Wrapped Space-Time Codes for Quasi-Static Multiple-Antenna Channels

Giuseppe Caire[1], *Giulio Colavolpe*[2]

[1]Mobile Communications Group, Institut EURECOM, Valbonne, FRANCE
[2]Dipartimento di Ingegneria dell'Informazione, Università di Parma, Parma, ITALY

Abstract

We propose a new space-time coding scheme for the quasi-static multiple-antenna channel with perfect channel state information at the receiver and no channel state information at the transmitter. The new scheme includes both trellis space-time codes and layered space-time codes as special cases. When the number of transmit antennas is not larger than the number of receive antennas our scheme can be efficiently decoded by minimum mean-square error (MMSE) decision-feedback interference cancellation coupled with Viterbi decoding through the use of per-survivor processing (PSP). We discuss the code design for the new scheme, and show that finding codes with optimal diversity is much easier than for conventional trellis space-time codes.

Our scheme yields a large performance gain with respect to coded V-BLAST of similar complexity, and it can be easily coupled with the recently proposed *linear-dispersion precoding* to handle the case of a number of receive antennas smaller than the number of transmit antennas.

1. Introduction

Transmission schemes based on multiple antennas have attracted much attention in the recent years as a viable solution to increase spectral efficiency and performance of wireless channels. Roughly speaking, works on multiple antennas can be classified depending on the assumptions on the channel state information (CSI) available at the transmitter and at the receiver. In practice, the assumption of perfect CSI at the receiver holds approximately when the channel varies very slowly with respect to the duration of a codeword (quasi-static assumption). This is a quite realistic assumption in several situations where the mobility of wireless terminals is limited or absent (e.g., indoor wireless local-area networks, wireless local loops). On the contrary, the assumption of perfect CSI at the transmitter holds only if a delay-free error-free feedback link from receiver to transmitter exists, or if time-division duplexing is used, where each end can estimate the channel from the incoming signal in the reverse direction. Motivated by the above considerations, we conclude that assuming perfect CSI at the receiver and no CSI at the transmitter is reasonable.

Coding schemes for the quasi-static multiple-antenna channel with perfect receiver CSI have been proposed in several works (see for example [1, 2, 3, 4, 5]). In this paper we propose a new scheme nicknamed "wrapped" space-time coding that includes both the trellis space-time codes of [1] and the layered space-time codes of [2] as special cases. For the case $r \geq t$, our scheme can be efficiently decoded by minimum mean-square error (MMSE) decision-feedback interference cancellation coupled with Viterbi decoding, through the use of per-survivor processing (PSP) [6]. We discuss the code design for the new scheme, and show that finding codes with optimal diversity is

actually much easier than for conventional trellis space-time codes. We show many examples where the maximum possible diversity is achieved by well-known trellis codes.

We compare our scheme with V-BLAST [7], which is also based on decision-feedback interference cancellation but it does not exploit PSP, and we show that the latter suffers severely from error propagation in the feedback decisions while the proposed scheme does not. Also, we consider the case $r < t$ and we show that our scheme can work as outer coding where inner coding (or "precoding") is obtained by the recently proposed linear dispersion codes [5].

2. Coding for Quasi-Static Multiple-Antenna Channels

The multiple-input multiple-output (MIMO) system with t transmitting (Tx) and r receiving (Rx) antennas considered in this paper is defined by [1, 8]

$$\mathbf{y}_n = \sqrt{\gamma}\mathbf{H}\mathbf{x}_n + \mathbf{z}_n, \quad n = 1, \ldots, N \qquad (1)$$

where $\mathbf{x}_n \in \mathcal{X}^t$ is the vector of modulation symbols transmitted in parallel at time n by the Tx antennas, $\mathcal{X} \subseteq \mathbb{C}$ denotes a complex modulation signal set with unit average energy, $\mathbf{z}_n \in \mathbb{C}^r$ is the noise vector i.i.d. $\sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$, $\mathbf{y}_n \in \mathbb{C}^r$ is the corresponding vector of received signal samples at the output of the Rx antennas, $\mathbf{H} \in \mathbb{C}^{r \times t}$ is the channel matrix, γ is the SNR per Tx antenna, and N is the code block length. We assume that the channel matrix is normalized such that $\frac{1}{rt}\text{trace}E[\mathbf{H}\mathbf{H}^H] = 1$, so that the average received SNR per Rx antenna is given by $t\gamma$. We consider the case where $\mathbf{H}$ is random but constant over $N \gg \max\{t, r\}$ channel uses, and we assume that the receiver knows $\mathbf{H}$ perfectly, while the transmitter has no knowledge of $\mathbf{H}$.

A space-time code (STC) for the above channel is a set $\mathcal{S} \subseteq \mathcal{X}^{t \times N}$ of $t \times N$ complex matrices (codewords). Codeword matrices $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$ are transmitted by columns, in N consecutive channel uses. The STC spectral efficiency is given by $\eta = \frac{1}{N}\log_2 |\mathcal{S}|$ bit/channel use. By definition, the information bit-energy over noise power spectral density ratio is given by $E_b/N_0 = t\gamma/\eta$.

In [1], Tarokh *et al.* found criteria to design STC. They considered the pairwise-error probability (PEP) with maximum-likelihood (ML) decoding and, in the case of Rayleigh/Rician fading coefficients, they indicated as the most important criterion for constructing STC, the maximization of the minimum rank ρ of the codeword difference matrix $\mathbf{D} = \mathbf{X}' - \mathbf{X}$ over all distinct $\mathbf{X}, \mathbf{X}' \in \mathcal{S}$. We shall refer to this minimum rank as the code *rank-diversity*.

For large N, STC can be constructed from multidimensional trellis codes (M-TCM) with trellis termination [1, 4]. Namely, consider a M-TCM code $\mathcal{C}$ over $\mathcal{X}$ of rate $R = b/t$
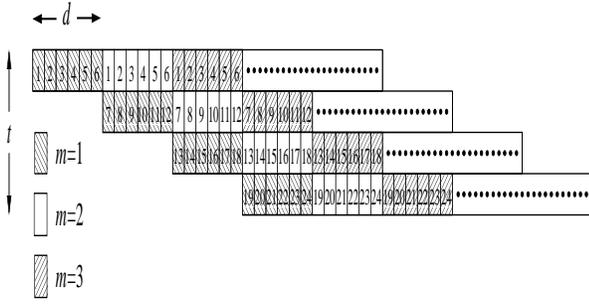
Figure 1: *Layered ST code with $t = 4$, $d = 6$ and $N' = 24$. The index m indicates the component codewords $\mathbf{c}^{(m)}$ and the integer entries in the array indicate the index of the ℓ-th element $c_\ell^{(m)}$ of $\mathbf{c}^{(m)}$.*

Figure 2: *Wrapped ST code with $t = 4$ and $d = 6$. The integer entries in the array indicates the index of the ℓ-th element c_ℓ of the component codeword $\mathbf{c}$. An example of the indexing rule $\ell_j(n)$ is given for $j = 3$ and $n = 20$.*

bit/symbol, where each trellis step corresponds to b information (input) bits and to t code (output) symbols, and the subcode of all $\mathbf{c} \in \mathcal{C}$ leaving a given trellis state s_0 and merging into a given trellis state s_N after N trellis steps. Then, a trellis STC can be obtained by simply formatting the codewords $\mathbf{c}$ as $t \times N$ matrices, i.e., by transmitting the t symbols produced by the trellis encoder at each trellis step in parallel on the t Tx antennas.

The difficulty in constructing these codes is that the rank-diversity is hard to evaluate and it is not easily related to the algebraic properties of the underlying M-TCM code. An exception is represented by the class of binary and quaternary trellis codes over $\mathbb{Z}_2$ and $\mathbb{Z}_4$ mapped onto BPSK and QPSK, respectively. In fact, in this case a condition on the underlying algebraic codes referred to as the *binary-rank* criterion is shown to imply the rank-diversity of the resulting STC $\mathcal{S}$ and it is used to construct STCs with full rank-diversity (i.e., with $\rho = t$) [8].

The ML decoder for trellis STC can be implemented by the Viterbi Algorithm (VA) applied on the trellis of the underlying M-TCM code. However, for fixed code rate R and rank-diversity ρ the decoder complexity grows exponentially with the number of Tx antennas t. Because of the difficulty in code design and because of decoding complexity, trellis STC are practically restricted to small t.

In [2], Foschini proposed a STC scheme suited to handle a very large number of Tx and Rx antennas. In this scheme, M codes $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(M)}$ over $\mathcal{X}$ produce independently codewords $\mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(M)}$ of length $N' = td$, where $d \geq 1$ and M are given integers. These codewords are diagonally interleaved (or "layered"),[1] as shown in Fig. 1, in order to form the $t \times N$ (with $N = d(M + t - 1)$) codeword matrix $\mathbf{X} = \mathcal{L}_d(\mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(M)})$, where $\mathcal{L}_d$ indicates the "layered" diagonal interleaver with interleaving delay d (see Fig. 1).

A reduced-complexity suboptimal decoder for this scheme is obtained by a linear front-end followed by decision-feedback interference cancellation [2]. The linear front-end, defined by a matrix $\mathbf{F} \in \mathbb{C}^{r \times t}$, produces the sequence of received vectors

$$\mathbf{v}_n = \mathbf{F}^H \mathbf{y}_n, \quad n = 1, \ldots, N$$

Then, each codeword $\mathbf{c}^{(m)}$ is decoded by taking as observable the sequence of samples

$$r_\ell^{(m)} = v_{j,n} - \sqrt{\gamma} \sum_{k=j+1}^{t} b_{j,k} \widehat{x}_{k,n}, \quad \ell = 1, \ldots, N' \quad (2)$$

---

[1] In [7] another layered scheme based on standard rectangular (or "vertical") interleaving is considered, and in [9] a comparison between "horizontal" and diagonal layered schemes is presented.
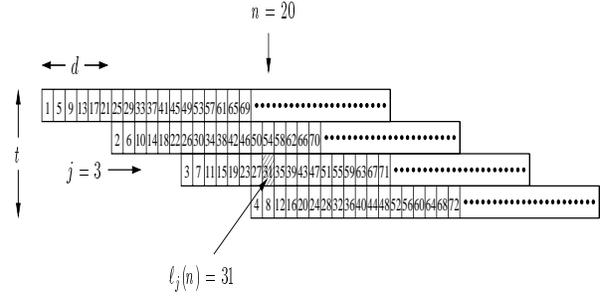
where the one-to-one index mapping $(m, \ell) \leftrightarrow (j, n)$ is induced by the interleaver $\mathcal{L}_d$, $v_{j,n}$ is the j-th element of $\mathbf{v}_n$, $b_{j,k}$ is the (j, k)-th element of the matrix $\mathbf{B} = \mathbf{F}^H \mathbf{H}$ and $\widehat{x}_{k,n}$ is a decision on the (k, n)-th symbol of $\mathbf{X}$. From Fig. 1, we see that the elements $x_{k,n}$ for $k = j + 1, \ldots, t$ correspond to either zeros (for which no decision is needed) or to symbols of codewords $\mathbf{c}^{(m')}$ with index $m' < m$. Therefore, by decoding the codewords in the order $m = 1, \ldots, M$, the decisions needed in (2) are provided by earlier decoded codewords.

3. "Wrapped" Space-Time Codes

In the original layered construction of [2], the length of each component codeword is $N' = dt$. Because of practical hardware complexity limitations t cannot be a very large number. This implies that in order to have long codewords, the interleaving delay d must be large. If interleaving delay is an issue, the layered scheme is forced to work with short component code block length N'. This might pose a serious problem for using trellis codes with a large number of states. In fact, the code memory might not be negligible with respect to N' thus yielding a non-negligible rate loss due to trellis termination.

For this reason, we propose a scheme which keeps the the simplicity of decision-feedback decoding while allows for arbitrarily long component codewords and small interleaving delay. Interestingly, trellis STC and layered STC are found to be special cases. In the proposed scheme, a single code $\mathcal{C}$ over $\mathcal{X}$ produces a codeword $\mathbf{c}$ of length N''. This codeword is diagonally interleaved in order to form the $t \times N$ codeword matrix $\mathbf{X} = \mathcal{D}_d(\mathbf{c})$, with $N = N''/t + (t - 1)d$. The diagonal interleaver $\mathcal{D}_d$ is defined by

$$x_{j,n} = \begin{cases} c_{\ell_j(n)} & \text{if } 1 \leq \ell_j(n) \leq N'' \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

for $1 \leq j \leq t$ and $1 \leq n \leq N$, where

$$\ell_j(n) = [n - 1 - (j - 1)d]t + j. \quad (4)$$

The codeword matrix $\mathbf{X}$ is filled by wrapping the codeword $\mathbf{c}$ around its diagonals, as illustrated by Fig. 2. Hence, this STC scheme shall be referred to as *wrapped* space-time code (WSTC). In this way, the interleaving delay d becomes a free parameter, independent of the component codeword block length N''.

As a limiting case, the interleaving delay may be also $d = 0$, i.e., a vertical interleaver may be used. For consistence with the case $d > 0$, where code symbols with lower index take the lower positions in each column of the codeword matrix $\mathbf{X}$

(see Fig. 2), for $d = 0$ we assume that the codeword matrix $\mathbf{X}$ is filled with the elements of the codeword $\mathbf{c}$ by starting from the lower positions of the codeword matrix. In this case, index $\ell_j(n)$ in (3) becomes

$$\ell_j(n) = (n-1)t + t - j + 1 . \qquad (5)$$

Remark 1 If $\mathcal{C}$ is a M-TCM code of rate $R = b/t$, then $\mathcal{S} = \mathcal{D}_0(\mathcal{C})$ is a trellis STC. If $\mathcal{C} = \mathcal{C}^{(1)} \times \cdots \times \mathcal{C}^{(M)}$ (i.e., $\mathcal{C}$ is given by the Cartesian product of codes $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(M)}$), then $\mathcal{S} = \mathcal{D}_d(\mathcal{C}) = \mathcal{L}_d(\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(M)})$ is a layered STC. Hence, WSTC is a generalization of both trellis and layered STC schemes. $\Diamond$

Remark 2 Because of the lower and upper triangles of zero symbols in the codeword matrix defined by (3), there is an inherent rate loss of $(t-1)d/N$. This is negligible if $N \gg td$. Moreover, if the transmission of a long sequence of codewords is envisaged, the codeword matrices can be concatenated in order to fill the leading and tailing triangles of zeros, so that no rate loss is incurred. $\Diamond$

4. Decoding of Wrapped Space-Time Codes

In the case where $\mathcal{C}$ is a trellis code, we propose a reduced-complexity suboptimal decoder obtained by combining decision-feedback interference cancellation with the VA, in analogy with the delayed decision-feedback detection (DDFD) approach for ISI channels with coding (see [10, 11, 12, 13] and references therein). The decoder takes as observable the sequence of samples

$$r_\ell = v_{j,n} - \sqrt{\gamma} \sum_{k=j+1}^{t} b_{j,k} \widehat{x}_{k,n} , \quad \ell = 1, \ldots, N'' \qquad (6)$$

where $v_{j,n}$ and $b_{j,k}$ are defined as in (2) and where $1 \le j \le t$ and $1 \le n \le N$ are the unique integers for which $\ell_j(n) = \ell$. From the index mapping (4) (or (5) for $d = 0$), we see that the elements $x_{k,n}$ for $k = j+1, \ldots, t$ correspond to either zeros (for which no decision is needed) or to symbols of $\mathbf{c}$ with index $\ell' \le \ell - td + 1$ ($\ell' \le \ell - 1$ for $d = 0$). These decisions are found in the survivor history according to *per-survivor processing* (PSP) [6].

   Choice of the linear front-end filter. For the sake of simplicity, the decoder treats the sequence r_ℓ as if it was produced by the *virtual* scalar-input additive-noise channel

$$r_\ell = \sqrt{\beta_j} c_\ell + \nu_\ell, \quad \ell = 1, \ldots, N'' \qquad (7)$$

where ν_ℓ is assumed i.i.d. $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$ and, from (4) or (5), we have that

$$j = \begin{cases} t - |\ell - 1|_t & \text{for } d = 0 \\ |\ell - 1|_t + 1 & \text{for } d > 0 \end{cases}$$

($|\cdot|_t$ denotes a modulo t operation). The SNR β_j of the above channel is given by

$$\beta_j = \frac{\gamma |b_{j,j}|^2}{|\mathbf{f}_j|^2 + \gamma \sum_{k=1}^{j-1} |b_{j,k}|^2} \qquad (8)$$

where $\mathbf{f}_j$ is the j-th column of the front-end filter $\mathbf{F}$. Because of diagonal interleaving, the codeword $\mathbf{c}$ is cyclically interleaved over t virtual additive white Gaussian noise (AWGN) channels with SNRs $\beta_1, \ldots, \beta_t$, so that exactly N''/t symbols are assigned to each channel j. Notice that (8) is the true SNR of

channel j if decisions in (6) are correct, i.e., if the contribution of past symbols is canceled exactly. Moreover, the noise samples ν_ℓ are not Gaussian and not independent, in general. However, provided that these assumptions hold, this scheme decomposes the MIMO channel (1) into t parallel channels with cyclic interleaving.

   Given the analogy between this scheme and decision-feedback equalization of ISI channels, standard choices for the front-end filter matrix $\mathbf{F}$ are also inspired by equalization [14]. If $\mathbf{H}$ has rank t, an information-lossless front-end is given by the WMF $\mathbf{F} = \mathbf{Q}$, where $\mathbf{H} = \mathbf{Q}\mathbf{B}$ is the "QR" decomposition [15] of the channel matrix $\mathbf{H}$, where $\mathbf{Q} \in \mathbb{C}^{r \times t}$ has orthonormal columns and $\mathbf{B} \in \mathbb{C}^{t \times t}$ is upper triangular. In this case, the j-th channel SNR is given by $\beta_j = \gamma |b_{j,j}|^2$ and the additive noise is exactly Gaussian i.i.d. (subject to the assumption of perfect feedback decisions). Another information-lossless front-end is the unbiased MMSE filter whose j-th column $\mathbf{f}_j$ is the solution of the SNR maximization problem

$$\begin{cases} \text{maximize} & \beta_j \\ \text{subject to} & |\mathbf{f}_j|^2 + \gamma \sum_{k=1}^{j-1} |b_{j,k}|^2 = 1 \end{cases}$$

and it is given explicitly by

$$\mathbf{f}_j = \frac{\left[\mathbf{I} + \gamma \sum_{k=1}^{j-1} \mathbf{h}_k \mathbf{h}_k^H \right]^{-1} \mathbf{h}_j}{\sqrt{\mathbf{h}_j^H \left[\mathbf{I} + \gamma \sum_{k=1}^{j-1} \mathbf{h}_k \mathbf{h}_k^H \right]^{-1} \mathbf{h}_j}} \qquad (9)$$

where $\mathbf{h}_j$ denotes the j-th column of $\mathbf{H}$. In this case, the j-th channel SNR is given by

$$\beta_j = \gamma \mathbf{h}_j^H \left[\mathbf{I} + \gamma \sum_{k=1}^{j-1} \mathbf{h}_k \mathbf{h}_k^H \right]^{-1} \mathbf{h}_j \qquad (10)$$

and the additive noise is neither Gaussian nor i.i.d. (even assuming perfect feedback decisions).

   If $\mathbf{H}$ has rank less than t, the WMF is not defined and the MMSE filter is information-lossy. Subject to mild conditions on the statistics of $\mathbf{H}$, the probability that $\mathbf{H}$ has rank less than t when $r \ge t$ is zero. Therefore, the above schemes can be practically applied whenever $r \ge t$. In the rest of this paper we restrict our treatment to this case, and we briefly address the case $r < t$ in Section 6.

5. Code Design for the WSTC

Assuming that the parallel channel model with cyclic interleaving (7) holds, the PEP between two codewords $\mathbf{c}, \mathbf{c}' \in \mathcal{C}$ for given channel SNRs $\beta_1, \ldots, \beta_t$ is given by

$$P(\mathbf{c} \to \mathbf{c}' | \beta_1, \ldots, \beta_t) = Q\left(\sqrt{2 \sum_{j=1}^{t} \beta_j w_j} \right) \qquad (11)$$

where $Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$ is the Gaussian tail function and where we define the *squared Euclidean weight* (SEW) w_j as

$$w_j = \frac{1}{4} \sum_{n=1}^{N} |x'_{j,n} - x_{j,n}|^2 \qquad (12)$$

(the correspondence between symbols of $\mathbf{c}$ and $\mathbf{c}'$ and symbols of $\mathbf{X} = \mathcal{D}_d(\mathbf{c})$ and $\mathbf{X}' = \mathcal{D}_d(\mathbf{c}')$ is given by (3)).

A sensible criterion for the design of the component code $\mathcal{C}$ is to maximize the code *block-diversity* δ, defined by

$$\delta = \min_{\mathbf{c},\mathbf{c}' \in \mathcal{C} : \mathbf{c}' \neq \mathbf{c}} |\{j \in \{1, \ldots, t\} \ : \ w_j \neq 0\}| \qquad (13)$$

that is, to maximize the minimum number of non-zero rows in the matrix difference $\mathbf{D} = \mathbf{X}' - \mathbf{X}$ for each pair of distinct codewords matrices $\mathbf{X}, \mathbf{X}' \in \mathcal{D}_d(\mathcal{C})$. The block-diversity criterion has been investigated in [16, 17, 18] for the design of trellis codes for cyclic interleaving and/or periodic puncturing. The relationship between the rank-diversity of a WSTC and the block-diversity of its component code is given by the following:

Proposition 1 *Consider a code $\mathcal{C}$ over $\mathcal{X}$ of rate R bit/symbol and block-diversity δ. Then, the rank-diversity ρ of the corresponding WSTC $\mathcal{S} = \mathcal{D}_d(\mathcal{C})$ satisfies:*

$$\rho \leq \delta \leq 1 + \left\lfloor t \left(1 - \frac{R}{\log_2 |\mathcal{X}|}\right) \right\rfloor . \qquad (14)$$

Moreover, there exist d for which $\rho = \delta$.

The proof can be found in [19].

Remark 3 From Theorem 3.3.1 of [1], we know that for any STC over $\mathcal{X}$ with t Tx antennas and spectral efficiency $\eta = tR$ the rank-diversity satisfies

$$\rho \leq 1 + \left\lfloor t \left(1 - \frac{R}{\log_2 |\mathcal{X}|}\right) \right\rfloor . \qquad (15)$$

Since this is the same upper bound on block-diversity given in Proposition 1, we get that the wrapping construction incurs no loss of optimality in terms of rank-diversity (for an appropriate choice of the delay d). As a matter of fact, while it is difficult to construct codes with rank-diversity equal to the upper bound (15), it is very easy to find trellis codes for which the upper bound (14) on δ is met with equality, for several coding rates and values of t. Examples of these codes are tabulated in [16, 18]. Therefore, the wrapping construction is a powerful tool to construct STC with maximum rank-diversity. ◊

Remark 4 From Lemma 3.3.1 of [1] we know that a trellis STC with rank-diversity ρ must have constraint length $L \geq \rho$.[2] This constraint does not apply to WSTC. For example, the binary 4-state convolutional code (CC) of rate $1/4$ with generators $(5, 7, 7, 7)$ (octal notation [14]) has constraint length 3, but the corresponding WSTC for $t = 4$ antennas with interleaver delay $d \geq 2$ achieves $\rho = 4$. This fact is explained by noticing that the diagonal interleaver *expands* the state space of the overall interleaved code. On the other hand, the PSP-based VA decoder proposed for WTSCs works on the trellis of the underlying code $\mathcal{C}$, i.e., it ignores the state space expansion.[3] Therefore, it is not a priori clear if WSTCs, even if optimal from the rank-diversity point of view, are going to pay a large penalty when PSP-based VA decoding is used instead of optimal ML decoding. In [19], we show that the penalty incurred by the MMSE front-end and sufficiently large interleaving delay d is practically negligible, while the penalty incurred by the WMF front-end and any d can be very large. ◊

---

[2]We define the constraint length of a trellis code as $L = \nu_{\max} + 1$, where $\nu_{\max}$ is the maximal length of the shift-registers in the canonical feedforward encoder.

[3]Again, we stress the analogy of the problem at hand with the case of trellis coding over a *finite-memory* ISI channel, where the optimal ML decoder requires a number of states generally larger than that of the code alone.
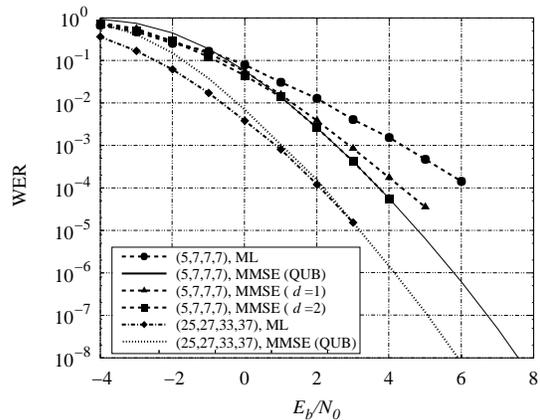

Figure 3: *WER of the WSTCs constructed from the CCs with generators (5,7,7,7) and (25,27,33,37) mapped onto BPSK, independent Rayleigh fading, $t = r = 4$.*

Remark 5 In [20], a computational efficient trellis-based algorithm for computing the block-diversity of trellis codes with cyclic interleaving is given. In [19], we give another computationally efficient method for calculating the block-diversity. By using this method, we computed the block-diversity for the binary codes of rate $1/4$ and $1/2$ tabulated in [14] and mapped onto BPSK and QPSK, for different values of t. Some of these results are reported in [19]. We observe that several codes achieve maximum block-diversity and therefore are good candidates for WSTC. ◊

Example 1 Fig. 3 shows the word-error rate (WER) vs. E_b/N_0 for WTSCs obtained from the binary CC with generators $(5, 7, 7, 7)$ mapped onto BPSK and transmitted over a $t = r = 4$ channel, with $d = 0, 1, 2$. The channel matrix has i.i.d. elements $\sim \mathcal{N}_{\mathbb{C}}(0, 1)$ (independent Rayleigh fading). Each transmitted codeword corresponds to 128 information bits. Decoding is performed by the PSP-based VA working on the underlying 4-state trellis, with MMSE front-end. As d increases, the slope of the WER curve becomes steeper and steeper. In fact, it can be checked that for $d = 0$ the resulting WSTC has $\rho = 2$, for $d = 1$ has $\rho = 3$ and for $d = 2$ has $\rho = 4$. The solid curve denoted as "QUB" is an analytical *Quasi-Upper Bound* developed in [19] assuming the parallel channel model induced by perfect decision-feedback.

For the sake of comparison, we show the WER performance of the trellis STC or, equivalently, WSTC with $d = 0$, obtained by formatting the 16-state binary CC with generators $(25, 27, 33, 37)$ mapped onto BPSK. This code is shown to achieve rank-diversity $\rho = 4$ in [8]. As expected, the slope of the error curve for this code is the same of the 4-state WSTC with $d = 2$. We also show the QUB for the 16-state trellis STC assuming the parallel channel model induced by perfect decision-feedback with MMSE front-end. The fact that the simulated ML curve gets asimptotically very close to the QUB makes us conjecture that the PSP-based VA with MMSE front-end pays almost no penalty with respect to optimal ML decoding. An information-theoretic explanation of this fact is provided in [19]. ◊

6. Performance Examples

In this Section, the WER performance of various WSTCs are assessed for quasi-static fading channels with independent Rayleigh fading. The union bound on the WER derived in [19]
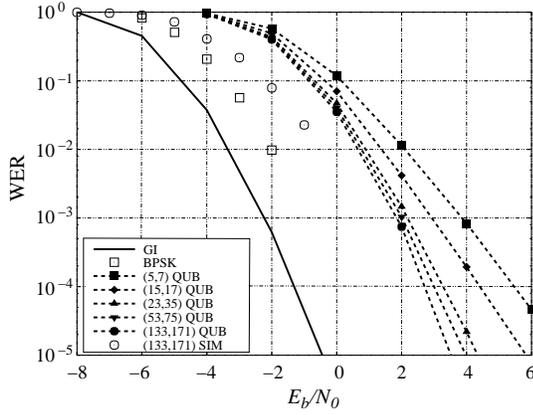
Figure 4: *WER vs.* E_b/N_0 *for WSTCs based on binary CCs of rate* $1/2$ *mapped onto BPSK, independent Rayleigh fading, with* $t = r = 4$.

Figure 5: *WER of WSTC and V-BLAST schemes with code (23,25) and QPSK. Independent Rayleigh fading,* $t = r = 4$. *Spectral efficiency* $\eta = 4$ *bit/channel use.*

is also employed (curves denoted by "QUB") and the results are compared with computer simulation of the full PSP-based VA decoder (curves denoted by "SIM"). In our simulations, each codeword corresponds to 128 information symbols. As a consequence, the codeword length is variable, depending on the code rate.

Results for known convolutional codes. Fig. 4 shows the QUB on the WER for some WSTC based on binary CCs of rate $1/2$ mapped onto BPSK, with $t = r = 4$. Some points obtained by simulating the full PSP-based VA decoder (with interleaving delay $d = 2$) is shown for the sake of comparison. Also, we included the outage probability with Gaussian inputs (GI) [21, 22] and with discrete-inputs (points labeled BPSK) [19]. The difference in the slope of the WER curves for the various codes reflects the different block-diversities. Remarkably, there is still a consistent gap (about 1.5 dB at WER = 10^{-3}) between these simple off-the-shelf codes and the outage curve. This calls for the design of good component codes for the WSTC scheme.

Comparison with V-BLAST. In [7], a simplified space-time decision-feedback detection scheme nicknamed "V-BLAST" is presented. This scheme is equivalent to a WSTC with $d = 0$, but decision-feedback interference cancellation is obtained by feeding back symbol-by-symbol decisions without per-survivor processing. Since the order of decisions is not dictated by the trellis time-ordering of the underlying code, decisions are made in an order that depends on $\mathbf{H}$, in order to limit error propagation in the feedback. Namely, the columns of $\mathbf{H}$ are permuted so that "QR" decomposition of the permuted matrix yields WMF channel "gains" such that $\min_{j=1,\dots,t} |b_{j,j}|^2$ is maximized. As in classical decision feedback-equalization, if the V-BLAST detector is concatenated with a decoder, hard decisions at the detector decision point are made only for decision-feedback purpose, but soft values are fed to the decoder.

Fig. 5 compares the WER vs. E_b/N_0 of a WSTC with either MMSE and WMF front-end obtained from the binary CC with generators $(23, 35)$ mapped onto QPSK, with $t = r = 4$ antennas with the scheme obtained by concatenating the same trellis code with the V-BLAST detector. Simulations of the PSP-based VA decoder for the WSTC (obtained for $d = 2$) are in perfect agreement with the corresponding QUB. The V-BLAST performance was obtained by simulation. For comparison, we show also the WER resulting from a genie-aided V-BLAST detector with ideal feedback decisions (curves denoted by "Id. F."), which is very similar to the WER achieved by the WSTC without genie. This shows that the large perfor-
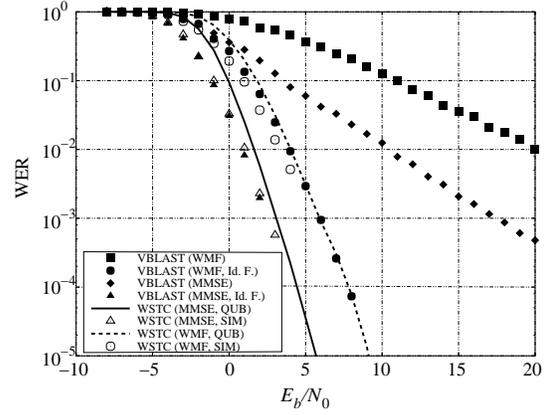
mance degradation of V-BLAST with respect to the WSTC is due to error propagation in the decision feedback, and that the PSP-based decoder is very effective in preventing such propagation while the special detection ordering of V-BLAST is not. Moreover, the complexity of the detection ordering algorithm is larger than the extra complexity of the VA due to PSP, for large t. Then, WSTC with PSP-based VA decoding is not only more effective, but might be also simpler than V-BLAST.

Handling the case $r < t$ **via LD precoding.** When $r < t$ the low-complexity PSP-based decoding scheme cannot be applied. Recently, Hochwald and Hassibi proposed a scheme called "linear dispersion" (LD) coding [5]. This scheme takes blocks of Q modulation symbols and map them onto the complex $t \times T$ matrix signals

$$\mathbf{S} = \sum_{q=1}^{Q} (x_q \mathbf{C}_q + x_q^* \mathbf{D}_q)$$

where $\mathbf{C}_q$ and $\mathbf{D}_q$ are complex $t \times T$ matrices defining the LD code. Then, $\mathbf{S}$ is transmitted column-by-column over T channel uses. The resulting spectral efficiency is $\eta = \frac{Q}{T}R$, where R is the rate of the outer code. We can think of LD coding as a precoder which shapes the original $t \times r$ complex MIMO channel into a virtual $2Q \times 2rT$ real MIMO channel. As long as Q and T are chosen such that $Q \leq rT$, WSTC with PSP-based decoding can be applied as an outer coding/decoding scheme to the LD-precoded channel.

As an example, Fig. 6 shows the WER vs. E_b/N_0 of a coding scheme obtained by concatenating the WSTC obtained from the binary CC with generators (23,35) mapped onto QPSK with a LD precoder designed in [5] for the $t = 4$, $r = 1$ channel, with $Q = 4$ and $T = 4$. The resulting spectral efficiency is $\eta = 1$ bit/channel use. For the sake of comparison, we show also the outage curve with GI on the original 4×1 channel (curve denoted by "No prec.") and the outage curve with GI of the LD-precoded channel. We notice that the gap between the GI outage of the original 4×1 channel and the WER of the concatenated WSTC with LD-precoding is almost entirely due to the LD precoder, since the simulated WER with PSP-based VA decoding is less than 1.5 dB away from the outage of the LD-precoded channel, while the gap between LD-precoded and original channels is about 5 dB for WER $\approx 10^{-4}$.

The LD precoders in [5], included the one used in this example, where designed in order to maximize the average mutual information at a given SNR, and not to minimize the outage
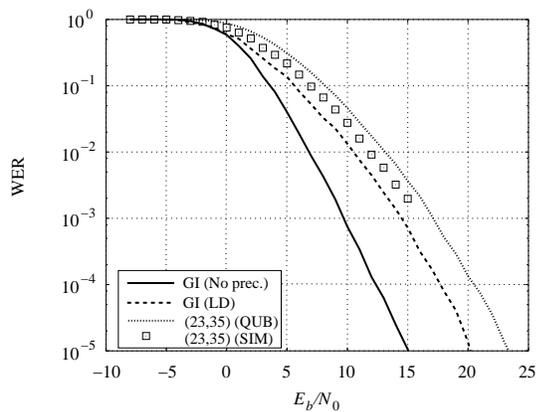
Figure 6: *WER of WSTC with code (23,25) and QPSK concatenated with LD-precoding. Independent Rayleigh fading, $t = 4$, $r = 1$. Spectral efficiency $\eta = 1$ bit/channel use.*

probability for a given spectral efficiency. We believe that a better WER performance for the LD-precoded channel can be achieved by explicitly designing the precoder in order to minimize the outage probability.

7. Conclusions

A new scheme, nicknamed "wrapped" STC was proposed. This scheme generalizes both trellis and layered STCs and it is suited to a large number of antennas and low complexity decoding, based on MMSE decision-feedback coupled with PSP-based Viterbi Decoding. We showed that any trellis codes with maximal block-diversity can be turned into a WSTC with maximal rank-diversity, with the advantage that block-diversity is easy to check and maximal block-diversity is easily achieved by several well-known trellis codes. We also showed via numerical experiments that the MMSE decision-feedback receiver coupled with the PSP-based VA, providing very reliable decisions, performs very close to optimal ML decoding.

Performance examples of the proposed scheme constructed from well-known binary linear convolutional codes were provided. WSTCs compare very favorably with respect to coded V-BLAST of similar complexity. In the case $r < t$, where the decision-feedback scheme cannot be used directly, our scheme is naturally suited to work as outer code where the inner code (pre-coder) is a linear dispersion code as studied in [5].

8. References

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

[2] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Syst. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[3] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. on Inform. Theory*, vol. 45, no. 5, pp. 1456, July 1999.

[4] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis-coded modulation," *IEEE Comm. Letters*, vol. 4, no. 1, pp. 20–22, January 2000.

[5] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," submitted to IEEE Trans. on Inform. Theory, August 2000.

[6] R. Raheli, A. Polydoros, and C.-K. Tzou, "Per-survivor processing: A general approach to MLSE in uncertain environments," *IEEE Trans. on Commun.*, vol. 43, no. 2, pp. 354–364, February 1995.

[7] J. Foschini, G. Golden, R. Valenzuela, and P. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, no. 11, pp. 1841–1852, November 1999.

[8] R. Hammons and H. El Gamal, "On the theory of space-time codes for PSK modulation," *IEEE Trans. on Inform. Theory*, vol. 46, no. 2, pp. 524–542, March 2000.

[9] D.-S. Shiu and J. Kahn, "Layered space-time codes for wireless communications using multiple transmit antennas," in *IEEE Intern. Conf. on Commun., ICC '99*, Vacouver, Canada, June 1999.

[10] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. on Commun.*, vol. 37, no. 5, pp. 428–436, May 1989.

[11] R. Ashkenazi and S. Shamai, "Suboptimal detection for intersymbol interference inflicted cable channels," *AEU*, vol. 51, no. 5, pp. 246–254, May 1997.

[12] V. Eyuboglu and S. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. on Commun.*, vol. 36, no. 1, pp. 13–20, January 1988.

[13] P. Chevillat and E. Eleftheriou, "Decoding of trellis-encoded signals in the presence of intersymbol interference and noise," *IEEE Trans. on Commun.*, vol. 37, no. 7, pp. 669–676, July 1989.

[14] J. Proakis, *Digital communications, 3rd Ed.*, McGraw-Hill, New York, 1995.

[15] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, New York, 1985.

[16] R. Knopp and P. Humblet, "On coding for block fading channels," *IEEE Trans. on Inform. Theory*, vol. 46, no. 1, pp. 189–205, January 2000.

[17] E. Malkamaki and H. Leib, "Evaluating the performance of convolutional codes over block fading channels," *IEEE Trans. on Inform. Theory*, vol. 45, no. 5, pp. 1643–1646, July 1999.

[18] R. Wesel, X. Liu, and W. Shi, "Trellis codes for periodic erasures," *IEEE Trans. on Commun.*, vol. 48, no. 6, pp. 938–947, June 2000.

[19] G. Caire and G. Colavolpe, "On space-time coding for quasi-static multiple-antenna channels," submitted to IEEE Trans. on Inform. Theory, March 2001.

[20] A. Lapidoth, "The performance of convolutional codes on the block erasure channel using various finite interleaving techniques," *IEEE Trans. on Inform. Theory*, vol. 40, no. 5, pp. 1459–1473, September 1994.

[21] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm. ETT*, vol. 10, no. 6, pp. 585–596, November 1999.

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