Linear Convolutive Space-Time Precoding for Spatial Multiplexing MIMO Systems

Abdelkader Medles, Dirk T.M. Slock

Eurecom Institute *
2229 route des Crêtes, B.P. 193
06904 Sophia Antipolis Cedex
FRANCE
Email: {medles,slock}@eurecom.fr

Abstract

The use of multiple transmitter and receiver antennas allows to transmit multiple signal streams in parallel and hence to increase communication capacity. To distribute the multiple signal streams over the MIMO channel, linear space-time codes have been shown to be a convenient way to reach high capacity gains with a reasonable complexity. The space-time codes that have been introduced so far are block codes, leading to the manipulation of possibly large matrices. To reduce complexity, we propose to introduce convolutive codes, associated with MIMO filters. We investigate capacity, error exponents, matched filter bounds and diversity for convolutive MIMO channels, with linear space-time coding systems based on MIMO filters. We consider full rate and lower rate coding systems for various scenarios of number of TX antennas versus number of RX antennas. Combined convolutive/block coding schemes are also introduced.

1 Introduction

Spatial multiplexing has been introduced independently in a 1994 Stanford University patent by A. Paulraj and by Foschini [1] at Bell Labs. Spatial multiplexing can be viewed as a limiting case of Spatial Division Multiple Access (SDMA) in which the various mobile users are colocated in one single user multi antenna mobile terminal. In that case, the various users are no longer distinguishable on the basis of their (main) direction (DOA) since all antennas are essentially colocated. Nevertheless, if the scattering environment is sufficiently rich, the antenna arrays at TX and RX can see the different DOAs of the multiple paths. One can then imagine transmitting multiple data streams, one stream per path. For this, the set of paths to be used should be resolvable in angle at both TX and RX. Without channel knowledge at the TX, the multiple streams to be transmitted just get mixed over the multiple paths in the matrix channel. They can generally be linearly recovered at the RX if the channel matrix rank equals or exceeds the number

*Eurécom's research is partially supported by its industrial partners: Ascom, Swisscom, France Télécom, La Fondation CEGETEL, Bouygues Télécom, Thales, ST Microelectronics, Motorola, Hitachi Europe and Texas Instruments. The work reported here was also sponsored in part by the RNRT (French National Network for Research in Telecommunications) project ERMITAGES.

of streams. This rank equals the number of paths that are simultaneously resolvable at TX and RX. The assumptions we shall adopt for the proposed approach are no channel knowledge at TX, perfect channel knowledge at RX, frequency-flat channels for most of the paper, full rate transmission ($N_s = N_{tx}$), $N_{rx} \geq N_{tx}$ such that the rank of the channel possibly equals the number of streams N_s.

2 Linear prefiltering approach

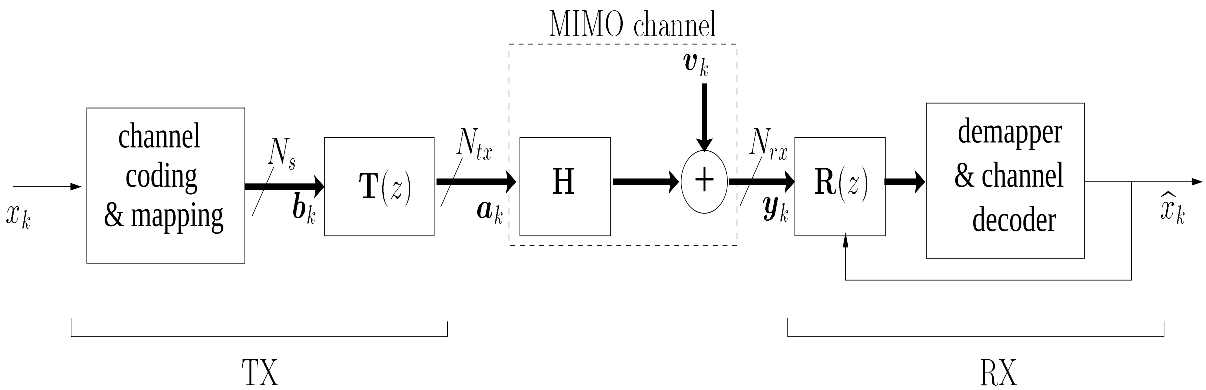

Figure 1: General ST coding setup.

A general ST coding setup is sketched in Fig. 1. The incoming stream of bits gets transformed to N_s symbol streams through a combination of channel coding, interleaving, symbol mapping and demultiplexing. The result is a vector stream of symbols $\boldsymbol{b}_k$ containing N_s symbols per symbol period. The N_s streams then get mapped linearly to the N_{tx} transmit antennas and this part of the transmission is called linear ST precoding. The output is a vector stream of symbols $\boldsymbol{a}_k$ containing N_{tx} symbols per symbol period. The linear precoding is spatiotemporal since an element of $\boldsymbol{b}_k$ may appear in multiple components (space) and multiple time instances (time) of $\boldsymbol{a}_k$. The vector sequence $\boldsymbol{a}_k$ gets transmitted over a MIMO channel $\mathbf{H}$ with N_{rx} receive antennas, leading to the symbol rate vector received signal $\boldsymbol{y}_k$ after sampling. The linear precoding can be considered to be an inner code, while the nonlinear channel coding etc. can be considered to be an outer code. As the number of streams is a factor in the overall bitrate, we shall call the case $N_s = N_{tx}$ the full rate case, while $N_s = 1$ corresponds to the single rate case. Instead of multiple antennas, more general multiple channels can be considered by oversampling, by using polarization diversity or other EM component variations, by working in beamspace, or by considering in phase and in quadrature (or equivalently complex and complex conjugate) components. In the case of oversampling, some excess bandwidth should be introduced at the transmitter, possibly involving spreading which would then be part of the linear precoding.

As we shall see below, channel capacity can be attained by a full rate system without precoding ($\mathbf{T}(z) = I$). In that case, the channel coding has to be fairly intense since it has to spread the information contained in each transmitted bit over space (across TX antennas) and time, see the left part in Fig. 2 and [13]. The goal of introducing the linear precoding is to simplify (possibly going as far as eliminating) the channel coding part [5]. In the case of linear dispersion codes [7],[8], transmission is not continuous but packet-wise (block-wise). In that case, a packet of T vector symbols $\boldsymbol{a}_k$ (hence a $N_{tx} \times T$ matrix) gets constructed as a linear combination of fixed matrices in which the combination coefficients are symbols b_k. A particular case is the Alamouti code which is a full diversity single rate code corresponding to block length $T = N_{tx} = 2,\ N_s = 1$.

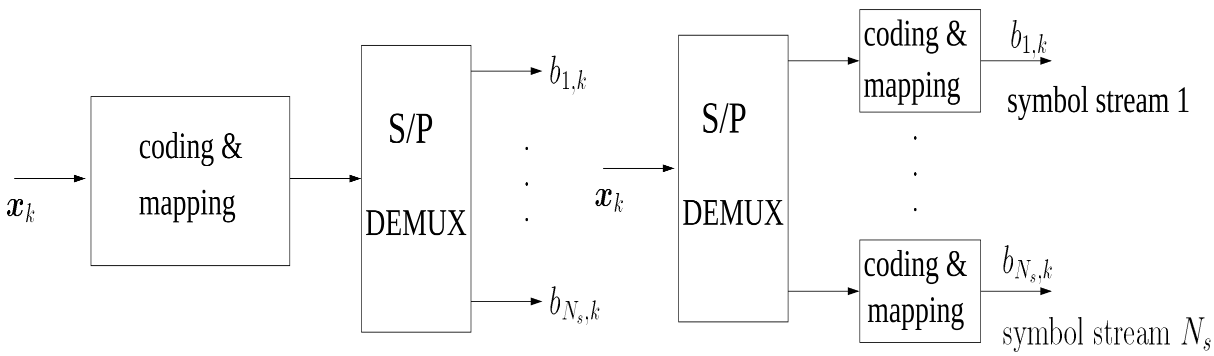

Figure 2: Two channel coding, interleaving, symbol mapping and demultiplexing choices.

In this paper we shall focus on continuous transmission in which linear precoding corresponds to MIMO prefiltering. This linear convolutive precoding can be considered as a special case of linear dispersion codes (making abstraction of the packet boundaries) in which the fixed matrices are time-shifted versions of the impulse responses of the columns of $\mathbf{T}(z)$, see Fig. 1. Whereas in the absence of linear precoding, the last operation of the encoding part is spatial demultiplexing (serial-to-parallel (S/P) conversion) (see left part of Fig. 2), this S/P conversion is the first operation in the case of linear precoding, see the right part of Fig. 2. After the S/P conversion, we have a mixture of channel coding, interleaving and symbol mapping, separately per stream. The existing BLAST systems are special cases of this approach. VBLAST is a full rate system with $\mathbf{T}(z) = \mathbf{I}_{N_{tx}}$ which leads to quite limited diversity. DBLAST is a single rate system with $\mathbf{T}(z) = [1\ z^{-1}, \ldots, z^{-(N_{tx}-1)}]^T$ which leads to full diversity (delay diversity). We would like to introduce a prefiltering matrix $\mathbf{T}(z)$ without taking a hit in capacity, while achieving full (spatial) diversity. The MIMO prefiltering will allow us to capture all diversity (spatial, and frequential for channels with delay spread) and will provide some coding gain. The optional channel coding per stream then serves to provide additional coding gain and possibly (with interleaving) to capture the temporal diversity (Doppler spread) if there is any. Finally, though time-invariant filtering may evoke continuous transmission, the prefiltering approach is also immediately applicable to block transmission by replacing convolution by circulant convolution.

3 Capacity

Consider the MIMO AWGN channel

$$\boldsymbol{y}_k = \mathbf{H}\,\boldsymbol{a}_k + \boldsymbol{v}_k = \mathbf{H}\,\mathbf{T}(q)\,\boldsymbol{b}_k + \boldsymbol{v}_k \tag{1}$$

where the noise power spectral density matrix is $S_{\boldsymbol{vv}}(z) = \sigma_v^2\, I,\ q^{-1}\boldsymbol{b}_k = \boldsymbol{b}_{k-1}$. The **ergodic capacity** when channel knowledge is absent at the TX and perfect at the RX is given by:

$$\begin{aligned} C(S_{\boldsymbol{aa}}) &= \mathrm{E}_H \tfrac{1}{2\pi j} \oint \tfrac{dz}{z} \ln\det(I + \tfrac{1}{\sigma_v^2}\mathbf{H}\,S_{\boldsymbol{aa}}(z)\,\mathbf{H}^H) \\ &= \mathrm{E}_H \tfrac{1}{2\pi j} \oint \tfrac{dz}{z} \ln\det(I + \tfrac{1}{\sigma_v^2}\mathbf{H}\,\mathbf{T}(z)\,S_{\boldsymbol{bb}}(z)\,\mathbf{T}^\dagger(z)\,\mathbf{H}^H) \\ &= \mathrm{E}_H \tfrac{1}{2\pi j} \oint \tfrac{dz}{z} \ln\det(I + \rho\,\mathbf{H}\,\mathbf{T}(z)\,\mathbf{T}^\dagger(z)\,\mathbf{H}^H) \end{aligned} \tag{2}$$

where we assume that the channel coding and interleaving per stream leads to spatially and temporally white symbols: $S_{\boldsymbol{bb}}(z) = \sigma_b^2 I$, and $\rho = \frac{\sigma_b^2}{\sigma_v^2}$. The expectation E_H is here w.r.t. the distribution of the channel. As in [4], we assume the entries $\mathbf{H}_{i,j}$ of the channel to be mutually independent zero mean complex Gaussian variables with unit variance (Rayleigh flat fading MIMO channel model). Teletar has shown [4] that for such a channel model, the optimization of the capacity subject to the TX power constraint $\frac{1}{2\pi j}\oint \frac{dz}{z}\mathrm{tr}(S_{\boldsymbol{aa}}(z)) \leq N_{tx}\sigma_b^2$ leads to the requirement of a white (and Gaussian) vector transmission signal $S_{\boldsymbol{aa}}(z) = \sigma_b^2 I$. Combined with the whiteness of the vector stream $\boldsymbol{b}_k$ resulting from the channel encoding, this leads to the requirement for the prefilter to be paraunitary: $\mathbf{T}(z)\mathbf{T}^{\dagger}(z) = \mathbf{I}$ to avoid capacity loss.

Motivated by the consideration of diversity also (see below), we propose to use the following paraunitary prefilter

$$\begin{aligned} \mathbf{T}(z) &= \mathbf{D}(z)\, Q \\ \mathbf{D}(z) &= diag\{1,\, z^{-1}, \ldots, z^{-(N_{tx}-1)}\}\,,\quad Q^H Q = I\,,\quad |Q_{ij}| = \tfrac{1}{\sqrt{N_{tx}}} \end{aligned} \tag{3}$$

where Q is a (constant) unitary matrix with equal magnitude elements. Note that for a channel with delay spread, the prefilter can be immediately adapted by replacing the elementary delay z^{-1} by z^{-L} for channel of length (delay spread) L. For the flat propagation channel $\mathbf{H}$ combined with the prefilter $\mathbf{T}(z)$ in (3), symbol stream n ($b_{n,k}$) passes through the equivalent SIMO channel

$$\sum_{i=1}^{N_{tx}} z^{-(i-1)} \mathbf{H}_{:,i} Q_{i,n} \tag{4}$$

which now has memory due to the delay diversity introduced by $\mathbf{D}(z)$. It is important that the different columns $\mathbf{H}_{:,i}$ of the channel matrix get spread out in time to get full diversity (otherwise the streams just pass through a linear combination of the columns, as in VBLAST, which offers limited diversity). The delay diversity only becomes effective by the introduction of the mixing/rotation matrix Q, which has equal magnitude elements for uniform diversity spreading. The prefilter introduced in [14] is essentially the same as the one in (3). However, the symbol stream $\boldsymbol{b}_k$ in [14] is a subsampled stream, subsampled by a factor N_{tx}. As a result, the system is single rate. The advantage in that case though is that no interference between the rotated substreams occurs, which simplifies detection.

4 Matched Filter Bound and Diversity

The Matched Filter Bound (MFB) is the maximum attainable SNR for symbol-wise detection, when the interference from all other symbols has been removed. Hence the multistream MFB equals the MFB for a given stream. For VBLAST ($\mathbf{T}(z) = \mathbf{I}$), the MFB for stream n is

$$\mathrm{MFB}_n = \rho ||\mathbf{H}_{:,n}||_2^2 \tag{5}$$

hence, diversity is limited to N_{rx}. For the proposed $\mathbf{T}(z) = \mathbf{D}(z)\, Q$ on the other hand, stream n has MFB

$$\mathrm{MFB}_n = \rho \frac{1}{N_{tx}}\, ||\mathbf{H}||_F^2 \tag{6}$$

hence this $\mathbf{T}(z)$ provides the same full diversity $N_{tx} N_{rx}$ for all streams. Larger diversity order leads to larger outage capacity.

5 Pairwise Probability of Error $\mathbf{P}_e$

The received signal is:

$$\boldsymbol{y}_k = \mathbf{H}\ \mathbf{T}(q)\ \boldsymbol{b}_k + \boldsymbol{v}_k = \mathbf{H}\ \mathbf{D}(q)\ Q\ \boldsymbol{b}_k + \boldsymbol{v}_k = \mathbf{H}\ \mathbf{D}(q)\ \boldsymbol{c}_k + \boldsymbol{v}_k \tag{7}$$

where $\boldsymbol{c}_k = Q\,\boldsymbol{b}_k = [c_1(k)\, c_2(k)\, \dots\, c_{N_{tx}}(k)]^T$. We consider now the transmission of the coded symbols over a duration of T symbol periods. The accumulated received signal is then:

$$\mathbf{Y} = \mathbf{H}\ \mathbf{C} + \mathbf{V} \tag{8}$$

where $\mathbf{Y}$ and $\mathbf{V}$ are $N_{rx} \times T$ and $\mathbf{C}$ is $N_{tx} \times T$. The structure of $\mathbf{C}$ will become clear below. Over a Rayleigh flat fading i.i.d. MIMO channel, the probability of deciding erroneously $\mathbf{C}'$ for transmitted $\mathbf{C}$ is upper bounded by (see [5]):

$$\mathbf{P}(\mathbf{C} \to \mathbf{C}') \leq (\prod_{i=1}^{r} \lambda_i)^{-N_{rx}} (\frac{\rho}{4})^{-N_{rx}\, r} \tag{9}$$

where $\boldsymbol{r}$ and λ_i are rank and eigenvalues of $(\mathbf{C} - \mathbf{C}')^H(\mathbf{C} - \mathbf{C}')$, and

$$\mathbf{C}-\mathbf{C}' = \frac{1}{\sigma_b}\begin{bmatrix} c_1(0) - {c'}_1(0) & c_1(1) - {c'}_1(1) & \dots & \dots & \dots & \dots \\ 0 & \ddots & \ddots & \dots & \dots & \dots \\ \vdots & \ddots & \ddots & \ddots & \dots & \dots \\ 0 & \dots & 0 & c_{N_{tx}}(0) - {c'}_{N_{tx}}(0) & c_{N_{tx}}(1) - {c'}_{N_{tx}}(1) & \dots \end{bmatrix} \tag{10}$$

Let i be the time index of the first error, and introduce $\boldsymbol{e}_k = \frac{1}{\sigma_b}(\boldsymbol{c}_k - \boldsymbol{c}'_k)$, then:

$$\mathbf{C} - \mathbf{C}' = \begin{bmatrix} 0 & \dots & 0 & e_1(i) & \dots & \dots & \dots & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 & e_{N_{tx}}(i) & \dots & \dots \end{bmatrix} . \tag{11}$$

Under the condition

$$\prod_{n=1}^{N_{tx}} e_n(i) \neq 0 \tag{12}$$

the upper bound on the pairwise error probability becomes (maximized for a single error event i):

$$\mathbf{P}(\mathbf{C} \to \mathbf{C}') \leq (\prod_{n=1}^{N_{tx}} |e_n(i)|^2)^{-N_{rx}} . (\frac{\rho}{4})^{-N_{rx}\, N_{tx}} . \tag{13}$$

Hence, full diversity $N_{rx}\, N_{tx}$ is guaranteed, and the coding gain is: $\min_{\boldsymbol{e}_i \neq 0} \prod_{n=1}^{N_{tx}} |e_n(i)|^2$. The condition (12) is well known in the design of lattice constellations (see [9], [10]), a field based on the theory of numbers. A solution that satisfies our criteria of unitary matrix and equal magnitude components of Q, is the Vandermonde matrix:

$$Q^s = \frac{1}{\sqrt{N_{tx}}}\begin{bmatrix} 1 & \theta_1 & \dots & {\theta_1}^{N_{tx}-1} \\ 1 & \theta_2 & \dots & {\theta_2}^{N_{tx}-1} \\ \vdots & \vdots & & \vdots \\ 1 & \theta_{N_{tx}} & \dots & {\theta_{N_{tx}}}^{N_{tx}-1} \end{bmatrix} \tag{14}$$

where the θ_i are the roots of $\theta^{N_{tx}} - j = 0\,,\ j = \sqrt{-1}$.

5.1 Optimality for QAM constellations in the case $N_{tx} = 2^{n_t}$

For $N_{tx} = 2^{n_t}(n_t \in \mathbb{Z})$, Q^s also leads to satisfaction of (12) [9], and guarantees for any constellation such that $b_n(i) - b'_n(i) = a + jb \in \mathbb{Z}[j]$ ($\mathbb{Z}[j] = \{a + jb \,|\, a, b \in \mathbb{Z}\}$), with $\boldsymbol{b}_i - \boldsymbol{b}'_i \in (\mathbb{Z}[j])^{N_{tx}}/0)$, that $({N_{tx}}^{N_{tx}/2} \prod_{n=1}^{N_{tx}} e_n(i)) \in \mathbb{Z}[j]/0$, and hence:

$$\min_{\boldsymbol{e}_i \neq 0} \prod_{n=1}^{N_{tx}} |e_n(i)|^2 \geq \left(\frac{1}{N_{tx}}\right)^{N_{tx}} \quad . \tag{15}$$

For finite QAM constellations with $(2M)^2$ points, any symbol can be written as: $b_n(i) = d\{(2l-1) + j(2p-1)\}$ where $d \in \mathbb{R}^{+*}$, $l, p \in \{-M+1, -M+2, \ldots, M\}$. Then $\frac{1}{\sigma_b}(b_n(i) - b'_n(i)) = \frac{2d}{\sigma_b}(l' + jp')$, $l', p' \in \{-2M+1, -2M+2, \ldots, 2M-1\}$ and $\sigma_b^2 = \frac{2(4M^2-1)d^2}{3}$. The lower bound of (15), which is valid in fact for any Vandermonde matrix Q of the form in (14) built with roots of a polynomial of order N_{tx} with coefficients in $\mathbb{Z}[j]$ and satisfying a certain number of conditions [9] (hence Q^s is a special case of this family), becomes

$$\min_{\boldsymbol{e}_i \neq 0} \prod_{n=1}^{N_{tx}} |e_n(i)|^2 \geq \left(\frac{4d^2}{\sigma_b^2}\right)^{N_{tx}} \left(\frac{1}{N_{tx}}\right)^{N_{tx}} = \left(\frac{4d^2}{N_{tx}\sigma_b^2}\right)^{N_{tx}} \quad . \tag{16}$$

In what follows, we consider an upper bound for the coding gain for any matrix Q with normalized columns. The minimal product of errors $\prod_n |e_n(i)|^2$ is upper bounded by a particular error instance corresponding to a single error in the b's, when $\frac{1}{\sigma_b}(\boldsymbol{b}_i - \boldsymbol{b}'_i) = \frac{2d}{\sigma_b}\boldsymbol{w}_{n_0}$, where $\boldsymbol{w}_{n_0}$ is the vector with one in the n_0^{th} coefficient and zeros elsewhere, hence

$$\min_{\boldsymbol{e}_i \neq 0} \prod_{n=1}^{N_{tx}} |e_n(i)|^2 \leq \left(\frac{4d^2}{\sigma_b^2}\right)^{N_{tx}} \prod_{n=1}^{N_{tx}} |Q_{n,n_0}|^2 \quad . \tag{17}$$

Now, given that $\sum_{n=1}^{N_{tx}} |Q_{n,n_0}|^2 = 1$, then by applying Jensen's inequality, we get

$$\prod_{n=1}^{N_{tx}} |Q_{n,n_0}|^2 \leq \left(\frac{1}{N_{tx}}\right)^{N_{tx}} \quad . \tag{18}$$

Hence,

$$\min_{\boldsymbol{e}_i \neq 0} \prod_{n=1}^{N_{tx}} |e_n(i)|^2 \leq \left(\frac{4d^2}{\sigma_b^2}\right)^{N_{tx}} \left(\frac{1}{N_{tx}}\right)^{N_{tx}} = \left(\frac{4d^2}{N_{tx}\sigma_b^2}\right)^{N_{tx}} \tag{19}$$

is an upper bound for the coding gain for any matrix Q with normalized columns. Now, the intersection of the sets of matrices that lead to the lower bound (16) and the upper bound (19) includes the unitary matrix Q^s given in (14), which hence achieves the upper bound on the coding gain: $\min_{\boldsymbol{e}_i \neq 0} \prod_{n=1}^{N_{tx}} |e_n(i)|^2 = \left(\frac{4d^2}{N_{tx}\sigma_b^2}\right)^{N_{tx}}$.

Remark1: The Jensen's inequality (18) becomes an equality if and only if all the coefficients Q_{n,n_0}, $n = 1, \ldots, N_{tx}$ have the same module $1/\sqrt{N_{tx}}$. This holds for any $n_0 = 1, \ldots, N_{tx}$. Hence we conclude that a necessay condition on any unitary matrix Q to maximize the coding gain is to have all equal magnitude coefficients. This is equivalent to our condition to achieve the same maximum MFB for all streams (full diversity).

Remark2: In the case when $N_{tx} \neq 2^{n_t}$ (and using Q^s), the coding gain is closely related to the size of the used QAM constellation, and is in general lower then the upper bound given above.

6 ML Reception

In principle, we can perform Maximum Likelihood reception since the delay diversity transforms the flat channel into a channel with finite memory. However, the number of states would be the product of the constellation sizes of the N_{tx} streams to the power $N_{tx}-1$. Hence, if all the streams have the same constellation size $|\mathcal{A}|$, the number of states would be $|\mathcal{A}|^{N_{tx}(N_{tx}-1)}$, which will be much too large in typical applications. Suboptimal ML reception can be performed in the form of sphere decoding [11]. The complexity of this can still be too large though and therefore suboptimal receiver structures will be considered in the next section. Before continuing however, we wish to make a comparison of the proposed approach with the full rate linear dispersion codes of Belfiore & Galliou [8], based on Galois theory:

- both schemes have full rate with full diversity,
- design of one unitary matrix to optimize coding gain in our approach, need at least two unitary matrices in their design and the coding gain optimization is more involved,
- ML decoding of similar complexity if minimal block length ($= N_{tx}$) is used in their design.

7 MIMO DFE Reception

Let $\mathbf{G}(z) = \mathbf{H}\,\mathbf{T}(z) = \mathbf{H}\,\mathbf{D}(z)\,Q$ be the cascade transfer function of channel and precoding. The matched filter RX is

$$\boldsymbol{x}_k = \mathbf{G}^\dagger(q)\,\boldsymbol{y}_k = \mathbf{G}^\dagger(q)\,\mathbf{G}(q)\,\boldsymbol{b}_k + \mathbf{G}^\dagger(q)\,\boldsymbol{v}_k = \mathbf{R}(q)\,\boldsymbol{b}_k + \mathbf{G}^\dagger(q)\,\boldsymbol{v}_k \tag{20}$$

where $\mathbf{R}(z) = \mathbf{G}^\dagger(z)\,\mathbf{G}(z)$, and the psdf of $\mathbf{G}^\dagger(q)\,\boldsymbol{v}_k$ is $\sigma_v^2\,\mathbf{R}(z)$. The DFE RX is then:

$$\widehat{\mathbf{b}}_k \;=\; -\underbrace{\overline{\mathbf{L}}(q)}_{\text{feedback}}\;\mathbf{b}_k + \underbrace{\mathbf{F}(q)}_{\text{feedback}}\;\mathbf{x}_k \tag{21}$$

where feedback $\overline{\mathbf{L}}(z)$ is strictly "causal". Two design criteria for feedforward and feedback filters are possible: MMSE ZF and MMSE.

7.1 MIMO MMSE ZF DFE RX

Consider the matrix spectral factorization:

$$\mathbf{G}^\dagger(z)\mathbf{G}(z) \;=\; \mathbf{R}(z) \;=\; \mathbf{L}^\dagger(z)\,\Sigma\,\mathbf{L}(z) \tag{22}$$

where $\mathbf{L}(z) = \sum_k \mathbf{L}_k\,z^{-k}$ with $diag(\mathbf{L}_0) = I$ (monic), $\Sigma > 0$ is diagonal and constant. Then $\mathbf{F}(z) = \Sigma^{-1}\,\mathbf{L}^{-\dagger}(z)$, $\overline{\mathbf{L}}(z) = \mathbf{L}(z) - I$. The total feedforward filter is a scaled Whitened Matched Filter (WMF)

$$\mathbf{F}(z)\mathbf{G}^\dagger(z) = \Sigma^{-\frac{1}{2}}\,\Sigma^{-\frac{1}{2}}\mathbf{L}^{-\dagger}(z)\mathbf{G}^\dagger(z) = \Sigma^{-\frac{1}{2}}\,\mathbf{U}(z) \tag{23}$$

where $\mathbf{U}(z)$ = paraunitary/lossless/WMF. The forward filter output

$$\mathbf{F}(q)\,\mathbf{x}_k = \mathbf{L}(q)\,\mathbf{b}_k + \mathbf{F}(q)\mathbf{G}^\dagger(q)\,\mathbf{v}_k = \mathbf{L}(q)\,\mathbf{b}_k + \mathbf{e}_k \tag{24}$$

where $\mathbf{S}_{\mathbf{ee}}(z) = \sigma_v^2\,\Sigma^{-1}$. At detector output i: $\text{SNR}_i = \rho\,\Sigma_{ii}$. We can detect the $\mathbf{b}_k$ elementwise by backsubstitution (feedback) and symbol-by-symbol detection.

7.2 MIMO MMSE DFE RX

Consider now the backward channel model based on LMMSE [3]:

$$\mathbf{b}_k \;=\; \widehat{\mathbf{b}}_k + \widetilde{\mathbf{b}}_k \;=\; \mathbf{S}_{\mathbf{bx}}(q)\,\mathbf{S}_{\mathbf{xx}}^{-1}(q)\,\mathbf{x}_k + \widetilde{\mathbf{b}}_k \tag{25}$$

where $\mathbf{S}_{\mathbf{bx}}(z) \;=\; \mathbf{S}_{\mathbf{bb}}(z)\,\mathbf{G}^{\dagger}(z)\mathbf{G}(z)$ and $\mathbf{S}_{\mathbf{xx}}(z) \;=\; \mathbf{G}^{\dagger}(z)\mathbf{G}(z)\,\mathbf{S}_{\mathbf{bb}}(z)\,\mathbf{G}^{\dagger}(z)\mathbf{G}(z) + \sigma_v^2\,\mathbf{G}^{\dagger}(z)\mathbf{G}(z)$. Hence $\mathbf{S}_{\mathbf{bx}}(z)\,\mathbf{S}_{\mathbf{xx}}^{-1}(z) \;=\; \mathbf{R}^{-1}(z)$ with $\mathbf{R}(z) \;=\; \mathbf{G}^{\dagger}(z)\mathbf{G}(z)+\sigma_v^2\,\mathbf{S}_{\mathbf{bb}}^{-1}(z) \;=\; \mathbf{G}^{\dagger}(z)\mathbf{G}(z)+\frac{1}{\rho}\,I$ so $\mathbf{b}_k \;=\; \mathbf{R}^{-1}(q)\mathbf{x}_k+\widetilde{\mathbf{b}}_k$. We get $\mathbf{S}_{\widetilde{\mathbf{b}}\widetilde{\mathbf{b}}}(z) \;=\; \mathbf{S}_{\mathbf{bb}}(z)-\mathbf{S}_{\mathbf{bx}}(z)\mathbf{S}_{\mathbf{xx}}^{-1}(z)\mathbf{S}_{\mathbf{xb}}(z) \;=\; \sigma_v^2\,\mathbf{R}^{-1}(z)$. Apply again matrix spectral factorization:

$$\mathbf{R}(z) \;=\; \mathbf{L}^{\dagger}(z)\,\Sigma\,\mathbf{L}(z) \tag{26}$$

then $\mathbf{b}_k \;=\; \mathbf{L}^{-1}(q)\,\Sigma^{-1}\,\mathbf{L}^{-\dagger}(q)\,\mathbf{x}_k + \widetilde{\mathbf{b}}_k$. We get

$$\mathbf{F}(q)\,\mathbf{x}_k = \Sigma^{-1}\,\mathbf{L}^{-\dagger}(q)\,\mathbf{x}_k \;=\; \mathbf{L}(q)\,\mathbf{b}_k - \mathbf{L}(q)\,\widetilde{\mathbf{b}}_k = \mathbf{L}(q)\,\mathbf{b}_k + \mathbf{e}_k \tag{27}$$

where $\mathbf{S}_{\mathbf{ee}}(z) \;=\; \mathbf{L}(z)\mathbf{R}^{-1}(z)\mathbf{L}^{\dagger}(z) \;=\; \sigma_v^2\,\Sigma^{-1}$. At detector output i again: $\text{SNR}_i = \rho\,\Sigma_{ii}$. In general

$$\Sigma^{MMSE} > \Sigma^{MMSEZF} \;\;\Rightarrow\;\; \text{SNR}_i^{MMSE} > \text{SNR}_i^{MMSEZF}$$

and even $\text{SNR}_i^{UMMSE} \;=\; \text{SNR}_i^{MMSE} - 1 \;>\; \text{SNR}_i^{MMSEZF}$ where UMMSE refers to Unbiased MMSE.

7.3 Capacity Decomposition

For a given channel realization

$$\begin{aligned} \mathbf{C} &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z}\log_2\det(I_{N_{rx}} + \rho\,\mathbf{G}(z)\mathbf{G}^{\dagger}(z)) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z}\log_2\det(I_{N_{tx}} + \rho\,\mathbf{G}^{\dagger}(z)\mathbf{G}(z)) \\ &= \tfrac{1}{2\pi j}\oint \tfrac{dz}{z}\log_2\det(\rho\,\mathbf{R}^{MMSE}(z)) \;=\; \log_2\det(\rho\,\Sigma^{MMSE}) \\ &= \sum_{n=1}^{N_{tx}}\log_2\text{SNR}_i^{MMSE} \;=\; \sum_{n=1}^{N_{tx}}\log_2(1+\text{SNR}_i^{UMMSE}) \end{aligned} \tag{28}$$

The total capacity = sum of capacities of N_{tx} substreams output by a UMMSE DFE, taken as independent AWGN channels (Gaussian approximation of UMMSE error signal).

7.4 Matrix Spectral Factorization Considerations

Conventionally: $\mathbf{L}(z) \;=\; \sum_{k=0}^{\infty}\mathbf{L}_k\,z^{-k}$, where $\mathbf{L}_0$ is unit diagonal and lower triangular. Consider a generalization with relative delays via linear prediction $\mathbf{P}(z) = \mathbf{L}^{-1}(z)$ applied to $\widetilde{\mathbf{b}}_k$: the conventional linear predictor $\mathbf{P}_c(z)$ applied to

$$\mathbf{Z}(q)\,\widetilde{\mathbf{b}}_k = \begin{bmatrix} \widetilde{b}_{1,k} \\ \widetilde{b}_{2,k-d_1} \\ \vdots \\ \widetilde{b}_{N_{tx},k-d_{N_{tx}-1}} \end{bmatrix} \tag{29}$$

leads to the generalized predictor: $\mathbf{P}(z) = \mathbf{Z}^{-1}(z)\,\mathbf{P}_c(z)\,\mathbf{Z}(z)$. We can obtain the triangular spectral factor or predictor as the limiting case as delays $\to \infty$. $\mathbf{P}(z)$: strictly lower triangular elements are noncausal Wiener filters to estimate a signal component in terms of the previous signal components, the diagonal elements are SISO prediction error filters of the resulting residual signals.

With triangular spectral factors and feedback filters: we detect one symbol stream over all time and then pass to the next symbol stream. With a conventional feedback filter: we process all symbols one after the other at a given time instant, and then pass to the next time instant. The advantage of triangular factors/feedback: can incorporate channel decoding in detection before use of symbols in feedback (leading to the stripping approach of Verdu & Müller or Varanasi & Guess..) $\Rightarrow$ much more reliable feedback. In practice: finite relative delays between substreams suffice.

7.5 Triangular MIMO DFE and VBLAST

With triangular feedback: MIMO DFE works as follows:
1. we apply a SIMO DFE to detect a substream, the design of the SIMO DFE considers the remaining substreams as colored noise.
2. we subtract the detected and decoded substream from the RX signal and pass on to the next substream.
For the first substream, all remaining streams are interferers, the last substream gets detected in the single stream scenario. Hence, triangular MIMO DFE = extension of VBLAST to the dynamic case. Here, the dynamics (temporal dispersion) have been introduced by linear convolutive precoding (introducing delay diversity). Advantages:

- no ordering issue: can process streams in any order,
- higher diversity order, less dispersion of substream SNRs.

7.6 Practical Implementation of MIMO DFE

Although the complexity of a suboptimal receiver like the MIMO DFE can still be considered quite high, a practical approximation is possible as follows. One should consider the Noise Predictive DFE form. In this case, the forward filter is in fact the Linear MMSE (LMMSE) receiver. The backward filter is then a MIMO noise prediction filter. We suggest to use the triangular MIMO predictor structure for reasons already mentioned. The complexity of the MIMO predictor can be adjusted by adjusting the prediction order. This gives performance in between that of the LMMSE receiver and that of the DFE. The LMMSE receiver/forward filter can be approximated by polynomial expansion.

References

[1] G.J. Foschini. "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when Using Multi-Element Antennas". *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41-59, 1996.

[2] P.W. Wolniansky, G.J. Foschini, G.D. Golden and R.A. Valenzuela. "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel". Invited paper, Proc. *ISSSE-98*, Pisa, Italy, Sept. 29, 1998.

[3] G. Ginis and J.M. Cioffi. "On the Relation Between V-BLAST and the GDFE". *IEEE Communications Letters*, vol. 5, no. 9, pp. 364-366, Sept. 2001.

[4] I.E. Telatar. "Capacity of Multi-antenna Gaussian Channels". *Eur. Trans. Telecom.*, vol. 10, no. 6, pp. 585-595, Nov/Dec 1999.

[5] V. Tarokh, N. Seshadri, and A.R. Calderbank. "Space-Time Codes for High Data Rates Wireless communication: Performance criterion and code construction". *IEEE Trans. Info. Theory*, vol. 44, no. 2, pp. 744-765, March 1998.

[6] D. Mary and D.T.M. Slock. "Vectorial DPCM Coding and Application to Wideband Coding of Speech". *Proc. ICASSP Conf.*, Salt Lake City, USA, May 2001.

[7] B. Hassibi and B. M. Hochwald. "High-Rate Codes that are Linear in Space and Time". Submitted to *IEEE Trans. Info. Theory*.

[8] S. Galliou and J.C. Belfiore. "A New Familly of Linear Full-Rate Space-Time Codes Based on Galois Theory". Draft paper.

[9] X. Giraud, E. Boutillon and J.C. Belfiore. "Algebraic Tools to Build Modulation Schemes for Fading Channels". *IEEE Trans. Info. Theory*, vol. 43, no. 3, pp. 938-952, May 1997.

[10] J. Boutros, E.Viterbo, C.Rastello and J.C. Belfiore. "Good Lattice Constellations for both Rayleigh Fading and Gaussian Channels". *IEEE Trans. Info. Theory*, vol. 42, no. 2, pp. 502-518, March 1996.

[11] M.O. Damen, A.Chkeif and J.C. Belfiore. "Sphere Decoding of Space-Time codes". In *IEEE ISIT 2000*, Sorrento, Italy, June 2000.

[12] S. Moshavi, E.G. Kanterakis, D.L. Schilling. "Multistage Linear Receivers for DS-CDMA Systems". *Int. J. Wireless Inform. Networks*, vol. 3, no. 1, pp. 1-17, Jan. 1996.

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

[14] Y. Xin, Z. Wang, G. Giannakis "Space-Time Diversity Systems Based on Unitary Constellation-Rotating Precoders". *Proc. ICASSP Conf.*, Salt Lake City, USA, May 2001.