EM-based Semi-blind Channel Estimation in Amplify-and-Forward Two-Way Relay Networks

Saeed Abdallah and Ioannis N. Psaromiligkos

***Abstract*—In this letter, we propose an expectation maximization (EM)-based algorithm for semi-blind channel estimation of reciprocal channels in amplify-and-forward (AF) two-way relay networks (TWRNs). By utilizing data samples as well as pilots, the proposed algorithm provides substantially higher estimation accuracy than the conventional training-based least squares (LS) estimator without incurring a significant computational cost. Simulation results also show that it performs very close to the corresponding semi-blind Cramer-Rao bound.**

***Index Terms*: Amplify and Forward, Expectation Maximization, Semi-blind Channel Estimation, Two-way Relays.**

I. Introduction

Two-way relay networks (TWRNs) [1] have recently been the subject of intense research efforts due to their high spectral efficiency compared to conventional one-way relay networks. TWRNs employing the amplify-and-forward (AF) relaying protocol are particularly appealing due to the minimal processing required at the relay. AF TWRNs, however, require highly accurate information about the channel at the terminals, for both self-interference cancellation and coherent decoding.

The problem of channel estimation for AF TWRNs has been considered in a number of works [2]–[5]. Most works adopt a training-based approach that requires each terminal to transmit a pilot sequence known to the other terminal. Unfortunately, the transmission of pilots consumes bandwidth resources, which undermines the spectral efficiency of TWRNs. One alternative to training-based estimation is semi-blind estimation [6] which, in addition to using pilots, also incorporates the received data samples into the estimation.

The potential of semi-blind channel estimation to outperform training-based estimation of nonreciprocal flat-fading channels for AF TWRNs was demonstrated through a Cramer-Rao bound (CRB) analysis in [7]. In particular, the exact semi-blind CRB was derived assuming square QAM and was compared with the pilot-based CRB, showing that semi-blind estimation can provide substantially higher estimation accuracy using a limited number of data samples. However, no specific semi-blind estimation algorithms were considered in [7].

In this letter we derive a semi-blind expectation-maximization (EM)-based channel estimator for AF TRWNs assuming reciprocal flat-fading channels. Each iteration of the proposed algorithm has a low computational cost and only a small number of iterations is needed to achieve convergence. Using simulations, we show that the EM algorithm performs very close to the semi-blind CRB and, even with a limited number of data samples, provides substantially better accuracy than the training-based least-squares (LS) estimator. We also show that the EM algorithm provides a significant improvement in throughput since a smaller number of pilots would be needed to achieve the same symbol-error rate (SER) as the LS estimator.

The rest of this letter is organized as follows. In Section II we present the system model. In Section III, we derive the EM algorithm. Simulation results are presented in Section IV. Finally, our conclusions are discussed in Section V.

II. System Model

We consider a half-duplex AF TWRN with two terminals $\mathcal{T}_1$, $\mathcal{T}_2$ and a single relay $\mathcal{R}$ operating in flat-fading channel conditions. Bidirectional communication between the two terminals takes place in two phases:

Phase 1: Each terminal transmits a block of L pilot symbols, followed by N data symbols. We denote by $\boldsymbol{t}_1 \triangleq [t_{11}, \ldots, t_{1L}]^T$ and $\boldsymbol{t}_2 \triangleq [t_{21}, \ldots, t_{2L}]^T$ the pilot symbol vectors of $\mathcal{T}_1$ and $\mathcal{T}_2$, and by $\boldsymbol{s}_1 \triangleq [s_{11}, \ldots, s_{1N}]^T$ and $\boldsymbol{s}_2 \triangleq [s_{21}, \ldots, s_{2N}]^T$ the transmitted data symbol vectors of $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively. The data symbols $s_{21}, \ldots, s_{2N}$ are drawn randomly from a constellation $S = \{\xi_1, \ldots, \xi_M\}$ of size M. The average transmission powers of $\mathcal{T}_1$ and $\mathcal{T}_2$ during pilot transmission are P_1 and P_2, respectively, i.e., $\boldsymbol{t}_1^H \boldsymbol{t}_1 = LP_1$ and $\boldsymbol{t}_2^H \boldsymbol{t}_2 = LP_2$. For simplicity, we assume that each terminal uses the same power during pilot and data transmission, i.e., $\mathbb{E}\left\{\boldsymbol{s}_1^H \boldsymbol{s}_1\right\} = NP_1$ and $\mathbb{E}\left\{\boldsymbol{s}_2^H \boldsymbol{s}_2\right\} = NP_2$.

The corresponding received pilot and data signal vectors at $\mathcal{R}$ are

$$\boldsymbol{r}_P = h\boldsymbol{t}_1 + g\boldsymbol{t}_2 + \boldsymbol{\omega} \tag{1}$$

and

$$\boldsymbol{r}_D = h\boldsymbol{s}_1 + g\boldsymbol{s}_2 + \boldsymbol{n} \tag{2}$$

where h and g are the complex coefficients of the flat fading channels $\mathcal{T}_1 \to \mathcal{R}$ and $\mathcal{T}_2 \to \mathcal{R}$, respectively, and $\boldsymbol{\omega}$ and $\boldsymbol{n}$ are additive white Gaussian noise vectors with mean zero and covariance matrices[1] $\sigma^2 \boldsymbol{I}_L$ and $\sigma^2 \boldsymbol{I}_N$, respectively. The channel coefficients h and g are modelled as independent circularly complex Gaussian random variables with mean 0 and variance γ^2 and are assumed to remain fixed during the transmission of the L pilots and N data symbols.

The authors are with the Electrical and Computer Engineering Dept., McGill University, Montreal, Quebec, Canada. Email: saeed.abdallah@mail.mcgill.ca; yannis@ece.mcgill.ca.

This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC).

[1] $\boldsymbol{I}_L$ denotes the $L \times L$ identity matrix.

Phase 2: The relay broadcasts the vectors $A\boldsymbol{r}_P$ and $A\boldsymbol{r}_D$ in sequence, where $A>0$ is the amplification factor. Assuming reciprocal channels, the corresponding received pilot signal vector at $\mathcal{T}_1$ is

$$\boldsymbol{y} = Ah^2\boldsymbol{t}_1 + Ahg\boldsymbol{t}_2 + Ah\boldsymbol{\omega} + \boldsymbol{\omega}_1 \tag{3}$$

and the received data signal vector is

$$\boldsymbol{z} = Ah^2\boldsymbol{s}_1 + Ahg\boldsymbol{s}_2 + Ah\boldsymbol{n} + \boldsymbol{n}_1 \tag{4}$$

where $\boldsymbol{\omega}_1$ and $\boldsymbol{n}_1$ are additive white Gaussian noise vectors with mean zero and covariance matrices $\sigma^2\boldsymbol{I}_L$ and $\sigma^2\boldsymbol{I}_N$, respectively. The average transmission power of the relay is maintained at P_r over the long term (i.e., over many transmitted blocks) by using the amplification factor

$$A = \sqrt{\frac{P_r}{\gamma^2(P_1+P_2)+\sigma^2}}. \tag{5}$$

We are interested in the estimation of the cascaded channel parameters $a \triangleq h^2$ and $b \triangleq hg$, which are sufficient for detection. The noise variance σ^2 is assumed to be known at $\mathcal{T}_1$.

III. Proposed Channel Estimation Algorithm

Let $\boldsymbol{\theta} \triangleq [a,b]^T$ be the vector of unknown parameters that we wish to estimate. The observed vectors $\{\boldsymbol{y},\boldsymbol{z}\}$ represent the incomplete data set and the data symbols $\boldsymbol{s}_2$ represent the hidden data set. Hence, the complete data set is $\{\boldsymbol{y},\boldsymbol{z},\boldsymbol{s}_2\}$ and the corresponding log-likelihood function (LLF) is

$$\mathcal{L}(\boldsymbol{y},\boldsymbol{z},\boldsymbol{s}_2;\boldsymbol{\theta}) = -N\log M - (N+L)\log(\pi\sigma^2(A^2|a|+1)) - \frac{1}{\sigma^2(A^2|a|+1)}\left(\left\|\boldsymbol{y}-Aa\boldsymbol{t}_1-Ab\boldsymbol{t}_2\right\|^2+\left\|\boldsymbol{z}-Aa\boldsymbol{s}_1-Ab\boldsymbol{s}_2\right\|^2\right). \tag{6}$$

An iteration of the EM algorithm, say the tth one, consists of two steps. The first step, called the expectation step (E-step) consists of evaluating the expectation

$$Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(t)}) = \mathbb{E}\left\{\mathcal{L}(\boldsymbol{y},\boldsymbol{z},\boldsymbol{s}_2;\boldsymbol{\theta})|\boldsymbol{y},\boldsymbol{z};\boldsymbol{\theta}^{(t)}\right\} \tag{7}$$

of the LLF of the complete data, $\mathcal{L}(\boldsymbol{y},\boldsymbol{z},\boldsymbol{s}_2;\boldsymbol{\theta})$, with respect to the conditional PMF $f(\boldsymbol{s}_2|\boldsymbol{y},\boldsymbol{z};\boldsymbol{\theta}^{(t)})$ of the hidden data given the observations and the current estimate $\boldsymbol{\theta}^{(t)} \triangleq [a^{(t)},b^{(t)}]^T$ of $\boldsymbol{\theta}$. The second step of the EM algorithm is the maximization step (M-step) which consists of maximizing $Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(t)})$ with respect to $\boldsymbol{\theta}$ to obtain an updated estimate $\boldsymbol{\theta}^{(t+1)}$, i.e.,

$$\boldsymbol{\theta}^{(t+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(t)}). \tag{8}$$

In our case, the E-step and M-step are as follows:
E-step: We have[2]

$$Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(t)}) = -(N+L)\log(\pi\sigma^2(A^2|a|+1)) - \frac{1}{\sigma^2(A^2|a|+1)}\times\left(\|\boldsymbol{y}-Aa\boldsymbol{t}_1-Ab\boldsymbol{t}_2\|^2+\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\left|z_i-Aas_{1i}-Ab\xi_j\right|^2\right) \tag{9}$$

[2]We ignore the term $N\log M$ since it has no effect on the solution.

where $\beta_{i,j}^{(t)}$ is the posterior PMF of the ith data symbol during the tth iteration, given by

$$\beta_{i,j}^{(t)} = \frac{e^{-\frac{1}{\sigma^2(A^2|a^{(t)}|+1)}|z_i-Aa^{(t)}s_{1i}-Ab^{(t)}\xi_j|^2}}{\sum_{k=1}^{M} e^{-\frac{1}{\sigma^2(A^2|a^{(t)}|+1)}|z_i-Aa^{(t)}s_{1i}-Ab^{(t)}\xi_k|^2}}. \tag{10}$$

M-step: We need to obtain the values $a^{(t+1)}$, $b^{(t+1)}$ such that

$$\left\{a^{(t+1)},b^{(t+1)}\right\} = \arg\max_{\boldsymbol{\theta}=[a,b]^T} Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(t)}\right). \tag{11}$$

Regarding $b^{(t+1)}$, it can be easily verified that the value of b that maximizes $Q(\boldsymbol{\theta};\boldsymbol{\theta}^{(t)})$ for a given value of a is

$$b_o(a) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\xi_j^*\left(z_i-Aas_{1i}\right)+\boldsymbol{t}_2^H\boldsymbol{y}-Aa\boldsymbol{t}_2^H\boldsymbol{t}_2}{A\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}+A\boldsymbol{t}_2^H\boldsymbol{t}_2}. \tag{12}$$

Substituting $b_o(a)$ in place of b in (9), we obtain the following updated objective function that depends only on a:

$$Q(a;\boldsymbol{\theta}^{(t)}) = -N\log(\pi\sigma^2(A^2|a|+1)) - \frac{1}{G^2\sigma^2(A^2|a|+1)}\times\left(\left\|G\boldsymbol{y}-AaG\boldsymbol{t}_1-A\mathcal{I}\boldsymbol{t}_2+A^2a\mathcal{X}\boldsymbol{t}_2\right\|^2+\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\left|Gz_i-AaGs_{1i}-A\mathcal{I}\xi_j+A^2a\mathcal{X}\xi_j\right|^2\right) \tag{13}$$

where

$$G = A\left(\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}|\xi_j|^2+\boldsymbol{t}_2^H\boldsymbol{t}_2\right),\ \mathcal{I} = \sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\xi_j^*\boldsymbol{z}_i+\boldsymbol{t}_2^H\boldsymbol{y},$$

and

$$\mathcal{X} = \sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\xi_j^*s_{1i}+\boldsymbol{t}_2^H\boldsymbol{t}_1.$$

In order to maximize (13) with respect to a, we will maximize it first with respect to the phase $\phi_a \triangleq \angle a$ and then the amplitude $|a|$. Maximizing (13) with respect to ϕ_a we obtain

$$\phi_a^{(t+1)} = \pi - \angle\left(\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\left(Gz_i-A\mathcal{I}\xi_j\right)^*\left(A^2\mathcal{X}\xi_j-AGs_{1i}\right)+\left(G\boldsymbol{y}-A\mathcal{I}\boldsymbol{t}_2\right)^H\left(A^2\mathcal{X}\boldsymbol{t}_2-AG\boldsymbol{t}_1\right)\right). \tag{14}$$

Substituting $\phi_a^{(t+1)}$ into (13), we obtain the following function that depends only on $|a|$:

$$Q(|a|;\boldsymbol{\theta}^{(t)}) = -N\log(\pi\sigma^2(A^2|a|+1)) - \frac{\breve{V}|a|^2-2\breve{W}|a|+\breve{U}}{\sigma^2(A^2|a|+1)}, \tag{15}$$

where

$$\breve{U} = \frac{1}{G^2}\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\left|Gz_i-A\mathcal{I}\xi_j\right|^2+\frac{1}{G^2}\left\|G\boldsymbol{y}-A\mathcal{I}\boldsymbol{t}_2\right\|^2,$$

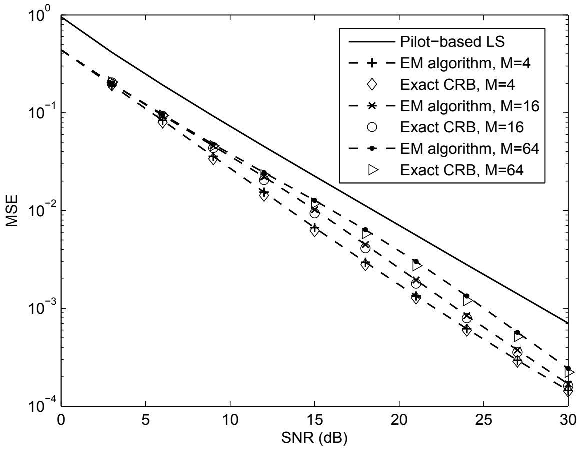

Fig. 1. MSE performance of the EM algorithm and the LS estimator along with the corresponding semi-blind CRB plotted versus SNR ($N = 32$ and $L = 8$, 4 EM iterations).

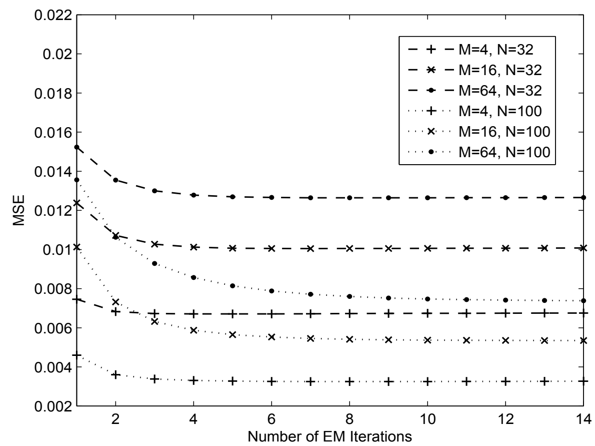

Fig. 2. MSE performance of the EM algorithm plotted versus the number of EM iterations ($N = 32, 100$, $L = 8$, SNR= 15dB).

TABLE I
DATA THROUGHPUT FOR THE EM AND LS ESTIMATORS.

SNR (dB)	SER	Throughput (EM/LS)
10	3×10^{-1}	0.88 / 0.78
15	1.5×10^{-1}	0.8 / 0.68
20	6×10^{-2}	0.8 / 0.63
25	1.5×10^{-2}	0.85 / 0.58
30	2×10^{-3}	0.95 / 0.73

$$\check{V}=\frac{1}{G^2}\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\left|A^2\mathcal{X}\xi_j-AGs_{1i}\right|^2+\frac{1}{G^2}\left\|A^2\mathcal{X}\boldsymbol{t}_2-AG\boldsymbol{t}_1\right\|^2$$

and

$$\check{W}=\frac{1}{G^2}\Bigg|\sum_{i=1}^{N}\sum_{j=1}^{M}\beta_{i,j}^{(t)}\left(Gz_i-A\mathcal{I}\xi_j\right)^*\left(A^2\mathcal{X}\xi_j-AGs_{1i}\right)+ \left(G\boldsymbol{y}-A\mathcal{I}\boldsymbol{t}_2\right)^H\left(A^2\mathcal{X}\boldsymbol{t}_2-AG\boldsymbol{t}_1\right)\Bigg|.$$

The derivative of (15) w.r.t. $|a|$ is

$$\frac{dQ(|a|;\boldsymbol{\theta}^{(t)})}{d|a|}=-\frac{NA^2}{A^2|a|+1}-\frac{A^2\check{V}|a|^2+2\check{V}|a|-A^2\check{U}-2\check{W}}{\sigma^2(A^2|a|+1)^2}. \quad (16)$$

Setting $\frac{dQ(|a|;\boldsymbol{\theta}^{(t)})}{d|a|}=0$, we obtain the quadratic equation

$$A^2\check{V}|a|^2+(2\check{V}+NA^4\sigma^2)|a|+NA^2\sigma^2-A^2\check{U}-2\check{W}=0. \quad (17)$$

Solving (17), we finally get

$$|a|^{(t+1)}=\frac{1}{2A^2\check{V}}\Bigg(-(2\check{V}+NA^4\sigma^2)+ \sqrt{(2\check{V}+NA^4\sigma^2)^2-4A^2\check{V}(NA^2\sigma^2-A^2\check{U}-2\check{W})}\Bigg). \quad (18)$$

As we can see from (12), (14) and (18), the computational complexity of each EM iteration is $O(MN)$, i.e., it is linear in the number of data samples for a given modulation order.

IV. SIMULATION RESULTS

In this section, we investigate through simulations the MSE performance of the derived EM algorithm. We model h and g as independent complex Gaussian RVs with mean zero and variance $\gamma^2 = 1$. We assume that $P_1 = P_2 = P_r$ and average our results over 100 independent channel realizations. The data symbols are generated from square QAM constellations. We consider the modulation orders $M = 4, 16, 64$. The pilot vectors $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are obtained using $M = 4$ and chosen to be orthogonal. In our plots, we consider the total MSE, which is the sum of the MSE for the estimation of a and b. As a benchmark on the MSE performance of the EM algorithm, we use the semi-blind Cramer Rao bound (CRB) for reciprocal channel estimation. This bound can be obtained by following the same approach as in [7].

In Fig. 1, we plot the MSE performance of the derived semi-blind EM algorithm versus SNR for $L = 8$ and $N = 32$, where the SNR is defined as $10\log_{10}\frac{P_2}{\sigma^2}$. The channel estimates are obtained after 4 iterations of the EM algorithm. For comparison, we also plot the MSE of the LS estimator that only uses the pilot samples, as well as the semi-blind CRB. As we can see from Fig. 1, the MSE performance of the EM algorithm is very close to the CRB for the whole SNR range. Moreover, the EM algorithm provides substantially higher accuracy than the LS estimator. The gain in accuracy depends on the modulation order: the lower the modulation order the higher the gain.

We next consider the convergence of the EM algorithm. In Fig. 2, we plot the MSE of the EM algorithm versus the number of iterations for $N = 32$ and $N = 100$, assuming 8 pilots and an SNR of 15dB. Fig. 2 shows that the number of iterations needed for convergence is small for all modulation orders. In all cases, convergence is achieved within at most 12 iterations (as few as 4 iterations are sufficient in some cases). Convergence becomes slightly slower as the modulation order increases and as the number of data samples increases.

Finally, in Table I, we compare the achievable throughput of the proposed EM algorithm and the LS estimator at different SNR and SER values, assuming that the channel is constant for the duration of 40 data samples and that QPSK modulation is used. To obtain the achievable throughput, we first determine through simulations the average number of pilots required by each algorithm to achieve a certain SER performance at a given SNR. The corresponding number of data symbols that can be transmitted is then divided by 40 to obtain the throughput. As we can see from Table I, substantial improvements in throughput (up to 27%) are possible by using the EM algorithm.

V. Conclusions

In this work, we derived the EM algorithm for semi-blind channel estimation in AF TWRNs assuming reciprocal flat-fading channel conditions. Our simulations showed that the derived EM algorithm outperforms the pilot-based LS estimator even with a limited number of data symbols and performs very close to the corresponding semi-blind CRB. The proposed algorithm requires only a small number of low-complexity iterations to converge. Hence, the accuracy gains of semi-blind estimation predicted in [7] through CRB analysis can indeed be realized for reciprocal channels at an affordable computational price. Finally, by virtue of its higher accuracy the EM algorithm requires a smaller number of pilots compared to the LS estimator, which allows for transmitting more data symbols, resulting in a significant improvement in throughput and spectral efficiency.

References

[1] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

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

[3] ——, "Channel estimation for OFDM modulated two-way relay networks," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4443–4455, Nov. 2009.

[4] B. Jiang, F. Gao, X. Gao, and A. Nallanathan, "Channel estimation and training design for two-way relay networks with power allocation," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2022–2032, Jun. 2010.

[5] F. Roemer and M. Haardt, "Tensor-based channel estimation (TENCE) and iterative refinements for two-way relaying with multiple antennas and spatial reuse," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5720–5735, Nov. 2010.

[6] E. de Carvalho and D. Slock, "Blind and semi-blind FIR multichannel estimation: (global) identifiability conditions," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1053–1064, Apr. 2004.

[7] S. Abdallah and I. N. Psaromiligkos, "Exact Cramer-Rao bounds for semi-blind channel estimation in amplify-and-forward two-way relay networks employing square QAM modulation," *ArXiv pre-print cs.IT/1207.5483*, Jul. 2012.