Proceedings of the 6th Nordic Signal Processing Symposium - NORSIG 2004
June 9 - 11, 2004
Espoo, Finland

Receiver Concepts for Differential Space-Time Modulation Schemes over Flat Time-Varying Channels

Xiao-Ming Chen and Peter A. Hoeher
Information and Coding Theory Lab, University of Kiel
Kaiserstr. 2, D-24143 Kiel, GERMANY, E-mail: {xc,ph}@tf.uni-kiel.de

Abstract*— This paper addresses receiver structures for unitary differential space-time modulation schemes, where channel coefficients are modelled as complex Gaussian-distributed random variables and where only second-order statistics about channel coefficients are available at the receiver. Specifically, we derive the maximum-likelihood (ML) receiver for flat time-varying channels. Since the complexity of the ML receiver increases exponentially with the block length, reduced-complexity receivers are also employed to obtain a compromise between performance and complexity. The conventional differential detector and the subblock-by-subblock detector can be interpreted as special cases of the proposed sequence estimator. Moreover, the performance of the sequence estimator under channel mismatch is investigated.

I. Introduction

For a reliable high-rate signal transmission, multiple-input multiple-output (MIMO) systems attract increasing attention in conjunction with space-time processing techniques. To facilitate the requirement of channel estimation, differential space-time modulation (DSTM) schemes have been proposed for flat fading channels [1], [2], [3], where data symbols can be recovered without knowing the channel coefficients. A differential detector has been proposed in [1], [2] taking two adjacent data symbol matrices into account. The channel coefficients are assumed to be constant for the considered time interval. For moderately or fast time-varying channels, so-called noncoherent receivers have been investigated in [4] for the diagonal DSTM proposed in [2]; the focus is on a multiple-symbol detector and a simplified version thereof.

In this paper, we investigate the block-wise transmission with a long block length in conjunction with DSTM schemes. We derive the maximum-likelihood (ML) receiver for unitary DSTM schemes over time-varying channels, where channel coefficients are modelled as complex zero-mean Gaussian-distributed random variables. Since the complexity of the ML receiver inhibits its application for practical systems with a long block length, reduced-complexity receivers are derived by putting additional constraints on the ML receiver. It is shown that previous works [1], [4], [5] may be interpreted as special cases of the derived ML receiver. Although the actual channel coefficients are not necessary for the proposed receivers, the noise variance and the normalized fade rate are needed for the data estimation, which motivates the investigation of the performance of the proposed receivers under the channel mismatch.

II. System Model

Throughout this paper we use the complex baseband notation.

A. Differential Space-Time Modulator

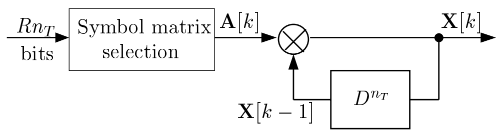

Fig. 1. Transmitter structure of DSTM schemes

Fig. 1 illustrates the transmitter structure of DSTM schemes, where R, n_T, $\mathbf{A}[k]$, and $\mathbf{X}[k]$ denote the transmission rate, the number of transmit antennas, the data symbol matrices before and after the differential space-time encoding, respectively. Note that a suitable Gray mapping is inherent in the selection of symbol matrices, following the design criteria for space-time codes [1], [6].

While DPSK modulation in single antenna systems is performed on a scalar basis, DSTM is carried out on a matrix basis:

$$\begin{aligned}\mathbf{X}[k] &= [\mathbf{x}[kn_T], \cdots, \mathbf{x}[(k+1)n_T-1]] \\ &= \mathbf{X}[k-1]\mathbf{A}[k] \in \mathcal{C}^{n_T \times n_T}, \end{aligned} \qquad (1)$$

where $\mathbf{x}[\kappa]$ denotes the data vector transmitted over n_T antennas at time index κ. Within this paper, focus is on unitary data symbol matrices, i.e., $\mathbf{A}[k]\mathbf{A}^H[k] = \mathbf{I}_{n_T}$ and $\mathbf{X}[k]\mathbf{X}^H[k] = \mathbf{I}_{n_T}$.

B. Time-Varying MIMO Channel Model

The time-varying flat MIMO channel model can be represented as follows:

$$\mathbf{Y}[k] = [\mathbf{H}[kn_T], \cdots, \mathbf{H}[(k+1)n_T-1]]\,\bar{\mathbf{X}}[k] + \mathbf{N}[k], \qquad (2)$$

where

$$\bar{\mathbf{X}}[k] = \mathrm{diag}\left\{\mathbf{x}[kn_T], \cdots, \mathbf{x}[(k+1)n_T-1]\right\} \in \mathcal{C}^{n_T^2 \times n_T},$$

and $\mathbf{H}[\kappa] \in \mathcal{C}^{n_R \times n_T}$ denotes the channel matrix at time index κ. Note that n_R denotes the number of receive antennas. Assuming a block transmission with N data symbol matrices pro block, the channel model can compactly be written as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{N} \qquad \in \mathcal{C}^{n_R \times N n_T}, \qquad (3)$$

where $\mathbf{Y} = [\mathbf{Y}[0], \cdots \mathbf{Y}[N-1]]$, and $\mathbf{H} \in \mathcal{C}^{n_R \times N n_T^2}$, $\mathbf{X} \in \mathcal{C}^{N n_T^2 \times N n_T}$ and $\mathbf{N}$ are suitably defined according to (2). The temporal/spatial correlation function of channel coefficients is assumed as

$$\mathrm{E}\left\{h^{i,j}[\kappa+\kappa'] \left(h^{p,q}[\kappa]\right)^*\right\} = \delta[i-p]\delta[j-q]\phi_h[\kappa'], \quad (4)$$

where $0 \le i,p \le n_R - 1,\ 0 \le j,q \le n_T - 1$, $\delta[\cdot]$ denotes the discrete-time delta function, and $h^{i,j}[\kappa]$ is the channel coefficient between the i-th receive antenna and the j-th transmit antenna at time index κ. Notice that (4) implies that channel coefficients have the identical time-correlation property and that there is no spatial correlation. Assuming 2-D isotropic scattering, the autocorrelation function of complex Gaussian-distributed channel coefficients reads

$$\phi_h[\kappa'] = \mathrm{E}\left\{h^{i,j}[\kappa+\kappa'] \left(h^{i,j}[\kappa]\right)^*\right\} = J_0\left(2\pi f_D T \kappa'\right), \quad (5)$$

where $J_0(\cdot)$ denotes the zeroth order Bessel function of first kind and $f_D T$ denotes the normalized fade rate. Moreover, the average power of channel coefficients is normalized to $\mathrm{E}\left\{|h^{i,j}[\kappa]|^2\right\} = 1$.

Based on the channel model (3), the likelihood function of the observation matrix $\mathbf{Y}$ can be described as [7]

$$p(\mathbf{Y} \mid \mathbf{X}, \mathbf{\Phi_H}) = \frac{1}{\left(\pi^{N n_T} \det\left(\frac{1}{n_R}\mathbf{C_Y}\right)\right)^{n_R}} \times \exp\left(\mathrm{Tr}\left\{-\mathbf{Y}\left(\frac{1}{n_R}\mathbf{C_Y}\right)^{-1}\mathbf{Y}^H\right\}\right). \quad (6)$$

The covariance matrix of channel outputs is defined as

$$\mathbf{C_Y} \triangleq \mathrm{E}\left\{\mathbf{Y}^H\mathbf{Y} | \mathbf{X}, \phi_h[\kappa]\right\} = n_R \mathbf{X}^H \mathbf{\Phi_H} \mathbf{X} + n_R \sigma_n^2 \mathbf{I}_{N n_T}, \quad (7)$$

where σ_n^2 denotes the noise variance of additive white Gaussian noise process at each receive antenna. Denoting $\Phi_h[\kappa] = \mathrm{diag}\left\{\phi_h[\kappa]\right\} \in \mathcal{C}^{n_T \times n_T}$, the channel correlation matrix can be evaluated as follows:

$$\begin{aligned}\mathbf{\Phi_H} &\triangleq \frac{1}{n_R}\mathrm{E}\left\{\mathbf{H}^H\mathbf{H}\right\} = \begin{bmatrix} \Phi_h[0] & \cdots & \Phi_h[N n_T - 1] \\ \vdots & \ddots & \vdots \\ \Phi_h[N n_T - 1] & \cdots & \Phi_h[0] \end{bmatrix} \\ &= \mathbf{\Phi}_h \otimes \mathbf{I}_{n_T} \in \mathcal{C}^{N n_T^2 \times N n_T^2}, \end{aligned} \quad (8)$$

where $\otimes$ denotes the Kronecker product and

$$\mathbf{\Phi}_h = \begin{bmatrix} \phi_h[0] & \cdots & \phi_h[N n_T - 1] \\ \vdots & \ddots & \vdots \\ \phi_h[N n_T - 1] & \cdots & \phi_h[0] \end{bmatrix} \in \mathcal{C}^{N n_T \times N n_T}.$$

C. Maximum-Likelihood Receiver for Unitary DSTM

According to the likelihood function (6), the maximum-likelihood sequence estimation can be written as

$$\widehat{\mathbf{X}} = \arg\min_{\widetilde{\mathbf{X}}} \left\{\mathrm{Tr}\left\{\mathbf{Y}\widetilde{\mathbf{C}}_\mathbf{Y}^{-1}\mathbf{Y}^H\right\} + \log\det\left(\widetilde{\mathbf{C}}_\mathbf{Y}\right)\right\}, \quad (9)$$

where the conditional covariance $\widetilde{\mathbf{C}}_\mathbf{Y} = \mathrm{E}\left\{\mathbf{Y}^H\mathbf{Y} \mid \widetilde{\mathbf{X}},\ \mathbf{\Phi_H}\right\}$, cf. (7), is essential for the sequence estimation.

For unitary data symbol matrices, $\tilde{\mathbf{x}}^H[k n_T + \kappa_1]\tilde{\mathbf{x}}[k n_T + \kappa_2] = \delta[\kappa_1 - \kappa_2]$. Let us define the signal containing term in $\widetilde{\mathbf{C}}_\mathbf{Y}$ as $\widetilde{\mathbf{C}}_\mathbf{X} \triangleq \widetilde{\mathbf{X}}^H \mathbf{\Phi_H} \widetilde{\mathbf{X}}$. Invoking (1), $\widetilde{\mathbf{C}}_\mathbf{X}$ can be evaluated as shown on the top of the next page, where $\odot$ denotes the element-wise multiplication (also called Hardmard product) and for $1 \le i \le N-1$

$$\Phi_i = \begin{bmatrix} \phi[i n_T] & \cdots & \phi[(i+1)n_T - 1] \\ \vdots & \ddots & \vdots \\ \phi[(i+1)n_T - 1] & \cdots & \phi[i n_T] \end{bmatrix}.$$

Moreover, $\widetilde{\mathbf{C}}_\mathbf{X}$ is independent of the phase reference $\mathbf{X}[0]$, as shown in (10).

Eqns. (10) and (9) together give the decision rule of the ML sequence estimation for unitary DSTM schemes. Note that the ML decision rules derived in [4] and [5] can be regarded as two special cases of the derived ML receiver, where the approach in [4] is for the diagonal DSTM and the approach in [5] is suitable for quasi-static channels (channel coefficients are constant over n_T symbol periods).

III. Reduced-Complexity Receivers for DSTM Schemes

The ML receiver (9) is an exhaustive search approach, where the complexity exponentially increases with the block length $N n_T$. This motivates an investigation of reduced-complexity receivers suitable for DSTM schemes.

A. Differential Detector (DD)

Under the quasi-static assumption, the differential detector in the context of DSTM has been derived in [1], [2]:

$$\begin{aligned}\widehat{\mathbf{A}}[k] &= \arg\min_{\widetilde{\mathbf{A}}[k]} \left\{-\mathrm{Tr}\left(\Re\{\mathbf{Y}[k-1]\widetilde{\mathbf{A}}[k]\mathbf{Y}^H[k]\}\right)\right\} \\ &= \arg\min_{\widetilde{\mathbf{A}}[k]} \left\{\| \mathbf{Y}[k] - \mathbf{Y}[k-1]\widetilde{\mathbf{A}}[k] \|^2\right\}, \end{aligned} \quad (11)$$

where the term $\mathbf{Y}[k-1]\widetilde{\mathbf{A}}[k] \approx \mathbf{H}[k]\widetilde{\mathbf{X}}[k]$ implies that (11) approximates the coherent ML estimate for $\mathbf{X}[k]$. If only two adjacent matrices are taken into consideration and the channel coefficients remain constant for $2n_T$ symbol periods, (11) has been shown to be the maximum-likelihood solution [1], [2].

Let us assume that the channel coefficients are approximately constant for $K_s n_T$ adjacent symbols. Consequently, the relevant term for detection in (11) can be generalized as follows:

$$\sum_{i=0}^{K_s-2}\sum_{j=1}^{K_s-1-i} \| \mathbf{Y}[k-i] - \mathbf{Y}[k-(i+j)] \times \prod_{l=j-1}^{0} \widetilde{\mathbf{A}}[k-(i+l)] \|^2, \quad (12)$$

where $\mathbf{Y}[k-(i+j)]\prod_{l=j-1}^{0}\widetilde{\mathbf{A}}[k-(i+l)]$ actually estimates $\mathbf{H}[k-i]\mathbf{X}[k-i]$. There are total $K_s - 1 - i$ estimates for $\mathbf{H}[k-i]\mathbf{X}[k-i]$ instead of one single estimate as in (11). For $K_s = 2$, (12) reduces to the argument in (11). Based on (12), $K_s - 1$ data symbols can jointly be estimated based on $K_s n_T$ channel outputs.

$$\widetilde{\mathbf{C}}_{\mathbf{X}} = \begin{bmatrix} \mathbf{I}_{n_T} & \Phi_1 \odot \left(\widetilde{\mathbf{A}}[1]\right) & \cdots & \Phi_{N-1} \odot \left(\prod_{m=1}^{N-1} \widetilde{\mathbf{A}}[m]\right) \\ \Phi_1 \odot \left(\widetilde{\mathbf{A}}[1]\right)^H & \mathbf{I}_{n_T} & \cdots & \Phi_{N-2} \odot \left(\prod_{m=2}^{N-1} \widetilde{\mathbf{A}}[m]\right) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{N-1} \odot \left(\prod_{m=1}^{N-1} \widetilde{\mathbf{A}}[m]\right)^H & \Phi_{N-2} \odot \left(\prod_{m=2}^{N-1} \widetilde{\mathbf{A}}[m]\right)^H & \cdots & \mathbf{I}_{n_T} \end{bmatrix} \in \mathcal{C}^{Nn_T \times Nn_T}. \quad (10)$$

B. Subblock-by-Subblock Detector (SSD)

Taking the time correlation of channel coefficients into account, the data estimation may be performed on a subblock-by-subblock basis. If K_s denotes the number of data matrices after DSTM within a subblock, the data estimation is performed for $K_s - 1$ data matrices before DSTM as follows:

$$\begin{aligned} &\left(\widehat{\mathbf{A}}[k-K_s+2], \cdots, \widehat{\mathbf{A}}[k]\right) \quad (13) \\ &= \arg \min_{\left\{\widetilde{\mathbf{A}}[k-K_s+2],\cdots,\widetilde{\mathbf{A}}[k]\right\}} \mathrm{L}\left(\widetilde{\mathbf{A}}[k-K_s+2], \cdots, \widetilde{\mathbf{A}}[k]\right). \end{aligned}$$

The conditional negative log-likelihood function is defined as

$$\begin{aligned} &\mathrm{L}\left(\widetilde{\mathbf{A}}[k-K_s+2], \cdots, \widetilde{\mathbf{A}}[k]\right) \quad (14) \\ &\triangleq \mathrm{Tr}\left\{\mathbf{Y}_{K_s}[k]\widetilde{\mathbf{C}}^{-1}_{\mathbf{Y}_{K_s}[k]}\mathbf{Y}^H_{K_s}[k]\right\} + \log\det\left(\widetilde{\mathbf{C}}_{\mathbf{Y}_{K_s}[k]}\right), \end{aligned}$$

where

$$\begin{aligned} \mathbf{Y}_{K_s}[k] &= [\mathbf{Y}[k-K_s+1], \cdots, \mathbf{Y}[k]] \\ &= \mathbf{H}_{K_s}[k]\mathbf{X}_{K_s}[k] + \mathbf{N}_{K_s}[k] \in \mathcal{C}^{n_R \times K_s n_T} \\ \widetilde{\mathbf{C}}_{\mathbf{Y}_{K_s}[k]} &= \mathrm{E}\left\{\mathbf{Y}^H_{K_s}[k]\mathbf{Y}_{K_s}[k] \mid \widetilde{\mathbf{X}}_{K_s}[k], \mathbf{\Phi}_{\mathbf{H}}\right\}. \end{aligned}$$

For the special case $f_D T = 0$, (14) coincides with (12), which implies that (12) is a maximum likelihood solution for time-invariant channels.

C. Sequence Estimator (SE)

The detectors presented in III-A and III-B ignore the dependence of data symbols between adjacent subblocks. Consequently, the sequence estimation decouples into isolated subblock detections. Taking the dependence of data symbols between adjacent subblocks into account, sequence estimation may be performed under the following assumption:

$$\begin{aligned} p\left(\mathbf{Y} \mid \widetilde{\mathbf{X}}, \mathbf{\Phi}_{\mathbf{H}}\right) &\approx \prod_{k=0}^{N_s-1} p\left(\mathbf{Y}_{K_s}[k] \mid \widetilde{\mathbf{X}}_{K_s}[k], \mathbf{\Phi}_{\mathbf{H}}\right) \\ &= \prod_{k=0}^{N_s-1} p\left(\mathbf{Y}_{K_s}[k] \mid \widetilde{\mathbf{A}}[k-K_s+2], \cdots, \widetilde{\mathbf{A}}[k], \mathbf{\Phi}_{\mathbf{H}}\right), \quad (15) \end{aligned}$$

where N_s denotes the total number of overlapping subblocks in each block and the second equation can easily be seen from (10). Fig. 2 illustrates how to divide a long burst into overlapping subblocks, where $M_s n_T$ symbols are overlapped between two adjacent subblocks. The parameter M_s actually indicates the dependence degree, where a larger M_s means a higher dependence and vice versa. For the maximal overlapping ($M_s = K_s - 1$), the number of overlapping subblocks can be determined as $N_s = N - K_s + 2$. For the minimal overlapping ($M_s = 1$), $N_s = N/(K_s - 1)$, and the sequence estimator reduces to the subblock-by-subblock detector (13), where the overlapping is necessary to resolve the phase ambiguity as obvious from II-C.

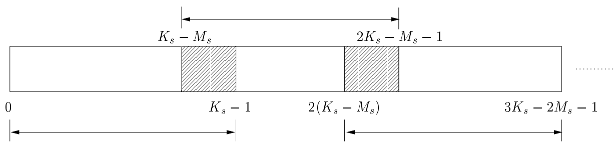

Fig. 2. Overlapping subblocks with parameters K_s and M_s

According to (15), sequence estimation is carried out as

$$\widehat{\mathbf{X}} = \arg\max_{\widetilde{\mathbf{X}}} \left\{ \prod_{k=0}^{N_s-1} p\left(\mathbf{Y}_{K_s}[k] \mid \widetilde{\mathbf{X}}_{K_s}[k], \mathbf{\Phi}_{\mathbf{H}}\right) \right\}. \quad (16)$$

Consequently, an evaluation of (16) can be accomplished by means of a trellis with M^{K_s-2} states ($M = 2^{Rn_T}$), where state transitions are determined by the data symbols corresponding to $\widetilde{\mathbf{A}}[k-K_s+2], \cdots, \widetilde{\mathbf{A}}[k]$, cf. (15). In the following, we consider the special case $M_s = K_s - 1$, which implies that there are M branches starting from each state and M branches merging into each state in the defined trellis. For other cases, i.e., $1 < M_s < K_s - 1$, suitable trellises can also be defined similar as in the context of noncoherent sequence estimation for DPSK signals [8].

IV. Numerical Results

The different receivers have been tested for two DSTM schemes, namely, the DSTM scheme proposed by Hughes [1] (referred to as DSTM-Hug in the sequel) and the diagonal DSTM proposed by Hochwald and Sweldens [2]. For simulations, we have selected the system parameters as $R = 1$, $n_T = 2$, $n_R = 1$, and $N = 66$. Accordingly, the data symbol matrices before DSTM for the DSTM-Hug scheme are chosen as follows: ($1 \le k \le N - 1$)

$$\mathbf{A}[k] \in \left\{ \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\},$$

and

$$\mathbf{X}[0] = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

serves as the phase reference.

Our simulation results show a similar performance between two DSTM schemes for quasi-static and time-varying channels. Therefore, only results for the DSTM-Hug scheme are presented within this paper. Fig. 3 shows the performance of different receivers for static channels, where no significant improvement is visible by means of SSD or SE compared to the conventional DD receiver. It can also be seen from Fig. 3 that all receivers achieve a diversity degree of two. By increasing K_s, the performance of SE/SSD will approach the performance of coherent detection, i.e., channel coefficients are perfectly known at the receiver (denoted as "known CIR"). Simulation results for different f_DTs are illustrated on the left hand side (LHS) of Fig. 4. For slowly fading channels ($f_DT \leq 5 \cdot 10^{-3}$), different receivers exhibit a similar performance despite different complexities. For moderately or fast time-varying channels, the DD receiver shows a significant performance degradation compared to SSD and SE. The SE shows a big advantage over the SSD for fast time-varying channels ($f_DT \geq 10^{-2}$). The error floor at high f_DT values can be reduced significantly by increasing the complexity of SE/SSD, as shown on the right hand side (RHS) of Fig. 4. While the diversity degree for the SE with $K_s = 4$ approximates one, the SE with $K_s = 5$ can achieve a diversity degree of about 1.5.

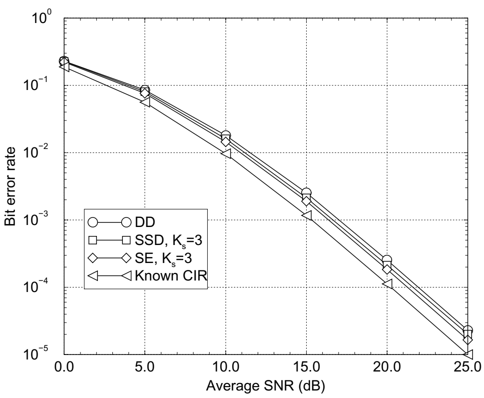

Fig. 3. BER performance of different receivers for $f_DT = 0$.

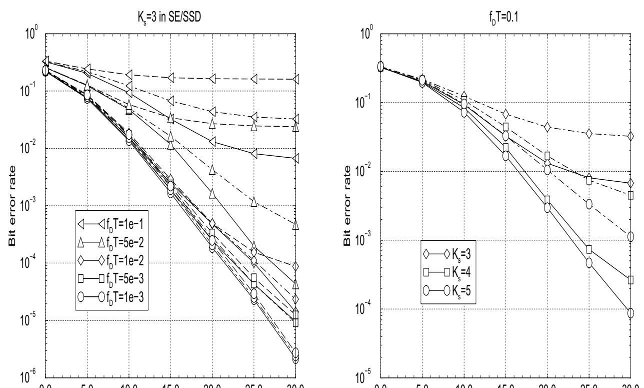

Fig. 4. Comparison of different receivers for different normalized fade rates. Solid lines, dot-dashed lines, and long-dashed lines correspond to SE, SSD, and DD, respectively.

Till this point, the SNR and f_DT values are perfectly known at the receiver. The effect of channel mismatch on the performance is shown in Fig. 5, where curves on the LHS and curves on the RHS correspond to a mismatch of noise variance and to a mismatch of f_DT, respectively. For large f_DTs, an under-determination of SNR causes a performance degradation, while an over-determination of SNR will be problematic for low f_DTs, e.g., $f_DT = 10^{-3}$. Concerning the mismatch of f_DT, an over-determination does not cause a significant degradation for $f_DT \geq 5 \cdot 10^{-3}$. The receivers under high f_DTs ($f_DT \geq 5 \cdot 10^{-2}$) is more robust against the channel mismatch than receivers under low f_DTs.

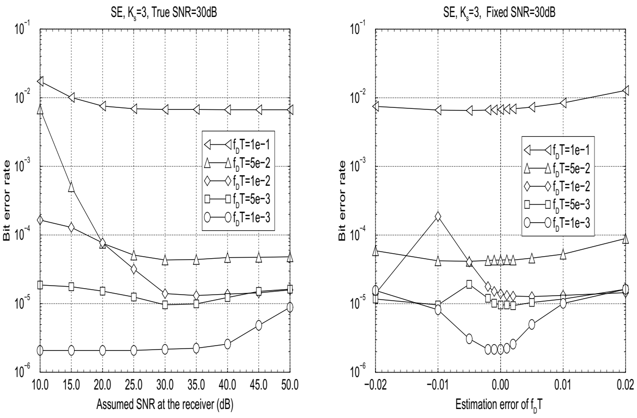

Fig. 5. Effect of channel mismatch on BER performance.

V. Conclusions

In this paper, we derived the ML receiver and reduced-complexity receivers thereof for unitary DSTM schemes, where only the correlation function of channel coefficients is available at the receiver. For the derived sequence estimator, the overlapping length M_s adjusts the computational complexity and the performance of the sequence estimator. The subblock-by-subblock detector and the differential detector have been shown as special cases of the sequence estimator. Increasing K_s in SE/SSD will improve the coding advantage of DSTM schemes for quasi-static channels and improve the diversity advantage for severe time-selective channels. The SE significantly outperforms the SSD for fast time-varying channels, where the DD fails to deliver reasonable results. Finally, the channel mismatch with respect to the noise variance and the normalized fade rate was investigated for the SE by means of simulations.

References

[1] B.L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[2] B.M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 2041–2052, Dec. 2000.

[3] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 1169–1175, July 2000.

[4] R. Schober and L. Lampe, "Noncoherent receivers for differential space-time modulation," *IEEE Trans. Commun.*, vol. 50, pp. 768–777, Apr. 2002.

[5] Y. Liu and X.-D. Wang, "Multiple-symbol decision-feedback space-time differential decoding in fading channels," *EURASIP J. Applied Signal Processing*, no. 3, pp. 297–304, Mar. 2002.

[6] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rates wireless communications: Performance analysis and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[7] H. Shin and J.H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2636–2647, Oct. 2003.

[8] R. Schober, *Noncoherent Detection and Equalization for MDPSK and MDAPSK Signals*, Dissertation, University of Erlangen-Nuremberg, Germany, May 2000.