Abstract

Space-time codes employ constellations of complex matrices. There are many advantages if this constellation forms a group under multiplication. Special cases have been considered to date, but in this paper all such constellations are analyzed. For a given constellation, the Chernoff bound for the probability of error is an explicit rational function in the SNR. By comparing these functions, our method gives a means to find, for a given number of transmit antennas, rate, and SNR, the best group constellation.

1 Introduction

Space-time coding is a recently discovered technique for implementing antenna diversity. There are two branches, space-time block codes (STBC’s) and space-time trellis codes (STTC’s). Historically, STTC’s came first, but Alamouti’s scheme [1] led research towards STBC’s, with their simpler decoders. Alamouti’s scheme is supported by third generation systems but has the drawback of being a single 2-antenna system.

The uniqueness of Alamouti’s scheme can be traced back to the fact that the only division algebra properly containing $\mathbb{C}$ is the quaternions $\mathbb{H}$. People have tried to generalize this in various ways, with two major developments, both highly algebraic.

The first, developed at Lucent by Hassibi, Hochwald, Shokrollahi, and Sweldens [10], is to take constellations closed under multiplication. Since the unit quaternions are isomorphic to $SU(2)$, this led to the use of various interesting subgroups of $SU(2)$ and higher-dimensional analogues. We focus on this case and make a systematic study that obtains for a given number of antennas, SNR, and rate, the optimal (in terms of probability of decoding error) space-time group code of this kind. This is analogous to finding among block codes the optimal linear block code of given length and rate. They may not turn out to be the best codes but they give a benchmark to beat.
The second and currently more popular method of producing ST BC’s is to use division algebras containing $F := \mathbb{Q}(\sqrt{-1})$. The point is that it suffices that they contain the modulation alphabet, which is typically QPSK or 16-QAM, implying that containing $F$ is enough. As opposed to the case over $\mathbb{C}$, many of these division algebras exist. The current work (of Belfiore, Sethuraman, and others [8], [9]) involves heavy use of algebraic number theory.

In [10], it was proposed that for space-time coding for an unknown channel, a constellation of unitary matrices that form a group $G$ should be used. This has the advantages of simplifying differential encoding (codewords are products of group elements and so stay in the group, reducing encoding to a simple lookup) and of homogeneity (similar to the advantages of linear block codes). Ensuring the codewords are well separated amounts to showing that $|\text{det}(A - B)|$ for any different $A$ and $B$ in the constellation $G$ is maximized, and in particular is nonzero. This means that the eigenvalues of any nontrivial matrix in $G$ never equal one, and such a group is called fixed-point-free (fpf). [10] classified all fpf constellations and studied their performances. In particular, a practical set-up of an fpf $G$ of order 63 with three antennas was tested in a corridor at Lucent Technologies and performed well.

Just as minimum distance is not the best evaluator of a linear block code, [11] found that by allowing just a few matrices with eigenvalue one, better performances than with fpf constellations are possible. By analyzing the probability of error, at asymptotically high SNR, fpf constellations will always beat non-fpf ones. However, there is no such guarantee for realistic SNR. This motivates us to look for more groups. Specifically for a given number of transmitter antennas $M$, rate $R$ (essentially $|G|$), and SNR $\rho$, we give a method that yields the group $G$ of $M$ by $M$ matrices of order (approximately) $R$ that yields the smallest probability of error at SNR $\rho$.

Mathematically, this requires us to know all finite groups $G$ consisting of $M$ by $M$ complex matrices. This problem has a long history and complete answers are known for small $n$ [3]. For example, for $M = 2$, Klein [6] showed that the projective linear quotient $H$ of $G$ is a cyclic group, a dihedral group, $A_4$, $S_4$, or $A_5$. For each $H$ there is an infinite family of corresponding $G$, yielding the complete list of 2 by 2 finite matrix groups.

On the engineering side, the Chernoff bound for probability of error as a function of SNR $\rho$ for a given constellation $G$ turns out to be a rational function $E(G, t)$, where $t$ is a simple increasing function of $\rho$ whose form depends on whether the channel is known to the receiver or not. This is a finer measure than both diversity product and diversity sum since we recover these from $E(G, t)$ in the asymptotically high and low SNR cases.

We indicate how $E(G, t)$ arises and how, extending [11], it can be computed just from character values, i.e. traces. We then give a thorough analysis in the $M = 2$ case of which are the best group constellations for a given rate and SNR and indicate how this can be carried out for any given $M, R$, and $t$. This then completes the study of space-time group codes begun in [10] and sets the bar for those searching for better non-group constellations.

2 Channel Model and Differential Space-Time Modulation

Consider a communication link with $M$ transmitter antennas and $N$ receiver antennas that operates in a Rayleigh flat-fading environment. Each receiver antenna responds
to each of transmitter antenna by multiplication by a statistically independent complex fading coefficient. Flat fading means the channel coefficients are constant with respect to time. The received signals at each antenna and each time are corrupted by additive independent identically distributed complex Gaussian noise. Assuming the coding is done in $T$ time intervals, we write the channel model in a compact form for each time interval as,

$$X = \sqrt{\frac{\rho T}{M}} SH + W$$  \hspace{1cm} (1)

where $\rho$ is the SNR, $S$ the signal (a $T \times M$ complex matrix), $H$ the channel (an $M \times N$ complex matrix giving the fades), and $W$ the noise (a $T \times N$ matrix, entries independent complex Gaussian random variables, mean 0, variance 1). The factor $\sqrt{T/M}$ normalizes the expected value of the vector of transmitter symbols.

### 2.1 Channel Unknown to Receiver

If the receiver does not have channel knowledge, [4] generalized differential phase-shift keying (DPSK) to multiple antennas by choosing $T = 2M$ and with signals of the form

$$\Phi_\ell = \frac{1}{\sqrt{2}} \begin{pmatrix} I \\ V_\ell \end{pmatrix}$$  \hspace{1cm} (2)

where the $V_\ell$ are $M$ by $M$ unitary matrices. The differential transmission is given by

$$S_\tau = V_\tau S_{\tau-1}, \quad \tau = 1, 2, ...$$  \hspace{1cm} (3)

where the $S_\tau$ are the $M$ by $M$ transmitter symbols in each $M$ time intervals. Similarly, the differential transmission is expressed as

$$X_\tau = V_\tau X_{\tau-1} + \sqrt{2}W'_\tau$$  \hspace{1cm} (4)

where the entries of $W'_\tau$ are $M$ by $N$ independent complex Gaussian additive noise with zero means and unit variances. According to [4], the ML decoder admits the form

$$(\hat{z}_\tau)_{ML} = \arg\min ||X_\tau - V_\ell X_{\tau-1}||$$  \hspace{1cm} (5)

and the Chernoff bound on the pairwise probability is given by

$$P_e \leq (1/2) \prod_{m=1}^{M} \left(1 + \frac{\rho^2}{4(1 + 2\rho)} \sigma_m^2(V_\ell - V_\ell)\right)^{-N}$$  \hspace{1cm} (6)

where $\sigma_m(V_\ell - V_\ell)$ is the $m$th singular value of the $M \times M$ matrix $V_\ell - V_\ell$.

### 2.2 Channel Known to Receiver

Since in differential modulation we consider signal constellations of $M$ by $M$ unitary matrices, here we should use the model

$$X_\tau = \sqrt{\rho T} SH + W_\tau$$  \hspace{1cm} (7)

with the $S_\tau$ $M$ by $M$ unitary matrices and the $W_\tau$ $M$ by $N$ complex white Gaussian noise.
The maximum-likelihood receiver computes
\[
(z_\tau)_{ML} = \arg \min_{z} ||X_\tau - V_\ell H_\tau||
\] (8)
and the Chernoff bound is given by
\[
P_e \leq (1/2) \prod_{m=1}^{M} \left( 1 + \frac{\rho}{4} \sigma_m^2 (V_\ell - V_{\ell'}) \right)^{-N} \] (9)

### 3 Performance Criterion

We consider three criteria for comparing different group constellations, namely diversity product, diversity sum and probability of error. The first is a good criterion for high SNR, the second for low SNR. The last is the most complete measure of how a constellation performs. Following [11], an upper bound for the probability of error is given by.

\[
P_e \leq \frac{1}{2} \sum_{C \neq I} |C| \prod_{m=1}^{M} \left( 1 + 2f(\rho)(1 - \Re(\lambda_m(C))) \right)^{-N} \] (10)

where \( C \) runs through the non-identity conjugacy classes of the group, \( \lambda_m(C) \) is the \( m \)th eigenvalue of \( C \), and \( f(\rho) = \rho/4 \) or \( \rho^2/(4(1+2\rho)) \) in the known and unknown channel cases respectively.

This bound is the most valuable for comparing codes, but as in [11] its computation is tedious. The computer algebra system Magma [2] provides a new tool to compute this bound quickly and so compare performances of various group constellations.

From (6) and (9) if the SNR is high, the upper bound pairwise probability of error is dominated by the term.

\[
\prod_{m=1}^{M} \sigma_m(V_\ell - V_{\ell'}) = |\det(V_\ell - V_{\ell'})| \] (11)

In order to be able to compare constellations with different \( M \) the diversity product is defined as

\[
\zeta_V = \frac{1}{2} \min |\det(V_\ell - V_{\ell'})| \] (12)

over all pairs \( \ell \neq \ell' \).

A constellation with higher diversity product is regarded as better. If it is nonzero, we say that the constellation has full diversity.

On the other hand, when SNR approaches zero, (6) and (9) can be written as

\[
P_e \leq \frac{1}{2} \prod_{m=1}^{M} \left( 1 + f(\rho)\sigma_m^2 (V_\ell - V_{\ell'}) \right)^{-N} \] (13)

where \( f(\rho) = \rho/4 \) when the channel is known and \( f(\rho) = \frac{\rho^2}{4(1+2\rho)} \) when the channel is unknown. In either case \( f(\rho) \) is a nondecreasing function for positive \( \rho \) and its value approaches zero as \( \rho \) approaches zero. We may also take \( N = 1 \) since the term \( 1 + f(\rho)\sigma_m^2 (V_\ell - V_{\ell'}) \) is always greater than one. Computing the Taylor series for the pairwise error probability as a function of \( f(\rho) \) around zero, we have
\[
P_\epsilon \leq \left( \frac{1}{2} + \frac{1}{2} \left( \sum_{m=1}^{M} (f(\rho) \sigma_m^2 (V_\ell - V_{\ell'}) + O(f(\rho)^2))^{-1} \right) \right)
\]  

(14)

As with the diversity product, to compare constellations with different \( M \) the diversity sum is defined as

\[
\frac{1}{2} \sqrt{\frac{\sum_{m=1}^{M} \sigma_m (V_\ell - V_{\ell'})}{M}}
\]  

(15)

The criterion we will use in this paper is what we will call \( E(G,t) \) which is really eq. (10) with \( t = f(\rho) \), where \( f(\rho) = \rho/4 \) or \( \rho^2/(4(1+2\rho)) \) in the known and unknown channel cases respectively. Note that in both cases \( f \) are monotonically increasing with \( \rho \); this means in order to compare different schemes (of course with the same side information available), it suffices to compare their \( E(G,t) \). As mentioned, the computation of \( E(G,t) \) is tedious and hence is carried out by MAGMA. There are some technical issues here, since in the formula for \( E(G,t) \) we are seeking eigenvalues over the complex field, which is not available in MAGMA (also in any sort of algebraic computer system). Hence we have to resort to finite field; however, we need a finite field which will give complete information for the character. From character theory, if \( G \) has exponent \( n \), a splitting field for \( G \) (the one containing all values of character of \( G \) over \( \mathbb{C} \)) needs to contain all the \( n \)th root of unity. Let \( p \equiv 1 \pmod{n} \), where \( p \) is prime, then \( GF(p) \) is a splitting field for \( G \). In MAGMA we will use the irreducible modules over such a finite field. Then to transfer the answer back to the complex numbers, we need to solve the discrete log problem. Interested readers should consult MAGMA manual.

An almost immediate consequence of adopting \( E(G,t) \) as criterion is that, we only have to search among irreducible groups, since if \( G \) is a direct sum of \( H \) and \( K \) which are groups of \( L \) and \( M \) square matrices respectively,

\[
E(G,t) = \sum_{C \neq \emptyset} |C| \prod_{m=1}^{L+M} (1+2f(\rho)(1-Re(\lambda_m(C))))^{-N}
\]

\[
= \sum_{C_H \neq \emptyset} \sum_{C_K \neq \emptyset} |C_H||C_K| \prod_{m=1}^{L} (1+2f(\rho)(1-Re(\lambda_m(C_H))))^{-N} \prod_{m=1}^{M} (1+2f(\rho)(1-Re(\lambda_m(C_K))))^{-N}
\]

\[
+ \sum_{C_H \neq \emptyset} |C_H| \prod_{m=1}^{L} (1+2f(\rho)(1-Re(\lambda_m(C_H))))^{-N}
\]

\[
+ \sum_{C_K \neq \emptyset} |C_K| \prod_{m=1}^{M} (1+2f(\rho)(1-Re(\lambda_m(C_K))))^{-N}
\]

\[
= E(H,t)E(K,t)+E(H,t)+E(K,t)
\]  

(16)

i.e., \( E(G,t) \) can be derived from those of \( E(H,t) \) and \( E(K,t) \).

4 Our Strategy

Since fixed-point-free groups do not occur in sufficient abundance, we take the systematic approach of finding ALL finite groups of complex matrices. Let \( GL(M,\mathbb{C}) \) denote the group of all \( M \) by \( M \) invertible complex matrices.

The main idea is that for any given \( M \), the finite subgroups \( G \) of \( GL(M,\mathbb{C}) \) are in principle known. For a given \( M, \rho \), and rate \( R \), we can then compute the probability of decoding error for each \( G \) of about the size corresponding to \( R \) and see which group gives the smallest bound on probability of error. Simulations confirm that these bounds are good. Thus we can obtain the optimal space-time group code for any given rate, SNR, and number of transmit antennas.
The scalar matrices inside $GL(M, \mathbb{C})$ form a normal subgroup and the quotient group is denoted $PGL(M, \mathbb{C})$. If $G$ is a subgroup of $GL(M, \mathbb{C})$, then its image in the quotient group is denoted $PG$. The problem of classifying all finite subgroups of $GL(M, \mathbb{C})$ breaks down into two steps.

1. Given $M$, find all finite subgroups $H$ of $PGL(M, \mathbb{C})$.
2. Given a finite subgroup $H$ of $PGL(M, \mathbb{C})$, find all finite subgroups $G$ of $GL(M, \mathbb{C})$ such that $PG = H$.

The case $M = 1$ is completely understood, since $GL(1, \mathbb{C})$ consists of the nonzero complex numbers, whose finite subgroups are the cyclic groups of $m$th roots of 1, one for each $m$. We illustrate the general method by explaining in detail the case $M = 2$ below.

5  Finite Subgroups of $GL(2, \mathbb{C})$

As for step one, by Klein [6], the finite subgroups of $PGL(2, \mathbb{C})$ are:

(a) cyclic groups;
(b) dihedral groups;
(c) $A_4$ of order 12, $S_4$ of order 24, $A_5$ of order 60.

Groups in case (a) are called reducible, since the corresponding subgroups of $GL(2, \mathbb{C})$ arise by direct sums of the 1-dimensional case. As noted above, we normally do not need separately to compute reducible cases.

Groups in case (b) are called imprimitive, since the corresponding subgroups of $GL(2, \mathbb{C})$ arise by induction from the 1-dimensional case. The essentially new cases are the primitive ones (case (c)).

As for step two, for any given subgroup $H$ of $PGL(2, \mathbb{C})$, there are infinitely many $G$ with $PG = H$, but they arise in an orderly fashion. For example, if $PG = A_5$, then $G$ belongs to a family of groups $G_k(k = 1, 2, ...)$, where $G_1 = SL(2, 5)$ and $G_k$ has order $120k$. In general the following can be shown.

**Theorem 1.** If $PG = A_5$, then there is one group $G$ of order $120k$ for each $k = 1, 2, ...$;
if $PG = S_4$, then there are two groups $G$ of order $48k$ for each $k = 1, 2, ...$;
if $PG = A_4$, then there are two groups $G$ of order $24k$ for each $k$ divisible by 3, otherwise only one;
if $PG = D_2$, then there are two groups $G$ of order $8k$ for each $k$;
if $PG = D_n$ ($n$ odd), then there are two groups $G$ of order $2nk$ for each even $k$, otherwise only one;
if $PG = D_n$ ($n$ even), then there are three groups $G$ of order $4nk$ for each $k$.

6  Example for $M = 2$

If $G$ is a finite group of $M \times M$ complex matrices, let $E(G, t)$ denote Shokrollahi’s formula for the upper bound for probability of decoding error for $t = f(\rho)$, where $f(\rho) = \rho/4$ or $\rho^2/(1 + 2\rho)$ in the known and unknown channel cases respectively.

**Example**: Suppose the desired rate means we want $|G|$ approximately 120. From the above analysis there is one group of order 120 in the $A_5$ family, none in the $S_4$ family, one in the $A_4$ family but there are two characters that give rise to different $E(G, t)$, two in the $D_2$ family, two in the $D_3$ family, two in the $D_5$ family, three in the $D_6$ family, and
three in the $D_{10}$ family but only two distinct $E(G, t)$. So there are just fourteen functions to compare.

These functions can then be plotted against $t$, as shown in figure 1 (the number in brackets after groups are the diversity products). We find that even for low SNR the best group turns out to be the group in the $A_5$ family, namely the well-known $SL(2, 5)$ example.

In general, asymptotically $E(G, t) \sim C/t^d$ where $d$ is largest (equal to $M$) for fixed point free groups. Thus for asymptotically large SNR, the fpf groups are best, but it can be that for realistic SNR other groups are better (as in Shokrollahi’s $SL(2, 17)$ example from [11], although a complete comparison with other groups of similar order is not included in that paper).

7 Finite Subgroups in Higher Dimensions

There has been a long history of classifying finite subgroups of $PGL(M, \mathbb{C})$. For a given $M$, there are only finitely many primitive subgroups (case (c) before) [5]. They are all computed at least for $M \leq 8$. In practice, it is doubtful that more than 8 transmit antennas will be employed.

We conducted some experiments below, comparing $E(G, t)$ for various groups $G$ in higher dimensions. For each degree $M$, our first aim is to see whether the fpf groups are really good and if so for what SNR range. In other words we compare their probability of error curves $E(G, t)$ as $t$ varies.

We started by look at Feit’s list [3], picking ones that are outstanding and comparing them with the fpf groups with comparable rate (i.e. comparable group order).

Specifically, looking at the probability of error from Feit’s list of different groups of odd degrees (3, 5, 7), we obtain the plots of the error bound as SNR varies as shown if figure 2, 3, 4 for groups of degree 3, 4, 5 respectively.

In order to see clearly, we start by an example in degree 3. In this case, from [10] the only family of fpf groups of degree 3 is that of groups denoted $G(m, r)$. Specifically we look at $A_5$ of order 60 and $PSL(2, 7)$ of order 168 which turn out to perform well. The fpf groups of order close to 60 of the $G(m, r)$ family are $G(21, 4)$ and $G(21, 16)$, and the fpf groups of order close to 168 are $G(57, 7)$ and $G(57, 49)$. The comparison between their error curves is shown figure 5.

As you can see, in this case the fpf groups perform very well. We can make similar comparisons for other degrees. For degree 5, there are two groups which perform well. The low-rate one is $S_5$ of order 120 and the high-rate one is $PSL(2, 11)$ of order 660. It turns out that there is no fpf group of degree 5 and order close to 120, so $S_5$ wins unchallenged. However, there are some with order close to 660, e.g., $G(110, 31), G(110, 71), G(110, 81)$, and $G(110, 91)$. The comparison of error curves is shown in figure 6.

There are patterns in the form of the rational function $E(G_k, t)$ as $k$ varies such that the groups $G_k$ have the same $PG_k$. It may be feasible but hard to obtain explicit formulae for $E(G_k, t)$ as functions of $k$ and $t$, which would clarify why some groups perform better than others. So far we have computed all $E(G, t)$ of numerator and denominator degree less than 300.
References


Figure 1: Pe vs f(SNR) for group codes for 2 ant. and size 120

Figure 2: Pe vs f(SNR) for group codes in Feit’s list of degree 3

Figure 3: Pe vs f(SNR) for group codes in Feit’s list of degree 5
Figure 4: \( Pe \) vs \( f(SNR) \) for group codes in Feit’s list of degree 7

Figure 5: \( Pe \) vs \( f(SNR) \) for group codes in Feit’s list of degree 3 compared to FPF groups of similar rate

Figure 6: \( Pe \) vs \( f(SNR) \) for group codes in Feit’s list of degree 5 compared to FPF groups of similar rate