Finally, Lemma 1 establishes that the asymptotic behavior of (28), while integrating over either of the two polyhedra, is the same. Each multiple integral in (28) is in the form of I_M of Lemma 1, i.e., polynomial in c_M with smallest exponent $M+\sum_{i=1}^{M}(N-M+k_i)=MN$. Therefore, the upper bound (16) decays with ρ^{-MN} in low spectral efficiency, indicating diversity order is no less than MN. At the same time, MN is actually the maximum possible diversity order, so the outage probability of MMSE receiver has diversity of MN.

We now proceed to show the high-rate result, where the developments parallel those for $M=2$ in (20). For $R>M\log M$ the outage region $\mathcal{A}$ can be upper and lower bounded with orthogonal slabs along the coordinates. The first set that is a subset of $\mathcal{A}$ has M orthogonal slabs where the jth slab is defined as $\lambda_j \le d_M$ and $\lambda_{i\neq j}\ge 0$, where $d_M=\rho^{-1}((1/M)2^{R/M}-1)$. The outage region $\mathcal{A}$ is a subset of the second set of slabs whose definition is the same as the first set with d_M replaced with $\tilde{d}_M=\rho^{-1}(2^{R/M}-1)$.

Therefore, the right-hand side of the bound (16) is the same as (28) with the exception that the integration region $\mathcal{A}$ could be either of above sets. Considering the possibility of some zero k_i in (27) and the unbounded shape of $\mathcal{A}$, there are dominating terms such as

$$\int_{\lambda_j\le d_M,\lambda_{i\neq j}\ge 0}\cdots\int e^{-\sum_i \lambda_i}\lambda_j^{N-M}\times\prod_{i\neq j}^{M}\lambda_i^{N-M+k_i}d\lambda_1\cdots d\lambda_M$$

which is polynomial in d_M with the minimum exponent of $N-M+1$. This indicates that the bound (16) decays with $\rho^{-(N-M+1)}$ in high spectral efficiency. This completes the proof of Theorem 2. ■

ACKNOWLEDGMENT

The authors would like to thank Dr. G. Caire and Dr. N. Al-Dhahir for their comments.

REFERENCES

[1] G. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *J. Sel. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.

[2] N. Prasad and M. K. Varanasi, "Outage analysis and optimization for multiaccess and V-BLAST architecture over MIMO Rayleigh fading channels," in *Proc. 41th Ann. Allerton Conf. Commun., Control, Comput.*, 2003, pp. 358–367.

[3] T. Guess, H. Zhang, and T. V. Kotchiev, "The outage capacity of BLAST for MIMO channels," in *Proc. IEEE ICC*, 2003, pp. 2628–2632.

[4] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 43, no. 1, pp. 1740–1750, Feb./Mar./Apr. 1994.

[5] M. Rupp, C. Mecklenbrauker, and G. Gritsch, "High diversity with simple space time block-codes and linear receivers," in *Proc. IEEE GLOBECOM*, 2003, pp. 302–306.

[6] E. K. Onggosanusi, A. G. Dabak, T. Schmidl, and T. Muharemovic, "Capacity analysis of frequency-selective MIMO channels with suboptimal detectors," in *Proc. IEEE ICASSP*, 2002, pp. 2369–2372.

[7] H. Gao, P. J. Smith, and M. V. Clark, "Theoretical reliability of MMSE linear diversity combining in Rayleigh-fading additive interference channels," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 666–672, May 1998.

[8] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[9] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed. San Diego, CA: Academic, 1994.

Joint MIMO Channel Tracking and Symbol Decoding Using Kalman Filtering

B. Balakumar, S. Shahbazpanahi, and T. Kirubarajan

***Abstract*—In this paper, the problem of channel tracking is considered for multiple-input multiple-output (MIMO) communication systems where the MIMO channel is time-varying. We consider a class of MIMO systems where orthogonal space–time block codes are used as the underlying space–time coding schemes. For such systems, a two-step MIMO channel tracking algorithm is proposed. As the first step, Kalman filtering is used to obtain an initial channel estimate for the current block based on the channel estimates obtained for previous blocks. Then, in the second step, the so-obtained initial channel estimate is refined using a decision-directed iterative method. We show that, due to specific properties of orthogonal space–time block codes, both the Kalman filter and the decision-directed algorithm can be significantly simplified. Simulation results show that the proposed tracking method can provide results in a symbol error rate performance that is 1 dB better than that of the differential receiver.**

***Index Terms*—Channel tracking, decision-directed channel equalization, Kalman filtering, multiple-input multiple-output (MIMO) communications, space–time coding.**

I. INTRODUCTION

Multiple-input multiple-output (MIMO) communications and space–time coding have been the focus of extensive research efforts. Among different space–time coding schemes presented in the literature, orthogonal space–time block codes (OSTBCs) [1], [2] are of particular interest because they achieve full diversity at a low receiver complexity. Indeed, given the MIMO channel, the maximum likelihood (ML) optimal receiver for OSTBCs consists of a linear receiver followed by a symbol-by-symbol decoder. Also, it has recently been shown in [3] that for a majority of OSTBCs, the MIMO channel is blindly identifiable. This interesting property of OSTBCs is based on the assumption that the channel is fixed during a long enough time interval. However, the channel may be time-varying in practice due to the mobility of the transmitter and/or receiver, as well as due to the carrier frequency mismatch between the transmitter and receiver. Therefore, channel tracking is essential in these cases.

In [4], Kalman filtering has been studied in application to channel tracking for MIMO communication systems. The method of [4] is based on two assumptions. First, the underlying space–time coding scheme is based on Alamouti code [1], and therefore its application is limited to the case of two transmit antennas. Second, the channel is assumed to be time-varying during the transmission of each block. The latter assumption implies that the linear ML receiver is optimal in a mean sense [4].

Kalman filtering has been applied to the problem of MIMO channel tracking in several other research reports [5], [6]. Also, in [7], a fre-

Manuscript received September 19, 2006; revised April 27, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Simon J. Godsill.This paper was presented in parts at the 2006 IEEE Workshop on Sensor Array and Multichannel Signal Processing (IEEE SAM'06), Waltham, MA, and the 2006 European Signal Processing Conference (EUSIPCO'06), Florence, Italy.

B. Balakumar and T. Kirubarajan are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: bala@grads.ece.mcmaster.ca; kiruba@mcmaster.ca).

S. Shahbazpanahi is with the Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, ON L1H 7K4, Canada (e-mail: shahram.shahbazpanahi@uoit.ca).

Digital Object Identifier 10.1109/TSP.2007.901663

1053-587X/$25.00 © 2007 IEEE

quency domain equalization method has been proposed for single carrier MIMO systems. Particle filtering has also been used in [8] for MIMO channel tracking. However, neither of the methods of [6] and [8] has been developed for OSTBCs. In [9], a Kalman filtering approach has been used in the maximization step of an expectation-maximization (EM) method to track the frequency selective MIMO channel when the underlying code is an OSTBC and when an orthogonal frequency division multiplexing (OFDM) is used. However, this method does not make use of the structure of the underlying OSTBC to simplify Kalman filtering.

In this paper, we extend the result of [4] for any type of OSTBCs and show that Kalman filtering can be significantly simplified due to the specific structure of OSTBCs. Unlike [4], we assume that the channel is fixed during the transmission of each block of data, and it can only change between blocks. Based on such an assumption, we develop a two-step channel tracking algorithm. In the first step, Kalman filtering is used at the beginning of each block to obtain an initial channel estimate for that block based on the channel estimate obtained for previous block. In the second step, to improve the quality of the channel estimate obtained by Kalman filtering, we propose a simple iterative channel estimation technique. This iterative method is in fact a decision-directed algorithm and it consists of sequential use of a linear receiver and a linear channel estimator.

We should mention that the idea of iterating the MIMO channel and symbol estimates is very common in the literature (see, for example, [10]). However, our paper presents two contributions to the field. First, we extend the results in [4] to any type of OSTBCs. Second, we use interesting properties of OSTBCs to show that the Kalman filtering based channel tracking can be significantly simplified.

The rest of the paper is organized as follows. Section II provides a background on the topic. We present our channel tracking algorithm in Section III. Simulation results are presented in Section IV and conclusions are drawn in Section V.

II. Background

Consider a MIMO system with N transmit and M receive antennas. We consider a block transmission scheme and assume that within the block period T the channel is fixed, i.e., the channel is assumed to be *quasi-static*. However, between different blocks the channel can change. Based on such an assumption, the nth received block can be written as

$$\mathbf{Y}(n) = \mathbf{X}(n)\mathbf{H}(n) + \mathbf{V}(n) \tag{1}$$

where $\mathbf{Y}(n)$ is the $T \times M$ matrix of the received signals, $\mathbf{X}(n)$ is the $T \times N$ matrix of transmitted signals, $\mathbf{V}(n)$ is the $T \times M$ matrix of noise, and $\mathbf{H}(n)$ is the $N \times M$ channel matrix during the nth block period. The noise $\mathbf{V}(n)$ is assumed to be zero-mean complex Gaussian and both spatially and temporally white with variance $\sigma_v^2/2$ per real dimension.

In space–time block coding, the matrix $\mathbf{X}(n)$ is a mapping that transforms a block of complex symbols to a $T \times N$ complex matrix. Hence, we hereafter replace $\mathbf{X}(n)$ with $\mathbf{X}(\mathbf{s}(n))$ where $\mathbf{s}(n)$ is the nth symbol vector of length K. Let us define $\mathbf{s}(n)$ as $\mathbf{s}(n) = [\,s_1(n) \quad s_2(n) \quad ,\ldots, \quad s_K(n)\,]^T$ where $(\cdot)^T$ denotes the transpose operator. The $T \times N$ matrix $\mathbf{X}(\mathbf{s}(n))$ is called an OSTBC [1], [2] if: 1) all elements of $\mathbf{X}(\mathbf{s}(n))$ are linear functions of the K complex variables $s_1(n), s_2(n), \ldots, s_K(n)$ and their complex conjugates and 2) for any arbitrary $\mathbf{s}$, $\mathbf{X}(\mathbf{s}(n))$ satisfies $\mathbf{X}^H(\mathbf{s}(n))\mathbf{X}(\mathbf{s}(n)) = \|\mathbf{s}(n)\|^2\mathbf{I}_N$ where $\mathbf{I}_N$ is the $N \times N$ identity matrix, $\|\cdot\|$ is the Euclidean norm, and $(\cdot)^H$ denotes Hermitian transpose.

It follows from the definition of OSTBC that matrix $\mathbf{X}(\mathbf{s}(n))$ can be written as

$$\mathbf{X}(\mathbf{s}(n)) = \sum_{k=1}^{K}(\mathbf{C}_k \mathrm{Re}\{s_k(n)\} + \mathbf{D}_k \mathrm{Im}\{s_k(n)\}) \tag{2}$$

where $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real and imaginary parts, respectively, and $\mathbf{C}_k$ and $\mathbf{D}_k$ matrices are defined as[1]

$$\mathbf{C}_k = \mathbf{X}(\mathbf{u}_k) \tag{3}$$
$$\mathbf{D}_k = \mathbf{X}(j\mathbf{u}_k) \tag{4}$$

where $\mathbf{u}_k$ is the kth column of identity matrix $\mathbf{I}_K$ and $j = \sqrt{-1}$. Let us define the "underline" operator for a matrix $\mathbf{P}$ as

$$\underline{\mathbf{P}} \triangleq \begin{bmatrix} \mathrm{vec}\{\mathrm{Re}(\mathbf{P})\} \\ \mathrm{vec}\{\mathrm{Im}(\mathbf{P})\} \end{bmatrix} \tag{5}$$

where $\mathrm{vec}\{\cdot\}$ refers to the vectorization operator stacking all the columns of a matrix on top of each other. Using (2) and (5), one can rewrite (1) as [11]

$$\tilde{\mathbf{y}}(n) \triangleq \underline{\mathbf{Y}(n)} = \mathbf{A}\left(\mathbf{H}(n)\right)\tilde{\mathbf{s}}_n + \tilde{\mathbf{v}}_n \tag{6}$$

where $\tilde{\mathbf{s}}_n \triangleq \underline{\mathbf{s}(n)}$, $\tilde{\mathbf{v}}_n \triangleq \underline{\mathbf{V}(n)}$ and the $2MT \times 2K$ *real* matrix $\mathbf{A}(\mathbf{H}(n))$ is given by

$$\mathbf{A}(\mathbf{H}(n)) = [\underline{\mathbf{C}_1\mathbf{H}(n)} \quad \ldots \quad \underline{\mathbf{C}_K\mathbf{H}(n)} \quad \underline{\mathbf{D}_1\mathbf{H}(n)} \quad \ldots \quad \underline{\mathbf{D}_K\mathbf{H}(n)}]. \tag{7}$$

It has been shown that for any channel matrix $\mathbf{H}(n)$, the matrix $\mathbf{A}(\mathbf{H}(n))$ satisfies the so-called *decoupling* property, i.e., its columns are orthogonal to each other and have identical norms [12]. More specifically, it satisfies

$$\mathbf{A}^T(\mathbf{H}(n))\mathbf{A}(\mathbf{H}(n)) = \|\mathbf{H}(n)\|_F^2\mathbf{I}_{2K} \tag{8}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Let us define the $2MN \times 1$ *time-varying channel vector* $\mathbf{h}(n)$ as $\mathbf{h}(\mathbf{n}) \triangleq \underline{\mathbf{H}(n)}$. With a small abuse of notation, we hereafter replace $\mathbf{A}(\mathbf{H}(n))$ with $\mathbf{A}(\mathbf{h}(n))$. Therefore, we rewrite (8) as

$$\mathbf{A}^T(\mathbf{h}(n))\mathbf{A}\left(\mathbf{h}(n)\right) = \|\mathbf{h}(n)\|^2\mathbf{I}_{2K}. \tag{9}$$

Since $\mathbf{A}(\mathbf{h}(n))$ is linear in $\mathbf{h}(n)$, there exists a unique $4KMT \times 2MN$ matrix $\mathbf{\Phi}$ such that $\mathrm{vec}\{\mathbf{A}(\mathbf{h}(n))\} = \mathbf{\Phi}\mathbf{h}(n)$ where $\mathbf{\Phi}$ is a $4KMT \times 2MN$ matrix whose kth column is given by

$$[\mathbf{\Phi}]_k = \mathrm{vec}\{\mathbf{A}(\mathbf{e}_k)\}. \tag{10}$$

Here, $[\cdot]_k$ denotes the kth column of a matrix and $\mathbf{e}_k$ is the kth column of identity matrix $\mathbf{I}_{2MN}$. Note that matrix $\mathbf{\Phi}$ can be written as $\mathbf{\Phi} = [\mathbf{\Phi}_1^T \quad \mathbf{\Phi}_2^T \quad \ldots \quad \mathbf{\Phi}_{2K}^T]^T$ where each submatrix $\mathbf{\Phi}_k$ $(k = 1, \ldots 2K)$ describes the linear relationship between the kth column of $\mathbf{A}(\mathbf{h}(n))$ and $\mathbf{h}(n)$, i.e., $[\mathbf{A}(\mathbf{h}(n))]_k = \mathbf{\Phi}_k\mathbf{h}(n)$.

Given the channel vector $\mathbf{h}(n)$, the optimal ML decoder for OSTBCs consists of a linear receiver followed by a symbol-by-symbol decoder [13]. Indeed, the linear receiver computes $\hat{\tilde{\mathbf{s}}}_n$, the estimate of $\tilde{\mathbf{s}}_n$ as

$$\hat{\tilde{\mathbf{s}}}_n = \frac{1}{\|\,\mathbf{h}(n)\,\|^2}\mathbf{A}^T\left(\mathbf{h}(n)\right)\tilde{\mathbf{y}}_n. \tag{11}$$

[1]In fact, any OSTBC is completely defined by the set of matrices $\{\mathbf{C}_k, \mathbf{D}_k\}_{k=1}^K$.

Then, the symbol-by-symbol decoder builds the estimate $\hat{\mathbf{s}}(n)$, of vector $\mathbf{s}(n)$ as $\hat{\mathbf{s}}(n) = [\,\mathbf{I}_K \quad j\mathbf{I}_K\,]\,\hat{\tilde{\mathbf{s}}}_n$. The kth element of $\hat{\mathbf{s}}(n)$ is compared with all points in the constellation corresponding to $s_k(n)$ and the closest point in this constellation to the kth element of $\hat{\mathbf{s}}(n)$ is accepted as the kth decoded symbol.

Note, however, that implementation of the ML decoder requires the knowledge of the time-varying channel. If the channel is fixed, one can use training to estimate the channel in a nonblind fashion. However, in practice, the channel is time-varying, and hence tracking of the MIMO channel is required. Recently, blind channel estimation has been studied in the literature (see, for example, [3]). The blind channel estimation of [3] is based on the assumption that the channel is fixed, and hence, it is not applicable to time-varying channels.

Without assuming any model for the MIMO channel, the problem of joint channel tracking and symbol detection is ill-posed. Fortunately, in many practical scenarios, the wireless channels can be modeled with a few parameters. It has been shown in [16] that the first-order autoregressive (AR) model can be used as a sufficiently precise method to describe the time-varying behavior of wireless channels. Based on this model, we assume that the channel variation between adjacent blocks is modeled as a first-order AR model, i.e.,

$$\mathbf{H}(n) = \alpha\mathbf{H}(n-1) + \mathbf{W}(n) \tag{12}$$

where $\mathbf{W}(n)$ is an $N \times M$ noise matrix that is assumed to be zero-mean complex Gaussian with independent entries and variance of $\sigma_w^2/2$ per real dimension. This implies that $\mathbf{W}(n)$, and consequently $\mathbf{H}(n)$, are zero-mean wide-sense stationary processes. The parameter α is a complex scalar that can be estimated using the method of [15], and hence, it is herein assumed to be known. The noise variance σ_w^2 and α are related as $\sigma_w^2 = \sigma_h^2(1-|\alpha|^2)$ where σ_h^2 is the variance of each element of $\mathbf{H}(n)$ and $|\cdot|$ denotes the amplitude of a complex number.

III. Kalman-Filter-Based Channel Tracking

In this section, we study the problem of channel tracking via Kalman filtering. We propose a two-step channel tracking algorithm. In the first step of this algorithm, Kalman filtering is used to obtain an initial channel estimate for each block based on the channel estimates obtained for the previous blocks. In the second step, the so-obtained initial channel estimate is refined using an iterative decision-directed technique, which involves a linear ML channel estimator based on the decoded symbols. In fact, the linearity of such an ML channel estimator follows from the interesting properties of OSTBCs. We will also show that due to the specific structure of OSTBCs, Kalman-filtering-based channel tracking can be significantly simplified.

To derive the two-step channel tracking algorithm, we rewrite (6) as

$$\tilde{\mathbf{y}}_n = \mathbf{B}(\tilde{\mathbf{s}}_n)\mathbf{h}(n) + \tilde{\mathbf{v}}_n \tag{13}$$

where the $2MT \times 2MN$ real matrix $\mathbf{B}(\tilde{\mathbf{s}}_n)$ is defined as

$$\mathbf{B}(\tilde{\mathbf{s}}_n) \triangleq [\,\mathbf{A}(\mathbf{e}_1)\tilde{\mathbf{s}}_n \quad \mathbf{A}(\mathbf{e}_2)\tilde{\mathbf{s}}_n \quad \ldots \quad \mathbf{A}(\mathbf{e}_{2MN})\tilde{\mathbf{s}}_n\,] \tag{14}$$

and $\mathbf{e}_k$, as defined earlier, is the kth column of the identity matrix $\mathbf{I}_{2MN}$. The following Lemma plays an important role in simplifying the forthcoming Kalman filtering algorithm.

Lemma 1: The matrix $\mathbf{B}(\tilde{\mathbf{s}}_n)$ satisfies

$$\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{B}(\tilde{\mathbf{s}}_n) = \|\mathbf{s}(n)\|^2\mathbf{I}_{2MN}. \tag{15}$$

Proof: We first show that the submatrices $\{\mathbf{\Phi}_k\}_{k=1}^{2K}$ satisfy the following equations:

$$\mathbf{\Phi}_l^T\mathbf{\Phi}_m = \begin{cases}\mathbf{I}_{2MN}, & \text{if } l = m\\ -\mathbf{\Phi}_m^T\mathbf{\Phi}_l, & \text{if } l \neq m.\end{cases} \tag{16}$$

To prove (16), we use the decoupling property in (9). Indeed, for any channel vector $\mathbf{h}$, the decoupling property in (9) implies that

$$[\mathbf{A}(\mathbf{h})]_l^T[\mathbf{A}(\mathbf{h})]_l = \|\mathbf{h}\|^2 \tag{17}$$

or

$$\mathbf{h}^T\mathbf{\Phi}_l^T\mathbf{\Phi}_l\mathbf{h} = \mathbf{h}^T\mathbf{h}. \tag{18}$$

Since (18) holds true for any $\mathbf{h}$ and because $\mathbf{\Phi}_l^T\mathbf{\Phi}_l$ is a symmetric matrix, we conclude that $\mathbf{\Phi}_l^T\mathbf{\Phi}_l = \mathbf{I}_{2MN}$. To prove the second part of (16), based on the fact that different columns of $\mathbf{A}(\mathbf{h})$ are orthogonal to each other, we can write

$$\left.\begin{aligned}[\mathbf{A}(\mathbf{h})]_l^T[\mathbf{A}(\mathbf{h})]_m &= \mathbf{h}^T\mathbf{\Phi}_l^T\mathbf{\Phi}_m\mathbf{h} = 0\\ [\mathbf{A}(\mathbf{h})]_m^T[\mathbf{A}(\mathbf{h})]_l &= \mathbf{h}^T\mathbf{\Phi}_m^T\mathbf{\Phi}_l\mathbf{h} = 0\end{aligned}\right\} \Rightarrow \mathbf{h}^T \times \left(\mathbf{\Phi}_l^T\mathbf{\Phi}_m + \mathbf{\Phi}_m^T\mathbf{\Phi}_l\right)\mathbf{h} = 0. \tag{19}$$

Since (19) holds true for any vector $\mathbf{h}$ and since $\mathbf{\Phi}_l^T\mathbf{\Phi}_m + \mathbf{\Phi}_m^T\mathbf{\Phi}_l$ is a symmetric matrix, we conclude that $\mathbf{\Phi}_l^T\mathbf{\Phi}_m + \mathbf{\Phi}_m^T\mathbf{\Phi}_l = 0$. This completes the proof of (16).

We now use (16) to prove (15). To do so, we can write the matrix $\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{B}(\tilde{\mathbf{s}}_n)$ as in (20), shown at the bottom of the page, where we have used (14). Note also that

$$\tilde{\mathbf{s}}_n^T\underbrace{\mathbf{A}^T(\mathbf{e}_l)\mathbf{A}(\mathbf{e}_l)}_{\|\mathbf{e}_l\|^2\mathbf{I}_{2K}}\tilde{\mathbf{s}}_n = \|\tilde{\mathbf{s}}_n^T\|^2 \tag{21}$$

which follows from the decoupling property. For $l \neq m$, the following set of equalities holds true:

$$\begin{aligned}\tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_l)\mathbf{A}(\mathbf{e}_m)\tilde{\mathbf{s}}_n &= \sum_{r=1}^{2K}\sum_{s=1}^{2K}\tilde{s}_{n,r}[\mathbf{A}(\mathbf{e}_l)]_r^T[\mathbf{A}(\mathbf{e}_m)]_s\tilde{s}_{n,s}\\ &= \sum_{r=1}^{2K}\sum_{s=1,s\neq r}^{2K}\tilde{s}_{n,r}\left(\mathbf{e}_l^T\mathbf{\Phi}_r^T\mathbf{\Phi}_s\mathbf{e}_m\right)\tilde{s}_{n,s}\\ &\quad + \sum_{r=1}^{2K}\tilde{s}_{n,r}\left(\mathbf{e}_l^T\overbrace{\mathbf{\Phi}_r^T\mathbf{\Phi}_r}^{\mathbf{I}_{2MN}}\mathbf{e}_m\right)\tilde{s}_{n,r}\end{aligned}$$

$$\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{B}(\tilde{\mathbf{s}}_n) = \begin{bmatrix}\tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_1)\mathbf{A}(\mathbf{e}_1)\tilde{\mathbf{s}}_n & \tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_1)\mathbf{A}(\mathbf{e}_2)\tilde{\mathbf{s}}_n & \cdots & \tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_1)\mathbf{A}(\mathbf{e}_{2MN})\tilde{\mathbf{s}}_n\\ \tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_2)\mathbf{A}(\mathbf{e}_1)\tilde{\mathbf{s}}_n & \tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_2)\mathbf{A}(\mathbf{e}_2)\tilde{\mathbf{s}}_n & \cdots & \tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_2)\mathbf{A}(\mathbf{e}_{2MN})\tilde{\mathbf{s}}_n\\ \vdots & \vdots & \ddots & \vdots\\ \tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_{2MN})\mathbf{A}(\mathbf{e}_1)\tilde{\mathbf{s}}_n & \tilde{\mathbf{s}}_n^T\mathbf{A}^T(\mathbf{e}_{2MN})\mathbf{A}(\mathbf{e}_2)\tilde{\mathbf{s}}_n & \cdots & \tilde{\mathbf{s}}_n\mathbf{A}^T(\mathbf{e}_{2MN})\mathbf{A}(\mathbf{e}_{2MN})\tilde{\mathbf{s}}_n\end{bmatrix} \tag{20}$$

$$= \sum_{r=1}^{2K} \sum_{s=1, s\neq r}^{2K} \tilde{s}_{n,r} \left(\mathbf{e}_l^T \mathbf{\Phi}_r^T \mathbf{\Phi}_s \mathbf{e}_m\right) \tilde{s}_{n,s} + \sum_{r=1}^{2K} \tilde{s}_{n,r} \overbrace{\mathbf{e}_l^T \mathbf{e}_m}^{0} \tilde{s}_{n,r}$$
$$= -\sum_{r=1}^{2K} \sum_{s=1, s\neq r}^{2K} \tilde{s}_{n,s} \left(\mathbf{e}_l^T \mathbf{\Phi}_s^T \mathbf{\Phi}_r \mathbf{e}_m\right) \times \tilde{s}_{n,r} + 0$$
$$= -\sum_{s=1}^{2K} \sum_{r=1, r\neq s}^{2K} \tilde{s}_{n,s} \left(\mathbf{e}_l^T \mathbf{\Phi}_s^T \mathbf{\Phi}_r \mathbf{e}_m\right) \times \tilde{s}_{n,r}$$
$$= -\tilde{\mathbf{s}}_n^T \mathbf{A}^T (\mathbf{e}_l) \mathbf{A} (\mathbf{e}_m) \tilde{\mathbf{s}}_n$$

where $\tilde{s}_{n,r}$ is the rth element of $\tilde{\mathbf{s}}_n$. Therefore, we obtain that, for $l \neq m$

$$\tilde{\mathbf{s}}_n^T \mathbf{A}^T(\mathbf{e}_l)\mathbf{A}(\mathbf{e}_m)\tilde{\mathbf{s}}_n = 0. \tag{22}$$

It follows from (21) and (22) that $\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{B}(\tilde{\mathbf{s}}_n) = \|\tilde{\mathbf{s}}_n\|^2\mathbf{I}_{2MN} = \|\mathbf{s}(n)\|^2\mathbf{I}_{2MN}$ and the proof is complete. ■

It follows from (13) and (15) that given $\tilde{\mathbf{s}}_n$, the ML estimate of the channel vector $\mathbf{h}(n)$ can be obtained as

$$\hat{\mathbf{h}}_{\mathrm{ML}}(n) = \frac{1}{\|\mathbf{s}(n)\|^2}\mathbf{B}^T(\tilde{\mathbf{s}}_n)\tilde{\mathbf{y}}_n. \tag{23}$$

Therefore, if the information symbols were available, the optimal ML channel estimation would involve a linear estimator as in (23). However, in practice, the information symbols are not available and they have to be estimated. To overcome this problem, one can use an iterative decision-directed channel estimation scheme. That is, given an initial channel estimate for the nth block, say $\hat{\mathbf{h}}^{(0)}(n)$, one can replace $\mathbf{h}(n)$ in (11) with $\hat{\mathbf{h}}^{(0)}(n)$ and obtain an estimate for $\tilde{\mathbf{s}}_n$, say $\hat{\tilde{\mathbf{s}}}_n^{(0)}$. This estimate of $\tilde{\mathbf{s}}_n$ will, in turn, be used in (23) instead of $\tilde{\mathbf{s}}(n)$ to obtain a new estimate for $\mathbf{h}(n)$, say $\hat{\mathbf{h}}^{(1)}(n)$. This new channel estimate will again be used in (11) instead of $\mathbf{h}(n)$ to obtain a new estimate of $\mathbf{s}(n)$. This procedure is repeated until the normalized difference between two consecutive channel estimates is negligible. The accuracy of this iterative decision-directed channel estimation scheme depends on the availability of a precise enough initial channel vector estimate $\hat{\mathbf{h}}^{(0)}(n)$. We propose to use Kalman filtering to obtain the initial channel estimate, $\hat{\mathbf{h}}^{(0)}(n)$, based on the channel estimates obtained for the previous blocks as well as the nth block received data. In what follows, we discuss the details of the Kalman filtering technique when applied to our MIMO channel tracking problem. We show that using Lemma 1, the Kalman filter can be simplified significantly. To show this, we use (13) as the observation model of the Kalman filter [14]. Note that the data model in (13) is real-valued. To obtain a real-valued state transition equation, we can rewrite (12) as

$$\mathbf{h}(n) = \mathbf{F}\mathbf{h}(n-1) + \mathbf{w}(n) \tag{24}$$

where $\mathbf{F} \triangleq \begin{bmatrix} \mathrm{Re}(\alpha)\mathbf{I}_{MN} & -\mathrm{Im}(\alpha)\mathbf{I}_{MN} \\ \mathrm{Im}(\alpha)\mathbf{I}_{MN} & \mathrm{Re}(\alpha)\mathbf{I}_{MN} \end{bmatrix}$ and $\mathbf{w}(n) = \underline{\mathbf{W}(n)}$ is the real-valued process noise with covariance matrix $\mathbf{Q} = \left(\sigma_w^2/2\right)\mathbf{I}_{2MN}$. We can use (24) as the real-valued state transition equation required for Kalman filtering.

The Kalman filtering problem for channel tracking in OSTBC-based MIMO communication system can now be formally stated as follows: Given the measurement-to-state matrix $\mathbf{B}(\tilde{\mathbf{s}}_n)$, use the observed data $\tilde{\mathbf{y}}_n$ to find the minimum mean squared error (MMSE) estimate of the components of the state vector $\mathbf{h}(n)$ for each $n \geq 1$.

Given the estimate of the state at time $n-1$, i.e., $\mathbf{h}(n-1|n-1)$, and the associated error covariance matrix $\mathbf{P}(n-1|n-1)$, the Kalman filter [14] is used to obtain the estimate of the state at time n, i.e., $\mathbf{h}(n|n)$ and the associated error covariance matrix $\mathbf{P}(n|n)$. The Kalman filtering algorithm can be summarized as follows:

$$\mathbf{h}(n|n-1) = \mathbf{F}\mathbf{h}(n-1|n-1) \tag{25}$$
$$\mathbf{P}(n|n-1) = \mathbf{F}\mathbf{P}(n-1|n-1)\mathbf{F}^T + \mathbf{Q} \tag{26}$$
$$\hat{\tilde{\mathbf{y}}}_n = \mathbf{B}(\tilde{\mathbf{s}}_n)\mathbf{h}(n|n-1) \tag{27}$$
$$\boldsymbol{\nu}(n) = \tilde{\mathbf{y}}_n - \hat{\tilde{\mathbf{y}}}_n \tag{28}$$
$$\mathbf{P}_\nu(n) = \mathbf{R} + \mathbf{B}(\tilde{\mathbf{s}}_n)\mathbf{P}(n|n-1)\mathbf{B}^T(\tilde{\mathbf{s}}_n) \tag{29}$$
$$\mathbf{G}(n) = \mathbf{P}(n|n-1)\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{P}_\nu^{-1}(n) \tag{30}$$
$$\mathbf{h}(n|n) = \mathbf{h}(n|n-1) + \mathbf{G}(n)\boldsymbol{\nu}(n) \tag{31}$$
$$\mathbf{P}(n|n) = \mathbf{P}(n|n-1) - \mathbf{G}(n)\mathbf{P}_\nu(n)\mathbf{G}^T(n) \tag{32}$$

where $\mathbf{h}(n|n-1)$ is the predicted state, $\mathbf{P}(n|n-1)$ is the covariance matrix of the predicted state, $\hat{\tilde{\mathbf{y}}}_n$ is the predicted observation, $\boldsymbol{\nu}(n)$ is the innovation process, $\mathbf{P}_\nu(n)$ is the innovation covariance matrix, $\mathbf{G}(n)$ is the Kalman gain [14], and $\mathbf{R} = E\{\tilde{\mathbf{v}}_n\tilde{\mathbf{v}}_n^T\}$ is the covariance matrix of the measurement noise $\tilde{\mathbf{v}}_n$. As we assumed that the measurement noise is spatio-temporally white with a variance of $\sigma_v^2/2$ per real dimension, $\mathbf{R} = (\sigma_v^2/2)\mathbf{I}_{2MT}$ holds true.

The following Lemma uses the result of Lemma 1 to reduce the computational complexity of finding $\mathbf{P}_\nu^{-1}(n)$ in (30).

Lemma 2: If $\mathbf{P}(n-1|n-1)$ is a diagonal matrix, then, $\mathbf{P}(n|n-1)$ in (26) and $\mathbf{P}(n|n)$ in (32) are also diagonal, i.e., if

$$\mathbf{P}(n-1|n-1) = \delta_{n-1}\mathbf{I}_{2MN} \tag{33}$$

then

$$\mathbf{P}(n|n-1) = \beta_n\mathbf{I}_{2MN} \tag{34}$$
$$\mathbf{P}(n|n) = \delta_n\mathbf{I}_{2MN} \tag{35}$$

where

$$\beta_n = \delta_{n-1}|\alpha|^2 + \frac{\sigma_w^2}{2} \text{ and } \delta_n = \frac{\sigma_v^2\beta_n}{2\|s(n)\|^2\beta_n + \sigma_v^2}. \tag{36}$$

Proof: Substituting (33) into the predicted state in (26), we can rewrite it as

$$\mathbf{P}(n|n-1) = \delta_{n-1}\mathbf{F}\mathbf{F}^T + \mathbf{Q} = |\alpha|^2\delta_{n-1}\mathbf{I}_{2MN} + \mathbf{Q} = \underbrace{\left(|\alpha|\delta_{n-1}^2 + \frac{\sigma_w^2}{2}\right)}_{\beta_n}\mathbf{I}_{2MN}. \tag{37}$$

Inserting (37) into (29) and using matrix inversion lemma, $\mathbf{P}_\nu^{-1}(n)$ in (29) can be written as

$$\mathbf{P}_\nu^{-1}(n) = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{B}(\tilde{\mathbf{s}}_n) \times \left(\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{R}^{-1}\mathbf{B}(\tilde{\mathbf{s}}_n) + \mathbf{P}^{-1}(n|n-1)\right)^{-1} \times \mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{R}^{-1}$$
$$= \frac{2}{\sigma_v^2}\mathbf{I}_{2MT} - \frac{4}{\sigma_v^4}\mathbf{B}(\tilde{\mathbf{s}}_n) \times \left(\frac{2}{\sigma_v^2}\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{B}(\tilde{\mathbf{s}}_n) + \frac{1}{\beta_n}\mathbf{I}_{2MN}\right)^{-1}\mathbf{B}^T(\tilde{\mathbf{s}}_n)$$
$$= \frac{2}{\sigma_v^2}\mathbf{I}_{2MT} - \left(\frac{4\beta_n}{2\|\mathbf{s}(n)\|^2\beta_n\sigma_v^2 + \sigma_v^4}\right) \times \mathbf{B}(\tilde{\mathbf{s}}_n)\mathbf{B}^T(\tilde{\mathbf{s}}_n) \tag{38}$$

where the fact that $\mathbf{B}^T(\tilde{\mathbf{s}}_n)\mathbf{B}(\tilde{\mathbf{s}}_n) = \|\mathbf{s}(n)\|^2\mathbf{I}_{2MN}$ has been used.

Using (30) and (38), we rewrite (32) as

$$\begin{aligned}\mathbf{P}(n|n) &= \mathbf{P}(n|n-1) \\ &\quad -\mathbf{P}(n|n-1)\mathbf{B}^T(\tilde{\mathbf{s}}_n) \\ &\quad \times \left(\frac{2}{\sigma_v^2}\mathbf{I}_{2MT} - \left(\frac{4\beta_n}{2\|\mathbf{s}(n)\|^2\beta_n\sigma_v^2+\sigma_v^4}\right) \right. \\ &\quad \left. \times \mathbf{B}(\tilde{\mathbf{s}}_n)\mathbf{B}^T(\tilde{\mathbf{s}}_n)\right)\mathbf{B}(\tilde{\mathbf{s}}_n)\mathbf{P}(n|n-1) \\ &= \left(\frac{\sigma_v^2\beta_n}{2\|s(n)\|^2\beta_n+\sigma_v^2}\right)\mathbf{I}_{2MN}. \end{aligned} \tag{39}$$

The proof is complete. ■

Based on Lemma 2, if $\mathbf{P}(0|0)$ is initialized as a diagonal matrix, $\mathbf{P}(n|n-1)$ and $\mathbf{P}(n|n)$ always take the form of (34) and (35), respectively. Hence, $\mathbf{P}_\nu^{-1}$ in (30) is simplified as in (38). It is also noteworthy that using (37) and (38) the Kalman filter gain $\mathbf{G}(n)$ in (30) can be written as

$$\mathbf{G}(n) = \underbrace{\beta_n\left(\frac{2}{\sigma_v^2} - \frac{4\beta_n\|\mathbf{s}(n)\|^2}{2\|\mathbf{s}(n)\|^2\beta_n\sigma_v^2+\sigma_v^4}\right)}_{\triangleq \mu_n}(\tilde{\mathbf{s}}_n)\mathbf{B}^T. \tag{40}$$

Using (27), (28), and (40), we can simplify (31) as

$$\begin{aligned}\mathbf{h}(n|n) &= \mathbf{h}(n|n-1) + \mu_n\mathbf{B}^T \\ &\quad \times(\tilde{\mathbf{s}}_n)\left(\tilde{\mathbf{y}}(n) - \mathbf{B}(\tilde{\mathbf{s}}_n)\mathbf{h}(n|n-1)\right) \\ &= (1-\mu_n\|\mathbf{s}(n)\|^2)\mathbf{h}(n|n-1) + \mu_n\mathbf{B}^T(\tilde{\mathbf{s}}_n)\tilde{\mathbf{y}}(n). \end{aligned} \tag{41}$$

Therefore, the Kalman filtering algorithm presented in (25)–(32) can be simplified as it follows:

$$\mathbf{h}(n|n-1) = \mathbf{F}\mathbf{h}(n-1|n-1) \tag{42}$$

$$\begin{aligned}\beta_n &= \delta_{n-1}\|\alpha\|^2 + \frac{\sigma_w^2}{2}, \\ \mu_n &= \beta_n\left(\frac{2}{\sigma_v^2} - \frac{4\beta_n\|\mathbf{s}(n)\|^2}{2\|\mathbf{s}(n)\|^2\beta_n\sigma_v^2+\sigma_v^4}\right)\end{aligned} \tag{43}$$

$$\begin{aligned}\hat{\mathbf{h}}^{(0)}(n) &= \mathbf{h}(n|n) = (1-\mu_n\|\mathbf{s}(n)\|^2)\mathbf{h}(n|n-1) \\ &\quad + \mu_n\mathbf{B}^T(\tilde{\mathbf{s}}_n)\tilde{\mathbf{y}}(n)\end{aligned} \tag{44}$$

$$\delta_n = \frac{\sigma_v^2\beta_n}{2\|s(n)\|^2\beta_n+\sigma_v^2}. \tag{45}$$

We then use the so-obtained $\hat{\mathbf{h}}^{(0)}(n)$ in the aforementioned iterative procedure to improve its accuracy.

Remark 1: Note that the simplified Kalman filter requires the knowledge of the symbol vector $\tilde{\mathbf{s}}_n$ (or $\mathbf{s}(n)$). However, the primary objective is to decode $\mathbf{s}(n)$. To overcome this obstacle, we propose to replace $\tilde{\mathbf{s}}_n$ in the Kalman filter (42)–(45), by its estimate, which is obtained by replacing the true channel vector in (11) by the predicted channel vector $\mathbf{h}(n|n-1)$ as

$$\hat{\tilde{\mathbf{s}}}_n = \frac{1}{\|\mathbf{h}(n|n-1)\|^2}\mathbf{A}^T\left(\mathbf{h}(n|n-1)\right)\tilde{\mathbf{y}}_n. \tag{46}$$

Note that given the predicted channel vector $\mathbf{h}(n|n-1)$, the symbol estimate in (46) is optimal in the ML sense.

Remark 2: To initiate the whole process, we also need to obtain an accurate enough channel estimate $\hat{\mathbf{h}}(0)$ as well as its initial covariance $\delta_0\mathbf{I}_{2MN}$. To obtain such an initial channel estimate, one can use a training block $\mathbf{s}(0)$, which is known at the receiver. At the beginning of the tracking process, the receiver can then use (23) to obtain the ML estimate of $\mathbf{h}(0)$ as f

$$\hat{\mathbf{h}}(0) = \frac{1}{\|\mathbf{s}(0)\|^2}\mathbf{B}^T(\tilde{\mathbf{s}}_0)\tilde{\mathbf{y}}_0 \tag{47}$$

where $\tilde{\mathbf{s}}_0 = \underline{\mathbf{s}(0)}$ is defined.

To find δ_0, we note that

$$\delta_0\mathbf{I}_{2MN} = E\left\{\left(\hat{\mathbf{h}}(0) - E\{\hat{\mathbf{h}}(0)\}\right)\left(\hat{\mathbf{h}}(0) - E\{\hat{\mathbf{h}}(0)\}\right)^T\right\}. \tag{48}$$

From (47) it follows that

$$\begin{aligned}E\{\hat{\mathbf{h}}(0)\} &= \frac{1}{\|\mathbf{s}(0)\|^2}\mathbf{B}^T(\tilde{\mathbf{s}}_0)E\{\tilde{\mathbf{y}}_0\} \\ &= \frac{1}{\|\mathbf{s}(0)\|^2}\mathbf{B}^T(\tilde{\mathbf{s}}_0)E\{\mathbf{B}(\tilde{\mathbf{s}}_0)\mathbf{h}(0)+\tilde{\mathbf{v}}_0\} \\ &= 0\end{aligned} \tag{49}$$

where we have used the assumption that the channel is zero-mean. Therefore, we can write

$$\begin{aligned}\delta_0\mathbf{I}_{2MN} &= E\left\{\hat{\mathbf{h}}(0)\hat{\mathbf{h}}^T(0)\right\} \\ &= \frac{1}{\|\mathbf{s}(0)\|^4}\mathbf{B}^T(\tilde{\mathbf{s}}_0)E\left\{\tilde{\mathbf{y}}_0\tilde{\mathbf{y}}_0^T\right\}\mathbf{B}(\tilde{\mathbf{s}}_0) \\ &= \frac{1}{\|\mathbf{s}(0)\|^4}\mathbf{B}^T(\tilde{\mathbf{s}}_0) \\ &\quad \times\left(\mathbf{B}(\tilde{\mathbf{s}}_0)\underbrace{E\{\mathbf{h}(0)\mathbf{h}^T(0)\}}_{(\sigma_h^2/2)\mathbf{I}_{2MN}}\mathbf{B}^T(\tilde{\mathbf{s}}_0)\right. \\ &\quad \left. + \left(\frac{\sigma_v^2}{2}\right)\mathbf{I}_{2MT}\right)\mathbf{B}(\tilde{\mathbf{s}}_0) \\ &= \left(\frac{\sigma_h^2}{2} + \frac{\sigma_v^2}{2\|\mathbf{s}(0)\|^2}\right)\mathbf{I}_{2MN} \\ &= \frac{1}{2}\left(\frac{\sigma_w^2}{1-|\alpha|^2} + \frac{\sigma_v^2}{\|\mathbf{s}(0)\|^2}\right)\mathbf{I}_{2MN}.\end{aligned}$$

Therefore, using the knowledge of α, σ_v^2, σ_w^2, and $\mathbf{s}(0)$, one can obtain δ_0 as

$$\delta_0 = \frac{1}{2}\left(\frac{\sigma_w^2}{1-|\alpha|^2} + \frac{\sigma_v^2}{\|\mathbf{s}(0)\|^2}\right). \tag{50}$$

Remark 3: To avoid error propagation, we need to repeat training once in a while. The training repetition period (TRP) determines the bandwidth efficiency of the system and it is defined as the distance, in terms of number of blocks, between two consecutive training blocks.

Remark 4: In terms of computational complexity, the proposed channel tracking method enjoys the low computational complexity of linear processing. More specifically, the first step requires the computation of $\mathbf{B}^T(\hat{\tilde{\mathbf{s}}}_n)\tilde{\mathbf{y}}(n)$, and therefore, $2MT$ real multiplications are required for computation of each entry of $\mathbf{h}(n|n)$. Taking into account that $\mathbf{h}(n|n)$ is of length $2MN$, the total computational complexity of the first step is of order $\mathcal{O}(M^2NT)$. The second step is indeed an iterative algorithm. In each iteration, we need to compute four quantities: $\|\mathbf{h}^{(k-1)}(n)\|^2$, $\mathbf{A}^T\left(\mathbf{h}^{(k-1)}(n)\right)\tilde{\mathbf{y}}_n$, $\|\hat{\tilde{\mathbf{s}}}_n^k\|^2$, and $\mathbf{B}^T\left(\hat{\tilde{\mathbf{s}}}_n^{(k)}\right)\tilde{\mathbf{y}}_n$. Computing these four quantities requires $2MN$, $4KMT$, $2K$, and $4M^2NT$ real multiplications, respectively. Therefore, the computational complexity of the second step is of the order $\mathcal{O}(M^2NT)$ per iteration of the first step. The traditional Kalman filtering method involves the computation of $\mathbf{P}_\nu^{-1}(n)$. This amounts to a computational complexity of the order $\mathcal{O}(M^3T^3)$ per iteration. Therefore, the proposed method significantly reduces the computational complexity of the traditional Kalman filtering.

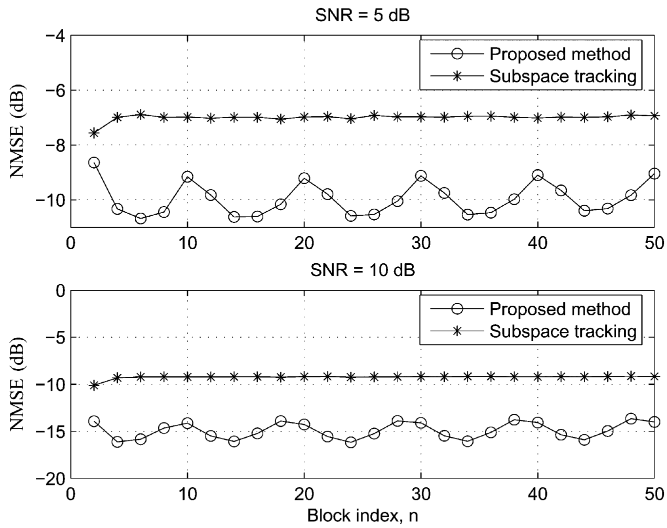

Fig. 1. NMSE of channel estimates versus block index n for $\mathrm{TRP} = 10$.

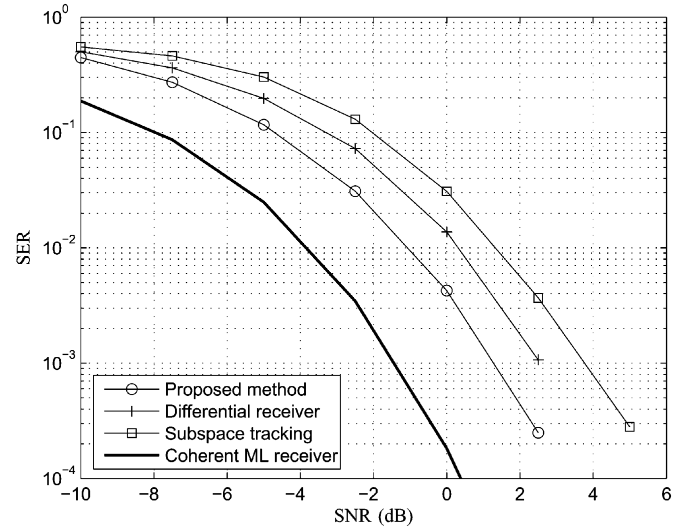

Fig. 2. SERs versus SNR for different methods and for $\mathrm{TRP} = 10$.

IV. Simulation Results

In our numerical example, we consider the 3/4 rate code of [12] with $N = M = T = 4$, and $K = 3$. The signal-to-noise ratio (SNR) is defined as σ_h^2/σ_v^2. In each simulation run, the elements of $\mathbf{H}(n)$ are generated according to Jakes model [17] corresponding to a carrier frequency of $F_o = 1.9$ GHz, terminal speed of 250 km/h and a transmission rate of 144 kb/s. This results in $\alpha = J_0(0.2\pi F_m T_s)e^{j2\pi F_o T_s} = 0.9998e^{j0.281}$ where F_m is the Doppler frequency, T_s is the sampling time, and $J_0(\cdot)$ is the zeroth-order Bessel function of first kind. In terms of channel estimation accuracy, we compare our Kalman filtering based channel tracking technique with the online implementation of the technique developed in [3]. In order to implement the method of [3] in an online manner, we have used the subspace tracking approach proposed in [3, Section III.G]. In our comparison, we use normalized mean squared error (NMSE) of the channel estimates defined as

$$\mathrm{NMSE} = E\left\{\frac{\|\mathbf{H}(n) - \hat{\mathbf{H}}(n)\|^2}{\|\mathbf{H}(n)\|^2}\right\}. \tag{51}$$

In terms of symbol-error rate (SER), we compare our method not only with the method of [3] but also with the differential space–time coding scheme [12]. The latter scheme does not require the channel to be estimated. It should be mentioned that the two methods with which the proposed method is compared do not require regular transmission of pilots, whereas the proposed method does.

Fig. 1 shows the NMSE of the channel estimates versus the block index n, for different methods and for two different values of SNR. In this figure, $\mathrm{TRP} = 10$ blocks is chosen. As can be seen from this figure, compared to the method of [3], the proposed channel tracking scheme has a lower NMSE as it tracks the channel between every two pilots. Fig. 2 illustrates the SERs of different methods, versus SNR, for $\mathrm{TRP} = 10$. In this figure, we have also plotted the SER for the (clairvoyant) coherent ML receiver that is aware of the time-varying channel. It is noteworthy that the latter receiver does not correspond to any practical application and it is herein considered only for the sake of comparison. We have also plotted the performance of a differential coding scheme which uses the same OTSTBC which we have used in our method. As can be seen from this figure, for $\mathrm{TRP} = 10$, our Kalman filtering based technique outperforms the differential space–time coding scheme by 1 dB.

It is interesting to observe that our technique outperforms the technique of [3] by more than 2 dB. In fact, when applied to track a time-varying MIMO channel, the algorithm of [3] performs even worse than the differential scheme. This is not surprising as [3] assumes that the MIMO channel is fixed within the observation interval. Therefore, the method of [3] is not applicable whenever the MIMO channel variations are fairly fast.

It is worth mentioning that each training block, the performance of our technique in terms of the NMSE degrades until the next training block is received. This explains the "periodic" behavior of our algorithm.

V. Conclusion

In this paper, we investigated the problem of channel tracking for MIMO communication systems where the MIMO channel is time-varying. We considered MIMO systems where orthogonal space–time block codes are used to encode the information symbols. For such systems, we presented a two-step MIMO channel tracking algorithm. As the first step, Kalman filtering is used to obtain an initial channel estimate for the current block based on the channel estimates obtained for previous blocks. Then, in the second step, the so-obtained initial channel estimate is refined using a decision-directed iterative method. We have shown that due to the interesting properties of orthogonal space–time block codes, both the Kalman filter and the decision-directed algorithm can be significantly simplified. To initiate this method and to avoid error propagation, one needs to use a training block once in a while. The number of information carrying blocks between two consecutive training blocks (called training repetition period) is a measure of bandwidth efficiency of our channel tracking scheme. Our simulation results show that with a training repetition period of 10 blocks, our channel tracking method can have a performance, in terms of symbol error rate, within 2 dB from the coherent ML receiver.

References

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 10, pp. 1451–1458, Oct. 1998.

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 1456–1467, Jul. 1999.

[3] S. Shabhazpanahi, A. B. Gershman, and J. H. Manton, "Closed-form blind MIMO channel estimation for orthogonal space–time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4506–4517, Dec. 2005.

[4] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filtering for time-selective fading channels," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 183–186, Feb. 2002.

[5] D. Schafhuber, G. Matz, and F. Hlawatsch, "Kalman tracking of time-varying channels in wireless MIMO-OFDM systems," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2003, pp. 1261–1265.

[6] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[7] J. Coon, S. Armour, M. Beach, and J. McGeehan, "Adaptive frequency-domain equalization for single-carrier multiple-input multiple-output wireless transmissions," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3247–3256, Aug. 2005.

[8] S. Haykin, K. Huber, and Z. Chen, "Bayesian sequential state estimation for MIMO wireless communications," *Proc. IEEE*, vol. 92, no. 3, pp. 439–454, Mar. 2004.

[9] T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, Dallas, TX, Nov. 29, 2004, vol. 4, pp. 2487–2492.

[10] Y. Jia, C. Andrieu, R. J. Piechocki, and M. Sandell, "PDA multiple model approach for joint channel tracking and symbol detection in MIMO systems," *Proc. IEE Commun.*, vol. 153, no. 4, pp. 501–507, Aug. 2006.

[11] M. Gharavi-Alkhansari and A. B. Gershman, "Constellation space invariance of space–time block codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 331–334, Jan. 2005.

[12] E. G. Larsson and P. Stoica, *Space Time Block Coding for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[13] G. Ganesan and P. Stoica, "Space-time block codes: A maximum SNR approach," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1650–1656, May 2001.

[14] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation, Tracking and Navigation: Principles, Techniques and Software*. New York: Wiley, 2001.

[15] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Process.*, vol. 53, no. 2/3, pp. 211–229, 1996.

[16] H. Wang and P. Chang, "On verifying the first-order markovian assumption for a Rayleigh fading channel model," *IEEE Trans. Vehic. Technol.*, vol. 45, no. 5, pp. 353–357, May 1996.

[17] W. C. Jakes, *Microwave Mobile Communication*. New York: Wiley, 1974.