Joint Semi-Blind Channel Estimation and Synchronization in Two-Way Relay Networks

Qiong Zhao, *Student Member, IEEE*, Zhendong Zhou, *Senior Member, IEEE*,
Jun Li, *Member, IEEE*, and Branka Vucetic, *Fellow, IEEE*

***Abstract*—In this paper, we propose a synchronization and channel estimation method for amplify-and-forward two-way relay networks (AF-TWRNs) based on a low-complexity maximum-likelihood (LCML) algorithm and a joint synchronization and channel estimation (JSCE) algorithm. For synchronous AF-TWRNs, the LCML algorithm blindly estimates general nonreciprocal flat-fading channels. We formulate the channel estimation as a convex optimization problem and obtain a closed-form channel estimator. Based on the mean square error (MSE) analysis of the LCML algorithm, we propose a generalized LCML (GLCML) algorithm to perform channel estimation in the presence of the timing offset. Based on the approximation of the LCML algorithm, the JSCE algorithm is proposed to estimate jointly the timing offset and channel parameters. The theoretical analysis shows that the closed-form LCML channel estimator is consistent and unbiased. The analytical MSE expression shows that the estimation error approaches zero in scenarios with either a high signal-to-noise ratio (SNR) or a large frame length. Monte Carlo simulations are employed to verify the theoretical MSE analysis of the LCML algorithm. In the absence of perfect timing synchronization, the GLCML algorithm selects an estimation sample, which produces the optimal channel estimation, according to the MSE analysis. Simulation results also demonstrate that the JSCE algorithm is able to achieve accurate timing offset estimation.**

***Index Terms*—Channel estimation, convex optimization, maximum-likelihood, synchronization, two-way relays.**

I. Introduction

RECENTLY two-way relay networks (TWRNs) [1] have attracted considerable research interest due to their potential for improving spectral efficiency and reliability in wireless communications. The amplify-and-forward (AF) relaying protocol [2] is commonly used in TWRNs since minimal signal processing is required at the relay node. In AF-TWRNs, synchronization and channel estimation are two essential issues for signal detection. Efficient channel estimation in AF-TWRNs poses a big challenge for doing coherent demodulation. In the case of asynchronous AF-TWRNs, the existence of relative frequency and timing offsets between signals from two source nodes makes the coherent demodulation more challenging.

Many algorithms have been proposed for estimating the channel in AF-TWRNs assuming perfect synchronization, whereas not much attention has been paid to studying the joint synchronization and channel estimation (JSCE) problem. The existing channel estimation methods use either training-based approaches [3], which estimate the channel by using training symbols known to both the transmitter and receiver, or blind approaches [4], which do not depend on training symbols. It has been stated in [5]–[7] that training-based channel estimation approaches achieve good performance and are practical. Nevertheless, the spectral efficiency is significantly reduced by training overhead. Unlike training-based channel estimation, blind channel estimation approaches [8], [10] remarkably decrease undesirable training overhead. Thus, they offer a superior tradeoff between estimation accuracy and spectral efficiency. In [8], under M-ary phase-shift keying (MPSK) modulation, a semi-blind channel estimation algorithm, which employs only one training symbol per estimation, is proposed for estimating reciprocal flat-fading channels in single-relay AF-TWRNs. Since reciprocal channels are not always practical, a deterministic maximum-likelihood (DML) channel estimator considering nonreciprocal channels is proposed in [10]. As the DML estimator is unsuitable for BPSK, an alternative estimator called the modified constrained maximum-likelihood (MCML) estimator is proposed for this case in [10]. The MCML channel estimator takes into account the BPSK structure and approaches the true channel with high probability at a high SNR. The approaches in [8] and [10] noticeably reduce training overhead and achieve accurate MSE performances. However, the proposed objective functions are nonconvex. Due to the nonconvex optimization function, the DML and MCML algorithms have to rely on numerical solutions by using optimization tools.

To consider channel estimation and synchronization jointly, a training-based estimation method is proposed in [7] for estimating channel parameters and the frequency offset in orthogonal-frequency-division-multiplexing-modulated TWRNs. However, perfect timing synchronization is assumed, and significant training overhead are still required for the estimation.

In this paper, we start by assuming perfect frequency synchronization, and we then study the channel estimation and timing synchronization problem jointly in MPSK-modulated AF-TWRNs.

First, we assume perfect timing synchronization and propose a semi-blind low-complexity maximum-likelihood (LCML)

Manuscript received April 12, 2013; accepted December 29, 2013. Date of publication January 21, 2014; date of current version September 11, 2014. This paper was presented in part at the 22nd IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Toronto, ON, Canada, September 2011. The review of this paper was coordinated by Dr. A. J. Al-Dweik.

The authors are with the Telecommunication Laboratory, School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: qiong.zhao@sydney.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2014.2301812

channel estimation algorithm for estimating general nonreciprocal flat-fading channels. In the LCML algorithm, we formulate a convex objective function [9] and derive a closed-form channel estimator. With the channel state information, only one training symbol is necessary for resolving phase ambiguity [11] in the signal demodulation.

In the asynchronous AF-TWRNs, we extend the LCML algorithm and develop a generalized LCML (GLCML) algorithm to perform channel estimation in the presence of a timing offset. We derive two channel estimations from the overlapped and nonoverlapped signals, respectively. Then, we devise an estimation sample selection criterion (ESSC) to choose the channel estimation with the minimum MSE.

Then, a JSCE algorithm is proposed for estimating the timing offset. We first propose a frame synchronization algorithm, which is referred to as the frame asynchronous channel estimation (FACE) algorithm, to estimate the frame offset (integral timing offset) by energy detection and the cross correlation of the received and transmitted signals. After frame boundaries are determined by the frame synchronization algorithm, the symbol synchronization is performed jointly with the channel estimation, by the symbol asynchronous channel estimation (SACE) algorithm, based on the overlapped signals to estimate the symbol offset (fractional timing offset).

The theoretical analysis shows that the LCML channel estimator is consistent (the estimation error approaches zero if the number of estimation samples approaches infinite) and unbiased (the expectation of the channel estimation equals the real channel parameter value) [12], and its computational complexity $\mathcal{O}(N)$ is linear in terms of samples size N. The theoretical MSE performance of the LCML algorithm shows that the derived LCML channel estimator approaches the real channel parameters in scenarios with either a high SNR or a large frame length. Monte Carlo simulations are employed to confirm the analytical MSE results of the LCML algorithms. Moreover, the simulation results demonstrate that the GLCML algorithm always select the optimal channel estimation in the cases of varying timing offsets and that the JSCE algorithm is able to achieve accurate timing offset estimation.

The main contributions of this paper are as follows.

1) In synchronous AF-TWRNs, a convex optimization function for blind channel estimation is not available in the open literature as far as we know. We propose the LCML algorithm with a convex optimization function that produces a closed-form channel estimator. Furthermore, the availability of the closed-form channel estimation enables us to derive the analytical estimation MSE. Based on the MSE analysis, we make the proposed algorithm more practical by relaxing the assumption of perfect synchronization and propose the GLCML algorithm for the asynchronous system, where there exists a relative timing offset between both the source nodes.
2) We formulate the channel estimation and timing synchronization as an overall maximum-likelihood estimation problem, which is solved by the JSCE algorithm. The analysis shows that the error probability of the FACE algorithm approaches zero in a scenario with a large frame length. The symbol offset is estimated jointly with the channel estimation by the SACE algorithm. The JSCE algorithm is capable of achieving accurate timing offset estimation and channel estimation even in the absence of perfect timing synchronization.
3) We analyze the performance of the LCML algorithm in the scenarios with a high SNR and a large frame length, respectively. In the case of a high SNR, the LCML channel estimations approach the real channel parameter values with the probability of $1-(1/M^{N-1})$ when $M=2$ and $1-(2/M)^{N-1}(M-1)$ if $M>2$, where M is the modulation order and N denotes the frame length; we further conclude that the LCML channel estimator is unbiased. On the other hand, in the case of a large frame length, the LCML channel estimator is consistent if channel parameters belong to compact sets (closed and bounded sets) [12].
4) We derive a closed-form MSE expression of the LCML channel estimator with respect to the SNR and N, MSE $\propto (1/\text{SNR}\ N)$ in BPSK $(M=2)$ and MSE $\propto (2/\text{SNR}\ N)$ in MPSK $(M>2)$, which is consistent with the simulation results. Both the theoretical and numerical MSE performances demonstrate that the LCML channel estimator approaches the true channel for scenarios with either a high SNR or a large frame length. Based on the analytical MSE expression, the optimal channel estimation is achieved in the GLCML algorithm by comparing the estimation MSE of the overlapped and nonoverlapped samples.

The remainder of this paper is organized as follows. Section II presents the system model of AF-TWRNs. In Section III, we formulate an overall maximum-likelihood optimization function in asynchronous TWRNs, assuming that the source node is able to synchronize with either of the transmitted signals. The LCML, GLCML, and JSCE algorithms are proposed in Sections IV–VI, respectively, to study the channel estimation problem formulated in Section III. The timing offset is estimated in Section VI. Then, we present the analysis of the LCML algorithm in Section VII. Simulation results are provided to verify the proposed algorithms and analysis in Section VIII. Finally, conclusions are drawn in Section IX.

Notations: $\|\mathbf{a}\|$ and $\mathbf{a}^*$ represent the norm and complex conjugate of vector $\mathbf{a}$, respectively. The conjugate transpose and transpose of vector $\mathbf{a}$ are denoted $\mathbf{a}^H$ and $\mathbf{a}^T$, respectively. The angle and norm of complex number a are denoted $\angle a$ and $|a|$, respectively. Expectation and variance of a are denoted $\mathrm{E}\{a\}$ and $\mathrm{Var}\{a\}$, respectively. $\mathbb{N}$ represents a nonnegative integer set. $\mathbb{C}$ and $\mathbb{R}$ stand for the complex and real number fields, respectively.

II. System Model

A typical half-duplex AF-TWRN over quasi-static flat-fading [1] channels is considered. In the scenario of a frequency-selective channel, the proposed channel estimation algorithms can be applied to estimate channel parameters of different frequency subbands individually. The system shown in Fig. 1

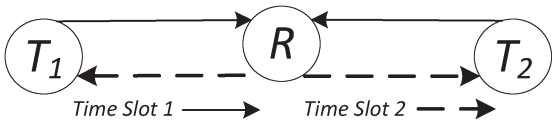


Fig. 1. Dual time-slot TWRN with two source nodes and one relay node.

is composed of three nodes: source nodes T_1 and T_2 and relay node R. Each node is equipped with a single antenna. The two source nodes are out of each other's transmission range.

Each signal transmission process consists of two time slots. In the first time slot, the source nodes T_1 and T_2 simultaneously transmit signals to the relay node. The transmitted signals of source nodes T_1 and T_2 are $s_1(t) = \sum_{i=1}^{N} e^{jw_1 Ti} s_{1,i} f(t - iT)$ and $s_2(t) = \sum_{i=1}^{N} e^{jw_2 Ti} s_{2,i} f(t - iT)$, respectively. Without loss of generality, both T_1 and T_2 are assumed to transmit N symbols in a frame. $f(t)$ is the time-invariant pulse-shaping transmit filter of source nodes T_1 and T_2. w_1 and w_2 are the carrier angular frequencies of the transmitters. The carrier frequency is assumed perfectly synchronized across source nodes T_1 and T_2. T is the symbol time interval, and $j \triangleq \sqrt{-1}$. As MPSK modulation is employed, we obtain

$$s_{1,i} = \begin{cases} \sqrt{P_1} e^{j\phi_{1i}}, & i = 1, 2, \ldots, N \\ 0, & \text{else} \end{cases}$$
$$s_{2,i} = \begin{cases} \sqrt{P_2} e^{j\phi_{2i}}, & i = 1, 2, \ldots, N \\ 0, & \text{else} \end{cases}$$

where P_1 and P_2 are the transmit power of T_1 and T_2, respectively, and ϕ_{1i} and ϕ_{2i} are MPSK-modulated phases, which are independent and uniformly distributed in set $S_M = \{(2\pi(l-1)/M), l = 1, \ldots, M\}$, with M being the modulation order.

During the first time slot, the overall signal received by the relay node is given by

$$r(t) = h_1 s_1(t + \tau_1) + g_1 s_2(t + \tau_2) + n_1(t).$$

Due to timing asynchrony, there exist timing offsets τ_1 and τ_2 of source nodes T_1 and T_2 relative to the relay node R, respectively. As the relay node only amplifies the received signal, the timing offsets τ_1 and τ_2 are not needed for synchronization or channel estimation. This will be proven in Sections III and VI. h_1 and g_1 are complex coefficients of flat-fading channels $T_1 \to R$ and $T_2 \to R$, respectively. Since nonreciprocal channels are considered in this paper, the complex channel coefficients of the links $R \to T_1$ and $R \to T_2$ are denoted h_2 and g_2, respectively. Channel coefficients h_1 and h_2 and g_1 and g_2 are modeled as independent and identically distributed (i.i.d.) in $\mathcal{CN}(0, \sigma_c^2)$ and remain fixed during one frame. Here, $\mathcal{CN}(0, \sigma_c^2)$ represents the complex normal distribution with a zero mean and the variance of σ_c^2. $n_1(t)$ is complex additive white Gaussian noise (AWGN) distributed in $\mathcal{CN}(0, \sigma_n^2)$.

In the second time slot, the relay node amplifies the received signal $r(t)$ and then broadcasts the amplified signal $A_r(t) = Kr(t)$, where K is the power scaling factor. To maintain average power of P_r at the relay node over a long term, the power scaling factor $K = \sqrt{(P_r/\sigma_c^2 P_1 + \sigma_c^2 P_2 + \sigma_n^2)}$ [5] is used in this paper, where P_r denotes the transmit power of R. Here, we assume that P_1, P_2, σ_c^2, and σ_n^2 are *a priori* known to relay node R. However, the rough knowledge of P_1 and P_2 is sufficient to obtain K. In this paper, all the estimations are performed at the source nodes, rather than at the relay node.

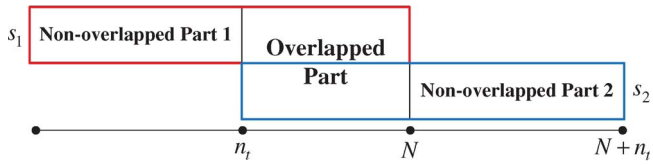


Fig. 2. Structure of the received signal in the asynchronous system.

Without loss of generality, signal detection at source node T_1 is considered. The received signal at T_1 is obtained as $r_1(t) = h_2 K r(t) + n_2(t)$, where $n_2(t)$ is AWGN distributed in $\mathcal{CN}(0, \sigma_n^2)$. Assuming perfect frequency synchronization, there exists a relative timing offset between T_1 and T_2 in the time asynchronous AF-TWRNs. We refer to the integral timing offset n_t as the frame offset and the fractional timing offset τ as the symbol offset.

In this paper, the joint timing synchronization and channel estimation problem in asynchronous AF-TWRNs is studied in two parts.

- Section III: Under the assumption that T_1 is able to synchronize with either $s_1(t)$ or $s_2(t)$, we formulate the channel estimation problem in the presence of a relative timing offset between both source nodes, which is solved in two steps.
  1) Section IV: The LCML channel estimator is proposed to estimate the channel in a time-synchronized system.
  2) Section V: The GLCML channel estimator is proposed by extending the LCML algorithm in the presence of a timing offset. As a result, the channel estimation problem formulated in Section III is solved.
- Section VI: We propose the JSCE algorithm that enables T_1 to synchronize with either $s_1(t)$ or $s_2(t)$. The JSCE algorithm enables the joint estimation of the timing offset and channel parameters.

III. Channel Estimation in Asynchronous Amplify-and-Forward Two-Way Relay Networks

Here, we study channel estimation in asynchronous AF-TWRNs under the assumption that T_1 is able to synchronize with either $s_1(t)$ or $s_2(t)$. We will relax this assumption in Section VI. Due to the timing offset, the transmitted signals $s_1(t)$ and $s_2(t)$ are not aligned. As a result, there is one overlapped part and two nonoverlapped parts in the received signal, as shown in Fig. 2. The overlapped part consists of two signals, and the nonoverlapped parts consist of only one signal. When T_1 is synchronized with the target signal, it filters $r_1(t)$ with the matched filter $f'(t) = f(T - t)$ and samples it at every T period. In the system, T_1 performs twice samplings to the received signal. First, T_1 synchronizes with $s_1(t)$ and obtains the nonoverlapped samples of $s_1(t)$ as

$$r_{\text{async}}(i) = Kh_1 h_2 s_{1,i} + Kh_2 n_{1,i} + n_{2,i}, \quad i = 1, \ldots, n_t. \quad (1)$$

Then, T_1 synchronizes with $s_2(t)$ and obtains the overlapped and nonoverlapped samples of $s_2(t)$ as

$$\begin{aligned} r_{\text{async}}(i) &= Kh_1h_2\left[f_1(\tau)s_{1,\,i} + f_2(\tau)s_{1,\,i+1}\right] \\ &\quad + Kg_1h_2s_{2,\,i-n_t} + Kh_2n_{1i} + n_{2i} \\ &\quad i = n_t+1,\ldots,N \end{aligned} \tag{2}$$

$$\begin{aligned} r_{\text{async}}(i) &= Kg_1h_2s_{2,\,i-n_t} + Kh_2n_{1,\,i} + n_{2,\,i}, \\ &\quad i = N+1,\ldots,N+n_t \end{aligned} \tag{3}$$

where n_{1i} and n_{2i} are AWGN distributed in $\mathcal{CN}(0,\,\sigma_n^2)$, and the frame length of the received samples is $N+n_t$. $f_1(\tau)$ and $f_2(\tau)$ are the factors resulting from the symbol offset τ and the use of a matched filter. The values of $f_1(\tau)$ and $f_2(\tau)$ are related to the filter type and symbol offset τ. The analytical expressions of $f_1(\tau)$ and $f_2(\tau)$ are derived in Appendix A. Equation (2) implies that the timing offsets τ_1 and τ_2 between the source nodes and relay node are not needed for channel estimation.

Based on the received discrete signal samples, T_1 needs to demodulate s_2. Due to the fact that T_1 knows its self-transmitted signal $s_1(t)$ and the employed phase modulation scheme, signal demodulation of $s_2(t)$ can be conducted based on the power scaling factor K and channel coefficients h_1, h_2, g_1, and g_2. However, it is complicated to estimate these unknown variables individually. By inspecting (1)–(3), it is sufficient for the signal demodulation purpose to estimate $f_1(\tau)$, $f_2(\tau)$, and the composite channel parameters $H \triangleq Kh_1h_2$, $G \triangleq Kg_1h_2$, and $\sigma^2 \triangleq (K^2|h_2|^2+1)\sigma_n^2$ jointly in MPSK-modulated AF-TWRNs. The power scaling factor K is not assumed known to both the source nodes but is estimated jointly with channel coefficients at both the source nodes. In fact, the estimation of K accounts for any signal power variations, including the free-space path loss.

Channel estimation is performed at T_1 by using $N+n_t$ received samples $r_{\text{async}}(i)$, $i=1,\ldots,N+n_t$, given in (1)–(3). Let $\mathbf{r} \triangleq [r_{\text{async}}(1),\ldots,r_{\text{async}}(N+n_t)]^T$ be the vector of the received signals. Here, we denote set $\mathbf{v} \triangleq [v_1,\,v_2,\ldots,\,v_n]$ and subset $\mathbf{v}_i^j \triangleq [v_i,\,v_{i+1},\ldots,\,v_j]$. The overlapped part is denoted $\mathbf{r}_{n_t+1}^N$, and the nonoverlapped parts of $s_1(t)$ and $s_2(t)$ are represented by $\mathbf{r}_1^{n_t}$ and $\mathbf{r}_{N+1}^{N+n_t}$, respectively. Due to the timing offset, we can take advantage of different expressions of the received signal and obtain channel estimations from the nonoverlapped and overlapped signals, respectively.

A. Nonoverlapped

Based on the nonoverlapped signal $\mathbf{r}_1^{n_t}$, channel parameter H can be estimated by a training-based method [3] since $s_1(t)$ is known to T_1. Therefore, we obtain the estimation of H as $\hat{H}_{\text{no}}$

$$\hat{H}_{\text{no}} = \frac{\sum_{i=1}^{n_t} r_{\text{async}}(i)s_1^*(i)}{An_t}. \tag{4}$$

Without loss of generality, the amplitude of s_1 and s_2 are assumed equal to A.

B. Overlapped

Overlapped samples in (2) can be expressed in vector form as

$$\mathbf{r}_{n_t+1}^N = H_1\mathbf{s}_1 + G\mathbf{s}_2 + H_2\mathbf{s}_3 + \mathbf{n} \tag{5}$$

where $H_1 \triangleq f_1(\tau)H$, $H_2 \triangleq f_2(\tau)H$, $\mathbf{s}_1 \triangleq [s_{1,n_t+1},\ldots,s_{1,N}]^T$, $\mathbf{s}_2 \triangleq [s_{2,1},\ldots,s_{2,N-n_t}]^T$, and $\mathbf{s}_3 \triangleq [s_{1,n_t+2},\ldots,s_{1,N},0]^T$. The noise term $\mathbf{n} \triangleq [n_{n_t+1},\ldots,n_N]^T$, where $n_i = Kh_2n_{1i}+n_{2i}$ for $i=1,\ldots,N+n_t$.

Assuming that $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$ are deterministic unknown vectors, the actual statistics of $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$ are considered in analyzing the behavior of the proposed algorithm theoretically in Section VII. The received vector $\mathbf{r}_{n_t+1}^N$ follows a complex Gaussian distribution with the expectation $\mathrm{E}\{\mathbf{r}_{n_t+1}^N\} = H_1\mathbf{s}_1 + G\mathbf{s}_2 + H_2\mathbf{s}_3$ and variance $\mathrm{Var}\{\mathbf{r}_{n_t+1}^N\} = \sigma^2\mathbf{I}$. The vector of unknown parameters $\Theta \triangleq [H_1,\,H_2,\,G,\,\sigma^2,\,\phi_{2,1},\ldots,\,\phi_{2,N-n_t}]^T$ can be estimated by maximizing the likelihood function of $\mathbf{r}_{n_t+1}^N$

$$\begin{aligned} &\Pr\left(\mathbf{r}_{n_t+1}^N;\Theta\right) \\ &= \frac{1}{\left(\pi\sigma^2\right)^{N-n_t}}\exp\left[-\frac{\left|\mathbf{r}_{n_t+1}^N - H_1\mathbf{s}_1 - G\mathbf{s}_2 - H_2\mathbf{s}_3\right|^2}{\sigma^2}\right] \end{aligned} \tag{6}$$

which can be simplified as the log-likelihood function (7) shown at the bottom of the page, where ϕ_g is the phase of G. Hereafter, $\hat{H}$, $\hat{H}_1$, $\hat{H}_2$, $\hat{G}$, and $\hat{\sigma}$ represent the estimated values of H, H_1, H_2, G, and σ, respectively.

With the knowledge of $\hat{n}_t$, the log-likelihood function (7) could be maximized if $|r_{\text{async}}(i) - H_1s_{1,\,i} - H_2s_{1,\,i+1} - A|G|e^{j(\phi_g+\phi_{2,\,i-n_t})}|$, $i=n_t+1,\ldots,N$ is minimized. Since $|r_{\text{async}}(i) - H_1s_{1,\,i} - H_2s_{1,\,i+1} - A|G|e^{j(\phi_g+\phi_{2,\,i-n_t})}|$ represents the *Euclidean distance* between $r_{\text{async}}(i) - H_1s_{1,\,i} - H_2s_{1,\,i+1}$ and $A|G|e^{j(\phi_g+\phi_{2,\,i-n_t})}$, the minimum distance could be achieved if (8) is met

$$\begin{aligned} &\angle\left\{r_{\text{async}}(i) - H_1s_{1,\,i} - H_2s_{1,\,i+1}\right\} \\ &= \angle\left\{A|G|e^{j\left(\phi_g+\phi_{2,\,i-n_t}\right)}\right\}, \qquad \forall\, i = n_t+1,\ldots,N. \end{aligned} \tag{8}$$

However, in general, (8) does not hold for all samples as the modulated phases ϕ_1 and $\phi_2 \in S_M$, and they cannot take continuous values in $[0, 2\pi)$. To derive a simple expression of $|G|$, we assume that ϕ_1 and ϕ_2 are continuously valued

$$L(\mathbf{r}_{n_t+1}^N;\Theta) = -(N-n_t)\log(\pi\sigma^2) - \frac{\sum_{i=n_t+1}^{N}\left|r_{\text{async}}(i) - H_1s_{1,\,i} - H_2s_{1,\,i+1} - A\,|G|\,e^{j\left(\phi_g+\phi_{2,\,i-n_t}\right)}\right|^2}{\sigma^2} \tag{7}$$

in $[0, 2\pi)$ [10], i.e., $M \to \infty$. Under the condition (8), (7) is approximated as

$$L\left(\mathbf{r}_{n_t+1}^{N};\Theta\right) = -(N-n_t)\log(\pi\sigma^2) - \frac{1}{\sigma^2} \times \sum_{i=n_t+1}^{N} \left(|r_{\text{async}}(i) - H_1 s_{1,i} - H_2 s_{1,i+1}| - A|G|\right)^2. \quad (9)$$

By maximizing (9), $\hat{G}$ is estimated as

$$|\hat{G}| = \frac{\left\|\mathbf{r}_{n_t+1}^{N} - H_1\mathbf{s}_1 - H_2\mathbf{s}_3\right\|}{A(N-n_t)},$$
$$\hat{\phi_g} = \angle\sum_{i=0}^{l-1}\left(r_{\text{async}}(i) - H_1 s_{1,i} - H_2 s_{1,i+1}\right)s_{2,i-n_t}^{*} \quad (10)$$

where l is the number of training symbols. As the channel is assumed fixed and flat over one frame, the channel fading and phase shift are the same for each symbol. In the proposed algorithms, we use only one training symbol to find $\hat{\phi}_g$, which will be applied to resolve the phase ambiguity [11] in the signal demodulation. It will be shown in Section VIII that one training symbol is sufficient to achieve a near optimal symbol error rate (SER).

Substituting (10) into (9), the log-likelihood function reduces to

$$L\left(\mathbf{r}_{n_t+1}^{N};\Theta\right) = -(N-n_t)\log(\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{i=n_t+1}^{N}\mathbb{F}_i(H_1,H_2) \quad (11)$$

where

$$\mathbb{F}_i(H_1,\,H_2) \triangleq \left(|r_{\text{async}}(i) - H_1 s_{1,i} - H_2 s_{1,i+1}| - \frac{\sum_{k=n_t+1}^{N}|r_{\text{async}}(k) - H_1 s_{1,k} - H_2 s_{1,k+1}|}{N-n_t}\right)^2.$$

Since (11) is differentiable with respect to σ^2, let $(\partial L(\mathbf{r}_{n_t+1}^{N};\,\Theta)/\partial\sigma^2) = 0$; thus, $\hat{\sigma}^2$ is obtained as

$$\hat{\sigma}^2 = \frac{\sum_{i=n_t+1}^{N}\mathbb{F}_i(H_1,H_2)}{N-n_t}. \quad (12)$$

Substituting (12) into (11), we obtain

$$\left[\hat{H},\hat{f}_1(\tau),\hat{f}_2(\tau)\right] = \underset{\alpha\in\mathbb{C},0<a<1,0<b<1}{\arg\min}\ \frac{\sum_{i=n_t+1}^{N}\mathbb{F}_i(a\alpha,\,b\alpha)}{N-n_t}. \quad (13)$$

Hence, $\hat{H}_1 = \hat{f}_1(\tau)\hat{H}$, and $\hat{H}_2 = \hat{f}_2(\tau)\hat{H}$. Equations (10) and (12) show that $\hat{G}$ and $\hat{\sigma}$ depend on $[\hat{H}_1,\,\hat{H}_2]$. Hence, the estimation of Θ can be obtained as long as $[\hat{H}_1,\,\hat{H}_2]$ is available. To solve (13), we propose the LCML and GLCML algorithms in Sections IV and V, respectively. In addition, the GLCML algorithm provides a criterion to select the estimation sample from overlapped and nonoverlapped signals, which produces the optimal channel estimation.

IV. Low-Complexity Maximum-Likelihood Channel Estimation in Synchronous Amplify-and-Forward Two-Way Relay Networks

To study the problem formulated in (13), we start from channel estimation in the synchronous system. In synchronous AF-TWRNs, all the signals are time synchronized. Hence, transmitted signals are fully aligned. In this case, the given derivations still hold when the symbol and frame offsets are zero ($\tau = n_t = 0$), $f_1(\tau) = 1$, and $f_2(\tau) = 0$ (see Appendix A). As a result, the received signal becomes

$$\mathbf{r} = H\mathbf{s}_1 + G\mathbf{s}_2 + \mathbf{n} \quad (14)$$

where $\mathbf{r} \triangleq [r_1,\ldots,r_N]^T$, $\mathbf{s}_1 \triangleq [s_{1,1},\ldots,s_{1,N}]^T$, $\mathbf{s}_2 \triangleq [s_{2,1},\ldots,s_{2,N}]^T$, and $\mathbf{n} \triangleq [n_1,\ldots,n_N]^T$. Here, the closed-form expression of H is derived as $\hat{H}_{\text{lcml}}^{\text{BPSK}}$ in BPSK ($M=2$) and $\hat{H}_{\text{lcml}}^{\text{MPSK}}$ in MPSK ($M>2$).

A. BPSK

Here, we propose the LCML channel estimation algorithm in the case of BPSK ($M=2$). The log-likelihood function is obtained by the maximum-likelihood estimation method as

$$L(\mathbf{r};\Theta) = -\frac{\sum_{i=1}^{N}|r_i - Hs_{1,i} - Gs_{2,i}|^2}{\sigma^2} - N\log(\pi\sigma^2). \quad (15)$$

By inspecting (15), the estimated parameters that satisfy the following condition:

$$r_i - Hs_{1,i} - Gs_{2,i} = 0, \quad \text{for} \quad i = 1,\ldots,N \quad (16)$$

will definitely maximize the log-likelihood function (15). Based on this fact, we make some approximations and propose an LCML channel estimation algorithm. In [10], an MCML estimator is proposed for BPSK. Different from the MCML estimator, the LCML estimator produces closed-form channel estimations.

Since $s_{2,i} = \pm 1$ in BPSK, we obtain $\hat{s}_{2,i} = \text{sgn}\{(r_i - Hs_{1,i})e^{-j\phi_g}\}$, where $\text{sgn}(x)$ denotes the signum function, and ϕ_g is the phase of G. Therefore, the signal detection of $s_{2,i}$ depends on the availability of $\hat{H}$ and $\hat{\phi}_g$. Then, we propose a convex optimization function to estimate channel parameters $\hat{H}$ and $\hat{\phi}_g$. As the estimated parameters satisfying condition (16) also make $(r_i - Hs_{1,i})^2 - (Gs_{2,i})^2 = 0$ hold for $i = 1,\ldots,N$, we eliminate $s_{2,i}$ in (15) and obtain the optimization function, i.e.,

$$L_{\text{sync}}^{\text{BPSK}}(\mathbf{r};\,\Theta) = -\frac{\sum_{i=1}^{N}\left|(r_i - Hs_{1,i})^2 - G^2\right|^2}{\sigma^2} - N\log(\pi\sigma^2). \quad (17)$$

As the proposed optimization function (17) is convex, we derive $\hat{G}$ and $\hat{\sigma}$ by letting the first partial derivative $(\partial L_{\text{sync}}^{\text{BPSK}}/\partial G^2) = 0$ and $(\partial L_{\text{sync}}^{\text{BPSK}}/\partial\sigma^2) = 0$, and we obtain

$$\hat{G}^2 = \frac{\sum_{i=1}^{N}(r_i - Hs_{1,i})^2}{N} \quad (18)$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N}\left|(r_i - Hs_{1,i})^2 - \frac{\sum_{k=1}^{N}(r_k - Hs_{1,k})^2}{N}\right|^2}{N}. \quad (19)$$

By substituting (18) and (19) into (17), $\hat{H}$ can be obtained by minimizing the following objective function:

$$F_{\text{sync}}^{\text{BPSK}}(H)=\frac{\sum_{i=1}^{N}\left|(r_i-Hs_{1,i})^2-\frac{\sum_{k=1}^{N}(r_k-Hs_{1,k})^2}{N}\right|^2}{N}. \quad (20)$$

Letting $(\partial F_{\text{sync}}^{\text{BPSK}}(H)/\partial H)=0$, the closed-form channel estimation is obtained as

$$\hat{H}_{\text{lcml}}^{\text{BPSK}}=\frac{\sum_{i=1}^{N}\left(r_i^2-\frac{\sum_{k=1}^{N}r_k^2}{N}\right)\left(r_i^*s_{1,i}^*-\frac{\sum_{k=1}^{N}r_k^*s_{1,k}^*}{N}\right)}{2\sum_{i=1}^{N}\left|r_is_{1,i}-\frac{\sum_{k=1}^{N}r_ks_{1,k}}{N}\right|^2}. \quad (21)$$

To estimate $\hat{\phi}_g$, we use l training symbols and obtain

$$\hat{\phi}_g^{\text{BPSK}}=\angle\sum_{i=1}^{l-1}(r_i-Hs_{1,i})s_{2i}^*. \quad (22)$$

B. MPSK($M>2$)

The optimization problem (13) is simplified as

$$\hat{H}=\arg\min_{\alpha\in\mathbb{C}}\frac{\sum_{i=1}^{N}\left(|r_i-\alpha s_{1,i}|-\frac{\sum_{k=1}^{N}|r_k-\alpha s_{1,k}|}{N}\right)^2}{N}. \quad (23)$$

In [10], a semi-blind channel estimation method, which is referred to as the DML channel estimator, is proposed to estimate H by solving (23). As an analytical solution to (23) is not available, the DML algorithm has to rely on numerical solutions by using optimization tools. As done in [13], the DML algorithm can be implemented by the steepest descent optimization method to obtain a solution. To achieve a closed-form channel estimator, we propose an LCML channel estimation algorithm, which is based on a convex optimization function to estimate nonreciprocal channels (channel coefficients h_1, h_2, g_1, and g_2 are i.i.d.) in synchronous AF-TWRNs.

By inspecting (23), we find it represents the variance of $|r_i-\alpha s_{1i}|$ if $N\to\infty$, and the presence of $|r_i-\alpha s_{1i}|$ makes (23) nonconvex. By replacing $|r_i-\alpha s_{1i}|$ by $|r_i-\alpha s_{1i}|^2$, we formulate a convex objective function to estimate H, and we denote $\hat{H}_{\text{lcml}}^{\text{MPSK}}$ as the channel estimation here, i.e.,

$$\hat{H}_{\text{lcml}}^{\text{MPSK}}=\arg\min_{\alpha\in\mathbb{C}}\text{f}(\mathbf{r};\alpha),$$
$$\text{f}(\mathbf{r};\alpha)=\frac{1}{N}\sum_{i=1}^{N}\left(|r_i-\alpha s_{1i}|^2-\frac{\sum_{k=1}^{N}|r_k-\alpha s_{1k}|^2}{N}\right)^2. \quad (24)$$

By letting the first partial derivative $(\partial f(\mathbf{r};\alpha)/\partial\Re\{\alpha\})=0$, where $\Re\{\alpha\}$ and $\Im\{\alpha\}$ denote real and imaginary parts of a complex number α, respectively, we get

$$\Re\left\{\hat{H}_{\text{lcml}}^{\text{MPSK}}\right\}=j\Im\left\{\hat{H}_{\text{lcml}}^{\text{MPSK}}\right\}\frac{(\mathbf{C}_2^T\mathbf{C}_2-\mathbf{C}_3^T\mathbf{C}_3)}{(\mathbf{C}_2+\mathbf{C}_3)^T(\mathbf{C}_2+\mathbf{C}_3)}-\frac{\mathbf{C}_1^T\mathbf{C}_2+\mathbf{C}_1^T\mathbf{C}_3}{(\mathbf{C}_2+\mathbf{C}_3)^T(\mathbf{C}_2+\mathbf{C}_3)} \quad (25)$$

where

$$\begin{cases}\mathbf{C}_1\triangleq[C_{1,1},\ldots,C_{1,N}]^T, & C_{1i}=|r_i|^2-\frac{\|\mathbf{r}\|^2}{N}\\ \mathbf{C}_2\triangleq[C_{2,1},\ldots,C_{2,N}]^T, & C_{2i}=\frac{\mathbf{s}_1^H\mathbf{r}}{N}-s_{1i}^*r_i\\ \mathbf{C}_3\triangleq[C_{3,1},\ldots,C_{3,N}]^T, & C_{3i}=\frac{\mathbf{s}_1^T\mathbf{r}^*}{N}-s_{1i}r_i^*,\\ & \forall i=1,\ldots,N.\end{cases}$$

Substituting (25) into (24), we derive a closed-form $\Im\{\hat{H}_{\text{lcml}}^{\text{MPSK}}\}$ from $(\partial f(\mathbf{r};\alpha)/\partial\Im\{\alpha\})=0$ as follows:

$$\Im\left\{\hat{H}_{\text{lcml}}^{\text{MPSK}}\right\}=\frac{\mathbf{C}_1^T\mathbf{C}_2\left(\mathbf{C}_3^T\mathbf{C}_3+\mathbf{C}_2^T\mathbf{C}_3\right)}{2j\left(\mathbf{C}_2^T\mathbf{C}_2\mathbf{C}_3^T\mathbf{C}_3-\left(\mathbf{C}_2^T\mathbf{C}_3\right)^2\right)}-\frac{\mathbf{C}_1^T\mathbf{C}_3\left(\mathbf{C}_2^T\mathbf{C}_2+\mathbf{C}_2^T\mathbf{C}_3\right)}{2j\left(\mathbf{C}_2^T\mathbf{C}_2\mathbf{C}_3^T\mathbf{C}_3-\left(\mathbf{C}_2^T\mathbf{C}_3\right)^2\right)}. \quad (26)$$

The channel estimation in synchronous AF-TWRNs is obtained as $\hat{H}_{\text{lcml}}^{\text{MPSK}}=\Re\{\hat{H}_{\text{lcml}}^{\text{MPSK}}\}+j\Im\{\hat{H}_{\text{lcml}}^{\text{MPSK}}\}$, where $j\triangleq\sqrt{-1}$.

V. Generalized Low-Complexity Maximum-Likelihood Algorithm in Asynchronous Amplify-and-Forward Two-Way Relay Networks

In the asynchronous system, the timing offset is nonzero, which results in partly aligned received signal, as shown in Fig. 2. The factors $f_1(\tau)$ and $f_2(\tau)$ are unknown, resulting in channel parameters H_1 and H_2 in (13), which needed to be estimated. Here, we extend the LCML algorithm to estimate channel parameters in the asynchronous system, by proposing a GLCML algorithm. In Section V-A, the estimation of H is obtained based on overlapped samples as $\hat{H}_o$. Then, a selection criterion is proposed in Section V-B to chose the optimal channel estimation from $\hat{H}_{\text{no}}$ and $\hat{H}_o$.

A. Case: $n_t\neq 0$, $\tau\neq 0$

To achieve channel estimation from the optimization function (13), we make a similar approximation as in Section IV by replacing $|r_{\text{async}}(i)-H_1s_{1,i}-H_2s_{1,i+1}|$ as $|r_{\text{async}}(i)-H_1s_{1,i}-H_2s_{1,i+1}|^2$. Then, (13) is updated as

$$\left[\hat{H}_o,\,\hat{f}_1(\tau),\,\hat{f}_2(\tau)\right]=\arg\min_{\alpha\in\mathbb{C},\,0<a<1,\,0<b<1}L\left(\mathbf{r}_{n_t+1}^N;\,\alpha,\,a,\,b\right)$$
$$L\left(\mathbf{r}_{n_t+1}^N;\alpha,a,b\right)=\frac{1}{N-n_t}\times\sum_{i=n_t+1}^{N}\left(|r_{\text{async}}(i)-a\alpha s_{1,i}-b\alpha s_{1,i+1}|^2-\frac{\sum_{k=n_t+1}^{N}|r_{\text{async}}(k)-a\alpha s_{1,k}-b\alpha s_{1,k+1}|^2}{N-n_t}\right)^2. \quad (27)$$

Note that we get another estimation of parameter H, which is denoted $\hat{H}_o$. Then, we obtain $\hat{H}_1=\hat{f}_1(\tau)\hat{H}_o$ and $\hat{H}_2=\hat{f}_2(\tau)\hat{H}_o$.

B. Estimation Sample Selection Criterion

In the asynchronous system, where $n_t \neq 0$ and $\tau \neq 0$, we obtain two estimations of channel parameter H from the overlapped signal as $\hat{H}_o$ and the nonoverlapped signal as $\hat{H}_{\rm no}$. The channel estimation $\hat{H}$ is selected from $\hat{H}_o$ and $\hat{H}_{\rm no}$ as the one with the minimum MSE by the ESSC. By taking into account the statistics of $\mathbf{s}_1$, the estimation MSE of $\hat{H}_{\rm no}$ is derived as

$$\text{MSE}_{\hat{H}_{\rm no}} = \frac{\sigma^2}{A^2 n_t} \tag{28}$$

under the condition that symbol offset $\tau = 0$, $\hat{H}_o = \hat{H}_{\rm lcml}^{\rm BPSK}$ $(M = 2)$, and $\hat{H}_o = \hat{H}_{\rm lcml}^{\rm MPSK}$ $(M > 2)$. The estimation MSE of $\hat{H}_o$ is approximated as (50) and (51) in Section VII.

Comparing (28) and (50) and (51), we find that the frame length N and the frame offset n_t affect the MSE performance of the two channel estimators. Let us define $n_s \triangleq N - n_t$; thus, we get the ESSC as follows:

$$\begin{cases} n_t > n_s, & M = 2 \\ n_t > \frac{n_s^4 - 5n_s^3 + 8n_s^2 - 4n_s}{2n_s^3 - 3n_s^2 + n_s - 1}, & M > 2. \end{cases} \tag{29}$$

If the ESSC (29) holds, which is equivalent to $\text{MSE}_{\hat{H}_{\rm no}} < \text{MSE}_{\hat{H}_o}$, then we use the nonoverlapped samples $\mathbf{r}_1^{n_t}$ to estimate the channel parameters as (4), i.e., $\hat{H} = \hat{H}_{\rm no}$. Otherwise, the overlapped samples are selected to obtain the channel estimations by the GLCML algorithm as $\hat{H} = \hat{H}_o$.

VI. Joint Synchronization and Channel Estimation Algorithm

Here, we propose a JSCE algorithm, which enables T_1 to synchronize with either $s_1(t)$ or $s_2(t)$. The basic idea is as follows. As $s_1(t)$ is the self-transmitted signal of source node T_1, T_1 can synchronize with $s_1(t)$ by a cross-correlation method [14] even without the knowledge of the timing offset. After synchronizing with $s_1(t)$, T_1 is able to synchronize with $s_2(t)$ if the knowledge of the relative timing offset between source nodes T_1 and T_2 is available. We propose the JSCE algorithm, which enables source node T_1 to estimate the relative timing offset, i.e., frame offset n_t and symbol offset τ. As a result, T_1 can synchronize with either $s_1(t)$ or $s_2(t)$.

First, source node T_1 tries to synchronize with its own transmitted signal $s_1(t)$ via the cross-correlation method. It calculates the cross correlation of the received sample $r_1(t)$ and the self-transmitted signal $s_1(t)$. Then, it filters $r_1(t)$ with a matched filter $f'(t) = f(T - t)$ and samples it at every T period. The frame length of the resulting signal samples is $N + n_t + 1$ due to the frame offset. As T_1 does not know n_t and $0 < n_t < N$, T_1 obtains $2N$ samples to detect the frame offset since. The received sample at T_1 is given in (30), shown at the bottom of the page, where $G_1 = f_1(\tau)G$, and $G_2 = f_2(\tau)G$. $f_1(\tau), f_2(\tau) < 1$ regardless of the type of filters (see Appendix A). Equation (30) implies that the timing offsets τ_1 and τ_2 between the source nodes and relay node are not needed for synchronization and channel estimation. The received discrete signal vector at T_1 is denoted $\mathbf{r}_t \triangleq [r_t(1), \ldots, r_t(2N)]^T$. The overlapped part is denoted $\mathbf{r}_{t\,n_t+1}^{\,N}$, and the nonoverlapped parts of $\mathbf{s}_1$ and $\mathbf{s}_2$ are represented by $\mathbf{r}_{t\,1}^{\,n_t}$ and $\mathbf{r}_{t\,N+1}^{\,2N}$, respectively.

As $G_1 = f_1(\tau)G$ and $G_2 = f_2(\tau)G$, we estimate symbol offset τ as

$$\hat{\tau} = \arg\min_{0<\mu<T} \left| \frac{f_1(\mu)}{f_2(\mu)} - \frac{\hat{G}_1}{\hat{G}_2} \right| \tag{31}$$

if $s_{2,1}$ and $s_{2,N}$ are the pilot symbol known to both the source nodes. We obtain the estimation of G_1 and G_2 from (30) as

$$\begin{aligned} \hat{G}_1 &= \left(r_t(n_t + 1) - H s_{1,n_t+1}\right) s_{2,1}^* \\ \hat{G}_2 &= r_t(N + n_t + 1) s_{2,N}^*. \end{aligned} \tag{32}$$

Substituting (32) into (31) yields

$$\hat{\tau} = \arg\min_{0<\mu<T} \left| \frac{f_1(\mu)}{f_2(\mu)} - \frac{\left(r_t(n_t + 1) - H s_{1,n_t+1}\right) s_{2,1}^*}{r_t(N + n_t + 1) s_{2,N}^*} \right| \tag{33}$$

which shows that $\hat{\tau}$ depends on the availability of the estimation of H. Note that (33) is applicable to any type of filters. Based on (30), we propose the FACE algorithm in Section VI-A to estimate the integral timing offset, i.e., the frame offset n_t. In the case that frame offset $n_t \neq 0$, the nonoverlapped signal $\mathbf{r}_{t\,1}^{\,n_t}$ can be used to derive one estimation of H as $\hat{H}_f$ (34). Then, $\hat{\tau}$ is the obtained from (33) based on $\hat{H}_f$ (34). On the other hand, in the case that frame offset $n_t = 0$, the SACE algorithm is proposed in Section VI-B to estimate channel H as (41) and (43) based on the overlapped part $\mathbf{r}_{t\,n_t+1}^{\,N}$ of the received sample. As a result, the fractional timing offset, i.e., the symbol offset τ, is obtained from (33).

A. Frame Synchronization Algorithm

The frame synchronization algorithm is described in Table I. In Stage 1, the cross correlation of the received sample $r_1(t)$ and transmitted signal $s_1(t)$ is calculated, and $s_1(t)$ begins when $\text{xcorr}(r_1(t), s_1(t))$ defined in Table I reaches its maximum. Under the assumption $N \to \infty$, we get $\max[\text{xcorr}(r_1(t), s_1(t)] = N|H|^2A^2$. In Stage 2, the total power of every N received samples is denoted P_R, as defined

$$r_t(i) = \begin{cases} Hs_{1,i} + Kh_2 n_{1i} + n_{2i}, & i = 1, \ldots, n_t \\ Hs_{1,i} + G_1 s_{2,i-n_t} + G_2 s_{2,i-n_t-1} + Kh_2 n_{1i} + n_{2i}, & i = n_t + 1, \ldots, N \\ G_1 s_{2,i-n_t} + G_2 s_{2,i-n_t-1} + Kh_2 n_{1i} + n_{2i}, & i = N + 1, \ldots, N + n_t + 1 \\ n_{2i}, & i = N + n_t + 2, \ldots, 2N \end{cases} \tag{30}$$

TABLE I
FRAME SYNCHRONIZATION ALGORITHM

Stage 1-Locate the beginning of $\mathbf{s}_1$
$xcorr(r_1(t), s_1(t))(x) = \lvert \int_{-\infty}^{\infty} r_1(x), s_1^*(t+x)dx \rvert$
$L_1 = \underset{x}{\operatorname{argmax}}\ xcorr(r_1(t), s_1(t))(x),$
Stage 2-Locate the beginning of $\mathbf{s}_2$
$P_R(i) = \sum_{j=0}^{N-1} \lvert r_t(i+j)\rvert^2 f(j+1), \quad f(i) = 1, i = 1, ..., N$
$a = [L_1, P_R(L_1)], \quad c = [L_1 + N, P_R(L_1 + N)],$
$b = \underset{a_x \le p_x \le c_x}{\operatorname{argmax}} \ D(p, a, c) \quad L_2 = b_x.$

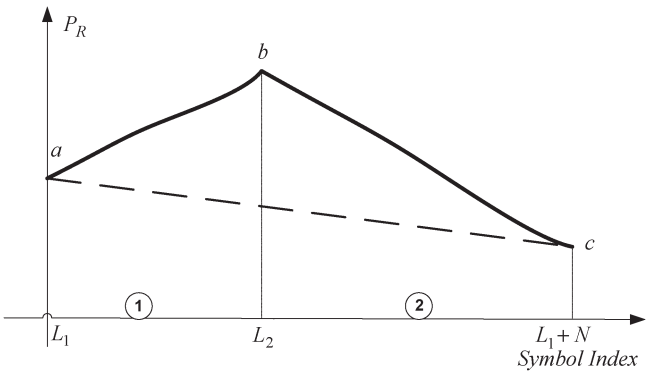


Fig. 3. Total power of N signals.

in Table I. If $N \to \infty$ and SNR $\to \infty$, the curve of P_R is as shown in Fig. 3, and its slope is represented as

$$\mathbf{slope} = \left[\left(|G_1|^2 + |G_2|^2 - |H|^2\right)A^2, -\left(|G_1|^2 + |G_2|^2 + |H|^2\right)A^2\right]$$

where $\mathbf{slope}(i)$ denotes the slope of P_R at intervals i and represents the rate at which signal power changes, where $i = 1, 2$. By analyzing the pattern of P_R, we derive the frame offset n_t. The beginning of $s_1(t)$ has been found in Stage 1, whose x-coordinate is L_1. Without loss of generality, we assume $n_t > 0$, and a is assumed the beginning of $s_1(t)$. c is defined as the point whose x-coordinate is $c = L_1 + N$ on curve P_R. As shown in Fig. 3, among all the points between a and c on curve P_R, the longest distance to the line ac occurs at b, which is the beginning of $\mathbf{s}_2$, whose x-coordinate is L_2. As a result, b is found by optimizing $D(x, y, z)$, which calculates the distance from point x to line yz. We obtain the frame offset estimation $\hat{n}_t = L_2 - L_1$.

The asymptotic performance of the frame synchronization algorithm is analyzed under the condition $N \to \infty$ and SNR $\to \infty$. In Stage 1, detection error occurs when $|H|$ is sufficiently small so that the maximum cross correlation could not be differentiated. In Stage 2, we utilize the P_R slope differences to locate the beginning of $s_2(t)$. When $|G_1|^2 + |G_2|^2$ is sufficiently small, there will be no slope difference between two intervals. Hence, detection error occurs if $|G| \to 0$. Since $H \triangleq Kh_1h_2$ and $G \triangleq Kg_1h_2$, $|H|$ and $|G|$ are i.i.d. with the cumulative probability function $\mathcal{F}(u) = 1 - e^{(-u/\sigma_c^2)}$, $u > 0$. Obviously, the probability that $|H|$ or that $|G|$ is arbitrarily small approaches zero. Therefore, the error probability of frame synchronization algorithm approaches zero under the condition of large N and SNR.

Based on the estimation of the frame offset $\hat{n}_t$, we propose the FACE algorithm, in which the channel parameter H can be estimated in a similar way as in Section III-A using the nonoverlapped signal $\mathbf{r}_{t1}^{n_t}$. Therefore, we obtain the estimation of H as $\hat{H}_f$ in the following:

$$\hat{H}_f = \frac{\sum_{i=1}^{n_t} r_t(i)s_1^*(i)}{A\hat{n}_t}, \quad \hat{n}_t \neq 0. \tag{34}$$

B. Symbol Synchronization Algorithm

If the frame offset $n_t = 0$, the vector of unknown parameters $\Theta_s \triangleq [H, \sigma^2, \phi_{2,1}, \ldots, \phi_{2,N-n_t}]$ can be estimated by maximizing the log-likelihood function of $\mathbf{r}_{t\,n_t+1}^{N}$, i.e.,

$$L\left(\mathbf{r}_{t\,n_t+1}^{N}; \Theta_s\right) = -(N - n_t)\log(\pi\sigma^2) - \frac{\sum_{i=n_t+1}^{N} \left|r_t(i) - Hs_{1,i} - G_1 s_{2,i-n_t} - G_2 s_{2,i-n_t-1}\right|^2}{\sigma^2}. \tag{35}$$

By inspecting (35), the estimated parameters that satisfy the following condition:

$$r_t(i) - Hs_{1,i} - G_1 s_{2,i-n_t} - G_2 s_{2,i-n_t-1} = 0, \quad \text{for} \quad i = n_t + 1, \ldots, N \tag{36}$$

will definitely maximize the log-likelihood function (35). Based on this fact, we make some approximations and propose the SACE algorithm.

1) BPSK: Condition (36) implies that the estimated parameters also satisfy $(r_t(i) - Hs_{1,i})^2 - (G_1 s_{2,i-n_t} + G_2 s_{2,i-n_t-1})^2 = 0$ for $i = n_t + 1, \ldots, N$. As $s_{2,i} = \pm 1$ in BPSK, the log-likelihood function (35) can be approximated as (37), shown at the bottom of the page. If channel parameters satisfy

$$(r_t(i) - Hs_{1,i})^2 - G_1^2 - G_2^2 - 2G_1G_2 s_{2,i-n_t} s_{2,i-n_t-1} = 0, \quad i = n_t + 1, \ldots, N \tag{38}$$

$$\begin{aligned} L\left(\mathbf{r}_{t\,n_t+1}^{N}; \Theta_s\right) &= -(N - n_t)\log(\pi\sigma^2) - \frac{\sum_{i=n_t+1}^{N} \left|(r_t(i) - Hs_{1,i})^2 - (G_1 s_{2,i-n_t} + G_2 s_{2,i-n_t-1})^2\right|^2}{\sigma_s^2} \\ &= -(N - n_t)\log(\pi\sigma^2) - \frac{\sum_{i=n_t+1}^{N} \left|(r_t(i) - Hs_{1,i})^2 - G_1^2 - G_2^2 - 2G_1G_2 s_{2,i-n_t} s_{2,i-n_t-1}\right|^2}{\sigma^2} \end{aligned} \tag{37}$$

then the approximated log-likelihood function (37) is maximized. As condition (38) suggests that channel parameters also make $\angle((r_t(i)-Hs_{1,i})^2-G_1^2-G_2^2)=\angle(2G_1G_2s_{2,i-n_t}s_{2,i-n_t-1})$ hold for $i=n_t+1,\ldots,N$, we further approximate (37) as

$$L\left(\mathbf{r}_{t\,n_t+1}^{N};\Theta_s\right)=-(N-n_t)\log(\pi\sigma^2)-\frac{\sum_{i=n_t+1}^{N}\left(\left|(r_t(i)-Hs_{1,i})^2-G_1^2-G_2^2\right|-2|G_1G_2|\right)^2}{\sigma^2}. \quad (39)$$

After two steps of approximations, we eliminate unknown signal $s_{2,i}$ from the objective function. Therefore, the semi-blind channel estimation is made possible by these approximations. Maximizing the given objective function (39) in terms of σ^2 yields

$$\hat{\sigma}_s^2=\frac{\sum_{i=n_t+1}^{N}\left(\left|(r_t(i)-Hs_{1,i})^2-G_1^2-G_2^2\right|-2|G_1G_2|\right)^2}{N-n_t}. \quad (40)$$

By substituting (32) and (40) into (39), we obtain $\hat{H}_s$ as

$$\hat{H}_s^{\mathrm{BPSK}}=\arg\min_{\alpha\in\mathbb{C}}\times\frac{\sum_{i=n_t+1}^{N}\left(\left|(r_t(i)-\alpha s_{1,i})^2-\hat{G}_1^2-\hat{G}_2^2\right|-2|\hat{G}_1\hat{G}_2|\right)^2}{A(N-n_t)}. \quad (41)$$

2) MPSK($M>2$): Here, we propose the SACE algorithm for MPSK, where modulation order $M>2$. Condition (36) implies that the estimated parameters also make $\angle(r_t(i)-Hs_{1,i})-\angle(G_1s_{2,i-n_t}+G_2s_{2,i-n_t-1})=0$ hold for $i=n_t+1,\ldots,N$. Therefore, (35) is approximated as

$$L\left(\mathbf{r}_{t\,n_t+1}^{N};\Theta_s\right)=-(N-n_t)\log(\pi\sigma^2)-\frac{\sum_{i=n_t+1}^{N}\left(|r_t(i)-Hs_{1,i}|^2-|G_1e^{j\phi_{2,i-n_t}}+G_2e^{j\phi_{2,i-n_t-1}}|^2\right)^2}{\sigma^2}.$$

Condition (36) also suggests that the estimated parameters satisfy $|r_t(i)-Hs_{1,i}|^2=|G_1s_{2,i-n_t}+G_2s_{2,i-n_t-1}|^2$ for $i=n_t+1,\ldots,N$. We make another approximation $|G_1s_{2,i-n_t}+G_2s_{2,i-n_t-1}|^2=(\sum_{i=n_t+1}^{N}|r_t(i)-Hs_{1,i}|^2/N-n_t)$ and obtain

$$L\left(\mathbf{r}_{t\,n_t+1}^{N};\Theta_s\right)=-(N-n_t)\log(\pi\sigma^2)-\frac{\sum_{i=n_t+1}^{N}\left(|r_t(i)-Hs_{1,i}|^2-\frac{\sum_{k=n_t+1}^{N}|r_t(k)-Hs_{1,k}|^2}{N-n_t}\right)^2}{\sigma^2}. \quad (42)$$

Maximizing the approximate objective function (42), we get $\hat{H}_s^{\mathrm{MPSK}}=\Re\{\hat{H}_s^{\mathrm{MPSK}}\}+j\Im\{\hat{H}_s^{\mathrm{MPSK}}\}$, where $\Re\{\hat{H}_s^{\mathrm{MPSK}}\}$ and $\Im\{\hat{H}_s^{\mathrm{MPSK}}\}$ denote the real and imaginary parts of the complex number $\hat{H}_s^{\mathrm{MPSK}}$, respectively, i.e.,

$$\Im\left\{\hat{H}_s^{\mathrm{MPSK}}\right\}=\frac{\mathbf{C}_1^T\mathbf{C}_2\left(\mathbf{C}_3^T\mathbf{C}_3+\mathbf{C}_2^T\mathbf{C}_3\right)}{2j\left(\mathbf{C}_2^T\mathbf{C}_2\mathbf{C}_3^T\mathbf{C}_3-\left(\mathbf{C}_2^T\mathbf{C}_3\right)^2\right)}-\frac{\mathbf{C}_1^T\mathbf{C}_3\left(\mathbf{C}_2^T\mathbf{C}_2+\mathbf{C}_2^T\mathbf{C}_3\right)}{2j\left(\mathbf{C}_2^T\mathbf{C}_2\mathbf{C}_3^T\mathbf{C}_3-\left(\mathbf{C}_2^T\mathbf{C}_3\right)^2\right)}$$

$$\Re\left\{\hat{H}_s^{\mathrm{MPSK}}\right\}=j\Im\{\hat{H}_{\mathrm{MPSK}}\}\frac{\left(\mathbf{C}_2^T\mathbf{C}_2-\mathbf{C}_3^T\mathbf{C}_3\right)}{(\mathbf{C}_2+\mathbf{C}_3)^T(\mathbf{C}_2+\mathbf{C}_3)}-\frac{\mathbf{C}_1^T\mathbf{C}_2+\mathbf{C}_1^T\mathbf{C}_3}{(\mathbf{C}_2+\mathbf{C}_3)^T(\mathbf{C}_2+\mathbf{C}_3)}. \quad (43)$$

where

$$\begin{cases}\mathbf{C}_1\triangleq[C_{1,n_t+1},\ldots,C_{1,N}]^T\\ \mathbf{C}_2\triangleq[C_{2,n_t+1},\ldots,C_{2,N}]^T\\ \mathbf{C}_3\triangleq[C_{3,n_t+1},\ldots,C_{3,N}]^T\quad\forall\, i=n_t+1,\ldots,N.\\ C_{1i}=|r_t(i)|^2-\frac{\sum_{k=n_t+1}^{N}|r_t(k)|^2}{N-n_t}\\ C_{2i}=\frac{\sum_{k=n_t+1}^{N}s_{1k}^*r_t(k)}{N-n_t}-s_{1i}^*r_t(i)\\ C_{3i}=\frac{\sum_{k=n_t+1}^{N}s_{1k}r_t^*(k)}{N-n_t}-s_{1i}r_t^*(i).\end{cases}$$

VII. Analysis of Low-Complexity Maximum-Likelihood Channel Estimation Algorithm

Here, we will analyze the computational complexity of the LCML algorithm and its behaviors in the cases of high SNR and large frame length, as well as its MSE performance.

A. Computational Complexity Analysis

To achieve the objective of the signal demodulation of $\mathbf{s}_2$ from (14) in synchronous AF-TWRNs, we first cancel the self-interference term $H\mathbf{s}_1$ in (14) with the channel state information of $\hat{H}_{\mathrm{lcml}}^{\mathrm{MPSK}}$ derived in (25) and (26), and the transmitted signal $\mathbf{s}_1$. Then, $\hat{\phi}_g$ is obtained from (10) with the help of only one training symbol, so that the phase ambiguity in MPSK demodulation could be resolved. Due to the phase modulation, $|\hat{G}|$ and $\hat{\sigma}^2$ are not needed for the signal demodulation [11]. Therefore, the demodulation of $\mathbf{s}_2$ depends only on the availability of $\hat{H}_{\mathrm{lcml}}^{\mathrm{MPSK}}$ and the training symbol. Equations (25) and (26) indicate that the calculation of the closed-form expression of channel estimation involves summation and multiplication operations. The computational complexity is considered in terms of the number of samples N. The summation operation leads to computational complexity of $\mathcal{O}(N)$, and the multiplication operation results in constant complexity. As a result, the computational complexity of the LCML estimation algorithm is $\mathcal{O}(N)$ in the case of $M>2$. Similarly, from (21), we obtain that the computational complexity of the LCML estimation algorithm as $\mathcal{O}(N)$ in the case of $M=2$.

Since the signal demodulation only depends on the availability of $\hat{H}_{\text{lcml}}^{\text{MPSK}}$ and the training symbol, the following analysis focuses on the channel estimation $\hat{H}_{\text{lcml}}^{\text{MPSK}}$. Hereafter, we denote $\hat{H}_{\text{lcml}}^{\text{MPSK}}$ as $\hat{H}$ for simplicity.

B. Large Sample Size

In the large-frame-length scenario, we will prove that, when the channel parameter spaces of H and G are restricted to compact sets [12], the LCML channel estimator is consistent [12]. In other words, if channel parameters H and G are bounded, $\hat{H}$ converges in probability to the real channel value H as the frame length N becomes larger.

For simplicity, we define the estimation error $v \triangleq H - \hat{H}$ and $V_N(v) \triangleq f(\mathbf{r}; H - v)$. Then, (24) is expressed as

$$V_N(v) = \frac{1}{N}\sum_{i=1}^{N}\left(|z_i(v)|^2 - \frac{\sum_{k=1}^{N}|z_k(v)|^2}{N}\right)^2 \tag{44}$$

where $z_i(v) \triangleq vs_{1i} + Gs_{2i} + n_i, i = 1, \ldots, N$. In (44), $V_N(v)$ represents the sample variance of random variable $|z_i(v)|^2$. Since the channel estimator (44) belongs to the class of extremum estimators [15], we use the following fundamental lemma for the consistency of extremum estimators [15], to prove the consistency of the proposed LCML channel estimator.

Lemma 1: If v belongs to a compact set Ω and $V_N(v)$ converges uniformly to $F(v)$, where $F(v)$ is continuous and uniquely minimized at $v = v_o$, then $\hat{v}$ converges in probability to v_o, where $\hat{v} = \arg\min_{v\in\Omega} V_N(v)$.

In (44), signal terms s_{1i} and s_{2i} and noise term n_i are i.i.d. for each index $i = 1, \ldots, N$. As a result, $z_i(v)$ is i.i.d., and we define $V(v)$ as the true variance of $|z_i(v)|^2$. $V(v)$ is considered the function $F(v)$ in Lemma 1. If Conditions 1.1, 1.2, and 1.3 are satisfied:

Conditions:

1) the channel parameter $H, G \in$ compact set Ω;
2) the optimization function $V_N(v)$ converges uniformly [16] to $V(v)$;
3) $V(v)$ is continuous and has a unique global minimum at $v_o = 0$;

then Lemma 1 could be applied to the LCML channel estimator, which implies that $\hat{v}$ converges in probability to v_o. Condition 1.3 shows that estimation error $v_o = 0$. Then, we could conclude that $\hat{H}$ converges in probability to the real channel value H (see Theorem 1), i.e., the LCML channel estimator is consistent.

Theorem 1: If H and $G \in \Omega$, the following channel estimator:

$$\hat{H} = \arg\min_{u\in\mathbb{C}} \frac{1}{N}\sum_{i=1}^{N}\left(|r_i - us_{1i}|^2 - \frac{\|\mathbf{r} - u\mathbf{s}_1\|^2}{N}\right)^2$$

is consistent.

Proof: See Appendix E. ■

The consistency of the LCML channel estimator suggests that, in each estimation process, the estimation becomes more accurate with larger frame length N.

C. High SNR

Theorem 2: For a fixed finite frame length N, the proposed LCML channel estimator approaches the true channel as SNR $\to \infty$ with probability $1 - (1/2^{N-1})$ for BPSK $(M = 2)$ and $1 - (M-1)(2/M)^{N-1}$ for MPSK $(M > 2)$.

Proof: It has been proven in [10, App. D] that the DML algorithm approaches the true channel with the probability $1 - (2/M)^{N-1}(M-1)$ in the scenario of high SNR. As the proposed LCML algorithm for MPSK $(M > 2)$ behaves exactly the same as the DML estimator, we use the conclusion of [10] and obtain that the LCML algorithm for MPSK $(M > 2)$ approaches the true channel with the probability $1 - (2/M)^{N-1}(M-1)$ in the scenario of high SNR. The LCML algorithm for BPSK $(M = 2)$ is different from that for MPSK $(M > 2)$. Hence, we analyze the behavior of the LCML channel estimator for BPSK $(M = 2)$ in high SNR scenarios in Appendix F and conclude that the LCML algorithm for BPSK approaches the true channel with the probability $1 - (1/2^{N-1})$ in the scenario of a high SNR. ■

D. MSE Performance

Here, the MSE performance of the proposed LCML channel estimator will be assessed analytically. We derive the expression for the MSE in terms of the SNR and frame length N. The definition of the channel estimation MSE is

$$\begin{aligned}
\text{MSE}_{\hat{H}} &= E\left\{|\hat{H} - H|^2\right\} = \text{MSE}_{\Re\{\hat{H}\}} + \text{MSE}_{\Im\{\hat{H}\}} \\
\text{MSE}_{\Re\{\hat{H}\}} &= E\left\{\left(\Re\{\hat{H}\} - \Re\{H\}\right)^2\right\} \\
\text{MSE}_{\Im\{\hat{H}\}} &= E\left\{\left(\Im\{\hat{H}\} - \Im\{H\}\right)^2\right\}.
\end{aligned} \tag{45}$$

We begin with the calculation of $\text{MSE}_{\Im\{\hat{H}\}}$. By expanding (26), we obtain $\Im\{\hat{H}\} = f(x, y) \triangleq (x/y)$, where

$$\begin{cases}
x = \displaystyle\sum_{i=1}^{N} C_{1i}C_{2i}\left(\sum_{i=1}^{N} C_{3i}^2 + \sum_{i=1}^{N} C_{2i}C_{3i}\right) \\
\qquad - \displaystyle\sum_{i=1}^{N} C_{1i}C_{3i}\left(\sum_{i=1}^{N} C_{2i}^2 + \sum_{i=1}^{N} C_{2i}C_{3i}\right) \\
y = 2j\left(\displaystyle\sum_{i=1}^{N} C_{2i}^2 \sum_{i=1}^{N} C_{3i}^2 - \left(\sum_{i=1}^{N} C_{2i}C_{3i}\right)^2\right).
\end{cases} \tag{46}$$

Equation (46) shows that x and y are the summations of N terms. According to the *central limit theorem* [16], x and y are asymptotically complex normal distributed random variables with expectations μ_x and μ_y and variances σ_x^2 and σ_y^2, respectively.

For the simplicity of $\text{MSE}_{\Im\{\hat{H}\}}$ calculation, we approximate $\Im\{\hat{H}\}$ as its first-degree Taylor polynomial representation [17] $(\mu_x/\mu_y) + (x/\mu_y) - (\mu_x y/\mu_y^2)$ at the point (μ_x, μ_y) (see Lemma 2).

Lemma 2: If SNR $\to \infty$, $f(x, y) = (x/y)$ equals its first-degree Taylor polynomial approximation at the point (μ_x, μ_y), namely

$$f(x, y) = \frac{\mu_x}{\mu_y} + \frac{x}{\mu_y} - \frac{\mu_x y}{\mu_y^2}.$$

Proof: See Appendix G. ■

As proven in Lemma 2, under the condition SNR $\to \infty$, we have

$$\Im\{\hat{H}\} = \frac{\mu_x}{\mu_y} + \frac{x}{\mu_y} - \frac{\mu_x y}{\mu_y^2}. \tag{47}$$

After calculating the expectation of (47), we obtain the first moment of $\Im\{\hat{H}\}$ as

$$E\left\{\Im(\hat{H})\right\} = \frac{\mu_x}{\mu_y}. \tag{48}$$

The expectations of x and y are calculated from (66), and we obtain $(\mu_x/\mu_y) = \Im\{H\}$. That is, $E\{\Im(\hat{H})\} = \Im\{H\}$. Similarly, we derive $E\{\Re(\hat{H})\} = \Re\{H\}$. Therefore, $E\{\hat{H}\} = H$, which suggests that the LCML channel estimator is unbiased if SNR $\to \infty$, and this helps to proves the following theorem.

Theorem 3: If SNR $\to \infty$, the following channel estimator:

$$\hat{H} = \arg\min_{u\in\mathbb{C}} \frac{1}{N}\sum_{i=1}^{N}\left(|r_i - us_{1i}|^2 - \frac{\|\mathbf{r} - u\mathbf{s}_1\|^2}{N}\right)^2$$

is unbiased.

It has been proven that $E\{\Im(\hat{H})\} = \Im\{H\}$ on the condition SNR $\to \infty$. Replacing $\Im\{H\}$ with $E\{\Im(\hat{H})\}$ in (45), we obtain $\text{MSE}_{\Im\{\hat{H}\}}$ as

$$\text{MSE}_{\Im\{\hat{H}\}} = E\left\{\left[\Im(\hat{H}) - E\left\{\Im(\hat{H})\right\}\right]^2\right\}. \tag{49}$$

By substituting (47) and (48) into (49), $\text{MSE}_{\Im\{\hat{H}\}}$ is expressed as

$$\text{MSE}_{\Im\{\hat{H}\}} = \frac{E\{x^2\}}{\mu_y^2} - 2\frac{\mu_x}{\mu_y}\frac{E\{xy\}}{\mu_y^2} + \left(\frac{\mu_x}{\mu_y}\right)^2\frac{E\{y^2\}}{\mu_y^2}.$$

Likewise, we derive $\text{MSE}_{\Re\{\hat{H}\}}$ and then obtain the expression of $\text{MSE}_{\hat{H}}$ by analyzing the moments of x and y from (66), i.e.,

$$\text{MSE}_{\hat{H}} = \frac{\sigma^2}{A^2N}\kappa \tag{50}$$

where

$$\kappa = \frac{4A^2|G|^2(N^4-8N^3+8N^2-4N+1)\left(\sigma^4+A^2|G|^2\right)^2}{(N^2-3N+2)^2(|G|^2A^2+\sigma^2)^4}$$

$$+\frac{2(N-1)(N^3-3N^2+N-1)\left(|G|^4A^4+\sigma^4\right)\left(\sigma^4+A^2|G|^2\right)^2}{(N^2-3N+2)^2(|G|^2A^2+\sigma^2)^4}.$$

Equation (50) shows $\kappa \propto 2 + (\sigma^2/A^2N)$ and $\text{MSE}_{\hat{H}} \propto (2\sigma^2/A^2N)$, which implies that the LCML channel estimations approach the true channel parameter values in a scenario with either a high SNR or a large frame length. Furthermore, we derive the MSE expression of the LCML algorithm in the case of BPSK $(M = 2)$ from the closed-form channel estimation (21) as

$$\text{MSE}_{H_{\text{lcml}}^{\text{BPSK}}} = \frac{\sigma^2}{A^2N}. \tag{51}$$

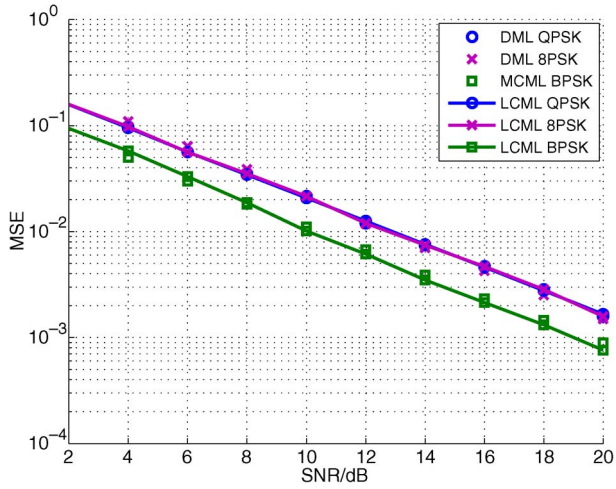


Fig. 4. Comparison of MSE performance of the LCML, MCML, and DML channel estimators versus SNR for $N = 45$.

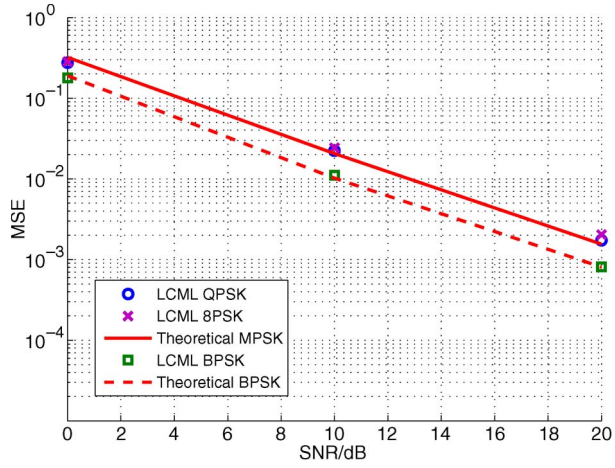


Fig. 5. MSE performance of the LCML channel estimator versus SNR for $N = 45$.

VIII. Simulation Results

Here, we will evaluate the performance of the proposed LCML, GLCML, and JSCE algorithms numerically by Monte Carlo simulations over flat-fading channels. In the simulations, we employ MPSK signal modulation and assume $P_r = P_1 = P_2 = P$. The SNR is defined as (P/σ_n^2), where σ_n^2 denotes the AWGN power. The channel parameters h_1, h_2, g_1, and g_2 are modeled as i.i.d. in $\mathcal{CN}(0, 1)$ and remain fixed during one frame.

We begin by comparing the MSE performance of the DML and MCML algorithms with the LCML channel estimators. The three algorithms are proposed for channel estimation in synchronous TWRNs. The MSE performance comparison for different modulation orders is plotted versus the SNR for frame length $N = 45$ in Fig. 4, which shows that the proposed LCML estimator achieves as good MSE performances as the DML and MCML estimators. Due to the nonconvex optimization function, the DML and MCML algorithms have to rely on numerical solutions by using optimization tools. In contrast, the LCML algorithm is based on a closed-form channel estimator.

Figs. 5 and 6 show the MSE performance of the LCML channel estimator versus SNR for frame length $N = 45$ and

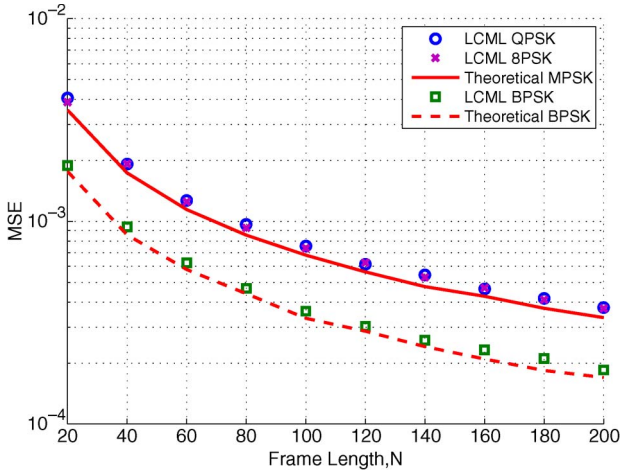


Fig. 6. MSE performance of the LCML channel estimator versus N for SNR $= 20$ dB.

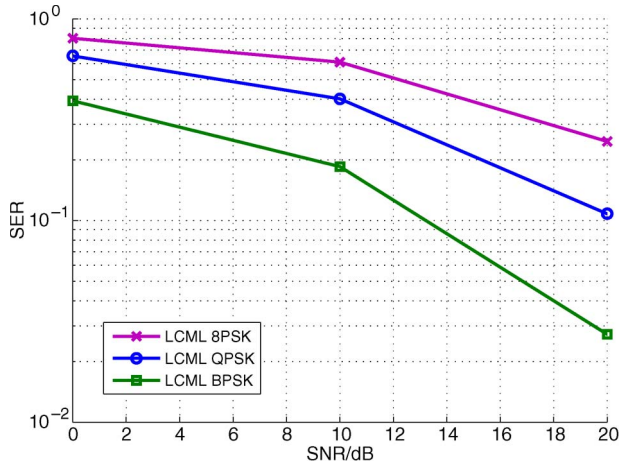


Fig. 7. SER performance of the LCML channel estimator versus SNR for $N = 20$ with 1 training symbol.

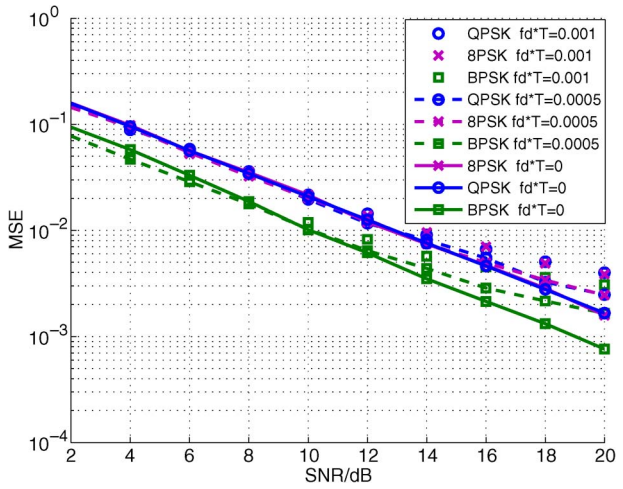


Fig. 8. MSE performance of the LCML channel estimator in time-varying channels versus SNR in the case of $N = 45$.

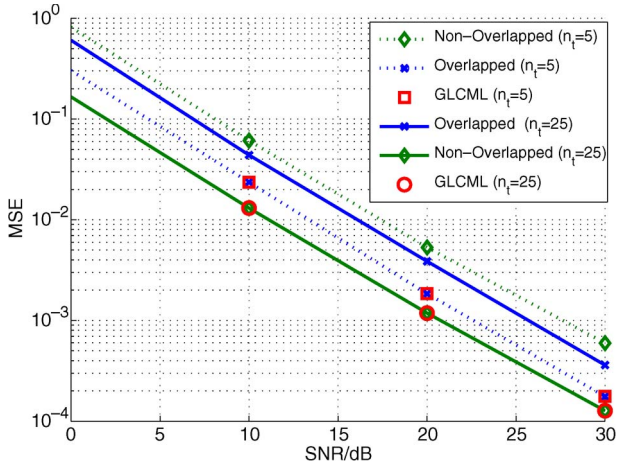


Fig. 9. MSE performance of the GLCML channel estimator versus SNR for $N = 45$ in QPSK.

versus N for SNR $= 20$ dB, respectively. The analytical MSE performance is included as a reference in both figures. The MSE performance of the LCML channel estimator improves with increasing SNR or N. The simulation results verify Theorems 1 and 2 in Section VII that the proposed LCML channel estimator approaches the true channel in the case of high SNR and is consistent in the case of a large frame length. The numerical MSE performance approaches the theoretical MSE, which verifies the MSE analysis in Section VII-D that the analytical expression of MSE is inversely proportional to SNR and N.

Next, the SER performance of the LCML is shown in Fig. 7. In the proposed algorithms, we use only one training symbol to find $\hat{\phi}_g$, which will be applied to resolve the phase ambiguity [16] in the signal demodulation. Fig. 7 shows that one training symbol is sufficient to achieve a near optimal SER.

In the scenarios of time-varying channels, the MSE performance of the proposed algorithm versus SNR is shown in Fig. 8. The Jakes' channel model [18] is used to simulate the time-varying channel with the normalized Doppler frequency f_dT, where f_d is the Doppler shift, and T is the input symbol period. In Fig. 8, the channel is static when $f_dT = 0$, which is the channel model assumed in this paper. In the case of time-varying channels, the MSE performance degrades as the LCML algorithm is not designed for time-varying channels. However, in the case where fade rate f_dT is 0.0005, the LCML algorithm achieves a similar MSE performance as in the static channel scenario. This demonstrates that the proposed LCML algorithm is applicable to some slow time-varying channels.

The MSE performance of the GLCML algorithm with timing offset $n_t = 5$ and $n_t = 25$ for QPSK is shown in Fig. 9. When $n_t = 25$, the ESSC (29) holds. Then, the nonoverlapped signal is selected to obtain channel estimation that gives a better estimation MSE than the overlapped signal. If $n_t = 5$, (29) no longer holds. Thus, the overlapped signal is used to estimate channel parameters. Fig. 9 shows that the GLCML algorithm always selects the samples producing a channel estimation with the minimum MSE in the timing asynchronous system.

Fig. 10 shows the frame synchronization error in BPSK and QPSK of the FACE algorithm, which is defined as $(|\hat{n}_t - n_t|/n_t)$, versus (n_t/N) for SNR $= 20$ dB. The synchronization performance improves with the frame length N, which confirms the theoretical result in Section VI-A that the synchronization error approaches zero if $N \to \infty$. As the signal

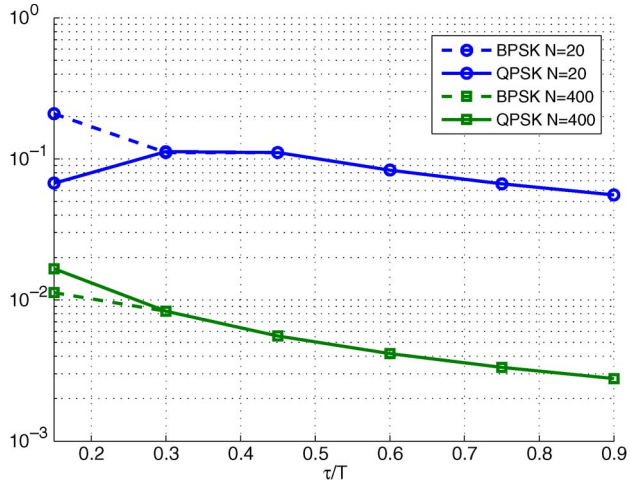


Fig. 10. Frame synchronization performance of the FACE algorithm versus frame offset for SNR = 20 dB.

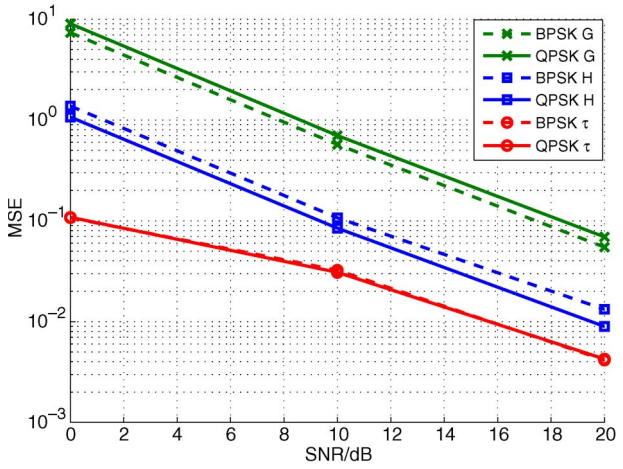


Fig. 11. MSE performance of the JSCE channel estimator versus SNR for $N = 45$ in the case of frame offset $n_t \neq 0$.

power detection and the cross-correlation method are employed in the FACE algorithm to determine the frame offset, the FACE algorithm achieves the same frame synchronization error performance in the cases of different modulation orders.

The MSE performance of the JSCE algorithm for different modulation orders is plotted versus SNR for frame length $N = 45$ in Figs. 11 and 12. For simplicity, the rectangular pulse-shaping filter is used. In both cases where the frame offset $n_t = 0$ and $n_t \neq 0$, the JSCE algorithm is able to achieve accurate symbol offset estimations. Figs. 10–12 show that the JSCE achieves joint channel and timing offset estimation in the asynchronous system.

IX. Conclusion

In this paper, we proposed an LCML algorithm and a JSCE algorithm for channel estimation in AF-TWRNs. In the LCML algorithm, we formulated a convex optimization objective function and obtained a closed-form channel estimation in synchronous AF-TWRNs. The GLCML algorithm was proposed in the asynchronous TWRNs by extending the LCML algorithm and the JSCE algorithm was proposed to estimate the timing

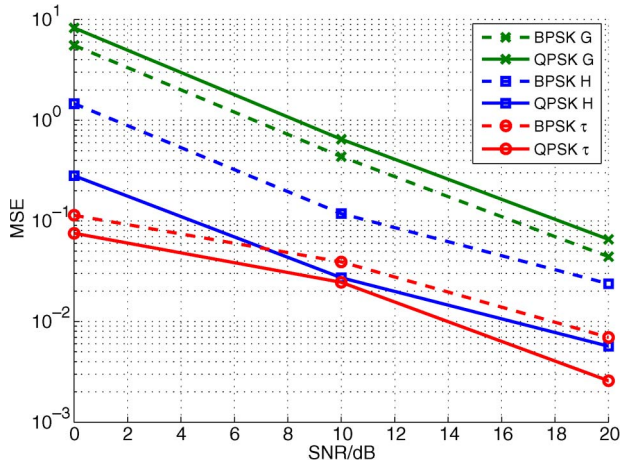


Fig. 12. MSE performance of the JSCE channel estimator versus SNR for $N = 45$ in the case of frame offset $n_t = 0$.

offset and channel parameters jointly. The LCML estimator achieves closed-form estimations. Theoretical analysis of the LCML algorithm proves that the derived channel estimates approach the real channel parameter values in scenarios with a high SNR or a large frame length. Both the analytical MSE expression and Monte Carlo simulations show that the average MSE performance of the LCML channel estimator improves as either SNR or frame length increases. The numerical results show that the GLCML algorithm achieves channel estimation performance as good as the LCML algorithm even in the presence of timing offset, and the JSCE algorithm is able to achieve accurate timing offset and channel estimations jointly.

Appendix A

As frequency is perfectly synchronized between both source nodes, source node T_1 downconverts the received signal to the baseband signal as follows:

$$r(t) = H \sum_{i=n_t+1}^{N} s_{1,i} f(t - iT - \tau) + G \sum_{i=n_t+1}^{N} s_{2,i} f(t - iT) + Kh_2 n_1(t) + n_2(t) \quad (52)$$

where $H = Kh_1h_2$, and $G = Kg_1h_2$. As T_1 is able to synchronize with $s_2(t)$, it filters the baseband signal with a matched filter $f'(t) = f(T - t)$ and then samples it at every T period to get

$$r_{\text{async}}(i) = H\left[f_1(\tau)s_{1,i} + f_2(\tau)s_{1,i+1}\right] + Gs_{2,i-n_t} + Kh_2 n_{1i} + n_{2i}, \quad i = n_t + 1, \ldots, N \quad (53)$$

where $f_1(\tau)$ and $f_2(\tau)$ are the factors resulting from the symbol offset τ and the use of a matched filter. $f_1(\tau)$ and $f_2(\tau)$ are estimated together with H by the GLCML algorithm.

In the following, we will show the derivation from (52) to (53) step by step and conclude that the values of $f_1(\tau)$ and $f_2(\tau)$ are related to the filter type and symbol offset τ. In

addition, it is proven that $f_1(\tau)$ and $f_2(\tau) < 1$, regardless of the filter type.

As different types of filters do not affect the noise part in the received samples, we consider that the noise-free signal is received by T_1 for simplicity, i.e.,

$$r(t) = H \sum_{i=n_t+1}^{N} s_{1,i} f(t - iT - \tau) + G \sum_{i=n_t+1}^{N} s_{2,i} f(t - iT).$$

Let $r(t)$ pass through the matched filter $f'(t)$, we obtain

$$\begin{aligned} r'(t) =& r(t) * f'(t) = \left(H \sum_{i=n_t+1}^{N} s_{1,i} f(t - iT - \tau) \right) * f'(t) \\ &+ \left(G \sum_{i=n_t+1}^{N} s_{2,i} f(t - iT) \right) * f'(t) \\ =& H \int_{-\infty}^{\infty} \left(\sum_{i=n_t+1}^{N} s_{1,i} f(\mu - iT - \tau) \right) f'(t - \mu)\, d\mu \\ &+ G \int_{-\infty}^{\infty} \left(\sum_{i=n_t+1}^{N} s_{2,i} f(\mu - iT) \right) f'(t - \mu)\, d\mu \\ =& H \int_{-\infty}^{\infty} \left(\sum_{i=1}^{N} s_{1,i} f(\mu - iT - \tau) \right) f(T - t + \mu)\, d\mu \\ &+ G \int_{-\infty}^{\infty} \left(\sum_{i=1}^{N} s_{2,i} f(\mu - iT) \right) f(T - t + \mu)\, d\mu \\ =& H \sum_{i=n_t+1}^{N} \left(s_{1,i} \int_{-\infty}^{\infty} f(\mu - iT - \tau) f(\mu + T - t)\, d\mu \right) \\ &+ G \sum_{i=n_t+1}^{N} \left(s_{2,i} \int_{-\infty}^{\infty} f(\mu - iT) f(\mu + T - t)\, d\mu \right). \end{aligned}$$

Define $x \triangleq \mu + T - t$; thus, $r'(t)$ is updated as

$$\begin{aligned} r'(t) =& H \sum_{i=n_t+1}^{N} \left(s_{1,i} \int_{-\infty}^{\infty} f(x) f\left(x - (i+1)T + t - \tau\right)\, dx \right) \\ &+ G \sum_{i=n_t+1}^{N} \left(s_{1,i} \int_{-\infty}^{\infty} f(x) f\left(x - (i+1)T + t\right)\, dx \right). \end{aligned}$$

Note that there exist autocorrelation terms of $f(x)$ in (54); thus, we denote the autocorrelation of $f(x)$ as $R(\Delta x) = \int_{-\infty}^{\infty} f(x) f(x + \Delta x)\, dx$ and obtain

$$\begin{aligned} r'(t) = H \sum_{i=n_t+1}^{N} R\left(t - (i+1)T - \tau\right) s_{1,i} \\ + G \sum_{i=n_t+1}^{N} R\left(t - (i+1)T\right) s_{2,i}. \end{aligned}$$

Then, sampling the signal every T period yields

$$\begin{aligned} r_{\text{async}}(i) = R(\tau) H s_{1,i} + R(T - \tau) H s_{1,i-1} + G s_{2,i}, \\ i = n_t + 1, \ldots, N. \end{aligned}$$

Hence, $f_1(\tau) = R(\tau)$ and $f_2(\tau) = R(T - \tau)$, whose values are related to the filter type and symbol offset τ. The normalized value of $R(\Delta x)$ has the following property: $R(\Delta x) = 1$ if $\Delta x = 0$; otherwise, $R(\Delta x) < 1$. Therefore, $f_1(\tau)$, $f_2(\tau) < 1$. In the case of rectangular pulse-shaping filters, $f_1(\tau) = ((T - \tau)/T)$, and $f_2(\tau) = (\tau/T)$.

APPENDIX B

Here, we will derive $V(v)$, which is the variance of $|z_i(v)|^2$, as stated in (57). $|z_i(v)|^2$ could be expanded as

$$\begin{aligned} |z_i(v)|^2 =& A^2|v|^2 + A^2|G|^2 + vG^* s_{1i} s_{2i}^* + v^* G s_{1i}^* s_{2i} \\ &+ |n_i|^2 + (v s_{1i} + G s_{2i}) n_i^* + (v^* s_{1i}^* + G^* s_{2i}^*)\, n_i. \end{aligned} \quad (54)$$

First, assuming that s_{1i} and s_{2i} are deterministic, the conditional expectation and variance [19] of $|z_i(v)|^2$ are obtained as

$$\begin{aligned} &\mathrm{E}\left\{ |z_i(v)|^2 \,|\, (s_{1i}, s_{2i}) \right\} \\ &= A^2|v|^2 + A^2|G|^2 + vG^* s_{1i} s_{2i}^* + v^* G s_{1i}^* s_{2i} + |h_2|^2 \sigma_n^2 + \sigma_n^2 \\ &\mathrm{Var}\left\{ |z_i(v)|^2 \,|\, (s_{1i}, s_{2i}) \right\} \\ &= \mathrm{E}\left\{ |z_i(v)|^4 \,|\, (s_{1i}, s_{2i}) \right\} - \mathrm{E}\left\{ |z_i(v)|^2 \,|\, (s_{1i}, s_{2i}) \right\}^2 \\ &= 2\left(K^2|h_2|^2\sigma_n^2 + \sigma_n^2\right) \\ &\quad \times \left(A^2|v|^2 + A^2|G|^2 + vG^* s_{1i} s_{2i}^* + v^* G s_{1i}^* s_{2i}\right). \end{aligned} \quad (55)$$

According to the total law of variance [19]

$$\begin{aligned} &\mathrm{Var}\left\{ |z_i(v)|^2 \right\} \\ &\quad = \mathrm{E}\left\{ \mathrm{Var}\left\{ |z_i(v)|^2 \,|(s_{1i}, s_{2i}) \right\} \right\} \\ &\qquad + \mathrm{Var}\left\{ \mathrm{E}\left\{ |z_i(v)|^2 \,|(s_{1i}, s_{2i}) \right\} \right\} \\ &\mathrm{E}\left\{ \mathrm{Var}\left\{ |z_i(v)|^2 \,|(s_{1i}, s_{2i}) \right\} \right\} \\ &\quad = 2\left(K^2|h_2|^2\sigma_n^2 + \sigma_n^2\right)\left(A^2|v|^2 + A^2|G|^2\right) \\ &\mathrm{Var}\left\{ \mathrm{E}\left\{ |z_i(v)|^2 \,|(s_{1i}, s_{2i}) \right\} \right\} \\ &\quad = \mathrm{E}\left\{ \mathrm{E}\left\{ |z_i(v)|^2 \,|(s_{1i}, s_{2i}) \right\}^2 \right\} \\ &\qquad - \mathrm{E}\left\{ \mathrm{E}\left\{ |z_i(v)|^2 \,|(s_{1i}, s_{2i}) \right\} \right\}^2 \\ &\quad = 2A^4|v|^2|G|^2. \end{aligned}$$

Expand $V(v)$ with respect to $(\Re\{v\}, \Im\{v\})$; then, we get

$$\begin{aligned} V(v) =& 2A^2|G|^2\left(K^2|h_2|^2\sigma_n^2 + \sigma_n^2\right) \\ &+ 2A^2\left(K^2A^2|h_2|^2\sigma_n^2 + A^2|G|^2 + \sigma_n^2\right)|v|^2 \end{aligned} \quad (56)$$

where $|v|^2 = (\Re\{v\}^2 + \Im\{v\}^2)$, and $(\Re\{v\}, \Im\{v\}) \in \mathbb{R}^2$.

APPENDIX C

Lemma 3: Assume H and G both belong to compact set Ω, then $V_N(v)$ converges uniformly to $V(v)$ as $N \to \infty$.

Proof: We use *the uniform law of large numbers* (see Lemma 4) [15] to prove Lemma 3. Let $g(x_i,\, \theta)$ be a function of the parameter $\theta \in \Omega$ and a sequence of i.i.d. random variables $x_i \in \mathbb{C}$, $i = 1, \ldots, N$.

Lemma 4: Suppose that, for all x_i, (a) $g(x_i, \theta)$ is continuous at each $\theta \in \Omega$ with a probability of 1, (b) $g(x_i,\, \theta)$ is dominated by a function $D(x_i)$ for all $\theta \in \Omega$, i.e., $|g(x_i,\, \theta)| \le D(x_i)$, $\forall\, \theta \in \Omega$, and (c) $E\{D(x_i)\} < \infty$, then $(1/N)\sum_{i=1}^{N} g(x_i, \theta)$ converges uniformly to $E\{g(x_i, \theta)\}$ when $N \to \infty$: $\mathbf{sup}_{\theta\in\Omega}|(1/N)\sum_{i=1}^{N} g(x_i,\, \theta) - E\{g(x_i,\, \theta)\}| \xrightarrow{p} 0$, $N \to \infty$.

First, we will prove that Lemma 4 is applicable to $|z_i(v)|^2$. Since $z_i(v) \triangleq vs_{1i} + Gs_{2i} + n_i$, $i = 1, \ldots, N$, and $|z_i(v)|^2$ is derived in Appendix B as follows:

$$|z_i(v)|^2 = A^2|v|^2 + A^2|G|^2 + vG^* s_{1i} s_{2i}^* + v^* G s_{1i}^* s_{2i} + |n_i|^2 + vs_{1i} n_i^* + Gs_{2i} n_i^* + v^* s_{1i}^* n_i + G^* s_{2i}^* n_i.$$

$|z_i(v)|^2$ is a function of fixed parameters v and G, and i.i.d. random variables s_{1i}, s_{2i}, and n_i. Define vector $\mathbf{z} \triangleq [A^2|v|^2,\, A^2|G|^2,\, vG^* s_{1i} s_{2i}^*,\, v^* G s_{1i}^* s_{2i},\, |n_i|^2,\, vs_{1i} n_i^*,\, Gs_{2i} n_i^*,\, v^* s_{1i}^* n_i,\, G^* s_{2i}^* n_i]$. Note that Lemma 4 can be extended to multivariate cases with multiple fixed parameters and random variables [10]. Condition (a) is met, as $|z_i(v)|^2$ is continuous at each $v, G \in \Omega$ with probability of 1. Using the *triangle inequality* [12], we obtain $|z_i(v)|^2 \le (|vs_{1i}| + |Gs_{2i}| + |n_i|)^2$. As $|H|,\ |G| < \xi$ is assumed in Condition 1.1, i.e., $|z_i(v)|^2 \le 4A^2\xi^2 + |n_i|^2 + 4A\xi|n_i|$, then condition (b) is satisfied. Define $d(n_i) \triangleq 4A^2\xi^2 + |n_i|^2 + 4A\xi|n_i|$; thus, we get $E\{d(n_i)\} = 4A^2\xi^2 + (K^2|h_2|^2+1)\sigma_n^2 + \sqrt{(2\sigma_n^2(K^2|h_2|^2+1)/\pi)}$. Since channel coefficient h_2 and noise power σ_n are bounded, $E\{d(n_i)\} < \infty$, and condition (c) is met. Therefore, Lemma 4 can be applied to $|z_i(v)|^2$ to prove that $\mathbf{sup}_{v\in\Omega}|(1/N)\sum_{i=1}^{N}|z_i(v)|^2 - E\{|z_i(v)|^2\}| \xrightarrow{p} 0$ as $N \to \infty$.

It has been proven that the sample mean $(1/N)\sum_{i=1}^{N} |z_i(v)|^2$ converges uniformly to the expectation $E\{|z_i(v)|^2\}$ as $N \to \infty$. If we define random i.i.d. $g_i(v) \triangleq (|z_i(v)|^2 - (\sum_{k=1}^{N}|z_k(v)|^2/N))^2$, $i = 1, \ldots, N$, then $g_i(v)$ converges uniformly to $g_i'(s_{1i}, s_{2i}, n_i; v, G) = (|z_i(v)|^2 - E\{|z_k(v)|^2\})^2$. Here, $g_i'(s_{1i}, s_{2i}, n_i; v, G)$ is a function of parameters v and G, and random variables s_{1i}, s_{2i}, and n_i. Then, we will validate that Lemma 4 can be applied to $g_i'(s_{1i}, s_{2i}, n_i; v, G)$. By inspection, $g_i'(s_{1i}, s_{2i}, n_i; v, G)$ is continuous at each v, $G \in \Omega$ with a probability of 1; hence, condition (a) is met. Using triangle inequality, we get the dominant function of $|g_i'(s_{1i}, s_{2i}, n_i; v, G)|$ from (54) (see Appendix B), i.e.,

$$\begin{aligned}
&|g_i'(s_{1i}, s_{2i}, n_i; v, G)| \\
&= \left|vG^* s_{1i} s_{2i}^* + v^* G s_{1i}^* s_{2i} + |n_i|^2 + (vs_{1i} + Gs_{2i}) n_i^* \right. \\
&\quad \left. + (v^* s_{1i}^* + G^* s_{2i}^*)\, n_i - K^2|h_2|^2\sigma_n^2 - \sigma_n^2\right|^2 \\
&\le \left(|vG^* s_{1i} s_{2i}^*| + |v^* G s_{1i}^* s_{2i}| + |n_i|^2 + |(vs_{1i} + Gs_{2i}) n_i^*| \right. \\
&\quad \left. + |(v^* s_{1i}^* + G^* s_{2i}^*)\, n_i| + K^2|h_2|^2\sigma_n^2 + \sigma_n^2\right)^2 \\
&< \left(2A^2\xi^2 + |n_i|^2 + 4A\xi|n_i| + K^2\xi^2\sigma_n^2 + \sigma_n^2\right)^2.
\end{aligned}$$

Define $D(n_i) \triangleq (2A^2\xi^2 + |n_i|^2 + 4A\xi|n_i| + K^2\xi^2\sigma_n^2 + \sigma_n^2)^2$; thus, condition (b) is met. Since noise power $\sigma^2 = (K^2|h_2|^2 + 1)\sigma_n^2$ and $|h_2| < \xi$, $E\{D(n_i)\} < \infty$, and condition (c) is met. As a result, we obtain $\mathbf{sup}_{v\in\Omega}|(1/N)\sum_{i=1}^{N} g_i'(s_{1i}, s_{2i}, n_i; v, G) - E\{g_i'(s_{1i}, s_{2i}, n_i; v, G)\}| \xrightarrow{p} 0$ as $N \to \infty$.

The conclusion that $\sup_{v\in\Omega}|(1/N)\sum_{i=1}^{N}|z_i(v)|^2 - E\{|z_i(v)|^2\}| \xrightarrow{p} 0$ as $N \to \infty$ implies that $V_N(v)$ (44) converges uniformly to $(1/N)\sum_{i=1}^{N} g_i'(s_{1i}, s_{2i}, n_i; v, G)$. Since $V(v) = E\{g_i'(s_{1i}, s_{2i}, n_i; v, G)\}$ and it has been proved that $(1/N)\sum_{i=1}^{N} g_i'(s_{1i}, s_{2i}, n_i; v, G)$ converges uniformly to $E\{g_i'(s_{1i}, s_{2i}, n_i; v, G)\}$ as $N \to \infty$, therefore, $V_N(v)$ converges uniformly to $V(v)$ as $N \to \infty$. ■

APPENDIX D

Lemma 5: $V(v)$ has a unique global minimum with respect to v occurring at $v_o = 0$.

Proof: To demonstrate Condition 1.3, we will prove that $v = 0$ is a local minimum of $V(v)$ first (see Lemma 6) and then the convexity of $V(v)$ (see Lemma 7).

Lemma 6: $v = 0$ is a local minimum of $V(v)$.

Proof: Expanding $V(v)$ and we obtain (see Appendix B)

$$\begin{aligned}
V\left(\Re\{v\}, \Im\{v\}\right) &= 2A^2|G|^2\left(|h_2|^2\sigma_n^2 + \sigma_n^2\right) \\
&\quad + 2A^2\left(A^2|h_2|^2\sigma_n^2 + A^2|G|^2 + \sigma_n^2\right)\left(\Re\{v\}^2 + \Im\{v\}^2\right) \qquad (57)
\end{aligned}$$

with $(\Re\{v\}, \Im\{v\}) \in \mathbb{R}^2$, where $\mathbb{R}$ denotes real number field. The first partial derivative [12] of $V(v)$ is

$$\frac{\partial V}{\partial \Re\{v\}} = 4A^2\left(A^2|h_2|^2\sigma_n^2 + A^2|G|^2 + \sigma_n^2\right)\Re\{v\}$$

$$\frac{\partial V}{\partial \Im\{v\}} = 4A^2\left(A^2|h_2|^2\sigma_n^2 + A^2|G|^2 + \sigma_n^2\right)\Im\{v\}$$

which are equal to zero at the critical point (0, 0), i.e., $v = 0$ is a local extremum. According to the *second partial derivative test* [12], the second derivative test discriminant is

$$\nabla^2 V(v) = \begin{vmatrix} \frac{\partial^2 V}{\partial \Re\{v\}^2} & \frac{\partial V}{\partial \Re\{v\}\partial \Im\{v\}} \\ \frac{\partial V}{\partial \Re\{v\}\partial \Im\{v\}} & \frac{\partial^2 V}{\partial \Im\{v\}^2} \end{vmatrix}$$

where the second partial derivatives of $V(v)$ are

$$\frac{\partial V}{\partial \Re\{v\}\partial \Im\{v\}} = 0$$

$$\frac{\partial^2 V}{\partial \Re\{v\}^2} = \frac{\partial^2 V}{\partial \Im\{v\}^2} = 4A^2\left(A^2|h_2|^2\sigma_n^2 + A^2|G|^2 + \sigma_n^2\right).$$

Then, we get $\nabla^2 V(v) = 16A^4(A^2|h_2|^2\sigma_n^2 + A^2|G|^2 + \sigma_n^2)^2 > 0$ and $(\partial^2 V/\partial \Re\{v\}^2) > 0$ at $v = 0$. Hence, the local extremum $v = 0$ is a local minimum of $V(v)$. ■

Then, we will prove this local minimum is also the global minimum of $V(v)$ by demonstrating the convexity of $V(v)$ according to Lemma 7 [9].

Lemma 7: Define $\mathbf{dom}V$ as the domain of V; then, $V(v)$ is convex if and only if $\mathbf{dom}V$ is convex and its Hessian is positive semidefinite: For all $v \in \mathbf{dom}V$, $\nabla^2 V(v) \geq 0$.

It has been proven in Lemma 6 that $\nabla^2 V > 0$ for all $(\Re\{v\}, \Im\{v\}) \in \mathbb{R}^2$. Since $\mathbf{dom}V = \mathbb{R}^2$ is convex, $V(v)$ is a convex function according to Lemma 7. For a convex function, its local minimum is also the global minimum; therefore, the local minimum point $v = 0$ is the global minimum of $V(v)$. We complete the proof based on Lemmas 6 and 7. ■

APPENDIX E

Theorem 1:

Proof: Regarding Condition 1.1, there are no upper bounds on H and G, strictly speaking, if we treat h_1, h_2, g_1, and g_2 as ideal complex Gaussian random variables. However, we can always choose a sufficiently large ξ such that $\Pr(|H|, |G| < \xi) = 1 - \epsilon$, where ϵ can be made arbitrarily small. Therefore, Condition 1.1 can be satisfied by assuming that the amplitude of channel coefficients h_1, h_2, g_1, and g_2 are bounded. Conditions 1.2 and 1.3 have been proven in Lemma 3 (see Appendix C) and Lemma 5 (see Appendix D), respectively. Since Conditions 1.1, 1.2, and 1.3 are satisfied, Theorem 1 holds. ■

APPENDIX F

Theorem 2:

Proof: For simplicity, we define the estimation error $v_2 \triangleq H - \hat{H}_{\mathrm{lcml}}^{\mathrm{BPSK}}$. Then, (20) is expressed as

$$\begin{aligned} F_{\mathrm{sync}}^{\mathrm{BPSK}}(v_2) &= \frac{1}{N}\sum_{i=1}^{N}\left| z_i^2(v_2) - \frac{\sum_{k=1}^{N} z_k^2(v_2)}{N}\right|^2 \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(\Re\left\{z_i^2(v_2)\right\} - \frac{\sum_{k=1}^{N}\Re\left\{z_k^2(v_2)\right\}}{N}\right)^2 \\ &\quad + \frac{1}{N}\sum_{i=1}^{N}\left(\Im\left\{z_i^2(v_2)\right\} - \frac{\sum_{k=1}^{N}\Im\left\{z_k^2(v_2)\right\}}{N}\right)^2 \end{aligned} \quad (58)$$

where $z_i(v_2) \triangleq v_2 s_{1i} + G s_{2i} + n_i,\ i = 1, \ldots, N$. Equation (58) represents the sum of sample variance of the real and imaginary parts of random variable $z_i^2(v_2)$.

$z_i(v_2)$ in (58) is approximated as $z_i(v_2) = v_2 s_{1i} + G s_{2i}$, $i = 1, \ldots, N$, as SNR $\to \infty$. As s_{1i}, $s_{2i} = \pm 1$ in BPSK, we obtain

$$z_i^2(v_2) = v_2^2 + G^2 + 2 v_2 G s_{1i} s_{2i}$$

$$\Re\left\{z_i^2(v_2)\right\} = \Re\left\{v_2^2 + G^2\right\} + 2\Re\{v_2 G\}\cos(\phi_{1i} + \phi_{2i}) \quad (59)$$

$$\Im\left\{z_i^2(v_2)\right\} = \Im\left\{v_2^2 + G^2\right\} + 2\Im\{v_2 G\}\cos(\phi_{1i} + \phi_{2i}). \quad (60)$$

Since the objective function (58) is the sum of the sample variance of $\Re\{z_i^2(v_2)\}$ and $\Im\{z_i^2(v_2)\}$, it is obvious that $F_{\mathrm{sync}}^{\mathrm{BPSK}}(v_2) \geq 0$ with equality if and only if the terms $z_i^2(v_2)$, $i = 1, \ldots, N$, are all equal. This is equivalent to either of the following two conditions given $G \neq 0$:

$$\begin{cases} \mathbf{Condition\ 1}: v_2 = 0, \quad \mathbf{Condition\ 2}: \quad x_i = x_j \\ x_i = \cos(\phi_{1i} + \phi_{2i}), \quad x_j = \cos(\phi_{1j} + \phi_{2j}), \\ \text{where} \quad i \neq j \quad \text{and } i, j = 1, \ldots, N. \end{cases} \quad (61)$$

Condition 1 indicates that there is an unique global minimum of the objective function at $v_2 = 0$, i.e., $\hat{H}_{\mathrm{lcml}}^{\mathrm{BPSK}} = H$. Whereas, on the other hand, there are infinite number of global minimum if Condition 2 holds. In BPSK, $x_i, x_j = \pm 1$ as $\phi_{1i}, \phi_{2i} = [0, \pi]$. As a result, $\Pr(\mathbf{Condition\ 2}) = (1/2^{N-1})$.

The probability that there exists an unique global minimum of the objective function (58) conditioned on SNR $\to \infty$ is $1 - (1/2^{N-1})$. ■

APPENDIX G

In the following, we will validate Lemma 2 by proving $\Im\{\hat{H}\}$ equals its Taylor series expansion $T(x, y)$ if either SNR or $N \to \infty$ (see Lemma 8) at first and then demonstrating $T(x, y) = (\mu_x/\mu_y) + (x/\mu_y) - (\mu_x y/\mu_y^2)$ if SNR $\to \infty$ (see Lemma 9).

Since the multivariate function $f(x, y) = (x/y)$, where $(x, y) \in \mathbb{C}^2$ is $k+1$ times continuously differentiable at the point $(\mu_x, \mu_y) \in \mathbb{C}^2$, its Taylor series expansion at the point (μ_x, μ_y) is expressed as

$$\begin{aligned} T(x, y) &= \sum_{n_x + n_y = 0}^{k} \frac{\partial^{n_x + n_y} f(\mu_x, \mu_y)}{\partial x^{n_x} \partial y^{n_y}} \frac{(x - \mu_x)^{n_x}(y - \mu_y)^{n_y}}{n_x! n_y!} \\ &\quad + \sum_{n_x' + n_y' = k+1} R_{n_x' + n_y'}(x, y)(x - \mu_x)^{n_x'}(y - \mu_y)^{n_y'} \end{aligned}$$

where n_x, n_y, n_x', $n_y' \in \mathbb{N}$, and $R_{k+1}(x, y)$ is the remainder of the kth-degree Taylor polynomial approximation [17].

Lemma 8: The Taylor series expansion $T(x, y)$ about the point (μ_x, μ_y) is equal to $f(x, y) = (x/y)$ for $(x, y) \in \mathbb{C}^2$ if and only if either SNR or frame length $N \to \infty$.

Proof: Since $f(x, y) = (x/y)$ is infinitely differentiable at the point (μ_x, μ_y), its Taylor series expansion is summed up as

$$T(x, y) = \lim_{k \to \infty} \frac{x}{y}\left(1 - \left(1 - \frac{y}{\mu_y}\right)^k\right)$$

which shows that, if $|y - \mu_y| < |\mu_y|$ holds, then $T(x, y) = f(x, y)$. Next, we will prove that $\Pr(|y - \mu_y| < |\mu_y|) = 1$ if either SNR or $N \to \infty$.

As y is a complex normal distributed variable in $\mathcal{CN}(\mu_y, \sigma_y^2)$, where

$$\mu_y = \frac{-j2(A^4A^2|G|^2+\sigma^2)^2(N^2-3N+2)}{N^2}$$
$$\sigma_y^2 = \frac{8A^7\sigma(A^2|G|^2+\sigma^2)^4(N^2-3N+2)^2}{|G|^4N^5} \quad (62)$$

according to its cumulative distribution function, we obtain

$$\Pr\left(|y-\mu_y| < |\mu_y|\right) = \operatorname{erf}\left(\frac{|\mu_y|}{\sqrt{2\sigma_y^2}}\right) \quad (63)$$

where $\operatorname{erf}(x)$ denotes the error function [11]. By substituting (62) into (63), we obtain $\Pr(|y-\mu_y| < |\mu_y|) = \operatorname{erf}((|G|^2/2)\sqrt{(AN/\sigma)})$. The values of the error function state that, if $(|G|^2/2)\sqrt{(AN/\sigma)} \to \infty$, then $\Pr(|y-\mu_y| < |\mu_y|) = 1$, which implies that Taylor series expansion $T(x, y)$ equals the function $f(x, y)$ for $(x, y) \in \mathbb{C}^2$ in a scenario with either a high SNR or a large frame length. ■

There is no need to meet the strict requirement of $(|G|^2/2)\sqrt{(AN/\sigma)} \to \infty$. As long as $(|G|^2/2)\sqrt{(AN/\sigma)}$ is not too small, $T(x, y) = f(x, y)$ holds. For example, $\operatorname{erf}((|G|^2/2)\sqrt{(AN/\sigma)}) = 0.9953$ when $(|G|^2/2)\sqrt{(AN/\sigma)} = 2$. In practical applications, it is possible to adjust SNR or frame length N, to satisfy the condition for $\Pr(|y-\mu_y| < |\mu_y|) = 1$.

Lemma 9: If SNR $\to \infty$, the Taylor series $T(x, y)$ equals its first-degree Taylor polynomial approximation at the point (μ_x, μ_y), namely

$$T(x, y) = \frac{\mu_x}{\mu_y} + \frac{x}{\mu_y} - \frac{\mu_x y}{\mu_y^2}. \quad (64)$$

Proof: Expand $T(x, y)$ to its first-order Taylor polynomial approximation; then, we obtain

$$T(x, y) = \frac{\mu_x}{\mu_y} + \frac{x}{\mu_y} - \frac{\mu_x y}{\mu_y^2} + R_2(x, y)$$

where $R_2(x, y)$ is the remainder [17]. According to *Taylor's theorem for multivariate functions* [17], the upper bound of the remainder $R_{k+1}(x, y)$ is

$$|R_{k+1}(x, y)| \le \max_{n_x'+n_y'=k+1} \max_{(x,y)\in\mathbb{R}^2} \frac{k+1}{n_x'!n_y'!} \left|\frac{\partial^{k+1} f(x, y)}{\partial x^{n_x'} \partial y^{n_y'}}\right| \quad (65)$$

where $R_{k+1}(x, y)$ represents the remainder of the kth-degree Taylor polynomial approximation. Equation (65) indicates that the upper bound of $|R_{k+1}(x, y)|$ is related to the maximum norm of the $k+1$th partial derivative of $f(x, y)$. As we expand Taylor series $T(x, y)$ to its first order, $k = 1$ in this case. Calculate the second-order partial derivative of $f(x, y)$; then, we obtain the upper bound of the remainder $|R_2(x, y)|$, i.e.,

$$|R_2(x, y)|_{up} = \max\left\{\left|\frac{2x}{y^3}\right|, \left|\frac{2}{y^2}\right|\right\}.$$

Expanding x and y from (26) and approximating x and y under the condition of SNR $\to \infty$ yield

$$x = -2j|G|^4\Im\{H\} * \left(A^8 - \frac{2A^4}{N^2}s_2^T s_1^* s_1^T s_2^* - \frac{\left(s_2^T\right)_\cdot^2 \left(s_1^*\right)_\cdot^2 \left(s_1^T\right)_\cdot^2 \left(s_2^*\right)_\cdot^2}{N^2} + \frac{\left(s_2^T\right)_\cdot^2 \left(s_1^*\right)_\cdot^2 \left(s_1^T s_2^*\right)^2}{N^3} + \frac{\left(s_1^T\right)_\cdot^2 \left(s_2^*\right)_\cdot^2 \left(s_2^T s_1^*\right)^2}{N^3}\right)$$
$$y = -2j|G|^4 * \left(A^8 - \frac{2A^4}{N^2}s_2^T s_1^* s_1^T s_2^* - \frac{\left(s_2^T\right)_\cdot^2 \left(s_1^*\right)_\cdot^2 \left(s_1^T\right)_\cdot^2 \left(s_2^*\right)_\cdot^2}{N^2} + \frac{\left(s_2^T\right)_\cdot^2 \left(s_1^*\right)_\cdot^2 \left(s_1^T s_2^*\right)^2}{N^3} + \frac{\left(s_1^T\right)_\cdot^2 \left(s_2^*\right)_\cdot^2 \left(s_2^T s_1^*\right)^2}{N^3}\right). \quad (66)$$

Here, we define $\mathbf{s}_\cdot^2 \triangleq [s_1^2, s_2^2, \ldots, s_N^2]$, in which $\mathbf{s} = [s_1, s_2, \ldots, s_N]$. Equation (66) shows that $|R_2(x, y)|_{up} \propto (1/|y|^2)$, which approaches zero as SNR $\to \infty$. Then, we conclude that $T(x, y) = (\mu_x/\mu_y) + (x/\mu_y) - (\mu_x y/\mu_y^2)$ if SNR $\to \infty$. ■

REFERENCES

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkley Symp. Math. Stat. Prob.*, Berkley, CA, USA, 1961, vol. 1, pp. 611–644.

[2] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," *Comput. Commun. Rev.*, vol. 37, no. 4, pp. 397–408, Oct. 2007.

[3] F. Gao, R. Zhang, and Y. Liang, "On channel estimation for OFDM based two-way relay networks," in *Proc. IEEE ICC*, 2009, pp. 1–5.

[4] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Tech.*, vol. 56, no. 2, pp. 670–685, Mar. 2007.

[5] B. Jiang, T. Cui, and A. Nallanathan, "On channel estimation and optimal training design for amplify and forward relay networks," *IEEE Trans. Wireless Commun*, vol. 7, no. 5, pp. 1907–1916, May 2008.

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

[7] G. Wang, F. Gao, and C. Tellambura, "Joint frequency offset and channel estimation methods for two-way relay networks," in *Proc. IEEE GLOBECOM*, 2009, pp. 1–5.

[8] S. Abdallah and I. N. Psaromiligkos, "Semi-blind channel estimation for amplify-and-forward two-way relay networks employing constant-modulus constellations," in *Proc. 44th Annu. CISS*, Princeton, NJ, USA, Mar. 2010, pp. 1–5.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[10] S. Abdallah and I. N. Psaromiligkos, "Blind channel estimation for amplify-and-forward two-way relay networks employing M-PSK modulation," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3604–3615, Jul. 2012.

[11] U. Mengali, *Synchronization Techniques for Digital Receivers*. New York, NY, USA: Springer-Verlag, 1997.

[12] J. M. Wooldridge, *Introductory Econometrics: A Modern Approach*. Mason, OH, USA: Cengage Learning, 2009.

[13] S. Abdallah and I. N. Psaromiligkos, "Partially-blind estimation of reciprocal channels for AF two-way relay networks employing M-PSK modulation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 5, pp. 1649–1654, May 2012.

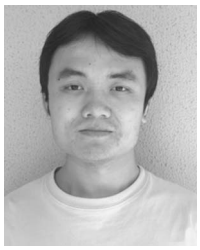
[14] W. Weisstein, Cross-Correlation Theorem, From MathWorld—A Wolfram Web Resource. [Online]. Available: http://mathworld.wolfram.com/Cross-Correlation.html

[15] W. Newey and D. McFadden, "Large sample estimation and hypothesis testing," in *Handbook of Econometrics*. New York, NY, USA: Elsevier, 1994, ch. 36.

[16] K. Ito, *Encyclopedic Dictionary of Mathematics: The Mathematical Society of Japan*. Cambridge, MA, USA: MIT Press, 1993.
[17] P. Dienes, *The Taylor Series: An Introduction to the Theory of Functions of a Complex Variable*. Oxford, U.K.: Clarendon, 1931.
[18] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited," *Electron. Lett.*, vol. 29, no. 13, pp. 1162–1163, Jun. 1993.
[19] A. Neil, *A Course in Probability*. Reading, MA, USA: Addison-Wesley, 2005, pp. 385–386.

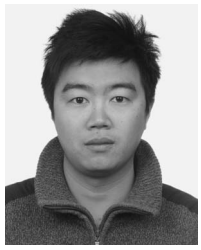
Qiong Zhao (S'11) received the B.S. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 2010. She is currently working toward the Ph.D. degree with the Telecommunications Laboratory, School of Electrical and Information Engineering, The University of Sydney, Sydney, Australia.

Her current research interests include channel estimation, synchronization, cooperative communications, and traffic scheduling for wireless communications.



Zhendong Zhou (S'03–M'07–SM'12) received the B.S. and M.S. degrees in electrical engineering from Zhejiang University, Hangzhou, China, in 2000 and 2003, respectively, and the Ph.D. degree from The University of Sydney, Sydney, Australia, in 2007.

Since 2006, he has been a Research Fellow with the Telecommunications Laboratory, School of Electrical and Information Engineering, The University of Sydney. His research interests include adaptive modulation and coding, multiple-input–multiple-output (MIMO) space–time processing, multiuser MIMO broadcast channels, distributed relay processing, and software-defined radios.



Jun Li (M'09) received the Ph.D. degree in electronic engineering from Shanghai Jiaotong University, Shanghai, China, in 2009.

From January 2009 to June 2009, he was a Research Scientist with the Department of Research and Innovation, Alcatel Lucent Shanghai Bell. From June 2009 to April 2012, he was a Postdoctoral Fellow with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia. Since April 2012, he has been a Research Fellow with Telecommunications Laboratory, School of Electrical and Information Engineering, The University of Sydney, Sydney. His research interests include network information theory, channel coding theory, wireless network coding and cooperative communications.

Dr. Li has served as a Technical Program Committee member for several international conferences, such as the 2008 International Symposium on Computer Science and Computational Technology, the 2008 IEEE International Conference on Circuits and Systems for Communications, the 2009, 2010, and 2013 Asia–Pacific Conferences on Communications, the 2011 Spring IEEE Vehicular Technology Conference, the 2011 and 2014 IEEE International Conferences on Communications, the 2012 IEEE Region 10 Conference, the 2013 International Workshop on High Mobility Wireless Communications, the 2013 International Conference on Connected Vehicles and Expo, and the 2014 Asia–Pacific Conference on Wireless and Mobile.

Branka Vucetic (F'03) received the B.S., M.S., and Ph.D. degrees from the University of Belgrade, Belgrade, Yugoslavia, in 1972, 1978, and 1982, respectively, all in electrical engineering.

She is currently the Peter Nicol Russell Chair of Telecommunications Engineering with the Telecommunications Laboratory, School of Electrical and Information Engineering, The University of Sydney, Sydney, Australia. She is an internationally recognized expert in wireless communications and coding. Her research has involved collaborations with industry and government organizations in Australia and several other countries.