Space-Time Coding for Capacity Enhancement in Future-Generation Wireless Communications Networks

D.D.N. Bevan (corresponding author), R. Tanner, C.R. Ward
Nortel Networks, Harlow Laboratories, London Road, Harlow, Essex, CM17 9NA U.K.

Keywords: Space-Time Coding (STC), MIMO Channels, BLAST, STTC, STTCM, STTD, Convolutional Encoder, QPSK.

ABSTRACT

In this paper we present a simple yet novel STC approach, based on a logical extension of simple convolutional codes. We compare this STC scheme with known transmission techniques, including Space-Time Transmit Diversity (STTD) and BLAST, in terms of code construction and performance. We present simulation results of actual performance, and discuss the reasons for, and implications of, the performance trends observed.

1. INTRODUCTION

It is envisaged that the applications and services which 3G mobile systems (e.g. [1][2]) will be required to carry will call for significantly higher capacities than are realised today (particularly on the system downlink). One possible way to achieve this would be to employ multiple receive antennas at the user terminal. This could be used to reduce interference, hence achieving power-efficient data transmission. Since interference and capacity have an inverse relationship [3], this would maximise system capacity. However, it could lead to a new capacity limit due to the limited spectral efficiency of the transmitted waveform (or the 'Walsh code exhaustion' problem as it pertains to CDMA systems). A possible solution to this would be to introduce additional antennas also at the transmitter end of the link, creating a multiple-input-multiple-output (MIMO) Basestation-Mobile channel. Using an information-theoretic approach, Telatar [4] has demonstrated that there are significant capacity gains to be exploited through the inherent parallelism of the MIMO channel, allowing us to 'squeeze' more information through the link for a given per-user signalling bandwidth (i.e. symbol rate or chip rate).

Space-Time Coding (STC) is a revolutionary development for exploiting the MIMO channel by using antenna array processing technology, which is currently stimulating considerable interest across the wireless industry. This new innovation stems from the pioneering work in the early 1990’s of Raleigh & Cioffi at Stanford [5], Wittneben at ASCOM in Switzerland [6], more recently by Foschini & Gans of Lucent Labs [7][8][9], and Tarokh and colleagues at AT&T Research Labs [10]. The Space-Time Coding ‘concept’ builds on the significant work by Winters in the mid-80’s which highlighted the importance of antenna diversity on the capacity of wireless systems [11].

The use of multiple antennas at both the transmitter and receiver is essential for the STC concept to work effectively, since STC exploits both the temporal and spatial dimensions for the construction of coding designs which effectively mitigate fading (for improved power efficiency) and are able to capitalise upon parallel transmission paths within the propagation channel (for improved bandwidth efficiency). These parallel transmission paths, also referred to as ‘spatial modes’, ‘spatial channels’ or
'data pipes', arise in those cases when the multipath environment is rich in scattering. For more discussion of the scattering environment, and its influence on theoretical MIMO system capacity see [12].

2. SPACE-TIME CODING

The fundamental principle of Space-Time Coding (STC), illustrated in Figure 1, is to take an input stream of information bits, and from them generate vector outputs of simultaneous transmissions at a number of transmit antennas. We term this simultaneous transmission of modulation symbols from different antennas a 'Space-Time Symbol' (STS) or 'Space-Time Vector Symbol' (STVS). Since modulation symbols can conventionally be represented by a complex number (using complex baseband representation [13]), an STS can be represented by a vector of complex numbers, with the number of complex elements in the vector equal to the number of transmit antennas. Note that in an FDMA/TDMA system these simultaneous symbols are using the same carrier frequency and same symboling waveform. In a CDMA system they also use an identical symboling waveform, in the form of an identical spreading code (i.e. Walsh code). Strictly-speaking, we usually consider that a 'pure' STC coder has no instantaneous knowledge of the MIMO channels. We assume that the receiver (decoder) does know the instantaneous channels (i.e. it is operating in a 'non-blind' mode).

![Space-Time Coding System Model](image)

Figure 1: Space-Time Coding - system model

For convenience, we will assume a non-dispersive channel. The receiver simultaneously detects all of the elements of a transmitted STS using a single symbol-matched filter per receiver antenna. These detection outputs are built up into a (vector) detection statistic. Thus we see in Figure 1 that every element of this vectorial receiver detection statistic is a superposition of the multiple simultaneous transmissions, as seen at each receiver antenna. Each element, viewed over many STS, will thus resemble a 'jumbled' QAM-type constellation (in the noise-free case), where the exact form of this constellation depends on the STC encoder structure, STC modulation alphabet and instantaneous channel.
In this paper we present simulation results for the comparative performance of a number of candidate STC schemes in a MIMO channel, and discuss reasons for the performance trends we identify.

3. CANDIDATE SPACE-TIME CODES

In this section we discuss, in general terms, a number of different families of Space-Time Codes. These are illustrated for ease of understanding in Figure 2.

Figure 2: Candidate approaches to Space-Time Coding

3.1. Orthogonal Space-Time Block Coding (OSTBC)

The two best-known OSTBC schemes are a) Space-Time Transmit Diversity (STTD), shown in Figure 2, which was first described by Alamouti [14], a modified version of which was subsequently proposed for the UMTS 3G W-CDMA standard by Texas Instruments [15], and b) Orthogonal Transmit Diversity (also described in [15]), as proposed by Motorola for inclusion in the cdma2000 3G CDMA standard. The principle of both schemes is to map coded modulation symbols to two transmit antennas. STTD maps all modulated symbols to both transmit antennas, whereas OTD maps each modulation symbol to only one of the antennas, and relies on the memory within the coder to obtain any diversity advantage. Note that both STTD and OTD have the desirable property that they don't expand the bandwidth of the signal\(^1\), and furthermore, the receiver can process the received signal using linear processing (the 'modulation symbol demapper' of Figure 2), prior to standard Viterbi decoding (see [14] [15] for a fuller description of OSTBC transmitter and receiver processing). Using STTD and OTD we are able to extract diversity without having to sacrifice spectral efficiency, since the codes are 'rate 1'. However, they don't offer any intrinsic 'coding gain'.

\(^1\) That is, the number of STS at the coder output is equal to the number of modulation symbols at the input.
3.2. Layered Space-Time Processing

The Layered Space-Time Processing approach to STC was first introduced by Lucent's Bell Labs, with their BLAST family of STC structures [8]. The basic concept behind layered STC is illustrated in Figure 2 (ii). The information bits are demultiplexed into individual streams, which are then fed into individual encoders. These coders may be binary convolutional coders, or even no coding at all. The outputs of the coders are modulated and fed to the separate antennas, from which they are transmitted, using the same carrier-frequency/symbol waveform (TDMA) or Walsh code (CDMA). At the receiver, a spatial beamforming/nulling (zero-forcing) process is used at the front end in order to separate the individual coded streams, and feed them to their individual decoders. The outputs of the decoders are multiplexed back to reconstruct the estimate of the original information bitstream.

There are a number of variations on the original BLAST theme. One is to modify the receiver antenna pre-processing to carry out MMSE beamforming rather than nulling, in order to improve the wanted-signal SNR at the expense of slightly-increased ISI (from one transmit antennas signal across to another). Both the MMSE and nulling approaches have the disadvantage that some of the diversity potential of the receiver antenna array is necessarily sacrificed in the beamforming process. In order to overcome this problem, the inventors of BLAST propose layering the receiver processing, such that after the strongest signal has been decoded (typically using the Viterbi MLSE algorithm), it is subtracted from the received antenna signals in order to remove the necessity to null it when detecting the weaker signals. This process is iterated down until the detection of the weakest signal requires no nulling at all, and its diversity performance is therefore maximised. The disadvantage with this layered approach, like all subtractive multi-user detection schemes, is that wrong detection decisions higher up in the chain can adversely affect performance lower down.

3.3. Space-Time Trellis Coding (STTC)

Space-Time Trellis Coding (STTC) was first discussed in the literature by Tarokh et al [10], and is itself a generalisation of Trellis-Coded Modulation (TCM) as first introduced by Ungerboeck [16]. The basic structure of the STTC encoder is illustrated in Figure 2. The encoder can be considered in the most general sense as a finite-state state-machine. The value of the latest group of information bits is used to determine the transitions which will occur between present state and next state. The result of this transition is the transmission of a 'Space-Time Symbol' (STS) which is actually equivalent to the simultaneous transmission of a group of symbols, one from each of the transmit antennas. The constituent symbols of the STS could in principle be chosen from any set, such as QPSK, 8-PSK or 16-QAM. However, in the simulations presented later in this paper we have only considered the QPSK case.

We now attempt briefly to explain the operation of the STTC decoder, which is based on the Viterbi algorithm. The task of the STTC decoder is to track the progress of the encoder state machine, based on the observable signal received on a multiplicity of receiver antennas. The approach usually taken is to employ the Viterbi algorithm in order to form the maximum-likelihood ('hard') estimate of the original information sequence. The Viterbi algorithm requires at its input, for each 'branch' of the decoding (i.e. for each transition from present state to next state), a set of log-likelihood ratios (LLR) for the set of all possible hypotheses of the received STS. This set of LLRs is generated by the preceding ‘demodulator’ stage2 (see Figure 2), based on squared Euclidean distances (in the $N_R$-ary complex vector space of the receiver baseband processor) between the hypothesised (‘noise-free’) received symbol (for the trellis branch under consideration) and the actual received signal (see Figure

2 Where, after Massey [17], we note that the ‘modulator’ must be a memoryless device, and we consequently define a demodulator as a device to ‘extract likelihood information about the possible transmitted signal in every symbol period’, this likelihood information subsequently being passed to the decoder.
1, which illustrates the 'noise-free' hypotheses). Note that as the channel changes (due to Doppler),
this set of hypotheses for the received STS will also change.

It can be seen that the complexity of such a maximum-likelihood receiver (for the non-dispersive
channel) is therefore a function of:

i) The number of states in the trellis

ii) The size of the set of possible STS (i.e. the set of possible received-STS hypotheses)

iii) The number of branches leaving a state (which also determines the code's spectral efficiency)

iv) The number of receiver antennas, $N_R$

v) The Doppler rate of the channel (which affects the rate of update of the hypotheses)

Tarokh et. al. have analysed STTC codes at length [10], and have devised heuristics to design STTCs
which they claim are in some sense optimal (for a given number of states). However, in [10] they
only published optimal STTC codes for $N_T=2$, and low numbers of states. Since then, other groups
have also investigated optimised STTC codes for different numbers of antennas and trellis states (see
for example [18][19]).

This paper presents a class of codes hereafter termed 'Modified-Convolutional Space-Time Trellis
Codes' (MC-STTC), which can be easily generalised to higher numbers of states and more than 2
transmit antennas. The technique is the rather intuitive one of taking a well-known optimal binary
convolutional code of a given constraint length [13], mapping the outputs to QPSK symbols, and
demultiplexing the result to the different transmit antennas (see Figure 3). Thus it becomes
immediately apparent that MC-STTC technique bears a similarity to BLAST, the fundamental
difference between them being that for BLAST the demultiplexing takes place before the coding,
whereas for MC-STTC it is the other way around.

![Figure 3: MC-STTC block diagram](image-url)
4. SIMULATION DESCRIPTION

We now describe the simulations carried out in order to compare the performance of a subset of different STC schemes, chosen from the classes described above. The simulation results have also been published elsewhere in [20]. The simulations were carried out for a simplified system at the symbol-level, whereby frames of STS were transmitted over a quasi-static flat Rayleigh fading MIMO channel (hence there is no interleaving of channel symbols). Thus the complex (Gaussian) random fading coefficient between each transmitter antenna and each receiver antenna is constant over a frame, and independent of every other fading coefficient for the same frame and for other frames. Notation $N_T : N_R$ is used to denote numbers of antennas at each end of the link. Each frame carries 400 information bits (800 in the case of uncoded BLAST), where an 'information bit' is defined as a data bit at the input to the Space-Time encoder (to which tail bits are added if necessary, to return the encoder to a known state). Performance is quantified in terms of required $E_b/N_0$ and $E_s/N_0$ at each receiver antenna for a mean frame-error-rate (FER) of 10%. $E_s$ is the mean energy (squared value) of the received STS at each receiver antenna, and $E_b$ is the mean energy per information bit at each receiver antenna (i.e. $E_s$ divided by the spectral efficiency). The 10% value of FER was chosen to be a suitable threshold at which the power efficiency of the code is maximised if an ARQ protocol is employed in a packet-data protocol. There is no simulation of fast transmit power control, so the transmitted power is constant over the duration of a simulation, and received $E_b$ is also equal to transmitted $E_b$, making the normalising assumption that each channel coefficient has a mean square value of unity.

In this section we provide a brief description of the various STC schemes which have been simulated, with details presented below. The scheme labelled 'benchmark' is a 1/2-rate $k=9$ maximum free distance binary convolutional encoder [13 Table 8-2-1], with output bits mapped to QPSK symbols. In a static 1:1 channel it achieves 10% FER for an $E_b/N_0$ of 2.2dB and 1% BER for an $E_b/N_0$ of 1.7dB. 'STTD' is a concatenation of the benchmark encoder with the Space-Time Transmit Diversity block code [14]. 'STTC-32' is the Tarokh-Seshradi-Calderbank (TSC) code ([10] Fig. 6) with 32 states and two information bits per STS. 'MC-STTC-256' is a concatenation of the ('benchmark') 1/2-rate $k=9$ maximum free distance binary convolutional encoder, a QPSK modulation mapper, and a demux mapper to the different transmit antennas. It thus has $2^{k-1} = 256$ states in the trellis, and a spectral efficiency equal to the number of transmit antennas employed. A Viterbi-based maximum-likelihood sequence estimation (MLSE) receiver structure was used for all of the coding schemes described here.

A number of layered space-time schemes were also simulated (labelled 'BLAST'). The variants of BLAST shown in the table use 'genie-aided' subtraction of the transmission from the stronger transmit antenna when detecting the transmission from the weaker. That is, the signal seen at the receiver antennas from the stronger transmit antenna (for this frame) is perfectly subtracted prior to detection of the signal from the weaker antenna. The 'uncoded' BLAST case uses no error correction code (it is raw QPSK), hence it achieves a high spectral efficiency of 4 bps/Hz for a 2:2 MIMO configuration. Separate frames of data (each carrying 400 information bits) are transmitted from each of the two transmit antennas. The stronger of the two frames is detected first, by spatially nulling out (zero forcing) the weaker transmission, before detecting the bits. A 'genie-aided' subtraction process is then applied to perfectly subtract this signal at the receiver, before Maximum-Ratio Combination (MRC) [21] beamforming to the weaker transmission (i.e. avoiding the requirement to null the stronger, since it has already been subtracted). Finally we carry out a detection of the noisy QPSK data at the output of this beamformer, which is our noisy received signal from the weaker transmit antenna.

The 'coded-BLAST' case uses two separate 'benchmark' encoders, one for each antenna, followed by a symbol mapper to QPSK prior to transmission. Receiver processing is carried out similarly to that applied for the uncoded BLAST case in that beamformer nulling is applied to isolate the signal from the stronger transmit antenna (along with a transfer of soft metrics from the beamformer pre-
processor to the respective decoder), followed by 'genie- aided' subtraction and MRC to generate soft metrics pertaining to the second antenna transmission.

5. SIMULATION RESULTS

The results of the simulations are shown in Table 1 showing the trade-off of $E_b/N_0$ performance (power efficiency) against spectral efficiency.

<table>
<thead>
<tr>
<th>STC-Algorithm (N_T:N_R)</th>
<th>$E_b/N_0$ dB for 10% FER</th>
<th>Spectral Efficiency (bps/Hz)</th>
<th>$E_S/N_0$ (SNR) dB for 10% FER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:1</td>
<td>11.0</td>
<td>1.0</td>
<td>11.0</td>
</tr>
<tr>
<td>1:2</td>
<td>4.5</td>
<td>1.0</td>
<td>4.5</td>
</tr>
<tr>
<td>1:4</td>
<td>-0.5</td>
<td>1.0</td>
<td>-0.5</td>
</tr>
<tr>
<td>STTD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:1</td>
<td>7.2</td>
<td>1.0</td>
<td>7.2</td>
</tr>
<tr>
<td>2:2</td>
<td>2.2</td>
<td>1.0</td>
<td>2.2</td>
</tr>
<tr>
<td>STTC-32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:1</td>
<td>11.0</td>
<td>2.0</td>
<td>14.0</td>
</tr>
<tr>
<td>2:2</td>
<td>4.6</td>
<td>2.0</td>
<td>7.6</td>
</tr>
<tr>
<td>MC-STTC-256</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:1</td>
<td>9.6</td>
<td>2.0</td>
<td>12.6</td>
</tr>
<tr>
<td>2:2</td>
<td>3.5</td>
<td>2.0</td>
<td>6.5</td>
</tr>
<tr>
<td>4:1</td>
<td>12.0</td>
<td>4.0</td>
<td>18.0</td>
</tr>
<tr>
<td>4:2</td>
<td>4.0</td>
<td>4.0</td>
<td>10.0</td>
</tr>
<tr>
<td>4:4</td>
<td>-1.5</td>
<td>4.0</td>
<td>4.5</td>
</tr>
<tr>
<td>BLAST</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>uncoded 2:2</td>
<td>12.0</td>
<td>4.0</td>
<td>18.0</td>
</tr>
<tr>
<td>coded 2:2</td>
<td>7.1</td>
<td>2.0</td>
<td>10.1</td>
</tr>
</tbody>
</table>

Table 1: Simulated performance of STC algorithms

Notes:

a) Notation $N_T:N_R$ is used to denote numbers of antennas at each end of the link.
b) $E_S$ denotes energy per transmitted STS
c) The benchmark is a 1/2-rate k=9 binary convolutional encoder [13], with output bits mapped to QPSK symbols. In a static 1:1 channel it achieves 10% FER for an $E_b/N_0$ of 2.2dB and 1% BER for an $E_b/N_0$ of 1.7dB.
d) The benchmark 1:4 number is approximate, as it has been extrapolated from results at a lower FER.
e) 'STTD' is a concatenation of the benchmark encoder with the STTD block code [14].

f) 'STTC-32' is the Tarokh-Seshradi-Calderbank STTC with 32 states, and two information bits per STS

g) 'MC-STTC-256' is a concatenation of the benchmark scheme, a QPSK modulation mapper, and a demux mapper to the different transmit antennas, as described above. It thus has $2^{k_1}=256$ states in the trellis, and a spectral efficiency which equals the number of transmit antennas.

h) The variants of BLAST shown in the table use 'genie-aided' subtraction of the transmission from the stronger transmit antenna when detecting the transmission from the weaker.

i) The 'uncoded' BLAST case uses no error correction code (it is raw QPSK), hence it achieves a high spectral efficiency of 4 bps/Hz. The 'coded' case uses two separate 'benchmark' encoders, one for each antenna.

Examining the results of Table 1 we see that the benchmark has poor performance in the 1:1 case, with a required $E_b/N_0$ of 11.0dB for 10% FER. This $E_b/N_0$ is 8.8dB higher (i.e. worse) than the static channel performance, and is due to the high fade margin required due to the lack of diversity. This performance can be improved either by increasing the number of receiver antennas (yielding both SNR gain and diversity gain) or by increasing the number of transmit antennas and using STTD, or both. In the latter case, for STTD 2:2, we see a required $E_b/N_0$ of only 2.2dB. For the case of two receiver antennas, this 2.2dB $E_b/N_0$ is the best result for the candidate algorithms considered. However, it is achieved at the expense of a poor spectral efficiency of only 1bps/Hz. We can double the spectral efficiency to 2bps/Hz using the STTC-32 code, but at the cost of a 2.4dB degradation in mean $E_b/N_0$. If we increase the number of code states to 256 using the MC-STTC-256, such that the number of states is now equal to that in the benchmark case, we improve our 2:2 performance to an $E_b/N_0$ of 3.5dB. Thus, compared to the STTD 2:2 case we have a degradation of only 1.3dB, but have achieved twice the spectral efficiency.

For the MC-STTC-256 case, if we now extend our receiver to have four antennas at each end of the link, we can further improve both our power efficiency, achieving an $E_b/N_0$ of -1.5dB, and our spectral efficiency, achieving 4bps/Hz. So even using this simple Space-Time code, we see that when we go from an STTD 2:1 scheme to an MC-STTC-256 4:4 we get 8.7dB $E_b/N_0$ reduction at a 4x increase in spectral efficiency.

The uncoded BLAST scheme shown in Table 1 performs poorly in terms of $E_b/N_0$, but has a high spectral efficiency due to the absence of redundancy, and the use of parallel transmission paths. Looking at the coded BLAST result in Table 1 we see that by application of coding the $E_b/N_0$ performance of BLAST has been improved by 4.9dB, down from 12.0dB to 7.1dB. However, this is achieved at the expense of a reduction in spectral efficiency. The notional 'order' of diversity is somewhere between 1 and 2, since the order of diversity is 1 for the detection of the stronger signal, and 2 for detection of the weaker. Therefore we would naturally expect the performance of coded BLAST to lie somewhere between that of the baseline 1:1 and 1:2 cases (11.0dB and 4.5dB $E_b/N_0$ respectively), and hence the simulation result of 7.1dB seems intuitively reasonable. Note that this result relies on subtraction performance approaching that of a 'magic genie'. Preliminary simulation investigations for non-genie-aided BLAST indicate that in practice the performance of coded BLAST 2:2 is in fact some 2-3dB worse than this, due to the effect of error propagation down the layers. This would suggest that the transmission from the stronger transmit antenna (for any given frame) is (despite its higher transmit power) actually more vulnerable (and therefore has a higher frame error probability) than the transmission from the weaker, due to the 'noise amplification' in the receiver antenna nulling process.
Based on the above reasoning, we would not expect even coded BLAST to reach the performance of baseline 1:2, although we must bear in mind that it is achieving twice the spectral efficiency. Finally, we note that further performance gains could possibly be achieved by adding more complexity into the BLAST transmitter and receiver processing, for example by a) time-switching the coded data frames across the two transmit antennas, as proposed by Foschini [8] and/or b) reiterating the decoding process by subtracting the transmission of weaker transmit antenna away from the received signal, and re-attempting detection of the transmission from the stronger transmit antenna (step and repeat until detection convergence).

Finally, we represent the results of Table 1 graphically in Figure 4. Since it is desirable to have codes which are at once both power efficient (low Eb/No) and spectrally efficient (high bps/Hz), we seek codes towards the top left hand side of the graph. However, to achieve this we will also need higher numbers of antenna elements at each end of the link.
6. CONCLUSIONS

For applications where good diversity performance on the Basestation-Mobile link is required, STTD is seen to offer good performance, and indeed of all the approaches considered for the 2:2 antenna configuration it performs the best in terms of its required Eb/No. However, it does not offer a particularly high spectral efficiency, and thus its use will be restricted to systems with a capacity which is fundamentally interference-limited rather than bandwidth-limited.

Of the more spectrally-efficient Space-Time Coding approaches simulated, the STTC techniques offer the best trade-off of performance versus complexity (when considered purely in terms of number of encoder/decoder states) for the applications of interest. A simply-derived STTC has been presented (namely the MC-STTC-256), based on the 1/2-rate convolutional encoder, which provides extremely good performance.

The layered (BLAST) codes, as described in this paper, offer poorer performance. This is largely due to the loss of diversity, which arises due to the antenna nulling process, and error-propagation effects due to imperfect subtraction. It is believed that coded BLAST schemes will always prove to be inferior in performance to STTC due to this diversity loss, and because of the fact that the decoders are inefficient, since they do not 'share' information in order to carry out joint estimation of the separate transmitted layers. However, we would note that when the number of antennas at each end of the link becomes large, these BLAST schemes will obtain certain advantages in terms of receiver processing complexity.

We note that the simulation results presented hereabove make a number of simplifying assumptions with regard to issues such as the fading correlation between antennas (a function of the 'richness' of the MIMO scattering channel), real-life system impairments (e.g. channel estimation error [22]) etc. In addition to exploring new and improved code designs, consideration of these associated real-world issues is also an area ripe for further investigation, and one which is key to the process of transferring Space-Time Codes from the academic domain into real-world application within future-generation wireless communications networks.

7. REFERENCES


