Algebraic Cayley Differential Space-Time Codes

Frédérique Oggier and Babak Hassibi
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125
{frederique,hassibi}@systems.caltech.edu

Abstract

Cayley differential codes have been proposed as a solution for coding over the noncoherent MIMO channel. Using the Cayley transform that maps the space of Hermitian matrices to the manifold of unitary matrices, they are particularly suitable for high data rate, since they have an easy encoding and can be decoded using a Sphere-Decoder algorithm. However, at high rate, the problem of evaluating if a Cayley code is fully diverse may become intractable, and previous work has focused instead on maximizing a mutual information criterion. The drawback of this approach is that it requires a lot of computations for a given chosen number of transmit antennas as well as given rate. In this work, we study Cayley codes in the context of division algebras, an algebraic tool that allows to get fully diverse codes. We present an algebraic construction of Cayley codes that are fully diverse, and show that this approach yields naturally, without further optimization, codes that perform similarly or closely to previous unitary differential codes, including previous Cayley codes, and codes built from Lie groups.

I. INTRODUCTION

It is now well known that multiple antennas at the transmitter and receiver ends of a wireless system help increasing the data rate. Both the scenarios where the receiver knows the channel (*coherent case*) or alternatively does not know the channel (*non coherent case*) have been largely studied. In particular, design of space-time codes to resist the fading of wireless systems have been paid a lot of attention. For the noncoherent MIMO channel, a popular strategy is to use unitary differential modulation [3], [4], which requires the design of unitary matrices that are fully diverse, that is, that satisfy that the difference of two matrices is full rank.

There has been a lot of work on the differential coding problem. In [7], a systematic parametrization for the two antennas case has been done. Among the algebraic approaches, in [15], representation of fixed

This work was supported in part by the Swiss National Science Foundation grant PBEL2-110209 and by NSF grant CCR-0133818, by Caltech's Lee Center for Advanced Networking and by a grant from the David and Lucille Packard Foundation.

point free groups has been considered, while in [6], [5], representation of Lie groups has been investigated. The drawback of all these approaches is that it is difficult to use them to get high rate codes. In [2], Cayley codes have been introduced to construct differential codes for high data rate. The idea is to use the Cayley transform that maps the space of Hermitian matrices to the manifold of unitary matrices. Cayley codes have thus an easy encoding, that makes them available at high rate, and can be decoded using a Sphere Decoder algorithm [17], [1]. However, this makes the full diversity criterion difficult to check, and instead, the authors in [2] suggest another design criterion, based on mutual information. The drawback of Cayley codes is that optimizing the mutual information criterion is based on heavy numerical optimization, and depends on the number of antennas and rate one would like to use. Thus each new requirement of antennas or rate asks for a lot of computations to optimize a new code.

Parallely, division algebras have been introduced for coding over coherent MIMO channels [14], [10]. These algebraic objects became of interest, since they naturally provide families of matrices that are fully diverse. The idea of this work is to combine the two approaches: to use division algebras to construct Cayley codes, in order to get fully diverse Cayley codes. We are interested in seeing if fully diverse Cayley codes can compare to mutual information optimized Cayley codes, and if yes, if the method is more general and asks for less computations. We will show that the new approach yields Cayley codes that naturally behave, for different number of transmit antennas, similarly or closely to previous differential codes: previous Cayley codes for 2, 3 or 4 transmit antennas, Lie groups based codes for 3 and 4 transmit antennas. This fulfills the challenge of getting a family of codes that simultaneously does as good or closely as good in several dimensions than previous codes optimized for a given dimension.

The rest of the paper is organized as follows: we recall first unitary differential modulation and the construction of Cayley codes in Section II. We then introduce division algebras and show how to use them for our purpose in Section III. Section IV contains the code constructions together with their simulations.

II. Cayley Differential Codes

We start by briefly recalling the idea behind unitary space-time modulation, before introducing Cayley codes.

A. *Differential unitary space-time modulation*

Consider a Rayleigh flat fading channel with M transmit antennas and N receive antennas, with *unknown channel information*. The channel is used in blocks of M channel uses, so that the transmitted signal can be represented as an $M \times M$ matrix $\mathbf{S}_t$, where $t = 0, 1\ldots$ represents the block channel use. If we assume that the channel is constant over M channel uses, we may write it as

$$\mathbf{Y}_t = \sqrt{\rho}\mathbf{S}_t\mathbf{H}_t + \mathbf{W}_t,\ t = 0, 1, \ldots \tag{1}$$

Here $\mathbf{H}_t$, the channel matrix, and $\mathbf{W}_t$, the noise matrix, are two $M \times N$ matrices with independent complex normal coefficients, and ρ is the expected SNR at each receive antenna.

We use *differential unitary space-time modulation* [3], [4]. The transmitted signal $\mathbf{S}_t$ is encoded using differential modulation, that is (assuming $\mathbf{S}_0 = \mathbf{I}$)

$$\mathbf{S}_t = \mathbf{X}_{z_t}\mathbf{S}_{t-1},\ t = 1, 2, \ldots, \tag{2}$$

where $z_t \in \{0, \ldots, L-1\}$ is the data to be transmitted, and $\mathcal{C} = \{\mathbf{X}_0, \ldots, \mathbf{X}_{L-1}\}$ the constellation to be designed. By definition of the scheme itself, the matrices $\mathbf{X}$ have to be unitary (so that the product $\mathbf{S}$ does not go to zero or infinity).

Note that since the channel is used M times, the transmission rate is

$$R = \frac{1}{M}\log_2 L. \tag{3}$$

The size $|\mathcal{C}|$ of the constellation is thus $L = 2^{MR}$.

If we further assume the channel constant for $2M$ consecutive uses, we get from Equations (1) and (2) that

$$\begin{aligned}
\mathbf{Y}_t &= \sqrt{\rho}\mathbf{X}_{z_t}\mathbf{S}_{t-1}\mathbf{H} + \mathbf{W}_t \\
&= \mathbf{X}_{z_t}(\mathbf{Y}_{t-1} - \mathbf{W}_{t-1}) + \mathbf{W}_t \\
&= \mathbf{X}_{z_t}\mathbf{Y}_{t-1} + \mathbf{W}'_t.
\end{aligned}$$

Since the matrix $\mathbf{H}$ does not appear in the last equation, this means differential modulation allows decoding without knowledge of the channel.

The maximum likelihood decoder is thus given by

$$\hat{z}_t = \arg\min_{l=0,\ldots,|\mathcal{C}|-1} \|\mathbf{Y}_t - \mathbf{X}_l\mathbf{Y}_{t-1}\|.$$

At high SNR, the pairwise block probability of error P_e can be upper bounded by [3], [4]

$$P_e \leq \left(\frac{1}{2}\right)\left(\frac{8}{\rho}\right)^{MN} \frac{1}{|\det(\mathbf{X}_i - \mathbf{X}_j)|^{2N}}.$$

Clearly, the bigger $\det(\mathbf{X}_i - \mathbf{X}_j)$, $i \neq j$, is, the better the code will perform. Thus the *diversity product*, given by,

$$\zeta_{\mathcal{C}} = \frac{1}{2} \min_{\mathbf{X}_i \neq \mathbf{X}_j} |\det(\mathbf{X}_i - \mathbf{X}_j)|^{1/M} \qquad \forall\ \mathbf{X}_i \neq \mathbf{X}_j \in \mathcal{C} \tag{4}$$

has been defined as a measure of the quality of the code. We say that *full diversity* is achieved when

$$\det(\mathbf{X}_i - \mathbf{X}_j) \neq 0 \qquad \forall\ \mathbf{X}_i \neq \mathbf{X}_j \in \mathcal{C}. \tag{5}$$

A design criterion can be summarized as follow: find a code constellation $\mathcal{C}$ of $L = 2^{MR}$ unitary matrices such that $\zeta_{\mathcal{C}}$ is maximized. Much of the work on differential coding has been on getting fully diverse codes, with the best possible diversity [7], [15], [5], [6]. However, the Cayley codes approach is different, as recalled in next section.

B. Cayley Differential Codes

Cayley codes for transmission over the noncoherent channel have been introduced in [2]. They are based on the *Cayley transform*, which maps the nonlinear Stiefel manifold to the linear space of skew-Hermitian matrices (and vice versa). Let $\mathbf{A}$ be an Hermitian matrix, and thus $i\mathbf{A}$ be skew-Hermitian. Its Cayley transform is given by

$$\mathbf{V} = (\mathbf{I} + i\mathbf{A})^{-1}(\mathbf{I} - i\mathbf{A}),$$

where the Hermitian matrix $\mathbf{A}$ is given by

$$\mathbf{A} = \sum_{q=1}^{Q} \alpha_q \mathbf{A}_q,$$

where $\alpha_1, \ldots, \alpha_Q$ are real scalars, chosen from a set $\mathcal{S}$ with r possible values, and where $\mathbf{A}_q$ are fixed $M \times M$ complex Hermitian matrices. Note that the inverse of $\mathbf{I} + i\mathbf{A}$ always exists since $i\mathbf{A}$ has strictly imaginary eigenvalues. It is easy to check that $\mathbf{V}$ is unitary.

The performance of a Cayley code depends on Q, the Hermitian basis matrices $\{\mathbf{A}_q\}$, and the set $\mathcal{S}$ from which each α_q is chosen. To choose the $\{\mathbf{A}_q\}$, the approach of [2] consists in optimizing a mutual information criterion, instead of the traditional diversity criterion (4). It is argued that at high rate, checking the diversity may become intractable. The $\mathbf{A}_1, \ldots, \mathbf{A}_Q$ are thus chosen such that the distribution on $\mathbf{V}$ is close to the distribution that maximizes the mutual information between it and the pair $(\mathbf{Y}_{\tau-1}, \mathbf{Y}_\tau)$. It has been showed that the optimal distribution on $\mathbf{V}$ is to choose $\mathbf{V}$ *isotropically distributed*. (An isotropically distributed unitary matrix $\mathbf{V}$ has a probability density function which is invariant to pre- and post-multiplication by an arbitrary unitary matrix.) Consequently, it is shown that the optimal distribution on $\mathbf{A}$ is

$$p(\mathbf{A}) = \frac{2^{M^2-M}(M-1)!\cdots 1!}{\pi^{M(M+1)/2}}\frac{1}{\det(I+\mathbf{A}^2)^M}.$$

The above probability density function is a generalization of the scalar Cauchy distribution, which is why $\mathbf{A}$ is said to be *Cauchy distributed*. Finally, using Cauchy matrices is shown to imply that a good choice for the scalars $\alpha_1,\ldots,\alpha_Q$ is to take them scalar Cauchy distributed. The optimizations are done numerically using a gradient method.

The drawback of this method is that the mutual information optimization depends on both the number of antennas and the rate, so that heavy computations have to be repeated each time one of these two parameters is changing.

In the following, we are interested in considering similar Cayley codes, except that we will replace the mutual information criterion by the original diversity criterion (4). In order to do so, we need the algebraic concept of *division algebras*. In the next section, we introduce division algebras and explain how to use them for getting Cayley codes.

III. Cyclic algebras

In [2], it has been shown that a set of unitary matrices obtained via the Cayley transform is fully diverse (5) if and only if the set of the skew-Hermitian matrices is. In [14], [10], families of matrices achieving full diversity for the coherent MIMO channel have been found using an algebraic object called *division algebras*, and more particularly *cyclic division algebras*. In this section, we recall what are these objects, and how they allow to get families of matrices that are fully diverse. Our aim is then to build a subset of Hermitian matrices inside the algebra, in order to get a fully diverse family of Hermitian matrices for building Cayley codes. For basic algebraic definitions and facts, we let the reader refer for example to [16], [8]. Also, a tutorial providing the background necessary to understand cyclic algebra based codes is available in [10].

A. Basic definitions

Let L/K be a Galois extension of degree n such that its Galois group $G=\mathrm{Gal}(L/K)$ is cyclic, with generator σ. Denote by K^* (resp. L^*) the non-zero elements of K (resp. L), and choose an element $\gamma\in K^*$. We construct a non-commutative algebra, denoted by $\mathcal{A}=(L/K,\sigma,\gamma)$, as follows:

$$\mathcal{A}=L\oplus eL\oplus\ldots\oplus e^{n-1}L$$

such that e satisfies

$$e^n = \gamma \quad \text{and} \quad \lambda e = e\sigma(\lambda) \text{ for } \lambda \in L.$$

Such an algebra is called a *cyclic algebra*. It is a right vector space over L, and as such has dimension $(\mathcal{A} : L) = n$.

Cyclic algebras have been considered for coding applications since they naturally provide families of matrices. Since each $x \in \mathcal{A}$ is expressible as

$$x = x_0 + ex_1 + \ldots + e^{n-1}x_{n-1},\ x_i \in L \text{ for all } i,$$

there is a correspondence between $x \in \mathcal{A}$ and the matrix of multiplication by x given by

$$\begin{pmatrix} x_0 & \gamma\sigma(x_{n-1}) & \gamma\sigma^2(x_{n-2}) & \ldots & \gamma\sigma^{n-1}(x_1) \\ x_1 & \sigma(x_0) & \gamma\sigma^2(x_{n-1}) & \ldots & \gamma\sigma^{n-1}(x_2) \\ \vdots & & \vdots & & \vdots \\ x_{n-2} & \sigma(x_{n-3}) & \sigma^2(x_{n-4}) & \ldots & \gamma\sigma^{n-1}(x_{n-1}) \\ x_{n-1} & \sigma(x_{n-2}) & \sigma^2(x_{n-3}) & \ldots & \sigma^{n-1}(x_0) \end{pmatrix}. \tag{6}$$

Example 1: For $n = 2$, we have

$$\mathcal{A} = L \oplus eL$$

such that e satisfies

$$e^2 = \gamma \quad \text{and} \quad \lambda e = e\sigma(\lambda) \text{ for } \lambda \in L.$$

Now, an element $x \in \mathcal{A}$ can be written $x = x_0 + ex_1$. Let us compute the multiplication by x of any element $y \in \mathcal{A}$.

$$\begin{aligned} xy &= (x_0 + ex_1)(y_0 + ey_1) \\ &= x_0y_0 + x_0ey_1 + ex_1y_0 + ex_1ey_1 \\ &= x_0y_0 + e\sigma(x_0)y_1 + ex_1y_0 + \gamma\sigma(x_1)y_1, \end{aligned}$$

since $e^2 = \gamma$ and using the noncommutativity rule $\lambda e = e\sigma(\lambda)$. In matrix form in the basis $\{1, e\}$, this yields

$$xy = \begin{pmatrix} x_0 & \gamma\sigma(x_1) \\ x_1 & \sigma(x_0) \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}.$$

There is thus a correspondence

$$x = x_0 + e x_1 \in \mathcal{A} \leftrightarrow \begin{pmatrix} x_0 & \gamma\sigma(x_1) \\ x_1 & \sigma(x_0) \end{pmatrix}.$$

Note that in particular

$$e \in \mathcal{A} \leftrightarrow \begin{pmatrix} 0 & \gamma \\ 1 & 0 \end{pmatrix}.$$

Cyclic algebras provide us with a family of matrices as follows

$$\left\{ \mathbf{X} = \begin{pmatrix} x_0 & \gamma\sigma(x_{n-1}) & \gamma\sigma^2(x_{n-2}) & \dots & \gamma\sigma^{n-1}(x_1) \\ x_1 & \sigma(x_0) & \gamma\sigma^2(x_{n-1}) & \dots & \gamma\sigma^{n-1}(x_2) \\ \vdots & & \vdots & & \vdots \\ x_{n-2} & \sigma(x_{n-3}) & \sigma^2(x_{n-4}) & \dots & \gamma\sigma^{n-1}(x_{n-1}) \\ x_{n-1} & \sigma(x_{n-2}) & \sigma^2(x_{n-3}) & \dots & \sigma^{n-1}(x_0) \end{pmatrix} \mid x_i \in L \text{ for all } i \right\}. \tag{7}$$

It is clearly linear (since σ is). Thus

$$\det(\mathbf{X}_1 - \mathbf{X}_2) = \det(\mathbf{X}),\ \mathbf{0} \neq \mathbf{X} \in \mathcal{C},\ \text{for all } \mathbf{X}_1 \neq \mathbf{X}_2 \in \mathcal{C},$$

so that the diversity (defined in (4)) simplifies to

$$\zeta_{\mathcal{C}} = \frac{1}{2} \min_{\mathbf{X} \neq \mathbf{0}} |\det(\mathbf{X})|^{1/M},\ \mathbf{X} \in \mathcal{C}.$$

If now the cyclic algebra we consider is also a *field*, that is, by definition, that every element in $\mathcal{A}$ is invertible, this guarantees that $\zeta_{\mathcal{C}} > 0$.

Definition 1: We call a non-commutative field a *division algebra*.

In particular, a cyclic algebra which has also a structure of field is called a *cyclic division algebra*. It is thus enough to consider cyclic division algebras to get $\zeta_{\mathcal{C}} > 0$.

B. An involution on the algebra

With the notations of Section III-A, let $\mathcal{A} = (L/K, \sigma, \gamma)$ be a cyclic division algebra. We now have a linear family $\mathcal{C}$ of invertible matrices (as described in (7)) among which we are looking for Hermitian matrices, i.e., $\mathbf{X} \in \mathcal{C}$ such $\mathbf{X} = \mathbf{X}^\dagger$, where † denotes the conjugate transpose.

In [9], [11], the problem of finding unitary matrices in a cyclic algebra has been studied. It has been shown that taking the transpose conjugate of a matrix $\mathbf{X}$ is equivalent to apply an involution α on the corresponding element $x \in \mathcal{A}$.

$\mathcal{A}$		$\mathcal{M}_n(L)$
x	$\leftrightarrow$	$\mathbf{X}$
$\alpha(x)$	$\leftrightarrow$	$\mathbf{X}^\dagger$

With this notation, finding a unitary matrix $\mathbf{X}$, such that $\mathbf{X}\mathbf{X}^\dagger = \mathbf{I}$, translates into finding an element $x \in \mathcal{A}$ such that $x\alpha(x) = 1$. Since this approach was successful, it is natural to look for Hermitian matrices by considering the equation $x = \alpha(x)$.

$\mathcal{A}$		$\mathcal{M}_n(L)$
$x\alpha(x) = 1$	$\leftrightarrow$	$\mathbf{X}\mathbf{X}^\dagger = \mathbf{I}$
$\alpha(x) = x$	$\leftrightarrow$	$\mathbf{X}^\dagger = \mathbf{X}$

The involution α gives us a convenient way to work in the algebra $\mathcal{A}$, instead of working in $\mathcal{M}_n(L)$ directly. Let us first recall how the involution was defined on $\mathcal{A}$. Let $\alpha_L : L \to L$ denote the complex conjugation, then $\alpha : \mathcal{A} \to \mathcal{A}$ is given by

$$\alpha(x_0 + ex_1 + \ldots + e^{n-1}x_{n-1}) = \alpha_L(x_0) + e^{-1}\sigma^{-1}(\alpha_L(x_1)) + \ldots + e^{-(n-1)}\sigma^{-(n-1)}(\alpha_L(x_{n-1})).$$

We refer to [9], [11] for the technical hypothesis, and the proof that α as defined above indeed gives the correspondence $\alpha(x) \leftrightarrow \mathbf{X}^\dagger$.

Example 2: Take $n = 3$ (to obtain matrices of size 3). Let $\zeta_3 = e^{2i\pi/3}$ be a primitive third root of unity, and similarly let $\zeta_7 = e^{2i\pi/7}$ be a primitive 7th root of unity. For short, denote $\theta = \zeta_7 + \zeta_7^{-1}$. We consider the number field

$$L = \mathbb{Q}(\theta, \zeta_3) = \{a + b\theta + c\theta^2 \mid a, b, c \in \mathbb{Q}(\zeta_3)\}$$

with cyclic Galois group Gal$(L/\mathbb{Q}(\zeta_3)) = \langle\sigma\rangle$, $\sigma : \zeta_7 + \zeta_7^{-1} \mapsto \zeta_7^2 + \zeta_7^{-2}$. We consider the algebra $\mathcal{A} = (L/K, \sigma, \gamma)$, where $\gamma = \zeta_3$.

Let $\tau : \mathbb{Q}(\zeta_3) \to \mathbb{Q}(\zeta_3)$ be defined by $\tau(b_0 + b_1\zeta_3) = b_0 + b_1\zeta_3^2$ and $\zeta_3^2 = -\zeta_3 - 1$. This is clearly the complex conjugation. Since θ is totally real, the complex conjugation α_L on L can be written as

$$\begin{array}{rlcl} \alpha_L : & L & \to & L \\ & a_0 + a_1\theta + a_2\theta^2 & \mapsto & \tau(a_0) + \tau(a_1)\theta + \tau(a_2)\theta^2. \end{array}$$

The involution on $\mathcal{A}$ is thus defined by

$$\begin{array}{rlcl} \alpha : & \mathcal{A} & \to & \mathcal{A} \\ & x_0 + ex_1 + e^2x_2 & \mapsto & \alpha_L(x_0) + e\zeta_3^2\sigma(\alpha_L(x_2)) + e^2\zeta_3^2\sigma^2(\alpha_L(x_1)). \end{array}$$

The matrix corresponding to x is

$$\mathbf{X} = \begin{pmatrix} x_0 & x_1 & x_2 \\ \zeta_3\sigma(x_2) & \sigma(x_0) & \sigma(x_1) \\ \zeta_3\sigma^2(x_1) & \zeta_3\sigma^2(x_2) & \sigma^2(x_0) \end{pmatrix}.$$

Since $\alpha_L\sigma = \sigma\alpha_L$ [1], we can rewrite $\alpha(x)$, denoting $\alpha_L = \bar{}$, as

$$\overline{x_0} + e\zeta_3^2\overline{\sigma(x_2)} + e^2\zeta_3^2\overline{\sigma^2(x_1)}.$$

The matrix of $\alpha(x)$ is thus

$$\mathbf{X} = \begin{pmatrix} \overline{x_0} & \zeta_3^2\overline{\sigma(x_2)} & \zeta_3^2\overline{\sigma^2(x_1)} \\ \overline{x_1} & \overline{\sigma(x_0)} & \zeta_3^2\overline{\sigma^2(x_2)} \\ \overline{x_2} & \overline{\sigma(x_1)} & \overline{\sigma^2(x_0)} \end{pmatrix},$$

which is clearly $\mathbf{X}^\dagger$.
The characterization we will use in this work is thus

$$\mathbf{X} = \mathbf{X}^\dagger \iff x = \alpha(x). \tag{8}$$

We show in next section how to use it to obtain families of Hermitian matrices which are fully-diverse.

IV. Constructions in small dimensions

In this section, we give the constructions of the Cayley codes for 2, 3 and 4 antennas, and the simulation results of these codes compared to known constructions.

Let us start with some preliminary remarks. As already noticed in [2], if we are interested in M-dimensional matrices, the basis of matrices will contain at most M^2 matrices. In the following, we will thus give a basis of M^2 matrices, $M = 2, 3, 4$, since it is an empirical observation that to get high rate, it is better to have first $Q = M^2$ matrices in the basis, so as to keep the coefficients of α_q small. For the simulations presented here, we take α_q distributed as Cauchy random variables, and the codewords are normalized to have Frobenius norm 1, as suggested in [2]. The algebras used are division algebras as shown in [10].

The general procedure we will use for all the code constructions is

1) Consider a cyclic division algebra.

[1]That α_L commutes with $Gal(L/K)$ is a general requirement [9], [11]

2) Solve $x = \alpha(x)$.
3) Use the correspondence $x \leftrightarrow \mathbf{X}$ to get an Hermitian matrix $\mathbf{X}$ which is fully diverse.
4) Decompose $\mathbf{X}$ in a basis of Hermitian matrices, all of them being fully diverse. This will give the basis to encode the Cayley code.

A. *For 2 transmit antennas*

Consider the cyclic algebra $\mathcal{A} = (\mathbb{Q}(i,\sqrt{5})/\mathbb{Q}(i), \sigma, i)$, where $\sigma : \sqrt{5} \mapsto -\sqrt{5}$. Let $x \in \mathcal{A}$,

$$x = x_0 + ex_1,\ \ x_0, x_1 \in \mathbb{Q}(i,\sqrt{5}).$$

Using the characterization (8), we have

$$x_0 + ex_1 = \alpha(x_0) + (-i)e\sigma(\alpha(x_1)).$$

By unicity of the coefficients in a given basis, we have

$$x_0 = \alpha(x_0) = \bar{x_0},$$

implying that $x_0 \in \mathbb{Q}(\sqrt{5})$. Namely

$$x_0 = a_0 + \frac{\sqrt{5}+1}{2}b_0.$$

Looking at the second coefficient yields

$$x_1 = a_1 + \frac{1+\sqrt{5}}{2}b_1 = -i(a_1 + \frac{1-\sqrt{5}}{2}b_1),$$

with $a_1, b_1 \in \mathbb{Q}(i)$. Let $\theta = \frac{1+\sqrt{5}}{2}$. It is easy to solve this equation and to see that

$$\begin{aligned} a_1 &= s(1-\theta) - t\theta \\ b_1 &= t(1-\theta) - s\theta, \end{aligned}$$

with $s, t \in \mathbb{Q}$. Thus, x can be written

$$x = [a_0 + \theta b_0] + e[(s(1-\theta) - t\theta) + i(t(1-\theta) - s\theta)],$$

or equivalently, in matrix equations

$$\mathbf{X} = a_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b_0 \begin{pmatrix} \theta & 0 \\ 0 & 1-\theta \end{pmatrix} + s \begin{pmatrix} 0 & 1-\theta-i\theta \\ i\theta + (1-\theta) & 0 \end{pmatrix} + t \begin{pmatrix} 0 & -\theta + i(1-\theta) \\ -i(1-\theta) - \theta & 0 \end{pmatrix}.$$

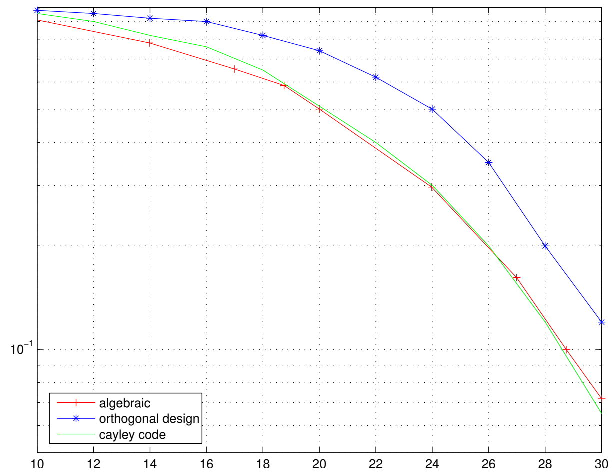

Fig. 1. $M = 2$, $N = 2$, $R = 6$, $Q = 4$

We thus get a basis of 4 matrices

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \theta & 0 \\ 0 & 1-\theta \end{pmatrix}, \begin{pmatrix} 0 & 1-\theta-i\theta \\ i\theta+(1-\theta) & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\theta+i(1-\theta) \\ -i(1-\theta)-\theta & 0 \end{pmatrix} \right\}.$$

Each matrix is easily checked to be Hermitian, and after normalizing by the Frobenius norm, we get

$$\left\{ \begin{matrix} \begin{pmatrix} 0.70711 & 0 \\ 0 & 0.70711 \end{pmatrix}, \begin{pmatrix} 0.93417 & 0 \\ 0 & -0.35682 \end{pmatrix}, \\ \begin{pmatrix} 0 & 0.66056-0.25231i \\ -0.66056-0.25231i & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0.25231-0.66056i \\ -0.25231-0.66056i & 0 \end{pmatrix} \end{matrix} \right\}.$$

In Figure 1, we can see that the Cayley code we obtain for $M = 2$ transmit antennas and $N = 2$ receive antennas is doing as good as the optimized Cayley code given in [2] for a rate of $R = 6$.

B. *For 3 transmit antennas*

We consider the algebra $\mathcal{A} = (\mathbb{Q}(\zeta_3, \zeta_7+\zeta_7^{-1})/\mathbb{Q}(\zeta_3), \sigma, \zeta_3)$, where $\sigma(\zeta_7+\zeta_7^{-1}) = \zeta_7^2+\zeta_7^{-2}$. Let $x \in \mathcal{A}$, $x = x_0 + ex_1 + e^2x_2$, with $x_i = a_i + \zeta_3 b_i$, $i = 1, 2, 3$. The condition $x = \alpha(x)$ gives

$$\begin{cases} x_0 &= \alpha(x_0) \\ a_1 &= -\sigma(a_2) \\ b_1 &= \sigma(b_2) - \sigma(a_2) \end{cases}$$

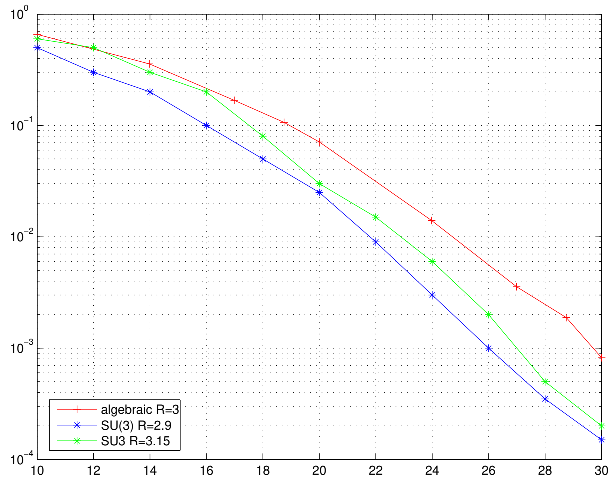

Fig. 2. $M = 3,\ N = 1,\ R = 3,\ Q = 9$

This yields $x = x_0 + ex_1 + e^2x_2$, with

$$\begin{aligned}
x_0 =& \quad a_0 + b_0\theta + c_0\theta^2 \\
x_1 =& \quad -s_0 - s_1(\theta^2 - 2) - s_2(3 - \theta - \theta^2) \\
& \quad +\zeta_3[t_0 + t_1(\theta^2 - 2) + t_2(3 - \theta - \theta^2) - s_0 - s_1(\theta^2 - 2) - s_2(3 - \theta - \theta^2)] \\
x_2 =& \quad s_0 + s_1\theta + s_2\theta^2 + \zeta_3(t_0 + t_1\theta + t_2\theta^2)
\end{aligned}$$

In matrix equations, this yields

$$\begin{aligned}
\mathbf{X} =& \; a_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + b_0 \begin{pmatrix} \theta & 0 & 0 \\ 0 & \sigma(\theta) & 0 \\ 0 & 0 & \sigma^2(\theta) \end{pmatrix} + c_0 \begin{pmatrix} \theta^2 & 0 & 0 \\ 0 & \sigma(\theta^2) & 0 \\ 0 & 0 & \sigma^2(\theta^2) \end{pmatrix} \\
&+ s_0 \begin{pmatrix} 0 & -1-\zeta_3 & 1 \\ \zeta_3 & 0 & -1-\zeta_3 \\ \zeta_3(-1-\zeta_3) & \zeta_3 & 0 \end{pmatrix} + s_1 \begin{pmatrix} 0 & -\theta^2+2-\zeta_3(\theta^2-2) & \theta \\ \zeta_3(-2+\theta^2) & 0 & \zeta_3^2(1-\theta-\theta^2) \\ \theta & \zeta_3(1-\theta-\theta^2) & 0 \end{pmatrix} \\
&+ s_2 \begin{pmatrix} 0 & \zeta_3^2(3-\theta-\theta^2) & \theta^2 \\ \zeta_3(3-\theta-\theta^2) & 0 & \zeta_3^2(2+\theta) \\ \theta^2 & \zeta_3(2+\theta) & 0 \end{pmatrix} + t_0 \begin{pmatrix} 0 & \zeta_3 & \zeta_3 \\ \zeta_3^2 & 0 & \zeta_3 \\ \zeta_3^2 & \zeta_3^2 & 0 \end{pmatrix} \\
&+ t_1 \begin{pmatrix} 0 & \zeta_3(\theta^2-2) & \zeta_3\theta \\ \zeta_3^2(\theta^2-2) & 0 & \zeta_3(1-\theta-\theta^2) \\ \zeta_3^2\theta & \zeta_3^2(1-\theta-\theta^2) & 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 & \zeta_3(3-\theta-\theta^2) & \zeta_3\theta^2 \\ \zeta_3^2(3-\theta-\theta^2) & 0 & \zeta_3(2+\theta) \\ \zeta_3^2\theta^2 & \zeta_3^2(2+\theta) & 0 \end{pmatrix} .
\end{aligned}$$

We clearly get a basis of 9 Hermitian matrices.

In Fig. 2, we compare the Cayley code with Lie group based codes proposed in [5]. The performances are close though the Lie group codes are a bit better. Note that the decoding algorithms are different, which can explain some of the loss, since Cayley codes are not decoded using true maximum likelihood (ML). See [2] for comparing the probability of error between true ML and the linearized Sphere Decoder.

C. *For 4 transmit antennas*

We consider the algebra $\mathcal{A} = (\mathbb{Q}(\zeta_{15}+\zeta_{15}^{-1})/\mathbb{Q}(i),\sigma,i)$, with $\sigma : \zeta_{15}+\zeta_{15}^{-1} \mapsto \zeta_{15}^{2}+\zeta_{15}^{-2}$. Let $x = x_0 + ex_1 + e^2x_2 + e^3x_3$ be an element of $\mathcal{A}$. We have

$$x = \alpha(x) \iff \begin{cases} x_0 &= \alpha(x_0) \\ x_1 &= -i\sigma(\alpha(x_3)) \\ x_2 &= -i\sigma^2(\alpha(x_2)) \\ x_3 &= -i\sigma^3(\alpha(x_1)) \end{cases}$$

The first equation, as previously, tells that

$$x_0 = a_0 + b_0\theta + c_0\theta^2 + d_0\theta^3,\ a_0,b_0,c_0,d_0 \in \mathbb{Q}.$$

This yields 4 diagonal matrices:

$$\begin{aligned}\mathbf{X}_0 \;=\; & a_0\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + b_0\begin{pmatrix} \theta & 0 & 0 & 0 \\ 0 & \sigma(\theta) & 0 & 0 \\ 0 & 0 & \sigma^2(\theta) & 0 \\ 0 & 0 & 0 & \sigma^3(\theta) \end{pmatrix} \\ & +c_0\begin{pmatrix} \theta^2 & 0 & 0 & 0 \\ 0 & \sigma(\theta^2) & 0 & 0 \\ 0 & 0 & \sigma^2(\theta^2) & 0 \\ 0 & 0 & 0 & \sigma^3(\theta^2) \end{pmatrix} + d_0\begin{pmatrix} \theta^3 & 0 & 0 & 0 \\ 0 & \sigma(\theta^3) & 0 & 0 \\ 0 & 0 & \sigma^2(\theta^3) & 0 \\ 0 & 0 & 0 & \sigma^3(\theta^3) \end{pmatrix}\end{aligned}$$

The third equation depends only on x_2. Writing $x_2 = a_2 + b_2\theta + c_2\theta^2 + d_2\theta^3$ and solving it gives

$$\begin{cases} a_2 &= s_0(i-1) - 5s_1 + s_2 + s_3 \\ b_2 &= s_1(1+i) + s_2(-1-4i) + s_3(-4-i) \\ c_2 &= (1+i)s_1 \\ d_2 &= is_2 + s_3 \end{cases}$$

The coefficient e^2x_2 can thus be written as the following matrix:

$$s_0\begin{pmatrix} 0 & 0 & i-1 & 0 \\ 0 & 1 & 0 & i-1 \\ i(i-1) & 0 & 0 & 0 \\ 0 & i(i-1) & 0 & 0 \end{pmatrix}$$

$$+s_1\begin{pmatrix} 0 & 0 & -5+(1+i)(\theta+\theta^2) & 0 \\ 0 & 0 & 0 & -5+(1+i)(\sigma(\theta)+\sigma(\theta^2)) \\ -5+(1-i)(\theta+\theta^2) & 0 & 0 & 0 \\ 0 & -5+(1-i)(\sigma(\theta)+\sigma(\theta^2)) & 0 & 0 \end{pmatrix}$$

$$+s_2\begin{pmatrix} 0 & 0 & 1+\theta(-1-4i)+i\theta^3 & 0 \\ 0 & 0 & 0 & 1+\sigma(\theta)(-1-4i)+i\sigma(\theta^3) \\ 1+\theta(-1+4i)-i\theta^3 & 0 & 0 & 0 \\ 0 & 1+\sigma(\theta)(-1+4i)-i\sigma(\theta^3) & 0 & 0 \end{pmatrix}$$

$$+s_3\begin{pmatrix} 0 & 0 & 1+\theta(-4-i)+\theta^3 & 0 \\ 0 & 0 & 0 & 1+\sigma(\theta)(-4-i)+\sigma(\theta^3) \\ 1+\theta(-4+i)+\theta^3 & 0 & 0 & 0 \\ 0 & 1+\sigma(\theta)(-4+i)+\sigma(\theta^3) & 0 & 0 \end{pmatrix}$$

The two equations left to solve are

$$\begin{cases} x_1 & = & -i\sigma(\alpha(x_3)) \\ x_3 & = & -i\sigma^3(\alpha(x_1)). \end{cases}$$

However, it is easy to see that they are redundant (for example, by applying twice σ on the first equation). It suffices thus to consider, for example, the first equation. Writing again $x_1 = a_1 + b_1\theta + c_1\theta^2 + d_1\theta^3$ and solving the first equation gives

$$\begin{cases} a_1 & = & -t_1+2t_3-3t_5+7t_7+i(-t_0+2t_2-3t_4+7t_6) \\ b_1 & = & 4t_5-3t_7+i(4t_4-3t_6) \\ c_1 & = & -t_3-3t_7+i(-t_2-3t_6) \\ d_1 & = & -t_7-t_5+i(-t_4+t_6) \end{cases}$$

The coefficient $ex_1 + e^3x_3$ can thus be written as the following matrix:

$$t_0 \begin{pmatrix} 0 & -i & 0 & 1 \\ i & 0 & -i & 0 \\ 0 & i & 0 & -i \\ 1 & 0 & i & 0 \end{pmatrix} + t_1 \begin{pmatrix} 0 & -1 & 0 & i \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \\ -i & 0 & -1 & 0 \end{pmatrix}$$

$$+t_2 \begin{pmatrix} 0 & i(2-\theta^2) & 0 & \theta \\ -i(2-\theta^2) & 0 & -i(1-4\theta+\theta^3) & 0 \\ 0 & i(1-4\theta+\theta^3) & 0 & -i(2+3\theta-\theta^2-\theta^3) \\ \theta & 0 & i(2+3\theta-\theta^2-\theta^3) & 0 \end{pmatrix}$$

$$+t_3 \begin{pmatrix} 0 & 2-\theta^2 & 0 & i\theta \\ 2-\theta^2 & 0 & -(1-4\theta+\theta^3) & 0 \\ 0 & -(1-4\theta+\theta^3) & 0 & -(2+3\theta-\theta^2-\theta^3) \\ \theta & 0 & -(2+3\theta-\theta^2-\theta^3) & 0 \end{pmatrix}$$

$$+t_4 \begin{pmatrix} 0 & -i(3-4\theta+\theta^3) & 0 & \theta^2 \\ i(3-4\theta+\theta^3) & 0 & -i(4+3\theta-\theta^2-\theta^3) & 0 \\ 0 & i(4+3\theta-\theta^2-\theta^3) & 0 & -i(2+\theta) \\ \theta^2 & 0 & i(2+\theta) & 0 \end{pmatrix}$$

$$+t_5 \begin{pmatrix} 0 & -(3-4\theta+\theta^3) & 0 & i\theta^2 \\ -(3-4\theta+\theta^3) & 0 & -(4+3\theta-\theta^2-\theta^3) & 0 \\ 0 & -(4+3\theta-\theta^2-\theta^3) & 0 & -(2+\theta) \\ -i\theta^2 & 0 & -(2+\theta) & 0 \end{pmatrix}$$

$$+t_6 \begin{pmatrix} 0 & -i(-7+3\theta+3\theta^2-\theta^3) & 0 & \theta^3 \\ i(-7+3\theta+3\theta^2-\theta^3) & 0 & -i(3-15\theta+4\theta^3) & 0 \\ 0 & i(3-15\theta+4\theta^3) & 0 & -i(5+12\theta-3\theta^2-4\theta^3) \\ \theta^3 & 0 & i(5+12\theta-3\theta^2-4\theta^3) & 0 \end{pmatrix}$$

$$+t_7 \begin{pmatrix} 0 & -(-7+3\theta+3\theta^2-\theta^3) & 0 & i\theta^3 \\ -(-7+3\theta+3\theta^2-\theta^3) & 0 & -(3-15\theta+4\theta^3) & 0 \\ 0 & -(3-15\theta+4\theta^3) & 0 & -(5+12\theta-3\theta^2-4\theta^3) \\ -i\theta^3 & 0 & -(5+12\theta-3\theta^2-4\theta^3) & 0 \end{pmatrix}$$

We thus get 16 Hermitian matrices. In Figure 3, we compare, for 4 transmit antennas and 1 receive

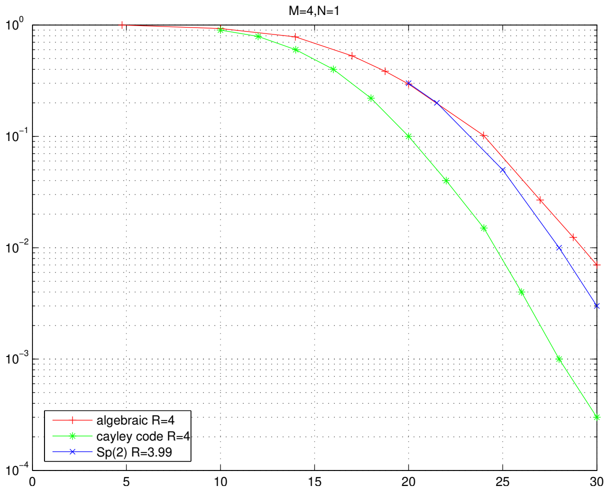

Fig. 3. $M = 4$, $N = 1$, $R = 4$,$Q = 16$

antennas, the new Cayley code to a Lie group code [6] and to an optimized Cayley code given in [2]. The new Cayley code and the Lie group code are very close. The previous Cayley code clearly performs better. However, again, one has to be careful about the decoding algorithm used. The performance of the previous Cayley code has been reported from [2], where the decoding was done this time using true maximum likelihood through exhaustive search. The Lie group has been decoded using a Sphere Decoder algorithm, while the new Cayley code has been decoded through a linearized Sphere Decoder.

V. Conclusion

In this work, we looked at Cayley codes for the differential MIMO noncoherent channel. We use division algebras in order to get fully diverse Cayley codes. This yields a family of codes that naturally fulfills the challenge of performing similarly or close to previous Cayley codes in dimension 2, 3 and 4, as well as Lie group based codes in dimension 3 and 4.

References

[1] B. Hassibi and H. Vikalo, “On sphere decoding algorithm. I. Expected complexity”, *IEEE Transactions on Signal Processing*, vol. 53, no. 8, August 2005, pp. 2806 - 2818.

[2] B. Hassibi and B. Hochwald,“Cayley Differential Unitary Space-Time Codes”, *IEEE Trans. on Information Theory*, vol.48, June 2002.

[3] B. Hochwald and W. Sweldens,“Differential unitary space time modulation”, *IEEE Trans. Commun.*, vol.48, Dec. 2000.

[4] B. Hughes, “Differential space-time modulation,” *IEEE Trans. Inform. Theory*, vol 46, Nov. 2000.

[5] Y. Jing, B. Hassibi, “Three-Transmit-Antenna Space-Time Codes Based on SU(3)”, *IEEE Trans. Signal Processing*, vol. 53, no 10, Oct. 2005.

[6] Y. Jing, B. Hassibi, “Design of full-diverse multiple-antenna codes based on Sp(2)”, *IEEE Trans. on Information Theory*, vol. 50, no 11, Nov. 2004.

[7] Xue-Bin Liang and Xiang-Gen Xia, “Unitary Signal Constellations for Differential Space-Time Modulation With Two Transmit Antennas: Parametric Codes, Optimal Designs, and Bounds,” *IEEE Trans. on Inform. Theory*, vol.48, no.8, August 2002.

[8] F. Oggier and E. Viterbo, “Algebraic number theory and code design for Rayleigh fading channels,” in *Foundations and Trends in Communications and Information Theory*, vol. 1, pp. 333–415, 2004.

[9] F. Oggier, E. Lequeu, “Families of unitary matrices achieving full diversity,”, *proceedings of the International Symposium on inf. theory, 2005, Adelaide*.

[10] F. E. Oggier, G. Rekaya, J.-C. Belfiore and E. Viterbo, “Perfect Space-Time Block Codes”, to appear in *IEEE Trans. on Information Theory*, September 2006.

[11] F. Oggier, “Cyclic Algebras for Noncoherent Differential Space-Time Coding”, submitted to *IEEE Trans. on Information Theory*, June 2006. .

[12] R. S. Pierce, “Associative Algebras”, Springer-Verlag, New York, 1982.

[13] I. Reiner,“Maximal Orders”, Academic Press, 1975.

[14] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, “Full-diversity, high-rate space-time block codes from division algebras,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2596– 2616, October 2003.

[15] A. Shokrollahi, B. Hassibi, B. M. Hochwald, W. Sweldens, “Representation Theory for High-Rate Multiple-Antenna Code Design”, *IEEE Trans. Information Theory*, vol. 47, no 6, Sept. 2001.

[16] I.N. Steward and D.O.Tall,“Algebraic Number Theory”, Chapman and Hall, 1979.

[17] E. Viterbo and J. Boutros: “A Universal Lattice Code Decoder for Fading Channels,” *IEEE Transactions on Information Theory*, vol. 45, n. 5, pp. 1639–1642, July 1999.