ML Detection with Blind Linear Prediction for Differential Space-Time Block Code Systems

Seree Wanichpakdeedecha, Kazuhiko Fukawa, Hiroshi Suzuki, Satoshi Suyama
Tokyo Institute of Technology
2-12-1, O-okayama, Meguro-ku, Tokyo, 152-8550, Japan
E-mail: {seree_w, fukawa, suzuki, ssuyama}@radio.ss.titech.ac.jp

***Abstract*— This paper proposes a maximum likelihood detection (MLD) method for the differential space-time block code (DSTBC) in cooperation with blind linear prediction (BLP) of fast frequency-flat fading channels. The method, which linearly predicts the fading complex envelope, determines the linear prediction coefficients by the method of Lagrange multipliers, and does not use decision-feedback nor require information on channel parameters such as the maximum Doppler frequency in contrast to conventional ones. Computer simulations under a fast fading condition demonstrate that the proposed method with a proper degree of polynomial approximation can be superior in BER performance to a conventional method that estimates the coefficients by the RLS algorithm using a training sequence.**

I. Introduction

Alamouti's space-time block code (STBC) has attracted much attention because it can achieve full spatial diversity without bandwidth expansion by using two transmit antennas [1]. However, STBC requires accurate channel estimation for coherent detection, although it can hardly perform such estimation under fast fading environments. In order to achieve the spatial diversity without channel state information (CSI), differential STBC (DSTBC) was proposed [2]. In a static channel, however, BER performance of DSTBC is approximately 3 dB worse than that of STBC with coherent detection. For improving the performance, some methods including multiple-symbol detection (MSD) were applied to the detection of DSTBC [3], [4]. Nevertheless, these schemes assume a time-invariant channel during each transmission block.

To cope with the fast fading, the methods of linear prediction were applied to DSTBC in [5]-[8]. The linear prediction scheme estimates a current complex envelope by using the past ones, and performs the maximum-likelihood sequences estimation (MLSE) to find the sequence that maximizes the log likelihood function. However, this conventional scheme needs information on both the maximum Doppler frequency and signal-to-noise ratio (SNR) in order to obtain the optimum prediction coefficients. Another conventional scheme estimates the coefficients by the RLS algorithm using decision-feedback, which may cause error propagation [9], [10].

For solving these problems, this paper proposes blind linear prediction (BLP) for DSTBC that employs the method of Lagrange multipliers to determine the prediction coefficients by using neither the data of decision-feedback nor information on the channel parameters.

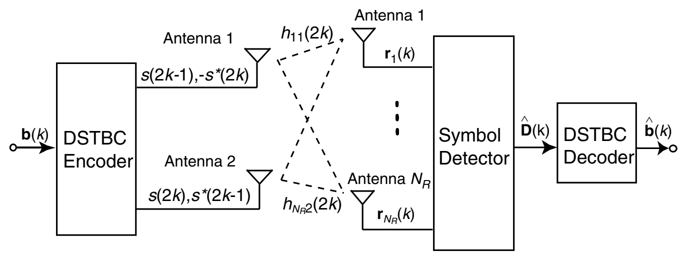


Fig. 1 The system model

II. System Model

A. Received signal

Fig. 1 shows the system model of DSTBC with two transmit antennas and N_R receive antennas. Let T denote the symbol duration. The channel is assumed to be time-varying frequency-flat Rayleigh-fading, where $h_{lp}(i)$ is a channel impulse response between the p-th transmit antenna (p = 1, 2) and the l-th receive antenna ($l = 1, ..., N_R$) at discrete time iT. The receiver for DSTBC [6] assumes that $h_{lp}(i)$ is constant during each STBC block (two symbols) as

$$h_{lp}(2k) = h_{lp}(2k-1), \tag{1}$$

where k is a block index. Note that this assumption is used below except the computer simulation section. It is also noteworthy that it is not satisfied under very fast fading conditions, which causes the receiver considerable intrablock interference [8].

$r_l(i)$ represents the signal received by the l-th receive antenna at discrete time iT, and a 2-by-1 received signal vector $\mathbf{r}_l(k)$ is defined as

$$\mathbf{r}_l^{\mathrm{H}}(k) = \big(r_l^*(2k-1) \quad r_l^*(2k)\big), \tag{2}$$

where the superscript $^{\mathrm{H}}$ and * denote Hermitian transposition and complex conjugation, respectively. On the assumption of (1), $\mathbf{r}_l(k)$ can be written as

$$\mathbf{r}_l(k) = \mathbf{D}(k)\mathbf{h}_l(k) + \mathbf{n}_l(k), \tag{3}$$

where the 2-by-1 impulse response vector $\mathbf{h}_l(k)$ and the 2-by-1 noise vector $\mathbf{n}_l(k)$ are given by

$$\mathbf{h}_l^{\mathrm{H}}(k) = \big(h_{l1}^*(2k) \quad h_{l2}^*(2k)\big), \tag{4}$$
$$\mathbf{n}_l^{\mathrm{H}}(k) = \big(n_l^*(2k-1) \quad n_l^*(2k)\big). \tag{5}$$

Here, $n_l(i)$ is an additive white Gaussian noise at discrete time iT with variance σ_n^2 and zero mean, and is statistically independent with respect to receive antenna index l.

The 2-by-2 STBC symbol matrix $\mathbf{D}(k)$ is defined as

$$\mathbf{D}^{\mathrm{H}}(k) = \big(\mathbf{s}_1^*(k) \;\; \mathbf{s}_2^*(k)\big), \tag{6}$$

where the 2-by-1 transmitted symbol vectors $\mathbf{s}_1(k)$ and $\mathbf{s}_2(k)$ are given by

$$\mathbf{s}_1^{\mathrm{H}}(k) = \big(s^*(2k{-}1) \;\; s^*(2k)\big), \tag{7}$$
$$\mathbf{s}_2^{\mathrm{H}}(k) = \big(-s(2k) \;\; s(2k{-}1)\big). \tag{8}$$

Here, $s(i)$ represents a complex symbol of 2^{N_d}-PSK modulation. $\mathbf{s}_1(k)$ and $\mathbf{s}_2(k)$, which are orthogonal to each other and have unit lengths, correspond to transmitted signals at discrete time $(2k-1)T$ and $2kT$, respectively. Note that the power of each transmitted symbol is fixed to $1/2$ in order to keep the transmit power per symbol constant ($= 1$).

B. DSTBC encoder and decoder

The DSTBC encoder in **Fig. 1** performs differential encoding with the STBC format, which can be expressed as [2]

$$\begin{aligned}\mathbf{s}_1^{\mathrm{T}}(k) &= A(k)\mathbf{s}_1^{\mathrm{T}}(k-1) + B(k)\mathbf{s}_2^{\mathrm{T}}(k-1)\\ &= \big(A(k) \;\; B(k)\big)\mathbf{D}(k-1),\end{aligned} \tag{9}$$

where the superscript $^{\mathrm{T}}$ denotes transposition. The complex vector $\big(A(k) \;\; B(k)\big)$ depends on $2N_d$ information bits and satisfies the constraint that the elements of $\mathbf{s}_1(k)$ should be a 2^{N_d}-PSK constellation symbol. Note that $\mathbf{s}_1(0)$ can be arbitrarily set, e.g. to $\big(1/\sqrt{2} \;\; 1/\sqrt{2}\big)$.

The DSTBC decoder in **Fig. 1** detects $\big(A(k) \;\; B(k)\big)$ from $\hat{\mathbf{D}}(k-1)$ and $\hat{\mathbf{D}}(k)$ which the symbol detector provides as estimated values of $\mathbf{D}(k{-}1)$ and $\mathbf{D}(k)$. The detection can be expressed as

$$(\hat{A}(k) \;\; \hat{B}(k)) = \hat{\mathbf{s}}_1^{\mathrm{T}}(k)\hat{\mathbf{D}}^{\mathrm{H}}(k-1), \tag{10}$$

where $\hat{A}(k)$, $\hat{B}(k)$, and $\hat{\mathbf{s}}_1(k)$ are estimated values of $A(k)$, $B(k)$, and $\mathbf{s}_1(k)$, respectively. The derivation of (10) uses the property that $\mathbf{D}(k)$ is an unitary matrix. From $(\hat{A}(k) \;\; \hat{B}(k))$, the $2N_d$ transmitted information bits can be determined.

III. Channel Model

A. Multipath propagation

Under the multipath propagation, $h_{lp}(i)$ can be written as

$$h_{lp}(i) = \sum_d a_{lp,d}e^{j(2\pi f_D T\cos\theta_{lp,d})i}, \tag{11}$$

where $a_{lp,d}$ and $\theta_{lp,d}$ represent a complex envelope and incident angle of the d-th propagation path, respectively. f_D is the maximum Doppler frequency and f_DT is generally much less than 1.

B. Autoregressive (AR) process

On the assumption that $h_{lp}(i)$ of (11) includes M dominant paths, $h_{lp}(i)$ has M dominant tones and then can be modeled as an AR process of order M. Considering this model and the assumption of (1), $h_{lp}(2k)$ is expressed as

$$h_{lp}(2k) = \sum_{m=1}^{M} c_{lp,m}h_{lp}[2(k-m)] + v_{lp}(2k), \tag{12}$$

where $c_{lp,m}$ is equal to an AR parameter multiplied by -1, and $v_{lp}(2k)$ that represents residual propagation paths is assumed to be a negligible white-noise process.

Next, let us consider Taylor series expansion of $h_{lp}[2(k-m)]$ that is given by

$$h_{lp}[2(k-m)] = \sum_{q=0}^{\infty}\frac{(-2m)^q}{q!}h_{lp}^{(q)}(2k), \tag{13}$$

where $h_{lp}^{(q)}(2k)$ is the q-th derivative of $h_{lp}(2k)$ and is derived from (11) as

$$h_{lp}^{(q)}(2k) = (f_DT)^q\Big[\sum_d a_{lp,d}(j2\pi\cos\theta_{lp,d})^q e^{j4\pi f_D T k\cos\theta_{lp,d}}\Big]. \tag{14}$$

Evidently, $h_{lp}^{(q)}(2k)$ is proportional to $(f_DT)^q$ and exponentially decays as q increases. Let us assume that $h_{lp}^{(q)}(2k)$ with $q > q_1$ can be neglected. The integer q_1 is referred to as a degree of polynomial approximation.

Substituting (13) into (12) and neglecting $h_{lp}^{(q)}(2k)$ with $q > q_1$ and the white noise $v_{lp}(2k)$ yield

$$h_{lp}(2k) \approx \sum_{q=0}^{q_1}\Big[\sum_{m=1}^{M} c_{lp,m}(-2m)^q(q!)^{-1}\Big]h_{lp}^{(q)}(2k). \tag{15}$$

The conditions for any $h_{lp}^{(q)}(2k), 0 \le q \le q_1$, to satisfy (15) are given by

$$\sum_{m=1}^{M} c_{lp,m} = 1, \tag{16}$$

$$\sum_{m=1}^{M} c_{lp,m}(-2m)^q = 0, \;\; q_1 \ge 1, \;\; 1 \le q \le q_1. \tag{17}$$

Since (16) and (17) are independent of indices l and p, we assume that $c_{lp,m} = c_m$ for all l and p. Accordingly, (12), (16), and (17) can be rewritten in simple vector forms as

$$\mathbf{h}_l(k) = \sum_{m=1}^{M} c_m\mathbf{h}_l(k-m) + \mathbf{v}_l(k), \tag{18}$$
$$\mathbf{c}^{\mathrm{H}}\mathbf{b}_0 = 1, \tag{19}$$
$$\mathbf{c}^{\mathrm{H}}\mathbf{b}_q = 0, \;\; q_1 \ge 1, \;\; 1 \le q \le q_1, \tag{20}$$

where the M-by-1 prediction coefficient vector $\mathbf{c}$ and the M-by-1 vector $\mathbf{b}_q$ are defined as

$$\mathbf{c}^{\mathrm{H}} = \big(c_1 \;\; c_2 \;\; \cdots \;\; c_M\big), \tag{21}$$
$$\mathbf{b}_q^{\mathrm{H}} = \big((-2)^q \;\; (-4)^q \;\; \cdots \;\; (-2M)^q\big), \;\; 0 \le q \le q_1. \tag{22}$$

$\mathbf{v}_l(k)$ is the 2-by-1 process noise vector having $v_{lp}(2k)$ as its elements.

IV. Symbol Detection

A. MLSE

The symbol detector in **Fig. 1** performs MLSE on the basis of the linear prediction. The log likelihood function is derived from (3) as follows:

First, multiplying the both hand sides of (3) by $\mathbf{D}^{-1}(k)$ results in

$$\mathbf{D}^{-1}(k)\mathbf{r}_l(k) = \mathbf{h}_l(k) + \mathbf{D}^{-1}(k)\mathbf{n}_l(k). \tag{23}$$

Since $\mathbf{D}(k)$ is an unitary matrix, this can be rewritten as

$$\mathbf{D}^{\mathrm{H}}(k)\mathbf{r}_l(k) = \mathbf{h}_l(k) + \tilde{\mathbf{n}}_l(k), \tag{24}$$

where the 2-by-1 vector $\tilde{\mathbf{n}}_l(k)$ is defined as

$$\tilde{\mathbf{n}}_l(k) = \mathbf{D}^{\mathrm{H}}(k)\mathbf{n}_l(k). \tag{25}$$

The autocorrelation matrix of $\tilde{\mathbf{n}}_l(k)$ is given by

$$\left\langle \tilde{\mathbf{n}}_l(k)\tilde{\mathbf{n}}_l^{\mathrm{H}}(k) \right\rangle = \left\langle \mathbf{n}_l(k)\mathbf{n}_l^{\mathrm{H}}(k) \right\rangle = \sigma_n^2 \mathbf{I}, \tag{26}$$

where $\langle\cdot\rangle$ denotes ensemble average and $\mathbf{I}$ is the 2-by-2 identity matrix. It is also evident from (25) that $\tilde{\mathbf{n}}_l(k)$ is statistically independent with respect to both time index k and receive antenna index l. Therefore, the log likelihood function can be expressed as [6]

$$L = \sum_{l=1}^{N_R} l(\mathbf{r}_l|\mathbf{s}) = -\sum_{l=1}^{N_R}\sum_{k} \left\| \mathbf{D}_n^{\mathrm{H}}(k)\mathbf{r}_l(k) - \hat{\mathbf{h}}_l(k) \right\|^2, \tag{27}$$

where $\mathbf{D}_n(k)$ is a candidate of $\mathbf{D}(k)$ and $\hat{\mathbf{h}}_l(k)$ is an estimate of the impulse response vector $\mathbf{h}_l(k)$.

B. Blind linear prediction (BLP)

1) Applying linear prediction to MLSE: Let us discuss the estimation of $\hat{\mathbf{h}}_l(k)$ by linear prediction. Considering (18) and (24), $\hat{\mathbf{h}}_l(k)$ can be approximated as

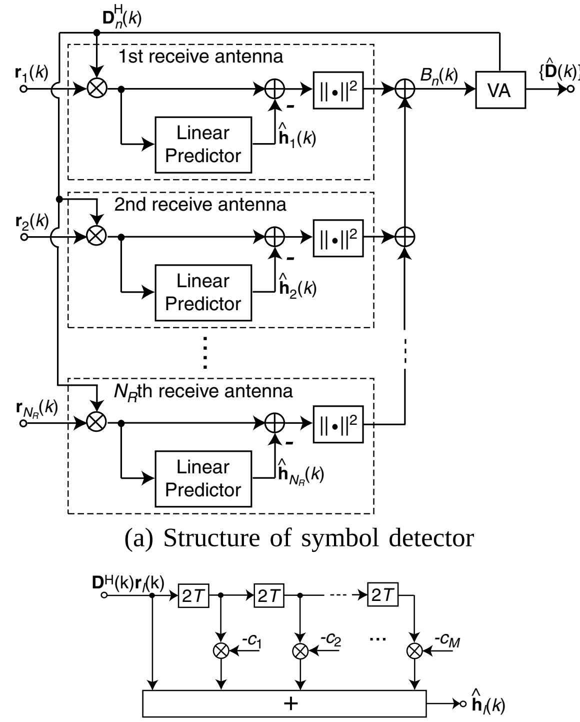
$$\hat{\mathbf{h}}_l(k) = \sum_{m=1}^{M} c_m \mathbf{D}_n^{\mathrm{H}}(k-m)\mathbf{r}_l(k-m). \tag{28}$$

Substituting (28) into (27) yields the branch metric given by

$$B_n(k) = -\sum_{l=1}^{N_R} \left\| \mathbf{D}_n^{\mathrm{H}}(k)\mathbf{r}_l(k) - \sum_{m=1}^{M} c_m \mathbf{D}_n^{\mathrm{H}}(k-m)\mathbf{r}_l(k-m) \right\|^2. \tag{29}$$

A block diagram of the symbol detector using this branch metric is shown in **Fig. 2** (a). For all l, $\mathbf{r}_l(k)$ is multiplied by $\mathbf{D}_n^{\mathrm{H}}(k)$ which the Viterbi algorithm (VA) processor provides. The resultant is fed into the linear predictor that generates $\hat{\mathbf{h}}_l(k)$ following (28) and of which structure is depicted in **Fig. 2** (b). Squaring $\mathbf{D}_n^{\mathrm{H}}(k)\mathbf{r}_l(k) - \hat{\mathbf{h}}_l(k)$ and combining the resultants with respect to l yield the branch metric, which is passed into the VA processor.

The VA processor searches the sequence of $\mathbf{D}_n(k)$ that maximizes the log likelihood function of (27) by VA, and



(a) Structure of symbol detector

(b) Structure of linear predictor

Fig. 2 Symbol detector

outputs the resultant as the detected signal $\hat{\mathbf{D}}(k)$. VA considers sequences of $\mathbf{D}_n(k)$ as those of states, and can effectively find the maximum likelihood sequence on a trellis diagram. In this case, the state at discrete time $2kT$ is expressed as $\{\mathbf{s}_{1n}(k-1), \mathbf{s}_{1n}(k-2), ..., \mathbf{s}_{1n}(k-M)\}$ where $\mathbf{s}_{1n}(k)$ is a candidate of $\mathbf{s}_1(k)$. Therefore, the numbers of states and state transitions are equal to 2^{2MN_d} and 2^{2N_d}, respectively. **Fig. 3** shows the trellis diagram for BPSK ($N_d = 1$) with $M = 2$.

2) Prediction coefficients: Conventional methods determine the prediction coefficients $\mathbf{c}$ by solving Yule-Walker equation [7] or using the RLS algorithm [9] which require information on $f_D T$ and SNR, or detected symbols. The proposed blind method can determine the coefficients without detected and training symbols and any channel information as follows:

A 2-by-1 prediction error vector $\mathbf{e}_l(k)$ is defined as

$$\begin{aligned}\mathbf{e}_l(k) &= \mathbf{D}^{\mathrm{H}}(k)\mathbf{r}_l(k) - \sum_{m=1}^{M} c_m \mathbf{D}^{\mathrm{H}}(k-m)\mathbf{r}_l(k-m) \\ &= \left[\mathbf{h}_l(k) - \sum_{m=1}^{M} c_m \mathbf{h}_l(k-m)\right] + \left[\tilde{\mathbf{n}}_l(k) - \sum_{m=1}^{M} c_m \tilde{\mathbf{n}}_l(k-m)\right].\end{aligned} \tag{30}$$

When $\mathbf{D}_n(k) = \mathbf{D}(k)$ for all k, the branch metric of (29) is given by

$$B_n(k) = -\sum_{l=1}^{N_R} \left\| \mathbf{e}_l(k) \right\|^2. \tag{31}$$

Since $\mathbf{e}_l(k)$ can be considered a virtual noise, minimizing $\left\langle \left\| \mathbf{e}_l(k) \right\|^2 \right\rangle$ results in the optimal BER performance. To de-

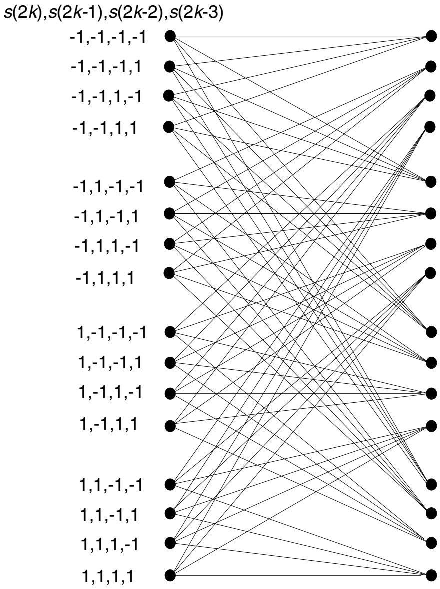


Fig. 3 Trellis diagram for BPSK with the prediction order of 2 ($M = 2$)

crease $\left\langle \left\| \mathbf{e}_l(k) \right\|^2 \right\rangle$ as much as possible, the proposed method forces the first term of (30) to be zero, and then minimizes the mean squared norm of the second term. This is because the first term can be zero due to the property that $\mathbf{h}_l(k)$ is a deterministic process, whereas the second term cannot vanish due to the property that $\tilde{\mathbf{n}}_l(k)$ is a stochastic process.

After forcing the first term to be zero, $\left\langle \left\| \mathbf{e}_l(k) \right\|^2 \right\rangle$ is expressed as

$$\left\langle \left\| \mathbf{e}_l(k) \right\|^2 \right\rangle = 2\sigma_n^2 \Big(1 + \sum_{m=1}^{M} \left| c_m \right|^2 \Big). \qquad (32)$$

Minimizing (32) is equivalent to

$$\sum_{m=1}^{M} \left| c_m \right|^2 = \mathbf{c}^{\mathrm{H}} \mathbf{c} \rightarrow min. \qquad (33)$$

The conditions for the first term to vanish are given by (19) and (20). Thus, the above minimization is equivalent to minimizing (33) under the constraints of (19) and (20). By applying the method of Lagrange multipliers, this problem becomes equivalent to obtaining $\mathbf{c}$ that minimizes the following cost function $f(\mathbf{c})$:

$$f(\mathbf{c}) = \begin{cases} \mathbf{c}^{\mathrm{H}}\mathbf{c} + \lambda_0(1 - \mathbf{c}^{\mathrm{H}}\mathbf{b}_0) + \lambda_0^*(1 - \mathbf{b}_0^{\mathrm{H}}\mathbf{c}), \\ \qquad \text{for} \;\; q_1 = 0 \\ \mathbf{c}^{\mathrm{H}}\mathbf{c} + \lambda_0(1 - \mathbf{c}^{\mathrm{H}}\mathbf{b}_0) + \lambda_0^*(1 - \mathbf{b}_0^{\mathrm{H}}\mathbf{c}) \\ - \sum_{q=1}^{q_1} \lambda_q \mathbf{c}^{\mathrm{H}}\mathbf{b}_q - \sum_{q=1}^{q_1} \lambda_q^* \mathbf{b}_q^{\mathrm{H}}\mathbf{c}, \;\; \text{for} \;\; q_1 \geq 1 \end{cases} \qquad (34)$$

where $\{\lambda_q\}$ is a set of Lagrange multipliers. By taking the derivatives of $f(\mathbf{c})$ with respect to $\mathbf{c}$ and $\{\lambda_q\}$, and equating its results to zero, the optimal $\mathbf{c}$ is derived as

$$\mathbf{c} = \mathbf{c}^* = \sum_{q=0}^{q_1} (\mathbf{B}_{q_1}^{-1})_{q+1,1} \mathbf{b}_q, \qquad (35)$$

TABLE I Simulation condition

Packet format	128 symbols
Constellation mapping	BPSK, QPSK
Receive antennas	1 branch
Channel	Fast flat Rayleigh fading
$f_D T$	$1{\times}10^{-3}$– $1{\times}10^{-1}$

TABLE II Proper value of the degree of the polynomial approximation (q_1) corresponding to prediction order (M)

M	q_1
2	1
3	1
4	1, 2

where $(\mathbf{B}_{q_1})_{iu} = \mathbf{b}_{i-1}^{\mathrm{H}} \mathbf{b}_{u-1}, \; i, u = 1, ..., (q_1 + 1)$.

Note that the optimal $\mathbf{c}$ does not depend on any channel characteristics and varies according to q_1. It is also noteworthy that the solution with $M = 2$ and $q_1 = 1$ is equivalent to linear prediction coefficients for SISO in [11].

V. Computer Simulations

A. Simulation condition

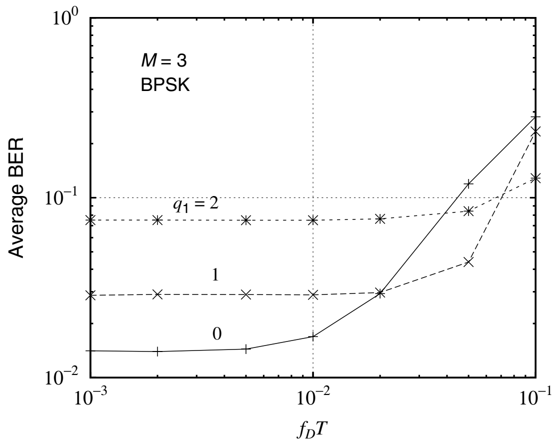
A series of computer simulations was conducted to evaluate the performance of BLP. The simulation condition is summarized in **Table I**. For comparison, the conventional methods, the solution of Yule-Walker equation using information on SNR and $f_D T$, and the RLS algorithm using N_{TS} training symbols instead of detected ones, were also evaluated. Note that the RLS algorithm using only detected symbols [9], [10] is obviously inferior to that using the training in BER performance, owing to the error propagation.

B. Average BER of BLP

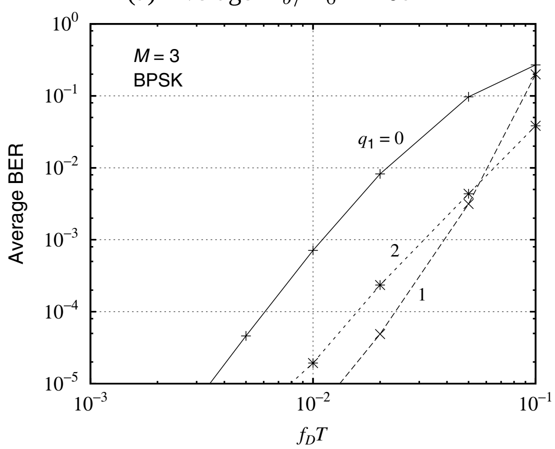
Fig. 4 shows the effects of $f_D T$ on BER performance of BLP with BPSK and $M = 3$. It can be seen that increasing q_1 can improve the tracking performance whereas it degrades the BER performance under low SNR. This is because larger q_1 makes the approximation of (15) more accurate but increases the number of constraints of (20), which enhances the noise power of (32) and causes the degradation in low SNR region. Note that the BER performance of BLP with $q_1 = 2$ at high SNR degrades due to the intrablock interference. Considering this trade-off, the proper values of q_1 from $f_D T = 10^{-3}$ to $5{\times}10^{-2}$ are listed in **Table II**.

C. Comparison with other methods

Figs. 5 (a) and (b) show average BER performance versus average E_b/N_0 of BLP and the conventional methods with $f_D T = 0.02$ for BPSK and QPSK, respectively. The result with the solution of Yule-Walker equation can be considered as a lower bound of the other methods. It can be seen that BLP is superior to the RLS method using the training in BER performance when average E_b/N_0 is greater than 20 dB and that the BER of BLP is very close to the lower bound. It is also found that the BER with QPSK is worse than that with

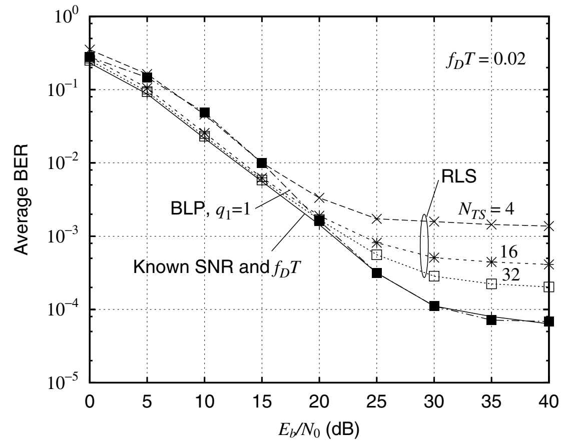


(a) Average $E_b/N_0 = 10$dB

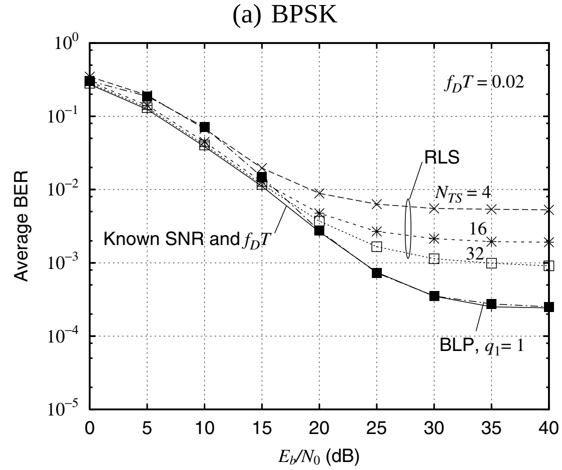


(b) Average $E_b/N_0 = 40$dB

Fig. 4 Average BER of blind linear prediction (BLP) versus f_DT with the polynomial approximation degree q_1 as a parameter



(a) BPSK



(b) QPSK

Fig. 5 Average BER versus average E_b/N_0 of proposed and conventional methods with the prediction order $M = 2$

BPSK. The reason is that the fading fluctuation often causes two bits error per symbol with QPSK while it causes at most one bit error per symbol with BPSK.

VI. Conclusion

ML detection with BLP for DSTBC on a fast fading channel has been proposed. BLP can determine the linear prediction coefficients without neither decision-feedback nor information on the channel parameters in contrast to the conventional ones. Computer simulation has shown that the BER performance of BLP with a proper degree of the polynomial approximation is superior to that of the conventional RLS method using the training method.

References

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

[2] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 18, no. 7, pp. 1169-1174, July 2000.

[3] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300-308, Mar. 1990.

[4] C. Gao and A. M. Haimovich, "Multiple-symbol differential detection for space-time block codes," in *Proc. Conf. Information Sciences and Systems*, Princeton, Mar. 2002.

[5] F. Adachi, "Adaptive differential detection using linear prediction for M-ary DPSK," *IEEE Trans. Veh. Technol.*, vol. 47, no. 3, pp. 909-918, Aug. 1998.

[6] E. Chiavaccini and G. M. Vitetta, "Further results on Tarokh's space-time differential technique," in *Proc. IEEE ICC'02*, vol. 3, pp. 1778-1782, Apr. 2002.

[7] P. Tarasak, H. Minn, and V. K. Bhargava, "Linear prediction receiver for differential space-time block codes with spatial correlation," *IEEE Commun. Lett.*, vol. 7, no. 11, pp. 543-545, Nov. 2003.

[8] P. Tarasak, H. Minn, and V. K. Bhargava, "Improved approximate maximum-likelihood receiver for differential space-time block codes over Rayleigh-fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 461-468, Mar. 2004.

[9] R. Schober and W. H. Gerstacker, "Decision-feedback differential detection based on linear prediction for MDPSK signals transmitted over Ricean fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, no. 3, pp. 391-402, Mar. 2000.

[10] R. Schober and Lutz H.-J. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 768-777, May 2002.

[11] K. Fukawa and H. Suzuki, "Recursive least squares adaptive algorithm maximum likelihood sequence estimation: RLS-MLSE - an application of maximum likelihood estimation theory to mobile radio -," *IEICE Transaction B-II*, vol. J76-B-II, no. 4, pp. 202-214, Apr. 1993.