Pilot Optimization, Channel Estimation, and Optimal Detection for Full-Duplex OFDM Systems With IQ Imbalances

Feng Shu, *Member, IEEE*, Jin Wang, Jun Li, *Senior Member, IEEE*, Riqing Chen, and Wen Chen, *Senior Member, IEEE*

***Abstract*—In full-duplex orthogonal frequency-division multiplexing systems with IQ imbalances, a frequency-domain least-squares (FD-LS) channel estimator is proposed to estimate both source-to-destination (intended) and destination-to-destination (self-interference) channels. Subsequently, an optimal closed-form pilot matrix is derived to minimize the sum of mean square errors (sum-MSE) of the proposed FD-LS channel estimator. Then, an improved FD-LS estimator is presented and proved to further improve the performance of the FD-LS by exploiting the time-domain property of the channel. In the presence of channel estimation error, an optimal low-complexity maximum likelihood (ML) detector is developed by using eigenvalue and singular value decompositions and whitening the residual self-interference plus noise. Simulation results show that given the proposed FD-LS estimator, the optimal pilot matrix performs much better than those nonsingular pilot matrices with larger conditional numbers (CN). To be specific, the former achieves about 10-dB signal-to-noise ratio (SNR) gain over the latter with CN $=$ 10, the improved FD-LS channel estimator provides about 6-dB SNR gain over the FD-LS estimator at a fixed bit error rate (BER), and the proposed whitening-filter ML detector performs at least 0.6 dB better than the conventional ML detector at a given BER of 10^{-3} in the medium-and high-SNR regions.**

***Index Terms*—Channel estimation, full-duplex, IQ imbalances, least-squares, orthogonal frequency-division multiplexing (OFDM), pilot design, sum of mean square errors (sum-MSE).**

Manuscript received April 19, 2016; revised October 30, 2016 and January 11, 2017; accepted February 7, 2017. Date of publication February 13, 2017; date of current version August 11, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61472190, Grant 61271230 and Grant 61501238; the Open Research Fund of National Key Laboratory of Electromagnetic Environment, China Research Institute of Radiowave Propagation, under Grant 201500013; the open research fund of the National Mobile Communications Research Laboratory, Southeast University, China, under Grant 2013D02; the Research Fund for the Doctoral Program of Higher Education of China under Grant 20113219120019; and the Foundation of Cloud Computing and Big Data for Agriculture and Forestry (117-612014063). The review of this paper was coordinated by Dr. H. Jiang.

F. Shu is with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China, and also with the College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China (e-mail: shufeng@njust.edu.cn).

J. Wang and J. Li are with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: jin.wang@njust.edu.cn; jun.li@njust.edu.cn).

R. Chen is with the College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China (e-mail: riqing.chen@fafu.edu.cn).

W. Chen is with the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: wenchen@sjtu.edu.cn).

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.2017.2667686

I. Introduction

WIRELESS communication systems are traditionally designed for half-duplex mode, where they can transmit and receive data on two different frequency bands or time slots. By allowing simultaneous transmission/reception over the same time and spectrum, compared to half-duplex, full-duplex technique may double the transmission rate and spectrum efficiency, and attracts considerable research interests from both academia and the industrial world [1]–[4]. In [5] and [6], the authors investigate the effect of channel estimation errors on bidirectional full-duplex relay networks with amplify-and-forward strategy. A least square (LS) channel estimator is implemented by iterating between channel estimation and intended signal detection to separate the received signal from the known transmitted signal in frequency domain [7], [8]. A semi-blind maximum-likelihood (ML) algorithm is proposed to jointly estimate the self-interference channels of MIMO full-duplex systems. The closed-form solution is derived and an iterative procedure is developed to further improve the estimation performance in high signal-to-noise ratio (SNR) region [9], [10]. In [11], a blind ML channel estimator is presented to simultaneously estimate both self-interference and intended channels in full-duplex massive system by the expectation maximization algorithm.

In conventional half-duplex OFDM systems with IQ-imbalances, channel estimation, pilot design, and IQ-imbalance compensation have been extensively investigated in [12]–[22]. IQ imbalances destroy the orthogonality between the signals in I and Q branches, and distort the received signal [12], [13]. In OFDM systems with transmit and receive IQ imbalances, the post-FFT LS channel estimation is constructed to estimate channel parameters, and then several compensations including pre-distortion scheme at the transmitter, and adaptive equalizer and a pre-FFT correction at the receiver are designed to reduce the IQ-imbalance distortion [15]. In [17], the pilot design of minimizing mean square error is proposed for MIMO-OFDM with frequency-dependent I/Q imbalances, and several typical pilot patterns are designed and perform well in terms of pilot overhead, estimation performance, and general applicability. In [18], a time-domain sparse channel estimator based on iterative shrinkage is proposed for OFDM system with IQ imbalance in sparse wireless channels and achieves a significant SNR gain over existing frequency-domain estimation. A novel time domain blind compensation is proposed for

frequency-independent and frequency-dependent IQ imbalances in direct-conversion receivers [20]. In [21], an adaptive efficient post-FFT equalizer is shown to achieve an ideal compensation. In massive MIMO systems with IQ-imbalances, the authors investigate the effect of IQ imbalances employing maximum-ratio combining (MRC) receivers and derives a linear minimum-mean-square-error channel estimator [22].

To the best of our knowledge, channel estimation and pilot design for OFDM system with both full-duplex and IQ-imbalance are still open problems. In this paper, we mainly focus on the investigation of channel estimation and pilot design in such an OFDM system. Our main contributions are as follows.

1) We propose a frequency-domain least square channel estimator of achieving the Cramer-Rao lower bound (CRLB) for full-duplex OFDM system with IQ-imbalances. This proposed FD-LS estimates both destination-to-destination ($D \rightarrow D$) and source-to-destination ($S \rightarrow D$) channels. The estimated $D \rightarrow D$ channel is for self-interference cancellation and the estimated $S \rightarrow D$ channel is used for detection. Here, it is noted the estimated channel parameters consist of the combined effect of IQ-imbalance and channel impulse response (CIR).
2) Following the above, we model the problem of optimal training design for the proposed FD-LS channel estimator as a convex optimization problem of minimizing the sum of variances given the power constraint. When the average transmit powers from both source and destination are identical, the closed-form expression of the optimal pilot matrix is proved to be any four columns of a unitary matrix and a scalar factor, where the factor depends on the average transmit power and the number of pilot OFDM symbols.
3) The performance of the proposed FD-LS channel estimator is further improved by taking advantage of the time-domain property of channel. The estimated frequency-domain channel gains are transformed to the time-domain channel by using FFT, and converted backed into the frequency-domain by IFFT after all the channel tap gains outside cyclic prefix (CP) are forced to zeros. This operation is proved to achieve an N/L times signal-to-noise gain for independent, identically distributed (iid) Gaussian noise in Appendix A.
4) Taking both residual self-interference and IQ imbalance into account an entity after serial full-duplex self-interference cancelation, their covariance matrix is derived, the whitening filter (WF) is designed to whiten the residual self-interference and noise to a white Gaussian vector in order to enhance the performance of maximum likelihood (ML) detector at destination receiver. In the WF process, eigen-value decomposition is used to compute the WF coefficients and singular value decomposition (SVD) is adopted to reduce the complexity of the whitening-filter-based maximum likelihood (WFML) detector.

This paper is organized as follows. The system model with full-duplex and IQ imbalance is described in Section II.

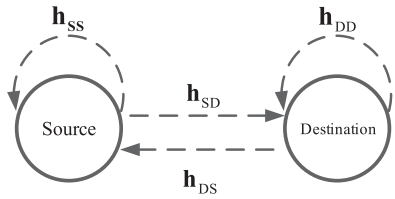


Fig. 1. Full-duplex system model.

Section III presents the design of the FD-LS channel estimator to estimate both $S \rightarrow D$ and $D \rightarrow D$ channels. Its sum of variances is computed by the Fisher information matrix. In Section IV, an optimal pilot pattern is proposed for the proposed FD-LS estimator by minimizing the sum of variances. In Section V, we show how to improve the performance of the proposed FD-LS by exploiting the time-domain property, and the WFML detector is designed to convert the colored noise vector plus residual full-duplex self-interference, whose covariance matrix is derived in Appendix B, into a white noise vector. Simulation results and discussions are presented in Section VI. Finally, Section VII concludes this paper.

Notations: Throughout the paper, matrices and vectors are denoted by letters of bold upper case and bold lower case, respectively. Signs $(\bullet)^H$, $(\bullet)^*$, $(\bullet)^T$, $(\bullet)^{-1}$, tr$(\bullet)$, $\| \bullet \|_F$, and det$(\bullet)$ denote matrix conjugate transpose, conjugate, transpose, inverse, trace, norm-2, and determinant, respectively. The notation $\mathcal{E}\{\bullet\}$ refers to the expectation operation. The symbol $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. $\mathbf{0}_{n \times m}$ denotes an all-zero matrix of size $n \times m$. $\otimes$ denotes the Kronecker product of two matrices. vec$(\mathbf{X})$ is an operation of stacking all columns of $\mathbf{X}$ to a large column vector. diag $\{\mathbf{a}\}$ denotes an operation of placing all elements of the vector $\mathbf{a}$ over the diagonal of diagonal matrix.

II. System Model

We consider a point-to-point full-duplex OFDM wireless communication link with IQ-imbalances as shown in Fig. 1. In the figure, there are two nodes: source and destination nodes. Both source and destination operate on full-duplex model that allows simultaneous transmission and reception over the same time and frequency band. Destination node is chosen as our research object. Moreover, both destination and source nodes use the direct transmitter and receiver. This means there exist both full-duplex self-interference and IQ-imbalance distortion at destination receiver.

The transmit vectors of data symbols corresponding to the nth OFDM symbol from source to destination, and destination to destination are expressed as

$$\mathbf{x}_S(n,:) = [x_S(n,1)\, x_S(n,2)\, \cdots\, x_S(n,N)]^T \tag{1}$$

and

$$\mathbf{x}_D(n,:) = [x_D(n,1)\, x_D(n,2)\, \cdots\, x_D(n,N)]^T \tag{2}$$

respectively, where N is the total number of subcarriers. Their IDFTs can be defined as

$$\overline{\mathbf{x}}_S(n,:) = \mathbf{F}^H_{N\times N}\mathbf{x}_S(n,:) \tag{3}$$

and

$$\overline{\mathbf{x}}_D(n,:) = \mathbf{F}^H_{N\times N}\mathbf{x}_D(n,:) \tag{4}$$

where $\mathbf{F}_{N\times N}$ is the normalized N-point discrete Fourier transform matrix

$$\mathbf{F}_{N\times N} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \exp\left\{\frac{-2\pi j(1\times 1)}{N}\right\} & \cdots & \exp\left\{\frac{-2\pi j[1\times(N-1)]}{N}\right\} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \exp\left\{\frac{-2\pi j[(N-1)\times 1]}{N}\right\} & \cdots & \exp\left\{\frac{-2\pi j[(N-1)^2]}{N}\right\} \end{pmatrix}. \tag{5}$$

After cyclic prefix (CP) is inserted before the signal $\overline{\mathbf{x}}_S(n,:)/\overline{\mathbf{x}}_D(n,:)$, where CP is larger than the sum of the maximum propagation delay and maximum channel-tap length of the desired and the interference channels in slow-fading wireless environments, the new formed signal experiences through transmit IQ-imbalance, multipath channel, receive IQ-imbalance, and full-duplex interference, finally we obtain the receive frequency-domain data-symbol vector corresponding to the nth OFDM symbol at destination as

$$\begin{aligned} \mathbf{y}(n,:) = {} & \mathrm{diag}\{\underbrace{\mu_{r,D}\mu_{t,S}\mathbf{H}^n_{SD} + \nu_{r,D}\nu^*_{t,S}\,(\mathbf{H}^n_{SD})^{\#}}_{\mathbf{H}^{a,n}_{SD}}\}\mathbf{x}_S(n,:) \\ & + \mathrm{diag}\{\underbrace{\mu_{r,D}\nu_{t,S}\mathbf{H}^n_{SD} + \nu_{r,D}\mu^*_{t,S}\,(\mathbf{H}^n_{SD})^{\#}}_{\mathbf{H}^{b,n}_{SD}}\}\mathbf{x}^{\#}_S(n,:) \\ & + \mathrm{diag}\{\underbrace{\mu_{r,D}\mu_{t,D}\mathbf{H}^n_{DD} + \nu_{r,D}\nu^*_{t,D}\,(\mathbf{H}^n_{DD})^{\#}}_{\mathbf{H}^{a,n}_{DD}}\}\mathbf{x}_D(n,:) \\ & + \mathrm{diag}\{\underbrace{\mu_{r,D}\nu_{t,D}\mathbf{H}^n_{DD} + \nu_{r,D}\mu^*_{t,D}\,(\mathbf{H}^n_{DD})^{\#}}_{\mathbf{H}^{b,n}_{DD}}\}\mathbf{x}^{\#}_D(n,:) \\ & + \mathbf{w}(n,:) \end{aligned} \tag{6}$$

where the superscript $\#$ is defined as

$$\begin{aligned} \mathbf{X}^{\#} = {} & [X^*(1)\,X^*(N)\,\cdots\,X^*(N/2+2)\,X^*(N/2+1) \\ & \times\, X^*(N/2)\,\cdots\,X^*(2)]^T \end{aligned} \tag{7}$$

with $\mathbf{X} = [X(1)\ X(2)\ \cdots\ X(N/2)\ X(N/2+1)\ X(N/2+2)\ \cdots\ X(N)]^T$

$$\mu_{t,S} = \cos(\theta_{t,S}/2) + j\alpha_{t,S}\sin(\theta_{t,S}/2)$$

$$\nu_{t,S} = \alpha_{t,S}\cos(\theta_{t,S}/2) - j\sin(\theta_{t,S}/2) \tag{8}$$

$$\mu_{t,D} = \cos(\theta_{t,D}/2) + j\alpha_{t,D}\sin(\theta_{t,D}/2)$$

$$\nu_{t,D} = \alpha_{t,D}\cos(\theta_{t,D}/2) - j\sin(\theta_{t,D}/2) \tag{9}$$

$$\mu_{r,D} = \cos(\theta_{r,D}/2) + j\alpha_{r,D}\sin(\theta_{r,D}/2)$$

$$\nu_{r,D} = \alpha_{r,D}\cos(\theta_{r,D}/2) - j\sin(\theta_{r,D}/2) \tag{10}$$

$$\mathbf{H}^n_{SD} = \mathbf{F}_{N\times N}\begin{pmatrix}\mathbf{h}^n_{SD}\\ \mathbf{0}_{(N-L)\times 1}\end{pmatrix} \tag{11}$$

$$\mathbf{H}^n_{DD} = \mathbf{F}_{N\times N}\begin{pmatrix}\mathbf{h}^n_{DD}\\ \mathbf{0}_{(N-L)\times 1}\end{pmatrix} \tag{12}$$

and

$$\mathbf{w}(n,:) = \mu_{r,D}\mathbf{v}(n,:) + \nu_{r,D}\mathbf{v}^{\#}(n,:) \tag{13}$$

where $\theta_{t,S}$ and $\alpha_{t,S}$ are phase and amplitude imbalance between I and Q branches at source transmitter, $\theta_{t,D}$ and $\alpha_{t,D}$, and $\theta_{r,D}$ and $\alpha_{r,D}$ are phase and amplitude imbalances between I and Q branches at destination transmitter and receiver, respectively. $\mathbf{h}^n_{SD} = [h^n_{SD}(1)\,h^n_{SD}(2)\,\cdots\,h^n_{SD}(L)]^T$ is the CIR from source to destination, $\mathbf{h}^n_{DD} = [h^n_{DD}(1)\,h^n_{DD}(2)\,\cdots\,h^n_{DD}(L)]^T$ is the CIR from destination to destination where L is the length of CP. $\mathbf{v}(n,:)$ is the additive white Gaussian noise (AWGN) vector over nth OFDM symbol with each element being zero mean and variance $\sigma^2_{\mathbf{v}}$ in frequency domain.

In the right-hand side of (6), the first term is the useful received signal, the second one is the IQ-imbalance distortion, and the third and fourth terms are the full-duplex self-interference.

III. Proposed Frequency-Domain Channel Estimator

Observing the system model as shown in (6), if we can estimate the four vectors of frequency-domain channel gain vectors (CGVs) $\mathbf{H}^{a,n}_{SD}$, $\mathbf{H}^{b,n}_{SD}$, $\mathbf{H}^{a,n}_{DD}$, and $\mathbf{H}^{b,n}_{DD}$ at destination, then we recover the transmit data vectors $\mathbf{x}_S(n,:)$ by some typical detection algorithms like zero-forcing, and maximum likelihood after the full-duplex interference, the third and fourth terms on the right-hand side of (6), are cancelled by using the estimated $\mathbf{H}^{a,n}_{DD}$ and $\mathbf{H}^{b,n}_{DD}$. In this section, we will show how to estimate $\mathbf{H}^{a,n}_{SD}$, $\mathbf{H}^{b,n}_{SD}$, $\mathbf{H}^{a,n}_{DD}$, and $\mathbf{H}^{b,n}_{DD}$. Here, $\mathbf{H}^{a,n}_{DD}$, and $\mathbf{H}^{b,n}_{DD}$ are called the self-interference channel while $\mathbf{H}^{a,n}_{SD}$, and $\mathbf{H}^{b,n}_{SD}$ are called the intended channel.

In terms of (6), the received data symbol over the kth subcarrier of the nth OFDM symbol can be written as follows:

$$\begin{aligned} y(n,k) = {} & H^{a,n}_{SD}(k)x_S(n,k) + H^{b,n}_{SD}(k)x^*_S(n,N-k+2) \\ & + H^{a,n}_{DD}(k)x_D(n,k) + H^{b,n}_{DD}(k)x^*_D(n,N-k+2) \\ & + w(n,k) \end{aligned} \tag{14}$$

with

$$H_{SD}^{a,n}(k) = \mu_{r,D}\mu_{t,S}H_{SD}^{n}(k) + \nu_{r,D}\nu_{t,S}^{*} \times \left(H_{SD}^{n}(N-k+2)\right)^{*} \quad (15)$$

$$H_{SD}^{b,n}(k) = \mu_{r,D}\nu_{t,S}H_{SD}^{n}(k) + \nu_{r,D}\mu_{t,S}^{*} \times \left(H_{SD}^{n}(N-k+2)\right)^{*} \quad (16)$$

$$H_{DD}^{a,n}(k) = \mu_{r,D}\mu_{t,D}H_{DD}^{n}(k) + \nu_{r,D}\nu_{t,D}^{*} \times \left(H_{DD}^{n}(N-k+2)\right)^{*} \quad (17)$$

and

$$H_{DD}^{b,n}(k) = \mu_{r,D}\nu_{t,D}H_{DD}^{n}(k) + \nu_{r,D}\mu_{t,D}^{*} \times \left(H_{DD}^{n}(N-k+2)\right)^{*}. \quad (18)$$

Similar to (14), we have

$$\begin{aligned} y^{*}(n,N-k+2) &= \left(H_{SD}^{a,n}(N-k+2)\right)^{*} x_{S}^{*}(n,N-k+2) \\ &+ \left(H_{SD}^{b,n}(N-k+2)\right)^{*} x_{S}(n,k) \\ &+ \left(H_{DD}^{a,n}(N-k+2)\right)^{*} x_{D}^{*}(n,N-k+2) \\ &+ \left(H_{DD}^{b,n}(N-k+2)\right)^{*} x_{D}(n,k) + w^{*}(n,N-k+2). \end{aligned} \quad (19)$$

In the following, we call the two subchannels k and $N-k+2$ as the kth subchannel/subcarrier pair. Combining (14) and (19) yields

$$\mathbf{z}^{n,k} = \mathbf{\Gamma}_{SD}^{k}\mathbf{x}_{S}^{n,k} + \mathbf{\Gamma}_{DD}^{k}\mathbf{x}_{D}^{n,k} + \mathbf{w}^{n,k} \quad (20)$$

where

$$\mathbf{z}^{n,k} = \begin{pmatrix} y(n,k) \\ y^{*}(n,N-k+2) \end{pmatrix} \quad (21)$$

$$\mathbf{x}_{S}^{n,k} = \begin{pmatrix} x_{S}(n,k) \\ x_{S}^{*}(n,N-k+2) \end{pmatrix} \quad (22)$$

$$\mathbf{x}_{D}^{n,k} = \begin{pmatrix} x_{D}(n,k) \\ x_{D}^{*}(n,N-k+2) \end{pmatrix} \quad (23)$$

$$\mathbf{\Gamma}_{SD}^{n,k} = \begin{pmatrix} H_{SD}^{a,n}(k) & H_{SD}^{b,n}(k) \\ \left(H_{SD}^{b,n}(N-k+2)\right)^{*} & \left(H_{SD}^{a,n}(N-k+2)\right)^{*} \end{pmatrix} \quad (24)$$

$$\mathbf{\Gamma}_{DD}^{n,k} = \begin{pmatrix} H_{DD}^{a,n}(k) & H_{DD}^{b,n}(k) \\ \left(H_{DD}^{b,n}(N-k+2)\right)^{*} & \left(H_{DD}^{a,n}(N-k+2)\right)^{*} \end{pmatrix} \quad (25)$$

and

$$\mathbf{w}^{n,k} = \begin{pmatrix} w(n,k) \\ w^{*}(n,N-k+2) \end{pmatrix}. \quad (26)$$

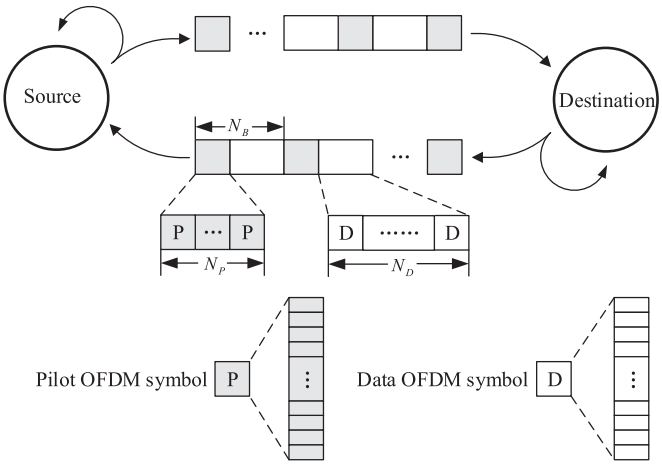


Fig. 2. Pilot pattern for the full-duplex OFDM system.

Applying the vec$(\bullet)$ operator to both sides of (20) yields

$$\begin{aligned} \mathbf{z}^{n,k} &= \text{vec}\left(\mathbf{z}^{n,k}\right) \\ &= \text{vec}\left(\mathbf{\Gamma}_{SD}^{n,k}\mathbf{x}_{S}^{n,k}\right) + \text{vec}\left(\mathbf{\Gamma}_{DD}^{n,k}\mathbf{x}_{D}^{n,k}\right) + \text{vec}\left(\mathbf{w}^{n,k}\right) \\ &= \text{vec}\left(\mathbf{I}_{2}\mathbf{\Gamma}_{SD}^{n,k}\mathbf{x}_{S}^{n,k}\right) + \text{vec}\left(\mathbf{I}_{2}\mathbf{\Gamma}_{DD}^{n,k}\mathbf{x}_{D}^{n,k}\right) + \text{vec}\left(\mathbf{w}^{n,k}\right) \end{aligned} \quad (27)$$

with $k \in \{2,3,\ldots,N/2\}$. By using the property of vec operator in [23]

$$\text{vec}\left(\mathbf{AYB}\right) = \left(\mathbf{B}^{T} \otimes \mathbf{A}\right)\text{vec}\left(\mathbf{Y}\right) \quad (28)$$

we can rewrite (27) as

$$\begin{aligned} \mathbf{z}^{n,k} &= \left[\left(\mathbf{x}_{S}^{n,k}\right)^{T} \otimes \mathbf{I}_{2}\right]\text{vec}\left(\mathbf{\Gamma}_{SD}^{n,k}\right) + \left[\left(\mathbf{x}_{D}^{n,k}\right)^{T} \otimes \mathbf{I}_{2}\right] \\ &\times \text{vec}\left(\mathbf{\Gamma}_{DD}^{n,k}\right) + \mathbf{w}^{n,k} = \underbrace{\left[\left(\mathbf{x}_{S}^{n,k}\right)^{T} \otimes \mathbf{I}_{2} \quad \left(\mathbf{x}_{D}^{n,k}\right)^{T} \otimes \mathbf{I}_{2}\right]}_{\mathbf{\Gamma}^{n,k}} \begin{bmatrix} \text{vec}\left(\mathbf{\Gamma}_{SD}^{n,k}\right) \\ \text{vec}\left(\mathbf{\Gamma}_{DD}^{n,k}\right) \end{bmatrix} + \mathbf{w}^{n,k}. \end{aligned} \quad (29)$$

Apparently, the above equation is under-determined because we use two equations to compute eight unknowns. Obviously, we need at least eight equations to calculate eight unknowns. This implies that at least four pilot OFDM symbols are required to do one-time frequency-domain channel estimation.

As shown in Fig. 2, each frame consists of N_F blocks and pilot sub-block of N_P pilot OFDM symbols. Each block is made up of N_B OFDM symbols. In each block, the first N_P OFDM symbols (pilot sub-block) are used as pilot OFDM symbols and the remaining N_D ones (data sub-block) are data OFDM symbols, where $N_B = N_P + N_D$, and N_P is larger than or equal to four according to the above analysis. In terms of N_B, N_P, and N_D, for the nth OFDM symbol, we may determine it belongs

to which block and whether data or pilot. By using the simple integer division, i.e., the Euclidean algorithm, we have

$$n = (m-1)N_B + q \tag{30}$$

where q and m denote the remainder and quotient when n is divided by N_B. Here, $m \in \{1, 2, \ldots, N_F\}$ is the block index and means that the nth OFDM symbol belongs to block m. If $q \in \{1, 2, \ldots, N_P\}$, then the corresponding OFDM symbol is one pilot OFDM symbol. Otherwise, it is one data OFDM symbol. Obviously, to achieve a high-spectrum efficiency, N_D should be taken to be far larger than N_P.

Below, the channel is assumed to be constant or slow fading during N_P continuous pilot or data OFDM symbols. Stacking all the kth subcarrier pairs of the N_P pilot OFDM symbols in block m forms a large receive vector

$$\begin{aligned}\mathbf{z}_P^{m,k} &= \begin{pmatrix} \mathbf{z}^{mN_B+1,k} \\ \mathbf{z}^{mN_B+2,k} \\ \vdots \\ \mathbf{z}^{mN_B+N_P,k} \end{pmatrix} \\ &= \begin{pmatrix} \left(\mathbf{x}_S^{mN_B+1,k}\right)^T \otimes \mathbf{I}_2 & \left(\mathbf{x}_D^{mN_B+1,k}\right)^T \otimes \mathbf{I}_2 \\ \left(\mathbf{x}_S^{mN_B+2,k}\right)^T \otimes \mathbf{I}_2 & \left(\mathbf{x}_D^{mN_B+2,k}\right)^T \otimes \mathbf{I}_2 \\ \vdots & \vdots \\ \left(\mathbf{x}_S^{mN_B+N_P,k}\right)^T \otimes \mathbf{I}_2 & \left(\mathbf{x}_D^{mN_B+N_P,k}\right)^T \otimes \mathbf{I}_2 \end{pmatrix} \\ &\quad \times \begin{pmatrix} \mathrm{vec}\left(\mathbf{\Gamma}_{SD,P}^{m,k}\right) \\ \mathrm{vec}\left(\mathbf{\Gamma}_{DD,P}^{m,k}\right) \end{pmatrix} + \begin{pmatrix} \mathbf{w}^{mN_B+1,k} \\ \mathbf{w}^{mN_B+2,k} \\ \vdots \\ \mathbf{w}^{mN_B+N_P,k} \end{pmatrix} \\ &= \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)\mathbf{\Gamma}_P^{m,k} + \mathbf{w}_P^{m,k} \end{aligned} \tag{31}$$

with

$$\mathbf{X}_P^{m,k} = \begin{pmatrix} \mathbf{x}_S^{mN_B+1,k} & \mathbf{x}_S^{mN_B+2,k} & \cdots & \mathbf{x}_S^{mN_B+N_P,k} \\ \mathbf{x}_D^{mN_B+1,k} & \mathbf{x}_D^{mN_B+2,k} & \cdots & \mathbf{x}_D^{mN_B+N_P,k} \end{pmatrix}^T \tag{32}$$

$$\mathbf{\Gamma}_P^{m,k} = \begin{pmatrix} \mathrm{vec}\left(\mathbf{\Gamma}_{SD,P}^{m,k}\right) \\ \mathrm{vec}\left(\mathbf{\Gamma}_{DD,P}^{m,k}\right) \end{pmatrix} \tag{33}$$

$$\mathbf{w}_P^{m,k} = \begin{pmatrix} \mathbf{w}^{mN_B+1,k} \\ \mathbf{w}^{mN_B+2,k} \\ \vdots \\ \mathbf{w}^{mN_B+N_P,k} \end{pmatrix} \tag{34}$$

where N_P should be chosen to be larger than or equal to four in order to successfully complete both $S \to D$ and $D \to D$ channel estimation, and $\mathbf{X}_P^{m,k}$ is called an $N_P \times 4$ pilot matrix. Given $\mathbf{X}_P^{m,k}$ and $\mathbf{\Gamma}_P^{m,k}$, the probability density function of the received vector $\mathbf{z}_P^{m,k}$ will be determined by the covariance of the noise vector $\mathbf{w}_P^{m,k}$. The noise covariance matrix is calculated as

$$\mathbf{C}_{\mathbf{w}_P^{m,k}\mathbf{w}_P^{m,k}} = \mathcal{E}\left[\mathbf{w}_P^{m,k}\left(\mathbf{w}_P^{m,k}\right)^H\right] = \mathbf{I}_{N_P} \otimes \mathbf{C}_{2\times 2} \tag{35}$$

with $\mathbf{C}_{2\times 2} = \mathbf{C}_{\mathbf{w}^{n,k}\mathbf{w}^{n,k}}$ denoted as

$$\begin{aligned}\mathbf{C}_{2\times 2} &= \mathcal{E}\left[\mathbf{w}^{n,k}\left(\mathbf{w}^{n,k}\right)^H\right] \\ &= \sigma_{\mathbf{v}}^2 \begin{pmatrix} \mu_{r,D}\mu_{r,D}^* + \nu_{r,D}\nu_{r,D}^* & 2\mu_{r,D}\nu_{r,D} \\ 2\mu_{r,D}^*\nu_{r,D}^* & \mu_{r,D}\mu_{r,D}^* + \nu_{r,D}\nu_{r,D}^* \end{pmatrix}. \end{aligned} \tag{36}$$

From (31), given $\mathbf{\Gamma}_P^{m,k}$ and $\mathbf{X}_P^{m,k}$, the likelihood function is

$$\begin{aligned}\mathrm{p}(\mathbf{z}_P^{m,k}|\mathbf{\Gamma}_P^{m,k}, \mathbf{X}_P^{m,k}) &= \frac{1}{[2\pi \det(\mathbf{C}_{2\times 2})]^{N_P}} \\ &\quad \times \exp\left\{-\frac{1}{2}\left[\mathbf{z}_P^{m,k} - \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)\mathbf{\Gamma}_P^{m,k}\right]^T \right. \\ &\quad \left. \times \left(\mathbf{I}_{N_P} \otimes \mathbf{C}_{2\times 2}^{-1}\right)\left[\mathbf{z}_P^{m,k} - \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)\mathbf{\Gamma}_P^{m,k}\right]^*\right\}. \end{aligned} \tag{37}$$

If we consider the natural logarithm of the PDF

$$\begin{aligned}\ln \mathrm{p}(\mathbf{z}_P^{m,k}|\mathbf{\Gamma}_P^{m,k}, \mathbf{X}_P^{m,k}) &= N_P \ln 2\pi \det(\mathbf{C}_{2\times 2}) - \frac{1}{2} \\ &\quad \times \left[\mathbf{z}_P^{m,k} - \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)\mathbf{\Gamma}_P^{m,k}\right]^T \left(\mathbf{I}_{N_P} \otimes \mathbf{C}_{2\times 2}^{-1}\right) \\ &\quad \times \left[\mathbf{z}_P^{m,k} - \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)\mathbf{\Gamma}_P^{m,k}\right]^* \end{aligned} \tag{38}$$

from which the first derivative follows as

$$\begin{aligned}\frac{\partial \ln \mathrm{p}(\mathbf{z}_P^{m,k}|\mathbf{\Gamma}_P^{m,k}, \mathbf{X}_P^{m,k})}{\partial\left(\mathbf{\Gamma}_P^{m,k}\right)^*} &= \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)^H \left(\mathbf{I}_{N_P} \otimes \mathbf{C}_{2\times 2}^{-1}\right) \\ &\quad \times \left[\mathbf{z}_P^{m,k} - \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)\mathbf{\Gamma}_P^{m,k}\right]. \end{aligned} \tag{39}$$

Assuming that $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ is invertible, setting the above partial derivative to zero produces

$$\hat{\mathbf{\Gamma}}_{P,FD-LS}^{m,k} = \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)^+ \mathbf{z}_P^{m,k} \tag{40}$$

which is called the frequency-domain least square (FD-LS) channel estimator below and is an unbiased estimator. If and only if the pilot matrix $\mathbf{X}_P^{m,k}$ in (40) is a full-rank matrix, the above FD-LS estimator can successfully perform channel estimation. Otherwise, channel estimation fails, where A^+ is the Moore–Penrose pseudoinverse of matrix A. When matrix A is rank full, $A^+ = (A^HA)^{-1}A^H$. Using the property

$$\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A})\,\mathrm{tr}(\mathbf{B}) \tag{41}$$

we have the Fisher information matrix [24]

$$\begin{aligned}\mathbf{I}\left(\mathbf{\Gamma}_P^{m,k}\right) &= \mathcal{E}\left(\frac{\partial^2 \ln \mathrm{p}(\mathbf{z}_P^{m,k}|\mathbf{\Gamma}_P^{m,k}, \mathbf{X}_P^{m,k})}{\partial\left(\mathbf{\Gamma}_P^{m,k}\right)^*\partial\left(\mathbf{\Gamma}_P^{m,k}\right)}\right) \\ &= \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)^H \left(\mathbf{I}_{N_P} \otimes \mathbf{C}_{2\times 2}^{-1}\right)\left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right) \\ &= \left[\left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}\right] \otimes \mathbf{C}_{2\times 2}^{-1}. \end{aligned} \tag{42}$$

Upon inverting the Fisher information, we have

$$\mathbf{I}^{-1}\left(\mathbf{\Gamma}_P^{m,k}\right) = \left[\left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}\right]^{-1} \otimes \mathbf{C}_{2\times 2}. \tag{43}$$

Using the above matrix, the Cramer-Rao lower bound (CRLB) of the sum of variances of all estimated elements of $\mathbf{\Gamma}_P^{m,k}$ is given by

$$\begin{aligned}\sum_{i=1}^{8} \text{var}(\mathbf{\Gamma}_P^{m,k}(i)) &\geq \text{tr}\left[\mathbf{I}^{-1}\left(\mathbf{\Gamma}_P^{m,k}\right)\right] \\ &= \text{tr}\left\{\left[\left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}\right]^{-1}\right\} \text{tr}\left(\mathbf{C}_{2\times 2}\right)\end{aligned} \tag{44}$$

where $\mathbf{\Gamma}_P^{m,k}(i)$ is the ith element of the channel vector $\mathbf{\Gamma}_P^{m,k}$. Considering the unbiased estimator as shown in (40), the CRLB of the sum of variances of all estimated elements of $\mathbf{\Gamma}_P^{m,k}$ is shown to be achievable by the sum of mean square error (MSE) of the proposed FD-LS in the next section.

IV. Joint Optimization Design of Pilot Pattern

In the previous section, an FD-LS channel estimator is constructed as indicated in (40). However, the pilot matrix $\mathbf{X}_P^{m,k}$ in (40) will have a profound impact on the performance of the FD-LS estimator. What we concern is the problem that the FD-LS minimizes the sum of square of estimation error under what condition the pilot matrix $\mathbf{X}_P^{m,k}$ satisfies. In what follows, we will address the problem of minimizing the sum of variances as a convex optimization. In the case of $P_S = P_D$, where P_S and P_D is the average transmit power per subcarrier of source and destination nodes, respectively, we derive a closed-form expression for the optimal pilot matrix.

A. Sum of MSE of the Proposed FD-LS

Substituting (31) in (40) yields

$$\begin{aligned}\hat{\mathbf{\Gamma}}_{P,FD-LS}^{m,k} &= \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)^+ \mathbf{z}_P^{m,k} \\ &= \mathbf{\Gamma}_P^{m,k} + \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)^+ \mathbf{w}_P^{m,k}.\end{aligned} \tag{45}$$

Let us define the channel estimation error as

$$\begin{aligned}\Delta\mathbf{\Gamma}_{P,FD-LS}^{m,k} &= \hat{\mathbf{\Gamma}}_{P,FD-LS}^{m,k} - \mathbf{\Gamma}_P^{m,k} \\ &= \left(\mathbf{X}_P^{m,k} \otimes \mathbf{I}_2\right)^+ \mathbf{w}_P^{m,k}\end{aligned} \tag{46}$$

which forms the sum of MSEs of the proposed FD-LS as follows

$$\begin{aligned}\text{Sum-MSE}^{m,k} &= \mathcal{E}\left\{\text{tr}\left[\Delta\mathbf{\Gamma}_{P,FD-LS}^{m,k}(\Delta\mathbf{\Gamma}_{P,FD-LS}^{m,k})^H\right]\right\} \\ &= \text{tr}\left\{\left(\left[\left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}\right]^{-1} \otimes \mathbf{I}_2\right) \mathcal{E}\left(\mathbf{w}_P^{m,k}\left(\mathbf{w}_P^{m,k}\right)^H\right)\right\}.\end{aligned} \tag{47}$$

Using (35) and the property

$$(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD}) \tag{48}$$

(47) is reduced to the simple form

$$\begin{aligned}\text{Sum-MSE}^{m,k} &= \text{tr}\left\{\left(\left[\left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}\right]^{-1} \otimes \mathbf{I}_2\right) \times \mathcal{E}\left[\mathbf{w}_P^{m,k}\left(\mathbf{w}_P^{m,k}\right)^H\right]\right\} \\ &= \text{tr}\left\{\left[\left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}\right]^{-1} \otimes \mathbf{C}_{2\times 2}\right\} \\ &= \text{tr}\left(\mathbf{C}_{2\times 2}\right) \text{tr}\left\{\left[\left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}\right]^{-1}\right\}\end{aligned} \tag{49}$$

which agrees with (44). Thus, the proposed FD-LS is a linear unbias minimum variance estimator.

B. Minimizing Sum of MSEs

Obviously, to make the proposed channel estimator in (40) achieve a good performance, we should minimize the sum of variances in (49) by optimizing the pilot matrix $\mathbf{X}_P^{m,k}$ with limit on transmit power of source and destination nodes. This can be formulated as the following convex optimization problem

$$\begin{aligned}&\text{minimize} \quad \text{Sum-MSE}^{m,k} \\ &\text{subject to } \text{tr}\left[\mathbf{E}_S^H \left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k} \mathbf{E}_S\right] \leq 2N_P P_S \\ &\qquad\quad \text{tr}\left[\mathbf{E}_D^H \left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k} \mathbf{E}_D\right] \leq 2N_P P_D\end{aligned} \tag{50}$$

where matrix $\mathbf{X}_P^{m,k}$ is the optimization variable,

$$\mathbf{E}_S = \begin{pmatrix}\mathbf{I}_2 & \mathbf{0}_{2\times 2}\end{pmatrix}^T \tag{51}$$

and

$$\mathbf{E}_D = \begin{pmatrix}\mathbf{0}_{2\times 2} & \mathbf{I}_2\end{pmatrix}^T. \tag{52}$$

If we define $\mathbf{Y} = \left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k}$, called pilot product matrix below, the optimization problem in (50) reduces to

$$\begin{aligned}&\text{minimize} \quad \text{tr}\left(\mathbf{C}_{2\times 2}\right) \text{tr}\left(\mathbf{Y}^{-1}\right) \\ &\text{subject to } \text{tr}\left[\mathbf{E}_S^H \mathbf{Y} \mathbf{E}_S\right] \leq 2N_P P_S \\ &\qquad\quad \text{tr}\left[\mathbf{E}_D^H \mathbf{Y} \mathbf{E}_D\right] \leq 2N_P P_D \\ &\qquad\quad \mathbf{Y} \succ \mathbf{0}\end{aligned} \tag{53}$$

which is a convex optimization and can be solved by interior-point method [25] because the function $\text{tr}\left(\mathbf{Y}^{-1}\right)$ is convex in accordance with [26, Th. 7.6.10]. From the definition of $\mathbf{Y}$, it is evident matrix $\mathbf{Y}$ is positive definite and can be eigen-value decomposed as

$$\mathbf{Y} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \tag{54}$$

where $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}_4$, and matrix $\mathbf{\Lambda}$ is a diagonal matrix with diagonal element i being $\lambda_i > 0$. Here, λ_i is the ith eigenvalue of matrix $\mathbf{Y}$. Then, the sum of MSE can be

rewritten as

$$\text{Sum-MSE}^{m,k} = \text{tr}\left(\mathbf{C}_{2\times 2}\right) \sum \frac{1}{\lambda_i}. \tag{55}$$

Plugging (54) in (53) and removing the constant in objective function yields

$$\begin{aligned} &\text{minimize} \quad \text{tr}(\mathbf{\Lambda}^{-1}) \\ &\text{subject to } \text{tr}\left[\left(\mathbf{U}^H \mathbf{E}_S \mathbf{E}_S^H \mathbf{U}\right) \mathbf{\Lambda}\right] \le 2N_P P_S \\ &\qquad \text{tr}\left[\left(\mathbf{U}^H \mathbf{E}_D \mathbf{E}_D^H \mathbf{U}\right) \mathbf{\Lambda}\right] \le 2N_P P_D \\ &\qquad \mathbf{\Lambda} \succ \mathbf{0}, \end{aligned} \tag{56}$$

which can be rewritten as

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{4} \frac{1}{\lambda_i} \\ &\text{subject to } \sum_{i=1}^{4} \left(\mathbf{U}^H \mathbf{E}_S \mathbf{E}_S^H \mathbf{U}\right)_{ii} \lambda_i \le 2N_P P_S \\ &\qquad \sum_{i=1}^{4} \left(\mathbf{U}^H \mathbf{E}_D \mathbf{E}_D^H \mathbf{U}\right)_{ii} \lambda_i \le 2N_P P_D \\ &\qquad \lambda_i > 0, \forall i \in \{1,2,3,4\}. \end{aligned} \tag{57}$$

It is not easy to find a closed-form solution to the optimization problem above. However, when $P_S = P_D$, the above optimization problem has a closed-form solution. Below, we present the deriving process of the closed-form solution. First, the above problem is relaxed as

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{4} \frac{1}{\lambda_i} \\ &\text{subject to } \sum_{i=1}^{4} \left[\left(\mathbf{U}^H \mathbf{E}_S \mathbf{E}_S^H \mathbf{U}\right)_{ii} + \left(\mathbf{U}^H \mathbf{E}_D \mathbf{E}_D^H \mathbf{U}\right)_{ii}\right] \lambda_i \\ &\qquad \le 2N_P \left(P_S + P_D\right) \\ &\qquad \lambda_i > 0, \forall i \in \{1,2,3,4\} \end{aligned} \tag{58}$$

which has the simple form as

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{4} \frac{1}{\lambda_i} \\ &\text{subject to } \sum_{i=1}^{4} \lambda_i \le 2N_P \left(P_S + P_D\right) \\ &\qquad \lambda_i > 0, \forall i \in \{1,2,3,4\}. \end{aligned} \tag{59}$$

Removing the second inequality in the above optimization yields

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{4} \frac{1}{\lambda_i} \\ &\text{subject to } \sum_{i=1}^{4} \lambda_i \le 2N_P \left(P_S + P_D\right). \end{aligned} \tag{60}$$

Solving the above optimization, we define the associated Lagrangian function [25]

$$f(\mathbf{\Lambda}, \mu) = \sum_{i=1}^{4} \frac{1}{\lambda_i} + \mu \left[\sum_{i=1}^{4} \lambda_i - 2N_P \left(P_S + P_D\right)\right]. \tag{61}$$

Taking the first-order derivative of the above function $f(\mathbf{\Lambda}, \mu)$ with respect to λ_i and setting it to zero, we have

$$\frac{f(\mathbf{\Lambda}, \mu)}{\partial \lambda_i} = -\frac{1}{\lambda_i^2} + \mu = 0 \tag{62}$$

then,

$$\lambda_i = \frac{1}{\sqrt{\mu}}, \forall i \in \{1,2,3,4\}. \tag{63}$$

Placing the above in the constraint of the problem (60) gives

$$\lambda_i = \frac{N_P}{2}(P_S + P_D), \forall i \in \{1,2,3,4\} \tag{64}$$

which means

$$\mathbf{Y} = \left(\mathbf{X}_P^{m,k}\right)^H \mathbf{X}_P^{m,k} = \frac{N_P}{2}(P_S + P_D)\mathbf{I}_4. \tag{65}$$

Obviously, the optimal matrix $\mathbf{Y}$ above is a feasible solution to the optimization problems (58) and (59). In the case of $P_S = P_D$, the above solution also satisfies the two tighter constraint conditions of the optimization problem (57). Thus, we claim that we obtain the closed-form solution to the original problem for $N_P \ge 4$ and $P_S = P_D$. However, $N_P < 4$ will lead to a singular pilot product matrix $(\mathbf{X}_P^{m,k})^H \mathbf{X}_P^{m,k}$.

C. Theorem, Application, and Analysis

We summarize the above result in an important theorem.

Theorem 1: In a full-duplex OFDM system with IQ imbalances, the optimal pilot matrix $\mathbf{X}_P^{m,k}$ satisfies the condition $(\mathbf{X}_P^{m,k})^H \mathbf{X}_P^{m,k} = \frac{N_P}{2}(P_S + P_D)\mathbf{I}_4$ in terms of the criterion of minimizing the sum of variances in the scenario $P_S = P_D$. In other words, the optimal pilot matrix $\mathbf{X}_P^{m,k}$ is equal to an $N_P \times 4$ matrix with all four columns composed of four independent columns of an $N_P \times N_P$ unitary matrix multiplied by a scalar number depending on modulation type and transmit power. ■

Application scenario 1: Given $N_P = 2^i$, and $P_S = P_D$, in terms of the above condition (65), the optimal pilot matrix can be designed to be any four rows of the $2^i \times 2^i$ Hadamard matrix multiplied by a constant, where the normalized constant is a factor of guaranteeing that the average transmit pilot power constraint is satisfied. For example, given 16 QAM and $i = 2$, the optimal pilot matrix may be chosen to be

$$\mathbf{X}_P^{m,k\star} = \frac{\sqrt{2(P_S + P_D)}(1+3i)}{2\sqrt{10}} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \tag{66}$$

which is actually a superimposed pilot pattern.

Application scenario 2: For any positive integer $N_P \ge 4$, and $P_S = P_D$, in terms of Theorem 1, we may construct an optimal

orthogonal pilot pattern as follows:

$$\mathbf{X}_P^{m,k\star} = \frac{\sqrt{2(P_S+P_D)}(1+3i)}{\sqrt{10}} \times \left(\mathbf{e}_{N_P,m_1}\ \mathbf{e}_{N_P,m_2}\ \mathbf{e}_{N_P,m_3}\ \mathbf{e}_{N_P,m_4}\right) \quad (67)$$

where $\mathbf{e}_{N_P,m_n}$ is chosen to be the m_n column of $N_P \times N_P$ identity matrix with $m_n \in \{1,2,\ldots,N_P\}$ and $m_1 \neq m_2 \neq m_3 \neq m_4$. More detailedly, when $m_1 = 1$, $m_2 = 2$, $m_3 = 3$, and $m_4 = 4$, the above pilot matrix becomes

$$\mathbf{X}_P^{m,k\star} = \frac{\sqrt{2(P_S+P_D)}(1+3i)}{\sqrt{10}}\left(\mathbf{e}_{N_P,1}\ \mathbf{e}_{N_P,2}\ \mathbf{e}_{N_P,3}\ \mathbf{e}_{N_P,4}\right) = \begin{pmatrix}\mathbf{I}_4 \\ \mathbf{0}_{(N_P-4)\times 4}\end{pmatrix}. \quad (68)$$

Theorem 1 means that the conditional number of the optimal $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ is equal to one. Now, we discuss the effect of the conditional number of $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ on the sum of MSE. If we assume that the maximum and minimum non-zero eigenvalues of $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ are $\lambda_{\max}$ and $\lambda_{\min}$, then

$$\lambda_{\max} = \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 = \lambda_{\min} \geq 0. \quad (69)$$

Let us define the conditional number of pilot product matrix $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ as follows:

$$\gamma = \frac{\lambda_{\max}}{\lambda_{\min}}. \quad (70)$$

Making use of (69), we have

$$\operatorname{tr}(\mathbf{C}_{2\times 2}) \bullet \frac{1}{\lambda_{\max}} \bullet (3+\gamma) \leq \text{MSE}^{m,k} = \operatorname{tr}(\mathbf{C}_{2\times 2}) \sum \frac{1}{\lambda_i} \leq \operatorname{tr}(\mathbf{C}_{2\times 2}) \bullet \frac{1}{\lambda_{\max}} \bullet (3\gamma+1), \quad (71)$$

which is further relaxed as

$$\frac{(3+\gamma)\operatorname{tr}(\mathbf{C}_{2\times 2})}{2N_P(P_S+P_D)} \leq \text{MSE}^{m,k} \leq \frac{2(1+3\gamma)\operatorname{tr}(\mathbf{C}_{2\times 2})}{N_P(P_S+P_D)} \quad (72)$$

due to

$$\frac{N_P(P_S+P_D)}{2} \leq \lambda_{\max} < 2N_P(P_S+P_D). \quad (73)$$

Observing the above squeeze inequality, we find the sum of MSE grows approximately linearly with the conditional number γ. As γ tends to positive infinity, that is, matrix $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ approaches singular, the sum-MSE performance of the proposed FD-LS will degrade seriously. In the special case the conditional number γ of $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ being 1, we have

$$\text{Sum-MSE}^{m,k} = \frac{4}{\lambda_{\max}}\operatorname{tr}(\mathbf{C}_{2\times 2}) = \frac{8\operatorname{tr}(\mathbf{C}_{2\times 2})}{N_P(P_S+P_D)}. \quad (74)$$

Dividing two sides of (72) by (74), we obtain that the gain attained by the optimal pattern relative to any pilot matrix with conditional number γ within the interval $[\frac{3+\gamma}{16}, \frac{1+3\gamma}{4}]$. Thus, the gain becomes more significant as the conditional number of matrix $(\mathbf{X}_P^{m,k})^H\mathbf{X}_P^{m,k}$ increases. Let us consider one scenario, $\gamma = 157$, the sum-MSE of the proposed FD-LS with such a pilot matrix is at least ten times that of one with the optimal pilot matrix. This will result in a severe performance loss on the proposed FD-LS. In Section VI, we will demonstrate how γ affects the performance of the proposed FD-LS estimator by simulations.

V. Improved Channel Estimation, Interference Cancellation, and Detection

In the previous section, we construct a frequency-domain channel estimator that doesn't take advantage of the time-domain property of the channel. In the following, we will first transform the combined frequency-domain channel estimated by FD-LS estimator into its time-domain channel impulse response (TD-CIR) by using IFFT. In the second step, the taps of TD-CIR outside CP will be set to be zeros. In the final step, we convert the new TD-CIR back to its frequency domain by using FFT.

A. Improved FD-LS Channel Estimator

Similar to [18], four ideal frequency-domain CGVs are related to their TD-CIRs by the following equations:

$$\mathbf{H}_{SD,P}^{a,m} = \mathbf{F}_{N\times N}\begin{pmatrix}\underbrace{\mu_{r,D}\mu_{t,S}\mathbf{h}_{SD,P}^{m} + \nu_{r,D}\nu_{t,S}^{*}\{\mathbf{h}_{SD,P}^{m}\}^{*}}_{\mathbf{h}_{SD,P}^{a,m}} \\ \mathbf{0}_{(N-L)\times 1}\end{pmatrix} = \mathbf{F}_{N\times L}\mathbf{h}_{SD,P}^{a,m} \quad (75)$$

$$\mathbf{H}_{SD,P}^{b,m} = \mathbf{F}_{N\times N}\begin{pmatrix}\underbrace{\mu_{r,D}\nu_{t,S}\mathbf{h}_{SD,P}^{m} + \nu_{r,D}\mu_{t,S}^{*}\{\mathbf{h}_{SD,P}^{m}\}^{*}}_{\mathbf{h}_{SD,P}^{b,m}} \\ \mathbf{0}_{(N-L)\times 1}\end{pmatrix} = \mathbf{F}_{N\times L}\mathbf{h}_{SD,P}^{b,m} \quad (76)$$

$$\mathbf{H}_{DD,P}^{a,m} = \mathbf{F}_{N\times N}\begin{pmatrix}\underbrace{\mu_{r,D}\mu_{t,D}\mathbf{h}_{DD,P}^{m} + \nu_{r,D}\nu_{t,D}^{*}\{\mathbf{h}_{DD,P}^{m}\}^{*}}_{\mathbf{h}_{DD,P}^{a,m}} \\ \mathbf{0}_{(N-L)\times 1}\end{pmatrix} = \mathbf{F}_{N\times L}\mathbf{h}_{DD,P}^{a,m} \quad (77)$$

and

$$\mathbf{H}_{DD,P}^{b,m} = \mathbf{F}_{N\times N}\begin{pmatrix}\underbrace{\mu_{r,D}\nu_{t,D}\mathbf{h}_{DD,P}^{m} + \nu_{r,D}\mu_{t,D}^{*}\{\mathbf{H}_{DD,P}^{m}\}^{*}}_{\mathbf{h}_{DD,P}^{b,m}} \\ \mathbf{0}_{(N-L)\times 1}\end{pmatrix} = \mathbf{F}_{N\times L}\mathbf{h}_{DD,P}^{b,m}. \quad (78)$$

After the four vectors of frequency-domain channel gain $\mathbf{H}_{SD,P}^{a,m}$, $\mathbf{H}_{SD,P}^{b,m}$, $\mathbf{H}_{DD,P}^{a,m}$, and $\mathbf{H}_{DD,P}^{b,m}$ are estimated by using the FD-LS channel estimator in Section III. We denote

the estimated four frequency-domain CGVs of block m as $\hat{\mathbf{H}}_{SD,P}^{a,m}$, $\hat{\mathbf{H}}_{SD,P}^{b,m}$, $\hat{\mathbf{H}}_{DD,P}^{a,m}$, and $\hat{\mathbf{H}}_{DD,P}^{b,m}$. Using (75)–(78), the corresponding estimated TD-CIRs are directly given by

$$\begin{aligned}\hat{\mathbf{h}}_{SD,P}^{a,m} &= \mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{SD,P}^{a,m}\\ \hat{\mathbf{h}}_{SD,P}^{b,m} &= \mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{SD,P}^{b,m}\\ \hat{\mathbf{h}}_{DD,P}^{a,m} &= \mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{DD,P}^{a,m}\\ \hat{\mathbf{h}}_{DD,P}^{b,m} &= \mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{DD,P}^{b,m}\end{aligned} \tag{79}$$

where

$$\begin{aligned}\mathbf{E}_{L\times N} &= \left(\mathbf{e}_{N,1}\ \mathbf{e}_{N,2}\cdots\ \mathbf{e}_{N,L}\right)^{T}\\ &= \left(\mathbf{I}_{L}\ \mathbf{0}_{L\times(N-L)}\right)\end{aligned} \tag{80}$$

with $\mathbf{e}_{N,i}$ being the ith column vector of the N by N identity matrix. Applying the FFT operations to both sides of (79), we have the resulting improved estimator

$$\begin{aligned}\tilde{\mathbf{H}}_{SD,P}^{a,m} &= \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{SD,P}^{a,m}\\ \tilde{\mathbf{H}}_{SD,P}^{b,m} &= \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{SD,P}^{b,m}\\ \tilde{\mathbf{H}}_{DD,P}^{a,m} &= \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{DD,P}^{a,m}\end{aligned}$$

and

$$\tilde{\mathbf{H}}_{DD,P}^{b,m} = \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^{H}\hat{\mathbf{H}}_{DD,P}^{b,m}. \tag{81}$$

Due to the above FFT/IFFT operation process of exploiting the time-domain property of CIR, the sum-MSE of the improved channel estimator will be approximately reduced to $\frac{L}{N}$ of that of FD-LS in Section IV, which is proved in Appendix A. In particular, it is noted that the window length is chosen as L, where $L \geq \lfloor t_{\max} \rfloor + 1$ with $t_{\max}$ being the normalized maximum channel delay. Additionally, a larger L means the less improvement on sum-MSE of the improved FD-LS.

In the case of the slow-fading channel, i.e., channel keeps constant during one block, we simply set the CGVs of the remaining data sub-block per block equal the estimated CGVs of the corresponding pilot sub-block. If channel varies slowly during one block, the CGV corresponding to data OFDM symbol n can be obtained by the time-direction interpolation among the CGVs of pilot sub-blocks of several adjacent blocks closest to it. This will be presented in our simulation section. However, in the following subsection, we still assume channel keeps constant during one block to derive conveniently.

B. Interference Cancellation and Whitening-Filter Detection

Since we have estimated the full-duplex self-interference channel by the proposed FD-LS and improved FD-LS estimators in Sections III and V, and at the same time, the destination receiver knows the transmit data from the destination transmitter well, the full-duplex self-interference in the received signal can be cancelled to form a new receive signal vector

$$\begin{aligned}\bar{\mathbf{z}}^{n,k} &= \mathbf{\Gamma}_{SD}^{n,k}\mathbf{x}_{S}^{n,k} + \mathbf{\Gamma}_{DD}^{n,k}\mathbf{x}_{D}^{n,k} + \mathbf{w}^{n,k} - \hat{\mathbf{\Gamma}}_{DD}^{n,k}\mathbf{x}_{D}^{n,k}\\ &= \mathbf{\Gamma}_{SD}^{n,k}\mathbf{x}_{S}^{n,k} + \underbrace{\left(\mathbf{\Gamma}_{DD}^{n,k} - \hat{\mathbf{\Gamma}}_{DD}^{n,k}\right)\mathbf{x}_{D}^{n,k}}_{\text{Residual self-interference } \Delta\mathbf{I}_{DD}^{n,k}} + \mathbf{w}^{n,k}\end{aligned} \tag{82}$$

where $\mathbf{\Gamma}_{SD}^{n,k}$ is the ideal source-to-destination channel corresponding to the kth subcarrier pair of OFDM symbol n, $\mathbf{\Gamma}_{DD}^{n,k}$ and $\hat{\mathbf{\Gamma}}_{DD}^{n,k}$ are the corresponding ideal and estimated destination-to-destination channels. Obviously, the last two terms in the right side of (82) are the residual interference and channel noise, and their sum is a colored vector of noise plus interference, which can be viewed as a colored noise vector. Thus, to realize an optimal maximum likelihood detection, due to IQ-imbalance, the colored noise vector should be whitened into a vector of white noise by using its covariance matrix even if the residual self-interference is completely removed. From Appendix B, if the proposed FD-LS channel estimator in Section III is used, then its covariance matrix is

$$\begin{aligned}\mathbf{C}_{IN} &= \mathcal{E}\left[\left(\Delta\mathbf{I}_{DD}^{n,k} + \mathbf{w}^{n,k}\right)\left(\Delta\mathbf{I}_{DD}^{n,k} + \mathbf{w}^{n,k}\right)^{H}\right]\\ &= \left(1 + \frac{2P_D}{N_P\left(P_S + P_D\right)}\right)\mathbf{C}_{2\times 2}.\end{aligned} \tag{83}$$

Quite similarly, if the improved FD-LS channel estimator in (81) is adopted at destination receiver, the corresponding covariance matrix of residual self-interference plus noise is

$$\mathbf{C}_{IN} = \left(1 + \frac{2LP_D}{NN_P\left(P_S + P_D\right)}\right)\mathbf{C}_{2\times 2}. \tag{84}$$

Theorem 2: In a full-duplex OFDM system with IQ imbalances, if the optimal pilot pattern as indicated in Theorem 1 is used at transmitters of both source and destination in Fig. 1, and the improved FD-LS channel estimator and serial interference cancellation are used at receiver, then the residual self-interference has the following form of covariance matrix

$$\mathbf{C}_{IN} = \left(1 + \frac{2LP_D}{NN_P\left(P_S + P_D\right)}\right)\mathbf{C}_{2\times 2} \tag{85}$$

due to channel estimation error, which presents the same form as the covariance matrix of the channel noise vector with only difference being a constant factor. Thus, the whitening-filter at receiver is determined by only the covariance matrix of channel noise vector including receive IQ imbalance. ■

Proof. The proof of Theorem 2 is similar to the proof in Appendix B. Thus, we omit it here. ■

Based on (83) and (85), let us define the whitening filter (WF) as

$$\mathbf{W}_{WF} = \mathbf{C}_{2\times 2}^{-1/2} = \mathbf{Q}\mathbf{\Sigma}_{\mathbf{C}}^{-\frac{1}{2}} \tag{86}$$

with

$$\mathbf{W}_{WF}\mathbf{W}_{WF}^{H} = \mathbf{C}_{2\times 2}^{-1} \tag{87}$$

where matrix $\mathbf{C}_{2\times 2}$ is a Hermitian and positive semi-definite matrix, and its eigenvalue decomposition (EVD) is given by

$$\mathbf{C}_{2\times 2} = \mathbf{Q}\mathbf{\Sigma}_{\mathbf{C}}\mathbf{Q}^{H} \tag{88}$$

with $\mathbf{\Sigma}_{\mathbf{C}}$ being a diagonal matrix and $\mathbf{Q}$ unitary.

Applying the above WF operation to two sides of (82), we have

$$\begin{aligned}\bar{\bar{\mathbf{z}}}^{n,k} &= \mathbf{W}_{WF}\bar{\mathbf{z}}^{n,k} = \mathbf{W}_{WF}\mathbf{z}^{n,k} - \mathbf{W}_{WF}\hat{\mathbf{\Gamma}}_{DD}^{n,k}\mathbf{x}_{D}^{n,k} \\ &= \mathbf{W}_{WF}\mathbf{\Gamma}_{SD}^{n,k}\mathbf{x}_{S}^{n,k} + \mathbf{W}_{WF}\left(\Delta\mathbf{I}_{DD}^{n,k} + \mathbf{w}^{n,k}\right). \end{aligned} \tag{89}$$

Making use of the new model, the corresponding ML and ZF detectors are represented as

$$\bar{\mathbf{x}}_{S,ML}^{n,k} = \arg\min||\mathbf{W}_{WF}\bar{\mathbf{z}}^{n,k} - \mathbf{W}_{WF}\hat{\mathbf{\Gamma}}_{SD}^{n,k}\mathbf{x}_{S}^{n,k}||^2 \tag{90}$$

and

$$\bar{\mathbf{x}}_{S,ZF}^{n,k} = \left(\mathbf{W}_{WF}\hat{\mathbf{\Gamma}}_{SD}^{n,k}\right)^{-1}\bar{\mathbf{z}}^{n,k} \tag{91}$$

which are abbreviated as WFML and WFZF, respectively. For comparison, in terms of (82) before whitening filter, the CML (conventional ML) and CZF (conventional ZF) detectors are given by

$$\bar{\mathbf{x}}_{S,CML}^{n,k} = \arg\min||\bar{\mathbf{z}}^{n,k} - \hat{\mathbf{\Gamma}}_{SD}^{n,k}\mathbf{x}_{S}^{n,k}||^2 \tag{92}$$

and

$$\bar{\mathbf{x}}_{S,CZF}^{n,k} = \left(\hat{\mathbf{\Gamma}}_{SD}^{n,k}\right)^{-1}\bar{\mathbf{z}}^{n,k} \tag{93}$$

respectively.

C. Proposed Low-Complexity WFML Detector

Firstly, the matrix $\mathbf{W}_{WF}\mathbf{\Gamma}_{SD}^{n,k}$ in (89) can be decomposed by SVD as follows:

$$\mathbf{W}_{WF}\mathbf{\Gamma}_{SD}^{n,k} = \mathbf{U}^{n,k}\mathbf{\Sigma}^{n,k}\left(\mathbf{V}^{n,k}\right)^{H} \tag{94}$$

where $\mathbf{U}^{n,k}$, $\mathbf{\Sigma}^{n,k}$, and $\left(\mathbf{V}^{n,k}\right)^{H}$ are the results of the SVD decomposition. We use $\mathbf{V}^{n,k}$ as the precoding matrix at the source, and $\left(\mathbf{U}^{n,k}\right)^{H}$ as the receive beamforming at the destination. After applying precoding and receive beamforming, (89) can be converted to

$$\begin{aligned}\tilde{\mathbf{y}}^{n,k} &= \left(\mathbf{U}^{n,k}\right)^{H}\mathbf{W}_{WF}\mathbf{\Gamma}_{SD}^{n,k}\mathbf{V}^{n,k}\mathbf{x}_{S}^{n,k} \\ &\quad + \left(\mathbf{U}^{n,k}\right)^{H}\mathbf{W}_{WF}\left(\Delta\mathbf{I}_{DD}^{n,k} + \mathbf{w}^{n,k}\right) \\ &= \mathbf{\Sigma}^{n,k}\mathbf{x}_{S}^{n,k} + \underbrace{\left(\mathbf{U}^{n,k}\right)^{H}\mathbf{W}_{WF}\left(\Delta\mathbf{I}_{DD}^{n,k} + \mathbf{w}^{n,k}\right)}_{\mathbf{n}_{R}''} \end{aligned} \tag{95}$$

where $\mathbf{n}_{R}''$ is the new formed noise plus residual self-interference vector after beamforming. Considering that $\left(\mathbf{U}^{n,k}\right)^{H}$ is an unitary matrix, all components of the vector of the residual self-interference plus noise vector $\mathbf{n}_{R}''$ are still mutually independent with the corresponding elements of covariance matrix $\mathbf{I}_2$. This means that the above process of WF-plus-SVD converts two coupled spatial subchannels in (89) into two independent parallel spatial subchannels in (95). Then, each substream can be individually detected or decoded. In a mathematical language, since both $\left(\mathbf{U}^{n,k}\right)^{H}$ and $\mathbf{V}^{n,k}$ are unitary matrices, the optimization problem in (90) can be formulated as the following problem:

$$\begin{aligned}\hat{\mathbf{x}}_{S,WF-SVD}^{n,k} &= \arg\min_{\mathbf{x}^{n,k}\in\mathbb{C}^2}\|\tilde{\mathbf{y}}^{n,k} - \mathbf{\Sigma}^{n,k}\mathbf{x}^{n,k}\|^2 \\ &= \arg\min_{\mathbf{x}^{n,k}(i)\in\mathbb{C}}\sum_{i=1}^{2}\|\tilde{\mathbf{y}}^{n,k}(i) - \alpha_i\mathbf{x}^{n,k}(i)\|^2 \end{aligned} \tag{96}$$

where α_i is the (i,i) diagonal element of matrix $\mathbf{\Sigma}^{n,k}$, and $\tilde{\mathbf{y}}^{n,k}(i)$ and $\mathbf{x}^{n,k}(i)$ are the i-th elements of $\tilde{\mathbf{y}}^{n,k}$ and $\mathbf{x}^{n,k}$, respectively. The above problem is divided into two sub-optimization problems

$$\hat{\mathbf{x}}_{S,WF-SVD}^{n,k}(i) = \arg\min_{\mathbf{x}^{n,k}(i)\in\mathbb{C}}\|\tilde{\mathbf{y}}^{n,k}(i) - \alpha_i\mathbf{x}^{n,k}(i)\|^2 \tag{97}$$

for $\forall i \in \{1,2\}$. In terms of the above simplification process, the complexity of the ML detector using exhaustive search is significantly reduced from $O(M^2)$ FLOPs to $O(2^3 + 2M)$ FLOPs, where $O(2^3)$ comes from the SVD decomposition and $O(2M)$ is the detection complexity. In summary, compared to the conventional ML detection methods with sub-optimal performance, the proposed detector in (96) makes a significant reduction on complexity and achieves the optimal detection performance as the WFML in (97) due to unitary transformation of preserving the system performance unchanged before and after.

VI. Simulations and Discussions

In this section, we evaluate the performance of the proposed FD-LS channel estimators, optimal pilot pattern, and WFML detector, and compare them with the conventional ones. The system parameters in our simulation are set as follows: OFDM symbol length $N = 512$, cyclic prefix $L = 32$, signal bandwidth $BW = 10$ MHz, the number of pilot OFDM symbols $N_P = 4$, $N_D = 8$, digital modulation 16QAM, carrier frequency $f_c = 2$ GHz, and $P_S = P_D$, where P_S and P_D are the transmit powers of source and destination, respectively. The typical EVA channel model with maximum path delay 2.51 μ s in LTE standard is employed in our simulation. It is noted that the IQ-imbalance parameters for all the remaining figures except for Fig. 9 are chosen to be $\theta_{t,S} = \theta_{t,D} = \theta_{r,D} = 2^o$ and $\alpha_{t,S} = \alpha_{t,D} = \alpha_{r,D} = 1$ dB.

A. Constant Channel During One Block

To see the influence of the conditional number of pilot product matrix on the sum-MSE performance of the proposed FD-LS channel estimator, Fig. 3 plots the curves of sum-MSE versus SNR of the proposed FD-LS for different conditional numbers of pilot product matrix. It is shown that the sum-MSE performance gradually grows as the conditional number of pilot product matrix approaches one. Observing this figure, it is obvious that the proposed optimal pilot product matrix with CN $= 1$ achieves the

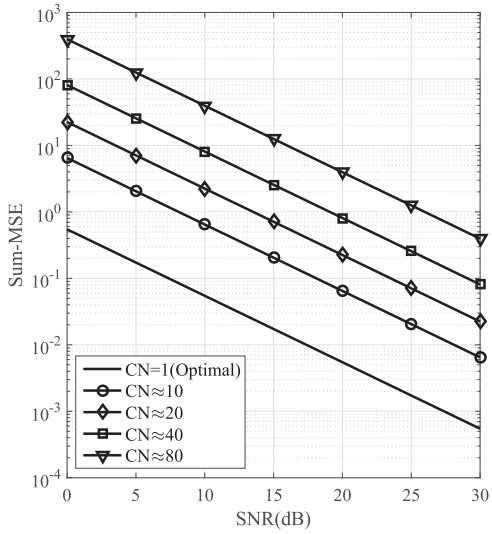


Fig. 3. Curves of average sum-MSE versus SNR of the proposed FD-LS for different conditional numbers of pilot product matrix (CN is short for conditional number).

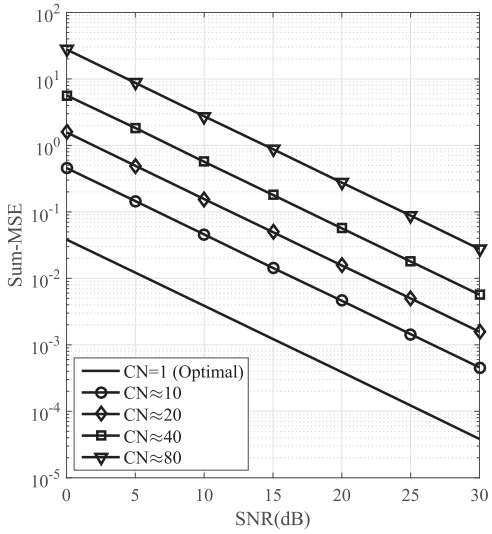


Fig. 5. Curves of average sum-MSE versus SNR of the proposed improved FD-LS for pilot matrices with different conditional numbers.

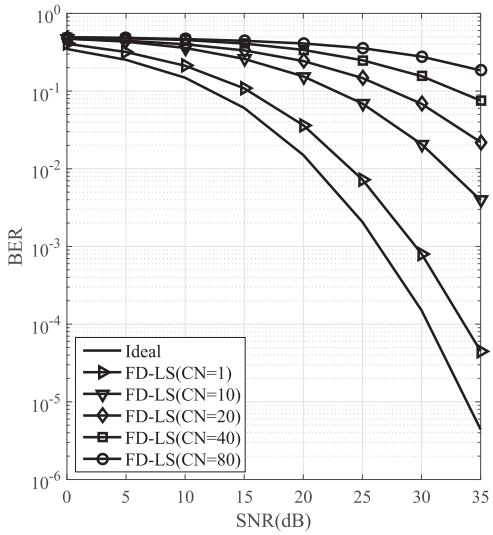


Fig. 4. Curves of BER versus SNR of the proposed FD-LS for pilot matrices with different conditional numbers (conventional ML detection).

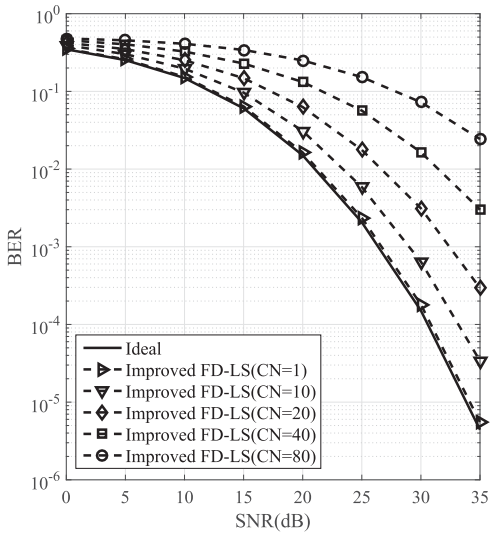


Fig. 6. Curves of BER versus SNR of the proposed improved FD-LS for pilot matrices with different conditional numbers (conventional ML detection).

best sum-MSE performance, and shows substantial SNR gains over other four different pilot matrices with larger values of CN being 10, 20, 40, and 80. For example, the FD-LS with CN $=$ 1 achieves around 10 dB SNR gain over that with CN $=$ 10 in terms of sum-MSE for almost all SNR regions. As CN grows up to 80, the SNR gain achieved by the FD-LS with CN $=$ 1 is about 30 dB. Thus, the CN of pilot product matrix makes a significant impact on the sum-MSE performance of FD-LS.

Fig. 4 demonstrates the curves of achievable BER versus SNR of the proposed FD-LS for pilot product matrices with different conditional numbers. Observing this figure, the proposed FD-LS with CN $=$ 1 achieves the best BER performance compared to ones with large CNs. As the CN of pilot product matrix increases, the BER performance becomes worse. For example, the BER approaches 10% at SNR $=$ 30 dB for CN $\geq$ 20. This reveals a fact that the conditional number of pilot product matrix also makes a profound impact on the BER performance of the proposed FD-LS channel estimator.

Figs. 5 and 6 illustrate the curves of sum-MSE and BER versus SNR of the proposed improved FD-LS channel estimator for pilot product matrices with different conditional numbers. Observing the two figures, the improved FD-LS demonstrates the similar trends of performance as the proposed FD-LS shown in Figs. 3 and 4. Now, let us compare the sum-MSE performance of the FD-LS and improved FD-LS. Fixing sum-MSE $=$ 0.01 and CN $=$ 1 in Figs. 3 and 5, the improved FD-LS shows about 12 dB SNR gain over the FD-LS. Comparing the BER curves in Figs. 4 and 6 yields the similar performance order as CN

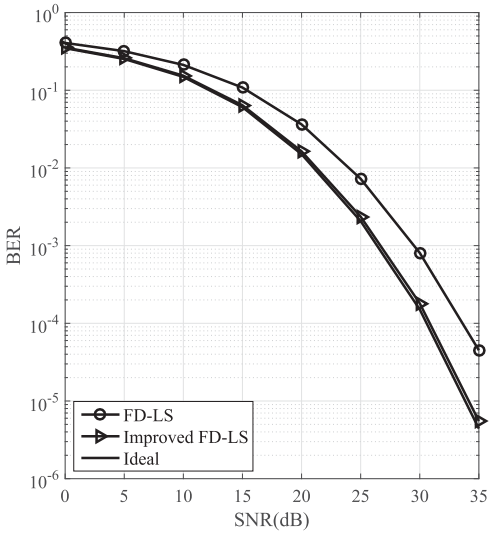


Fig. 7. Curves of average BER versus SNR of different channel estimators (CN = 1).

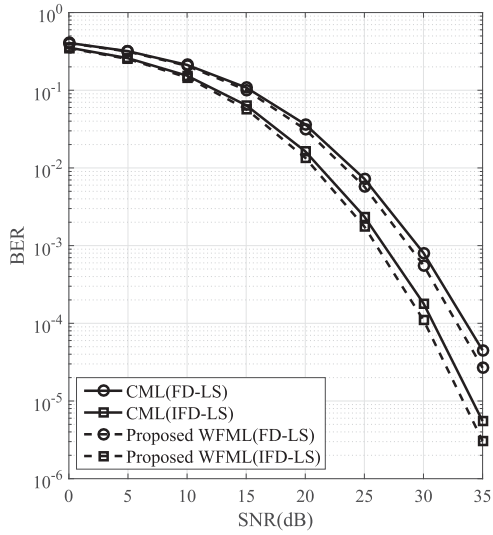


Fig. 8. Curves of BER versus SNR of different ML detectors with optimal pilot and fixed channel estimator where CML and IFD-LS are short for conventional ML detection and improved FD-LS, respectively (CN = 1).

approaches one. In summary, the improved FD-LS performs much better than the FD-LS in BER and sum-MSE senses.

To gain insight into the detailed BER performance gap between the FD-LS and its improved version, Fig. 7 indicates the curves of achievable BER versus SNR of the FD-LS and improved FD-LS for optimal pilot product matrix (CN = 1), where perfect channel state information (CSI) is used as a reference. Obviously, the improved FD-LS performs better than the proposed FD-LS and is close to the perfect CSI. Measuring roughly, the improved FD-LS makes an about $2 \sim 3$ dB SNR gain improvement over the proposed FD-LS at BER$= 10^{-2}$.

Now, we turn to evaluate the performance of the proposed WFML detector. Fig. 8 shows the curves of BER versus SNR of CML, and proposed WFML when different channel estimators are adopted. The experimental results show that the proposed WFML detector attains about 0.6 dB/0.7 dB SNR gain over the conventional ML one for given BER $= 10^{-2}/10^{-3}$. The achievable SNR gain is about 0.7 dB for both BER $= 10^{-3}$ and BER $= 10^{-4}$ provided that the improved FD-LS is used. The main reason is that not only the residual self-interference but also the additive noise vector are colored due to the receive IQ imbalances although the improved FD-LS can provide a higher estimated precision over the original FD-LS to reduce the residual self-interference.

To inspect the effect of the change in IQ-imbalance parameters on the BER performance of the proposed WFML detector, the six IQ-imbalance parameters are set as follows: $\theta_{t,S} = \theta_{t,D} = \theta_{r,D} = 2^\circ$ and $\alpha_{t,S} = \alpha_{t,D} = \alpha_{r,D} = 2$ dB. Fig. 9 plots the curves of BER versus SNR of WFML and CML detectors under such parameters. Compared to Fig. 8, the SNR gain in Fig. 9 due to whitening-filter increases up to about $3 \sim 4$ dB for a given BER $= 10^{-2}$ when the FD-LS is used. If the improved FD-LS is used, the SNR gain is also $3 \sim 4$ dB for a given BER $= 10^{-3}$. Eventually, even when the ideal channel state information is available, the full-duplex self-interference is completely cancelled, the WFML detector achieves a substantial SNR gain over CML due to the colored channel noise from IQ-imbalances. Doubling the IQ-imbalance amplitude distortion will make the covariance of noise and residual self-interference more color. Thus, the achievable SNR gain produced by whitening-filter is substantially improved.

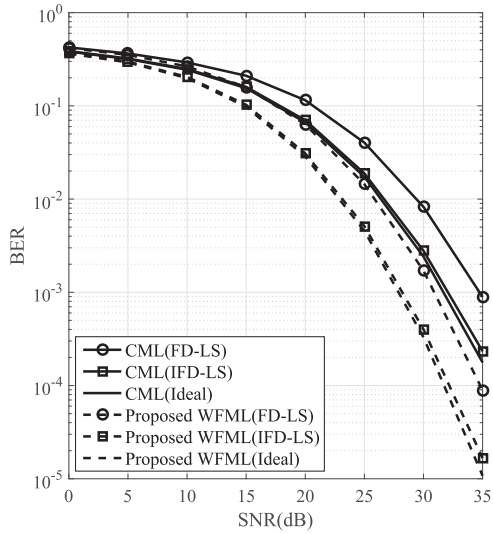


Fig. 9. Curves of BER versus SNR of different ML detectors with optimal pilot and fixed channel estimator (CN = 1).

B. Interpolation for Time-Varying Channel During One Block

In the above, we assume that channel keeps approximately constant during one block, the CGV $\hat{\mathbf{H}}_{SD}^{a,n}$ is simply taken to be $\hat{\mathbf{H}}_{SD,P}^{a,m}$. If channel varies during one block but still keeps constant during each pilot sub-block of N_P pilot OFDM

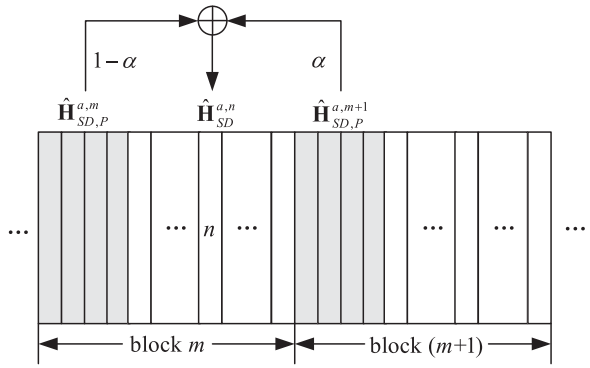


Fig. 10. Demonstration of linear interpolation.

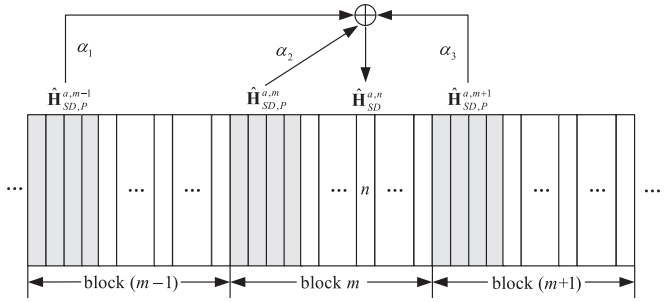


Fig. 11. Second-order polynomial interpolation.

symbol, then the linear interpolation (LI) or second-order polynomial interpolation (SOPI) in [27] are adopted to estimate the CGVs associated with OFDM symbol n. In Fig. 10, we plot the frame structure detailedly. Since the nth OFDM data symbol is between the pilot symbols of block m and $m+1$, the pilot channel gain vectors $\hat{\mathbf{H}}_{SD,P}^{a,m}$ and $\hat{\mathbf{H}}_{SD,P}^{a,m+1}$ of the two adjacent blocks m and $m+1$ are used to linearly interpolate the CGV corresponding to the nth OFDM data symbol by the following LI:

$$\hat{\mathbf{H}}_{SD}^{a,n} = \hat{\mathbf{H}}_{SD}^{a,mN_B+q} = (1-\alpha)\hat{\mathbf{H}}_{SD,P}^{a,m} + \alpha\hat{\mathbf{H}}_{SD,P}^{a,m+1}, \quad (98)$$

where

$$\alpha = \frac{q - N_P/2 - 1/2}{N_B}. \quad (99)$$

Combining the improved FD-LS (IFD-LS) with the LI above forms one new channel estimator IFD-LS-LI for time-variant wireless channel.

If the second-order polynomial interpolation (SOPI) method is used, a linear combination of the three estimated pilot CGVs of three adjacent pilot sub-blocks closet to OFDM symbol n, as shown in Fig. 11, yields the CGV of OFDM symbol n as the following form:

$$\hat{\mathbf{H}}_{SD}^{a,n} = \hat{\mathbf{H}}_{SD}^{a,mN_B+q} = \alpha_1\hat{\mathbf{H}}_{SD,P}^{a,m-1} + \alpha_2\hat{\mathbf{H}}_{SD,P}^{a,m} + \alpha_3\hat{\mathbf{H}}_{SD,P}^{a,m+1} \quad (100)$$

where

$$\alpha_1 = \frac{\alpha(\alpha-1)}{2} \quad (101)$$

$$\alpha_2 = -(\alpha-1)(\alpha+1) \quad (102)$$

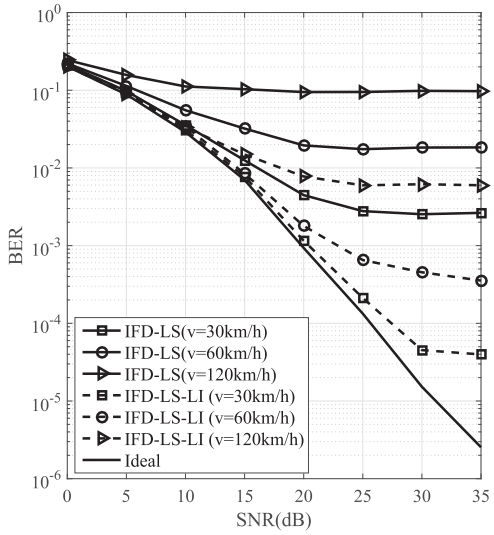


Fig. 12. Curves of BER versus SNR for different car speeds using LI (CN = 1).

and

$$\alpha_3 = \frac{\alpha(\alpha+1)}{2}. \quad (103)$$

Combining the proposed IFD-LS with the SOPI above forms the new channel estimator IFD-LS-SOPI for time-variant wireless channel.

To evaluate the impact of the Doppler spread or the time-varying of channel on BER performance, Fig. 12 plots the curves of BER versus SNR of the IFD-LS and IFD-LS-LI estimators for three different car speeds: 30, 60, and 120 km/h, where QPSK modulation at transmitter and the WFML detector at receiver are used. From this figure, as car speed increases from 30 km/h to 120 km/h, the proposed IFD-LS degrades greatly. Compared to the IFD-LS, the IFD-LS-LI presents a strong ability of combating mobility. At car speed v = 120 km/h and SNR = 35 dB, the BERs of IFD-LS-LI and IFD-LS estimators are about 6×10^{-3} and 1×10^{-1}, respectively. Due to linear interpolation, the performance improvement achieved by IFD-LS-LI over IFD-LS is significant in mobile channel.

Fig. 13 demonstrates the curves of BER versus SNR of the IFD-LS-SOPI and IFD-LS for different car speeds: 30, 60, and 120 km/h. From this figure, we find there is the same performance trend as Fig. 12. Inspecting Fig. 13 and Fig. 12, it follows that the performance difference between IFD-LS-LI and IFD-LS-SOPI is trivial even at car speed = 120 km/h. In summary, SOPI and LI greatly improve the ability of IFD-LS estimator to combat the channel mobility. In other words, IFD-LS-LI and IFD-LS-SOPI can be applied to the full-duplex OFDM systems with IQ-imbalance in time-variant wireless channel.

VII. Conclusion

In this paper, we propose a novel FD-LS channel estimator to estimate both self-interference and intended channels. Subsequently, the problem of optimizing the training matrix is

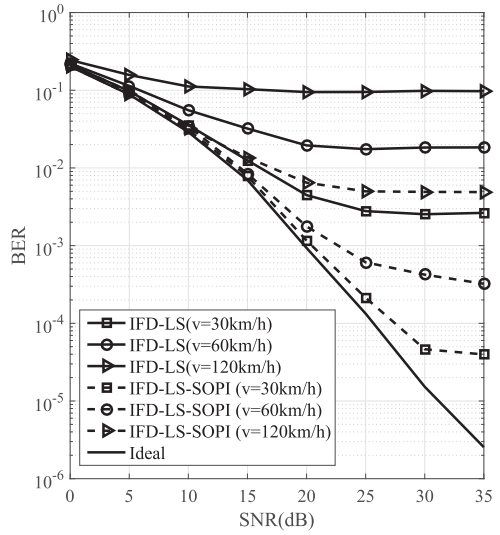


Fig. 13. Curves of BER versus SNR for different car speeds using SOPI (CN = 1).

modelled as a convex optimization problem. Given the same average transmit powers of source and destination, the closed-form expression of the optimal pilot product matrix is derived to be a multiplier of unitary matrix. In other words, the conditional number of the optimal pilot product matrix is equal to one. Following this, an improved FD-LS channel estimator is presented with a significant sum-MSE performance improvement over the previous proposed FD-LS by exploiting the time-domain property of channels. We also find the performance of both the proposed FD-LS and improved FD-LS are intimately related to the conditional number of the pilot product matrix. Finally, a low-complexity WFML detector is developed by applying whitening filer, EVD, and SVD beamforming. The simulation results show that the proposed WFML detector achieves a better BER performance than the conventional ML one. Additionally, its computational saving over the conventional ML one is substantial with the former's complexity being $O(M)$ FLOPs and the latter one $O(M^2)$ FLOPs. Finally, the improved FD-LS estimator is combined with the linear and second-order interpolations to improve its ability of combating mobility. The schemes proposed by us can be applied to the future D2D, V2V, and unmanned-aerial-vehicle networking.

APPENDIX A
DERIVATION OF MSE OF THE IMPROVED FD-LS ESTIMATOR

Proof: Using the FD-LS channel estimator proposed by us, we model the estimated channel gain as follows:

$$\hat{\mathbf{H}}^{a,m}_{SD,P} = \mathbf{H}^{a,m}_{SD,P} + \mathbf{e}^{a,m}_{SD,P} \tag{104}$$

where $\mathbf{e}^{a,m}_{SD,P}$ denotes the error due to the FD-LS. Then, the corresponding MSE is given by

$$\mathrm{MSE}_{SD,P} = \mathcal{E}\left(\mathrm{tr}\left\{\mathbf{e}^{a,m}_{SD,P}\left(\mathbf{e}^{a,m}_{SD,P}\right)^H\right\}\right). \tag{105}$$

In accordance with the error model in (104), and the definition of the improved FD-LS estimator in (81), we have the estimation error model of the improved FD-LS estimator

$$\begin{aligned}\tilde{\mathbf{H}}^{a,m}_{SD,P} &= \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}^H_{N\times N}\hat{\mathbf{H}}^{a,m}_{SD,P} \\ &= \mathbf{H}^{a,m}_{SD,P} + \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}^H_{N\times N}\mathbf{e}^{a,m}_{SD,P}.\end{aligned} \tag{106}$$

The corresponding MSE of the improved FD-LS is written as follows:

$$\begin{aligned}\mathrm{MSE}'_{SD,P} = \mathcal{E}\Big(\mathrm{tr}\Big\{&(\mathbf{e}^a_{SD})^H\mathbf{F}_{N\times N}\mathbf{E}^H_{L\times N}\mathbf{F}^H_{N\times L}\mathbf{F}_{N\times L} \\ &\times \mathbf{E}_{L\times N}\mathbf{F}^H_{N\times N}\mathbf{e}^{a,m}_{SD,P}\Big\}\Big).\end{aligned} \tag{107}$$

Considering $\mathbf{F}^H_{N\times L}\mathbf{F}_{N\times L} = \mathbf{I}_L$, the above MSE reduces to

$$\begin{aligned}\mathrm{MSE}'_{SD,P} = \mathcal{E}\Big(\mathrm{tr}\Big\{&\left(\mathbf{e}^{a,m}_{SD,P}\right)^H\mathbf{F}_{N\times N}\mathbf{E}^H_{L\times N} \\ &\times \mathbf{E}_{L\times N}\mathbf{F}^H_{N\times N}\mathbf{e}^{a,m}_{SD,P}\Big\}\Big).\end{aligned} \tag{108}$$

In terms of the definition $\mathbf{E}_{L\times N}$, we have

$$\mathbf{F}_{N\times N}\mathbf{E}^H_{L\times N} = \mathbf{F}_{N\times L} \tag{109}$$

and

$$\mathbf{E}_{L\times N}\mathbf{F}^H_{N\times N} = \mathbf{F}^H_{N\times L}. \tag{110}$$

Substituting the above two expressions in (108) yields the reduction form

$$\mathrm{MSE}_{SD,P} = \mathcal{E}\Big(\mathrm{tr}\Big\{\left(\mathbf{e}^{a,m}_{SD,P}\right)^H\mathbf{F}_{N\times L}\mathbf{F}^H_{N\times L}\mathbf{e}^{a,m}_{SD,P}\Big\}\Big). \tag{111}$$

Using the matrix identity

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA}) \tag{112}$$

we have

$$\begin{aligned}\mathrm{MSE}'_{SD,P} &= \mathcal{E}\Big(\mathrm{tr}\Big\{\mathbf{F}^H_{N\times L}\mathbf{e}^{a,m}_{SD,P}\left(\mathbf{e}^{a,m}_{SD,P}\right)^H\mathbf{F}_{N\times L}\Big\}\Big) \\ &= \mathrm{tr}\Big\{\mathbf{F}^H_{N\times L}\mathcal{E}\left(\mathbf{e}^{a,m}_{SD,P}\left(\mathbf{e}^{a,m}_{SD,P}\right)^H\right)\mathbf{F}_{N\times L}\Big\}.\end{aligned} \tag{113}$$

If $\mathcal{E}(\mathbf{e}^{a,m}_{SD,P}(\mathbf{e}^{a,m}_{SD,P})^H) = \sigma^2_{SD,e}\mathbf{I}_N$, where $\sigma^2_{SD,e}$ stands for the estimated error variance, then the above MSE can be simplified as

$$\mathrm{MSE}'_{SD,P} = \sigma^2_{SD,e}\mathrm{tr}\left\{\mathbf{F}_{N\times L}\mathbf{F}^H_{N\times L}\right\}. \tag{114}$$

Because of

$$\mathrm{tr}\left\{\mathbf{F}_{N\times L}\mathbf{F}^H_{N\times L}\right\} = L \tag{115}$$

we have

$$\mathrm{MSE}'_{SD,P} = \sigma^2_{SD,e}L = \frac{L}{N}\mathrm{MSE}_{SD,P} \tag{116}$$

which implies that the improved FD-LS in Section V can achieve an approximate L/N improvement over the FD-LS in Section III in terms of MSE. This completes the approximate proof of the improved factor L/N of MSE. ■

APPENDIX B
COVARIANCE MATRIX OF RESIDUAL SELF-INTERFERENCE PLUS NOISE

Proof: Because channel is assumed to be constant during one block, so $\hat{\mathbf{\Gamma}}^{n,k}$ is taken to be $\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS}$, and $\mathbf{\Gamma}^{m,k}_{P}$=$\mathbf{\Gamma}^{n,k}$, where $\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS}$ is given by (40). Substituting (65) in (40) yields

$$\hat{\mathbf{\Gamma}}^{n,k} = \begin{pmatrix} \mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{n,k}_{SD}\right) \\ \mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{n,k}_{DD}\right) \end{pmatrix} = \underbrace{\begin{pmatrix} \mathrm{vec}\left(\mathbf{\Gamma}^{n,k}_{SD}\right) \\ \mathrm{vec}\left(\mathbf{\Gamma}^{n,k}_{DD}\right) \end{pmatrix}}_{\mathbf{\Gamma}^{n,k}} + \frac{1}{N_P P_S}\left[\left(\mathbf{X}^{m,k}_{P}\right)^{H} \otimes \mathbf{I}_2\right]\mathbf{w}^{m,k}_{P} \tag{117}$$

where $\mathbf{w}^{m,k}_{P}$ denotes channel noise vector over all k th-pair subchannels of all pilot OFDM symbols per block, and is independent of $\mathbf{w}^{n,k}$. From (117), we extract the estimated $D \rightarrow D$ (self-interference) channel

$$\mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{n,k}_{DD}\right) = \mathbf{E}_{DD}\hat{\mathbf{\Gamma}}^{n,k}_{FD-LS} = \mathrm{vec}\left(\mathbf{\Gamma}^{n,k}_{DD}\right) + \frac{1}{N_P P_S}\mathbf{E}_{DD}\left[\left(\mathbf{X}^{m,k}_{P}\right)^{H} \otimes \mathbf{I}_2\right]\mathbf{w}^{m,k}_{P} \tag{118}$$

with

$$\mathbf{E}_{DD} = \left(\mathbf{0}_{4\times 4}\ \mathbf{I}_4\right). \tag{119}$$

The relationship between $\mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{n,k}_{DD}\right)$ and $\hat{\mathbf{\Gamma}}^{n,k}_{DD}$ is formulated as

$$\hat{\mathbf{\Gamma}}^{n,k}_{DD} = \left[\mathbf{E}_{DA}\mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{n,k}_{DD}\right)\ \mathbf{E}_{DB}\mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{n,k}_{DD}\right)\right] \tag{120}$$

with

$$\mathbf{E}_{DA} = \left(\mathbf{I}_2\ \mathbf{0}_{2\times 2}\right) \tag{121}$$

and

$$\mathbf{E}_{DB} = \left(\mathbf{0}_{2\times 2}\ \mathbf{I}_2\right). \tag{122}$$

Plugging (118) in (120) forms

$$\hat{\mathbf{\Gamma}}^{n,k}_{DD} = \left[\mathbf{E}_{DA}\mathbf{E}_{DD}\mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS}\right)\ \mathbf{E}_{DB}\mathbf{E}_{DD}\mathrm{vec}\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS}\right)\right]. \tag{123}$$

In the same manner

$$\mathbf{\Gamma}^{n,k}_{DD} = \left[\mathbf{E}_{DA}\mathbf{E}_{DD}\mathrm{vec}\left(\mathbf{\Gamma}^{m,k}_{P}\right)\ \mathbf{E}_{DB}\mathbf{E}_{DD}\mathrm{vec}\left(\mathbf{\Gamma}^{m,k}_{P}\right)\right]. \tag{124}$$

Subtracting (124) from (123), we have

$$\hat{\mathbf{\Gamma}}^{n,k}_{DD} - \mathbf{\Gamma}^{n,k}_{DD} = \left[\mathbf{E}_{DA}\mathbf{E}_{DD}\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS} - \mathbf{\Gamma}^{m,k}_{P}\right) \times \mathbf{E}_{DB}\mathbf{E}_{DD}\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS} - \mathbf{\Gamma}^{m,k}_{P}\right)\right]. \tag{125}$$

Now, we first compute the covariance matrix of the residual self-interference $\Delta\mathbf{I}^{n,k}_{DD}$

$$\begin{aligned}\mathbf{C}_{\Delta\mathbf{I}_{DD}} &= \mathcal{E}\left[\Delta\mathbf{I}^{n,k}_{DD}\left(\Delta\mathbf{I}^{n,k}_{DD}\right)^{H}\right] \\ &= \mathcal{E}\left[\left(\hat{\mathbf{\Gamma}}^{n,k}_{DD} - \mathbf{\Gamma}^{n,k}_{DD}\right)\mathbf{x}^{n,k}_{D}\left(\mathbf{x}^{n,k}_{D}\right)^{H} \times \left(\hat{\mathbf{\Gamma}}^{n,k}_{DD} - \mathbf{\Gamma}^{n,k}_{DD}\right)^{H}\right].\end{aligned} \tag{126}$$

Considering $\mathcal{E}[\mathbf{x}^{n,k}_{D}(\mathbf{x}^{n,k}_{D})^{H}] = P_D\mathbf{I}_2$, the above covariance matrix reduces to

$$\mathcal{E}\left[\Delta\mathbf{I}^{n,k}_{DD}\left(\Delta\mathbf{I}^{n,k}_{DD}\right)^{H}\right] = P_D\mathcal{E}\left[\left(\hat{\mathbf{\Gamma}}^{n,k}_{DD} - \mathbf{\Gamma}^{n,k}_{DD}\right) \times \left(\hat{\mathbf{\Gamma}}^{n,k}_{DD} - \mathbf{\Gamma}^{n,k}_{DD}\right)^{H}\right]. \tag{127}$$

Placing (125) in (127) produces

$$\begin{aligned}\mathcal{E}\left[\Delta\mathbf{I}^{n,k}_{DD}\left(\Delta\mathbf{I}^{n,k}_{DD}\right)^{H}\right] &= P_D\mathbf{E}_{DA}\mathbf{E}_{DD} \\ &\times \mathcal{E}\left[\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS} - \mathbf{\Gamma}^{m,k}_{P}\right)\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS} - \mathbf{\Gamma}^{m,k}_{P}\right)^{H}\right] \\ &\times \mathbf{E}^{T}_{DD}\mathbf{E}^{T}_{DA} + P_D\mathbf{E}_{DB}\mathbf{E}_{DD}\bullet \\ &\times \mathcal{E}\left[\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS} - \mathbf{\Gamma}^{m,k}_{P}\right)\left(\hat{\mathbf{\Gamma}}^{m,k}_{P,FD-LS} - \mathbf{\Gamma}^{m,k}_{P}\right)^{H}\right] \\ &\times \mathbf{E}^{T}_{DD}\mathbf{E}^{T}_{DB}.\end{aligned} \tag{128}$$

Using the results of (49), the above covariance matrix can be simplified as

$$\mathcal{E}\left[\Delta\mathbf{I}^{n,k}_{DD}\left(\Delta\mathbf{I}^{n,k}_{DD}\right)^{H}\right] = \frac{2P_D}{N_P(P_D+P_S)}\mathbf{C}_{2\times 2}. \tag{129}$$

Considering the covariance matrix of channel noise vector $\mathbf{w}^{n,k}$ is

$$\mathbf{C}_{\mathbf{w}^{n,k}} = \mathcal{E}\left[\mathbf{w}^{n,k}\left(\mathbf{w}^{n,k}\right)^{H}\right] = \mathbf{C}_{2\times 2} \tag{130}$$

and $\mathbf{w}^{n,k}$ is independent of $\mathbf{w}^{m,k}_{P}$ and $\mathbf{x}^{n,k}_{D}$, we have

$$\begin{aligned}\mathbf{C}_{IN} &= \mathcal{E}\left[\left(\Delta\mathbf{I}^{n,k}_{DD} + \mathbf{w}^{n,k}\right)\left(\Delta\mathbf{I}^{n,k}_{DD} + \mathbf{w}^{n,k}\right)^{H}\right] \\ &= \mathbf{C}_{\Delta\mathbf{I}_{DD}} + \mathbf{C}_{\mathbf{w}^{n,k}} = \left(1 + \frac{2P_D}{N_P(P_S+P_D)}\right)\mathbf{C}_{2\times 2}.\end{aligned} \tag{131}$$

Until now, we completes the deriving of the covariance matrix of the residual self-interference and noise. ■

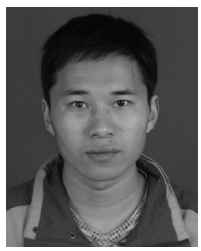
ACKNOWLEDGMENT

The authors would like to thank all anonymous reviewers and Prof. H. Minn for their astute and constructive technical comments. Their comments significantly improve the quality of this paper.

REFERENCES

[1] T. Riihonen, S. Werner, and R. Wichman, "Optimized gain control for single-frequency relaying with loop interference," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2801–2806, Jun. 2009.
[2] H. Ju, E. Oh, and D. Hong, "Catching resource-devouring worms in next generation wireless relay systems: Two-way relay and full duplex relay," *IEEE Commun. Mag.*, vol. 47, no. 9, pp. 58–65, Sep. 2009.
[3] T. M. Kim, H. J. Yang, and A. J. Paulraj, "Distributed sum-rate optimization for full-duplex MIMO system under limited dynamic range," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 555–558, Jun. 2013.
[4] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.
[5] D. Kim, H. Ju, S. Park, and D. Hong, "Effects of channel estimation error on full-duplex two-way networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4666–4672, Nov. 2013.
[6] R. Hu, M. Peng, Z. Zhao, and X. Xie, "Investigation of full-duplex relay networks with imperfect channel estimation," in *Proc. IEEE/CIC Int. Conf. Commun. China*, Oct. 2014, pp. 576–580.
[7] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.
[8] R. Hu, M. Peng, Z. Zhao, and X. Xie, "Full-duplex wireless communication using transmitter output based echo cancellation," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2011, pp. 1–5.
[9] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator in MIMO full-duplex systems," in *Proc. IEEE 80th Veh. Technol. Conf.*, Sep. 2014, pp. 1–5.
[10] X. F. Li and C. Tepedelenlioglu, "Maximum likelihood channel estimation for residual self-interference cancellation in full duplex relays," in *Proc. 49th Asilomar Conf. Signals, Syst., Comput.*, Feb. 2015, pp. 807–811.
[11] A. Koohian, H. Mehrpouyan, M. Ahmadian, and M. Azarbad, "Bandwidth efficient channel estimation for full-duplex communication systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 4710–4714.
[12] B. Razavi, *RF Microelectronics*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.
[13] Y. Li, *In-Phase and Quadrature Imbalance-Modeling, Estimation, and Compensation*. New York, NY, USA: Springer-Verlag, 2014.
[14] A. Tarighat and A. H. Sayed, "Compensation schemes and performance analysis of I/Q imbalances in OFDM receiver," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3257–3268, Aug. 2005.
[15] A. Tarighat and A. H. Sayed, "Joint compensation of transmitter and receiver impairments in OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 240–247, Jan. 2007.
[16] E. Lopez-Estraviz, S. De Rore, F. Horlin, and A. Bourdoux, "Pilot design for joint channel and frequency-dependent transmit/receive IQ imbalance estimation and compensation in OFDM-based transceivers," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2007, pp. 4861–4866.
[17] H. Minn and D. Munoz, "Pilot designs for channel estimation of MIMO-OFDM systems with frequency-dependent I/Q imbalances," *IEEE Trans. Commun.*, vol. 58, no. 8, pp. 2252–2264, Jun. 2010.
[18] F. Shu, J. Zhao, X. You, M. Wang, Q. Chen, and B. Stevan, "An efficient sparse channel estimator combining time-domain LS and iterative shrinkage for OFDM systems with IQ-imbalances," *Sci. China Inf. Sci.*, vol. 53, no. 11, pp. 2604–2610, Nov. 2012.
[19] Y. Liang, F. Shu, Y. Zhang, and J. Zhao, "High-performance compensation scheme for frequency-dependent IQ imbalances in OFDM transmitter and receiver," *J. Syst. Eng. Electron.*, vol. 24, no. 2, pp. 204–208, Apr. 2013.
[20] H. Lin and K. Yamashita, "Time domain blind I/Q imbalance compensation based on real-valued filter," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4342–4350, Dec. 2013.
[21] M. Beheshti, M. J. Omidi, and A. M. Doost-Hoseini, "Joint compensation of transmitter and receiver IQ imbalance for MIMO-OFDM over doubly selective channels," *Wireless Pers. Commun.*, vol. 70, no. 2, pp. 537–559, May 2013.
[22] N. Kolomvakis, M. Matthaiou, J. Li, M. Coldrey, and T. Svensson, "Massive MIMO with IQ imbalance: Performance analysis and compensation," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 1703–1709.
[23] A. K. Gupta and D. K. Nagar, *Matrix Variate Distributions* (Monographs & Surveys in Pure & Applied Mathematics). London, U.K.: Chapman & Hall/CRC, 2000.
[24] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Upper Saddle River, NJ, USA: Prentice-Hall, 1992.
[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2013.
[27] M. H. Hsieh and C. H. Wei, "Channel estimation for OFDM systems based on comb-type pilot arrangement in frequency selective fading channels," *IEEE Trans. Consum. Electron.*, vol. 44, no. 1, pp. 217–225, Aug. 1998.

Feng Shu (M'07) received the B.S. degree from Fuyang Teaching College, Fuyang, China, in 1994; the M.S. degree from Xidian University, Xi'an, China, in 1997; and the Ph.D. degree from Southeast University, Nanjing, China, in 2002. From September 2009 to September 2010, he was a Visiting Postdoctoral Researcher at the University of Texas at Dallas. In October 2005, he joined the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, where he is currently a Professor and Ph.D. supervisor. He has authored or coauthored about 200 papers, of which more than 80 are in archival journals, including more than 20 papers in IEEE Journals and 43 SCI-indexed papers. He holds four Chinese patents. His research interests include wireless networks, wireless location, and array signal processing.



Jin Wang received the B.S. degree in 2012 from Nanjing University of Science and Technology, Nanjing, China, where he is currently working toward the Ph.D. degree with the School of Electronic and Optical Engineering. His research interests include wireless communications and signal processing.

Jun Li (M'09–SM'16) received the Ph.D. degree in electronic engineering from Shanghai Jiao Tong 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. Since 2015, he has been with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China. His research interests include network information theory, channel coding theory, wireless network coding, and cooperative communications.

Riqing Chen received the B.Eng. degree in communication engineering from Tongji University, Shanghai, China, in 2001; the M.Sc. degree in communications and signal processing from Imperial College London, London, U.K., in 2004; and the Ph.D. degree in engineering science from the University of Oxford, Oxford, U.K., in 2010. He is currently a Lecturer at the Faculty of Computer Science and Information Technology, Fujian Agriculture and Forestry University, Fujian, China. His research interests include network security, spread coding, signal processing, and wireless sensor technologies.

Wen Chen (M'03–SM'11) received the B.S. and M.S. degrees from Wuhan University, Wuhan, China, in 1990 and 1993, respectively, and the Ph.D. degree from the University of Electro-Communications, Tokyo, Japan, in 1999. From 1999 to 2001, he was a Researcher with the Japan Society for the Promotion of Sciences. Since 2006, he has been a Full Professor with the Department of Electronic Engineering, Shanghai Jiao Tong University (SJTU), Shanghai, China, where he is also the Director of the Institute for Signal Processing and Systems. From 2014 to 2015, he was the Dean of the School of Electronic Engineering and Automation with the Guilin University of Electronic Technology, China. Since 2016, he has been the Chairman of SJTU Intelligent Property Management Corporation. His research interests include network coding, cooperative communications, multiple-access technique, and multiple-input multiple-output orthogonal frequency-division multiplexing systems.