Performance Analysis of Concatenated Space-Time Coding with Two Transmit Antennas

Henrik Schulze, *Member, IEEE*

Abstract

In this paper, we show how Alamouti's [1] simple but useful transmit diversity scheme for two antennas can be combined with a standard outer error correcting code to achieve a stronger concatenated space-time coding (STC) scheme. By introducing a matrix formalism that allows to interpret the transmission channel as a rotation in an Euclidean space, it can be easily shown that this scheme with 2 transmit (TX) and L_r receive (RX) antennas is equivalent to a simple receive antenna setup with $2L_r$ receive antennas. Analytical formulas for pair error probabilities will be derived for the time and/or frequency flat fading and for the ideally interleaved Rayleigh fading channel as well as for the correlated fading channel. As a practical example, we study how the performance of a Walsh-Hadamard coded multi-carrier (MC-) CDMA system depends on the correlation bandwidth of the channel and the number of RX and TX antennas.

Keywords

Channel coding, diversity, fading channels, space-time coding, wireless communications.

I. Introduction

RELIABLE transmission over a fading channel requires some kind of *diversity*, i.e. the information must be spread over the channel in such a way that one deep fade will not cause an information loss. The simplest method is *repetition diversity*, i.e. the same information will be received at L different frequencies, time slots, or antennas. Channel coding together with interleaving is more efficient. For a code with Hamming distance d, the same power law SNR^{-d} for the bit error rate (BER) as a function of the signal-to-noise ratio (SNR) can be reached with less bandwidth requirements.

It has to be kept in mind that the code can only exploit its maximum diversity degree d if the code symbols are affected by statistically independent fading amplitudes. This means that the information has to be spread in time and/or frequency by a sufficiently large interleaver. In many cases, time interleaving is limited by the delay that can be tolerated by the application. For multi-carrier systems, frequency interleaving is very useful, but it provides full diversity only if the signal bandwidth B is significantly larger than the correlation bandwidth B_{corr} of the channel. On the other hand, for high frequencies as used by most wireless communication systems, a sufficient spatial separation of antennas is often possible, at least at the base station site. In the last few years, much work has been spent on the concept of transmitter diversity and space-time codes (STC) to take advantage of this possibility, see e.g. [1], [2], [3], [4], [5], [6], [7], [8], [9] and references therein.

In such a system, the information is encoded and the code symbols are split up on L_t transmit antennas in such a way that the receiver with L_r receive antennas can recover the information. The maximum diversity degree that can be obtained by this setup is given by $L = L_t L_r$ (see e.g. [4]). In contrast to a simple receive antenna repetition diversity setup, there exist space-time codes with multiple transmit antennas that yield a positive coding gain. The price for that gain is an additional bandwidth expansion.

One very simple and remarkable STC scheme for two transmit antennas with the maximum diversity gain, but with no coding gain and no bandwidth expansion has been introduced by Alamouti [1]. In fact, it can be shown that the Alamouti scheme with $L_t = 2$ transmit antenna and L_r receive antenna is mathematically equivalent to a setup with one transmit and $2L_r$ receive antennas. We will discuss this in detail below.

Instead of looking for STCs with a high coding gain, one may use standard error correcting codes to obtain the coding gain and then combine it with the Alamouti scheme to increase the diversity degree. The benefit

The Author is with the University of Applied Sciences Südwestfalen, D-59872 Meschede, Germany. Telephone: (+49) 291 9910-300, Fax: (+49) 291 9910-40. E-mail: schulze@fh-meschede.de.

then is a wide choice among all the well-known highly efficient error correcting codes. Furthermore, the usual decoders can be applied. Our goal is to investigate the performance of such a system.

In this paper, we will first re-derive the Alamouti scheme in a different formalism. We will describe the channel by a matrix $\mathbf{C}$ that depends on the complex fading coefficients c_l and which acts on the vector $\mathbf{s}$ of a transmitted pair (s_1, s_2) of complex symbols. The underlying structure of the transmission scheme can be understood by the properties of the channel matrix. For the simplest case of $L_t = 2$ and $L_r = 1$, $\mathbf{C}$ can be shown to be a unitary 2×2-matrix multiplied by a factor $\sqrt{|c_1|^2 + |c_2|^2}$. This means that the transmission channel can be interpreted as a rotation in the signal space of four real (i.e. two complex) dimensions, followed by an attenuation factor whose power corresponds to the sum of the two transmission branch powers. It turns out that this is a straightforward generalization of the analysis of the two-antenna maximum ratio combining (MRC) receiver.

A simple method to achieve a coding gain is to enhance the Euclidean distance between the possible signal vectors $\mathbf{s}$ by reducing the set of possible signals to a smaller subset. For example, one can reduce the set of 16 possible Gray labeled QPSK symbol pairs to the subset of 8 with odd parity of their labeling bits. This signal set reduction can be interpreted as an additional outer coding with a simple (4,3) single parity check (SPC) code. As an interesting by-product we can show that this code is equivalent to the so-called *quaternion* space-time code introduced by Hughes [6]. This simple example is the starting point for the construction of more powerful space-time codes based on concatenation.

Without any additional channel coding, the Alamouti scheme alone does not show an impressive performance because the number of receive antennas is restricted for practical reasons, so that only a small diversity degree can be achieved. Furthermore, it provides no coding gain. Therefore, one should regard it as an aid to a conventional coding scheme to overcome the problems due to insufficient interleaving.

We observe that a setup which combines the Alamouti transmission scheme with any familiar error correcting code is equivalent to the combination of that code with repetition diversity. Also, we derive expressions for the pair error probabilities of such a concatenated space-time coding scheme in a Rayleigh fading channel for three cases: For flat fading, i.e. the fading amplitude does not vary in time or frequency during the transmission of one code word; for independent fading amplitudes, which corresponds to ideal interleaving; and, finally, we give an expression for the intermediate and most realistic case of correlated fading amplitudes. This can be achieved by a Karhunen-Loeve transform of the vector of correlated fading amplitudes to a vector of uncorrelated fading amplitudes. For simplicity, this analysis has been restricted to BPSK and QPSK transmission, but the same methods can be extended, e.g. to QAM transmission.

As a concrete application, we consider multi-carrier (MC-) CDMA transmission [10], [11] with Walsh-Hadamard (WH) codes [12], [13], [14] in a frequency-correlated fading channel.

This paper is organized as follows: In Section II, we introduce a matrix formalism that is especially suited for the Alamouti TX diversity scheme. It allows a simple geometrical analysis that gives a visual idea how the optimum receiver has to be constructed and how performance curves can be derived. In Section III, we introduce the concatenated STC setup and analyze the performance for different types of fading channels. In Section IV, a practical example for such a concatenated STC scheme will be given and analytical performance curves will be presented. In Section V, some conclusions will be drawn.

II. Analysis of Alamouti's Space-Time Coding Scheme

A. Basic Transmission Scheme

We first consider a transmission scheme with one antenna at the receiver and two antennas at the transmitter as proposed in [1]. The antennas at the transmitter are assumed to be spatially separated so that fading of the two transmission paths can be regarded as statistically independent. We consider a Rayleigh fading channel, i.e. the corresponding complex fading amplitudes c_1 and c_2 are complex Gaussian random variables with mean zero and average power one.

A pair of modulated complex symbols (s_1, s_2) is transmitted in the following way: At time slot 1 the symbol s_1 is transmitted from antenna 1 and s_2 is transmitted from antenna 2. The received signal (without noise) at time slot 1 is then given by $c_1 s_1 + c_2 s_2$. At time slot 2, the symbol s_2^* is transmitted from antenna 1 and $-s_1^*$

is transmitted from antenna 2. We assume that the fading coefficients remain constant during the two time slots. The received signal at time slot 2 is then given by $-c_2 s_1^* + c_1 s_2^*$. For our convenience and to introduce the matrix formalism, we take the complex conjugate of the received symbol in the second time slot before any other further processing at the receiver is done. We can therefore say that, at time slot 2, s_1 and s_2 have been transmitted over the channel branches with fading coefficients $-c_2^*$ and c_1^*, respectively. Fig. 1 shows this equivalent setup.

We now introduce a vector/matrix formalism and write the two symbols s_1 and s_2 as a column vector $\mathbf{s}$ and the received symbols r_1 and r_2 as a column vector $\mathbf{r}$. This vector can be written as

$$\mathbf{r} = \mathbf{C}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{n}$ is an additive white Gaussian noise vector with noise variance $\sigma^2 = N_0/2$ per real dimension and

$$\mathbf{C} = \begin{pmatrix} c_1 & c_2 \\ -c_2^* & c_1^* \end{pmatrix}. \tag{2}$$

is the *channel matrix*. We observe that this matrix has the property

$$\mathbf{C}^\dagger \mathbf{C} = \mathbf{C}\mathbf{C}^\dagger = \left(|c_1|^2 + |c_2|^2 \right) \mathbf{I}_2, \tag{3}$$

where

$$\mathbf{C}^\dagger = \begin{pmatrix} c_1^* & -c_2 \\ c_2^* & c_1 \end{pmatrix} \tag{4}$$

is the Hermitian conjugate (=transposed and complex conjugate) of the matrix $\mathbf{C}$, and $\mathbf{I}_2$ is the 2×2 identity matrix. $|c_1|^2 + |c_2|^2$ is the total power of the composed fading channel belonging to two transmit antennas. Equation (3) means that the channel matrix $\mathbf{C}$ is of the form

$$\mathbf{C} = \sqrt{|c_1|^2 + |c_2|^2}\, \mathbf{U}, \tag{5}$$

where $\mathbf{U}$ is a unitary matrix (i.e. a matrix of the form $\mathbf{U}^\dagger \mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = \mathbf{I}_2$). Unitary matrices (like orthogonal matrices for real vector spaces) are invertible matrices that leave Euclidean distances invariant. They can be visualized as rotations (possibly combined with a reflection) in an Euclidean space. This means that the transmission channel given by eq. (1) can be separated into three parts:

1. A rotation in 2 complex (= 4 real) dimensions.
2. An attenuation by the composed fading amplitude $\sqrt{|c_1|^2 + |c_2|^2}$.
3. An AWGN channel.

Keeping in mind that multiplicative fading is just a phase rotation together with an attenuation by a real fading amplitude, we can now interpret this transmission according to eq. (1) with a matrix of type (2) as a generalization of the familiar multiplicative fading from one to two complex dimensions. Because a rotation leaves Euclidian distances invariant, this transmission scheme behaves just like generalized multiplicative fading with a higher-dimensional rotation and an attenuation given by $\sqrt{|c_1|^2 + |c_2|^2}$.

Before exploiting this idea in more detail, we will generalize the matrix formalism to more than one receive antenna.

B. Extension for Multiple Receive Antennas

We now consider L_r receive antennas and the same transmission scheme. This situation can be described by eq. (1) as well, if we define

$$\mathbf{r} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \cdots \\ \mathbf{r}_{L_r} \end{pmatrix}, \tag{6}$$

where $\mathbf{r}_1 = (r_1, r_2)^T$ is the vector of symbols received at the first antenna, $\mathbf{r}_2 = (r_3, r_4)$ is the vector of symbols received at the second antenna, and so forth. The matrix $\mathbf{C}$ has the following block structure

$$\mathbf{C} = \begin{pmatrix} \mathbf{C_1} \\ \mathbf{C_2} \\ \dots \\ \mathbf{C}_{L_r} \end{pmatrix}, \tag{7}$$

with each sub-matrix given by

$$\mathbf{C}_l = \begin{pmatrix} c_{2l-1} & c_{2l} \\ -c_{2l}^* & c_{2l-1}^* \end{pmatrix}. \tag{8}$$

Here, c_{2l-1} is the fading coefficient of transmission branch $2l-1$ from transmit antenna 1 to receive antenna l, and c_{2l} is the fading coefficient of transmission branch $2l$ from transmit antenna 2 to receive antenna l. Fig. 2 shows this numbering of transmission branches for $L_r = 2$. In the following, we will always assume that the fading coefficients of different branches are statistically independent. This is justified by a sufficient spatial separation of the antennas. Each sub-matrix $\mathbf{C}_l$ has the property

$$\mathbf{C}_l^\dagger \mathbf{C}_l = \mathbf{C}_l \mathbf{C}_l^\dagger = \left(|c_{2l-1}|^2 + |c_{2l}|^2\right) \mathbf{I}_2 \tag{9}$$

and can therefore be written as a scalar multiplied by a unitary matrix.

We now also include the case of only $L_t = 1$ transmit antenna into the formalism and define $L = L_r L_t$, with $L_t \in \{1, 2\}$. We can still describe this situation by eq. (1) with $\mathbf{r} = (r_1, r_2, ..., r_L)^T$. For only one transmit antenna, $\mathbf{s}$ reduces to a scalar s, and we thus have to consider only one time slot at the receiver. Then r_l is the symbol received at antenna number l. In either case, the channel matrix is given by eq. (7): For $L_t = 2$, $\mathbf{C}_l$ is given by eq. (8), while for only one transmit antenna, $\mathbf{C}_l$ reduces to the scalar c_l, the complex fading amplitude for receive antenna number l.

This common formal treatment allows us to show that the Alamouti space time coding scheme is symmetric with respect to transmit and receive antennas. We will see that the performance depends only on the number L: One receive and two transmit antennas are equivalent to one transmit and two receive antennas, and two transmit and two receive antennas are equivalent to one transmit and four receive antennas. This is due to the fact that

$$\mathbf{C}^\dagger \mathbf{C} = \sum_{l=1}^{L} |c_l|^2 \cdot \mathbf{I}_2 \tag{10}$$

for either case.

C. Diversity Combining at the Receiver

We consider the general case that the transmitted signals are given by column vectors $\mathbf{s}$ and the received signals by column vectors $\mathbf{r}$ that can be written as

$$\mathbf{r} = \mathbf{C}\mathbf{s} + \mathbf{n}, \tag{11}$$

where $\mathbf{C}$ is a matrix that describes the channel and $\mathbf{n}$ is the AWGN vector with noise variance $\sigma^2 = N_0/2$ per real dimension.

It follows from the maximum likelihood principle that the most probable transmitted signal $\mathbf{s}$ is the one that minimized the squared Euclidean distance $\|\mathbf{r} - \mathbf{C}\mathbf{s}\|^2$, where the minimum has to be taken over the set of all possible transmitted signals $\mathbf{s}$. This is equivalent to maximizing the metric $m(\mathbf{r}, \mathbf{s}|\mathbf{C})$

$$m(\mathbf{r}, \mathbf{s}|\mathbf{C}) := \Re\{< \mathbf{C}^\dagger \mathbf{r}|\mathbf{s} >\} - \frac{1}{2}\|\mathbf{C}\mathbf{s}\|^2 \tag{12}$$

Here $< .|. >$ denotes the scalar product in the (complex) vector space of signals. The first term in the metric can be interpreted as a cross correlation of the linearly preprocessed receive vector $\mathbf{C}^\dagger \mathbf{r}$ with all possible transmit vectors $\mathbf{s}$, and the second term is just half the received symbol energy (without noise).

In our special case where we have a channel matrix that satisfies the condition (10), the energy term can be simplified, and the metric can then be written as

$$m(\mathbf{r},\mathbf{s}|\mathbf{C}) = \Re\{< \mathbf{C}^\dagger\mathbf{r}|\mathbf{s} >\} - \frac{1}{2}\sum_{l=1}^{L}|c_l|^2\|\mathbf{s}\|^2. \tag{13}$$

If all possible transmit vectors $\mathbf{s}$ have the same length (as is the case for PSK signaling), the second (energy) term does not influence the decision so that it can be ignored. The first term is the real part of a scalar product in a 2-dimensional complex vector space which can be interpreted as a scalar product in a 4-dimensional real vector space (note that with $\mathbf{z} = \mathbf{z}_\Re + \mathrm{j}\mathbf{z}_\Im$ and $\mathbf{v} = \mathbf{v}_\Re + \mathrm{j}\mathbf{v}_\Im$, the relation $\Re\{< \mathbf{z}|\mathbf{v} >\} = \mathbf{z}_\Re \cdot \mathbf{v}_\Re + \mathbf{z}_\Im \cdot \mathbf{v}_\Im$ holds, where the scalar product between real vectors is denoted by a dot). The first term can be interpreted as follows: $\mathbf{C}^\dagger$ acts as a diversity combiner on the received vector $\mathbf{r}$. The combiner output

$$\mathbf{C}^\dagger\mathbf{r} = \sum_{l=1}^{L_r}\mathbf{C}_l^\dagger\mathbf{r}_l \tag{14}$$

will be compared with all possible transmit vectors $\mathbf{s}$ and, for equal symbol energy, the one with the greatest cross-correlation (scalar product)

$$\Re\{< \mathbf{C}^\dagger\mathbf{r}|\mathbf{s} >\} = (\mathbf{C}^\dagger\mathbf{r})_\Re \cdot \mathbf{s}_\Re + (\mathbf{C}^\dagger\mathbf{r})_\Im \cdot \mathbf{s}_\Im. \tag{15}$$

will be taken. It is obvious that eq. (14) is a generalization of the well-known maximum-ratio combiner (MRC) for PSK: For one transmit antenna, each term $\mathbf{C}_l^\dagger\mathbf{r}_l$ reduces to $c_l^* r_l$, i.e. the received phasors r_l have to be rotated back and weighted with the channel amplitude, while for two transmit antennas,

$$\mathbf{C}^\dagger{}_l\mathbf{r}_l = \begin{pmatrix} c_{2l-1}^* r_{2l-1} - c_{2l} r_{2l} \\ c_{2l}^* r_{2l-1} + c_{2l-1} r_{2l} \end{pmatrix} \tag{16}$$

can be interpreted as the back rotation of a generalized phasor $\mathbf{r}_l$ in 4 real dimensions, weighted with the amplitude $\sqrt{|c_{2l-1}|^2 + |c_{2l}|^2}$. We have, thus, demonstrated that this combiner, ingeniously hand-crafted by Alamouti [1], is the natural generalization of the MRC receiver and can be derived from general principles.

Fig. 3 shows the block diagram of this generalized MRC receiver. Here, the case of a non-constant energy term, as would be necessary e.g. for QAM signaling, has been included. At the combiner input, the receive vectors (generalized phasors) $\mathbf{r}_l$ are back-rotated and weighted by the multiplication with matrices $\mathbf{C}_l^\dagger$ and then summed up. The sum will be cross-correlated with all possible transmit vectors $\mathbf{s}$ and, for the case of non-constant energy, half the vector energy has to be subtracted, weighted with the total transmission power $\sum_{l=1}^{L}|c_l|^2$. Then the maximum of the resulting metric values has to be taken. In many typical practical cases, decision variables for individual bits can be obtained more easily. For QPSK with Gray mapping, each bit corresponds to the sign for one real dimension, so four decision variables have to be considered instead of 16 metric values. For BPSK, the situation is even simpler. For higher level PSK and for QAM signaling, simple decision thresholds can be derived as well by using geometrical arguments. We note that this generalized MRC combiner does not only allow symbol decisions. It also provides soft outputs that may be used for an outer decoder. In the general case, these have to be calculated from all metric values as log-likelihood ratios (LLRs) for the bits. For BPSK and QPSK signaling, these LLRs reduce to the soft MRC output values for each dimension, each corresponding to the respective bit.

D. Error Probability Analysis

We consider the optimum receiver derived above and want to calculate the pair error probability (PEP) $P(\mathbf{s} \to \hat{\mathbf{s}})$ of an error that happens, if the signal vector $\mathbf{s}$ has been transmitted, but the receiver decides for another vector $\hat{\mathbf{s}}$. The conditional error probability for a fixed channel matrix $\mathbf{C}$ is given by the Gaussian probability integral (see e.g. [17], Chap. 4 and 13)

$$P(\mathbf{s} \to \hat{\mathbf{s}}|\mathbf{C}) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{1}{4N_0}\|\mathbf{C}\mathbf{s} - \mathbf{C}\hat{\mathbf{s}}\|^2}\right). \tag{17}$$

$P(\mathbf{s} \to \hat{\mathbf{s}})$ will be obtained by averaging over the fading amplitudes. Using eq. (10) and defining

$$\Delta = \|\mathbf{s} - \hat{\mathbf{s}}\|/2 \tag{18}$$

we can write this as

$$P(\mathbf{s} \to \hat{\mathbf{s}}|\mathbf{C}) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\sum_{l=1}^{L} |c_l|^2 \frac{\Delta^2}{N_0}}\right) \tag{19}$$

and the pair error probability is given by

$$P(\mathbf{s} \to \hat{\mathbf{s}}) = \mathrm{E}_c\left\{\frac{1}{2}\mathrm{erfc}\left(\sqrt{\sum_{l=1}^{L} |c_l|^2 \frac{\Delta^2}{N_0}}\right)\right\} \tag{20}$$

where $\mathrm{E}_c\{\cdot\}$ is the expectation over all fading coefficients inside the matrix $\mathbf{C}$. We assume independent and identically distributed Rayleigh fading amplitudes $|c_l|^2$ with average power normalized by $\mathrm{E}\{|c_l|^2\} = 1$. The sum of squared amplitudes can be shown to be χ^2- distributed and the average can be performed by evaluating one integral, leading to the well established formula for diversity with L independent branches, see [18]. Recently, it has been pointed out by [15] that this problem can also be treated in a very elegant way by using the polar representation of the Gaussian probability integral given by

$$\frac{1}{2}\mathrm{erfc}(x) = \frac{1}{\pi}\int_0^{\pi/2} \exp\left(-\frac{x^2}{\sin^2\theta}\right) d\theta. \tag{21}$$

With this formula, the averaging over the L Rayleigh distributed fading amplitudes c_l can be easily performed as described in detail in [15], leading to the expression

$$P(\mathbf{s} \to \hat{\mathbf{s}}) = \frac{1}{\pi}\int_0^{\pi/2} \left(\frac{1}{1 + \frac{1}{\sin^2\theta}\frac{\Delta^2}{N_0}}\right)^L d\theta. \tag{22}$$

This integral over a finite interval can be solved using the residue theorem. For practical purposes, it is also very easy to evaluate it numerically. It turns out that this integral representation of error probabilities is often more flexible. For this reason, we prefer it for our subsequent analysis. We also note that a Chernoff bound can easily be obtained from eq. (22) by upper-bounding the integrand by the maximum at $\theta = \pi/2$, leading to

$$P(\mathbf{s} \to \hat{\mathbf{s}}) \le \frac{1}{2}\left(\frac{1}{1 + \frac{\Delta^2}{N_0}}\right)^L. \tag{23}$$

For high signal-to-noise ratios, we can use the inequality

$$\frac{1}{1 + \frac{1}{\sin^2\theta}\frac{\Delta^2}{N_0}} \le \sin^2\theta \left(\frac{\Delta^2}{N_0}\right)^{-1}$$

to upper bound the integrand in eq. (22). The integral can then be solved resulting in the asymptotically tight upper bound

$$P(\mathbf{s} \to \hat{\mathbf{s}}) \le \frac{1}{2}\frac{1}{4^L}\binom{2L}{L}\left(\frac{\Delta^2}{N_0}\right)^{-L} \tag{24}$$

which means that the error probability decays with the power L of the inverse SNR for L- fold repetition diversity.

Now let E_b be the total energy per data bit available at the receiver and E_S the energy per complex transmit symbol s_1 or s_2. We assume M-ary modulation, so each of them carries $\log_2(M)$ data bits. Both symbols are assumed to be of equal (average) energy which means that $E_S = \mathrm{E}\{|s_1|^2\} = \mathrm{E}\{|s_2|^2\} = \mathrm{E}\{\|\mathbf{s}\|^2/2\}$. We have assumed that $\mathrm{E}\{|c_l|^2\} = 1$. Then, for each time slot, the total energy $L_t E_S$ is transmitted at all L_t antennas together and the same (average) energy is available at each of the L_r receive antennas. Therefore, the total energy available at the receiving site for that time slot is

$$L_r L_t E_S = \log_2(M) E_b. \tag{25}$$

For only one transmit antenna, $SNR = E_S/N_0$ is the SNR at the receive antenna for Nyquist pulse forming as well as for OFDM multi-carrier transmission. For L_t transmit antennas, this generalizes to $SNR = L_t E_S/N_0$. All the above expressions for the error probability as a function of E_b/N_0 depend only on $L = L_r L_t$, i.e. using Alamouti's transmit diversity scheme with 2 antennas is equivalent to one transmit antenna and twice as many receive antennas [1]. This means that the scheme achieves the full diversity degree, but no coding gain. For uncoded BPSK or QPSK transmission, the value of each data bit affects only one real dimension. The event of an erroneous bit decision corresponds to the squared half-distance

$$\Delta^2 = E_S/\log_2(M) = E_b/L, \tag{26}$$

which means that BPSK ($M = 2$) and QPSK ($M = 4$) have the same E_b/N_0 performance. Inserting this into eq. (22), we observe that the error rate for the AWGN channel given by the Gaussian probability integral (21) will be reached in the limit $L \to \infty$ since $(1 + x/L)^L \to e^x$ for all x and the limit can be performed under the integral.

E. The Quaternion Code as a Simple Example for Concatenated Coding

One possibility to achieve a coding gain is to increase the squared half-distance Δ^2 by reducing the signal alphabet of possible vectors to be transmitted according to the Alamouti scheme. If, for QPSK transmission with conventional Gray mapping, only those pairs of QPSK symbols (s_1, s_2) will be used that correspond to labeling bit quadruples (a_1, a_2, a_3, a_4) with even parity, the symbol alphabet reduces from 16 to 8. Since every two vectors $\mathbf{s}$ and $\hat{\mathbf{s}}$ now differ at least in two real dimensions, it can easily be seen that the squared half-distance for the closest error event increases by a factor of two from $\Delta^2 = E_S/2$ to $\Delta^2 = E_S$. Since there are only 1.5 data bits per transmitted QPSK symbol, $E_S = 1.5 E_b/L$ holds. The coding gain of this SPC(4,3,2) code with Hamming distance $d = 2$ and rate $R_c = 3/4$ is just $dR_c = 1.5$ or 1.76 dB.

It is interesting to note that this simple concatenation of Alamouti's STC and an outer SPC code is equivalent to the so-called *quaternion* space-time code investigated by Hughes [6] in the context of group codes. This can be seen as follows: If we rotate a conventional QPSK constellation with Gray mapping by $\pi/4$, i.e. use $s_1, s_2 \in \{\pm 1, \pm j\}$, the possible 8 transmit vectors of even parity are

$$\mathbf{s} \in \left\{ \pm \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \pm \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \pm \begin{pmatrix} j \\ j \end{pmatrix}, \pm \begin{pmatrix} j \\ -j \end{pmatrix}, \right\}, \tag{27}$$

which are just the possible transmit vectors of the quaternion code introduced in [6] [2].

III. Concatenated Space-Time Coding

A. The Basic Transmission Scheme

We generalize the idea of the above example and consider arbitrary linear error correcting outer codes. The encoded bits of a given encoder will be mapped on a sequence $\{\mathbf{s}(i)\}_{i=1}^{N}$ of two-dimensional vectors of BPSK or QPSK symbols with Gray mapping. If the channel is time- and/or frequency-variant, the diversity of the code can be exploited better if two bits of the same code word are mapped on different vectors $\mathbf{s}(i)$. For QPSK, e.g., this can be realized by multiplexing four code words of length N in such a way that each vector $\mathbf{s}(i)$ is labeled by a quadruple (a_1, a_2, a_3, a_4) of bits from different code words. If an additional delay is allowed, this can be combined with an interleaver to take more advantage of statistical independence. To apply this idea to

[1]Note that this holds only if we plot the bit error rates as a function of E_b/N_0. There is a 3 dB difference between the curves of twofold receive and twofold transmit diversity if they are plotted as a function of $SNR = (log_2 M/L_r) \cdot E_b/N_0$, as done by [1]

[2]In [6] a 3 dB gain has been stated, which is due to the fact that SNR and not E_b/N_0 is considered.

the quaternion code of the above example, four parallel SPC(4,3,2) code words are multiplexed and mapped on four vectors $\mathbf{s}(1)$, $\mathbf{s}(2)$, $\mathbf{s}(3)$, and $\mathbf{s}(4)$. If these vectors can be ideally separated by an interleaver, the code exploits its diversity degree $d = 2$. Together with the diversity degree 2 of the Alamouti scheme for one RX antenna, this leads to a total diversity degree of 4 instead of 2 for the quaternion code.

The basic transmission scheme is depicted in Fig. 4. The bits that leave the outer encoder are multiplexed by MUX (with an optional interleaver) as described above. Then the TX mapper maps each pair (quadruple) of bits on a 2-dimensional vector of BPSK (QPSK) symbols and transmits it over two antennas according to the Alamouti scheme. The channel is given by

$$\mathbf{r}(i) = \mathbf{C}(i)\mathbf{s}(i) + \mathbf{n}(i), \tag{28}$$

where the index i stands for the time/frequency variance. We assume that there is no such variance during the transmission of one 2-dimensional vector $\mathbf{s}(i)$. Then the channel matrix $\mathbf{C}(i)$ with block structure as in eq. (7) has sub-matrices

$$\mathbf{C}_l(i) = \begin{pmatrix} c_{2l-1}(i) & c_{2l}(i) \\ -c_{2l}^*(i) & c_{2l-1}^*(i) \end{pmatrix} \tag{29}$$

that lead to the property

$$\mathbf{C}^\dagger(i)\mathbf{C}(i) = \sum_{l=1}^{L} |c_l(i)|^2 \cdot \mathbf{I}_2. \tag{30}$$

At the receiver, the generalized MRC combiner decides for the sequence with the maximum metric

$$\sum_{i=1}^{N} m(\mathbf{r}(i), \mathbf{s}(i)|\mathbf{C}(i)), \tag{31}$$

where $m(\mathbf{r}(i), \mathbf{s}(i)|\mathbf{C}(i))$ is the MRC metric of the Alamouti scheme given by eq. (13). This means that for BPSK (QPSK), the two (four) decision variables obtained from the generalized weighted back-rotation of the diversity combiner as discussed in Section II-C are the optimum soft inputs to the outer channel decoder. For higher level modulation schemes and/or other mappings (e.g. TCM with 8-PSK), the metric computation is more involved and requires the computation of log-likelihood ratios. These expressions can be easily derived from the formulas in Section II-C.

B. Error Probability Analysis

The pair error probability is given by

$$P\left(\{\mathbf{s}(i) \to \hat{\mathbf{s}}(i)\}_{i=1}^{N}\right) = \mathrm{E}_c\left\{\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{1}{4N_0}\sum_{i=1}^{N}\|\mathbf{C}(i)(\mathbf{s}(i)-\hat{\mathbf{s}}(i))\|^2}\right)\right\}, \tag{32}$$

where $\mathrm{E}_c\{\cdot\}$ denotes the expectation over all channel matrices $\mathbf{C}(i)$. For brevity, we will now denote the sequence $\{\mathbf{s}(i)\}_{i=1}^{N}$ by the column vector $\mathbf{x} = (s_1(1), s_2(1), s_1(2), s_2(2), ..., s_1(N), s_2(N))^T$ of length $2N$. $\hat{\mathbf{x}}$ is defined similarly. Defining

$$\Delta(i) = \|\mathbf{s}(i) - \hat{\mathbf{s}}(i)\|/2 \tag{33}$$

and using eq. (30) we obtain

$$P\left(\mathbf{x} \to \hat{\mathbf{x}}\right) = \mathrm{E}_c\left\{\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{1}{N_0}\sum_{i=1}^{N}\sum_{l=1}^{L}|c_l(i)|^2\Delta^2(i)}\right)\right\}. \tag{34}$$

There are two extreme cases that we want to discuss separately before we treat the general case: If the channel is absolutely time-frequency flat, the whole code word experiences the same fading and the fading coefficients do not depend on the index i. In the other extreme, the fading coefficients for different values of i are statistically independent. We call this an ideally interleaved channel. The typical real situation will, of course, be somewhere in between.

B.1 Error Analysis for Flat Fading Channels

For this case we assume that the channel does not vary during the transmission of one code word of length N, i.e. $\mathbf{C}(i) = \mathbf{C}$ and $c_l(i) = c_l$ are the same for all values of i. Then eq. (34) simplifies to

$$P(\mathbf{x} \to \hat{\mathbf{x}}) = \mathrm{E}_c \left\{ \frac{1}{2} \mathrm{erfc} \left(\sqrt{\frac{1}{N_0} \sum_{l=1}^{L} |c_l|^2 \sum_{i=1}^{N} \Delta^2(i)} \right) \right\}. \tag{35}$$

This is just equation (20) in Section II-D with eq. (18) replaced by the definition

$$\Delta^2 = \sum_{l=1}^{N} \Delta^2(i). \tag{36}$$

Thus, all the formulas in Section II-D for the Rayleigh fading channel are valid with this replacement. For BPSK or QPSK and an error event of Hamming distance d, the sequences $\{\mathbf{s}(i)\}_{i=1}^{N}$ and $\{\hat{\mathbf{s}}(i)\}_{i=1}^{N}$ differ in d real dimensions and $\Delta^2 = dE_S/log_2(M)$. Using $LE_S = R_c \log_2(M)E_b$, we obtain

$$\Delta^2 = \frac{1}{L} dR_c E_b. \tag{37}$$

We observe a coding gain of $\lg(dR_c)$ decibels compared to (Alamouti) diversity, cf. eq. (26), while the diversity degree remains L. This is due to the fact that the code can not exploit any channel diversity. Therefore, the multiplexer/interleaver described above is useless for flat fading channels.

B.2 Error Analysis for Ideally Interleaved Channels

Let us now assume that all coefficients $c_l(i)$ are statistically independent complex Gaussian random variables. Let the transmitted and the detected sequences $\{\mathbf{s}(i)\}_{i=1}^{N}$ and $\{\hat{\mathbf{s}}(i)\}_{i=1}^{N}$ differ in exactly d positions $i = i_1, ..., i_d$, corresponding with an error event of Hamming distance d. Due to the multiplexing, for $i = i_1, ..., i_d$, the vectors $\mathbf{s}(i)$ and $\hat{\mathbf{s}}(i)$ differ in one real dimension by the Euclidean half-distance $\Delta := \|\mathbf{s}(i) - \hat{\mathbf{s}}(i)\|/2$ given by $\Delta^2 = E_S/log_2(M)$. For all other values of i, the difference equals zero by definition. The pair error probability is then given by

$$P(\{\mathbf{x} \to \hat{\mathbf{x}}) = \mathrm{E}_c \left\{ \frac{1}{2} \mathrm{erfc} \left(\sqrt{\frac{1}{N_0} \sum_{n=1}^{d} \sum_{l=1}^{L} |c_l(i_n)|^2 \Delta^2} \right) \right\}. \tag{38}$$

Since the fading coefficients are independent, this is the same as eq. (20), but with the sum of dL independent identically distributed coefficients $|c_l(i)|^2$ instead of the sum of L identically distributed coefficients $|c_l|^2$. All the formulas for the pair error probability in Section II-D apply with L replaced by dL and eq. (26) replaced by

$$\Delta^2 = \frac{1}{L} R_c E_b = \frac{1}{dL} dR_c E_b. \tag{39}$$

Obviously, the diversity degree of the concatenated scheme is $d \cdot L$, which the product of the diversity degree of the code and the antenna diversity degree. The coding gain is given by dR_c which is the gain of the outer code compared to a simple outer $d-$fold repetition code.

B.3 Error Analysis for Correlated Rayleigh Fading Channels

For the general correlated Rayleigh fading channel, we define the vector

$$\mathbf{c}_l = \begin{pmatrix} c_l(i_1) \\ c_l(i_2) \\ \cdots \\ c_l(i_d) \end{pmatrix}, \tag{40}$$

of the fading coefficients for transmission path number $l \in \{1, 2, ..., L\}$. The vector entries $c_l(i)$ are identically distributed, but correlated complex Gaussian random variables with mean zero and average power one. They are completely characterized by their autocorrelation. We have assumed that fading coefficients for different transmission paths l (i.e. different pairs of transmit and receive antennas) are independent and identically distributed, i.e.

$$\mathrm{E}\left\{\mathbf{c}_k\mathbf{c}_l^\dagger\right\} = \mathrm{E}\left\{\mathbf{c}_l\mathbf{c}_l^\dagger\right\}\delta_{kl} \tag{41}$$

holds and

$$\mathbf{R_{cc}} = \mathrm{E}\left\{\mathbf{c}_l\mathbf{c}_l^\dagger\right\} \tag{42}$$

is independent of l. Since the autocorrelation matrix $\mathbf{R_{cc}}$ is Hermitian, there exists a unitary matrix $\mathbf{V}$ such that $\mathbf{R_{cc}} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^\dagger$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, ..., \lambda_d)$ is the diagonal matrix of the eigenvalues. This is just a special case of the Karhunen-Loeve transform, see e.g. [19]. It can easily be shown that $\mathbf{b}_l = \mathbf{V}^\dagger\mathbf{c}_l$ is a vector of uncorrelated (and therefore independent) Gaussian random variables with mean zero. These unitary equivalent fading vectors $\mathbf{b}_l$ are statistically independent for different values of l and have the same statistics. For each fixed vector $\mathbf{b}_l$, the components $b_l(i)$ are uncorrelated and have unequal powers $\lambda_i = \mathrm{E}\left\{|b_l(i)|^2\right\}$ for different indeces i, but independent of l. Since $\mathbf{V}$ is a unitary matrix, it leaves the length of vectors invariant, i.e. $\|\mathbf{c}_l\| = \|\mathbf{b}_l\|$ or

$$\sum_{n=1}^{d}|c_l(i_n)|^2 = \sum_{n=1}^{d}|b_l(i_n)|^2 \tag{43}$$

for all $l \in \{1, ..., L\}$. We insert this into eq. (38) and obtain

$$P\left(\mathbf{x} \to \hat{\mathbf{x}}\right) = \mathrm{E}_c\left\{\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{1}{N_0}\sum_{n=1}^{d}\sum_{l=1}^{L}|b_l(i_n)|^2\Delta^2}\right)\right\}, \tag{44}$$

with Δ given by eq. (39). Because of the statistical independence of the coefficients $b_l(i_n)$, we can again use the method described in [15] and eventually obtain

$$P(\mathbf{x} \to \hat{\mathbf{x}}) = \frac{1}{\pi}\int_0^{\pi/2}\prod_{i=1}^{d}\left(\frac{1}{1+\frac{\lambda_i}{\sin^2\theta}\frac{\Delta^2}{N_0}}\right)^L d\theta. \tag{45}$$

The maximum diversity degree is $d\cdot L$, which corresponds to the optimum extreme case for independent fading. In this ideal case, all eigenvalues λ_i will be one. Typically, the fading coefficients may be highly correlated, so that some or many of the eigenvalues λ_i are zero or close to zero, which lowers the effective diversity degree. If there is only one nonzero eigenvalue, e.g. $\lambda_1 = d$, we obtain the extreme case of a flat channel with antenna diversity degree L.

IV. A Practical Example

A. System Setup

In this example, we concentrate on Walsh-Hadamard (WH) codes which have a wide range of applications in CDMA systems, both for multi-carrier (MC) and for direct sequence (DS) CDMA, see e.g. [10], [11], [12], [13], [14]. Here we consider an MC-CDMA system with K carriers. We now consider a $\mathrm{WH}(N, \log_2 N, N/2)$ code of N different code words, code rate $R_c = \log_2(N)/N$, and constant weight $N/2$, which means that each pair of code words differs in $d = N/2$ symbol positions. For BPSK modulation, we assume that an integer number of code words of length N will be transmitted in one time slot, i.e. K is an integer multiple of N. For QPSK modulation, we assume that an even number of code words of length N will be transmitted in one time slot.

The multiplexing, thus, can be realized in the following way: For QPSK, the first and second code word of length N will be mapped on the real and the imaginary part of $\{s_1(i)\}_{i=1}^N$, resp., and the third and the fourth code word will be mapped on the real and the imaginary part of $\{s_2(i)\}_{i=1}^N$, resp., to arrange a set of N vectors $\{\mathbf{s}(i)\}_{i=1}^N$ that carry $4N$ code words. This set of vectors will be transmitted during two time slots according to the Alamouti scheme for each fixed index i that has to be understood as a sub-carrier frequency index.

For BPSK, the multiplexing will be similar, but the imaginary parts will not be modulated. For $2K = mN$, m such sets of vectors $\{\mathbf{s}(i)\}_{i=1}^{N}$ will be multiplexed and frequency interleaved on K sub-carriers during two time slots. Each code word is assumed to be multiplexed and modulated in such a way that the encoded bits are equally spaced in frequency and are spread over a bandwidth B.

We assume that the channel is frequency selective, but so slowly time variant that the fading coefficients can be assumed to be constant over two time slots. Inter-symbol interference is assumed to be absorbed, e.g. by a guard interval. We may now work with a frequency-discrete channel model as introduced above. We note that, depending on the length of the guard interval, an additional loss in E_b/N_0 has to be taken into account.

The pairwise error probability is given by

$$P(\mathbf{x} \to \hat{\mathbf{x}}) = \frac{1}{\pi} \int_0^{\pi/2} \prod_{i=1}^{d} \left(\frac{1}{1 + \frac{\log_2 N}{LN} \frac{\lambda_i}{\sin^2 \theta} \frac{E_b}{N_0}} \right)^L d\theta. \tag{46}$$

This is different for different pairs $\mathbf{x}$ and $\hat{\mathbf{x}}$ and depends on the d positions where the two code words differ. The autocorrelation matrix and therefore the eigenvalues λ_i depend on these positions. The correlations of the channel sampled at these positions can be very different. Assume, for example, that for BPSK modulation the all-zero code word has been sent, which corresponds to the vector $\mathbf{x_1}$ with the symbol $+\sqrt{E_S}$ in each position. If $\hat{\mathbf{x}}$ is the vector with symbol $+\sqrt{E_S}$ in position $i = 1, 2, .., N/2$ and $-\sqrt{E_S}$ in position $i = N/2 + 1, ..., N$, there is a much higher correlation of the fading amplitudes than for $\hat{\mathbf{x}}$ being the vector with an alternating sign (i.e. with components $\hat{x}_i = (-1)^{i-1}\sqrt{E_S}$). The numerical evaluation of the pairwise error probabilities shows that the curves for different pairs can differ by several decibels. We found that this effect - that is unknown for uncorrelated fading - can be mitigated significantly if we introduce an *intra code word interleaving* which is just a pseudo random permutation of the coded symbols.

For practical considerations, the bit error probability P_b is much more interesting than the pairwise error probability. The union bound for detecting a wrong code word (block error probability) is given by

$$P_{Block} \leq \sum_{n=2}^{N} P_{N/2}(\mathbf{x_1} \to \mathbf{x_n}), \tag{47}$$

where we have assumed that the code word $\mathbf{x_1}$ has been sent. For the bit error probability it follows the union bound

$$P_b \leq \frac{N/2}{N-1} \sum_{n=2}^{N} P_{N/2}(\mathbf{x_1} \to \mathbf{x_n}). \tag{48}$$

The decay depends on the eigenvalues $\{\lambda_i\}_{i=1}^{d}$. To calculate them, the correlation properties of the channel have to be specified. For our numerical investigations, we will assume an exponential delay power spectrum

$$S_D(\tau) = \frac{1}{\tau_m} e^{-\tau/\tau_m} \epsilon(\tau), \tag{49}$$

where τ_m is the mean delay and $\epsilon(\tau)$ is the Heaviside function. The corresponding frequency autocorrelation function is given by

$$K(f) = \frac{1}{1 + \mathrm{j}2\pi f \tau_m}. \tag{50}$$

The correlation bandwidth of the channel is given by $B_{corr} = \tau_m^{-1}$.

B. Bit Error Rates

We choose the WH code of length 16 and Hamming distance 8. Figure 5 shows the BERs (union bounds) for this code for the AWGN and for the ideally interleaved Rayleigh fading channel with independent fading for each code symbol.

Figure 6 shows the upper bounds for the BER for $L_r L_t = 1, 2, 4, 8$ and a channel with $B\tau_m = 0.5$. This means that the signal bandwidth is only half the coherence bandwidth of the channel with the frequency variance over the bandwidth beeing quite small. As an example, the GSM delay power profile *typical urban* is exponential with $\tau_m = 1\,\mu s$, so $B\tau_m = 0.5$ corresponds to a signal bandwidth $B = 500\,kHz$. Due to the poor frequency selectivity, there is a heavy performance degradation of 8 decibels at 10^{-5} for one receive and one transmit antenna compared to the ideally interleaved Rayleigh fading channel. Antenna diversity at the transmitter and/or at the receiver is able to compensate this degradation. Only one antenna more ($L_r L_t = 2$) allows us to retrieve about 5 decibels of this loss. With $L_r L_t = 4$, e.g. two antennas on both sides, the performance of the independently fading Rayleigh channel will be reached. With $L_r L_t = 8$, e.g. two transmit and four receive antennas, the performance will be even better.

Figure 7 shows the upper bounds for the BER for $L_r L_t = 1, 2, 4, 8$ and a channel with $B\tau_m = 2$. This means that the transmission bandwidth is twice the coherence bandwidth of the channel. The performance degradation compared to independent Rayleigh fading is only 2 decibels at 10^{-5} for one receive and one transmit antenna. For two transmit and two receive antennas, the performance is even slightly better compared to independent fading. For two transmit and four receive antennas, there is an overcompensation of 3 decibels.

Figure 8 shows the upper bounds for the BER for $L_r L_t = 1, 2, 4, 8$ and a channel with $B\tau_m = 8$. This means that the signal bandwidth is 8 times the coherence bandwidth of the channel. The performance degradation compared to independent Rayleigh fading can be neglected. It is interesting to note that with $L_r L_t = 8$, the AWGN limit is nearly reached.

The question now is how much bandwidth is needed without diversity to get close to the independently fading Rayleigh channel. We have seen that for the WH(16,4,8) code, $B\tau_m = 2$ is already good, and for $B\tau_m = 8$ the ideal case is practically reached. The Hamming distance of this code is given by $d = 4$. One must keep in mind that a stronger code with a higher distance will need more bandwidth to exploit its higher diversity degree. For the WH(64,6,32) code one might assume that $B\tau_m = 8$ is already good, and for $B\tau_m = 32$ the ideal case is practically reached. This can be confirmed by numerical evaluation.

V. Conclusions

We have shown that the Alamouti transmit diversity scheme with two transmit antennas is equivalent to a receive antenna diversity scheme with twice the number of receive antennas. The optimum receiver turns out to be a generalization of the conventional MRC receiver for a higher dimensional Euclidean space. This TX diversity scheme can easily be combined with standard outer coding methods to achieve a high coding gain. Antenna diversity can improve the performance of the code significantly, especially if interleaving is not sufficient.

References

[1] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications" , *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451-1458, Oct. 1998.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction" , *IEEE Trans. Inform. Theory*, vol. IT-44, pp. 744-765, Mar. 1998.

[3] G.J. Foschini, M.J. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas", *Wireless Personal Communications,* vol. 6, pp. 311-335, 1998.

[4] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-Time Block Codes for Wireless Communications: Performance Results" , *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451-460, Mar. 1999.

[5] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-Time Block Codes from Orthogonal Designs" , *IEEE Trans. Inform. Theory*, vol. IT-45, pp. 1456-1467, Jul. 1999.

[6] B. L. Hughes , "Differential Space-Time Modulation" , *IEEE Trans. Inform. Theory*, vol. IT-46, pp. 2567-2578, Nov. 2000.

[7] G. Ganesan, P. Stoica, "Space-Time Block Codes: A Maximum SNR Approach", *IEEE Trans. on Inf. Theory,* vol. IT-47, pp. 1650-1656, May 2001.

[8] G. Ganesan, P. Stoica, "Utilizing Space-Time Block Diverity for Wireless Communication", *Wireless Personal Communications,* vol. 18, pp. 149-163, 2001.

[9] Z. Liu, B. Giannakis, "Space-TIme Block-Coded Multiple Access Through Frequency-Selective Fading Channels", *IEEE Trans. on Comm.* vol COM-49, pp. 1033-1044, June 2001.

[10] K. Fazel, S. Kaiser, M. Schnell, "A Flexible and High Performance Cellular Mobile Communications System Based on Orthogonal Multi-Carrier SSMA", *Wireless Personal Communications* vol. 2, pp. 121-144, 1995.

[11] G. Fettweis, K. Anvari, A. Shaikh Bahai, "On Multi-Carrier Code Division Multiple Access (MC-CDMA) Modem Design, *Proc. IEEE Vehicular Technology Conference*, Stockholm, pp. 1670-1674, 1994.

[12] A. Dekorsy, S. Fischer, K.-D. Kammeyer, "Maximum Likelihood Decoding of M-ary Modulated Signals for Multi-Carrier Spread-Spectrum Systems", *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications* (PIMRC), Boston, September 1998.

[13] A. Dekorsy, K.-D. Kammeyer, "A new OFDM-CDMA Uplink Concept with M-ary Orthogonal Modulation", *European Trans. on Telecommunications*, vol. 10, pp. 377-390, 1999.

[14] V. Kuehn, A. Dekorsy, K.-D. Kammeyer, "Channel Coding Aspects in an OFDM-CDMA System", *Proc. 3rd ITG Conf. on Source and Channel Coding*, Munich 2000, 31-36.

[15] M.K. Simon, D. Divsalar, "Some New Twists to Problems Involving the Gaussian Probability Integral", *IEEE Trans. on Comm.*, vol. COM-46(2), Feb. 1998.

[16] M.K. Simon, Mohamed-Slim Alouini, *Digital Communications over Fading Channels.* Wiley, New York, 2000.

[17] S. Benedetto, E. Biglieri *Principles of Digital Transmission With Wireless Applications*, KLuwer, New York, 1999.

[18] A. Proakis, *Digital Communcations.* 3rd ed., McGraw-Hill, New York, 1995.

[19] S. Haykin, Adaptive Filter Theory. 3rd ed., Prentice Hall, 1996.

Figure Captions

Figure 1: Block diagram for the Alamouti scheme with two TX and one RX antennas with transmission branch channel coefficients c_1, c_2.

Figure 2: Block diagram for two TX and two RX antennas with transmission branch channel coefficients c_1, c_2, c_3, c_4.

Figure 3: The generalized MRC receiver.

Figure 4: Block diagram for the concatenated STC scheme.

Figure 5: Bit error curves (union bounds) of the WH(16,4,8) code for the AWGN and for the independently fading Rayleigh channel.

Figure 6: Bit error curves (union bounds) of the WH(16,4,8) code for the Rayleigh fading channel with $B\tau_m = 0.5$.

Figure 7: Bit error curves (union bounds) of the WH(16,4,8) code for the Rayleigh fading channel with $B\tau_m = 2$.

Figure 8: Bit error curves (union bounds) of the WH(16,4,8) code for the Rayleigh fading channel with $B\tau_m = 8$.

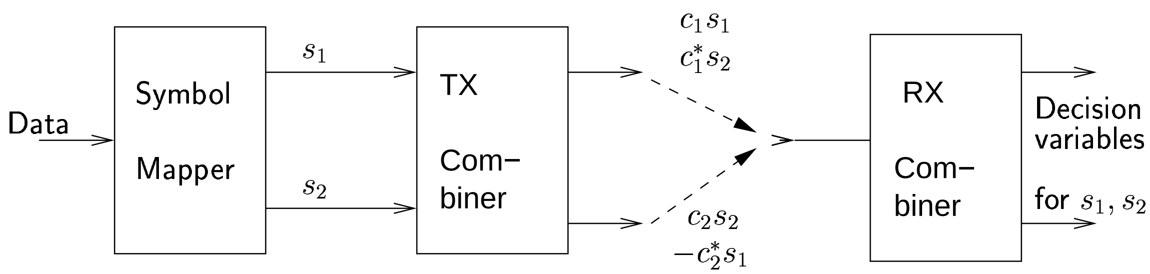

Fig. 1. Block diagram for the Alamouti scheme with two TX and one RX antennas with transmission branch channel coefficients c_1, c_2.

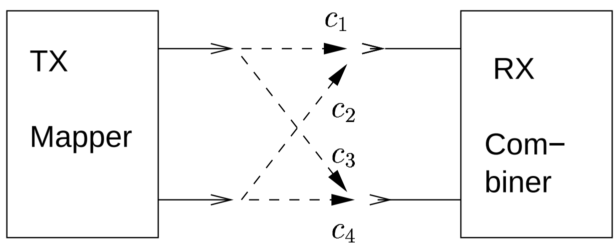

Fig. 2. Block diagram for two TX and two RX antennas with transmission branch channel coefficients c_1, c_2, c_3, c_4.

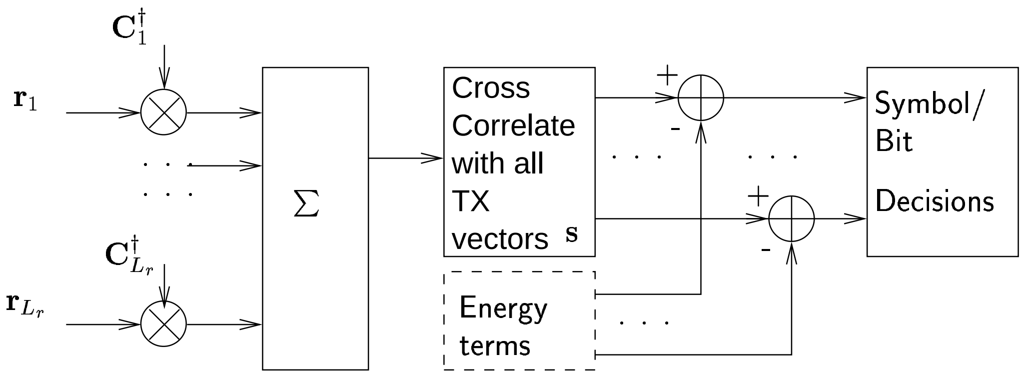

Fig. 3. The generalized MRC receiver.

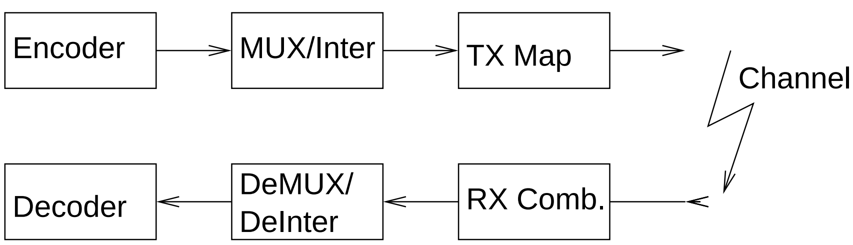

Fig. 4. Block diagram for the concatenated STC scheme.

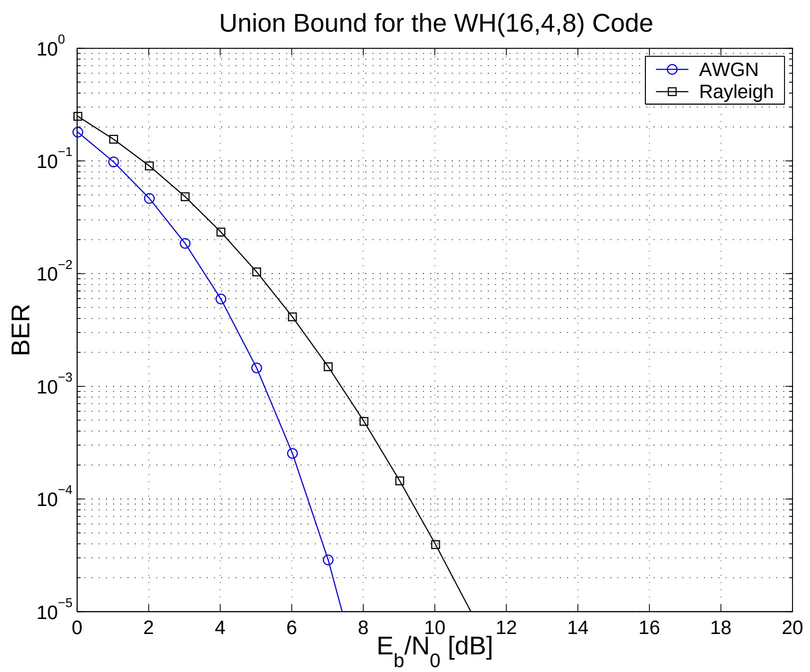

Fig. 5. Bit error curves (union bounds) of the WH(16,4,8) code for the AWGN and for the independently fading Rayleigh channel.

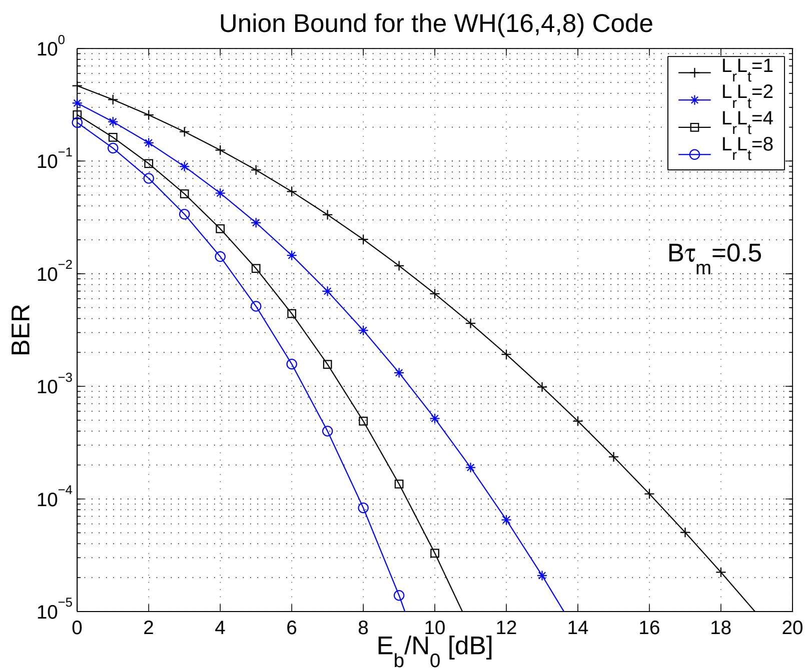

Fig. 6. Bit error curves (union bounds) of the WH(16,4,8) code for the Rayleigh fading channel with $B\tau_m = 0.5$.

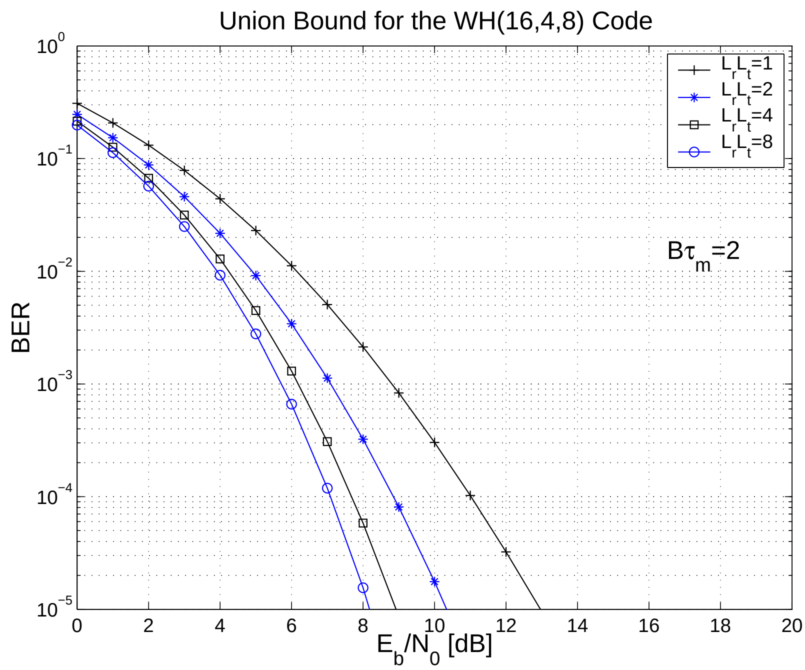

Fig. 7. Bit error curves (union bounds) of the WH(16,4,8) code for the Rayleigh fading channel with $B\tau_m = 2$.

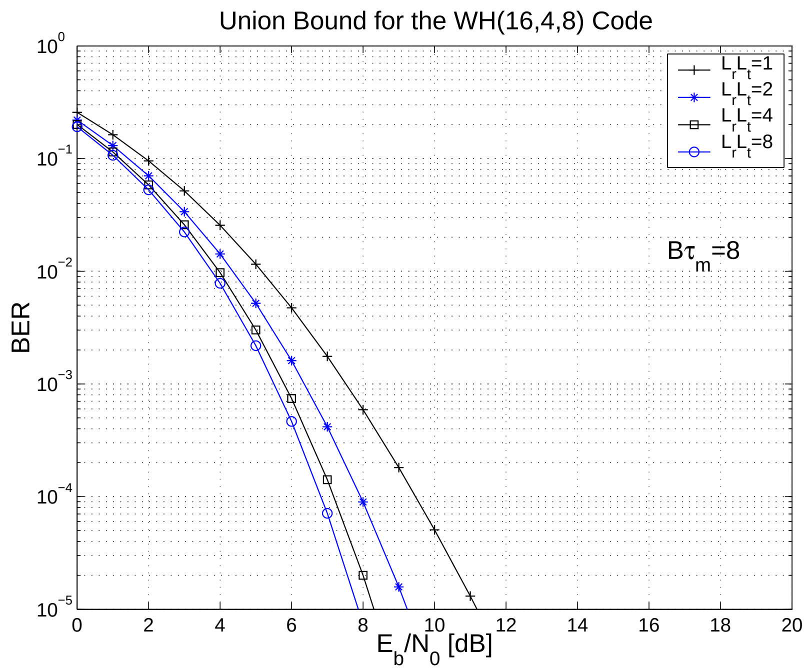

Fig. 8. Bit error curves (union bounds) of the WH(16,4,8) code for the Rayleigh fading channel with $B\tau_m = 8$.