EM Algorithm based Channel Estimation for Amplify-and-Forward Relay Networks with Unknown Noise Correlation

Chao Zhang, Suhua Tang[1] Pinyi Ren
School of Electronic and Information Engineering, Xi'an Jiaotong University, China.
Email: {chaozhang, pyren}@mail.xjtu.edu.cn
[1]Department of Smart Networks, ATR Adaptive Communications Research Laboratories, Japan
Email: shtang@atr.jp

***Abstract*—Due to common interference or noise propagation, noise correlation between relays could occur in Amplify-and-Forward relay networks. To estimate channel coefficient, the noise covariance is required in traditional channel estimations. On the other hand, we also need channel coefficient to estimate the noise covariance. Therefore, traditional channel estimators can not be utilized when noise correlation is unknown. In this paper, we propose an Expectation-Maximization algorithm based iterative channel estimator to solve this problem. Moreover, we analyze the modified Cramér-Rao bound to show the performance of the proposed channel estimation. Finally, simulation results show that the proposed channel estimator can work well in Amplify-and-Forward relay networks with unknown noise correlation.**

I. Introduction

Relay-assisted communication is a promising strategy that exploits the spatial diversity available among a collection of distributed single antenna terminals for both centralized and decentralized wireless networks. In most relay networks, a two-stage relaying strategy is used. In the first stage, a source transmits and all relays listen; in the second stage, relays cooperatively forward the source symbols to the destination. Amplify-and-Forward (AF) and Decode-and-Forward (DF) are two most common relaying schemes [1]. With lower processing complexity at relays, AF based relay network can achieve full diversity as DF scheme at high power regime [3]. Therefore, AF based relay network recently has attracted many attentions [4]-[6]. We thus focus our attention on AF based relay networks.

Most work in literature [2]-[6] always assumes that the receiver noise at relays is independent in spatial. However, Gomadam and Jafar [7] showed that noise correlation between relays could occur in AF relay networks due to common interference or noise propagation. They also found that noise correlation is beneficial regardless of whether the relays know the noise covariance. Consider noise correlation at the relays, an optimized relay scheme under a receive power constraint was proposed in [8]. For AF relay networks without noise correlation, [9]-[11] had proposed various effective channel estimators which require noise covariance as an input variable. Thus, above channel estimators can not be utilized in the noise correlation case. To the best of our knowledge, the channel estimation problem in AF relay networks with noise correlation has not been investigated yet.

In this paper, we intend to propose an Expectation-Maximization (EM) algorithm based channel estimation for AF relay networks, where noise correlation at relays is considered and not known by the destination. If the noise is independent between relays, the destination can compute the noise variance easily. If there exists noise correlation at relays, the receiver with channel estimators designed by [9]-[11] can not work well. In order to achieve good performance, estimating the noise covariance also needs the knowledge of channel coefficients. As a result, these channel estimators fall into a dead loop. Thanks to the EM algorithm, which can perform maximum likelihood (ML) estimation in the presence of unobserved data, we treat the unknown data as unobserved data and propose an iterative estimator for both channel coefficient and noise covariance. Moreover, we analyze the modified Cramér-Rao bound to show the performance of the proposed channel estimation. Finally, the validity of channel estimator is also verified by simulation results.

II. System Model

We consider a wireless network with N relays, one source and one destination. Every node has a single antenna that can not transmit and receive simultaneously. Denote the channel coefficient from the source to the ith relay as f_i and the channel coefficient from the ith relay to the destination as g_i. Assume that all f_i and g_i are independent complex Gaussian random variables with zero-mean and variance δ_{si}^2 and δ_{id}^2, respectively [4]. Note that we suppose there is no direct channel between the source node and the destination node. We further assume a block fading channel model, where channel gain stays constant during a time block and changes from block to block. It is assumed that the instantaneous channel is unknown to the transmitting node but perfectly known at receiving node [1]. Moreover, all relays are synchronized during relaying phase. The impact of synchronization error between relays is beyond the scope of our discussion [3].

Suppose the source node has to transmit T symbols to the destination node. The signal transmission is divided into two phases. During the first phase, the source broadcasts

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

its signal $s[t]$ ($t = 1, ..., T$) to all relays. Assume $s[t]$ is randomly selected from a MPSK modulation constellation $\mathcal{A} = \{s_1, s_2, ..., s_M\}$ and $E\{|s[t]|^2\} = 1$. Then the ith relay receives

$$r_i[t] = \sqrt{P_s} f_i s[t] + n_i[t] \tag{1}$$

where P_s is the source transmit power and $n_i[t]$ is the receiver noise which follows the complex Gaussian distribution with $n_i[t] \sim \mathcal{CN}(0, N_0)$. As stated in [7], common interference or multiple-hop noise propagation could incur noise correlation between relays. So suppose $E\{n_i[t]n_j[t]^*\} \neq 0$ if $i \neq j$. After that, the ith relay will amplify the received signal as

$$x_i[t] = \sqrt{\frac{P_r}{P_s\delta_{si}^2 + N_0}} r_i[t] \tag{2}$$

where P_r is the relay power. Then, the ith relay would forward $x_i[t]$ to the destination during the ith time-slot of the second phase. Hence the received signal of the ith time-slot at the destination can be expressed as

$$y_i[t] = \sqrt{\frac{P_s P_r}{P_s\delta_{si}^2 + N_0}} f_i g_i s[t] + \sqrt{\frac{P_r}{P_s\delta_{si}^2 + N_0}} g_i n_i[t] + v_i[t] \tag{3}$$

where $v_i[t]$ is the receiver noise at the destination and $v_i[t] \sim \mathcal{CN}(0, N_0)$. After all N time-slots of phase 2, the receiver at the destination could collect $\{y_1[t], ..., y_N[t]\}$. To express clearly, we write the received signal in form of vector and matrix

$$\mathbf{y}[t] = \mathbf{P}_s \mathbf{h} s[t] + \mathbf{P}_r \mathbf{G} \mathbf{n}[t] + \mathbf{v}[t] \tag{4}$$

where

$$\mathbf{P}_s = \text{diag}\left\{\sqrt{\frac{P_s P_r}{P_s\delta_{s1}^2 + N_0}}, \sqrt{\frac{P_s P_r}{P_s\delta_{s2}^2 + N_0}}, ..., \sqrt{\frac{P_s P_r}{P_s\delta_{sN}^2 + N_0}}\right\}$$

$$\mathbf{P}_r = \text{diag}\left\{\sqrt{\frac{P_r}{P_s\delta_{s1}^2 + N_0}}, \sqrt{\frac{P_r}{P_s\delta_{s2}^2 + N_0}}, ..., \sqrt{\frac{P_r}{P_s\delta_{sN}^2 + N_0}}\right\}$$

$$\mathbf{h} = [f_1 g_1, f_2 g_2, ..., f_N g_N]^T$$

$$\mathbf{G} = \text{diag}\{g_1, g_2, ..., g_N\}$$

$$\mathbf{y}[t] = [y_1[t], y_2[t], ..., y_N[t]]^T$$

$$\mathbf{n}[t] = [n_1[t], n_2[t]..., n_N[t]]^T$$

$$\mathbf{v}[t] = [v_1[t], v_2[t], ..., v_N[t]]^T$$

Obviously, $\mathbf{n}[t] \sim \mathbf{CN}(0, \mathbf{K}_n)$ is additive white Gaussian noise with covariance matrix $\mathbf{K}_n$ [7] and $[\mathbf{K}_n]_{i,j} = E\{n_i[t]n_j[t]^*\}$. Similarly, we also assume $\mathbf{v}[t] \sim \mathbf{CN}(0, \mathbf{K}_v)$ is the additive white Gaussian noise with covariance matrix $\mathbf{K}_v$. Note that we assume there is no correlation between the destination noise $\mathbf{v}[t]$ and relay noise $\mathbf{n}[t]$ like [7]. Then the equivalent receiver noise is

$$\mathbf{w}[t] = \mathbf{P}_r \mathbf{G} \mathbf{n}[t] + \mathbf{v}[t] \tag{5}$$

Therefore, there is $\mathbf{w}[t] \sim \mathcal{CN}(0, \mathbf{P}_r^2 \mathbf{D}_g \mathbf{K}_n + \mathbf{K}_v)$, where $\mathbf{D}_g = \text{diag}\{|g_1|^2, |g_2|^2, ..., |g_N|^2\}$. For convenience, we denote the equivalent noise covariance as $\mathbf{\Sigma} = \mathbf{P}_r^2 \mathbf{D}_g \mathbf{K}_n + \mathbf{K}_v$.

Our goal in this paper is to estimate the unknown channel $\mathbf{h}$ and noise parameter $\mathbf{\Sigma}$ jointly. To allow unique estimation of the channel $\mathbf{h}$ (i.e., to resolve the phase ambiguity incurred by M-PSK modulation), we also assume that there are T_p pilot symbols, i.e., $s_p[t]$, $1 \leq t \leq T_p$, are inserted at the beginning of the transmission of T data symbols. Assume all pilots are processed by the same procedure like data symbols. And we denote the corresponding received signal as $\mathbf{y}_p[t]$. Generally, T_p is so small that pilot symbols could not provide reliable channel estimation for data detecting. To make the channel estimation work well, we assume channel $\{f_i\}$, $\{g_i\}$ and noise covariance keep constant within each $T + T_p$ symbol intervals, and change from one block to another block.

III. EM Algorithm based Channel Estimation

In this section, we introduce an EM-based channel estimation scheme for AF relay networks with unknown noise correlation. As stated in system model, given the known pilot symbols $s_p[1], s_p[2], ..., s_p[T_p]$ and the received $\mathbf{y}[1], \mathbf{y}[2], ..., \mathbf{y}[T]$ and $\mathbf{y}_p[1], \mathbf{y}_p[2], ..., \mathbf{y}_p[T_p]$, we intend to compute the ML estimates of the equivalent channel $\mathbf{h}$ and noise covariance matrix $\mathbf{\Sigma}$ for each block.

The EM algorithm is a general iterative method for computing ML estimates in the case where directly maximizing the likelihood function of observed data is not easily performed. Each EM iteration consists of two steps: expectation (E) step, which computes the expectation of the log-likelihood with respect to the conditional distribution of the unobserved data given the observed data, and maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. Generally speaking, we prefer to choose an unobserved data which lets the M step easy [15].

In this paper, the unknown symbols from source node are modeled as the unobserved data, i.e., $s[t], t = 1, ..., T$, and the parameters to be estimated are $\mathbf{h}$ and $\mathbf{\Sigma}$. As $\mathbf{y}[t]$ follows a complex multivariate Gaussian distribution, then the probability density function (pdf) given $\mathbf{h}$, $\mathbf{\Sigma}$ and $s[t]$ is

$$p\left(\mathbf{y}[t] | s[t], \mathbf{h}, \mathbf{\Sigma}\right) = \frac{1}{\pi^N |\Sigma|} \exp\left\{-(\mathbf{y}[t] - \mathbf{P}_s \mathbf{h} s[t])^H \mathbf{\Sigma}^{-1} (\mathbf{y}[t] - \mathbf{P}_s \mathbf{h} s[t])\right\} \tag{6}$$

For pilot symbols, i.e., $s_p[t], t = 1..., T_p$, above expression also holds for $p(\mathbf{y}_p[t] | s_p[t], \mathbf{h}, \mathbf{\Sigma})$, with $\mathbf{y}[t]$ and $s[t]$ replaced by $\mathbf{y}_p[t]$ and $s_p[t]$. Thus, the expected observed data likelihood function is

$$\prod_{t=1}^{T} p(s[t]) p\left(\mathbf{y}[t] | s[t], \mathbf{h}, \mathbf{\Sigma}\right) \prod_{t=1}^{T_p} p(\mathbf{y}_p[t] | s_p[t], \mathbf{h}, \mathbf{\Sigma}) \tag{7}$$

which is also called by complete-data likelihood function [12]. By taking the natural logarithm of (7) and omitting the terms independent of $\mathbf{h}$ and $\mathbf{\Sigma}$, we have

$$\begin{aligned} L(\mathbf{h}, \mathbf{\Sigma}) = &-(T + T_p)\left\{\ln|\mathbf{\Sigma}| + \text{Tr}\left\{\mathbf{\Sigma}^{-1}\left(\mathbf{R}_{yy} - \mathbf{R}_{ys}\mathbf{h}^H\mathbf{P}_s^H \right.\right.\right. \\ &\left.\left.\left. - \mathbf{P}_s\mathbf{h}\mathbf{R}_{ys}^H + R_{ss}\mathbf{h}\mathbf{P}_s\mathbf{P}_s^H\mathbf{h}^H\right)\right\}\right\} \end{aligned} \tag{8}$$

where $|.|$ denotes the determinant and

$$\mathbf{R}_{yy} = \frac{1}{T+T_p}\left(\sum_{t=1}^{T}\mathbf{y}[t]\mathbf{y}[t]^H + \sum_{t=1}^{T_p}\mathbf{y}_p[t]\mathbf{y}_p^H[t]\right)$$

$$\mathbf{R}_{ys} = \frac{1}{T+T_p}\left(\sum_{t=1}^{T}\mathbf{y}[t]s[t]^* + \sum_{t=1}^{T_p}\mathbf{y}_p[t]s_p[t]^*\right)$$

$$R_{ss} = \frac{1}{T+T_p}\left(\sum_{t=1}^{T}|s[t]|^2 + \sum_{t=1}^{T_p}|s_p[t]|^2\right)$$

Note that

$$\begin{aligned}&(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}s[t])^H\mathbf{\Sigma}^{-1}(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}s[t])\\ &=\mathrm{Tr}\{\mathbf{\Sigma}^{-1}(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}s[t])(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}s[t])^H\}\end{aligned} \tag{9}$$

is used in above derivations. In EM algorithm, the E step computes the conditional expectation of the observed data log-likelihood function given the observed data $\mathbf{y}[1], \mathbf{y}[2], ..., \mathbf{y}[T]$ and $\mathbf{y}_p[1], \mathbf{y}_p[2], ..., \mathbf{y}_p[T_p]$ and the estimated parameters from last iteration. For the kth iteration, the conditional expectation is

$$\begin{aligned}&Q(\mathbf{h},\mathbf{\Sigma}|\mathbf{h}^{(k)},\mathbf{\Sigma}^{(k)}) = E_{s[t]}\left\{L(\mathbf{h},\mathbf{\Sigma})\right\}\\ &=-(T+T_p)\left\{\ln|\mathbf{\Sigma}| + \mathrm{Tr}\left\{\mathbf{\Sigma}^{-1}\left(\mathbf{R}_{yy} - \left(\mathbf{R}_{ys}^{(k)}\right)\mathbf{h}^H\mathbf{P}_s^H\right.\right.\right.\\ &\left.\left.\left.-\mathbf{P}_s\mathbf{h}\left(\mathbf{R}_{ys}^{(k)}\right)^H + R_{ss}^{(k)}\mathbf{h}\mathbf{P}_s\mathbf{P}_s^H\mathbf{h}^H\right)\right\}\right\}\end{aligned} \tag{10}$$

where

$$\begin{aligned}\mathbf{R}_{ys}^{(k)} &= E_{s[t]|\mathbf{y}[t]}\left\{\mathbf{R}_{ys}|\mathbf{h}^{(k)},\mathbf{\Sigma}^{(k)}\right\}\\ &=\frac{1}{T+T_p}\left(\sum_{t=1}^{T}\mathbf{y}[t]\left(\sum_{m=1}^{M}\rho_m^{(k)}[t]s_m^*\right) + \sum_{t=1}^{T_p}\mathbf{y}_p[t]s_p[t]^*\right)\\ R_{ss}^{(k)} &= E_{s[t]|\mathbf{y}[t]}\left\{R_{ss}|\mathbf{h}^{(k)},\mathbf{\Sigma}^{(k)}\right\}\\ &=\frac{1}{T+T_p}\left(\sum_{t=1}^{T}\sum_{m=1}^{M}\left(|s_m|^2\rho_m^{(k)}[t]\right) + \sum_{t=1}^{T_p}|s_p[t]|^2\right)\end{aligned}$$

and

$$\rho_m^{(k)}[t] = \frac{\exp\left\{-\left(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}^{(k)}s_m\right)^H(\mathbf{\Sigma}^{(k)})^{-1}\left(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}^{(k)}s_m\right)\right\}}{\sum_{n=1}^{M}\exp\left\{-\left(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}^{(k)}s_n\right)^H(\mathbf{\Sigma}^{(k)})^{-1}\left(\mathbf{y}[t]-\mathbf{P}_s\mathbf{h}^{(k)}s_n\right)\right\}} \tag{11}$$

Then the M step maximizes $Q(\mathbf{h},\mathbf{\Sigma}|\mathbf{h}^{(k)},\mathbf{\Sigma}^{(k)})$ with respect to $\mathbf{h}$ and $\mathbf{\Sigma}$ to obtain the $k+1$th iterative parameters:

$$\left(\mathbf{h}^{(k+1)},\mathbf{\Sigma}^{(k+1)}\right) = \arg\max_{\mathbf{h},\mathbf{\Sigma}}\left\{Q(\mathbf{h},\mathbf{\Sigma}|\mathbf{h}^{(k)},\mathbf{\Sigma}^{(k)})\right\} \tag{12}$$

Due to the results in Appendix of [12], the solution of (12) are

$$\mathbf{h}^{(k+1)} = \mathbf{P}_s^{-1}\frac{\mathbf{R}_{ys}^{(k)}}{R_{ss}^{(k)}} \tag{13}$$

and

$$\mathbf{\Sigma}^{(k+1)} = \mathbf{R}_{yy} - \frac{\mathbf{R}_{ys}^{(k)}\left(\mathbf{R}_{ys}^{(k)}\right)^H}{R_{ss}^{(k)}} \tag{14}$$

if

$$\mathbf{R}_{yy} - \frac{\mathbf{R}_{ys}^{(k)}\left(\mathbf{R}_{ys}^{(k)}\right)^H}{R_{ss}^{(k)}}$$

is a positive definite matrix. To ensure positive definiteness with probability 1 of $\mathbf{\Sigma}^{(k+1)}$, we refer to the Theorem 10.1.1 in [14] and obtain following condition:

$$T+T_p \geq N+1 \tag{15}$$

Because the number of relays can not be too large, this condition is always guaranteed in practice.

Therefore, we obtain the following EM algorithm based channel estimation iteration:

- Step 1)

$$\mathbf{h}^{(k+1)} = \mathbf{P}_s^{-1}\frac{\sum_{t=1}^{T}\mathbf{y}[t]\left(\sum_{m=1}^{M}\rho_m^{(k)}[t]s_m^*\right) + \sum_{t=1}^{T_p}\mathbf{y}_p[t]s_p[t]^*}{\sum_{t=1}^{T}\sum_{m=1}^{M}\left(|s_m|^2\rho_m^{(k)}[t]\right) + \sum_{t=1}^{T_p}|s_p[t]|^2} \tag{16}$$

- Step 2)

$$\begin{aligned}\mathbf{\Sigma}^{(k+1)} =&\frac{1}{T+T_p}\left(\sum_{t=1}^{T}\mathbf{y}[t]\mathbf{y}[t]^H + \sum_{t=1}^{T_p}\mathbf{y}_p[t]\mathbf{y}_p^H[t]\right)\\ &-\mathbf{P}_s\mathbf{h}^{(k+1)}\left(\mathbf{h}^{(k+1)}\right)^H\mathbf{P}_s^H\end{aligned} \tag{17}$$

- Step 3) Use (11), derive the $\rho_m^{(k+1)}[t]$.

If $|\mathbf{h}^{(k)}-\mathbf{h}^{(k+1)}| < \epsilon_h$ or $|\mathbf{\Sigma}^{(k)}-\mathbf{\Sigma}^{(k+1)}| < \epsilon_\Sigma$ is not satisfied, continue running step 1) to step 3) again. As the convergence properties of the EM Algorithm, we just need rough estimates of $\mathbf{h}$ and $\mathbf{\Sigma}$ as initial parameters. As each pilot symbol can produce one estimate of $\mathbf{h}$, the average Least-Square (LS) estimate [18] of $\mathbf{h}$ is $\mathbf{h}^{(0)} = \frac{1}{T_p}\sum_{t=1}^{T_p}\left(|s_p[t]|^2\mathbf{P}_s^2\right)^{-1}(s_p[t]\mathbf{P}_s)^H\mathbf{y}_p[t]$. By (17), we can derive the initial noise covariance $\mathbf{\Sigma}^{(0)}$. As a result, we have initialized the EM algorithm. Note that by our assumptions, $R_{ss}=1$, $|s[t]|=1$ and $|s_t[t]|=1$ can be substituted into (16) and (17) to simplify expressions. However, the purpose that we give these complete expressions is to show that our channel estimation is not limited by the MPSK modulation.

IV. Performance Analysis

To illustrate the performance of proposed iterative channel estimation, we usually compute the Cramér-Rao Bound (CRB). In our channel estimation, channel coefficient and noise covariance matrix should be jointly estimated in the presence of unknown data symbols (nuisance parameters). Thus the standard CRB may be quite tedious to evaluate [16][17]. To avoid complicated computation, we may use the modified CRB (MCRB) [17], which is easier to compute and a lower bound on the standard CRB [13].

Definition: Set the observed vector as $\mathbf{x}$, the parameter vector to be estimated as $\mathbf{u}$ and the nuisance vector as $\mathbf{v}$. Also denote the conditional pdf of $\mathbf{x}$ as $p(\mathbf{x}|\mathbf{v};\mathbf{u})$. Then, the MCRB of the estimate on $\mathbf{u}$ is $\mathcal{I}_M(\mathbf{u})^{-1}$ and it meets

$$\begin{aligned}[\mathcal{I}_M(\mathbf{u})]_{i,j} &= -E_{\mathbf{x},\mathbf{u}}\left\{\frac{\partial^2 \log p(\mathbf{x}|\mathbf{v};\mathbf{u})}{\partial u_i \partial u_j}\right\} \\ &= E_{\mathbf{v}}\{[\mathcal{I}(\mathbf{v};\mathbf{u})]_{i,j}\}\end{aligned} \tag{18}$$

where $[\mathcal{I}(\mathbf{v};\mathbf{u})]_{i,j} = -E_{\mathbf{x}|\mathbf{v}}\left\{\frac{\partial^2 \log p(\mathbf{x}|\mathbf{v};\mathbf{u})}{\partial u_i \partial u_j}\right\}$. Note that $\mathcal{I}$ is the Fisher Information Matrix (FIM) for the estimation $\mathbf{u}$ when $\mathbf{v}$ is known [18]. In other words, we can obtain the MCRB through calculating the expectation with respect to $\mathbf{v}$ of the corresponding FIM.

For convenience, we also follow suggestions of [13], [12] and [18] to define the unknown channel and noise parameters as $\mathbf{\Theta} = [\tilde{\mathbf{h}}^T, \tilde{\mathbf{\Sigma}}^T]^T$, where $\tilde{\mathbf{h}} = [\text{Re}(\mathbf{h})^T, \text{Im}(\mathbf{h})^T]^T$ and $\tilde{\mathbf{\Sigma}} = [\text{Re}(\text{vech}(\mathbf{\Sigma}))^T, \text{Im}(\text{vechd}(\mathbf{\Sigma}))^T]^T$. Herein vech(.) means staking elements below the main diagonal columnwise into a column vector and vechd(.) means staking elements below the main diagonal columnwise and the diagonal elements into a column vector [12]. For the expectation of unknown symbol is unit, the MCRB for parameter $\mathbf{\Theta}$ is identical to the standard CRB if the data symbols $s[t]$ are known [13][18]. Due to the Example 3.6 in [18], we have FIM for $\mathbf{\Theta}$ given $\mathbf{y}$

$$\mathcal{I}(\mathbf{y};\mathbf{\Theta}) = \begin{bmatrix} -E\left\{\frac{\partial^2 \ln p(\mathbf{y};\mathbf{\Theta})}{\partial \tilde{\mathbf{h}}^2}\right\} & -E\left\{\frac{\partial^2 \ln p(\mathbf{y};\mathbf{\Theta})}{\partial \tilde{\mathbf{h}} \partial \tilde{\mathbf{\Sigma}}}\right\} \\ -E\left\{\frac{\partial^2 \ln p(\mathbf{y};\mathbf{\Theta})}{\partial \tilde{\mathbf{\Sigma}} \partial \tilde{\mathbf{h}}}\right\} & -E\left\{\frac{\partial^2 \ln p(\mathbf{y};\mathbf{\Theta})}{\partial \tilde{\mathbf{\Sigma}}^2}\right\} \end{bmatrix} \tag{19}$$

By (8), there are

$$E\left\{\frac{\partial^2 \ln p(\mathbf{y};\mathbf{\Theta})}{\partial \tilde{\mathbf{h}} \partial \tilde{\mathbf{\Sigma}}}\right\} = 0 \quad \text{and} \quad E\left\{\frac{\partial^2 \ln p(\mathbf{y};\mathbf{\Theta})}{\partial \tilde{\mathbf{\Sigma}} \partial \tilde{\mathbf{h}}}\right\} = 0$$

Therefore,

$$\text{MCRB}_\Theta = \begin{bmatrix} \text{MCRB}_{\tilde{h}} & 0 \\ 0 & \text{MCRB}_{\tilde{\Sigma}} \end{bmatrix} \tag{20}$$

Correspondingly, we can derive the MCRB for $\mathbf{\Theta}$ through solving the MCRBs for $\tilde{\mathbf{h}}$ and $\tilde{\mathbf{\Sigma}}$. For MCRB of $\tilde{\mathbf{h}}$, we can follow the Example 15.9 in [18] and obtain

$$\text{MCRB}_{\tilde{\mathbf{h}}} = \frac{\tilde{\mathbf{P}}_s^{-2}}{2(T+T_p)} \begin{bmatrix} \text{Re}\,(\mathbf{\Sigma}) & -\text{Im}\,(\mathbf{\Sigma}) \\ \text{Im}\,(\mathbf{\Sigma}) & \text{Re}\,(\mathbf{\Sigma}) \end{bmatrix} \tag{21}$$

where $\tilde{\mathbf{P}}_s = \text{diag}\{\mathbf{P}_s, \mathbf{P}_s\}$. For MCRB of $\tilde{\mathbf{\Sigma}}$, by the equation (15.52) in [18], there is

$$\text{MCRB}_{\tilde{\mathbf{\Sigma}}} = \frac{1}{T+T_p}\mathcal{I}(\tilde{\mathbf{\Sigma}})^{-1} \tag{22}$$

The (i,j)th element of $\mathcal{I}(\tilde{\mathbf{\Sigma}})$ is

$$\left[\mathcal{I}(\tilde{\mathbf{\Sigma}})\right]_{i,j} = \text{Tr}\left\{\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial \tilde{\Sigma}_i}\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial \tilde{\Sigma}_j}\right\} \tag{23}$$

where $\tilde{\Sigma}_i$ is the ith element of $\tilde{\mathbf{\Sigma}}$ and

$$\frac{\partial \mathbf{\Sigma}}{\partial \tilde{\Sigma}_i} = \begin{bmatrix} \frac{\partial [\mathbf{\Sigma}]_{11}}{\partial \tilde{\Sigma}_i} & \frac{\partial [\mathbf{\Sigma}]_{12}}{\partial \tilde{\Sigma}_i} & \cdots & \frac{\partial [\mathbf{\Sigma}]_{1N}}{\partial \tilde{\Sigma}_i} \\ \frac{\partial [\mathbf{\Sigma}]_{21}}{\partial \tilde{\Sigma}_i} & \frac{\partial [\mathbf{\Sigma}]_{22}}{\partial \tilde{\Sigma}_i} & \cdots & \frac{\partial [\mathbf{\Sigma}]_{2N}}{\partial \tilde{\Sigma}_i} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial [\mathbf{\Sigma}]_{N1}}{\partial \tilde{\Sigma}_i} & \frac{\partial [\mathbf{\Sigma}]_{N1}}{\partial \tilde{\Sigma}_i} & \cdots & \frac{\partial [\mathbf{\Sigma}]_{NN}}{\partial \tilde{\Sigma}_i} \end{bmatrix} \tag{24}$$

According to the definition of $\tilde{\mathbf{\Sigma}}$, we can derive the MCRB for $\tilde{\mathbf{\Sigma}}$. Finally, the MCRB for $\mathbf{\Theta}$ is obtained.

V. Simulation Results

In this section, we evaluate the performance of the proposed estimation and detection algorithms through computer simulations. In our simulations, we set all channels and receiver noise follow standard complex Gaussian distribution, i.e., $\delta_{g_i}^2 = \delta_{f_i}^2 = 1$ and $N_0 = 1$. The transmitted symbols were generated from an uncoded BPSK modulated constellation. Set the total power as $P_t = P_s + NP_r$, then the system signal-noise ratio (SNR) is $SNR = P_t/N_0$. Without lose of generality, let $\mathbf{K}_v = \mathbf{I}$ and

$$[\mathbf{K}_n]_{i,j} = N_0 \times 0.9^{|i-j|} \times \exp\left\{j\left(\frac{\pi}{2}\right)(i-j)\right\} \tag{25}$$

which also is used in [12][13] as correlated noise covariance. Each data block consists of $T_p = 2$ pilot symbols and $T = 32$ data symbols. We assume that channel coefficients and noise covariance do not change within one data block. We set the iterative stop condition as $\epsilon_h = \epsilon_\Sigma = 10^{-3}$.

Fig. 1 and Fig. 2 show the average mean square error (MSE) of estimators introduced in above context. To show the tendency of the EM based iterative channel estimation, we simulate EM based estimators with different numbers of iterations. Note that 0-iteration estimator just employs the two pilot symbols to perform LS estimation, i.e., $\hat{\mathbf{h}} = \mathbf{h}^{(0)}$. We plot the MCRB of proposed estimator in both figures to show the performance lower bounds of these estimators. Obviously, 0-iteration estimator has the worst performance among all considered estimators. In other words, the performance of 0-iteration estimator is the upper bound of these estimators. Observing both figures, we can conclude that the estimator with more iterations has smaller MSE and better performance. Therefore, if the number of iterations is large enough, the MSE of proposed channel estimator can approach the MCRB closely. In addition, higher SNR also incurs lower MSE. Summarily, both figures show that our proposed channel estimation works well if the number of iterations or the SNR is large enough.

We show the bit error rate (BER) performances of two-relay network in Fig. 3 when the proposed joint channel estimation is used. We assume that the ideal receiver used in our simulations can access the exact channel coefficient and noise covariance. Therefore, the BER performance of ideal receiver is the lower bound of practical receivers with channel estimation. We can see that the ideal receiver outperforms the 0-iteration receiver about 2.5 dB. The 2-iteration receiver has a slightly better performance than 1-iteration receiver,

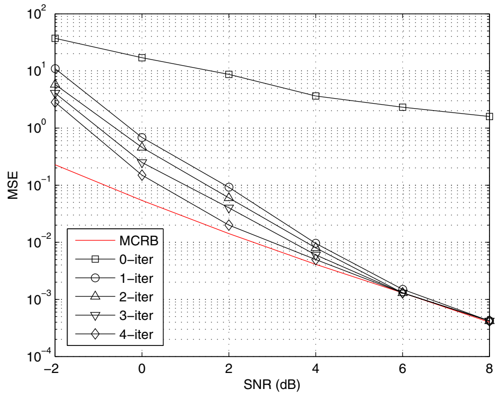

Fig. 1. Average MSE versus SNR in two-relay network (N=2)

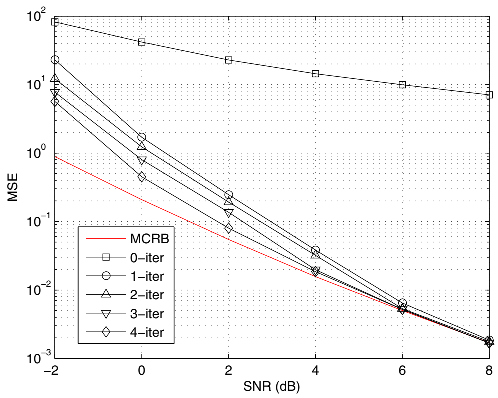

Fig. 2. Average MSE versus SNR in four-relay network (N=4)

which outperforms the 0-iteration receiver about 1.6 dB. The performance of 2-iteration receiver approaches that of ideal receiver closely, especially when SNR is larger than 5 dB. We also can infer that receiver with more iterations can produce better BER performance.

VI. Conclusions

In this paper, we proposed an EM algorithm based channel estimation scheme for Amplify-and-Forward relay networks with unknown noise correlation.The validity of proposed channel estimator is verified by simulation results. It shows that the larger the number of iterations is, the better performance the proposed receiver can achieve.

Acknowledge

This paper was supported by the National Natural Science Foundation of China (No. 61102082) and National High-tech R&D Program of China (863 Program) (Grant No. 2011AA01A105), and National Science and Technology Major Project of China (Grant No. 2012ZX03001032-002).

References

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior,"*IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004

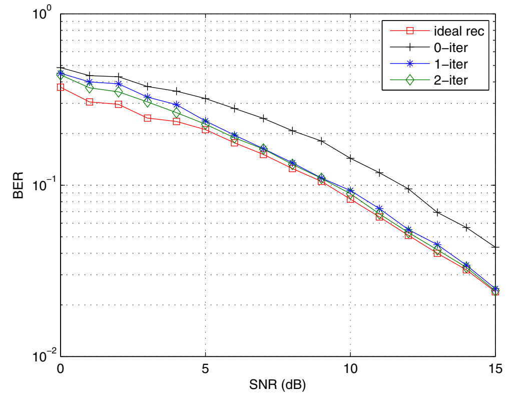

Fig. 3. BER performance of two-relay network with different channel estimations (0 dB$\leq$ SNR $\leq$ 15 dB)

[2] A. Sabharwal, "Impact of Half-duplex Radios and Decoding Latencies on MIMO Relay Channel ,"Proc. of the Allerton Conference, 2004

[3] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks,"*IEEE Trans. Wireless Commu.*vol. 5, no. 12, pp. 3524-3536, Dec. 2006.

[4] C. Zhang, H. Yin, W. Wang and G. Wei, "Distributed space-time diversity system using linear constellation precoding"Proc. IEEE WCNC, Budpest, Hungary, April 2009.

[5] Yi Zhao, Raviraj Adve, and Teng Joon Lim, "Improving Amplify-and-Forward Relay Networks: Optimal Power Allocation versus Selection "*IEEE Trans. Wireless Commu.*vol. 6, no. 8, pp. 3524-3536, Aug. 2007.

[6] S. Jin, M.R. McKay, C. Zhong, K.-K. Wong,"Ergodic Capacity Analysis of Amplify-and-Forward MIMO Dual-Hop Systems "*IEEE Transactions on Information Theory*,vol.56, no.5, pp.2204-2224, May 2010.

[7] K.S. Gomadam, S.A. Jafar, "The Effect of Noise Correlation in Amplify-and-Forward Relay Networks,", IEEE Transactions on Information Theory, vol.55, no.2, pp.731-745, Feb. 2009.

[8] A.S. Behbahani and A. M. Eltawil, "Amplify-and-forward relay networks under received power constraint ,"IEEE Transactions on Wireless Communications. vol. 8 no 11, Nov. 2009

[9] F. Gao, T. Cui, A. Nallanathan,, "On channel estimation and optimal training design for Amplify and Forward relay networks, " *IEEE Trans on Wireless Commu.,* vol. 7, no. 5, pp. 1907-1916, May. 2008.

[10] C. S. Patel, G. L. Stüber, "Channel estimation for Amplify and Forward relay based cooperation diversity systems, " *IEEE Trans. on Wireless Commu.,*vol. 6, no. 6, pp. 2348-2356, June 2007.

[11] Jeng-Shin Sheu and Wern-Ho Sheen, "An EM Algorithm-Based Channel Estimation for OFDM Amplify-And-Forward Relaying Systems", in Proceeding of IEEE ICC 2010.

[12] A. Dogandzic, W. Mo, Z. Wang, "Semi-blind SIMO flat-fading channel estimation in unknown spatially correlated noise using the EM algorithm ,"IEEE Transactions on Signal Processing. Vol. 52, no. 6, pp. 1791-1797. June 2004.

[13] W. Mo, Z. Wang and A. Dogandzic. (2007), EM-based iterative receiver for coded MIMO systems in unknown spatially correlated noise. Wireless Communications and Mobile Computing, 7: 81-89

[14] R.J. Muirhead. Aspects of Multivariate Statistical Theory. Wiley: New York, USA, 1982.

[15] G. J. McLachlan andT . Krishnan, The EM Algorithm and Extensions. New York: Wiley, 1997.

[16] C. Herzet, V. Ramon, A. Renaux, and L. Vandendorpe, "A Cramér-Rao Bound Characterization of the EM-Algorithm Mean Speed of Convergence ,"IEEE Transactions on Signal Processing vol.56, no.6 pp.2218-2228. 2008

[17] F. Gini, R. Reggiannini, and U. Mengali, "The Modified CramérCRao Bound in Vector Parameter Estimation, "IEEE Trans. on Commun., Vol. 46, No. 1, pp.52-60. Jan. 1998

[18] S. M. Kay, Fundamentals of statistical signal processing Volume I: Estimation theory, Prentice Hall PTR, New Jersey, 1998