A suboptimum iterative decoder for space-time trellis codes

Alberto Tarable, Guido Montorsi and Sergio Benedetto
CERCOM, Dipartimento di Elettronica e delle Telecomunicazioni,
Politecnico di Torino, Italy
email: {tarable, montorsi, benedetto}@polito.it

Abstract—The main problem of space-time trellis codes is constituted by their complexity, which grows exponentially with the number of transmit antennas. To avoid this shortcoming, one can think to a suboptimum decoder in which a simplified metric, together with a preliminary filtering step, is used. In this paper, we develop this idea and give some possible choices for the filter design. These different choices are compared with one another by means of analytical tools and simulations.

In two independent papers, Telatar [6] and Foschini [3] have shown that, for high signal-to-noise ratio, the capacity of a multi-input multi-output (MIMO) fading channel with transmit and receive antennas grows linearly with min (r, t). Since then, space-time codes (STCs), i.e., the coding schemes for MIMO channels, have received considerable attention. In the family of STCs, block STCs (STBCs), like the Alamouti code, have been until now more popular in applications, because of their low decoding complexity. Instead, trellis STCs (STTCs) have a decoding complexity that is exponential with the number of transmit antennas. To avoid this shortcoming, one can think to a suboptimum decoder in which a simplified metric, first interleaved according to permutation \( \pi \), and then sent to the \( j \)-th transmit antenna. The symbols transmitted from each antenna belong to a constellation \( S = \{ s_0, s_1, ..., s_{2^m-1} \} \) with size \( 2^m \) and average energy \( E_s \). The transmitted signal can then be represented by a \( t \times L \) complex matrix \( X \) with elements in \( S \), where \( L \) is the block length. See Fig. 1 for a picture of the encoder. The received signal can be written as:

\[
R = HX + Z,
\]

where \( H = [h_{ij}] \) is the \( r \times t \) complex matrix of the channel realization. In our model, the \( h_{ij} \)'s are independent complex random variables with Rayleigh-distributed modulus and uniformly-distributed phase. We assume that the fading is slow, i.e., it remains constant for the whole block length, and that the receiver has perfect knowledge of the channel coefficients. The \( r \times L \) complex matrix \( Z \) is the Gaussian additive noise, whose elements are independent circular complex Gaussian random variables with zero mean and variance of the real and imaginary part \( \sigma^2/2 \).

At the receiver side, the received signal is first passed through a maximal-ratio combiner (MRC) whose output is:

\[
Y = H^H R = H^H HX + N,
\]

where \( N = H^H Z \) is colored Gaussian noise. In the \( i \)-th symbol interval, we can write:

\[
y[i] = H^H Hx[i] + n[i],
\]
where \( y[i], \, x[i] \) and \( n[i] \) are the \( i \)-th columns of \( Y \), \( X \) and \( N \), respectively.

The MRC output enters then the iterative decoder. Let us consider the \( t \)-th iteration in the decoding process. The SISO-AS has two inputs, the MRC output and the feedback from the SISO trellis decoder, which was computed in the \((l-1)\)-th iteration. We refer to the latter as the current symbol statistics, in the sense that the SISO-AS interprets it as an \( a \) priori knowledge on the symbols emitted by each antenna. For the \( j \)-th antenna and \( i \)-th symbol interval, there are \( 2^n - 1 \) inputs:

\[
LLR_j^{(i)}[i, n] = \log \frac{Pr \{ x_j[i] = s_n \}}{Pr \{ x_j[i] = \bar{s}_0 \}} \quad n = 1, ..., 2^n - 1,
\]

where \( LLR \) stands for log-likelihood ratio and \( x_j[i] \) is the \( j \)-th element of \( x[i] \). In the first iteration, when there is no a priori knowledge on the transmitted symbols, the \( a \) priori LLRs are all set to zero. We say that the SISO-AS is linear if:

- the MRC output is linearly filtered in the following way:
  \[
  \tilde{y}^{(t)}[i] = M^{(t)}[i]y[i],
  \]
  where \( M^{(t)}[i] \) is a \( t \times t \) complex matrix, and
- the output of the SISO-AS for the \( j \)-th antenna and \( i \)-th symbol interval is a set of \( 2^n - 1 \) LLRs, computed element-by-element from \( \tilde{y}^{(t)}[i] \):
  \[
  LLR_j^{(i)}[i, n] = \log \frac{Pr \{ \tilde{y}_j^{(t)}[i] \mid x_j[i] = s_n \}}{Pr \{ \tilde{y}_j^{(t)}[i] \mid x_j[i] = \bar{s}_0 \}},
  \]
  \[
  n = 1, ..., 2^n - 1,
  \]
  where \( \tilde{y}_j^{(t)}[i] \) is the \( j \)-th element of \( \tilde{y}^{(t)}[i] \).

We remark that the concept of linearity only involves the way the MRC output is processed to obtain \( \tilde{y}^{(t)}[i] \), and is not related with the operations involved in (6). From (5) and (3), we can write

\[
\tilde{y}_j^{(t)}[i] = \beta_j^{(t)}[i] x_j[i] + \sum_{j' \neq j} \beta_{jj'}^{(t)}[i] x_{j'}[i] + \tilde{n}_j^{(t)}[i],
\]

where

\[
\beta_{jj'}^{(t)}[i] = \left( M^{(t)}[i] H_i H_i^H \right)_{jj'}
\]

and \( \tilde{n}_j^{(t)}[i] \sim \mathcal{N} \left( 0, \tilde{\sigma}_j^{(t)}[i] \right) \) is complex Gaussian noise, being

\[
\tilde{\sigma}_j^{(t)}[i] = \sigma^2 \left( M^{(t)}[i] H_i H_i^H M^{(t)}[i] H_i^H \right)_{jj'}.
\]

Since the exact computation of the probabilities in (6) would be exponentially complex with \( t \), the so-called Gaussian approximation (GA) is invoked, which consists in approximating the sum (residual IAI + noise) with a Gaussian random variable with the same mean and variance. The GA consists in giving an approximate expression to \( \tilde{y}_j^{(t)}[i] \):

\[
\tilde{y}_j^{(t)}[i] \simeq \nu_j^{(t)}[i] x_j[i] + \rho_j^{(t)}[i],
\]

where \( \nu_j^{(t)}[i] \sim \mathcal{N} \left( \mu_j^{(t)}[i], \sigma_j^{(t)}[i] \right) \), \( \mu_j^{(t)}[i] \) and \( \rho_j^{(t)}[i] \) being mean and variance, respectively, of the (IAI + noise) term in (7). Assuming that the current symbol statistics for different antennas are independent, it can be easily shown that:

\[
\mu_j^{(t)}[i] = \sum_{j' \neq j} \beta_{jj'}^{(t)}[i] x_{j'}[i],
\]

\[
\rho_j^{(t)}[i] = \tilde{\sigma}_j^{(t)}[i] + \sum_{j' \neq j} \beta_{jj'}^{(t)}[i] \sigma_{j'}^{(t)}[i],
\]

where

\[
\frac{1}{n} \sum_{n=0}^{2^n-1} e^{LLR_j^{(i)}[i, n]} s_n \]

is the current average of \( x_j[i] \) and

\[
\text{var}(t)[i] = E \left[ \left| x_j[i] \right|^2 \mid t \text{-th iteration} \right] - \left| \tilde{z}_j^{(t)}[i] \right|^2
\]

is its variance.

Invoking the GA, from (10) and (6), the LLRs output by the SISO-AS will be given by (15), at the top of next page.

The output LLRs for the \( j \)-th antenna are then deinterleaved according to permutation \( \pi^{-1} \) and sent to the SISO trellis decoder. This block performs the so-called BCJR or forward-backward algorithm [1] and, after interleaving, the current symbol statistics for the next iteration \( LLR_j^{(i+1)}[i, n] \) (see (4)) is obtained by subtracting from the decoder output the corresponding input \( LLR_j^{(i)}[i, n] \). Then, another iteration starts. In the last iteration, the SISO trellis decoder must also supply a hard estimate of the information bits, which represents the output of the whole space-time decoder. See Fig. 2 for a picture of the iterative decoder.

Although suboptimal, this decoder has the strong advantage that it permits to introduce an interleaver for each antenna, as shown in Fig. 1, which often results in a coding gain also w.r.t. to the optimal decoder without the interleavers [2]. Different iterative decoders can be obtained by choosing different linear SISO-ASs, depending on \( M^{(t)}[i] \) in (5). In the next section, we will describe some possible choices.
II. DESIGN OF THE LINEAR SISO-AS

A. WP SISO-AS

One possible choice for $M^{(l)}[i]$ in (5) is based on a work by Wang and Poor on turbo multiuser detection [7]. If we define $m_k^{(l)}[i]$ as the $k$-th row of $M^{(l)}[i]$, it consists in the solution to the following minimum mean-square-error (MMSE) problem:

$$m_k^{(l)}[i] = \arg \min_{m \in \mathbb{C} \times K} E \left[ \left| x_k[i] - m \left( y[i] - \mathbf{H} \mathbf{x}_k^{(l)}[i] \right) \right|^2 \right],$$

(16)

where $\mathbf{x}_k^{(l)}[i]$ is the average value of IAI seen by the signal coming from the $k$-th transmit antenna:

$$\mathbf{x}_k^{(l)}[i] = \mathbf{H}^H \mathbf{x}_k^{(l)}[i],$$

(17)

where $\mathbf{x}_k^{(l)}[i]$ is a $l$-long vector, whose $j$-th element is

$$\left( \mathbf{x}_k^{(l)}[i] \right)_j = \left\{ \begin{array}{ll} x_k[i], & j \neq k \\ 0, & j = k \end{array} \right.,$$

(18)

This SISO-AS will be called hereafter WP SISO-AS. The solution to the MMSE problem in (16) is given by [7]:

$$m_k^{(l)}[i] = \mathbf{e}_k^T \left( \mathbf{V}_k^{(l)}[i] + \sigma^2 (\mathbf{H}^H \mathbf{H})^{-1} \right)^{-1} \left( \mathbf{H}^H \mathbf{y} \right)^{-1},$$

(19)

where

$$\mathbf{V}_k^{(l)}[i] = \text{diag} \left( \text{var} \left( x_1[i] \right), \ldots, \text{var} \left( x_{k-1}[i] \right), \right.$$  
$$E_s, \text{var} \left( x_{k+1}[i] \right), \ldots, \text{var} \left( x_l[i] \right) \bigg),$$

(20)

and the variance of all interfering symbols is computed on the basis of the a priori information fed back from the SISO trellis decoder.

In the first iteration, $\mathbf{x}_k^{(1)}[i] = 0$ and $\mathbf{V}_k^{(1)}[i] = E_s \mathbf{I}$, because there is no a priori information on the transmitted symbols. Then (apart from unessential constants):

$$m_k^{(1)}[i] = \mathbf{e}_k^T (\mathbf{H}^H \mathbf{H} + \delta_s \mathbf{I})^{-1},$$

(21)

being $\delta_s = \sigma^2/E_s$, which is equal to the filter introduced in [2].

In the asymptotic case in which IAI is perfectly known, $\mathbf{V}_k^{(\infty)}[i] = E_s \mathbf{e}_k \mathbf{e}_k^T$, and (19) becomes:

$$m_k^{(\infty)}[i] = \mathbf{e}_k^T,$$

(22)

i.e. no filtering at all. Since the MMSE filter maximizes the signal-to-(interference+noise) ratio (SINR) at the filter output, the WP SISO-AS can be considered in this sense the best linear SISO-AS. Unfortunately, it needs one size-$l$ matrix inversion

(1)This is an alternative description to the one given in [7], in which soft interference cancellation is performed before filtering. The two operations can be easily exchanged, provided that the IAI is computed accordingly. Put in this way, it fits into our general framework for linear SISO-ASs

(see (19)) for every antenna and every symbol interval, at each iteration. Thus, it has a complexity $O(l^2)$ per antenna per decoded symbol per iteration.

B. S SISO-AS and derivatives

A simpler strategy can be devised following the same ideas of [4]. When the reliability on the interfering symbols is low, i.e., in the first few iterations of the space-time decoder, the filter is a fixed MMSE filter as in (21). When the current symbol statistics is supposedly good, i.e., after a certain number of iterations, there is no filtering at all, like in (22). In formulas:

$$m_k^{(l)}[i] = \left\{ \begin{array}{ll} \mathbf{e}_k^T \left( \mathbf{H}^H \mathbf{H} + \delta_s \mathbf{I} \right)^{-1}, & l \leq m_k[i] \\ \mathbf{e}_k^T, & l > m_k[i] \end{array} \right.,$$

(23)

where $m_k[i]$ is an integer representing the switching iteration for antenna $k$ at symbol time $i$. We will refer to this structure as the switched SISO-AS (S SISO-AS).

Since the MMSE filter in (26) can be computed once and for all, and is the same for all antennas, unlike in the WP SISO-AS, which requires a different filter for every antenna, time instant, iteration, the S SISO-AS is much simpler than the WP SISO-AS. Its complexity per decoded symbol per antenna per iteration grows only linearly with $t$.

Notice also that, while the filter matrix changes only once along the iterations, the output LLRs change at every iteration, because the current symbol statistics, which is updated by the SISO trellis decoder at each iteration, determines the quantities $\mu_j^{(l)}[i]$ and $\rho_j^{(l)}[i]$ in (10).

The choice of the switching iteration is a degree of freedom that can be used to find a trade-off between performance and complexity. We can further differentiate according to the switching strategy:

- **SINR-maximizing strategy** [4]: The S SISO-AS switches the first time that the output SINR is greater for no filtering than for MMSE filtering. In formulas, we have, with the symbols introduced in the previous section (see (10)):

$$m_k[i] = \min \left\{ l : \frac{\beta_{kk, \text{MRC}}^{(l)} \rho_{kk,MRC}^{(l)}}{\beta_{kk, \text{MMSE}}^{(l)} \rho_{kk, \text{MMSE}}^{(l)}} > \frac{\beta_{kk, \text{MRC}}^{(l)} \rho_{kk, \text{MRC}}^{(l)}}{\beta_{kk, \text{MMSE}}^{(l)} \rho_{kk, \text{MMSE}}^{(l)}} \right\},$$

(24)

where the subscript MRC refers to the filter matrix in (22), while the subscript MMSE refers to the filter in (26). This can be considered to be the best among the known switching strategies. Since the quantities in (24) must be computed anyway, the additional complexity needed to implement such strategy is small.

- **Fixed switching strategy**: In this case, the switching iteration is fixed to some predetermined value:

$$m_k[i] \equiv m, \forall k, i$$

(25)
which may range from 0 to \( \infty \).

If \( m_k[i] = \infty \), i.e., if the S SISO-AS never switches, we have the SISO-AS described in [2], which always filters according to (21), called MMSE SISO-AS in the following:

\[
{\mathbf{m}}^{(l)}_{k,MMSE}[i] = {\mathbf{e}}^T_k \left( {\mathbf{H}}^H {\mathbf{H}} + \delta_l {\mathbf{I}}_s \right)^{-1}
\]

which does not depend on the symbol interval index \( i \), nor on the iteration number \( l \).

Instead, if \( m_k[i] = 0 \), there is no filtering at all, not even in the first iterations:

\[
{\mathbf{m}}^{(l)}_{k,MRC}[i] = {\mathbf{e}}^T_k,
\]

Because the LLRs are computed directly from the MRC output, this will be called the MRC SISO-AS.

C. Performance analysis

To compare the different space-time decoders described above, we will compute their asymptotic performance when the iterations grow to infinity and for \( \delta_s \to 0 \). When the WP SISO-AS is used, for \( l \to \infty \), the filter tends to \( {\mathbf{e}}^T_k \), as shown in (22), IAI is perfectly cancelled when extracting the LLR, and the instantaneous SNR for the \( k \)-th antenna is then

\[
\text{SNR}_k \left( \frac{\left( {\mathbf{H}}^H {\mathbf{H}} \right)^{-1}}{\delta_s} \right) \downarrow \text{SNR}_{k,MRC}.
\]

The same is true for the S SISO-AS, because it behaves like the WP SISO-AS for \( l \to \infty \) (see (23)). Instead, the MMSE SISO-AS, for \( \delta_s \to 0 \), tends to

\[
{\mathbf{m}}^{(l)}_{k,MMSE}[i] \to {\mathbf{e}}^T_k \left( {\mathbf{H}}^H {\mathbf{H}} \right)^{-1},
\]

i.e., it completely suppress IAI. The instantaneous SNR for the \( k \)-th antenna will then be

\[
\text{SNR}_k \left( \frac{1}{\left( {\mathbf{H}}^H {\mathbf{H}} \right)^{-1}} \right) \downarrow \text{SNR}_{k,MMSE}.
\]

By using the matrix inversion lemma, we can show that:

\[
\text{SNR}_{k,MRC} > \text{SNR}_{k,MMSE},
\]

i.e., the WP SISO-AS and the S SISO-AS have better asymptotic performance that the MMSE SISO-AS.

III. Simulations

In this section, we show some simulation results, depicting the performance of the different receivers described in the previous section.

In the first simulation, the STTC used is obtained from the two-antennas 4-state rate-1/2 STTC in [5] by interleaving the symbol stream at the second antenna. The interleaver used is a spread interleaver, which has been kept fixed for all simulations. The block length at the encoder input is equal to 128 bits. Fading is fixed during a single frame, while fading coefficients in different frame intervals are independent. At the receiver, there are two receive antennas. Five different SISO-AS are used. The first is the WP SISO-AS, whose filter is reported in (19). The second is the S SISO-AS (see (23)), where the switching iteration is chosen according to the SINR-maximization criterion. The third is the S SISO-AS again, but the switching iteration is fixed, i.e., \( m_k[i] = 5, \forall k, i \), in (23). The fourth is the MMSE SISO-AS of (26). The fifth is the MRC SISO-AS, is of (27).

The results are shown in Fig. 3. The dashed lines show the performance after the first iteration. The solid lines show the performance after ten iterations. It can be seen that the best SISO-AS, as it could be expected, is the WP SISO-AS. However, the S SISO-AS with the SINR-maximizing switching strategy has practically the same performance. The S SISO-AS with the fixed switching strategy, instead, loses up to 2 dB at the tenth iteration, for bit error rate (BER) between \( 10^{-4} \) and \( 10^{-5} \). The MMSE SISO-AS loses 2 dB for BER = \( 10^{-4} \). It can be seen that the MRC SISO-AS, which is the worst in the first iteration already, is asymptotically the worst one, losing almost 4 dB for BER = \( 10^{-4} \).

Fig. 4 shows the impact of the GA in this two-antenna scenario. The performance of the S SISO-AS with \( m_k[i] = 5, \forall k, i \), is compared with the same S SISO-AS, when LLRs are extracted without the GA, i.e., computing (6) with the exact distribution of \( {\mathbf{y}}^T[i] \) in (7). It can be seen that the GA has a strong impact on the receiver performance. The GA-based receiver loses 2 dB for BER = \( 10^{-4} \) w.r.t. to the receiver without the GA, and the gap seems to increase with higher SNR. This can be explained with the fact that there are only two transmit antennas, so that IAI contains the contributions of a single antenna. The loss of the GA is likely to decrease with increasing number of transmit antennas.

IV. Conclusions

The design of a linear interface (the antenna separator) for the suboptimum iterative decoding of space-time trellis codes has been dealt with in this paper. Several choices, which are
different in complexity and performance, have been compared through simulations.

It is worth noting that, because of interleaving on each antenna, the structure of the trellis code is lost. This means that new design rules for the space-time code have to be found. The overall architecture seems to suggest the adoption of a pragmatic approach, in which a binary trellis code with high free distance is mapped on the constellation points. However, this has not been treated in this paper and deserves further investigation in a future work.

ACKNOWLEDGMENT

This work has been partially funded by the Italian Ministry of Education and Research under CERCOM and FIRB-PRIMO funds.

REFERENCES