SVD-Based Channel Estimation for MIMO Relay Networks

Xinwei Yu
Mathematical and Statistical Science
University of Alberta, Edmonton, Canada
Email: xinweiyu@math.ualberta.ca

Yindi Jing
Electrical and Computer Engineering
University of Alberta, Edmonton, Canada
Email: yindi@ualberta.ca

***Abstract*—For a general multi-input-multi-output (MIMO) relay network, an estimation method for the receiver to obtain the end-to-end channels is proposed. Instead of straightforwardly estimating entries of the end-to-end channel matrix, the proposed scheme takes into consideration the special structure of the end-to-end channel matrix. By parameterizing the channel matrix with its singular values and singular vectors using singular value decomposition (SVD), the proposed scheme estimates the singular values and left and right singular vectors, which are then combined to form an estimation of the overall channel matrix. The proposed estimation follows the maximum-likelihood (ML) estimation method. Simulations on the mean square error (MSE) of the channel estimation are presented, which show the advantage of the proposed scheme over straightforward estimation of the channel entries for networks whose transmitter and receiver are equipped with multiple antennas.**

I. Introduction

Spatial diversity is one of the most effective techniques in wireless communications to mitigate the fading effect of wireless channels and thus improve the performance. In recent years, cooperative relay network, where relay nodes cooperate to establish a multi-input-multi-output (MIMO) communication link between a transmitter and a receiver, has been shown to be a promising infrastructure to achieve spatial diversity in a wireless network. Early researches in cooperative relay network focus on coherent cooperative schemes, e.g., decode-and-forward, amplify-and-forward (AF), distributed space-time coding (DSTC), distributed relay beamforming, in which perfect channel state information (CSI) is usually assumed at some or all of the network nodes.

In reality, a training process, sometimes, also a CSI feedback and feed-forward process, is required for network nodes to obtain their required CSI. Due to the existence of noise in training, information feedback, and information forward, the obtained CSI is always imperfect. Thus, the training design and the channel estimation rule that can lead to optimal performance, e.g., minimum mean square error (MSE) on channel information, are important and practical problems in wireless network communications.

These problems of training design and estimation methods have drawn increasing attention recently for networks with both regenerative and non-regenerative relays [1]–[10]. The training problem for non-regenerative relay network [2]–[10] is particularly interesting. For non-regenerative schemes such as AF and DSTC, the receiver is usually required to know the end-to-end equivalent channel coefficients from antennas of the transmitter to antennas of the receiver through antennas of the relays. Due to the concatenation of the channels of different transmission phases, the end-to-end channel coefficients are correlated with each other, which can lead to training models and estimation results that are very different from those of a point-to-point MIMO system [11]–[13].

For a single-antenna network with a single non-regenerative relay, the training and channel estimation problems are investigated in [2]–[7]. In [2]–[4], the channel training and the diversity performance with channel estimation error using mismatched maximum likelihood (ML) receiver are studied. In both [2] and [3], channel estimation is performed at the receiver. In [2], ML and linear-minimum-mean-square-error (LMMSE) estimators are employed, and DSTC is used for the data transmission. It is shown that full diversity can be achieved with both estimators. A similar conclusion is drawn in [3] where Golden code is used for data transmission. In [4], a scheme, where the transmitter-relay channel is estimated at the relay and forwarded to the receiver, is proposed. In [5], the optimal power allocations between training and data transmission and between the broadcasting phase and the relaying phase in training that maximize the signal-to-noise ratio (SNR) are investigated. In [6], the joint optimization of the training time, the power allocation between training and data transmission, and the power allocation between the transmitter and the relay that maximizes a mutual information lower bound is investigated. For a relay network with multiple relays but a single antenna at each node, the training problem is studied in [7]–[9]. In [7], a training scheme for the receiver to estimate the end-to-end channel coefficient is proposed for an AF scheme with matched filter. [8] and [9] are on the estimation of the end-to-end channel matrix at the receiver using DSTC.

The aforementioned papers are on the training problem for single-antenna relay networks, i.e., networks whose nodes have single antenna only. For the general MIMO relay networks with multiple transmit and receive antennas and multiple relays, in [14], training schemes for the receiver to estimate both the transmitter-relay channel matrix and relay-receiver channel matrix are proposed. The LMMSE estimations and the optimal pilot designs that minimize the MSE are derived.

978-1-4673-1881-5/12/$31.00 ©2012 IEEE

The requirement on the training time that can lead to full diversity in data transmission with mismatched ML decoding is also investigated. In [15], both mismatched and matched decodings for networks with multiple relays and multiple receive antennas under channel estimation error are analyzed. Then an adaptive decoding is proposed to balance the network performance and decoding complexity.

In this paper, we consider the general MIMO relay network and study the estimation of the end-to-end channel matrix. We first clarify the difference of this work to previous ones. The end-to-end channel matrix estimation is investigated in [8], [9] for single-antenna relay network only, in which entries of the end-to-end channel vector are estimated directly. In this paper, we investigate networks whose nodes have multiple antennas. The proposed estimation is via an approach totally different to previous ones. Instead of directly estimating entries of the end-to-end channel matrix, we take into consideration the special structure of the channel matrix. By using the singular value decomposition (SVD), the channel matrix is parameterized by its largest singular value and left and right singular vectors. Estimations of the largest singular value and the left and right singular vectors are first derived based on the ML criterion. They are then used to construct an estimation of the channel matrix. Compared with the entry-based estimation, which ignores the special structure of the end-to-end channel matrix and estimates each channel entry, the proposed estimation is shown by simulation to be superior in MSE.

The rest of the paper is organized as follows. In Section II, the cooperative relay network model and the channel estimation problems are explained. Section III is on the proposed SVD-based ML estimation of the end-to-end channel matrix at the receiver. Numerically simulated MSEs are shown in Section IV. Section V contains the conclusions.

Here are the notation used in this paper. We use bold upper case letters to denote matrices and bold lower case letters to denote vectors, which can be either row vectors or column vectors. For a matrix $\mathbf{A}$, its conjugate, transpose, Hermitian, Frobenius norm, determinant, trace, and inverse are denoted by $\overline{\mathbf{A}}$, $\mathbf{A}^t$, $\mathbf{A}^*$, $\|\mathbf{A}\|_F$, $\det\mathbf{A}$, $\mathrm{tr}\mathbf{A}$, and $\mathbf{A}^{-1}$, respectively. $\mathrm{vec}(\mathbf{A})$ is the column vector formed by stacking the columns of $\mathbf{A}$. $\mathbf{I}_n$ is the $n \times n$ identity matrix. $\mathbf{0}_n$ is the n-dimensional vector of all zeros. $\otimes$ denotes the Kronecker product. $\mathbb{E}(\cdot)$ is the average operator for a random variable, a random vector, or a random matrix. We use $\hat{a}$ to denote the estimation of a.

II. MIMO Relay Network Model and Channel Estimation Problem

We consider the general MIMO multiple-relay network as shown in Figure 1. Assume that the transmitter has M antennas, the receiver has N antennas, and there are in total R antennas at the relays. The relay antennas can be either co-located at one relay or distributively located at different relays. Denote the channel from the mth antenna of the transmitter to the rth relay antenna as f_{mr} and the channel from the rth relay antenna to the nth antenna of the receiver as g_{rn}. The channels are assumed to be i.i.d. $\mathcal{CN}(0,1)$, i.e., independent

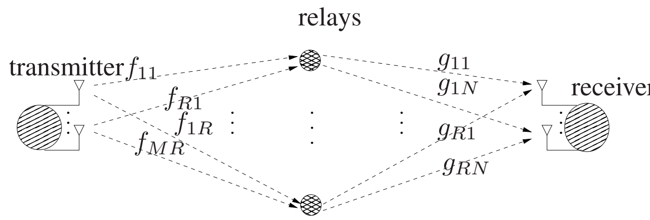

Fig. 1. MIMO multiple-relay network.

and identically distributed Rayleigh flat-fading channels. The channels are also assumed to remain constant during the training process.

For some cooperation schemes such as non-regenerative DSTC [16], [17] and AF with fixed gain relay power coefficient [18], [19], the end-to-end channel matrix is required to be known at the receiver. The end-to-end channel from the mth transmit antenna to the nth receive antenna via the rth relay antenna is $f_{mr}g_{rn}$. By using matrix and vector representation, for the MIMO relay network, the $M \times N$ end-to-end channel matrix from the transmitter to the receiver via the rth relay antenna is

$$\mathbf{H}_r \triangleq \begin{bmatrix} f_{1r}g_{r1} & f_{1r}g_{r2} & \cdots & f_{1r}g_{rN} \\ f_{2r}g_{r1} & f_{2r}g_{r2} & \cdots & f_{2r}g_{rN} \\ \vdots & \vdots & \ddots & \vdots \\ f_{Mr}g_{r1} & f_{Mr}g_{r2} & \cdots & f_{Mr}g_{rN} \end{bmatrix}$$

Define

$$\mathbf{f}_r \triangleq \begin{bmatrix} f_{1r} & f_{2r} & \cdots & f_{Mr} \end{bmatrix}^t$$
$$\mathbf{g}_r \triangleq \begin{bmatrix} g_{r1} & g_{r2} & \cdots & g_{rN} \end{bmatrix}.$$

$\mathbf{f}_r$ and $\mathbf{g}_r$ are thus the channel vector from the transmitter to the rth relay antenna and the channel vector from the rth relay antenna to the receiver. We have,

$$\mathbf{H}_r = \mathbf{f}_r\mathbf{g}_r. \tag{1}$$

To represent all end-to-end channel coefficients from the transmitter antennas to the receiver antennas via all relay antennas, we can write $\mathbf{H}_1, \cdots, \mathbf{H}_R$ into a column to obtain the $MR \times N$ end-to-end channel matrix:

$$\mathbf{H}_{\mathrm{E2E}} \triangleq \begin{bmatrix} \mathbf{H}_1^t & \mathbf{H}_2^t & \cdots & \mathbf{H}_R^t \end{bmatrix}^t.$$

The end-to-end channel estimation problem is thus to find an estimation of $\mathbf{H}_{\mathrm{E2E}}$ at the receiver.

Since channels related to different relay antennas are independent, we can solve the channel estimation problem for a network with multiple relay antennas by solving the estimation problem related to each relay antenna one by one. In other words, we consider the estimation of $\mathbf{H}_r$ for different r separately and sequentially. Thus, without loss of generality, we only need to consider the estimation of the channels corresponding to one relay antenna, i.e., a network with single relay antenna. In what follows, we investigate the estimation of $\mathbf{H}_r$. For simplicity of presentation, we omit the index r for relay antennas.

III. SVD-Based ML Channel Estimation

In this section, we propose the training and estimation of the end-to-end channel matrix at the receiver. The training process is presented first, followed by the ML estimation.

A. Training

We conduct a two-step training process, where in the first step, space-time coding is used at the transmitter and in the second step, the relay antenna conducts AF with fixed gain power coefficient. Each training step takes M symbol transmissions. Thus, the total training time is $2M$.

During the first step, the transmitter sends $\sqrt{P_1}\mathbf{S}$, where the $M\times M$ matrix $\mathbf{S}$ is the pilot, normalized as $\mathrm{tr}\{\mathbf{S}^*\mathbf{S}\}=M$. This normalization implies that the average transmit power of the transmitter is P_1 per transmission. The relay receives:

$$\mathbf{r}=\sqrt{P_1}\mathbf{S}\mathbf{f}+\mathbf{u},$$

where $\mathbf{u}$ is the noise vector at the relay. We assume that the noises are i.i.d. circularly symmetric complex Gaussian (CSCG) with zero-mean and unit-variance, i.e., $\mathbf{u}\sim\mathcal{CN}(\mathbf{0}_M,\mathbf{I}_M)$. During the second step, the relay antenna amplifies its received signal and forwards to the receiver with power P_2. The fixed gain power coefficient $\sqrt{\frac{P_2}{P_1+1}}$ is used. Define

$$\alpha\triangleq\frac{P_2}{P_1+1}\qquad\beta\triangleq\frac{P_1P_2}{P_1+1}.$$

The signal matrix received at the receiver, denoted as $\mathbf{X}$, can be calculated to be [17]:

$$\mathbf{X}=\sqrt{\beta}\mathbf{S}\mathbf{H}+\mathbf{W},\tag{2}$$

where

$$\mathbf{W}\triangleq\sqrt{\alpha}\mathbf{u}\mathbf{g}+\mathbf{V}$$

with $\mathbf{V}$ the $M\times N$ noise matrix at the receiver. Again, we assume that entries of $\mathbf{V}$ are i.i.d. CSCG with zero-mean and unit-variance, i.e., $\mathrm{vec}(\mathbf{V})\sim\mathcal{CN}(\mathbf{0}_{MN},\mathbf{I}_{NM})$. For a given $\mathbf{g}$, $\mathrm{vec}(\mathbf{W})$ can be shown to be a CSCG random vector with zero mean. Its covariance matrix can be calculated to be

$$\mathbf{R}_{\mathbf{W}}\triangleq\left(\mathbf{I}_N+\alpha\mathbf{g}^t\bar{\mathbf{g}}\right)\otimes\mathbf{I}_M.\tag{3}$$

B. SVD-Based ML Estimation

The end-to-end channel estimation problem is to estimate $\mathbf{H}$ base on the observation $\mathbf{X}$ in (2). The training equation (2) has the same format as a Gaussian observation model or the training equation for a multiple-antenna direct communication system without relaying [11]. However, it has two main differences to a traditional Gaussian estimation problem. First, in a traditional Gaussian model, the noise term is assumed to be independent to the parameters to be estimated. Here, however, $\mathbf{R}_{\mathbf{W}}$, the covariance matrix of the noise $\mathbf{W}$, is a function of $\mathbf{g}$, which depends on $\mathbf{H}$, the matrix to be estimated. Second, in a traditional Gaussian model, entries of the vector/matrix to be estimated are assumed to be independent. Here, for the MIMO relay network, due to the special structure of $\mathbf{H}$ defined in (1), its rank is 1. Thus, entries of $\mathbf{H}$ are related. Of the total MN entries in $\mathbf{H}$, only $M+N-1$ of them are independent. These make the estimation problem of the MIMO relay network different and more challenging.

In this paper, instead of representing $\mathbf{H}$ by its $M+N-1$ independent entries, we decompose $\mathbf{H}$ as

$$\mathbf{H}=a\tilde{\mathbf{f}}\tilde{\mathbf{g}},\tag{4}$$

where

$$a\triangleq\|\mathbf{f}\|_F\|\mathbf{g}\|_F,\quad\tilde{\mathbf{f}}\triangleq\frac{\mathbf{f}}{\|\mathbf{f}\|_F},\quad\text{and}\quad\tilde{\mathbf{g}}\triangleq\frac{\mathbf{g}}{\|\mathbf{g}\|_F}.\tag{5}$$

$\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}$ are the directions of $\mathbf{f}$ and $\mathbf{g}$, respectively. They have unit-norm. The decomposition in (5) is the SVD of $\mathbf{H}$. a is the largest singular value of $\mathbf{H}$, which is also the only non-zero singular value. $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{g}}$ are the corresponding left and right singular vectors.

Since entries of $\mathbf{f}$ and $\mathbf{g}$ are i.i.d. $\mathcal{CN}(0,1)$, $a,\tilde{\mathbf{f}},\tilde{\mathbf{g}}$ are mutually independent. The estimation of $\mathbf{H}$ can thus be transformed into the estimation of $(a,\tilde{\mathbf{f}},\tilde{\mathbf{g}})$. The following theorem on the ML estimation of $\mathbf{H}$ is proved.

Theorem 1: When the pilot matrix $\mathbf{S}$ is nonsingular, let $\mathbf{X}=\sum_{i=1}^{\min\{M,N\}}\sigma_{\mathbf{X},i}\mathbf{u}_i\mathbf{v}_i^*$ be the singular value decomposition, where $\sigma_{\mathbf{X},1}\geq\sigma_{\mathbf{X},2}\geq\cdots\geq\sigma_{\mathbf{X},\min\{M,N\}}\geq 0$ are the ordered singular values of $\mathbf{X}$ and $\mathbf{u}_i$ and $\mathbf{v}_i$ are the left and right singular vectors corresponding to $\sigma_{\mathbf{X},i}$. The ML estimation of $(a,\tilde{\mathbf{f}},\tilde{\mathbf{g}})$ is

$$\hat{a}=\frac{1}{\sqrt{\beta}}\sigma_{\mathbf{X},1}\|\mathbf{S}^{-1}\mathbf{u}_1\|_F,\quad\hat{\tilde{\mathbf{f}}}=\frac{\mathbf{S}^{-1}\mathbf{u}_1}{\|\mathbf{S}^{-1}\mathbf{u}_1\|_F},\quad\hat{\tilde{\mathbf{g}}}=\mathbf{v}_1^*.\tag{6}$$

The ML estimation of the end-to-end channel matrix is thus

$$\hat{\mathbf{H}}=\frac{\sigma_{\mathbf{X},1}}{\sqrt{\beta}}\mathbf{S}^{-1}\mathbf{u}_1\mathbf{v}_1^*.\tag{7}$$

Proof: For a given channel realization, the probability density function (PDF) of $\mathbf{X}|\mathbf{f},\mathbf{g}$ is

$$p(\mathbf{X}|\mathbf{f},\mathbf{g})=\frac{1}{(2\pi)^{TN}\det\mathbf{R}_{\mathbf{W}}}e^{-\mathrm{vec}\left(\mathbf{X}-\sqrt{\beta}\mathbf{S}\mathbf{H}\right)^*\mathbf{R}_{\mathbf{W}}^{-1}\mathrm{vec}\left(\mathbf{X}-\sqrt{\beta}\mathbf{S}\mathbf{H}\right)}.$$

Using the decomposition of $\mathbf{H}$ in (4), we obtain an estimation for $\mathbf{H}$ by estimating a, $\tilde{\mathbf{f}}$, and $\tilde{\mathbf{g}}$. Notice that

$$\det\mathbf{R}_{\mathbf{W}}=\det{}^T\left(\mathbf{I}_N+\alpha\mathbf{g}^t\bar{\mathbf{g}}\right)\det{}^N\mathbf{I}_T=\left(1+\alpha\|\mathbf{g}\|_F^2\right)^T,$$

which does not dependent on the values of $a,\tilde{\mathbf{f}},\tilde{\mathbf{g}}$. The ML estimation finds the parameter values that maximize the conditional PDF, i.e.,

$$\left(\hat{a},\hat{\tilde{\mathbf{f}}},\hat{\tilde{\mathbf{g}}}\right)=\arg\max_{(a,\tilde{\mathbf{f}},\tilde{\mathbf{g}})}p(\mathbf{X}|\mathbf{f},\mathbf{g})=$$
$$\arg\min_{(a,\tilde{\mathbf{f}},\tilde{\mathbf{g}})}\mathrm{vec}\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right)^*\mathbf{R}_{\mathbf{W}}^{-1}\mathrm{vec}\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right).\tag{8}$$

From (3),

$$\mathbf{R}_{\mathbf{W}}=\left(\mathbf{I}_N+\alpha\|\mathbf{g}\|_F^2\tilde{\mathbf{g}}^t\bar{\tilde{\mathbf{g}}}\right)\otimes\mathbf{I}_M.$$

Thus,

$$\mathbf{R}_{\mathbf{W}}^{-1}=\left(\mathbf{I}_N+\alpha\|\mathbf{g}\|_F^2\tilde{\mathbf{g}}^t\bar{\tilde{\mathbf{g}}}\right)^{-1}\otimes\mathbf{I}_M.$$

We recall the following identities for Kronecker product and vectorization:

$$\mathrm{vec}\left(\mathbf{ABC}\right)=\left(\mathbf{C}^t\otimes\mathbf{A}\right)\mathrm{vec}\left(\mathbf{B}\right),$$
$$\mathrm{vec}\left(\mathbf{A}\right)^*\mathrm{vec}\left(\mathbf{B}\right)=\mathrm{tr}\left(\mathbf{A}^*\mathbf{B}\right).$$

Using these equalities, after some calculations, we have

$$\mathrm{vec}\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right)^*\mathbf{R}_{\mathbf{W}}^{-1}\mathrm{vec}\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right)=$$
$$\mathrm{tr}\left[\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right)^*\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right)\left(\mathbf{I}_N+\alpha\|\mathbf{g}\|_F^2\tilde{\mathbf{g}}^*\tilde{\mathbf{g}}\right)^{-1}\right].\quad(9)$$

Also notice that

$$\begin{aligned}&\left(\mathbf{I}_N+\alpha\|\mathbf{g}\|_F^2\tilde{\mathbf{g}}^*\tilde{\mathbf{g}}\right)^{-1}\\&=\mathbf{I}_N-\tilde{\mathbf{g}}^*\left(\left(\alpha\|\mathbf{g}\|_F^2\right)^{-1}+\tilde{\mathbf{g}}\tilde{\mathbf{g}}^*\right)^{-1}\tilde{\mathbf{g}}\\&=\mathbf{I}_N-m\tilde{\mathbf{g}}^*\tilde{\mathbf{g}},\end{aligned}\quad(10)$$

where

$$m\triangleq\frac{\alpha\|\mathbf{g}\|_F^2}{1+\alpha\|\mathbf{g}\|_F^2}.$$

The object function in (8) can thus be simplified to

$$\begin{aligned}&F\left(a,\tilde{\mathbf{f}},\tilde{\mathbf{g}}\right)\\&\triangleq\mathrm{tr}\left[\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right)^*\left(\mathbf{X}-a\sqrt{\beta}\mathbf{S}\tilde{\mathbf{f}}\tilde{\mathbf{g}}\right)\left(\mathbf{I}_N-m\tilde{\mathbf{g}}^*\tilde{\mathbf{g}}\right)\right]\\&=\left(1-m\right)\beta\|\mathbf{S}\tilde{\mathbf{f}}\|_F^2\left(a-\frac{\Re\left(\tilde{\mathbf{f}}^*\mathbf{S}^*\mathbf{X}\tilde{\mathbf{g}}^*\right)}{\sqrt{\beta}\|\mathbf{S}\tilde{\mathbf{f}}\|_F^2}\right)^2\\&\quad-(1-m)\frac{\Re^2\left(\tilde{\mathbf{f}}^*\mathbf{S}^*\mathbf{X}\tilde{\mathbf{g}}^*\right)}{\|\mathbf{S}\tilde{\mathbf{f}}\|_F^2}+\|\mathbf{X}\|_F^2-m\|\mathbf{X}\tilde{\mathbf{g}}^*\|_F^2.\end{aligned}$$

Thus to minimize $F\left(a,\tilde{\mathbf{f}},\tilde{\mathbf{g}}\right)$, we need

$$a=\frac{\Re\left(\tilde{\mathbf{f}}^*\mathbf{S}^*\mathbf{X}\tilde{\mathbf{g}}^*\right)}{\sqrt{\beta}\|\mathbf{S}\tilde{\mathbf{f}}\|_F^2},\quad(11)$$

from the first term of $F\left(a,\tilde{\mathbf{f}},\tilde{\mathbf{g}}\right)$.

With this choice of a, the optimization problem in (8) reduces to

$$\left(\hat{\tilde{\mathbf{f}}},\hat{\tilde{\mathbf{g}}}\right)=\arg\max_{(\tilde{\mathbf{f}},\tilde{\mathbf{g}})}G\left(\tilde{\mathbf{f}},\tilde{\mathbf{g}}\right),$$

where

$$G\left(\tilde{\mathbf{f}},\tilde{\mathbf{g}}\right)\triangleq(1-m)\frac{\Re^2\left[\left(\mathbf{S}\tilde{\mathbf{f}}\right)^*\mathbf{X}\tilde{\mathbf{g}}^*\right]}{\|\mathbf{S}\tilde{\mathbf{f}}\|_F^2}+m\|\mathbf{X}\tilde{\mathbf{g}}^*\|_F^2.$$

We have

$$G\left(\tilde{\mathbf{f}},\tilde{\mathbf{g}}\right)\le(1-m)\frac{\|\mathbf{S}\tilde{\mathbf{f}}\|_F^2\|\mathbf{X}\tilde{\mathbf{g}}^*\|_F^2}{\|\mathbf{S}\tilde{\mathbf{f}}\|_F^2}+m\|\mathbf{X}\tilde{\mathbf{g}}^*\|_F^2=\|\mathbf{X}\tilde{\mathbf{g}}^*\|_F^2$$

with equality when $\mathbf{S}\tilde{\mathbf{f}}=\gamma\mathbf{X}\tilde{\mathbf{g}}^*$ for some real γ. This is always possible for non-singular pilot $\mathbf{S}$ and the solution is $\tilde{\mathbf{f}}=\gamma\mathbf{S}^{-1}\mathbf{X}\tilde{\mathbf{g}}^*$. Since $\tilde{\mathbf{f}}$ has unit-norm, we have

$$\tilde{\mathbf{f}}=\frac{\mathbf{S}^{-1}\mathbf{X}\tilde{\mathbf{g}}^*}{\|\mathbf{S}^{-1}\mathbf{X}\tilde{\mathbf{g}}^*\|_F}.\quad(12)$$

With optimal choice of $\tilde{\mathbf{f}}$, the optimization becomes the maximization of $\|\mathbf{X}\tilde{\mathbf{g}}^*\|_F^2$. With the SVD of $\mathbf{X}$ given in Theorem 1, it is clear that $\|\mathbf{X}\tilde{\mathbf{g}}^*\|_F^2$ is maximized when $\tilde{\mathbf{g}}^*=\mathbf{v}_1$, the right singular vector corresponding to the largest singular value of $\mathbf{X}$. That is, the ML estimation of $\tilde{\mathbf{g}}$ is $\hat{\tilde{\mathbf{g}}}=\mathbf{v}_1^*$. Further, using this result in (12) then (11), the ML estimation is obtained as in (6) and (7). ■

The estimation $\hat{\mathbf{H}}$ in (7) is rank-one, which conforms to the structure of the end-to-end channel matrix in (1).

In particular, when the pilot matrix is designed as the identity matrix, $\mathbf{S}=\mathbf{I}_M$, that is, the transmitter antennas send unit sequentially, following Theorem 1, the ML estimation is

$$\hat{a}=\frac{1}{\sqrt{\beta}}\sigma_{\mathbf{X},1},\ \hat{\tilde{\mathbf{f}}}=\mathbf{u}_1,\ \hat{\tilde{\mathbf{g}}}=\mathbf{v}_1^*,\ \hat{\mathbf{H}}=\frac{\sigma_{\mathbf{X},1}}{\sqrt{\beta}}\mathbf{u}_1\mathbf{v}_1^*.$$

The ML estimation on a is a scaled version of the largest singular value of the observation matrix $\mathbf{X}$; while the ML estimations on the directions of $\mathbf{f}$ and $\mathbf{g}$ are the left and right singular vectors corresponding to the largest single value.

IV. Simulation Results

In this section, we simulate the MSE of the proposed estimation. Since the focus of this paper is to propose a new estimation rule that takes into consideration the special structure of the end-to-end channel matrix, optimizations of the training pilot, power allocation, and training time are not considered. They are interesting and important issues but are out the focus of this paper, and left for future work. We set $\mathbf{S}=\mathbf{I}_M$ and $P_1=P_2=P$ to avoid optimizations of the pilot and the power allocation.

The MSE of the end-to-end channel estimation is defined as $\mathrm{MSE}(\hat{\mathbf{H}})\triangleq\mathbb{E}\|\mathbf{H}-\hat{\mathbf{H}}\|_F^2$. Note that in traditional ML estimation, the MSE is the power of the estimation error averaged over the noises. In our estimation model, in addition to the noises, the channels are also random, the average is over not only the noises but also the random channels. That is, in the conducted Monte-Carlo simulation, a distinct channel realization is used for each iteration. The channels are modeled as CSCG random variables with zero-mean and unit variance, whose amplitude is Rayleigh distributed. The calculated MSE is the average power of the estimation error over a large number of channel realizations.

We compare the proposed SVD-based estimation in Theorem 1 with an entry-based estimation scheme explained in what follows. From (2), if neither the special structure of $\mathbf{H}$ nor the relation between $\mathbf{H}$ and the covariance matrix of the noise term $\mathbf{W}$ is taken into account, the estimation problem can be seen as a traditional Gaussian one and the ML estimation is

$$\hat{\mathbf{H}}=\frac{1}{\sqrt{\beta}}(\mathbf{S}^*\mathbf{S})^{-1}\mathbf{S}^*\mathbf{X}.\quad(13)$$

In Figure 2, we show the MSE of $\mathbf{H}$ as a function of the training power P of relay networks with $M=N=2$ and $M=N=5$. We can see that for both networks, the proposed SVD-based estimation is better than the entry-based estimation

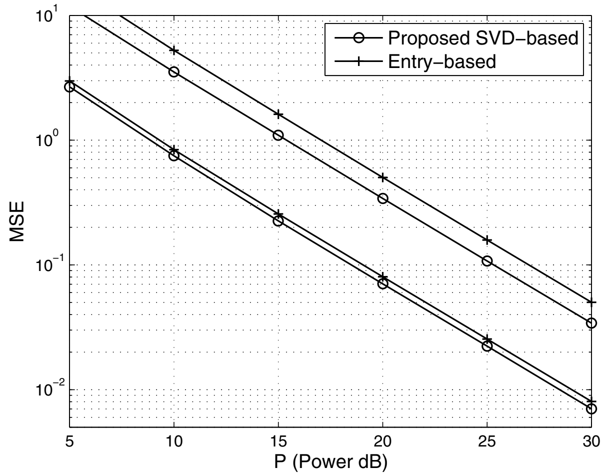

Fig. 2. MSE of $\hat{\mathbf{H}}$ for networks with $M = N = 2$ and $M = N = 5$.

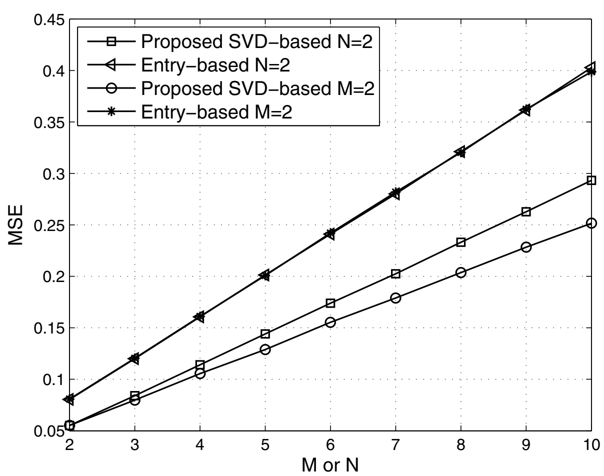

Fig. 3. MSE of $\hat{\mathbf{H}}$ of a network with different M or N at $P = 20$dB.

by about 2dB. For both the proposed SVD-based and the entry-based estimations, the MSEs have a linear decreasing rate with respect to the training power. This is important in achieving full diversity in data transmission [15].

In Figure 3, we show how the MSE behaves as the numbers of transmitter and receiver antennas change. We consider networks with $M = 2$ while N changes from 2 to 10 and also networks with $N = 2$ while M changes from 2 to 10. P is set to be 20dB. We can see that for both estimations, the MSE increases as M or N increases since the dimension of $\mathbf{H}$ increases, but the proposed SVD-based estimation has a smaller increase rate. Its advantage over the entry-based estimation is prominent, especially for large values of transmitter antennas M or receiver antennas N.

V. Conclusions

In this paper, we considered the channel estimation for a general MIMO relay network. The proposed scheme follows the ML estimation method. By using the singular value decomposition, estimations on the largest singular value and the corresponding left and right singular vectors are found, which are combined to obtain the ML estimation of the end-to-end channel matrix at the receiver. The proposed scheme requires no channel state information feedback or feed-forward and takes advantage the special structure of the end-to-end channel matrix of MIMO relay networks. Simulation results on the mean square error of the proposed estimation are shown and its advantage over the entry-based estimation is demonstrated.

References

[1] F. Gao, T. Cui, and A. Nallanathan, "Optimal training design for channel estimation in decode-and-forward relay networks with individual and total power constraints," *IEEE Trans. on Signal Processing*, vol. 56, pp. 5937-5949, Dec. 2008.

[2] H. Mheidat and M. Uysal, "Non-coherent and mismatched-coherent receivers for distributed STBCs with amplify-and-forward relaying," *IEEE Trans. on Wireless Commu.*, vol. 6, pp. 4060-4070, Nov. 2007.

[3] B. Gedik and M. Uysal, "Impact of imperfect channel estimation on the performance of amplify-and-forward relaying," *IEEE Trans. on Wireless Commu.*, vol. 8, pp. 1468-1479, Mar. 2009.

[4] O. Amin, B. Gedik, and M. Uysal, "Channel estimation for amplify-and-forward relaying: Cascaded against disintegrated estimators," *IET Commu.*, vol. 4, pp. 1207-1216, Jan. 2010.

[5] B. Gedik, and M. Uysal, " Training power optimization for amplify-and-forward cooperative systems," *Proc. of IEEE VTC 2008-Fall*, Sept. 2008, pp. 1-5.

[6] Y. Jia and A. Vosoughi, "Transmission resource allocation for training based amplify-and-forward relay systems," *IEEE Trans. on Wireless Commu.*, vol. 10, pp. 450-455, Feb. 2011.

[7] C. M. Yetis, and A. H. Kayran, "A new training protocol for channel state estimation in wireless relay networks," *IEEE Trans. on Signal Processing*, vol. 58, pp. 5950-5955, Nov. 2010.

[8] F. Gao, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. on Wireless Commu.*, vol. 7, pp. 1907-1916, May 2008.

[9] G. S. Rajan and B. S. Rajan, "Leveraging coherent distributed space-time codes for noncoherent communication in relay networks via training," *IEEE Trans. on Wireless Commu.*, vol. 8, pp. 683-688, Feb. 2009,

[10] F. Gao, R. Zhang, and Y. Liang, "On channel estimation for amplify-and-forward two-way relay networks," *Proc. of IEEE GLOBECOM*, Nov. 2008, pp. 1-5.

[11] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. on Information Theory*, vol. 49, pp. 951-963, Apr. 2003.

[12] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. on Signal Processing*, vol. 54, pp. 884893, Mar. 2006.

[13] A. Vosoughi, and A. Scaglione, "Everything you always wanted to know about training: Guidelines derived using the affine precoding framework and the CRB," *IEEE Trans. on Signal Processing,* vol. 54, pp. 940-954, Mar. 2006.

[14] S. Sun and Y. Jing, "Channel training design in amplify-and-forward MIMO relay networks," *IEEE Trans. on Wireless Commu.*, vol. 10, pp. 3380-3391, Oct. 2011.

[15] S. Sun and Y. Jing, "Training and decoding for cooperative network with multiple relays and receive antennas," *IEEE Trans. on Commu.*, 2012, Accepted.

[16] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. on Wireless Commun.*, vol. 5, pp. 3524-3536, Dec. 2006.

[17] Y. Jing and B. Hassibi, "Diversity analysis of distributed space-time codes in relay networks with multiple transmit/receive antennas," *EURASIP Jour. on Advanced Signal Processing*, 2008.

[18] R. U. Nabar, H. Bölcskei, and F. W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE Journ. on Selected Areas in Commu.* , pp. 1099-1109, Aug. 2004.

[19] C. Patel, G. Stüber, and T. Pratt, "Statistical properties of amplify and forward relay fading channels," *IEEE Trans. on Vehicular Technology,* vol. 55, pp. 19, Jan. 2006.