Decision-Directed Estimation of MIMO Time-Varying Rayleigh Fading Channels

Jie Gao, *Member, IEEE*, and Huaping Liu, *Member, IEEE*

***Abstract*—This paper presents a decision-directed (DD) maximum *a posteriori* probability (MAP) channel-estimation scheme for multiple-input multiple-output (MIMO) time-varying fading channels. With the estimate of the channel matrix for the current symbol interval, a zero-forcing (ZF) receiver is applied to detect the spatially multiplexed data on a symbol-by-symbol basis. Symbol decisions are then fed to the channel predictor for estimation of channel coefficients in future symbol intervals. Simulated error performance of a ZF receiver with the DD MAP and perfect channel estimates is provided and compared.**

***Index Terms*—Decision feedback, MAP channel estimation, multiple-input multiple-output (MIMO) systems, time-varying fading.**

I. Introduction

MULTIPLE-input multiple-output (MIMO) communication systems have been shown to provide high spectral efficiencies [1]. If perfect channel coefficients are available at the receiver, a linear increase in ergodic capacity is achievable with MIMO systems [2], [3]. Perfect channel estimates, however, can be obtained only if the channel is either static for a long time (noise can be averaged out) or perfect (no noise). The rapid phase and amplitude variations inherent in a time-varying fading channel render perfect estimates impossible, regardless of the type of channel-estimation method used. Channel estimation has been studied extensively for single-antenna systems (e.g., [4]–[8]). For MIMO systems, channel-estimation schemes have been mostly based on pilot-assisted approaches, assuming a quasi-static fading model that allows the channel to be constant for a block of symbols and change independently to a new realization. In [9], a pilot-embedding method, where low-level pilots are transmitted concurrently with data, was proposed for turbo decoding in an MIMO system. The effects of pilot-assisted channel estimation on the achievable data rates (capacity lower bound) over a frequency-nonselective, quasi-static fading channel were analyzed in [10]. In this scheme, periodic pilot signals assigned to different transmit antennas are assumed to be mutually orthogonal. Although it avoids interantenna interference within the pilot periods, such scheme could significantly lower the spectral efficiency of the system. Throughput of a system with a maximum-likelihood (ML) channel estimator that employs periodic optimal training sequences for block and continuous flat-fading channels was studied in [11]. In [12], an iterative method was derived to improve the estimation of channel parameters for an MIMO system, based on the assumption that data decisions have already been made. This method needs to invert a matrix of size $L \times L$, with L being the frame length per transmit antenna, for every frame. With practical frame lengths (e.g., $L = 130$, as applied in simulations in [12]), the computational load could be prohibitively high.

In this paper, we derive a decision-directed (DD) maximum *a posteriori* probability (MAP) channel-estimation scheme for MIMO systems over time-varying fading channels. A zero-forcing (ZF) receiver is applied to detect the spatially multiplexed symbols transmitted in the current symbol interval. The estimated symbols are then incorporated in the DD MAP channel predictor to obtain estimates of the channel coefficients in future symbol intervals. The proposed scheme does not rely on the assumption of a quasi-static fading model and can be applied in a time-varying environment. Compared to most existing schemes, it has a lower complexity and is capable of operating with significantly less pilot symbols.

Manuscript received October 18, 2003; revised January 8, 2004; accepted June 16, 2004. The editor coordinating the review of this paper and approving it for publication is D. Gesbert.

The authors are with the School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR 97331 USA (e-mail: hliu@eecs.orst.edu; gaoji@engr.orst.edu).

Digital Object Identifier 10.1109/TWC.2005.852131

II. System Model

Consider a communication system with M transmit and N receive antennas, denoted as an (M, N) system, over a time-varying, frequency-nonselective Rayleigh fading channel. In the transmitter, data are serial-to-parallel converted and sent to M transmit antennas for simultaneous transmission. Each receive antenna responds to each transmit antenna through a statistically independent fading coefficient. The received signals are corrupted by additive white Gaussian noise, which is statistically independent among different receive antennas. We focus on the baseband model of a system, which employs M-ary phase-shift keying (PSK) with zero intersymbol-interference design. The results can be easily extended to an MIMO system employing a more general pulse-amplitude modulation scheme. The mth transmitted data stream (the signal from the mth transmit antenna) is expressed as

$$x_m(t) = \sum_{i=-\infty}^{\infty} \sqrt{E_s} s_m(i) g(t - iT), \qquad m = 1, \ldots, M \quad (1)$$

where $s_m(i)$ is the ith symbol of the mth data stream, E_s is the energy per symbol, T is the symbol interval, and $g(t)$ is the transmitted Nyquist pulse applied to all transmitted data streams. Energy of $g(t)$ is normalized to unity, i.e., $\int_{-\infty}^{\infty} g^2(t) \mathrm{d}t = 1$.

The time-varying fading channel introduces a random amplitude and phase shift to the transmitted signal. The fading-channel process $h(t)$ is modeled as a normalized, zero-mean complex wide-sense stationary Gaussian process with a spaced-time correlation function $\Phi(\Delta t)$ expressed as $\Phi(\Delta t) = E\{h(t)h^*(t+\Delta t)\}$, where $E\{\cdot\}$ denotes statistical expectation and $(\cdot)^*$ represents complex conjugate. In a typical mobile-communication environment, the spaced-time correlation function of the channel can be modeled as $\Phi(\Delta t) = J_0(2\pi f_\mathrm{d}\Delta t)$ [13], where f_d represents the maximum Doppler shift of the channel and $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind.

The received signal of the nth antenna $r_n(t)$, $n = 1,\ldots,N$, is the sum of signals transmitted from M transmit antennas and is expressed as

$$r_n(t) = \sum_{m=1}^{M} \sqrt{E_s} h_{n,m}(t) x_m(t) + \nu_n(t), \qquad n = 1,\ldots,N \tag{2}$$

where $h_{n,m}(t)$ represents the fading process for signals from the mth transmit antenna to the nth receive antenna and $\nu_n(t)$ is a complex zero-mean white Gaussian noise process with power spectral density N_0. The received signal $r_n(t)$ is filtered by a matched filter, matched to $g(t)$, and then sampled at the symbol rate of each data stream.

Let the $M \times 1$ transmitted signal vector in the ith symbol interval be $\boldsymbol{s}(i) = [s_1(i)s_2(i)\cdots s_M(i)]^\mathrm{T}$, where $[\cdot]^\mathrm{T}$ denotes transpose. The $N \times 1$ received signal vector at the ith discrete-time interval is obtained as

$$\boldsymbol{r}(i) = \sqrt{E_s}\boldsymbol{\mathcal{H}}(i)\boldsymbol{s}(i) + \boldsymbol{\nu}(i) \tag{3}$$

where $\boldsymbol{\nu}(i)$ is the complex zero-mean noise vector. The channel matrix $\boldsymbol{\mathcal{H}}(i)(N \times M)$ is expressed as

$$\boldsymbol{\mathcal{H}}(i) = [\,\boldsymbol{h}_1(i) \quad \boldsymbol{h}_2(i) \quad \ldots \quad \boldsymbol{h}_M(i)\,] = \begin{bmatrix} h_{1,1}(i) & h_{1,2}(i) & \cdots & h_{1,M}(i) \\ h_{2,1}(i) & h_{2,2}(i) & \cdots & h_{2,M}(i) \\ \vdots & & & \vdots \\ h_{N,1}(i) & h_{N,2}(i) & \cdots & h_{N,M}(i) \end{bmatrix} \tag{4}$$

where $N \times 1$ column vectors $\boldsymbol{h}_m(i) = [h_{1,m}(i) \;\; h_{2,m}(i) \ldots h_{N,m}(i)]^\mathrm{T}$, $m = 1,\ldots,M$, represent the channel coefficients from the mth transmit antenna to all N receive antennas. Each element of $\boldsymbol{\mathcal{H}}(i)$ is a zero-mean complex Gaussian random variable of unit variance. In the discrete-time channel formulation adopted above, the fading process is piecewise constant approximated in each symbol interval. It is assumed that the temporal variations of the fading processes $h_{n,m}(t)$ are such that the piecewise-constant, discrete-time approximation is valid. The nth component of $\boldsymbol{r}(i)$ represents the received signal from the nth receive antenna and is expressed as $r_n(i) = \sum_{m=1}^{M} \sqrt{E_s} h_{n,m}(i) s_m(i) + \nu_n(i)$. Let us assume that $\hat{\boldsymbol{\mathcal{H}}}(i-L),\ldots,\hat{\boldsymbol{\mathcal{H}}}(i-1)$, estimates of $\boldsymbol{\mathcal{H}}(i-L),\ldots,\boldsymbol{\mathcal{H}}(i-1)$, and $\hat{\boldsymbol{s}}(i-L),\ldots,\hat{\boldsymbol{s}}(i-1)$, estimates of $\boldsymbol{s}(i-L),\ldots,\boldsymbol{s}(i-1)$, have been obtained. At the beginning of the transmission, these channel coefficients could be obtained by using pilot symbols. In the proposed channel estimation and data detection scheme, $\boldsymbol{\mathcal{H}}(i)$ is obtained using $\hat{\boldsymbol{\mathcal{H}}}(i-L),\ldots,\hat{\boldsymbol{\mathcal{H}}}(i-1)$ and $\hat{\boldsymbol{s}}(i-L),\ldots,\hat{\boldsymbol{s}}(i-1)$. Then, $\boldsymbol{s}(i)$ is detected using $\hat{\boldsymbol{\mathcal{H}}}(i)$. After that, $\hat{\boldsymbol{\mathcal{H}}}(i-L+1),\ldots,\hat{\boldsymbol{\mathcal{H}}}(i)$ and $\hat{\boldsymbol{s}}(i-L+1),\ldots,\hat{\boldsymbol{s}}(i)$ are used to estimate $\boldsymbol{\mathcal{H}}(i+1)$. Periodic pilot blocks can be inserted in the data stream to improve estimation quality and to stop error propagation when the receiver is operating in the DD mode.

III. Channel Estimation and Data Detection

A. DD MAP Channel Estimation

Due to interantenna interference, it is impossible to solve for $\hat{\boldsymbol{\mathcal{H}}}(i)$ based on the received signal model given in (3), even if an estimate of symbol vector $\boldsymbol{s}(i)$ is available. Let us assume that estimates of previous symbols $\hat{s}_m(i-L), \hat{s}_m(i-L+1), \ldots, \hat{s}_m(i-1)$ and channel coefficients in previous symbol intervals $\hat{h}_{n,m}(i-L), \hat{h}_{n,m}(i-L+1),\ldots,\hat{h}_{n,m}(i-1)$ have been made. To estimate $h_{n,m}(i)$, $m = 1,\ldots,M$, $n = 1,\ldots,N$, we consider a sliding-window approach in which $\hat{h}_{n,m}(i)$ is derived from the received signals $\boldsymbol{r}(l)$, $l = i-L,\ldots,i-1$, and symbol decisions within a window of L symbols preceding the current symbol. Specifically, $y_{n,m}(l)$ is constructed as

$$y_{n,m}(l) = \frac{\left[r_n(l) - \sum\limits_{\substack{k=1 \\ k\neq m}}^{M} \sqrt{E_s}\hat{h}_{n,k}(l)\hat{s}_k(l) \right]}{\hat{s}_m(l)}$$
$$m = 1,\ldots,M, \quad n = 1,\ldots,N, \quad l = (i-L),(i-L+1),\ldots,(i-1). \tag{5}$$

Note that the nth element of $\boldsymbol{r}(l)$, $r_n(l)$, consists of the desired signal, the interantenna interference, and a noise term. Ideally, if the channel is noiseless, feedback symbol decisions are correct, and channel estimates are perfect, then $y_{n,m}(l)$ in (5) equals exactly $\sqrt{E_s}h_{n,m}(l)$, the desired component needed for the DD channel estimation. In a practical time-varying fading environment, there will be decision and channel-estimation errors, and $y_{n,m}(l)$ does not perfectly represent the desired signal component $\sqrt{E_s}h_{n,m}(l)$.

Let us define an $L \times 1$ vector $\boldsymbol{y}_{nm}(i,L)$ and an $(L+1) \times 1$ vector $\boldsymbol{x}(i,L)$ as

$$\boldsymbol{y}_{nm}(i,L) = [y_{n,m}(i-L) \; y_{n,m}(i-L+1) \; \cdots \; y_{n,m}(i-1)]^\mathrm{T} \tag{6}$$

$$\boldsymbol{x}(i,L) = \begin{bmatrix} \boldsymbol{y}_{nm}(i,L) \\ h_{n,m}(i) \end{bmatrix}. \tag{7}$$

The covariance matrix of zero-mean vector $\boldsymbol{x}(i,L)$ can be written as $\boldsymbol{F}_x = E\{\boldsymbol{x}(i,L)\boldsymbol{x}^\mathrm{H}(i,L)\} = \begin{bmatrix} \boldsymbol{F}_{x11} & \boldsymbol{f}_{x12} \\ \boldsymbol{f}_{x21} & \boldsymbol{f}_{x22} \end{bmatrix}$, where $(\cdot)^\mathrm{H}$ denotes conjugate transpose and

$$\boldsymbol{F}_{x11} = E\left\{\boldsymbol{y}_{nm}(i,L)\boldsymbol{y}_{nm}^\mathrm{H}(i,L)\right\} \tag{8a}$$
$$\boldsymbol{f}_{x12} = E\left\{\boldsymbol{y}_{nm}(i,L)h_{n,m}^\mathrm{H}(i)\right\} \tag{8b}$$
$$\boldsymbol{f}_{x21} = E\left\{h_{n,m}(i)\boldsymbol{y}_{nm}^\mathrm{H}(i,L)\right\} = \boldsymbol{f}_{x12}^\mathrm{H} \tag{8c}$$
$$f_{x22} = E\left\{h_{n,m}(i)h_{n,m}^\mathrm{H}(i)\right\}. \tag{8d}$$

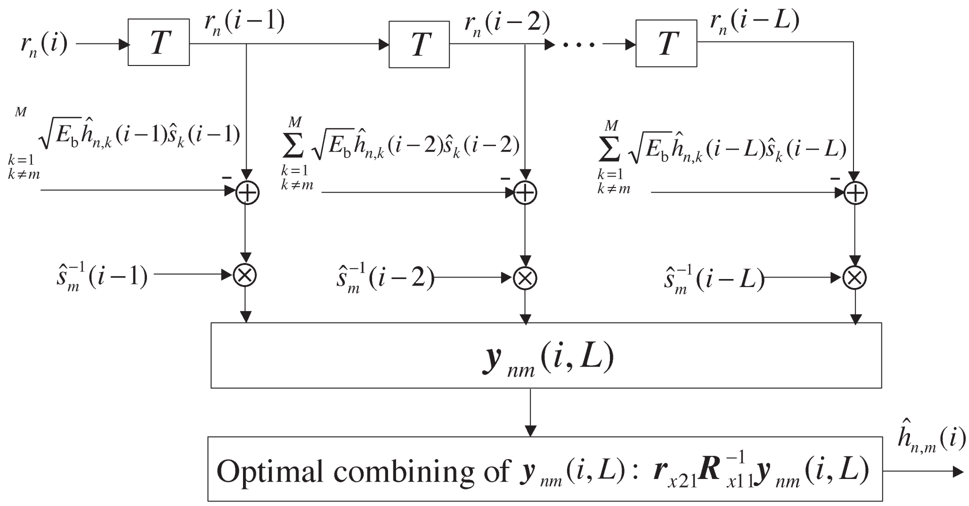

Fig. 1. The DD MAP channel predictor for MIMO systems.

Given the availability of $\boldsymbol{y}_{nm}(i,L)$, the estimate of $h_{n,m}(i)$ can be obtained by maximizing its conditional-probability density function $p\{h_{n,m}(i)|\boldsymbol{y}_{nm}(i,L)\}$ as

$$\hat{h}_{n,m}(i) \triangleq \max_{h_{n,m}(i)} p\left\{h_{n,m}(i)|\,\boldsymbol{y}_{nm}(i,L)\right\}. \tag{9}$$

For the Rayleigh channels being considered, the $(L+1)\times 1$ vector $\boldsymbol{x}(i,L)$ is complex Gaussian.[1] Therefore, the conditional-probability density function can be written as

$$\begin{aligned} p\left\{h_{n,m}(i)|\boldsymbol{y}_{nm}(i,L)\right\} &= \frac{p\left\{\boldsymbol{y}_{nm}(i,L), h_{n,m}(i)\right\}}{p\left\{\boldsymbol{y}_{nm}(i,L)\right\}} \\ &= \frac{\frac{1}{\pi^{L+1}|\boldsymbol{F}_x|}\exp\left[-\boldsymbol{x}^{\mathrm{H}}(i,L)\boldsymbol{F}_x^{-1}\boldsymbol{x}(i,L)\right]}{\frac{1}{\pi^{L}|\boldsymbol{F}_{x11}|}\exp\left[-\boldsymbol{y}_{nm}^{\mathrm{H}}(i,L)\boldsymbol{F}_{x11}^{-1}\boldsymbol{y}_{nm}(i,L)\right]} \end{aligned} \tag{10}$$

where $|\cdot|$ denotes the determinant of a matrix. By using the matrix-inversion lemma [14], $\boldsymbol{F}_x^{-1}$ is expressed as

$$\boldsymbol{F}_x^{-1} = \begin{bmatrix} \boldsymbol{F}_{x11} & \boldsymbol{f}_{x12} \\ \boldsymbol{f}_{x21} & f_{x22} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{F}_{xi11} & \boldsymbol{f}_{xi12} \\ \boldsymbol{f}_{xi21} & f_{xi22} \end{bmatrix} \tag{11}$$

where

$$\boldsymbol{F}_{xi11} = \left(\boldsymbol{F}_{x11} - \boldsymbol{f}_{x12}\, f_{x22}^{-1}\boldsymbol{f}_{x21}\right)^{-1} \tag{12a}$$

$$f_{xi22} = \left(f_{x22} - \boldsymbol{f}_{x21}\boldsymbol{F}_{x11}^{-1}\,\boldsymbol{f}_{x12}\right)^{-1} \tag{12b}$$

$$\boldsymbol{f}_{xi12} = -\,\boldsymbol{F}_{xi11}\,\boldsymbol{f}_{x12}\, f_{x22}^{-1} \tag{12c}$$

$$\boldsymbol{f}_{xi21} = -\, f_{xi22}\,\boldsymbol{f}_{x21}\boldsymbol{F}_{x11}^{-1}. \tag{12d}$$

Note that $\boldsymbol{F}_x$ and $\boldsymbol{F}_{x11}$ are fixed, and thus, independent of $\boldsymbol{x}(i,L)$. Therefore, maximizing the conditional probability density function is equivalent to minimizing the following quadratic function:

$$\lambda = \boldsymbol{x}^{\mathrm{H}}(i,L)\boldsymbol{F}_x^{-1}\boldsymbol{x}(i,L) - \boldsymbol{y}_{nm}^{\mathrm{H}}(i,L)\boldsymbol{F}_{x11}^{-1}\boldsymbol{y}_{nm}(i,L). \tag{13}$$

By letting the conjugate derivative of λ with respect to $h_{n,m}(i)$ be equal to zero, we obtain the DD MAP estimate of $h_{n,m}(i)$ as[2]

$$\hat{h}_{n,m}(i) = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{y}_{nm}(i,L) \tag{14}$$

where $\boldsymbol{w} = (\,\boldsymbol{f}_{x21}\boldsymbol{F}_{x11}^{-1})^{\mathrm{H}}$ is the $L\times 1$ tap-weight vector. This procedure needs to be done for all elements of $\boldsymbol{\mathcal{H}}(i)$ to form the estimated channel matrix $\hat{\boldsymbol{\mathcal{H}}}(i)$. Because all elements of $\boldsymbol{\mathcal{H}}(i)$ are identically distributed, tap weight $\boldsymbol{w}$ is common for all coefficients (any combination of n and m). If significant changes in the channel statistics (e.g., the Doppler shift) have occurred, however, $\boldsymbol{f}_{x21}$ and $\boldsymbol{F}_{x11}$ (and thus $\boldsymbol{w}$) must be updated to reflect such changes.

The DD MAP channel-prediction procedure is illustrated in Fig. 1. When applied to the special case of a single-antenna system, the channel estimate derived in this paper is similar to the linear minimum mean-square error (MMSE) estimate [4], [6]. The tap weight is the same, but the MAP predictor estimate derived in this paper combines previous received signals scaled by the corresponding symbol decisions that form $\hat{h}_{n,m}(i)$, whereas the MMSE estimate given in [4] and [6] combines estimates (e.g., obtained via an ML approach) of past channel trajectory up to time $i-1$.

[1]Under normal-operation conditions, there will be occasional erroneous decisions on previously sent symbols $\hat{s}_m(l)$. However, a decision error does not affect the Gaussian distribution of $\boldsymbol{x}(i,L)$. This is because both the channel and noise components are zero-mean complex Gaussian (Rayleigh magnitude and uniform phase between 0 to 2π) and a feedback-decision error only introduces a rotation to the phase of the channel coefficient and noise component of $y_{n,m}(l)$.

[2]Because received signals and decisions of symbols in previous symbol periods are used to predict the current channel state, the scheme derived is actually a channel predictor. Although the term "channel estimation" is usually used to broadly refer to the procedure from which the channel state is obtained through either prediction or estimation [4], it is more precise to describe the $\hat{h}_{n,m}(i)$ derived in this paper as a "predictor channel estimate," a term adopted in [5].

B. Detection

Given the received signal in (3), $\boldsymbol{s}(i)$ can be detected using several algorithms, such as the ML detection [15], [16], MMSE detection, ZF detection [17], [18], and the Bell Labs layered space–time architecture (BLAST) scheme [19], [20]. The ZF scheme has the lowest complexity and supports orthogonal-matrix triangularization (QR decomposition) implementation. Moreover, at high signal-to-noise ratios (SNR), performance of the ZF scheme approaches that of the MMSE scheme. For these reasons, the ZF scheme will be adopted in this paper for data detection.

In the ZF scheme, the decision vector for the M spatially multiplexed symbols in the ith interval is written as

$$\boldsymbol{\beta}(i) = \hat{\boldsymbol{\mathcal{H}}}^{+}(i)\boldsymbol{r}(i) \tag{15}$$

where $(\cdot)^{+}$ denotes the pseudoinverse. Because we consider an overdetermined system $(N \geq M)$, $\hat{\boldsymbol{\mathcal{H}}}^{+}(i)$, omitting symbol index i, can be calculated as $\hat{\boldsymbol{\mathcal{H}}}^{+} = [\hat{\boldsymbol{\mathcal{H}}}^{\mathrm{H}}\hat{\boldsymbol{\mathcal{H}}}]^{-1}\hat{\boldsymbol{\mathcal{H}}}^{\mathrm{H}}$. If channel estimates are perfect (i.e., $\hat{\boldsymbol{\mathcal{H}}} = \boldsymbol{\mathcal{H}}$), then $\boldsymbol{\beta}(i) = \sqrt{E_s}\boldsymbol{s}(i) + \boldsymbol{\xi}(i)$, where $\boldsymbol{\xi}(i) = \boldsymbol{\mathcal{H}}^{+}(i)\boldsymbol{\nu}(i)$ is the noise component after the ZF operation.

IV. Numerical Examples and Discussion

For all numerical examples, binary phase-shift keying (BPSK) with a data rate of $R_{\mathrm{b}} = 1$ Mb/s is chosen. Bit decisions for the BPSK system are obtained by slicing the real part of $\boldsymbol{\beta}(i)$ given in (15) as $\hat{s}(i) = \mathrm{sgn}\{\Re[\boldsymbol{\beta}(i)]\}$, where $\Re(\cdot)$ denotes the real part. The Doppler shift is calculated based on a center frequency $f_{\mathrm{c}} = 2.0$ GHz. Fading processes among all transmit and receive antennas are assumed independent and identically distributed; their first- and second-order statistics do not change over the entire transmission horizon. Although the proposed DD MAP scheme does not require periodic pilot bits in principle, errors introduced in applying (5) will accumulate over bits. Therefore, periodic pilot bits are added, but with a large block length of $K = 600$ bits, unless explicitly specified otherwise. Let P $(M \leq P \ll K)$ represent the number of pilot bits in each pilot period. For small values of P, the fading rates of interest are such that the channel remains approximately constant during one pilot period. The received signals in the pilot period are written in an $N \times P$ matrix as $\boldsymbol{Y}_p = [\boldsymbol{r}(1)\boldsymbol{r}(2)\cdots\boldsymbol{r}(P)] = \boldsymbol{\mathcal{H}}_p\boldsymbol{S}_p + \boldsymbol{V}_p$, where $\boldsymbol{r}(p)$ was given in (3), $\boldsymbol{\mathcal{H}}_p$ is the channel coefficient matrix in the pilot period, $\boldsymbol{S}_p = [\boldsymbol{s}(1)\boldsymbol{s}(2)\ldots\boldsymbol{s}(P)]$, and $\boldsymbol{V}_p = [\boldsymbol{\nu}(1)\boldsymbol{\nu}(2)\ldots\boldsymbol{\nu}(P)]$. Hence, channel estimates in the pilot period are obtained as $\hat{\boldsymbol{\mathcal{H}}}_p = \boldsymbol{Y}_p\boldsymbol{S}_p^{+}$.

Memory depth (window length) L for the DD MAP channel predictor affects the error performance. Fig. 2 shows the bit-error rate (BER) versus bit-energy-to-noise-density ratio (E_{b}/N_0) curves of a (2, 3) system with different values of L. The maximum Doppler shift is $f_{\mathrm{d}} = 74$ Hz ($f_{\mathrm{d}}T_{\mathrm{b}} = 7.4 \times 10^{-5}$), which is obtained based on a vehicular speed of $v = 40$ km/h. BER curves shown are for memory depths of $L = 3$, 5, 7, and 9. For comparison purposes, the error-rate curve with perfect channel estimates is also shown in the aforementioned figure. With the set of system parameters applied, error performance improves significantly when L increases from 3 to 7. However, when L increases to over 7, performance improvement is negligible. With $L = 9$ and other parameters adopted for a target BER of 10^{-3}, the proposed scheme performs approximately 2.5 dB worse than the case when all coefficients of the matrix channel are perfectly known to the receiver.

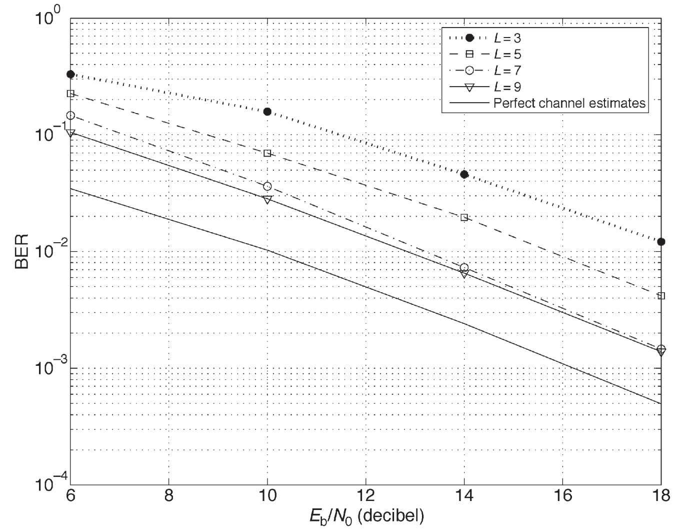

Fig. 2. BER versus E_{b}/N_0 with different memory depth L.

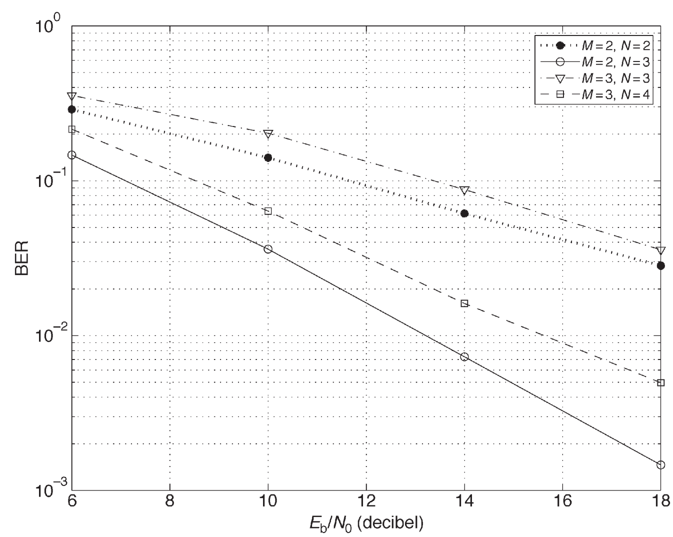

Fig. 3. BER versus E_{b}/N_0 for a different number of transmit and receive antennas.

Since channel estimation depends on the accuracy of inter-antenna interference cancellation using (5), it is expected that the performance will degrade when the number of transmit antennas increases. Fig. 3 shows the BER versus E_{b}/N_0 curves with $L = 7$ and $(N, M) = (2, 2)$, (2, 3), (3, 3), and (3, 4). Other parameters applied are the same as adopted for Fig. 2. Under ideal conditions, the ZF detection should yield the same error performance for cases of $(M, N) = (2, 2)$ and $(M, N) = (3, 3)$, if the received signal energy per symbol per antenna is the same [21]. With actual channel estimates, it is observed that a (3, 3) system performs worse than a (2, 2) system.

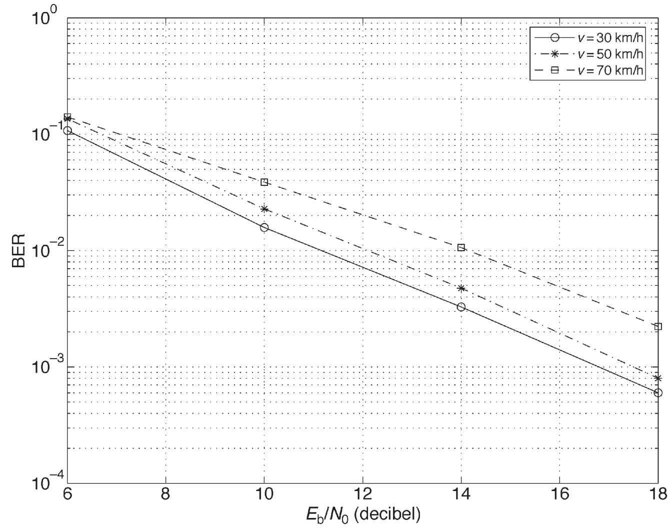

Fig. 4. BER versus $E_{\rm b}/N_0$ with different fading rates.

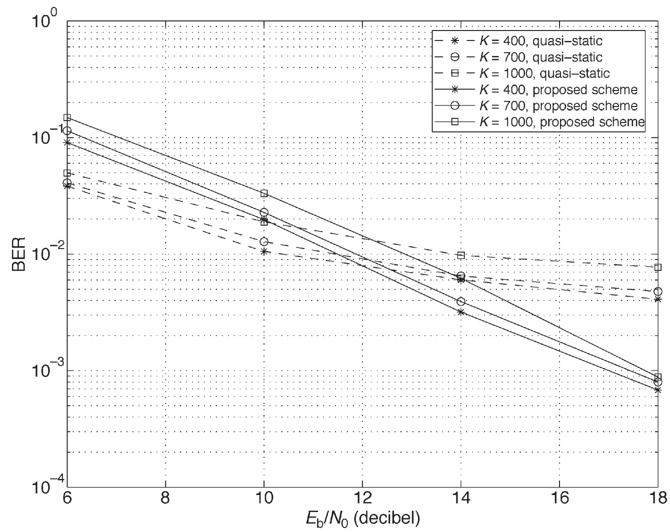

Fig. 5. Performance comparison of the proposed scheme and the scheme with the channel treated as block fading.

The fading rate quantified by the Doppler shift affects the performance of any channel-estimation schemes. If channel phase and magnitude remain constant over a number of bits, accurate channel estimates are possible. As the fading rate increases, estimation quality deteriorates. Fig. 4 shows BER versus $E_{\rm b}/N_0$ curves of a (2, 3) system with $L = 7$ and $v = 30$, 50, and 70 km/h. Performance degradation from $v = 30\ \rm km/h$ to $v = 50\ \rm km/h$ is considerably less significant than that from $v = 50\ \rm km/h$ to $v = 70\ \rm km/h$.

The rationale behind the quasi-static fading model is that the channel remains approximately constant over one block of data. If this were true, a simple method would be to apply the channel estimates obtained using pilot symbols embedded with data for data detection in the whole block. This scheme may not work well when the system is operating over a time-varying fading channel. Fig. 5 compares the performance of a (2, 3) system employing the proposed scheme with that of the scheme based on the quasi-static fading model described above. Fading rate is calculated based on $v = 50\ \rm km/h$ and block lengths of $K = 400$, 700, and 1000 are evaluated. The proposed scheme performs worse in the low $E_{\rm b}/N_0$ region (high BER values), but the scheme based on the quasi-static model reaches an error floor between 10^{-2} to 10^{-3} with the set of system parameters applied. The major factor causing this behavior of the proposed scheme in the low $E_{\rm b}/N_0$ region is that higher error rates result in worse estimates of $\boldsymbol{y}_{nm}(i, L)$.

V. Conclusion

A DD MAP channel-estimation scheme for symbol-by-symbol detection in MIMO systems has been derived. This scheme has low complexity and can be applied to time-varying Rayleigh fading channels with an arbitrary spaced-time correlation function. Numerical results indicate that a long memory depth is unnecessary for a system to work well. The channel-estimation quality deteriorates as the number of transmit antennas increases. The fading rate seems to have a high impact on system performance, and the proposed scheme is more appropriate for channels with low to medium normalized Doppler shifts. Large block length between adjacent pilot blocks can be deployed with the proposed scheme. This results in minimum overhead for pilot symbols. The scheme based on the quasi-static channel model may reach an error floor whereas the proposed scheme works very well at high $E_{\rm b}/N_0$ values.

Acknowledgment

The authors wish to thank the anonymous reviewers for their helpful comments.

References

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

[2] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

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

[4] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. New York: Wiley, 1998, ch. 14.

[5] A. Svensson, "1 and 2 stage decision feedback coherent detectors for DQPSK in fading channels," in *Proc. IEEE Vehicular Technology Conf. (VTC)*, Chicago, IL, Jul. 1995, pp. 644–648.

[6] B. Lindoff, C. Östberg, and H. Eriksson, "Channel estimation for the W-CDMA system, performance and robustness analysis from a terminal perspective," in *Proc. IEEE Vehicular Technology Conf. (VTC)*, Houston, TX, May 1999, vol. 2, pp. 1565–1569.

[7] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1697–1705, Sep. 2001.

[8] P. D. Alexander and A. J. Grant, "Iterative decoding and channel estimation," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, Jun. 2000, p. 171.

[9] H. Zhu, B. Farhang-Boroujeny, and C. Schlegel, "Pilot embedding for joint channel estimation and data detection in MIMO communication systems," *IEEE Commun. Lett.*, vol. 7, no. 1, pp. 30–32, Jan. 2003.

[10] D. Samardzija and N. Mandayam, "Pilot assisted estimation of MIMO fading channel response and achievable data rates," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2882–2890, Nov. 2003.

[11] Q. Sun, D. C. Cox, H. C. Huang, and A. Lozano, "Estimation of continuous flat fading MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 549–553, Oct. 2002.

[12] X. Deng, A. M. Haimovich, and J. Garcia-Frias, "Decision directed iterative channel estimation for MIMO systems," in *Proc. Int. Conf. Communications (ICC)*, Anchorage, AK, May 2003, pp. 2326–2329.

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

[14] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[15] X. Zhu and R. D. Murch, "Performance analysis of maximum likelihood detection in a MIMO antenna system," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 187–191, Feb. 2002.

[16] N. Prasad and M. K. Varanasi, "Decision feedback detection using modulation constraints for MIMO Rayleigh fading channels," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, Jun./Jul. 2002, p. 331.

[17] D. A. Gore, R. W. Heath, Jr., and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov. 2002.

[18] D. W. Waters and J. R. Barry, "Noise-predictive decision-feedback detection for multiple-input multiple-output channels," *IEEE Trans. Signal Processing*, vol. 53, no. 5, pp. 1852–1859, May 2005.

[19] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-blast space–time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[20] K. Liu and A. M. Sayeed, "Improved layered space–time processing in multiple antenna systems," in *Proc. 39th Annu. Allerton Conf. Communications, Control and Computers*, Monticello, IL, Oct. 2001, pp. 757–767.

[21] H. Liu, "Error performance of MIMO systems in frequency selective Rayleigh fading channels," in *Proc. IEEE Global Telecommunications (GLOBECOM)*, San Francisco, CA, Dec. 2003, pp. 2104–2108.