

# Reduced-Complexity Space-Time Turbo-Equalization for Frequency-Selective MIMO Channels

Gerhard Bauch and Naofal Al-Dhahir

G. Bauch is with the Institute for Communications Engineering, Munich University of Technology, Germany.  
E-mail: Gerhard.Bauch@ei.tum.de.

Contact author. N. Al-Dhahir is with AT&T Shannon Laboratory, Florham Park, NJ 07932, USA. E-mail:  
naofal@research.att.com.

### Abstract

We consider turbo equalization of space-time-coded transmission over frequency-selective fading multiple-input multiple-output channels. A multiple-input multiple-output finite-impulse-response pre-filter is proposed and shown to reduce the turbo equalizer complexity significantly at a small performance loss. Advantages of the proposed scheme are that we do not alter the equalization algorithm or require the channel to be minimum phase. The prefiltered turbo equalizer is an attractive receiver structure for broadband wireless transmission using spectrally-efficient high-order modulation schemes as in EDGE.

### Keywords

Space-Time Codes, Transmit Diversity, Turbo Equalization, Channel-Shortening Filters.

## I. INTRODUCTION

Future wireless communication systems promise to offer a variety of multimedia services which require reliable transmission at high data rates. In order to achieve such high data rates, transmission over frequency-selective fading multiple-input multiple-output (MIMO) channels is of great interest since the capacity of MIMO channels shows enormous gains over that of their constituent single-input single-output (SISO) channels [1], [2]. Two main approaches have been proposed on how to make use of multiple antennas in a wireless system. The first approach is to transmit independent symbols simultaneously from different antennas as in Bell-Labs-Layered-Space-Time-Architecture (BLAST) [3]. The second approach is transmit diversity with space-time-coded modulation [4].

In this paper, we consider the general equalization problem in MIMO channels. We use interleaved space-time-coded transmissions as an example. However, the ideas are more general and can be applied to any coded transmission scheme over MIMO channels.

On frequency-selective channels, an equalizer has to be applied to resolve time dispersion. To achieve the best performance, a BCJR-MAP [5] or Viterbi equalizer is used since it does not suffer from noise enhancement or error propagation like linear or decision feedback equalizers (DFE), respectively. However, the complexity of a BCJR-MAP or Viterbi equalizer, measured by the number of states, increases exponentially with the number of transmit antennas and the channel memory. Even for moderate time dispersion, the complexity becomes prohibitive for practical systems, especially when high-order constellations (e.g. 8-PSK) are used to achieve higher spectral efficiencies.

In single-input channels, the equalization complexity can be reduced using standard reduced-complexity methods as described in [6], [7]. However, in MIMO channels, such methods require the channels from all transmit antennas to all receive antennas to be minimum phase simultaneously. In general, it is not possible to design a finite-length MIMO prefilter to meet the minimum-phase requirement exactly.

In this paper, we follow a different approach in that we propose to design a set of receive prefilters which concentrate the energy of the channel in a small number of adjacent taps. The equalizer sees a shortened channel and therefore needs a smaller number of states. In a coded system, powerful detection techniques like turbo equalization can be applied to the shortened MIMO channel to achieve high performance at reduced complexity. Two additional advantages of our proposed scheme are that we do not alter the ML or MAP equalization algorithm itself and we do not require the MIMO channel to be minimum phase. The main contribution of this paper is the design of a prefiltered reduced-complexity joint turbo equalization and decoding scheme for space-time-coded transmissions over MIMO frequency-selective channels. This work is distinct from [12] and [13] in that we consider space-time-coded signals and MIMO channels, we add a prefilter to reduce equalization complexity, and we consider high-order spectrally-efficient modulation schemes such as 8-PSK. Our focus in this paper is on turbo equalization for space-time trellis codes. Other equalization schemes for space-time trellis codes as well as equalization for space-time block codes are studied in [10].

This paper is organized as follows. We start by presenting the channel model in Section II. Section III includes background material on transmit antenna diversity with space-time codes and discusses the application of the turbo principle to joint equalization and decoding (turbo equalization). In Section IV, the problem of prohibitive complexity of turbo equalization in MIMO channels and the difficulties in applying standard reduced-complexity equalization methods are discussed. This is followed by proposing a reduced-complexity turbo equalization scheme with finite impulse response (FIR) channel shortening filters. Finally, simulation results are given in Section V.

## II. CHANNEL MODEL AND ASSUMPTIONS

We consider a MIMO wireless system with  $n_T$  transmit antennas and  $n_R$  receive antennas. Due to multipath propagation and time dispersion, the wireless MIMO channel consists of a frequency-selective fading channel between each transmit and receive antenna pair. The antennas on both ends are separated far enough to ensure independently fading channels from each transmit to each receive antenna. The memory  $D$  per transmit antenna is assumed to be the same for all channels. If the transmit antennas are physically co-located at the same base station, then this assumption is justified by the fact that the number of multipath components with different delays is dictated by large structures and reflecting objects.

The channel model considered in this paper is shown in Figure 1. The  $d$ -th tap of the frequency-selective channel from transmit antenna  $i$  to receive antenna  $j$  at time  $k$  is denoted  $h_k^{(ij)}(d)$ ,  $d = 0, \dots, D$ . All taps  $h_k^{(ij)}(d)$  are independent complex Gaussian random variables with zero mean and equal mean power and satisfy the channel energy normalization condition

$$\sum_{d=0}^D E\{|h_k^{(ij)}(d)|^2\} = 1. \quad (1)$$

The total mean energy transmitted per use of the MIMO channel is  $E_s$ , i.e. the mean energy per symbol  $x_k^{(i)}$  transmitted from antenna  $i$  is

$$E_s^{(i)} = \frac{E_s}{n_T} \quad : \quad \text{for } i = 1, \dots, n_T. \quad (2)$$

We consider quasi-static fading only. This implies transmission in bursts of  $L$  symbols per transmit antenna, where the channel is assumed to be constant over a burst and changes independently from burst to burst. This is a reasonable assumption for a mobile system like GSM or EDGE<sup>1</sup>, which employs frequency hopping and where the coherence-time is much larger than the burst duration. Due to delay restrictions, channel coding is applied to each burst separately. In the following mathematical formulation, we drop the index  $k$  from the channel taps since they are assumed constant over the block under consideration. The symbol  $y_k^{(j)}$  observed at the receive antenna  $j$  at time  $k$  is a superposition of the signals transmitted from all transmit antennas perturbed by noise. A vector of  $N$  symbols

<sup>1</sup>EDGE: Enhanced Data Rates for GSM Evolution.

per receive antenna is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (3)$$

where

$$\mathbf{y} = \begin{bmatrix} y_{k+N-1}^{(1)} \\ \vdots \\ y_{k+N-1}^{(n_R)} \\ \vdots \\ y_k^{(1)} \\ \vdots \\ y_k^{(n_R)} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_{k+N-1}^{(1)} \\ \vdots \\ x_{k+N-1}^{(n_T)} \\ \vdots \\ x_{k-D}^{(1)} \\ \vdots \\ x_{k-D}^{(n_T)} \end{bmatrix}, \mathbf{n} = \begin{bmatrix} n_{k+N-1}^{(1)} \\ \vdots \\ n_{k+N-1}^{(n_R)} \\ \vdots \\ n_k^{(1)} \\ \vdots \\ n_k^{(n_R)} \end{bmatrix}, \quad (4)$$

and

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}(\mathbf{0}) & \mathbf{H}(\mathbf{1}) & \cdots & \mathbf{H}(\mathbf{D}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}(\mathbf{0}) & \mathbf{H}(\mathbf{1}) & \cdots & \mathbf{H}(\mathbf{D}) & \mathbf{0} & \cdots \\ \vdots & & \ddots & & & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}(\mathbf{0}) & \mathbf{H}(\mathbf{1}) & \cdots & \mathbf{H}(\mathbf{D}) \end{bmatrix}. \quad (5)$$

The matrix

$$\mathbf{H}(\mathbf{d}) = \begin{bmatrix} h^{(11)}(d) & h^{(21)}(d) & \cdots & h^{(n_T 1)}(d) \\ \vdots & & & \vdots \\ h^{(1n_R)}(d) & h^{(2n_R)}(d) & \cdots & h^{(n_T n_R)}(d) \end{bmatrix} \quad (6)$$

consists of the  $d$ -th taps of all underlying SISO channels. Furthermore, we assume additive white Gaussian noise (AWGN) with variance  $\sigma^2 = N_0/2$  per real dimension and receive antenna, i.e. the autocorrelation matrix of the noise is given by

$$\mathbf{R}_{\mathbf{nn}} = N_0 \mathbf{I}_{n_R N}, \quad (7)$$

where  $\mathbf{I}_{n_R N}$  is the  $n_R N \times n_R N$  identity matrix.

### III. BACKGROUND MATERIAL

#### A. Principle and Examples of Space-Time Trellis Codes

Space-time coding schemes [4] use multiple transmit antennas to increase the diversity level, i.e. the asymptotic slope of the BER curve over SNR, and possibly achieve additional coding gain. The simplest example of a space-time trellis code is delay diversity where the

same symbol is transmitted from antenna  $i$  ( $i = 1, 2, \dots, n_T$ ) with a time delay of  $(i - 1)T_s$ , where  $T_s$  is the symbol duration. Delay diversity can be viewed as a rate  $1/n_T$  repetition code where the code symbols are spread in space. If we use a more powerful code instead of the repetition code in conjunction with delay diversity, a coding gain can be obtained in addition to the maximum diversity gain. Delay diversity and the best known space-time trellis code for 8-PSK modulation and  $n_T = 2$  transmit antennas taken from [4] illuminate how space-time trellis codes work. The trellis structures are shown in Figure 2. The tables in Figure 2 determine the symbols transmitted from antennas one and two, respectively. The first (second) symbol in row  $i$  and column  $j$  is the symbol transmitted from the first (second) antenna in the transition from state  $i$  to state  $j$ , where natural mapping of bits to constellation points is considered and a state is determined by the symbol in the memory element. It can be seen that the only difference between delay diversity and the space-time trellis code is in the symbol transmitted from the second antenna in odd-indexed states. The space-time trellis code is obtained from delay diversity by multiplying the symbol transmitted from antenna two by  $-1$  in odd-indexed states.

### B. Space-Time Trellis Codes in Frequency Selective Channels

A space-time trellis code with  $M^{n_T-1}$  states<sup>2</sup> transforms a frequency-selective MIMO channel with memory  $D$  per transmit antenna to a frequency-selective single-input multiple-output (SIMO) channel with memory  $D + n_T - 1$ . Diversity gain is obtained by the increased number of independently fading taps compared to a SISO system with the same channel memory  $D$ . The receiver sees a SIMO channel whose impulse response length is increased by  $n_T - 1$  taps. If the resulting impulse response length exceeds the length resolvable by the equalizer, standard reduced complexity techniques such as described in [6], [7] can be applied. DFE which is integrated in the equalizer metric.

Assuming no interleaver  $\Pi_c$ , it is clear from Figures 3 and 4, that even on frequency-selective channels, the space-time trellis code of Figure 2 can be detected with the same complexity as delay diversity. Now, the channel taps depend on the data which can be easily accounted for by proper labeling of the trellis transitions of the BCJR-MAP or Viterbi equalizer with channel taps. However, in this case, the application of standard reduced-

<sup>2</sup>Examples include delay diversity [25] and the 8-state 8-PSK space-time trellis code [4] illustrated in Figure 2.

complexity techniques is not possible since the allpass filter that converts the channel into its minimum-phase equivalent now depends on the data. This is a first motivation to investigate alternative reduced-complexity equalization techniques.

Especially in fast-fading environments, the performance can be further improved if an interleaver  $\Pi_c$  is used before transmission. The interleaving rule has to be the same for all transmit antennas in order not to destroy the diversity effect of the space-time code. This complicates the task of finding a simple single-input equivalent channel model in order to perform joint equalization and decoding in a single trellis. However, the turbo principle can now be applied to perform joint iterative equalization and decoding, as described in the next section.

### C. Turbo-Equalization

The principle of turbo equalization [12], [13] for interleaved transmission of coded data over a frequency-selective channel is illustrated in Figure 5. The equalizer computes the a-posteriori information  $\log P(x_k^{(i)}|\mathbf{y})$  about the code symbols based on the receive filter outputs  $\mathbf{y}$ . The a-priori input is set to  $\log P_a(x_k^{(i)}) = 0$  in the first equalization step. The a-posteriori information is deinterleaved according to the mapping rule  $\Pi_c^{-1}$  and passed to the channel input of the decoder. The decoder delivers not only an a-posteriori information about the information symbols but also about the code symbols. From this a-posteriori information  $\log P(x_k^{(i)}|\mathbf{y})$  we can compute an extrinsic information  $\log P_e(\hat{x}_k^{(i)})$  by subtracting the soft input from the corresponding soft output. The extrinsic information can be considered as independent from previous equalizer decisions. It is fed back to the a-priori input of the equalizer. Equalization of the stored received values is done again using the decoder's extrinsic information as an additional a-priori information in order to obtain less erroneous soft equalizer decisions. Several iterations can be performed in order to decrease the BER. We denote by *iteration* a complete detection step including one equalization and decoding step. Since detection becomes iterative in the second equalization step, we denote the first equalization and decoding step by *iteration 0*.

#### IV. REDUCED-COMPLEXITY TURBO EQUALIZATION

For the general case of transmission over MIMO frequency-selective channels, the complexity of a BCJR-MAP or Viterbi equalizer, measured by the number of states, grows exponentially not only with the per transmit antenna channel memory  $D$  and the spectral efficiency in bit/s/Hz but also with the number  $n_T$  of transmit antennas. For example, an equalizer for EDGE with 8-PSK modulation,  $n_T = 2$  transmit antennas, and the common per transmit antenna memory  $D = 4$  would need to deal with  $M^{n_T \cdot D} = 8^{2 \cdot 4} = 16777216$  states. This is clearly prohibitive for today's technology. Therefore, reduced-complexity techniques have to be applied.

Previous approaches include reduced-state and delayed decision feedback sequence estimation algorithms as proposed in [6] and [7], respectively, where the channel is usually made minimum phase by an allpass prefilter in order to concentrate the energy of the impulse response in its leading taps. The taps exceeding the resolvable length are either neglected or taken into account by a DFE which is integrated in the equalizer metric. However, for MIMO channels, the impulse responses from all transmit antennas have to be made minimum phase simultaneously. Generally, this is not possible with a FIR prefilter. Therefore, some adaptive compromise receive filter has to be found. Design criteria for such filters are still open research problems. Another approach would be the application of sequential algorithms like the M-BCJR algorithm [14].

Here, we consider an approach which does not alter the BCJR-MAP or Viterbi equalizer. We use a prefilter which concentrates the energy of the impulse response in a small number of adjacent taps. The cascade of the original MIMO channel and the prefilters can be viewed as a MIMO channel with smaller memory  $D_s < D$ . The succeeding equalizer can be designed for a smaller channel memory, which reduces the number of states by a factor of  $M^{n_T(D-D_s)}$ . Furthermore, as we will show in the next section, using this approach, part of the loss in time diversity due to channel shortening can be recovered by a higher level of "receive antenna diversity".

### A. MIMO Channel Shortening with FIR Filters

In order to shorten the effective channel memory, each of the received signals  $y_k^{(j)}$ ,  $j = 1, \dots, n_R$ , is filtered by  $n_{R_s}$  FIR filters, which have the structure shown in Figure 6. Each of the  $n_{R_s}$  filters has  $N$  taps. The cascade of the receive filters and the original MIMO channel with  $n_T$  transmit antennas and  $n_R$  receive antennas can be viewed as a MIMO channel with  $n_T$  transmit and  $n_{R_s}$  receive antennas. In order to reduce the number of states of the succeeding BCJR-MAP or Viterbi equalizer, the memory of the resulting channel should be  $D_s < D$ . For the special case of  $n_{R_s} = n_T$ , channel-shortening filters have been derived in [15]. In this paper, we derive design criteria for shortening filters followed by a BCJR-MAP or Viterbi equalizer and relax the condition on  $n_{R_s}$ .

The filtered signal is given by

$$\tilde{\mathbf{y}} = \mathbf{W}\mathbf{y}, \quad (8)$$

where

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n_R \cdot N} \\ \vdots & & & \vdots \\ w_{n_{R_s},1} & w_{n_{R_s},2} & \cdots & w_{n_{R_s},n_R \cdot N} \end{bmatrix} \quad (9)$$

is the  $n_{R_s} \times n_R \cdot N$  filter taps matrix. The  $n_{R_s} \times n_T(D_s + 1)$  matrix  $\mathbf{H}^{(s)}$

$$\mathbf{H}^{(s)} = \begin{bmatrix} h_{(s)}^{(1,1)}(0) & \cdots & h_{(s)}^{(n_T,1)}(0) & \cdots & h_{(s)}^{(1,1)}(D_s) & \cdots & h_{(s)}^{(n_T,1)}(D_s) \\ \vdots & & & & & & \vdots \\ h_{(s)}^{(1,n_{R_s})}(0) & \cdots & h_{(s)}^{(n_T,n_{R_s})}(0) & \cdots & h_{(s)}^{(1,n_{R_s})}(D_s) & \cdots & h_{(s)}^{(n_T,n_{R_s})}(D_s) \end{bmatrix} \quad (10)$$

defines the shortened channel. The diversity advantage of the shortened system is upper bounded by  $n_T \cdot n_{R_s} \cdot (D_s + 1)$ . Because only  $D_s + 1 < D + 1$  channel taps are considered, the filtering error vector is given by

$$\mathbf{e} = \tilde{\mathbf{H}}^{(s)}\mathbf{x} - \mathbf{W}\mathbf{y}. \quad (11)$$

The  $n_{R_s} \times n_T(N + D)$  matrix  $\tilde{\mathbf{H}}^{(s)}$  is defined by

$$\tilde{\mathbf{H}}^{(s)} = \begin{bmatrix} \mathbf{0}_{n_{R_s} \times n_T \Delta} & \mathbf{H}^{(s)} & \mathbf{0}_{n_{R_s} \times n_T \cdot s} \end{bmatrix} \quad (12)$$

$$= \mathbf{H}^{(s)} \begin{bmatrix} \mathbf{0}_{n_T(D_s+1) \times n_T \Delta} & \mathbf{I}_{n_T(D_s+1)} & \mathbf{0}_{n_T(D_s+1) \times n_T \cdot s} \end{bmatrix}, \quad (13)$$

where  $s = N + D - D_s - \Delta - 1$ . The relative delay  $\Delta$ ,  $0 \leq \Delta \leq N + D - D_s - 1$ , between the receive filters and the shortened channel impulse response is to be optimized. We restrict  $\Delta$  to be equal for all transmit antennas. A generalization to different  $\Delta$  was considered in [15] to reduce the error energy of the prefilter. Such a generalization would not increase the number of equalizer states but the assignment of transmitted symbols to trellis transitions would have to be implemented adaptively, which might not be desirable in practice.

For a given relative delay  $\Delta$  and a given shortened channel  $\tilde{\mathbf{H}}^{(s)}$ , the optimum (according to a minimum mean square error criterion) filter tap settings  $\mathbf{W}$  are computed using the orthogonality principle [16], which states that the optimum error sequence is orthogonal to the observed data, i.e.  $E\{\mathbf{e}\mathbf{y}^H\} = \mathbf{0}$  (where  $(\cdot)^H$  denotes the conjugate transpose and  $E[\cdot]$  denotes the expected value). Using (11) we obtain

$$E\{(\tilde{\mathbf{H}}^{(s)}\mathbf{x} - \mathbf{W}\mathbf{y})\mathbf{y}^H\} = \tilde{\mathbf{H}}^{(s)}\mathbf{R}_{xy} - \mathbf{W}\mathbf{R}_{yy} = \mathbf{0}, \quad (14)$$

where  $\mathbf{R}_{xy}$  and  $\mathbf{R}_{yy}$  are the input–output cross–correlation and the output auto–correlation matrices, respectively. Therefore, the optimum filter tap settings are given by

$$\mathbf{W}_{\text{opt}} = \tilde{\mathbf{H}}^{(s)}\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}. \quad (15)$$

Since the additive white Gaussian noise  $\mathbf{n}$  and the data  $\mathbf{x}$  are uncorrelated, i.e.  $E\{\mathbf{x}\mathbf{n}^H\} = \mathbf{0}$ , we obtain the cross-correlation matrix

$$\mathbf{R}_{xy} = E\{\mathbf{x}\mathbf{y}^H\} = \mathbf{R}_{xx}\mathbf{H}^H, \quad (16)$$

and the autocorrelation matrix

$$\begin{aligned} \mathbf{R}_{yy} &= E\{\mathbf{y}\mathbf{y}^H\} \\ &= E\{(\mathbf{H}\mathbf{x} + \mathbf{n})(\mathbf{x}^H\mathbf{H}^H + \mathbf{n}^H)\} \\ &= \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{R}_{nn}. \end{aligned} \quad (17)$$

Combining (15), (16), and (17), we have

$$\begin{aligned} \mathbf{W}_{\text{opt}} &= \tilde{\mathbf{H}}^{(s)}\mathbf{R}_{xx}\mathbf{H}^H(\mathbf{H}\mathbf{R}_{xx}\mathbf{H}^H + \mathbf{R}_{nn})^{-1} \\ &= \tilde{\mathbf{H}}^{(s)}(\mathbf{R}_{xx}^{-1} + \mathbf{H}^H\mathbf{R}_{nn}^{-1}\mathbf{H})^{-1}\mathbf{H}^H\mathbf{R}_{nn}^{-1}, \end{aligned} \quad (18)$$

where we used the *Matrix Inversion Lemma*<sup>3</sup> in the second equality above. Equation (18) clearly shows that the MIMO prefilter  $\mathbf{W}_{\text{opt}}$  includes a whitening matched filter component given by  $\mathbf{H}^H \mathbf{R}_{\text{nn}}^{-1}$  in addition to the channel shortening component (see also [15]).

In transmit diversity using space-time codes, the data  $\mathbf{x}$  is correlated. However, if an interleaver is applied before transmission the data can be assumed to be uncorrelated in time and space and the auto-correlation matrix is given by

$$\mathbf{R}_{\text{xx}} = E\{\mathbf{x}\mathbf{x}^H\} = \frac{E_s}{n_T} \mathbf{I}_{n_T(N+D)}. \quad (19)$$

Now we have to determine the optimum shortened channel impulse response  $\tilde{\mathbf{H}}^{(s)}$ . The error autocorrelation matrix  $\mathbf{R}_{\text{ee}}$  of the error vector  $\mathbf{e}$  in (11) is

$$\begin{aligned} \mathbf{R}_{\text{ee}} &= E\{\mathbf{e}\mathbf{e}^H\} = E\{(\tilde{\mathbf{H}}^{(s)}\mathbf{x} - \mathbf{W}\mathbf{y})(\mathbf{x}^H \tilde{\mathbf{H}}^{(s)H} - \mathbf{y}^H \mathbf{W}^H)\} \\ &= \tilde{\mathbf{H}}^{(s)} \mathbf{R}_{\text{xx}} \tilde{\mathbf{H}}^{(s)H} - \tilde{\mathbf{H}}^{(s)} \mathbf{R}_{\text{xy}} \mathbf{W}^H - \mathbf{W} \mathbf{R}_{\text{xy}}^H \tilde{\mathbf{H}}^{(s)H} + \mathbf{W} \mathbf{R}_{\text{yy}} \mathbf{W}^H. \end{aligned} \quad (20)$$

Using (15) and (13) we further obtain

$$\begin{aligned} \mathbf{R}_{\text{ee}} &= \tilde{\mathbf{H}}^{(s)} \underbrace{(\mathbf{R}_{\text{xx}} - \mathbf{R}_{\text{xy}} \mathbf{R}_{\text{yy}}^{-1} \mathbf{R}_{\text{xy}}^H)}_{\mathbf{R}^\perp} \tilde{\mathbf{H}}^{(s)H} \\ &= \tilde{\mathbf{H}}^{(s)} \mathbf{R}^\perp \tilde{\mathbf{H}}^{(s)H} \\ &= \mathbf{H}^{(s)} \begin{bmatrix} \mathbf{0}_{n_T(D_s+1) \times n_T \Delta} \mathbf{I}_{n_T(D_s+1)} \mathbf{0}_{n_T(D_s+1) \times n_T s} \\ \mathbf{0}_{n_T \Delta \times n_T(D_s+1)} \\ \mathbf{I}_{n_T(D_s+1)} \\ \mathbf{0}_{n_T s \times n_T(D_s+1)} \end{bmatrix} \mathbf{H}^{(s)H} \\ &= \mathbf{H}^{(s)} \tilde{\mathbf{R}} \mathbf{H}^{(s)H}. \end{aligned} \quad (21)$$

The SNR at the filter output  $j$ , denoted by  $SNR_j$ , is given by the  $(j, j)$  diagonal entry of the matrix  $\mathbf{R}_{\tilde{\mathbf{r}}\tilde{\mathbf{r}}} \mathbf{R}_{\text{ee}}^{-1}$  where  $\mathbf{R}_{\tilde{\mathbf{r}}\tilde{\mathbf{r}}}$  is the auto-correlation matrix of the shortened channel outputs with no noise, i.e.

$$\tilde{\mathbf{r}} = \mathbf{H}^{(s)} \mathbf{x}. \quad (22)$$

The performance of the BCJR-MAP equalizer depends on the SNRs at the outputs of the shortened channel. Therefore, the FIR filters are designed such that the sum of the

<sup>3</sup> $\mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{D} \mathbf{A}^{-1} \mathbf{B} + \mathbf{C}^{-1})^{-1} \mathbf{D} \mathbf{A}^{-1} = (\mathbf{A} + \mathbf{B} \mathbf{C} \mathbf{D})^{-1}$ .

equalizer's input SNRs is maximized, i.e.

$$\sum_{j=1}^{n_{R_s}} SNR_j = \text{trace}(\mathbf{R}_{\tilde{\mathbf{r}}\tilde{\mathbf{r}}}\mathbf{R}_{\mathbf{e}\mathbf{e}}^{-1}) \rightarrow \max. \quad (23)$$

The constraint

$$\mathbf{H}^{(s)}\mathbf{H}^{(s)\mathbf{H}} = \tilde{\mathbf{H}}^{(s)}\tilde{\mathbf{H}}^{(s)\mathbf{H}} = \mathbf{I}_{n_{R_s}} \quad (24)$$

is imposed to exclude the trivial solution  $\mathbf{H}^{(s)} = \mathbf{0}$  and was shown in [15] to result in a lower prefiltering error energy than the more traditional constraint  $\mathbf{H}^{(s)}(k) = \mathbf{I}$  for some  $0 \leq k \leq D_s$ . Using (24), (22) and (19) the design criterion in (23) simplifies to

$$\text{trace}(\mathbf{R}_{\mathbf{e}\mathbf{e}}^{-1}) \rightarrow \max, \quad (25)$$

which is maximized by choosing  $\mathbf{H}^{(s)}$  to *diagonalize*  $\mathbf{R}_{\mathbf{e}\mathbf{e}}$ , as defined in (21), subject to the constraint in (24). To perform this diagonalization, consider the eigen-decomposition

$$\tilde{\mathbf{R}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathbf{H}} \quad (26)$$

of  $\tilde{\mathbf{R}}$  in (21) where  $\mathbf{U}$  is a unitary matrix containing the eigenvectors of  $\tilde{\mathbf{R}}$  and  $\mathbf{\Lambda}$  is a diagonal matrix with entries  $\lambda_i$  denoting the eigenvalues of  $\tilde{\mathbf{R}}$  in *increasing order*, i.e.,  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{n_T(D_s+1)}$ . Using this decomposition we have

$$\mathbf{R}_{\mathbf{e}\mathbf{e}} = \mathbf{H}^{(s)}\tilde{\mathbf{R}}\mathbf{H}^{(s)\mathbf{H}} = \mathbf{H}^{(s)}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathbf{H}}\mathbf{H}^{(s)\mathbf{H}}. \quad (27)$$

If we define  $\mathbf{e}_i^{\mathbf{H}}$  to be the  $i^{\text{th}}$  unit row vector (has a one in its  $i^{\text{th}}$  entry and zeros everywhere else), then we can write the  $(i, i)$  diagonal element of  $\mathbf{R}_{\mathbf{e}\mathbf{e}}$  as follows

$$\mathbf{R}_{\mathbf{e}\mathbf{e}}(\mathbf{i}, \mathbf{i}) = \mathbf{e}_i^{\mathbf{H}}\mathbf{H}^{(s)}\tilde{\mathbf{R}}\mathbf{H}^{(s)\mathbf{H}}\mathbf{e}_i \stackrel{\text{def}}{=} \mathbf{g}_i^{\mathbf{H}}\tilde{\mathbf{R}}\mathbf{g}_i, \quad (28)$$

where the row vector  $\mathbf{g}_i^{\mathbf{H}}$  has unit norm since  $\mathbf{g}_i^{\mathbf{H}}\mathbf{g}_i = \mathbf{e}_i^{\mathbf{H}}\mathbf{H}^{(s)}\mathbf{H}^{(s)\mathbf{H}}\mathbf{e}_i = \mathbf{1}$ . To maximize (25), we should choose  $\mathbf{g}_i$  to minimize  $\mathbf{R}_{\mathbf{e}\mathbf{e}}(\mathbf{i}, \mathbf{i})$  for all  $1 \leq i \leq n_{R_s}$  subject to the constraint  $\mathbf{g}_i^{\mathbf{H}}\mathbf{g}_i = 1$ . It is well known that this minimization is achieved by choosing the  $\mathbf{g}_i$ 's to be the eigenvectors of  $\tilde{\mathbf{R}}$  corresponding to its  $n_{R_s}$  smallest eigenvalues. Therefore, the columns of  $\mathbf{H}^{(s)\mathbf{H}}$  have to be chosen as eigenvectors of  $\tilde{\mathbf{R}}$ . With this choice, it can be readily checked that the matrix  $\mathbf{R}_{\mathbf{e}\mathbf{e}}$  becomes diagonal and we have

$$\mathbf{H}^{(s)}\mathbf{U} = \left[ \mathbf{I}_{n_{R_s}} \mathbf{0}_{n_{R_s} \times n_T(D_s+1) - n_{R_s}} \right]. \quad (29)$$

The relative delay  $\Delta$  is optimized by choosing the matrix  $\bar{\mathbf{R}}$  to maximize

$$\sum_{i=1}^{n_{R_s}} \frac{1}{\lambda_i} \rightarrow \max, \quad (30)$$

where  $\lambda_i, i = 1, \dots, n_{R_s}$ , are the smallest eigenvalues of  $\bar{\mathbf{R}}$ . The size of  $\bar{\mathbf{R}}$  is  $n_T(D_s + 1) \times n_T(D_s + 1)$ . Therefore, the shortened channel  $\mathbf{H}^{(s)}$  has a maximum of

$$n_{R_s} \leq n_T(D_s + 1) \quad (31)$$

outputs. If the original channel has

$$n_R < n_T(D_s + 1) \quad (32)$$

receive antennas, then the equivalent shortened MIMO channel can have a larger number of outputs. This can be viewed as a transformation of time diversity to receive antenna diversity. We emphasize that having  $n_{R_s} > n_R$  does not increase the number of *states* in the turbo equalizer/decoder. Instead, only the branch metric calculation complexity is increased as if the number of receive antennas has increased from  $n_R$  to  $n_{R_s}$ .

In summary, in order to calculate the FIR filter taps  $\mathbf{W}$  and the shortened channel  $\mathbf{H}^{(s)}$ , we first have to determine  $\mathbf{H}^{(s)}$  from the eigenvectors corresponding to the  $n_{R_s}$  smallest eigenvalues of  $\bar{\mathbf{R}}$ . Thereby, the relative delay  $\Delta$  is optimized. Then, the filter tap settings can be calculated according to (15).

Since  $\mathbf{R}_{ee}$  is a diagonal matrix, the noise at different outputs of the shortened channel is spatially uncorrelated. However, due to the prefiltering, the noise is temporally colored at each of the filter outputs and the noise power is different across the outputs of the shortened channel. Better performance is achieved by incorporating knowledge of the temporal noise auto-correlation matrix in the equalizer metric calculation. In this paper, we assume white Gaussian noise for simplicity and account only for the different noise powers at the prefilter outputs in the BCJR equalizer metric.

### B. Fractionally-Spaced FIR Filters

Using the Matrix Inversion Lemma, the matrix  $\mathbf{R}^\perp$  in (21) can be expressed as follows

$$\begin{aligned} \mathbf{R}^\perp &= (\mathbf{R}_{xx}^{-1} + \mathbf{H}^H \mathbf{R}_{nn}^{-1} \mathbf{H})^{-1} \\ &= \left( \frac{n_T}{E_s} \mathbf{I}_{n_T(D+N)} + \frac{1}{N_0} \mathbf{H}^H \mathbf{H} \right)^{-1}. \end{aligned} \quad (33)$$

Since the dimension of  $\mathbf{H}$  is  $n_R N \times n_T(N + D)$  it follows that

$$\text{rank } \mathbf{H}^H \mathbf{H} = \min\{n_R N, n_T(N + D)\}. \quad (34)$$

Hence, in the case of a small number  $n_R < n_T$  of receive antennas, the number of zero eigenvalues of  $\mathbf{H}^H \mathbf{H}$  increases. Equivalently, the number of large eigenvalues of  $\mathbf{R}^\perp$ , and hence  $\mathbf{R}_{ee}$ , increases since  $\tilde{\mathbf{H}}^{(s)}$  is a unitary matrix. This results in lower SNRs at the FIR filter outputs and worse equalizer performance. For small  $n_R$ , the rank of  $\mathbf{H}^H \mathbf{H}$  can be increased using fractionally-spaced FIR filters. Assuming an oversampling factor of  $l$ , we obtain

$$\mathbf{y} = \begin{bmatrix} y_{k+N-1,l}^{(1)} \\ \vdots \\ y_{k+N-1,1}^{(n_R)} \\ \vdots \\ y_{k,l}^{(1)} \\ \vdots \\ y_{k,1}^{(n_R)} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_{k+N-1}^{(1)} \\ \vdots \\ x_{k+N-1}^{(n_T)} \\ \vdots \\ x_{k-D}^{(1)} \\ \vdots \\ x_{k-D}^{(n_T)} \end{bmatrix}, \mathbf{n} = \begin{bmatrix} n_{k+N-1,l}^{(1)} \\ \vdots \\ n_{k+N-1,1}^{(n_R)} \\ \vdots \\ n_{k,l}^{(1)} \\ \vdots \\ n_{k,1}^{(n_R)} \end{bmatrix} \quad (35)$$

instead of (4) and the channel matrix  $\mathbf{H}$  now becomes of size  $n_R l N \times n_T(N + D)$ . Hence, provided that the transmitted signal has excess bandwidth (beyond the minimum Nyquist bandwidth), temporal oversampling increases the rank of  $\mathbf{H}^H \mathbf{H}$  from that given in (34) to

$$\text{rank } \mathbf{H}^H \mathbf{H} = \min(ln_R N, n_T(N + D)). \quad (36)$$

It is also clear from (36) that, even with sufficient excess bandwidth, temporal oversampling by a factor more than  $\frac{n_T}{n_R}(1 + \frac{D}{N})$  would not increase the rank of  $\mathbf{H}^H \mathbf{H}$  beyond the value  $n_T(N + D)$ .

### C. MIMO Prefilter Computational Complexity

In this subsection, we enumerate the computational tasks required to compute  $\mathbf{H}_{\text{opt}}^{(s)}$  and  $\mathbf{W}_{\text{opt}}$  and their complexities as measured by the number of complex multiplies. We assume uncorrelated input and noise processes.

1. Computation of  $\bar{\mathbf{R}}$  using (21) requires inversion of the *block-Toeplitz* matrix  $\mathbf{R}_{yy}$  with a complexity of  $O(l^3 n_R^3 N^2)$  operations using the efficient algorithm in [17]. Computation of the Hermitian matrix product  $\mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx}$  requires  $\frac{1}{2} n_T n_R l N (N + D) (n_R l N + n_T (N + D))$  complex multiplies.
  2. The complexity of computing  $\mathbf{H}^{(s)}$  is dominated by the complexity of the eigen-decomposition in Equation (26) which requires  $O(n_T^3 (n_{R_s} + 1)^3)$  operations.
  3. Finally, computing  $\mathbf{W}_{opt}$  using (15) requires  $n_{R_s} n_T n_R l N (D_s + 1)$  complex multiplies.
- It is worth mentioning that this complexity estimate assumes a given value of  $\Delta$ , i.e., the complexity of searching for the optimum  $\Delta$  is not included.

In Figure 7, we plot the computational complexity in millions of instructions per second (MIPS) versus  $N$  and  $D$  for  $n_T = n_R = n_{R_s} = 2$ ,  $l = 2$ , and  $D_s = \frac{D}{2}$ . We assume that each complex multiply can be executed in one instruction. Furthermore, this figure assumes that the MIMO prefilter coefficients are updated once every 0.5 msec which is approximately equal to the burst duration in EDGE. It can be seen from the figure that for  $D = 4$ , as in a Typical Urban EDGE channel, and  $N = 10$ , the computational complexity of the  $\frac{T}{2}$ -spaced MIMO prefilter is  $\approx 80$  MIPS which is within the processing power of state-of-the-art programmable DSP chips.

## V. SIMULATION RESULTS

In this section, we present simulation results for BPSK and 8-PSK delay diversity with  $n_T = 2$  transmit antennas. Results for the 8-state 8-PSK space-time trellis code (see Figure 2) are also given. We use the ‘‘BCJR Symbol-by-Symbol MAP algorithm’’ for both equalization and decoding and assume that the receiver has perfect knowledge of the channel taps and the noise variance. Before transmission, the code symbols are reordered by a pseudo random interleaver. As stated before, the interleaving rule is the same for all transmit antennas. The channel is considered time invariant during transmission of a block of size 450 information bits but changes independently from one block to the next (quasi-static fading). The channel power delay profile is assumed uniform as described in Section II.

For each block, only one turbo iteration is performed. It is a typical result for turbo detection in block fading channels that an improvement in Bit Error Rate (BER) is achieved

only in the first iteration [18].

Unless otherwise stated, we use the maximum possible number  $n_{R_s}$  of channel shortening filters according to (31) with  $N = 10$  taps per receive antenna and an oversampling factor  $l = 1$ . First, we consider BPSK delay diversity for a simple illustration of the main ideas. Figure 8 compares the BER for different channel memories  $D$  and  $D_s$ , respectively, and  $n_R = 2$  receive antennas. The performance of the full turbo equalizer without channel shortening ( $D = D_s$ ) is also included. It can be seen from the Figure that channel shortening performs well. At a BER of  $10^{-5}$ , the degradation is about 2 dB only if the channel is shortened from  $D = 2$  to  $D_s = 1$ . In the case of  $D = 4$  and  $D_s = 1$  the degradation increases to 3.5 dB. However, in this case the equalizer has to deal with only 4 states instead of 256, which translates to a significant reduction in complexity. If the channel is only shortened from  $D = 4$  to  $D_s = 2$ , the degradation reduces to 2.1 dB, but the equalizer has to deal with 16 states. The BPSK curves show only small gains achieved by the turbo iteration due to the short block length and the fact that BPSK delay diversity is a weak outer code with a constraint length of only two. However, our objective here is to show the principle. For 8-PSK, we obtain a considerably larger gain by a turbo iteration as will be shown later. The impact of the number  $N$  of FIR filter taps is depicted in Figure 9 for  $D = 2$ ,  $D_s = 1$ , and  $n_R = n_T = 2$  transmit and receive antennas. The performance improves appreciably with increasing number of taps up to about  $N = 10$ . Therefore, we chose  $N = 10$  for our simulations. Figure 10 demonstrates the benefit of using the maximum number  $n_{R_s}$  of FIR filters according to (31). For the example of  $D = 2$  and  $D_s = 1$ , the shortened channel has a maximum of  $n_{R_s} = 4$  outputs. If we consider only  $n_{R_s} = n_R = n_T = 2$ , there is a significant loss in terms of the diversity level achieved. However, part of the loss is recovered by the turbo iteration. In case of  $n_{R_s} = 4$ , part of the loss in terms of time diversity is transformed to a higher level of “receive antenna diversity”. The turbo iteration now achieves an additional gain of about 0.5 dB but does not increase the diversity level. The gain over  $n_{R_s} = 2$  is about 1 dB at a BER of  $10^{-5}$ . Figure 11 shows the BER for  $D = 2$  and  $D_s = 1$  if only  $n_R = 1$  receive antenna is used. As explained in Section IV-B, we observe a high error floor due to the low effective SNR. However, channel shortening with  $n_R = 1$  works well if fractionally-spaced FIR filters with

an oversampling factor  $l = 2$  are used. In order to demonstrate the advantage of full turbo equalization, we also show results for MAP detection in Figure 12, where joint equalization and decoding is performed in a single trellis. For MAP detection no interleaving and no turbo iterations are possible. Figure 13 gives simulation results for 8-PSK modulation with delay diversity and the space-time trellis code of Figure 2. The channel is shortened from  $D = 2$  to  $D_s = 1$ . Unlike BPSK, we obtain a significant gain due to the turbo iteration. Furthermore, we observe a much larger difference between delay diversity and the space-time code than in flat-fading channels, where delay diversity performs about 2 dB worse than the space-time trellis code for  $n_R = 2$ . The coding gain of a space-time trellis code seems to become more important when prefiltering techniques are applied, especially for  $n_R \geq n_T$ . All simulation results presented thusfar adopted BER as a performance measure. In some scenarios, block error rate (BLER) is also of interest. In Figure 14 we repeat the comparisons made in Figure 13 but with BLER as a performance measure to demonstrate that the same general conclusions continue to hold.

## VI. CONCLUSIONS

We proposed the use of an FIR MIMO prefilter to reduce the complexity of space-time turbo equalization for broadband wireless channels. The complexity reductions become very dramatic for high-order signal constellation and long channel delay spread. Part of the performance loss due to the channel shortening performed by the MIMO prefilter is recovered using iterative detection and a transformation of time diversity to spatial diversity. Main advantages of the proposed scheme are that we do not alter the MAP or ML equalization algorithm itself and do not require the channel to be minimum phase. We believe that the combination of space-time coding, MIMO prefiltering, and turbo equalization offers a viable solution to the problem of high data rate transmission over broadband wireless channels.

## ACKNOWLEDGMENTS

We would like to thank Dr. A.R. Calderbank of AT&T Shannon Laboratory and Prof. J. Hagenauer of Munich University of Technology for their continued support of this work. The first author would also like to thank members of AT&T Shannon Laboratory for their

support during his visit. The second author would like to thank members of Institute for Communications Engineering for their hospitality during his short visit to the Munich University of Technology.

## REFERENCES

- [1] G. J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.
- [2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications (ETT)*, vol. 10, no. 6, November/December 1999.
- [3] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, Nr.2, pp. 41–59, 1996.
- [4] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.
- [5] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 284–287, March 1974.
- [6] M.V. Eyuboglu and S.U. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Transactions on Communications*, vol. 36, pp. 12–20, January 1988.
- [7] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Transactions on Communications*, vol. 37, no. 5, pp. 428–436, May 1989.
- [8] G. Bauch, J. Hagenauer, and N. Seshadri, "Turbo-TCM and transmit antenna diversity in multipath fading channels," in *2nd International Symposium on Turbo Codes*, September 2000, pp. 307–312.
- [9] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting and decoding: Turbo-codes (1)," in *International Conference on Communications (ICC)*. IEEE, May 1993, pp. 1064–1070.
- [10] N. Al-Dhahir, "Overview of equalization schemes for space-time-coded transmission with application to EDGE," in *Vehicular Technology Conference (VTC)*. IEEE, October 2001, pp. 1053–1057.
- [11] J. Hagenauer, "The turbo principle: Tutorial introduction and state of the art," in *International Symposium on Turbo Codes*. ENST de Bretagne, September 1997, pp. 1–11.
- [12] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of inter-symbol interference: Turbo-equalization," *European Transactions on Telecommunications*, vol. 6, no. 5, pp. 507–511, September–October 1995.
- [13] G. Bauch, H. Khorram, and J. Hagenauer, "Iterative equalization and decoding in mobile communications systems," in *The Second European Personal Mobile Communications Conference (2.EPMCC'97) together with 3. ITG-Fachtagung "Mobile Kommunikation"*. VDE/ITG, September/October 1997, pp. 307–312.
- [14] V. Franz and J. Anderson, "Concatenated decoding with a reduced-search BCJR algorithm," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 186–195, February 1998.
- [15] N. Al-Dhahir, "FIR channel-shortening equalizers for MIMO ISI channels," *IEEE Transactions on Communications*, pp. 213–218, February 2001.
- [16] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 2nd edition, 1991.
- [17] H. Akaike, "Block toeplitz matrix inversion," *SIAM J. Appl. Math.*, vol. 24, no. 2, March 1973.
- [18] G. Bauch, "*Turbo-Entzerrung*" und *Sendeantennen-Diversity* mit "*Space-Time-Codes*" im *Mobilfunk*, phd

thesis, Department of Communications Engineering, Munich University of Technology, January 2001, (in German).

- [19] N. Al-Dhahir and J. Cioffi, "Efficiently computed reduced-parameter input-aided MMSE equalizers for ML detection: A unified approach," *IEEE Transactions on Information Theory*, vol. 42, pp. 903–915, May 1996.
- [20] N. Al-Dhahir and A.H.Sayed, "The finite-length MIMO MMSE-DFE," *IEEE Transactions on Signal Processing*, pp. 2921–2936, October 2000.
- [21] G. Bauch and V. Franz, "A comparison of soft-in/soft-out algorithms for "turbo-detection"," in *International Conference on Telecommunications (ICT)*, June 1998.
- [22] G. Bauch and A. Naguib "MAP equalization of space-time coded signals over frequency selective channels," in *Wireless Communications and Networking Conference (WCNC)*, September 1999.
- [23] V. Franz and G. Bauch, "Turbo-detection for enhanced data for GSM evolution," in *IEEE Conference on Vehicular Technology (VTC)*, September 1999.
- [24] A. Glavieux, C. Laot, and J. Labat, "Turbo-equalization over a frequency selective channel," in *International Symposium on Turbo Codes*. ENST de Bretagne, September 1997, pp. 96–102.
- [25] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *International Conference on Communications (ICC)*. IEEE, 1993, pp. 1630–1633.

## LIST OF FIGURES

1	MIMO ISI channel model. . . . .	21
2	Delay diversity and space-time trellis code with 8 states for 8-PSK modulation and $n_T = 2$ transmit antennas. . . . .	21
3	8-PSK 8 state space-time trellis code for $n_T = 2$ with interleaving in frequency- selective channels. . . . .	22
4	Equivalent channel model for 8-PSK 8 state space-time trellis code with $n_T =$ 2 in frequency-selective channels. . . . .	22
5	Turbo equalization. . . . .	23
6	FIR Channel Shortening Filter. . . . .	24
7	Computational Complexity. . . . .	25
8	BPSK, $n_T = n_R = 2$ , maximum $n_{R,s}$ , $N = 10$ . . . . .	25
9	BPSK, $n_T = n_R = 2$ , $n_{R,s} = 4$ , $D = 2$ , $D_s = 1$ . . . . .	26
10	BPSK, $n_T = n_R = 2$ , $D = 2$ , $D_s = 1$ , $N = 10$ . . . . .	27
11	BPSK, $n_T = 2$ , $n_R = 1$ , $D = 2$ , $D_s = 1$ , $N = 10$ . . . . .	28
12	BPSK, $n_T = 2$ , $D = 2$ , $D_s = 1$ , $N = 10$ . . . . .	28
13	8-PSK, $n_T = n_R = 2$ , $n_{R,s} = 4$ , $D = 2$ , $D_s = 1$ , $N = 10$ . . . . .	29
14	8-PSK, $n_T = n_R = 2$ , $n_{R,s} = 4$ , $D = 2$ , $D_s = 1$ , $N = 10$ . . . . .	29

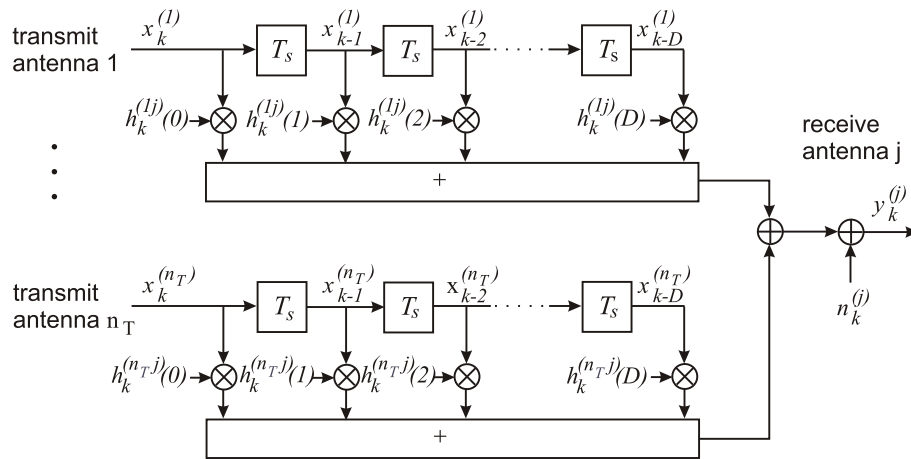


Fig. 1. MIMO ISI channel model.

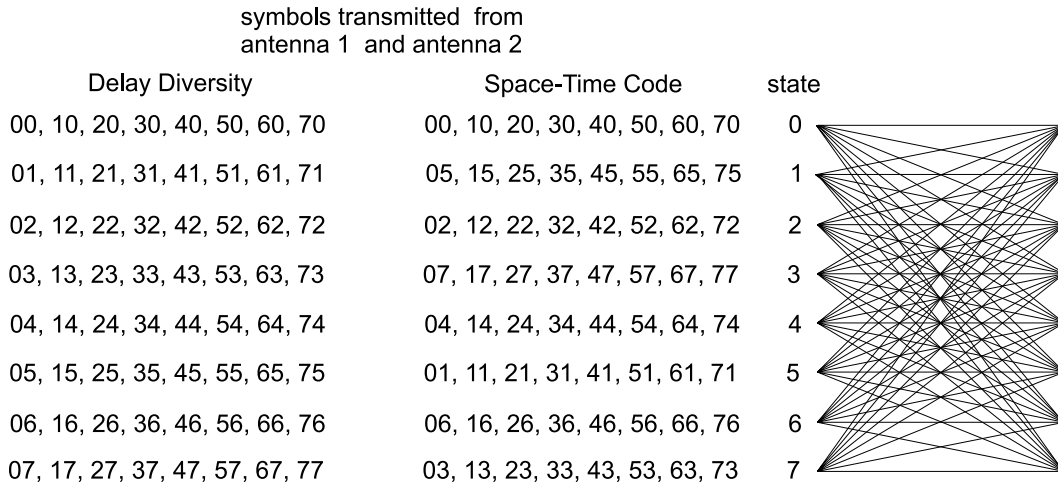


Fig. 2. Delay diversity and space-time trellis code with 8 states for 8-PSK modulation and  $n_T = 2$  transmit antennas.

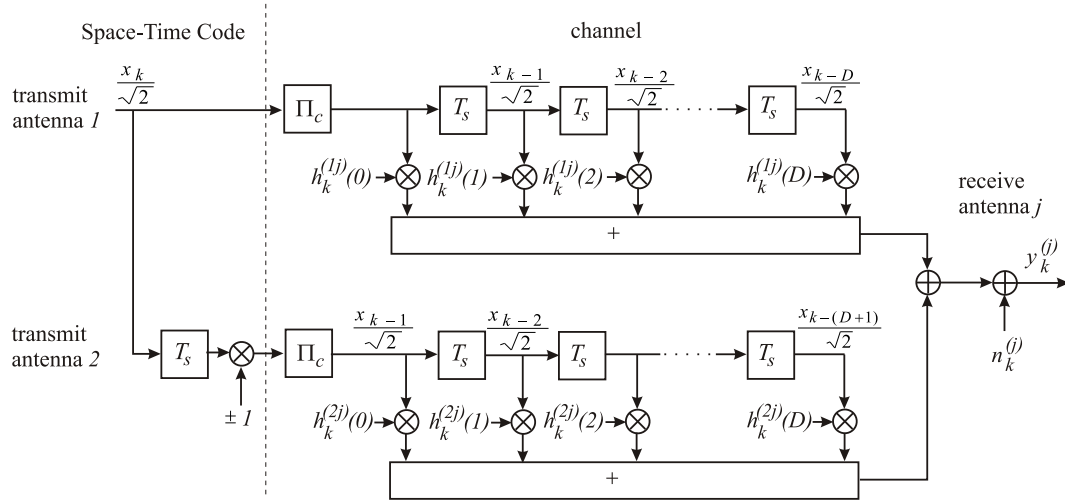


Fig. 3. 8-PSK 8 state space-time trellis code for  $n_T = 2$  with interleaving in frequency-selective channels.

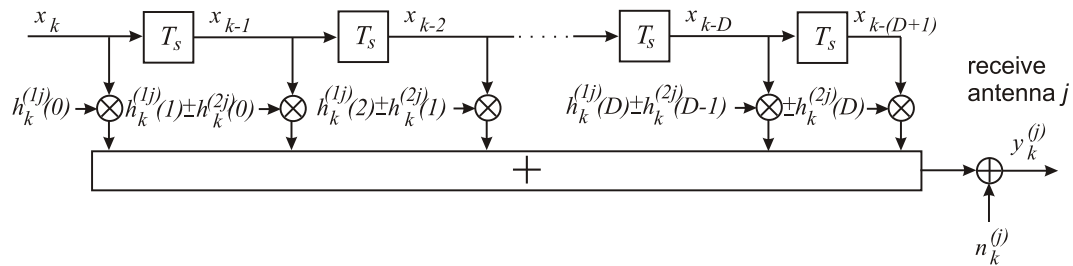


Fig. 4. Equivalent channel model for 8-PSK 8 state space-time trellis code with  $n_T = 2$  in frequency-selective channels.

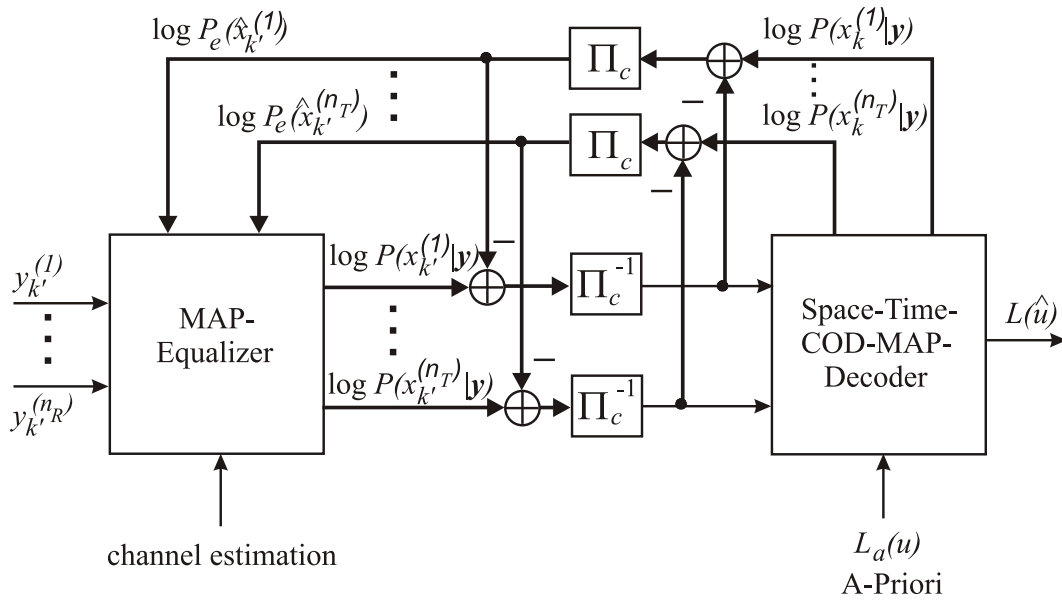


Fig. 5. Turbo equalization.

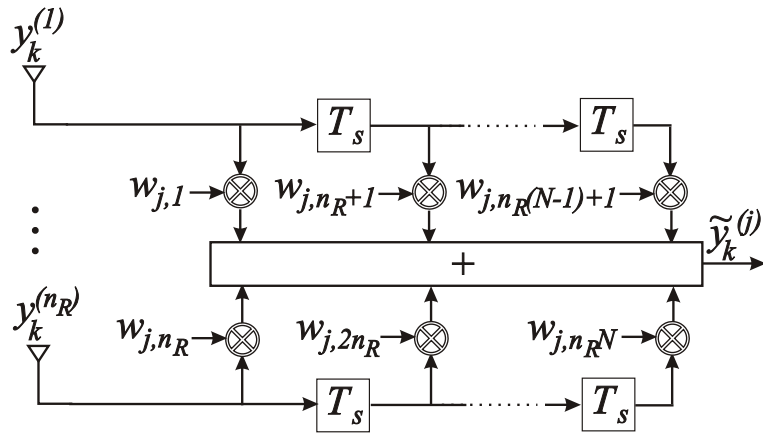


Fig. 6. FIR Channel Shortening Filter.

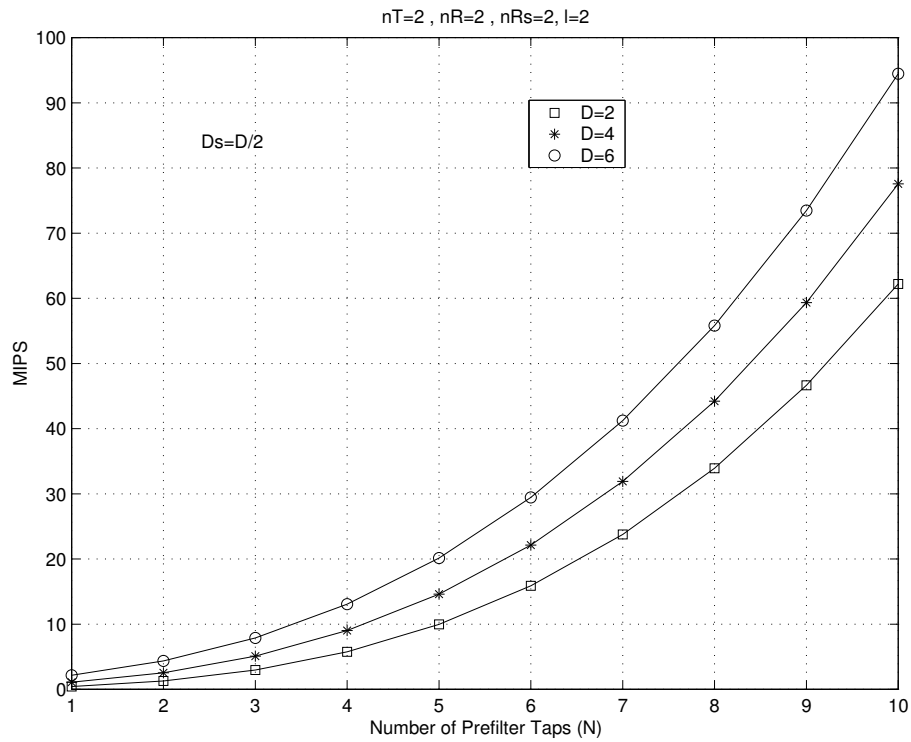


Fig. 7. Computational Complexity.

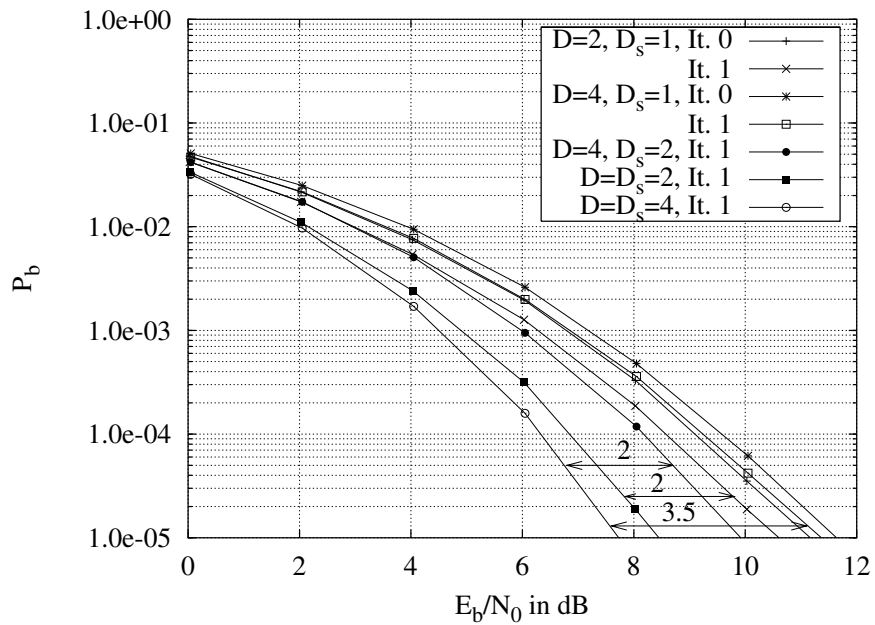


Fig. 8. BPSK,  $n_T = n_R = 2$ , maximum  $n_{R_s}$ ,  $N = 10$ .

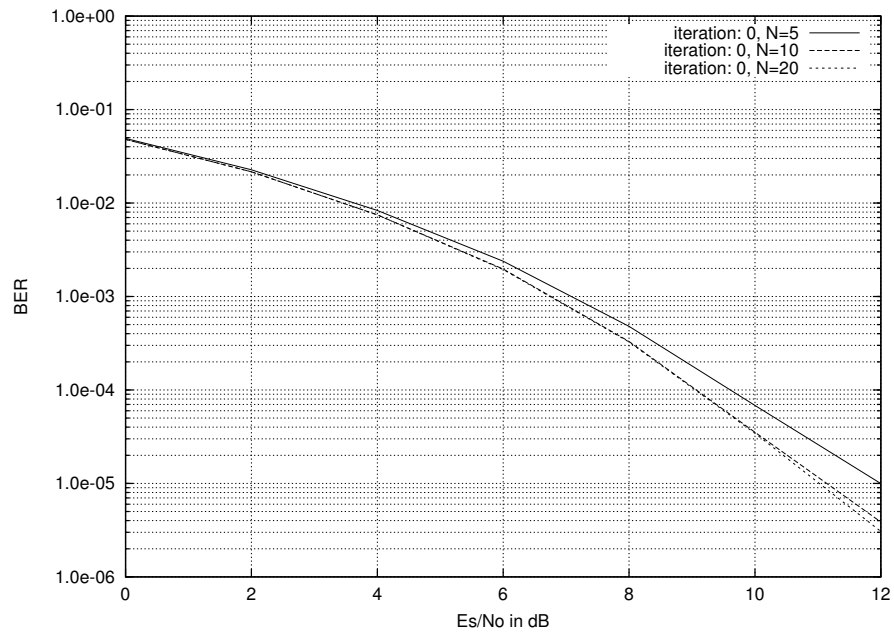


Fig. 9. BPSK,  $n_T = n_R = 2$ ,  $n_{R,s} = 4$ ,  $D = 2$ ,  $D_s = 1$ .

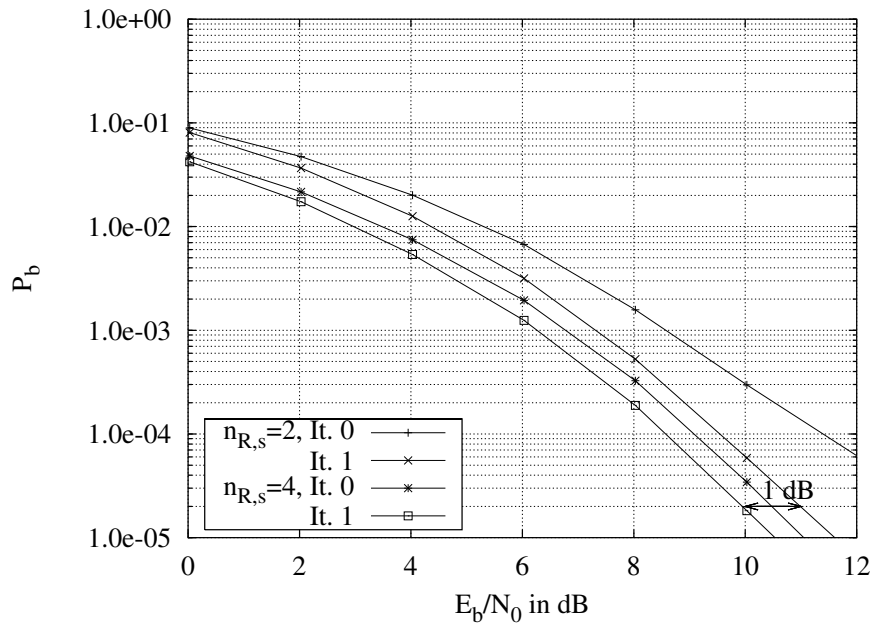


Fig. 10. BPSK,  $n_T = n_R = 2$ ,  $D = 2$ ,  $D_s = 1$ ,  $N = 10$ .

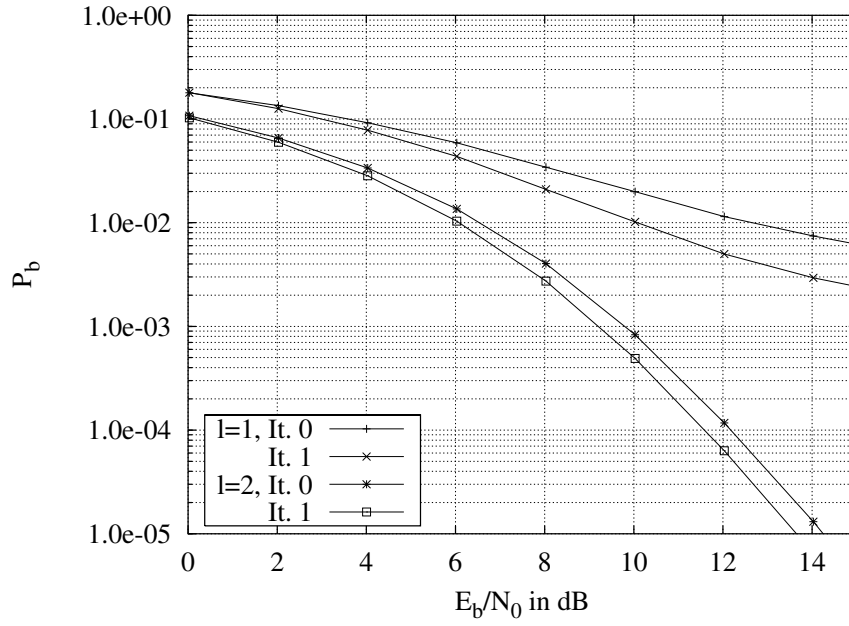


Fig. 11. BPSK,  $n_T = 2$ ,  $n_R = 1$ ,  $D = 2$ ,  $D_s = 1$ ,  $N = 10$ .

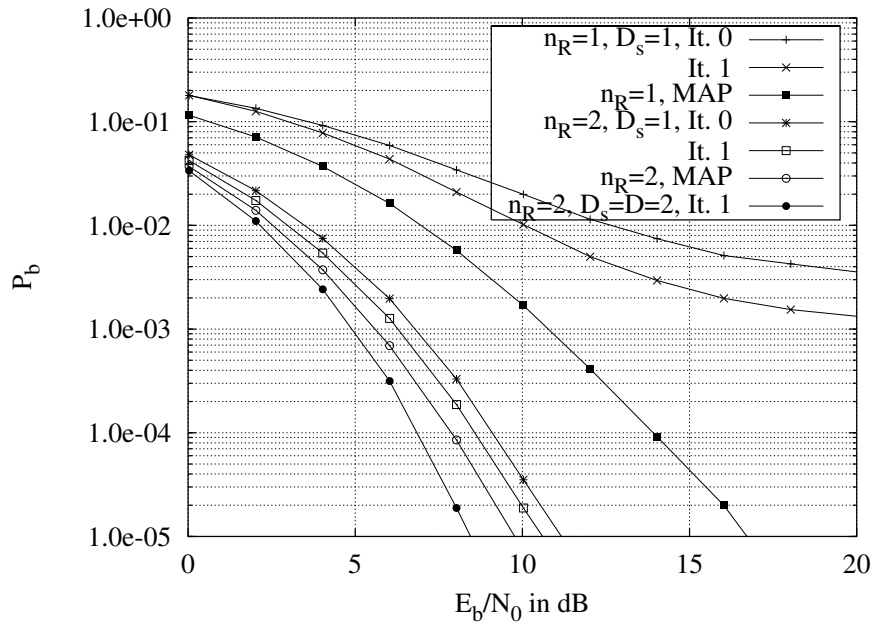


Fig. 12. BPSK,  $n_T = 2$ ,  $D = 2$ ,  $D_s = 1$ ,  $N = 10$ .

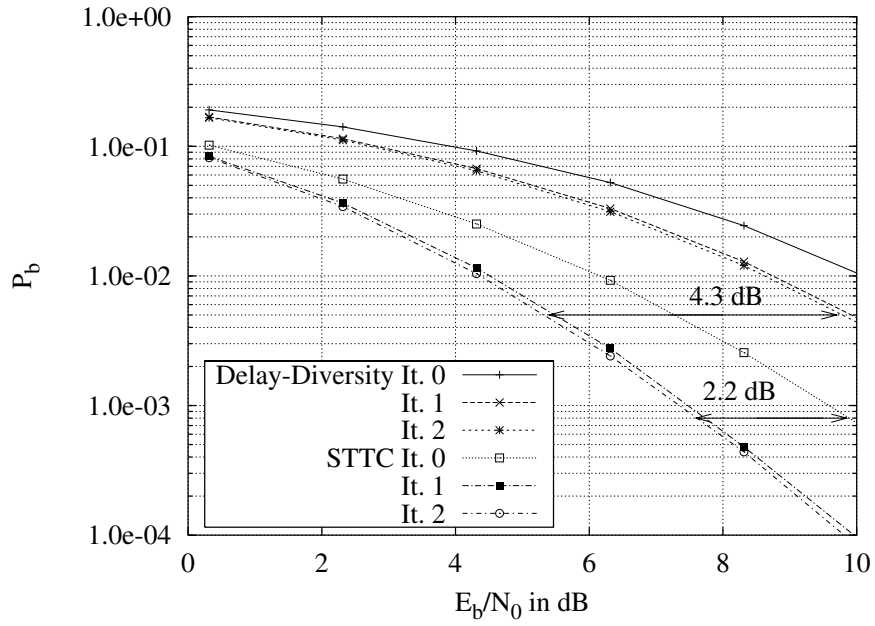


Fig. 13. 8-PSK,  $n_T = n_R = 2$ ,  $n_{R,s} = 4$ ,  $D = 2$ ,  $D_s = 1$ ,  $N = 10$ .

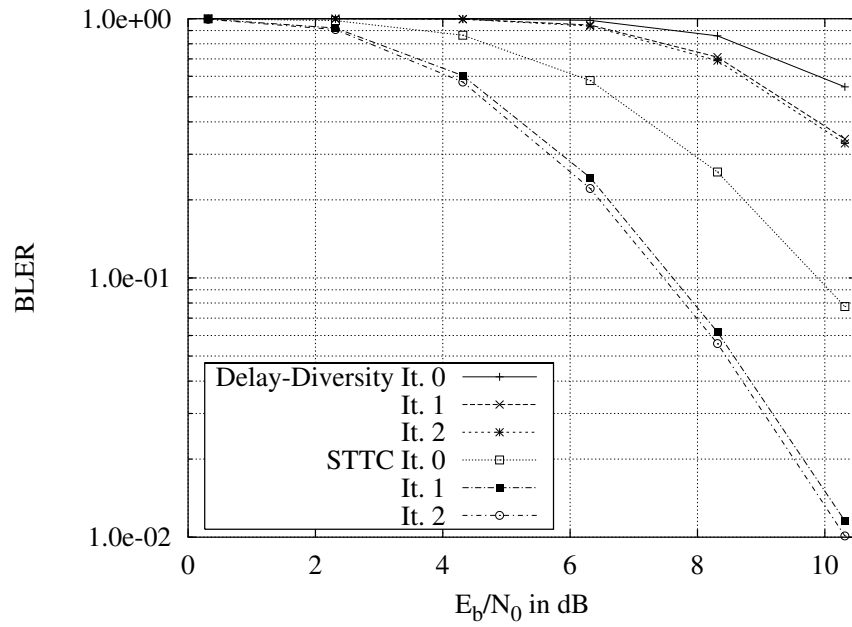


Fig. 14. 8-PSK,  $n_T = n_R = 2$ ,  $n_{R,s} = 4$ ,  $D = 2$ ,  $D_s = 1$ ,  $N = 10$ .