Efficient Channel Estimation for MIMO Single-Carrier Block Transmission With Dual Cyclic Timeslot Structure

Xiqi Gao, *Senior Member, IEEE*, Bin Jiang, *Student Member, IEEE*, Xiaohu You, *Member, IEEE*, Zhiwen Pan, Yisheng Xue, *Member, IEEE*, and Egon Schulz

***Abstract*—We investigate channel estimation for timeslot-structured single-carrier block transmission (SCBT) over space-, time-, and frequency-selective fading multiple-input multiple-output (MIMO) channels. A MIMO-SCBT with a dual cyclic timeslot structure is presented first. Then, an optimal channel estimation in the minimal mean square error (MMSE) sense on the timeslot basis is investigated. It is shown that the optimal pilots for the timeslot-based MMSE channel estimation are related to the statistical channel state information in eigenmode. Under the assumption that the transmit correlation is unknown at the transmitter, the optimal pilots satisfy the same condition as reported for the block-based least-square (LS) channel estimation in literature, and the channel estimation can be simplified to initial block-based LS channel estimation followed by space–time postprocessing. Particularly, for spatially uncorrelated channels, the space–time postprocessing can be reduced to pathwise processing. A new design of the pilot sequences is given, which leads to an efficient implementation of the channel estimation. Later on, a more efficient implementation for the initial channel estimation is obtained by using the structure of the pilot sequences, and discrete cosine transform (DCT)-based implementation is developed for the space–time postprocessing to approximate the optimal solution with low implementation complexity. Finally, the performance of the proposed channel estimation is verified via simulations.**

***Index Terms*—Channel estimation, multiple-input multiple-output (MIMO), single-carrier block transmission (SCBT).**

I. Introduction

HIGH data rate wireless transmission with high spectral efficiency and power efficiency is required to meet the rapidly increasing demands for mobile communications [1], [2]. Orthogonal frequency division multiplexing (OFDM) technique has attracted much attention and found many applications due to its simple implementation, robustness against frequency-selective fading channels, and relative easiness to employ multiple-antenna transmission techniques [3], [4]. The aforementioned advantages of the OFDM technique mainly benefit from the cyclic prefix inserted in each OFDM symbol that converts the linear convolution channel to a cyclic one and keeps the orthogonality of the OFDM signal even in multipath channels. A single-carrier block transmission (SCBT) with a cyclic prefix or a zero-padded suffix has appeared to be an alternative promising technique and has also received much attention [5]–[10]. The SCBT with a cyclic prefix is often referred to as single-carrier frequency-domain equalization (SC-FDE) when FDE is used, and it is also known as cyclic-prefixed single-carrier (CP-SC) transmission. It almost keeps the same advantages as the OFDM technique, while avoiding the shortcomings of the OFDM in peak-to-average power ratio (PAPR) and sensitivity to frequency shifts. A generalized multicarrier transmission that combines a set of SCBTs through multicarrier filter banks has been proposed to support broadband wireless transmission and asynchronous multiple access in cellular systems [2], [11].

In wireless communication environments, accurate channel estimation plays an important role in improving the overall system performance, especially when the channels are fast fading [12]–[17]. For practical applications, pilot-aided channel estimation is more attractive due to its simplicity and reliability, compared with blind and decision-directed methods [17], [18]. When multiple transmit and multiple receive antennas are employed to increase the link capacity or reliability, the number of the channel parameters to be estimated increases in proportion to that of the transmit antennas and the receive antennas. It is, therefore, very important, challenging as well, to design accurate and efficient channel estimation schemes for real multiple-input multiple-output (MIMO) systems. In this paper, we investigate the channel estimation for SCBT over triply selective, i.e., space-, time-, and frequency-selective MIMO channels [19].

In recent years, several works on the pilot (training)-aided channel estimation for SCBT over single-antenna and multiantenna channels have been reported [16]–[23]. The theoretical design of optimal training for SCBT over doubly selective (i.e., time and frequency selective) single-antenna

Paper approved by A. Lozano, the Editor for Wireless Network Access and Performance of the IEEE Communications Society. Manuscript received December 8, 2005; revised August 19, 2006 and December 28, 2006. This work was supported in part by the National Natural Science Foundation of China under Grant 60572072 and Grant 60496311, in part by the China High-Tech 863-FuTURE Project under Grant 2006AA01Z264 and Grant 2003AA123310, and in part by the International Cooperation Project on Beyond 3G Mobile of China under Grant 2005DFA10360.

X. Q. Gao, B. Jiang, X. H. You, and Z. W. Pan are with the National Mobile Communication Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: xqgao@seu.edu.cn; bjiang@seu.edu.cn; xhyu@seu.edu.cn; pzw@seu.edu.cn).

Y. S. Xue is with the RTS, Corporate Technology, Siemens Ltd. China, Beijing 100094, China (e-mail: yisheng.xue@siemens.com).

E. Schulz is with the Future Radio Concepts, Information and Communication Mobile Networks, Siemens AG, Munich D-81541, Germany (e-mail: egon.schulz@siemens.com).

Digital Object Identifier 10.1109/TCOMM.2007.908551

channels is presented in [16] and extended to MIMO channels in [17]. It is shown that the optimal training design obtained by maximizing a tight lower bound on the average channel capacity is equivalent to that obtained by minimizing the minimum mean-square channel estimation error, and the optimal training strategy consists of equispaced and equipowered pilot symbols inserted periodically with the period dictated by the channels' Doppler spread. Training-aided frequency-domain-equalized single-carrier (TASC) scheme is proposed in [20], where the cyclic prefix is replaced by a training sequence. This approach is also discussed in [5] and [6]. The TASC scheme enables better synchronization and possible channel estimation with additional known symbols. By adding a known suffix to each symbol block before performing the cyclic prefixing, a pilot cyclic prefixed single-carrier (PCP-SC) scheme is proposed in [21]. With the aid of the CP and the suffix, the least square (LS) channel estimation can be performed by the fast Fourier transform (FFT). It is shown that the additional known symbols also ensure the unknown symbols recoverable regardless of the channel null locations. The LS channel estimation for MIMO SC-FDE is proposed in [22] and [23], where the optimal pilot sequences for multiple transmit antennas are designed to achieve an optimal LS estimate in minimal mean square error (MMSE) sense. Armed with the optimal pilots, the LS channel estimation for MIMO SC-FDE can be fast implemented via FFT as that in [21]. In our paper, we are interested in a timeslot-structured SCBT over MIMO channels, where each timeslot consists of several transmission blocks. A dual cyclic timeslot structure is designed for the SCBT. The channel estimation and the fast implementation are investigated with the optimality on the timeslot basis.

The rest of the paper is outlined as follows. In Section II, we present the MIMO-SCBT with dual cyclic timeslot structure and establish the signal model for channel estimation. The optimal channel estimation in the MMSE sense on the timeslot basis is investigated in Section III. We first deal with the channel estimation for the pilot segments only, and then, extend to the data segments. It is shown that the optimal pilots for the timeslot-based MMSE channel estimation are related to the statistical channel state information (CSI) in eigenmode. Under the assumption that the transmit correlation is unknown at the transmitter in designing pilots for practical applications, the optimal pilots satisfy the same condition as reported in [22] and [23] for the block-based LS channel estimation, and the channel estimation can be simplified to initial block-based LS channel estimation followed by space–time postprocessing, which can be reduced to the pathwise processing for spatially uncorrelated MIMO channel. The design of the pilot sequences is also involved in Section III. Fast implementations for the initial channel estimation and the space-time postprocessing are proposed in Section IV. A more efficient implementation for the initial channel estimation is obtained by exploiting the structure of the pilot sequences, and discrete cosine transform (DCT)-based implementation for the space–time postprocessing is developed to approximate the optimal solution with low implementation complexity. Simulation results are provided in Section V, and Section VI concludes the paper.

Notations: We will use upper (lower) boldface letters to denote matrices (column vectors), sometimes with subscripts to emphasize their sizes. An $N \times N$ identity matrix will be denoted as $\mathbf{I}_N$, and an all-zero matrix will be denoted as $\mathbf{0}$. The kth column vector of an identity matrix is a unit vector and is denoted by $\mathbf{e}_k$. Superscript H will denote Hermitian transpose, T transpose, and $\dagger$ pseudoinverse. We will reserve $\otimes$ for Kronecker product and $E\{\cdot\}$ for expectation with respect to all random variables within the brackets. We will use $[\mathbf{A}]_{k,l}$ to denote the (k,l)th entry of a matrix $\mathbf{A}$, $\mathrm{tr}\{\mathbf{A}\}$ for its trace, and $[\mathbf{a}]_k$ for the kth entry of the column vector $\mathbf{a}$. $\mathrm{diag}\{\mathbf{x}\}$ will stand for a diagonal matrix with $\mathbf{x}$ on its main diagonal, and $\mathrm{circ}\{\mathbf{x}\}$ will stand for a column circular matrix with the first column $\mathbf{x}$, and $\mathbf{v} = \mathrm{vec}\{\mathbf{V}\}$ is the stacking operator. The notation $((n))_N$ stands for n modulo N, $\lfloor x \rfloor$ ($\lceil x \rceil$) denotes the largest (smallest) integer no more (less) than x. $\mathbf{\Pi}_N^n$ stands for the $N \times N$ permutation matrix defined as

$$\mathbf{\Pi}_N^n = \begin{bmatrix} \mathbf{0} & \mathbf{I}_n \\ \mathbf{I}_{N-n} & \mathbf{0} \end{bmatrix}.$$

II. MIMO-SCBT With Dual Cyclic Timeslot Structure

A. System Model

In an SCBT system, the signal is transmitted block by block, where each block typically consists of a data sequence and a cyclic prefix or a zero-padded suffix [5]–[10]. In this section, we present a timeslot-structured SCBT over MIMO channels with dual cyclic timeslot structure. The transmitted signal sequence is composed of one-by-one timeslots, which are supposed to support time-division multiple access (TDMA). In each timeslot, there are several transmission blocks that we refer to as subtimeslots.

A MIMO link with N_T transmit and N_R receive antennas over a triply selective fading channel constitutes the scenario of interest of this paper. Fig. 1 shows the block diagram of the MIMO SCBT. At the transmitter, the input bit stream is passed through error control coding and modulation symbol mapping followed by space–time transmit processing, yielding the N_T transmitted data sequences $d_{n,i}(l)$, where $n = 0, 1, \ldots, N_T - 1$. The variable $d_{n,i}(l)$ represents the transmitted data sequence on the nth transmit antenna during the ith timeslot. The space–time processing could be space–time coding, spatial multiplexing, or linear precoding with CSI feedback. The N_T transmitted digital baseband signals $x_{n,i}(l)$ are generated by intermittently inserting pilot and guard signals in the transmitted data sequence. At the receiver, from the N_R received digital baseband signals $y_{m,i}(l)$, where $m = 0, 1, \ldots, N_R - 1$, the impulse response of the MIMO channels are obtained through pilot-aided channel estimation, and the reconstructed bit stream is obtained by space–time coherent detection followed by demapping and decoding. The space–time coherent detection could be MIMO channel-equalization- or interference-cancelation-based detection with or without feedback information from the decoder [24], [25].

Fig. 2 shows the dual cyclic timeslot structure. Each timeslot consists of K subtimeslots and a tail. Each subtimeslot consists of a cyclic guard G, a pilot P, and user data D, and the tail

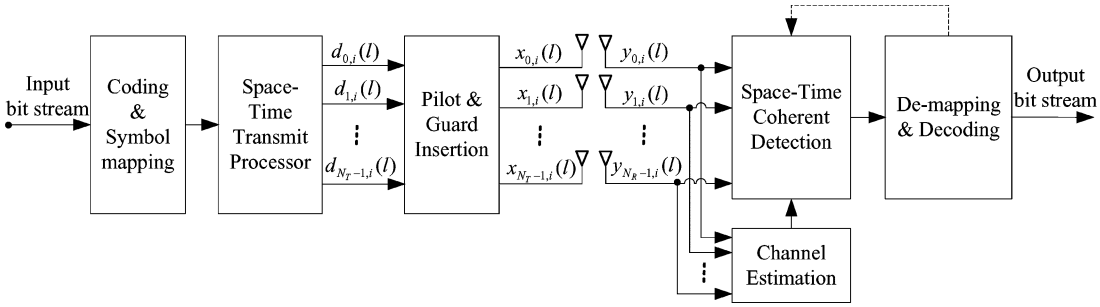

Fig. 1. MIMO SCBT with N_T transmit and N_R receive antennas.

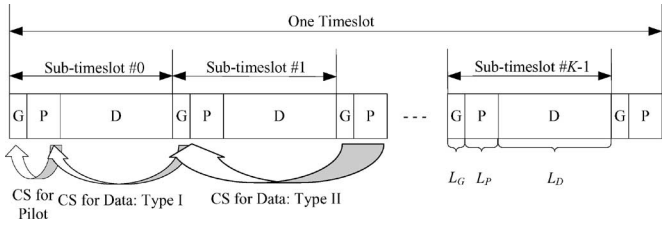

Fig. 2. Dual cyclic timeslot structure for the MIMO SCBT, where CS represents the cyclic structure, type I CS is for a segment consisting of D plus G, and type II CS is for a segment consisting of D plus G and P.

consists of a cyclic guard G and a pilot P. The length of the cyclic guard L_G should be no less than the maximum channel delay spread N_P, and less than that of the pilot sequence L_P of which the last L_G data constitute the cyclic guard. The pilot sequences are the same for different subtimeslots. According to the sampling theorem, the spacing between two successive pilot segments that is equal to the length of the subtimeslot must satisfy the following: $L_S = L_G + L_P + L_D \leq 1/(2f_D T_S)$, where f_D is the maximum Doppler frequency offset, T_S is the symbol interval, and L_D is the length of the user data. This allows one to obtain the channel parameters of user data segments from that of the pilot segments.

From Fig. 2, it can be observed that there is a cyclic guard prior to each pilot segment, and there is also a "cyclic guard" G (and P) prior to each segment consisting of D plus G (and P). This dual cyclic structure is helpful for both channel estimation and signal detection in the receiver. With a cyclic guard prior to each pilot segment, the received signal in each pilot segment appears to be a circular convolution of the transmitted pilots with the channel impulse response. In this way, the interpath interference can be avoided, which allows low-complexity channel estimation with high performance. This will become clear in the subsequent sections. With the implicit cyclic guard prior to each segment consisting of *D* plus *G* (and *P*), the received signal in such a segment also appears to be a circular convolution of the corresponding transmitted signal with the channel impulse response, and thus, discrete Fourier transform (DFT)-based detection can be performed when the channel approximates constant in such a segment [5], [6], [24]. This requires that L_S should be much less than $1/(2f_D T_S)$. In a fast fading channel, the channel may change in each data segment, and there exists a performance loss with DFT-based detection. One solution is to separate a data segment into several subblocks with a cyclic prefix inserted to each subblock, which implies a loss in spectral efficiency. An efficient way is to deploy interference-cancelation-based detection [25]. This is beyond the scope of this paper.

The implicit cyclic structure for unknown user data employs a known constant data block as a cyclic prefix instead of the one taken from the unknown user data in the conventional cyclic prefixed SCBT. Similar signaling has also been suggested in [20] and [21]. In [20], the known data block is proposed to serve as a training sequence and a cyclic prefix, and it is shown that the known symbols enable better synchronization in addition to saving the overhead of the cyclic prefix. The known data blocks are proposed in [21] to serve as a training sequence, a cyclic prefix, and a suffix to each data block, and it is shown that channel responses can be resolved by the FFT, and the known symbols ensure data symbols recoverable regardless of the channel null location. The timeslot structure shown in Fig. 2 gives a compact design for the timeslot-structured SCBT supporting the TDMA. Different from [21], the suffix to each data block is also implicit, which removes the additional overhead.

B. Signal Model for Channel Estimation

The channel is assumed to be time invariant during each pilot segment, and this assumption is neither true nor necessary for the channel between different pilot segments. Depending on the rate at which the channel is changing with time, there may or may not be a need to further track the channel variations during the data segment. In this paper, we are concerned with the problem of channel estimation during both the pilot segments and the data segments.

Let $\mathbf{s}_n = [s_n(0), s_n(1), \ldots, s_n(L_P - 1)]^T$ be the block of L_P pilot symbols transmitted from the *n*th transmit antenna. We can write the received signal vector from the *m*th receive antenna for the kth pilot segment after the removal of CP as

$$\mathbf{y}_m(k) = \sum_{n=0}^{N_T-1} \sum_{p=0}^{N_P-1} h_{m,n}^p(k)\mathbf{s}_n^p + \mathbf{z}_m(k), \quad k = 0, 1, \ldots, K \tag{1}$$

where $\mathbf{s}_n^p = \mathbf{\Pi}_{L_p}^p \mathbf{s}_n$ is the *p*th cyclic shift version of $\mathbf{s}_n$; $h_{m,n}^p(k)$ is the *p*th multipath coefficient of the MIMO channel from transmit antenna *n* to receive antenna *m*; and $\mathbf{z}_m(k)$ is the additive white Gaussian noise (AWGN) vector with zero mean and covariance matrix $\sigma_z^2 \mathbf{I}_{L_P}$. Let $\mathbf{y}(k) =$

$[\mathbf{y}_0^T(k), \mathbf{y}_1^T(k), \ldots, \mathbf{y}_{N_R-1}^T(k)]^T$, then we have

$$\mathbf{y}(k) = (\mathbf{I}_{N_R} \otimes \mathbf{S})\mathbf{h}(k) + \mathbf{z}(k) \tag{2}$$

where $\mathbf{h}(k)$ is a vector of size $N_R N_T N_P$ with $[\mathbf{h}(k)]_{mN_TN_P+nN_P+p} = h_{m,n}^p(k)$, $\mathbf{S} = [\mathbf{S}_0, \mathbf{S}_1, \ldots, \mathbf{S}_{N_T-1}]$ with $\mathbf{S}_n = [\mathbf{s}_n^0, \mathbf{s}_n^1, \ldots, \mathbf{s}_n^{N_P-1}]$, and $\mathbf{z}(k) = [\mathbf{z}_0^T(k), \mathbf{z}_1^T(k), \ldots, \mathbf{z}_{N_R-1}^T(k)]^T$.

Conventionally, the channel estimation can be performed to obtain initial channel estimates from (2) and enhanced ones by using the correlation in the time domain and/or spatial domain. To achieve optimal channel estimation in the whole timeslot, one needs to deal with the received pilot signals simultaneously. Stacking $\mathbf{y}(k), k = 0, 1, \ldots, K$, into a vector $\mathbf{y} = [\mathbf{y}^T(0), \mathbf{y}^T(1), \ldots, \mathbf{y}^T(K)]^T$, we have the expression

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{z} \tag{3}$$

where $\mathbf{X} = \mathbf{I}_{(K+1)N_R} \otimes \mathbf{S}, \mathbf{h} = [\mathbf{h}^T(0), \mathbf{h}^T(1), \ldots, \mathbf{h}^T(K)]^T$, and $\mathbf{z} = [\mathbf{z}^T(0), \mathbf{z}^T(1), \ldots, \mathbf{z}^T(K)]^T$.

III. Channel Estimation and Pilot Design

A. Channel Estimation for Pilot Segments

We start from the channel estimation for pilot segments, of which the task is to recover $\mathbf{h}$ from the observation $\mathbf{y}$ and the known $\mathbf{X}$. The channel estimation will be extended to the data segments in the Section III-B.

From (3), the linear MMSE solution of $\mathbf{h}$ is [26]

$$\hat{\mathbf{h}} = \mathbf{C}_{\text{opt}}\mathbf{y} \tag{4}$$

where $\mathbf{C}_{\text{opt}}$ is a $(K+1)N_R N_T N_P \times (K+1)N_R L_P$ matrix given by

$$\begin{aligned}\mathbf{C}_{\text{opt}} &= \arg\min_{\mathbf{C}} E\{\|\mathbf{h} - \hat{\mathbf{h}}\|^2\} \\ &= \mathbf{R}_{\mathbf{h}}\mathbf{X}^H(\mathbf{X}\mathbf{R}_{\mathbf{h}}\mathbf{X}^H + \sigma_z^2\mathbf{I}_{(K+1)N_R L_P})^{-1} \\ &= \mathbf{R}_{\mathbf{h}}(\mathbf{X}^H\mathbf{X}\mathbf{R}_{\mathbf{h}} + \sigma_z^2\mathbf{I}_{(K+1)N_R N_T N_P})^{-1}\mathbf{X}^H \end{aligned} \tag{5}$$

where $\mathbf{R}_{\mathbf{h}} = E\{\mathbf{h}\mathbf{h}^H\}$, and the noise correlation matrix is assumed to be $\mathbf{R}_{\mathbf{z}} = \sigma_z^2\mathbf{I}_{(K+1)N_R L_P}$.

The spatial, path, and temporal correlations are involved in the channel correlation matrix $\mathbf{R}_{\mathbf{h}}$. Spatial correlation can often be characterized by the Kronecker product of transmit and receive correlation matrices [19]. We assume that the path correlation and the temporal correlation are independent of the transmit and receive correlations. Then, the channel correlation matrix can be expressed by a Kronecker-type structure as [19]

$$\mathbf{R}_{\mathbf{h}} = \mathbf{R}_{DPR} \otimes \mathbf{R}_{RX} \otimes \mathbf{R}_{TX} \otimes \mathbf{R}_{ISI} \tag{6}$$

where $\mathbf{R}_{TX}$ and $\mathbf{R}_{RX}$ are the transmit and receive antenna correlation matrices measured at the first path, $\mathbf{R}_{ISI}$ is the path correlation matrix measured at any fixed subchannel, and $\mathbf{R}_{DPR}$ is the temporal correlation matrix that is determined by the Doppler frequency offset and the length of the subtimeslot.

Assume that the spectrum decompositions of Hermitian matrices $\mathbf{R}_{DPR}$, $\mathbf{R}_{RX}$, and $\mathbf{R}_{TX}$ can be expressed as $\mathbf{R}_{DPR} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, $\mathbf{R}_{RX} = \mathbf{U}_R\mathbf{\Lambda}_R\mathbf{U}_R^H$, and $\mathbf{R}_{TX} = \mathbf{U}_T\mathbf{\Lambda}_T\mathbf{U}_T^H$, respectively, where $\mathbf{\Lambda} = \text{diag}\{\lambda_0, \ldots, \lambda_K\}$, $\mathbf{\Lambda}_T = \text{diag}\{\lambda_{T,0}, \ldots, \lambda_{T,N_T-1}\}$, and $\mathbf{\Lambda}_R = \text{diag}\{\lambda_{R,0}, \ldots, \lambda_{R,N_R-1}\}$. Furthermore, we assume that the channel parameters are uncorrelated for different paths [27], i.e., $\mathbf{R}_{ISI} = \mathbf{\Lambda}_{ISI} = \text{diag}\{\rho_0^2, \rho_1^2, \ldots, \rho_{N_P-1}^2\}$, where ρ_p^2 is the average power of the pth path and $\sum_{p=0}^{N_P-1}\rho_p^2 = 1$. Then, the spectrum decomposition of $\mathbf{R}_{\mathbf{h}}$ is

$$\mathbf{R}_{\mathbf{h}} = \tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{U}}^H \tag{7}$$

where $\tilde{\mathbf{U}} = \mathbf{U} \otimes \mathbf{U}_R \otimes \mathbf{U}_T \otimes \mathbf{I}_{N_P}$, and $\tilde{\mathbf{\Lambda}} = \mathbf{\Lambda} \otimes \mathbf{\Lambda}_R \otimes \mathbf{\Lambda}_T \otimes \mathbf{\Lambda}_{ISI}$.

The mean square error of the channel estimate given by (4) can be expressed as

$$\begin{aligned}\sigma_{\hat{\mathbf{h}}}^2 &= E\{\|\mathbf{h} - \hat{\mathbf{h}}\|^2\} \\ &= \text{tr}\left\{\mathbf{R}_{\mathbf{h}}\left(\frac{1}{\sigma_z^2}\mathbf{X}^H\mathbf{X}\mathbf{R}_{\mathbf{h}} + \mathbf{I}_{(K+1)N_R N_T N_P}\right)^{-1}\right\}. \end{aligned} \tag{8}$$

Optimal pilot is assumed to achieve minimal $\sigma_{\hat{\mathbf{h}}}^2$ with the constraint on the total power of the transmitted pilot signals. In [22] and [23], it has been shown that with the LS channel estimation performed on each cyclic prefixed pilot segment, the optimal pilot should satisfy the orthogonal condition, i.e., $\mathbf{S}^H\mathbf{S}$ is equal to an identity matrix up to a factor. In our case, with the MMSE channel estimation performed on the whole timeslot, we have the following theorem.

Theorem 1: Let $\text{tr}\{\mathbf{S}^H\mathbf{S}\} = N_T N_P L_P$. For a triply selective MIMO channel with the correlation matrix given by (6) and known at the transmitter, the mean square error of the linear MMSE channel estimate given by (4) can achieve the following lower bound:

$$\sigma_{\hat{\mathbf{h}}}^2 \geq \breve{\sigma}_{\hat{\mathbf{h}}}^2 = \tilde{\sigma}_{\hat{\mathbf{h}}}^2(\mathbf{\Gamma}_{\text{opt}}) \tag{9}$$

if and only if the following condition is satisfied:

$$\mathbf{S}^H\mathbf{S} = L_P(\mathbf{U}_T\mathbf{\Gamma}_{\text{opt}}\mathbf{U}_T^H \otimes \mathbf{I}_{N_P}) \tag{10}$$

where

$$\tilde{\sigma}_{\hat{\mathbf{h}}}^2(\mathbf{\Gamma}) = \sum_{k,m,n,p} \lambda_k\lambda_{R,m}\lambda_{T,n}\rho_p^2\left(1 + \frac{L_P}{\sigma_z^2}\lambda_k\lambda_{R,m}\lambda_{T,n}\rho_p^2\gamma_n\right)^{-1} \tag{11}$$

with $\mathbf{\Gamma} := \text{diag}\{\gamma_0, \ldots, \gamma_{N_T-1}\}$, and $\mathbf{\Gamma}_{\text{opt}}$ is determined by

$$\mathbf{\Gamma}_{\text{opt}} = \arg\min_{\mathbf{\Gamma}:\ \text{tr}\{\mathbf{\Gamma}\} \leq N_T, \mathbf{\Gamma} \geq 0} \tilde{\sigma}_{\hat{\mathbf{h}}}^2(\mathbf{\Gamma}). \tag{12}$$

For proof, see Appendix I.

Theorem 1 implies that the optimal pilots are related to the statistical CSI in eigenmode as explained later. Let $\tilde{\mathbf{S}} = [\tilde{\mathbf{s}}_0, \tilde{\mathbf{s}}_1, \ldots, \tilde{\mathbf{s}}_{N_T-1}]$ and $\tilde{\mathbf{S}}_p = \mathbf{\Pi}_{L_p}^p\tilde{\mathbf{S}}$. From (10), we have that $\tilde{\mathbf{S}}^H\tilde{\mathbf{S}} = L_P\mathbf{U}_T\mathbf{\Gamma}_{\text{opt}}\mathbf{U}_T^H$ and $\tilde{\mathbf{S}}_p^H\tilde{\mathbf{S}}_q = \mathbf{0}$ for $p \neq q$. Therefore, the transmit pilot matrix $\tilde{\mathbf{S}}$ can be expressed as $\tilde{\mathbf{S}} = \breve{\mathbf{S}}\mathbf{\Gamma}_{\text{opt}}^{1/2}\mathbf{U}_T^H$, where $\breve{\mathbf{S}}$ satisfies the orthogonal conditions $\breve{\mathbf{S}}^H\breve{\mathbf{S}} = L_P\mathbf{I}_{N_T}$ and $(\mathbf{\Pi}_{L_p}^p\breve{\mathbf{S}})^H(\mathbf{\Pi}_{L_p}^q\breve{\mathbf{S}}) = \mathbf{0}$ for $p \neq q$. This means that the optimal pilots should be orthogonal sequences transmitted along the eigendirections of the transmit antenna correlation matrix with the powers specified by (12). The constrained nonlinear optimizations in (12) is the so-called water filling problem. The

closed-form solution for the general case of $\mathbf{R_h}$ has not been found, even for flat fading MIMO channels [28].

Corollary 1: Let $\mathrm{tr}\{\mathbf{S}^H\mathbf{S}\} = N_T N_P L_P$. For a triply selective MIMO channel with the correlation matrix given by (6) and the transmit antenna correlation unknown at the transmitter, the mean square error of the linear MMSE channel estimate given by (4) can achieve the following lower bound:

$$\sigma^2_{\hat{\mathbf{h}}} \geq \breve{\sigma}^2_{\hat{\mathbf{h}}} = \sum_{k,m,n,p} \lambda_k \lambda_{R,m} \lambda_{T,n} \rho_p^2 \left(1 + \frac{L_P}{\sigma_z^2} \lambda_k \lambda_{R,m} \lambda_{T,n} \rho_p^2\right)^{-1} \tag{13}$$

if and only if the following condition is satisfied:

$$\mathbf{S}^H\mathbf{S} = L_P \mathbf{I}_{N_T N_P}. \tag{14}$$

Without the knowledge about the transmit correlation at the transmitter, the transmit antenna correlation matrix can be set to be an identity matrix, and thus, the proof of Corollary 1 is straightforward by using Theorem 1. The orthogonal condition (14) that the optimal pilots satisfy is the same condition as reported in [22] and [23] for the block-based LS channel estimation.

Although the correlations between transmit antennas may or may not exist in realistic environments, it is reasonable to set the transmit antenna correlation matrix to be identity when designing the optimal pilots for practical applications. First of all, to obtain the optimal training sequences determined by (10) and (12), the estimated statistical CSI obtained at the receiver should be fed back to the transmitter, which complicates the system realizations. Second, the transmitted training signals may suffer from high PAPR due to the "water filling" operation. Moreover, the SCBT technique is often suggested for uplink transmission [5] where rich scattering is assumed so that transmit antennas are usually considered as uncorrelated. These reasons motivate us to design pilots without the knowledge about the transmit antenna correlation at the transmitter.

With the relations (7) and (14), the linear MMSE channel estimate can be rewritten as

$$\hat{\mathbf{h}} = \tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\left(L_P\tilde{\mathbf{\Lambda}} + \sigma_z^2 \mathbf{I}_{(K+1)N_R N_T N_P}\right)^{-1} \tilde{\mathbf{U}}^H \mathbf{X}^H \mathbf{y}. \tag{15}$$

Let $\hat{\mathbf{h}}_{\mathrm{ini}} = (1/L_P)\mathbf{X}^H\mathbf{y}$ denote the initial channel estimate, with optimal pilot satisfying the orthogonal condition of (14), where $\hat{\mathbf{h}}_{\mathrm{ini}}$ is the LS channel estimate. Then, (15) can be rewritten as

$$\hat{\mathbf{h}} = \tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\left(\tilde{\mathbf{\Lambda}} + \frac{\sigma_z^2}{L_P}\mathbf{I}_{(K+1)N_R N_T N_P}\right)^{-1} \tilde{\mathbf{U}}^H \hat{\mathbf{h}}_{\mathrm{ini}}. \tag{16}$$

From the block diagonal structure of $\mathbf{X}$, the initial channel estimation can be performed on each pilot segment and each receive antenna separately. With $\hat{\mathbf{h}}_{\mathrm{ini},m}(k)$ representing the corresponding initial channel estimate on the kth pilot segment and the mth receive antenna, we have

$$\hat{\mathbf{h}}_{\mathrm{ini},m}(k) = \frac{1}{L_P}\mathbf{S}^H \mathbf{y}_m(k). \tag{17}$$

Let $\mathbf{h}^p$ be a vector of size $(K+1)N_R N_T$ with $[\mathbf{h}^p]_{kN_R N_T + mN_T + n} = h^p_{m,n}(k)$, and $\hat{\mathbf{h}}^p$ and $\hat{\mathbf{h}}^p_{\mathrm{ini}}$ be the corresponding MMSE and initial estimates. Then, the following relation holds:

$$\hat{\mathbf{h}}^p = \mathbf{U}_{st}\mathbf{\Lambda}_{st,p}\mathbf{U}_{st}^H \hat{\mathbf{h}}^p_{\mathrm{ini}} \tag{18}$$

where $\mathbf{U}_{st} = \mathbf{U} \otimes \mathbf{U}_R \otimes \mathbf{U}_T$ and $\mathbf{\Lambda}_{st,p} = (\rho_p^2 \mathbf{\Lambda} \otimes \mathbf{\Lambda}_R \otimes \mathbf{\Lambda}_T)(\rho_p^2 \mathbf{\Lambda} \otimes \mathbf{\Lambda}_R \otimes \mathbf{\Lambda}_T + (\sigma_z^2/L_P)\mathbf{I}_{(K+1)N_R N_T})^{-1}$. The previous two equations imply that the MMSE channel estimation can be performed by initial channel estimation on each pilot segment and each receive antenna followed by space–time filtering.

For spatially uncorrelated MIMO channels, the transmit and receive antenna correlation matrices $\mathbf{R}_{TX}$ and $\mathbf{R}_{RX}$ are identity, i.e., $\mathbf{R}_{TX} = \mathbf{I}_{N_T}$ and $\mathbf{R}_{RX} = \mathbf{I}_{N_R}$. In this case, we have that $\mathbf{U}_{st} = \mathbf{U} \otimes \mathbf{I}_{N_R N_T}$ and $\mathbf{\Lambda}_{st,p} = \mathbf{\Lambda}_p \otimes \mathbf{I}_{N_R N_T}$, where $\mathbf{\Lambda}_p = \rho_p^2 \mathbf{\Lambda}(\rho_p^2 \mathbf{\Lambda} + (\sigma_z^2/L_P)\mathbf{I}_{K+1})^{-1}$. Let $\mathbf{h}^p_{m,n} = [h^p_{m,n}(0), h^p_{m,n}(1), \ldots, h^p_{m,n}(K)]^T$, and $\hat{\mathbf{h}}^p_{m,n}$ and $\hat{\mathbf{h}}^p_{\mathrm{ini},m,n}$ be the corresponding MMSE and initial estimates. From (18), we have

$$\hat{\mathbf{h}}^p_{m,n} = \mathbf{U}\mathbf{\Lambda}_p\mathbf{U}^H \hat{\mathbf{h}}^p_{\mathrm{ini},m,n}. \tag{19}$$

This means that the space–time filtering as in (18) can be reduced to pathwise processing.

B. Extension to the Data Segments

With the channel estimates of all the pilot segments in a timeslot, one could perform channel interpolation in the time domain to get the channel parameters of data segments. In this section, we deal with the optimal channel estimation to cover data segments in addition to the pilot segments.

As mentioned in Section II, we assume that the channel is time invariant during each pilot segment and possibly time variant between different pilot segments. Now, we are supposed to estimate the channel responses of each segment of length L_S/M starting from the first pilot segment in a timeslot. Here, M is an integer. We assume that L_S is a multiple of M, and the channel is time invariant in each segment of length L_S/M. We denote the $M-1$ channel vectors between the kth and $(k+1)$th pilot segments as $\mathbf{h}(k + l/M)$, where $l = 1, 2, \ldots, M-1$. Stacking all the $KM+1$ channel vectors into an extended channel vector $\underline{\mathbf{h}}$ in the increasing order of the time index, we have the expression

$$\mathbf{h} = \mathbf{A}\underline{\mathbf{h}} \tag{20}$$

where $\mathbf{A} = \mathbf{B} \otimes \mathbf{I}_{N_T N_R N_P}$, and $\mathbf{B}$ is composed of the (kM)th rows of $\mathbf{I}_{KM+1}$ with $k = 0, 1, \ldots, K$. Throughout this paper, the lower bar of variables is reserved for the extensions. Substituting (20) into (3) yields

$$\mathbf{y} = \mathbf{X}\mathbf{A}\underline{\mathbf{h}} + \mathbf{z}. \tag{21}$$

Now, the task of channel estimation is to recover $\underline{\mathbf{h}}$ from the observation $\mathbf{y}$ and the known $\mathbf{X}$.

From (21), the linear MMSE solution of $\underline{\mathbf{h}}$ is

$$\hat{\underline{\mathbf{h}}} = \underline{\mathbf{C}}_{\mathrm{opt}}\mathbf{y} \tag{22}$$

where $\underline{\mathbf{C}}_{\rm opt} = \mathbf{R}_{\underline{\mathbf{h}}}\mathbf{A}^H(\mathbf{X}^H\mathbf{X}\mathbf{R}_{\mathbf{h}} + \sigma_z^2\mathbf{I}_{(K+1)N_RN_TN_P})^{-1}\mathbf{X}^H$. When the pilots satisfy the condition as in (14), $\underline{\mathbf{C}}_{\rm opt}$ can be reexpressed as

$$\underline{\mathbf{C}}_{\rm opt} = \mathbf{R}_{\underline{\mathbf{h}}}\mathbf{A}^H(L_P\mathbf{R}_{\mathbf{h}} + \sigma_z^2\mathbf{I}_{(K+1)N_TN_RN_P})^{-1}\mathbf{X}^H. \quad (23)$$

With the same assumptions as in Section III-A, the channel correlation matrix $\mathbf{R}_{\underline{\mathbf{h}}} = E\{\underline{\mathbf{h}}\,\underline{\mathbf{h}}^H\}$ can be expressed as

$$\mathbf{R}_{\underline{\mathbf{h}}} = \mathbf{R}_{DPR} \otimes \mathbf{R}_{RX} \otimes \mathbf{R}_{TX} \otimes \mathbf{R}_{ISI} \quad (24)$$

where $\underline{\mathbf{R}}_{DPR}$ is the extended temporal correlation matrix of $\mathbf{R}_{DPR}$.

Let $\underline{\mathbf{R}}_{DPR} = \underline{\mathbf{U}}\,\underline{\mathbf{\Lambda}}\,\underline{\mathbf{U}}^H$, where $\underline{\mathbf{U}}$ is the eigenmatrix of $\underline{\mathbf{R}}_{DPR}$ and $\underline{\mathbf{\Lambda}}$ is a diagonal matrix containing the eigenvalues $\underline{\lambda}_k, k = 0, 1, \ldots, KM+1$ of $\underline{\mathbf{R}}_{DPR}$. Then, we have the following spectrum decomposition of $\mathbf{R}_{\underline{\mathbf{h}}}$:

$$\mathbf{R}_{\underline{\mathbf{h}}} = \tilde{\underline{\mathbf{U}}}\,\tilde{\underline{\mathbf{\Lambda}}}\,\tilde{\underline{\mathbf{U}}}^H \quad (25)$$

where $\tilde{\underline{\mathbf{U}}} = \underline{\mathbf{U}} \otimes \mathbf{U}_R \otimes \mathbf{U}_T \otimes \mathbf{I}_{N_P}$, and $\tilde{\mathbf{\Lambda}} = \underline{\mathbf{\Lambda}} \otimes \mathbf{\Lambda}_R \otimes \mathbf{\Lambda}_T \otimes \mathbf{\Lambda}_{ISI}$. Therefore, we have

$$\underline{\mathbf{C}}_{\rm opt} = \tilde{\underline{\mathbf{U}}}\,\tilde{\underline{\mathbf{\Lambda}}}\,\tilde{\underline{\mathbf{U}}}^H\mathbf{A}^H\tilde{\mathbf{U}}(L_P\tilde{\mathbf{\Lambda}} + \sigma_z^2\mathbf{I}_{(K+1)N_TN_RN_P})^{-1}\tilde{\mathbf{U}}^H\mathbf{X}^H \quad (26)$$

and the linear MMSE estimate of $\underline{\mathbf{h}}$ can be rewritten as

$$\hat{\underline{\mathbf{h}}} = \tilde{\underline{\mathbf{U}}}\,\tilde{\underline{\mathbf{\Lambda}}}\,\tilde{\underline{\mathbf{U}}}^H\mathbf{A}^H\tilde{\mathbf{U}}\left(\tilde{\mathbf{\Lambda}} + \frac{\sigma_z^2}{L_P}\mathbf{I}_{(K+1)N_TN_RN_P}\right)^{-1}\tilde{\mathbf{U}}^H\hat{\mathbf{h}}_{\rm ini}. \quad (27)$$

Let $\underline{\mathbf{h}}^p$ be the extended version of $\mathbf{h}^p$ with the relation $\mathbf{h}^p = (\mathbf{B} \otimes \mathbf{I}_{N_RN_T})\underline{\mathbf{h}}^p$, and $\hat{\underline{\mathbf{h}}}^p$ be the corresponding optimal estimate. From (27), we have

$$\hat{\underline{\mathbf{h}}}^p = \mathbf{U}_{st,E}\mathbf{\Lambda}_{st,p}\mathbf{U}_{st}^H\hat{\mathbf{h}}_{\rm ini}^p \quad (28)$$

where $\mathbf{U}_{st,E} = \mathbf{U}_E \otimes \mathbf{U}_R \otimes \mathbf{U}_T$ and

$$\mathbf{U}_E = \underline{\mathbf{U}}\,\underline{\mathbf{\Lambda}}\,\underline{\mathbf{U}}^H\mathbf{B}^H\mathbf{U}\mathbf{\Lambda}^\dagger. \quad (29)$$

This implies that the MMSE estimation of $\underline{\mathbf{h}}$ can be performed by initial channel estimation on each pilot segment and each receive antenna followed by space–time filtering and "interpolation".

Let $\underline{\mathbf{h}}_{m,n}^p$ be the extended version of $\mathbf{h}_{m,n}^p$ as defined in the Section III-A, i.e., $\mathbf{h}_{m,n}^p = \mathbf{B}\underline{\mathbf{h}}_{m,n}^p$, and $\hat{\underline{\mathbf{h}}}_{m,n}^p$ be the corresponding MMSE estimates. For the spatially uncorrelated MIMO channels, we have that $\mathbf{U}_{st,E} = \mathbf{U}_E \otimes \mathbf{I}_{N_RN_T}$. From (28), the following relation holds:

$$\hat{\underline{\mathbf{h}}}_{m,n}^p = \mathbf{U}_E\mathbf{\Lambda}_p\mathbf{U}^H\hat{\mathbf{h}}_{{\rm ini},m,n}^p. \quad (30)$$

Therefore, the space–time filtering and interpolation given by (28) can be reduced to pathwise filtering and interpolation.

From the expression of $\mathbf{U}_E$ given by (29), it appears that $\mathbf{U}_E$ is related to the eigenvalues and the eigenvectors of the correlation matrices. To simplify the implementation of the space–time postprocessing given by (28), it is valuable to have an insight into the matrix $\mathbf{U}_E$. At first, from (29), it is easy to verify that the following relation holds:

$$\mathbf{B}\mathbf{U}_E = \mathbf{U}. \quad (31)$$

This gives that the channel estimate $\hat{\mathbf{h}}^p$ in (18) is related to $\hat{\underline{\mathbf{h}}}^p$ in (28) as $\hat{\mathbf{h}}^p = (\mathbf{B} \otimes \mathbf{I}_{N_RN_T})\hat{\underline{\mathbf{h}}}^p$. Equation (31) also implies that the (*Mi*)th row vector of $\mathbf{U}_E$, $0 \le i \le K$ is independent of the eigenvalues of the channel correlation matrix. Furthermore, we have the following theorem about the dependency of $\mathbf{U}_E$.

Theorem 2: Suppose that $\underline{\mathbf{h}}$ is linearly recoverable from $\mathbf{h}$, i.e., there exists a $(KM+1)N_RN_TN_P \times (K+1)N_RN_TN_P$ matrix $\mathbf{C}$ so that $\underline{\mathbf{h}} = \mathbf{C}\mathbf{h}$. Let $\mathbf{R}_{DPR}$ be of full rank. Then, the matrix $\mathbf{U}_E$ given by (29) is independent of the eigenvalues of channel correlation matrices.

For proof, see Appendix II. In Section IV, the previous statement will be employed for developing fast implementation of the channel estimation.

C. Design of Optimal Pilot

From the orthogonal condition (14), we can see that the training sequences transmitted from multiple antennas must have impulse-like autocorrelation and zero cross correlation. The problem of designing multiple pilot sequences can be reduced to a much easier and well-understood problem of designing a single sequence with impulse-like autocorrelation [13], [22], [23].

A sequence is said to be perfect if all its out-of-phase periodic autocorrelation terms are equal to zero [13]. Here, we would like to refer to such a sequence as cyclic orthogonal sequence. Over the past 40 years, numerous constructions of cyclic orthogonal sequence with constant magnitude have been proposed [29]–[32]. The constant magnitude property precludes PAPR problem in practical applications. The cyclic orthogonal sequences with constant magnitude are known as perfect root-of-unity sequences whose elements are complex roots of unity. The well-known Chu sequence of arbitrary length L_P has the alphabet size L_P, and the Frank sequence and the Milewski sequence of specific lengths have the reduced alphabet size [29]–[31]. General constructions for sequences with arbitrary length or minimum alphabet size can be found, e.g., in [32]. In this paper, we are interested in the Chu sequence and a dedicated class of sequences drawn from the DFT matrix.

The Chu sequence of length L_P is defined as

$$a(l) = \begin{cases} e^{j\pi r l^2/L_P}, & \text{for even} \quad L_P \\ e^{j\pi r l(l+1)/L_P}, & \text{for odd} \quad L_P \end{cases} \quad (32)$$

where $l = 0, 1, \ldots, L_P - 1$, and r is relatively prime to L_P.

The dedicated class of sequences of length $L_P = 2^{2n}$ or $L_P = 2^{2n-1}$ is directly drawn from the DFT matrix. We refer to such a sequence as the DFT-based sequence. Let $\mathbf{W}_N$ be the N-point nonnormalized DFT matrix with the (k,l)th entry $[\mathbf{W}_N]_{k,l} = W_N^{kl}$, where $W_N = e^{-j2\pi/N}$ and $N = 2^n$. Assume that

$$\mathbf{W}_N = \begin{bmatrix} \mathbf{W}_N^{(0)} & \mathbf{W}_N^{(1)} \end{bmatrix}^T$$

where $\mathbf{W}_N^{(0)}$ and $\mathbf{W}_N^{(1)}$ are the submatrices of $\mathbf{W}_N$ with size $N \times N/2$. Then, the DFT-based sequence of length L_P is given

by

$$\mathbf{a} = [a(0), a(1), \ldots, a(L_P - 1)]^T = \mathrm{vec}\{\tilde{\mathbf{W}}_N\} \tag{33}$$

where $\tilde{\mathbf{W}}_N$ is equal to $\mathbf{W}_N$ and $(\mathbf{W}_N^{(0)} + j * \mathbf{W}_N^{(1)})/(1+j)$ for $L_P = 2^{2n}$ and $L_P = 2^{2n-1}$, respectively. It can be verified that the resulting DFT-based sequence is a special case of the Frank sequence and the Milewski sequence when $L_P = 2^{2n}$ and $L_P = 2^{2n-1}$, respectively. The alphabet size of the sequence is equal to N.

In [22] and [23], repeated and phase-rotated Chu sequences are suggested for the pilot sequences for N_T transmit antennas. A base sequence is first designed from a length L_P/N_T Chu sequence by repeating it N_T times. Then, the base sequence and $N_T - 1$ phase-rotated versions (frequency-shifted version) of the base sequence are used as pilots for N_T transmit antennas. In this paper, we consider an alternative design that gives a fast implementation of the initial channel estimation, as discussed in the Section IV. The pilot sequence for each transmit antenna is derived by cyclically right-shifting a single base sequence $a(l)$, where $0 \le l \le L_P - 1$. From the base sequence $a(l)$, the pilots for the N_T transmit antennas are generated as

$$s_n(l) = a(((l - nN_P))_{L_P}) \tag{34}$$

where $n = 0, 1, \ldots, N_T - 1$. To satisfy the orthogonal condition given by (14), L_P should be no less than $N_T N_P$. In this case, the matrix $\mathbf{S}$ can be expressed as

$$\mathbf{S} = \bar{\mathbf{S}}\,[\,\mathbf{I}_{N_T N_P} \quad \mathbf{0}\,]^T \tag{35}$$

where

$$\bar{\mathbf{S}} = \mathrm{circ}\{\mathbf{a}\}. \tag{36}$$

The $\bar{\mathbf{S}}$ is orthogonal if and only if the base sequence $a(l)$ is cyclically orthogonal. For any case with $L_P \ge N_T N_P$, the cyclic orthogonality of $a(l)$ is sufficient for the $\mathbf{S}$ satisfying the orthogonal condition given by (14).

IV. Fast Implementation

From the earlier discussion, the channel estimation can be performed by the initial channel estimation as in (17) followed by space–time filtering as in (18) for pilot segments only, or by space–time filtering and interpolation as in (28) for both pilot and data segments. Now, we turn to the fast implementation for the initial channel estimation and the space–time postprocessing.

A. Initial Channel Estimation

The direct implementation of the initial channel estimation given by (17) needs $N_T N_P L_P$ complex multiplications and $N_T N_P (L_P - 1)$ additions. In practice, the factor $1/L_P$ in (17) can be moved to the space–time postprocessing without any increase in operations. For the pilots with the alphabet size no larger than 4, the complex multiplications can be replaced by additions. With a larger alphabet size for achieving an arbitrary length sequence, the direct implementation could be far from efficient.

Substituting (35) into (17), we have

$$\hat{\mathbf{h}}_{\mathrm{ini},m}(k) = \frac{1}{L_P}[\mathbf{I}_{N_T N_P} \ \mathbf{0}]\bar{\mathbf{S}}^H \mathbf{y}_m(k). \tag{37}$$

With the pilot design given by (34), the $\bar{\mathbf{S}}$ is a circular matrix and can be expressed as [33]

$$\bar{\mathbf{S}} = \mathbf{W}_{L_P}^H \mathbf{\Lambda}_a \mathbf{W}_{L_P} \tag{38}$$

where $\mathbf{W}_{L_P}$ is the L_P-point DFT matrix and $\mathbf{\Lambda}_a = (1/L_P)\mathrm{diag}\{\mathbf{W}_{L_P}\mathbf{a}\}$. Note that for cyclic orthogonal pilot sequences, the diagonal elements of $\mathbf{\Lambda}_a$ are of constant magnitude. Substituting (38) into (37) yields

$$\hat{\mathbf{h}}_{\mathrm{ini},m}(k) = \frac{1}{L_P}[\mathbf{I}_{N_T N_P} \ \mathbf{0}]\mathbf{W}_{L_P}^H \mathbf{\Lambda}_a^H \mathbf{W}_{L_P} \mathbf{y}_m(k). \tag{39}$$

This gives a fast implementation that needs two L_P-point DFT and L_P additional complex multiplications. Such implementation of the blockwise initial LS channel estimation is equivalent to that in [22] and [23], where frequency-domain channel estimate is reconstructed from the time-domain channel estimate. Notice that the implementation applies to any case with cyclic orthogonal pilot sequences, and the structure of a specific pilot sequence is not exploited.

When the Chu sequence and the DFT-based sequence given in Section III-A serve as the base sequence $a(l)$, we have the following theorems.

Theorem 3: Let $a(l)$ be the Chu sequence of length L_P. Then, the matrix $\bar{\mathbf{S}}$ defined as in (36) can be factorized as

$$\bar{\mathbf{S}} = \mathbf{\Phi}\mathbf{P}_\gamma \mathbf{W}_{L_P} \mathbf{\Psi} \tag{40}$$

where $\mathbf{\Phi}$ is a diagonal matrix with the kth diagonal entry as $e^{j2\pi rk^2/L_P}$ and $e^{j2\pi rk(k+1)/L_P}$ for even and odd L_P, respectively; $\mathbf{\Psi}$ is a diagonal matrix with the lth diagonal entry as $e^{j2\pi rl^2/L_P}$ and $e^{j2\pi rl(l-1)/L_P}$ for even and odd L_P, respectively; and $\mathbf{P}_\gamma = [\mathbf{e}_{\gamma(0)}, \mathbf{e}_{\gamma(1)}, \ldots, \mathbf{e}_{\gamma(L_P-1)}]^T$ with $\gamma(k) = ((rk))_{L_P}$.

Theorem 4: Let $a(l)$ be the DFT-based sequence given by (33). Then, the matrix $\bar{\mathbf{S}}$ defined as in (36) can be factorized as

$$\bar{\mathbf{S}} = \mathbf{P}_\alpha^H (\mathbf{I}_Q \otimes \mathbf{W}_N^H)\mathbf{\Sigma}\mathbf{P}_\beta (\mathbf{I}_Q \otimes \mathbf{W}_N)\mathbf{P}_\alpha \tag{41}$$

where Q is equal to N and $N/2$ for $L_P = 2^{2n}$ and $L_P = 2^{2n-1}$, respectively; $\mathbf{P}_\alpha = [\mathbf{e}_{\alpha(0)}, \mathbf{e}_{\alpha(1)}, \ldots, \mathbf{e}_{\alpha(L_P-1)}]^T$ with $\alpha(k_1 N + k_0) = k_0 Q + k_1$ for $0 \le k_0 \le N-1$ and $0 \le k_1 \le Q-1$; $\mathbf{P}_\beta = [\mathbf{e}_{\beta(0)}, \mathbf{e}_{\beta(1)}, \ldots, \mathbf{e}_{\beta(L_P-1)}]^T$ with $\beta(k) = ((k + N((k))_N))_{L_P}$; and $\mathbf{\Sigma}$ is a diagonal matrix.

The proofs of Theorems 3 and 4 are provided in Appendixes III and IV, respectively.

Substituting (40) and (41) into (37), we have

$$\hat{\mathbf{h}}_{\mathrm{ini},m}(k) = \frac{1}{L_P}[\mathbf{I}_{N_T N_P} \ \mathbf{0}]\mathbf{\Psi}^H \mathbf{W}_{L_P}^H \mathbf{P}_\gamma^H \mathbf{\Phi}^H \mathbf{y}_m(k) \tag{42}$$

$$\hat{\mathbf{h}}_{\mathrm{ini},m}(k) = \frac{1}{L_P}[\mathbf{I}_{N_T N_P} \ \mathbf{0}]\mathbf{P}_\alpha^H (\mathbf{I}_Q \otimes \mathbf{W}_N^H)\mathbf{P}_\beta^H \mathbf{\Sigma}^H * (\mathbf{I}_Q \otimes \mathbf{W}_N)\mathbf{P}_\alpha \mathbf{y}_m(k). \tag{43}$$

From (42) and (43), the initial channel estimation can be implemented more efficiently than that given by (39). Fig. 3(a) and

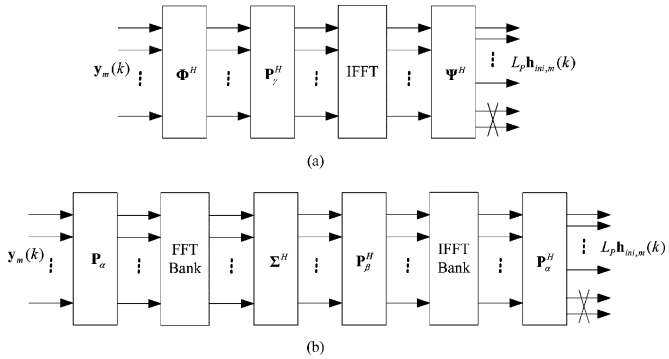

Fig. 3. Efficient implementation structure for initial channel estimation with the following sequences. (a) Chu sequence. (b) DFT-based sequence.

TABLE I
COMPUTATIONAL COMPLEXITY OF THE INITIAL CHANNEL ESTIMATION

Type of pilot sequences and its corresponding computation formula	Number of complex multiplications	Number of complex additions
General orthogonal sequence (17)	$L_P N_T N_P$	$(L_P-1)N_T N_P$
General cyclic orthogonal sequence (39)	$L_P(\log_2 L_P+1)$	$2L_P \log_2 L_P$
Chu sequence (42)	$L_P(\frac{1}{2}\log_2 L_P+2)$	$L_P \log_2 L_P$
DFT-based cyclic orthogonal sequence (43)	$L_P(\lceil\frac{1}{2}\log_2 L_P\rceil+1)$	$2L_P\lceil\frac{1}{2}\log_2 L_P\rceil$

(b) shows the implementation structures for the initial channel estimation with the Chu sequence and that with the DFT-based sequence, respectively. For the former case, only one L_P-point DFT and $2L_P$ additional complex multiplications are needed. For the latter case, $2Q$ N-point DFT and L_P additional complex multiplications are needed. When the pilot length is a power of 2, FFT algorithms are available. For the radix-2 FFT algorithm, an L-size DFT requires $(1/2)L\ \log_2 L$ complex multiplications and $L\ \log_2 L$ complex additions. The computational complexity of the initial channel estimation in terms of the number of complex multiplications and complex additions per receive antenna is listed in Table I for the four implementations given by (17), (39), (42), and (43). Although only two kinds of sequences are involved in the earlier discussion, similar fast implementations can be obtained for the initial channel estimation with other existing sequences through similar procedures. For example, one can obtain an implementation similar to (42) with the Zadoff–Chu sequence or to (43) with the generalized chirp-like sequence of minimum alphabets [32].

B. Space–Time Postprocessing for Spatially Uncorrelated Channels

The space–time postprocessing as in (18) for pilot segments only and in (28) for both pilot and data segments are related to the eigendecomposition of the temporal correlation matrices $\mathbf{R}_{\mathbf{h}}$ and $\mathbf{R}_{\underline{\mathbf{h}}}$. This means that the online estimation of the correlation matrices and the online eigenvalue decomposition (EVD) should be performed, which is computationally complicated. In this section, we discuss the fast implementation of space–time postprocessing for spatially uncorrelated channels, and then, extend the algorithm to triply selective channels in the Section IV-C.

For spatially uncorrelated channels, the space–time postprocessing as in (18) and (28) can be reduced to the pathwise processing as in (19) and (30). The pathwise processing can be explained as a three-step procedure. First, the initial channel estimate $\hat{\mathbf{h}}^p_{\mathrm{ini}\,m,n}$ is optimally decorrelated with the eigenmatrix $\mathbf{U}$. Due to the orthogonality of the pilot matrix $\mathbf{S}$, $\hat{\mathbf{h}}^p_{\mathrm{ini}\,m,n}$ can be expressed as $\hat{\mathbf{h}}^p_{\mathrm{ini}\,m,n}=\mathbf{h}^p_{m,n}+\mathbf{z}^p_{\mathrm{ini}\,m,n}$, where $\mathbf{z}^p_{\mathrm{ini}\,m,n}$ is an AWGN vector with zero mean and covariance matrix $(\sigma_z^2/L_P)\mathbf{I}_{K+1}$. The correlation matrix of $\mathbf{U}^H\hat{\mathbf{h}}^p_{\mathrm{ini}\,m,n}$ can be expressed as $\rho_p^2\mathbf{\Lambda}+\sigma_z^2/L_P\mathbf{I}_{K+1}$. Second, the optimal filtering in the MMSE sense is performed on each element of the decorrelated vector with the eigenvalues of the correlation matrix $\mathbf{R}_{DPR}$. Finally, the MMSE channel estimates $\hat{\mathbf{h}}^p_{m,n}$ and $\hat{\underline{\mathbf{h}}}^p_{m,n}$ are reconstructed with the eigenmatrix $\mathbf{U}$ and $\mathbf{U}_E$, respectively. The matrices $\hat{\mathbf{h}}^p_{m,n}$ and $\hat{\underline{\mathbf{h}}}^p_{m,n}$ are linear combinations of the column vectors of $\mathbf{U}$ and $\mathbf{U}_E$, respectively. From (31), the ith column vector of $\mathbf{U}_E$ can be considered as an extended version of that of $\mathbf{U}$.

To simplify the implementation, the optimal decorrelation transform of a correlated signal is often replaced by a fast one in many applications such as image coding. DCT is one of the popular fast transforms. The basis set of DCT provides a good approximation to the eigenvectors of the class of Toeplitz matrices that constitute the correlation matrices of finite-order stationary Markov processes [34]. In other words, the correlation matrix of the correlated signal vector can be well diagonalized by the DCT matrix. We fix the matrix $\mathbf{U}$ in (19) and (30) to be the transpose of the $(K+1)$-point type II DCT matrix. The (k,l)th entry of the $(K+1)$-point type II DCT matrix $\mathbf{C}^{II}_{K+1}$ is [35]

$$[\mathbf{C}^{II}_{K+1}]_{k,l}=\kappa_k\ \cos\frac{\pi k(l+0.5)}{K+1} \tag{44}$$

where

$$\kappa_k=\begin{cases}1/\sqrt{K+1}, & k=0\\ \sqrt{2}/\sqrt{K+1}, & k\neq 0.\end{cases} \tag{45}$$

From Theorem 2, we know that the $\mathbf{U}_E$ given by (29) is independent of the eigenvalues of channel correlation matrices if $\underline{\mathbf{h}}$ is linearly recoverable from $\mathbf{h}$. The recoverability of $\underline{\mathbf{h}}$ from $\mathbf{h}$ can be met approximately in practice. Therefore, the $\mathbf{U}_E$ can be approximated with a matrix that is not related to the eigenvalues of the channel correlation matrices. We propose to replace $\mathbf{U}_E$ in (30) by the transpose of an extended DCT matrix $\underline{\mathbf{C}}^{II}_{K+1}$, which is defined as

$$[\underline{\mathbf{C}}^{II}_{K+1}]_{k,l}=\kappa_k\ \cos\frac{\pi k(l/M+0.5)}{K+1} \tag{46}$$

where $0\le k\le K$ and $0\le l\le MK$. It can be verified that the following relations hold

$$\underline{\mathbf{C}}^{II}_{K+1}\mathbf{B}^H=\mathbf{C}^{II}_{K+1} \tag{47}$$

$$\underline{\mathbf{C}}^{II}_{K+1}=[\mathbf{I}_{K+1}\ \mathbf{0}]\mathbf{D}\mathbf{P} \tag{48}$$

where $\mathbf{P}=[\mathbf{0}_{(MK+1)\times\lfloor M/2\rfloor}\ \mathbf{I}_{MK+1}\ \mathbf{0}_{(MK+1)\times\lfloor M/2\rfloor}]^T$, and $\mathbf{D}$ is the $(MK+M+1)$-point type I DCT matrix $\mathbf{C}^{I}_{M(K+1)+1}$

for even M and the $(MK+M)$-point type II DCT matrix $\mathbf{C}^{II}_{M(K+1)}$ for odd M. The definition of the type I DCT matrix $\mathbf{C}^{I}_{M(K+1)+1}$ can be found, e.g., in [35]. The relation in (47) is consistent with that in (31).

With $\mathbf{U}$ and $\mathbf{U}_E$ replaced by $(\mathbf{C}^{II}_{K+1})^T$ and $(\underline{\mathbf{C}}^{II}_{K+1})^T$, respectively, (19) and (30) become

$$\hat{\mathbf{h}}^p_{m,n} = (\mathbf{C}^{II}_{K+1})^T \mathbf{\Lambda}_p \mathbf{d}^p_{m,n} \tag{49}$$

$$\underline{\hat{\mathbf{h}}}^p_{m,n} = (\underline{\mathbf{C}}^{II}_{K+1})^T \mathbf{\Lambda}_p \mathbf{d}^p_{m,n} \tag{50}$$

where

$$\mathbf{d}^p_{m,n} = \mathbf{C}^{II}_{K+1} \hat{\mathbf{h}}^p_{\mathrm{ini}\,m,n}. \tag{51}$$

Let $d^p_{m,n}(k)$ denote the kth element of $\mathbf{d}^p_{m,n}$. Under the condition that the basis set of the DCT approximates to the eigenvectors, we have

$$E\{|d^p_{m,n}(k)|^2\} \approx \lambda_k \rho_p^2 + \sigma_z^2/L_P. \tag{52}$$

Thus, the kth diagonal entry of $\mathbf{\Lambda}_p$ can be approximated by $E\{|d^p_{m,n}(k)|^2\} - \sigma_z^2/L_P$.

From (49)–(52), we can see that the calculation of channel correlation matrix and its EVD have been avoided. Thus, the pathwise processing can be implemented efficiently. First, the initial channel estimates $\hat{\mathbf{h}}^p_{\mathrm{ini}\,m,n}$ are transformed to nearly decorrelated vectors $\mathbf{d}^p_{m,n}$ via the $(K+1)$-point DCT as in (51). Then, from the resulting $\mathbf{d}^p_{m,n}$, the diagonal matrix $\mathbf{\Lambda}_p$ can be estimated online by using the relationship shown in (52). The expectation can be replaced by the sample average over $N_T \times N_R$ transmit–receive channels, and possibly, over timeslots. With the estimate of $\mathbf{\Lambda}_p$, filtering is performed on each element of $\mathbf{d}^p_{m,n}$. Finally, the filtered vectors are transformed to $\hat{\mathbf{h}}^p_{m,n}$ via the DCT as in (49), or $\underline{\hat{\mathbf{h}}}^p_{m,n}$ via the extended DCT as in (50).

In the implementation, the main computational load relies on that of the DCTs. For the pilot segments only, one $(K+1)$-point type II DCT and one $(K+1)$-point type III DCT, which is the inverse of type II DCT, are needed for each path of every transmit–receive channel. For both the pilot and data segments, one $(K+1)$-point type II DCT and one $(MK+M+1)$-point type I DCT or $(MK+M)$-point type II DCT (for even or odd M, respectively) are needed. When the number of pilot segments in a timeslot $K+1$ is a power of two, as well as M, fast DCT algorithms are available, e.g., in [35]. The numbers of real multiplications needed for the fast implementations of the (K+1)-point type II DCT and $(MK+M)$-point type II DCT are $(K+1)((3/4)\log_2(K+1)-1)+3$ and $(MK+M)((3/4)\log_2(MK+M)-1)+3$, respectively.

C. Space–Time Filtering for Triply Selective Channels

For triply selective channels, the space–time postprocessing cannot be simplified to pathwise processing due to the spatial correlation. When the matrices $\mathbf{R}_{TX}$ and $\mathbf{R}_{RX}$ have the following structure:

$$[\mathbf{R}_{TX}]_{n_1,n_2} = \rho_T^{|n_1-n_2|}, \qquad 0 \le \rho_T < 1 \tag{53}$$

$$[\mathbf{R}_{RX}]_{m_1,m_2} = \rho_R^{|m_1-m_2|}, \qquad 0 \le \rho_R < 1 \tag{54}$$

TABLE II
SIMULATION PARAMETERS

Parameter	Value
Transmit antenna number	4
Receive antenna number	4
Transmit correlation factor	0 or 0.2
Receive correlation factor	0 or 0.5
Carrier frequency	3.5GHz
Symbol rate	1.28Msps
Velocity speed	250km/h or 3km/h
Power spectral density	Jakes' model
Multipath delay profile	Eight paths uniformly distributed over the interval $0 \sim 5.46875\mu s$
Multipath power profile	Exponentially distributed
Sub-timeslot number K	7
Guard length L_G	8
Pilot length L_P	32
User data length L_D	312

the spatial–domain decorrelation can also be performed by using DCT. This covariance matrix model is frequently used in literature (see, e.g., [36] and the references therein). With the spatial correlation matrices $\mathbf{U}_R$ and $\mathbf{U}_T$ replaced by $(\mathbf{C}^{II}_{N_R})^T$ and $(\mathbf{C}^{II}_{N_T})^T$, respectively, in addition to the temporal correlation matrices replaced by DCT matrices, as shown in Section IV-B, the space–time postprocessing as in (18) and (28) become

$$\hat{\mathbf{h}}^p = (\mathbf{C}^{II}_{K+1} \otimes \mathbf{C}^{II}_{N_R} \otimes \mathbf{C}^{II}_{N_T})^T \mathbf{\Lambda}_{st,p} \mathbf{d}^p \tag{55}$$

$$\underline{\hat{\mathbf{h}}}^p = (\underline{\mathbf{C}}^{II}_{K+1} \otimes \mathbf{C}^{II}_{N_R} \otimes \mathbf{C}^{II}_{N_T})^T \mathbf{\Lambda}_{st,p} \mathbf{d}^p \tag{56}$$

where

$$\mathbf{d}^p = (\mathbf{C}^{II}_{K+1} \otimes \mathbf{C}^{II}_{N_R} \otimes \mathbf{C}^{II}_{N_T}) \hat{\mathbf{h}}^p_{\mathrm{ini}}. \tag{57}$$

Let $d^p_{m,n}(k)$ denote the $(kN_RN_T + mN_T + n)$th element of $\mathbf{d}^p$. Under the condition that the basis set of the DCT approximates to the eigenvectors, we have

$$E\{|d^p_{m,n}(k)|^2\} \approx \lambda_k \lambda_{R,m} \lambda_{T,n} \rho_p^2 + \frac{\sigma_z^2}{L_P}. \tag{58}$$

Thus, the $(kN_RN_T + mN_T + n)$th diagonal entry of $\mathbf{\Lambda}_{st,p}$ can be approximated by $E\{|d^p_{m,n}(k)|^2\} - \sigma_z^2/L_P$.

From (55) to (58), the space–time postprocessing can be implemented efficiently. First, the initial channel estimates $\hat{\mathbf{h}}^p_{\mathrm{ini}}$ are transformed to nearly decorrelated vectors $\mathbf{d}^p$ via DCTs, as in (57). Then, from the resulting $\mathbf{d}^p$, the diagonal matrix $\mathbf{\Lambda}_{st,p}$ can be estimated online by using the relationship shown in (58). With the estimate of $\mathbf{\Lambda}_{st,p}$, filtering is performed on each element of $\mathbf{d}^p$. Finally, the filtered vectors are transformed to $\hat{\mathbf{h}}^p$ as in (55) or $\underline{\hat{\mathbf{h}}}^p$ as in (56).

V. SIMULATIONS

To evaluate the performance of the channel estimation for the MIMO-SCBT with dual cyclic timeslot structure, simulations have been performed in Rayleigh fading MIMO channels. Four transmit and four receive antennas are assumed. The channels

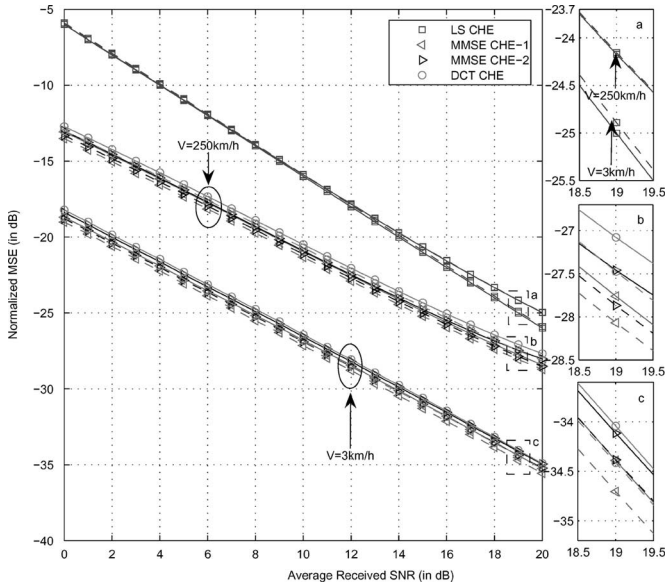

Fig. 4. NMSE performance of the channel estimation methods for the pilot segments only, where solid and dashes lines represent spatially uncorrelated and correlated channels, respectively.

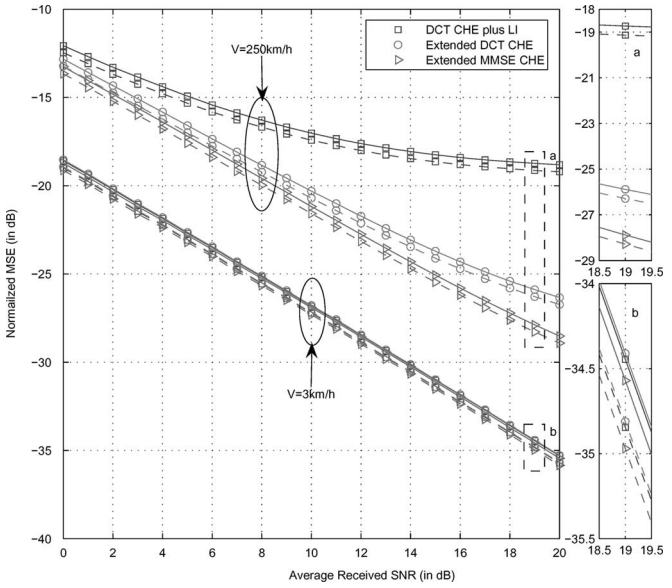

Fig. 5. NMSE performance of the channel estimation methods for both the pilot and data segments, where solid and dashed lines represent spatially uncorrelated and correlated channels, respectively.

corresponding to different transmit and receive antennas have the same statistics. The power spectral density satisfies Jakes' model, i.e., the (i, j)th entry of $\mathbf{R}_{DPR}$ is $J_0(2\pi f_D(i-j)L_S T_S)$. The power delay profile is assumed to be exponentially distributed. The channel time delay is uniformly distributed over the interval 0–5.46875 μs for $N_P = 8$ paths. The additive channel noise is spatially and temporally white Gaussian with zero mean and the variance determined by the SNR. The transmitted signals are of 1.28 MHz 3 dB bandwidth, i.e., the symbol rate is 1.28 Ms/s. The carrier frequency is set to be 3.5 GHz. The vehicle speed is set to be 250 or 3 km/h. Both the spatially correlated and uncorrelated channels are considered. For the spatially correlated channel, the spatial correlation factors are set to be $\rho_T = 0.2$ and $\rho_R = 0.5$. The parameters in the timeslot structure are as follows: the subtimeslot number $K = 7$, the guard length $L_G = 8$, the pilot length $L_P = 32$, the user data length $L_D = 312$, and the spacing between two successive pilot segments $L_S = 352$. We summarize these parameters in Table II.

The channel estimation for pilot segments only and for both the pilot and data segments are simulated separately. Normalized MSE (NMSE) performances of the channel estimation are evaluated. The NMSE (in decibels) is defined as $10\log(E\{\| \mathbf{h} - \hat{\mathbf{h}} \|^2\}/E\{\| \mathbf{h} \|^2\})$ for pilot segments only and $10\log(E\{\| \underline{\mathbf{h}} - \underline{\hat{\mathbf{h}}} \|^2\}/E\{\| \underline{\mathbf{h}} \|^2\})$ for both the pilot and data segments.

For pilot segments only, the following four channel estimation methods are compared via simulations.

1) *LS CHE:* LS channel estimation is the same as the initial channel estimation in (17), which can be fast implemented as in (39), (42), and (43) for an arbitrary cyclic orthogonal pilot sequence, the Chu sequence, and the DFT-based sequence, respectively.
2) *MMSE CHE-1:* MMSE channel estimation with the initial channel estimation as in (17) and the space–time postprocessing as in (18) and (19) for spatially correlated and uncorrelated channels, respectively, where the $\mathbf{U}_{st}$, $\mathbf{\Lambda}_{st,p}$, $\mathbf{U}$, and $\mathbf{\Lambda}_p$ are obtained from known channel correlation and known power delay profile.
3) *MMSE CHE-2:* MMSE channel estimation with the initial channel estimation as in (17) and the space–time postprocessing as in (18) and (19) for spatially correlated and uncorrelated channels, respectively, where the $\mathbf{U}_{st}$, $\mathbf{\Lambda}_{st,p}$, $\mathbf{U}$, and $\mathbf{\Lambda}_p$ are obtained from the online estimates of channel correlation and power delay profile. The channel correlation and the power delay profile are estimated from the samples of the initial channel estimates.
4) *DCT CHE:* DCT-based channel estimation with the initial channel estimation as in (17) and the space–time postprocessing as in (49), (51), and (52) for the spatially uncorrelated channel and as in (55), (57), and (58) for the spatially correlated channel.

For both the pilot and data segments, we compare the following three channel estimation methods with $M = 4$.

1) *Extended MMSE CHE:* Extended MMSE channel estimation with the initial channel estimation as in (17) and the space–time postprocessing as in (28) and (30) for the spatially correlated and the uncorrelated channels, respectively, where $\mathbf{U}_{st,E}$, $\mathbf{U}_{st}$, $\mathbf{\Lambda}_{st,p}$, $\mathbf{U}_E$, $\mathbf{U}$, and $\mathbf{\Lambda}_p$ are obtained from known channel correlation and known power delay profile.
2) *DCT CHE plus LI:* DCT-based channel estimation, as defined before, followed by linear interpolation to obtain channel estimates in user data segments.
3) *Extended DCT CHE:* Extended DCT-based channel estimation with the initial channel estimation as in (17) and

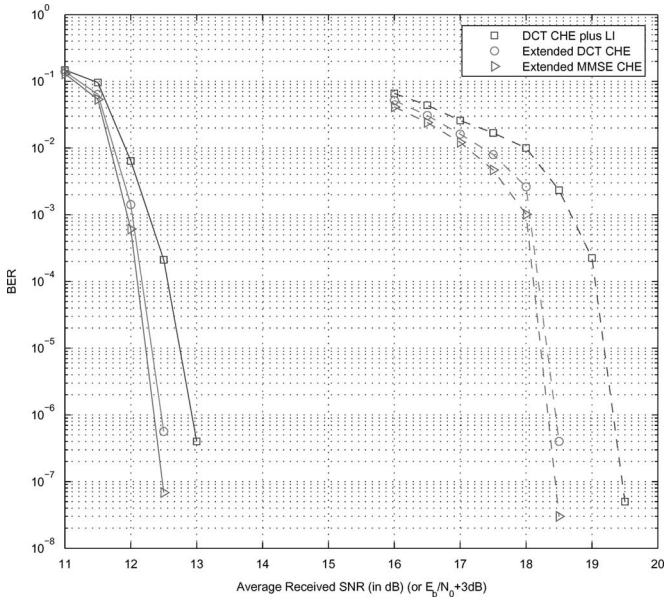

Fig. 6. BER performance with difference channel estimation methods for 1/2-rate turbo-encoded and 16-QAM mapped signals, where solid and dashed lines represent spatially uncorrelated and correlated channels, respectively.

the space–time postprocessing as in (50)–(52) for the spatially uncorrelated channel and as in (56)–(58) for the spatially correlated channel, respectively.

The bit error rate (BER) performances with the previous three channel estimation methods are also compared through the link simulation. In the simulations, the spatial multiplexing transmission scheme and a practical iterative receiver as introduced in [11] and [25] are employed.

Fig. 4 shows the NMSE performance of the channel estimation methods for the pilot segments only. From the results, we can see that the DCT-based channel estimation can well approximate the MMSE channel estimation with estimated or known channel correlation and known power delay profile, while it outperforms the LS channel estimation by 2.7–6.7 dB for the high mobility and 9.2–12.5 dB for the low mobility in the range of the measured SNRs. Furthermore, we can see that about 0.4 dB gain benefits from the spatial correlations.

Fig. 5 compares the NMSE performance of the channel estimation methods for both the pilot and data segments. From the results, it can be observed that the extended DCT-based channel estimation still well approximates the MMSE channel estimation with known channel correlation and known power delay profile. The extended DCT-based channel estimation outperforms the linear interpolated DCT-based channel estimation by up to 7.5 dB in fast fading channels, and the performance gain gets more significant as the received SNR increases. In slow fading channels, the linear interpolated DCT-based channel estimation has a similar performance as the extended DCT-based channel estimation. Furthermore, we can see that about 0.3-dB gain benefits from the spatial correlations.

Fig. 6 shows the BER performance of the channel estimation methods for 1/2-rate turbo-encoded data using 16 quadratic-amplitude modulation (QAM) mapped signals in fast fading channels. From the results, it can be observed that the extended DCT-based channel estimation still well approximates the MMSE channel estimation with the known channel correlation and power delay profile, while it outperforms the linear interpolated DCT-based channel estimation by up to 1.0 and 0.5 dB at BER $= 10^{-6}$ for spatially correlated and uncorrelated channels, respectively.

VI. Conclusion

An efficient channel estimation for a timeslot-structured SCBT over triply selective fading MIMO channels has been investigated in this paper. A MIMO-SCBT with dual cyclic timeslot structure is presented first. Then, we derive the optimal channel estimation in the MMSE sense on the timeslot basis. It is shown that the optimal pilots for the timeslot-based MMSE channel estimation are related to the statistical CSI in eigenmode. Under the assumption that the transmit correlation is unknown at the transmitter, the optimal pilots satisfy the same condition as reported in [22] and [23] for the block-based LS channel estimation, and the channel estimation can be simplified to the initial block-based LS channel estimation followed by space–time postprocessing. When the channel experiences spatially uncorrelated fading, the space–time postprocessing can be reduced to a more tractable pathwise processing. A new design of the pilot sequences is suggested, where the pilot sequence for each transmit antenna is derived by cyclically right-shifting a single base sequence. The Chu sequence of arbitrary length and the DFT-based sequence with small alphabet size can be used as the base sequence. A more efficient implementation for the initial channel estimation is obtained by using the structure of the pilot sequence. A DCT-based implementation for the space–time postprocessing is developed to approximate the optimal solution with low implementation complexity. The simulation results have verified the performance of the proposed channel estimation.

Appendix I
Proof of Theorem 1

Substituting (7) into (8) and using the fact that $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$, yield

$$\begin{aligned}\sigma_{\tilde{\mathbf{h}}}^2 &= \mathrm{tr}\left\{\tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{U}}^H\left(\frac{1}{\sigma_z^2}\mathbf{X}^H\mathbf{X}\tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{U}}^H + \mathbf{I}_{(K+1)N_R N_T N_P}\right)^{-1}\right\} \\ &= \mathrm{tr}\left\{\tilde{\mathbf{\Lambda}}\left(\frac{1}{\sigma_z^2}\tilde{\mathbf{U}}^H\mathbf{X}^H\mathbf{X}\tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}} + \mathbf{I}_{(K+1)N_R N_T N_P}\right)^{-1}\right\} \\ &= \mathrm{tr}\left\{\tilde{\mathbf{\Lambda}}\left(\frac{1}{\sigma_z^2}(\mathbf{I}_{(K+1)N_R} \otimes (\mathbf{U}_T \otimes \mathbf{I}_{N_P})^H\mathbf{S}^H \right.\right. \\ &\qquad \left.\left. \times \mathbf{S}(\mathbf{U}_T \otimes \mathbf{I}_{N_P}))\tilde{\mathbf{\Lambda}} + \mathbf{I}_{(K+1)N_R N_T N_P}\right)^{-1}\right\}. \end{aligned} \quad (59)$$

From [15], we know that the mean square error $\sigma^2_{\hat{\mathbf{h}}}$ is lower bounded as in (60), shown at the bottom of the page, where the equality holds if and only if $(\mathbf{U}_T \otimes \mathbf{I}_{N_P})^H \mathbf{S}^H \mathbf{S}(\mathbf{U}_T \otimes \mathbf{I}_{N_P})$ is diagonal.

Let $\mathbf{\Phi}_{p_1,p_2}$ be an $N_T \times N_T$ matrix with the (n_1, n_2)th entry defined as $[\mathbf{\Phi}_{p_1,p_2}]_{n_1,n_2} = [\mathbf{S}^H \mathbf{S}]_{n_1 N_P + p_1, n_2 N_P + p_2}$, where $0 \le p_1, p_2 \le N_P - 1$ and $0 \le n_1, n_2 \le N_T - 1$. Since $\mathbf{s}^{p_1}_{n_1} = \mathbf{\Pi}^{p_1}_{L_P} \mathbf{s}_{n_1}$ and $\mathbf{s}^{p_2}_{n_2} = \mathbf{\Pi}^{p_2}_{L_P} \mathbf{s}_{n_2}$, it follows that $[\mathbf{\Phi}_{p_1,p_2}]_{n_1,n_2} = \mathbf{s}^H_{n_1} \mathbf{\Pi}^{((p_1-p_2))_{L_P}}_{L_P} \mathbf{s}_{n_2}$. A sufficient and necessary condition for $(\mathbf{U}_T \otimes \mathbf{I}_{N_P})^H \mathbf{S}^H \mathbf{S}(\mathbf{U}_T \otimes \mathbf{I}_{N_P})$ to be diagonal is

$$\mathbf{U}^H_T \mathbf{\Phi}_{p_1,p_2} \mathbf{U}_T = \begin{cases} L_P \mathbf{\Gamma}, & p_1 = p_2 \\ \mathbf{0}, & p_1 \ne p_2 \end{cases} \tag{61}$$

where $\mathbf{\Gamma} = \mathrm{diag}\{\gamma_0, \ldots, \gamma_{N_T-1}\}$, and $\sum_{n=0}^{N_T-1} \gamma_n = N_T$. From (61), $\mathbf{S}^H \mathbf{S}$ can be expressed as

$$\mathbf{S}^H \mathbf{S} = L_P (\mathbf{U}_T \mathbf{\Gamma} \mathbf{U}^H_T) \otimes \mathbf{I}_{N_P}. \tag{62}$$

Therefore, the lower bound $\tilde{\sigma}^2_{\hat{\mathbf{h}}}$ is a function of $\mathbf{\Gamma}$, as shown in (11). This bound can be further lower bounded by

$$\breve{\sigma}^2_{\hat{\mathbf{h}}} = \tilde{\sigma}^2_{\hat{\mathbf{h}}}(\mathbf{\Gamma}_{\mathrm{opt}}). \tag{63}$$

This completes the proof.

APPENDIX II
PROOF OF THEOREM 2

We say that $\underline{\mathbf{h}}$ is linearly recoverable from $\mathbf{h}$, if there exists a $(KM+1)N_R N_T N_P \times (K+1)N_R N_T N_P$ matrix $\mathbf{C}$ so that $\underline{\mathbf{h}} = \mathbf{C}\mathbf{h}$. From (20), the optimal solution of $\mathbf{C}$ is $\mathbf{C} = \mathbf{R}_{\underline{\mathbf{h}}} \mathbf{A}^H \mathbf{R}^{-1}_{\mathbf{h}}$. The recoverability means that

$$\underline{\mathbf{h}} = \mathbf{R}_{\underline{\mathbf{h}}} \mathbf{A}^H \mathbf{R}^{-1}_{\mathbf{h}} \mathbf{h}. \tag{64}$$

From (64) and (20), we have

$$\mathbf{R}_{\underline{\mathbf{h}}} = \mathbf{R}_{\underline{\mathbf{h}}} \mathbf{A}^H \mathbf{R}^{-1}_{\mathbf{h}} \mathbf{A} \mathbf{R}_{\underline{\mathbf{h}}} \tag{65}$$

$$\mathbf{R}_{\mathbf{h}} = \mathbf{A} \mathbf{R}_{\underline{\mathbf{h}}} \mathbf{A}^H. \tag{66}$$

Substituting (6), (24), and the relation $\mathbf{A} = \mathbf{B} \otimes \mathbf{I}_{N_T N_R N_P}$ into (65) and (66) yields

$$\underline{\mathbf{R}}_{DPR} = \underline{\mathbf{R}}_{DPR} \mathbf{B}^H \mathbf{R}^{\dagger}_{DPR} \mathbf{B} \underline{\mathbf{R}}_{DPR} \tag{67}$$

$$\mathbf{R}_{DPR} = \mathbf{B} \underline{\mathbf{R}}_{DPR} \mathbf{B}^H. \tag{68}$$

Therefore, the rank of $\underline{\mathbf{R}}_{DPR}$ is equal to that of $\mathbf{R}_{DPR}$. With the assumption that $\mathbf{R}_{DPR}$ is of full rank, we have

$$rank(\underline{\mathbf{R}}_{DPR}) = rank(\mathbf{R}_{DPR}) = K + 1. \tag{69}$$

Let

$$\underline{\mathbf{\Lambda}} = \begin{bmatrix} \underline{\mathbf{\Lambda}}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

where the diagonal matrix $\underline{\mathbf{\Lambda}}_0$ is of the order of $K+1$ and contains $K+1$ nonzero eigenvalues of $\underline{\mathbf{R}}_{DPR}$. Let

$$\underline{\mathbf{U}} = [\, \underline{\mathbf{U}}_0 \quad \underline{\mathbf{U}}_1 \,]$$

where $\underline{\mathbf{U}}_0$ is composed of the first $K+1$ columns. Then, from (67) and the spectrum decompositions of $\underline{\mathbf{R}}_{DPR}$ and $\mathbf{R}_{DPR}$, we have

$$\underline{\mathbf{U}}\, \underline{\mathbf{\Lambda}}\, \underline{\mathbf{U}}^H \mathbf{B}^H \mathbf{U} \mathbf{\Lambda}^{-1} = \mathbf{U}_0 (\mathbf{B}\mathbf{U}_0)^{-1} \mathbf{U}. \tag{70}$$

$$\sigma^2_{\mathbf{h}} \ge \tilde{\sigma}^2_{\hat{\mathbf{h}}} = \sum_m \frac{1}{\left[\frac{1}{\sigma^2_z} (\mathbf{I}_{(K+1)N_R} \otimes (\mathbf{U}_T \otimes \mathbf{I}_{N_P})^H \mathbf{S}^H \mathbf{S} (\mathbf{U}_T \otimes \mathbf{I}_{N_P})) + \tilde{\mathbf{\Lambda}}^{\dagger} \right]_{m,m}} \tag{60}$$

$$\begin{aligned} [\bar{\mathbf{S}}]_{k,l} &= a(((k-l))_{L_P}) = a(k-l) \\ &= \begin{cases} e^{j\pi r k^2/L_P} e^{-j2\pi r k l/L_P} e^{j\pi r l^2/L_P}, & \text{for even } L_P \\ e^{j\pi r k(k+1)/L_P} e^{-j2\pi r k l/L_P} e^{j\pi r l(l-1)/L_P}, & \text{for odd } L_P. \end{cases} \end{aligned} \tag{71}$$

$$a(l) = \begin{cases} W_N^{\lfloor l/Q \rfloor ((l))_Q}, & L_P = 2^{2n} \\ (W_N^{\lfloor l/Q \rfloor ((l))_Q} + j W_N^{\lfloor l/Q \rfloor (Q + ((l))_Q)})/(1+j), & L_P = 2^{2n-1}. \end{cases} \tag{74}$$

$$[\tilde{\mathbf{S}}_{k_1,l_1}]_{k_0,l_0} = \begin{cases} W_N^{(k_0-l_0)((k_1-l_1))_Q}, & L_P = 2^{2n} \\ (W_N^{(k_0-l_0)((k_1-l_1))_Q} + j W_N^{(k_0-l_0)(Q+((k_1-l_1))_Q)})/(1+j), & L_P = 2^{2n-1} \end{cases} \tag{78}$$

$$\tilde{\mathbf{\Sigma}}_{k_1,l_1} = \begin{cases} \omega_0(k_1,l_1)\mathrm{diag}\{\mathbf{e}_{((k_1-l_1))_Q}\}, & L_P = 2^{2n} \\ (\omega_0(k_1,l_1)\mathrm{diag}\{\mathbf{e}_{((k_1-l_1))_Q}\} + j\omega_1(k_1,l_1)\mathrm{diag}\{\mathbf{e}_{Q+((k_1-l_1))_Q}\})/(1+j), & L_P = 2^{2n-1}. \end{cases} \tag{80}$$

Therefore, $\mathbf{U}_E$ is independent of the eigenvalues of the channel correlation matrices.

APPENDIX III
PROOF OF THEOREM 3

From (32), it can be verified that the Chu sequence $a(l)$ is periodic, i.e., $a(((l))_{L_P}) = a(l)$. The (k,l)th entry of $\overline{\mathbf{S}}$ can be expressed as (71) shown at the bottom of the previous page. Therefore, $\overline{\mathbf{S}}$ can be expressed as

$$\bar{\mathbf{S}} = \mathbf{\Phi F \Psi} \tag{72}$$

where $\mathbf{F}$ is an $L_P \times L_P$ matrix with the (k,l)th entry as $e^{-j2\pi rkl/L_P}$. Furthermore, it can be verified that the following relation holds:

$$\mathbf{F} = \mathbf{P}_\gamma \mathbf{W}_{L_P}. \tag{73}$$

Substituting (73) into (72) yields (40). This completes the proof.

APPENDIX IV
PROOF OF THEOREM 4

The lth element of the sequence given by (33) can be expressed as (74), shown at the bottom of the previous page. The sequence is periodic, and thus, the (k,l)th entry of $\overline{\mathbf{S}}$ can be expressed as

$$[\bar{\mathbf{S}}]_{k,l} = a(k-l). \tag{75}$$

Let $\tilde{\mathbf{S}} = \mathbf{P}_\alpha \bar{\mathbf{S}} \mathbf{P}_\alpha^H$, $\tilde{\mathbf{S}}_{k_1,l_1}$ be an $N \times N$ submatrix of $\tilde{\mathbf{S}}$ with the (k_0,l_0)th entry defined as

$$[\tilde{\mathbf{S}}_{k_1,l_1}]_{k_0,l_0} = [\tilde{\mathbf{S}}]_{k_1N+k_0,l_1N+l_0} \tag{76}$$

where $0 \le k_0, l_0 \le N-1$ and $0 \le k_1, l_1 \le Q-1$. Then, we have

$$[\tilde{\mathbf{S}}_{k_1,l_1}]_{k_0,l_0} = [\bar{\mathbf{S}}]_{k_0Q+k_1,l_0Q+l_1} = a((k_0-l_0)Q+k_1-l_1). \tag{77}$$

Substituting (74) into (77) yields (78), shown at the bottom of the previous page. From (78), we have

$$\tilde{\mathbf{S}}_{k_1,l_1} = \mathbf{W}_N^H \tilde{\mathbf{\Sigma}}_{k_1,l_1} \mathbf{W}_N \tag{79}$$

where $\tilde{\mathbf{\Sigma}}_{k_1,l_1}$ is defined as in (80), shown at the bottom of the previous page, where $\omega_0(k_1,l_1) = W_N^{\lfloor (k_1-l_1)/Q \rfloor ((k_1-l_1))_Q}$ and $\omega_1(k_1,l_1) = W_N^{\lfloor (k_1-l_1)/Q \rfloor)(Q+((k_1-l_1))_Q)}$. Define an $L_P \times L_P$ matrix with the entries as

$$[\tilde{\mathbf{\Sigma}}]_{k_1N+k_0,l_1N+l_0} = [\tilde{\mathbf{\Sigma}}_{k_1,l_1}]_{k_0,l_0}. \tag{81}$$

Then, we have

$$\tilde{\mathbf{S}} = (\mathbf{I}_Q \otimes \mathbf{W}_N^H) \tilde{\mathbf{\Sigma}} (\mathbf{I}_Q \otimes \mathbf{W}_N). \tag{82}$$

Note that $\tilde{\mathbf{\Sigma}}$ has only one nonzero element in each row and each column, which can be converted to a diagonal matrix by using a row or column permutation. It can be verified that $\mathbf{\Sigma} = \tilde{\mathbf{\Sigma}} \mathbf{P}_\beta^H$ is diagonal. Substituting $\tilde{\mathbf{\Sigma}} = \mathbf{\Sigma P}_\beta$ into (82), and then, into $\bar{\mathbf{S}} = \mathbf{P}_\alpha^H \tilde{\mathbf{S}} \mathbf{P}_\alpha$ yields (41). This completes the proof.

REFERENCES

[1] *Framework and Overall Objectives of the Future Development of IMT-2000 and Systems Beyond IMT-2000*, Recommendation ITU-R M.1645, 2003.

[2] X. H. You, G. A. Chen, M. Chen, and X. Q. Gao, "Toward beyond 3G: The FuTURE project of China," *IEEE Commun. Mag.*, vol. 43, no. 1, pp. 70–75, Jan. 2005.

[3] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications—Where Fourier meets Shannon," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 29–48, May 2000.

[4] G. L. Stuber, J. R. Barry, S. W. Mclaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.

[5] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equaliza for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.

[6] N. Benvenuto and S. Tomasin, "On the comparison between OFDM and single carrier modulation with a DFE using a frequency-domain feedforward filter," *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 947–955, Jun. 2002.

[7] A. Gusmão , R. Dinis, and N. Esteves, "On frequency-domain equalization and diversity combining for broadband wireless communications," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1029–1033, Jul. 2003.

[8] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space–time block-coded transmissions over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 304–306, Jul. 2001.

[9] J. Louveaux, L. Vandendorpe, and T. Sartenaer, "Cyclic prefixed single carrier and multicarrier transmission: Bit rate comparison," *IEEE Commun. Lett.*, vol. 7, no. 4, pp. 180–182, Apr. 2003.

[10] S. L. Zhou and G. B. Giannakis, "Single-carrier space–time block-coded transmissions over frequency-selective fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 164–179, Jan. 2003.

[11] X. Q. Gao, X. H. You, B. Jiang, Z. W. Pan, and X. Y. Wang, "Generalized multi-carrier transmission technique for beyond 3G mobile communications," in *Proc. IEEE PIMRC 2005*, Sep., vol. 2, pp. 972–976.

[12] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.

[13] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 384–391, Mar. 2003.

[14] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[15] S. Ohno and G. B. Giannakis, "Capacity maximizing MMSE—Optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2138–2145, Sep. 2004.

[16] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly-selective wireless fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[17] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453–466, Mar. 2005.

[18] C. Tepedelenlioglu 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.

[19] C. S. Xiao, J. X. Wu, S. Y. Leong, Y. R. Zheng, and K. B. Letaief, "A discrete-time model for triply selective MIMO Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1678–1688, Sep. 2004.

[20] L. Deneire, B. Gyselinckx, and M. Engels, "Training sequence versus cyclic prefix—A new look on single carrier communication," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 292–294, Jul. 2001.

[21] Y. H. Zeng and T. S. Ng, "Pilot cyclic prefixed single carrier communication: Channel estimation and equalization," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 56–59, Jan. 2005.

[22] J. Coon, M. Beach, and J. McGeehan, "Optimal training sequences for channel estimation in cyclic-prefix-based single-carrier systems with transmit diversity," *IEEE Signal Process. Lett.*, vol. 11, no. 9, pp. 729–732, Sep. 2004.

[23] J. Siew, J. Coon, R. J. Piechocki, A. Dowler, A. Nix, M. Beach, S. Armour, and J. McGeehan, "A channel estimation algorithm for MIMO-SCFDE," *IEEE Commun. Lett.*, vol. 8, no. 9, pp. 555–557, Sep. 2004.

[24] D. M. Wang, J. Y. Hua, X. Q. Gao, X. H. You, M. Weckerle, and E. Costa, "Turbo detection and decoding for single-carrier block transmission systems," in *Proc. IEEE PIMRC 2004*, Sep., vol. 2, pp. 1163–1167.
[25] D. M. Wang, X. Q. Gao, and X. H. You, "Low complexity turbo receiver for multi-user STBC block transmission systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2625–2632, Oct. 2006.
[26] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[27] H. Bolcskei, M. Borgmann, and A. J. Paularj, "Impact of the propagation environment on the performance of space–frequency coded MIMO-OFDM," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 427–439, Apr. 2003.
[28] J. H. Kotecha and A. M. Sayeed, "Transmit signal design of optimal estimation of correlated MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 546–557, Feb. 2004.
[29] D. C. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.
[30] R. L. Frank and S. A. Zadoff, "Phase shift pulse codes with good periodic correlation properties," (Correspondence), *IRE Trans. Inf. Theory*, vol. IT-8, no. 6, pp. 381–382, Oct. 1962.
[31] A. Milewski, "Periodic sequences with optimal properties for channel estimation and fast start-up equalization," *IBM J. Res. Dev.*, vol. 27, no. 5, pp. 426–431, Sep. 1983.
[32] B. M. Popovic, "Generalized chirp-like polyphase sequences with optimal correlation properties," *IEEE Trans. Inf. Theory*, vol. 38, no. 4, pp. 1406–1409, Jan. 1992.
[33] P. J. Davis, *Circulant Matrices*. New York: Wiley, 1979.
[34] V. Sanchez, P. Garcia, A. M. Peinado, J. C. Segura, and A. J. Rubio, "Diagonalizing properties of the discrete cosine transforms," *IEEE Trans. Signal Process.*, vol. 43, no. 11, pp. 2631–2641, Nov. 1995.
[35] Z. D. Wang, "Fast algorithms for the discrete W transform and for the discrete Fourier transform," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. Assp-32, no. 4, pp. 803–816, Aug. 1984.
[36] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

Xiqi Gao (S'92–A'96–M'02–SM'07) received the Ph.D. degree in electrical engineering from the Southeast University, Nanjing, China, in 1997.

In April 1992, he joined the Department of Radio Engineering, Southeast University, where he is currently a Professor of information systems and communications. From September 1999 to August 2000, he was a Visiting Scholar at the Massachusetts Institute of Technology, Cambridge, and Boston University, Boston, MA. His current research interests include broadband multicarrier transmission for beyond 3G mobile communications, space–time wireless communications, iterative detection/decoding, and signal processing for wireless communications.

Dr. Gao received the Science and Technology Progress Awards of the State Education Ministry of China in 1998 and 2006. He is currently an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

Bin Jiang (S'06) received the B.S. degree in electrical engineering in 2002 from the Southeast University, Nanjing, China, where he is currently working toward the Ph.D. degree in communication and signal processing at the School of Information Science and Engineering.

His current research interests include the area of signal processing and wireless communication for multiinput multioutput (MIMO) systems.

Xiaohu You (M'91) received the M.S. and Ph.D. degrees in electrical engineering from the Southeast University, Nanjing, China, in 1985 and 1988, respectively.

Since 1990, he has been with the National Mobile Communications Research Laboratory, Southeast University, where he is currently a Professor and the Director. His current research interests include mobile communications, advanced signal processing, and applications. He is the author or coauthor of two published books and more than 20 articles published in IEEE journals. From 1993 to 1997, he was a Team Leader, and was engaged in the development of China's first Global System for Mobile Communications (GSM) and code-division multiple-access (CDMA) trial systems. During 1998, he was the Premier Foundation Investigator of the China National Science Foundation. From 1999 to 2001, on leave from the Southeast University, he was the Chief Director of the China's Third-Generation (C3G) Mobile Communications R&D Project. He is currently responsible for organizing China's B3G R&D activities under the National 863 High-Tech Program.

Prof. You is the Chairman of the China 863-FuTURE Expert Committee.

Zhiwen Pan received the Ph.D. degree from Nanjing University, Nanjing, China, in 1998.

From 1998 to 2000, he was a Postdoctoral Fellow at the Southeast University, Nanjing, where he was an Associate Professor since 2000 and a Professor after 2004. Since 2002, he has been involved in the research and development of Beyond Third-Generation (3G) demo system. He is the author or coauthor of more than 20 published papers, and holds 10 patents. His current research interests include space–time signal processing in mobile communications, transceiver design in multi-input multi-output (MIMO) systems, and design and implementation of wireless communications systems.

Yisheng Xue (S'01–M'02) received the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2002.

From August 2002 to October 2004, he was a Postdoctoral Researcher at the University Duisburg-Essen, Germany. From November 2004 to June 2005, he was a Research Associate at McGill University, Canada. Since August 2005, he has been a Senior Research Scientist with the Corporate Technology, Siemens Ltd. China, Beijing. His current research interests include protocol design and signal processing for wireless communications.

Egon Schulz received the Diploma in physics from the University of Siegen, Siegen, Germany, in 1982, and the Ph.D. degree from the Department of Electrical Engineering, Technical University of Darmstadt, Darmstadt, Germany, in 1988.

From 1982 to 1988, he was a Research Assistant at the Technical University of Darmstadt, where he was working on the design and analysis of channel coding for communication systems. In 1988, he joined the Mobile Network Division, Siemens AG, Munich, Germany, and investigated and developed radio link control strategies for mobile radio systems. From 1992 to 1993, he was a Professor of electrical and telecommunication engineering at the University of Darmstadt. In 1993, he returned to Siemens as the Director for System Engineering for radio in the local and for wireless private branch exchange systems based on Digital Enhanced Cordless Telecommunications (DECT) and Wideband Code-Division Multiple Access (WB-CDMA). Since 1998, he has been the Director for Radio Access Network for Third-Generation Mobile Systems. Since 2000, he has been an In Charge of future radio beyond third generation (3G). He is involved in national and international projects for Beyond 3G/4G Radio Access Network.

Prof. Schulz was a member of the European Telecommunications Standards Institute (ETSI) Standardization Group for the Global System for Mobile Communications (GSM).