BEM-Based Estimation for Time-Varying Channels and Training Design in Two-Way Relay Networks

Gongpu Wang[†], Feifei Gao[*], and Chintha Tellambura[†]

[†]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada,
Email: {gongpu, chintha}@ece.ualberta.ca
[*]School of Engineering and Science, Jacobs University, Bremen, Germany
Email: feifeigao@ieee.org

Abstract—In this paper, channel estimation for two-way relay networks (TWRNs) over time-varying channels is investigated. We consider the *amplify-and-forward* (AF) relaying scheme and adopt the complex-exponential basis expansion model (CE-BEM) that represents the time-varying channel by a finite number of parameters. We develop the estimation methods for both the cascaded channels and the individual channels and also apply the total least square (TLS) algorithm to improve the estimation accuracy. Moreover, the training design is discussed and a heuristic criterion is proposed to minimize the condition number of the estimation matrix. The simulation results verify the goodness of the criterion.

I. INTRODUCTION

Bidirectional relay networks [1] have attracted much attention recently due to their enhanced spectral efficiency over unidirectional relay networks [2]. Typically, two source terminals will simultaneously send data to a relay node, at which a "network coding"-like process is applied [3]. The relay then forwards the resultant data to both source terminals. This system is also named as a two-way relay network (TWRN).

In [4] the optimal mapping function at the relay node that minimizes the transmission bit-error rate (BER) was proposed while in [5], the distributed space-time code (STC) was designed for both AF and DF TWRN. Moreover, the optimal beamforming at the multi-antenna relay that maximizes the capacity of AF-based TWRN was developed in [6] and the suboptimal resource allocation in an orthogonal frequency division multiplexing (OFDM) based TWRN was derived in [7]. On the other hand, channel estimation is critically important for those works [4]-[7] that assume perfect channel knowledge at the relay node and/or the source terminals. The first two channel estimation algorithm for TWRN were designed in [8] and [9] for flat and frequency-selective channels, respectively. However, [8] and [9] consider static channels only.

In this paper, we address the problem of estimating the time-varying TWRN channels. With the complex-exponential basis expansion model (CE-BEM), the number of the channel parameters to be estimated is substantially reduced. We propose a new channel estimation approach that is to first estimate the cascaded channels from both source terminals to relay
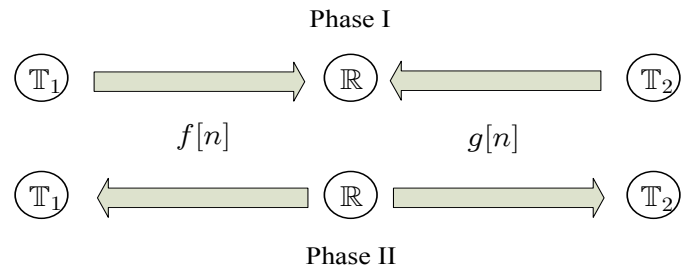
Fig. 1. System model for two-way relay network with time varying channel.

and then to recover the individual channels. The estimation results are further refined by total least square method (TLS). Due to the complexity of the noise structure, we currently design two training sequences from minimizing the condition number of the estimation matrix. Nonetheless, these training sequences are numerically shown to give superior performance than random training.

II. SYSTEM MODEL

Consider a TWRN with two source nodes $\mathbb{T}_1$ and $\mathbb{T}_2$, and one relay node $\mathbb{R}$, as shown in Fig. 1. Each node has only one antenna that can transmit or receive in half-duplex. The baseband channels from $\mathbb{T}_1$ to $\mathbb{R}$, and from $\mathbb{T}_2$ to $\mathbb{R}$ are denoted by $f[n]$ and $g[n]$ respectively, where n is the discrete time index. Due to reciprocity, the channels from $\mathbb{R}$ to $\mathbb{T}_1$, and from $\mathbb{R}$ to $\mathbb{T}_2$ are $f[n]$ and $g[n]$, too.

The channel statistics depend on the movement of the three nodes. For example, if $\mathbb{T}_1$ and $\mathbb{T}_2$ are static but $\mathbb{R}$ is moving, then we can assume [11]

$$E\{f[n+m]f^*[n]\} = J_0(2\pi f_{d1} m T_s), \qquad (1)$$

$$E\{g[n+m]g^*[n]\} = J_0(2\pi f_{d2} m T_s), \qquad (2)$$

where J_0 is the zero-th order Bessel function of the first kind, f_{d1} is the maximum Doppler shift associated with $f[n]$, f_{d2} is the maximum Doppler shift associated with $g[n]$, and T_s is the symbol sampling time. If $\mathbb{T}_1$ is static but $\mathbb{T}_2$ and $\mathbb{R}$ are moving, then [11]

$$E\{f[n+m]f^*[n]\} = J_0(2\pi f_{d1} m T_s), \qquad (3)$$

$$E\{g[n+m]g^*[n]\} = J_0(2\pi f_{d1} m T_s)J_0(2\pi f_{d2} m T_s). \qquad (4)$$

Assume that the peak powers for $\mathbb{T}_1$ and $\mathbb{T}_2$ are P_s. During Phase I, $\mathbb{T}_1$ and $\mathbb{T}_2$ simultaneously transmit M symbols $s_1[n]$ and $s_2[n]$, respectively. Relay $\mathbb{R}$ then receives

$$r[n] = f[n]s_1[n] + g[n]s_2[n] + w_r[n], \quad n = 0, \ldots, M-1, \tag{5}$$

where $w_r[n]$ is the circularly symmetric complex Gaussian (CSCG) noise with the variance σ_w^2. In Phase II, $\mathbb{R}$ scales $r[n]$ by a factor α and then forwards it to both $\mathbb{T}_i$. Due to symmetry, we only presents the estimation at $\mathbb{T}_1$. The received signal is

$$y[n + M + \Delta] = f[n + M + \Delta]\alpha r[n] + w_1[n], \tag{6}$$
$$n = 0, \ldots, M-1,$$

where Δ denotes the process delay at the relay node, and $w_1[n]$ is the CSCG noise with variance σ_w^2.

From [12], the BEM of $f[n]$ and $g[n]$ can be expressed as

$$f[n] = \sum_{q=0}^{Q} f_q e^{jw_q n}, \quad 0 \le n \le N-1 \tag{7}$$

$$g[n] = \sum_{q=0}^{Q} g_q e^{jw_q n}, \quad 0 \le n \le N-1 \tag{8}$$

where $N \triangleq 2M + \Delta$, Q is a properly selected integer (usually greater than 2), and $e^{jw_q n}$'s are basis. Following [12], we choose $w_q \triangleq 2\pi(q - Q/2)/N$.

Instead of estimating all $f[n]$ and $g[n]$, the task now becomes estimating $4Q + 2$ coefficients f_q and g_q.

III. PROPOSED CHANNEL ESTIMATION

A. System Model Reformulation

We first rewrite (6) as

$$y[n + M + \Delta]$$
$$= \alpha \left(\sum_{q=0}^{Q} f_q e^{jw_q(n+M+\Delta)} \right) \left(\sum_{q=0}^{Q} f_q e^{jw_q n} \right) s_1[n]$$
$$+ \alpha \left(\sum_{q=0}^{Q} f_q e^{jw_q(n+M+\Delta)} \right) \left(\sum_{q=0}^{Q} g_q e^{jw_q n} \right) s_2[n]$$
$$+ \alpha f[n + M + \Delta] w_r[n] + w_1[n + M + \Delta]. \tag{9}$$

In order to put (9) in matrix-form, we define

$$\mathbf{y} = [y[M+\Delta], y[M+\Delta+1], ..., y[2M+\Delta-1]]^T,$$
$$\mathbf{w}_r = [w_r[0], w_r[1], ..., w_r[M-1]]^T,$$
$$\mathbf{w}_1 = [w_1[M+\Delta], w_1[M+\Delta+1], ..., w_1[2M+\Delta-1]]^T,$$
$$\mathbf{f} = [f[M+\Delta], f[M+\Delta+1], ..., f[2M+\Delta-1]]^T,$$
$$\mathbf{s}_i = [s_i[0], s_i[1], ..., s_i[M-1]]^T, \quad i = 1, 2$$
$$\mathbf{S}_i = \text{diag}\{\mathbf{s}_i\}.$$

It can be verified that

$$\mathbf{y} = \alpha \mathbf{S}_1 \mathbf{A} \mathbf{h}_1 + \alpha \mathbf{S}_2 \mathbf{A} \mathbf{h}_2 + \underbrace{\alpha \mathbf{f} \odot \mathbf{w}_r + \mathbf{w}_1}_{\mathbf{w}}, \tag{10}$$

where $\odot$ denotes the Hadamard product, $\mathbf{A}$ is a $M \times (Q+1)^2$ Vandermonde matrix with the $(i(Q+1)+k)$th column

$$\mathbf{A}[:, i(Q+1)+k] = e^{jw_i(M+\Delta)} \begin{bmatrix} 1 \\ e^{j(w_i+w_k)} \\ \vdots \\ e^{j(M-1)(w_i+w_k)} \end{bmatrix}, \tag{11}$$
$$i, k = 0, \ldots, Q,$$

$\mathbf{h}_1$ is the $(Q+1)^2 \times 1$ vector with the $(i(Q+1)+k)$th entry $f_i f_k$, $\mathbf{h}_2$ is the $(Q+1)^2 \times 1$ vector with the $(i(Q+1)+k)$th entry $f_i g_k$, and $\mathbf{w}$ denotes the equivalent combined noise. Bearing in mind that $J_0(0) = 1$, the noise covariance matrix can be computed as

$$\mathbf{R} = E(\mathbf{w}\mathbf{w}^H) = (|\alpha|^2 + 1)\sigma_w^2 \mathbf{I}_{M \times M}. \tag{12}$$

Lemma 1: Matrix $\mathbf{A}$ defined in (11) has rank $2Q + 1$.

Proof: From (11), it is easily known that the $(i(Q+1)+k)$th column of $\mathbf{A}$ is the same with the $(k(Q+1)+i)$th column of $\mathbf{A}$ except a constant factors, i.e., they are linearly dependent. In fact, since $\mathbf{A}$ is a Vandermonde matrix, the rank of $\mathbf{A}$ is equal to the number of different $(w_i + w_k)$ values. From the definition of w_q, we know that $(w_i + w_k)$ varies from $2w_0$ to $2w_Q$, so there are $2Q+1$ equally spaced $(w_i + w_k)$, with the space $2\pi q/N$. Proved. ∎

Combining those linearly dependent columns in $\mathbf{A}$, we can rewrite (10) as

$$\mathbf{y} = \alpha \mathbf{S}_1 \mathbf{A}_0 \mathbf{x}_1 + \alpha \mathbf{S}_2 \mathbf{A}_0 \mathbf{x}_2 + \mathbf{w}, \tag{13}$$

where $\mathbf{x}_i$ is a $(2Q+1) \times 1$ vector with entries

$$\mathbf{x}_1(m) = \sum_{i+k=m} e^{jw_i(M+\Delta)} f_i f_k, \quad m = 0, \ldots, 2Q \tag{14}$$

$$\mathbf{x}_2(m) = \sum_{i+k=m} e^{jw_i(M+\Delta)} f_i g_k, \tag{15}$$

and $\mathbf{A}_0$ is an $M \times (2Q+1)$ Vandermonde matrix shown in (16) on the top of the next page.

Remark 1: Note that $\mathbf{x}_i$, $i = 1, 2$ represents the equivalent channel vector that contains the mixed channel information. It is sufficient to estimate $\mathbf{x}_i$ for data detection. Nonetheless, another important problem for channel estimation in TWRN is to estimate the individual channel knowledge [9] so that some optimization actions can be performed.

B. Least Square Estimation

Equation (13) can be rewritten as

$$\mathbf{y} = \underbrace{[\alpha \mathbf{S}_1 \mathbf{A}_0, \alpha \mathbf{S}_2 \mathbf{A}_0]}_{\mathbf{C}} \underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}}_{\mathbf{x}} + \mathbf{w}, \tag{17}$$

where $\mathbf{C}$ and $\mathbf{x}$ are defined as the corresponding items. When $M \ge (4Q+2)$, the LS estimation of $\mathbf{x}$ can be immediately obtained as

$$\check{\mathbf{x}} = (\mathbf{C}^H \mathbf{C})^{-1} \mathbf{C}^H \mathbf{y}. \tag{18}$$

$$\mathbf{A}_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ e^{j2w_0} & e^{j(w_0+w_1)} & \cdots & e^{j(w_{Q-1}+w_Q)} & e^{j2w_Q} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ e^{j2w_0(M-1)} & e^{j(w_0+w_1)(M-1)} & \cdots & e^{j(w_{Q-1}+w_Q)(M-1)} & e^{j2w_Q(M-1)} \end{bmatrix} \tag{16}$$

Due to the special structure of $\mathbf{A}_0$ and $\mathbf{C}$, the condition number of $\mathbf{C}$ can be large and results in a noise enhancement the inverse of $\mathbf{C}^H\mathbf{C}$ is taken. This phenomenon becomes even severe when Q gets large. To overcome this, we apply the total least square method (TLS) [13] to refine our estimation results.

By assuming that $\mathbf{y}$ and $\mathbf{C}$ contain disturbance $\mathbf{e}$ and $\mathbf{E}$, respectively, we obtain:

$$(\mathbf{C} + \mathbf{E})\mathbf{x} = \mathbf{y} + \mathbf{e}, \tag{19}$$

or equivalently

$$\underbrace{[-\mathbf{y}, \mathbf{C}]}_{\mathbf{J}}\mathbf{z} + \underbrace{[-\mathbf{e}, \mathbf{E}]}_{\mathbf{D}}\mathbf{z} = \mathbf{0}_{M\times 1}, \tag{20}$$

where $\mathbf{J}$ and $\mathbf{D}$ are defined as the corresponding items, and $\mathbf{z} = [1, \mathbf{x}^T]^T$. The TLS finds $\mathbf{x}$, $\mathbf{E}$, and $\mathbf{e}$ from the following optimization:

$$\min_{\mathbf{D},\mathbf{x}} ||\mathbf{Dz}||_F^2, \tag{21}$$

$$s.t. \quad (\mathbf{y} + \mathbf{e}) \in \text{Range}(\mathbf{C} + \mathbf{E}), \tag{22}$$

where $|| \cdot ||_F^2$ denotes the Frobenius norm of a matrix. From (20), we can skip the estimation of $\mathbf{E}$ and $\mathbf{e}$ and solve the following optimization:

$$\min_{\mathbf{z}} ||\mathbf{Jz}||_F^2, \tag{23}$$

The solution to (23) is easily proved to be the eigenvector of $\mathbf{J}^H\mathbf{J}$ that corresponds the smallest eigenvalue. Since $\mathbf{C}$ is a full column rank, matrix $\mathbf{J}^H\mathbf{J}$ cannot have more than one zero eigenvalue. Bearing in mind that the first entry of $\mathbf{z}$ must be 1, the solution to $\mathbf{z}$ is unique.

Thus we choose $\mathbf{J} = [-\mathbf{y}, \mathbf{C}]$ whose singular value decomposition (SVD) is

$$\mathbf{J} = \mathbf{U} \begin{bmatrix} \sigma_1(\mathbf{J}) & & & 0 \\ & \sigma_2(\mathbf{J}) & & \\ & & \ddots & \\ 0 & & & \sigma_{4Q+3}(\mathbf{J}) \end{bmatrix} \mathbf{V}^H, \tag{24}$$

where $\mathbf{U}$, $\mathbf{V}$ are two unitary matrices and $\sigma_i(\mathbf{J})$'s are the singular values of $\mathbf{J}$ that are arranged in decreasing order. The estimation of $\mathbf{x}$ can be expressed as

$$\hat{\mathbf{x}} = \frac{1}{\mathbf{V}(1, 4Q+3)} \begin{bmatrix} \mathbf{V}(2, 4Q+3) \\ \vdots \\ \mathbf{V}(4Q+3, 4Q+3) \end{bmatrix}, \tag{25}$$

where $\mathbf{V}(m_1, m_2)$ is the (m_1, m_2)th entry of $\mathbf{V}$.

C. Recovering the Individual Channel Knowledge.

The estimate of $\mathbf{x}_1$ and $\mathbf{x}_2$, denoted by $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ can be extracted from $\hat{\mathbf{x}}$ as the first $2Q+1$ and the last $2Q+1$ elements, respectively. The next task is to recover the individual channel coefficient $\hat{f}_q$ and $\hat{g}_q$ from the mixed channel estimates.

Observing the structure in (14) and (15), a straightforward way is to estimate f_q sequentially. We first estimate f_0 from

$$\hat{f}_0 = \pm \sqrt{\hat{\mathbf{x}}_1(0)/e^{jw_0(M+\Delta)}}. \tag{26}$$

Choosing any of the positive or the negative sign in (26), we can compute $\hat{f}_1$ as

$$\hat{f}_1 = \frac{\hat{\mathbf{x}}_1(1)}{\hat{f}_0 e^{j\omega_0(M+\Delta)} + \hat{f}_0 e^{j\omega_1(M+\Delta)}}. \tag{27}$$

We then sequentially compute $\hat{f}_q$ from $\hat{\mathbf{x}}_1(q)$ with the previous estimate of $\hat{f}_0, \ldots, \hat{f}_{q-1}$. The detailed steps are straightforward and are omitted here.

Similarly, we can obtain $\hat{g}_q$ sequentially with the estimated $\hat{f}_q$ and $\hat{\mathbf{x}}_1$. For example

$$\hat{g}_0 = \frac{\hat{\mathbf{x}}_2[0]}{e^{jw_0(M+\Delta)}\hat{f}_0}, \tag{28}$$

$$\hat{g}_1 = \frac{\hat{\mathbf{x}}_2[1] - e^{jw_1(M+\Delta)}\hat{f}_1\hat{g}_0}{e^{jw_0(M+\Delta)}\hat{f}_0}. \tag{29}$$

Note that there exists simultaneous sign ambiguity (SSA) in the estimated results, i.e., either $\{\hat{f}_q, \hat{g}_q\}$ or $\{-\hat{f}_q, -\hat{g}_q\}$ are found. Nonetheless, SSA does not affect the data detection and the system optimization as pointed out in [9].

IV. TRAINING SEQUENCE DESIGN

In this section, we will design the training sequence that can achieve better estimation results. Since the LS and TLS methods are used, one natural choice is to minimize the condition number of $\mathbf{C}$ that represents the stability of our estimation method and substantially influence our estimation performance.

Consider the optimization problem

$$\arg\min_{\mathbf{s}_1, \mathbf{s}_2} \quad \kappa\big(\mathbf{C}(\mathbf{s}_1, \mathbf{s}_2)\big) = \frac{\sigma_1(\mathbf{C})}{\sigma_{4Q+2}(\mathbf{C})}, \tag{30}$$

where $\kappa(\cdot)$ denotes the condition number. When $s_i[n]$ are chosen from finite constellations, an exhaustive search to find the optimal training sequence is a possible option.

However, since the computational complexity of searching is prohibitively high when M gets larger, we propose two heuristic training sequences here.

Type-1 Training: The training symbol from $\mathbb{T}_1$ and $\mathbb{T}_2$ appear alternatively with maximum power P_s and arbitrary

phase. The idea behind is to avoid the interference between two training sequences, so a better estimation is expected.

Type-2 Training: We choose $|s_1[n]| = P_s$ with arbitrary phase and set $s_2[n] = (-1)^n s_1[n]$.

Denote $\mathbf{C}$ from type-1 training as $\mathbf{C}_1$, and that from type-2 training as $\mathbf{C}_2$. We provide the following theorem.

Theorem _1:_ The condition numbers of $\mathbf{C}_1$ and $\mathbf{C}_2$ are the same.

Proof: Instead of proving the theorem directly, we prove

$$\kappa(\mathbf{C}_1^H\mathbf{C}_1) = \kappa(\mathbf{C}_2^H\mathbf{C}_2). \tag{31}$$

It can be readily checked that

$$\kappa(\mathbf{C}^H\mathbf{C}) = \kappa\left(\mathbf{A}_D^H \begin{bmatrix} \mathbf{S}_1^H\mathbf{S}_1, & \mathbf{S}_1^H\mathbf{S}_2 \\ \mathbf{S}_2^H\mathbf{S}_1, & \mathbf{S}_2^H\mathbf{S}_2 \end{bmatrix} \mathbf{A}_D\right), \tag{32}$$

where

$$\mathbf{A}_D = \begin{bmatrix} \mathbf{A}_0, & \mathbf{0}_{M\times(2Q+1)} \\ \mathbf{0}_{M\times(2Q+1)}, & \mathbf{A}_0 \end{bmatrix}. \tag{33}$$

Suppose the SVD of $\mathbf{A}_0$ is

$$\mathbf{A}_0 = \mathbf{U}_1\mathbf{\Xi}_1\mathbf{V}_1^H, \tag{34}$$

where $\mathbf{U}_1$ is the $M \times M$ unitary matrix, $\mathbf{V}_1$ is the $(2Q+1) \times (2Q+1)$ matrix, and $\mathbf{\Xi}_1$ is the $M \times (2Q+1)$ diagonal matrix.

Let $\mathbf{U}_0$ be the first $2Q+1$ columns of $\mathbf{U}_1$, and $\mathbf{\Xi}_0$ be the first $2Q+1$ rows of $\mathbf{\Xi}_1$. Then $\mathbf{A}_0$ can be rewritten as

$$\mathbf{A}_0 = \mathbf{U}_0\mathbf{\Xi}_0\mathbf{V}_1^H. \tag{35}$$

Since for any unitary matrix $\mathbf{V}_1$, there is

$$\kappa(\mathbf{V}_1^H\mathbf{X}\mathbf{V}_1) = \kappa(\mathbf{X}), \tag{36}$$

we can rewrite (32) as

$$\kappa(\mathbf{C}^H\mathbf{C}) = \kappa\left(\mathbf{\Xi}^H \begin{bmatrix} \mathbf{U}_0^H\mathbf{S}_1^H\mathbf{S}_1\mathbf{U}_0, & \mathbf{U}_0^H\mathbf{S}_1^H\mathbf{S}_2\mathbf{U}_0 \\ \mathbf{U}_0^H\mathbf{S}_2^H\mathbf{S}_1\mathbf{U}_0, & \mathbf{U}_0^H\mathbf{S}_2^H\mathbf{S}_2\mathbf{U}_0 \end{bmatrix} \mathbf{\Xi}\right), \tag{37}$$

where

$$\mathbf{\Xi} = \begin{bmatrix} \mathbf{\Xi}_0, & \mathbf{0} \\ \mathbf{0}, & \mathbf{\Xi}_0 \end{bmatrix},$$

and $\mathbf{0}$ represents the all-zero matrix with suitable dimensions. To further simplify (37), we take the following steps. Decompose $\mathbf{U}_0$ as

$$\mathbf{U}_0 = [\mathbf{u}_1 \ \mathbf{u}_2 \ \ldots \ \mathbf{u}_M]^H,$$

where $\mathbf{u}_i^H$ is ith row of $\mathbf{U}_0$, and define

$$\bar{\mathbf{U}}_{01} = \sum_{i=1}^{M/2} \mathbf{u}_{2k-1}\mathbf{u}_{2k-1}^H, \qquad \bar{\mathbf{U}}_{02} = \sum_{i=1}^{M/2} \mathbf{u}_{2k}\mathbf{u}_{2k}^H.$$

For proposed type-1 and type-2 training, it can be verified that

$$\kappa(\mathbf{C}_1^H\mathbf{C}_1) = \kappa\left(\mathbf{\Xi}^H \begin{bmatrix} \bar{\mathbf{U}}_{01}, & \mathbf{0} \\ \mathbf{0}, & \bar{\mathbf{U}}_{02} \end{bmatrix} \mathbf{\Xi}\right), \tag{38}$$

$$\kappa(\mathbf{C}_2^H\mathbf{C}_2) = \kappa\left(\mathbf{\Xi}^H \begin{bmatrix} \mathbf{I}, & \bar{\mathbf{U}}_0 \\ \bar{\mathbf{U}}_0, & \mathbf{I} \end{bmatrix} \mathbf{\Xi}\right), \tag{39}$$
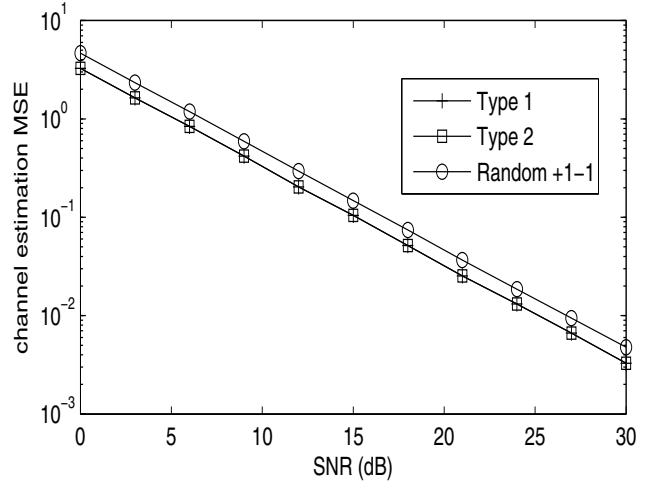

Fig. 2. Channel estimation MSE of $\mathbf{x}$

where $\bar{\mathbf{U}}_0 = \bar{\mathbf{U}}_{01} - \bar{\mathbf{U}}_{02}$.

From

$$\bar{\mathbf{U}}_{01} + \bar{\mathbf{U}}_{02} = \mathbf{U}_0^H\mathbf{U}_0 = \mathbf{I}, \tag{40}$$

we obtain the following equality

$$\mathbf{M}^H \begin{bmatrix} \bar{\mathbf{U}}_{01}, & \mathbf{0} \\ \mathbf{0}, & \bar{\mathbf{U}}_{02} \end{bmatrix} \mathbf{M} = \begin{bmatrix} \mathbf{I}, & \bar{\mathbf{U}}_0 \\ \bar{\mathbf{U}}_0, & \mathbf{I} \end{bmatrix}, \tag{41}$$

where $\mathbf{M} = \begin{bmatrix} \mathbf{I}, & \mathbf{I} \\ -\mathbf{I}, & \mathbf{I} \end{bmatrix}$. Since $\mathbf{M}$ is a scale of unitary matrix, we know

$$\kappa\left(\begin{bmatrix} \bar{\mathbf{U}}_{01}, & \mathbf{0} \\ \mathbf{0}, & \bar{\mathbf{U}}_{02} \end{bmatrix}\right) = \kappa\left(\begin{bmatrix} \mathbf{I}, & \bar{\mathbf{U}}_0 \\ \bar{\mathbf{U}}_0, & \mathbf{I} \end{bmatrix}\right). \tag{42}$$

From (42), it can also be proved that

$$\kappa\left(\mathbf{\Xi}^H \begin{bmatrix} \bar{\mathbf{U}}_{01}, & \mathbf{0} \\ \mathbf{0}, & \bar{\mathbf{U}}_{02} \end{bmatrix} \mathbf{\Xi}\right) = \kappa\left(\mathbf{\Xi}^H \begin{bmatrix} \mathbf{I}, & \bar{\mathbf{U}}_0 \\ \bar{\mathbf{U}}_0, & \mathbf{I} \end{bmatrix} \mathbf{\Xi}\right), \tag{43}$$

Therefore, combining (31), (43), (38) and (39), we can establish $\kappa(\mathbf{C}_1) = \kappa(\mathbf{C}_2)$. ∎

V. SIMULATION RESULTS

Computer simulations are provided to verify the effectiveness of the proposed algorithms as well as the proposed training sequences. We select $\Delta = 2$, $Q = 2$, and the time-varying channels are generated by using two basis functions, as was done in recent papers on time-varying channel estimation. For all simulations, 10000 Monte-Carlo runs are performed. We compare three different types of the training sequences, i.e., the proposed type-1 training, the proposed type-2 training, and the random training sequence with symbol values $\pm\sqrt{P_s}$. The type-1 training has peak power $2P_s$ so that the total power for all three different training are kept the same.
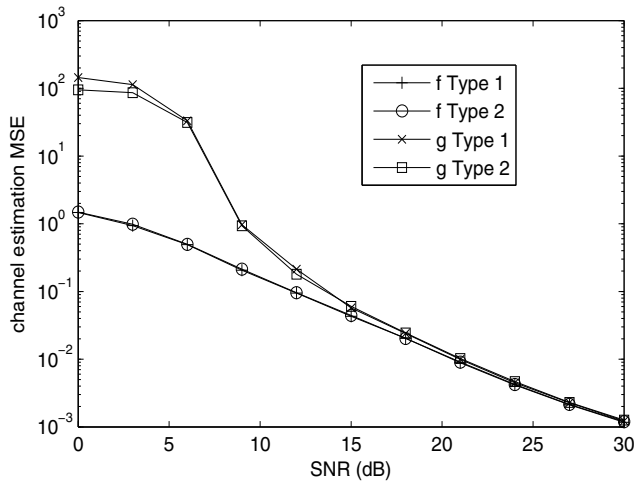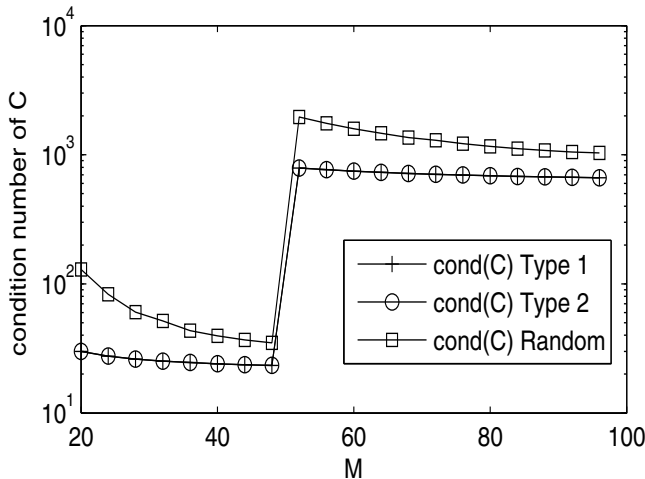
Fig. 3. Channel estimation MSE of **f** and **g**

Fig. 4.   Condition number of A_0

A. Estimation MSE

In the first example, the normalized estimation mean square error (MSE) of $\mathbf{x}$ is chosen as the figure of merit. The phase length is taken as $M = 48$. The MSEs of $\mathbf{x}$ versus SNR for three different training sequences are shown in Fig. 2. It is seen that the proposed two training designs yield the same MSE because they both have the same training power and the conditional number in $\mathbf{C}$. Moreover, the proposed training outperforms the random training sequence by 2 dB. Nonetheless, the proposed type-2 training is preferable than type-1 since the former has a low peak power and hence is more suitable for amplification by a linear power amplifier.

Next we examine the estimation accuracy of the basis coefficient f_q and g_q. The corresponding MSE curves versus SNR are given in Fig. 3 for the proposed two different training sequences. Once again, both training sequences yield the very similar estimation accuracy. We further observe a better performance in estimating the coefficient f_q. This is a direct

result that g_q is obtained from the estimates of f_q, where error propagation occurs.

B. Comparison of the Condition number

Next we vary the value M from 20 to 98 and compare the condition number of $\mathbf{C}$ from the three different training sequences. The corresponding curves are shown in Fig. 4. Clearly, the condition numbers of $\mathbf{C}$ of the proposed training sequences are much lower than that of the random training, indicating a better performance.

VI. CONCLUSION

In this paper, we studied channel estimation for time-varying TWRN channels. We built the BEM channel model and proposed the estimation algorithms. Both the cascaded channels and the individual channel parameters were found. The conditional number was proposed to design the training sequences and two different training designs were developed. Interestingly, if the overall training powers are the same, the proposed two training sequences yield the same estimation accuracy. Nonetheless, more sophisticated training design can be found in the journal version of this paper [10].

REFERENCES

[1] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Proc. Annual Conference on Signals, Systems, and Computers*, Pacific Grove, USA, Oct. 2005, pp. 1066–1071.
[2] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
[3] Li, S.-Y.R. and Yeung, R.W. and Ning Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49 , no. 2, pp.371–381, Feb. 2003.
[4] T. Cui, T. Ho, and J. Kliewer, "Memoryless relay strategies for two-way relay channels," in *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3132–3143, Oc. 2009.
[5] T. Cui, F. Gao, T. Ho, and A. Nallanathan, "Distributed space-time coding for two-way wireless relay networks", *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 658-671, Feb. 2009.
[6] R. Zhang, Y.-C. Liang, C. C. Chai, and S. G. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Select. Areas Commun.*, vol. 27, no. 5, pp. 699–712, June 2009.
[7] C. K. Ho, R. Zhang and Y.-C. Liang, "Two-way relaying over OFDM: optimized tone permutation and power allocation," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 3908–3912.
[8] F. Gao, R. Zhang, and Y.-C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.
[9] ——, "Channel estimation for OFDM modulated two-way relay networks," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4443–4455, Nov. 2009.
[10] G. Wang, F. Gao, W. Chen and C. Tellambura, " Channel estimation and training design for time varying channel in two-way relay networks," to be submitted.
[11] W. C. Jakes, *Microwave Mobile Communications*.   New York: Wiley, 1974.
[12] X. Ma and G. B. Giannakis, "Maximum-diversity transmissions over doubly-selective wireless channels," *IEEE Trans. Inf. Theory*, vol.49, pp. 1832-1840, Jul. 2003.
[13] G. H. Golub and V. Loan, "An analysis of the total least squares problem," SZAM J. Numer. Anal., vol. 17, no. 6, pp. 883-893, Dec. 1980.
[14] R.A. Horn, C. R. Johnson, "Topics in Matrix Analysis." Cambridge University Press, 1991.