{w}}_H)(\mathbf{w}_H - \hat{\mathbf{w}}_H)^H\right] \\ =& \text{Tr}\left[\mathbf{C}_{0,H}(\mathbf{w}_H - \mathbf{T}_R\mathbf{y}_R)(\mathbf{w}_H - \mathbf{T}_R\mathbf{y}_R)^H\right],\end{aligned}$$

where $\mathbf{C}_{0,H} \triangleq \mathbf{C}_{t,H}^H\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}^H\mathbf{C}_{r,H}$. Substituting $\mathbf{T}_R$ into e_R and using the matrix identity $(\mathbf{I} + \mathbf{AB})^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{I} + \mathbf{BA})^{-1}\mathbf{B}$, we obtain the following more compact form

$$\begin{aligned}e_R =& \text{Tr}\Big[\mathbf{C}_{0,H}\Big(\mathbf{I} + \left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)^H\mathbf{K}_R^{-1} \\ & \times \left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)\Big)^{-1}\Big].\end{aligned} \quad (9)$$

Note that since the channel estimation model in (7) is linear and Gaussian, the proposed LMMSE estimator is equivalent to the optimal MMSE estimator.[1]

IV. Training Sequence Design for MAC Phase

To design the training sequence in the MAC phase, the $\mathbf{S}_1$ and $\mathbf{S}_2$ are optimized subject to two source power constraints aiming at minimizing the total estimation MSE, i.e., e_R in (9). The corresponding optimization problem can be formulated as

$$\begin{aligned}\min_{\mathbf{S}_1,\mathbf{S}_2} \quad & e_R \text{ in } (9) \\ \text{s.t.} \quad & \text{Tr}(\mathbf{S}_i\mathbf{S}_i^H) \le P_i,\ i = 1, 2.\end{aligned} \quad (10)$$

Before solving (10), we first introduce the following lemma that deals with the minimum length of $\mathbf{S}$.[2]

Lemma 1. *To achieve an arbitrary small MSE with infinite power at the source nodes, the minimum length of the source training sequence should be set to $L_S = N_1 + N_2$. Otherwise, even with infinite power at the source nodes, the total MSE is lower bounded by $\sum_{n=1}^{M}\sigma_{r,H,n}\sum_{m=L_S+1}^{N_1+N_2}\sigma_{t,H,m}$ with $\sigma_{t,H,m}$ being the m-th eigenvalue of $\mathbf{Z}_{r,H}$ and $\mathbf{Z}_{t,H} = \text{Blkdiag}(\mathbf{Z}_{t,H_1}, \mathbf{Z}_{t,H_2})$. Moreover, for $\mathbf{K}_{q,R} = q\mathbf{I}$ and any power constraint at the source node, if the optimal solution of $\mathbf{S}$ in* (10) *has a rank r, the minimum length of source training sequence can be set to $L_S = r$.*

Although the objective function in (10) has a similar form to that of point-to-point systems, there are two power constraints in (10) that make the problem of solving this non-convex optimization problem more difficult than that of point-to-point systems in [7], [8]. Also, as we consider the colored disturbance, (10) cannot be solved as in [5], [6] by deriving the optimal structures of the training sequences. In order to proceed, we first note that e_R in (9) can be obtained by substituting (8) into (9). Thus, to make the problem tractable, we propose an iterative algorithm, which decouples the primal problem into two sub-problems and solves each of them in an alternating approach. Let us rewrite (9) as

$$\begin{aligned}\tilde{e}_R =& \text{Tr}\left[\mathbf{C}_{0,H}(\mathbf{w}_H - \mathbf{T}_R\mathbf{y}_R)(\mathbf{w}_H - \mathbf{T}_R\mathbf{y}_R)^H\right] \\ =& \text{Tr}\Big[\mathbf{C}_{0,H} - \left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)^H\mathbf{T}_R^H\mathbf{C}_{0,H}^H - \\ & \mathbf{C}_{0,H}\mathbf{T}_R\left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right) + \mathbf{C}_{0,H}\mathbf{T}_R\left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right) \\ & \times \left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)^H\mathbf{T}_R^H + \mathbf{C}_{0,H}\mathbf{T}_R\mathbf{K}_R\mathbf{T}_R^H\Big].\end{aligned}$$

Then, the optimization problem in (10) is equivalent to

$$\begin{aligned}\min_{\mathbf{T}_R,\mathbf{S}_1,\mathbf{S}_2} \quad & \tilde{e}_R \\ \text{s.t.} \quad & \text{Tr}(\mathbf{S}_i\mathbf{S}_i^H) \le P_i,\ i = 1, 2.\end{aligned} \quad (11)$$

In the first subproblem, we intend to optimize the LMMSE estimator matrix $\mathbf{T}_R$ for a given $\mathbf{S}_1$ and $\mathbf{S}_2$. Since $\mathbf{T}_R$ is not related to the power constraint, the problem simplifies to an unconstrained optimization problem given by

$$\min_{\mathbf{T}_R} \quad \tilde{e}_R. \quad (12)$$

Given that (12) is convex with respect to $\mathbf{T}_R$, by setting its gradient to zero, we obtain the optimal $\mathbf{T}_R$ as given in (8).

In the second subproblem, the training sequences $\mathbf{S}_1$ and $\mathbf{S}_2$ need to be optimized for a given $\mathbf{T}_R$ by solving the following optimization problem

$$\begin{aligned}\min_{\mathbf{S}_1,\mathbf{S}_2} \quad & \tilde{e}_R \\ \text{s.t.} \quad & \text{Tr}(\mathbf{S}_i\mathbf{S}_i^H) \le P_i,\ i = 1, 2.\end{aligned} \quad (13)$$

Next, it is shown that the optimization problem in (13) can be transformed into a convex quadratically constrained quadratic programable (QCQP) problem. To achieve this goal, we first reformulate the last term in $\tilde{e}_R$ as

$$\begin{aligned}& \text{Tr}\left[\mathbf{C}_{0,H}\mathbf{T}_R\left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)\left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)^H\mathbf{T}_R^H\right] \\ \overset{(a)}{=}& \text{Tr}\left[\mathbf{T}_R^H\mathbf{C}_{0,H}\mathbf{T}_R(\mathbf{S}^T \otimes \mathbf{I})(\mathbf{C}_{t,H}\mathbf{C}_{t,H}^H \otimes \mathbf{C}_{r,H}\mathbf{C}_{r,H}^H)(\mathbf{S}^* \otimes \mathbf{I})\right] \\ \overset{(b)}{=}& \text{vec}(\mathbf{S} \otimes \mathbf{I})^H\left(\mathbf{T}_R^H\mathbf{C}_{0,H}\mathbf{T}_R \otimes \mathbf{C}_{tr}^T\right)\text{vec}(\mathbf{S} \otimes \mathbf{I}) \\ \overset{(c)}{=}& \mathbf{s}^H\mathbf{E}^H\left(\mathbf{T}_R^H\mathbf{C}_{0,H}\mathbf{T}_R \otimes \mathbf{C}_{tr}^T\right)\mathbf{E}\mathbf{s},\end{aligned} \quad (14)$$

[1] Please refer to [10] for the details of the optimal LMMSE estimator and the MSE derivation in the BC phase.

[2] Please refer to [10] for the omitted proofs of this paper.

where $\mathbf{C}_{tr} \triangleq \mathbf{C}_{t,H}\mathbf{C}_{t,H}^H \otimes \mathbf{C}_{r,H}\mathbf{C}_{r,H}^H$, $\mathbf{s} \triangleq \text{vec}(\mathbf{S})$, $\mathbf{E} \triangleq \text{Blkdiag}(\tilde{\mathbf{E}}_{(1)}, \tilde{\mathbf{E}}_{(2)}, \cdots, \tilde{\mathbf{E}}_{(L_S)})$, $\tilde{\mathbf{E}}_{(i)} = \tilde{\mathbf{E}}$, $\tilde{\mathbf{E}} \triangleq [\bar{\mathbf{E}}_{(1)}; \bar{\mathbf{E}}_{(2)}; \cdots; \bar{\mathbf{E}}_{(M)}]$, $\bar{\mathbf{E}}_{(i)} \triangleq \text{Blkdiag}(\underbrace{\mathbf{e}_i, \mathbf{e}_i, \cdots, \mathbf{e}_i}_{N_1+N_2 \text{ elements}})$, and $\mathbf{e}_i$ is the unit vector with $e_{i,i} = 1$ and $e_{i,j} = 1, j \neq i$. In (14), Eq. (a) is obtained by using the circular property $\text{Tr}(\mathbf{AB}) = \text{Tr}(\mathbf{BA})$ and the matrix identity $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$, Eq. (b) is obtained by using the identity $\text{Tr}(\mathbf{ABCD}) = \text{vec}(\mathbf{D})^T(\mathbf{A} \otimes \mathbf{C}^T)\text{vec}(\mathbf{B}^T)$ and $(\mathbf{A} \otimes \mathbf{B})^H = \mathbf{A}^H \otimes \mathbf{B}^H$, and Eq. (c) is obtained by using $\text{vec}(\mathbf{S} \otimes \mathbf{I}) = \mathbf{Es}$. Similarly, the term $\text{Tr}\left[\mathbf{C}_{0,H}\mathbf{T}_R(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H})\right]$ can be expressed as

$$\text{Tr}\left[\mathbf{C}_{0,H}\mathbf{T}_R(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H})\right] = \text{vec}(\mathbf{C}_T)^T\mathbf{Es}, \quad (15)$$

where $\mathbf{C}_T \triangleq (\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H})\mathbf{C}_{0,H}\mathbf{T}_R$. To obtain (15), we use the fact $\text{Tr}(\mathbf{A}^T\mathbf{B}) = \text{vec}(\mathbf{A})^T\text{vec}(\mathbf{B})$. The source power constrain in (13) can be rewritten as

$$\text{Tr}(\mathbf{S}_i\mathbf{S}_i^H) = \text{Tr}(\mathbf{E}_i\mathbf{S}\mathbf{S}^H) = \mathbf{s}^H(\mathbf{I} \otimes \mathbf{E}_i)\mathbf{s}, \quad (16)$$

where $\mathbf{E}_1 \triangleq \text{Blkdiag}(\mathbf{I}_{N_1}, \mathbf{0}_{N_2 \times N_2})$ and $\mathbf{E}_2 \triangleq \text{Blkdiag}(\mathbf{0}_{N_1 \times N_1}, \mathbf{I}_{N_2})$, and the second equation in (16) is obtained by using $\text{Tr}(\mathbf{ABCD}) = \text{vec}(\mathbf{D}^T)^T(\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$. According to (14), (15), and (16), the optimization problem in (13) can be transformed into

$$\begin{aligned} \min_{\mathbf{s}} \quad & \mathbf{s}^H\mathbf{E}^H\left(\mathbf{T}_R^H\mathbf{C}_{0,H}\mathbf{T}_R \otimes \mathbf{C}_{tr}^T\right)\mathbf{Es} - 2\Re(\text{vec}(\mathbf{C}_T)^T\mathbf{Es}) \\ \text{s.t.} \quad & \mathbf{s}^H(\mathbf{I} \otimes \mathbf{E}_i)\mathbf{s} \leq P_i, \; i = 1, 2. \end{aligned} \quad (17)$$

Since both $\mathbf{E}^H\left(\mathbf{T}_R^H\mathbf{T}_R \otimes \mathbf{C}_{tr}^T\right)\mathbf{E}$ and $\mathbf{I} \otimes \mathbf{E}_i$ are positive semidefinite matrices, we conclude that the optimization problem in (17) is a convex QCQP problem, which can be easily solved by applying the available software package.

In summary, we outline the proposed iterative training design algorithm as follows:

Algorithm 1

- **Initialize** $\mathbf{S}_1$, $\mathbf{S}_2$
- **Repeat**
  - Update the LMMSE estimator matrix $\mathbf{T}_R$ using (8) for fixed $\mathbf{S}_1$ and $\mathbf{S}_2$;
  - For fixed $\mathbf{T}_R$, solve the convex QCQP problem in (17) to get the optimal $\mathbf{S}_1$ and $\mathbf{S}_2$;
- **Until** The difference between the MSE from one iteration to another is smaller than a certain predetermined threshold.

THEOREM 1. *The proposed iterative precoding design in Algorithm 1 is convergent and the limit point of the iteration is a stationary point of* (11).

To this point, it is shown that the joint source training design can be solved via *Algorithm 1*. In the following, we illustrate that for some special cases, the optimal solution of (10) can be obtained in closed-form.

A. When $\mathbf{K}_{r,R} = \mathbf{Z}_{r,H}$

We first consider the case with $\mathbf{K}_{r,R} = \mathbf{Z}_{r,H}$, which corresponds to a scenario where the disturbance is dominated by the interference from neighboring users as shown in [9]. Accordingly, the LMMSE estimator in (8) can be rewritten as

$$\begin{aligned} \mathbf{T}_R &= \left[\mathbf{C}_{t,H}^H\mathbf{S}^* \otimes \mathbf{C}_{r,H}^H\right] \\ &\quad \times \left[\mathbf{S}^T\mathbf{C}_{t,H}\mathbf{C}_{t,H}^H\mathbf{S}^* \otimes \mathbf{C}_{r,H}\mathbf{C}_{r,H}^H + \mathbf{K}_R\right]^{-1} \\ &= \underbrace{\mathbf{C}_{t,H}^H\mathbf{S}^*\left(\mathbf{S}^T\mathbf{C}_{t,H}\mathbf{C}_{t,H}^H\mathbf{S}^* + \mathbf{K}_{q,R}\right)^{-1}}_{\triangleq \mathbf{T}_{R,1}} \otimes \mathbf{C}_{r,H}^{-1}, \end{aligned}$$

which further leads to $\mathbf{T}_R^H\mathbf{C}_{0,H}\mathbf{T}_R = (\mathbf{T}_{R,1} \otimes \mathbf{C}_{r,H}^{-1})^H (\mathbf{C}_{t,H}^H\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}^H\mathbf{C}_{r,H})(\mathbf{T}_{R,1} \otimes \mathbf{C}_{r,H}^{-1}) = \mathbf{T}_{R,1}^H\mathbf{C}_{t,H}^H\mathbf{C}_{t,H} \mathbf{T}_{R,1} \otimes \mathbf{I}$, and

$$\begin{aligned} &\text{Tr}\left[\mathbf{C}_{0,H}\mathbf{T}_R\left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)\left(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H}\right)^H\mathbf{T}_R^H\right] \\ &\overset{(a)}{=} \text{Tr}(\mathbf{Z}_{r,H})\text{Tr}\left[\mathbf{S}^T\mathbf{C}_{t,H}\mathbf{C}_{t,H}^H\mathbf{S}^*\mathbf{T}_{R,1}^H\mathbf{C}_{t,H}^H\mathbf{C}_{t,H}\mathbf{T}_{R,1}\right] \\ &\overset{(b)}{=} \text{Tr}(\mathbf{Z}_{r,H})\sum_{i=1}^{2}\text{Tr}\left(\mathbf{S}_i^T\mathbf{Z}_{t,H_i}\mathbf{S}_i^*\mathbf{T}_{R,1}^H\mathbf{C}_{t,H}^H\mathbf{C}_{t,H}\mathbf{T}_{R,1}\right). \end{aligned} \quad (18)$$

In (18), Eq. (a) is obtained by using $\text{Tr}(\mathbf{A} \otimes \mathbf{B}) = \text{Tr}(\mathbf{A})\text{Tr}(\mathbf{B})$, and Eq. (b) is obtained as $\mathbf{C}_{t,H}$ is a block diagonal matrix as shown in (6). In addition, the term $\text{Tr}\left[\mathbf{C}_{0,H}\mathbf{T}_R(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H})\right]$ can be reexpressed as

$$\begin{aligned} &\text{Tr}\left[\mathbf{C}_{0,H}\mathbf{T}_R(\mathbf{S}^T\mathbf{C}_{t,H} \otimes \mathbf{C}_{r,H})\right] \\ &= \text{Tr}\left[(\mathbf{S}^T\mathbf{C}_{t,H}\mathbf{C}_{t,H}^H\mathbf{C}_{t,H}\mathbf{T}_{R,1}) \otimes \mathbf{C}_{r,H}\mathbf{C}_{r,H}^H\right] \\ &= \text{Tr}(\mathbf{Z}_{r,H})\sum_{i=1}^{2}\text{Tr}\left(\mathbf{S}_i^T\mathbf{Z}_{t,H_i}\mathbf{C}_{t,H_i}\mathbf{T}_{R,1,i}\right), \end{aligned} \quad (19)$$

where $\mathbf{T}_{R,1,1} \triangleq \mathbf{T}_{R,1}(1 : N_1, :)$ and $\mathbf{T}_{R,1,2} \triangleq \mathbf{T}_{R,1}(N_1 + 1 : N_1 + N_2, :)$. Based on (18) and (19), (13) is equivalent to the following optimization problem

$$\begin{aligned} \min_{\mathbf{S}_1, \mathbf{S}_2} \quad & \sum_{i=1}^{2}\left\{\text{Tr}\left(\mathbf{S}_i^T\mathbf{Z}_{t,H_i}\mathbf{S}_i^*\mathbf{T}_{R,1}^H\mathbf{C}_{t,H}^H\mathbf{C}_{t,H}\mathbf{T}_{R,1}\right) - \right. \\ & \left. \text{Tr}\left(\mathbf{S}_i^T\mathbf{Z}_{t,H_i}\mathbf{C}_{t,H_i}\mathbf{T}_{R,1,i}\right) - \text{Tr}\left(\mathbf{T}_{R,1,i}^H\mathbf{C}_{t,H_i}^H\mathbf{Z}_{t,H_i}^H\mathbf{S}_i^*\right)\right\} \\ \text{s.t.} \quad & \text{Tr}(\mathbf{S}_i\mathbf{S}_i^H) \leq P_i, \; i = 1, 2. \end{aligned} \quad (20)$$

We note that compared to (13), (20) has a simpler form and can be solved in closed-form via the KKT conditions. Thus, the optimal $\mathbf{s}_i \triangleq \text{vec}(\mathbf{S}_i)$ can be obtained as

$$\mathbf{s}_i = \left[\mathbf{X}_{s,1} \otimes \mathbf{X}_{s,2,i} + \lambda_i\mathbf{I}\right]^{-1}\mathbf{x}_{s,3,i}, \quad (21)$$

where λ_i is the lagrangian multiplier associated with the power constraint at S_i, $\mathbf{X}_{s,1} \triangleq \mathbf{T}_{R,1}^H\mathbf{C}_{t,H}^H\mathbf{C}_{t,H}\mathbf{T}_{R,1}$, $\mathbf{X}_{s,2,i} \triangleq \mathbf{Z}_{t,H_i}^T$, and $\mathbf{x}_{s,3,i} \triangleq \text{vec}\left(\left(\mathbf{T}_{R,1,i}^H\mathbf{C}_{t,H_i}^H\mathbf{Z}_{t,H_i}^H\right)^T\right)$. The optimal λ_i in (21) can be zero or should be chosen to activate the power constraint. For the case where $\lambda_i \neq 0$, the following lemma is introduced.

LEMMA 2. *The function* $g(\lambda_i) = \text{Tr}\left(\mathbf{S}_i\mathbf{S}_i^H\right) = \text{Tr}\left(\mathbf{s}_i\mathbf{s}_i^H\right)$, *with* $\mathbf{s}_i$ *defined above* (21), *is monotonically decreasing with respect to* λ_i *and the optimal* λ_i *is upper-bounded by* $\sqrt{\frac{\sigma_{s,3,i}}{P_i}} - \sigma_{s,min,i}$. *Here,* $\sigma_{s,3,i}$ *denotes the smallest eigenvalue of* $\mathbf{X}_{s,1} \otimes \mathbf{X}_{s,2,i}$ *and* $\sigma_{s,min,i} = ||\mathbf{x}_{s,3,i}||_2^2$.

By applying *Lemma 2*, the optimal λ_i that meets the condition $\mathrm{Tr}\left(\mathbf{S}_i\mathbf{S}_i^H\right) = P_i$ can be readily obtained via the bisection search algorithm.

B. When $\mathbf{K}_{q,R} = q\mathbf{I}$

This scenario corresponds to a practical case, where the disturbance consists of both the additive white Gaussian noise and the temporally uncorrelated interference. To proceed, the total MSE is first rewritten as

$$\begin{aligned} e_R =& \mathrm{Tr}\left[\mathbf{C}_{0,H}\left(\mathbf{I}+\frac{1}{q}\mathbf{C}_{t,H}^H\mathbf{S}^*\mathbf{S}^T\mathbf{C}_{t,H}\otimes\mathbf{C}_{r,H}^H\mathbf{K}_{r,R}^{-1}\mathbf{C}_{r,H}\right)^{-1}\right] \\ =& \sum_{n=1}^{M}\sigma_{r,H,n}\mathrm{Tr}\left[\left(\mathbf{Z}_{t,H}^{-1}+\alpha_n\mathbf{S}^*\mathbf{S}^T\right)^{-1}\right], \end{aligned}$$

where $\alpha_n \triangleq \dfrac{\sigma_{r,H,n}}{q\delta_{r,R,n}}$. Subsequently, the optimization problem in (10) can be rewritten as

$$\begin{aligned} \min_{\mathbf{S}_1,\mathbf{S}_2} \quad & \sum_{n=1}^{M}\sigma_{r,H,n}\mathrm{Tr}\left[\left(\mathbf{Z}_{t,H}^{-1}+\alpha_n\mathbf{S}^*\mathbf{S}^T\right)^{-1}\right] \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{E}_i\mathbf{S}^*\mathbf{S}^T) \le P_i,\ i=1,2 \end{aligned} \tag{22}$$

where $\mathbf{E}_i$ is defined in (16). Although the optimization problem in (22) is non-convex with respect to $\mathbf{S}_i$, it is noted that one may optimize (22) with respect to the positive semidefinite matrix $\mathbf{S}^*\mathbf{S}^T$ instead of the training sequence $\mathbf{S}_i$. Accordingly, after solving for the optimum $\mathbf{S}^*\mathbf{S}^T$, the solution can be decomposed to obtain $\mathbf{S}_i$. This approach is preferable since in (22), the objective function and the constraint both depend on $\mathbf{S}^*\mathbf{S}^T$ and not $\mathbf{S}_i$. Hence, by defining $\mathbf{Q}_S \triangleq \mathbf{S}^*\mathbf{S}^T$, the following equivalent problem can be obtained

$$\begin{aligned} \min_{\mathbf{Q}_S\succeq\mathbf{0}} \quad & \sum_{n=1}^{M}\sigma_{r,H,n}\mathrm{Tr}\left[\left(\mathbf{Z}_{t,H}^{-1}+\alpha_n\mathbf{Q}_S\right)^{-1}\right] \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{E}_i\mathbf{Q}_S) \le P_i,\ i=1,2. \end{aligned} \tag{23}$$

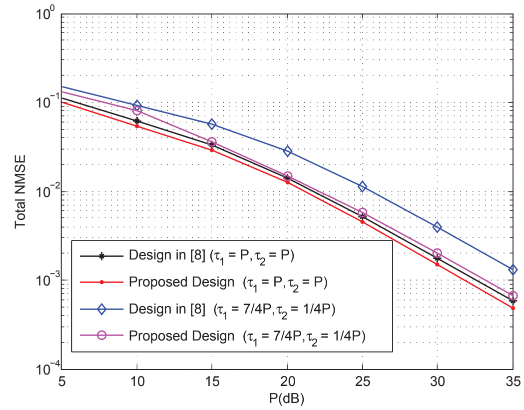
THEOREM 2. *The optimization problem in* (23) *is convex with respect to the positive semidefinite matrix* $\mathbf{Q}_S$.

Next, it is shown that the optimization problem in (23) can be solved by transforming it into a semidefinite programming (SDP) problem. By introducing the variables $\mathbf{X}_n$, the problem in (23) can be rewritten in an equivalent form as

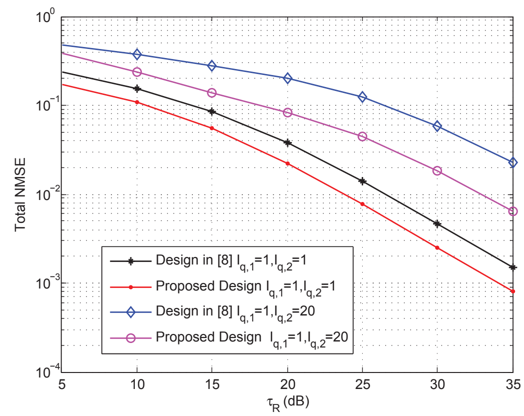
$$\begin{aligned} \min_{\mathbf{Q}_S\succeq\mathbf{0},\mathbf{X}_n} \quad & \sum_{n=1}^{M}\sigma_{r,H,n}\mathrm{Tr}\left(\mathbf{X}_n\right) \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{E}_i\mathbf{Q}_S) \le P_i,\ i=1,2 \\ & \left(\mathbf{Z}_{t,H}^{-1}+\alpha_n\mathbf{Q}_S\right)^{-1} \preceq \mathbf{X}_n, \forall n \end{aligned} \tag{24}$$

By using the Schur complement, (24) can be further transformed into the following SDP problem

$$\begin{aligned} \min_{\mathbf{Q}_S\succeq\mathbf{0},\mathbf{X}_n} \quad & \sum_{n=1}^{M}\sigma_{r,R,n}\mathrm{Tr}\left(\mathbf{X}_n\right) \\ \text{s.t.} \quad & \mathrm{Tr}(\mathbf{E}_i\mathbf{Q}_S) \le P_i,\ i=1,2 \\ & \begin{bmatrix} \mathbf{Z}_{t,H}^{-1}+\alpha_n\mathbf{Q}_S & \mathbf{I} \\ \mathbf{I} & \mathbf{X}_n \end{bmatrix} \succeq \mathbf{0},\ \forall n \end{aligned} \tag{25}$$



(a) MAC phase: $d_{t,H_1} = 1.5$, $d_{t,H_2} = 1.8$, $d_{r,G} = 1.3$, $\eta_{q,R} = 0.9$ and $I_{q,R} = 1$.



(b) BC phase: $d_{t,G} = 1.9$, $d_{r,G_1} = 1.95$, $d_{r,G_2} = 0.3$, $\eta_{q,1} = 0.9$, $\eta_{q,2} = -0.9$.

Fig. 2. Performance comparison with the existing design in [8].

By solving the SDP problem in (25), the optimal solution to the optimization problem in (23) is obtained. However, this numerical method of solving this optimization problem has a relatively high computational complexity. To obtain the optimal structure of $\mathbf{S}_i$ and gain a better understanding of the optimization in (22), the following theorem is introduced.

THEOREM 3. *With the minimum training sequence length requirement* $L_S \ge N_1 + N_2$*, the optimal training sequence* $\mathbf{S}_i$ *in* (22) *should satisfy the condition* $\mathbf{S}_1^*\mathbf{S}_2^T = \mathbf{0}$*. In addition, the optimal* $\mathbf{S}_i$ *has a form of* $\mathbf{S}_i = \mathbf{U}_{t,H_i}^*\mathbf{\Sigma}_{s_i}\mathbf{V}_{s_i}^H$*, where* $\mathbf{V}_{s_i}$ *is chosen such that* $\mathbf{V}_{s_1}^H\mathbf{V}_{s_2} = \mathbf{0}$ *and* $\mathbf{\Sigma}_{s_i}$ *is a diagonal eigenvalue matrix with* $[\mathbf{\Sigma}_{s_i}]_{n,n} = \sigma_{s_i,n}$*.* $\mathbf{\Sigma}_{s_i}$ *can be obtained by solving the following water-filling problem of* $\sum_{n=1}^{M}\frac{\alpha_n\sigma_{r,H,n}\sigma_{t,H_i,m}^2}{(1+\alpha_n\sigma_{t,H_i,m}\sigma_{s_i,m}^2)^2} = \lambda_i$*, where the optimal* λ_i *should be selected such that* $\sum_{m=1}^{N_i}\sigma_{s_i,m}^2 = P_i$*.*

Remark 1: Due to the space limitation, we omit the training design for the BC phase. The details can be found in [10].

V. SIMULATION RESULTS

The total normalized MSE (NMSE), defined as either $\frac{1}{M(N_1+N_2)}\sum_{i=1}^{2}\mathbb{E}\{||\mathbf{H}_i - \hat{\mathbf{H}}_i||_F^2\}$ or $\frac{1}{M(N_1+N_2)}\sum_{i=1}^{2}\mathbb{E}\{||\mathbf{G}_i - \hat{\mathbf{G}}_i||_F^2\}$, is utilized to illustrate the

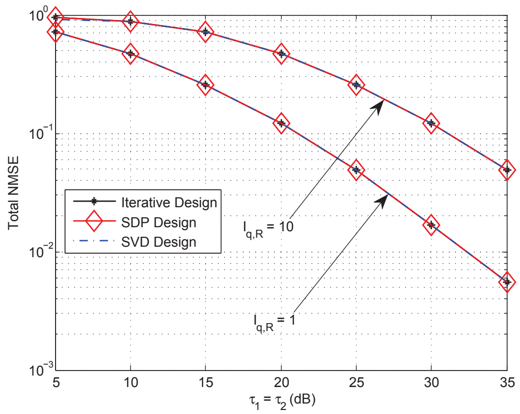


Fig. 3. Performance illustration of the case with $\mathbf{K}_{q,R} = q\mathbf{I}$ for MAC phase channel estimation.

performance of the proposed algorithms. In all simulations, the channel covariances are assumed to have the following structures $[\mathbf{Z}_{t,b}]_{n,m} = z_{t,b}J_0(d_{t,b}|n-m|),\ b \in \{H_1, H_2, G\}$, and $[\mathbf{Z}_{r,b}]_{n,m} = z_{r,b}J_0(d_{r,b}|n-m|),\ b \in \{H, G_1, G_2\}$, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, $d_{t,b}$ and $d_{r,b}$ are proportional to the carrier frequency and the antenna separation vectors at the transmitter and the receiver, respectively [7]. Moreover, the scalars $z_{t,b}$ and $z_{r,b}$ are normalization factors such that $\mathrm{Tr}(\mathbf{Z}_{t,H_i}) = N_i$, $\mathrm{Tr}(\mathbf{Z}_{r,H}) = M$, $\mathrm{Tr}(\mathbf{Z}_{t,G}) = M$ and $\mathrm{Tr}(\mathbf{Z}_{r,G_i}) = N_i$. The temporal covariance of the disturbance is assumed to be modeled via a first order autoregressive (AR) filter, i.e., $[\mathbf{K}_{q,b}]_{n,m} = I_{q,b}k_{q,b}\eta_{q,b}^{|n-m|}$ for $b \in \{1, 2, R\}$ [7]. Here, the scalar $k_{q,b}$ is a normalization factor similar to $\mathbf{Z}_{t,b}$ and $\mathbf{Z}_{r,b}$. Moreover, $I_{q,b}$ indicates the strength of the interference from the nearby users. For simplicity, the lengths of the source and relay training sequences are assumed to be $L_S = N_1 + N_2$ and $L_R = M$, respectively. The sum power at the two sources are assumed to be $P_1 + P_2 = 2P$. If not specified otherwise, we assume that $N_1 = N_2 = M = 3$.

In Fig. 2, we compare the total NMSE of the proposed iterative training design algorithm with that of [8], which is intended for point-to-point systems. The plots in Figs. 2(a) and 2(b) illustrate that compared to the approach in [8], the proposed training design can significantly improve the accuracy of channel estimation in TWR systems. This gain is even more pronounced when the two source nodes operate at different transmit power levels during the MAC phase and when the strengths of the disturbances at the two source nodes are asymmetric, i.e., $I_{q,1} \neq I_{q,2}$ during the BC phase. This can be mainly attributed to the fact that the proposed training design algorithm, i.e., *Algorithm 1*, takes into account the temporal correlation of the disturbances at the relay node in the MAC phase, while ensuring that the training sequences transmitted from the relay simultaneously match the channels corresponding to relay-to-source links during the BC phase.

In Fig. 3, the performance of the proposed training sequence design algorithms and channel estimators in the MAC phase for $\mathbf{K}_{q,R} = q\mathbf{I}$ is demonstrated. Three training sequence design approaches are taken into consideration: 1) The iterative design based on *Algorithm 1*; 2) The SDP design based on (25); and 3) The SVD design based on *Theorem 3*. As shown in *Theorem 2*, in this case, the optimization problem for finding the optimal training sequences is convex. Hence, it is well-known that both the SDP and the SVD design schemes can achieve optimal channel estimation performance. This outcome is also verified by the results in Fig. 3. However, it is interesting to note that the proposed iterative algorithm can also achieve optimal performance, which further verifies its effectiveness for designing the training sequences.

VI. Conclusions

In this paper, we considered the channel estimation for MIMO TWR systems. The impact of the interference from neighboring devices and the effect of antenna correlations on the design of training sequences and channel estimation performance were taken into consideration. To realize channel estimation of four individual links, we proposed to decompose the bidirectional link into MAC phase and BC phase, and conducted the channel estimation in each phase. The optimal LMMSE estimators were first derived. Then the corresponding training design problems were formulated and solved by using different optimization skills.

References

[1] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.

[2] R. Wang and M. Tao, "Joint source and relay precoding designs for MIMO two-way relaying based on MSE criterion," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1352–1365, March 2012.

[3] J. Zhao, M. Kuhn, A. Wittneben, and G. Bauch, "Self-interference aided channel estimation in two-way relaying systems," in *IEEE GLOBECOM 2008*, New Orleans, USA, Dec. 2008

[4] T.-H. Pham, Y.-C. Liang, A. Nallanathan, and H. K. Garg, "Optimal training sequences for channel estimation in bi-directional relay networks with multiple antennas," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 474–479, Feb. 2010.

[5] D.-H. Kim, M. Ju, and H.-M. Kim, "Optimal training signal design for estimation of correlated MIMO channels in two-way amplify-and-forward relay systems," *IEEE Commun. Lett.*, vol. 17, no. 3, pp. 491–494, March 2013.

[6] C. Chiong, Y. Rong, and Y. Xiang, "Channel training algorithms for two-way MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 3988–3998, Aug. 2013.

[7] M. Biguesh, S. Gazor, and M. Shariat, "Optimal training sequence for MIMO wireless systems in colored environments," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3144–3153, Aug. 2009.

[8] E. Björnson and B. E.Bjornson, "A framework for training-based estimation in arbitrarily correlated rician MIMO channels with rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, March 2010.

[9] Y. Liu, T. Wong, and W. W. Hager, "Training signal design for estimation of correlated MIMO channels with colored interference," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1486–1497, April 2007.

[10] R. Wang, M. Tao, H. Mehrpouyan, and Yingbo Hua "Channel estimation and optimal training design for correlated MIMO two-way relay systems in colored environment," *https://arxiv.org/submit/1025766.*, July 2014.