Channel Estimation for Two-Way Relay Networks under Time-Selective Environment

Gongpu Wang, Feifei Gao, Wen Chen, and Chintha Tellambura

Abstract—In this paper, we consider the problem of channel estimation for two-way relay networks (TWRN) under time-selective environment. We first parameterize the time-varying channels by the basis expansion model (BEM) and then propose a novel pilot symbol aided modulation (PSAM) for TWRN. A linear approach to estimate the cascaded channels is designed and the optimal training sequences are derived based on minimizing the mean-square error (MSE) criterion. Moreover, we develop an algorithm to recover the individual channel knowledge with which both the channel estimation accuracy and the system performance can be improved. Various simulations are provided to corroborate the proposed studies.

I. INTRODUCTION

Recently, two-way relay network (TWRN) has attracted lots of attention due to its enhanced spectral efficiency compared to its one-way counterpart [1]. Like any other wireless communication systems, a TWRN performs better with better channel estimates.

Assuming the block-fading channels, [2] and [3] addressed the problems of channel estimation for TWRN under flat-fading and frequency-selective fading environment, respectively. However, moving relays or moving mobiles are more typical scenarios for TWRN, since the relative motion between any two nodes doubles the Doppler effect [4]. This places the demand of estimating the time-varying channels that is usually achieved by with the aid of the periodical training scheme, also known as pilot symbol aided modulation (PSAM) [5].

The time-varying Rayleigh fading channels are usually described by Jakes model [6]. To reduce the number of parameters, two approximations of Jakes model have been proposed in the literatures: autoregressive (AR) process or basis expansion model (BEM) [7]. AR model describes channel variation through a symbol-by-symbol updating manner. On the other hand, BEM expresses the time-varying taps into superposition of the time-varying basis functions weighted by time-invariant coefficients.

In this paper, we will focus on BEM and address the problems of channel estimation as well as training design for TWRN. We propose a novel PSAM that is particularly suitable for TWRN under time-selective environment. We then developed the channel estimation prototype such that a linear estimator can be applied for the cascaded channels. The optimal training sequences are then derived based on minimizing the channel mean square error (MSE) criterion. Moreover, we propose to recover the individual channel knowledge and then

re-construct the cascaded channels. Finally, simulations are provided to corroborate the proposed studies.

II. PROBLEM FORMULATION

Consider a TWRN with two source nodes $\mathbb{T}_1$ and $\mathbb{T}_2$ and one relay node $\mathbb{R}$. Each node has only one antenna that cannot transmit and receive simultaneously. The baseband channel from $\mathbb{T}_i$, $i = 1, 2$ to $\mathbb{R}$ is assumed as time-selective flat-fading and is denoted by $h_i(n)$, where n is the discrete time index. Moreover, the channels are modeled as wide-sense stationary (WSS) zero mean complex Gaussian (ZMCG) random processes with variances $\sigma_{h_i}^2$. Due to reciprocity, the channel from $\mathbb{R}$ to $\mathbb{T}_i$ is also $h_i(n)$.

A. Time-Varying Channel in Relay Network

The channel statistics in relay network depend on the mobility of the three nodes. Denote f_{d1}, f_{d2} and f_{dr} as the maximum Doppler shifts due to the motion of $\mathbb{T}_1$, $\mathbb{T}_2$ and $\mathbb{R}$, respectively. The discrete autocorrelation functions of $h_i(n)$'s can be computed as [8]

$$R_{h_i}(m) = E\{h_i(n+m)h_i^*(n)\}$$
$$= \sigma_{h_i}^2 J_0(2\pi f_{di} m T_s) J_0(2\pi f_{dr} m T_s), i = 1, 2, \quad (1)$$

where $J_0(\cdot)$ is the zero-th order Bessel function of the first kind and T_s is the symbol sampling time. The power spectrum of $h_i(n)$ then has the bandwidth $f_i = f_{di} + f_{dr}$ [8].

With the knowledge of the spectrum bandwidth, the parsimonious finite-parameter BEM can be applied to approximate the time-varying channels, and $h_i(n)$'s during any NT_s interval are modeled as

$$h_1(n) = \sum_{q=0}^{Q_1} \lambda_q w_1(q), \qquad h_2(n) = \sum_{q=0}^{Q_2} \mu_q w_2(q), \quad (2)$$

where $0 \leq n \leq N-1$, λ_q's and μ_q's are the BEM coefficients that remain invariant within one interval of NT_s but will change in the next interval, while $w_i(q)$'s are the bases that capture the time variation and will remain the same for any time interval. The most widely adopted bases are the Fourier bases [9], and the channels can be explicitly modeled as

$$h_1(n) = \sum_{q=0}^{Q_1} \lambda_q e^{j2\pi(q-Q_1/2)n/N}, \quad 0 \leq n \leq N-1, \quad (3a)$$

$$h_2(n) = \sum_{q=0}^{Q_2} \mu_q e^{j2\pi(q-Q_2/2)n/N}, \quad 0 \leq n \leq N-1. \quad (3b)$$
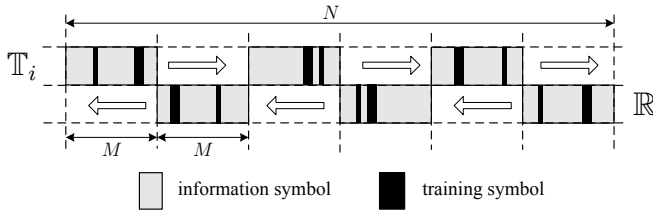
Fig. 1. Structure of one transmission frame.

The number of bases Q_i should be at least $2\lceil f_i N T_s \rceil$. To simplify the discussion, we assume $f_1 = f_2 = f_d$ and $Q_1 = Q_2 = Q$. Nevertheless, the proposed principle can be straightforwardly extended to more general cases. The bases then contain $Q + 1$ carriers $\{e^{j\omega_q n}\}_{q=0}^{Q}$ with $\omega_q \triangleq 2\pi(q - Q/2)/N$.

B. Proposed PSAM Strategy

Fig. 1 depicts our propose transmission model for time-varying TWRN, where a time interval NT_s is divided into several equi-length sub-blocks. Specifically, $\mathbb{T}_i$ divides its data block, containing both the information symbols and the pilot symbols, into K sub-blocks, each with length $M = N/2K$. After sending the kth sub-block, $\mathbb{T}_i$ will receive the sub-block sent back from $\mathbb{R}$ before it sends the $(k+1)$th sub-block. We will name the kth sub-block sent from $\mathbb{T}_i$ as the kth transmitted sub-block (TSB) while name the kth sub-block received at $\mathbb{T}_i$ as the kth received sub-block (RSB).

Symbol-wisely, we would like to use an unique time $n \in \{0, \ldots, N-1\}$ to represent the time slot for both the transmitting and the receiving period. For the ease of discussion, we define the overall information symbol and pilot index sets for TSB as $\mathcal{D}_t$ and $\mathcal{T}_t$. Similarly, define $\mathcal{D}_r$ and $\mathcal{T}_r$ as the overall information and pilot symbol index sets for RSB.

Represent the data symbols sent from $\mathbb{T}_i$ by $s_i(n)$, $n \in \mathcal{D}_t \bigcup \mathcal{T}_t$. Assume that the information symbols from $\mathbb{T}_i$ have the average power P_i, i.e., $\mathrm{E}\{|s_i(n)|^2\} = P_i$, $\forall n \in \mathcal{D}_t$, while the total training power from $\mathbb{T}_i$ is $P_{i,t}$, i.e., $\sum_{n \in \mathcal{T}_t} |s_i(n)|^2 = P_{i,t}$. With perfect synchronization, $\mathbb{R}$ receives

$$r(n) = h_1(n)s_1(n) + h_2(n)s_2(n) + w_r(n), \quad n \in \mathcal{D}_t \bigcup \mathcal{T}_t,$$

where $w_r(n)$ is the circularly symmetric complex Gaussian (CSCG) noise with variance σ_r^2. If the average transmit power of $\mathbb{R}$ is P_r, then the received signal will be scaled by a factor

$$\alpha(n) = \begin{cases} \sqrt{\frac{P_r}{\sigma_{h_1}^2 P_1 + \sigma_{h_2}^2 P_2 + \sigma_r^2}} & n \in \mathcal{D}_r \\ \sqrt{\frac{P_r}{\sigma_{h_1}^2 P_{1,t}/|\mathcal{T}_t| + \sigma_{h_2}^2 P_{2,t}/|\mathcal{T}_t| + \sigma_r^2}} & n \in \mathcal{T}_r \end{cases} \quad (4)$$

to keep the power constraint, where $|\mathcal{T}_t|$ is the cardinality of the set $\mathcal{T}_t$. Due to symmetry, we only present the estimation procedure at $\mathbb{T}_1$. The received signal is

$$\begin{aligned} y(n) &= \alpha(n)h_1(n)r(n-M) + w_1(n) \\ &= \alpha(n)h_1(n)h_1(n-M)s_1(n-M) \\ &\quad + \alpha(n)h_1(n)h_2(n-M)s_2(n-M) \\ &\quad + \underbrace{\alpha(n)h_1(n)w_r(n-M) + w_1(n)}_{w(n)}, \quad n \in \mathcal{D}_r \bigcup \mathcal{T}_r, \quad (5) \end{aligned}$$

where $w_1(n)$ is the CSCG noise at $\mathbb{T}_1$ with variance σ_1^2, and $w(n)$ denotes the overall noise.

C. On Channel Estimation

Applying (3), we rewrite (5) as

$$\begin{aligned} y(n) &= \alpha(n)\underbrace{\sum_{p=0}^{Q}\sum_{q=0}^{Q}\lambda_p\lambda_q e^{-j\omega_q M}e^{j(\omega_p+\omega_q)n}}_{b_1(n)} s_1(n-M) \\ &\quad + \alpha(n)\underbrace{\sum_{p=0}^{Q}\sum_{q=0}^{Q}\lambda_p\mu_q e^{-j\omega_q M}e^{j(\omega_p+\omega_q)n}}_{b_2(n)} s_2(n-M) + w(n) \\ &= \alpha(n)\sum_{m=0}^{2Q}\Big(\underbrace{\sum_{p+q=m}\lambda_p\lambda_q e^{-j\omega_q M}}_{x_1(m)}\Big)e^{j\theta_m n}s_1(n-M) \\ &\quad + \alpha(n)\sum_{m=0}^{2Q}\Big(\underbrace{\sum_{p+q=m}\lambda_p\mu_q e^{-j\omega_q M}}_{x_2(m)}\Big)e^{j\theta_m n}s_2(n-M) + w(n), \end{aligned}$$

$$n \in \mathcal{D}_r \bigcup \mathcal{T}_r, \quad (6)$$

where $\theta_m = 2\pi(m-Q)/N$, while $b_i(m)$ and $x_i(m)$ are defined as the corresponding items. When deriving the last equality in (6), we use the property that $\omega_p + \omega_q = \omega_{p'} + \omega_{q'}$ whenever $p + q = p' + q'$.

Define

$$\begin{aligned} \mathbf{x}_i &= [x_i(0), x_i(1), \ldots, x_i(2Q)]^T, \\ \mathbf{\Gamma} &= \mathrm{diag}\{e^{-j\omega_0 M}, e^{-j\omega_1 M}, e^{-j\omega_Q M}\}, \\ \boldsymbol{\lambda} &= [\lambda_0, \lambda_1, \ldots, \lambda_Q]^T, \quad \boldsymbol{\mu} = [\mu_0, \mu_1, \ldots, \mu_Q]^T, \end{aligned}$$

and define $\mathbf{\Lambda}$ as the $(2Q+1) \times (Q+1)$ Toeplitz matrix with the first column $[\boldsymbol{\lambda}^T, \mathbf{0}_{1 \times Q}]^T$. We can explicitly express the convolution as

$$\mathbf{x}_1 = \boldsymbol{\lambda} \otimes (\mathbf{\Gamma}\boldsymbol{\lambda}) = \mathbf{\Lambda}\mathbf{\Gamma}\boldsymbol{\lambda}, \quad \mathbf{x}_2 = \boldsymbol{\lambda} \otimes (\mathbf{\Gamma}\boldsymbol{\mu}) = \mathbf{\Lambda}\mathbf{\Gamma}\boldsymbol{\mu}. \quad (7)$$

Since directly estimating λ_q and μ_q from (6) requires complicated non-linear processing, we propose to first estimate the equivalent BEM coefficients $x_i(m)$ (with $4Q+2$ unknowns) and then recover the original BEM coefficients λ_q, μ_q.

III. CHANNEL ESTIMATION ALGORITHMS

Let us specifically define the indices in $\mathcal{T}_r$ as $n_0 < n_1 \ldots < n_T$, where $T+1 = |\mathcal{T}_r|$ denotes the total number of pilot symbols in RSB. The corresponding pilot symbol indices in TSB are then $n_i - M$. Define $\mathbf{y}_t = [y(n_0), y(n_1), \ldots, y(n_T)]^T$, $\mathbf{w}_t = [w(n_0), w(n_1), \ldots, w(n_T)]^T$, $\mathbf{t}_i = [s_i(n_0-M), s_i(n_1-M), \ldots, s_i(n_T-M)]^T$, $\mathbf{T}_i = \mathrm{diag}\{\mathbf{t}_i\}$, $i = 1, 2$ where $\mathbf{t}_i$ contains all the pilot symbols from $\mathbb{T}_i$. For notation simplicity, the mth entry of $\mathbf{t}_i$ is also denoted by $t_i(m)$, $m = 0, \ldots, T$.

With the aid of (6), we can express $\mathbf{y}_t$ in a matrix form

$$\mathbf{y}_t = \alpha\mathbf{T}_1\mathbf{A}\mathbf{x}_1 + \alpha\mathbf{T}_2\mathbf{A}\mathbf{x}_2 + \mathbf{w}_t, \quad (8)$$

where we omit the index n in $\alpha(n)$ because the scaling is constant for all pilot symbols, and $\mathbf{A}$ is the $(T+1) \times (2Q+1)$ matrix

$$\mathbf{A} = \begin{bmatrix} e^{j\theta_0 n_0} & e^{j\theta_1 n_0} & \cdots & e^{j\theta_{2Q} n_0} \\ e^{j\theta_0 n_1} & e^{j\theta_1 n_1} & \cdots & e^{j\theta_{2Q} n_1} \\ \vdots & \vdots & \cdots & \vdots \\ e^{j\theta_0 n_T} & e^{j\theta_1 n_T} & \cdots & e^{j\theta_{2Q} n_T} \end{bmatrix}. \tag{9}$$

A. Channel Estimation for Equivalent BEM

When $T + 1 \geq 4Q + 2$, we have sufficient observations to estimate all the unknown $x_i(m)$'s. Define

$$\mathbf{T} = [\alpha \mathbf{T}_1 \mathbf{A}, \alpha \mathbf{T}_2 \mathbf{A}], \qquad \mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T. \tag{10}$$

The lease square (LS) estimation of $\mathbf{x}$ is immediately obtained as

$$\hat{\mathbf{x}} = \mathbf{T}^\dagger \mathbf{y} = (\mathbf{T}^H \mathbf{T})^{-1} \mathbf{T}^H \mathbf{y} \tag{11}$$

with the error covariance matrix

$$\mathbf{W} = \mathbf{T}^\dagger \mathbf{D} (\mathbf{T}^\dagger)^H, \tag{12}$$

where $\mathbf{T}^\dagger$ is the pseudo-inverse of the matrix $\mathbf{T}$, and

$$\mathbf{D} = \begin{bmatrix} \alpha^2 \sigma_r^2 |h_1(n_0)|^2 + \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha^2 \sigma_r^2 |h_1(n_T)|^2 + \sigma_1^2 \end{bmatrix}.$$

B. Training Design

The channel estimation MSE is defined as $tr(\mathbf{W})$ and is, unfortunately, related to the instant CSI. In order to design pilot symbols that, on average, yield the best performance, we propose to minimize the average MSE that is defined as

$$\text{AMSE} = E\{tr(\mathbf{W})\} = (\alpha^2 \sigma_{h_1}^2 \sigma_r^2 + \sigma_1^2) tr((\mathbf{T}^H \mathbf{T})^{-1}), \tag{13}$$

where the property $J_0(0) = 1$ is used.

The training design for the proposed scheme includes selecting the number of the pilot symbols per block, the placement of pilot symbols, as well as the power allocation for each pilot, based on the AMSE.

The optimal $\mathbf{t}_i$, as well as pilot index set $\mathcal{T}_r$ should be obtained from

$$(P1): \quad \min_{\mathbf{t}_1, \mathbf{t}_2, \mathcal{T}_r} \quad (\alpha^2 \sigma_{h_1}^2 \sigma_r^2 + \sigma_1^2) tr((\mathbf{T}^H \mathbf{T})^{-1}) \tag{14}$$
$$s.t. \quad \mathbf{t}_i^H \mathbf{t}_i \leq P_{i,t}, \quad i = 1, 2.$$

Since α is only related with T, we can first solve the problem for a given T, and it becomes

$$(P2): \quad \min_{\substack{\mathbf{t}_1, \mathbf{t}_2, \\ n_i: 0 \leq i \leq T}} \quad tr((\mathbf{T}^H \mathbf{T})^{-1}) \tag{15}$$
$$s.t. \quad \mathbf{t}_i^H \mathbf{t}_i \leq P_{i,t}, \quad i = 1, 2.$$

The challenging part here lies in that $\mathbf{T}_i$ must be a diagonal matrix.

From [10], we know that

$$tr((\mathbf{T}^H \mathbf{T})^{-1}) \geq \sum_{i=0}^{4Q+1} \frac{1}{[\mathbf{T}^H \mathbf{T}]_{i,i}} = \sum_{i=1}^{2} \frac{2Q+1}{\sum_{m=0}^{T} |t_i(m)|^2} \tag{16}$$

and the equality holds when $\mathbf{T}^H \mathbf{T}$ is diagonal. However, this inequality does not directly testify that the diagonality must hold for the optimal $\mathbf{T}^H \mathbf{T}$. Let us first formulate a new optimization problem:

$$(P3): \quad \min_{\mathbf{t}_1, \mathbf{t}_2} \quad \sum_{i=1}^{2} \frac{2Q+1}{\sum_{m=0}^{T} |t_i(m)|^2} \tag{17}$$
$$s.t. \quad \mathbf{t}_i^H \mathbf{t}_i \leq P_{i,t}, \quad i = 1, 2.$$

Obviously, the optimal objective of (P3) serves as a lower bound for (P2). Since (P3) is a simple convex optimization, any training sequences that satisfy $\mathbf{t}_i^H \mathbf{t}_i = P_{i,t}$ would be optimal. Therefore, if we could find $\mathbf{t}_i$'s that make $\mathbf{T}^H \mathbf{T}$ diagonal, then they must also be the optimal solution for problem (P2). Namely, the optimal solutions to (P2), if existing, should satisfy

$$\mathbf{A}^H \mathbf{T}_i^H \mathbf{T}_i \mathbf{A} = P_{i,t} \mathbf{I}_{2Q+1}, \quad i = 1, 2, \tag{18a}$$
$$\mathbf{A}^H \mathbf{T}_1^H \mathbf{T}_2 \mathbf{A} = \mathbf{0}_{2Q+1}. \tag{18b}$$

Observing the Vandemonde structure of $\mathbf{A}$ and the structure of θ_m, we know that if pilot symbols are equi-powered and equi-spaced over $\{0, N-1\}$, then (18a) is satisfied, i.e.,

$$C1): \quad |t_i(m)|^2 = \frac{P_{i,t}}{T+1}, \quad \forall m = 0, 1, \ldots, T, \quad i = 1, 2,$$
$$C2): \quad n_m = mL + l_0, \quad \forall l_0 \in [M, L-1],$$

where $L = \frac{N}{T+1}$ is an integer, and the property that $n_0 \geq M$ is taken into consideration.

Observing from Fig. 1, receiving equi-spaced pilot at $\mathbb{T}_i$ is only possible if there are $T+1$ TSB (RSB), each containing only one pilot symbol that is put at the same position. Meanwhile, $T+1$ should divide K, i.e., $K = \tau(T+1)$ for an integer τ. For the rest $K - (T+1)$ TSB (RSB), only information symbols are transmitted.

Remark 1: We need $T + 1 \geq 4Q + 2$ sub-blocks with pilot symbol to provide sufficient observations for channel estimation. For time-varying channel, $Q \geq 1$, then one data frame at least contains $K = 6$ TSB and RSB, respectively. Therefore, the traditional two-way relaying scheme [1] ($K = 1$) is not suitable for time-selective environment. We here re-emphasize the novelty of our proposed new scheme in Fig. 1.

Moreover, under C1) and C2), the following condition can guarantee (18b):

$$C3): \quad \sum_{m=0}^{T} t_2^*(m) t_1(m) e^{j2\pi mk/(T+1)} = 0,$$

for $\forall k = -2Q, -2Q+1, \ldots, 2Q$. One example of pilot sequences that satisfies conditions C1)–C3) is

$$
\mathbf{t}_1 = \sqrt{\frac{P_{1,t}}{(T+1)}}[+1, +1, +1, \ldots, +1, +1]^T, \tag{19a}
$$

$$
\mathbf{t}_2 = \sqrt{\frac{P_{2,t}}{(T+1)}}[1, e^{j2\pi v/(T+1)}, \ldots, e^{j2\pi Tv/(T+1)}]^T, \tag{19b}
$$

$$
\forall v = 2Q+1, \ldots, (T+1) - 2Q - 1.
$$

With the optimal training, the minimum $tr((\mathbf{T}^H\mathbf{T})^{-1})$ is $(2Q+1)(1/P_{1,t} + 1/P_{2,t})$ and is independent from T. Thanks to this property, the optimal T can be independently obtained from

$$
|\mathcal{T}_r| = \arg\min_K \alpha^2. \tag{20}
$$

Since α is an increasing function of T, the optimal T should be made as large as possible. However, increasing T would reduce the efficiency of the data transmission and, consequently, the system throughput. Besides, the constant $\sigma_{h_1}^2 \sigma_r^2$ will dominate the summation from $\left(\sigma_{h_1}^2 \sigma_r^2 + \frac{\sigma_1^2}{\alpha^2}\right)$ when T is greater than a certain threshold. Therefore, increasing T beyond a certain value cannot apparently improve the channel estimation MSE, but the throughput will be linearly decreased.

More meaningful design of T should be obtained from maximizing the transmission throughput criterion as did in [9]. Due to page limit, in this paper we simply consider the achieving the minimum amount of training as the goal of optimal training length. Therefore, T should be selected as the minimum value that satisfies $T+1 \geq 4Q+2$ and $(T+1)\tau = K$ for some τ.

C. Recovery of BEM Coefficients

Once $\mathbf{x}_1$ and $\mathbf{x}_2$ are estimated, it is possible to recover the original BEM coefficients λ_q and μ_q. Afterwards, one may reconstruct $\mathbf{x}_i$'s from the estimated λ_q and μ_q. By doing so, the estimation accuracy can be improved since we reinforce the BEM length into the consideration [3]. Such a process is also known as *denoising*. It is, meanwhile, important to find a low complex way to retrieve $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ from $\mathbf{x}_i$ that avoids the nonlinear searching of multi-variables.

Let $\tilde{\boldsymbol{\lambda}}$ be the Z-point discrete Fourier transform (DFT) of $\boldsymbol{\lambda}$, whose mth entry is defined as

$$
\tilde{\lambda}_m = \sum_{q=0}^{Q} \lambda_q e^{-j2\pi qm/Z}. \tag{21}
$$

The m-th element of the Z-point DFT of $\boldsymbol{\Gamma}\boldsymbol{\lambda}$ is

$$
\xi_m = \sum_{q=0}^{Q} \lambda_q e^{\frac{-j2\pi(q-Q/2)}{2K}} e^{\frac{-j2\pi qm}{Z}} = e^{j\frac{\pi Q}{2K}} \sum_{q=0}^{Q} \lambda_q e^{-j\frac{2\pi q(\frac{Z}{2K}+m)}{Z}}. \tag{22}
$$

If $R = \frac{Z}{2K}$ is an integer, then (22) becomes $e^{j\frac{\pi Q}{2K}}\tilde{\lambda}_{\langle m+R\rangle_Z}$, where $\langle \cdot \rangle_Z$ denotes the modulo-Z operation. Then the mth

element of the Z-point NDFT of $\mathbf{x}_1(m)$ is

$$
\tilde{x}_1(m) = \xi_m\tilde{\lambda}_m = e^{j\frac{\pi Q}{2K}}\tilde{\lambda}_m\tilde{\lambda}_{\langle m+R\rangle_Z}. \tag{23}
$$

Our target is to retrieve Z unknown $\tilde{\lambda}_m$ from Z equations

$$
\tilde{\lambda}_m\tilde{\lambda}_{\langle m+R\rangle_Z} = \tilde{x}_1(m)e^{-j\frac{\pi Q}{2K}} = c_m, \quad \forall m = 0, \ldots, Z-1. \tag{24}
$$

Let us first provide the following lemma and theorem for general values of Z and R.

Lemma 1: If Z and R are co-prime, then the index set $\mathcal{I} = \{\langle m+uR\rangle_Z\}_{u=0}^{Z-1}$ is the same as the universal set $\{0, \ldots, Z-1\}$, or equivalently

$$
\langle m+uR\rangle_Z \neq \langle m+vR\rangle_Z, \quad 0 \leq u < v \leq Z-1. \tag{25}
$$

Proof: Let us first assume the contrary holds, i.e.,

$$
\langle m+uR\rangle_Z = \langle m+vR\rangle_Z, \quad \exists u < v. \tag{26}
$$

Then we derive

$$
(v-u)R = kZ \tag{27}
$$

for some integer $k \neq 0$. Since Z and R are co-prime, their least common multiple must be ZR. However in (27) $(v-u) < Z$, so equality (27) cannot hold. By the contradiction, we proved Lemma 1. ∎

Theorem 1: If Z is odd and is co-prime with R, then $\tilde{\lambda}_m$'s can be found from (24) as

$$
\tilde{\lambda}_m = \frac{\left(\prod_{i=0}^{Z-1} c_i\right)^{1/2}}{\prod_{i=0}^{\frac{Z-3}{2}} c_{\langle m+(2i+1)R\rangle_Z}}, \quad \forall m = 0, \ldots, Z-1, \tag{28}
$$

with only common sign ambiguity.

Proof: Let us now define a new variable

$$
c = \left(\prod_{i=0}^{Z-1} c_i\right)^{1/2} = I_s\prod_{i=0}^{Z-1}\tilde{\lambda}_i, \tag{29}
$$

where $I_s = \pm 1$ representing the sign ambiguity.

When Z is odd, the denominator in (28) can be expanded as

$$
\prod_{i=0}^{\frac{Z-3}{2}} c_{\langle m+(2i+1)R\rangle_Z} = \prod_{i=1}^{Z-1}\tilde{\lambda}_{\langle m+iR\rangle_Z} = \prod_{\substack{i=0 \\ i\neq m}}^{Z-1}\tilde{\lambda}_i, \tag{30}
$$

where Lemma 1 is applied in the last equality.

Dividing (29) by (30) proves Theorem 1, where I_s serves as the common sign ambiguity for all $\tilde{\lambda}_m$. ∎

Unfortunately, in our problem $Z = 2RK$ can only be an even number. We then propose to choose $Z = 2RK+1$, which is odd and must be co-prime with R. In this case,

$$
\xi_m = e^{j\frac{\pi Q}{2K}}\sum_{q=0}^{Q}\lambda_q e^{-j\frac{2\pi q(R+m)}{Z}}e^{-j\frac{2\pi q\frac{1}{2K}}{Z}} \approx \tilde{\lambda}_{\langle m+R\rangle_Z}, \tag{31}
$$

so (24) approximately holds.

After obtaining $\tilde{\boldsymbol{\lambda}}$ from Theorem 1, we can recover $\boldsymbol{\lambda}$ from the first $Q+1$ elements of the Z-point inverse Fourier transform (IDFT) of $\tilde{\boldsymbol{\lambda}}$. Then , $\boldsymbol{\mu}$ can be obtained from $\boldsymbol{\mu} = \boldsymbol{\Gamma}^H\boldsymbol{\Lambda}^{\dagger}\mathbf{x}_2$.
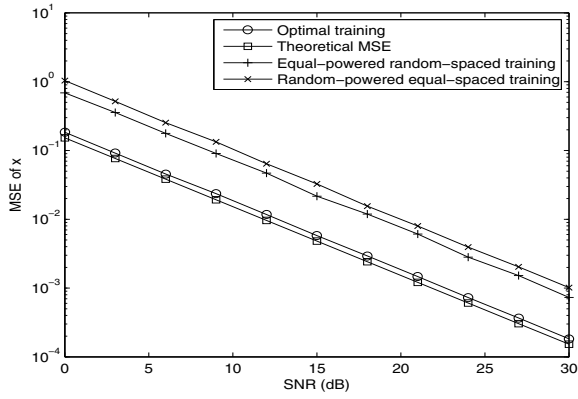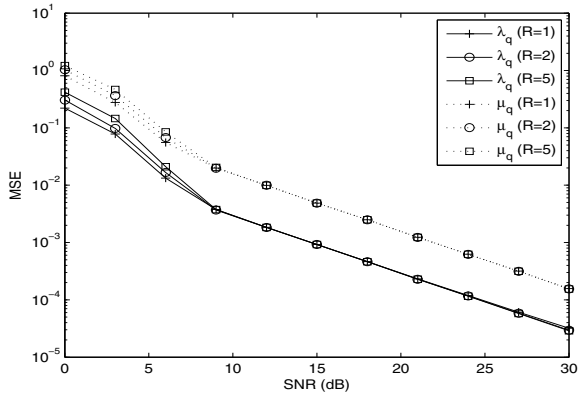
Fig. 2. Channel estimation MSEs versus SNR for $\mathbf{x}$

Fig. 3. Channel estimation MSEs versus SNR for $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ with different choices of R.

IV. SIMULATION RESULTS

Parameters are taken as $Q = 4$, $N = 352$, $M = 8$, and $K = 22$. Totally 10000 Monte-Carlo trials are used for averaging. For channel estimation part, we directly generate the time-varying channels from the BEM model (3) as did in [9]. We compare the optimal training with two other choices: i) equi-powered but randomly spaced pilots; ii) equi-spaced but random powered pilots. The estimation MSEs of the equivalent channel $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T]^T$ from the three different types of training are shown in Fig. 2. The theoretical MSE is also displayed for comparison. Clearly, the designed training sequence achieves the best performance and its MSE is very close to the theoretical one.

Next we verify the performance of recovering the coefficients $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$. From (22), we take $R = 1, 2, 5$, respectively and $Z = 2RK + 1$ are $45, 89, 221$, respectively. The estimation MSEs of both $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ versus SNR are shown in Fig. 3. The choice of different R does not affect the performance much. Compared with the estimated $\boldsymbol{\lambda}$, larger error in $\boldsymbol{\mu}$ is observed. This is expected since the errors in $\boldsymbol{\lambda}$ propagate to the estimates of $\boldsymbol{\mu}$. It is also observed that the performance degrades severely when SNR is relatively low, say smaller than 8 dB. This is because the the product in both the numerator and the denominator of (28) is susceptible to error propagation.
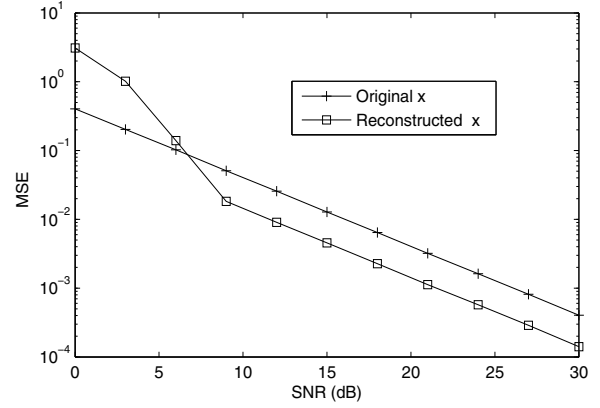

Fig. 4. Channel estimation MSEs versus SNR for the estimated $\mathbf{x}$ and the recovered $\mathbf{x}$.

After obtaining the estimated coefficients $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, we can re-construct $\mathbf{x}$ whose performance is shown in Fig.4. It can be seen that recovered $\mathbf{x}$ is more accurate than the originally estimated $\mathbf{x}$ at relatively high SNR.

V. CONCLUSION

In this paper, we studied the channel estimation and training design for TWRN. We proposed a novel PSAM scheme that is applicable for time selective environment and developed estimation algorithms for both the cascaded channels and individual channels. We also derived the optimal training strategy based on minimizing the estimation MSE criterion. Finally simulations were conducted to demonstrate the effectiveness of the proposed study.

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] 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.

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

[4] C. S. Patel and G. L. Stuber, "Channel estimation for amplify and forward relay based cooperation diversity systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 3348–3356, June 2007.

[5] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686-693, Nov. 1991.

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

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

[8] A. S. Akki and F. Haber, "A statistical model for mobile-to-mobile land communication channel," *IEEE Trans. Veh. Technol.*, vol. VT-35, no. 1, pp. 2–7, Feb. 1986.

[9] 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.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.