Iterative Channel Estimation Using Virtual Pilot Signals for MIMO-OFDM Systems

Sunho Park, *Student Member, IEEE*, Byonghyo Shim, *Senior Member, IEEE*, and Jun Won Choi, *Member, IEEE*

***Abstract*—The number of transmit and receive antennas in multi-input multi-output (MIMO) systems is increasing rapidly to enhance the throughput and reliability of next-generation wireless systems. Benefits of large size MIMO systems, however, can be realized only when the quality of estimated channels is ensured at the transmitter and receiver side alike. In this paper, we introduce a new decision-directed channel estimation technique dealing with pilot shortage in the MIMO-OFDM systems. The proposed channel estimator uses soft symbol decisions obtained by iterative detection and decoding (IDD) scheme to enhance the quality of channel estimate. Using the soft information from the decoders, the proposed channel estimator selects reliable data tones, subtracts interstream interferences, and performs re-estimation of the channels. We analyze the optimal data tone selection criterion, which accounts for the reliability of symbol decisions and correlation of channels between the data tones and pilot tones. From numerical simulations, we show that the proposed channel estimator achieves considerable improvement in system performance over the conventional channel estimators in realistic MIMO-OFDM scenarios.**

***Index Terms*—Channel estimation, decision directed channel estimation, iterative detection and decoding, joint channel estimation and detection, multi-input multi-output (MIMO), orthogonal frequency division multiplexing (OFDM).**

I. Introduction

MULTI-INPUT multi-output orthogonal frequency division multiplexing (MIMO-OFDM) is a technology widely used in various commercial systems including 3GPP long term evolution (LTE), and IEEE wireless LAN standard. Recently, MIMO-OFDM systems with large number of antennas, referred to as massive MIMO, are of great interest to further improve throughput and reliability of next generation wireless systems. One of major issues in realizing the massive MIMO-OFDM systems is that the amount of pilot signals needed for channel estimation is proportional to the number of the transmit antennas. When the number of transmit antennas is large, pilot signals occupy significant portion of downlink resources, eating out the data throughput significantly. One can consider the reduction of pilot signal density but this is undesirable since it will simply cause the degradation of channel estimation quality, affecting link performance and data throughput eventually.

When the pilot resources are depleted, one can naturally consider the option of using the data signals for the channel estimation purpose. This approach, often referred to as *decision-directed channel estimation* (DD-CE), uses the decided (sliced) data symbols in the re-estimation of channels. There have been a number of approaches related to DD-CE techniques [1]–[11]. These include maximum a posteriori (MAP) based channel estimation [1], expectation maximization (EM)-based joint detection and channel estimation [2]–[5], Kalman filter based soft decision channel estimator [6]–[8], and channel estimation with interference suppression in MIMO channels[9]–[11]. Two main concerns of the DD-CE for the MIMO-OFDM systems, not thoroughly addressed in the previous efforts, are the reliability of the detected data symbols and the interstream interference caused by the MIMO transmission. First, the quality of the estimated channel would not be appealing if the soft statistics of the data tones used for channel re-estimation are not reliable. Second, unlike the pilot symbols, the data symbols in the MIMO-OFDM systems are transmitted simultaneously through multiple transmit antennas so that channel estimation is interfered by the data symbols coming from other transmit antennas. Clearly, proper control of these interstream interferences is crucial for effective decision-directed channel estimation.

The primary goal of this paper is to demonstrate that the selective use of data signals, in conjunction with the state of the art *iterative detection and decoding* (IDD) technique [13], is effective means to improve channel estimation quality of the MIMO-OFDM systems. The proposed scheme selects the reliable data tones (henceforth referred to *virtual pilots*) among all possible data tones, purifies the chosen virtual pilots via interstream interference cancellation, and then performs soft-decision-directed channel estimation. By employing deliberately chosen virtual pilot signals as well as the pilot signals, channel estimation quality and subsequent detection and decoding quality can be improved substantially. Note that when the observation of (reliable) data tones is added to the observation of the pilot tone (see Fig. 1), the channel estimator

Manuscript received July 18, 2014; revised January 15, 2015; accepted March 09, 2015. Date of publication March 25, 2015; date of current version May 12, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Walaa Hamouda. This work was supported by the Communications Research Team, DMC R&D Center, Samsung Electronics Co. Ltd., the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (no. 2014049151, 2014R1A5A1011478 and 2014R1A1A2055805), and the research grant from Hanyang University (HY-2013). This paper was presented in part at the Information Theory and Application Workshop (ITA), San Diego, CA, USA, February 2014.

S. Park is with School of Information and Communication, Korea University, Seoul, Korea, and also with the Department of Electrical and Computer Engineering, Seoul National University, Seoul 136-701, Korea (e-mail: shpark@islab.snu.ac.kr).

B. Shim is with the Institute of New Media and Communications and School of Electrical and Computer Engineering, Seoul National University, Seoul 136-701, Korea (e-mail: bshim@snu.ac.kr).

J. W. Choi is with the Department of Electrical Engineering, Hanyang University, Seoul 133-791, Korea (e-mail: junwchoi@hanyang.ac.kr).

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/TSP.2015.2416684

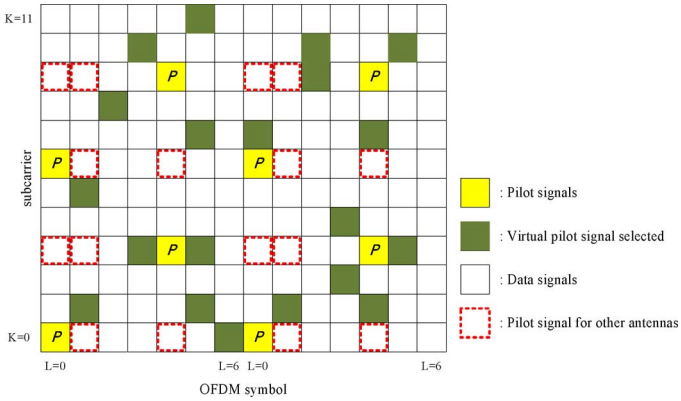


Fig. 1. Illustration of virtual pilot signal selection and soft decision-directed channel estimation.

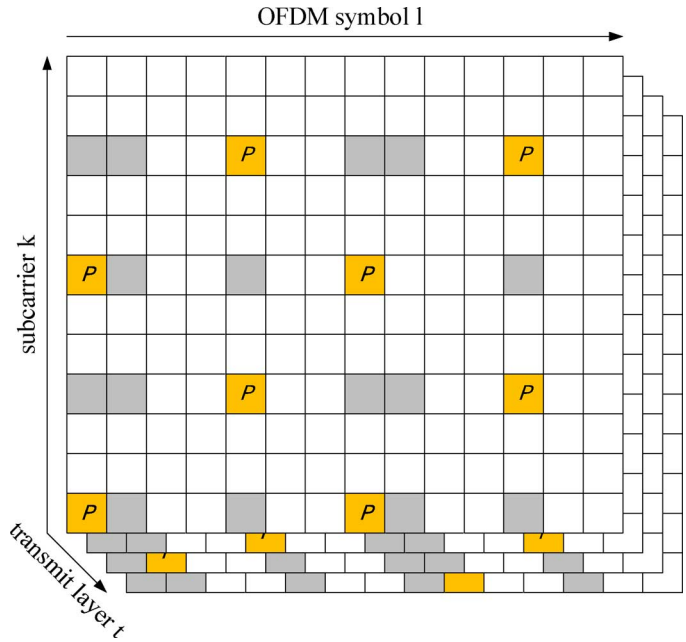


Fig. 2. Subcarrier, OFDM symbol, and transmit layer indices over three-dimensional resource grid.

can capture time-frequency variations of channels better and also filter out the interference and noise more effectively. In our analysis, we show that data tones maximizing both the *reliability of soft decisions* (derived from the *a posteriori* log-likelihood ratio of data tone) and the *correlations of channel gains between pilot tones and data tones* improves the channel estimation quality substantially. We also show from numerical simulations in realistic MIMO-OFDM scenarios that the proposed method outperforms the conventional channel estimation methods in terms of the MSE and link level BER performance.

The remainder of this paper is organized as follows. In Section II, we provide the brief summary of the MIMO-OFDM systems and also review the conventional channel estimation schemes. In Section III, we present the proposed virtual pilot based channel estimation algorithm and the analysis of the virtual pilot selection criterion. In Section IV, we present the simulation results and conclude the paper in Section V.

We briefly summarize notations used in this paper. Upper and lower case letters written in boldface denote matrices and vectors, respectively. Superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and conjugate transpose (hermitian operator), respectively. $\|\cdot\|^2$ indicates an L_2-norm of a vector. diag $(\cdot)$ is a diagonal matrix that has elements on the main diagonal. tr $(\cdot)$ is matrix trace that the sum of the elements on diagonal. $\mathcal{CN}(m, \sigma^2)$ denotes a circular symmetric complex Gaussian density with mean m and variance σ^2. $E_x[\cdot]$ denotes expectation over the random variable x. $\mathbf{A} \odot \mathbf{B}$ denotes the Hadamard (element-by-element) product between the matrices $\mathbf{A}$ and $\mathbf{B}$. $\mathrm{Cov}(\mathbf{A}, \mathbf{B}) = E[(\mathbf{A} - E[\mathbf{A}])(\mathbf{B} - E[\mathbf{B}])^H]$ denotes the covariance matrix between matrices $\mathbf{A}$ and $\mathbf{B}$.

II. System Description

A. MIMO-OFDM Systems

In this section, we review the MIMO-OFDM systems briefly. Our system model is based on the LTE-Advanced cellular standard [12] but the model can be readily applied to any MIMO-OFDM systems with minor modifications. The information bits $\{b_i\}$ are encoded by a rate C channel encoder, producing the coded bit sequence $\{c_i\}$. This sequence $\{c_i\}$ is permuted via a random interleaver and then interleaved bits are mapped to a finite alphabet symbol $\{d_i\}$ in 2^Q-ary quadrature amplitude modulation (QAM) constellations. Pilot sequence is generated via quadrature phase shift keying (QPSK) modulation. Pilot and data symbols are allocated over three-dimensional resource grid indexed by (k, l, t), where k, l, and t correspond to subcarrier (frequency), OFDM symbol (time), and transmit layer (space) indices, respectively (see Fig. 2). Each code block occupies K subcarriers, L OFDM symbols, and T transmit layers. The pilot signals are allocated to the predetermined positions (gray-colored parts in Fig. 2) and data symbols fill the rest.

The symbols generated for one OFDM symbol index l are transformed to the time-domain via inverse discrete Fourier transform (IDFT). After an addition of the cyclic prefix (CP), the time-domain OFDM signal is modulated and transmitted through the transmit antennas. The transmit signals are propagated through frequency selective MIMO channels, whose power and delay profiles are specified by the vectors, $[\rho_0, \ldots, \rho_{\Lambda-1}]$ and $[\tau_0, \ldots, \tau_{\Lambda-1}]$, respectively ($\Lambda$ is the number of distinct paths). Received signals, after the removal of the CP, are transformed to the frequency domain OFDM symbols via the DFT operation. When the channel delay spread is less than the CP length, the transmit data at each subcarrier experiences MIMO flat fading channel so that the received signal for a resource element (a small rectangle in Fig. 2) is modeled as

$$\begin{bmatrix} y_{k,l}^{(0)} \\ \vdots \\ y_{k,l}^{(R-1)} \end{bmatrix} = \begin{bmatrix} h_{k,l}^{(0,0)} & \cdots & h_{k,l}^{(0,T-1)} \\ \vdots & \ddots & \vdots \\ h_{k,l}^{(R-1,0)} & \cdots & h_{k,l}^{(R-1,T-1)} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{\eta_{d,0}} d_{k,l}^{(0)} \\ \vdots \\ \sqrt{\eta_{d,T-1}} d_{k,l}^{(T-1)} \end{bmatrix} + \begin{bmatrix} n_{k,l}^{(0)} \\ \vdots \\ n_{k,l}^{(R-1)} \end{bmatrix}, \quad (1)$$

where T and R are the number of transmit and receive antennas, $y_{k,l}^{(r)}$ and $d_{k,l}^{(t)}$ are the signal received at the r-th receive antenna

and the symbol transmitted from the t-th transmit antenna, respectively, $h_{k,l}^{(r,t)}$ is the channel response from the t-th transmit to the r-th receive antennas, $\eta_{d,t}$ is the transmit signal power from the t-th transmit antenna, and $n_{k,l}^{(r)}$ is the complex Gaussian noise at the r-th receive antenna (i.e., $n_{k,l}^{(r)} \sim \mathcal{CN}(0,1)$). Note that subscript indices k and l indicate the subcarrier and OFDM symbol indices, respectively.

B. Iterative Detection and Decoding

As an effective decoding scheme for coded MIMO systems, IDD techniques have received much attention in recent years. In a nutshell, the IDD improves the quality of the MIMO detector and the channel decoder by exchanging the soft information in a form of log likelihood ratio (LLR). In the MIMO detector, *a posteriori* LLRs are generated using the observation and *a priori* LLRs delivered from the channel decoder. The *a posteriori* LLR, which is the sum of the *extrinsic* and *a priori* LLRs, is expressed as [13]–[15]

$$L(c_i) = \ln \frac{Pr(c_{i+}|\mathbf{y}_{k,l})}{Pr(c_{i-}|\mathbf{y}_{k,l})} \tag{2}$$

$$= \underbrace{\ln \frac{Pr(c_{i+})}{Pr(c_{i-})}}_{\text{a priori}} + \underbrace{\ln \frac{\sum_{\mathbf{c}_{i+}} Pr(\mathbf{y}_{k,l}|\mathbf{d}_{k,l,i+}) Pr(\mathbf{c}_{\bar{i}})}{\sum_{\mathbf{c}_{i-}} Pr(\mathbf{y}_{k,l}|\mathbf{d}_{k,l,i-}) Pr(\mathbf{c}_{\bar{i}})}}_{\text{extrinsic}}, \tag{3}$$

where $Pr(c_{i+})$ and $Pr(c_{i-})$ are the probabilities that $c_i = +1$ and $c_i = -1$, $\mathbf{c}_{i+} = [c_1, \ldots, c_{i-1},\ c_{i+}, c_{i+1}, \ldots, c_{QT}]$, $\mathbf{c}_{\bar{i}} = [c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_{QT}]$, $\mathbf{y}_{k,l} = [y_{k,l}^{(0)}, \ldots, y_{k,l}^{(R-1)}]$ where $\mathbf{d}_{k,l,i+} = f(\mathbf{c}_{i+})$, and $f(\cdot)$ is the function mapping the encoded bit vector into the symbol vector ($\mathbf{c}_{i-}$ and $\mathbf{d}_{k,l,-}$ are defined in the same way), respectively. In particular, the extrinsic LLR, the second term in the right-hand side of (3), is expressed as [13]

$$L_E(c_i) = \ln \frac{\sum_{\mathbf{d}_{k,l,i+}} \exp\left(-\frac{\|\mathbf{y}_{k,l} - \mathbf{H}_{k,l}\mathbf{d}_{k,l,i+}\|^2}{2} + \frac{1}{2}\mathbf{c}_{\bar{i}}^T \mathbf{L}_{A,\bar{i}}\right)}{\sum_{\mathbf{d}_{k,l,i-}} \exp\left(-\frac{\|\mathbf{y}_{k,l} - \mathbf{H}_{k,l}\mathbf{d}_{k,l,i-}\|^2}{2} + \frac{1}{2}\mathbf{c}_{\bar{i}}^T \mathbf{L}_{A,\bar{i}}\right)}, \tag{4}$$

where the channel matrix $\mathbf{H}_{k,l}$ is expressed as

$$\mathbf{H}_{k,l} = \begin{bmatrix} \sqrt{\eta_{d,0}} h_{k,l}^{(0,0)} & \cdots & \sqrt{\eta_{d,T-1}} h_{k,l}^{(0,T-1)} \\ \vdots & \ddots & \vdots \\ \sqrt{\eta_{d,0}} h_{k,l}^{(R-1,0)} & \cdots & \sqrt{\eta_{d,T-1}} h_{k,l}^{(R-1,T-1)} \end{bmatrix} \tag{5}$$

and $\mathbf{L}_{A,\bar{i}} = [L_A(c_1), \ldots, L_A(c_{i-1}), L_A(c_{i+1}), \ldots, L_A(c_{QT})]^T$ where $L_A(c_i) = \ln \frac{Pr(c_{i+})}{Pr(c_{i-})}$ is *a priori* LLR, respectively. After the deinterleaving, these extrinsic LLRs are used as *a priori* LLRs for the channel decoder. Using this information, the channel decoder generates the extrinsic LLRs. These extrinsic LLRs are interleaved and then sent back to the MIMO detector. For each code block, these operations are repeated until a suitably chosen termination condition is satisfied. From (4) and (5), it is clear that an accurate estimation of channel is crucial to ensure the reliable performance of the IDD systems.

C. Channel Estimation for MIMO-OFDM Systems

In the conventional MIMO-OFDM systems, the channel is estimated using the observations of the pilot signals (in the sequel we denote it by pilot observations) in a local window. The pilot observations corresponding to the t-th transmit antenna can be expressed as

$$z_0^{(r)} = \sqrt{\eta_{p,t}} h_0^{(r,t)} p_0^{(t)} + n_0^{(r)} \tag{6}$$

$$\vdots$$

$$z_{N_p-1}^{(r)} = \sqrt{\eta_{p,t}} h_{N_p-1}^{(r,t)} p_{N_p-1}^{(t)} + n_{N_p-1}^{(r)}, \tag{7}$$

where $z_j^{(r)}$ and $n_j^{(r)}$ are j-th observation and noise received at r-th receiver antenna, $p_j^{(t)}$ is the pilot signal in t-th transmit antenna, N_p is the total number of pilot observations in the window, and $\eta_{p,t}$ is the power allocated to the pilot signals at t-th transmit antenna. A vector form of the pilot observations is

$$\mathbf{z}_r = \sqrt{\eta_{p,t}} \mathbf{P}_t \mathbf{h}_{r,t} + \mathbf{n}_r, \tag{8}$$

where $\mathbf{z}_r = [z_0^{(r)}, \ldots, z_{N_p-1}^{(r)}]^T$, $\mathbf{h}_{r,t} = [h_0^{(r,t)}, \ldots, h_{N_p-1}^{(r,t)}]^T$, $\mathbf{n}_r = [n_0^{(r)}, \ldots, n_{N_p-1}^{(r)}]^T$, and $\mathbf{P}_t = \text{diag}([p_0^{(t)}, \ldots, p_{N_p-1}^{(t)}])$.

The MMSE estimate $\hat{\mathbf{h}}_{r,t}$ of $\mathbf{h}_{r,t}$ is

$$\hat{\mathbf{h}}_{r,t} = \text{Cov}(\mathbf{h}_{r,t}, \mathbf{z}_r)\, \text{Cov}(\mathbf{z}_r, \mathbf{z}_r)^{-1} \mathbf{z}_r \tag{9}$$

$$= \left(\sqrt{\eta_{p,t}} C_{h_t,h_t} \mathbf{P}_t^H\right) \left(\eta_{p,t} \mathbf{P}_t C_{h_t,h_t} \mathbf{P}_t^H + \mathbf{I}\right)^{-1} \mathbf{z}_r \tag{10}$$

$$= \left(\sqrt{\eta_{p,t}} C_{h_t,h_t} \mathbf{P}_t^H\right) \left(\eta_{p,t} \mathbf{P}_t C_{h_t,h_t} \mathbf{P}_t^H + \mathbf{P}_t \mathbf{P}_t^H\right)^{-1} \mathbf{z}_r \tag{11}$$

$$= \underbrace{\frac{1}{\sqrt{\eta_{p,t}}} C_{h_t,h_t} \left(C_{h_t,h_t} + \frac{1}{\eta_{p,t}} \mathbf{I}\right)^{-1}}_{\mathbf{w}_{r,t}^H} \mathbf{P}_t^H \mathbf{z}_r, \tag{12}$$

where $C_{h_i,h_j} = \text{Cov}(\mathbf{h}_i, \mathbf{h}_j)$ and (11) is from the fact that the pilot signals $p_i^{(t)}$ are chosen from QPSK modulation ($\mathbf{P}_t \mathbf{P}_t^H = \mathbf{P}_t^H \mathbf{P}_t = \mathbf{I}$). From (12), we observe that the MMSE channel estimation consists of two steps: 1) descrambling of the received signal using the pilot sequence (i.e., $\mathbf{P}_t^H \mathbf{z}_r$) and 2) application of the filter $\mathbf{w}_{r,t}^H$ to the descrambled signal. The covariance matrix C_{h_i,h_j}, required for the generation of $\mathbf{w}_{r,t}^H$, is computed parametrically by considering the relative location among pilot signals in frequency and time (symbol) domain. In the well-known Jakes' model, for example, the correlation of the channel gains spaced by Δk subcarriers and Δl OFDM symbol times is given by [16]

$$E\left[h_{k,l}^{(r,t)} h_{k+\Delta k,l+\Delta l}^{(r,t)}\right] = \left(\sum_{i=0}^{\Lambda-1} \rho_i e^{j2\pi \Delta k f_s \tau_i}\right) J_0(2\pi f_d T_s \Delta t), \tag{13}$$

where f_s and f_d are the spacing between adjacent subcarriers and Doppler frequency, respectively, T_s is the period for one OFDM symbol, and $J_0(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(m+1)} \left(\frac{x}{2}\right)^{2m}$ is the zero-th order Bessel function. Once the channels $\hat{\mathbf{h}}_{r,t}$ corresponding to the pilot signals are estimated, the channels for the

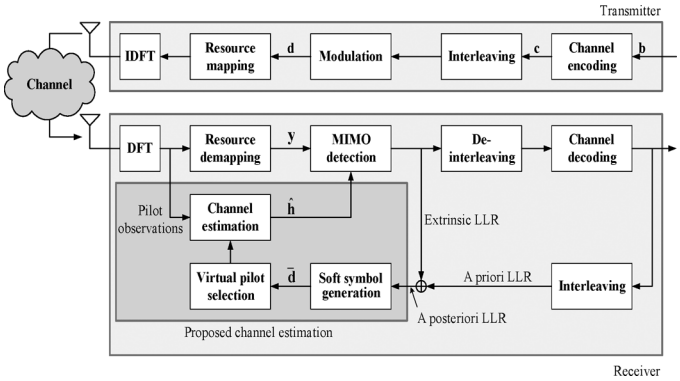


Fig. 3. Block diagram of the proposed channel estimation scheme using virtual pilot signals.

data tones are generated via proper interpolation among pilot channels.

III. Channel Estimation Using Virtual Pilot Signal

In this section, we discuss the channel estimation algorithm that exploits the *virtual pilot signals* as well as pilot signals in the re-estimation of channels. Fig. 3 depicts the block diagram of the proposed scheme. Firstly, *a posteriori* LLRs, obtained by adding *an extrinsic* LLRs of the MIMO detector and *a prior* LLRs of the decoder, is converted to the soft symbols $\bar{\mathbf{d}}$. Next, among data tones available in the window, virtual pilot signals are chosen. Virtual pilot signals used for the channel estimation should meet two conditions. First, the quality of virtual pilot signals should be good enough to be used for channel estimation. To meet this condition, clearly, the magnitude of *a posteriori* LLRs should be large. Second, the channels for the virtual pilot signals should be highly correlated with those for the pilot signals since otherwise virtual pilot signals would not be helpful in improving the quality of channel estimates. Using the selected virtual pilot signals together with the original pilot signals, channel is re-estimated and the newly generated channel estimate is used for the MIMO detection in the next iteration. The proposed algorithm is repeated until suitably chosen termination condition is satisfied (see Section IV.A).

A. Channel Re-Estimation Using Virtual Pilot Signal

Let N_d be the number of the virtual pilot signals used for the channel estimation, then the virtual pilot observations are expressed as

$$y_0^{(r)} = \sum_{i=0}^{T-1} \sqrt{\eta_{d,i}} g_0^{(r,i)} d_0^{(i)} + v_0^{(r)} \tag{14}$$

$$\vdots$$

$$y_{N_d-1}^{(r)} = \sum_{i=0}^{T-1} \sqrt{\eta_{d,i}} g_{N_d-1}^{(r,i)} d_{N_d-1}^{(i)} + v_{N_d-1}^{(r)} \tag{15}$$

where $y_j^{(r)}$ and $v_j^{(r)}$ are the observation and noise received at the r-th received antenna, $g_j^{(r,i)}$ is the channel gain from the i-th transmit antenna to the r-th receive antenna, $d_j^{(i)}$ is the data signal from the i-th transmit antenna, and $\eta_{d,i}$ is the transmitted power of the data symbols at i-th transmit antenna. A vector form of the virtual pilot observations is

$$\mathbf{y}_r = \sum_{i=0}^{T-1} \sqrt{\eta_{d,i}} \mathbf{D}_i \mathbf{g}_{r,i} + \mathbf{v}_r, \tag{16}$$

where $\mathbf{y}_r = [y_0^{(r)}, \ldots, y_{N_d-1}^{(r)}]^T$, $\mathbf{D}_i = \mathrm{diag}\left([d_0^{(i)}, \ldots, d_{N_d-1}^{(i)}]^T\right)$, $\mathbf{g}_{r,i} = [g_0^{(r,i)}, \ldots, g_{N_d-1}^{(r,i)}]^T$, and $\mathbf{v}_r = [v_0^{(r)}, \ldots, v_{N_d-1}^{(r)}]^T$, respectively. We note that in contrast to the pilot symbols $\mathbf{P}_t$ in (8), the data symbols in $\{\mathbf{D}_i\}_{i=0,\ldots,T-1}$ is unknown in a priori to the receiver and hence should be chosen among all available data tones (see Section III.B). By stacking pilot observation vector $\mathbf{z}_r$ and virtual pilot observation vector $\mathbf{y}_r$, we obtain the observation vector for the channel estimation

$$\begin{bmatrix} \mathbf{z}_r \\ \mathbf{y}_r \end{bmatrix} = \begin{bmatrix} \sqrt{\eta_{p,t}} \mathbf{P}_t \mathbf{h}_{r,t} \\ \sum_{i=0}^{T-1} \sqrt{\eta_{d,i}} \mathbf{D}_i \mathbf{g}_{r,i} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_r \\ \mathbf{v}_r \end{bmatrix} \tag{17}$$

$$= \begin{bmatrix} \sqrt{\eta_{p,t}} \mathbf{P}_t \mathbf{h}_{r,t} \\ \sqrt{\eta_{d,t}} \mathbf{D}_t \mathbf{g}_{r,t} \end{bmatrix} + \sum_{i=0, i\neq t}^{T-1} \begin{bmatrix} 0 \\ \sqrt{\eta_{d,i}} \mathbf{D}_i \mathbf{g}_{r,i} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_r \\ \mathbf{v}_r \end{bmatrix}. \tag{18}$$

Recall that $\mathbf{h}_{r,t}$ and $\mathbf{g}_{r,t}$ are the channel vectors for the pilot and data signals, respectively. Note that the second term in the right-hand side of (18) is the signal not coming from the transmit antenna t and hence becomes the interferences in channel estimation of the t-th transmit antenna. Since the estimation of channels using this interference-corrupted observation vector is undesirable, using newly updated soft information of data symbols, we cancel the interferences

$$\begin{bmatrix} \widetilde{\mathbf{z}}_r \\ \widetilde{\mathbf{y}}_r \end{bmatrix} = \begin{bmatrix} \mathbf{z}_r \\ \mathbf{y}_r \end{bmatrix} - \sum_{i=0, i\neq t}^{T-1} \begin{bmatrix} 0 \\ \sqrt{\eta_{d,i}} \overline{\mathbf{D}}_i \hat{\mathbf{g}}_{r,i} \end{bmatrix} \tag{19}$$

where the estimate of the interstream interference $\sqrt{\eta_{d,i}} \overline{\mathbf{D}}_i \hat{\mathbf{g}}_{r,i}$ is constructed from the soft estimate of the data symbols $\overline{\mathbf{D}}_i (\triangleq \mathrm{diag}([\bar{d}_0^{(i)}, \ldots, \bar{d}_{N_d-1}^{(i)}]^T)$ and the channel estimate $\hat{\mathbf{g}}_{r,i}$ obtained from the previous iteration. To be specific, if $L(c_{j,0}^{(i)}), \ldots, L(c_{j,Q-1}^{(i)})$ are the a posteriori LLRs of Q (coded) bits mapped to a data symbol $d_j^{(i)}$, then the first order moment of $d_j^{(i)}$ is [13]

$$\bar{d}_j^{(i)} = E[d_j^{(i)}] = \sum_{\theta\in\Theta} \theta \prod_{m=0}^{Q-1} Pr\left(c_{j,m}^{(i)}\right)$$

$$= \sum_{\theta\in\Theta} \theta \prod_{m=0}^{Q-1} \frac{1}{2}\left(1 + c_{j,m}^{(i)} \tanh\left(\frac{1}{2} L(c_{j,m}^{(i)})\right)\right), \tag{20}$$

where the set Θ includes all possible constellation points. In a similar way, the second order moment of $d_j^{(i)}$ is computed as

$$\lambda_j^{(i)} = E[|d_j^{(i)}|^2]$$

$$= \sum_{\theta\in\Theta} |\theta|^2 \prod_{m=0}^{Q-1} \frac{1}{2}\left(1 + c_{j,m}^{(i)} \tanh\left(\frac{1}{2} L(c_{j,m}^{(i)})\right)\right) \tag{21}$$

From (18) and (19), we have

$$\begin{aligned}\begin{bmatrix}\widetilde{\mathbf{z}}_r\\ \widetilde{\mathbf{y}}_r\end{bmatrix} &= \begin{bmatrix}\sqrt{\eta_{p,t}}\mathbf{P}_t\mathbf{h}_{r,t}\\ \sqrt{\eta_{d,t}}\mathbf{D}_t\mathbf{g}_{r,t}\end{bmatrix} \\ &\quad + \sum_{i=0,i\neq t}^{T-1}\begin{bmatrix}0\\ \sqrt{\eta_{d,i}}\mathbf{D}_i\mathbf{g}_{r,i}-\sqrt{\eta_{d,i}}\overline{\mathbf{D}}_i\hat{\mathbf{g}}_{r,i}\end{bmatrix}+\begin{bmatrix}\mathbf{n}_r\\ \mathbf{v}_r\end{bmatrix} \quad (22)\\ &= \begin{bmatrix}\sqrt{\eta_{p,t}}\mathbf{P}_t\mathbf{h}_{r,t}\\ \sqrt{\eta_{d,t}}\mathbf{D}_t\mathbf{g}_{r,t}\end{bmatrix} \\ &\quad + \sum_{i=0,i\neq t}^{T-1}\begin{bmatrix}0\\ \sqrt{\eta_{d,i}}\mathbf{D}_i(\mathbf{g}_{r,i}-\hat{\mathbf{g}}_{r,i})+\sqrt{\eta_{d,i}}\left(\mathbf{D}_i-\overline{\mathbf{D}}_i\right)\hat{\mathbf{g}}_{r,i}\end{bmatrix} \\ &\quad + \begin{bmatrix}\mathbf{n}_r\\ \mathbf{v}_r\end{bmatrix}. \quad (23)\end{aligned}$$

The MMSE estimates of $\mathbf{h}_{r,t}$ and $\mathbf{g}_{r,t}$ using the interference-cancelled observation $[\,\widetilde{\mathbf{z}}_r^T \quad \widetilde{\mathbf{y}}_r^T\,]^T$ in (23) are expressed as

$$\begin{aligned}\begin{bmatrix}\hat{\mathbf{h}}_{r,t}\\ \hat{\mathbf{g}}_{r,t}\end{bmatrix} &= \mathrm{Cov}\left(\begin{bmatrix}\mathbf{h}_{r,t}\\ \mathbf{g}_{r,t}\end{bmatrix},\begin{bmatrix}\widetilde{\mathbf{z}}_r\\ \widetilde{\mathbf{y}}_r\end{bmatrix}\right) \\ &\quad \times \mathrm{Cov}\left(\begin{bmatrix}\tilde{\mathbf{z}}_r\\ \tilde{\mathbf{y}}_r\end{bmatrix},\begin{bmatrix}\tilde{\mathbf{z}}_r\\ \tilde{\mathbf{y}}_r\end{bmatrix}\right)^{-1}\begin{bmatrix}\widetilde{\mathbf{z}}_r\\ \widetilde{\mathbf{y}}_r\end{bmatrix} \\ &= \Omega_{r,t}\Sigma_{r,t}^{-1}\begin{bmatrix}\widetilde{\mathbf{z}}_r\\ \widetilde{\mathbf{y}}_r\end{bmatrix}, \quad (24)\end{aligned}$$

where

$$\begin{aligned}\Omega_{r,t} &= \mathrm{Cov}\left(\begin{bmatrix}\mathbf{h}_{r,t}\\ \mathbf{g}_{r,t}\end{bmatrix},\begin{bmatrix}\widetilde{\mathbf{z}}_r\\ \widetilde{\mathbf{y}}_r\end{bmatrix}\right) \\ &= \begin{bmatrix}\sqrt{\eta_{p,t}}C_{h_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{h_t,g_t}\overline{\mathbf{D}}_t^H\\ \sqrt{\eta_{p,t}}C_{g_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{g_t,g_t}\overline{\mathbf{D}}_t^H\end{bmatrix} \quad (25)\end{aligned}$$

where $C_{h_t,g_t}=\mathrm{Cov}\,(\mathbf{h}_{r,t},\mathbf{g}_{r,t})$, $C_{g_t,g_t}=\mathrm{Cov}\,(\mathbf{g}_{r,t},\mathbf{g}_{r,t})$. In (25), we assumed that correlations of the channel frequency response across different antenna pairs are negligible when compared to those within the same antenna pair (i.e., $C_{h_i,h_i}\gg C_{h_i,h_j}$ and $C_{h_i,g_i}\gg C_{h_i,g_j}$ for $i\neq j$)[1].

Also,

$$\begin{aligned}\Sigma_{r,t} &= \mathrm{Cov}\left(\begin{bmatrix}\tilde{\mathbf{z}}_r\\ \tilde{\mathbf{y}}_r\end{bmatrix},\begin{bmatrix}\tilde{\mathbf{z}}_r\\ \tilde{\mathbf{y}}_r\end{bmatrix}\right) \\ &= \begin{bmatrix}\eta_{p,t}\mathbf{P}_tC_{h_t,h_t}\mathbf{P}_t^H+I & \sqrt{\eta_{p,t}\eta_{d,t}}\mathbf{P}_tC_{h_t,g_t}\overline{\mathbf{D}}_t^H\\ \sqrt{\eta_{p,t}\eta_{d,t}}\overline{\mathbf{D}}_tC_{g_t,h_t}\mathbf{P}_t^H & B\end{bmatrix} \quad (26)\end{aligned}$$

where

$$\begin{aligned}B &= \eta_{d,t}\left(\Lambda_t\odot C_{g_t,g_t}\right)+\sum_{i=0,i\neq t}^{T-1}\eta_{d,i}\left(\Lambda_i\odot C_{e_g,e_g}\right) \\ &\quad + \sum_{i=0,i\neq t}^{T-1}\eta_{d,i}\left(\Lambda_i-\overline{\mathbf{d}}_i\overline{\mathbf{d}}_i^H\right)\odot\hat{\mathbf{g}}_{r,i}\hat{\mathbf{g}}_{r,i}^H+I, \quad (27)\end{aligned}$$

where Λ_i=diag $\left([\lambda_0^{(i)},\ldots,\lambda_{N_d-1}^{(i)}]^T\right)$, C_{e_g,e_g}=Cov$(\mathbf{e}_{r,t},\mathbf{e}_{r,t})$, and $\mathbf{e}_{r,i}=\mathbf{g}_{r,i}-\hat{\mathbf{g}}_{r,i}$ is the estimation error of the channel $\mathbf{g}_{r,i}$. The derivation of (25) and (26) is presented in Appendix A.

[1]In fact, it is not too difficult to obtain the expression of the channel estimation when the correlations across different transmit antennas are non-negligible. However, the resulting expression is complicated and hence not so practical.

Note that C_{e_g,e_g} in (27) can be obtained from the error covariance matrix (see Appendix B). Note also that C_{h_t,h_t}, C_{h_t,g_t}, and C_{g_t,g_t} are computed once using (13) while C_{e_g,e_g} needs to be computed for each and every iteration. Once the channel estimates at pilot tones and data tones corresponding to the virtual pilot signals are obtained, the suitable interpolation scheme (e.g., linear interpolation) is applied to obtain the channel estimate for the rest of data tones.

It is worth mentioning that if the soft information on data tones is not available, which is true for the first iteration, then $\overline{\mathbf{D}}_i=\mathbf{0}$ and $\Lambda_i=\mathbf{I}$ and hence $\Omega_{r,t}$ in (25) and $\Sigma_{r,t}$ in (26) are simplified to

$$\Omega_{r,t} = \begin{bmatrix}\sqrt{\eta_{p,t}}C_{h_t,h_t}\mathbf{P}_t^H & 0\\ \sqrt{\eta_{p,t}}C_{g_t,h_t}\mathbf{P}_t^H & 0\end{bmatrix} \quad (28)$$

$$\Sigma_{r,t} = \begin{bmatrix}\eta_{p,t}\mathbf{P}_tC_{h_t,h_t}\mathbf{P}_t^H+I & 0\\ 0 & \sum_{i=0}^{T-1}\eta_{d,i}\left(\Lambda_i\odot C_{g_i,g_i}\right)+I\end{bmatrix}. \quad (29)$$

In this case, the channel estimate $\hat{\mathbf{h}}_{r,t}$ returns to the conventional MMSE based channel estimate in (12) since

$$\begin{aligned}&\begin{bmatrix}\hat{\mathbf{h}}_{r,t}\\ \hat{\mathbf{g}}_{r,t}\end{bmatrix} \\ &= \Omega_{r,t}\Sigma_{r,t}^{-1}\begin{bmatrix}\tilde{\mathbf{z}}_{r,t}\\ \tilde{\mathbf{y}}_{r,t}\end{bmatrix} \\ &= \begin{bmatrix}\sqrt{\eta_{p,t}}C_{h_t,h_t}\mathbf{P}_t^H & 0\\ \sqrt{\eta_{p,t}}C_{g_t,h_t}\mathbf{P}_t^H & 0\end{bmatrix} \\ &\quad \begin{bmatrix}\eta_{p,t}\mathbf{P}_tC_{h_t,h_t}\mathbf{P}_t^H+I & 0\\ 0 & \sum_{i=0}^{T-1}\eta_{d,i}\left(\Lambda_i\odot C_{g_i,g_i}\right)+I\end{bmatrix}^{-1} \\ &\quad \begin{bmatrix}\tilde{\mathbf{z}}_{r,t}\\ \tilde{\mathbf{y}}_{r,t}\end{bmatrix} \\ &= \begin{bmatrix}\sqrt{\eta_{p,t}}C_{h_t,h_t}\mathbf{P}_t^H\left(\eta_{p,t}\mathbf{P}_tC_{h_t,h_t}\mathbf{P}_t^H+I\right)^{-1}\tilde{\mathbf{z}}_{r,t}\\ \sqrt{\eta_{p,t}}C_{g_t,h_t}\mathbf{P}_t^H\left(\eta_{p,t}\mathbf{P}_tC_{h_t,h_t}\mathbf{P}_t^H+I\right)^{-1}\tilde{\mathbf{y}}_{r,t}\end{bmatrix} \\ &= \begin{bmatrix}\frac{1}{\sqrt{\eta_{p,t}}}C_{h_t,h_t}\left(C_{h_t,h_t}+\frac{1}{\eta_{p,t}}\mathbf{I}\right)^{-1}\mathbf{P}_t^H\mathbf{z}_r\\ \frac{1}{\sqrt{\eta_{p,t}}}C_{g_t,h_t}\left(C_{h_t,h_t}+\frac{1}{\eta_{p,t}}\mathbf{I}\right)^{-1}\mathbf{P}_t^H\mathbf{z}_r\end{bmatrix}, \quad (30)\end{aligned}$$

where $\tilde{\mathbf{z}}_r=\mathbf{z}_r$ from (19). This shows that the proposed channel estimator subsumes the conventional MMSE channel estimator when the soft symbol information is unavailable. It is worth mentioning that the channel estimate $\hat{\mathbf{g}}_{r,t}$ in (24) is the function of the covariance matrix C_{g_t,h_t} and the soft symbol estimate $\overline{\mathbf{D}}_i$. As the soft symbol estimate $\overline{\mathbf{D}}_i$ becomes more accurate, therefore, we expect better channel estimation performance.

B. Virtual Pilot Signal Selection Criterion

From our discussion thus far, it is clear that the choice of virtual pilot signals directly affects the quality of channel estimation. Perhaps an ideal way is to compare the performance metric (typically expressed in terms of MSE) for all possible choices of the virtual pilot signals $\mathbf{d}_t$ and then choose the set of data tones (say N_d data tones) minimizing the MSE. Since this approach

is computationally demanding and hence not so practical, we instead consider an approach of computing the MSE using only one virtual pilot tone. In this approach, a single data observation $y_n^{(r)} = \sum_{i=0}^{T-1} \sqrt{\eta_{d,i}} g_n^{(r,i)} d_n^{(i)} + v_n^{(r)}$ together with N_p pilot observations is used to compute the MSE of the data tone. Once the MSE is computed, we choose N_d-best candidates (data tones) with the smallest MSE as virtual pilots. For hypothetical selection of the n-th data tone, the MSE metric $\phi(n)$ is defined as

$$\phi(n) \triangleq E \left\| \begin{bmatrix} \mathbf{h}_{r,t} \\ g_n^{(r,t)} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{h}}_{r,t} \\ \hat{g}_n^{(r,t)} \end{bmatrix} \right\|^2 \tag{31}$$

$$= \mathrm{tr}\left(\mathrm{Cov}\left(\begin{bmatrix} \mathbf{h}_{r,t} \\ g_n^{(r,t)} \end{bmatrix} - \widetilde{\Omega}_{r,t} \widetilde{\Sigma}_{r,t}^{-1} \begin{bmatrix} \mathbf{z}_r \\ \tilde{y}_n^{(r)} \end{bmatrix} \right) \right) \tag{32}$$

$$= \mathrm{tr}\left(\mathrm{Cov}\left(\begin{bmatrix} \mathbf{h}_{r,t} \\ g_n^{(r,t)} \end{bmatrix} \right) - \mathrm{Cov}\left(\widetilde{\Omega}_{r,t} \widetilde{\Sigma}_{r,t}^{-1} \begin{bmatrix} \mathbf{z}_r \\ \tilde{y}_n^{(r)} \end{bmatrix} \right) \right) \tag{33}$$

$$= \mathrm{tr}\left(\mathrm{Cov}\left(\begin{bmatrix} \mathbf{h}_{r,t} \\ g_n^{(r,t)} \end{bmatrix} \right) - \widetilde{\Omega}_{r,t} \widetilde{\Sigma}_{r,t}^{-1} \widetilde{\Sigma}_{r,t} \widetilde{\Sigma}_{r,t}^{-1} \widetilde{\Omega}_{r,t}^H \right) \tag{34}$$

$$= \mathrm{tr}\left(\mathrm{Cov}\left(\begin{bmatrix} \mathbf{h}_{r,t} \\ g_n^{(r,t)} \end{bmatrix} \right) - \widetilde{\Omega}_{r,t} \widetilde{\Sigma}_{r,t}^{-1} \widetilde{\Omega}_{r,t}^H \right), \tag{35}$$

where

$$\widetilde{\Omega}_{r,t} = \mathrm{Cov}\left(\begin{bmatrix} \mathbf{h}_{r,t} \\ g_n^{(r,t)} \end{bmatrix}, \begin{bmatrix} \mathbf{z}_r \\ \tilde{y}_n^{(r)} \end{bmatrix} \right) = \begin{bmatrix} \sqrt{\eta_{p,t}} C_{h_t,h_t} \mathbf{P}_t^H & \sqrt{\eta_{d,t}} C_{h_t,g_t} \left(\overline{d}_n^{(t)} \right)^* \\ \sqrt{\eta_{p,t}} C_{g_t,h_t} \mathbf{P}_t^H & \sqrt{\eta_{d,t}} C_{g_t,g_t} \left(\overline{d}_n^{(t)} \right)^* \end{bmatrix} \tag{36}$$

and

$$\widetilde{\Sigma}_{r,t} = \mathrm{Cov}\left(\begin{bmatrix} \mathbf{z}_r \\ \tilde{y}_n^{(r)} \end{bmatrix} \right) = \begin{bmatrix} \eta_{p,t} \mathbf{P}_t C_{h_t,h_t} \mathbf{P}_t^H + I & \sqrt{\eta_{p,t}\eta_{d,t}} \mathbf{P}_t C_{h_t,g_t} \left(\overline{d}_n^{(t)} \right)^* \\ \sqrt{\eta_{p,t}\eta_{d,t}} \overline{d}_n^{(t)} C_{g_t,h_t} \mathbf{P}_t^H & \sum_{i=0}^{T-1} \eta_{d,i} \lambda_n^{(i)} + 1 \end{bmatrix}, \tag{37}$$

where $\tilde{y}_n^{(r)} = \sqrt{\eta_{d,t}} g_n^{(r,t)} d_n^{(t)} + \sum_{i=0,\neq t}^{T-1} \sqrt{\eta_{d,i}} g_n^{(r,i)} (d_n^{(i)} - \overline{d}_n^{(i)}) + v_n^{(r)}$, $C_{h_t,g_t} = \mathrm{Cov}\left(\mathbf{h}_{r,t}, g_n^{(r,t)} \right)$, and $C_{g_t,g_t} = E[|g_n^{(r,t)}|^2] = 1$. Recall that the definitions of $\overline{d}_n^{(t)}$ and $\lambda_n^{(i)}$ are provided in (20) and (21), respectively. Also, we ignored the correlations across different antenna pairs in (36) and (37) as we did in Section III.A.

From (35), (36), and (37), $\phi(n)$ is expressed as (see Appendix C)

$$\phi(n) = \Psi - \frac{\frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 \frac{1}{\eta_{p,t}^2} C_{g_t,h_t} \Sigma^{-2} C_{h_t,g_t}}{1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 C_{g_t,h_t} \Sigma^{-1} C_{h_t,g_t}} + \frac{\left(1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2\right) \left(1 - C_{g_t,h_t} \Sigma^{-1} C_{h_t,g_t}\right)}{1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 C_{g_t,h_t} \Sigma^{-1} C_{h_t,g_t}}, \tag{38}$$

where

$$\Sigma = C_{h_t,h_t} + \frac{1}{\eta_{p,t}} I \tag{39}$$

$$\Psi = \mathrm{tr}\left(\frac{1}{\eta_{p,t}} C_{h_t,h_t} \left(C_{h_t,h_t} + \frac{1}{\eta_{p,t}} I \right)^{-1} \right) \tag{40}$$

$$\xi = \sum_{i=0}^{T-1} \eta_{d,i} \lambda_n^{(i)} + 1. \tag{41}$$

Interestingly, we see from (38) that $|\overline{d}_n^{(t)}|^2/\xi$ and C_{h_t,g_t} are two main parameters affecting the data tone selection. Since the evaluation of $\phi(n)$ in (38) is expensive, we consider the scenario where pilot signals are spaced apart (typical scenario for large-scale MIMO systems). Since the correlation between pilot signals is weak in this case, we approximately have $C_{h_t,h_t} \approx I$ and hence we get the expression of $\phi(n)$ [see (42)–(43) shown at the bottom of the page]. In the high signal to noise ratio (SNR) regime ($\eta_{p,t} \gg 1$), we have $\frac{\eta_{p,t}}{\eta_{p,t}+1} \approx 1$ and $\frac{\eta_{p,t}^2 + \eta_{p,t} - 1}{(\eta_{p,t}+1)^2} \approx 1$, and thus

$$\phi(n) \approx \Psi + \frac{\left(1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2\right) \left(1 - \|C_{h_t,g_t}\|^2\right)}{1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 \|C_{h_t,g_t}\|^2} = \Psi + \frac{1}{\frac{1}{1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2} + \frac{1}{1 - \|C_{h_t,g_t}\|^2} - 1} \tag{44}$$

Noting that Ψ is not related to the data tone selection, one can observe that the data tone minimizing $\phi(n)$ is one maximizing the denominator of the second term in (44). In light of this, the resulting cost metric becomes

$$\tilde{\phi}(n) \triangleq \frac{1}{1 - \frac{|\overline{d}_n^{(t)}|^2}{\sum_{i=0}^{T-1} \eta_{d,i} \lambda_n^{(i)} + 1}} + \frac{1}{1 - \|C_{h_t,g_t}\|^2} \tag{45}$$

$$\phi(n) \approx \Psi - \frac{\frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 \frac{1}{(\eta_{p,t}+1)^2} \|C_{h_t,g_t}\|^2 - \left(1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2\right) \left(1 - \frac{\eta_{p,t}}{\eta_{p,t}+1} \|C_{h_t,g_t}\|^2\right)}{1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 \frac{\eta_{p,t}}{\eta_{p,t}+1} \|C_{h_t,g_t}\|^2} \tag{42}$$

$$= \Psi - \frac{-1 + \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 + \frac{\eta_{p,t}}{\eta_{p,t}+1} \|C_{h_t,g_t}\|^2 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 \frac{\eta_{p,t}^2 + \eta_{p,t} - 1}{(\eta_{p,t}+1)^2} |C_{h_t,g_t}\|^2}{1 - \frac{1}{\xi}\eta_{d,t} |\overline{d}_n^{(t)}|^2 \frac{\eta_{p,t}}{\eta_{p,t}+1} \|C_{h_t,g_t}\|^2}. \tag{43}$$

TABLE I
SUMMARY OF THE PROPOSED CHANNEL ESTIMATION ALGORITHM

Input: The LLRs on the transmitted symbols $\{d_n^{(t)}\}$, the received vectors $\{y_n^{(r)}\}$, the pilot sequence $\{\mathbf{P}_t\}$, the transmitted power for the pilot and data symbols $\{\eta_{p,t}\}$ and $\{\eta_{d,i}\}$

STEP 1: (**Virtual pilot signal selection**) Select the data signals according to the MSE metric in (45).
STEP 2: (**Interference subtraction**) Subtract the effect of interfering symbols from other antennas using (23).
STEP 3: (**Covariance matrix update**) Update the MMSE error covariance matrices from (52).
STEP 4: (**Channel estimate update**) Calculate the MMSE estimation using (25) to (26) and apply the MMSE filter to the observation using (24).
STEP 5: (**Interpolation**) Perform interpolation using the channel estimates both at the pilot tones and virtual pilot tones.

Note that the modified cost function $\tilde{\phi}(n)$ depends on the reliability of soft decisions $\left(\frac{|\overline{d}_n^{(t)}|^2}{\sum_{i=0}^{T-1}\eta_{d,i}\lambda_n^{(i)}+1}\right)^2$ and the correlation (of channel gains) between the data tones and the pilot tones ($\|C_{h_t,g_t}\|^2$). Hence, data tones with high reliability $\left(\frac{|\overline{d}_n^{(t)}|^2}{\sum_{i=0}^{T-1}\eta_{d,i}\lambda_n^{(i)}+1}\right)$ and correlation $\|C_{h_t,g_t}\|^2$ are chosen as the virtual pilots. In summary, for each transmit antenna, N_d-best data tones maximizing the simplified cost metric $\tilde{\phi}(n)$ are chosen. Observations of these tones, together with pilot observations, are used for the channel re-estimation. The proposed channel estimation algorithm is summarized in Table I.

C. Comments on Complexity

In this subsection, we analyze the computational complexity of the proposed channel estimator. Since the detection and decoding operations at the receiver are common, we focus only on computations associated with the channel estimation. While the major operations of the conventional MMSE estimation include the inversion and multiplication of a matrix in (12), those of the proposed scheme include interference cancellation in (19) and calculation of $\Omega_{r,t}$ and $\Sigma_{r,t}$ in (25) and (26). We denote the number of floating-point operations (flops) required for the conventional MMSE estimator as C_{conv} and those for the proposed scheme as C_{prop}. Then, C_{prop} can be factored into C_{IC} for interference cancellation, C_Ω for calculation of $\Omega_{r,t}$, and C_Σ for calculation of $\Sigma_{r,t}$. The complexity for each operation can be obtained as [18]:

- C_{conv} includes $\frac{4}{3}N_p^3+\frac{3}{2}N_p^2+\frac{1}{6}N_p$ flops for computing $\left(C_{h_t,h_t}+\frac{1}{\eta_{p,t}}\mathbf{I}\right)^{-1}$ and $4N_p^2$ flops for performing the rest of multiplications.
- C_{IC} includes N_tN_d flops for computing $\mathbf{y}_r-\sum_{i=0,i\neq t}^{T-1}\sqrt{\eta_{d,i}}\overline{\mathbf{D}}_i\hat{\mathbf{g}}_{r,i}$.
- C_Ω includes N_pN_d flops for the matrix computation for $\hat{\mathbf{h}}_{r,t}$.
- C_Σ includes $2N_p^2+2N_pN_d+5(N_t-1)N_d^2+(2N_t-1)N_d$ flops for matrix computation and $\frac{4}{3}(N_p+N_d)^3+\frac{3}{2}(N_p+N_d)^2-\frac{5}{6}(N_p+N_d)$ flops for inverting $\Sigma_{r,t}$.
- C_{prop} is the sum of C_{IC}, C_Ω, C_Σ, and $2(N_p+N_d)^2-(2N_p-1)(N_p+N_d)-N_p$ flops for the multiplication of $\Omega_{r,t}$, $\Sigma_{r,t}^{-1}$, and received vector in (24).

As observed in this summary, the computation complexity of the proposed scheme is a cubic function in the number of all pilot signals being used (real and virtual pilot signals). In a real implementation, therefore, we need to choose the number of the virtual pilots so that the resulting system does not violate the computational requirements.

[2]As the decisions on symbols are more reliable, the squared mean estimate $|\overline{d}_n^{(t)}|^2$ increases while the variances $\{\lambda_n^{(i)}\}_{i=0,\ldots,T-1}$ reduces.

IV. SIMULATIONS

In this section, we compare the performance of the proposed channel estimation technique with the conventional approaches through numerical simulations.

A. Simulation Setup

The simulation setup is based on MIMO-OFDM system with QPSK modulation. We consider 4×4 and 8×8 MIMO systems on Extended Vehicular A (EVA) channel model [19]. We consider pilot allocation strategy used for common reference signal (CRS) in current LTE standard [12]. Note that the CRS pilots are spaced uniformly over the time and frequency grid since they are shared among users in a cell (see Fig. 4). Note also that since pilot assignment is undefined for 8×8 systems, we assign one pilot symbol for each transmit antenna and each slot.

For the channel encoding, we use the half rate ($C=\frac{1}{2}$) Turbo code with feedback polynomial $1+D+D^2$ and feedforward polynomial $1+D^2$. Each code block spans a length of 14 OFDM symbols. As a MIMO detection algorithm, we employ a linear MMSE detector with parallel interference canceller (MMSE-PIC) [21] and as a soft-input soft-output channel decoding, we use the BCJR algorithm [22]. Between the MIMO detector and the channel decoder, a random interleaver is employed. We run maximum seven outer iteration but stop the iteration whenever the packet passes the CRC check (i.e., packet is decoded correctly) or no improvement in the channel estimation quality at pilot tones is observed (i.e., $\|\hat{\mathbf{h}}_{r,t}^{(j+1)}-\hat{\mathbf{h}}_{r,t}^{(j)}\|_2<\epsilon$ where j is the iteration index). The details of the simulation setup are summarized in Table II.

As a metric to measure the performance, we use the normalized MSE as well as bit error rate (BER) at the channel decoder output. These metrics are measured as a function of E_b/N_0, which is defined as [13]

$$\left.\frac{E_b}{N_0}\right|_{\text{dB}}=SNR|_{\text{dB}}+10\log_{10}\frac{R}{TCQ} \tag{46}$$

where R and T are the number of receive and transmit antennas, C is the code rate, Q is the number of bits per symbol, respectively. In our simulations, we consider the following algorithms:

1) Conventional MMSE channel estimator [20]: pilot-based channel estimation is performed based on the MMSE criteria.

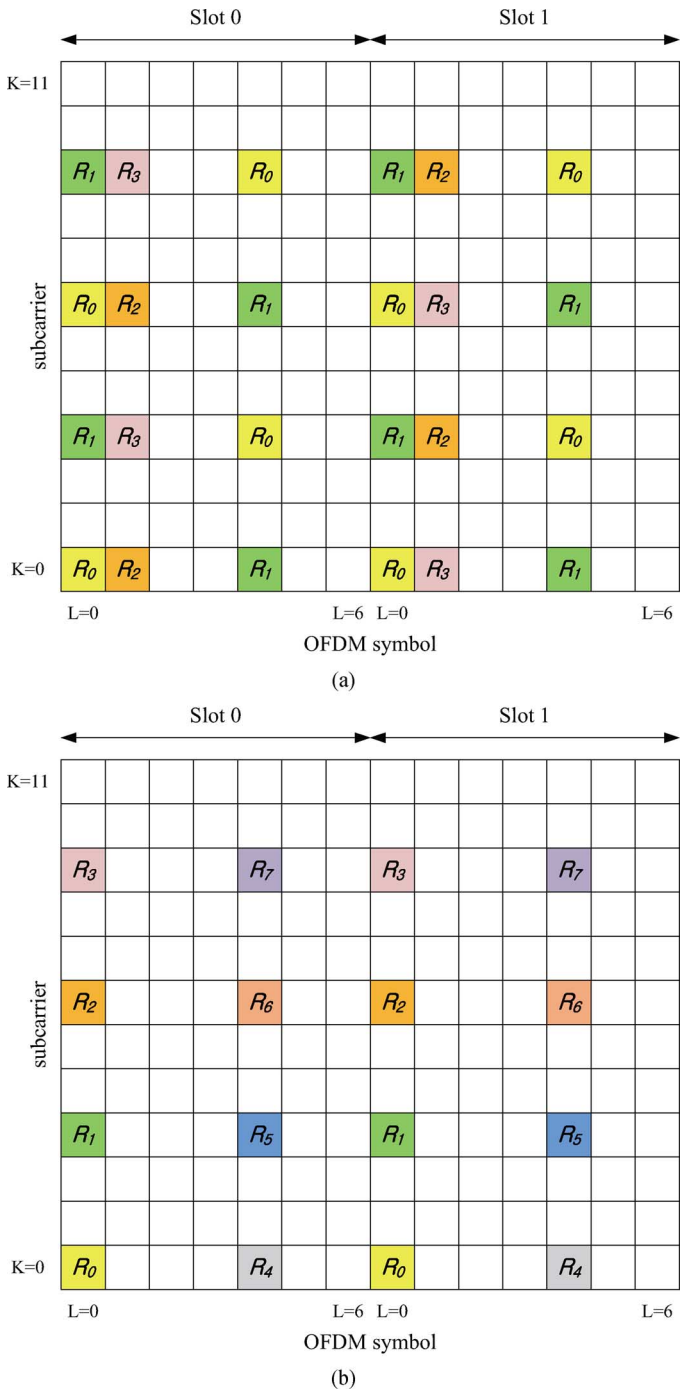


Fig. 4. Mapping of downlink pilot signals: (a) 4×4 and (b) 8×8 MIMO case.

2) Iterative decision-directed estimator [11]: In the first iteration, the MMSE channel estimation is performed. In the next iteration, iterative compensated MMSE (IC-MMSE) algorithm is conducted using the soft symbol estimates. To the best of our knowledge, the IC-MMSE seems to be the best one among DD-CE algorithms developed for the MIMO-OFDM systems.
3) Proposed channel estimator: In each iteration, a total of 32 virtual pilot signals are used. Through intensive simulations, we observe that 32 virtual pilots are enough to achieve maximal performance gain with reasonable complexity.
4) Proposed channel estimator with ideal interference cancellation: The interstream interference in virtual pilot signal (second term in the right-hand of (23)) is removed perfectly and then the proposed channel estimator is applied. This scheme provides the lower bound of the proposed method.

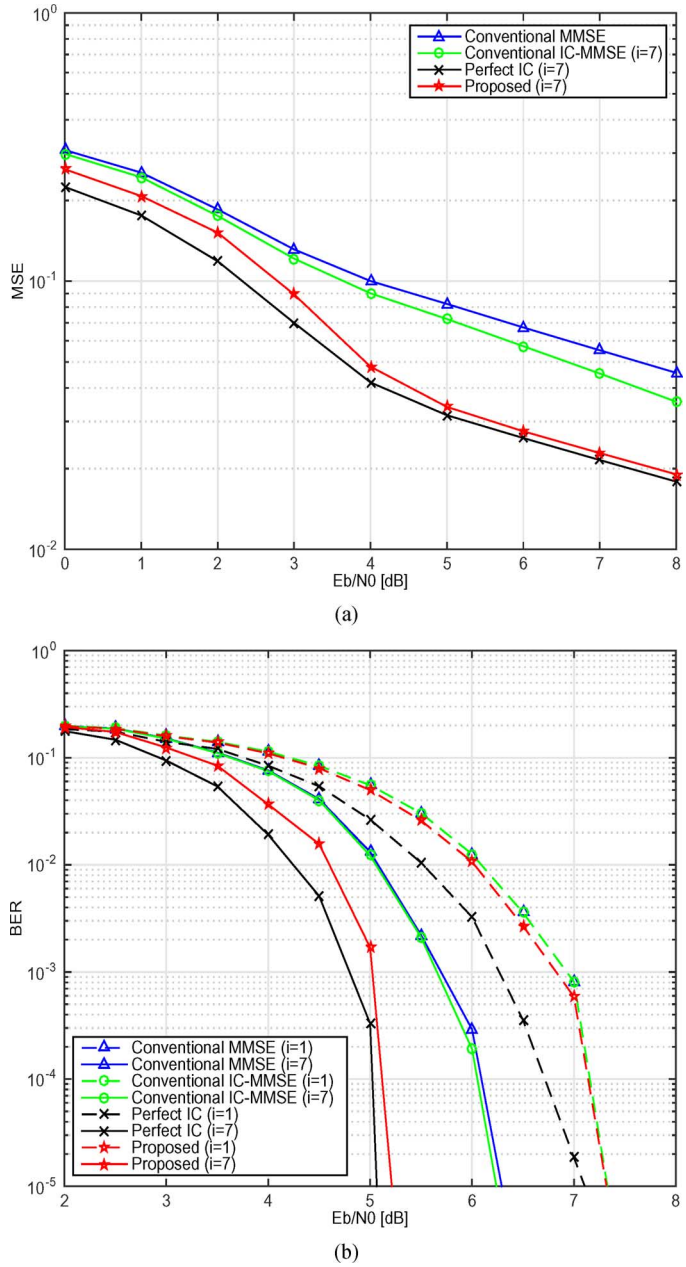


Fig. 5. Performance of the conventional and proposed scheme for 4×4 MIMO systems: (a) MSE and (b) BER.

B. Simulation Results

In Fig. 5(a), we investigate the MSE performance of the channel estimation algorithms in 4×4 MIMO systems. As shown in Fig. 5(a), the proposed method outperforms the conventional MMSE channel estimator and IC-MMSE channel estimator, yielding about 1.5 dB gain at 10% MSE. Whereas, the performance gain of the IC-MMSE over the conventional

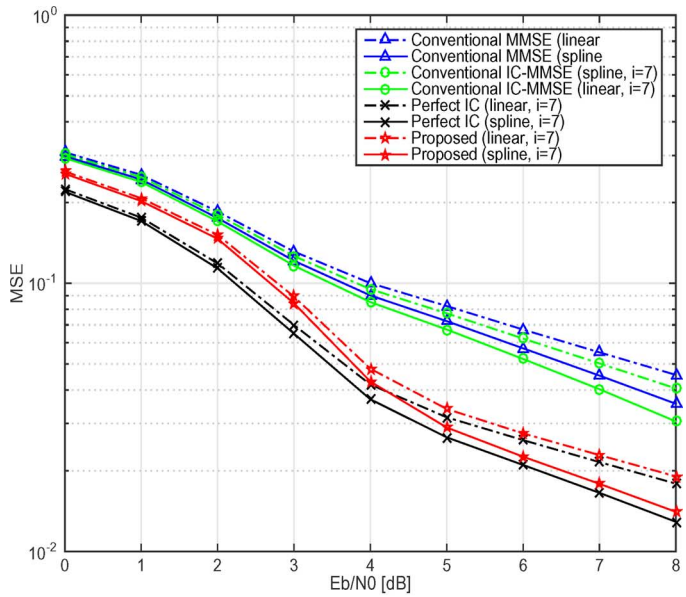
TABLE II
DETAILS OF THE SIMULATION SETUP

Parameter	Test1	Test2	Test3	Test4
Dimension of MIMO systems	4×4	4×4	8×8	4×4
System bandwidth	5Mhz	5Mhz	3Mhz	5Mhz
Channel model	EVA	EVA	EVA	EVA
Doppler frequency	70 Hz	70 Hz	70 Hz	70 Hz
MIMO correlation	low	low	low	high
Pilot overhead	14.29%	14.29%	19.05%	14.29%
Number of source bits	14,400	14,400	18,240	14,400
Modulation	QPSK	QPSK	QPSK	QPSK
Channel coding	Turbo code	Turbo code	Turbo code	Turbo code
Code rate	1/2	1/2	1/2	1/2
MIMO detector	MMSE-PIC	MMSE-PIC	MMSE-PIC	MAP
Number of inner iterations (Turbo decoder)	8	8	8	8
Number of outer iterations (IDD iteration)	7	7	7	7
Interpolation method	linear	spline cubic	linear	linear

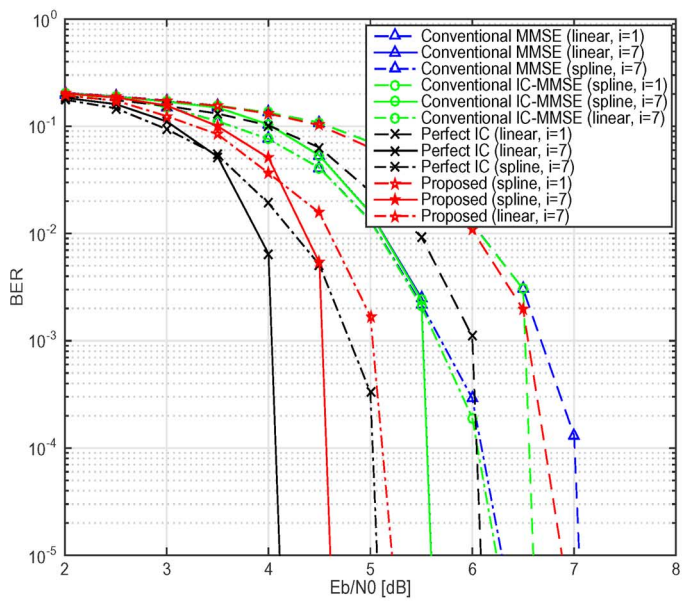
MMSE channel estimator is a bit marginal. This is because the IC-MMSE method does not account for the statistics of the residual interstream interference in the channel estimation. Note that while the IC-MMSE provides considerable performance gain over the conventional scheme for 2×2 systems [11], the gain fastly diminishes for 4×4 systems due to the improper control of the interstream interferences. Since the virtual pilot signals become more reliable as the SNR increases, it is no wonder that the proposed method performs close to the ideal interference cancellation scheme in high SNR regime. In Fig. 5(b), we plot the BER performance of channel estimation schemes for 4×4 MIMO systems. Although the gain of the proposed method over the conventional scheme is a bit too small in the first iteration ($i = 1$), the gain increases and becomes more than 1 dB in the waterfall regime when the number of iterations increases. Due to the reason stated above, IC-MMSE provides only marginal gain over the conventional MMSE channel estimator. Similar to the MSE case, the BER performance of the proposed channel estimator becomes close to that of the ideal interference cancellation scheme for high SNR regime.

Since the channel estimates at data tones are obtained via the interpolation of the channel estimates at pilot tones, it is of interest to investigate the impact of interpolation scheme on performance. In Fig. 6, we plot the MSE and BER performances of the channel estimation algorithms under test when the spline cubic interpolation method is used after channel estimation. While the linear interpolation literally offers a linear curve fitting to the data points provided, the spline cubic interpolation (spline function in *MATLAB*) provides more smooth and continuous polynomial fitted to the given data points. As shown in Fig. 6, we see that the performance gain of the proposed method is maintained when spline cubic interpolation is used. This is because the proposed scheme improves the channel estimation quality at both pilot and data tones so that it will also be helpful to obtain better interpolated values at data tones.

We next consider the performance of the proposed method for 8×8 MIMO systems. We observe from Fig. 7(a) that the proposed method outperforms the conventional LS and MMSE channel estimators by about 3 dB at 10% MSE. When compared to 4×4 MIMO systems, 8×8 MIMO systems have lower pilot density (67% of 4×4 systems) so that the proposed method

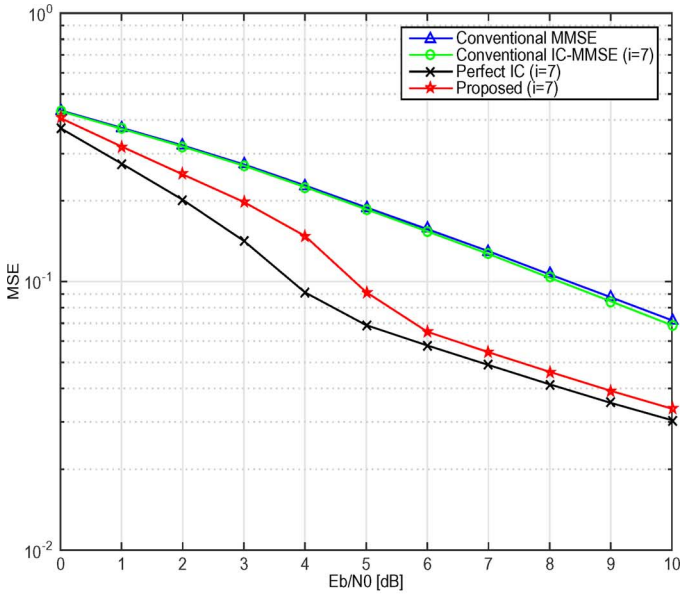


(a)

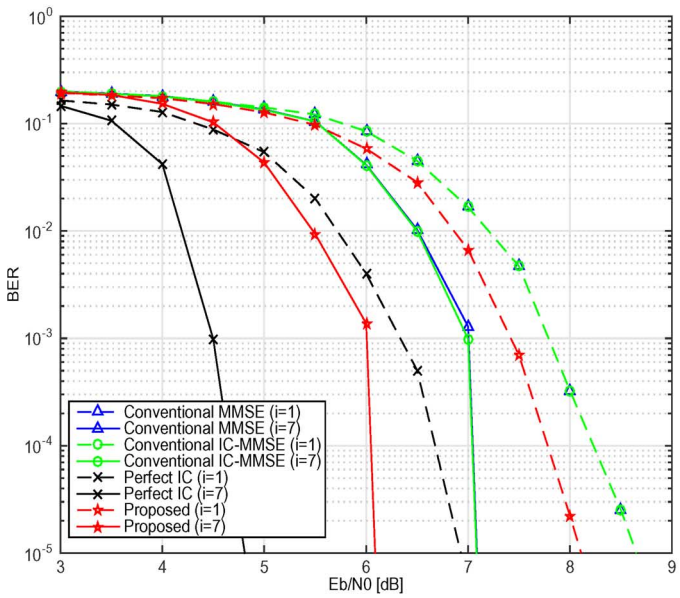


(b)

Fig. 6. Performance of the conventional and proposed scheme with spline cubic interpolation: (a) MSE and (b) BER.



(a)



(b)

Fig. 7. Performance of the conventional and proposed scheme for 8×8 MIMO systems: (a) MSE and (b) BER.

using both pilot signals and virtual pilot signals has clear benefit over the IC-MMSE as well as conventional scheme using pilot signals only. In Fig. 7(b), we plot the BER performance for 8×8 MIMO systems. Due to strong interstream interferences in the 8×8 system, we observe a considerable performance gap between the ideal interference cancellation scheme and the rest. While the performance of the IC-MMSE channel estimator is not so appealing in this setting, due to the benefit of using virtual pilot signals, the proposed method offers 0.5 dB performance gain (at BER $= 10^{-4}$) over the IC-MMSE channel estimator at the first iteration. The gain becomes more than 1 dB after seven iterations.

In Fig. 8, we investigate the MSE as a function of the number of virtual pilot signals for 8×8 MIMO systems. In general,

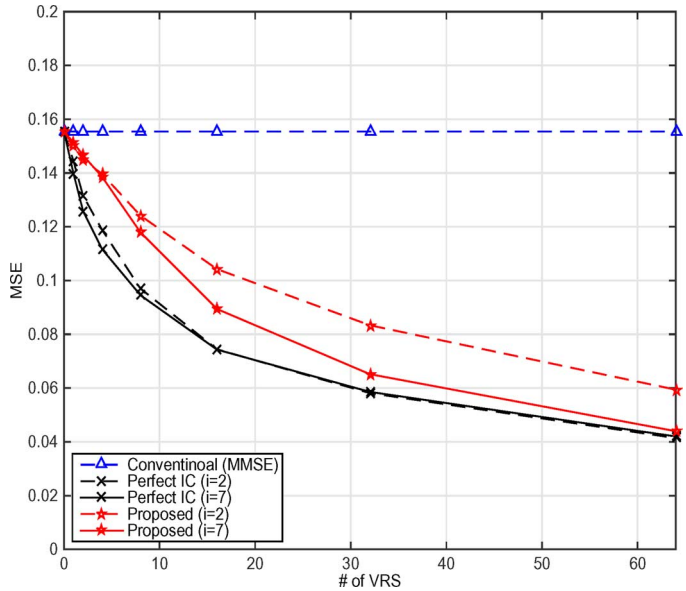


Fig. 8. MSE performance of conventional and proposed schemes as a function of number of virtual RS.

the MSE of the proposed method improves with the number of virtual pilot signals. We see that as the number of virtual pilot symbols increases, the performance of the proposed method approaches to the performance of the ideal interference cancellation scheme. We also observe that the gain obtained by the virtual pilot signals diminishes as the number of virtual pilot symbols increases.

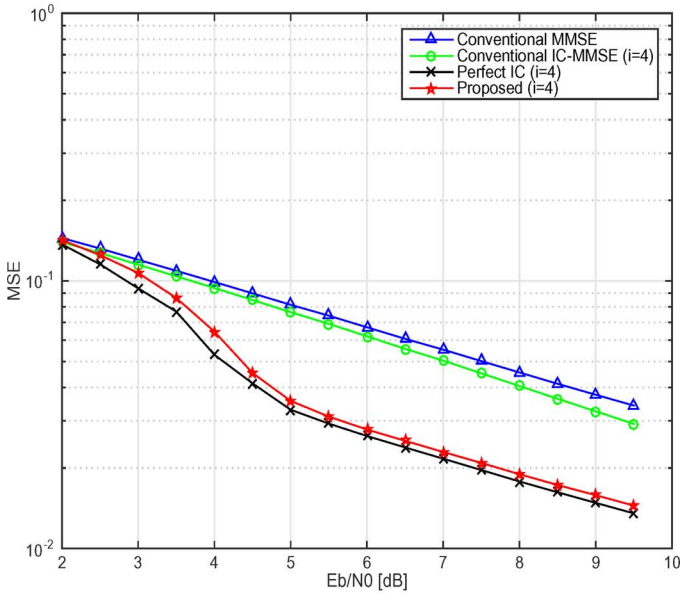
Finally, we consider the case where the channel matrix is spatially correlated. In this case, overall channel matrix is expressed as $\mathbf{H} = \mathbf{H}_{iid}\mathbf{C}^{1/2}$ where $\mathbf{H}_{iid}$ is the matrix whose entries follows complex Gaussian $\mathcal{CN}(0,1)$ and where $\mathbf{C}$ is the covariance matrix modeling the correlation in the transmit antenna array

$$\mathbf{C} = \begin{bmatrix} 1 & \alpha^{\frac{1}{9}} & \alpha^{\frac{4}{9}} & \alpha \\ \alpha^{\frac{1}{9}*} & 1 & \alpha^{\frac{1}{9}} & \alpha^{\frac{4}{9}} \\ \alpha^{\frac{4}{9}*} & \alpha^{\frac{1}{9}*} & 1 & \alpha^{\frac{1}{9}} \\ \alpha^{*} & \alpha^{\frac{4}{9}*} & \alpha^{\frac{1}{9}*} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.7169 & 0.2641 & 0.05 \\ 0.7169 & 1 & 0.7169 & 0.2641 \\ 0.2641 & 0.7169 & 1 & 0.7169 \\ 0.05 & 0.2641 & 0.7169 & 1 \end{bmatrix},$$

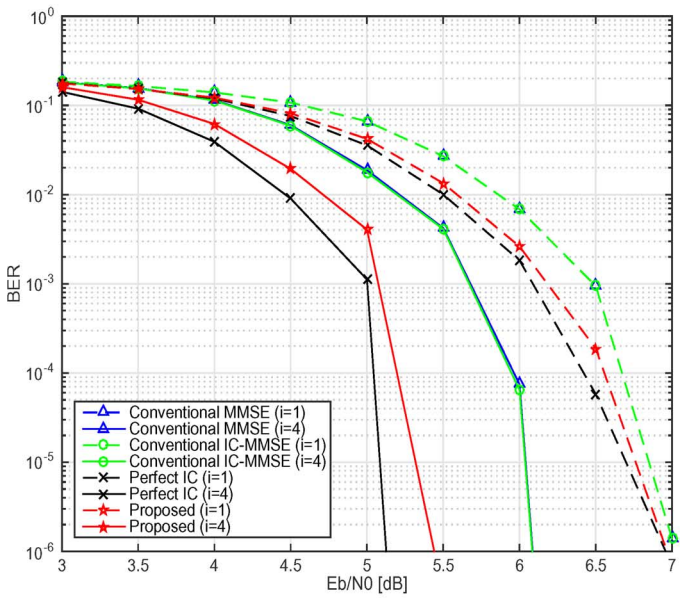
where we set $\alpha = 0.05$ in our simulations. Fig. 9(a) and (b) compare the MSE and BER performances of the channel estimation schemes for the correlated channel condition. We see that the proposed method exhibits 0.7 dB gain in BER performance over the conventional scheme. Although the gain is slightly worse than the uncorrelated scenario, we still see that the proposed method is effective and maintains a decent performance gain over the conventional ones.

V. Conclusions

In this paper, we proposed a new channel estimation algorithm to deal with the pilot shortage problem in MIMO-OFDM systems. By using the *reliable* data tones highly correlated with the pilot signals for the channel estimation, the proposed method



(a)



(b)

Fig. 9. Performance of the conventional and proposed scheme for the correlated channels: (a) MSE and (b) BER.

achieves better channel estimation quality at data tone and eventual performance improvement.

While our focus in this work was primarily on the OFDM-based MIMO system, the main concept can be readily applied to non-OFDM various wireless systems such as single-carrier MIMO systems, channel estimation under interference-limited environments (e.g., hot spots). Also, the extension of the proposed scheme to the multi-user MIMO (MU-MIMO) scenario employing demodulation reference signals (DM-RS) is an interesting direction worth investigation. Note that since the DM-RS, used for the demodulation of MU-MIMO in LTE-Advanced systems, is present only for the user being serviced, it is compactly assigned to the small number of resource blocks. Besides, since the DM-RS is mainly targeted for the demodulation of data symbols for MU-MIMO scenario, the precoding to remove the interuser interference is applied. Due to these reasons, we expect that the residual interference in the received signal at data tones would be smaller than the residual interference (interstream interference) for the single-user MIMO (SU-MIMO) system and the proposed scheme might provide better gain than what SU-MIMO offers. We leave these interesting explorations for our future work.

Appendix A
Derivation of (25) and (26)

Assuming that the cross-antenna correlations are negligible, i.e., $C_{h_i,h_j} = C_{h_i,g_j} = C_{g_i,g_j} = 0$ for $i \neq j$ and using (23), the cross-covariance matrix $\Omega_{r,t}$ is expressed as

$$\Omega_{r,t} = E\begin{bmatrix} \mathbf{h}_{r,t}\tilde{\mathbf{z}}_r^H & \mathbf{h}_{r,t}\tilde{\mathbf{y}}_r^H \\ \mathbf{g}_{r,t}\tilde{\mathbf{z}}_r^H & \mathbf{g}_{r,t}\tilde{\mathbf{y}}_r^H \end{bmatrix} = \begin{bmatrix} \sqrt{\eta_{p,t}}C_{h_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{h_t,g_t}\overline{\mathbf{D}}_t^H \\ \sqrt{\eta_{p,t}}C_{g_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{g_t,g_t}\overline{\mathbf{D}}_t^H \end{bmatrix}. \quad (47)$$

The auto-covariance matrix $\Sigma_{r,t}$ is also written by

$$\Sigma_{r,t} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} E[\tilde{\mathbf{z}}_r\tilde{\mathbf{z}}_r^H] & E[\tilde{\mathbf{z}}_r\tilde{\mathbf{y}}_r^H] \\ E[\tilde{\mathbf{y}}_r\tilde{\mathbf{z}}_r^H] & E[\tilde{\mathbf{y}}_r\tilde{\mathbf{y}}_r^H] \end{bmatrix} \quad (48)$$

where the elements of the matrix $\Sigma_{r,t}$ are obtained by

$$\begin{aligned} &A_{11} \\ &= E\left[(\sqrt{\eta_{p,t}}\mathbf{P}_t\mathbf{h}_{r,t} + \mathbf{n}_r)(\sqrt{\eta_{p,t}}\mathbf{P}_t\mathbf{h}_{r,t} + \mathbf{n}_r)^H\right] \\ &= \eta_{p,t}\mathbf{P}_t C_{h_t,h_t}\mathbf{P}_t^H + I \quad (49) \\ &A_{12} \\ &= A_{21}^H = E\Bigg[(\sqrt{\eta_{p,t}}\mathbf{P}_t\mathbf{h}_{r,t} + \mathbf{n}_r)(\sqrt{\eta_{d,t}}\mathbf{D}_t\mathbf{g}_{r,t} \\ &\quad + \sum_{i=0,i\neq t}^{T-1}(\sqrt{\eta_{d,i}}\mathbf{D}_i\mathbf{g}_{r,i} - \sqrt{\eta_{d,i}}\overline{\mathbf{D}}_i\hat{\mathbf{g}}_{r,i}) + \mathbf{v}_r)^H\Bigg] \\ &= \sqrt{\eta_{p,t}\eta_{d,t}}\mathbf{P}_t C_{h_t,g_t}\overline{\mathbf{D}}_t^H \quad (50) \\ &A_{22} \\ &= E\Bigg[\Bigg(\sqrt{\eta_{d,t}}\mathbf{D}_t\mathbf{g}_{r,t} \\ &\quad + \sum_{i=0,i\neq t}^{T-1}\left(\sqrt{\eta_{d,i}}\mathbf{D}_i\mathbf{e}_{r,i} + \sqrt{\eta_{d,i}}\left(\mathbf{D}_i - \overline{\mathbf{D}}_i\right)\hat{\mathbf{g}}_{r,i}\right) + \mathbf{v}_r\Bigg) \\ &\quad \cdot \Bigg(\sqrt{\eta_{d,t}}\mathbf{D}_t\mathbf{g}_{r,t} \\ &\quad + \sum_{i=0,i\neq t}^{T-1}\left(\sqrt{\eta_{d,i}}\mathbf{D}_i\mathbf{e}_{r,i} + \sqrt{\eta_{d,i}}\left(\mathbf{D}_i - \overline{\mathbf{D}}_i\right)\hat{\mathbf{g}}_{r,i}\right) + \mathbf{v}_r\Bigg)^H\Bigg] \\ &= \eta_{d,t}\left(\Lambda_t \odot C_{g_t,g_t}\right) + \sum_{i=0,i\neq t}^{T-1}\eta_{d,i}\left(\Lambda_i \odot C_{e_g,e_g}\right) \\ &\quad + \sum_{i=0,i\neq t}^{T-1}\eta_{d,i}\left(\Lambda_i - \overline{\mathbf{d}}_i\overline{\mathbf{d}}_i^H\right) \odot \hat{\mathbf{g}}_{r,i}\hat{\mathbf{g}}_{r,i}^H + I, \quad (51) \end{aligned}$$

where $\Lambda_i = \mathrm{diag}\left([\lambda_0^{(i)}, \ldots, \lambda_{N_d-1}^{(i)}]^T\right)$, $C_{e_g,e_g} = \mathrm{Cov}(\mathbf{e}_{r,t}, \mathbf{e}_{r,t})$, and $\mathbf{e}_{r,i} = \mathbf{g}_{r,i} - \hat{\mathbf{g}}_{r,i}$ is the estimation error of

the channel $\mathbf{g}_{r,i}$. In (51), we used the orthogonality principle of the MMSE estimation (i.e., $E[\mathbf{e}_{r,i}\hat{\mathbf{g}}_{r,i}^H]=0$).

Appendix B
Derivation of Error Covariance Matrix C_{e_g,e_g}

Let $\mathbf{e}_h=\mathbf{h}_{r,t}-\hat{\mathbf{h}}_{r,t}$ and $\mathbf{e}_g=\mathbf{g}_{r,t}-\hat{\mathbf{g}}_{r,t}$ be the estimation error of channel $\mathbf{h}_{r,t}$ and $\mathbf{g}_{r,t}$. Note that $\hat{\mathbf{h}}_{r,t}$ and $\hat{\mathbf{g}}_{r,t}$ are the MMSE estimates of $\mathbf{h}_{r,t}$ and $\mathbf{g}_{r,t}$ obtained in the previous IDD iteration. Using the orthogonality principle, the error covariance matrix becomes [17]

$$\begin{aligned}
&\mathrm{Cov}\left(\begin{bmatrix}\mathbf{e}_h\\ \mathbf{e}_g\end{bmatrix},\begin{bmatrix}\mathbf{e}_h\\ \mathbf{e}_g\end{bmatrix}\right)\\
&=\begin{bmatrix}C_{e_h,e_h} & C_{e_h,e_g}\\ C_{e_g,e_h} & C_{e_g,e_g}\end{bmatrix}\\
&=\mathrm{Cov}\left(\begin{bmatrix}\mathbf{h}_{r,t}\\ \mathbf{g}_{r,t}\end{bmatrix}-\begin{bmatrix}\hat{\mathbf{h}}_{r,t}\\ \hat{\mathbf{g}}_{r,t}\end{bmatrix},\begin{bmatrix}\mathbf{h}_{r,t}\\ \mathbf{g}_{r,t}\end{bmatrix}-\begin{bmatrix}\hat{\mathbf{h}}_{r,t}\\ \hat{\mathbf{g}}_{r,t}\end{bmatrix}\right)\\
&=\begin{bmatrix}C_{h_t,h_t} & C_{h_t,g_t}\\ C_{g_t,h_t} & C_{g_t,g_t}\end{bmatrix}-\Omega_{r,t}\Sigma_{r,t}^{-1}\Omega_{r,t}^H
\end{aligned}\tag{52}$$

where $\Omega_{r,t}$ and $\Sigma_{r,t}$ are the quantities obtained from (25) and (26) in the previous IDD iteration. Note that C_{e_g,e_g} is obtained from the submatrix of the error covariance matrix.

Appendix C
Derivation of (38)

From (35), we have

$$\phi(n)=\mathrm{tr}\left(\mathrm{Cov}\left(\begin{bmatrix}\mathbf{h}_{r,t}\\ g_n^{(r,t)}\end{bmatrix},\begin{bmatrix}\mathbf{h}_{r,t}\\ g_n^{(r,t)}\end{bmatrix}\right)-\widetilde{\Omega}_{r,t}\widetilde{\Sigma}_{r,t}^{-1}\widetilde{\Omega}_{r,t}^H\right)\tag{53}$$

where

$$\begin{aligned}
\widetilde{\Omega}_{r,t}&=\mathrm{Cov}\left(\begin{bmatrix}\mathbf{h}_{r,t}\\ g_n^{(r,t)}\end{bmatrix},\begin{bmatrix}\mathbf{z}_r\\ y_n^{(r)}\end{bmatrix}\right)\\
&=\begin{bmatrix}\sqrt{\eta_{p,t}}C_{h_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{h_t,g_t}\left(\overline{d}_n^{(t)}\right)^*\\ \sqrt{\eta_{p,t}}C_{g_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{g_t,g_t}\left(\overline{d}_n^{(t)}\right)^*\end{bmatrix}
\end{aligned}\tag{54}$$

and

$$\begin{aligned}
\widetilde{\Sigma}_{r,t}&=\mathrm{Cov}\left(\begin{bmatrix}\mathbf{z}_r\\ y_n^{(r)}\end{bmatrix},\begin{bmatrix}\mathbf{z}_r\\ y_n^{(r)}\end{bmatrix}\right)\\
&=\begin{bmatrix}\eta_{p,t}\mathbf{P}_tC_{h_t,h_t}\mathbf{P}_t^H+I & \sqrt{\eta_{p,t}\eta_{d,t}}\mathbf{P}_tC_{h_t,g_t}\left(\overline{d}_n^{(t)}\right)^*\\ \sqrt{\eta_{p,t}\eta_{d,t}}\overline{d}_n^{(t)}\mathbf{P}_t^H & \sum_{i=0}^{T-1}\eta_{d,i}E[|d_n^{(i)}|^2]+1\end{bmatrix}.
\end{aligned}\tag{55}$$

For convenience, we let

$$\begin{aligned}
&\begin{bmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{bmatrix}\\
&=\begin{bmatrix}\sqrt{\eta_{p,t}}C_{h_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{h_t,g_t}\left(\overline{d}_n^{(t)}\right)^*\\ \sqrt{\eta_{p,t}}C_{g_t,h_t}\mathbf{P}_t^H & \sqrt{\eta_{d,t}}C_{g_t,g_t}\left(\overline{d}_n^{(t)}\right)^*\end{bmatrix},
\end{aligned}\tag{56}$$

$$\begin{aligned}
&\begin{bmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{21} & \Gamma_{22}\end{bmatrix}\\
&=\begin{bmatrix}\eta_{p,t}\mathbf{P}_tC_{h_t,h_t}\mathbf{P}_t^H+I & \sqrt{\eta_{p,t}\eta_{d,t}}\mathbf{P}_tC_{h_t,g_t}\left(\overline{d}_n^{(t)}\right)^*\\ \sqrt{\eta_{d,t}\eta_{p,t}}\overline{d}_n^{(t)}C_{g_t,h_t}\mathbf{P}_t^H & \sum_{i=0}^{T-1}\eta_{d,i}E[|d_n^{(i)}|^2]+1\end{bmatrix},
\end{aligned}\tag{57}$$

and

$$\begin{bmatrix}\Theta_{11} & \Theta_{12}\\ \Theta_{21} & \Theta_{22}\end{bmatrix}=\begin{bmatrix}\Gamma_{11} & \Gamma_{12}\\ \Gamma_{21} & \Gamma_{22}\end{bmatrix}^{-1}.\tag{58}$$

Then, (53) is

$$\begin{aligned}
\phi(n)=&\mathrm{tr}\Bigg(\begin{bmatrix}C_{h_t,h_t} & C_{h_t,g_t}\\ C_{g_t,h_t} & C_{g_t,g_t}\end{bmatrix}\\
&-\begin{bmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{bmatrix}\begin{bmatrix}\Theta_{11} & \Theta_{12}\\ \Theta_{21} & \Theta_{22}\end{bmatrix}\begin{bmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{bmatrix}^H\Bigg)\\
=&\mathrm{tr}(C_{g_t,g_t}-\Omega_{21}\Theta_{11}\Omega_{21}^H-\Omega_{22}\Theta_{21}\Omega_{21}^H\\
&-\Omega_{21}\Theta_{12}\Omega_{22}^H-\Omega_{22}\Theta_{22}\Omega_{22}^H)\\
&+\mathrm{tr}(C_{h_t,h_t}-\Omega_{11}\Theta_{11}\Omega_{11}^H-\Omega_{12}\Theta_{21}\Omega_{11}^H\\
&-\Omega_{11}\Theta_{12}\Omega_{12}^H-\Omega_{12}\Theta_{22}\Omega_{12}^H).
\end{aligned}\tag{59}$$

Denoting the first and second terms in the right-hand side of (59) as $\phi_1(n)$ and $\phi_2(n)$, respectively, then $\phi(n)=\phi_1(n)+\phi_2(n)$ where

$$\begin{aligned}
\phi_1(n)=&\mathrm{tr}(C_{g_t,g_t}-\Omega_{21}\Theta_{11}\Omega_{21}^H-\Omega_{22}\Theta_{21}\Omega_{21}^H\\
&-\Omega_{21}\Theta_{12}\Omega_{22}^H-\Omega_{22}\Theta_{22}\Omega_{22}^H)\\
=&\mathrm{tr}\Big[C_{g_t,g_t}-\eta_{p,t}C_{g_t,h_t}\mathbf{P}_t^H\Theta_{11}\mathbf{P}_tC_{h_t,g_t}\\
&-\sqrt{\eta_{p,t}\eta_{d,t}}C_{g_t,h_t}\mathbf{P}_t^H\Theta_{12}C_{g_t,g_t}\overline{d}_n^{(t)}\\
&-\sqrt{\eta_{p,t}\eta_{d,t}}C_{g_t,g_t}\Theta_{21}\mathbf{P}_tC_{h_t,g_t}\left(\overline{d}_n^{(t)}\right)^*\\
&-\eta_{d,t}|\overline{d}_n^{(t)}|^2C_{g_t,g_t}\Theta_{22}C_{g_t,g_t}\Big]
\end{aligned}\tag{60}$$

$$\begin{aligned}
\phi_2(n)=&\mathrm{tr}(C_{h_t,h_t}-\Omega_{11}\Theta_{11}\Omega_{11}^H-\Omega_{12}\Theta_{21}\Omega_{11}^H\\
&-\Omega_{11}\Theta_{12}\Omega_{12}^H-\Omega_{12}\Theta_{22}\Omega_{12}^H)\\
=&\mathrm{tr}\Big[C_{h_t,h_t}-\eta_{p,t}C_{h_t,h_t}\mathbf{P}_t^H\Theta_{11}\mathbf{P}_tC_{h_t,h_t}\\
&-\sqrt{\eta_{p,t}\eta_{d,t}}C_{h_t,h_t}\mathbf{P}_t^H\Theta_{12}C_{g_t,h_t}\overline{d}_n^{(t)}\\
&-\sqrt{\eta_{p,t}\eta_{d,t}}C_{h_t,g_t}\Theta_{21}\mathbf{P}_tC_{h_t,h_t}\left(\overline{d}_n^{(t)}\right)^*\\
&-\eta_{d,t}|\overline{d}_n^{(t)}|^2C_{h_t,g_t}\Theta_{22}C_{g_t,h_t}\Big].
\end{aligned}\tag{61}$$

Using the inversion formula of partitioned matrices [17][3], we have

$$\begin{aligned}
\Theta_{11}=&\left(\Gamma_{11}-\frac{\eta_{p,t}\eta_{d,t}|\overline{d}_n^{(t)}|^2\mathbf{P}_tC_{h_t,g_t}C_{g_t,h_t}\mathbf{P}_t^H}{\Gamma_{22}}\right)^{-1}\\
=&\Gamma_{11}^{-1}+\frac{\eta_{p,t}\eta_{d,t}|\overline{d}_n^{(t)}|^2\Gamma_{11}^{-1}\mathbf{P}_tC_{h_t,g_t}C_{g_t,h_t}\mathbf{P}_t^H\Gamma_{11}^{-1}}{\Gamma_{22}-\eta_{p,t}\eta_{d,t}|\overline{d}_n^{(t)}|^2C_{g_t,h_t}\mathbf{P}_t^H\Gamma_{11}^{-1}\mathbf{P}_tC_{h_t,g_t}}
\end{aligned}\tag{62}$$

[3]Let $\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}^{-1}$, $\det(A_{11})\neq 0$, and $\det(A_{22})\neq 0$, then $B_{11}=(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}$, $B_{12}=-A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}$, $B_{21}=-A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}$, and $B_{22}=(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}$.

$$\Theta_{12} = -\frac{1}{\Gamma_{22}}\Bigg(\sqrt{\eta_{p,t}\eta_{d,t}}\left(\overline{d}_n^{(t)}\right)^* \Gamma_{11}^{-1}\mathbf{P}_t C_{h_t,g_t} + (\eta_{p,t}\eta_{d,t})^{3/2}|\overline{d}_n^{(t)}|^2\left(\overline{d}_n^{(t)}\right)^* \frac{\Gamma_{11}^{-1}\mathbf{P}_t C_{h_t,g_t} C_{g_t,h_t}\mathbf{P}_t^H\Gamma_{11}^{-1}\mathbf{P}_t C_{h_t,g_t}}{\Gamma_{22}-\eta_{p,t}\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\mathbf{P}_t^H\Gamma_{11}^{-1}\mathbf{P}_t C_{h_t,g_t}}\Bigg) \quad (63)$$

$$\Theta_{21} = -\frac{\sqrt{\eta_{p,t}\eta_{d,t}}\overline{d}_n^{(t)} C_{g_t,h_t}\mathbf{P}_t^H\Gamma_{11}^{-1}}{\Gamma_{22}-\eta_{p,t}\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\mathbf{P}_t^H\Gamma_{11}^{-1}\mathbf{P}_t C_{h_t,g_t}} \quad (64)$$

$$\Theta_{22} = \frac{1}{\Gamma_{22}-\eta_{p,t}\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\mathbf{P}_t^H\Gamma_{11}^{-1}\mathbf{P}_t C_{h_t,g_t}} \quad (65)$$

where (62) is from the matrix inversion lemma ($(A - BD^{-1}C) = A^{-1} + A^{-1}B(D - CA^{-1}B)CA^{-1}$). Using (62)–(65) and also noting that

$$\eta_{p,t}\mathbf{P}_t^H\Gamma_{11}^{-1}\mathbf{P}_t = \Sigma^{-1} \quad (66)$$

$$\Sigma \triangleq C_{h_t,h_t} + \frac{1}{\eta_{p,t}}I, \quad (67)$$

we have

$$\begin{aligned}\phi_1(n) =& \operatorname{tr}\Bigg[C_{g_t,g_t} - C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t} \\ &- \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}} \\ &+ \frac{1}{\Gamma_{22}}\left(\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,g_t}\right) \\ &+ \frac{1}{\Gamma_{22}}\frac{\eta_{d,t}^2|\overline{d}_n^{(t)}|^4 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}} \\ &+ \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,g_t}C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}} \\ &- \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,g_t}C_{g_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}\Bigg]. \end{aligned} \quad (68)$$

Note that the fourth term and fifth term in the right-hand side of (68) are reduced to

$$\begin{aligned}&\operatorname{tr}\Bigg[\frac{1}{\Gamma_{22}}\left(\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,g_t}\right) \\ &\quad+\frac{1}{\Gamma_{22}}\frac{\eta_{d,t}^2|\overline{d}_n^{(t)}|^4 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}\Bigg] \\ &= \operatorname{tr}\left[\frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}\right]. \end{aligned} \quad (69)$$

After some manipulations, we further have

$$\begin{aligned}\phi_1(n) =& \operatorname{tr}\Bigg[C_{g_t,g_t} - C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t} \\ &- \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}} \\ &- \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,g_t}C_{g_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}} \\ &+ \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}C_{g_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}} \\ &+ \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,g_t}C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}}\Bigg] \end{aligned} \quad (70)$$

$$\begin{aligned}=& C_{g_t,g_t} - C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t} \\ &- \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2\left|C_{g_t,g_t} - C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}\right|^2}{\Gamma_{22}-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}} \end{aligned} \quad (71)$$

$$\begin{aligned}=& \frac{\left(1-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,g_t}/\Gamma_{22}\right)}{1-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}/\Gamma_{22}} \\ &\cdot\left(C_{g_t,g_t} - C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}\right) \end{aligned} \quad (72)$$

where (71) is because $\operatorname{tr}(a) = a$ for scalar a. In a similar way, we have

$$\begin{aligned}\phi_2(n) =& \operatorname{tr}\left[\frac{1}{\eta_{p,t}}C_{h_t,h_t}\left(C_{h_t,h_t}+\frac{1}{\eta_{p,t}}I\right)^{-1}\right] \\ &- \frac{\frac{1}{\eta_{p,t}^2}\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\left(C_{h_t,h_t}+\frac{1}{\eta_{p,t}}I\right)^{-2}C_{h_t,g_t}/\Gamma_{22}}{1-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\left(C_{h_t,h_t}+\frac{1}{\eta_{p,t}}I\right)^{-1}C_{h_t,g_t}/\Gamma_{22}}. \end{aligned} \quad (73)$$

Finally, by adding (71) and (73), we have

$$\begin{aligned}\phi(n) =& \operatorname{tr}\left(\frac{1}{\eta_{p,t}}C_{h_t,h_t}\Sigma^{-1}\right) \\ &- \frac{\eta_{d,t}|\overline{d}_n^{(t)}|^2\frac{1}{\eta_{p,t}^2}C_{g_t,h_t}\Sigma^{-2}C_{h_t,g_t}/\Gamma_{22}}{1-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}/\Gamma_{22}} \\ &+ \frac{\left(1-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,g_t}/\Gamma_{22}\right)\left(C_{g_t,g_t}-C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}\right)}{1-\eta_{d,t}|\overline{d}_n^{(t)}|^2 C_{g_t,h_t}\Sigma^{-1}C_{h_t,g_t}/\Gamma_{22}}. \end{aligned} \quad (74)$$

Note that the first term in the right-hand side of (74) is a function of C_{h_t,h_t} and hence independent of data tones.

REFERENCES

[1] J. Gao and H. Liu, "Decision-directed estimation of MIMO time-varying Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1412–1417, Jul. 2005.

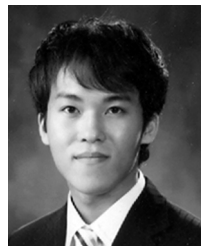
[2] C. Georghiades and J. Han, "Sequence estimation in the presence of random parameters via EM algorithm," *IEEE Trans. Commun.*, vol. 45, pp. 300–308, Mar. 1997.

[3] H. Zamiri-Jafarian and S. Pasupathy, "EM-based recursive estimation of channel parameters," *IEEE Trans. Commun.*, vol. 47, pp. 1297–1302, Sep. 1999.

[4] C. H. Aldana, E. de Carvalho, and J. M. Cioffi, "Channel estimation for multicarrier multiple input single output systems using the EM algorithm," *IEEE Trans. Signal Process.*, vol. 51, pp. 3280–3292, Dec. 2003.

[5] J. Ylioinas and M. Juntti, "Iterative joint detection, decoding, and channel estimation in Turbo-Coded MIMO-OFDM," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1784–1796, May 2009.

[6] S. Song, A. Singer, and K. Sung, "Soft input channel estimation for turbo equalization," *IEEE Trans. Signal Process.*, vol. 52, pp. 2885–2894, Oct. 2004.
[7] R. Otnes and M. Tuchler, "Soft iterative channel estimation for turbo equalization: Comparison of channel estimation algorithms," in *Proc. Int. Conf. Commun. Syst.*, Singapore, Nov. 2002, pp. 72–76.
[8] D. Yoon and J. Moon, "Soft-decision-driven channel estimation for pipelined turbo receivers," *IEEE Trans. Commun.*, vol. 59, pp. 2141–2151, Aug. 2011.
[9] M. Munster and L. Hanzo, "Parallel-interference-cancellation-assisted decision-directed channel estimation for OFDM systems using multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2148–2162, Sep. 2005.
[10] M. Jiang, J. Akhtman, and L. Hanzo, "Iterative joint channel estimation and multi-user detection for multiple-antenna aided OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2904–2914, Aug. 2007.
[11] Y. Liu and S. Sezginer, "Iterative compensated MMSE channel estimation in LTE systems," in *Proc. IEEE Int. Conf. Commun.*, 2012, pp. 4862–4866.
[12] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (Release 10)*, 3GPP TS 36.211 V10.7.0 (2013-03).
[13] B. Hochwald and S. T. Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.
[14] J. Choi, B. Shim, and A. Singer, "Efficient soft-input soft-output tree detection via an improved path metric," *IEEE Trans. Inf. Theory*, vol. 58, pp. 1518–1533, Mar. 2012.
[15] J. Lee, B. Shim, and I. Kang, "Soft-Input soft-Output list sphere detection with a probabilistic radius tightening," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2848–2857, Aug. 2012.
[16] W. C. Jakes, *Microwave Mobile Communications*. New York, NY, USA: Wiley, 1975.
[17] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
[18] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 4th ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 2013.
[19] *Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception (Release 10)*, 3GPP TS 36.101 V10.10.0 (2010-03).
[20] M. K. Ozdemir and H. Arslan, "Channel estimation for wireless OFDM systems," *IEEE Comm. Surveys Tuts.*, vol. 9, pp. 18–48, 2007.
[21] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: Theory and experiments," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2538–2546, Oct. 2002.
[22] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, pp. 284–287, Mar. 1974.
[23] S. Park, J. Choi, J. Seol, and B. Shim, "Virtual pilot signal for massive MIMO-OFDM systems," in *Proc. Inf. Theory Appl. Workshop (ITA)*, 2014.



Sunho Park (S'11) received the B.S. and M.S. degrees from the School of Information and Communication from Korea University, Seoul, in 2008 and 2010, respectively, where he is currently working toward the Ph.D. degree.

His research interests include wireless communications and signal processing.

Byonghyo Shim (SM'09) received the B.S. and M.S. degrees in control and instrumentation engineering from Seoul National University, Korea, in 1995 and 1997, respectively. He received the M.S. degree in mathematics and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign (UIUC), USA, in 2004 and 2005, respectively.

From 1997 and 2000, he was with the Department of Electronics Engineering, Korean Air Force Academy as an Officer (First Lieutenant) and an Academic Full-time Instructor. From 2005 to 2007, he was with Qualcomm Inc., San Diego, CA, USA, as a Staff Engineer. From 2007 to 2014, he was with the School of Information and Communication, Korea University, Seoul, as an Associate Professor. Since September 2014, he has been with the Department of Electrical and Computer Engineering, Seoul National University, where he is presently an Associate Professor. His research interests include wireless communications, statistical signal processing, estimation and detection, compressive sensing, and information theory.

Dr. Shim was the recipient of the 2005 M. E. Van Valkenburg Research Award from the Electrical and Computer Engineering Department of the University of Illinois and 2010 Hadong Young Engineer Award from IEIE. He is currently an Associate Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS, *Journal of Communications and Networks*, and a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC).

Jun Won Choi (M'08) received the B.S. and M.S. degrees in Electrical and Computer Engineering, Seoul National University, and earned the Ph.D. degree in Electrical and Computer Engineering, University of Illinois at Urbana- Champaign, respectively. In 2010, he joined Qualcomm, San Diego, USA. In Qualcomm, he participated in wireless communication system/algorithm design for developing LTE and LTEAdvanced modem chipsets. Since 2013, he has been a faculty of Department of Electrical Engineering in Hanyang University, Seoul, Korea. His research area includes signal processing, wireless communication, optimization theory, and machine learning. He is a member of IEEE and IEEE Signal Processing Society (SPS).