Doubly Selective Channel Estimation for Amplify-and-Forward Relay Networks

Gongpu Wang†, Feifei Gao[1,2], Jiaru Lin‡ and Chintha Tellambura*
†School of Computer and Information Technology, Beijing Jaiotong University, China
[1]National Mobile Communications Research Laboratory, Southeast University, Nanjing, China
[2]State Key Laboratory of Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Automation, Tsinghua University, Beijing, 100084, China
‡Key Laboratory of Universal Wireless Communications, Ministry of Education, China
*Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada
Email: gongpuwang@hotmail.com, feifeigao@ieee.org, jrlin@bupt.edu.cn, chintha@ece.ualberta.ca.

***Abstract*—In this paper, the estimation of doubly selective channel is considered for *amplify-and-forward* (AF) relay networks. The complex exponential basis expansion model (CE-BEM) is chosen to describe the time-varying channel, from which the infinite channel parameters are mapped onto finite ones. Since direct estimation of these coefficients encounters high computational complexity and large spectral cost, we develop an efficient estimator targeting at some specially defined channel parameters. The training sequence design that can minimize the channel estimation mean-square error is also proposed.**

I. Introduction

Wireless relay networks have attracted a lot of attention since the pioneer work [1]. Like any other wireless communication system, the relay network performs better with better channel estimates. Channel state information (CSI) is usually estimated and tracked by periodic training signals [2]. Flat-fading channel estimation was discussed in [3] and frequency-selective channel estimators were developed in [4], respectively. All these estimators assumed that channels are time-invariant during a certain period.

However in many practical cases, the source node, the relay node and the destination node can all be mobile. The relative motion between any two nodes causes Doppler shift and thus makes the channel time-varying [5]. Therefore, the relay network is expected to operate under doubly selective channels. To our best knowledge, estimation techniques for this case have not yet been developed. This motivates our current work.

The doubly selective channel is typically approximated in two ways: the autoregressive (AR) process [2] or the basis expansion model (BEM) [6]. AR model describes the channel variation through a symbol-by-symbol update manner, while BEM expresses the doubly selective channel as the superpositions of time-varying basis functions weighted by time-invariant coefficients. In this paper, we adopt the complex exponential BEM (CE-BEM) [7] and develop an efficient estimator to find the channel parameters, which can sufficiently aid the data detection. The optimal training sequence design that minimizes parameter mean-square error (MSE) is also proposed.

This work was supported in part by National Basic Research Program of China (973 Program) 2009CB320401 and by National Natural Science Foundation of China 61171100. The work of F. Gao was supported by the open research fund of National Mobile Communications Research Laboratory, Southeast University (No. 2011D02) and by the Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20110002120059).

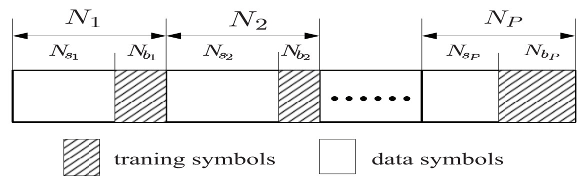

Fig. 1. Structure of one transmission block.

II. System model

Consider an *amplify-and-forward* (AF) relay network with one source node $\mathbb{S}$, one relay node $\mathbb{R}$ and one destination node $\mathbb{D}$. Let $h(i;l)$ denote the channel between $\mathbb{S}$ and $\mathbb{R}$, $g(i;l)$ denote the channel between $\mathbb{R}$ and $\mathbb{D}$, respectively. The lengths of both channels are assumed as $L+1$ without loss of generality. According to the CE-BEM analysis in [6], [7], we can express the doubly selective channels as

$$h(i;l)=\sum_{q=0}^{Q_1} h_q(l)e^{j2\pi(q-Q_1/2)i/N}, \tag{1}$$

$$g(i;l)=\sum_{q=0}^{Q_2} g_q(l)e^{j2\pi(q-Q_2/2)i/N}, \tag{2}$$

where $0\leq i\leq N-1$, $0\leq l\leq L$, $Q_m (m=1,2)$ is the number of basis, and N is the number of symbols during one transmission. The value of Q_m is $2\lceil f_{d_m}NT_s\rceil$ where f_{d_1} means the maximum Doppler shift of the link $\mathbb{S}\rightarrow\mathbb{R}$, while f_{d_2} denotes the maximum Doppler shift of the link $\mathbb{R}\rightarrow\mathbb{D}$. The CE-BEM coefficients $h_q(l)$ and $g_q(l)$ are assumed as zero-mean, complex Gaussian random variables with variance $\sigma^2_{h,q,l}$ and $\sigma^2_{g,q,l}$ respectively.

To simplify the notation, we assume $f_{d_1} = f_{d_2}$ and $Q_1 = Q_2 = Q$. Further denote $w_q = 2\pi(q - Q/2)/N$ and define

$$\mathbf{h}_q = [h_q(0), h_q(1), \cdots, h_q(L)]^T, \tag{3}$$
$$\mathbf{g}_q = [g_q(0), g_q(1), \cdots, g_q(L)]^T, \quad q \in [0, Q]. \tag{4}$$

We propose a new transmission scheme as shown in Fig. 1. Each transmission block that contains N symbols is divided into P subblocks. Assume the kth subblock contains N_k symbols, of which N_{s_k} symbols are data and are represented by $\mathbf{s}_k$, while N_{b_k} symbols are pilots and are represented by $\mathbf{b}_k$. The total number of data symbols is $N_s = \sum_{k=1}^{P} N_{s_k}$ and the total number of pilots is $N_p = \sum_{k=1}^{P} N_{b_k}$. With such a structure, we can represent the whole block as a vector $\mathbf{x} = [\mathbf{s}_1^T, \mathbf{b}_1^T, \cdots, \mathbf{s}_P^T, \mathbf{b}_P^T]$.

During the first phase, $\mathbb{R}$ receives

$$r(i) = \sum_{l=0}^{L} h(i;l)x(i-l) + w_1(i), \tag{5}$$

where $w_1(i)$ is the additive complex white Gaussian noise with mean zero and variance $\sigma_{w_1}^2$. During the second phase, $\mathbb{R}$ amplifies $r(i)$ with a constant factor α and then re-transmits it to $\mathbb{D}$. The signal obtained by $\mathbb{D}$ is

$$\begin{aligned} y(i) =& \alpha \sum_{l=0}^{L} g(i;l) r(i-l) + w_2(i) \\ =& \alpha \sum_{l=0}^{L} g(i;l) \left(\sum_{l=0}^{L} h(i;l)x(i-l) \right) \\ & + \underbrace{\alpha \sum_{l=0}^{L} g(i;l) w_1(i-l) + w_2(i)}_{w(i)}, \end{aligned} \tag{6}$$

where $w(i)$ is the combined noise.

III. Channel Estimation and Data Detection

Let us construct $N \times 1$ vectors $\mathbf{r}$, $\mathbf{y}$, and $N \times N$ matrices $\mathbf{H}$, $\mathbf{G}$ from $g(i;l)$ in the following way:

$$\mathbf{r} = [r(0), r(1), \cdots, r(N-1)]^T, \tag{7}$$
$$\mathbf{y} = [y(0), y(1), \cdots, y(N-1)]^T, \tag{8}$$
$$\mathbf{H}_{i,j} = h(i; i-j), \quad \mathbf{G}_{i,j} = g(i; i-j), \tag{9}$$

for $i, j = 1, 2, \cdots, N$. We can write (5) and (6) as

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{w}_1, \tag{10}$$
$$\mathbf{y} = \alpha \mathbf{G}\mathbf{r} + \mathbf{w}_2 = \alpha \mathbf{G}\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{11}$$

where $\mathbf{w}_i = [w_i(0), w_i(1), \cdots, w_i(N-1)]^T, i = 1, 2$ and $\mathbf{w} = [w(0), w(1), \cdots, w(N-1)]^T$.

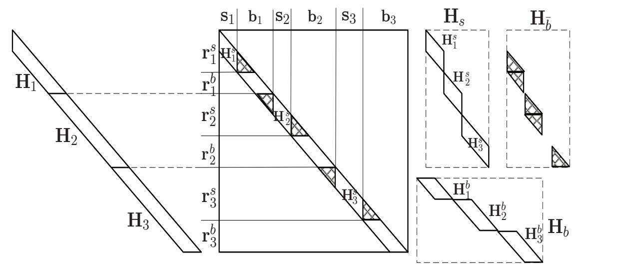

Fig. 2. Partition of the matrix $\mathbf{H}$ into $\mathbf{H}_s$, $\mathbf{H}_b$, and $\mathbf{H}_{\bar{b}}$ that are shown in dashed line on the right side of the figure.

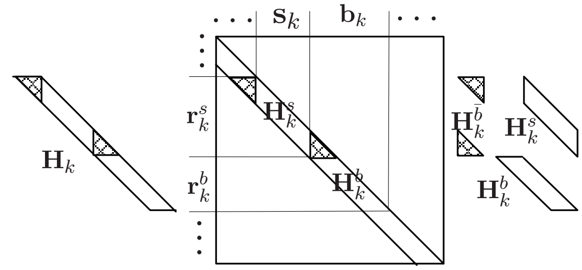

Fig. 3. Partition of the matrix $\mathbf{H}_k$.

A. Channel Partition

Following the channel partition method in [7], we can split the channel matrix $\mathbf{H}$ into three matrices, namely, $\mathbf{H}_s$, $\mathbf{H}_b$, and $\mathbf{H}_{\bar{b}}$, which are shown in Fig. 2. Similarly, the channel $\mathbf{H}_k$, the kth ($1 \leq k \leq P$) part of $\mathbf{H}$ corresponding to the kth sub-block input of $[\mathbf{s}_k, \mathbf{b}_k]$, can also be partitioned into three matrices $\mathbf{H}_k^s$, $\mathbf{H}_k^b$ and $\mathbf{H}_k^{\bar{b}}$, as shown in Fig. 3.

We then have

$$\mathbf{r}_s = \mathbf{H}_s \mathbf{s} + \mathbf{H}_{\bar{b}} \bar{\mathbf{b}} + \mathbf{w}_1^s, \tag{12}$$
$$\mathbf{r}_b = \mathbf{H}_b \mathbf{b} + \mathbf{w}_1^b, \tag{13}$$

where $\mathbf{r}_s = [(\mathbf{r}_1^s)^T, \cdots, (\mathbf{r}_P^s)^T]^T$, $\mathbf{r}_b = [(\mathbf{r}_1^s)^T, \cdots, (\mathbf{r}_P^s)^T]^T$, $\bar{\mathbf{b}}$ contains the first L and the last L entries of $\mathbf{b}_k$ for all $1 \leq k \leq P$, and $\mathbf{w}_1^s$, $\mathbf{w}_1^b$ denote the corresponding noise vectors.

Repeat the same partition process for $\mathbf{G}$ and $\mathbf{G}_k$, that is, split $\mathbf{G}$ into $\mathbf{G}_s$, $\mathbf{G}_{\bar{b}}$ and $\mathbf{G}_b$, while split $\mathbf{G}_k$, the kth component of $\mathbf{G}$, into $\mathbf{G}_k^s$, $\mathbf{G}_k^{\bar{b}}$ and $\mathbf{G}_k^b$. We then obtain

$$\mathbf{y}_s = \alpha \mathbf{G}_s \mathbf{r}_s + \alpha \mathbf{G}_{\bar{b}} \mathbf{r}_{\bar{b}} + \mathbf{w}_2^s, \tag{14}$$
$$\mathbf{y}_b = \alpha \mathbf{G}_b \mathbf{r}_b + \mathbf{w}_2^b, \tag{15}$$

where $\mathbf{y}_s = [(\mathbf{y}_1^s)^T, \cdots, (\mathbf{y}_P^s)^T]^T$, $\mathbf{y}_b = [(\mathbf{y}_1^b)^T, \cdots, (\mathbf{y}_P^b)^T]^T$, $\mathbf{r}_{\bar{b}}$ contains the first L and the last L entries of $\mathbf{r}_k^b$ for all $1 \leq K \leq P$, $\mathbf{w}_2^s$ and $\mathbf{w}_2^b$ denote the corresponding noise vectors.

Combining (13) and (15) produces

$$\mathbf{y}_b = \alpha \mathbf{G}_b \mathbf{H}_b \mathbf{b} + \underbrace{\alpha \mathbf{G}_b \mathbf{w}_1^b + \mathbf{w}_2^b}_{\mathbf{w}_b}. \tag{16}$$

where $\mathbf{w}_b$ is defined as the corresponding item. It can be readily checked that (16) is equivalent to

$$\mathbf{y}_b = \begin{bmatrix} \mathbf{y}_1^b \\ \vdots \\ \mathbf{y}_P^b \end{bmatrix} = \begin{bmatrix} \alpha \mathbf{G}_1^b \mathbf{H}_1^b \mathbf{b}_1 \\ \vdots \\ \alpha \mathbf{G}_P^b \mathbf{H}_P^b \mathbf{b}_P \end{bmatrix} + \mathbf{w}_b. \quad (17)$$

Since $\mathbf{H}_k^b$ is an $(N_{b_k} - L) \times N_{b_k}$ matrix and $\mathbf{G}_k^b$ is an $(N_{b_k} - 2L) \times (N_{b_k} - L)$ matrix, the training length for the kth sub-block should satisfy $N_{b_k} \geq 2L + 1$.

B. Estimation Algorithm

Let us define $\mathbf{\Lambda}_M^{(w_q)} = \text{diag}\{1, e^{jw_q}, \cdots, e^{jw_q(M-1)}\}$. For any $(L+1) \times 1$ vector $\mathbf{a} = [a_0, a_1, \cdots, a_L]^T$, define an $M \times (M+L)$ Toeplitz matrix $\mathbf{T}_{M+L}^{(\mathbf{a})}$ as

$$\mathbf{T}_{M+L}^{(\mathbf{a})} = \underbrace{\begin{bmatrix} a_L & \cdots & a_0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a_L & \cdots & a_0 \end{bmatrix}}_{M+L \text{ columns}}. \quad (18)$$

Based on these definitions, we then have the following two lemmas:

Lemma 1:

$$\mathbf{T}_{M+L}^{(\mathbf{a})} \mathbf{\Lambda}_{M+L}^{(w_q)} = \mathbf{\Lambda}_M^{(w_q)} \mathbf{T}_{M+L}^{(\boldsymbol{\mu}_a)}. \quad (19)$$

where $\boldsymbol{\mu}_a = [a_0 e^{jw_q L}, a_1 e^{jw_q(L-1)}, \cdots, a_L]$.

Proof: Proved from straightforward calculations. ■

Lemma 2: Suppose $\mathbf{a}_i = [a_{i,0}, a_{i,1}, \cdots, a_{i,L}]^T, i = 1, 2$. There is

$$\mathbf{T}_{M+L}^{(\mathbf{a}_1)} \mathbf{T}_{M+2L}^{(\mathbf{a}_2)} = \mathbf{T}_{M+2L}^{(\mathbf{a}_1 * \mathbf{a}_2)}, \quad (20)$$

where $*$ denotes linear convolution.

Proof: Proved from straightforward calculations. ■

According to these definitions and (1), we obtain

$$\mathbf{H} = \sum_{q=0}^{Q} \mathbf{\Lambda}_N^{(w_q)} \mathbf{\Phi}_q, \quad (21)$$

$$\mathbf{H}_k^b = \sum_{q=0}^{Q} e^{jw_q(N_{s_k} + L + \sum_{i=1}^{k-1} N_i)} \mathbf{\Lambda}_{N_{b_k}-L}^{(w_q)} \mathbf{T}_{N_{b_k}}^{(\mathbf{h}_q)}, \quad (22)$$

where $\mathbf{\Phi}_q$ is a lower triangular Toeplitz matrix with the first column $[h_q(0), \cdots, h_q(L), 0, \cdots, 0]^T$.

Similarly, based on (2), we obtain

$$\mathbf{G} = \sum_{q=0}^{Q} \mathbf{\Lambda}_N^{(w_q)} \mathbf{\Omega}_q, \quad (23)$$

$$\mathbf{G}_k^b = \sum_{q=0}^{Q} e^{jw_q(N_{s_k} + 2L + \sum_{i=1}^{k-1} N_i)} \mathbf{\Lambda}_{N_{b_k}-2L}^{(w_q)} \mathbf{T}_{N_{b_k}-L}^{(\mathbf{g}_q)}, \quad (24)$$

where $\mathbf{\Omega}_q$ is a lower triangular Toeplitz matrix with the first column $[g_q(0), \cdots, g_q(L), 0, \cdots, 0]^T$ and $\mathbf{T}_{N_{b_k}-L}^{(\mathbf{g}_q)}$ is a $(N_{b_k} - 2L) \times (N_{b_k} - L)$ Toeplitz matrix as defined in (18).

Combining (22) and (24) gives

$$\mathbf{G}_k^b \mathbf{H}_k^b = \sum_{m=0}^{Q} \sum_{n=0}^{Q} \theta_{m,n,k} \underbrace{\mathbf{\Lambda}_{N_{b_k}-2L}^{(w_m)} \mathbf{T}_{N_{b_k}-L}^{(\mathbf{g}_m)} \mathbf{\Lambda}_{N_{b_k}-L}^{(w_n)} \mathbf{T}_{N_{b_k}}^{(\mathbf{h}_n)}}_{\mathbf{\Xi}_{m,n,k}} \quad (25)$$

where

$$\theta_{m,n,k} = e^{jw_m(N_{s_k} + 2L + \sum_{i=1}^{k-1} N_i) + jw_n(N_{s_k} + L + \sum_{i=1}^{k-1} N_i)}, \quad (26)$$

and $\mathbf{\Xi}_{m,n,k}$ is defined as the corresponding item. Using Lemma 1 and Lemma 2, $\mathbf{\Xi}_{m,n,k}$ can be simplified as

$$\begin{aligned} \mathbf{\Xi}_{m,n,k} &= \mathbf{\Lambda}_{N_{b_k}-2L}^{(w_m)} \mathbf{\Lambda}_{N_{b_k}-2L}^{(w_n)} \mathbf{T}_{N_{b_k}-L}^{(\boldsymbol{\mu}_{g_m})} \mathbf{T}_{N_{b_k}}^{(\mathbf{h}_n)} \\ &= \mathbf{\Lambda}_{N_{b_k}-2L}^{(w_m+w_n)} \mathbf{T}_{N_{b_k}}^{(\boldsymbol{\lambda}_{m,n})}, \end{aligned} \quad (27)$$

where

$$\boldsymbol{\mu}_{g_m} = [g_m(0) e^{jw_n L}, g_m(1) e^{jw_n(L-1)}, \cdots, g_m(L)]^T, \quad (28)$$

$$\boldsymbol{\lambda}_{m,n} = \boldsymbol{\mu}_{g_m} * \mathbf{h}_n. \quad (29)$$

Since $\mathbf{T}_{N_{b_k}}^{(\boldsymbol{\lambda}_{m,n})}$ is a Toeplitz matrix, we obtain

$$\mathbf{G}_k^b \mathbf{H}_k^b \mathbf{b}_k = \sum_{m=0}^{Q} \sum_{n=0}^{Q} \theta_{m,n,k} \mathbf{\Lambda}_{N_{b_k}-2L}^{(w_m+w_n)} \mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)} \boldsymbol{\lambda}_{m,n}, \quad (30)$$

where $\mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)}$ is defined as

$$\mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)} = \begin{bmatrix} b_k(2L) & \cdots & b_k(0) \\ b_k(2L+1) & \cdots & b_k(1) \\ \vdots & \vdots & \vdots \\ b_k(N_{b_k} - 1) & \cdots & b_k(N_{b_k} - 2L - 1) \end{bmatrix}. \quad (31)$$

Unfortunately, it remains challenging to estimate $\boldsymbol{\lambda}_{m,n}$ from (30). A direct way to estimate all $\boldsymbol{\lambda}_{m,n}$ requires N_b to be no less than $2PL + (Q+1)^2(2L+1)$, which is too large and the transmission efficiency will be reduced. To solve this problem, we choose to estimate other type of channel information that requires smaller training length but can guarantee the data detection. Let us introduce two variables $\zeta_{q,k}$ and ϖ_q defined as

$$\varpi_q = w_m + w_n = 2\pi(q - Q)/N, \quad m, n \in [0, Q], \quad (32)$$

$$\zeta_{q,k} = e^{j\varpi_q(N_{s_k} + 2L + \sum_{i=1}^{k-1} N_i)}, \quad k \in [1, P], q \in [0, 2Q], \quad (33)$$

respectively. It can be readily checked that $\theta_{m,n,k} = \zeta_{m+n,k} e^{-jw_n L}$. Then we can combine those items that satisfy $m + n = q$ in (30) and obtain

$$\mathbf{G}_k^b \mathbf{H}_k^b \mathbf{b}_k = \sum_{q=0}^{2Q} \zeta_{q,k} \mathbf{\Lambda}_{N_{b_k}-2L}^{(\varpi_q)} \mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)} \boldsymbol{\eta}_q, \quad (34)$$

where $\boldsymbol{\eta}_q = \sum_{m+n=q} e^{-jw_n L} \boldsymbol{\lambda}_{m,n}$. Define

$$\boldsymbol{\eta} = [\boldsymbol{\eta}_0^T, \boldsymbol{\eta}_1^T, \cdots, \boldsymbol{\eta}_{2Q}^T]^T. \quad (35)$$

Substituting (34) into (17) provides a simpler model

$$\mathbf{y}_b = \alpha\boldsymbol{\Psi}_b\boldsymbol{\eta} + \mathbf{w}_b, \tag{36}$$

where

$$\boldsymbol{\Psi}_b = \begin{bmatrix} \zeta_{0,1}\boldsymbol{\Lambda}^{(\varpi_0)}_{N_{b_1}-2L}\mathbf{B}^{(\mathbf{b}_1)}_{N_{b_1}}, & \cdots, & \zeta_{2Q,1}\boldsymbol{\Lambda}^{(\varpi_{2Q})}_{N_{b_1}-2L}\mathbf{B}^{(\mathbf{b}_1)}_{N_{b_1}} \\ \vdots & \vdots & \vdots \\ \zeta_{0,P}\boldsymbol{\Lambda}^{(\varpi_0)}_{N_{b_P}-2L}\mathbf{B}^{(\mathbf{b}_P)}_{N_{b_P}}, & \cdots, & \zeta_{2Q,P}\boldsymbol{\Lambda}^{(\varpi_{2Q})}_{N_{b_P}-2L}\mathbf{B}^{(\mathbf{b}_P)}_{N_{b_P}} \end{bmatrix}, \tag{37}$$

Instead of estimating the coefficients $\mathbf{h}_q$ and $\mathbf{g}_q$, we could estimate $\boldsymbol{\eta}$ from

$$\hat{\boldsymbol{\eta}} = \frac{1}{\alpha}\left(\boldsymbol{\Psi}_b^H\boldsymbol{\Psi}_b\right)^{-1}\boldsymbol{\Psi}_b^H\mathbf{y}_b, \tag{38}$$

and $\hat{\boldsymbol{\eta}}_q$ is obtained from the corresponding section in $\hat{\boldsymbol{\eta}}$ for each $q \in [0, 2Q]$.

C. Data Detection

Substituting (12) into (14) yields

$$\mathbf{y}_s = \alpha\mathbf{G}_s\mathbf{H}_s\mathbf{s} + \alpha\mathbf{G}_s\mathbf{H}_{\bar{b}}\bar{\mathbf{b}} + \alpha\mathbf{G}_s\mathbf{w}_1^s + \alpha\mathbf{G}_{\bar{b}}\mathbf{r}_{\bar{b}} + \mathbf{w}_2^s. \tag{39}$$

Lemma 3: Among all training choices that lead to identical covariance matrix of the channel estimation error, if the training length N_{b_k} is greater than $4L+1$ and if the training has the first $2L$ and the last $2L$ entries equal to zero, then the interference to the data detection is minimized.

Proof: By setting the first $2L$ and the last $2L$ entries zero, the second item in (39) becomes zero and the fourth item in (39) will only contain the noise item $\alpha\mathbf{w}_1^{r_{\bar{b}}}$, which indicates the minimum interference for data detection. ■

Following Lemma 3, we can simplify (39) as

$$\mathbf{y}_s = \alpha\mathbf{G}_s\mathbf{H}_s\mathbf{s} + \underbrace{\alpha\mathbf{G}_s\mathbf{w}_1^s + \alpha\mathbf{w}_1^{r_{\bar{b}}} + \mathbf{w}_2^s}_{\mathbf{w}_s}, \tag{40}$$

which is equivalent to

$$\mathbf{y}_s = \begin{bmatrix}\mathbf{y}_1^s \\ \vdots \\ \mathbf{y}_P^s\end{bmatrix} = \begin{bmatrix}\alpha\mathbf{G}_1^s\mathbf{H}_1^s\mathbf{s}_1 \\ \vdots \\ \alpha\mathbf{G}_P^s\mathbf{H}_1^s\mathbf{s}_P\end{bmatrix} + \mathbf{w}_s. \tag{41}$$

Define $\mathbf{U}_M^{(\mathbf{h}_q)}$ as a Toeplitz matrix generated by the vector $\mathbf{h}_q$ in the following way:

$$\mathbf{U}_M^{(\mathbf{h}_q)} = \underbrace{\begin{bmatrix} h_q(0), & \cdots, & 0 \\ \vdots & \ddots & \vdots \\ h_q(L), & \ddots, & h_q(0) \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_q(L) \end{bmatrix}}_{M\text{columns}}. \tag{42}$$

Define $\boldsymbol{\mu}_{h_q} = [h_q(0)e^{jw_qL}, h_q(1)e^{jw_q(L-1)}, \cdots, h_q(L)]^T$, we can have

$$\mathbf{U}_M^{(\mathbf{h}_q)}\boldsymbol{\Lambda}_M^{(w_q)} = e^{-jw_qL}\boldsymbol{\Lambda}_{M+L}^{(w_q)}\mathbf{U}_M^{(\boldsymbol{\mu}_{h_q})}, \tag{43}$$

$$\mathbf{U}_{M+L}^{(\mathbf{g}_q)}\mathbf{U}_M^{(\mathbf{h}_q)} = \mathbf{U}_M^{(\mathbf{g}_q*\mathbf{h}_q)} \tag{44}$$

According to (1) and (2), we obtain

$$\mathbf{G}_k^s = \sum_{q=0}^{Q} e^{jw_q\sum_{i=1}^{k-1}N_i}\boldsymbol{\Lambda}^{(w_q)}_{N_{s_k}+2L}\mathbf{U}^{(\mathbf{g}_q)}_{N_{s_k}+L}, \tag{45}$$

$$\mathbf{H}_k^s = \sum_{q=0}^{Q} e^{jw_q\sum_{i=1}^{k-1}N_i}\boldsymbol{\Lambda}^{(w_q)}_{N_{s_k}+L}\mathbf{U}^{(\mathbf{h}_q)}_{N_{s_k}}. \tag{46}$$

Next it can be found that

$$\mathbf{G}_k^s\mathbf{H}_k^s = \sum_{m=0}^{Q}\sum_{n=0}^{Q}\phi_{m,n,k}\boldsymbol{\Lambda}^{(w_m)}_{N_{s_k}+2L}\mathbf{U}^{(\mathbf{g}_m)}_{N_{s_k}+L}\boldsymbol{\Lambda}^{(w_n)}_{N_{s_k}+L}\mathbf{U}^{(\mathbf{h}_n)}_{N_{s_k}} \tag{47}$$

where $\phi_{m,n,k} = e^{j(w_m+w_n)\sum_{i=1}^{k-1}N_i}$. Using (43) and (44), it can be derived that

$$\mathbf{U}^{(\mathbf{g}_m)}_{N_{s_k}+L}\boldsymbol{\Lambda}^{(w_n)}_{N_{s_k}+L}\mathbf{U}^{(\mathbf{h}_n)}_{N_{s_k}} = e^{-jw_nL}\boldsymbol{\Lambda}^{(w_n)}_{N_{s_k}+2L}\mathbf{U}^{(\boldsymbol{\lambda}_{m,n})}_{N_{s_k}}, \tag{48}$$

where $\boldsymbol{\mu}_{g_m}$ and $\boldsymbol{\lambda}_{m,n}$ are defined in (28) and (29) respectively. Substituting (48) into (47), we can obtain

$$\mathbf{G}_k^s\mathbf{H}_k^s = \sum_{q=0}^{2Q} e^{j\varpi_q\sum_{i=1}^{k-1}N_i}\boldsymbol{\Lambda}^{(\varpi_q)}_{N_{s_k}+2L}\mathbf{U}^{(\boldsymbol{\eta}_q)}_{N_{s_k}}. \tag{49}$$

Clearly, given the estimates of $\boldsymbol{\eta}_q$, $\mathbf{G}_k^s\mathbf{H}_k^s$ can be reconstructed from (49). Then, the data detection for $\mathbf{s}_k$ can be performed.

D. Training Sequence Design

The estimation error of $\boldsymbol{\eta}$ can be expressed as

$$\mathbf{e} = \hat{\boldsymbol{\eta}} - \boldsymbol{\eta} = \left(\boldsymbol{\Psi}_b^H\boldsymbol{\Psi}_b\right)^{-1}\boldsymbol{\Psi}_b^H\mathbf{w}_b. \tag{50}$$

The correlation matrix of $\mathbf{w}_b$ is found from (24) as

$$\mathbf{R}_{w_b} = E\left(\mathbf{w}_b\mathbf{w}_b^H\right) = \left(\sigma_{w_2}^2\sum_{q=0}^{Q}\sum_{l=0}^{L}|g_q(l)|^2 + \sigma_{w_1}^2\right)\mathbf{I}_{N_b-2PL}. \tag{51}$$

Thus the mean square error of $\mathbf{e}$ is

$$\sigma_e^2 = \mathrm{tr}\left(E(\mathbf{e}\mathbf{e}^H)\right) = C_e\mathrm{tr}\left(\boldsymbol{\Psi}_b^H\boldsymbol{\Psi}_b\right)^{-1} \tag{52}$$

where $C_e = \left(\sigma_{w_2}^2\sum_{q=0}^{Q}\sum_{l=0}^{L}|g_q(l)|^2 + \sigma_{w_1}^2\right)/\alpha^2$. According to [8], we know that σ_e^2 in (52) is lower bounded as

$$C_e\mathrm{tr}\left(\boldsymbol{\Psi}_b^H\boldsymbol{\Psi}_b\right)^{-1} \geq \sum_m \frac{C_e}{[\boldsymbol{\Psi}_b^H\boldsymbol{\Psi}_b]_{m,m}}, \tag{53}$$

where the equality holds if and only if $(\boldsymbol{\Psi}_b^H\boldsymbol{\Psi}_b)$ is a diagonal matrix. We then need to design the training sequence that can diagonalize $(\boldsymbol{\Psi}_b^H\boldsymbol{\Psi}_b)$.

Based on the definition of $\boldsymbol{\Psi}_b$ (37), the optimal training sequence that can minimizes the σ_e^2 requires the following

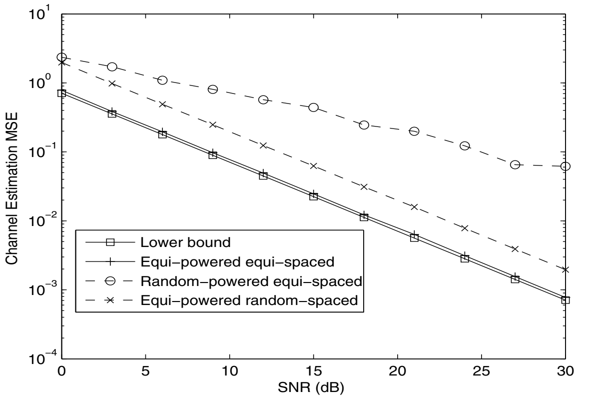

Fig. 4. Channel MSE versus the SNR.

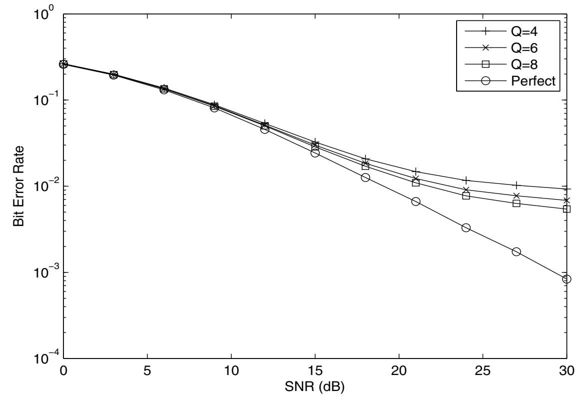

Fig. 5. BER versus the SNR: real channel.

conditions to be satisfied for $\forall q_1 \neq q_2, q_1, q_2 \in [0, 2Q]$:

$$\sum_{k=1}^{P} \left(\mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)}\right)^H \mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)} = \mathcal{P}_b \mathbf{I}_{2L+1}, \tag{54}$$

$$\sum_{k=1}^{P} \left(\mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)}\right)^H \mathbf{\Lambda}_{N_{b_k}-2L}^{(-\varpi_{q_1})} \zeta_{q_1,k}^H \zeta_{q_2,k} \mathbf{\Lambda}_{N_{b_k}-2L}^{(\varpi_{q_2})} \mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)} = \mathbf{0}_{2L+1}, \tag{55}$$

where $\mathcal{P}_b$ is the power allocated to the training sequence and $\mathbf{0}_{2L+1}$ is a $(2L+1) \times (2L+1)$ matrix with all zero entries.

Observing the structure of $\mathbf{B}_{N_{b_k}}^{(\mathbf{b}_k)}$, we know that (54) can be fulfilled if the following conditions are satisfied:

$$\text{(C1):} \quad N_{b_k} = 4L+1, \quad \forall k \in [1, P], \tag{56}$$

$$\text{(C2):} \quad \mathbf{b}_k = \sqrt{\mathcal{P}_b/P}[0, \cdots, 0, 1, 0, \cdots, 0]^T. \tag{57}$$

With conditions (C1) and (C2), we can further simplify (55) as

$$\frac{\mathcal{P}_b}{P} \sum_{k=1}^{P} \mathbf{\Lambda}_{2L+1}^{(-\varpi_{q_1})} \zeta_{q_1,k}^H \zeta_{q_2,k} \mathbf{\Lambda}_{2L+1}^{(\varpi_{q_2})} = \mathbf{0}_{2L+1}, \tag{58}$$

It can be readily checked that the sufficient conditions to achieve (58) is

$$\text{(C3):} \quad N = P(N_{s_k} + 4L + 1), \quad N_{s_k} = N_s/P, \quad \forall k \in [1, P].$$

IV. Simulation Results

We assume that the carrier frequency $f_c = 900$ MHz, one symbol period $T_s = 50\mu s$ and the mobility speed is 90 km/hour. Thus the maximum Doppler shift f_d is 75 Hz and $f_d T_s = 3.75 \times 10^{-3}$. Suppose one block contains 360 symbols, i.e., $N = 360$. Then $Q = 2\lceil N f_d T_s \rceil = 4$. Assume that $h(i;l)$ and $g(i;l)$ has 3 taps, i.e., $L = 2$. The doubly selective channels are generated directly from the CE-BEM (1) and (2). Thus we know that $P \geq (2Q+1) = 9$ and $N_b \geq P(4L+1) = 81$.

First we set the total number of pilots $N_b = 120$ and use three types of training: (i) equi-powered and equi-spaced (our optimal design); (ii) equi-powered but with random length; (iii) equi-spaced but with random power. For performance comparison, the total power of each type of training is the same. For each type of training, we find the MSE of our specially defined channel $\boldsymbol{\eta}$. The estimation MSEs for all three types of training versus SNR are plotted in Fig. 4. The lower bound of σ_e^2 (53) is also displayed for comparison.

We also examine the performance of the suggested estimation and detection methods under real channel situations. Three different number of bases Q are chosen as 4, 6, and 8 respectively, and hence the corresponding number of data symbols N_s is 279, 243, and 207. The BER versus SNR is plotted in Fig. 5. For comparison, the BER curve under perfect channel knowledge at the receiver is also displayed. Clearly, the proposed methods yield effective data detection.

V. Conclusion

In this paper, doubly selective channel estimation was considered for AF-based relay networks. Based on CE-BEM, we designed an efficient method to estimate equivalent channel parameters. The optimal training sequence that can minimize the estimation MSE was also derived.

References

[1] J. Laneman and G. Wornell, “Distributed space time block coded protocols for exploiting cooperative diversity in wireless networks,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[2] L. Tong, B. M. Sadler, and M. Dong, “Pilot-assisted wireless transmissions,” *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 12-25, 2004.

[3] F. Gao, T. Cui, and A. Nallanathan, “On channel estimation and optimal training design for amplify and forward relay network”, *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1907–1916, May 2008.

[4] A. S. Lalos, A. A. Rontogiannis, K. Berberidis, “Frequency Domain Channel Estimation for Cooperative Communication Networks,” *IEEE Trans. Signal Processing*, vol.58, no.6, pp.3400-3405, June 2010.

[5] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[6] G. B. Giannakis and C. Tepedelenlioǧlu, “Basis expnsion models and diversity techniques for blind indentification and equalization of time-varying channels,” *Proc. IEEE*, vol.86, pp. 1969-1986, Nov. 1998.

[7] X. Ma, G. B. Giannakis, and S. Ohno, “Optimal training for block transmissions over doubly selective wireless fading channels,” *IEEE Trans. Signal Processing,* vol. 41, no. 5, pp. 1351–1366, May 2003.

[8] S. Ohno and G. B. Giannakis, “Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels”, *IEEE Trans. Inform. Theory*, vol.50, no. 9, pp. 2138–2145, Sept. 2004.