Blind Recursive Subspace-Based Identification of Time-Varying Wideband MIMO Channels

Chao-Cheng Tu and Benoît Champagne, *Senior Member, IEEE*

Abstract—We present a blind recursive algorithm for tracking rapidly time-varying wireless channels in precoded multiple-input–multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) systems. Subspace-based tracking is normally considered for slowly time-varying channels only. Due to the frequency correlation of the wireless channels, the proposed scheme can collect data not only from the time but from the frequency domain as well to speed up the update of the required second-order statistics. After each such update, the subspace information is recomputed using the orthogonal iteration, and then, a new channel estimate is obtained. We also investigate choices of precoder in terms of the tradeoff between the symbol recovery capability and the channel estimation performance and demonstrate the convergence properties of our approach. The proposed algorithm is evaluated in a Third-Generation Partnership Project (3GPP) Spatial Channel Model suburban macro scenario, in which a mobile station is allowed to move in any direction with a speed up to 100 km/h, corresponding to a maximum Doppler shift of about 230 Hz in this case. Numerical experiments show that the normalized mean square error of the channel estimates converges to a level of -30 dB within less than five OFDM symbols when the signal-to-noise ratio (SNR) (per symbol) is ≥ 20 dB.

Index Terms—Blind, multiple-input–multiple-output (MIMO)-orthogonal frequency-division multiplexing (OFDM), recursive, subspace, time-varying.

I. INTRODUCTION

TRACKING time-varying (TV) channels with large Doppler spreads is a critical task, regardless of whether a nonblind or a blind approach is used [1]. In general, with a nonblind approach, training must be applied more frequently, because the channel estimates become obsolete shortly after the training period ends. On the contrary, a blind approach eliminates the need of large amount of training data and, therefore, is favored if complexity is not the main concern.

TV channels are mainly tracked using the following two categories of *blind* approaches: 1) batch processing to estimate the unknown parameters of an underlying TV channel model and 2) an adaptive processing algorithm that is sufficiently

fast to track the channel variations. Among various blind approaches in the first category, a basis expansion model has been proposed to convert a TV single-input–single-output (SISO) channel into a time-invariant (TI) single-input–multiple-output (SIMO) channel, followed by a standard second-order statistics (SOS)-based subspace method for blind channel estimation [2]. The idea of basis expansion was further extended for a TV-SIMO channel [3]–[5], and a generalized orthogonal frequency-division multiplexing (OFDM) system over a TV-SISO channel [6]. Similarly, an interpolation model was proposed to convert a TV-SISO channel into fixed parameters, for application to code-division multiple access (CDMA) systems [7].

Recently, there has been much research interest on adaptive algorithms in the second category. A zero-padding SISO-OFDM system using either the recursive least squares (RLS) or the least mean square (LMS) method for blind adaptive channel estimation was considered in [8]. It was reported that, for an inverse fast Fourier transform (IFFT) size of 64 and a padding length of 16, the relative channel estimation error converges to -27 dB in 500 symbols when the maximum Doppler shift is limited to 100 Hz and the signal-to-noise ratio (SNR) is 20 dB. By properly choosing the so-called repetition index, a cyclic prefixing (CP) SISO-OFDM system was also proposed in [9], where it was reported that, for a 64-point IFFT with a CP length of 16, the bit error rate (BER) can reach a level of 10^{-2} within 12 received blocks when the maximum Doppler shift is 50 Hz and the SNR is ≥ 25 dB. Recently, an adaptive filter using an autoregressive moving average process has been proposed for the blind deconvolution of multiple-input–multiple-output (MIMO) channels [10]. Alternative approaches to adaptive channel estimation based on reduced-rank processing have been considered in [11]–[13], where the main focus is on CDMA systems.

Although the aforementioned adaptive approaches offer interesting capabilities in tracking TV channels with high spectrum efficiency, they may not be adequate for applications in emerging and future generations of broadband mobile wireless systems, in which there is a need to provide high-rate transmission, such as a real-time video stream, between a user terminal and an access node whose relative positions rapidly vary over time. For example, the Third-Generation Partnership Project Long Term Evolution (3GPP LTE) specifications already call for high-performance broadband transmission with mobile speed up to 120 km/h, corresponding to a maximum Doppler shift of 220 Hz, and additional provision to support much higher speeds up to 350 km/h for high-speed trains [14]. The requirements set forth for International Mobile Telecommunications-Advanced (IMT-Advanced) are even more exacting.

Fig. 1. Precoded MIMO-OFDM system model.

Based on these considerations, there is a need to extend the range of the application of adaptive channel tracking to even faster TV channels. In addition, achieving this goal should not result in the loss of bandwidth efficiency or place unwanted restrictions on the number of transmit or receive antennas. We note that there currently exist algorithms for channel estimation in mobile MIMO-OFDM systems with large Doppler shifts [15]–[19]; however, they all require the use of extensive preambles or training sequences in the time–frequency domain.

In this paper, to avoid the need of training sequence in the estimation of TV channels, the blind subspace-based estimation method using frequency correlation in MIMO-OFDM systems, which was originally proposed for TI channels in [20], [21], is extended to the case of TV scenarios. Although this approach in the TI case requires a larger dimension of the ambiguity matrix and high-complexity singular value decomposition, these limitations are overcome in the TV case by using a precoder at the transmitter side and a computationally efficient orthogonal iteration for subspace tracking at the receiver side, respectively. The resulting approach can track TV MIMO wireless channels with large Doppler spread, which may change at each OFDM symbol time. In addition, it offers flexibility in choosing the number of transmit and receive antennas, enabling enhanced bandwidth efficiency through spatial processing. For a 256-point IFFT, the proposed algorithm is evaluated in a 3GPP Spatial Channel Model (SCM) suburban macro scenario. Our simulation results show that the normalized mean square error (NMSE) converges to less than -30 dB within five OFDM symbols, and the BER can reach a level of 10^{-2} when the maximum Doppler shift is about 230 Hz and the symbol SNR is ≥ 20 dB, which outperforms the approach in [8] and [9] in terms of estimation performance.

This paper is organized as follows. Section II is devoted to problem formulation, including a description of the MIMO-OFDM system model under consideration. The precoded blind subspace-based approach is introduced in Section III, and its proposed extension into an adaptive channel tracking algorithm is developed in Section IV. The precoder design is studied in Section V, whereas the numerical experiments and discussions

are presented in Section VI. Finally, conclusions are drawn in the last section.

The following notations are used. $\mathbf{X} = \mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix whose main diagonal is constructed from the entries of vector $\mathbf{x}$. The range of $\mathbf{A} \in \mathbb{C}^{m \times n}$ is defined by $\mathfrak{R}(\mathbf{A}) \stackrel{\text{def}}{=} \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{C}^{n \times 1}\}$. $\|\mathbf{x}\|_p$ and $\|\mathbf{X}\|_p$ represent the p-norms of a vector $\mathbf{x}$ and a matrix $\mathbf{X}$, respectively. $\mathfrak{D}_r(\mathbf{X})$ denotes the subspace spanned by the eigenvectors corresponding to the r largest eigenvalues of $\mathbf{X}$. $\lambda_r(\mathbf{X})$ represents the rth largest eigenvalue of the matrix $\mathbf{X}$. $\mathbf{I}_n$ is used to represent an identity matrix of order n. The symbols $\otimes$, $\oslash$, and $\odot$ represent the Kronecker product, the element-wise division, and the Hadamard product, respectively.

II. PROBLEM FORMULATION

Conventional blind subspace-based estimators are generally not favored when a fast TV channel is considered, because there may not be sufficient data samples to estimate the required statistics. The situation is even worse in the context of MIMO-OFDM systems, where a large dimension of the correlation matrix (up to thousands) is normally required. In [20] and [21], we have shown that the TI requirement in blind subspace-based estimation for MIMO-OFDM systems can significantly be relaxed by making use of subcarrier grouping to exploit frequency correlation among adjacent subcarriers. However, the ensuing reduction in the time averaging period comes at the price of a higher dimension of the ambiguity matrix. Here, to overcome this problem, we consider a precoded MIMO-OFDM system, as described below.

The system under consideration employs N_C subcarriers, N_T transmit antennas, and N_R receive antennas as per the block diagram shown in Fig. 1. To exploit the frequency correlation through the concept of subcarrier grouping, we assume that the frequency span of P adjacent subcarriers reside in the coherence bandwidth of the wireless channel, which are defined as the range of frequencies over which the frequency response matrix of the MIMO channel does not appreciably change [22]. We partition the subcarrier index set $\Omega \stackrel{\text{def}}{=} \{0, 1, \ldots, N_C - 1\}$
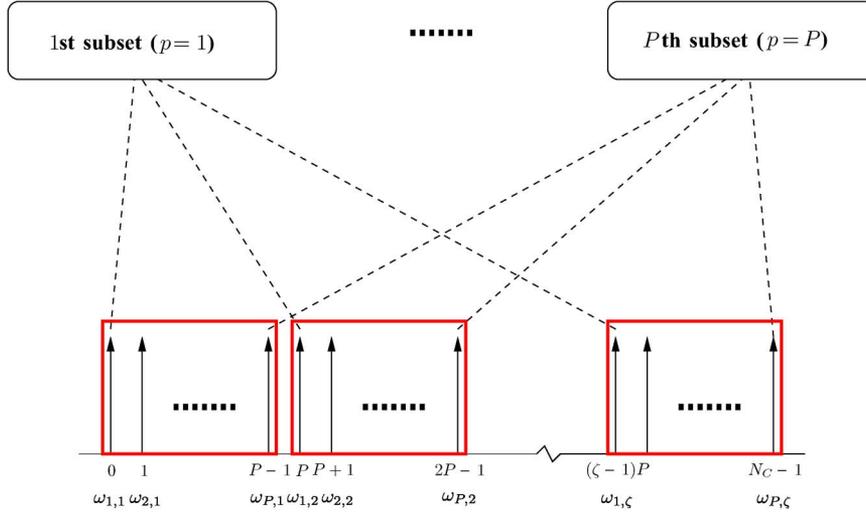
Fig. 2.   Schematic of the partitioning of the subcarrier index set $\Omega = \{0, 1, \ldots, N_C - 1\}$ into P disjoint subsets, i.e., $\Omega_p = \{\omega_{p,1}, \omega_{p,2}, \ldots, \omega_{p,\zeta}\}$, $p = 1, 2, \ldots, P$.

into P disjoint subsets (assuming $N_C/P = \zeta \in \mathbb{Z}^+$), with the pth subset denoted as $\Omega_p \overset{\text{def}}{=} \{\omega_{p,1}, \omega_{p,2}, \ldots, \omega_{p,\zeta}\}$, where $\omega_{p,i} \overset{\text{def}}{=} p - 1 + (i-1)P$ for $i = 1, 2, \ldots, \zeta$ and $p = 1, 2, \ldots, P$ (see Fig. 2). Let $\mathbf{x}_p^m \overset{\text{def}}{=} [\mathbf{x}_{1,p}^{m\ T} \ \mathbf{x}_{2,p}^{m\ T} \ \cdots \ \mathbf{x}_{N_T,p}^{m\ T}]^T$, where

$$\mathbf{x}_{j,p}^m \overset{\text{def}}{=} \left[x_j^m[\omega_{p,1}] \ x_j^m[\omega_{p,2}] \ \cdots \ x_j^m[\omega_{p,\zeta}] \right]^T \quad (1)$$

with $x_j^m[k]$ denoting the signal that was transmitted at the kth subcarrier, the jth transmit antenna, and mth OFDM symbol. In addition, let $\mathbf{y}_p^m \overset{\text{def}}{=} [\mathbf{y}_{1,p}^{m\ T} \ \mathbf{y}_{2,p}^{m\ T} \ \cdots \ \mathbf{y}_{N_R,p}^{m\ T}]^T$ and $\mathbf{n}_p^m \overset{\text{def}}{=} [\mathbf{n}_{1,p}^{m\ T} \ \mathbf{n}_{2,p}^{m\ T} \ \cdots \ \mathbf{n}_{N_R,p}^{m\ T}]^T$, where $\mathbf{y}_{i,p}^m \overset{\text{def}}{=} [y_i^m[\omega_{p,1}] \ y_i^m[\omega_{p,2}] \ \cdots \ y_i^m[\omega_{p,\zeta}]]^T$, $\mathbf{n}_{i,p}^m \overset{\text{def}}{=} [n_i^m[\omega_{p,1}] \ n_i^m[\omega_{p,2}] \ \cdots \ n_i^m[\omega_{p,\zeta}]]^T$, with $y_i^m[k]$ and $n_i^m[k]$ denoting the signal and noise received at the kth subcarrier, the ith received antenna, and the mth OFDM symbol, respectively. We have the following assumptions: 1) The length of the CP that was appended to each OFDM symbol is longer than the maximum excess delay of the channel, and 2) the average power of the transmit symbol alphabet is normalized to unity, i.e., $E[|x_j^m[k]|^2] = 1$.

Suppose that each input vector $\mathbf{x}_{j,p}^m$ in (1) is left-multiplied by a nonredundant precoding matrix $\mathbf{\Psi} \in \mathbb{C}^{\zeta \times \zeta}$ (whose choice is considered in Section V). Then, the input–output relationship for the pth frequency subset can be written as

$$\mathbf{y}_p^m = \mathcal{H}_p^m(\mathbf{I}_{N_T} \otimes \mathbf{\Psi})\mathbf{x}_p^m + \mathbf{n}_p^m \quad (2)$$

with the channel matrix $\mathcal{H}_p^m$ defined as

$$\mathcal{H}_p^m = \begin{bmatrix} \mathcal{H}_{1,1,p}^m & \cdots & \mathcal{H}_{1,N_T,p}^m \\ \vdots & \ddots & \vdots \\ \mathcal{H}_{N_R,1,p}^m & \cdots & \mathcal{H}_{N_R,N_T,p}^m \end{bmatrix}. \quad (3)$$

In (3), $\mathcal{H}_{i,j,p}^m = \text{diag}(\mathbf{h}_{i,j,p}^m)$, where $\mathbf{h}_{i,j,p}^m \overset{\text{def}}{=} [h_{i,j}^m[\omega_{p,1}], h_{i,j}^m[\omega_{p,2}], \ldots, h_{i,j}^m[\omega_{p,\zeta}]]^T$, and $h_{i,j}^m[k]$ denotes the equivalent frequency response between the ith receive and the jth transmit antennas over the kth subcarrier and the mth OFDM symbol.

In this paper, because the frequency span of P adjacent subcarriers is supposed to reside in the coherence bandwidth, we shall assume that the variations in the channel matrices $\mathcal{H}_p^m$ across these P subcarriers are negligible, although large changes in the channel responses over a wider bandwidth are possible [21]. Accordingly, we can define a new representative channel matrix, $\mathcal{H}^m \overset{\text{def}}{=} \mathcal{H}_1^m \cong \mathcal{H}_2^m \cong \cdots \cong \mathcal{H}_P^m$, and drop the index p for all channel-related quantities, including $\mathcal{H}_{i,j,p}^m$ and $\mathbf{h}_{i,j,p}^m$. Note that the aforementioned assumption of near equality is exploited only to derive our new algorithm; however, the new algorithm will be tested over realistic channel models for which equality does not hold. Additional details are given in Section VI.

In this paper, our first interest lies in the blind estimation and tracking of rapidly TV-MIMO channels with normalized Doppler frequencies that may reach significant values, e.g., up to 2% or more [23], and for which the channel matrix $\mathcal{H}^m$ is allowed to change at each OFDM symbol time. To this end, we seek to develop a blind subspace-based estimator $\hat{\mathcal{H}}^m$, which is a function of the observed data up to the current symbol time m, i.e., $\{\mathbf{y}_p^l, p = 1, 2, \ldots, P; l = 1, 2, \ldots, m\}$, and can recursively be updated in a computationally efficient manner. That is, we seek to develop a practical updating algorithm $\phi(\cdot)$ in which the channel estimate at the mth symbol time can be represented as $\hat{\mathcal{H}}^m = \phi(\hat{\mathcal{H}}^{m-1}, \{\mathbf{y}_p^m\}_{p=1}^P)$. In addition, because the precoder is placed at the transmitter side without having any feedback of channel knowledge from the receiver side, our second interest lies in determining a suitable precoder scheme to further enhance the estimation performance.

III. PRECODED BLIND SUBSPACE-BASED APPROACH

In this section, we first introduce a subspace-based channel estimation approach that exploits the frequency correlation among adjacent subcarriers in the precoded MIMO-OFDM system. Our presentation initially assumes a block processing framework; the proposed approach will then be extended to the recursive estimation and tracking of TV channels in Section IV.

For simplicity, let us temporarily drop the time index m from all the channel-related coefficients. Based on (2) and assuming that channel variations over P adjacent subcarriers are negligible, the correlation matrix $\mathbf{R_y} \overset{\text{def}}{=} E[\mathbf{y}_1 \mathbf{y}_1^H] \cong \cdots \cong E[\mathbf{y}_P \mathbf{y}_P^H]$, which is of size $\zeta N_R \times \zeta N_R$, can be written as

$$\mathbf{R_y} = \mathcal{H}(\mathbf{I}_{N_T} \otimes \boldsymbol{\Psi}\boldsymbol{\Psi}^H)\mathcal{H}^H + \sigma_n^2 \mathbf{I}_{\zeta N_R} \qquad (4)$$

where we have assumed that $\mathbf{n}_p$ and $\mathbf{x}_p$ are uncorrelated, $E[\mathbf{n}_p \mathbf{n}_p^H] = \sigma_n^2 \mathbf{I}_{\zeta N_R}$, and $E[\mathbf{x}_p \mathbf{x}_p^H] = \mathbf{I}_{\zeta N_T}$. Similar to [24], by partitioning $\mathbf{R_y}$ into submatrices of size $\zeta \times \zeta$, we may express its (u, v)th submatrix as

$$\mathbf{R}_{\mathbf{y},uv} = \sum_{j=1}^{N_T} \mathcal{H}_{u,j} \boldsymbol{\Psi}\boldsymbol{\Psi}^H \mathcal{H}_{v,j}^H + \delta_{uv}\sigma_n^2 \mathbf{I}_\zeta$$
$$= \left(\sum_{j=1}^{N_T} \mathbf{h}_{u,j}\mathbf{h}_{v,j}^H \right) \odot (\boldsymbol{\Psi}\boldsymbol{\Psi}^H) + \delta_{uv}\sigma_n^2 \mathbf{I}_\zeta \qquad (5)$$

where $u, v \in \{1, 2, \ldots, N_R\}$, $\delta_{uv} = 1$ if $u = v$; otherwise, it is zero. Let

$$\mathbf{W}_{uv} \overset{\text{def}}{=} \left[\mathbf{R}_{\mathbf{y},uv} - \delta_{uv}\sigma_n^2 \mathbf{I}_\zeta \right] \oslash \boldsymbol{\Psi}\boldsymbol{\Psi}^H \qquad (6)$$

be the (u, v)th submatrix of a new matrix $\mathbf{W}$. Then, based on (5), we can arrive at $\mathbf{W} = \mathbf{H}\mathbf{H}^H$, where matrix $\mathbf{H} \in \mathbb{C}^{(\zeta N_R) \times N_T}$ is defined in terms of the channel coefficients as

$$\mathbf{H} \overset{\text{def}}{=} \begin{bmatrix} \mathbf{h}_{1,1} & \cdots & \mathbf{h}_{1,N_T} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{N_R,1} & \cdots & \mathbf{h}_{N_R,N_T} \end{bmatrix}. \qquad (7)$$

Assuming that $\mathbf{H}$ has full column rank, it can be expressed in the form

$$\mathbf{H} = \mathbf{QA} \qquad (8)$$

where the columns of matrix $\mathbf{Q} \in \mathbb{C}^{(\zeta N_R) \times N_T}$ are the orthonormal eigenvectors of $\mathbf{W}$ associated with its N_T largest eigenvalues, and $\mathbf{A} \in \mathbb{C}^{N_T \times N_T}$ is an ambiguity matrix. In theory, the latter is constrained such that $\mathbf{AA}^H = \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix that contains these eigenvalues. As shown in [24] for a similarly structured subspace problem, the matrix of interest $\mathbf{H}$ is identifiable as long as it is a tall matrix, i.e., $\zeta N_R > N_T$. Therefore, this approach, indeed, offers flexibility in choosing the number of transmit and receive antennas, because $\zeta > 1$ is normally fulfilled, which means that $N_T \geq N_R$ is also applicable.

In practice, the channel estimate $\hat{\mathbf{H}}$ can be obtained from $\hat{\mathbf{H}} = \hat{\mathbf{Q}}\mathbf{A}$, where the columns of matrix $\hat{\mathbf{Q}}$ are the eigenvectors that correspond to the N_T largest eigenvalues of an estimated matrix $\hat{\mathbf{W}}$, with its (u, v)th submatrix denoted as $\hat{\mathbf{W}}_{uv}$, which can be obtained as

$$\hat{\mathbf{W}}_{uv} = \left[\hat{\mathbf{R}}_{\mathbf{y},uv} - \delta_{uv}\hat{\sigma}_n^2 \mathbf{I}_\zeta \right] \oslash \boldsymbol{\Psi}\boldsymbol{\Psi}^H \qquad (9)$$

where $\hat{\mathbf{R}}_{\mathbf{y},uv}$ denotes the (u, v)th submatrix of the sampled correlation matrix $\hat{\mathbf{R}}_\mathbf{y}$, and $\hat{\sigma}_n^2$ is an estimate of the noise vari-

ance. Therefore, the accuracy of the channel estimates largely depends on the accuracy of the estimates $\hat{\mathbf{R}}_\mathbf{y}$ and $\hat{\sigma}_n^2$.

In general, to achieve satisfactory performance in the channel estimation step, i.e., $\hat{\mathbf{H}} = \hat{\mathbf{Q}}\mathbf{A}$, the time averaging period T_{av} needed for the aforementioned estimation of the correlation matrix $\hat{\mathbf{R}}_\mathbf{y}$ must be no less than ζN_R, i.e., $T_{av} \geq \zeta N_R$ [21], [25]. However, by exploiting the concept of frequency averaging within the coherence bandwidth, which is denoted as B_c, the required averaging period can effectively be reduced by a factor P, i.e., $T_{av} \geq \zeta N_R/P$. In particular, the estimate of the correlation matrix $\mathbf{R}_\mathbf{y}$ can now be obtained as

$$\hat{\mathbf{R}}_\mathbf{y} = \frac{1}{PT_{av}} \sum_{n=1}^{T_{av}} \sum_{p=1}^{P} \mathbf{y}_p^n \mathbf{y}_p^{nH}. \qquad (10)$$

The merits of this approach in practical MIMO-OFDM systems have been demonstrated for the TI case in [21], where the effect of varying P on the estimation accuracy of frequency-selective channels is also studied.

In practice, B_c can be related to the root-mean-square (rms) delay spread σ_{rms} through $B_c \geq (\arccos(c)/2\pi\sigma_{\text{rms}})$ [22], where $c \in [0, 1]$ is the desired coherence level. Therefore, given the subcarrier spacing Δf for the OFDM system, σ_{rms} for the radio propagation environment, and $c = 0.54$, we can set $P = (2\pi\sigma_{\text{rms}}\Delta f)^{-1}$. We note that there exist many definitions of the coherence bandwidth [26]; therefore, relying on this formula to accurately calculate P may not be suitable for all cases. In addition, our algorithm only requires that variations in channel responses over P adjacent subcarriers remain small but not necessarily zero. Therefore, we generally find that the formula provides a useful guideline for initializing P but further refinement may be needed to achieve the desired tradeoff between smoothing and modeling errors in (10).

Note that, without employing the precoders $\boldsymbol{\Psi}$'s at the transmitter side, the dimension of the ambiguity matrix in [20] is $\zeta N_T \times N_T$, with $\zeta > 1$. Here, in contrast, the use of the precoder matrix $\boldsymbol{\Psi}$ makes reducing this dimension to $N_T \times N_T$ possible.

IV. RECURSIVE CHANNEL TRACKING

We now consider a TV scenario in which the wireless channels could change at each OFDM symbol time. Accordingly, we shall reintroduce the time index m for all the channel-related quantities, including the quantities associated with the aforementioned block-based subspace estimation. Clearly, the matrix $\mathbf{Q}^m$, i.e., $\mathbf{Q}$ in (8) at the mth symbol time, needs to be updated as new data samples become available to properly reflect changes in the unknown channel. Instead of applying an eigenvalue decomposition (EVD) on $\mathbf{W}^m$ at each time step, we can recursively update the EVD through an efficient subspace tracking algorithm to reduce the computational load. We notice that most fast subspace trackers with low complexity assume a rank-1 update [27], [28] and, hence, are not applicable here. Based on (8), we propose a new algorithm that combines the well-known orthogonal iteration with a joint time–frequency averaging to track the wideband TV MIMO channels, without repeatedly incurring EVD operations.

A. Recursive Approach Based on Orthogonal Iteration

Under the current model assumption, an estimate of the TV correlation matrix $\mathbf{R}_{\mathbf{y}}^m$ at the mth OFDM symbol time can be obtained by combining traditional window-based time averaging with frequency averaging over the P frequency subsets Ω_p for $p \in \{1, \ldots, P\}$. This approach results in

$$\hat{\mathbf{R}}_{\mathbf{y}}^m = \sum_{n=m-l+1}^{m} \sum_{p=1}^{P} \beta^{m-n} \mathbf{y}_p^n \mathbf{y}_p^{nH}$$

$$= \beta \hat{\mathbf{R}}_{\mathbf{y}}^{m-1} + \sum_{p=1}^{P} \mathbf{y}_p^m \mathbf{y}_p^{mH} - \sum_{p=1}^{P} \beta^l \mathbf{y}_p^{m-l} \mathbf{y}_p^{m-l\,H} \quad (11)$$

where $l \in \mathbb{N}$, and $0 \leq \beta \leq 1$ denotes the window length and the forgetting factor, respectively. The main challenge still lies in whether we can estimate the required second-order statistics within a sufficiently short processing window.

Considering a scenario in which no windowing is applied, i.e., $l = 1$, we can still collect $\{\mathbf{y}_p^m : p = 1, 2, \ldots, P\}$ at the mth OFDM symbol time without referring to the OFDM symbols of the previous time instances, i.e., $\mathbf{y}_p^n$ for $n < m$. Hence, we can conclude that it is possible to track TV channels that change at each OFDM symbol time, provided that $P > \zeta N_R$. In practice, this condition is not stringent; for example, the choices $(P, \zeta N_R) = (32, 24)$ and $(64, 12)$ were reported in [21], where both the Worldwide Interoperability for Microwave Access (WiMAX) specification and the 3GPP SCM are considered. Of course, by using $l > 1$ in combination with a suitable choice of β, the effective window length can be increased if the aforementioned condition is not met or if it is desired to obtain better smoothing of the channel estimates in the case of mobile terminals with low speed. Clearly, the effective window length cannot significantly exceed the channel coherence time. The choice of the parameters l and β is further discussed in Section VI.

Let $\hat{\mathbf{W}}^m$ be an estimate of the matrix $\mathbf{W}^m$, with its (u, v)th submatrix given as

$$\hat{\mathbf{W}}_{uv}^m \overset{\text{def}}{=} \left[\hat{\mathbf{R}}_{\mathbf{y},uv}^m - \delta_{uv} \hat{\sigma}_n^2 \mathbf{I}_\zeta \right] \oslash \mathbf{\Psi\Psi}^H \quad (12)$$

where $\hat{\mathbf{R}}_{\mathbf{y}}^m$ is now obtained through (11). In this paper, to track the TV channel $\mathcal{H}^m$ with low complexity, we propose to recursively update the principal eigenvectors of $\hat{\mathbf{W}}^m$, which is represented by matrix $\hat{\mathbf{Q}}^m$, and use them to estimate the unknown channel coefficient matrix $\mathbf{H}^m$ according to (8).

Several techniques are available to efficiently compute or update the dominant subspace of a matrix. Among these techniques, *orthogonal iteration* and its variants have been considered for adaptive estimation to a great extent (see [8], [29], and references therein). When applied to a constant matrix, this method is known to exponentially converge to the desired invariant subspace as the number of iterations increases [30]. It can be initialized with any matrix with orthonormal columns; in addition, it is well structured and suitable for real-time processing. The application of orthogonal iteration to the problem at hand is now described.

In particular, given a tall column orthonormal matrix $\hat{\mathbf{Q}}_0^m \in \mathbb{C}^{\zeta N_R \times N_T}$ at the mth OFDM symbol time, the method of *orthogonal iteration* generates a sequence of matrices $\hat{\mathbf{Q}}_\mu^m$, whose column span is assumed to approximate the span of the N_T-D dominant subspace of the matrix $\hat{\mathbf{W}}^m \in \mathbb{C}^{\zeta N_R \times \zeta N_R}$, according to the following recurrence [31], [32]:

$$\hat{\mathbf{Z}}_\mu^m = \hat{\mathbf{W}}^m \hat{\mathbf{Q}}_{\mu-1}^m, \quad \mu = 1, 2, \cdots, n_d \quad (13)$$

$$\hat{\mathbf{Q}}_\mu^m \hat{\mathbf{R}}_\mu^m = \hat{\mathbf{Z}}_\mu^m \quad \text{(QR decomposition)} \quad (14)$$

where $\hat{\mathbf{Q}}_\mu^m$, with an orthonormal column, and $\hat{\mathbf{R}}_\mu^m$, upper triangular, are the factors of a *skinny* QR decomposition of the intermediate matrix product $\hat{\mathbf{Z}}_\mu^m$. Note that, in practice, we choose $\hat{\mathbf{Q}}_0^m = \hat{\mathbf{Q}}_{n_d}^{m-1}$, except when $m = 0$ (the initial condition). The final accuracy depends on the chosen value of n_d; usually, only a few iterations are necessary, as will be shown in Section VI.

Finally, the desired TV channel estimate is obtained by

$$\hat{\mathbf{H}}^m = \hat{\mathbf{Q}}_{n_d}^m \mathbf{A}^m. \quad (15)$$

In (15), the columns of $\hat{\mathbf{Q}}_{n_d}^m$ are the *approximate* principal eigenvectors of $\hat{\mathbf{W}}^m$ that result from the application of the n_dth orthogonal iterations at the mth OFDM symbol time, and $\mathbf{A}^m$ represents the corresponding ambiguity matrix.

Based on orthogonal iteration, we also notice that estimating the dominant subspace of a slowly TV correlation matrix was considered in [33]; here, we extend the use of orthogonal iteration by allowing $n_d \geq 1$ to track a TV MIMO channel.

B. Convergence Properties

To motivate the use of the proposed recursive method in a TV wireless environment, we investigate its convergence properties as follows. Let us first assume that

$$\hat{\mathbf{U}}^{mH} \hat{\mathbf{W}}^m \hat{\mathbf{U}}^m = \hat{\mathbf{\Lambda}}^m = \text{diag}\left(\hat{\Lambda}_i^m \right) \quad (16)$$

is an EVD of $\hat{\mathbf{W}}^m$, where $\hat{\lambda}_1^m \geq \hat{\lambda}_2^m \geq \cdots \geq \hat{\lambda}_{\zeta N_R}^m \geq 0$, and let an EVD of $\mathbf{W}^m$ be similarly defined. Considering the partitions $\hat{\mathbf{\Lambda}}^m = \text{diag}(\hat{\mathbf{\Lambda}}_1^m \hat{\mathbf{\Lambda}}_2^m)$ and $\hat{\mathbf{U}}^m = [\hat{\mathbf{U}}_1^m \ \hat{\mathbf{U}}_2^m]$, where $\hat{\mathbf{U}}_1^m \in \mathbb{C}^{\zeta N_R \times N_T}$, $\hat{\mathbf{U}}_2^m \in \mathbb{C}^{\zeta N_R \times (\zeta N_R - N_T)}$, $\hat{\mathbf{\Lambda}}_1^m \in \mathbb{C}^{N_T \times N_T}$, and $\hat{\mathbf{\Lambda}}_2^m \in \mathbb{C}^{(\zeta N_R - N_T) \times (\zeta N_R - N_T)}$, we can define the distance between the two subspaces $\mathfrak{D}_{N_T}(\hat{\mathbf{W}}^m)$ and $\mathfrak{R}(\hat{\mathbf{Q}}_\mu^m)$ by [30]

$$\text{dist}\left(\mathfrak{D}_{N_T}(\hat{\mathbf{W}}^m), \mathfrak{R}(\hat{\mathbf{Q}}_\mu^m) \right) = \left\| \left(\hat{\mathbf{U}}_2^m \right)^H \hat{\mathbf{Q}}_\mu^m \right\|_2 \quad (17)$$

where $\|.\|_2$ denotes the spectral norm. Let the angle $\theta^m \in [0, \pi/2]$ be defined to provide a measure of the closeness of the two subspaces $\mathfrak{D}_{N_T}(\hat{\mathbf{W}}^m)$ and $\mathfrak{R}(\hat{\mathbf{Q}}_0^m)$ through

$$\cos(\theta^m) \overset{\text{def}}{=} \min_{\mathbf{u} \in \mathfrak{D}_{N_T}(\hat{\mathbf{W}}^m), \mathbf{v} \in \mathfrak{R}(\hat{\mathbf{Q}}_0^m)} \frac{|\mathbf{u}^H \mathbf{v}|}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2}. \quad (18)$$

Then, according to [34], we can arrive at

$$\text{dist}\left(\mathfrak{D}_{N_T}(\hat{\mathbf{W}}^m), \mathfrak{R}\left(\hat{\mathbf{Q}}_\mu^m \right) \right) \leq \tan(\theta^m) \left(\frac{\lambda_{N_T+1}^m}{\lambda_{N_T}^m} \right)^\mu \quad (19)$$
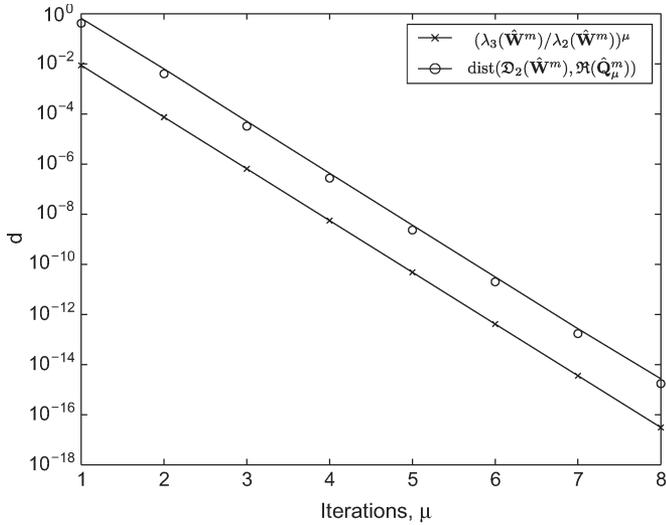
Fig. 3. Demonstration on the rate of convergence in subspace estimation by using orthogonal iteration.

for $\mu = 1, 2, \ldots, n_d$. According to (19), as long as the ratio $(\lambda_{N_T+1}^m / \lambda_{N_T}^m) < 1$, the iterated subspace $\mathfrak{R}(\hat{\mathbf{Q}}_\mu^m)$ exponentially converges to $\mathfrak{D}_{N_T}(\hat{\mathbf{W}}^m)$ with an arbitrary initial condition $\hat{\mathbf{Q}}_0^m = \hat{\mathbf{Q}}_{n_d}^{m-1}$. The convergence behavior of the orthogonal iteration scheme (13) and (14) as a function of μ is well predicted by (19) in the current application. To illustrate this point, Fig. 3 shows a plot of the subspace distance (17) as a function of μ when (13) and (14) is used to approximate the 2-D dominant subspace of a particular matrix $\hat{\mathbf{W}}_m$, appearing at a given symbol time m in one of our simulations. We also show a plot of $(\lambda_3(\hat{\mathbf{W}}^m) / \lambda_2(\hat{\mathbf{W}}^m))^\mu$ for reference.

Although the tracking performance can be improved by increasing n_d in a general sense, the iterated subspace converges to $\mathfrak{D}_{N_T}(\hat{\mathbf{W}}^m)$ instead of $\mathfrak{D}_{N_T}(\mathbf{W}^m)$. Therefore, the performance largely depends on whether we can obtain a good estimate of $\mathbf{W}^m$ at each OFDM symbol time. Although the use of additional frequency-domain samples as shown in (11) helps improve the quality of this estimation in a TV environment, some errors are inevitable. To investigate this effect, let us define the *residual estimation error* at time mth as

$$\text{Res}(m) \overset{\text{def}}{=} \text{dist}\left(\mathfrak{D}_{N_T}(\mathbf{W}^m), \mathfrak{R}\left(\hat{\mathbf{Q}}_{n_d}^m\right)\right)$$
$$= \left\| (\mathbf{U}_2^m)^H \hat{\mathbf{Q}}_{n_d}^m \right\|_2. \tag{20}$$

For a sufficiently large n_d and invoking the exponential convergence properties of orthogonal iteration, (20) can be expressed as

$$\text{Res}(m) \cong \left\| (\mathbf{U}_2^m)^H \hat{\mathbf{U}}_1^m \right\|_2 \tag{21}$$

where $\hat{\mathbf{U}}_1^m = \mathbf{U}_1^m + \Delta \mathbf{U}_1^m$, with $\Delta \mathbf{U}_1^m = \mathbf{U}_2^m (\mathbf{U}_2^m)^H \Delta \mathbf{W}^m \mathbf{U}_1^m (\mathbf{\Lambda}_1^m)^{-1}$ representing the first-order perturbed signal subspace due to estimation error $\Delta \mathbf{W}^m = \hat{\mathbf{W}}^m - \mathbf{W}^m$ [35]. By invoking the orthogonality between $\mathbf{U}_1^m$ and $\mathbf{U}_2^m$, we can obtain

$$\left\| (\mathbf{U}_2^m)^H \hat{\mathbf{U}}_1^m \right\|_2 = \left\| (\mathbf{U}_2^m)^H \Delta \mathbf{W}^m \mathbf{U}_1^m (\mathbf{\Lambda}_1^m)^{-1} \right\|_2$$
$$\leq \|\Delta \mathbf{W}^m\|_2 \left\| (\mathbf{\Lambda}_1^m)^{-1} \right\|_2. \tag{22}$$
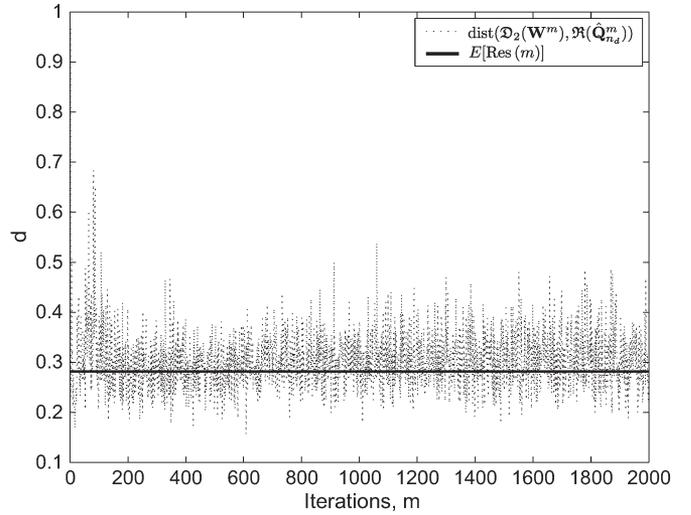
(a)
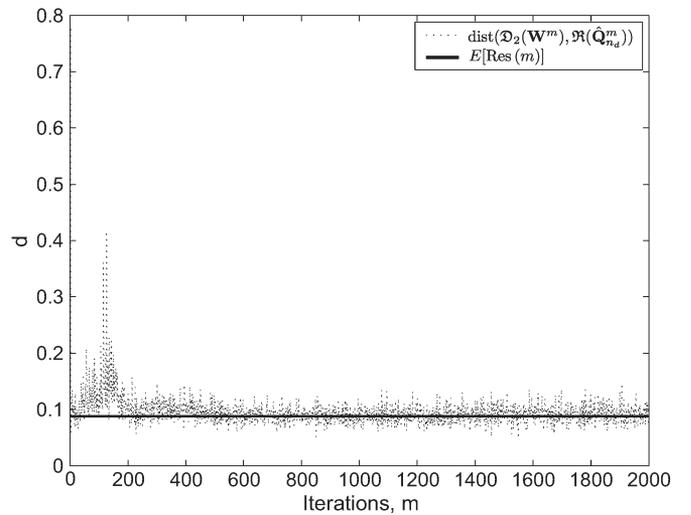
(b)

Fig. 4. Demonstration of $\text{dist}(\mathfrak{D}_2(\mathbf{W}^m), \mathfrak{R}(\hat{\mathbf{Q}}_{n_d}^m))$ and $E[\text{Res}(m)] = 8\sigma_w E[\lambda_1^m / \lambda_2^m]$ based on (26) for $\sigma_w^2 = 10^{-3}$ and 10^{-4}. Note that $E[\lambda_1^m / \lambda_2^m]$ is obtained by time averaging over 2000 iterations. (a) $\sigma_w^2 = 10^{-3}$. (b) $\sigma_w^2 = 10^{-4}$.

Consider a particular application of the orthogonal iteration scheme (13) and (14) in which a random estimation error $\Delta \mathbf{W}^m$ is added to $\mathbf{W}^m$ at the mth OFDM symbol time. In particular, let $[\Delta \mathbf{W}^m]_{i,j} \sim \mathcal{N}(0, \sigma_w^2)$, $\forall i, j$, be independent and identically distributed (i.i.d.) random variables. Then, considering that $E\|\Delta \mathbf{W}\|_2 < 2\sigma_w \sqrt{\zeta N_R}$ according to [36] and $\|(\mathbf{\Lambda}_1^m)^{-1}\|_2 = 1/\lambda_{N_T}^m$, where λ_i^m denotes the ith eigenvalue of $\mathbf{W}^m$, we can arrive at

$$E[\text{Res}(m)] \cong 2\sigma_w \sqrt{\zeta N_R} / \lambda_{N_T}^m. \tag{23}$$

To illustrate the accuracy of this formula, we consider a particular realization of $\mathbf{W}^m \in \mathbb{C}^{16 \times 16}$ with a 2-D dominant subspace (i.e., $N_T = 2$), with the additive estimation error $\Delta \mathbf{W}^m$ modeled as aforementioned. In Fig. 4(a) and (b), we plot $\text{Res}(m)$ (20) versus time m for $\sigma_w^2 = 10^{-3}$ and 10^{-4}, respectively, along with the expected value computed from (23). For small values of σ_w^2, we find that (23) accurately predicts the observed

mean level of Res (m), whereas it becomes less accurate for larger σ_w^2, mainly due to the effect of higher order terms in the perturbed signal subspace, i.e., $\Delta \mathbf{U}_1^m$. In practice, σ_w^2 is related to the estimation error of $\mathbf{R_y}$, resulting from additive white Gaussian channel noise, insufficient time averaging, and improper choice of P.

Based on (23), we can conclude that, to reduce the bias of the estimation, we should keep the dimension of $\mathbf{W}^m$, i.e., ζN_R, as small as possible. This condition coincides with our design goal to reduce the time averaging period by exploiting the frequency correlation, which, in effect, amounts to reducing the size of $\hat{\mathbf{R}}_\mathbf{y}^m$ and $\mathbf{W}^m$.

C. Proposed Blind Recursive Estimation Algorithm

We summarize our precoded subspace-based tracking algorithm and its corresponding complexity in Algorithm 1. The total computational complexity (flops) of the proposed algorithm for each iteration is $\mathcal{O}(\zeta^4) + \mathcal{O}(\zeta^2 N_R^2 N_T)$. For moderate choices of the number of transmit and receive antennas (i.e., $\zeta^2 > N_R^2 N_T$), this figure can be approximated as $\mathcal{O}(\zeta^4)$ and, thus, is generally much smaller than the competing algorithms in [8] and [9]. In practice, the ambiguity matrix $\mathbf{A}^m$ can be acquired by employing a small number of pilot symbols, resulting in the so-called semiblind approach. In this case, it is of interest to compare the number of pilot symbols required for estimating $\mathbf{A}^m$ with the number required for estimating the full channel matrix $\mathbf{H}^m$ through a conventional pilot-based approach. In particular, the estimation of $\mathbf{H}^m$ requires $\zeta N_R N_T$ complex symbols, whereas the estimation of $\mathbf{A}^m$ (taking the constraint $\mathbf{A}^m(\mathbf{A}^m)^H = \mathbf{\Lambda}^m$ into account) requires $(1/2)N_T^2$ complex symbols. Consequently, the semiblind version of the proposed algorithm will result in a reduction in the number of required pilots by a factor of $\varrho = 2\zeta N_R / N_T > 1$.

In theory, the proposed method can be applied to OFDM systems with arbitrary bandwidth $N_c \Delta f$. However, under a constraint of fixed coherence bandwidth, i.e., fixed P, special care needs to be taken in the case of extremely large system bandwidth. The first issue relates to the computational complexity of the proposed approach, which scales with the fourth power of $\zeta = N_c / P$. Second, to have sufficient samples from the frequency domain for estimating the correlation matrix in (11), we need $P > \zeta N_R$. Thus, increasing N_c without bound while keeping P fixed will eventually result in a problematic situation. In this case, one practical way of applying our proposed approach to wider system bandwidths is to break down the frequency band of interest into smaller contiguous subbands and use the algorithm with a reduced value of N_c in each subband. This approach avoids the aforementioned issues and achieves estimation accuracy similar to the estimation accuracy obtained for a smaller system bandwidth, as we have verified.

Algorithm 1: Blind recursive subspace-based identification of TV-MIMO channels

    Initialization: $\hat{\mathbf{Q}}_{n_d}^{l-1} = \mathbf{I}(:,1:N_T)$, $\hat{\mathbf{R}}_\mathbf{y}^0 = \mathbf{0}$, $\tilde{\mathbf{R}}_\mathbf{y}^0 = \mathbf{0}$

    for $m = 1, 2, \ldots$ **do**

        Input vector: $\mathbf{y}_1^m, \ldots, \mathbf{y}_P^m$

        $\tilde{\mathbf{R}}_\mathbf{y}^m = \sum_{p=1}^P \mathbf{y}_p^m \mathbf{y}_p^{m\,H}$

        if $m < l$ **then**

            $\hat{\mathbf{R}}_\mathbf{y}^m = \beta \hat{\mathbf{R}}_\mathbf{y}^{m-1} + \tilde{\mathbf{R}}_\mathbf{y}^m$

        else

            $\hat{\mathbf{R}}_\mathbf{y}^m = \beta \hat{\mathbf{R}}_\mathbf{y}^{m-1} + \tilde{\mathbf{R}}_\mathbf{y}^m - \beta^l \tilde{\mathbf{R}}_\mathbf{y}^{m-l}$ $\mathcal{O}(\zeta^2 N_R^2)$

            $\hat{\mathbf{W}}_{uv}^m = [\hat{\mathbf{R}}_{\mathbf{y},uv}^m - \delta_{uv}\hat{\sigma}_n^2 \mathbf{I}] \oslash \mathbf{\Psi}\mathbf{\Psi}^H$ $\mathcal{O}(\zeta^4)$

            $\hat{\mathbf{Q}}_0^m = \hat{\mathbf{Q}}_{n_d}^{m-1}$

            for $\mu = 1, 2, \ldots, n_d$ **do**

                $\hat{\mathbf{Z}}_\mu^m = \hat{\mathbf{W}}^m \hat{\mathbf{Q}}_{\mu-1}^m$ $\mathcal{O}(\zeta^2 N_R^2 N_T)$

                $\hat{\mathbf{Q}}_\mu^m \hat{\mathbf{R}}_\mu^m = \hat{\mathbf{Z}}_\mu^m$ (QR   factorization   on $\hat{\mathbf{Z}}_\mu^m$)

                                       $\mathcal{O}(\zeta^2 N_R^2 N_T)$

            end for

            $\hat{\mathbf{H}}^m = \hat{\mathbf{Q}}_{n_d}^m \mathbf{A}^m$ $\mathcal{O}(\zeta N_R N_T^2)$

        end if

    end for

V. PRECODER DESIGN

Various precoding techniques have been proposed in the context of OFDM systems. For example, a Vandermonde-matrix-based precoder was proposed to maximize diversity and coding gains when channel knowledge is available [37]. A block triangular precoding matrix was applied for every pair of OFDM symbols to assist blind channel estimation in [38]. However, none of these techniques can suit our special needs.

To simplify the notation for the following discussions, let us define $\mathbf{\Gamma} \stackrel{\text{def}}{=} \mathbf{\Psi}\mathbf{\Psi}^H$. Let $\psi_{i,j}$ and $\gamma_{i,j}$ denote the (i,j)th entry of matrices $\mathbf{\Psi}$ and $\mathbf{\Gamma}$, respectively. According to (12), the choice of $\mathbf{\Psi}$ does not appear to be restricted, except for the trivial constraint that the entries of $\mathbf{\Gamma}$ cannot be zeros, i.e., $\gamma_{i,j} \neq 0, \forall\, i, j$. Therefore, we can judiciously choose the precoder matrix to simplify the channel estimator and optimize its performance.

First, we note that, if the diagonal entries of $\mathbf{\Gamma}$ are identical, i.e., $\gamma_{i,i}$ is a constant for $i = 1, 2, \ldots, \zeta$, the additional estimation of the noise variance in (12) can be avoided. To be more specific, let us define a new matrix $\hat{\mathbf{T}}^m$, with its (u,v)th submatrix given as

$$\hat{\mathbf{T}}_{uv}^m \stackrel{\text{def}}{=} \hat{\mathbf{R}}_{\mathbf{y},uv}^m \oslash \mathbf{\Gamma}, \; u, v \in \{1, 2, \cdots, N_R\}. \qquad (24)$$

Then, we can arrive at $\hat{\mathbf{T}}^m = \hat{\mathbf{W}}^m + \rho\, \mathbf{I}$ (for some $\rho \in \mathbb{R}$). Because $\hat{\mathbf{T}}^m$ has the same invariant subspaces as $\hat{\mathbf{W}}^m$, we can simply apply $\hat{\mathbf{T}}^m$ instead of $\hat{\mathbf{W}}^m$ in Algorithm 1 to eliminate the noise variance estimation.

Second, letting $\Delta \mathbf{R}_\mathbf{y}^m$ denote the difference between the estimated and the *true* correlation matrices, i.e., $\hat{\mathbf{R}}_\mathbf{y}^m = \mathbf{R}_\mathbf{y}^m + \Delta \mathbf{R}_\mathbf{y}^m$, we may express $\hat{\mathbf{T}}_{uv}^m$ as follows:

$$\hat{\mathbf{T}}_{uv}^m = \underbrace{\mathbf{R}_{\mathbf{y},uv}^m \oslash \mathbf{\Gamma}}_{\stackrel{\text{def}}{=}\mathbf{T}_{uv}^m} + \underbrace{\Delta \mathbf{R}_{\mathbf{y},uv}^m \oslash \mathbf{\Gamma}}_{\stackrel{\text{def}}{=}\Delta\mathbf{T}_{uv}^m}. \qquad (25)$$

Then, it becomes clear that the choice of the precoder should focus on eliminating the error term $\Delta \mathbf{T}_{uv}^m$. Matrix $\Delta \mathbf{R}_{\mathbf{y},uv}^m$ in (25) has a random nature, which results from the effects of the TV channels, additive noise, and insufficient number of data samples. Let $J(\mathbf{\Psi}) \stackrel{\text{def}}{=} \sum_{u,v} E\|\Delta \mathbf{T}_{uv}^m\|_F^2 = \sum_{u,v} E\|\Delta \mathbf{R}_{\mathbf{y},uv}^m \oslash$

$\boldsymbol{\Psi}\boldsymbol{\Psi}^H\|_F^2$. Given that $\oslash$ is an element-wise division, minimizing $J(\boldsymbol{\Psi})$ is equivalent to maximizing (in a weighted sense) the entries of $\boldsymbol{\Gamma} = \boldsymbol{\Psi}\boldsymbol{\Psi}^H$. Nevertheless, the choice of a precoder is subject to a fixed transmit power and, thus, cannot arbitrarily be large; the entries of $\boldsymbol{\Gamma}$ should therefore be maximized based on the statistics of $\Delta\mathbf{R}_{\mathbf{y}}^m$ *and* subject to an adequate normalization.

In summary, the choice of the precoder can be optimized through the objective function, i.e.,

$$\min_{\boldsymbol{\Psi}} J(\boldsymbol{\Psi}) \qquad (26)$$

subject to the following two constraints.
(C1) To guarantee that the element-wise division in (12) and (24) is feasible, the precoder must fulfill the condition: $|\gamma_{i,j}| > \delta, \forall i, j$, for some small $\delta > 0$;
(C2) To normalize the average transmit power, we require that $\sum_j |\psi_{i,j}|^2 = 1, \forall i$.

Constraint C2 has interesting ramifications. First, it implies that the diagonal entries of $\boldsymbol{\Gamma}$ are identical, i.e.,; $\gamma_{i,i} = 1, \forall\ i$. Hence, there is no need for noise variance estimation, as explained. Second, it follows from the relation $\boldsymbol{\Gamma} = \boldsymbol{\Psi}\boldsymbol{\Psi}^H$ and the Cauchy–Schwartz inequality that

$$|\gamma_{i,j}|^2 \leq \gamma_{i,i}\gamma_{j,j} = 1. \qquad (27)$$

Therefore, the off-diagonal entries of $\boldsymbol{\Gamma}$ cannot be made arbitrarily large. Thus, the optimization problem (26) cannot be replaced by a simplified optimization over $\boldsymbol{\Gamma}$ with constraint $\gamma_{i,i} = 1$.

Note that the seminal precoding matrices in [37] and [38] do not fulfill C1 and, hence, are not applicable here. In particular, the optimized precoder in [38] is constructed for the case of asymptotic performance (i.e., SNR $\rightarrow \infty$) and does not provide clear insights for scenarios when the OFDM system is operated at low to moderate SNRs. In the following sections, we propose two precoders that fulfill the aforementioned constraints, and we resort to numerical experiments in Section VI to demonstrate the choices of precoders in various practical applications.

A. Design 1

In the absence of a more specific model, we consider a worst case situation and assume that the entries of $\Delta\mathbf{R}_{\mathbf{y}}^m$ are i.i.d. random variables with zero mean and equal variance; this choice is further supported by our numerical observations. Based on this assumption, the objective function in (26) becomes a standard optimization problem and can be solved using a Lagrange multiplier. Accordingly, the optimal precoder, assuming $\gamma_{i,j} \in \mathbb{R}^+$, is obtained as $\boldsymbol{\Psi}_0 = (1/\sqrt{\zeta})\mathbf{1}_{\zeta\times\zeta}\mathbf{B}$, where $\mathbf{1}_{\zeta\times\zeta}$ denotes a $\zeta \times \zeta$ matrix of all ones, and $\mathbf{B}$ is an arbitrary $\zeta \times \zeta$ unitary matrix.[1] In turn, this choice yields $\boldsymbol{\Gamma}_0 = \boldsymbol{\Psi}_0\boldsymbol{\Psi}_0^H = \mathbf{1}_{\zeta\times\zeta}$. This result coincides with the optimal choice of the precoder in terms of estimation performance, obtained from the numerical considerations in [39].

Nevertheless, the aforementioned precoder $\boldsymbol{\Psi}_0$ has rank 1 (condition number $= \infty$) and, thus, is not a good choice from the perspective of symbol recovery. To make $\boldsymbol{\Psi}_0$ nonsingular while keeping the estimation performance close to the optimum, we can perturb the entries of $\boldsymbol{\Psi}_0$ in the following manner: $\boldsymbol{\Psi}_0 \rightarrow \boldsymbol{\Psi}$, where the diagonal entries of $\boldsymbol{\Psi}$ now slightly exceed the off-diagonal entries. This approach is motivated by the following property: $\text{rank}(\boldsymbol{\Psi}) \geq \sum_{i=1}^{\zeta} |\psi_{i,i}|/b_i$, where $b_i \overset{\text{def}}{=} \sum_{j=1}^{\zeta} |\psi_{i,j}|$ [40]. This property suggests that, under C2, we can increase $\text{rank}(\boldsymbol{\Psi})$ from 1 by boosting the ratios $|\psi_{i,i}|/b_i$. Here, we propose to use a simple Toeplitz matrix to accomplish this goal. That is, we define

$$\boldsymbol{\Psi} = \boldsymbol{\Psi}(\nu) \overset{\text{def}}{=} \frac{1}{\sqrt{1+(\zeta-1)\nu^2}} \begin{bmatrix} 1 & \cdots & \nu \\ \vdots & \ddots & \vdots \\ \nu & \cdots & 1 \end{bmatrix}_{\zeta\times\zeta} \qquad (28)$$

where ν is shown as the common perturbed value of the off-diagonal entries of $\boldsymbol{\Psi}_0$. We note that a structure similar to (28) was employed for *block-based* channel estimation in [24].

Intuitively, there exists an optimal tradeoff in terms of ν between the symbol recovery and channel estimation performance for a given SNR. Although the analysis for determining an optimal value of ν for this combined objective appears difficult, some insight can be obtained as follows. Using (28), it can be verified that $\boldsymbol{\Gamma}$ will contain ones on its diagonal and

$$\phi(\nu) = 1 - \frac{(\nu-1)^2}{1+(\zeta-1)\nu^2} \leq 1 \qquad (29)$$

on its off-diagonal. We can also express the condition number of $\boldsymbol{\Gamma}$ in terms of ν as

$$\kappa_{\boldsymbol{\Gamma}}(\nu) = \frac{\zeta\left(1+(\zeta-1)\nu^2\right)}{(1-\nu)^2} - (\zeta-1). \qquad (30)$$

On the one hand, to minimize channel estimation errors, we seek to maximize $\phi(\nu)$, which is achieved when $\nu = 1$. On the other hand, to improve symbol recovery, we seek to minimize $\kappa_{\boldsymbol{\Gamma}}(\nu)$. To explain the behavior of our algorithm around $\nu = 1$, consider $\nu = 1 \pm \epsilon$, where $0 < \epsilon \ll 1$. It can be shown that $\phi(1 \pm \epsilon) \cong 1 - \epsilon^2/\zeta$, whereas $\kappa_{\boldsymbol{\Gamma}}(1+\epsilon) \gg \kappa_{\boldsymbol{\Gamma}}(1-\epsilon)$. Thus, for a given ϵ, the choice $\nu = 1 - \epsilon < 1$ is advantageous. In fact, we have confirmed through simulations that $\nu > 1$ invariably leads to inferior performance. In practice, we can impose some constraint $\kappa_{\boldsymbol{\Gamma}}(\nu) \leq \kappa^*$ for some finite κ^* and use (29) and (30) to find the corresponding $\nu \in (0,1)$ that maximizes $\phi(\nu)$.[2]

B. Design 2

We propose another precoder that is shown as a generalization of the previous precoder. To be more specific, $\boldsymbol{\Gamma}$ is a convex combination of the identity and all-one matrices that fulfill constraints C1 and C2 and is defined as follows:

$$\boldsymbol{\Gamma} = a\mathbf{I}_\zeta + (1-a)\mathbf{1}_{\zeta\times\zeta}, \; a \in [0, \zeta/(\zeta-1)] \qquad (31)$$

---

[1]It can be verified that this solution corresponds to a stationary point of the Lagrangian function; furthermore, because it leads to equality in (27), it achieves the global maximum.

[2]We note, based on (1) and (2), that the precoder outputs are allocated to OFDM subchannels that are $P\Delta f$ apart. Here, this separation is chosen to roughly correspond to the coherence bandwidth of the MIMO channel; therefore, the seemingly correlated precoder outputs will decorrelate as they go through the wideband channel. Therefore, except for the rank-deficient case $\nu \cong 1$, the spectral efficiency will not significantly be affected.

where the limits on a follow from (27) and the semidefinite positive property of $\mathbf{\Gamma}$. In particular, $a = 0$ represents the scenario where optimal channel estimation is achieved, whereas $a = 1$ corresponds to the optimal symbol recovery. Note that, for $\mathbf{\Psi}(\nu)$ in (28), the product $\mathbf{\Gamma} = \mathbf{\Psi}(\nu)\mathbf{\Psi}^H(\nu)$ is a special case of (31), with $a = 1 - \phi(\nu)$.

We can also represent (31) as circulant, i.e., $\mathbf{\Gamma} = \text{circ}\,(1, 1 - a, \ldots, 1 - a)$, and thus, the corresponding precoder matrix can conveniently be expressed as $\mathbf{\Psi} = \mathbf{S}\mathbf{P}^{1/2}$, where $\mathbf{S}$ is the unitary DFT matrix with entries $[\mathbf{S}]_{k,l} = (1/\sqrt{\zeta})e^{-j2\pi kl/\zeta}$ and $\mathbf{P} = \text{diag}[p_0 \cdots p_{\zeta-1}]$ with diagonal entries $p_0 = (1 - a)\zeta + a$ and $p_i = a$ for $i \neq 0$ obtained as the DFT of the first column of $\mathbf{\Gamma}$. Now, let $\boldsymbol{\chi}, \boldsymbol{\chi}' \in \mathbb{C}^{\zeta \times 1}$ be realizations of $\mathbf{x}_{j,p}^m$ and χ_i, and $\chi_i' \in \mathcal{A}$ be their ith entries, respectively. The distance between the corresponding precoder outputs can be written as

$$D(a; \boldsymbol{\chi}, \boldsymbol{\chi}') \overset{\text{def}}{=} \|\mathbf{\Psi}(\boldsymbol{\chi} - \boldsymbol{\chi}')\|_2^2$$
$$= (1 - a)\zeta \,|\chi_0 - \chi_0'|^2 + a \sum_{i=1}^{\zeta-1} |\chi_i - \chi_i'|^2 \geq 0.$$

Thus, the precoder that achieves the best tradeoff between symbol recovery and channel estimation can be found by optimizing the following combined objective function:

$$\max_a \left(-\epsilon a^2 + \min_{\boldsymbol{\chi} \neq \boldsymbol{\chi}'} D(a; \boldsymbol{\chi}, \boldsymbol{\chi}') \right) \tag{32}$$

where the choice of weighting factor $\epsilon \geq 0$ controls the final optimal tradeoff. Because the closed form of $\min_{\boldsymbol{\chi} \neq \boldsymbol{\chi}'} D(a; \boldsymbol{\chi}, \boldsymbol{\chi}')$ is intractable, we resort to computer search for solving (32). An example for quaternary phase-shift keying (QPSK) modulation (i.e., $\mathcal{A} = (1/\sqrt{2})\{\pm 1 \pm j\}$, with $|\mathcal{A}| = 4$) and $\zeta = 4$ will be demonstrated in the next section.

VI. Numerical Experiments

The MIMO system under consideration consists of $N_T = 2$ transmit and $N_R = 3$ receive antennas. The number of OFDM subcarriers N_C is set to 256, with a CP length of 32. For each time epoch, the incoming data streams are i.i.d. QPSK symbols. The OFDM useful symbol duration is 91.4 μs (11.4 μs for the CP), resulting in a subcarrier spacing of $\Delta f = 10.94$ kHz and a total system bandwidth of 2.8 MHz. Modeling of the TV MIMO channel is based on the 3GPP-SCM setup [41]. We consider a suburban macro scenario with carrier frequency $f_c = 2.5$ GHz and $E[\sigma_{\text{rms}}] = 0.17\,\mu$s. The mobile station (MS) is allowed to roam in a random direction at a constant speed of 100 km/h. Hence, the maximum Doppler shift is 231.5 Hz, and the normalized Doppler frequency is 0.02. For this given scenario, we find $B_c \cong 1/(2\pi\sigma_{\text{rms}}) = 936.2$ kHz $= 85.6\,\Delta f$; experimentally, we have found that a suitable value of P is 64. In the following discussion, we present the BER and NMSE of the proposed algorithm, where the ambiguity matrix is obtained from $\mathbf{A}^m = (\hat{\mathbf{Q}}_{n_d}^m)^\dagger \mathbf{H}_{\lceil P/2 \rceil}^m$. For the BER calculations, symbol recovery is implemented in the frequency domain. The NMSE for the mth channel estimate is defined as $\sum_{i,j,k} E[|\hat{h}_{i,j}^m[k] - h_{i,j}^m[k]|^2]/(\sum_{i,j,k} E[|h_{i,j}^m[k]|^2])$, where the sum over k runs
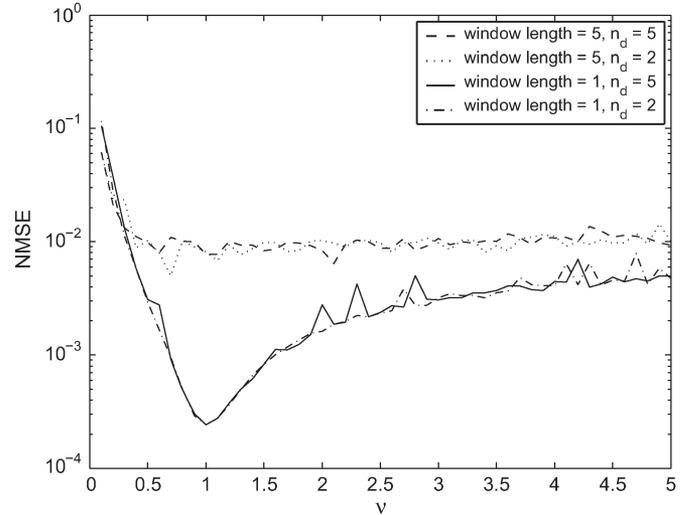
Fig. 5. NMSE versus precoder coefficient ν when the MS speed is 100 km/h ($E_b/N_0 = 14$ dB).
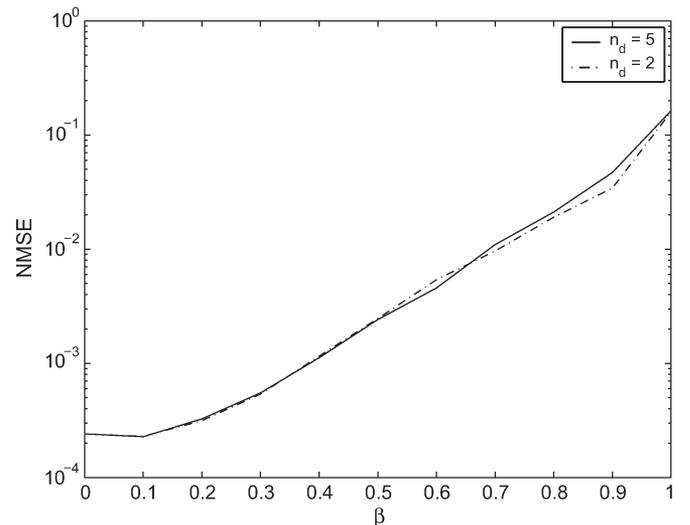
Fig. 6. NMSE versus forgetting factor β when the MS speed is 100 km/h ($\nu = 1$ and $E_b/N_0 = 14$ dB).

from 0 to $N_c - 1$, and the ensemble average is taken over 200 independent runs.[3]

Considering a rectangular window (i.e., $\beta = 1$) of length $l = 1$ and 5, we first investigate the choice of the precoder design 1 coefficient ν from the perspective of channel estimation performance. Fig. 5 shows the NMSE of the channel estimates versus ν after 100 iterations when the SNR per bit $E_b/N_0 = 14$ dB, where E_b and N_0 denote the energy per bit and the one-sided noise power spectral density, respectively. Choosing $l = 1$ gives the best performance, because the wireless channel very rapidly changes in this case. In particular, we observe that the NMSE reaches its minimum, i.e., 2.5×10^{-4} when $\nu = 1$, which coincides with our analysis in Section V. In Fig. 6, we investigate the choice of the forgetting factor β

[3]Although frequency-domain interpolation was not applied in this paper, it could be employed as a postprocessing step to smooth out the estimated channel responses and further reduce the NMSE.
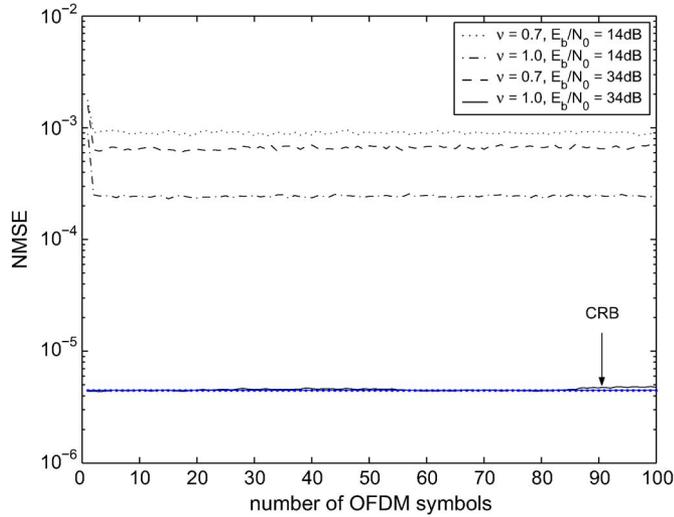
Fig. 7. NMSE versus the number of OFDM symbols when the MS speed is 100 km/h ($n_d = 2$).

Fig. 8. BER versus precoder parameter ν when the MS speed is 100 km/h ($n_d = 2$).

when an exponential window of infinite length is considered under the same condition. We can see that the NMSE reaches a minimum, a value comparable with the rectangular window with $l = 1$, i.e., around NMSE $= 2 \times 10^{-4}$ when $\beta \in [0, 0.1]$, which means that previous data samples are of little use for the estimation of the current channel statistics. Thus, in this TV channel, employing an exponential window cannot gain additional estimation accuracy. We also notice that there is no significant improvement in the estimation performance when n_d is increased from 2 to 5 in both figures. Hence, we simply assign $n_d = 2$ and employ a rectangular window of length $l = 1$ in the following experiments.

Fig. 7 presents the NMSE of the channel estimates versus OFDM symbol time when $\nu \in \{1, 0.7\}$ and $E_b/N_0 \in \{14, 34\}$ dB. We can see that the proposed algorithm can track the TV channel in less than five OFDM symbols in all cases and maintain its performance over time, despite the rapid variations in the channel coefficients. This case shows that reducing ν from 1 to a smaller value to achieve a compromise between estimation performance and symbol recovery will not affect the convergence rate. We notice that, for $\nu = 1$ at $E_b/N_0 = 34$ dB, the proposed blind estimator reaches the Cramer–Rao bound,[4] as computed for a TI channel observed over five OFDM symbols at the same SNR per bit [21].

In Fig. 8, we show the BER versus the precoder coefficient ν for various E_b/N_0's. We consider both the least squares (LS) and the total least squares (TLS) [42] estimation for symbol recovery. We observe that the higher the E_b/N_0, the larger the optimal choice of ν, and hence, the lower the BER. This case can be explained as follows. For a less noisy scenario, a shorter distance between any pair of the precoder outputs is allowed, and thus, we can increase the value of ν to gain better estimation performance and, by doing so, achieve a lower BER. We also observe in Fig. 8 that, for a given E_b/N_0,
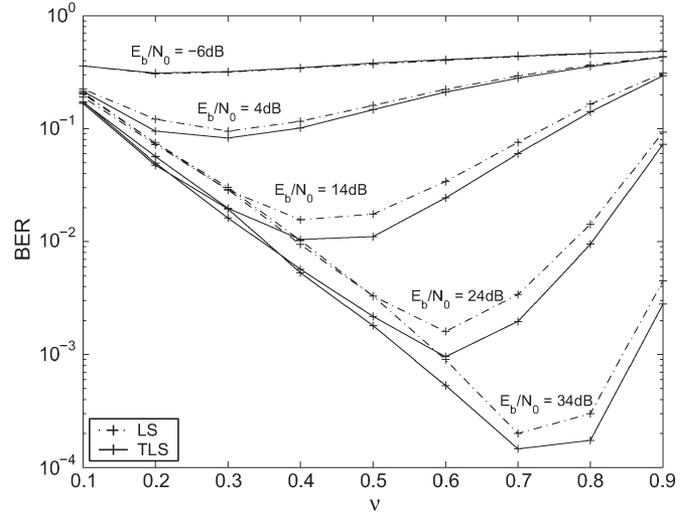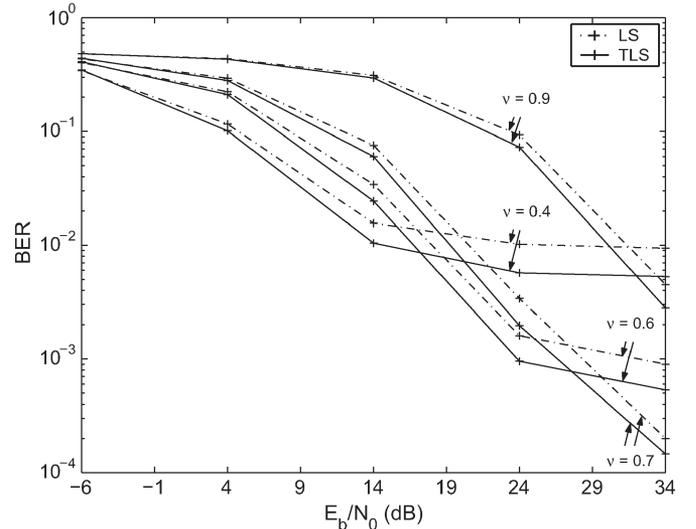
Fig. 9. BER versus E_b/N_0 when the MS speed is 100 km/h ($n_d = 2$).

the TLS estimation (solid lines) generally outperforms the LS estimation (dashed–dotted lines) estimation, because the TLS estimation takes the channel estimation errors into account while performing the symbol recovery.

Fig. 9 demonstrates the BER versus E_b/N_0 for various ν's, considering both the LS and TLS estimation for symbol recovery. We can see that, when $\nu = 0.4$, the proposed algorithm performs best in the low E_b/N_0 region due to its largest distance between any pair of precoder outputs; however, it performs the worst in the high E_b/N_0 region, because less accurate channel information is used. The case $\nu = 0.9$ yields the worst performance for almost all E_b/N_0 values, although the best estimation performance is achieved. A good choice of ν should fall between 0.6 and 0.7 when the E_b/N_0 is moderate to high, and a gain of 1–2 dB can be achieved by using the TLS instead of the LS estimation at the $E_b/N_0 = 19$ dB. Because we consider QPSK modulation with $\zeta = 4$, we can verify that the joint objective function (32) for design 2 reaches its maximum when $a = 1/\epsilon$. Fig. 10 demonstrates the BER versus E_b/N_0

---

[4]To make the comparison relevant, the CRB is calculated basis on the approximate staircase channel model with reduced dimensionality, i.e., $\mathcal{H}_1^m = \cdots = \mathcal{H}_P^m$.
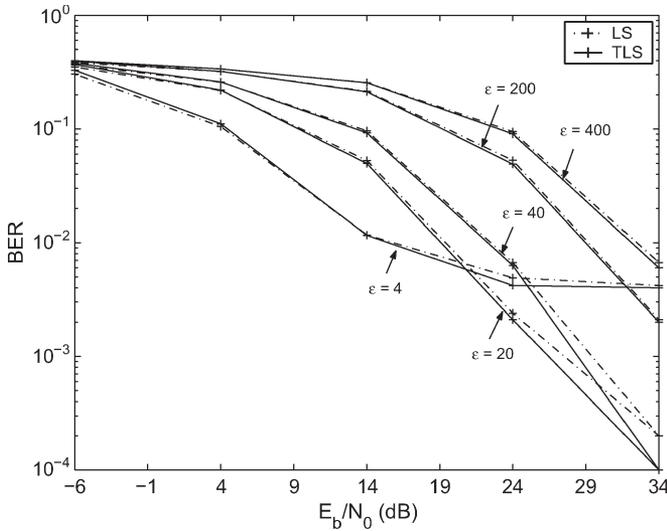
Fig. 10.   BER versus E_b/N_0 when the MS speed is 100 km/h ($n_d = 2$).
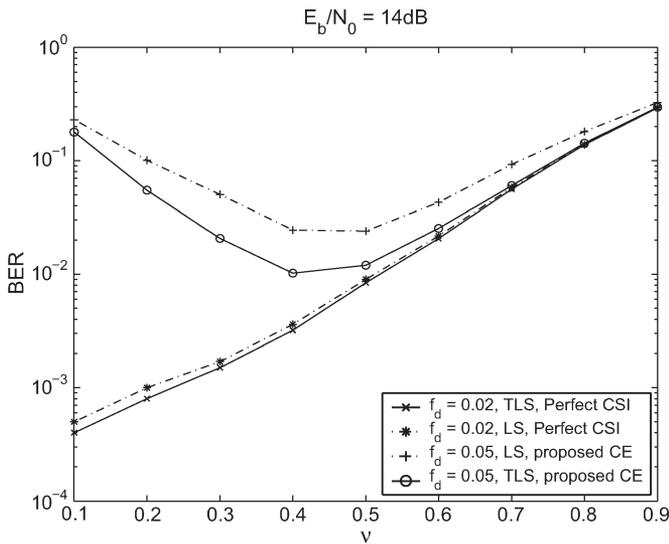
Fig. 11.   BER versus ν for cases of perfect CSI and a higher normalized Doppler frequency.

Fig. 12.   Performance comparison of the pilot-based and the proposed blind channel estimation when the MS speed is 100 km/h.

for various ϵ's, considering both the LS and TLS estimation for symbol recovery. For the given scenario, we can see that a good choice of ϵ should fall between 4 and 20. Here, the advantages of using the TLS instead of the LS are not significant.

One important measure of performance for OFDM systems is the so-called peak-to-average power ratio (PAPR) [22]. Comparing Figs. 9 and 10, we notice that, for a specific E_b/N_0, the achievable performance of the two proposed precoders are similar. However, precoder design 2 induces about 1.5 dB lower PAPR than design I, mainly due to the additional phase terms introduced by the Fourier matrix. This reduction is important, because an OFDM system with a lower PAPR can generally achieve a better BER and higher capacity.

In Fig. 11, we study the tradeoff between symbol recovery and estimation performance from additional angles, with E_b/N_0 set to 14 dB. The top part shows BER versus ν for the proposed method compared to recovery with perfect channel-state information (CSI). We can see that, for $\nu \leq 0.4$, the

BER with perfect CSI improves because the performance is no longer limited by the estimation error. However, when $\nu \geq 0.5$, the BER is restricted by the symbol recovery rather than the channel estimation accuracy, and in this case, the performance gap between the proposed algorithm and perfect CSI becomes negligible. The bottom part illustrates the effect of increasing the normalized Doppler frequency f_d from 0.02 to 0.05 (i.e., MS speed of 250 km/h). Here, the proposed algorithm only suffers modest performance degradation for both LS and TLS estimation approaches.

To demonstrate that the bandwidth efficiency can be increased by the proposed method, we provide a performance benchmark based on pilot-based channel estimation. For every d_T OFDM symbols in the time domain, a comb of pilot symbols is inserted in the frequency domain, with a regular spacing of 16 subcarriers. At the receiver side, the channel transfer function is first extracted at times and frequencies at which pilot symbols have been inserted. Then, the missing values of the transfer function between the pilot symbols are interpolated through filtering [43]. In this case, the performance of pilot-based channel estimation solely depends on the parameter d_T, and thus, we can observe how frequent the pilot symbols should be inserted to track the channel variations and meet the performance of the proposed blind approach.

In Fig. 12, we can see that, for $E_b/N_0 = -6$ dB with $d_T \leq 16$ and for $E_b/N_0 = 4$ dB with $d_T \leq 14$, the pilot-based method outperforms the proposed method by a small margin. However, when E_b/N_0 is increased to 14 dB, the pilot-based method must employ $d_T \leq 4$ and therefore sacrifice bandwidth efficiency to achieve a similar performance as in the proposed blind method. In particular, 1.56% of the bandwidth is then consumed by pilots with the pilot-based method compared to 0.39% with the proposed method when implemented in semiblind form. Frequency correction techniques, often used with pilot-based approaches to improve estimation accuracy, could also be applied to the proposed blind approach, although we did not explicitly consider this aspect. With regard to the implementation of these methods, we have

found that the computational complexity of the proposed blind algorithm is generally comparable to an efficient pilot-based algorithm that incorporates a low-pass filter and spline cubic interpolation.

VII. Conclusion

To estimate wideband TV channels with large Doppler spreads, we typically resort to pilot placements at consecutive OFDM symbol times over specific subcarriers, followed by different interpolation schemes. Traditionally, blind channel estimation has not been applied in such situation, because it normally requires a long observation interval and tends to exhibit a slow convergence rate. The main contribution of this paper has been in developing a new scheme to blindly track a wideband TV MIMO wireless channel that may change at each OFDM symbol time, without using extensive preambles or training sequences. Our approach offers flexibility in choosing the number of transmit and receive antennas and can achieve high bandwidth efficiency with low computational complexity. In particular, compared with a pilot-based approach, the semiblind version of our proposed algorithm consumes significantly less bandwidth to achieve a given BER at moderate to high SNR. Many avenues remain open for future work, including the use of spectral windows in the frequency averaging step, the application of frequency correction techniques, and the development of online schemes for adjusting the algorithm parameters based on available knowledge of channel delay spread and user mobility.

References

[1] K.-D. Kammeyer, V. Kühn, and T. Petermann, "Blind and nonblind turbo estimation for fast-fading GSM channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1718–1728, Sep. 2001.

[2] M. K. Tsatsanis and G. B. Giannakis, "Subspace methods for blind estimation of time-varying FIR channels," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 3084–3093, Dec. 1997.

[3] H. Liu and G. B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 3003–3013, Nov. 1998.

[4] J. K. Tugnait and W. Luo, "Blind identification of time-varying channels using multistep linear predictors," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1739–1749, Jun. 2004.

[5] B. Champagne, A. El-Keyi, and C.-C. Tu, "A subspace method for the blind identification of multiple time-varying FIR channels," in *Proc. IEEE Global Commun. Conf.*, Honolulu, HI, Dec. 2009, pp. 1–6.

[6] C. Tepedelenlioolu and G. B. Giannakis, "Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 2029–2043, Jul. 2000.

[7] Y. Sung and L. Tong, "Tracking of fast-fading channels in long-code CDMA," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 786–795, Mar. 2004.

[8] X. G. Doukopoulos and G. V. Moustakides, "Blind adaptive channel estimation in OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1716–1725, Jul. 2006.

[9] B. Su and P. P. Vaidyanathan, "Subspace-based blind channel identification for cyclic prefix systems using few received blocks," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4979–4993, Oct. 2007.

[10] M. S. Radenkovic, T. Bose, and B. Ramkumar, "Blind adaptive equalization of MIMO systems: New recursive algorithms and convergence analysis," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 57, no. 7, pp. 1475–1488, Jul. 2010.

[11] H. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 986–994, Jun. 2002.

[12] R. C. de Lamare and R. Sampai-Neto, "Reduced-rank adaptive filtering based on joint iterative optimization of adaptive filters," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 980–983, Dec. 2007.

[13] R. C. de Lamare and R. Sampai-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation and filtering," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2503–2513, Jul. 2009.

[14] "Technical Specification Group radio access network: Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)," Third-Generation Partnership Project, Sophia-Antipolis, France, 3GPP TR 25.913 Tech. Rep. V8.0.0 (2008-12), 2008.

[15] H.-G. Jeon and E. Serpedin, "A novel simplified channel tracking method for MIMO-OFDM systems with null subcarriers," *Signal Process.*, vol. 88, no. 4, pp. 1002–1016, Apr. 2008.

[16] J. Gao and H. Liu, "Low-complexity MAP channel estimation for mobile MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 774–780, Mar. 2008.

[17] H. M. Karkhanechi and B. C. Levy, "An efficient adaptive channel estimation algorithm for MIMO OFDM systems—Study of Doppler," *J. Signal Process. Syst.*, vol. 56, no. 2/3, pp. 261–271, Sep. 2009.

[18] D. Angelosante, E. Biglieri, and M. Lops, "Sequential estimation of multipath MIMO-OFDM channels," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3167–3181, Aug. 2009.

[19] X. Dai, H. Zhang, and D. Li, "Linearly time-varying channel estimation for MIMO/OFDM systems using superimposed training," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 681–693, Feb. 2010.

[20] C.-C. Tu and B. Champagne, "Subspace blind MIMO-OFDM channel estimation with short averaging periods: Performance analysis," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2008, pp. 24–29.

[21] C.-C. Tu and B. Champagne, "Subspace-based blind channel estimation for MIMO-OFDM systems with reduced time averaging," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1539–1544, Mar. 2009.

[22] A. F. Molisch, *Wireless Communications*. Chichester, U.K.: Wiley, 2005.

[23] Y.-S. Choi, P. J. Voltz, and F. A. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.

[24] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 784–789, Jan. 2007.

[25] X. Xu, Y. Jing, and X. Yu, "Subspace-based noise variance and SNR estimation for OFDM systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2005, vol. 1, pp. 23–26.

[26] A. Goldsmith, Ed., *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[27] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 95–107, Jan. 1995.

[28] B. Champagne and Q.-G. Liu, "Plane-rotation-based EVD updating schemes for efficient subspace tracking," *IEEE Trans. Signal Process.*, vol. 46, no. 7, pp. 1886–1900, Jul. 1998.

[29] D.-Z. Feng and W. X. Zheng, "Fast approximate inverse power iteration algorithm for adaptive total least squares FIR filtering," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 4032–4039, Oct. 2006.

[30] G. H. Golub and C. F. V. Loan, *Matrix Computation*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[31] D. S. Watkins, *Fundamentals of Matrix Computations*. New York: Wiley, 2002.

[32] I. C. F. Ipsen, *Numerical Matrix Analysis: Linear Systems and Least Squares*. Philadelphia, PA: SIAM, 2009.

[33] N. L. Owsley, "Adaptive data orthogonalization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 1978, vol. 3, pp. 109–112.

[34] C.-C. Tu and B. Champagne, "On convergence properties of subspace trackers based on orthogonal iteration," in *Proc. IEEE Pac. Rim Conf. Commun., Comput., Signal Process.*, 2009, pp. 65–70.

[35] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: Unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.

[36] P. C. Hansen, "The 2-norm of random matrices," *J. Comput. Appl. Math.*, vol. 23, no. 1, pp. 117–120, Jul. 1988.

[37] Z. Liu, Y. Xin, and G. B. Giannakis, "Linear constellation precoding for OFDM with maximum multipath diversity and coding gains," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 416–427, Mar. 2003.

[38] R. Lin and A. P. Petropulu, "Linear precoding-assisted blind channel estimation for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 54, no. 3, pp. 983–995, May 2005.

[39] C.-C. Tu and B. Champagne, "Subspace tracking of fast time-varying channels in precoded MIMO-OFDM systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2009, pp. 2565–2568.

[40] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas With Application to Linear Systems Theory*. Princeton, NJ: Princeton Univ. Press, 2005.

[41] "Technical Specification Group radio access network: Spatial channel model for multiple-input–multiple-output simulations (Rel. 6)," Third-

Generation Partnership Project, Sophia-Antipolis, France, 3GPP TR 25.996 Tech. Rep. V6.0.0 (2003-09), 2003.

[42] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications, and Contol.* Upper Saddle River, NJ: Prentice-Hall, 1995.

[43] 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.

Chao-Cheng Tu received the B.S. and M.S. degrees in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, in 1996 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering from McGill University, Montréal, QC, Canada, in 2010.

After the two-year obligatory military service in Taiwan, he was with ACER Communications and Multimedia Inc. as a Hardware Research and Development Engineer. In 2002, he was with the Department of Electronics Engineering, Southern Taiwan University of Technology, Tainan, as a part-time Lecturer. He has been with the Department of Electrical and Computer Engineering, McGill University, as a Research Assistant and a Teaching Assistant. He is currently with Interdigital Canada Ltée, Montréal, as a System Engineer. His research interests include wireless communications and signal processing for communications systems.

Dr. Tu received both the Best Student Paper Award at the IEEE Wireless Telecommunication Symposium in 2006 and the Student Travel Grant from the IEEE International Conference on Acoustics, Speech, and Signal Processing in 2009.

Benoît Champagne (S'87–M'89–SM'03) was born in Joliette, QC, Canada, in 1961. He received the B.Ing. degree in engineering physics from the Ecole Polytechnique de Montréal, Montréal, QC, Canada, in 1983, the M.Sc. degree in physics from the Université de Montréal, in 1985, and the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1990.

From 1990 to 1999, he was an Assistant Professor and then an Associate Professor with the Institut National de la Recherche Scientifique (INRS)–Énergie, Matériaux et Télécommunications, Université du Québec, Montréal. In September 1999, he was an Associate Professor with the Department of Electrical and Computer Engineering, McGill University, Montréal, where he was the Associate Chairman of Graduate Studies from January 2004 to August 2007 and is currently a Full Professor. He has been an Associate Editor of the *EURASIP Journal on Applied Signal Processing*. His research interests include statistical signal processing, including signal/parameter estimation, sensor array processing, and adaptive filtering, and applications thereof to digital communications systems.

Dr. Champagne has served as an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS and is currently an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.