Superimposed Training and Channel Estimation for Two-Way Relay Networks

Xiaoyan Xu*, Jianjun Wu*, Shubo Ren*, Xi Luan* and Haige Xiang*
*EECS, Peking University, 100871, China
Email: g.ericaxu@gmail.com, just@pku.edu.cn

***Abstract*—In this paper, the superimposed training strategy is introduced into the OFDM modulated amplify-and-forward (AF) two-way relay network (TWRN) to simplify the channel estimation at the destination, and the closed-form Bayesian Cramér-Rao lower bound (CRLB) is derived for the estimation of block-fading frequency-selective channels, which is used to guide the optimal training design. Through the superposition of an additional training vector at the relay under certain power allocation scheme, the separated channel information can be obtained directly at the destination. The Bayesian CRLB is derived for the random channel parameters, and orthogonal training vectors from the two source nodes are required to keep the Bayesian CRLB practical, due to the self-interference in the TWRN. A set of training vectors obtained from the minimization of the Bayesian CRLB are applied in a specific suboptimal channel estimation algorithm, and the mean-square error (MSE) performance is provided to verify the Bayesian CRLB results.**

Keywords*—Two-way relay, channel estimation, Bayesian Cramér-Rao lower bound(CRLB), training design, mean-square error.

I. Introduction

With the combination of cooperative communication, two-way relay network (TWRN) emerged a few years ago, and has attracted a great deal of interest recently [1], [2], due to its improved spectral efficiency over one-way relay network (OWRN). In a TWRN, a major difficulty lies in how to effectively recover the data transmitted over an unknown fading channel from the other source terminal. Channel estimation in TWRN has been studied in [3]–[8]. Specifically, in [4] and [5], the cascaded source-relay-source channels were estimated using block-based training under the assumption of time-invariant frequency-selective fading channels. Different from [4] and [5], where the relay only amplifies and forwards the received signal, [6] allowed the relay to first estimate the channel parameters and then allocated the powers for these parameters. The channel estimation problem was extended to the TWRN with multiple antennas at all the three nodes in [7]. A blind channel estimation algorithm based on the second order statistics of the received signal was proposed in [8] for AF TWRN. Inspired by the superimposed training in point-to-point communications, [9] and [10] designed a superimposed training strategy in AF OWRN.

In this work, we introduce the superimposed training s-trategy into OFDM modulated AF TWRN to simplify the channel estimation at the destination, and derive the closed-form Bayesian Cramér-Rao lower bound (CRLB) for the estimation of block-fading frequency-selective channels. The relay superimposes its own training signal over the received signal before forwarding it out, which provides the separated channel information, and the total relay power is allocated between the two parts reasonably. Due to the self-interference in the TWRN, the Bayesian CRLB is different from its counterpart in the OWRN, which requires orthogonal training vectors from the two source nodes and more complex constraints for the optimal training design. The simulation is provided to verify the Bayesian CRLB results by the MSE performance of a specific suboptimal channel estimation algorithm.

The structure of the rest of the paper is as follows. The system model of superimposed training in the TWRN is given in Section II. In Section III, the Bayesian CRLB of superimposed training based channel estimation is derived, the training design from the Bayesian CRLB and a suboptimal channel estimation algorithm are described. Finally, the simulation results are provided in Section IV and conclusions are drawn in Section V.

II. System Model

Consider a TWRN network where two nodes, T_1 and T_2, exchange information through one relay node R, as shown in Fig. 1. The transmission is divided into two phases. During Phase I, both T_1 and T_2 send a signal frame to R via an uplink manner, whereas during Phase II, R processes the received signals and broadcasts them to T_1 and T_2.

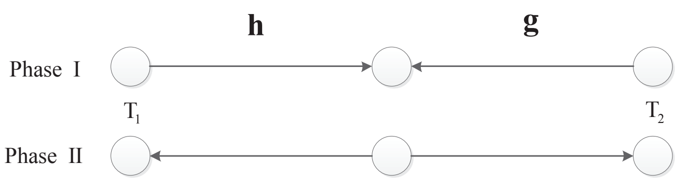


Fig. 1. Two-way relay network

The baseband channel between T_1 and R is denoted by $\mathbf{h} = [h_0, h_1, \cdots, h_{L_h-1}]^T$, and the one between T_2 and R is denoted by $\mathbf{g} = [g_0, g_1, \cdots, g_{L_g-1}]^T$, where L_h, L_g represents the number of the taps of the corresponding channel. Both $\mathbf{h}$ and $\mathbf{g}$ are assumed as zero-mean circularly symmetric complex Gaussian random vectors and remain unchanged at least for one round of data exchange. For simplicity, consider that $h_l \in CN(0, \sigma_{h,l}^2)$ and $g_l \in CN(0, \sigma_{g,l}^2)$ are independent from each other. For time-division-duplexing (TDD), the channel can be considered reciprocal. The average transmission powers of $\mathrm{T}_1, \mathrm{T}_2$ and R are denoted as P_s, P_s and P_r, respectively.

In the OFDM modulated TWRN, the first block in one data frame is devoted to training, as shown in Fig. 2.

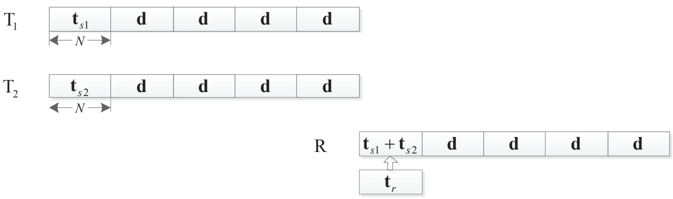


Fig. 2. Structure of one data frame with superimposed training

Suppose the OFDM block length is N, and denote the training vectors from T_1 and T_2 in frequency domain as $\tilde{\mathbf{t}}_{s1} = \left[\tilde{t}_{s1,0}, \tilde{t}_{s1,1}, \cdots, \tilde{t}_{s1,N-1}\right]^T$ and $\tilde{\mathbf{t}}_{s2} = \left[\tilde{t}_{s2,0}, \tilde{t}_{s2,1}, \cdots, \tilde{t}_{s2,N-1}\right]^T$, respectively. The training power is constrained by

$$\mathbf{t}_{s1}^H \mathbf{t}_{s1} = \tilde{\mathbf{t}}_{s1}^H \tilde{\mathbf{t}}_{s1} \leq NP_s, \quad \mathbf{t}_{s2}^H \mathbf{t}_{s2} = \tilde{\mathbf{t}}_{s2}^H \tilde{\mathbf{t}}_{s2} \leq NP_s. \tag{1}$$

To avoid the inter-block interference (IBI), both T_1 and T_2 insert the cyclic prefix (CP) of length $L_{cp,T_1,T_2} \geq \max\{L_h - 1, L_g - 1\}$ in the front of the OFDM block before transmission. Relay R superimposes a new training $\mathbf{t}_r$ over the received signal The following constraint should be satisfied for the training block,

$$\mathrm{E}\left\{\|\mathbf{r}_t\|^2\right\} = \alpha^2\left(\sigma_h^2 \mathbf{t}_{s1}^H \mathbf{t}_{s1} + \sigma_g^2 \mathbf{t}_{s2}^H \mathbf{t}_{s2} + N\sigma_n^2\right) + \mathbf{t}_r^H \mathbf{t}_r \leq NP_r, \tag{2}$$

where $\sigma_h^2 = \sum_{i=0}^{L_h-1} \sigma_{h,i}^2$ and $\sigma_g^2 = \sum_{i=0}^{L_g-1} \sigma_{g,i}^2$. We can prove when the optimal channel estimation is achieved, both the equalities in (1) and (2) must hold, regardless of the channel estimation algorithm used. Let $P_t = \frac{\mathbf{t}_r^H \mathbf{t}_r}{N}$ be the average power assigned for superimposed training at R, then,

$$P_t = P_r - \alpha^2\left(\sigma_h^2 P_s + \sigma_g^2 P_s + \sigma_n^2\right).$$

P_t can be seen as the function of α while the range of α is $\left(0, \sqrt{\frac{P_r}{\sigma_h^2 P_s + \sigma_g^2 P_s + \sigma_n^2}}\right)$. Relay R then adds a new CP that consists of the last $L_h - 1$ entries in $\mathbf{r}_t$ and forwards the superimposed signal to T_1 and T_2. For symmetry, only the process at T_1 is discussed. Let $\mathbf{w}_1 = \mathbf{h} \otimes \mathbf{h}$, and $\mathbf{w}_2 = \mathbf{h} \otimes \mathbf{g}$, then the received signal at T_1, after CP removal, is

$$\begin{aligned} \mathbf{y}_1 = & \alpha \Phi_{2L_h-1}(\mathbf{t}_{s1}) \mathbf{w}_1 + \alpha \Phi_{L_h+L_g-1}(\mathbf{t}_{s2}) \mathbf{w}_2 \\ & + \Phi_{L_h}(\mathbf{t}_r) \mathbf{h} + \alpha \mathbf{H} \mathbf{n}_r + \mathbf{n}_1, \end{aligned} \tag{3}$$

where $\mathbf{n}_1$ is the $N \times 1$ AWGN vector with variance σ_n^2 on each entries. The task of channel estimation in the TWRN is to find $\mathbf{h}$ and $\mathbf{g}$ from $\mathbf{y}_1$.

III. Training Design From Bayesian Cramér-Rao Bound and Suboptimal Channel Estimation

A. *Bayesian Cramér-Rao Bound*

The Cramér-Rao bound (CRB) for the estimation of deterministic parameters is given by the inverse of the FIM, and Van Trees derived an analogous bound to the CRB for random variables, referred to as "Bayesian CRB" (BCRB) [11]. With the assist of BCRB, the performance of the suboptimal estimators in the TWRN can be assessed, and the optimal training design could be obtained. Let $\theta = \left[\mathbf{h}^T, \mathbf{g}^T\right]^T$ be the Gaussian random vector to be estimated, and the FIM is defined as [11],

$$\mathcal{J} = \mathrm{E}\left\{\frac{\partial \ln p(\mathbf{y}_1, \theta)}{\partial \theta^*}\left(\frac{\partial \ln p(\mathbf{y}_1, \theta)}{\partial \theta^*}\right)^H\right\} = \begin{bmatrix} \mathcal{J}_{11} & \mathcal{J}_{12} \\ \mathcal{J}_{12}^H & \mathcal{J}_{22} \end{bmatrix},$$

where the expectation is taken over the joint probability density function $p(\mathbf{y}_1, \theta)$. The BCRB is the lower bound of any error covariance matrix $\mathrm{E}\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^H\}$.

The FIM is computed as

$$\begin{aligned} \mathcal{J}_{11} &= 4\alpha^2 c_2 \Phi_{L_h}^H(\mathbf{t}_{s1}) \Phi_{L_h}(\mathbf{t}_{s1}) + \alpha^2 \sigma_g^2 c_1 \Phi_{L_h}^H(\mathbf{t}_{s2}) \Phi_{L_h}(\mathbf{t}_{s2}) \\ &\quad + c_1 \Phi_{L_h}^H(\mathbf{t}_r) \Phi_{L_h}(\mathbf{t}_r) + N\sigma_n^4 \alpha^4 c_3 \mathbf{I} + \mathbf{R}_{\mathbf{h}}^{-1}, \\ \mathcal{J}_{12} &= 2\alpha^2 c_2 \Phi_{L_h}^H(\mathbf{t}_{s1}) \Phi_{L_g}(\mathbf{t}_{s2}), \\ \mathcal{J}_{21} &= 2\alpha^2 c_2 \Phi_{L_g}^H(\mathbf{t}_{s2}) \Phi_{L_h}(\mathbf{t}_{s1}), \\ \mathcal{J}_{22} &= \alpha^2 c_2 \Phi_{L_g}^H(\mathbf{t}_{s2}) \Phi_{L_g}(\mathbf{t}_{s2}) + \mathbf{R}_{\mathbf{g}}^{-1}, \end{aligned}$$

where $c_i, i = 1, 2, 3$ are defined in the following, $\mathbf{R}_{\mathbf{h}}$ and $\mathbf{R}_{\mathbf{g}}$ are the covariance matrices of $\mathbf{h}$ and $\mathbf{g}$, respectively.

As we can see, $\mathcal{J}_{12}$ and $\mathcal{J}_{21}$ are zero matrices as long as $\mathbf{t}_{s1}$ and $\mathbf{t}_{s2}$ are orthogonal. Since $\mathbf{t}_{s1}$ and $\mathbf{t}_{s2}$ are known training vectors from T_1 and T_2, the orthogonality can be guaranteed. Since $\mathcal{J}_{12} = \mathbf{0}$ and $\mathcal{J}_{21} = \mathbf{0}$, the BCRBs for $\mathbf{h}$ and $\mathbf{g}$ can be separately expressed as

$$\mathrm{BCRB}_{\mathbf{h}} = \mathcal{J}_{11}^{-1}, \qquad \mathrm{BCRB}_{\mathbf{g}} = \mathcal{J}_{22}^{-1},$$

and the channel error covariances are lower bounded by

$$\mathrm{Cov}_{\mathbf{h}} = \mathrm{E}\left\{\Delta\mathbf{h}\Delta\mathbf{h}^H\right\} \succeq \mathrm{BCRB}_{\mathbf{h}},$$

$$\mathrm{Cov}_{\mathbf{g}} = \mathrm{E}\left\{\Delta\mathbf{g}\Delta\mathbf{g}^H\right\} \succeq \mathrm{BCRB}_{\mathbf{g}}.$$

The optimal training vectors in the TWRN can be designed from the Bayesian CRLB through the minimization of $\mathrm{MSE}_{\mathbf{h}}$ and $\mathrm{MSE}_{\mathbf{g}}$, under the power constraints in (1) and (2) and the orthogonality constraint of $\mathbf{t}_{s1}$ and $\mathbf{t}_{s2}$.

B. *Training Design From Bayesian Cramér-Rao Bound*

The two optimizations of $\mathbf{h}$ and $\mathbf{g}$ from BCRB are formulated as

$$\begin{aligned} (\mathrm{P1}): \min_{\mathbf{t}_{s1}, \mathbf{t}_{s2}, \mathbf{t}_r} &\ \mathrm{tr}(\mathrm{BCRB}_{\mathbf{h}}) \\ \text{s.t.} &\ \mathbf{t}_{s1}^H \mathbf{t}_{s1} = NP_s, \\ &\ \mathbf{t}_{s2}^H \mathbf{t}_{s2} = NP_s, \\ &\ \mathbf{t}_r^H \mathbf{t}_r = NP_t. \end{aligned} \qquad \begin{aligned} (\mathrm{P2}): \min_{\mathbf{t}_{s2}} &\ \mathrm{tr}(\mathrm{BCRB}_{\mathbf{g}}) \\ \text{s.t.} &\ \mathbf{t}_{s2}^H \mathbf{t}_{s2} = NP_s. \end{aligned}$$

Lemma 1: The training sequences satisfying

$$\begin{aligned}
&(\mathrm{C1}): && \Phi_{L_h}^H(\mathbf{t}_{s1})\,\Phi_{L_h}(\mathbf{t}_{s1}) = NP_s\mathbf{I},\\
&(\mathrm{C2}): && \Phi_{\max\{L_h,L_g\}}^H(\mathbf{t}_{s2})\,\Phi_{\max\{L_h,L_g\}}(\mathbf{t}_{s2}) = NP_s\mathbf{I},\\
&(\mathrm{C3}): && \Phi_{L_h}^H(\mathbf{t}_r)\,\Phi_{L_h}(\mathbf{t}_r) = NP_t\mathbf{I},\\
&(\mathrm{C4}): && \Phi_{L_h}^H(\mathbf{t}_{s1})\,\Phi_{L_g}(\mathbf{t}_{s2}) = \mathbf{0},
\end{aligned}$$

are optimal solutions to (P1) and (P2).

Proof: For any positive-definite matrix $\mathbf{X}$, there is

$$\mathrm{tr}\left(\mathbf{X}^{-1}\right) \geq \sum_{i=0}^{N-1}\frac{1}{[\mathbf{X}]_{ii}},$$

and the equality holds when $\mathbf{X}$ is diagonal. For (P1), $\mathbf{R}_{\mathbf{h}}^{-1}$ is a constant diagonal matrix, and the diagonal elements of $\Phi_{L_h}^H(\mathbf{t}_{s1})\,\Phi_{L_h}(\mathbf{t}_{s1})$, $\Phi_{L_h}^H(\mathbf{t}_{s2})\,\Phi_{L_h}(\mathbf{t}_{s2})$, and $\Phi_{L_h}^H(\mathbf{t}_r)\,\Phi_{L_h}(\mathbf{t}_r)$ are constant NP_s, NP_s, and NP_t, respectively. Therefore,

$$\mathrm{tr}\left(\mathrm{BCRB}_{\mathbf{h}}\right) \geq \sum_{i=0}^{L_h-1}\frac{1}{N\left[P_s\left(4\alpha^2c_2+\alpha^2\sigma_g^2c_1\right)+c_1P_t+\sigma_n^4\alpha^4c_3+\frac{1}{N\sigma_{h,i}^2}\right]},$$

and the RHS is the lower bound of the $\mathrm{tr}\left(\mathrm{BCRB}_{\mathbf{h}}\right)$, since it is not related with any training vector. The lower bound is achieved when $\Phi_{L_h}^H(\mathbf{t}_{s1})\,\Phi_{L_h}(\mathbf{t}_{s1}) = NP_s\mathbf{I}$, $\Phi_{L_h}^H(\mathbf{t}_{s2})\,\Phi_{L_h}(\mathbf{t}_{s2}) = NP_s\mathbf{I}$ and $\Phi_{L_h}^H(\mathbf{t}_r)\,\Phi_{L_h}(\mathbf{t}_r) = NP_t\mathbf{I}$. Similarly, as for (P2), the lower bound is achieved when $\Phi_{L_g}^H(\mathbf{t}_{s2})\,\Phi_{L_g}(\mathbf{t}_{s2}) = NP_s\mathbf{I}$. From the two proofs we can obtain (C1), (C2) and (C3). (C4) is required to keep the BCRB compact. ∎

An example of such training sequences is provided here.

$$\left|\tilde{t}_{s1,i}\right|^2 = P_s,\ \left|\tilde{t}_{s2,i}\right|^2 = P_s,\ \left|\tilde{t}_{r,i}\right|^2 = P_t,\quad i = 0,\cdots,N-1,$$

$$\tilde{t}_{s1,i}^*\tilde{t}_{s2,i} = P_s\sqrt{N}\,[\mathbf{F}]_{ik} = P_se^{-j2\pi ik/N},\quad \forall k\in\{(L_h+L_g-1),\cdots,(N-L_g)\},$$

$$\tilde{t}_{s1,i}^*\tilde{t}_{r,i} = \sqrt{P_sP_tN}\,[\mathbf{F}]_{ik} = \sqrt{P_sP_t}e^{-j2\pi ik/N},\quad \forall k\in\{(2L_h-1),\cdots,(N-L_h)\},$$

$$\tilde{t}_{s2,i}^*\tilde{t}_{r,i} = \sqrt{P_sP_tN}\,[\mathbf{F}]_{ik} = \sqrt{P_sP_t}e^{-j2\pi ik/N},\quad \forall k\in\{(L_h+L_g-1),\cdots,(N-L_h)\}.$$

C. Suboptimal Channel Estimation

Using the optimal training vectors, we refer to the suboptimal estimators to verify the Bayesian CRLB results, since channel statistics are assumed known. Consider the case when the training length is sufficiently large, $N \geq 4L_h+L_g-2$.

Define the new $(4L_h+L_g-2)\times 1$ channel vector as $\mathbf{u} = \left[\mathbf{w}_1^T,\mathbf{w}_2^T,\mathbf{h}^T\right]^T$, whose covariance matrix $R_{\mathbf{u}}$ is a diagonal matrix

$$R_{\mathbf{u}} = \mathrm{E}_\theta\left\{\mathbf{u}\mathbf{u}^H\right\} = \begin{bmatrix} R_{\mathbf{w}_1} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & R_{\mathbf{w}_2} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & R_{\mathbf{h}}\end{bmatrix},$$

where $R_{\mathbf{w}_1}$, $R_{\mathbf{w}_2}$, $R_{\mathbf{h}}$ are the covariance matrix of $\mathbf{w}_1$, $\mathbf{w}_2$ and $\mathbf{h}$, respectively. The received training at T_1 is then

$$\mathbf{y}_1 = \Sigma\mathbf{u}+\mathbf{n},$$

where $\Sigma = \left[\alpha\Phi_{2L_h-1}(\mathbf{t}_{s1}),\ \alpha\Phi_{L_h+L_g-1}(\mathbf{t}_{s2}),\ \Phi_{L_h}(\mathbf{t}_r)\right]$. The standard Linear MMSE (LMMSE) estimate of $\mathbf{u}$ is,

$$\hat{\mathbf{u}} = R_{\mathbf{u}}\Sigma^H\left(\Sigma R_{\mathbf{u}}\Sigma^H+R_{\mathbf{n}}\right)^{-1}\mathbf{y}_1,$$

whose error covariance is

$$\mathrm{Cov}_{\mathbf{u}} = \left(R_{\mathbf{u}}^{-1}+\Sigma^HR_{\mathbf{n}}\Sigma\right)^{-1},$$

and the corresponding MSE is $\mathrm{MSE}_{\mathbf{u}} = \mathrm{tr}\left(\mathrm{Cov}_{\mathbf{u}}\right)$.

With the estimates $\hat{\mathbf{w}}_1$, $\hat{\mathbf{w}}_2$ and $\hat{\mathbf{h}}$, the initial estimate of $\hat{\mathbf{g}}$ can be computed from the de-convolution approach as

$$\hat{\mathbf{g}} = \Psi_{L_g}^\dagger(\hat{\mathbf{h}})\hat{\mathbf{w}}_2,$$

where $\Psi_{L_g}^\dagger(\hat{\mathbf{h}})$ is the $(L_h+L_g-1)\times L_g$ column-wise circulant matrix with the first column $[\hat{\mathbf{h}}^T\ \ \mathbf{0}_{1\times(L_g-1)}]^T$.

After obtaining the initial channel estimates, an iterative method is applied to updates $\hat{\mathbf{h}}$ and $\hat{\mathbf{g}}$ by turns. In each iteration, first $\hat{\mathbf{h}}$ is substituted back into (3) to update the LMMSE estimate $\hat{\mathbf{g}}$ using interference cancellation, and then the updated $\hat{\mathbf{g}}$ is substituted back into (3) to update the LMMSE estimate $\hat{\mathbf{h}}$ using interference cancellation. The iteration goes on until a certain stopping criterion is satisfied.

IV. Simulations

In this section, we provide some numerical results to illustrate our studies. The optimal training sequences are used and only the channel estimation at T_1 is considered. Let $L_h = L_g = 6$, all channel taps have unit variances, and the noise variance is also set as 1. For simplicity, $P_s = P_r$ is assumed and the SNR is defined as $P_s/\sigma_n^2 = P_s$. The total training power is the same in all the scenarios. The OFDM block length N is taken as 64, following IEEE 802.11a.

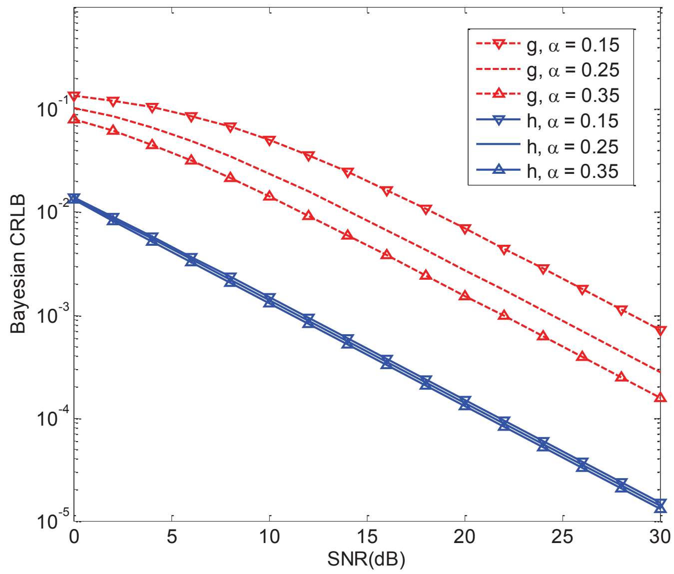


Fig. 3. Bayesian CRLBs versus SNR

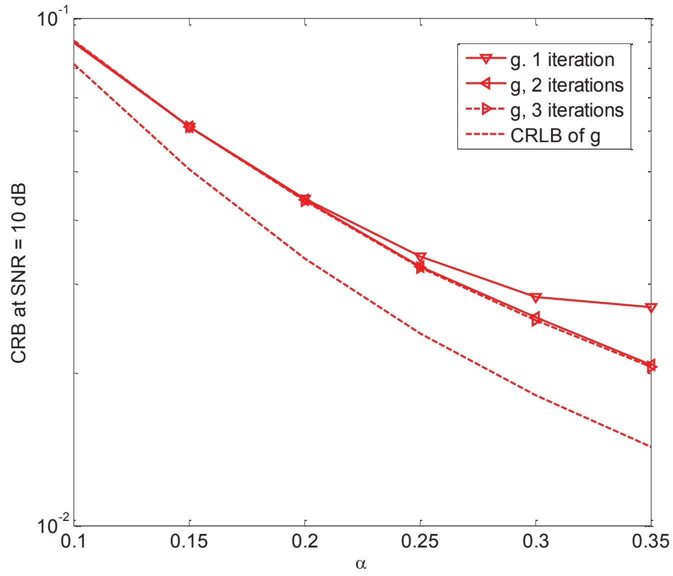


Fig. 5. channel estimation NMSEs versus α at SNR = 10 dB

The theoretical Bayesian CRLBs is calculated for $\mathbf{h}$ and $\mathbf{g}$ as the function of both SNR and α, as in Fig. 3. We can see that the Bayesian CRLB of $\mathbf{h}$ is always smaller than the Bayesian CRLB of $\mathbf{g}$ for any given α.

The NMSE performance of the suboptimal estimator versus SNR is displayed together with the theoretical Bayesian CRLBs in Fig. 4 for channel estimation of $\mathbf{h}$ and $\mathbf{g}$ with α fixed as 0.25. The NMSE performance of channel estimation for $\mathbf{h}$ is always better than the counterpart for $\mathbf{g}$, which is consistent with the comparison between the two CRLBs.

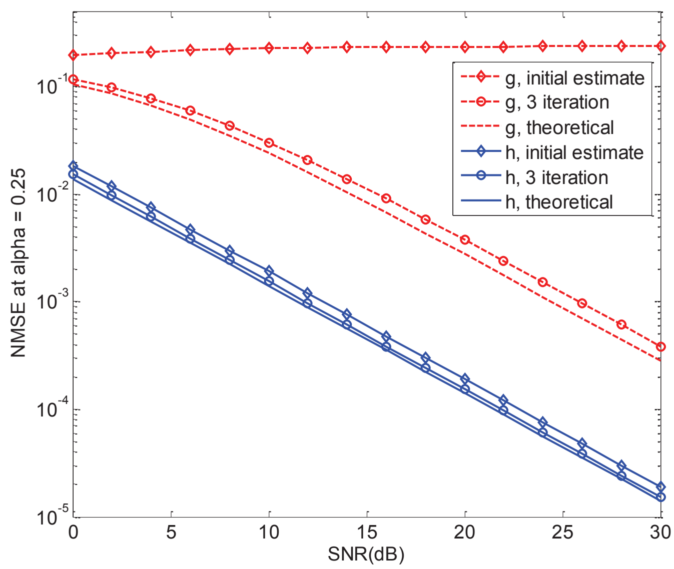


Fig. 4. Channel estimation NMSEs versus SNR for α = 0.25

It is also found that the iterative algorithm converges in three iterations, as Fig. 5, the NMSE curve after the third iteration almost coincides with the NMSE curve after the previous iteration.

V. Conclusion

In this paper, we have introduced the superimposed training strategy into OFDM modulated AF TWRN, which superimposes a new training vector at the relay, and provides the separated channel information at the destination to simplify the channel estimation. Under the circumstances that the random parameter vector to be estimated contains only fading channel coefficients, the closed-form Bayesian CRLB has been derived for the estimation of the block-fading frequency-selective channels, and then used to guide the optimal training design, which is more complicated than OWRN due to the self-interference in TWRN.

Acknowledgment

This work is partly supported by the National Science Foundation of China (Grant No. NFSC #61071083, #61371073) and the National High-Tech Research and Development Program of China (863 Program), No.2012AA01A506. Corresponding author: Jianjun Wu, E-mail: just@pku.edu.cn.

References

[1] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 379-389, 2007.

[2] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," *in Proceedings of the IEEE International Conference on Communications (ICC'07)*, pp. 707-712, Glasgow, Scotland, June 2007.

[3] F. Gao, R. Zhang, and Y. C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Transactions on Communications*, vol. 57, no. 10, pp. 3024-3033, 2009.

[4] F. Gao, R. Zhang, and Y. C. Liang, "Channel estimation for OFDM modulated two-way relay networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4443-4455, 2009.

[5] W. Yang, Y. Cai, J. Hu, and W. Yang, "Channel estimation for two-way relay OFDM networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, 2010.

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

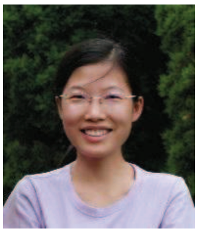
[7] T. H. Pham, Y. C. Liang, Nallanathan, A., and Garg, H.K., "Optimal training sequences for channel estimation in bi-directional relay networks with multiple antennas," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 474-479, 2010.

[8] X. Liao, L. Fan, and F. Gao, "Blind channel estimation for OFDM modulated two-way relay network," *in Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC'10)*, pp. 1-5, Sydney, NSW, April 2010.

[9] F. Gao, B. Jiang, X. Gao, and X. D. Zhang, "Superimposed training based channel estimation for OFDM modulated AF relay networks," *in Proceedings of the IEEE International Conference on Communications (ICC'11)*, pp. 1-5, Kyoto, Japan, June 2011.

[10] F. Gao, B. Jiang, X. Gao, and X. D. Zhang, "Superimposed training based channel estimation for OFDM modulated amplify-and-forward relay networks," *IEEE Transactions on Communications*, vol. 59, no. 7, pp. 2029-2039, 2011.

[11] H. L. Van Trees, *Detection Estimation and Modulation Theory*, pt. 1. Wiley, 1968.



Xiaoyan Xu received her bachelor degree in electrical engineering from Peking University, Beijing, China, in 2008. Since 2008, she has been a PhD candidate in Institution of Advanced Communications, Peking University, China. Her research interests are in the area of satellite mobile communications and wireless communications, such as MSS based on LTE, compressive sensing, cooperative communications. Email: g.ericaxu@gmail.com.

Haige Xiang received the bachelor of Engineering in Electronics from Peking University, Beijing, P. R. China, in 1964. Since 1964, he has been a professor at Electronics Department at Peking University. His research interests include information theory, wireless communications, channel coding, signal processing in communications and satellite communications. Professor Xiang is a member in IEEE. Email: xianghg@pku.edu.cn.