Blind Channel Estimation and Signal Retrieving for MIMO Relay Systems

Choo W. R. Chiong, Yue Rong, and Yong Xiang

Abstract

In this paper, we propose a blind channel estimation and signal retrieving algorithm for two-hop multiple-input multiple-output (MIMO) relay systems. This new algorithm integrates two blind source separation (BSS) methods to estimate the individual channel state information (CSI) of the source-relay and relay-destination links. In particular, a first-order Z-domain precoding technique is developed for the blind estimation of the relay-destination channel matrix, where the signals received at the relay node are pre-processed by a set of precoders before being transmitted to the destination node. With the estimated signals at the relay node, we propose an algorithm based on the constant modulus and signal mutual information properties to estimate the source-relay channel matrix. Compared with training-based MIMO relay channel estimation approaches, the proposed algorithm has a better bandwidth efficiency as no bandwidth is wasted for sending the training sequences. Numerical examples are shown to demonstrate the performance of the proposed algorithm.

Index Terms

Blind channel estimation, MIMO relay, Z-domain, constant modulus, mutual information.

I. INTRODUCTION

In an effort to provide reliable and high rate wireless communications, multiple-input multiple-output (MIMO) relay communication systems have attracted great research interests in the last decade [1]-[3]. For the MIMO relay systems in [1]-[3], the knowledge of the instantaneous channel state information (CSI) is necessary for the retrieval of the source signals at the destination node. The individual instantaneous CSI for both the source-relay and relay-destination links is also important for the optimization of MIMO relay systems through precoding matrices design and power allocation [1]-[5]. However, the instantaneous CSI is unknown in real wireless communication systems, and thus, has to be estimated at the destination node.

One of the possible solutions is by transmitting known training sequences to assist the estimation of the instantaneous CSI [6]-[13]. In [6], a channel estimation algorithm based on the least-squares (LS) fitting is proposed for MIMO relay systems. The performance of the algorithm in [6] is further analyzed and improved by using the weighted least-squares (WLS) fitting in [7]. A tensor-based channel estimation algorithm is developed in [8] for a two-way MIMO relay system. Since the algorithm in [8] exploits the channel reciprocity in a two-way relay system, its application in one-way MIMO relay systems is not straightforward. A superimposed training based channel estimation algorithm has been developed recently for orthogonal frequency-division multiplexing (OFDM) modulated relay systems in [9]. A two-stage linear minimum mean-squared error (LMMSE)-based channel training algorithm was proposed in [10]. The source-relay link CSI estimation in [10] was improved in [11] by taking into account the mismatch between the estimated and true CSI of the relay-destination link. In [12], a superimposed channel training algorithm for two-way MIMO relay systems was proposed, where the channel estimation is done in one stage through superimposing a training sequence at the relay node. A parallel factor (PARAFAC) analysis based MIMO relay channel estimation algorithm was developed in [13].

The main drawback of the training-based channel estimation algorithms is the high cost involved in sending the training sequences, considering the limited bandwidth available for wireless communication. Moreover, in some applications such as asynchronous wireless network and message interception, training-based algorithms are unrealistic and not suitable for implementation [14], [15]. In these applications, blind channel estimation techniques, which do not require training sequences, become important. Recursive least squares (RLS) and least mean squares (LMS) subspace-based adaptive algorithms were proposed in [16] for blind channel estimation in code-division multiple access (CDMA) systems. A subspace-based blind channel estimation algorithm with reduced time averaging was proposed in [17] for MIMO-OFDM systems. However, the algorithms in [16] and [17] were developed for point-to-point (single-hop) communication systems, and the extension to MIMO relay systems is not straightforward. A blind channel estimation algorithm based on the deterministic maximum likelihood (DML) approach was developed in [18] for two-way relay networks with constant-modulus (CM) signaling. In [19], non-redundant linear precoders are applied at the source nodes to blindly estimate the channels for two-way relay networks operating under OFDM modulation.

In this paper, we develop a blind channel estimation algorithm for two-hop MIMO relay communication systems by exploiting the link between blind source separation (BSS) and channel estimation. BSS techniques are able to separate a mixture of signals into individual source signals, without the knowledge (or little knowledge) of the source signals or the channel between the source and receiver. The proposed algorithm integrates two BSS methods to estimate the instantaneous CSI for the individual source-relay and relay-destination links. We would like to note that channel matrices of both the first-hop and second-hop are estimated at the destination node. The advantage of directly estimating both channel matrices at the destination node is to avoid sending the CSI from the relay node to the destination node [12], [13]. As the blind channel estimation algorithm we propose uses the communication data for channel estimation, unlike [10], there is no need for sending training signals from the relay node to the destination node.

In particular, we develop a first-order Z-domain precoding technique for the blind estimation of the relay-destination channel matrix using signals received at the destination node. In this algorithm, the signals received at the relay node are filtered by properly designed precoders before being transmitted to the destination node. By utilizing the Z-domain properties of the precoded signals, an estimation criterion is derived to recover the relay-destination channel matrix and signals received at the relay node. Note that in this algorithm, the order of the precoders is fixed to one, while a second-order Z-domain precoding algorithm was developed in [20] for blind separation of spatially correlated signals. Obviously, the computational complexity of the first-order precoder is smaller than that of the second-order precoder.

With the estimated received signals at the relay node, we then develop a blind channel estimation algorithm based on the constant modulus and signal mutual information (MI) properties to estimate the source-relay channel matrix. The constant modulus property of many modulated communication signals such as phase-shift keying (PSK) is exploited in this blind estimation algorithm. However, using the constant modulus property of signals alone does not guarantee the complete separation of the source signals and the channel matrix, as the constant modulus algorithm might capture the same signal even though there are multiple signal streams. To overcome this difficulty, we minimize a cost function which includes the MI of the estimated signals in addition to the constant modulus property, to ensure that all estimated signals are distinct. This algorithm does not have the problem of estimation error propagation as in [21] and [22]. A similar method was adopted in [15] for the extraction of unknown source signals, essentially in single-hop (point-to-point) MIMO wireless networks. However, in this paper, we apply this algorithm for channel estimation in dual-hop MIMO relay communication systems.

Comparing the proposed blind channel estimation algorithm with the training-based channel estimation techniques, the former one has a better bandwidth efficiency as all the bandwidth is used for the transmission of the communication signals. Simulation results show that the proposed blind channel estimation algorithm yields a better system bit-error-rate (BER) than that of the training-based algorithm at low signal-to-noise ratio (SNR) due to a better utilization of the power available at the source and relay nodes for channel estimation. We would like to note that the proposed algorithm can be applied in dual-hop MIMO relay systems with multiple distributed source nodes and multiple distributed relay nodes.

The rest of this paper is organized as follows. The system model of a three-node two-hop MIMO relay system is presented in Section II. In Section III, the first-order Z-domain precoding technique is developed to estimate the relay-destination channel matrix, while the signal MI modified constant modulus algorithm is proposed in Section IV to estimate the source-relay channel matrix. Section V shows numerical simulations to demonstrate the performance of the proposed algorithm. Finally, conclusions are drawn in Section VI.

II. SYSTEM MODEL

Let us consider a three-node two-hop MIMO communication system where the source node transmits information to the destination node through a relay node as shown in Fig. 1. The source, relay, and destination nodes are equipped with n_S, n_R, and n_D antennas, respectively. In this paper, we assume that the direct link between the source node and the destination node is sufficiently weak and thus can be ignored. This scenario occurs when the direct link is blocked by obstacles, such as tall buildings or mountains.

The communication process is completed in two time slots. In the first time slot, the source signal vector $\mathbf{s}(n) = [s_1(n), s_2(n), \cdots, s_{n_S}(n)]^T$ is transmitted from the source node, where $(\cdot)^T$ denotes the vector (matrix) transpose. The signal vector received at the relay node can be expressed as

$$\mathbf{y}_r(n) = \mathbf{H}_1 \mathbf{s}(n) + \mathbf{v}(n) \tag{1}$$

where $\mathbf{y}_r(n)$ is the $n_R \times 1$ received signal vector, $\mathbf{H}_1$ is the $n_R \times n_S$ MIMO channel matrix between the source node and the relay node, and $\mathbf{v}(n)$ is the $n_R \times 1$ noise vector at the relay node.

In the second time slot, each received signal stream in $\mathbf{y}_r(n)$ is preprocessed separately by a first-order precoder $p_i(z)$ as

$$p_i(z) = 1 - r_i z^{-1}, \quad i = 1, \cdots, n_R \tag{2}$$

where r_i is the zero of the precoder $p_i(z)$. Note that all zeros are distinct and satisfy $0 < |r_i| < 1$, for $i = 1, \cdots, n_R$, and are known at the destination node. Here $|\cdot|$ denotes the modulus of a scalar and the determinant of a matrix. From (2), the

ith precoded signal at the relay node can be written as

$$x_i(n) = p_i(z)y_{r,i}(n)$$
$$= y_{r,i}(n) - r_i y_{r,i}(n-1), \quad i = 1, \cdots, n_R \qquad (3)$$

where $y_{r,i}(n)$ is the ith element of $\mathbf{y}_r(n)$. It is worth noting that the precoding operation (3) can be readily implemented at physically distributed relay nodes, as there is no need for cooperation among different signal streams. The first-order precoding operation in (3) serves for the blind estimation of the relay-destination channel matrix, where the estimation criterion will be derived by exploiting the Z-domain properties of the precoders as shown in Section III.

The precoded signal vector $\mathbf{x}(n) = [x_1(n), x_2(n), \cdots, x_{n_R}(n)]^T$ is transmitted to the destination node, and the received signal vector at the destination node can be expressed as

$$\mathbf{y}(n) = \mathbf{H}_2 \mathbf{x}(n) + \mathbf{w}(n) \qquad (4)$$

where $\mathbf{H}_2$ is the $n_D \times n_R$ channel matrix between the relay node and the destination node and $\mathbf{w}(n) = [w_1(n), w_2(n), \cdots, w_{n_D}(n)]^T$ is the noise vector at the destination node. We assume that:

1) All noises are independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN).
2) The source signals in $\mathbf{s}(n)$ are temporally white, have constant modulus, and linearly independent with each other.
3) The noises are independent of the source signals.
4) The number of antennas at the receiving sides is equal or greater than that of the transmitting sides, i.e., $n_D \geq n_R \geq n_S$.

We would like to mention that the algorithm developed in this paper can be easily extended to MIMO relay systems with multiple source and relay nodes. With M source nodes and N relay nodes, each equipped with n_{Si} and n_{Rj} antennas, respectively, $i = 1, \cdots, M$, $j = 1, \cdots, N$, the received signal at the relay node can be rewritten as

$$\mathbf{y}_{r,j}(n) = \sum_{i=1}^{M} \mathbf{H}_{j,i} \mathbf{s}_i(n) + \mathbf{v}_j(n), \qquad j = 1, \cdots, N \qquad (5)$$

where $\mathbf{y}_{r,j}(n)$ is the $n_{R_j} \times 1$ received signal vector, $\mathbf{H}_{j,i}$ is the $n_{R_j} \times n_{S_i}$ MIMO channel matrix between the ith source node and the jth relay node, and $\mathbf{v}_j(n)$ is the $n_{R_j} \times 1$ noise vector at the jth relay node. We can rewrite (5) as

$$\mathbf{y}_r(n) = \mathbf{H}_1 \mathbf{s}(n) + \mathbf{v}(n) \qquad (6)$$

where $\mathbf{y}_r(n) \triangleq \left[\mathbf{y}_{r,1}^T(n), \mathbf{y}_{r,2}^T(n), \cdots, \mathbf{y}_{r,N}^T(n) \right]^T$, $\mathbf{s}(n) \triangleq \left[\mathbf{s}_1^T(n), \mathbf{s}_2^T(n), \cdots, \mathbf{s}_M^T(n) \right]^T$, $\mathbf{v}(n) \triangleq \left[\mathbf{v}_1^T(n), \mathbf{v}_2^T(n), \cdots, \mathbf{v}_N^T(n) \right]^T$, and

$$\mathbf{H}_1 \triangleq \begin{bmatrix} \mathbf{H}_{1,1} & \cdots & \mathbf{H}_{1,M} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{N,1} & \cdots & \mathbf{H}_{N,M} \end{bmatrix}.$$

Equation (6) is equivalent to (1), and the same analysis can be applied to MIMO relay systems with multiple source and relay nodes. Note that the first-order precoder $p_i(z)$ is redefined as

$$p_i(z) = 1 - r_i z^{-1}, \quad i = 1, \cdots, n_R$$

where $n_R = \sum_{j=1}^{N} n_{R_j}$ for the later case. All precoders are distinct and are known at the destination node. We would like to note that a narrow-band frequency-flat channel model is used in (1) and (4). For broadband orthogonal frequency-division multiplexing (OFDM) based communication systems such as 4G LTE, the proposed algorithms in this paper can be applied to each subcarrier of the OFDM system, where the channel fading is frequency-flat.

The model in (4) has a similar structure to the classical BSS problem. In BSS techniques, signal separation is usually achieved by exploiting the statistical properties of the source signals, either based on the higher-order statistics (HOS) or second-order statistics (SOS). Independent component analysis (ICA) is one example of the HOS-based BSS methods, and is generally applied for non-Gaussian source signals. One of the drawbacks of the HOS-based methods is the large number of data samples required for a satisfactory result. On the contrary, the number of data samples required by the SOS-based BSS methods is generally much smaller than the HOS-based BSS techniques. However, the SOS-based BSS methods usually require the source signals to be mutually uncorrelated. This limits the application of the SOS-based BSS methods in MIMO relay communication systems as the signals received at the relay node ($\mathbf{y}_r$ in (1)) are mutually correlated.

The algorithms in [18] and [19] only estimate the cascaded source-relay-destination channel in a single-input single-output (SISO) relay system, and does not provide the estimation of the individual second-hop channel in MIMO relay systems. The extension of these algorithms to MIMO relay case is not straightforward. Note that the information on the individual second-hop channel is important for the optimization of the receiver design at the destination node. For example, the MMSE receiver [3], [4] requires the second-hop channel information.

A second-order precoding-based BSS algorithm has been developed in [20] to separate mutually correlated sources. However, this algorithm might not be applicable to MIMO relay systems. This is because the algorithm in [20] does not allow any source signal to be linear combination of the other source signals (see [20], the paragraph after Assumptions A1)-A4)), while in a MIMO relay system, the signal component at the relay node (i.e. $\mathbf{H}_1\mathbf{s}(n)$) is a linear combination of the source signals $\mathbf{s}(n)$. Thus, when the noise at the relay node is sufficiently small, the signal at the relay node (1) does not satisfy the requirement of the second-order precoding method. This motivates us to develop the first-order precoding technique for blind channel estimation in MIMO relay systems as presented in the next section.

III. First-Order Z-Domain Precoding Based Channel Estimation

In this section, we develop a first-order Z-domain precoding algorithm for the blind estimation of the relay-destination channel matrix $\mathbf{H}_2$. The main idea of this approach is to preprocess the received signals at the relay node with the first-order Z-domain precoders before retransmitting them to the destination node. Then, by utilizing the Z-domain properties of the precoders, this blind channel estimation algorithm aims to find a separation matrix $\mathbf{B}_1$ to separate $\mathbf{x}(n)$ and $\mathbf{H}_2$ in (4) with only the observable output at the destination node $\mathbf{y}(n)$. Compared with [20], the first-order precoding technique requires less transmission time at the relay node and simplifies the implementation of the precoders at the relay node in practical MIMO relay systems.

Let $\mathbf{B}_1 = [\mathbf{b}_{1,1}, \mathbf{b}_{1,2}, \cdots, \mathbf{b}_{1,n_R}]$ be an $n_D \times n_R$ matrix, the desired outcome of the blind channel estimation algorithm is given by

$$\hat{\mathbf{x}}(n) = \mathbf{B}_1^H \mathbf{y}(n) = \mathbf{\Lambda}\mathbf{x}(n) + \mathbf{B}_1^H \mathbf{w}(n) \tag{7}$$

where $\hat{\mathbf{x}}(n)$ is an estimation of the precoded signal vector, $(\cdot)^H$ denotes complex conjugate transpose, and $\mathbf{\Lambda} \triangleq \mathbf{B}_1^H \mathbf{H}_2$ is a diagonal matrix of scaling ambiguity inherited in the blind estimation algorithm. Note that the permutation ambiguity usually associated with BSS methods does not exist in (7) as proved in Theorem 1 later on. Intuitively, this is due to the filtering operation (3) at the relay node before retransmitting the signals, as each signal stream in $\mathbf{y}_r(n)$ is preprocessed by a distinct precoder. The scaling ambiguity can be resolved and will be discussed later. Once the separation matrix $\mathbf{B}_1$ is obtained, $\mathbf{H}_2$ and $\mathbf{y}_r(n)$ can be efficiently estimated as shown later on. In the following subsection, we will first propose an estimation criterion by exploiting the Z-domain properties of the precoders and find the separation matrix $\mathbf{B}_1$ based on this criterion.

A. Estimation Criterion

Let us define the autocorrelation matrix of $\mathbf{y}_r(n)$ at time lag k as

$$\begin{aligned}\mathbf{C}_{\mathbf{y}_r\mathbf{y}_r}(k) &= \mathrm{E}\big[\mathbf{y}_r(n)\mathbf{y}_r(n-k)^H\big] \\ &= \mathbf{H}_1\mathbf{C}_{\mathbf{ss}}(k)\mathbf{H}_1^H + \mathbf{C}_{\mathbf{vv}}(k)\end{aligned} \tag{8}$$

where $\mathbf{C}_{\mathbf{ss}}(k) = \mathrm{E}\big[\mathbf{s}(n)\mathbf{s}(n-k)^H\big]$ and $\mathbf{C}_{\mathbf{vv}}(k) = \mathrm{E}\big[\mathbf{v}(n)\mathbf{v}(n-k)^H\big]$ are the autocorrelation matrices of $\mathbf{s}(n)$ and $\mathbf{v}(n)$, respectively, and $\mathrm{E}[\cdot]$ stands for the statistical expectation. Note that $\mathbf{C}_{\mathbf{vv}}(k) = \mathbf{0}$ for $k \neq 0$ as the noises are temporally independent. Based on (8), the power spectral matrix of $\mathbf{y}_r(n)$ is defined as

$$\mathbf{Q}_{\mathbf{y}_r\mathbf{y}_r}(z) = \sum_{k=-\infty}^{\infty} \mathbf{C}_{\mathbf{y}_r\mathbf{y}_r}(k)z^{-k}. \tag{9}$$

When the noise covariance matrix at the relay node $\mathbf{C}_{\mathbf{vv}}(0)$ is of full rank, it is easy to see that matrix $\mathbf{Q}_{\mathbf{y}_r\mathbf{y}_r}(z)$ is of full rank at $z = r_i$, $i = 1, \cdots, n_R$. For the case where the noise at the relay node is arbitrarily small, i.e., $\mathbf{C}_{\mathbf{vv}}(0)$ is a rank-deficient matrix, we assume that the number of antennas at the source and relay nodes are the same, i.e., $n_S = n_R$. With this assumption, the matrices $\mathbf{H}_1$ and $\mathbf{H}_1^H$ are of full rank. Since any source signal $\mathbf{s}(n)$ is not a linear combination of the other source signals, the following proposition is established.

Proposition 1: The power spectral matrix $\mathbf{Q}_{\mathbf{y}_r\mathbf{y}_r}(z)$ is of full rank at $z = r_i$ for $i = 1, \cdots, n_R$.

Let us denote the autocorrelation matrices of $\mathbf{y}(n)$ and $\mathbf{w}(n)$ as $\mathbf{C}_{\mathbf{yy}}(k)$ and $\mathbf{C}_{\mathbf{ww}}(k)$, respectively. It follows from (4) that

$$\begin{aligned}\mathbf{C}_{\mathbf{yy}}(k) &= \mathrm{E}\big[\mathbf{y}(n)\mathbf{y}(n-k)^H\big] \\ &= \mathbf{H}_2\mathbf{C}_{\mathbf{xx}}(k)\mathbf{H}_2^H + \mathbf{C}_{\mathbf{ww}}(k)\end{aligned} \tag{10}$$

where $\mathbf{C}_{\mathbf{xx}}(k) = \mathrm{E}\big[\mathbf{x}(n)\mathbf{x}(n-k)^H\big]$ is the autocorrelation matrix of $\mathbf{x}(n)$ and $\mathbf{C}_{\mathbf{ww}}(k) = \mathrm{E}\big[\mathbf{w}(n)\mathbf{w}(n-k)^H\big] = \mathbf{0}$ for $k \neq 0$ as the noises are temporally independent. Similarly, the power spectral matrix of $\mathbf{y}(n)$ can be derived based on (3), (9), and (10) as

$$\begin{aligned}\mathbf{Q}_{\mathbf{yy}}(z) &= \sum_{k=-\infty}^{\infty} \mathbf{C}_{\mathbf{yy}}(k)z^{-k} \\ &= \mathbf{H}_2\mathbf{Q}_{\mathbf{xx}}(z)\mathbf{H}_2^H + \mathbf{Q}_{\mathbf{ww}}(z) \\ &= \mathbf{H}_2\mathbf{P}(z)\mathbf{Q}_{\mathbf{y}_r\mathbf{y}_r}(z)\mathbf{P}(z^{-1})^H\mathbf{H}_2^H + \mathbf{Q}_{\mathbf{ww}}(z)\end{aligned} \tag{11}$$

where $\mathbf{Q_{xx}}(z) = \sum_{k=-\infty}^{\infty} \mathbf{C_{xx}}(k)z^{-k}$ and $\mathbf{Q_{ww}}(z) = \sum_{k=-\infty}^{\infty} \mathbf{C_{ww}}(k)z^{-k}$ are the power spectral matrices of $\mathbf{x}(n)$ and $\mathbf{w}(n)$, respectively, and $\mathbf{P}(z) = \mathrm{diag}(p_1(z), p_2(z), \cdots, p_{n_R}(z))$ is a diagonal matrix.

Let us introduce

$$\mathbf{T}_i(z) = \mathbf{P}_i(z)\mathbf{Q_{y_r y_r}}(z)\mathbf{P}(z^{-1})^H, \quad i = 1, \cdots, n_R$$

where $\mathbf{P}_i(z)$ is the matrix $\mathbf{P}(z)$ with the ith diagonal entry replaced by zero, i.e.,

$$\mathbf{P}_i(z) = \mathrm{diag}(p_1(z), \cdots, p_{i-1}(z), 0, p_{i+1}(z), \cdots, p_{n_R}(z)). \tag{12}$$

LEMMA 1: The rank of $\mathbf{T}_i(r_i)$ is $n_R - 1$, for $i = 1, \cdots, n_R$, and all rows of $\mathbf{T}_i(r_i)$ except for the ith row are linearly independent.

PROOF: It can be shown that for any r_i,

$$\mathrm{rank}(\mathbf{P}_i(r_i)) = n_R - 1$$

while the matrix $\mathbf{P}(r_i^{-1})^H$ is of full rank, since r_i^{-1} is not a zero of any precoder. It can be shown using (12) that all elements in the ith row of $\mathbf{T}_i(r_i)$ are zero. Using these results and Proposition 1, we obtain the rank of $\mathbf{T}_i(r_i)$ as $n_R - 1$. We can further deduce that all rows of $\mathbf{T}_i(r_i)$ except for the ith row are linearly independent. $\square$

Let $\mathbf{H}_{2,i}$ be equal to $\mathbf{H}_2$ with the ith column replaced by a zero vector, i.e.,

$$\mathbf{H}_{2,i} = [\mathbf{h}_{2,1}, \cdots, \mathbf{h}_{2,i-1}, \mathbf{0}, \mathbf{h}_{2,i+1}, \cdots, \mathbf{h}_{2,n_R}]. \tag{13}$$

We can rewrite (11) as

$$\begin{aligned} \mathbf{Q_{yy}}(r_i) &= \mathbf{H}_2 \mathbf{P}(r_i)\mathbf{Q_{y_r y_r}}(r_i)\mathbf{P}(r_i^{-1})^H \mathbf{H}_2^H + \mathbf{Q_{ww}}(r_i) \\ &= \mathbf{H}_{2,i}\mathbf{P}_i(r_i)\mathbf{Q_{y_r y_r}}(r_i)\mathbf{P}(r_i^{-1})^H \mathbf{H}_2^H + \mathbf{C_{ww}}(0) \\ &= \mathbf{H}_{2,i}\mathbf{T}_i(r_i)\mathbf{H}_2^H + \mathbf{C_{ww}}(0). \end{aligned} \tag{14}$$

Assuming that $\mathbf{C_{ww}}(0)$ can be estimated, which will be shown later, and removed from (14), we have

$$\bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) = \mathbf{H}_{2,i}\mathbf{T}_i(r_i)\mathbf{H}_2^H. \tag{15}$$

The following theorem establishes the estimation criterion for our blind channel estimation algorithm.

THEOREM 1: For $i = 1, \cdots, n_R$, $\mathbf{b}_{1,i}$ is an $n_D \times 1$ separation vector ensuring

$$\mathbf{b}_{1,i}^H \mathbf{H}_2 = [0, \cdots, 0, c_i, 0, \cdots, 0], \quad c_i \neq 0 \tag{16}$$

if and only if

$$\begin{cases} \mathbf{b}_{1,i}^H \bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) = \mathbf{0} & (17) \\ \mathbf{b}_{1,i}^H \mathbf{C_{yy}}(1)\mathbf{b}_{1,i} \neq 0. & (18) \end{cases}$$

PROOF: See Appendix A. $\square$

Theorem 1 holds when the autocorrelation matrix of $\mathbf{y}(n)$ has a time lag of $\tau = 1$, i.e., $\mathbf{C_{yy}}(1)$. Interestingly, it is shown in the following corollary that Theorem 1 is not valid for $\mathbf{C_{yy}}(\tau)$ with other time lag values.

COROLLARY 1: Theorem 1 does not hold for $\mathbf{C_{yy}}(\tau)$, $\tau \neq 1$.

PROOF: See Appendix B. $\square$

It can be seen that the proposed first-order precoding algorithm has different requirements on the selection of parameters compared with the second-order precoding algorithm in [20]. For example, τ can be 1 or 2 in the second-order precoding method but can only be 1 in the proposed first-order precoding algorithm. The implementation of the first-order Z-domain precoding based blind channel estimation algorithm is shown in the following subsection.

B. Algorithm Implementation

The following blind channel estimation procedures are applied to obtain the relay-destination channel matrix $\mathbf{H}_2$.

1) Compute the estimated autocorrelation matrix of $\mathbf{y}(n)$ as

$$\hat{\mathbf{C}}_{\mathbf{yy}}(k) \approx \frac{1}{L}\sum_{n=0}^{L-1} \mathbf{y}(n)\mathbf{y}(n-k)^H \tag{19}$$

where $L \geq n_D$ is the number of samples of the received signal.

2) Compute the estimated power spectral matrix of $\mathbf{y}(n)$ as

$$\hat{\mathbf{Q}}_{\mathbf{yy}}(r_i) \approx \sum_k \mathbf{C_{yy}}(k)r_i^{-k}, \quad i = 1, \cdots, n_R \tag{20}$$

where k is a finite integer.

3) Estimate the noise covariance matrix $\mathbf{C_{ww}}(0)$. It follows from (10) that

$$\mathbf{C_{yy}}(0) = \mathbf{H}_2\mathbf{C_{xx}}(0)\mathbf{H}_2^H + \mathbf{C_{ww}}(0). \tag{21}$$

Since the noises are assumed to be i.i.d. white Gaussian, we have

$$\mathbf{C_{ww}}(0) = \sigma_w^2\mathbf{I}_{n_D} \tag{22}$$

where σ_w^2 is the noise variance and $\mathbf{I}_n$ is an $n \times n$ identity matrix. Let us introduce the eigenvalue decomposition (EVD) of

$$\mathbf{C_{yy}}(0) = \mathbf{U}_Y\mathbf{\Lambda}_Y\mathbf{U}_Y^H \tag{23}$$

where $\mathbf{U}_Y$ is the unitary eigenvector matrix and $\mathbf{\Lambda}_Y$ is the diagonal eigenvalue matrix with descending diagonal elements. Obviously, from (21) there is

$$\mathbf{\Lambda}_Y = \mathbf{\Lambda}_X + \sigma_w^2\mathbf{I}_{n_D} \tag{24}$$

where $\mathbf{\Lambda}_X$ is the eigenvalue matrix of $\mathbf{H}_2\mathbf{C_{xx}}(0)\mathbf{H}_2^H$ with descending diagonal elements.
If $n_D > n_R$, i.e., $\mathbf{H}_2$ is a tall matrix, from (24), we have

$$\begin{aligned} \lambda_{y,i} &= \lambda_{x,i} + \sigma_w^2, & i = 1,\cdots,n_R \\ \lambda_{y,i} &= \sigma_w^2, & i = n_R+1,\cdots,n_D \end{aligned} \tag{25}$$

where $\lambda_{y,i}$, $i = 1,\cdots,n_D$, and $\lambda_{x,j}$, $j = 1,\cdots,n_R$, are the diagonal elements of $\mathbf{\Lambda}_Y$ and $\mathbf{\Lambda}_X$, respectively. From (25), we can estimate σ_w^2 as

$$\hat{\sigma}_w^2 = \frac{1}{n_D - n_R}\sum_{i=n_R+1}^{n_D}\hat{\lambda}_{y,i} \tag{26}$$

where $\hat{\lambda}_{y,i}$ is estimated $\lambda_{y,i}$ obtained from the EVD of $\hat{\mathbf{C}}_{\mathbf{yy}}(0)$. The estimated noise covariance matrix is given by

$$\hat{\mathbf{C}}_{\mathbf{ww}}(0) = \hat{\sigma}_w^2\mathbf{I}_{n_D}. \tag{27}$$

If $n_D = n_R$, i.e., $\mathbf{H}_2$ is a square matrix, the noise covariance matrix can be estimated prior to the transmission of data, i.e., when $\mathbf{y}(n) = \mathbf{w}(n)$, $n = 1,\cdots,J$, we have

$$\hat{\mathbf{C}}_{\mathbf{ww}}(0) \approx \frac{1}{J}\sum_{n=0}^{J-1}\mathbf{y}(n)\mathbf{y}(n)^H.$$

4) Estimate $\bar{\mathbf{Q}}_{\mathbf{yy}}(r_i)$ as

$$\bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) \triangleq \hat{\mathbf{Q}}_{\mathbf{yy}}(r_i) - \hat{\mathbf{C}}_{\mathbf{ww}}(0), \quad i = 1,\cdots,n_R. \tag{28}$$

5) Obtain separation matrix $\mathbf{B}_1$ as follows. From Lemma 1, it can be seen that $\bar{\mathbf{Q}}_{\mathbf{yy}}(r_i)$ has a rank of $n_R - 1$. Since $\bar{\mathbf{Q}}_{\mathbf{yy}}(r_i)$ is an $n_D \times n_D$ matrix, there are $n_D - n_R + 1$ zero singular values. As we assume $n_D \geq n_R$, there exists at least one zero singular value. Let $\mathbf{V}_i$ be an $n_D \times (n_D - n_R + 1)$ matrix whose columns consist of the $n_D - n_R + 1$ left singular vectors corresponding to the zero singular values of $\bar{\mathbf{Q}}_{\mathbf{yy}}(r_i)$, and column vector $\mathbf{u}_i$ be the eigenvector corresponding to any nonzero eigenvalue λ of $\mathbf{V}_i^H\hat{\mathbf{C}}_{\mathbf{yy}}(1)\mathbf{V}_i$. It can be proven that

$$\mathbf{u}_i^H\mathbf{V}_i^H\bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) = \mathbf{0}$$

and

$$\mathbf{u}_i^H\mathbf{V}_i^H\hat{\mathbf{C}}_{\mathbf{yy}}(1)\mathbf{V}_i\mathbf{u}_i = \lambda\mathbf{u}_i^H\mathbf{u}_i \neq 0.$$

Then, the separation vector $\mathbf{b}_{1,i}$ can be selected as $\mathbf{b}_{1,i}^H = \mathbf{u}_i^H\mathbf{V}_i^H$. The operations in this step are carried out for $i = 1,\cdots,n_R$.

6) The precoded signals can be estimated by

$$\hat{\mathbf{x}}(n) = \mathbf{B}_1^H\mathbf{y}(n), \quad n = 1,\cdots,L. \tag{29}$$

7) The relay-destination channel matrix is estimated as

$$\hat{\mathbf{H}}_2 = \mathbf{Y}\hat{\mathbf{X}}^\dagger \tag{30}$$

where $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \cdots, \mathbf{y}(L)]$ and $\hat{\mathbf{X}} = [\hat{\mathbf{x}}(1), \hat{\mathbf{x}}(2), \cdots, \hat{\mathbf{x}}(L)]$. Note that since $L \geq n_R$, we have the right inverse of $\hat{\mathbf{X}}$ as

$$\hat{\mathbf{X}}^\dagger = \hat{\mathbf{X}}^T(\hat{\mathbf{X}}\hat{\mathbf{X}}^T)^{-1} \tag{31}$$

where $(\cdot)^{-1}$ stands for matrix inversion.

IV. Channel Estimation Based on Signal MI Modified Constant Modulus Algorithm

In this section, we develop a signal MI modified constant modulus algorithm to estimate the first-hop channel matrix $\mathbf{H}_1$. Based on the estimated precoded signals $\hat{x}_i(n)$, $i = 1, \cdots, n_R$, the signals received at the relay node can be estimated by

$$\hat{y}_{r,i}(n) = \hat{x}_i(n) + r_i \hat{y}_{r,i}(n-1), \quad i = 1, \cdots, n_R. \tag{32}$$

Let us introduce an $n_R \times n_S$ separation matrix $\mathbf{B}_2$ and let

$$\hat{\mathbf{s}}(n) = \mathbf{B}_2^H \hat{\mathbf{y}}_r(n) = \mathbf{C}\mathbf{s}(n) + \mathbf{B}_2^H \mathbf{v}(n) \tag{33}$$

where $\hat{\mathbf{s}}(n)$ is the estimated source signal vector and $\mathbf{C} \triangleq \mathbf{B}_2^H \mathbf{H}_1$. This blind channel estimation algorithm aims to obtain the separation matrix $\mathbf{B}_2$ in order to recover the first-hop channel $\mathbf{H}_1$, only from the estimated relay channel output signals $\hat{\mathbf{y}}_r(n)$. Obviously, the estimation of $\mathbf{H}_1$ is affected by the accuracy of the estimation of $\mathbf{y}_r(n)$. As the source signals are unknown at the destination, there are inherent scaling and permutation ambiguities in this algorithm, i.e.,

$$\mathbf{C} = \mathbf{B}_2^H \mathbf{H}_1 = \mathbf{P}\boldsymbol{\Delta}$$

where $\mathbf{P}$ is a permutation matrix and $\boldsymbol{\Delta}$ is a diagonal matrix.

A. Development of the Algorithm

The general cost function for the constant modulus algorithm is given by

$$\sum_{i=1}^{n_S} \mathrm{E}\big[(|\hat{s}_i(n)|^2 - \gamma)^2\big]$$

where $\hat{s}_i(n)$ is the ith element of $\hat{\mathbf{s}}(n)$ and γ is a constant. As mentioned earlier, the constant modulus algorithm is capable of retrieving one source signal at a time. However, it does not guarantee the extraction of all source signals as the constant modulus algorithm might extract the same signal.

Similar to [15], we propose to exploit the MI property of the estimated signals, along with the constant modulus algorithm, to ensure that the channel matrix and source signals are completely separated. In particular, the following cost function with the addition of the MI term is minimized

$$J(\mathbf{B}_2) = \sum_{i=1}^{n_S} \mathrm{E}\big[(|\hat{s}_i(n)|^2 - \gamma)^2\big] + \beta\left[\sum_{i=1}^{n_S} \log(r_{ii}) - \log|\mathbf{R}_{\hat{s}\hat{s}}|\right] \tag{34}$$

where β is a positive real number that balances the constant modulus term and the MI term, r_{ii} is the ith diagonal element of $\mathbf{R}_{\hat{s}\hat{s}}$, and $\mathbf{R}_{\hat{s}\hat{s}} \triangleq \mathrm{E}\big[\hat{\mathbf{s}}(n)\hat{\mathbf{s}}(n)^H\big]$ is the covariance matrix of $\hat{\mathbf{s}}(n)$. From [15], we have the following proposition.

PROPOSITION 2: The MI term is zero when $\mathbf{R}_{\hat{s}\hat{s}}$ is a diagonal matrix, i.e., when the elements of $\hat{\mathbf{s}}(n)$ are uncorrelated.

Proposition 2 is important to ensure that all source signals are separated from the channel matrix $\mathbf{H}_1$ at the destination node. The cost function (34) can be rewritten as

$$J(\mathbf{B}_2) = \mathrm{E}\left[\sum_{i=1}^{n_S}\big(\mathbf{e}_i^T \mathbf{B}_2^H \hat{\mathbf{y}}_r(n)\hat{\mathbf{y}}_r(n)^H \mathbf{B}_2 \mathbf{e}_i - \gamma\big)^2\right] + \beta\left[\sum_{i=1}^{n_S} \log\big(\mathbf{e}_i^T \mathbf{B}_2^H \mathbf{R}_{\hat{y}_r\hat{y}_r}\mathbf{B}_2 \mathbf{e}_i\big) - \log\big|\mathbf{B}_2^H \mathbf{R}_{\hat{y}_r\hat{y}_r}\mathbf{B}_2\big|\right]$$

where $\mathbf{R}_{\hat{y}_r\hat{y}_r} \triangleq \mathrm{E}\big[\hat{\mathbf{y}}_r(n)\hat{\mathbf{y}}_r(n)^H\big]$ is the covariance matrix of $\hat{\mathbf{y}}_r(n)$ and $\mathbf{e}_i$ is an $n_S \times 1$ column vector whose elements are zero except for the ith element which is one. The gradient of $J(\mathbf{B}_2)$ is given by

$$\nabla J(\mathbf{B}_2) = \frac{\partial J(\mathbf{B}_2)}{\partial \mathbf{B}_2^*}$$
$$= 2\sum_{i=1}^{n_S} \mathrm{E}\Big[\big(|\hat{s}_i(n)|^2 - \gamma\big)\hat{\mathbf{y}}_r(n)\mathbf{e}_i^T\big(\hat{\mathbf{y}}_r(n)^H \mathbf{B}_2 \mathbf{e}_i\big)\Big] + \beta \mathbf{R}_{\hat{y}_r\hat{y}_r}\mathbf{B}_2\Big[(\mathrm{diag}(\mathbf{R}_{\hat{s}\hat{s}}))^{-1} - \mathbf{R}_{\hat{s}\hat{s}}^{-1}\Big]. \tag{35}$$

B. Algorithm Implementation

The procedure of applying the signal MI modified constant modulus algorithm to estimate the source-relay channel matrix $\mathbf{H}_1$ is listed below.

1) Initialize $\mathbf{B}_2^{(0)}$ and $\mathbf{R}_{\hat{y}_r\hat{y}_r}^{(0)}$; Set $i = 1$.
2) Update $\mathbf{R}_{\hat{y}_r\hat{y}_r}^{(i)}$ through

$$\mathbf{R}_{\hat{y}_r\hat{y}_r}^{(i)} = (1-\kappa)\mathbf{R}_{\hat{y}_r\hat{y}_r}^{(i-1)} + \kappa\hat{\mathbf{y}}_r(i)\hat{\mathbf{y}}_r(i)^H \tag{36}$$

where $0 < \kappa < 1$ is a small positive real number.
3) Estimate $\hat{\mathbf{s}}(i) = \big(\mathbf{B}_2^{(i-1)}\big)^H \hat{\mathbf{y}}_r(i)$.

4) Calculate $\mathbf{R}_{\hat{\mathbf{s}}\hat{\mathbf{s}}}^{(i)} = \left(\mathbf{B}_2^{(i-1)}\right)^H \mathbf{R}_{\hat{\mathbf{y}}_r\hat{\mathbf{y}}_r}^{(i)} \mathbf{B}_2^{(i-1)}$.

5) From Steps 1–4, an estimation of (35) is obtained by removing the expectation operator E in the equation. Let us denote this estimation as $\hat{\nabla} J(\mathbf{B}_2)$.

6) Update the separation matrix $\mathbf{B}_2$ as

$$\mathbf{B}_2^{(i)} = \mathbf{B}_2^{(i-1)} - \mu \hat{\nabla} J(\mathbf{B}_2)|_{\mathbf{B}_2=\mathbf{B}_2^{(i-1)}}. \tag{37}$$

7) Repeat Steps 2–6 for $i = 2, 3, \cdots, L$ to obtain $\mathbf{B}_2 = \mathbf{B}_2^{(L)}$.

8) The source signals are estimated as

$$\hat{\mathbf{s}}(n) = \mathbf{B}_2^H \hat{\mathbf{y}}_r(n), \quad n = 1, \cdots, L \tag{38}$$

9) Estimate the source-relay channel matrix as

$$\hat{\mathbf{H}}_1 = \hat{\mathbf{Y}}_r \hat{\mathbf{S}}^\dagger \tag{39}$$

where $\hat{\mathbf{Y}}_r = [\hat{\mathbf{y}}_r(1), \hat{\mathbf{y}}_r(2), \cdots, \hat{\mathbf{y}}_r(L)]$ and $\hat{\mathbf{S}} = [\hat{\mathbf{s}}(1), \hat{\mathbf{s}}(2), \cdots, \hat{\mathbf{s}}(L)]$. Note that since $L \geq n_S$, we have the right inverse of $\hat{\mathbf{S}}$ as

$$\hat{\mathbf{S}}^\dagger = \hat{\mathbf{S}}^T (\hat{\mathbf{S}}\hat{\mathbf{S}}^T)^{-1}. \tag{40}$$

We would like to note that the algorithm proposed in [15] was developed for blind signal separation in one-hop systems, whereas in this paper we apply this algorithm for channel estimation in dual-hop MIMO relay communication systems.

V. NUMERICAL EXAMPLES

In this section, we study the performance of the proposed blind MIMO relay channel estimation algorithm through numerical simulations. We consider a three-node two-hop MIMO relay system with n_S, n_R, and n_D antennas equipped at the source, relay, and destination nodes, respectively. For the proposed first-order Z-domain precoding based channel estimation algorithm, the zeros of the precoders in (2) are chosen as

$$r_i = \eta_i e^{\frac{j\pi(2i-1)}{2n_R}}, \quad i = 1, \cdots, n_R \tag{41}$$

where $j = \sqrt{-1}$ and $0 < \eta_i < 1$, $i = 1, \cdots, n_R$. This model ensures that all zeros are distinct and satisfy $0 < |r_i| < 1$, $i = 1, \cdots, n_R$, and the angles of zeros are equally spaced on the Z-plane. For the signal MI modified constant modulus based channel estimation algorithm, unless explicitly mentioned, the matrices $\mathbf{B}_2^{(n)}$ and $\mathbf{R}_{\hat{\mathbf{y}}_r\hat{\mathbf{y}}_r}^{(n)}$ are initialized as $\mathbf{B}_2^{(0)} = \left[\mathbf{I}_{n_S}, \mathbf{0}_{n_S \times (n_R-n_S)}\right]^H$ and $\mathbf{R}_{\hat{\mathbf{y}}_r\hat{\mathbf{y}}_r}^{(0)} = \mathbf{I}_{n_R}$, respectively, where $\mathbf{0}_{m \times n}$ is an $m \times n$ zero matrix. We choose $\mu = 0.0005$, $\kappa = 0.05$, $\beta = 1$, and $\gamma = 1$ based on the following reasons. The step size of the gradient descent algorithm μ is chosen to be small enough to ensure the convergence of the algorithm, while γ is chosen to be 1 as the absolute value of the source signals has a constant unit value. We apply the quadrature phase-shift keying (QPSK) modulation scheme in all our simulations. We assume that the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are complex Gaussian distributed with zero mean and unit variance, and channels do not change within L symbols of transmission. All simulation results are averaged over 1000 random channel realizations. The SNR of the source-relay and relay-destination link is denoted as $\text{SNR}_{\text{s-r}}$ and $\text{SNR}_{\text{r-d}}$, respectively.

In the first example, we evaluate the performance of the proposed blind channel estimation algorithm at various number of samples L of the received signal. Fig. 2 shows the BER of the proposed algorithm versus L for various n_S and n_R with $n_D = 4$ and $\text{SNR}_{\text{s-r}}$ and $\text{SNR}_{\text{r-d}}$ fixed at 20dB. It can be seen from Fig. 2 that the BER performance of the proposed algorithm improves when L increases. This is because in the proposed first-order Z-domain precoding based channel estimation algorithm, the accuracy of estimating the autocorrelation matrix $\mathbf{C}_{\mathbf{yy}}(k)$ is affected by L, i.e., the estimated $\mathbf{C}_{\mathbf{yy}}(k)$ approaches its theoretical value at a large L. Moreover, the performance of the signal MI modified constant modulus algorithm improves when a larger L is used as more iterations are involved in finding the separation matrix. In the following simulation examples, the number of samples is chosen as $L = 5000$ to achieve a good tradeoff between the performance and the computational complexity. For indoor wireless relay channels, channel fading is often relatively slow whenever the mobility of the nodes is relatively low, and for static nodes, the CSI can be almost constant [23], [24]. In this case, the required number of samples can be collected within the channel coherent time. Therefore, our algorithm can be applied to wireless systems with static nodes such as indoor machine-to-machine (M2M) communication.

In the second example, we study the performance of the proposed blind channel estimation algorithm in finding the separation matrix. For each channel realization, the mean interference rejection level (MIRL) for the first-order Z-domain precoding algorithm is calculated as

$$\text{MIRL}_{\mathbf{H}_2} = 10\log_{10}\left(\frac{1}{n_R(n_R-1)} \sum_{i=1}^{n_R} \sum_{j=1, j\neq i}^{n_R} \frac{|(\mathbf{B}_1^H\mathbf{H}_2)_{ij}|^2}{|(\mathbf{B}_1^H\mathbf{H}_2)_{ii}|^2}\right) \tag{42}$$

while the MIRL of the signal MI modified constant modulus algorithm is given by

$$\text{MIRL}_{\mathbf{H}_1} = 10\log_{10}\left(\frac{1}{n_S(n_S-1)}\sum_{i=1}^{n_S}\frac{\sum_{j=1}^{n_S}|(\mathbf{C})_{ij}|^2 - \max_{j}(|(\mathbf{C})_{ij}|^2)}{\max_{j}(|(\mathbf{C})_{ij}|^2)}\right). \tag{43}$$

Note that a smaller value of MIRL indicates a better performance of the blind channel estimation algorithm.

Fig. 3 shows the MIRL for the proposed blind channel estimation algorithm versus $\text{SNR}_{\text{r}-\text{d}}$ with $n_S = n_R = 3$, $n_D = 4$, and $\text{SNR}_{\text{s}-\text{r}} = 20\text{dB}$. It can be seen from Fig. 3 that the MIRL performance of the proposed blind channel estimation algorithm improves with the increase of $\text{SNR}_{\text{r}-\text{d}}$. Interestingly, the first-order Z-domain precoding technique performs better than the signal MI modified constant modulus algorithm, as the latter algorithm is affected by the accuracy of the estimation of $\mathbf{y}_r(n)$. Note that for the first-order Z-domain precoding technique, theoretically the derivation of the separation matrix is not affected by the noise at the destination node, thus only a small improvement is observed when $\text{SNR}_{\text{r}-\text{d}}$ increases. A plot of the MIRL of the proposed blind channel estimation algorithm versus $\text{SNR}_{\text{r}-\text{d}}$ for $\text{SNR}_{\text{s}-\text{r}} = 20\text{dB}$, $n_D = 4$, and various n_S and n_R is shown in Fig. 4. It can be seen from Fig. 4 that when the number of antennas at the source node and relay node increases, the MIRL also increases.

In the third example, we demonstrate the performance of the proposed blind channel estimation algorithm in terms of the normalized mean-squared error (NMSE). For the relay-destination channel, the NMSE is calculated as

$$\text{NMSE}_{\mathbf{H}_2} = \frac{\|\mathbf{H}_2 - \hat{\mathbf{H}}_2\|_F^2}{n_R n_D} \tag{44}$$

where $\|\cdot\|_F$ denotes the matrix Frobenius norm. Similarly, the NMSE for the estimation of the source-relay channel matrix is given by

$$\text{NMSE}_{\mathbf{H}_1} = \frac{\|\mathbf{H}_1 - \hat{\mathbf{H}}_1\|_F^2}{n_S n_R}. \tag{45}$$

Similar to [20], the scaling ambiguity in estimating $\mathbf{H}_2$ is removed by minimizing the MSE between $\mathbf{x}(n)$ and $\hat{\mathbf{x}}(n)$. The scaling and permutation ambiguity in the estimation of $\mathbf{H}_1$ is removed by minimizing the MSE between $\mathbf{s}(n)$ and $\hat{\mathbf{s}}(n)$.

Fig. 5 shows the NMSE of the proposed blind estimation algorithm versus $\text{SNR}_{\text{r}-\text{d}}$ for various n_S and n_R with $\text{SNR}_{\text{s}-\text{r}}$ fixed at 20dB and $n_D = 4$. It can be seen from Fig. 5 that the NMSE of estimating $\mathbf{H}_1$ and $\mathbf{H}_2$ decreases when the number of antennas at the source and relay nodes decreases. Note that only small improvement is observed in the estimation of $\mathbf{H}_2$ at high $\text{SNR}_{\text{r}-\text{d}}$ due to the error floor introduced in the estimation of the scaling ambiguity. We also investigate the performance of signal MI modified constant modulus channel estimation scheme when this algorithm is initialized with random matrices. It can be seen from Fig. 5 that the NMSE of the first-hop channel estimation with random matrices initialization is very similar to the NMSE when the scheme is initialized with identity matrix.

In the fourth example, we compare the proposed blind MIMO relay channel estimation algorithm with the training-based MIMO relay channel estimation algorithm developed in [10], where the training sequences are optimized with proper adjustment of the power available at the source and relay nodes for a fair comparison[1]. The channel correlation matrices used in the training-based algorithm [10] are set to identity matrices to have the same statistical distribution as the channel model used in the proposed blind channel estimation algorithm. Fig. 6 shows the NMSE performance of estimating $\mathbf{H}_2$ and $\mathbf{H}_1$ versus $\text{SNR}_{\text{r}-\text{d}}$ with $\text{SNR}_{\text{s}-\text{r}} = 20\text{dB}$, $n_S = n_R = 2$, and $n_D = 4$. The NMSE performance of two algorithms versus $\text{SNR}_{\text{r}-\text{d}}$ with $\text{SNR}_{\text{s}-\text{r}} = 20\text{dB}$, $n_S = n_R = 3$, and $n_D = 4$ is demonstrated in Fig. 7. It can be seen from Figs. 6 and 7 that at low SNR, the performance of the proposed algorithm is comparable to that of the training-based algorithm in estimating $\mathbf{H}_1$. Note that the proposed algorithm performs better than the training-based algorithm in the estimation of $\mathbf{H}_2$ at low SNR. This is because in the training-based approach, a large portion of the transmission power is allocated to broadcast information, while only a small part of the transmission power is used for channel estimation. In contrast, our proposed algorithm utilizes the transmitted information to estimate the channels, thus, all the power available for transmission is used for channel estimation. However, at high SNR, the training-based algorithm outperforms the proposed algorithm at the expense of bandwidth efficiency. The MSE error floors of the proposed algorithm at high $\text{SNR}_{\text{r}-\text{d}}$ are mainly due to the error introduced in the estimation of the ambiguities. For a fair comparison, both the proposed blind channel estimation algorithm and the training-based algorithm are adjusted to have the same transmission power. This limits the performance of the training-based algorithm as the power available for transmission is essentially fixed. Thus, error floors appear in the simulation results of the training-based algorithm.

Fig. 8 illustrates the BER performance of two algorithms versus $\text{SNR}_{\text{r}-\text{d}}$ when $n_S = n_R = 2$, $n_D = 4$, and $\text{SNR}_{\text{s}-\text{r}}$ is fixed at 20dB. As a benchmark, we also show the BER performance of the MIMO relay system where the channel matrices are perfectly known. It can be seen from Fig. 8 that the BER performance of the proposed blind channel estimation algorithm is close to the performance of the training-based algorithm. Due to the error floors, further increment of SNR would not result in a better BER performance.

---

[1]From (3), the transmission power consumed by the relay node for the proposed blind channel estimation algorithm is higher than the one in [10] by a factor of $(1 + |r_i|^2)$, due to the precoder applied at the relay node.

Finally, we compare the computational complexity of the proposed blind channel estimation algorithm and the training-based channel estimation technique [10]. The complexity of the first-order Z-domain precoding based channel estimation algorithm is governed by the EVD and the singular value decomposition (SVD) operations required in deriving the separation matrix, while the complexity of the signal MI modified constant modulus algorithm is governed by the matrix inversion operation in the gradient descent method. Thus, the computational complexity of the proposed blind channel estimation algorithm can be estimated as $\mathcal{O}(n_D^3 + L n_S^3)$, where the first term represents the complexity of the first-order Z-domain precoding based channel estimation algorithm, and the second term is the complexity of the signal MI modified constant modulus algorithm.

The complexity of the training-based channel estimation technique [10] can be estimated as $\mathcal{O}(d_\nu d_c n_R^2 + d_a d_{\mu_1} d_{\lambda_F} n_S + d_a d_{\mu_2} d_{\mu_3} d_{\lambda_S} n_D)$, where $d_\nu, d_{\mu_1}, d_{\mu_2}$, and d_{μ_3} stand for the number of iterations required to obtain the optimal Lagrangian multipliers associated with the optimization problem in [10], d_c and d_{λ_S} represent the number of bisection operations required to obtain the optimal training sequences, d_{λ_F} is the number of bisection operations required to derive the optimal relay amplification matrix, and d_a stands for the number of iterations required to find the local optimal solution to the problem.

The implementation of channel coding and decoding will benefit both the proposed algorithm and the training-based algorithm. However, channel coding and decoding are not included in this paper as the focus of this paper is on the channel estimation of MIMO relay networks.

VI. CONCLUSIONS

We have developed a new blind channel estimation algorithm for two-hop MIMO relay systems. The proposed algorithm is able to estimate the individual source-relay and relay-destination CSI at the destination node, which is necessary for retrieving the source signals at the destination node. In particular, a novel first-order Z-domain precoding technique has been developed for the blind estimation of the relay-destination channel matrix. The proposed algorithm has a similar BER performance to the training-based channel estimation algorithm, and better bandwidth efficiency as all the bandwidth is used for sending communication signals. The proposed algorithm can be extended to other MIMO relay communication systems such as multiuser MIMO relay systems with multiple relay nodes.

APPENDIX A
PROOF OF THEOREM 1

We prove Theorem 1 through verifying the necessity and sufficiency conditions. Assuming that (16) is satisfied, we prove the necessity of (17) as follow

$$\mathbf{b}_{1,i}^H \bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) = \mathbf{b}_{1,i}^H \mathbf{H}_{2,i} \mathbf{T}_i(r_i) \mathbf{H}_2^H = \mathbf{0}. \tag{46}$$

Since we assumed the source signals to be temporally white, from (3), (10), and (16), we prove the necessity condition for (18) as

$$\begin{aligned}
\mathbf{b}_{1,i}^H \mathbf{C}_{\mathbf{yy}}(1) \mathbf{b}_{1,i} &= \mathbf{b}_{1,i}^H \mathbf{H}_2 \mathbf{C}_{\mathbf{xx}}(1) \mathbf{H}_2^H \mathbf{b}_{1,i} \\
&= c_i c_i^* \mathrm{E}\big[x_i(n) x_i(n-1)^* \big] \\
&= |c_i|^2 \mathrm{E}\Big[\big(y_{r,i}(n) - r_i y_{r,i}(n-1) \big) \big(y_{r,i}(n-1) - r_i y_{r,i}(n-2) \big)^* \Big] \\
&= |c_i|^2 \mathrm{E}\big[-r_i y_{r,i}(n-1) y_{r,i}(n-1)^* \big] \\
&= -|c_i|^2 r_i \sigma_i^2 \\
&\neq 0
\end{aligned}$$

where $\sigma_i^2 \triangleq \mathrm{E}[y_{r,i}(n-1) y_{r,i}(n-1)^*]$.

Now we prove the sufficiency of (17) and (18). Since $\mathbf{b}_{1,i}^H \bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) = \mathbf{0}$, from (15) we have

$$\mathbf{b}_{1,i}^H \mathbf{H}_{2,i} \mathbf{T}_i(r_i) \mathbf{H}_2^H = \mathbf{0}. \tag{47}$$

The matrix $\mathbf{H}_2^H$ is of full row rank, and thus implying that

$$\mathbf{b}_{1,i}^H \mathbf{H}_{2,i} \mathbf{T}_i(r_i) = \mathbf{0}. \tag{48}$$

From Lemma 1, all the rows of the matrix $\mathbf{T}_i(r_i)$ excluding the ith row are linearly independent, and therefore we obtain that

$$\mathbf{b}_{1,i}^H \mathbf{h}_{2,j} = 0, \quad j = 1, \cdots, n_R, \quad j \neq i. \tag{49}$$

Subsequently, from (13) and (49), we have

$$\mathbf{b}_{1,i}^H \mathbf{H}_2 = \big[0, \cdots, 0, \mathbf{b}_{1,i}^H \mathbf{h}_{2,i}, 0, \cdots, 0 \big]. \tag{50}$$

Next, we consider $\mathbf{b}_{1,i}^H \mathbf{C}_{\mathbf{yy}}(1) \mathbf{b}_{1,i} \neq 0$. From (10), we have

$$\mathbf{b}_{1,i}^H \mathbf{H}_2 \mathbf{C}_{\mathbf{xx}}(1) \mathbf{H}_2^H \mathbf{b}_{1,i} \neq 0 \tag{51}$$

which implies that $\mathbf{b}_{1,i}^H \mathbf{H}_2 \neq 0$, and from (50), we can infer that

$$\mathbf{b}_{1,i}^H \mathbf{H}_2 = [0, \cdots, 0, c_i, 0, \cdots, 0] \tag{52}$$

where $c_i = \mathbf{b}_{1,i}^H \mathbf{h}_{2,i} \neq 0$. $\qquad\square$

APPENDIX B
PROOF OF COROLLARY 1

A. For $\tau = 0$

For the case of $n_D > n_R$, the channel matrix $\mathbf{H}_2$ has a row-rank deficiency, i.e., the rows of $\mathbf{H}_2$ are linearly dependent. Subsequently, an $n_D \times 1$ non-zero vector $\mathbf{b}_i$ exists such that

$$\mathbf{b}_i^H \mathbf{H}_2 = \mathbf{0}. \tag{53}$$

From (15) and (53), we have

$$\mathbf{b}_i^H \bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) = \mathbf{b}_i^H \mathbf{H}_{2,i} \mathbf{T}_i(r_i) \mathbf{H}_2^H = \mathbf{0}. \tag{54}$$

Based on (10) and (53), there is

$$
\begin{aligned}
\mathbf{b}_i^H \mathbf{C}_{\mathbf{yy}}(0) \mathbf{b}_i &= \mathbf{b}_i^H \mathbf{H}_2 \mathbf{C}_{\mathbf{xx}}(0) \mathbf{H}_2^H \mathbf{b}_i + \mathbf{b}_i^H \mathbf{C}_{\mathbf{ww}}(0) \mathbf{b}_i \\
&= \mathbf{b}_i^H \mathbf{C}_{\mathbf{ww}}(0) \mathbf{b}_i \\
&= \sigma_w^2 \mathbf{b}_i^H \mathbf{b}_i \\
&\neq 0.
\end{aligned}
\tag{55}
$$

It can be observed from (53)-(55) that $\mathbf{b}_i^H \bar{\mathbf{Q}}_{\mathbf{yy}}(r_i) = \mathbf{0}$ and $\mathbf{b}_i^H \mathbf{C}_{\mathbf{yy}}(0) \mathbf{b}_i \neq 0$ do not guarantee (16).

B. For $\tau \geq 2$

Assuming (16) is satisfied, we have

$$
\begin{aligned}
\mathbf{b}_i^H \mathbf{C}_{\mathbf{yy}}(\tau) \mathbf{b}_i &= \mathbf{b}_i^H \mathbf{H}_2 \mathbf{C}_{\mathbf{xx}}(\tau) \mathbf{H}_2^H \mathbf{b}_i \\
&= c_i c_i^* \mathrm{E}\big[x_i(n) x_i(n-\tau)^*\big] \\
&= |c_i|^2 \mathrm{E}\Big[\big(y_{r,i}(n) - r_i y_{r,i}(n-1)\big)\big(y_{r,i}(n-\tau) - r_i y_{r,i}(n-\tau-1)\big)^*\Big] \\
&= 0.
\end{aligned}
$$

This indicates that no separation vector $\mathbf{b}_i$ can satisfy the condition $\mathbf{b}_i^H \mathbf{C}_{\mathbf{yy}}(\tau) \mathbf{b}_i \neq 0$ for time lag $\tau \geq 2$. $\qquad\square$

ACKNOWLEDGMENT

REFERENCES

[1] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1398-1407, Apr. 2007.

[2] I. Hammerström and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2798-2802, Aug. 2007.

[3] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837-4851, Dec. 2009.

[4] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A tutorial on transceiver design for amplify-and-forward MIMO relay systems," *IEEE J. Sel. Areas Commun.*, vol. 30, pp. 1331-1346, Sep. 2012.

[5] Y. Rong, "Joint source and relay optimization for two-way linear non-regenerative MIMO relay communications," *IEEE Trans. Signal Process.*, vol. 60, pp. 6533-6546, Dec. 2012.

[6] P. Lioliou and M. Viberg, "Least-squares based channel estimation for MIMO relays," in *Proc. Int. ITG Workshop on Smart Antennas*, pp. 90-95, Feb. 2008.

[7] P. Lioliou, M. Viberg, and M. Coldrey, "Efficient channel estimation techniques for amplify and forward relaying systems," *IEEE Trans. Commun.*, vol. 60, pp. 3150-3155, Nov. 2012.

[8] F. Roemer and M. Haardt, "Tensor-based channel estimation and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Trans. Signal Process.*, vol. 58, pp. 5720-5735, Nov. 2010.

[9] F. Gao, B. Jiang, X. Gao, and X.-D. Zhang, "Superimposed training based channel estimation for OFDM modulated amplify-and-forward relay networks," *IEEE Trans. Commun.*, vol. 59, pp. 2029-2039, July 2011.

[10] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Process.*, vol. 59, pp. 4438-4446, Sep. 2011.

[11] C. W. R. Chiong, Y. Rong, and Y. Xiang, "Robust channel estimation algorithm for dual-hop MIMO relay channels," in *Proc. 23rd IEEE PIMRC*, pp. 2376-2381, Sep. 2012.

[12] C. W. R. Chiong, Y. Rong, and Y. Xiang, "Channel training algorithms for two-way MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 61, pp. 3988-3998, Aug. 2013.

[13] Y. Rong, M. R. A. Khandaker, and Y. Xiang, "Channel estimation of dual-hop MIMO relay systems via parallel factor analysis," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 2224-2233, June 2012.

[14] A. Ikhlef and D. L. Guennec, "A simplified constant modulus algorithm for blind recovery of MIMO QAM and PSK signals: A criterion with convergence analysis," *EURASIP Journal Wireless Commun. Network.*, article ID 90401, Sep. 2007.

[15] Y. Xiang, N. Gu, and K. L. Wong, "Adaptive blind source separation using constant modulus criterion and signal mutual information," in *Proc. IEEE Int. Conf. Industrial Technol.*, pp. 1371-1375, Dec. 2005.

[16] X. G. Doukopoulos and G. V. Moustakides, "Adaptive power techniques for blind channel estimation in CDMA systems," *IEEE Trans. Signal Process.*, vol. 53, pp. 1110-1120, Mar. 2005.

[17] C.-C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 1539-1544, Mar. 2010.

[18] S. Abdallah and I. Psaromiligkos, "Blind channel estimation for amplify-and-forward two-way relay networks employing M-PSK modulation," *IEEE Trans. Signal Process.*, vol. 60, pp. 3604-3615, July 2012.

[19] X. Liao, L. Fan, and F. Gao, "Blind channel estimation for OFDM modulated two-way relay network," in *Proc. IEEE Wireless Commun. Network. Conf.*, Sydney, Australia, Apr. 18-21, 2010.

[20] Y. Xiang, D. Peng, Y. Xiang, and S. Guo, "Novel Z-domain precoding method for blind separation of spatially correlated signals," *IEEE Trans. Neural Networks and Learning Systems*, vol. 24, pp. 94-105, Jan. 2013.

[21] H. Furukawa, Y. Kamio, and H. Sasaoka, "Cochannel interference reduction and path-diversity reception technique using CMA adaptive array antenna in digital land mobile communications," *IEEE Trans. Veh. Technol.*, vol. 50, pp. 605-616, Mar. 2001.

[22] J. J. Shynk, A. V. Keerthi, and A. Mathur, "Steady-state analysis of the multistage constant modulus array," *IEEE Trans. Signal Process.*, vol. 44, pp. 948-962, Apr. 1996.

[23] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Singapore: John Wiley & Sons (Asia), 2010.

[24] S. Noh, Y. Sung, and M. D. Zoltowski, "A new precoder design for blind channel estimation in MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 7011-7024, Dec. 2014.
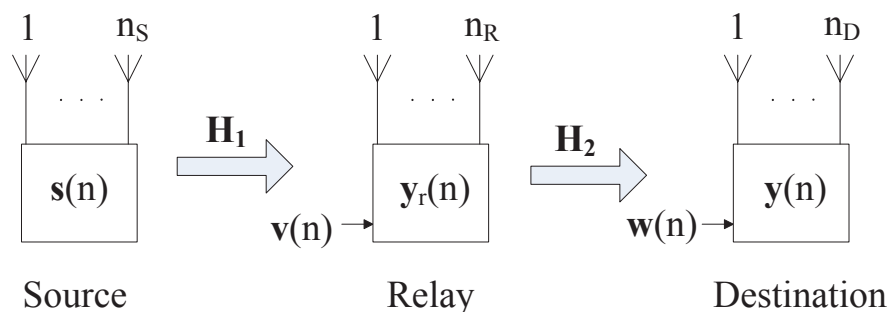
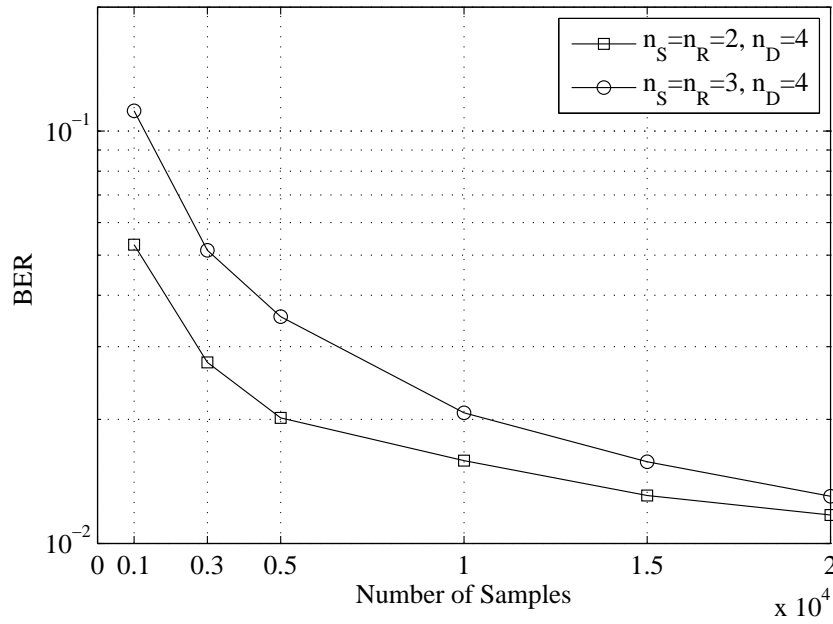Fig. 1. Block diagram of a general two-hop MIMO relay communication system.

Fig. 2. Example 1: BER versus number of samples for various n_S and n_R with $\text{SNR}_{r-d} = \text{SNR}_{s-r} = 20$dB and $n_D = 4$.
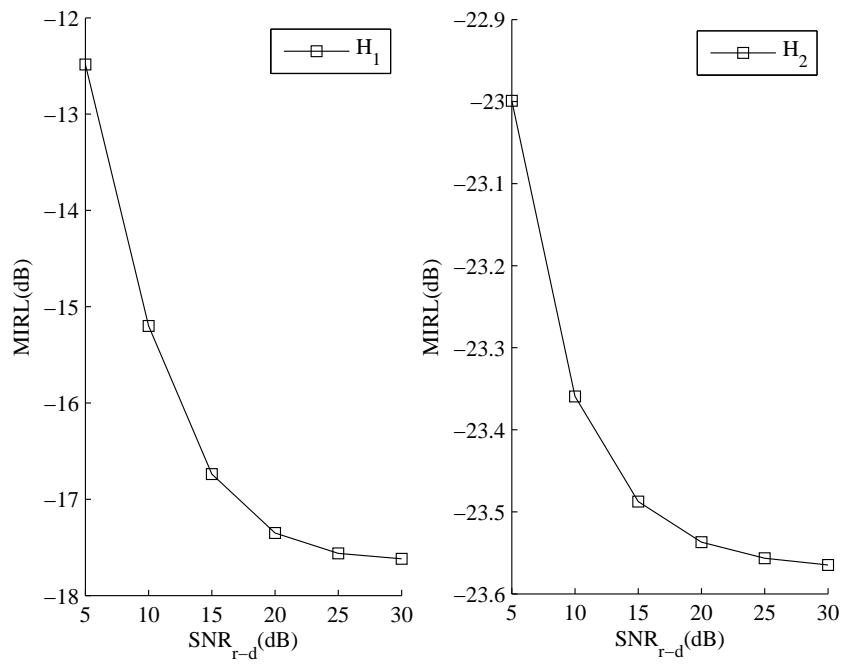

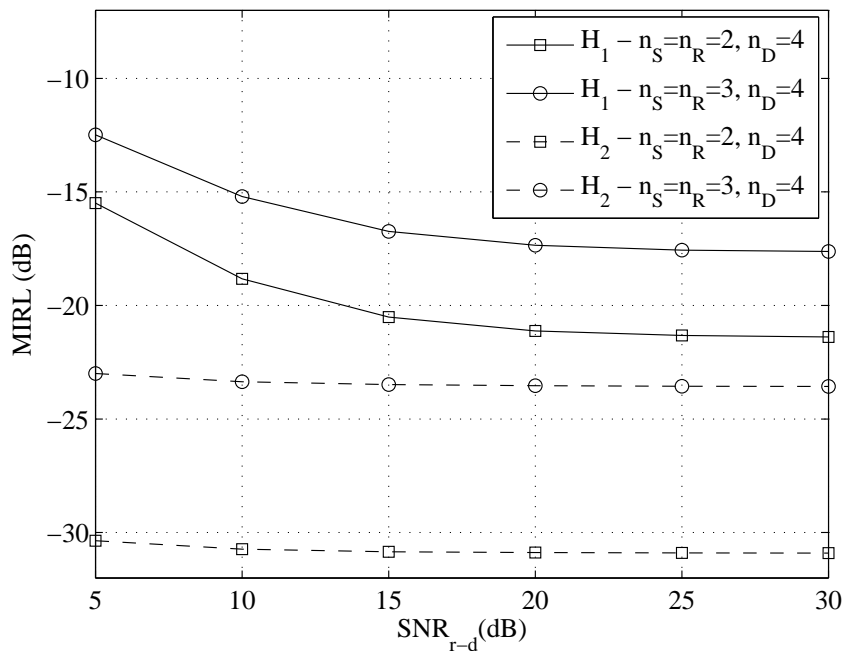Fig. 3. Example 2: MIRL versus SNR_{r-d} for $n_S = n_R = 3$, $n_D = 4$, and $\text{SNR}_{s-r} = 20$dB.

Fig. 4. Example 2: MIRL versus $\mathrm{SNR_{r-d}}$ for various n_S and n_R with $\mathrm{SNR_{s-r}} = 20$dB and $n_D = 4$.
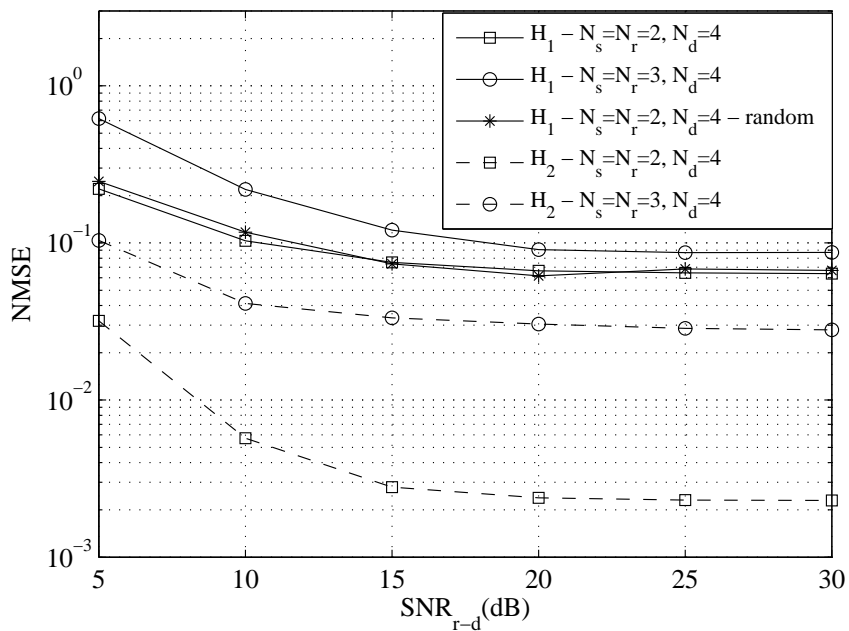
Fig. 5. Example 3: Normalized MSE versus $\mathrm{SNR_{r-d}}$ for various n_S and n_R with $\mathrm{SNR_{s-r}} = 20$dB and $n_D = 4$.
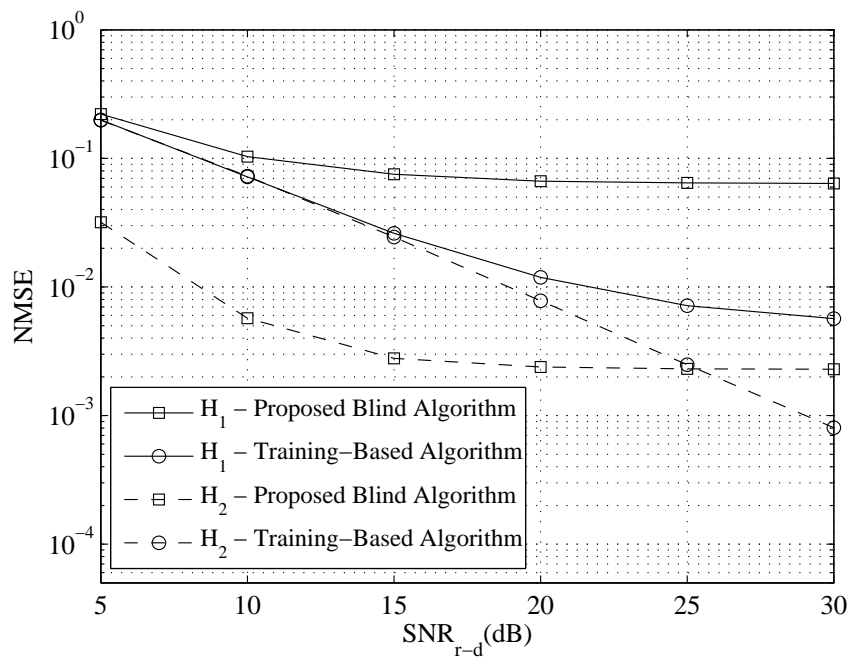
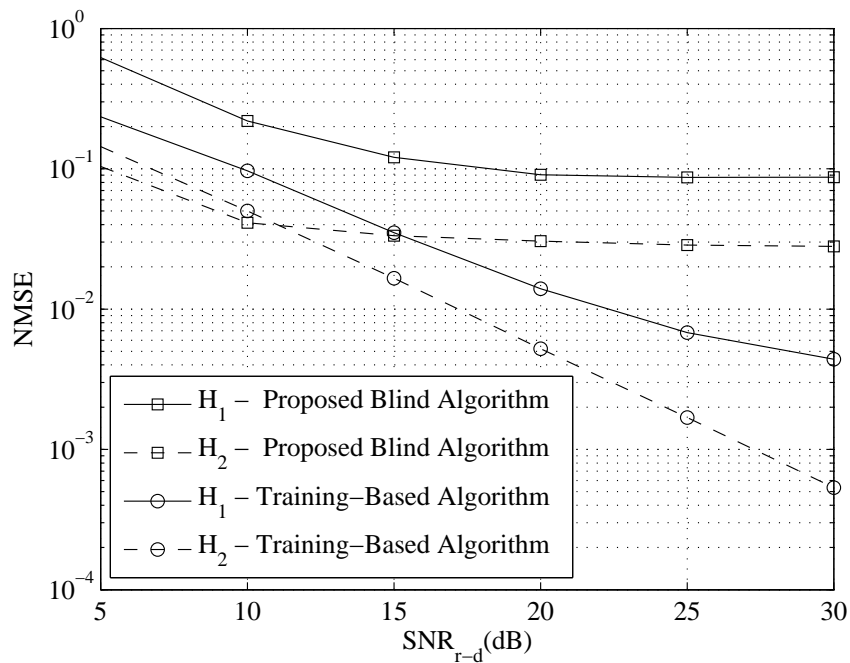Fig. 6. Example 4: Normalized MSE versus $\text{SNR}_{\text{r-d}}$ for $n_S = n_R = 2$ and $n_D = 4$ with $\text{SNR}_{\text{s-r}} = 20\text{dB}$.

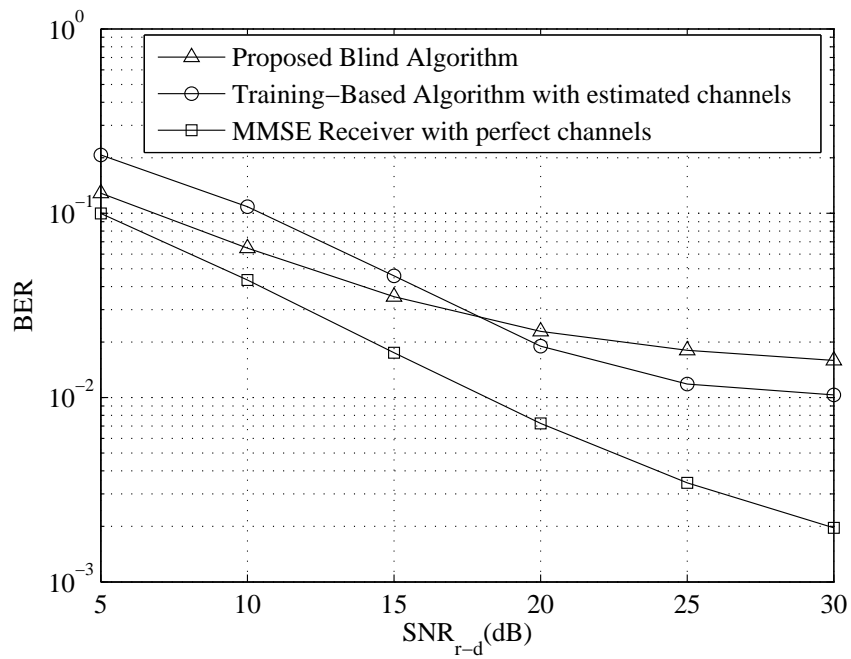Fig. 7. Example 4: Normalized MSE versus $\text{SNR}_{\text{r-d}}$ for $n_S = n_R = 3$ and $n_D = 4$ with $\text{SNR}_{\text{s-r}} = 20\text{dB}$.

Fig. 8. Example 4: BER versus $\text{SNR}_{\text{r}-\text{d}}$ for $n_S = n_R = 2$ and $n_D = 4$ with $\text{SNR}_{\text{s}-\text{r}} = 20\text{dB}$.