Nonintersecting Subspaces Based on Finite Alphabets

Frédérique E. Oggier[*]
Département de Mathématiques
Ecole Polytechnique Fédérale de Lausanne
1015 Lausanne - Switzerland
(Email: frederique.oggier@epfl.ch)

N. J. A. Sloane
Information Sciences Research Center
AT&T Shannon Labs
Florham Park, NJ 07932–0971, USA
(Email: njas@research.att.com)

Suhas N. Diggavi
School of Computer and Communication Sciences
Ecole Polytechnique Fédérale de Lausanne
1015 Lausanne - Switzerland
(Email: suhas.diggavi@epfl.ch)

A. R. Calderbank
Program in Applied and Computational Mathematics
Princeton University
Princeton, NJ 08540, USA
(Email: calderbank@math.princeton.edu)

Abstract

Two subspaces of a vector space are here called "nonintersecting" if they meet only in the zero vector. Motivated by the design of noncoherent multiple-antenna communications systems, we consider the following question. How many pairwise nonintersecting M_t-dimensional subspaces of an m-dimensional vector space V over a field $\mathbb{F}$ can be found, if the generator matrices for the subspaces may contain only symbols from a given finite alphabet $\mathcal{A} \subseteq \mathbb{F}$? The most important case is when $\mathbb{F}$ is the field of complex numbers $\mathbb{C}$; then M_t is the number of antennas. If $\mathcal{A} = \mathbb{F} = GF(q)$ it is shown that the number of nonintersecting subspaces is at most $(q^m - 1)/(q^{M_t} - 1)$, and that this bound can be attained if and only if m is divisible by M_t. Furthermore these subspaces remain nonintersecting when "lifted" to the complex field. Thus the finite field case is essentially completely solved. In the

[*]This work was carried out during F. E. Oggier's visit to AT&T Shannon Labs during the summer of 2003. She thanks the Fonds National Suisse, Bourses et Programmes d'Échange for support.

case when $\mathbb{F} = \mathbb{C}$ only the case $M_t = 2$ is considered. It is shown that if $\mathcal{A}$ is a PSK-configuration, consisting of the 2^r complex roots of unity, the number of nonintersecting planes is at least $2^{r(m-2)}$ and at most $2^{r(m-1)-1}$ (the lower bound may in fact be the best that can be achieved).

1. Introduction

The problem studied in this paper is motivated by communication over multiple antenna channels [10, 21, 22]. In [6], [22] it was shown that the capacity of the multiple-antenna channel grows linearly as a function of the minimum of the numbers of transmitting and receiving antennas. The proof assumed that the receiver has complete information about the channel. In [21] the emphasis was placed on reducing error probability by introducing correlation between signals transmitted from different antennas. These points of view can be combined by observing that there is a trade-off between rate and reliability [21], [23].

Most of the early work on multiple-antenna communications assumed that the receiver was able to track the channel perfectly—*i.e.,* used coherent detection. If coherent detection is difficult or too expensive, one can use noncoherent detection, as studied in [10]. The main result from this work is that the capacity is still (almost) linear in the minimum number of transmitting or receiving antennas [10], [26]. Hence, both in the coherent and noncoherent cases, it was established that the use of multiple-antennas leads to a gain in information transmission rate.

In [11], the error probability of multiple-antenna noncoherent communication channels was investigated. It was shown there (and in [26]) that if the channel is not known to the receiver, the coding problem is equivalent to one of packing subspaces (which represent codewords) according to a certain notion of distance. If an orthonormal basis for the subspaces were used, the diversity order (the slope of the error probability with respect to signal-to-noise ratio) was shown to depend on the dimension of the intersection of the subspaces.

In particular, to obtain maximal diversity, one wishes to construct a family of subspaces which intersect only at the origin. By a slight abuse of notation we will say that two vector spaces are "nonintersecting" if their only common point is the zero vector. An extensive characterization and classification of such group differential space-time code constructions was given in [19]. The focus of much of this work is on constructing codes which have the nonintersecting subspace property without imposing any constraints on the number of different symbols used to define the codewords—that is, the codewords are allowed to use a signal constellation that is larger than the minimum possible.

Motivated by these observations, the main question addressed in the present paper is the construc-

tion of nonintersecting subspaces, subject to the constraint that the codewords are defined using symbols from a fixed, small constellation. We focus on two cases: one in which the symbols are taken from a finite field and the other where they are taken from a PSK arrangement, *i.e.*, are complex roots of unity. Our aim is to find constructions that give the largest number of nonintersecting subspaces (*i.e.*, have the highest rate) subject to these constraints. Note that the constructions given in this paper could result in non-unitary space-time codes.

It is worth remarking that a recent paper by Lusina et al. [16] discusses an analogous problem for the case of coherent decoders. Another related paper is Lu and Kumar [15] explores code constructions with fixed alphabet constraints for achieving different points on the rate-diversity trade-off. Again, only coherent decoders are considered. A very recent paper by Kammoun and Belfiore [13] directly addresses the problem of constructing codes for non-coherent systems with a large value of $\Lambda(\mathbf{X}, \mathbf{X}')$ (see (7)) between subspaces. However, their approach is quite different from ours.

The present paper is organized as follows. In Section 2, we establish notation and formalize the question being studied. In Section 3, we study the case when the symbols are taken from a finite field, and in Section 4 when they are complex roots of unity (*i.e.*, PSK constellations). Section 5 compares the different constructions and mentions some directions for further research.

2. Preliminaries

In order to motivate the question studied in this paper and establish notation, consider a multiple antenna channel. Let the number of transmitting antennas be M_t and the number of receiving antennas be M_r. If $\mathbf{y}(k) \in \mathbb{C}^{M_r}$ is the received (column) vector at time k, we can write

$$\mathbf{y}(k) = \sqrt{E_s}\,\mathbf{H}(k)\,\mathbf{x}(k) + \mathbf{z}(k)\,, \tag{1}$$

where the matrix $\mathbf{H}(k) \in \mathbb{C}^{M_r \times M_t}$ represents the channel, the column vector $\mathbf{x}(k) \in \mathbb{C}^{M_t}$ is the channel input, E_s is the signal power per transmitting antenna, and $\mathbf{z}(k) \in \mathbb{C}^{M_r}$ is zero mean i.i.d. Gaussian noise with $\mathbb{E}[\mathbf{z}(k)\mathbf{z}(k)^H] = N_0\mathbf{I}$. We assume a Rayleigh flat fading model, *i.e.*, that the elements of $\mathbf{H}(k)$ are i.i.d. with a zero mean complex Gaussian distribution of unit variance. The channel is assumed to be block time-invariant, that is, $\mathbf{H}(k)$ is independent of k over a transmission block of m symbols, say $\mathbf{H}(k) = \mathbf{H}$ (although $\mathbf{H}(k)$ may vary from block to block). Looking at a single block of length m, during which the channel is assumed to be time-invariant, we can write

$$\mathbf{Y} = [\mathbf{y}(1), \ldots, \mathbf{y}(m)] = \sqrt{E_s}\,\mathbf{H}\,[\mathbf{x}(1), , \ldots, \mathbf{x}(m)] + [\mathbf{z}(1), \ldots, \mathbf{z}(m)] = \sqrt{E_s}\mathbf{H}\mathbf{X} + \mathbf{Z}\,. \tag{2}$$

The focus of this paper is on constructing the space-time codewords $\mathbf{X}$, subject to the constraint that the elements of $\mathbf{X}$ are selected from a particular alphabet $\mathcal{A}$.

2.1. Criteria for code design

In this paper we assume that the receiver will not attempt to estimate the channel matrix $\mathbf{H}$, *i.e.,* that we have a noncoherent receiver. Therefore, the maximum likelihood detection rule without using the channel state information ([11]) is that we should decode $\mathbf{Y}$ as that codeword $\hat{\mathbf{X}}$ which maximizes

$$\frac{\exp(-\operatorname{Trace}[\mathbf{Y}\Psi^{-1}\mathbf{Y}^H])}{|\pi\Psi|^{M_r}}\,, \tag{3}$$

where $\Psi = \mathbf{I}+E_s\mathbf{X}^H\mathbf{X}$, H denotes the transposed complex conjugate or adjoint matrix, and $|\cdot|$ denotes a determinant. Using the matrix inversion lemma [12, p. 19] and the property that $|\mathbf{I}+\mathbf{AB}| = |\mathbf{I}+\mathbf{BA}|$, the detection criterion can be rewritten as

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{X}} \left\{ E_s \operatorname{Trace}[\mathbf{Y}\mathbf{X}^H(\mathbf{I}+E_s\mathbf{X}\mathbf{X}^H)^{-1}\mathbf{X}\mathbf{Y}^H] - \log|\mathbf{I}+E_s\mathbf{X}\mathbf{X}^H| \right\}\,. \tag{4}$$

In the absence of channel state information at the receiver, Hochwald and Marzetta [11] argue that, at high signal-to-noise ratio, one should use unitary codewords $\mathbf{X}$, satisfying $\mathbf{X}\mathbf{X}^H = m\mathbf{I}$. Using this in (4), it follows that $\hat{\mathbf{X}}$ should be chosen to maximize

$$\operatorname{Trace}[\mathbf{Y}\mathbf{X}^H\mathbf{X}\mathbf{Y}^H]\,. \tag{5}$$

This implies that, for unitary codewords, the decoder should project the received signal onto the subspace defined by each of the codewords and declare the codeword with the maximal projection to be the winner. In [11] it is shown that, for unitary codewords, the probability that a transmitted codeword $\mathbf{X}$ is decoded as the codeword $\hat{\mathbf{X}}$ is bounded above by

$$\frac{1}{|\mathbf{I}_{M_t} + \frac{\rho^2 m^2}{4(1+\rho m)}[\mathbf{I}_{M_t} - \frac{1}{m^2}\hat{\mathbf{X}}\mathbf{X}^H\mathbf{X}\hat{\mathbf{X}}^H]|^{M_r}}\,, \tag{6}$$

where $\rho = \frac{E_s}{N_0}$ is the signal-to-noise ratio. If the signal-to-noise ratio is large, this pairwise error probability behaves like $(\frac{\Lambda\rho}{4})^{-M_r\nu}$, where ν is the rank of $[\mathbf{I}_{M_t} - \frac{1}{m^2}\hat{\mathbf{X}}\mathbf{X}^H\mathbf{X}\hat{\mathbf{X}}]$,

$$\Lambda = \Lambda(\mathbf{X},\hat{\mathbf{X}}) = |m\mathbf{I}_{M_t} - \frac{1}{m}\hat{\mathbf{X}}\mathbf{X}^H\mathbf{X}\hat{\mathbf{X}}^H|_+^{\frac{1}{\nu}}\,,$$

and $|\cdot|_+$ denotes the product of the nonzero eigenvalues. Note that

$$\left|\begin{bmatrix}\mathbf{X}\\ \hat{\mathbf{X}}\end{bmatrix}\begin{bmatrix}\mathbf{X}^H & \hat{\mathbf{X}}^H\end{bmatrix}\right| = |m^2\mathbf{I}_{M_t} - \hat{\mathbf{X}}\mathbf{X}^H\mathbf{X}\hat{\mathbf{X}}^H|\,,$$

which shows that $\nu = M_t$ is equivalent to the condition that the rows of $\mathbf{X}, \hat{\mathbf{X}}$ are linearly independent [11]. For this to happen we must have $m \geq 2M_t$.

Another interpretation can be given in terms of the principal angles between subspaces corresponding to pairs of codewords. The principal angles between subspaces $\mathbf{X}$ and $\mathbf{X}'$ are given by $\cos\theta_i = \frac{1}{m}\sigma_i(\mathbf{X}'\mathbf{X}^H)$ where $\sigma_i(\cdot)$ is the i-th singular value of the matrix ([4], [7]). Using this we obtain

$$\Lambda(\mathbf{X}, \mathbf{X}') = m \prod_{i=1}^{\nu} [1 - \cos^2\theta_i] = m \prod_{i=1}^{\nu} \sin^2\theta_i \, . \tag{7}$$

This provides a better measure of how good a code is: not only should the subspaces be nonintersecting, the value of $\Lambda(\mathbf{X}, \mathbf{X}')$ should be large for every pair $\mathbf{X}$, $\mathbf{X}'$ of distinct subspaces. The error probability will be dominated by the pair of codewords with the least rank ν and the least "distance" $\Lambda(\mathbf{X}, \mathbf{X}')$. For well separated subspaces this "distance" can also be approximated by

$$\sum_{i=1}^{\nu} \sin^2\theta_i \, , \tag{8}$$

which is the the notion of distance between subspaces used in [4] and [2].

Another way to compare these codes is by using the notion of diversity order (cf. [21]).

Definition 2.1. *If the average error probability $\bar{P}_e(\rho)$ as a function of the signal-to-noise ratio ρ satisfies*

$$\lim_{\rho \to \infty} \frac{\log(\bar{P}_e(\rho))}{\log(\rho)} = -d \, , \tag{9}$$

the coding scheme is said to have diversity order *d.*

It follows from (6) that the diversity order of the unitary space-time coding scheme is equal to $M_r\nu$. The maximal diversity order that can be achieved is therefore M_rM_t. We call codes that achieve this bound *fully diverse* codes.

In brief, to get a diversity order of M_rM_t, we need to construct nonintersecting subspaces which are far apart in the metric defined by (7). Though this property of nonintersecting subspaces yielding fully diverse codes was demonstrated for unitary codewords [11], non-unitary constructions with this property also yield fully diverse codes using the appropriate detection rule given in (4). This can be seen from the following argument. Consider a single receive antenna ($M_r = 1$). In the absence of noise, the received signal $\mathbf{Y} = \sqrt{E_s}\mathbf{H}\mathbf{X}$, and since $\mathbf{H} \in \mathbb{C}^{1\times M_t}$ the received vector $\mathbf{Y} \in \mathrm{rowspan}(\mathbf{X})$. Therefore, unless $\mathbf{H}$ is zero, we can distinguish the different codewords as long as the subspaces are nonintersecting. In applying this argument to the high signal-to-noise ratio regime, we just need

$||\mathbf{H}||^2 \gg 1/\mathrm{SNR}$ for us to be able to distinguish nonintersecting subspaces in the presence of noise. Since for $\mathbf{H} \sim \mathbb{C}\mathcal{N}(0, \mathbf{I}_{M_t})$ and high signal-to-noise ratio, we have $\mathbb{P}[||\mathbf{H}||^2 < 1/\mathrm{SNR}] \approx 1/\mathrm{SNR}^{M_t}$ [23], we see that nonintersecting subspaces would yield fully diverse codes. This argument can be made precise, but that is beyond the scope of this paper since we only use non-coherent multiple antenna systems to motivate the problem of nonintersecting subspaces with a finite alphabet.

Given this motivation, we will focus on obtaining maximal diversity order by constructing families of subspaces which are nonintersecting. Note that some of the constructions could yield non-unitary space-time codes. In order to further improve performance we need to maximize $\Lambda(\mathbf{X}, \mathbf{X}')$ over all pairs $\mathbf{X}$, $\mathbf{X}'$ of distinct subspaces. The rate of a code C is $R = \frac{1}{m}\log(|\mathcal{C}|)$. In trying to construct the maximal number of nonintersecting subspaces, we attempt to get the highest rate codes that achieve maximal diversity order.

2.2. Statement of the problem

Definition 2.2. *Let $\mathbb{F}$ be a field. A* codeword *or* subspace *will mean an M_t-dimensional subspace of $\mathbb{F}^m$. Two subspaces Π_1 and Π_2 are said to be* nonintersecting *over $\mathbb{F}$ if their intersection is trivial,* i.e., *if $\Pi_1 \cap \Pi_2 = \{0\}$.*

Suppose Π_1 is generated by (row) vectors $u_1, \ldots, u_{M_t} \in \mathbb{F}^m$, and Π_2 is generated by vectors $v_1, \ldots, v_{M_t} \in \mathbb{F}^m$. Let $P := \begin{bmatrix} \Pi_1 \\ \Pi_2 \end{bmatrix}$ denote the $2M_t \times m$ matrix with rows $u_1, \ldots, u_{M_t}, v_1, \ldots, v_{M_t}$. Then the following lemma is readily established.

Lemma 2.1. *The following properties are equivalent: (i) Π_1 and Π_2 are nonintersecting, (ii) P has rank $2M_t$ over $\mathbb{F}$, and (iii) if $m = 2M_t$ the determinant of P is nonzero.*

Suppose now that instead of allowing the entries in the matrices Π_1 and Π_2 to be arbitrary elements of $\mathbb{F}$, we restrict them to belong to a finite subset $\mathcal{A} \subseteq \mathbb{F}$, called the *alphabet*. In other words, the vectors $u_1, \ldots, u_{M_t}, v_1, \ldots, v_{M_t}$ must belong to $\mathcal{A}^m$. The question that we address is the following: given M_t, m and a finite alphabet $\mathcal{A} \subseteq \mathbb{F}$, how many subspaces can we find which are generated by vectors from $\mathcal{A}^m$ and which are pairwise nonintersecting over $\mathbb{F}$? Furthermore, if the size of $\mathcal{A}$ is specified in advance, which choice of $\mathcal{A}$ permits the biggest codes?

We first dispose of the trivial case when $M_t = 1$. Two nonzero vectors u, v are said to be *projectively distinct* over a field $\mathbb{F}$ if there is no $a \in F$ such that $u = av$. Then if $M_t = 1$, the maximum number of nonintersecting subspaces is simply the maximum number of projectively distinct vectors in $\mathcal{A}^m$.

In the following sections we will investigate the first question for two kinds of alphabets: (a) $\mathcal{A}$ is a finite field $\mathbb{F}$ (Section 3), and (b) $M_t = 2$ and $\mathcal{A} \subseteq \mathbb{C}^m$ is a set of complex roots of unity (Section 4).

Of course, for the application to multiple-antenna code design, the subspaces need to be disjoint over $\mathbb{C}$. In Theorem 3.4 of Section 3 we translate the results obtained over $\mathbb{F}$ to this case by "lifting" the subspaces to the complex field. Furthermore, for this application, the case $m = 2M_t$ is the most important.

3. Finite Fields

In this section we assume that the alphabet $\mathcal{A}$ and the field $\mathbb{F}$ are both equal to the finite field $GF(q)$, where q is a power of a prime p. At the end of the section we show how to "lift" these planes to the complex field (see Theorem 3.4). In this case there is an obvious upper bound which can be achieved in infinitely many cases. Let V denote the vector space $GF(q)^m$.

Theorem 3.1. *The number of pairwise nonintersecting M_t-dimensional subspaces of V is at most*

$$\frac{q^m - 1}{q^{M_t} - 1} \,. \tag{10}$$

Proof: There are $q^m - 1$ nonzero vectors in V and each subspace contains $q^{M_t} - 1$ of them. No nonzero vector can appear in more than one subspace. ■

It is convenient here to use the language of projective geometry, c.f. [17, Appendix B]. Recall that the points of the projective space $P(s, q)$ are equivalence classes of nonzero vectors from $GF(q)^{s+1}$, where two vectors are regarded as equivalent if one is a nonzero scalar multiple of the other.

A *spread* [9] in $PG(s, q)$ is a partition of the points into copies of $PG(r, q)$.

Theorem 3.2. *Such a spread exists if and only if $r + 1$ divides $s + 1$.*

Proof: This is a classical result, due to André ([1]; [9, Theorem 4.1.1]). ■

Corollary 3.3. *The bound* (10) *can be attained whenever M_t divides m, and only in those cases.*

Proof: This is immediate from the theorem, since a set of points in a projective space represents a set of projectively distinct lines in the corresponding vector space. ■

Note that the condition is independent of q. If a set of nonintersecting subspaces meeting (10) exists over one finite field then it exists over every finite field.

Furthermore, it is straightforward to construct the nonintersecting subspaces meeting the bound in (10), as we now show. The nonzero elements of a finite field $\mathbb{F}$ form a multiplicative group which will be denoted by $\mathbb{F}^*$. This is a cyclic group [14, Chap. 2].

Suppose M_t divides m, and consider the fields $F_0 = GF(q)$, $F_1 = GF(q^{M_t})$, $F_2 = GF(q^m)$. Then $F_0 \subseteq F_1 \subseteq F_2$. By regarding $GF(q^m)$ as a vector space of dimension m over $GF(q)$ we can identify F_2 with V. Similarly we can regard F_1 as a M_t-dimensional subspace of V. The desired spread is now obtained by partitioning F_2^* into (multiplicative) cosets of F_1^*.

Example 3.1. We consider the case $M_t = 2$, $m = 4$ and $\mathcal{A} = GF(2) = \{0, 1\}$. Then $F_0 = GF(2), F_1 = GF(4), F_2 = GF(16)$. Each plane in $GF(2)^4$ contains three nonzero vectors, and $GF(2)^4$ itself contains 15 nonzero vectors. We wish to find a spread of $PG(1, 2)$'s inside $PG(3, 2)$, that is, a partitioning of the 15 vectors into five disjoint sets of three, where each set of three adds to the zero vector.

Let $GF(16) = GF(2)[\alpha]$, where $\alpha^4 + \alpha + 1 = 0$. A table of the elements of this field and their binary representations can be found for example in [17, Fig. 3.3]. Then $GF(4)$ is the subfield $\{1, \alpha^5, \alpha^{10}\}$, so $F_1^* = \{\alpha^5, \alpha^{10}\}$, and we obtain the desired partition

$$F_2^* = \bigcup_{j=0}^{4} \alpha^j F_1^* \, .$$

Only two of the three vectors are needed to define each plane, so we have the following generators for the five planes:

$$(1, \alpha),\ (\alpha, \alpha^6),\ (\alpha^2, \alpha^7),\ (\alpha^3, \alpha^8),\ (\alpha^4, \alpha^9) \, .$$

Using the table in [17], we convert these to explicit generator matrices for the five nonintersecting planes:

$$\left[\begin{array}{c} 1000 \\ 0110 \end{array}\right], \left[\begin{array}{c} 0100 \\ 0011 \end{array}\right], \left[\begin{array}{c} 0010 \\ 1101 \end{array}\right], \left[\begin{array}{c} 0001 \\ 1010 \end{array}\right], \left[\begin{array}{c} 1100 \\ 0101 \end{array}\right].$$

The problem is therefore essentially solved as long as M_t divides m. If not, we can use partial spreads–see the surveys in [5] and [20].

We end this section by observing that a set of nonintersecting subspaces over a finite field $\mathcal{A} = GF(q)$, $q = p^k$, p prime, can always be "lifted" to a set of nonintersecting subspaces over a complex alphabet $\bar{\mathcal{A}}$ of the same size.

This can be done as follows. Suppose $GF(q) = GF(p)[\alpha]$, where α is a root of a primitive irreducible polynomial $f(X) \in GF(p)[X]$. Let $n = p^k - 1$ and let $\mu_n = e^{2\pi i/n}$. Adjoining μ_n to the

rational numbers $\mathbb{Q}$, we obtain the cyclotomic field $\mathbb{Q}(\mu_n)$, with ring of integers $\mathbb{Z}[\mu_n]$. It is a classical result from number theory that the ideal (p) in $\mathbb{Z}[\mu_n]$ factors into $g = \varphi(n)/k$ distinct maximal prime ideals $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_g$, where $\varphi(\cdot)$ is the Euler totient function. Furthermore, for each $\mathfrak{p}_j$, the residue class ring $\mathbb{Z}[\mu_n]/\mathfrak{p}_j \cong GF(q)$ (see for example [3, Theorem 10.45], [18, Chap. 10, §3B], [24, Theorem 2.13], [25, Theorem 7-2-4]). If we choose $\mathfrak{p}_j$ to be the ideal generated by p and $f(\mu_n)$, then $\mathbb{Z}[\mu_n]/\mathfrak{p}_j$ is exactly the version of $GF(q)$ that we started with. Note that since $\mathfrak{p}_\mathfrak{j}$ contains (p), it acts as reduction mod p on $\mathbb{Z}$. We therefore have a ring homomorphism from $\mathbb{Z}[\mu_n]$ to $GF(q)$ given by

$$\phi : \mathbb{Z}[\mu_n] \stackrel{\text{mod } p}{\rightarrow} \mathbb{Z}[\mu_n]/\mathfrak{p}_\mathfrak{j} \stackrel{\cong}{\rightarrow} GF(q) \ . \tag{11}$$

In this way we can lift vectors over $GF(q)$ to vectors over the alphabet $\bar{A}$ consisting of 0 and the $q-1$ powers of μ_n.

Example: Let $GF(8) = GF(2)[\alpha]$ where α is a root of $X^3 + X + 1$. Then $q = 8$, $n = 7$, $\mu_7 = e^{2\pi i/7}$. To lift $GF(8)$ to $\mathbb{C}$ we write $GF(8) = \{0, 1, \alpha, \alpha^2, \ldots, \alpha^6\}$, and lift 0 to 0 and α^j to μ_7^j for $j = 0, \ldots, 6$.

Let Π be an M_t-dimensional subspace of $GF(q)^m$. By lifting each element of a generator matrix we obtain an M_t-dimensional subspace $\bar{\Pi} \subseteq \mathbb{C}^m$, defined over an alphabet $\bar{\mathcal{A}}$ of size q.

Theorem 3.4. *If two subspaces Π_1, Π_2 of $GF(q)^m$ are nonintersecting, so are their lifts $\bar{\Pi}_1, \bar{\Pi}_2$.*

Proof: Let $P := \begin{bmatrix} \Pi_1 \\ \Pi_2 \end{bmatrix}$ and $\bar{P} := \begin{bmatrix} \bar{\Pi}_1 \\ \bar{\Pi}_2 \end{bmatrix}$. By Lemma 2.1, P has a $2M_t \times 2M_t$ invertible submatrix. Since ϕ is a ring homomorphism, the lift of this submatrix is also invertible. ■

It follows that the subspaces constructed in Corollary 3.3 are also nonintersecting when lifted to the complex field.

This construction gives full diversity order non-coherent space-time codes when the elements of the codewords are restricted to belong to a finite field. Their rate is

$$R = \frac{1}{m}\log(q^m - 1) - \frac{1}{m}\log(q^{M_t} - 1) < \log(q) \ ,$$

which according to Theorem 3.1 is the maximal achievable rate for diversity order $M_t M_r$. Moreover, the above relationship implies that for fully diverse codes constructed from a finite field, we cannot achieve a rate higher than $\log(|\mathcal{A}|)$.

4. PSK constellations

Throughout this section we assume that the alphabet $\mathcal{A}$ consists of the set of complex 2^r-th roots of unity, that is, $\mathcal{A} = \{e^{2\pi i j/2^r}, 0 \le j < 2^r\}$, for some $r \ge 1$. Let $\mu = e^{2\pi i/2^r}$ be a primitive 2^r-th root

of unity; $\mathcal{A}$ is a cyclic multiplicative group with generator μ. In this section we assume that $M_t = 2$, that is, the code consists of a set of pairwise nonintersecting planes.

Example 4.1. Some examples of roots of unity:

1. If $r = 1$, $\mu = -1$ and the alphabet is $\mathcal{A} = \{1, -1\}$.
2. If $r = 2$, $\mu = i$ and the alphabet is $\mathcal{A} = \{1, i, -1, -i\}$.
3. If $r = 3$, $\mu = (1 + i)/\sqrt{2}$ and the alphabet is $\mathcal{A} = \{e^{\pi ij/4}, 0 \leq j \leq 7\}$. This is the 8-PSK constellation.

There is a trivial upper bound.

Theorem 4.1. *Let $\mathcal{A}$ be the set of 2^r roots of unity, $r \geq 1$. Then the number of pairwise nonintersecting planes is at most $\frac{1}{2}|\mathcal{A}|^{m-1} = 2^{(m-1)r-1}$.*

Proof: If $v_1, v_2 \in \mathcal{A}^m$ are the generators for a plane, that plane also contains all multiples $\mu^j v_1$ and $\mu^j v_2$, a total of $2|\mathcal{A}|$ vectors. Since these sets of vectors must all be disjoint, the number of planes is at most $|\mathcal{A}|^m/(2|\mathcal{A}|)$. ■

The same argument shows that there are at most $\frac{1}{M_t}|\mathcal{A}|^{m-1}$ nonintersecting M_t-dimensional subspaces of complex m-dimensional space for any finite alphabet A. The implication of this in terms of rate is that

$$R \leq \frac{m-1}{m}\log(|\mathcal{A}|) - \frac{1}{m}\log(M_t) < \log(|\mathcal{A}|)\ .$$

Hence, for fully diverse codes constructed from PSK constellations, we cannot achieve a rate exceeding $\log(|\mathcal{A}|)$.

Example 4.2. Let $\mathcal{A}$ be the set $\{1, i, -1, -i\}$ and take $m = 4$. The total number of vectors in $\mathcal{A}^4$ is 4^4. Each vector has 4 multiples, so each plane accounts for at least 8 vectors. Therefore there are at most $\frac{4^4}{8} = 32$ planes.

In the other direction we will prove:

Theorem 4.2. *Assume $r \geq 1$ and that $m \geq 2$ is even. There exist $N = |\mathcal{A}|^{m-2} = 2^{(m-2)r}$ pairwise nonintersecting planes in $\mathbb{C}^m$ defined using the complex 2^r-th roots of unity.*

Note that the upper and lower bounds coincide in the case $r = 1$, that is, when $\mathcal{A} = \{1, -1\}$.

The proof is simplified by the use of valuations (cf. [8]). If $x \in \mathbb{Q}$, $x = 2^a \frac{b}{c}$ with $a, b, c \in \mathbb{Z}, c \neq 0$, b and c odd, then the 2-adic valuation of x is $\nu_2(x) = a$. Similarly, suppose x belongs to the cyclotomic field $\mathbb{Q}(\mu)$. Since $1 - \mu$ is a prime in $\mathbb{Z}[\mu]$, we can write x uniquely as $(1 - \mu)^a \frac{b}{c}$ with $a \in \mathbb{Z}$, $b, c \in \mathbb{Z}[\mu], c \neq 0$, b and c relatively prime to $1-\mu$. The $(1-\mu)$-adic valuation of x is then $\nu_{1-\mu}(x) = a$. It is easy to check that for $k \in \mathbb{Z}$, $k \neq 0$, $\nu_{1-\mu}(1 - \mu^k) = 2^{\nu_2(k)}$. In particular, if $k \in \mathbb{Z}$ is odd, $\nu_{1-\mu}(1 - \mu^k) = 1$.

We will also need a lemma:

Lemma 4.3. *Let Π be a plane in $\mathbb{C}^m$ generated by vectors v_1, v_2, and denote by*

$$\tilde{\Pi}_1 = \begin{bmatrix} v_1 & x_{11} & x_{12} \\ v_2 & x_{21} & x_{22} \end{bmatrix}$$

and

$$\tilde{\Pi}_2 = \begin{bmatrix} v_1 & y_{11} & y_{12} \\ v_2 & y_{21} & y_{22} \end{bmatrix}$$

two different embeddings of Π into $\mathbb{C}^{m+2}$. Then $\tilde{\Pi}_1 \cap \tilde{\Pi}_2 = \{0\}$ if and only if

$$\begin{vmatrix} y_{11} - x_{11} & y_{12} - x_{12} \\ y_{21} - x_{21} & y_{22} - x_{22} \end{vmatrix} \neq 0.$$

Proof: By Lemma 2.1, it is necessary and sufficient that the matrix $P := \begin{bmatrix} \tilde{\Pi}_1 \\ \tilde{\Pi}_2 \end{bmatrix}$ have rank 4. Subtracting the first and second rows of P from the third and fourth rows, we get the matrix

$$\begin{bmatrix} v_1 & x_{11} & x_{12} \\ v_2 & x_{21} & x_{22} \\ 0 & y_{11} - x_{11} & y_{12} - x_{12} \\ 0 & y_{21} - x_{21} & y_{22} - x_{22} \end{bmatrix}.$$

and the result follows. ■

We now give the proof of the theorem, for which we use induction on even values of m. For $m = 2$ we take the single plane

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Suppose the result is true for m. For each of the $|\mathcal{A}|^{m-2}$ pairwise nonintersecting planes in $\mathbb{C}^m$ we will construct $|\mathcal{A}|^2$ planes in $\mathbb{C}^{m+2}$, such that full set of planes so obtained is pairwise nonintersecting; this will establish the desired result.

If two planes are nonintersecting in $\mathbb{C}^m$ then they are certainly nonintersecting when embedded in any way in $\mathbb{C}^{m+2}$. So we need only show that the $|\mathcal{A}|^2$ embeddings of any single plane are pairwise nonintersecting.

Let Π be a plane in $\mathbb{C}^m$ generated by vectors v_1, v_2, and denote by $\tilde{\Pi}(a,b)$ the plane in $\mathbb{C}^{m+2}$ with generator matrix

$$\begin{bmatrix} v_1 & \mu^a & \mu^b \\ v_2 & \mu^{a+b} & \mu^{a+2b+1} \end{bmatrix},$$

for $a, b = 0, 1, \ldots, 2^r - 1$.

We will use Lemma 4.3 to show that all the planes $\{\tilde{\Pi}(a,b) \mid a \in \mathcal{A}, b \in \mathcal{A}\}$ are pairwise nonintersecting. For this we must show that

$$\begin{vmatrix} \mu^c - \mu^a & \mu^d - \mu^b \\ \mu^{c+d} - \mu^{a+b} & \mu^{c+2d+1} - \mu^{a+2b+1} \end{vmatrix} = 0$$

if and only if $a = c$ and $b = d$.

The above determinant is equal to

$$\mu^{2c+2d+1}(1-\mu^{a-c})(1-\mu^{(a-c)+2(b-d)}) - \mu^{c+2d}(1-\mu^{b-d})(1-\mu^{(a-c)+(b-d)}) . \tag{12}$$

If the determinant is zero, the $(1-\mu)$-adic valuations of the two terms on the right must be equal, that is,

$$2^{\nu_2(a-c)} + 2^{\nu_2(a-c+2(b-d))} = 2^{\nu_2(b-d)} + 2^{\nu_2(a-c+b-d)} . \tag{13}$$

We must show that this is true if and only if $a = c$ and $b = d$. We consider four cases, depending on the parity of $a-c$ and $b-d$. If $a-c \equiv 1, b-d \equiv 1 \pmod 2$ then (12) reads $1+1 = 1+2^{\nu_2(a-c+b-d)} \geq 3$ (since $a - c + b - d$ is even), a contradiction. Similarly, if $a - c \equiv 1, b - d \equiv 0 \pmod 2$ we get $1 + 1 = 2^{\nu_2(b-d)} + 1$, and if $a - c \equiv 0, b - d \equiv 1 \pmod 2$ we get $2^{\nu_2(a-c)} + 2^{\nu_2(a-c+2(b-d))} = 1 + 1$, which are also contradictions. The fourth possibility is $a - c \equiv b - d \equiv 0 \pmod 2$. Let $a - c = 2^s x$ and $b - d = 2^t y$, where x and y are odd, $s, t \geq 1$. We have

$$\nu_2(a - c + 2(b - d)) = \begin{cases} s & \text{if } s < t \\ s & \text{if } s = t \\ \geq t & \text{if } s > t \end{cases}$$

and

$$\nu_2(a - c + 2(b - d)) = \begin{cases} s & \text{if } s < t \\ \geq s & \text{if } s = t \\ t & \text{if } s > t \end{cases}$$

Substituting these valuations in (13) again gives a contradiction. This concludes the proof of Theorem 4.2.

5. Discussion

The following table compares the codes constructed in Sections 3 and 4 in the case $M_t = 2$, *i.e.*, codes which are pairwise nonintersecting 2-dimensional subspaces of $\mathbb{C}^m$, for $m = 4, 6$ and 8, and alphabets $\mathcal{A}$ of sizes 2, 4 and 8. The top entry in each cell gives the number of planes obtained from the finite field construction (Corollary 3.3). The bottom entry gives the lower and upper bounds obtained using complex $|\mathcal{A}|$-th roots of unity, from Theorem 4.2 and Theorem 4.1. Asymptotically, the rates of the two constructions are very similar. Both satisfy $\log(\text{ number of codewords })/m \approx \log(|\mathcal{A}|)$, for m large, and so both asymptotically achieve the maximal rate possible for fully diverse codes.

Note that the construction via finite fields results in codes for which alphabet consists of 0 and the complex $(|A| - 1)$-st roots of unity, whereas the construction via PSK constellations produces codes in which the symbols are the complex $|A|$-th roots of unity (and 0 is not used).

	$m = 4$	$m = 6$	$m = 8$
$\|\mathcal{A}\| = 2$	5 $4 - 4$	21 $16 - 16$	85 $64 - 64$
$\|\mathcal{A}\| = 4$	17 $16 - 32$	273 $256 - 512$	4369 $4096 - 8192$
$\|\mathcal{A}\| = 8$	65 $64 - 256$	4161 $4096 - 16384$	266305 $262144 - 1048576$

Table I. Number of pairwise nonintersecting planes in $\mathbb{C}^m$ for various sizes of the alphabet $|\mathcal{A}|$ (see text for details).

We end by mentioning some topics for further research.

- We also used clique-finding algorithms to search for larger sets of planes than those given in Theorem 4.2, again taking $\mathcal{A}$ to be the set of 2^r-th complex roots of unity. These searches were unsuccessful, and so we have not mentioned them elsewhere in the paper. These negative results lead us to conjecture, albeit weakly, that the lower bounds in Theorem 4.2 cannot be improved. It would be nice to have a better upper bound than that in Theorem 4.1 for the case $r > 1$. It would also be a worthwhile project to do a more extensive computer search for better codes, both for the above alphabet and for other alphabets.

  It is straightforward to formulate the search as a clique-finding problem. The first step is to prepare a list of candidate subspaces, making sure that the generator matrices use only symbols

from $\mathcal{A}$, and that the subspaces have the specified dimension and are distinct (a subspace may have many different generator matrices: only one version is placed on the list of candidates). Then a graph is constructed with the candidate subspaces as vertices, and with an edge joining two vertices if and only if the subspaces are nonintersecting. Then a good code is a maximal clique in this graph.

- Can the construction in Theorem 4.2 be generalized to the case when M_t is larger than 2? In particular, it would be interesting to do a computer search in the case $M_t = 3$ and $m = 6$.

- This paper has focused only on the existence and construction of finite alphabet codes which achieve maximal diversity order, and we did not consider decoding complexity. The decoding problem involves weighted projections of the received matrix $\mathbf{Y}$ onto the candidate subspaces (see (4)). In general this may require a search over 2^{mR} codewords, where R is the rate of the code. Since this number grows exponentially with the code length, a natural question to ask is whether there are codes which are optimally decodable in polynomial time, or have polynomial time sub-optimal decoders which perform satisfactorily.

- In [4] (see also [2]) a large number of optimal or putatively optimal packings of subspaces in $\mathbb{C}^m$ were constructed using (8) as a measure of "distance" between subspaces. It would be worthwhile repeating these calculations using (7) instead.

References

[1] J. André, Uber nicht-Desarguessche Ebenen mit transitiver Translationsgruppe, *Math. Z.*, **60** (1954), 156–186.

[2] A. R. Calderbank, R. H. Hardin, E. M. Rains, P. W. Shor and N. J. A. Sloane, A group-theoretic framework for the construction of packings in Grassmannian spaces, *J. Algebraic Combinatorics*, **9** (1999), 129–140.

[3] H. Cohn, *A Classical Invitation to Algebraic Numbers and Class Fields*, Springer-Verlag, NY, 1978.

[4] J. H. Conway, R. H. Hardin and N. J. A. Sloane, Packing lines, planes, etc.: packings in Grassmannian space, *Experimental Math.*, **5** (1996), 139–159.

[5] J. Eisfeld and L. Storme, (Partial) t-spreads and minimal t-covers in finite projective spaces, in *Lecture notes from the Socrates Intensive Course on Finite Geometry and its Applications, Ghent, April 2000*, Published electronically at http://www.maths.qmul.ac.uk/∼leonard/partialspreads/eisfeldstorme.ps.

[6] G. J. Foschini, Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas, *Bell Labs Technical Journal*, **1** (No. 2, 1996), 41–59.

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins Univ. Press, 2nd ed., 1989.

[8] F. Q. Gouvêa, *p-adic Numbers*, Springer-Verlag, NY, 1993.

[9] J. W. P. Hirschfeld, *Projective Geometries over Finite Fields*, Oxford Univ. Press, 1979.

[10] B. M. Hochwald and T. L. Marzetta, Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading, *IEEE Transactions on Information Theory*, **45** (No. 1, 1999), 139–157.

[11] B. M. Hochwald and T. L. Marzetta, Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading, *IEEE Transactions on Information Theory*, **46** (No. 2, 2000), 543–564.

[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, 1985.

[13] I. Kammoun and J.-C. Belfiore, A new family of Grassmann space-time codes for non-coherent MIMO systems, *IEEE Communications Letters*, **7** (No. 11, 2003), 528–530.

[14] R. Lidl and H. Niederreiter, *Finite Fields*, Addison-Wesley, Reading, MA, 1983.

[15] H.-F. Lu and P V. Kumar, Rate-diversity trade-off of space-time codes with fixed alphabet and optimal constructions for PSK modulation, *IEEE Transactions on Information Theory*, **49** (No. 10, 2003), 2747–2752.

[16] P. Lusina, E. M. Gabidulin and M. Bossert, Maximum rank distance codes as space-time codes, *IEEE Transactions on Information Theory*, **49** (No. 10, 2003), 2757–2760.

[17] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977, 10-th impression, 1998.

[18] P. Ribenboim, *Algebraic Numbers*, Wiley-Interscience, NY, 1972.

[19] A. Shokrollahi and B. Hassibi and B. M. Hochwald and W. Sweldens, Representation theory for high-rate multiple-antenna code design, *IEEE Transactions on Information Theory*, **47** (No. 6, 2001), 2335–2367.

[20] L. Soicher, *Computation of partial spreads*, Published electronically at http://www.maths.qmul.ac.uk/∼leonard/partialspreads/.

[21] V. Tarokh, N. Seshadri and A. R. Calderbank, Space-time codes for high data rate wireless communications: Performance criterion and code construction, *IEEE Transactions on Information Theory*, **44** (No. 2, 1998) 744–765.

[22] E. Telatar, Capacity of multi-antenna Gaussian channels, *European Transactions on Telecommunications*, **10** (No. 6, 1999), 585–596.

[23] D N C. Tse and P. Viswanath, *Principles of Wireless Communications*, Cambridge University Press, 2005.

[24] L. C. Washington, *Introduction to Cyclotomic Fields*, Springer-Verlag, NY, 1982.

[25] E. Weiss, *Algebraic Number Theory*, McGraw-Hill, NY, 1963.

[26] L. Zheng and D. N. C. Tse, Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel, *IEEE Transactions on Information Theory*, **48** (No. 2, 2002), 359–383.