Sum-MSE Gain of DFT-Based Channel Estimator Over Frequency-Domain LS One in Full-Duplex OFDM Systems

Jin Wang, Hai Yu, Feng Shu, *Member, IEEE*, Jinhui Lu, Riqing Chen, Jun Li, and Dushantha Nalin K. Jayakody, *Member, IEEE*

***Abstract*—In this paper, we make an investigation on the sum-mean-square-error (Sum-MSE) performance gain in full-duplex orthogonal frequency-division multiplexing (OFDM) systems in the presence of colored interference-plus-noise (IPN). This gain is defined as the ratio of Sum-MSE of frequency-domain least-square (LS) channel estimator to that of DFT-based LS one. The closed-form formula of the gain is derived. And, its simple upper and lower bounds are given using inequalities of matrix eigenvalues. The exact value of Sum-MSE gain depends heavily on the correlation factor of the IPN covariance matrix. More importantly, we also find that the Sum-MSE performance gain grows from 1 to N/L as the correlation factor gradually decreases from 1 to 0, where N and L denote the number of total subcarrier and the length of cyclic prefix, respectively. Also, via theoretical analysis, the exact Sum-MSE gain degenerates into 1 and N/L in two extreme scenarios: fully-correlated and white, respectively. The former 1 means there is no performance gain, while the latter N/L corresponds to the maximum Sum-MSE performance gain achievable. Numerical simulation further validates the above results. Additionally, the derived lower bound is shown to be closer to the exact value of Sum-MSE gain compared to the upper bound.**

***Index Terms*—Channel estimation, full duplex (FD), least squares (LSs), orthogonal frequency-division multiplexing (OFDM), sum-MSE performance gain, upper/lower bound.**

Manuscript received November 24, 2017; revised April 10, 2018; accepted June 23, 2018. Date of publication July 13, 2018; date of current version May 31, 2019. This work was supported in part by the National Natural Science Foundation of China under Grants 61472190, 61771244, 61702258, 61501238, and 61602245, in part by the China Postdoctoral Science Foundation under Grant 2016M591852, in part by the Postdoctoral research funding program of Jiangsu Province under Grant 1601257C, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20150791, and in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, China under Grant 2013D02. *(Corresponding authors: Feng Shu; Jinhui Lu; Riqing Chen.)*

J. Wang, H. Yu, J. Lu, and J. Li are with School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: jin.wang@njust.edu.cn; yuhai_njust@163.com; ljh_412@sina.com; jun.li@njust.edu.cn).

F. Shu is with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China, with the College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China, and also with the National Key Laboratory of Electromagnetic Environment, China Research Institute of Radiowave Propagation, Qingdao 266107, China (e-mail: shufeng0101@163.com).

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

D. N. K. Jayakody is with the Department of Control System Optimization, Institute of Cybernetics, National Research Tomsk Polytechnic University, Tomsk 634050, Russia (e-mail: nalin.jayakody@ieee.org).

Digital Object Identifier 10.1109/JSYST.2018.2850934

I. Introduction

In THE recent decade, full-duplex (FD) technique has become a hot research field in wireless communications, satellite communications, and mobile networks [1]–[5]. Compared to time-division-duplex (TDD) and frequency-division-duplex (FDD) techniques, FD has the potential to double data transmission rate by simultaneously transmitting and receiving signals over the same frequency band and time slot [6]–[10]. The major problem concerning FD is that the weak fading received signal is severely interfered by the strong FD self-interference (SI) [9]–[11]. In [12], the SI cancellation schemes are divided into three categories: propagation-domain, analog-circuit-domain, and digital-domain methods. In such a system, the high-performance channel estimator becomes extremely important in order to dramatically reduce the effect of SI [13]–[15].

Channel estimation has been extensively investigated in conventional TDD/FDD mode [16]–[18]. In general, the mean square error (MSE) is adopted as a performance metric for estimators [19]. For linear systems, several typical estimation schemes, such as minimum MSE (MMSE) estimator [19], Kalman filter [20], recursive least-square (LS) estimator [21], have been used to attain channel state information. Particle filter and sequential Monte Carlo methods are commonly applied to track channel parameters in nonlinear or dynamic systems [22]–[25]. Pilot sequences are very important for data-assisted estimators, and various literature have discussed the problems of pilot design and arrangement [26]–[28].

However, due to the existence of FD SI in FD systems, channel estimation and pilot optimization become a challenging problem and should be restudied. A digitally assisted analog channel estimator is designed to estimate SI channel for in-band FD radios [29]. Based on the expectation maximization method, authors in [30] propose a blind channel estimator to do simultaneous estimation for SI and intended channels in FD wireless systems. Using the maximum-likelihood criterion, the SI and intended channels are jointly estimated with the known transmitted symbols from itself and the pilot symbols from intended transceiver [31]. Besides, an iterative procedure is constructed to further enhance the estimation performance in the high signal-to-noise ratio region [32]. To exploit and measure the sparsity of SI channel and intended channel, authors in [33] propose a time-domain LS (TD-LS) channel estimator realized by adaptive

orthogonal matching pursuit scheme. In FD orthogonal frequency-division multiplexing (OFDM) systems with IQ imbalances, two channel estimators, frequency-domain and DFT-based LS, are presented. Subsequently, the corresponding optimal pilot product matrix is proved to be an identity matrix multiplied by a constant [34]. Also, the Sum-MSE performance gain of the DFT-based LS channel estimator over the frequency-domain one is derived to be N/L in white Gaussian noise scenario, where N is the total number of subcarriers and L is the length of cyclic prefix (CP).

How about the Sum-MSE performance gain in the colored interference/noise scenarios? Is it still equal to N/L? We will focus on the aspect research and make our effort to address this problem, our main contributions are summarized as follows.

1) In FD OFDM systems with colored interference-plus-noise (IPN), the optimal pilot pattern of minimizing the Sum-MSE per subcarrier is proposed and derived for frequency-domain LS (FD-LS) channel estimator. First, we combine the pilot symbols from source and destination nodes into a pilot matrix as the optimization variable. Then, an optimization problem is established to minimize the sum of MSEs of SI and intended channel per subcarrier, which is subject to two constraints of transmit powers at both source and destination nodes. By using the inverse matrix trace inequality in matrix theory, the column vectors of the optimal pilot matrix are proved to be orthogonal to each other.
2) We derive and analyze the Sum-MSE performance gain achieved by the DFT-based LS channel estimator over the FD-LS estimator in FD OFDM systems. The exact expression of Sum-MSE performance gain is defined as the Sum-MSE ratio of FD-LS channel estimator to DFT-based one. Due to the complex form of this gain, its simple lower bound and upper bounds are also proved to simplify the performance analysis and discussion in our simulation part by making use of eigenvalue inequalities in matrix theory. Additionally, the derived lower bound is tighter than the upper bound in several typical extreme scenarios.
3) From simulation results and analysis, we also find several important facts: a) In general, the DFT-based LS channel estimator is better than or equal to the FD-LS one in terms of Sum-MSE performance. b) As colored IPN goes white, the Sum-MSE performance gain increases up to N/L, which is the achievable maximum gain value. c) As the colored IPN tends to be fully-correlated, the Sum-MSE performance gain approaches one, i.e., no gain. In other words, the FD-LS can achieve the same Sum-MSE performance as DFT-based LS in the fully-correlated case.

The remainder is organized as follows. The FD system model is described in Section II. In Section III, the FD-LS estimator is adopted to estimate both intended and SI channels, and the corresponding optimal pilot matrix is derived. In Section IV, the Sum-MSE performance gain of the DFT-based channel estimator over the FD-LS one is defined, and its upper bound and lower bounds are derived. Simulation results and discussions are presented in Section V. Finally, Section VI concludes this whole paper.

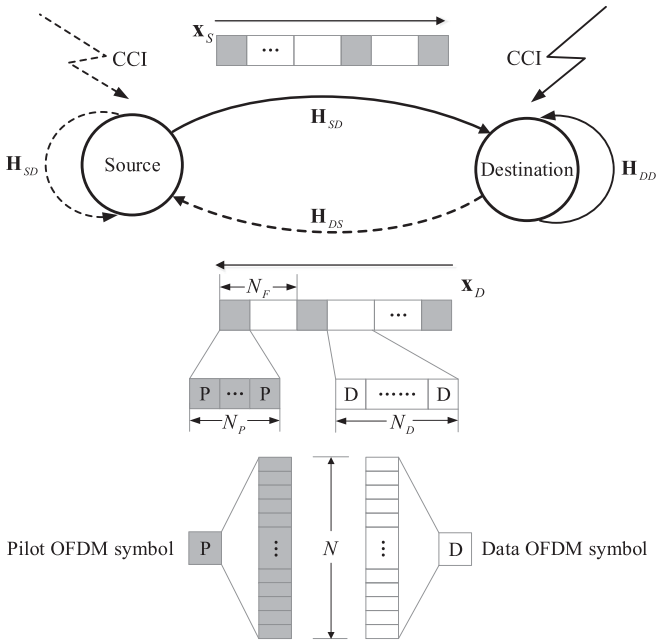


Fig. 1. Schematic diagram of FD OFDM system model.

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 $n \times m$.

II. System Model

Fig. 1 sketches the schematic diagram of a point-to-point FD OFDM system. Here, the destination node is used as a reference. The received vector at destination node is composed of three parts: the signal from source node via intended channel $\mathbf{H}_{\rm SD}$, the signal from local transmitter via SI channel $\mathbf{H}_{\rm DD}$, and the cochannel interference (CCI) from other nodes. Both channels from source to destination (S2D) and destination to destination (D2D) are assumed to be time-invariant during each frame, where each frame consists of N_F OFDM symbols, but vary from one frame to another one. Here, each frame consists of N_P pilot OFDM symbols and N_D data OFDM symbols, which is shown in Fig. 1. Usually, in a practical system, N_D is taken to be far larger than N_P to achieve a high spectrum-efficiency. In Fig. 1, a typical block-type pilot pattern is adopted to estimate both D2D and S2D channels.

Similar to [35], the ideal channel frequency responses (CFR) has the following relationship with its channel impulse responses (CIR):

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

and

$$\mathbf{H}_{\text{DD}} = \mathbf{F}_{N\times N}\begin{pmatrix}\mathbf{h}_{\text{DD}}\\ \mathbf{0}_{(N-L)\times 1}\end{pmatrix} = \mathbf{F}_{N\times L}\,\mathbf{h}_{\text{DD}} \tag{2}$$

where $\mathbf{h}_{\text{SD}}$ and $\mathbf{h}_{\text{DD}}$ are the S2D and D2D CIRs defined by

$$\mathbf{h}_{\text{SD}} = [h_{\text{SD}}(1)\, h_{\text{SD}}(2)\ \ldots\ h_{\text{SD}}(L)]^T \tag{3}$$

and

$$\mathbf{h}_{\text{DD}} = [h_{\text{DD}}(1)\, h_{\text{DD}}(2)\ \ldots\ h_{\text{DD}}(L)]^T \tag{4}$$

respectively. N is the total number of subcarriers, L is the length of the CP, and $\mathbf{F}_{N\times N}$ is the normalized discrete Fourier transform matrix as

$$\mathbf{F}_{N\times N} = \frac{1}{\sqrt{N}}\begin{pmatrix}1 & 1 & \cdots & 1\\ 1 & W^1 & \cdots & W^{N-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & W^{N-1} & \cdots & W^{(N-1)(N-1)}\end{pmatrix} \tag{5}$$

with $W = e^{-j\frac{2\pi}{N}}$.

The transmit vectors corresponding to the nth OFDM symbol from source and destination are denoted by

$$\mathbf{x}_S(n,:) = (x_S(n,1)\, x_S(n,2)\ \ldots\ x_S(n,N))^T \tag{6}$$

and

$$\mathbf{x}_D(n,:) = (x_D(n,1)\, x_D(n,2)\ \ldots\ x_D(n,N))^T \tag{7}$$

with average power limits P_S and P_D per subcarrier, respectively. After experiencing multipath channel $\mathbf{H}_{\text{SD}}$ and $\mathbf{H}_{\text{DD}}$, the nth received OFDM symbol at destination node is given by

$$\begin{aligned}\mathbf{y}(n,:) &= \text{diag}\{\mathbf{x}_S(n,:)\}\mathbf{H}_{\text{SD}} + \text{diag}\{\mathbf{x}_D(n,:)\}\mathbf{H}_{\text{DD}}\\ &\quad + \mathbf{w}_{\text{cci}}(n,:) + \mathbf{w}_n(n,:)\\ &= \begin{pmatrix}\text{diag}\{\mathbf{x}_S(n,:)\} & \text{diag}\{\mathbf{x}_D(n,:)\}\end{pmatrix}\begin{pmatrix}\mathbf{H}_{\text{SD}}\\ \mathbf{H}_{\text{DD}}\end{pmatrix}\\ &\quad + \mathbf{w}_{\text{cci}}(n,:) + \mathbf{w}_n(n,:),\end{aligned} \tag{8}$$

where $\mathbf{w}_{\text{cci}}(n,:)$ denotes the CCI vector and $\mathbf{w}_n(n,:)$ is the noise vector in frequency domain. For convenience of the following derivation and analysis, the sum of CCI and additive noise vector will be viewed as the new colored IPN vector as follows:

$$\mathbf{w}(n,:) = \mathbf{w}_{\text{cci}}(n,:) + \mathbf{w}_n(n,:) \tag{9}$$

where the above IPN vector is assumed to be independent along time direction n [36].

Obviously, (8) is an under-determined linear equation. Thus, at least two pilot OFDM symbols, $N_P \geq 2$, are required for source node and destination node to estimate both $\mathbf{H}_{\text{SD}}$ and $\mathbf{H}_{\text{DD}}$, which is expressed as follows:

$$\begin{pmatrix}\mathbf{y}(n,:)\\ \mathbf{y}(n+1,:)\\ \vdots\\ \mathbf{y}(n+N_P-1,:)\end{pmatrix} = \begin{pmatrix}\text{diag}\{\mathbf{x}_S(n,:)\} & \text{diag}\{\mathbf{x}_D(n,:)\}\\ \text{diag}\{\mathbf{x}_S(n+1,:)\} & \text{diag}\{\mathbf{x}_D(n+1,:)\}\\ \vdots & \vdots\\ \text{diag}\{\mathbf{x}_S(n+N_P-1,:)\} & \text{diag}\{\mathbf{x}_D(n+N_P-1,:)\}\end{pmatrix}\begin{pmatrix}\mathbf{H}_{\text{SD}}\\ \mathbf{H}_{\text{DD}}\end{pmatrix} + \underbrace{\begin{pmatrix}\mathbf{w}(n,:)\\ \mathbf{w}(n+1,:)\\ \vdots\\ \mathbf{w}(n+N_P-1,:)\end{pmatrix}}_{\tilde{\mathbf{w}}}. \tag{10}$$

If the above IPN vector $\tilde{\mathbf{w}}$ is colored, and its correlation function along time and frequency are assumed to be independent, the covariance matrix of colored IPN vector has the following form:

$$\mathbf{R}_{\text{cci}} = \mathcal{E}\{\tilde{\mathbf{w}}\tilde{\mathbf{w}}^H\} = \mathbf{R}_t \otimes \mathbf{R}_w \tag{11}$$

where matrix $\mathbf{R}_w$ denotes the $N\times N$ frequency-domain covariance matrix given by

$$\mathbf{R}_w = \begin{pmatrix}r_w(0) & r_w(-1) & \cdots & r_w(1-N)\\ r_w(1) & r_w(0) & \cdots & r_w(2-N)\\ \vdots & \vdots & \ddots & \vdots\\ r_w(N-1) & r_w(N-2) & \cdots & r_w(0)\end{pmatrix} \tag{12}$$

and matrix $\mathbf{R}_t$ denotes the $N_P \times N_P$ time-direction covariance matrix given by

$$\mathbf{R}_t = \begin{pmatrix}r_t(0) & r_t(-1) & \cdots & r_t(1-N_P)\\ r_t(1) & r_t(0) & \cdots & r_t(2-N_P)\\ \vdots & \vdots & \ddots & \vdots\\ r_t(N_P-1) & r_t(N_P-2) & \cdots & r_t(0)\end{pmatrix} \tag{13}$$

with $r_w(\Delta k) = \mathcal{E}\{w(n,k)w(n,k+\Delta k)^*\}$ and $r_t(\Delta n) = \mathcal{E}\{w(n,k)w(n+\Delta n,k)^*\}$.

III. FD-LS Estimator and the Corresponding Optimal Pilot Matrix

In this section, we derive a closed-form expression for the optimal pilot matrix per subcarrier. For the convenience of deriving below, we stack all kth subcarriers of N_P received

pilot OFDM symbols as follows:

$$\mathbf{y}_k = \mathbf{P}_k \mathbf{H}_k + \mathbf{w}_k \tag{14}$$

with

$$\mathbf{y}_k = \left(y(n,k) \cdots y(n+N_P-1,k) \right)^T \tag{15}$$

$$\mathbf{P}_k = \begin{pmatrix} x_S(n,k) & x_D(n,k) \\ \vdots & \vdots \\ x_S(n+N_P-1,k) & x_D(n+N_P-1,k) \end{pmatrix} \tag{16}$$

$$\mathbf{H}_k = \begin{pmatrix} H_{\mathrm{SD}}(k) & H_{\mathrm{DD}}(k) \end{pmatrix}^T \tag{17}$$

and

$$\mathbf{w}_k = \left(w(n,k) \cdots w(n+N_P-1,k) \right)^T. \tag{18}$$

Note that the $\mathbf{H}_k$ in (17) is just the channel parameter to be estimated.

In terms of (14), the FD-LS estimator can be readily given by

$$\hat{\mathbf{H}}_k = \left(\mathbf{P}_k^H \mathbf{P}_k \right)^{-1} \mathbf{P}_k^H \mathbf{y}_k = \mathbf{H}_k + \left(\mathbf{P}_k^H \mathbf{P}_k \right)^{-1} \mathbf{P}_k^H \mathbf{w}_k \tag{19}$$

which yields the channel estimation error

$$\Delta \mathbf{H}_k = \hat{\mathbf{H}}_k - \mathbf{H}_k = \left(\mathbf{P}_k^H \mathbf{P}_k \right)^{-1} \mathbf{P}_k^H \mathbf{w}_k \tag{20}$$

which forms the MSE of the FD-LS estimator over subcarrier k as follows:

$$\begin{aligned} \mathrm{MSE}_k &= \mathcal{E}\left\{ \mathrm{tr}\left[\Delta \mathbf{H}_k (\Delta \mathbf{H}_k)^H \right] \right\} \\ &= \mathrm{tr}\left[\mathbf{P}_k \left(\mathbf{P}_k^H \mathbf{P}_k \right)^{-2} \mathbf{P}_k^H \mathcal{E}\left\{ \mathbf{w}_k \mathbf{w}_k^H \right\} \right]. \end{aligned} \tag{21}$$

Now, the IPN vector $\mathbf{w}_k$ is assumed to be independent across different OFDM symbols in order to optimize the performance of FD-LS, the corresponding time-direction covariance matrix will be

$$\mathbf{R}_t = \mathcal{E}\left\{ \mathbf{w}_k \mathbf{w}_k^H \right\} = \sigma_I^2 \mathbf{I}_{N_P} \tag{22}$$

where σ_I^2 represents the average power of IPN [36]. Thus, the MSE will be simplified as

$$\mathrm{MSE}_k = \sigma_I^2 \mathrm{tr}\left[\left(\mathbf{P}_k^H \mathbf{P}_k \right)^{-1} \right]. \tag{23}$$

By minimizing the above MSE_k to design the pilot matrix $\mathbf{P}_k$ subject to the constraint of transmit power at source and destination nodes, we establish the following optimization problem:

$$\begin{aligned} \min \quad & \mathrm{tr}\left[\left(\mathbf{P}_k^H \mathbf{P}_k \right)^{-1} \right] \\ \text{s.t.} \quad & (\mathbf{P}_k^H \mathbf{P}_k)_{11} \le N_P P_S \\ & (\mathbf{P}_k^H \mathbf{P}_k)_{22} \le N_P P_D \end{aligned} \tag{24}$$

where P_S and P_D are the average transmit power per subcarrier of source and destination nodes, respectively.

Define $\mathbf{X}_k = \mathbf{P}_k^H \mathbf{P}_k$, the optimization problem in (24) is equivalent to

$$\begin{aligned} \min \quad & \mathrm{tr}(\mathbf{X}_k^{-1}) \\ \text{s.t.} \quad & (\mathbf{X}_k)_{11} \le N_P P_S \\ & (\mathbf{X}_k)_{22} \le N_P P_D. \end{aligned} \tag{25}$$

According to [37, Lemma 1], for positive definite matrix $\mathbf{Y} \in \mathbb{C}^{N \times N}$ with its diagonal element being $\{Y_{ii}\}_{i=1}^N$, the inverse matrix trace inequality $tr(\mathbf{Y}^{-1}) \ge \sum_{i=1}^N Y_{ii}^{-1}$ holds, and the equality holds if $\mathbf{Y}$ is diagonal. It implies that if the positive definite matrix $\mathbf{X}_k$ can be diagonally designed with the power constraints in (25), we claim that the optimal solution is obtained.

Therefore, it is evident that the optimal solution of problem (25) should be

$$\mathbf{X}_k = \begin{pmatrix} N_P P_S & 0 \\ 0 & N_P P_D \end{pmatrix}. \tag{26}$$

We summarize the above result in a theorem as follows.

Theorem 1: For an OFDM-based FD system, the optimal pilot matrix $\mathbf{P}_k$ should satisfy $\mathbf{P}_k^H \mathbf{P}_k = \mathrm{diag}\{N_P P_S, N_P P_D\}$ so as to minimize the Sum-MSE for the FD-LS estimator. It follows immediately that the optimal $\mathbf{P}_k$ needs to be column orthogonal matrix.

We can construct $\mathbf{P}_k$ as any two columns of $N_P \times N_P$ unitary matrix multiplied by any predefined constant value. For example, given $N_P = 4$ and 16QAM constellation, the optimal $\mathbf{P}_k$ is

$$\mathbf{P}_k{}^\star = \frac{(1+3i)}{\sqrt{10}} \begin{pmatrix} \sqrt{P_S} & \sqrt{P_D} \\ \sqrt{P_S} & \sqrt{P_D} \\ \sqrt{P_S} & -\sqrt{P_D} \\ \sqrt{P_S} & -\sqrt{P_D} \end{pmatrix}. \tag{27}$$

IV. DFT-Based LS Channel Estimator and Sum-MSE Performance Gain Analysis

In this section, we derive the Sum-MSE performance expressions of both FD-LS and DFT-based LS channel estimators. The Sum-MSE performance gain is defined as the ratio of the MSE of FD-LS channel estimator to that of DFT-based one. Its upper bound and lower bound are derived jointly. In two extreme situations: independent and fully-correlated, the simple expressions of the corresponding Sum-MSE performance gains are directly given and discussed together.

When the optimal pilot matrix satisfying (26) is adopted, the FD-LS estimator in (19) will become

$$\hat{\mathbf{H}}_k = \mathbf{H}_k + \mathbf{X}_k^{-1} \mathbf{P}_k^H \mathbf{w}_k \tag{28}$$

which is equivalently written in the following form

$$\begin{aligned} \hat{H}_{\mathrm{SD}}(k) &= H_{\mathrm{SD}}(k) \\ &\quad + \frac{1}{N_P P_S} \sum_{p=0}^{N_P-1} x_S^*(n+p,k) w(n+p,k) \end{aligned} \tag{29}$$

and

$$\begin{aligned} \hat{H}_{\mathrm{DD}}(k) &= H_{\mathrm{DD}}(k) \\ &\quad + \frac{1}{N_P P_D} \sum_{p=0}^{N_P-1} x_D^*(n+p,k) w(n+p,k). \end{aligned} \tag{30}$$

Due to the similar forms of the above two equations, we take $\mathbf{H}_{\mathrm{DD}}$ as an example below. Stacking all subcarriers, we can

obtain the estimated channel gain vector as follows:

$$\hat{\mathbf{H}}_{\text{DD}} = \mathbf{H}_{\text{DD}} + \mathbf{e}_{\text{DD}} \tag{31}$$

where

$$\mathbf{e}_{\text{DD}} = \frac{1}{N_P P_D} \sum_{p=0}^{N_P-1} \text{diag}\{\mathbf{x}_D^*(n+p,:)\}\mathbf{w}(n+p,:) \tag{32}$$

denotes the estimation error. Thus, the corresponding MSE of $\mathbf{H}_{\text{DD}}$ is given by

$$\begin{aligned}\text{MSE}_{\text{DD}} &= \mathcal{E}\{\text{tr}[\mathbf{e}_{\text{DD}}(\mathbf{e}_{\text{DD}})^H]\} \\ &= \text{tr}[\mathcal{E}\{\mathbf{e}_{\text{DD}}(\mathbf{e}_{\text{DD}})^H\}].\end{aligned} \tag{33}$$

Because of

$$\mathcal{E}\{\mathbf{x}_D^*(n+p,:)\mathbf{x}_D(n+p,:)^T\} = P_D\mathbf{I}_N \tag{34}$$

we have

$$\mathcal{E}\{\mathbf{e}_{\text{DD}}(\mathbf{e}_{\text{DD}})^H\} = \frac{\sigma_I^2}{2P_D}\mathbf{A} \tag{35}$$

where $\mathbf{A}$ denotes the normalized frequency-domain covariance matrix of $\mathbf{w}(n,:)$ with

$$\mathbf{R}_w = \sigma_I^2\mathbf{A}. \tag{36}$$

Obviously, the above matrix $\mathbf{A}$ is a positive semidefinite Hermitian matrix and all its diagonal elements are 1. Therefore, the above MSE of $\mathbf{H}_{\text{DD}}$ can be simplified as

$$\text{MSE}_{\text{DD}} = \frac{N\sigma_I^2}{N_P P_D}. \tag{37}$$

In the same fashion, we can obtain the MSE_{SD} corresponding to $\mathbf{H}_{\text{SD}}$ as follows:

$$\text{MSE}_{\text{SD}} = \frac{N\sigma_I^2}{N_P P_S}. \tag{38}$$

In terms of the above two MSEs, we define the Sum-MSE performance of the FD-LS channel estimator as follows:

$$\begin{aligned}\text{SumMSE}_{\text{FD-LS}} &= \text{MSE}_{\text{DD}} + \text{MSE}_{\text{SD}} \\ &= \left(\frac{1}{N_P P_S} + \frac{1}{N_P P_D}\right) N\sigma_I^2.\end{aligned} \tag{39}$$

Using the transformation relationship in (2), the corresponding estimated time-domain CIR channel gain vectors will be given by

$$\hat{\mathbf{h}}_{\text{DD}} = \mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^H\hat{\mathbf{H}}_{\text{DD}} \tag{40}$$

where

$$\mathbf{E}_{L\times N} = \begin{pmatrix}\mathbf{I}_L & \mathbf{0}_{L\times(N-L)}\end{pmatrix}. \tag{41}$$

Performing the DFT operation to both sides of (40) yields the following DFT-based channel estimator:

$$\tilde{\mathbf{H}}_{\text{DD}} = \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^H\hat{\mathbf{H}}_{\text{DD}} \tag{42}$$

which, in combination with the error model in (31), forms the estimation error model of the DFT-based LS estimator

$$\begin{aligned}\tilde{\mathbf{H}}_{\text{DD}} &= \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^H\hat{\mathbf{H}}_{\text{DD}} \\ &= \mathbf{H}_{\text{DD}} + \mathbf{F}_{N\times L}\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^H\mathbf{e}_{\text{DD}}.\end{aligned} \tag{43}$$

From (43), the corresponding MSE of $\mathbf{H}_{\text{DD}}$ by DFT-based LS is represented as

$$\begin{aligned}\text{MSE}'_{\text{DD}} = \mathcal{E}\Big\{\text{tr}\Big[&(\mathbf{e}_{\text{DD}})^H\mathbf{F}_{N\times N}\mathbf{E}_{L\times N}^H\mathbf{F}_{N\times L}^H\mathbf{F}_{N\times L} \\ &\mathbf{E}_{L\times N}\mathbf{F}_{N\times N}^H\mathbf{e}_{\text{DD}}\Big]\Big\}.\end{aligned} \tag{44}$$

Since $\mathbf{F}_{N\times N}\mathbf{E}_{L\times N}^H = \mathbf{F}_{N\times L}$ and $\mathbf{F}_{N\times L}^H\mathbf{F}_{N\times L} = \mathbf{I}_L$, the above MSE reduces to

$$\begin{aligned}\text{MSE}'_{\text{DD}} &= \mathcal{E}\{\text{tr}[(\mathbf{e}_{\text{DD}}^n)^H\mathbf{F}_{N\times L}\mathbf{F}_{N\times L}^H\mathbf{e}_{\text{DD}}]\} \\ &= \text{tr}[\mathcal{E}\{\mathbf{e}_{\text{DD}}(\mathbf{e}_{\text{DD}})^H\}\mathbf{F}_{N\times L}\mathbf{F}_{N\times L}^H].\end{aligned} \tag{45}$$

Substituting (35) into the above equation yields

$$\text{MSE}'_{\text{DD}} = \text{tr}\left(\frac{\sigma_I^2}{N_P P_D}\mathbf{A}\mathbf{F}_{N\times L}\mathbf{F}_{N\times L}^H\right) = \frac{\sigma_I^2}{N_P P_D}\text{tr}(\mathbf{AB}) \tag{46}$$

with

$$\mathbf{B} = \mathbf{F}_{N\times L}\mathbf{F}_{N\times L}^H \tag{47}$$

and $\mathbf{B}$ can be further decomposed as

$$\mathbf{B} = \mathbf{F}_{N\times N}\begin{pmatrix}\mathbf{I}_{L\times L} & \mathbf{0}_{(N-L)\times L} \\ \mathbf{0}_{L\times(N-L)} & \mathbf{0}_{(N-L)\times(N-L)}\end{pmatrix}\mathbf{F}_{N\times N}^H \tag{48}$$

which implies the singular value of matrix $\mathbf{B}$ is 0 or 1. In the same manner, we can obtain the MSE corresponding to $\mathbf{H}_{\text{SD}}$ as follows:

$$\text{MSE}'_{\text{SD}} = \text{tr}\left(\frac{\sigma_I^2}{N_P P_S}\mathbf{A}\mathbf{F}_{N\times L}\mathbf{F}_{N\times L}^H\right) = \frac{\sigma_I^2}{N_P P_S}\text{tr}(\mathbf{AB}). \tag{49}$$

Adding (46) and (49) forms the Sum-MSE performance of the DFT-based LS channel estimator as follows:

$$\begin{aligned}\text{SumMSE}_{\text{DFT}} &= \text{MSE}'_{\text{DD}} + \text{MSE}'_{\text{SD}} \\ &= \left(\frac{1}{N_P P_S} + \frac{1}{N_P P_D}]\right)\sigma_I^2\text{tr}(\mathbf{AB}).\end{aligned} \tag{50}$$

To evaluate the Sum-MSE performance gain achieved by the DFT-based channel estimator over the FD-LS one, let us define the following ratio:

$$\gamma = \frac{\text{SumMSE}_{\text{FD-LS}}}{\text{SumMSE}_{\text{DFT}}}. \tag{51}$$

In unit dB, the Sum-MSE performance gain achieved by the DFT-based channel estimator is $10\log_{10}\gamma$ dB.

Theorem 2: Matrices $\mathbf{A}$ and $\mathbf{B}$ are $N\times N$ positive semidefinite, where $\lambda_1(\mathbf{A}),\ldots,\lambda_N(\mathbf{A})$ denote the eigenvalues of

matrix $\mathbf{A}$, arranged in nondecreasing order, i.e., $\lambda_1(\mathbf{A}) \leq \lambda_2(\mathbf{A}) \leq \cdots \leq \lambda_N(\mathbf{A})$. If matrix $\mathbf{B}$ has the following form:

$$\mathbf{B} = \mathbf{U}_B \begin{pmatrix} \mathbf{I}_{L\times L} & \mathbf{0}_{L\times(N-L)} \\ \mathbf{0}_{(N-L)\times L} & \mathbf{0}_{(N-L)\times(N-L)} \end{pmatrix} \mathbf{U}_B^H \tag{52}$$

then, we have the following inequality:

$$\sum_{i=1}^{L} \lambda_i(\mathbf{A}) \leq \text{tr}(\mathbf{AB}) = \text{tr}\,(\mathbf{BA}) \leq \sum_{i=N-L+1}^{N} \lambda_i(\mathbf{A}). \tag{53}$$

Proof: Please see Appendix VII. ■

As a result, the Sum-MSE of the DFT-based LS will be bounded by

$$\left(\frac{1}{N_P P_S} + \frac{1}{N_P P_D}\right) \sigma_I^2 \sum_{i=1}^{L} \lambda_i(\mathbf{A}) \leq \text{SumMSE}_{\text{DFT}}$$

$$\leq \left(\frac{1}{N_P P_S} + \frac{1}{N_P P_D}\right) \sigma_I^2 \sum_{i=N-L+1}^{N} \lambda_i(\mathbf{A}). \tag{54}$$

Using the inequality (54), the inverse of the Sum-MSE performance gain γ is bounded by the following double side approximation

$$\underbrace{\frac{\sum_{i=1}^{L} \lambda_i(\mathbf{A})}{N}}_{\text{Lower bound}} \leq \gamma^{-1} \leq \underbrace{\frac{\sum_{i=N-L+1}^{N} \lambda_i(\mathbf{A})}{N}}_{\text{Upper bound}}. \tag{55}$$

Equivalently, the exact Sum-MSE performance gain γ is bounded by the following double side approximation:

$$\underbrace{\frac{N}{\sum_{i=N-L+1}^{N} \lambda_i(\mathbf{A})}}_{\text{Lower bound}} \leq \gamma \leq \underbrace{\frac{N}{\sum_{i=1}^{L} \lambda_i(\mathbf{A})}}_{\text{Upper bound}}. \tag{56}$$

Clearly, the inverse of the upper bound of γ^{-1} in (55) corresponds to the lower bound of γ in (56). Conversely, the inverse of the lower bound of γ^{-1} in (56) corresponds to the upper bound of γ in (56). When matrix $\mathbf{A}$ is singular, its smallest L eigenvalues could be zeroes, then the upper bound of γ on the right-hand side of (56) tends to infinity. Thus, in simulation section, we evaluate the bounds of γ^{-1} instead of γ.

In a practical wireless communication system, the covariance matrix $\mathbf{A}$ of the IPN vector can be estimated in advance. We will discuss how the different matrix $\mathbf{A}$ affects the MSE performance gains achieved by the DFT-based LS channel estimator in two extreme scenarios.

Scenario 1: When $\mathbf{A} = \mathbf{I}_N$, we have

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{B}) = \text{tr}\left(\mathbf{F}_{N\times L}\mathbf{F}_{N\times L}^H\right) = L \tag{57}$$

then

$$\text{SumMSE}_{\text{DFT}} = \left(\frac{1}{N_P P_S} + \frac{1}{N_P P_D}\right) L\sigma_I^2. \tag{58}$$

Making use of the definition of (51), we have

$$\gamma = \frac{N}{L} \tag{59}$$

which means that the Sum-MSE of DFT-based LS channel estimator will be reduced to L/N of that of the FD-LS one when the IPN vector is independent identically distributed (i.i.d.).

Scenario 2: Considering $\mathbf{A}$ is a matrix of all ones

$$\mathbf{A} = \mathbf{1}_N \mathbf{1}_N^H \tag{60}$$

where $\mathbf{1}_N$ is an N-D column vector of all ones. Such type of colored IPN vectors can be expressed as

$$\mathbf{w}(n,:) = g\mathbf{1}_N \tag{61}$$

where g is any random variable obeying some typical random distribution. Substituting the matrix in (60) in the trace $\text{tr}(\mathbf{AB})$, we have

$$\text{tr}(\mathbf{AB}) = \text{tr}\left(\mathbf{1}\mathbf{1}^H \mathbf{F}_{N\times L}\mathbf{F}_{N\times L}^H\right) = \text{tr}\left(\mathbf{t}\mathbf{t}^H\right) \tag{62}$$

where $\mathbf{t} = \mathbf{1}^H \mathbf{F}_{N\times L}$ is given by

$$t_l = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} W^{(l-1)(n-1)} = \begin{cases} \sqrt{N}, & l = 1[8pt] \\ 0, & 2 \leq l \leq L \end{cases}. \tag{63}$$

Substituting the above equation into (62) yields

$$\text{tr}(\mathbf{AB}) = N \tag{64}$$

then

$$\text{SumMSE}_{\text{DFT}} = \left(\frac{1}{N_P P_S} + \frac{1}{N_P P_D}\right) N\sigma_I^2. \tag{65}$$

Plugging the above expression and (39) into (51) gives

$$\gamma = 1 \tag{66}$$

which implies that there is no Sum-MSE performance gain achievable by the DFT-based LS channel estimator in the fully-correlated IPN case. In other words, the DFT-based LS channel estimator has the same Sum-MSE performance as the FD-LS one. From the above two special scenarios, we conclude that the Sum-MSE performance gains achieved by the DFT-based LS channel estimator are $10\log_{10}\frac{N}{L}$ dB, and 0 dB in the i.i.d. and fully-correlated cases, respectively. The covariance matrix of the colored interference and noise vector will have a significant influence on the Sum-MSE performance gain achieved by the DFT-based LS channel estimator. Obviously, the DFT-based LS channel estimator may harvest a larger Sum-MSE performance gain over the FD-LS one by increasing the value of N/L in the case of i.i.d. IPN vector.

V. Simulations and Discussions

In this section, we provide numerical results and analysis to evaluate the exact Sum-MSE performance gain, its upper bound and its lower bound by changing the values of correlation factor ρ and ratio N/L. To make the evaluation of Sum-MSE performance easy below, the covariance matrix $\mathbf{A}$ of IPN vector

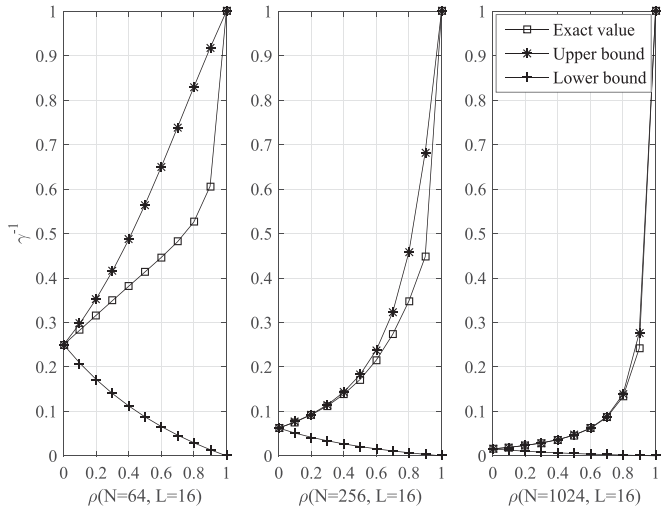


Fig. 2. γ^{-1} and its bounds in (55) versus ρ with fixed $L = 16$.

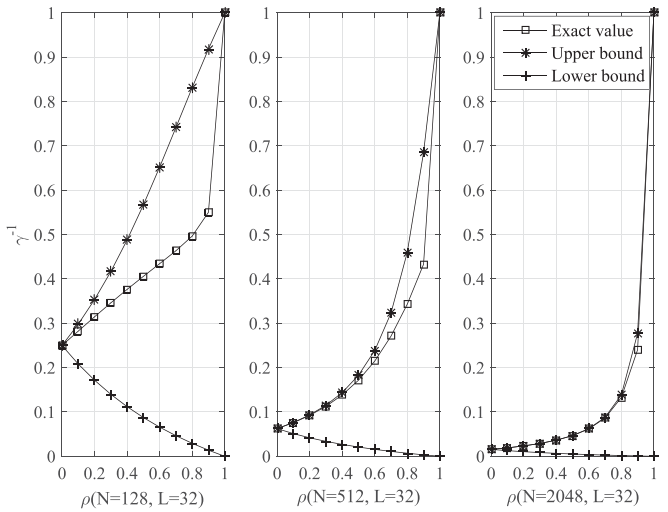


Fig. 3. γ^{-1} in terms of (51) and its bounds in (55) versus ρ with fixed $L = 32$.

is chosen as follows:

$$\mathbf{A} = \begin{pmatrix} 1 & \rho & \cdots & \rho^{N-1} \\ \rho & 1 & \cdots & \rho^{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \cdots & 1 \end{pmatrix} \tag{67}$$

where $\rho \in [0, 1]$ denotes the correlation factor between adjacent subcarrier of IPN vector. The correlation factor corresponding to two distinct subcarriers k_1 and k_2 with $\Delta k = k_2 - k_1$ is defined as

$$\rho^k = r_w(\Delta k). \tag{68}$$

For $\rho = 0$, all elements of the IPN vector are i.i.d., whereas all elements of the IPN vector are fully-correlated for $\rho = 1$. In other words, all elements of this vector obey the same random distribution as shown in (61). Changing the value of ρ will make a change in correlation degree among elements of IPN vector.

Fig. 2 demonstrates the curves of γ^{-1} and its bounds, including upper and lower bounds, versus ρ for different ratios N/L with fixed $L = 16$. Observing this figure, we can find the following results: given a fixed L, as N increases, the upper bound converges to the exact value γ^{-1}. When both L and N are fixed, the lower bound is far away from the exact value γ^{-1} with increase in the value of ρ. Consequently, we can make a conclusion that the upper bound is tighter compared to the lower bound.

Now, let us consider two extreme situations. At $\rho = 1$, the upper bound and the exact value γ^{-1} are equal to one. This means that there is no performance gain in the fully-correlated-IPN scenario. Conversely, at $\rho = 0$, i.e., the case of i.i.d. IPN vector, the Sum-MSE gain γ is equal to N/L. This is also the maximum performance gain achievable by the DFT-based LS channel estimator over the FD-LS one. In particular, from Fig. 2, we also find that decreasing the value of ρ will increase the performance gain. Additionally, at two extreme-value points

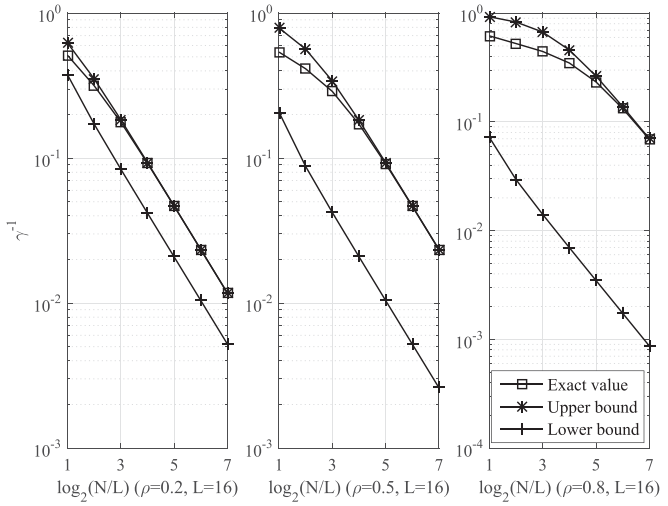


Fig. 4. γ^{-1} and its bounds in (55) versus $\log_2(N/L)$ with fixed $L = 16$.

$(0,\ L/N)$ and $(1,\ 1)$, the derived upper bound is in agreement with the exact Sum-MSE value for all three parts of Fig. 2.

Fig. 3 illustrates the curves of γ^{-1} and its bounds versus ρ for different ratios N/L with fixed $L = 32$. It is evident that Fig. 3 has the same tendency as Fig. 2.

Fig. 4 plots the curves of γ^{-1} and its bounds versus different ratios $\log_2(N/L)$ for different ρ and fixed $L = 16$. It intuitively follows from this figure that given L and ρ, as N increases, γ^{-1} will get closer and closer to its upper bound and even begins to overlap for a sufficiently large N/L. Specifically, for a smaller value of ρ, the exact value of γ^{-1} and its upper bound overlaps with each other starting from a relatively smaller value of N. For example, we can find that the curves of γ^{-1} and its upper bound begin to overlap at the value of N/L being 8, 16, and 64 for ρ increases from 0.2, 0.5 to 0.8, respectively. Consequently, we may claim that the upper bound is a good approximation to the exact value of γ^{-1} for almost all situations compared with lower bound.

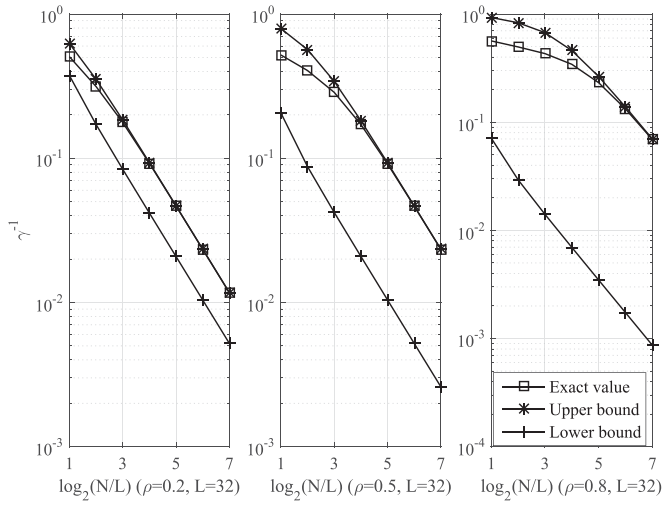


Fig. 5. γ^{-1} and its bounds in (55) versus $\log_2(N/L)$ with fixed $L = 32$.

Fig. 5 shows the curves of γ^{-1} and its bounds versus different ratios $\log_2(N/L)$ for different values of correlation factor ρ with fixed $L = 32$. Observing Fig. 5, it is very evident that similar to Fig. 4, we obtain the same results and trends.

In summary, the values of N/L and ρ have a dramatic impact on the Sum-MSE performance gain. Increasing the value of N/L improves the Sum-MSE performance gain, whereas reducing the value results in a loss on the Sum-MSE performance gain.

VI. Conclusion

In this paper, we investigate and analyze the Sum-MSE performance gain of DFT-based LS channel estimator over FD-LS one in the FD OFDM system with colored interference and noise. The exact value, upper bound, and lower bound of the Sum-MSE performance gain are derived, discussed, and verified. From numerical simulations and analysis, it follows that the lower bound is closer to the exact Sum-MSE performance gain compared to the upper bound. In two extreme scenarios: fully-correlated and i.i.d., the Sum-MSE performance gains are shown to be 1, and N/L, respectively. The correlation factor ρ of IPN covariance matrix has an obvious impact on the Sum-MSE performance gain. Roughly speaking, a large Sum-MSE performance gain increases from 1 to N/L as the correlation factor ρ decreases from 1 to 0. The above results can be applied to provide a guidance for pilot pattern optimization of channel estimation design in future FD wireless networks such as mobile communications, satellite communications, cooperative communications, V2V, unmanned-aerial-vehicles networks, internet of things (IoT), etc.

Appendix A
Proof of Theorem 2

Substituting (52) in $\mathrm{tr}(\mathbf{AB})$ yields

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}\left(\mathbf{U}_B^H \mathbf{A}\mathbf{U}_B \Lambda_B\right) \tag{69}$$

where

$$\mathbf{\Lambda_B} = \begin{pmatrix} \mathbf{I}_{L\times L} & \mathbf{0}_{L\times(N-L)} \\ \mathbf{0}_{(N-L)\times L} & \mathbf{0}_{(N-L)\times(N-L)} \end{pmatrix}. \tag{70}$$

Let us define

$$\tilde{\mathbf{A}} = \mathbf{U}_B^H \mathbf{A}\mathbf{U}_B \tag{71}$$

where $\tilde{\mathbf{A}}$ has the same set of eigenvalues as $\mathbf{A}$ due to the property of unitary transformation. Furthermore,

$$\mathrm{tr}\,(\mathbf{A}) = \mathrm{tr}(\tilde{\mathbf{A}}). \tag{72}$$

The identity (69) is rewritten as

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\tilde{\mathbf{A}}\Lambda_B). \tag{73}$$

Similar to (70), $\tilde{\mathbf{A}}$ is represented in the block matrix

$$\tilde{\mathbf{A}} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{21} \\ \tilde{\mathbf{A}}_{12} & \tilde{\mathbf{A}}_{22} \end{pmatrix}. \tag{74}$$

Using the above expression

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\tilde{\mathbf{A}}_{11}). \tag{75}$$

Making use of [38, Th. 4.3.28], we have

$$\begin{aligned} \lambda_1(\tilde{\mathbf{A}}) \le \lambda_1(\tilde{\mathbf{A}}_{11}) \le \lambda_{N-L+1}(\tilde{\mathbf{A}}) \\ \lambda_2(\tilde{\mathbf{A}}) \le \lambda_2(\tilde{\mathbf{A}}_{11}) \le \lambda_{N-L+2}(\tilde{\mathbf{A}}) \\ \vdots \\ \lambda_L(\tilde{\mathbf{A}}) \le \lambda_L(\tilde{\mathbf{A}}_{11}) \le \lambda_N(\tilde{\mathbf{A}}). \end{aligned} \tag{76}$$

Adding all the above L inequalities results in the fact that the trace in (73) is bounded by

$$\sum_{i=1}^{L} \lambda_i(\mathbf{A}) \le \sum_{i=1}^{L} \lambda_i(\tilde{\mathbf{A}}_{11}) = \mathrm{tr}(\tilde{\mathbf{A}}_{11}) \le \sum_{i=N-L+1}^{N} \lambda_i(\mathbf{A}). \tag{77}$$

■

References

[1] Z. Zhang, X. Chai, K. Long, and A. V. Vasilakos, "Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 128–137, May 2015.

[2] 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.

[3] C. Qi, L. Wu, Y. Huang, and A. Nallanathan, "Joint design of pilot power and pilot pattern for sparse cognitive radio systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5384–5390, Nov. 2015.

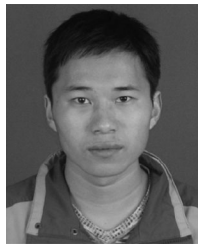
[4] X. Sun, K. Xu, W. Ma, and Y. Xu, "Multi-pair two-way massive MIMO AF full-duplex relaying with imperfect CSI over Ricean fading channels," *IEEE Access*, vol. 4, pp. 4933–4945, Jul. 2016.

[5] A. Aijaz and P. Kulkarni, "Protocol design for enabling full-duplex operation in next-generation IEEE 802.11 WLANs," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3438–3449, Dec. 2018.

[6] H. Chen, G. Li, and J. Cai, "Spectral-energy efficiency tradeoff in full-duplex two-way relay networks," *IEEE Syst. J.*, vol. 12, no. 1, pp. 583–592, Mar. 2018.

[7] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks?," in *Proc. IEEE Conf. Comput. Commun*, Apr. 2014, pp. 253–261.

[8] J. Wang *et al.*, "Pilot optimization and power allocation for OFDM-based full-duplex relay networks with IQ-imbalances," *IEEE Access*, vol. 5, pp. 24344–24352, Oct. 2017.
[9] 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.
[10] 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.
[11] 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.
[12] A. Sabharwal, P. Schniter, D. Guo, and D. W. Bliss, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.
[13] 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.
[14] 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.
[15] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702–3713, Jul. 2012.
[16] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.
[17] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "Analysis of DFT-based channel estimators for OFDM," *Wireless Pers. Commun.*, vol. 12, no. 1, pp. 55–70, Jan. 2000.
[18] S. A. Mohajeran and S. M. S. Sadough, "On the interaction between joint Tx/Rx IQI and channel estimation errors in DVB-T systems," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3271–3278, Dec. 2018.
[19] S. M. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice Hall, 1993.
[20] D. Schafhuber, G. Matz, and F. Hlawatsch, "Kalman tracking of time-varying channels in wireless MIMO-OFDM systems," in *Proc. 37th Asilomar Conf. Signals Syst. Comput.*, vol. 2, May. 2004, pp. 1261–1265.
[21] Y. Liang, H. Luo, and J. Huang, "Adaptive RLS channel estimation in MIMO-OFDM systems," in *Proc. IEEE Int. Symp. Commun. Inf. Technol.*, Oct. 2005, pp. 79–82.
[22] S. Barembruch, A. Garivier, and E. Moulines, "On approximate maximum-likelihood methods for blind identification: How to cope with the curse of dimensionality," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4247–4259, Nov. 2009.
[23] M. Imani and U. M. Braga-Neto, "Particle filters for partially-observed Boolean dynamical systems," *Automatica*, vol. 87, pp. 238–250, Jan. 2018.
[24] R. Douc, A. Garivier, E. Moulines, and J. Olsson, "Sequential Monte Carlo smoothing for general state space hidden Markov models," *Ann. Appl. Probab.*, vol. 21, no. 6, pp. 2109–2145, Dec. 2011.
[25] M. Imani and U. M. Braga-Neto, "Maximum-likelihood adaptive filter for partially observed Boolean dynamical systems," *IEEE Trans. Signal Process.*, vol. 65, no. 2, pp. 359–371, Jan. 2017.
[26] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.
[27] F. Shu, J. Lee, L. N. Wu, and G. L. Zhao, "Time-frequency channel estimation for digital amplitude modulation broadcasting systems based on OFDM," *IEE Proc.-Commun.*, vol. 150, no. 4, pp. 259–264, Aug. 2003.
[28] Y. Kang, K. Kim, and H. Park, "Efficient DFT-based channel estimation for OFDM systems on multipath channels," *IET Commun.*, vol. 1, no. 2, pp. 197–202, Apr. 2007.
[29] Y. Liu, X. Quan, W. Pan, and Y. Tang, "Digitally assisted analog interference cancellation for in-band full-duplex radios," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1079–1082, May 2017.
[30] 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.
[31] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator in MIMO full-duplex systems," in *Proc. IEEE 80th Veh. Technol. Conf. (Fall)*, Sep. 2014, pp. 1–5.
[32] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator for self-interference cancellation in full-duplex systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5122–5132, Jul. 2016.
[33] H. Yu *et al.*, "Compressed sensing-based time-domain channel estimator for full-duplex OFDM systems with IQ-imbalances," *Sci. China Inform. Sci.*, vol. 60, no. 8, p. 082303, Aug. 2017.
[34] F. Shu, J. Wang, J. Li, R. Chen, and W. Chen, "Pilot optimization, channel estimation and optimal detection for full-duplex OFDM systems with IQ-imbalances," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6993–7009, Aug. 2017.
[35] 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 Inform. Sci.*, vol. 53, no. 11, pp. 2604–2610, Nov. 2012.
[36] H. B. Mishra and K. Vasudevan, "Design of superimposed training sequence for spatially correlated multiple-input-correlated-multiple-output channels under interference-limited environments," *IET Commun.*, vol. 9, no. 10, pp. 1259–1268, Jun. 2015.
[37] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.
[38] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2013.



Jin Wang received the B.S. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2012, and is currently working toward the Ph.D. degree from the School of Electronic and Optical Engineering, Nanjing University of Science and Technology.

His research interests include wireless communications and signal processing.

Hai Yu received the B.S. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2000, and is currently working toward the Ph.D. degree from the School of Electronic and Optical Engineering, Nanjing University of Science and Technology.

His research interests include wireless communications and signal processing.

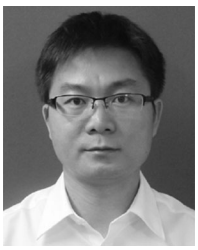
Feng Shu (M'16) was born in 1973. He received the B.S. degree from Fuyang Teaching College, Fuyang, China, in 1994, the M.S. degree from XiDian University, Xian, China, in 1997, and the Ph.D. degree from the Southeast University, Nanjing, in 2002.

From October 2003 to 2005, he was a Postdoctoral Researcher with the National Key Mobile Communication Lab, the Southeast University. From September 2009 to 2010, he was a Visiting Postdoctor with the University of Texas at Dallas, Richardson, TX, USA. In October 2005, he joined the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China, where he is currently a Professor and Supervisor of Ph.D. and graduate students. He is also with Fujian Agriculture and Forestry University, Fuzhou, China. He has authored or coauthored about 200 scientific and conference papers, of which more than 100 are in archival journals including more than 20 papers on IEEE Journals and 43 SCI-indexed papers. He holds four Chinese patents. His research interests include wireless networks, wireless location, and array signal processing.

Dr. Shu was the recipient of the Mingjian Scholar Chair Professor in Fujian Province. He is currently an Editor for IEEE ACCESS. He was a Session Chair or Technical Program Committee Member for various international conferences, such as IEEE WCSP 2016, IEEE VTC 2016, etc.

Jinhui Lu received the M.S. degree in communication and information system from the Nanjing University of Science and Technology, Nanjing, China, in 1985.

He is currently a Professor in the School of Electronic and Optical Engineering, Nanjing University of Science and Technology. He is a Senior Member of the Chinese Institute of Electronics, Beijing, China and a Director of China Education Society of Electronics. His research interests including Millimeter wave monitoring system, echo simulation of radar and passive location technology.



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 University of Oxford, Oxford, U.K., in 2010.

He is currently a Lecturer with the Faculty of Computer Science and Information Technology, Fujian Agriculture and Forestry University, Fuzhou, China. His research interests include network security, spread coding, signal processing, and wireless sensor technologies.

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.

Dushantha Nalin K. Jayakody (M'14) received the B.S. degree from the Dawood University of Engineering and Technology, Pakistan and was ranked as the merit position holder of the University (under SAARC Scholarship), the M.Sc. degree in electronics and communications engineering from the Department of Electrical and Electronics Engineering, Eastern Mediterranean University, Famagusta, Cyprus, in 2010 (under the University full graduate scholarship) and ranked as the first merit position holder of the department, and the Ph. D. degree in electronics and communications engineering, in 2014, from the University College Dublin, Dublin, Ireland.

From September 2014 to 2016, he was a Postdoc with the Coding Information Transmission group, University of Tartu, Estonia and University of Bergen, Norway. Since 2016, he has been a Professor with the Department of Control System Optimization, Institute of Cybernetics, National Research Tomsk Polytechnic University, Tomsk, Russia.

Dr. Jayakody was a Session Chair or Technical Program Committee Member for various international conferences, such as IEEE PIMRC 2013/2014, IEEE WCNC 2014/2016, and IEEE VTC 2015.