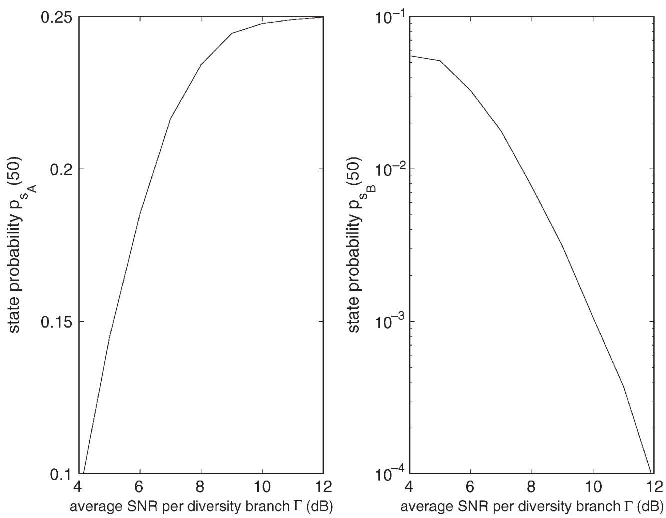

Fig. 4. State probabilities $p_{\mathrm{s_A}}(i)$ and $p_{\mathrm{s_B}}(i)$ versus Γ for the Alamouti code, two transmit antennas ($N = 2$), two symbol intervals ($L = 2$), number of receive antennas $M = 3$, time shift index $i = 50$.

in DEP decreases with time, and the DEP tends to saturate to a steady-state value. Further we see a crossover between the DFCE curve with the number of receive antennas $M = 3$ at time shift index $i = 5$ and the perfect channel estimation curve with $M = 2$ at $\Gamma = 13$ dB. This implies that the degradation in performance due to DFCE can be overcome by increasing the number of receive antennas.

Plots of the DEP, obtained both by computation and simulation versus time shift with $M = 3$ for different values of Γ are shown in Fig. 2. For the simulations, the MMSE estimate of the channel matrix is used. The computed and simulated values are seen to agree well. We again observe that the DEP increases with time shift but tends to saturate to a steady-state value. Fig. 3 shows plots of the transition probabilities p_a and p_b with respect to Γ. It is observed that p_a tends to increase and p_b tends to decrease with increase in Γ. We also find that p_a reaches a steady-state value of 0.25 implying that for higher SNR, the probability of transition to an error-free state increases.

Fig. 4 shows the plots of the state probabilities $p_{\mathrm{s_A}}(i)$ and $p_{\mathrm{s_B}}(i)$ for $i = 50$ with respect to the Γ. As with the transition probabilities, the state probability $p_{\mathrm{s_A}}(50)$ tends to increase with an increase in Γ but tends to saturate to a steady-state value of 0.25. However, $p_{\mathrm{s_B}}(50)$ decreases with an increase in Γ. This implies that with the increase in SNR, there is a higher probability to be in an error-free state.

References

[1] D. Sreedhar and A. Chockalingam, "BER analysis of space-time block codes from generalized complex orthogonal designs for M-PSK," in *Proc. IEEE Veh. Technol. Conf.—Fall*, Sep. 25–28, 2005, vol. 2, pp. 1065–1069.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[3] G. V. V. Sharma and A. Chockalingam, "Performance analysis of maximum-likelihood multiuser detection in space-time coded CDMA with imperfect channel estimation," in *Proc. IEEE Veh. Technol. Conf.—Spring*, May 17–19, 2004, vol. 3, pp. 1664–1668.

[4] Y. Gong and K. B. Letaief, "Low complexity channel estimation for space-time coded wideband OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 876–882, Sep. 2003.

[5] M. Bossert, A. Donder, and V. Zyablov, "Improved channel estimation with decision feedback for OFDM systems," *Electron. Lett.*, vol. 34, no. 11, pp. 1064–1065, May 28, 1998.

[6] M. Rim, "Optimally combining decision-directed and pilot-symbol-aided channel estimators," *Electron. Lett.*, vol. 39, no. 6, pp. 558–560, Mar. 20, 2003.

[7] Y. Ma, R. Schober, S. Pasupathy, and T. J. Lim, "Efficient BER evaluation of linear multiuser detectors with imperfect channel estimation for CDMA fading channels," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 242–250, Feb. 2006.

[8] J. E. Smee and N. C. Beaulieu, "Error-rate evaluation of linear equalization and decision feedback equalization with error propagation," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 656–665, May 1998.

[9] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[10] R. K. Mallik, D. Singh, and A. Kumari, "Analysis of rake reception with multiple symbol weight estimation for antipodal signaling," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1721–1729, Oct. 2003.

Blind Multi-Input–Multi-Output Channel Tracking Using Decision-Directed Maximum-Likelihood Estimation

Ebrahim Karami and Mohsen Shiva

***Abstract*—In this paper, a new channel-estimation algorithm based on maximum-likelihood (ML) algorithm for estimation and tracking of the multiple-input–multiple-output (MIMO) channels is presented. The ML algorithm presents the optimum estimation when the exact channel model is known. The derived channel-estimation algorithm is very efficient, with a computational complexity comparable to the least mean square and much lower than the recursive least squares and the Kalman algorithms. The proposed algorithm is analyzed, and the effect of the channel-tracking error is applied as a modifying component for the derived algorithm. The proposed algorithm is simulated for half- and full-rank flat-fading time-varying MIMO channels for the different values of $f_{\mathrm{D}}T$, E_b/N_0, and training lengths via Monte Carlo simulation technique. The minimum mean-square-error (mmse) joint detector is considered as the detection algorithm. The output of the mmse receiver is considered as the virtual training data in the blind mode of operation: the same as in the decision-directed algorithm. By various simulations, the bit error rate and the mse of tracking the proposed algorithm for different values of $f_{\mathrm{D}}T$, presenting the speed of channel variations, are evaluated and compared with the Kalman filtering approach. By simulating the proposed algorithm for different values of the training length, the minimum training length required for different channel conditions is extracted.**

***Index Terms*—Blind, channel tracking, decision-directed algorithm (DDA), maximum likelihood (ML), mean square error (mse), minimum mse (mmse), multiple-input multiple-output (MIMO), training.**

I. Introduction

In recent years, multiple-input–multiple-output (MIMO) channels have been introduced to achieve high data rates required by the next-generation wireless communication systems [1]. The use of MIMO channels, when bandwidth is limited, has much higher spectral efficiency versus single-input–single-output (SISO), single-input–multiple-output, and multiple-input–single-output channels [2].

Manuscript received July 20, 2004; revised November 17, 2004, August 6, 2005, December 10, 2005, December 27, 2005, and January 11, 2006. This work was supported in part by Center of Excellence on Electromagnetic Systems, Department of Electrical and Computer Engineering, University of Tehran. The review of this paper was coordinated by Prof. C. Xiao.

E. Karami is with the Centre for Wireless Communications, University of Oulu, Oulu 90014, Finland (e-mail: ebkarami@ee.oulu.fi).

M. Shiva is with the Department of Electrical and Computer Engineering, University of Tehran, Tehran 14399, Iran.

Digital Object Identifier 10.1109/TVT.2007.891479

Moreover, the diversity gain of MIMO channels is nearly of second order when channel matrix has full rank. Therefore, by employing the MIMO channels, not only does the mobility of wireless communications increase but also their robustness against fading, thus, making it efficient for the requirements of the next-generation wireless services.

To achieve the maximum capacity and diversity gain in MIMO channels, some optimization problems such as power [3], rate adaptation [4], code design [5], [6], joint detection [7], [8], channel estimation [9], [10], and tracking [11], [12] should be considered. Except for channel estimation that by itself is independent of the other aforementioned problems and is obligatory, the previously cited problems are dependent on the structure of the system for which channel estimation is prerequisite in their implementation. Joint detection is when all substreams are detected jointly, without which intersubstream interference occurs. Joint detection algorithms used in MIMO channels are developed based on multiuser-detection algorithms in code-division multiple-access systems. The optimum joint detection algorithm is the maximum-likelihood (ML) sequence estimation [13]. The computational complexity of the optimum receiver is impracticable when the number of transmitting substreams is large [14]. Moreover, with inaccurate channel information occurring when the channel-estimator tracking speed is not sufficient for accurate tracking of the channel variations, the implementation of the optimum receiver is more complex. Therefore, the suboptimum joint detection algorithms are more efficient solutions. In this paper, the minimum mean-square-error (mmse) detector is used as the joint detector [15], [16] because of its reasonable complexity and its soft output.

In SISO channels, especially in the flat-fading case, the channel estimation and its precision do not have a drastic impact on the performance of the receiver. Whereas, in MIMO channels, especially in outdoor MIMO channels, the precision and the convergence speed of the channel estimator have a very critical effect on the performance of the receiver [17], [18]. In SISO communications, channel estimators may or may not use the training sequence. When using the training data in the beginning, center, end, or both the beginning and the end of each block of the transmitting data, a predefined sequence is added. If training sequence is not used, the estimator is called the blind channel estimator. The blind channel estimator uses information latent in statistical properties of the transmitting data [19]. The derivation of the statistical properties of the data can be performed directly or indirectly. The scope of indirect blind methods is based on soft [20] or hard [21] decision-directed algorithms (DDAs) that use the previous estimation of the channel for the detection of data and apply it for the estimation of the channel in the last snapshot. Therefore, with decision directing, most of the nonblind algorithms can be implemented as blind. In full-rank MIMO channels, the use of an initial training data is mandatory, and without it, the channel estimator does not converge. In most of the previous works, block-fading channels are assumed, i.e., the assumption of a nearly constant channel state in the length of a block of data [22], [23]. In these works, the MIMO channel state is estimated using the training data in the beginning of the block which is applied for the detection of data in its remaining part. With the nonblock fading assumption, channel tracking must be performed in the nontraining part of the data where the corresponding algorithms are called semiblind algorithms. One of the most well-known tracking algorithms is the Kalman filtering estimation proposed by Komninakis *et al.* [11], [12]. In these papers, a Kalman filter is used as a MIMO channel tracker. The performance of this algorithm is shown to be relatively acceptable for Ricean channels where a part of the channel, due to line-of-sight components, is deterministic, but this algorithm has a high complexity. In [24], the ML estimator is proposed for tracking the MIMO channels. This algorithm extracts equations for the ML estimation of a time-invariant channel and extends it to a time-variant channel. Therefore, it does not have a desirable performance for time-varying channels. In [25], the ML algorithm that has an efficient tracking performance is derived for time-varying MIMO channels.

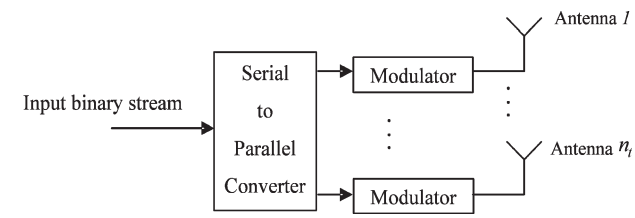

Fig. 1. Block diagram of a simple spatial multiplexed MIMO transmitter.

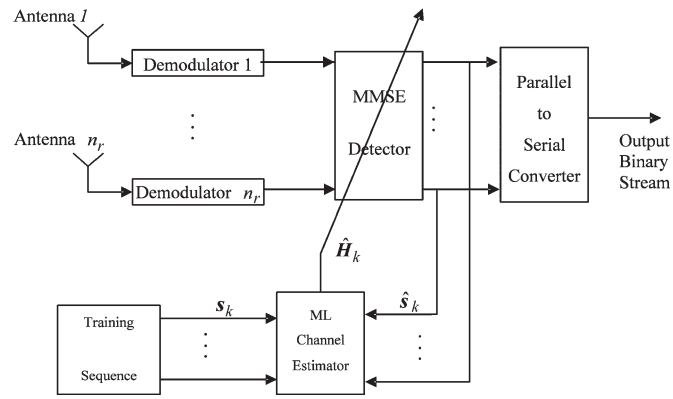

Fig. 2. Block diagram of the receiver.

In this paper, the ML MIMO channel-tracking algorithm is derived and analyzed, and the mean square of channel-tracking error that is used as a modifying component in the proposed algorithm is computed. Therefore, an efficient low-complexity MIMO channel-tracking algorithm is derived.

The rest of this paper is organized as follows. In Section II, models used for signal transmission and channel are introduced. In Section III, the proposed algorithm is derived for the first-order flat-fading MIMO channels. In Section IV, the blind-mode operation of the proposed algorithm is presented. In Section V, simulation results of the proposed receiver are presented. Concluding remarks are stated in Section VI. Note that in Appendix, the analysis of the proposed algorithm is presented.

II. System Model

Block diagrams of the transmitter and the receiver are shown in Figs. 1 and 2, respectively. In a spatial multiplexed MIMO system with n_t transmitter and n_r receiver with omnidirectional antennas, as shown in Fig. 1, a data stream which is multiplexed with training data is demultiplexed into n_t substreams where each substream is transmitted by one of the transmitting antennas. In the receiver, considering the flat Rayleigh fading channel, the linear combination of all transmitted data streams is observed under additive white Gaussian noise (AWGN). Specifically, the observed signal r_k^i from receiver i (with $i = 1, \ldots, n_\mathrm{r}$) at discrete time index k is

$$r_k^i = \sum_{j=1}^{n_\mathrm{t}} h_k^{i,j} s_k^j + w_k^i \tag{1}$$

where s_k^j is the transmitted symbol in the time index k, w_k^i is the AWGN in the ith received element, and $h_k^{i,j}$ is the channel coefficient between the jth input and the ith output of the MIMO channel.

Therefore, in each time instance, $n_{\mathrm{t}}n_{\mathrm{r}}$ channel parameters must be estimated. In a Rice channel, each of these channel parameters is a combination of a constant part and a time varying part, but in Rayleigh channel considered in this paper, $h_k^{i,j}$ is time varying because of non-LoS components that severely vary in the duration of data-block transmission with the autocorrelation given by the following equation [26]:

$$E\left\{h_k^{i,j}\left[h_l^{i,j}\right]^*\right\} \cong J_0\left(2\pi f_{\mathrm{D}}^{i,j}T|k-l|\right). \tag{2}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, superscript $*$ denotes the complex conjugate, $f_{\mathrm{D}}^{i,j}$ is the Doppler frequency shift for channel coefficient between the jth transmitter and the ith receiver, and T is the duration of each symbol. According to the wide sense stationary uncorrelated scattering model of Bello [27], all the channel taps are independent, namely, all $h_k^{i,j}$s vary independently according to the autocorrelation model of (2). The normalized spectrum for each tap $h_k^{i,j}$ is

$$S_k(f) = \begin{cases} \dfrac{1}{\pi f_{\mathrm{D}}^{i,j}T\sqrt{1-\left(\frac{f}{f_{\mathrm{D}}^{i,j}}\right)^2}}, & |f| < f_{\mathrm{D}}^{i,j}T \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The exact modeling of the vector process $h_k^{i,j}$ with a finite length auto-regressive (AR) model is impossible. For the implementation of a channel estimator, MIMO channel variations $h_k^{i,j}$ can be approximated by the following AR process of order L

$$h_k^{i,j} = \sum_{l=1}^{L} \alpha_{i,j,l}\, h_{k-l}^{i,j} + v_{i,j,k} \tag{4}$$

where $\alpha_{i,j,l}$ is the lth coefficient of the channel coefficient between the jth transmitter and the ith receiver, and $v_{i,j,k}$s are zero-mean independent identically distributed (i.i.d.) complex Gaussian processes with variances given by

$$E\left(v_{i,j,k}[v_{i,j,k}]^*\right) = \sigma_{v_{i,j,k}}^2. \tag{5}$$

The optimum selection of channel AR model parameters from correlation functions can be derived by solving the L following Wiener equations:

$$\begin{aligned} J_0\left(2\pi f_{\mathrm{D}}^{i,j}T|k-t|\right) &= \sum_{l=1}^{L} J_0\left(2\pi f_{\mathrm{D}}^{i,j}T|k-l-t|\right)\alpha_{i,j,l} \\ t &= k-L, k-L+1, \ldots, k-1. \end{aligned} \tag{6}$$

The length of the channel model must be chosen to a minimum of 90% of the energy spectrum of each channel coefficient contained in the frequency range of $|f| < f_{\mathrm{D}}^{i,j}T$.

The speed of channel variations is dependent on the Doppler shifts or equivalently on the relative velocity between the transmitters and the receivers' elements. A reasonable assumption, conventional in most of the scenarios, is the equal Doppler shifts, i.e., $f_{\mathrm{D}}^{i,j} = f_{\mathrm{D}}$. This assumption does not make any change in the derived algorithm. With this assumption, the matrix coefficients of the AR model can be replaced by scalar coefficients. For the first-order flat-fading Rayleigh channel model used in the next section, (1) and (4) are simplified in matrix forms as follows:

$$\mathbf{r}_k = \mathbf{H}_k\mathbf{s}_k + \mathbf{w}_k \tag{7}$$

$$\mathbf{H}_k = \alpha\mathbf{H}_{k-1} + \mathbf{V}_k \tag{8}$$

where $\mathbf{H}_k$ is the MIMO channel matrix, $\mathbf{V}_k$ is a matrix with i.i.d. Rayleigh elements with variance $\sigma_{\mathbf{V}}^2$, $\mathbf{w}_k$ is the noise vector with i.i.d. AWGN elements with variance $\sigma_{\mathbf{w}}^2$, and α is a constant parameter which can be calculated by solving the Wiener equation as follows:

$$\alpha = \mathbf{J}_0(2\pi f_{\mathrm{D}}T). \tag{9}$$

It is obvious that larger Doppler rates lead to smaller αs and, therefore, faster channel variations. Because of the orthogonality between the channel state and the additive random part in the first-order AR channel model, the power of time-varying part of each tap is

$$P_k = E\left|h_k^{i,j}\right|^2 = \frac{\sigma_{\mathbf{V}}^2}{1-\alpha^2}. \tag{10}$$

III. Proposed Algorithm

The proposed receiver is based on the ML algorithm. For estimation of the channel matrix in each time instance, the previous values of the channel matrix and the transmitted and received vectors in that time instance are available. For estimation of the channel matrix in time instance k, the following conditional log probability must be maximized:

$$\begin{aligned} \hat{\mathbf{H}}_k &= \mathrm{ArgMax}_{\mathbf{H}_k} \\ &\times \left\{\mathrm{Log}\, P(\mathbf{H}_k|\mathbf{H}_{k-1},\ldots,\mathbf{H}_{k-L},\mathbf{s}_k,\mathbf{s}_{k-1},\ldots,\mathbf{s}_{k-p},\mathbf{r}_k)\right\}. \end{aligned} \tag{11}$$

Of course, in the blind version design of the proposed algorithm, the transmitted vector or data vector is unknown, and therefore, some changes are needed to improve (11). Without a loss of generality, the algorithm is derived for the first-order flat-fading channel model which can be extended for more complex channel models. Consequently

$$\begin{aligned} &P(\mathbf{H}_k|\mathbf{H}_{k-1},\mathbf{s}_k,\mathbf{r}_k) \\ &= \frac{P(\mathbf{H}_k,\mathbf{H}_{k-1},\mathbf{s}_k,\mathbf{r}_k)}{P(\mathbf{H}_{k-1},\mathbf{s}_k,\mathbf{r}_k)} \\ &= \frac{P(\mathbf{r}_k|\mathbf{H}_k,\mathbf{H}_{k-1},\mathbf{s}_k)P(\mathbf{H}_k,\mathbf{H}_{k-1},\mathbf{s}_k)}{P(\mathbf{r}_k|\mathbf{H}_{k-1},\mathbf{s}_k)P(\mathbf{H}_{k-1},\mathbf{s}_k)} \\ &= \frac{P(\mathbf{r}_k|\mathbf{H}_k,\mathbf{H}_{k-1},\mathbf{s}_k)P(\mathbf{H}_k,\mathbf{H}_{k-1})}{P(\mathbf{r}_k|\mathbf{H}_{k-1},\mathbf{s}_k)P(\mathbf{H}_{k-1})} \\ &= \frac{P(\mathbf{r}_k|\mathbf{H}_k,\mathbf{H}_{k-1},\mathbf{s}_k)P(\mathbf{H}_k|\mathbf{H}_{k-1})}{P(\mathbf{r}_k|\mathbf{H}_{k-1},\mathbf{s}_k)}. \end{aligned} \tag{12}$$

Therefore, we have

$$\begin{aligned} \mathrm{Log}\, P(\mathbf{H}_k|\mathbf{H}_{k-1},\mathbf{s}_k,\mathbf{r}_k) &= \mathrm{Log}\, P(\mathbf{r}_k|\mathbf{s}_k,\mathbf{H}_k,\mathbf{H}_{k-1}) \\ &\quad + \mathrm{Log}\, P(\mathbf{H}_k|\mathbf{H}_{k-1}) - \mathrm{Log}\, P(\mathbf{r}_k). \end{aligned} \tag{13}$$

The last term in (13) is independent of the maximization operands and, therefore, has no effect on the maximization. Due to the first-order Markov assumption on the channel-variation model, the first log-probability term on the right-hand side of (13) can be simplified as

$$\begin{aligned} &\mathrm{Log}\, P(\mathbf{r}_k|\mathbf{s}_k,\mathbf{H}_k,\mathbf{H}_{k-1}) \\ &= \mathrm{Log}\, P(\mathbf{r}_k|\mathbf{s}_k,\mathbf{H}_k) \\ &= -\frac{1}{2\sigma_{\mathbf{w}}^2}\|\mathbf{r}_k - \mathbf{H}_k\mathbf{s}_k\|^2 - \frac{1}{2}\mathrm{Log}\left(2\pi\sigma_{\mathbf{w}}^2\right) \end{aligned} \tag{14}$$

and the second term on the right-hand side of (13) can be computed from the channel model as

$$\operatorname{Log} P(\mathbf{H}_k|\mathbf{H}_{k-1}) = -\frac{1}{2\sigma_{\mathbf{V}}^2}\|\mathbf{H}_k - \alpha\mathbf{H}_{k-1}\|^2 - \frac{1}{2}\operatorname{Log}\left(2\pi\sigma_{\mathbf{V}}^2\right). \tag{15}$$

By applying (13)–(15) to (11), the maximization of log probability in (11) is equivalent to the following minimization:

$$\hat{\mathbf{H}}_k = \operatorname{ArgMin}_{\mathbf{H}_k} C(\mathbf{H}_\mathrm{k}) \tag{16}$$

where $C(\mathbf{H}_k)$ is a cost function as follows:

$$C(\mathbf{H}_k) = \frac{1}{2\sigma_{\mathbf{w}}^2}\|\mathbf{r}_k - \mathbf{H}_k\mathbf{s}_k\|^2 + \frac{1}{2\sigma_{\nu}^2}\|\mathbf{H}_k - \alpha\mathbf{H}_{k-1}\|^2. \tag{17}$$

Consequently, by setting the gradient of $C(\mathbf{H}_k)$ to zero, the best estimate of the channel matrix is achieved as

$$\nabla_{\mathbf{H}_k} C(\hat{\mathbf{H}}_k) = 0 \tag{18}$$

$$\nabla_{\mathbf{H}_k} C(\mathbf{H}_k) = -\frac{1}{\sigma_{\mathbf{w}}^2}(\mathbf{r}_k - \mathbf{H}_k\mathbf{s}_k)\mathbf{s}_k^H + \frac{1}{\sigma_{\mathbf{V}}^2}(\mathbf{H}_k - \alpha\mathbf{H}_{k-1}) \tag{19}$$

$$\hat{\mathbf{H}}_k = \left(\frac{\sigma_{\mathbf{V}}^2}{\sigma_{\mathbf{w}}^2}\mathbf{r}_k\mathbf{s}_k^H + \alpha\mathbf{H}_{k-1}\right)\left(\frac{\sigma_{\mathbf{V}}^2}{\sigma_{\mathbf{w}}^2}\mathbf{s}_k\mathbf{s}_k^H + \mathbf{I}_{n_\mathrm{t}}\right)^{-1}. \tag{20}$$

By applying the matrix inversion lemma, the inverse of the matrix can be omitted as follows:

$$\hat{\mathbf{H}}_k = \left(\frac{\sigma_{\mathbf{V}}^2}{\sigma_{\mathbf{w}}^2}\mathbf{r}_k\mathbf{s}_k^H + \alpha\mathbf{H}_{k-1}\right)\left(\mathbf{I}_{n_\mathrm{t}} - \frac{\frac{\sigma_{\mathbf{V}}^2}{\sigma_{\mathbf{w}}^2}\mathbf{s}_k\mathbf{s}_k^H}{1 + n_\mathrm{t}\frac{\sigma_{\mathbf{V}}^2}{\sigma_{\mathbf{w}}^2}}\right). \tag{21}$$

Equation (21) presents an iterative algorithm for tracking a time variant MIMO channel. The remaining problem in the proposed algorithm is due to the error propagation of the channel-estimation error. In the proposed algorithm, the estimated channel matrix is derived from the real value of the previous state of the channel matrix, since the estimated values of the previous states are available. For modifying the channel-estimation algorithm, the model of the channel must be modified to show the relationship between the real value and the previous estimated value of the channel. Therefore

$$\begin{aligned}\mathbf{H}_k &= \alpha\mathbf{H}_{k-1} + \mathbf{V}_k\\ &= \alpha\hat{\mathbf{H}}_{k-1} + (\alpha\mathbf{H}_{k-1} - \alpha\hat{\mathbf{H}}_{k-1} + \mathbf{V}_k).\end{aligned} \tag{22}$$

The second part of (22) is renamed as $\mathbf{V}'_k$ with covariance matrix as

$$\sigma_{\mathbf{V}'_k}^2 = \alpha^2\sigma_{\mathbf{H}_{k-1}}^2 + \sigma_{\mathbf{V}}^2 \tag{23}$$

where $\sigma_{\mathbf{H}_{k-1}}^2$ is the covariance of channel-estimation error at time instant $k-1$. Because of the identical distribution assumption for elements of channel matrix, the channel-estimation error is independent of the channel-matrix indexes. Consequently, (21) can be easily modified by replacing the variance of the channel variation with its modified version given in (23), and therefore

$$\hat{\mathbf{H}}_k = \left(\frac{\sigma_{\mathbf{V}'_k}^2}{\sigma_{\mathbf{w}}^2}\mathbf{r}_k\mathbf{s}_k^H + \alpha\hat{\mathbf{H}}_{k-1}\right)\left(\mathbf{I}_{n_\mathrm{t}} - \frac{\frac{\sigma_{\mathbf{V}'_k}^2}{\sigma_{\mathbf{w}}^2}\mathbf{s}_k\mathbf{s}_k^H}{1 + n_\mathrm{t}\frac{\sigma_{\mathbf{V}'_k}^2}{\sigma_{\mathbf{w}}^2}}\right). \tag{24}$$

Assuming that the normalized symbol power $\mathbf{s}_k^H\mathbf{s}_k = n_\mathrm{t}$, (24) is thus simplified as

$$\hat{\mathbf{H}}_k = \alpha\hat{\mathbf{H}}_{k-1} + \beta_k(\mathbf{r}_k - \alpha\hat{\mathbf{H}}_{k-1}\mathbf{s}_k)\mathbf{s}_k^H \tag{25}$$

where

$$\beta_k = \frac{\frac{\sigma_{\mathbf{V}'_k}^2}{\sigma_{\mathbf{w}}^2}}{1 + n_\mathrm{t}\frac{\sigma_{\mathbf{V}'_k}^2}{\sigma_{\mathbf{w}}^2}}. \tag{26}$$

Equation (25) presents the least mean square (LMS)-like structure whose complexity versus the number of transmitter and receiver elements is in the order of 2. For finalizing the modified channel-estimation algorithm, the channel-estimation error must be estimated; therefore, a recursive equation for estimation of the channel-estimation error is derived with variance

$$\sigma_{\mathbf{H}_k}^2 = E\left[\left|\mathbf{H}_k^i - \hat{\mathbf{H}}_k^i\right|^2\right] \tag{27}$$

where $\mathbf{H}_k^i$ is an arbitrary element of the channel matrix in the time index k, and $\hat{\mathbf{H}}_k^i$ is its estimate. The following equation is derived (see the Appendix) for the variance of the channel-estimation error:

$$\sigma_{\mathbf{H}_k}^2 = \left\{1 - \frac{\frac{\sigma_{\mathbf{V}'_k}^2}{\sigma_{\mathbf{w}}^2}}{1 + n_\mathrm{t}\frac{\sigma_{\mathbf{V}'_k}^2}{\sigma_{\mathbf{w}}^2}}\right\}\sigma_{\mathbf{V}'_k}^2. \tag{28}$$

Therefore, using (23) and (28), $\sigma_{\mathbf{V}'_k}^2$ can be calculated iteratively and substituted in (26) for extraction of the modified version of the algorithm.

For the steady-state condition, where the time index k approaches infinity, $\sigma_{\mathbf{H}_k}^2$ can be assumed approximately independent of the time index. Therefore, with this assumption, $\sigma_{\mathbf{H}_\infty}^2$ can be calculated by solving (28) for $\sigma_{\mathbf{H}_k}^2 = \sigma_{\mathbf{H}_{k-1}}^2 = \sigma_{\mathbf{H}_\infty}^2$ as (29), shown at the bottom of the page.

The value of α is usually very close to unity, because $f_\mathrm{D}T$ is very close to zero. Thus, (29) can be approximated as follows:

$$\sigma_{\mathbf{H}_\infty}^2 = \frac{1}{2}\left(n_\mathrm{t}\sigma_{\mathbf{V}}^2 + \sqrt{n_\mathrm{t}^2\sigma_{\mathbf{V}}^4 + 4\sigma_{\mathbf{V}}^2\sigma_{\mathbf{w}}^2}\right). \tag{30}$$

IV. Blind Mode of Operation

Equation (25) is derived for the training-based channel-tracking problem. Therefore, use of the proposed algorithm misses the

$$\sigma_{\mathbf{H}_\infty}^2 = \frac{1}{2\left(n_\mathrm{t}(1-\alpha^2)+\alpha^2\right)} \times \left(\sigma_{\mathbf{w}}^2(1-\alpha^2) + n_\mathrm{t}\sigma_{\nu}^2 + \sqrt{\sigma_{\mathbf{w}}^4(1-\alpha^2)^2 + 2n_\mathrm{t}\sigma_{\mathbf{V}}^2\sigma_{\mathbf{w}}^2(1-\alpha^2) + n_\mathrm{t}^2\sigma_{\mathbf{V}}^4 + 4\alpha^2\sigma_{\mathbf{V}}^2\sigma_{\mathbf{w}}^2}\right) \tag{29}$$

information contents in the nontraining part (data part) of each of the transmitted block. Thus, in this section, (11) is rewritten to cover the information for channel tracking of the nontraining case. Consequently, (11) can be rewritten as follows for the nontraining symbols:

$$[\hat{\mathbf{H}}_k, \hat{\mathbf{s}}_k] = \mathrm{ArgMax}_{\mathbf{H}_k, \mathbf{s}_k} \left\{ \mathrm{Log}\, P(\mathbf{H}_k, \mathbf{s}_k | \mathbf{H}_{k-1}, \ldots, \mathbf{H}_{k-L}, \mathbf{r}_k) \right\}. \tag{31}$$

Equation (31) displays the general relationship for the joint ML channel tracking and the data detection for memoryless MIMO channels with uncorrelated transmit vectors. When MIMO channels have memory or the transmitted vector is correlated to the previous transmitted vectors, for instance in coded signals, the previous values of transmitted and received vectors must also be considered.

This joint optimization problem can be performed in a way similar to the proposed algorithm. For this purpose, the cost function presented in (17) must be minimized for all possible combinations of the transmitted vector. Therefore, the computational complexity of this algorithm is related exponentially to the number of elements in the transmitted vector. In this paper, DDA is proposed for extending (25) to the blind case.

For the use of DDA, in each snapshot, first, the transmit vector is estimated by assuming that the channel matrix is equal to the previous snapshot. This assumption is valid because of the small value of the $f_{\mathrm{D}}T$. Then, by using the estimated transmitted vector similar to the training vector in (25), the channel tracking is done. DDA algorithms can be summarized in the following detection and estimation equations:

$$\hat{\mathbf{s}}_k = g\left[\left(\hat{\mathbf{H}}_{k-1}^H \hat{\mathbf{H}}_{k-1} + \sigma_{\mathbf{w}}^2 \mathbf{I}_{n_t} \right)^{-1} \hat{\mathbf{H}}_{k-1}^H \mathbf{r}_k \right] \tag{32}$$

$$\hat{\mathbf{H}}_k = \alpha \hat{\mathbf{H}}_{k-1} + \beta_k (\mathbf{r}_k - \alpha \hat{\mathbf{H}}_{k-1} \hat{\mathbf{s}}_k) \hat{\mathbf{s}}_k^H \tag{33}$$

where $g(\cdot)$ is a function modeling of the decision device, and β_k can be calculated from (26). In the special case of BPSK signaling, $g(\cdot)$ is a signum function. Equation (33) is like (25) where $\mathbf{s}_k$ is replaced by $\hat{\mathbf{s}}_k$ which is the output of the decision device. In semiblind applications, a multiplex of $\mathbf{s}_k$ and $\hat{\mathbf{s}}_k$ is fed to the channel estimator (Fig. 2).

The channel-estimation algorithm in both the training-based and the DDA mode is summarized as follows.

Step 1) Initialization: $\hat{\mathbf{H}}_k = \mathbf{0}_{n_r \times n_t}$ and $\sigma_{\mathbf{H}}^2 = 1$, where $\mathbf{0}_{n_r \times n_t}$ is a $n_r \times n_t$ zero matrix.
Step 2) If the algorithm is in the DDA mode, the data vector is estimated using (32).
Step 3) Calculate β_k using (23) and (26).
Step 4) Update the channel vector estimate using (33) in the blind mode and (25) in the training mode.
Step 5) Update channel-tracking error $\sigma_{\mathbf{H}_k}^2$ using (28).
Step 6) Return to 2) for the next snapshot.

The main problem in using DDA for the proposed algorithm is the error that occurs in the calculation of the covariance of the channel-estimation error and its effect on the estimate of the channel matrix given in (25). In the training-based tracking, as shown in Section V, the channel-matrix error covariance given by (28) is very precise; therefore, the operation of the proposed algorithm is optimum for symbol-by-symbol tracking of the memoryless MIMO channels. But in the nontraining-based tracking, the error in computation of the covariance of the channel-estimation error causes some degradation from optimality. This degradation is added to the degradation due to the use of DDA which will be considered in analysis and optimization of the proposed algorithm in the blind mode of operation in future works.

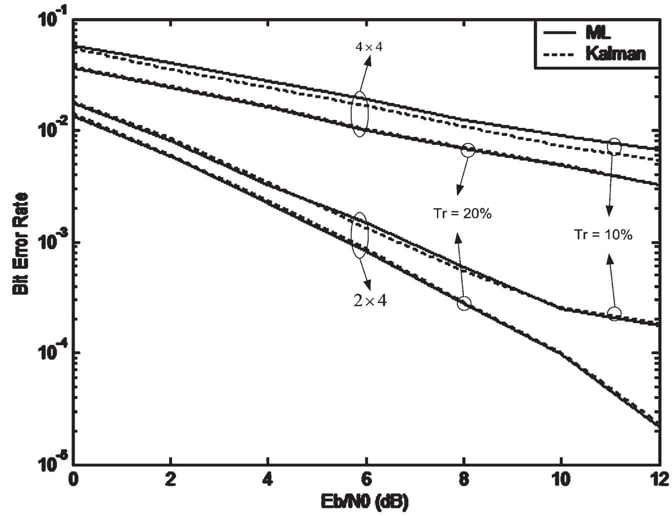

Fig. 3. BER of the proposed algorithm and the Kalman algorithm while $f_{\mathrm{D}}T = 0.004$.

V. Simulation Results

The proposed algorithm is simulated for the first-order flat-fading MIMO channels for different values of $f_{\mathrm{D}}T$ and training-data percentages by Monte Carlo simulation technique. The emphasis of this section is on the Rayleigh channel being the worst case. Of course, the effect of the fixed part of the Rice channel is the same as decreasing the $f_{\mathrm{D}}T$, which in turn decreases the share of time variations in the total channel strength.

The length of the data blocks transmitted on each transmitting antenna is considered 100 BPSK symbols, and the first section of each block is considered as the training part, and the remaining part is considered as the unknown nontraining data. In simulations for the full blind case, only 20 symbols are assumed as the initializing training symbols, and the others are assumed as the data symbols.

In the first part of simulations, the bit error rate (BER) of the proposed algorithm and the Kalman filtering approach are compared for the different values of training-data percentages and $f_{\mathrm{D}}T$ with the results shown in Figs. 3 and 4 corresponding to 2×4 and 4×4 MIMO channels, respectively. It must be noted that these results are for the uncoded MIMO systems shown in Figs. 1 and 2. In real applications, much lower BER values are obtained using high-gain error control coding with or without space–time coding.

For $f_{\mathrm{D}}T = 0.004$, shown in Fig. 3, for all cases, the ML algorithm presents a BER comparable to the Kalman filtering approach—of course while using 20% training—the ML algorithm presents a slightly better performance than the Kalman filtering approach, and vice versa for 10% training case.

For $f_{\mathrm{D}}T = 0.01$, which is shown in Fig. 4, the ML algorithm provides a slightly better BER than the Kalman filtering approach except for 10% training length in 4×4 channel. In this $f_{\mathrm{D}}T$ value, except for 20% training in 2×4 channel, bit error floors are observed.

In the second part of the simulations, the mse of tracking the proposed algorithm and the Kalman algorithm for different values of training-data percentages and $f_{\mathrm{D}}T$ for 2×4 and 4×4 MIMO channels with $E_b/N_0 = 10$ dB are shown in Figs. 5 and 6, respectively. In all simulations, the proposed algorithm and the Kalman algorithm have very close performances. In Fig. 5, i.e., for 2×4 channel case, the mse

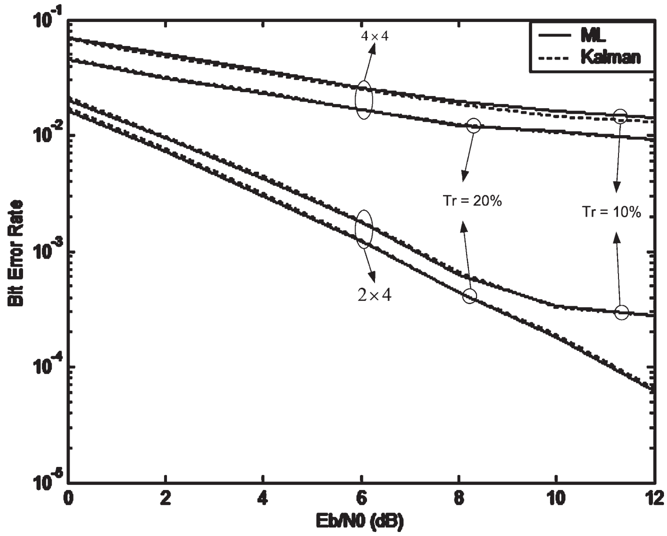

Fig. 4. BER of the proposed algorithm and the Kalman algorithm while $f_{\mathrm{D}}T = 0.01$.

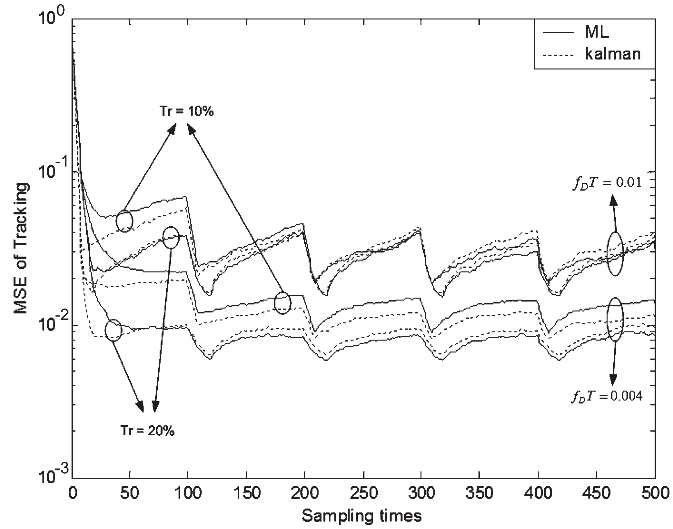

Fig. 6. MSE of tracking for the proposed algorithm and the Kalman algorithm for a 4 × 4 MIMO channel.

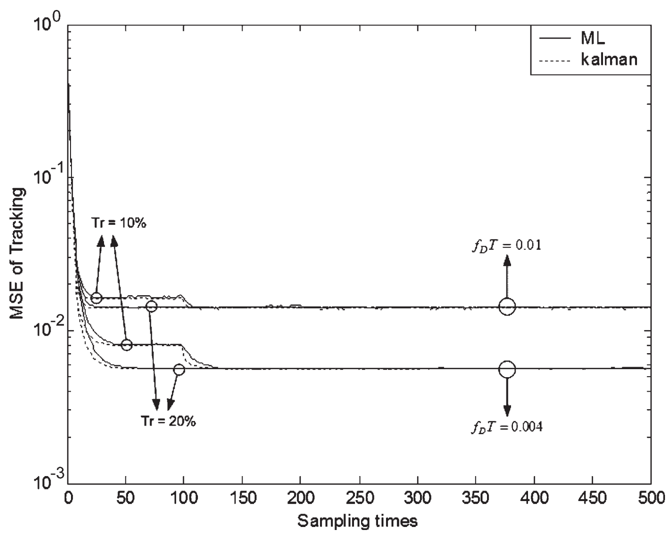

Fig. 5. MSE of tracking for the proposed algorithm and the Kalman algorithm for a 2 × 4 MIMO channel.

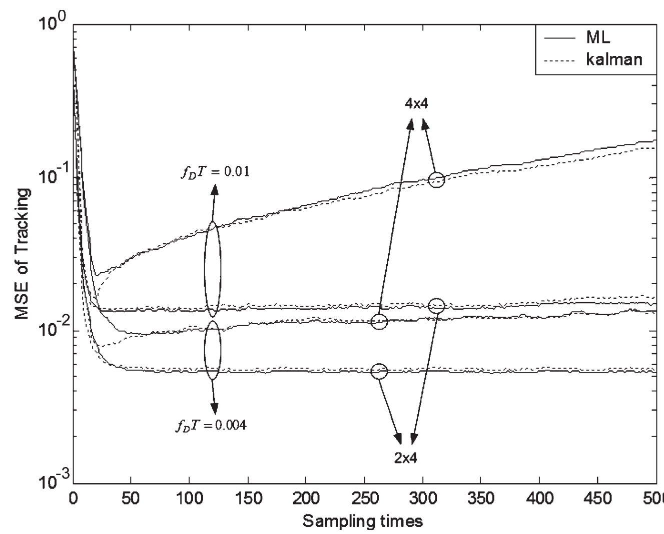

Fig. 7. MSE of tracking while using only 20 symbols initialization training.

has a near exponential decrease while operating in the training mode and is nearly constant while operating in the DD mode. Of course, the decreasing behavior of the mse occurs only once in the training mode and twice in the 20% and 10% training cases, respectively. As it can be seen from Fig. 5, in all cases, both algorithms present negligible mse of tracking.

For 4 × 4 MIMO channel, as it can be seen from Fig. 6, the results are worse. The performance with $f_{\mathrm{D}}T = 0.004$ is yet acceptable, and the mse of tracking is negligible, but with $f_{\mathrm{D}}T = 0.01$, the observed mse of tracking is relatively larger. In this channel, the exponential-like decreasing and increasing mses are observed in the training and the DD-based modes.

The performance of the ML algorithm and the Kalman filtering approaches is shown, when only 20 symbols of initial training are used, in Fig. 7. As it can be observed, except for $f_{\mathrm{D}}T = 0.01$ in 4 × 4 MIMO channel, in other cases, both algorithms can operate for a long time in the DD mode without any restriction.

Performance of the proposed algorithm versus E_b/N_0 for $f_{\mathrm{D}}T = 0.004$ and 0.01 is presented in Figs. 8 and 9, respectively. As it can be seen from these figures, for 2 × 4 MIMO channel, the mse computed from (28) nearly matches with the tracking mse measured from the simulations that is due to the low-decision error rate observed in this channel rank, but in a 4 × 4 MIMO channel, the performance of the proposed algorithm for both 10% and 20% trainings is different in full training or theoretical bound. Of course, in this case when $f_{\mathrm{D}}T = 0.004$, the mse of tracking decreases linearly. For $f_{\mathrm{D}}T = 0.01$, error floors depending on the applied training percents are observed.

Fig. 10 presents the mse of tracking versus the training percent while $E_b/N_0 = 10$ dB. The training length for very close to optimum operation (where with minimum training length, close to mmse of tracking is achieved) of the proposed algorithm is the settling point of these curves. Proper selection of these points seems to be 20 and 14 symbols for $f_{\mathrm{D}}T = 0.004$ and 0.01, respectively. In training length higher than this point, the difference between the lower bound of tracking and the mse presented by the proposed algorithm with DDA is very small, and it presents very close to optimum performance.

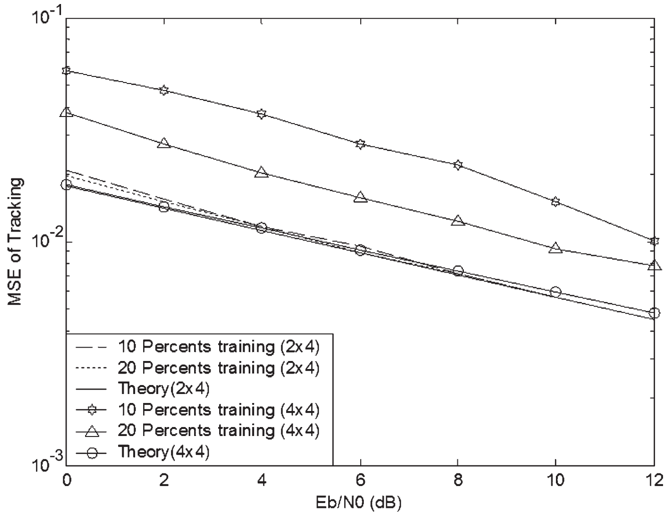

Fig. 8. MSE of the proposed algorithm versus E_b/N_0 while $f_{\mathrm{D}}T = 0.01$.

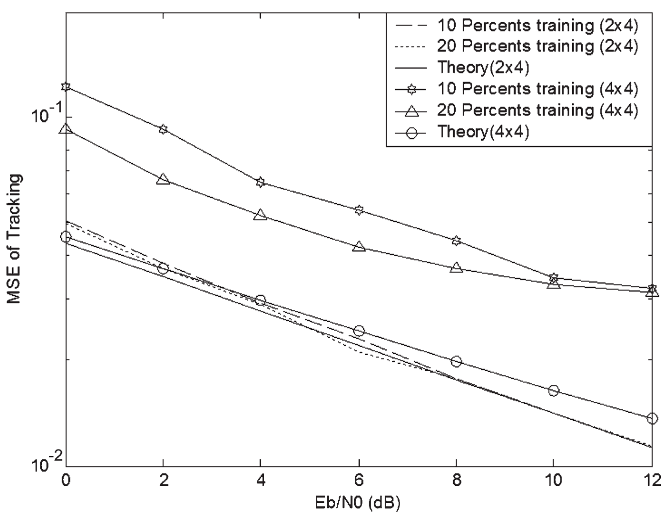

Fig. 9. MSE of the proposed algorithm versus E_b/N_0 while $f_{\mathrm{D}}T = 0.01$.

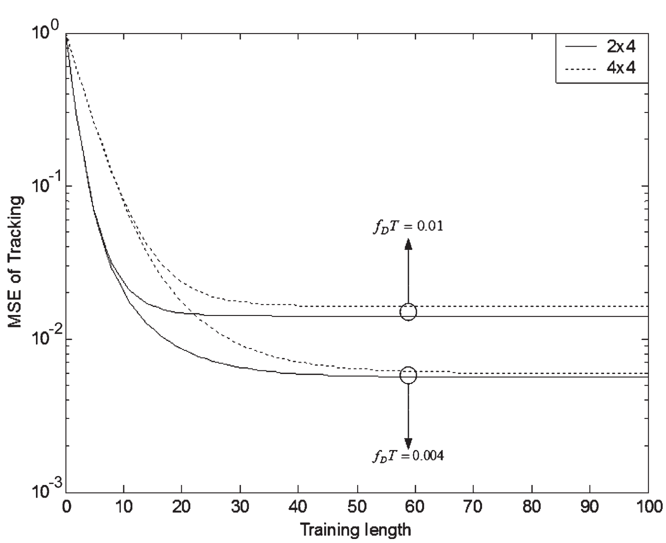

Fig. 10. MSE of the proposed algorithm versus training percent with $E_b/N_0 = 10$ dB.

VI. Conclusion

In this paper, the estimation and tracking algorithm of spatial multiplex MIMO channels based on the ML algorithm, which is the optimum channel-tracking algorithm when the channel model is known, is derived for the Rayleigh flat-fading channels. Through the analysis of the proposed algorithm, the channel-tracking error is computed, and its effect is applied to rectify the derived algorithm iteratively. The proposed algorithm is combined with the DDA to be extended as a blind tracking algorithm. The mmse joint detector is considered as the detection algorithm. The output of the mmse receiver is considered as the virtual training data in the blind mode of operation, as is the DDA. Estimation or initial tracking of the channel is assumed to be performed via training bits, while its tracking is assumed to be performed via both training and nontraining data symbols.

After simplification, it is shown that the structure and, therefore, the complexity of the proposed algorithm versus the number of transmitter and receiver elements are in the order of 2 that is similar to the LMS algorithm [28], which has much lower performance in comparison to the proposed algorithm. The BER and the mse of tracking the proposed algorithm are compared with the Kalman filtering approach, and it is shown that the performance of the proposed algorithm is slightly better than the Kalman filtering, especially in the blind mode of operation. Also, the BER presented by the ML algorithm is better than the Kalman algorithm except for $f_{\mathrm{D}}T = 0.01$ and 10% training length.

Moreover, the complexity of Kalman filtering is in the order of 5 which is much more complex than the proposed algorithm.

In all simulations, both the BER and the mse of tracking are considered as the comparison criteria. Of course, the mse of tracking can be regarded as the additional noise in the front end of the system, and therefore, its effect can be considered in the mmse receiver with modification in the noise variance.

Also, the proposed algorithm is simulated for half- and full-rank flat-fading time-varying MIMO channels for different $f_{\mathrm{D}}T$ and E_b/N_0. It is shown that in half-rank channel, the proposed algorithm presents close to lower bound performance, which is for the full training case, and is an essential property of the DDA technique. In 4×4 MIMO channel, the worse loss is observed from the full training lower bound which depends on the percentage of the applied training. This is due to the low performance of the mmse detector in full-rank channels.

Also, the proposed algorithm is simulated for different training percents in $E_b/N_0 = 10$ dB, and it is shown that the minimum of 20 and 14 training percents for $f_{\mathrm{D}}T = 0.004$ and 0.01, respectively, is sufficient for very close to optimum performance.

Appendix
Proof of (28)

Channel-estimation error $\boldsymbol{\xi}_k$ is computed as

$$\boldsymbol{\xi}_k = \hat{\mathbf{H}}_k - \mathbf{H}_k. \tag{A.1}$$

Because of i.i.d. assumption on the elements of the channel matrix, the channel-matrix error $\boldsymbol{\xi}_k$ has the following relationship with the variance of channel tracking:

$$\sigma^2_{\mathbf{H}_k} = \frac{1}{n_{\mathrm{r}} n_{\mathrm{t}}} \operatorname{trace}\left\{ E\left(\boldsymbol{\xi}_k^H \boldsymbol{\xi}_k \right) \right\}. \tag{A.2}$$

By applying (11), (12), and (25) in (A.1), we have

$$\begin{aligned}\boldsymbol{\xi}_k &= \alpha\hat{\mathbf{H}}_{k-1} + \beta_k(r_k - \alpha\hat{\mathbf{H}}_{k-1}\mathbf{s}_k)\mathbf{s}_k^H - \mathbf{H}_k \\ &= \alpha\hat{\mathbf{H}}_{k-1} + \beta_k(\mathbf{H}_k\mathbf{s}_k + \mathbf{w}_k - \alpha\hat{\mathbf{H}}_{k-1}\mathbf{s}_k)\mathbf{s}_k^H - \alpha\mathbf{H}_{k-1} - \mathbf{V}_k \\ &= \alpha(\hat{\mathbf{H}}_{k-1} - \mathbf{H}_{k-1}) \\ &\quad + \beta_k(\alpha\mathbf{H}_{k-1}\mathbf{s}_k + \mathbf{V}_k\mathbf{s}_k + \mathbf{w}_k - \alpha\hat{\mathbf{H}}_{k-1}\mathbf{s}_k)\mathbf{s}_k^H - \mathbf{V}_k \\ &= \alpha\boldsymbol{\xi}_{k-1} + \beta_k(-\alpha\boldsymbol{\xi}_{k-1}\mathbf{s}_k + \mathbf{V}_k\mathbf{s}_k + \mathbf{w}_k)\mathbf{s}_k^H - \mathbf{V}_k \\ &= \alpha\boldsymbol{\xi}_{k-1}\left(1 - \beta_k\mathbf{s}_k\mathbf{s}_k^H\right) + \beta_k\mathbf{w}_k\mathbf{s}_k^H - \mathbf{V}_k\left(1 - \beta_k\mathbf{s}_k\mathbf{s}_k^H\right).\end{aligned}$$

Therefore, because of the orthogonality between the above three terms, we get

$$\begin{aligned}E\left(\boldsymbol{\xi}_k^H\boldsymbol{\xi}_k\right) &= E\left[\alpha^2\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\boldsymbol{\xi}_{k-1}^H\boldsymbol{\xi}_{k-1}\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\right] \\ &+ \beta_k^2 E\left(\mathbf{s}_k\mathbf{w}_k^H\mathbf{w}_k\mathbf{s}_k^H\right) + E\left[\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\mathbf{V}_k^H\mathbf{V}_k\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\right].\end{aligned}$$

The above expectation operator is taken versus data vector $\mathbf{s}_k$, noise vector $\mathbf{w}_k$, and channel-variation matrix $\mathbf{V}_k$. Considering the commutative property of the trace of the two multiplied matrices, we have

$$\begin{aligned}&\text{trace}\left\{E\left(\boldsymbol{\xi}_k^H\boldsymbol{\xi}_k\right)\right\} \\ &\quad = \alpha^2 n_{\mathrm{r}}\sigma^2_{\mathbf{H}_{k-1}}\text{trace}\left\{E\left\lfloor\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\right\rfloor\right\} \\ &\qquad + \beta_k^2 n_{\mathrm{r}}\sigma^2_{\mathbf{w}}\text{trace}\left\{E\left(\mathbf{s}_k\mathbf{s}_k^H\right)\right\} \\ &\qquad - \alpha^2\sigma^2_{\mathbf{V}_k}\text{trace}\left\{E\left[\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\right]\right\}.\end{aligned}$$

With the assumption of normalized power in each transmitted antenna, $\mathbf{s}_k^H\mathbf{s}_k = n_{\mathrm{t}}$, and $E(\mathbf{s}_k\mathbf{s}_k^H) = \mathbf{I}_{n_{\mathrm{t}}}$, and therefore

$$\text{trace}\left\{E\left[\left(1-\beta\mathbf{s}_k\mathbf{s}_k^H\right)\left(1-\beta_k\mathbf{s}_k\mathbf{s}_k^H\right)\right]\right\} = 1 - 2\beta_k + n_{\mathrm{t}}\beta_k^2.$$

Consequently

$$\begin{aligned}\text{trace}\left\{E\left(\boldsymbol{\xi}_k^H\boldsymbol{\xi}_k\right)\right\} &= n_{\mathrm{r}}n_{\mathrm{t}}\alpha^2\left(1-2\beta_k+n_{\mathrm{t}}\beta_k^2\right)\sigma^2_{\mathbf{H}_{k-1}} \\ &\quad + n_{\mathrm{r}}n_{\mathrm{t}}\beta_k^2\sigma^2_{\mathbf{w}} + n_{\mathrm{r}}n_{\mathrm{t}}\left(1-2\beta_k+n_{\mathrm{t}}\beta_k^2\right)\sigma^2_{\mathbf{V}_k}.\end{aligned}$$

By using (24), we have $\sigma^2_{\mathbf{H}_k} = (1-2\beta_k+n_{\mathrm{t}}\beta_k^2)\sigma^2_{\mathbf{V}'_k} + \beta_k^2\sigma^2_{\mathbf{w}}$ or in another form, $\sigma^2_{\mathbf{H}_k} = (1-2\beta_k)\sigma^2_{\mathbf{V}'_k} + \beta_k^2(n_{\mathrm{t}}\sigma^2_{\mathbf{V}'_k} + \sigma^2_{\mathbf{w}})$.

Finally, by simplification of (A.2), we have

$$\sigma^2_{\mathbf{H}_k} = \left\{1 - \frac{\frac{\sigma^2_{\mathbf{V}'_k}}{\sigma^2_{\mathbf{w}}}}{1 + n_{\mathrm{t}}\frac{\sigma^2_{\mathbf{V}'_k}}{\sigma^2_{\mathbf{w}}}}\right\}\sigma^2_{\mathbf{V}'_k}. \tag{A.3}$$

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wirel. Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] M. A. Beach, R. L. Davies, P. Guemas, H. Xue, and J. P. McGeehan, "Capacity and service extension for future wireless networks using adaptive antennas," in *Proc. IEEE 9th ICAP*, Apr. 4–7, 1995, vol. 1, pp. 125–129.

[3] A. Zanella, M. Chiani, and M. Z. Win, "Analytical evaluation of MIMO systems with unequal power transmission in a Rayleigh fading environment," in *Proc. IEEE GLOBECOM*, Dec. 2003, no. 1, pp. 1837–1841.

[4] T. A. Thomas, K. L. Baum, and F. W. Vook, "Modulation and coding rate selection to improve successive cancellation reception in OFDM and spread OFDM MIMO systems," in *Proc. IEEE ICC*, May 2003, no. 1, pp. 2842–2846.

[5] J. H. Kotecha, Z. Hong, and A. M. Sayeed, "Coding and diversity gain tradeoff in space–time codes for correlated MIMO channels," in *Proc. IEEE GLOBECOM*, Dec. 2003, no. 1, pp. 646–650.

[6] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space–time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[7] S. Sfar, R. D. Murch, and K. B. Letaief, "Layered space–time multiuser detection over wireless uplink systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 653–668, Jul. 2003.

[8] X. Zhu and R. D. Murch, "MIMO-DFE based BLAST over frequency selective channels," in *Proc. IEEE GLOBECOM*, Nov. 2001, no. 1, pp. 499–503.

[9] R. Visoz and A. O. Berthet, "Iterative decoding and channel estimation for space–time BICM over MIMO block fading multipath AWGN channel," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1358–1367, Aug. 2003.

[10] V. Pohl, P. H. Nguyen, V. Jungnickel, and C. Von Helmolt, "How often channel estimation is needed in MIMO systems," in *Proc. IEEE GLOBECOM*, Dec. 2003, no. 1, pp. 814–818.

[11] 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.

[12] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Adaptive multi-input multi-output fading channel equalization using kalman estimation," in *Proc. ICC*, New Orleans, LA, Jun. 18–22, 2000, vol. 3, pp. 1655–1659.

[13] A. Abrardo, "Non-coherent MLSE detection for CDMA multiple-antenna systems," in *Proc. IEEE ICC*, Apr. 2002, no. 1, pp. 846–850.

[14] S. Verdú, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, no. 3, pp. 303–312, 1989.

[15] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code division multiple access channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 123–136, Jan. 1989.

[16] T. Ojanpera, "Overview of multiuser detection, interference cancellation for DS-CDMA," in *Proc. IEEE Int. Conf. Pers. Wirel. Commun.*, Bombay, India, Dec. 17–19, 1997, pp. 115–119.

[17] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 5, pp. 684–702, Jun. 2003.

[18] J. Baltersee, G. Flock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 12, pp. 2358–2368, Dec. 2001.

[19] L. Tong, "Blind sequence estimation," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2986–2994, Dec. 1995.

[20] Y. H. Kim and S. Shamsunder, "Adaptive algorithms for channel equalization with soft decision feedback," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 9, pp. 1660–1669, Dec. 1998.

[21] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a 'stop-and-go' decision-directed algorithm," *IEEE Trans. Commun.*, vol. COM-35, no. 9, pp. 877–887, Sep. 1987.

[22] A. Vosoughi and A. Scaglione, "Channel estimation for precoded MIMO systems," in *Proc. IEEE Workshop SSP*, St. Louis, MO, Sep. 28–Oct. 1, 2003, pp. 442–445.

[23] A. Stefanov and T. M. Duman, "Turbo-coded modulation for systems with transmit and receive antenna diversity over block-fading channels: System model, decoding approaches and practical considerations," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 958–968, May 2001.

[24] X. Deng, A. M. Haimovich, and J. G. Frias, "Decision directed iterative channel estimation for MIMO systems," in *Proc. IEEE ICC*, Anchorage, AK, vol. 4, no. 1, pp. 2326–2329.

[25] E. Karami and M. Shiva, "Maximum likelihood MIMO channel tracking," in *Proc. 59th IEEE VTC—Spring*, Milan, Italy, May 2004, pp. 876–879.

[26] W. C. Jakes, Jr., *Microwave Mobile Communications*. New York: Wiley, 1974.

[27] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst.*, vol. CS-11, no. 4, pp. 360–393, Dec. 1963.

[28] E. Karami, M. Shiva, and M. Khansari, "An efficient MIMO receiver structure for coded signals," in *Proc. 8th IEEE ICCS*, Nov. 2002, vol. 1, pp. 416–420.