arXiv:math/0401045v1 [math.CO] 6 Jan 2004

Unitary Space Time Constellation Analysis: An Upper Bound for the Diversity *

Guangyue Han, Joachim Rosenthal
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556.
e-mail: Han.13@nd.edu, Rosenthal.1@nd.edu
URL: http://www.nd.edu/~eecoding/

January 6, 2004

Abstract

The diversity product and the diversity sum are two very important parameters for a good-performing unitary space time constellation. A basic question is what the maximal diversity product (or sum) is. In this paper we are going to derive general upper bounds on the diversity sum and the diversity product for unitary constellations of any dimension n and any size m using packing techniques on the compact Lie group $U(n)$.

1 Introduction

Let A be a matrix with complex entries. A^* denotes the conjugate transpose of A. Let $\| \ \|$ denote the Frobenius norm of a matrix, i.e.,

$$\|A\| = \sqrt{\mathrm{tr}\,(AA^*)}.$$

A square matrix A is called unitary if $A^*A = AA^* = I$, where I denotes the identity matrix. We denote by $U(n)$ the set of all $n \times n$ unitary matrices. $U(n)$ is a real algebraic variety and a smooth manifold of real dimension n^2. For the purpose of this paper a unitary space time constellation (or code) $\mathcal{V}$ is simply a finite subset of $U(n)$,

$$\mathcal{V} = \{A_1, A_2, \cdots, A_m\} \subset U(n).$$

We say $\mathcal{V}$ has dimension n and size m. Unitary space time codes have been intensely studied in recent years and we refer the interested readers to [1, 9, 10, 13] and the references of these papers. The readers will find the motivation and engineering applications of such kind of codes. The quality of a unitary space time code is governed by two important parameters, the diversity product and the diversity sum.

*Both authors were supported in part by NSF grants DMS-00-72383 and CCR-02-05310. The first author was also supported by a fellowship from the Center of Applied Mathematics at the University of Notre Dame.

Definition 1.1. The *diversity product* [9] of a unitary space time code $\mathcal{V}$ is defined through

$$\prod \mathcal{V} := \frac{1}{2}\min\{|\det(A-B)|^{\frac{1}{n}} | A, B \in \mathcal{V}, A \neq B\}.$$

The *diversity sum* [11] is defined as

$$\sum \mathcal{V} := \frac{1}{2\sqrt{n}}\min\{\|A-B\| | A, B \in \mathcal{V}, A \neq B\}.$$

$\mathcal{V}$ is called fully diverse if $\prod \mathcal{V} > 0$. As explained in [7], a space time code with large diversity sum tends to perform well at low signal to noise ratios whereas a code with a large diversity product tends to perform well at high signal to noise ratios. A major coding design problem is the construction of unitary space time codes where the diversity sum (or product) is optimal or near optimal inside the set of all the space time codes with the same parameters n, m. We would like to remark that for every positive integer n and m, a Haar distributed random space time code is fully diverse with probability 1. A simple proof can be found in [7].

The purpose of this paper is to derive for n and m tight upper bounds for the diversity product $\prod \mathcal{V}$ and the diversity sum $\sum \mathcal{V}$. When $n = 1$ then trivially $|\det(A-B)| = \|A-B\|$ and it follows that $\sum \mathcal{V} = \prod \mathcal{V}$ in this situation. The following lemma states that for every space time code $\mathcal{V}$, $\sum \mathcal{V}$ is an upper bound for $\prod \mathcal{V}$ and by having an upper bound for $\sum \mathcal{V}$ we immediately also have an upper bound for $\prod \mathcal{V}$. The readers can find the statements about the relationship between $\prod \mathcal{V}$ and $\sum \mathcal{V}$ in [11], for completeness we include a detailed proof.

Lemma 1.2. *For any unitary space time code* $\mathcal{V}$*,*

$$\prod \mathcal{V} \leq \sum \mathcal{V}.$$

Proof. Let C be an $n \times n$ complex matrix with singular value decomposition

$$C = U\mathrm{diag}\,(c_1, c_2, \cdots, c_n)V,$$

where U, V are unitary matrices and $c_j \geq 0$ for $j = 1, 2, \cdots, n$ are the singular values of C. First we are going to prove

$$\frac{1}{2}|\det(C)|^{\frac{1}{n}} \leq \frac{1}{2\sqrt{n}}\|C\|.$$

If $c_j = 0$ for some j, then the inequality is trivial. Hence we assume $c_j > 0$ for all j's. Because U, V are unitary matrices, it follows that

$$\frac{1}{2}|\det(C)|^{\frac{1}{n}} = \frac{1}{2}\left(\prod_{j=1}^{n} c_j\right)^{\frac{1}{n}}.$$

Similarly one verifies that

$$\frac{1}{2\sqrt{n}}\|C\| = \frac{1}{2\sqrt{n}}\sqrt{\sum_{j=1}^{n} c_j^2}.$$

Applying Cauchy-Schwarz inequality, we have

$$\left(\prod_{j=1}^{n} c_j\right)^{\frac{1}{n}} \leq \frac{\sum_{j=1}^{n} c_j}{n} \leq \frac{1}{\sqrt{n}}\sqrt{\sum_{j=1}^{n} c_j^2}.$$

Hence one concludes that for an $n \times n$ square matrix C,

$$\frac{1}{2}|\det(C)|^{\frac{1}{n}} \leq \frac{1}{2\sqrt{n}}\|C\|.$$

By the definition of $\prod \mathcal{V}$, $\sum \mathcal{V}$ and the above inequality one gets

$$\prod \mathcal{V} \leq \sum \mathcal{V}.$$

$\square$

Of course it would be desirable to know for every n and m what the largest possible value of $\sum \mathcal{V}$ is. This is the motivation of the following definition.

Definition 1.3. Let $\Delta(n,m)$ be the infimum of all numbers such that for every unitary space time code $\mathcal{V}$ of dimension n and size m, one has

$$\sum \mathcal{V} \leq \Delta(n,m).$$

Remark 1.4. As pointed out by Liang and Xia [11] there exists a constellation $\mathcal{V}$ of dimension n and size m with $\sum \mathcal{V} = \Delta(n,m)$. This is due to the fact that $U(n)^m$ is a compact manifold.

The exact values of $\Delta(n,m)$ are only known in very few special cases. In the case $n = 1$, one checks that $\Delta(1,m) = \sin\frac{\pi}{m}$ for $m \geq 2$. When $n \geq 2$ and $m = 3$, one has $\Delta(n,3) = \frac{\sqrt{3}}{2}$. When $m = 2$, we have $\Delta(n,2) = 1$ for $n \geq 2$. For $n = 2$, the following values were computed in [11].

m	2	3	4	5	6	7	8	9	10 through 16
$\Delta(2,m)$	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{3}\sqrt{6}$	$\frac{1}{4}\sqrt{10}$	$\frac{1}{5}\sqrt{15}$	$\frac{1}{6}\sqrt{21}$	$\frac{1}{7}\sqrt{28}$	$\frac{1}{8}\sqrt{36}$	$\frac{1}{2}\sqrt{2}$

Liang and Xia [11] observed the connection between a unitary constellation and an Euclidean sphere code and beautifully derived an upper bound for 2 dimensional unitary constellations which is very tight when $m \leq 100$. In this paper we present a new general upper bound for $\Delta(n,m)$ for every dimension n and every size m while improving certain results in [11]. To the best of our knowledge the new upper bounds we derived are tighter than any previously published bounds as soon as m is sufficiently large.

2 Upper Bound Analysis

In this section we are going to study the packing problem on $U(n)$ and derive three upper bounds for the numbers $\Delta(n,m)$. All the resulted bounds are derived by differential geometric means and all bounds can be viewed as certain sphere packing bounds.

From a differential geometry point of view we can view $U(n)$ as a n^2-dimensional compact Lie group. $U(n)$ is also naturally a submanifold of the Euclidean space $\mathbb{R}^{2n^2}$. In this way $U(n)$ will have the induced geometry of the standard Euclidean geometry of $\mathbb{R}^{2n^2}$. Finally there is a third way to see $U(n)$ as a submanifold of another Riemannian manifold $S(n)$ and we will say more later.

The basic strategy for computing the upper bounds for $\Delta(n, m)$ is as follows. Given a unitary space time code $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, around each matrix A_j we can choose a neighborhood $N_r(A_j)$ with radius r (the radius will be specified later). Let $V_j = V(N_r(A_j))$ be the volume of the neighborhood $N_r(A_j)$. If all the neighborhoods are non-overlapping, then necessarily we will have

$$\sum_{j=1}^{m} V_j \leq V(U(n)),$$

where $V(U(n))$ denotes the total volume of unitary group $U(n)$. This inequality in turn will result in an upper bound for the numbers $\Delta(n, m)$. By employing different metrics (Euclidean or Riemannian) and by considering different embeddings of $U(n)$, we derive three different upper bounds for $\Delta(n, m)$.

Let $\mathcal{M}_1$ be the manifold consisting of all the $n \times n$ Hermitian matrices, i.e.

$$\mathcal{M}_1 = \{H | H = H^*\}.$$

$\mathcal{M}_1$ has dimension n^2 and can be viewed isometrically as Euclidean space $\mathbb{R}^{n^2}$. Assume that $H = (H_{jk})$ and assume that $H_{jk} = x_{jk} + iy_{jk}$. We use (dH) to denote the volume element of $\mathcal{M}_1$, where

$$(dH) = \left(\frac{i}{2}\right)^{n(n-1)/2} \bigwedge_{l=1}^{n} dH_{ll} \bigwedge_{j<k} dH_{jk} \bigwedge_{j<k} d\bar{H}_{jk} = \bigwedge_{l=1}^{n} dx_{ll} \bigwedge_{j<k} dx_{jk} \bigwedge_{j<k} dy_{jk}. \tag{2.1}$$

With a small abuse of the notation, one can check that the volume element of $\mathcal{M}_2$, the manifold consisting of all the $n \times n$ skew-Hermitian matrices, can be written as

$$(dH) = \left(\frac{i}{2}\right)^{n(n-1)/2} \left(\frac{1}{i}\right)^{n} \bigwedge_{l=1}^{n} dH_{ll} \bigwedge_{j<k} dH_{jk} \bigwedge_{j<k} d\bar{H}_{jk} = \bigwedge_{l=1}^{n} dy_{ll} \bigwedge_{j<k} dx_{jk} \bigwedge_{j<k} dy_{jk}. \tag{2.2}$$

For a unitary matrix U, if we differentiate $U^*U = I$, we will have

$$U^*dU + dU^*U = 0.$$

Therefore U^*dU is skew-Hermitian. The following lemma will characterize the volume element of $U(n)$. For the terminologies in this lemma, we refer to a standard differential geometry or integral geometry book, e.g. [8, 12].

Lemma 2.1. *The volume element of $U(n)$ induced by the Euclidean space $\mathbb{R}^{2n^2}$ is bi-invariant and the volume element can be written as (U^*dU) up to a scalar constant.*

Proof. The bi-invariance comes from the orthonormality of $U(n)$. (U^*dU) is left-invariant according to the definition. Indeed for a fixed yet arbitrary unitary matrix V,

$$(VU)^*d(VU) = U^*V^*VdU = U^*dU.$$

Since $U(n)$ is a compact Lie group and any compact Lie group is unimodular, (U^*dU) is also right-invariant. Because the bi-invariant n^2 differential forms are unique up to a scalar, one concludes that the volume element can be written as (U^*dU). ∎

The following theorem will represent the volume element of $U(n)$ in another way. One will see that it is closely related to the eigenvalues of unitary matrices.

Theorem 2.2. *For the Schur decomposition of a unitary matrix Θ:*

$$\Theta = U\mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})U^*, \tag{2.3}$$

we will have

$$(\Theta^* d\Theta) = \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 \wedge d\theta_2 \wedge \cdots \wedge d\theta_n \wedge (U^*dU - \mathrm{diag}\,(U^*dU)). \tag{2.4}$$

Proof. Let $D = \mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})$ and take the differential of Equation (2.3),

$$d\Theta = dUDU^* + UdDU^* + UDdU^*.$$

It follows that,

$$\Theta^* d\Theta = UD^*U^*dUDU^* + UD^*dDU^* + UdU^* = U(D^*U^*dUD + D^*dD)U^* + UdU^*.$$

Due to the right-invariance of the volume element in $U(n)$, it follows that

$$(\Theta^* d\Theta) = (U^*\Theta^* d\Theta U) = (D^*U^*dUD - U^*dU + i\mathrm{diag}\,(d\theta_1, d\theta_2, \cdots, d\theta_n)).$$

Note that $(D^*U^*dUD - U^*dU)_{jk} = (e^{i\theta_j} - e^{i\theta_k})U_{jk}$, therefore the diagonal elements of $D^*U^*dUD - U^*dU$ are all zeros and the off diagonal elements are scaled version of the ones of U^*dU. According to formula (2.2), the claim in the theorem follows. □

The following theorem calculates the volume of a small neighborhood with Euclidean distance r. Because of the homogeneity of $U(n)$, the center of this small "ball" is chosen to be I without loss of generality. For a unitary matrix U, we assume $e^{i\theta_j}$'s are its eigenvalues, i.e., $U \sim \mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})$. For a fixed unitary matrix A, let

$$U_r^E(n, A) = \{U \in U(n) | \|U - A\| \leq r\}.$$

Again because of the homogeneity of $U(n)$, $V(U_r^E(n, A))$ does not depend on the choice of A. In the sequel $V(U_r^E(n))$ will be used to denote $V(U_r^E(n, A))$ for any unitary matrix A. Let $S(n)$ denote a $2n^2 - 1$ dimensional sphere centered at the origin with radius $\sqrt{n}$, i.e.,

$$S(n) = \{(x_1, x_2, \cdots, x_{2n^2}) | x_1^2 + x_2^2 + \cdots + x_{2n^2}^2 = n\}.$$

Apparently $U(n)$ is a submanifold of $S(n)$. For a particular point $S_0 \in S(n)$, let

$$S_r(n, S_0) = \{S \in S(n) | \|S - S_0\| \leq r\}.$$

Theorem 2.3. *Let*

$$D_1 = \{(\theta_1, \theta_2, \cdots, \theta_n) | -\pi \leq \theta_j < \pi \text{ for } j = 1, 2, \cdots, n\} \tag{2.5}$$

and

$$D_2 = \left\{ (\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq \frac{r^2}{4} \right\}, \tag{2.6}$$

then

$$V(U_r^E(n)) = \frac{\iint_{D_1 \cap D_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} V(U(n)). \tag{2.7}$$

Proof. Note that $\|I-U\|_2 \leq r$ is equivalent to $\sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq \frac{r^2}{4}$. For a given unitary matrix Θ, the Schur decomposition $\Theta = U^* \mathrm{diag}\,(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})U$ is unique if θ_j's are strictly ordered. So if we take the integral of formula (2.4) over the integration region disregarding the order of θ_j's, we will obtain $n!$ times the volume of $V(U_r^E(n))$. Thus the volume of $U_r^E(n)$ will be

$$V(U_r^E(n)) = \frac{1}{n!} \iint_{D_1 \cap D_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \iint_{U(n)} (U^* dU - \mathrm{diag}\,(U^* dU)).$$

Using the same argument, we will derive the volume of $U(n)$:

$$V(U(n)) = \frac{1}{n!} \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \iint_{U(n)} (U^* dU - \mathrm{diag}\,(U^* dU)).$$

Compare the two derived volume formula, the claim in the theorem follows. □

Remark 2.4. By the Weyl denominator formula [5] one can replace

$$\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n$$

with $(2\pi)^n n!$. We keep it as it is to make the formula literally understandable.

There are several approaches to derive upper bounds for the diversity sum. The first approach considers $U(n)$ as a submanifold of $S(n)$, then chooses the non-overlapping neighborhoods to be small balls with radius r (with regard to the Euclidean distance). This will result in the first upper bound (B1) which we derive in this paper.

Theorem 2.5. *Let D_1 and D_2 be defined as in* (2.5) *and* (2.6)*. Assume $r_0^E = r_0^E(n, m)$ is the solution to the following equation (with variable r):*

$$m \iint_{D_1 \cap D_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n = \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n, \tag{2.8}$$

then

$$\Delta(n, m) \leq \sqrt{\frac{(r_0^E)^2}{n} - \frac{(r_0^E)^4}{4n^2}}. \tag{B1}$$

Proof. For a fixed yet arbitrary unitary constellation $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, consider m small non-overlapping neighborhoods $S_r(n, A_j)$ in $S(n)$. We can increase r such that there exist l, k such that $S_r(n, A_l)$ and $S_r(n, A_k)$ are tangent to each other. Apparently

$$U_r^E(n, A_j) = S_r(n, A_j) \cap U(n),$$

for any j. Since $S_r(n, A_j)$'s are non-overlapping, we conclude that $U_r^E(n, A_j)$'s are non-overlapping. Therefore we have

$$\sum_{j=1}^{m} V(U_r^E(n, A_j)) \leq V(U(n)),$$

that is

$$mV(U_r^E(n)) \leq V(U(n)).$$

One can check that $V(U_r^E(n))$ is an increasing function of r, so any r satisfying the above inequality will be less than the solution to the equality:

$$mV(U_r^E(n)) = V(U(n)),$$

which is essentially Equality (2.8). So we conclude that $r \leq r_0^E$.

Note that any two points $S_0, S_1 \in S(n)$ with two non-overlapping neighborhoods $S_r(n, S_0)$ and $S_r(n, S_1)$ will have distance $\|S_0 - S_1\| \geq 2\sqrt{r^2 - r^4/(4n)}$, where the equality holds only if $S_r(n, S_0)$ and $S_r(n, S_1)$ are tangent to each other. Apply the argument to A_j's and note that A_l and A_k are the closest pair of points with $\|A_l - A_k\| = 2\sqrt{r^2 - r^4/(4n)}$, we reach the conclusion of the theorem. $\square$

For a fixed $S_0 \in S(n)$, consider $S_r(n, S_0) \subset S(n)$. Let $\tau = \tau(n, r)$ denote the maximal number τ such that $S_r(n, S_1), S_r(n, S_2), \cdots, S_r(n, S_\tau)$ are non-overlapping and $S_r(n, S_j)$ is tangent to $S_r(n, S_0)$ for $j = 1, 2, \cdots, n$. One checks that $\tau(n, r)$ does not depend on the choice of S_0. In this sense $\tau(n, r)$ can be viewed as generalized kissing number [3] on an Euclidean sphere. For a fixed n dimensional unitary constellation $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, let $r(\mathcal{V})$ denote the maximal radius r such that $S_r(n, A_1), S_r(n, A_2), \cdots, S_r(n, A_m)$ are non-overlapping. Let $r_{opt} = r_{opt}(n, m)$ denote the maximal $r(\mathcal{V})$ over all possible n dimensional unitary constellation $\mathcal{V}$ with cardinality m. One checks $\Delta(n, m) = r_{opt}(n, m)/\sqrt{2n}$. The following theorem and corollary give a lower bound for the optimal diversity sum $\Delta(n, m)$.

Theorem 2.6. *Let D_1 be defined as in* (2.5) *and assume that $r_0^E = r_0^E(n, m)$ is the solution to the equation* (2.8). *Let*

$$\tilde{D}_2 = \left\{ (\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq \frac{(r_0^E)^2}{4} \right\}$$

and let

$$D_3 = \left\{ (\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \sin^2 \frac{\theta_j}{2} \leq r_{opt}(n, m)^2 - r_{opt}(n, m)^4/(4n) \right\}.$$

Then

$$\iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n$$
$$\leq (\tau(n, r_{opt}(n, m)) + 1) \iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n.$$

Proof. According to the derivation of r_0^E, we have

$$m \iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n = \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n. \tag{2.9}$$

Assume that $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$ is an n dimensional unitary constellation reaching $r_{opt}(n,m)$, i.e., $r(\mathcal{V}) = r_{opt}(n,m)$. For simplicity let $r = r(\mathcal{V})$. Let m' denote the maximal number such that $S_r(n, A_1), S_r(n, A_2), \cdots, S_r(n, A_m), \cdots, S_r(n, A_{m'})$ are non-overlapping. Let $r_1 = 2\sqrt{r^2 - r^2/(4n)}$, we claim that

$$U(n) \subset \bigcup_{j=1}^{m'} U_{r_1}^E(n, A_j).$$

Otherwise suppose there is a unitary matrix $A_0 \notin \bigcup_{j=1}^{m'} U_{r_1}^E(n, A_j)$, then $\|A_0 - A_j\| > r_1$ (see Theorem 2.5). Thus $S_r(n, A_0)$ does not intersect with $S_r(n, A_j)$ for $j = 1, 2, \cdots, m'$. Therefore one can find $m' + 1$ small balls with radius r which are non-overlapping. This contradicts the maximality of m'. Thus we have $\sum_{j=1}^{m'} V(U_{r_1}^E(n, A_j)) \geq V(U(n))$, that is

$$m' \iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \geq \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n. \quad (2.10)$$

We further claim that

$$m' \leq (m-1)(\tau(n,r)+1). \quad (2.11)$$

By contradiction assume that $m' \geq (m-1)(\tau(n,r)+1)+1$. Let

$$\mathrm{tang}\,(j) = \{l | 1 \leq l \leq m', S_r(n, A_l) \quad \text{tangent to} \quad S_r(n, A_j)\}.$$

According to the definition of $\tau(n,r)$, we know the cardinality of $\mathrm{tang}\,(j)$ is less than $\tau(n,r)$. We first pick j_1 from $\{0, 1, \cdots, m'\}$, then pick j_2 from $\{0, 1, \cdots, m'\} - \mathrm{tang}\,(j_1)$. And we continue this process by always picking j_{k+1} from

$$\{0, 1, \cdots, m'\} - \bigcup_{l=1}^{k} \mathrm{tang}\,(j_l).$$

Since the cardinality of the above set is strictly greater than 0 when $k \leq m-1$, we can pick $j_1, j_2, \cdots, j_m$ from the index set $\{1, 2, \cdots, m'\}$ such that $S_r(n, A_{j_1}), S_r(n, A_{j_2}), \cdots, S_r(n, A_{j_m})$ are non-overlapping and every two of them are not tangent to each other. Then we can find a small enough real number $\varepsilon > 0$ and increase the radius r to $r + \varepsilon$ such that

$$S_{r+\varepsilon}(n, A_{j_1}), S_{r+\varepsilon}(n, A_{j_2}), \cdots, S_{r+\varepsilon}(n, A_{j_m})$$

are still non-overlapping. However this contradicts the maximality of $r = r_{opt}(n,m)$.

The combination of the three formulas (2.9), (2.10), (2.11) will lead to

$$\frac{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} \leq \left(\frac{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} - 1 \right) (\tau(n,r)+1). \quad (2.12)$$

Note that the inequality above is in fact stronger than the claim in the theorem. We can reach the conclusion of the theorem by relaxing the right hand side of the inequality (by ignoring -1). $\square$

Corollary 2.7. *When $m \to \infty$, asymptotically we have*

$$\Delta(n,m) \geq 2\sqrt{n} r_0^E(n,m) \frac{1}{2} (\tau(2n^2-1)+1)^{-1/n^2}.$$

Proof. We only sketch the idea of the proof. Intuitively $U_r^E(n, A_0)$ looks more "flat" when $m \to \infty$ (consequently $r \to 0$), so $V(U_r^E(n, A_0))$ can be approximated by the volume of $U_r^E(n, A_0)$'s projection to the tangent space of $U(n)$ at A_0:

$$\iint_{D_1 \cap \tilde{D}_2} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \sim C(r_0^E)^{n^2}$$

for some constant C. The same argument will lead to

$$\iint_{D_1 \cap D_3} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n \sim C(2r_{opt})^{n^2}$$

for the same constant C. For any fixed n, $\tau(n, r)$ will approach to the standard kissing number in Euclidean space $\tau(2n^2-1)$ when r goes to zero. Combining the three approximations, we reach the claim according to the previous theorem.

□

$U(n)$ is a compact Lie group equipped with a Riemannian metric. Given two points $A_0, A_1 \in U(n)$, one can always find a geodesic $\gamma(t)$ (mapping from $[0,1]$ to $U(n)$) which will connect these two points, i.e. $\gamma(0) = A_0$ and $\gamma(1) = A_1$. Recall that the Euclidean distance of A_0 and A_1 is defined to be $\|A_0 - A_1\|$. We further define the Riemannian distance between A_0 and A_1 to be:

$$\mathrm{dist}\,(A_0, A_1) = \int_0^1 \|\gamma'(t)\| dt.$$

As a Lie group $U(n)$ is homogeneous. In particular one has that

$$\mathrm{dist}\,(A_0, A_1) = \mathrm{dist}\,(UA_0, UA_1) = \mathrm{dist}\,(A_0 U, A_1 U)$$

for any $U \in U(n)$. The following theorem utilizes the homogeneity and the relationship between the Riemannian distance and Euclidean distance to derive another upper bound for the diversity sum in general and it is the base of the second approach.

Theorem 2.8. *Let $f(\cdot)$ and $g(\cdot)$ be two fixed monotone increasing real functions. If*

$$g(\|A_0 - A_1\|) \leq \mathrm{dist}\,(A_0, A_1) \leq f(\|A_0 - A_1\|)$$

for any two unitary matrices A_0 and A_1, then

$$\Delta(n,m) \leq g^{-1}(2f(r_0^E(n,m)))/(2\sqrt{n}).$$

Proof. For a fixed unitary constellation $\mathcal{V} = \{A_1, A_2, \cdots, A_m\}$, consider

$$U_r^E(n, A_1), U_r^E(n, A_2), \cdots, U_r^E(n, A_m)$$

for $r > 0$. We can increase r until there exist j and k such that $U_r^E(n, A_j)$ and $U_r^E(n, A_k)$ are tangent to each other at a point A_0. As examined in Theorem 2.5, one can make a conclusion that $r \le r_0^E(n, m)$. Accordingly we have

$$\begin{aligned} \operatorname{dist}(A_j, A_k) &\le \operatorname{dist}(A_j, A_0) + \operatorname{dist}(A_k, A_0) \\ &\le f(\|A_j - A_0\|) + f(\|A_k - A_0\|) = 2f(r) \le 2f(r_0^E(n, m)). \end{aligned}$$

On the other hand since g is monotonically increasing one has:

$$\|A_j - A_k\| \le g^{-1}(\operatorname{dist}(A_j, A_k)).$$

The combination of the above two inequalities will lead to

$$\|A_j - A_k\| \le g^{-1}(2f(r_0^E(n, m))).$$

Immediately we will have

$$\sum \mathcal{V} \le g^{-1}(2f(r_0^E(n, m)))/(2\sqrt{n}).$$

Since $\mathcal{V}$ is an arbitrary unitary constellation, the claim in the theorem follows. $\square$

Based on the above theorem, the following corollary gives the second upper bound (B2).

Corollary 2.9. *For a real number r, let $\lfloor r \rfloor$ denote the greatest integer less than or equal to r, then*

$$\Delta(n, m) \le \sin\sqrt{\frac{\pi^2}{n}\left\lfloor\frac{(r_0^E)^2(n, m)}{4}\right\rfloor + \frac{4}{n}\arcsin^2\sqrt{\frac{(r_0^E)^2(n, m)}{4} - \left\lfloor\frac{(r_0^E)^2(n, m)}{4}\right\rfloor}}. \tag{B2}$$

Proof. Consider I and another point $U = V\operatorname{diag}(e^{i\theta_1}, e^{i\theta_2}, \cdots, e^{i\theta_n})V^*$, where $-\pi \le \theta_j < \pi$. It is known that [4] the geodesic from I to U can be parameterized by

$$\gamma(t) = V\operatorname{diag}(e^{i\theta_1 t}, e^{i\theta_2 t}, \cdots, e^{i\theta_n t})V^*,$$

where $0 \le t \le 1$. The Riemannian distance from I to U is

$$\operatorname{dist}(I, U) = \sqrt{\theta_1^2 + \theta_2^2 + \cdots + \theta_n^2}.$$

We want to derive $g(\cdot), f(\cdot)$ as in Theorem 2.8. Suppose the Euclidean distance between I and U is r, i.e.,

$$\sin^2\frac{\theta_1}{2} + \sin^2\frac{\theta_2}{2} + \cdots + \sin^2\frac{\theta_n}{2} = r^2/4.$$

After substituting with $x_j = \sin^2\theta_j/2$ and denoting $G(x) = \arcsin^2\sqrt{x}$, we convert the above problem to the following optimization problem:

Find the minimum and maximum of the function

$$F(x_1, x_2, \cdots, x_n) = \theta_1^2 + \theta_2^2 + \cdots + \theta_n^2 = 4(G(x_1) + G(x_2) + \cdots + G(x_n))$$

with the constraints $x_1+x_2+\cdots+x_n = r^2/4$ and $0 \le x_j \le 1$ for $j = 1, 2, \cdots, n$. Since $G(x)$ is a convex function on $[0, 1]$, we derive the lower bound of $F(x_1, x_2, \cdots, x_n)$,

$$4n \arcsin^2(r/(2\sqrt{n})) \le F(x_1, x_2, \cdots, x_n). \tag{2.13}$$

In the sequel we are going to calculate the upper bound of $F(x_1, x_2, \cdots, x_n)$. Without loss of generality, we assume $0 \le x_1 \le x_2 \le \cdots \le x_n \le 1$. Let $k = \lfloor r^2/4 \rfloor$ and $\alpha = r^2/4 - k$, we claim that $F(x_1, x_2, \cdots, x_n)$ will reach its maximum when

$$x_j = \begin{cases} 0 & 1 \le j \le n-k-1 \\ \alpha & j = n-k \\ 1 & n-k+1 \le j \le n \end{cases}$$

Suppose by contradiction that F reaches its maximum at $(x_1, x_2, \cdots, x_n)$ with $x_1 > 0$. Now from

$$x_1 + x_{n-k} + x_{n-k+1} + \cdots + x_n \le r^2/4 = k + \alpha,$$

surely one can find $x'_{n-k}, x'_{n-k+1}, \cdots, x'_n$ such that

$$x_1 + x_{n-k} + x_{n-k+1} + \cdots + x_n = x'_{n-k} + x'_{n-k+1} + \cdots + x'_n,$$

with $x'_j \ge x_j$ for $j = n-k, n-k+1, \cdots, n$. Now set $x_1^* = 0$, $x_j^* = x_j$ for $j = 2, 3, \cdots, n-k-1$ and $x_j^* = x'_j$ for $j = n-k, n-k+1, \cdots, n$. By the mean value theorem, there exist ζ_j's with $x_1^* = 0 \le \zeta_1 \le x_1$ and $x_j \le \zeta_j \le x_j^*$ for $j = 2, 3, \cdots, n$ such that

$$F(x_1^*, x_2^*, \cdots, x_n^*) - F(x_1, x_2, \cdots, x_n) = \sum_{j=1}^{n} G'(\zeta_j)(x_j^* - x_j).$$

Since $G(x)$ is a strictly convex function, we have

$$0 < G'(\zeta_1) < G'(\zeta_2) < \cdots < G'(\zeta_n).$$

Now

$$\begin{aligned} F(x_1^*, x_2^*, \cdots, x_n^*) - F(x_1, x_2, \cdots, x_n) &\ge G'(\zeta_2)(\sum_{j=2}^{n}(x_j^* - x_j)) - G'(\zeta_1)(x_1 - x_1^*) \\ &= (G'(\zeta_2) - G'(\zeta_1))(x_1 - x_1^*) = (G'(\zeta_2) - G'(\zeta_1))x_1 > 0. \end{aligned}$$

This contradicts the maximality of F at $(x_1, x_2, \cdots, x_n)$. Applying exactly the same analysis to $x_2, x_3, \cdots, x_{n-k-1}, x_{n-k}$ we deduce that $x_j = 0$ for $j = 2, 3, \cdots, n-k-1$ and $x_{n-k} = \alpha$. So the upper bound of F can be given as

$$F(x_1, x_2, \cdots, x_n) \le 4\left(k\frac{\pi^2}{4} + \arcsin^2(\sqrt{\alpha})\right).$$

Take $g(r) = 2\sqrt{n}\arcsin(r/(2\sqrt{n}))$ and $f(r) = 2\sqrt{k\pi^2/4 + \arcsin^2\sqrt{\alpha}}$, the corollary follows according to the previous theorem. $\square$

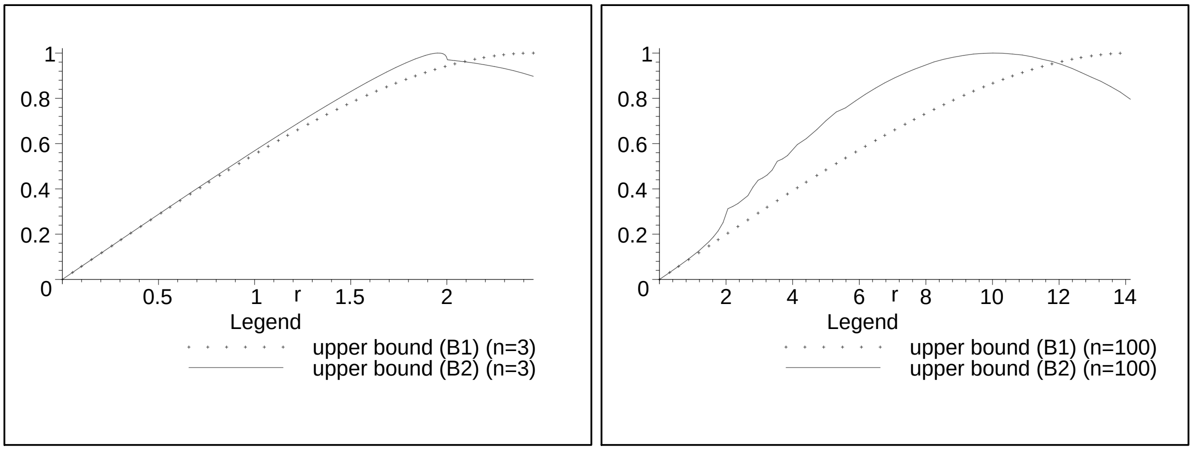

Figure 1: The comparisons of two upper bounds as functions for $n = 3$ and $n = 100$

Note that both upper bound (B1) and upper bound (B2) depend on $r_0^E(n, m)$. In Figure 1 we plot both upper bounds as functions of $r_0^E(n, m)$ for 3 and 100 dimensions. One can see that if and only if $r_0^E(3, m) > 2.0881$, the upper bound (B2) is tighter than the upper bound (B1). While for the 100 dimension case, the upper bound (B1) is tighter than the upper bound (B2) if and only if $r_0^E(100, m) > 11.9155$. In fact it can be checked that asymptotically when n is large enough, upper bound (B2) is tighter than upper bound (B1) if and only if $r_0^E(n, m) > 1.1892\sqrt{n}$.

For a packing problem on a manifold, alternatively one can choose the neighborhood to be a small "ball" with Riemannian radius r. This will be our third approach to derive an upper bound for the diversity sum. For a particular $A \in U(n)$, let

$$U_r^R(n, A) = \{U \in U(n) | \mathrm{dist}\,(U, A) \leq r\}.$$

Note that the constraint $\mathrm{dist}\,(U, I) \leq r$ is equivalent to

$$\theta_1^2 + \theta_2^2 + \cdots + \theta_n^2 \leq r^2.$$

Therefore we apply the same argument as in the proof of Theorem 2.3 and conclude that:

$$V(U_r^R(n)) = \frac{\iint_{D_1 \cap D_4} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n}{\iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n} V(U(n)),$$

where D_1 was defined in (2.5) and

$$D_4 := \{(\theta_1, \theta_2, \cdots, \theta_n) | \sum_{j=1}^{n} \theta_j^2 \leq r^2\}. \tag{2.14}$$

Instead of considering the Euclidean neighborhoods $U_r^E(n, A_1), U_r^E(n, A_2), \cdots, U_r^E(n, A_m)$, we can consider the Riemannian neighborhood $U_r^R(n, A_1), U_r^R(n, A_2), \cdots, U_r^R(n, A_m)$. Utilizing the fact that the Euclidean distance $\|A_j - A_k\|$ and the Riemannian distance $\mathrm{dist}\,(A_j, A_k)$ are related (compare with Formula (2.13)):

$$4n \arcsin^2(\|A_j - A_k\|/(2\sqrt{n})) \leq \mathrm{dist}\,(A_j, A_k)$$

for any two unitary matrices A_j and A_k, we can derive the third upper bound (B3). The proof of the following theorem is very similar to the one of Theorem 2.8 and for the sake of brevity we omit it.

Theorem 2.10. *Let D_1 and D_4 be defined as in* (2.5) *and* (2.14) *and assume $r_0^R(n,m)$ is the solution to the following equation (with variable r):*

$$m \iint_{D_1 \cap D_4} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n = \iint_{D_1} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 d\theta_1 d\theta_2 \cdots d\theta_n, \tag{2.15}$$

then

$$\Delta(n,m) \le \sin\left(\frac{r_0^R(n,m)}{\sqrt{n}}\right). \tag{B3}$$

We gave three approaches to derive upper bounds for the diversity sum and hence also for the diversity product. All of them involve the calculation of $r_0^E(n,m)$ or $r_0^R(n,m)$, which are the solutions of equation (2.8) and equation (2.15), respectively. Fortunately we are dealing with finding a root of a monotone increasing function (recall that both $V(U_r^E(n,m))$ and $V(U_r^R(n,m))$ are monotone increasing functions with respect to r), the bisection method [2] will be highly effective to solve this kind of problem. Our numerical experiments for small size constellations with small dimensions show that upper bound (B3) is looser than the first two upper bounds. However when m goes to infinity, these three upper bounds give almost the same estimation. This makes sense because asymptotically the small balls look like a n^2 dimensional ball in Euclidean space. One can see the derived upper bounds for 2 and 3 dimensional constellations in Figure 2.

We compare the derived upper bounds with the currently existing one presented in [11]. For $n = 2$ the upper bounds derived by Liang and Xia [11] tend to be better when $m \le 100$ and our bounds become tighter when $m \ge 100$ (see the following Table 1). For $n \ge 3$ Liang and Xia [11] outlined a method by considering a sphere packing computation in $S(n)$. It is our belief that this method will result in a weaker bound than the upper bounds we derived in this paper. For the sample programs to do the upper bound calculation, we refer to [6].

Table 1. For $n = 2$ the following table compares the upper bounds in [11] with our new bounds (B1) and (B2).

m	24	48	64	80	100	120	128	1000
upper bounds in [11]	0.6746	0.6193	0.5969	0.5799	0.5632	0.5499	0.5452	
upper bound (B1)	0.7598	0.6603	0.6131	0.5932	0.5578	0.5425	0.5347	0.3270
upper bound (B2)	0.7794	0.6734	0.6235	0.6026	0.5654	0.5496	0.5415	0.3285

One interesting fact about the limiting behavior of $\Delta(n,m)$ (when $m \to \infty$) is its connection to the Kepler problem [3]. Certainly one can use Kepler density [3] to obtain a tighter bound of the diversity sum asymptotically.

3 Conclusions and Future Work

We presented three approaches to derive upper bounds for the diversity sum of unitary constellations of any dimension n and any size m. The derived bounds seem to improve the

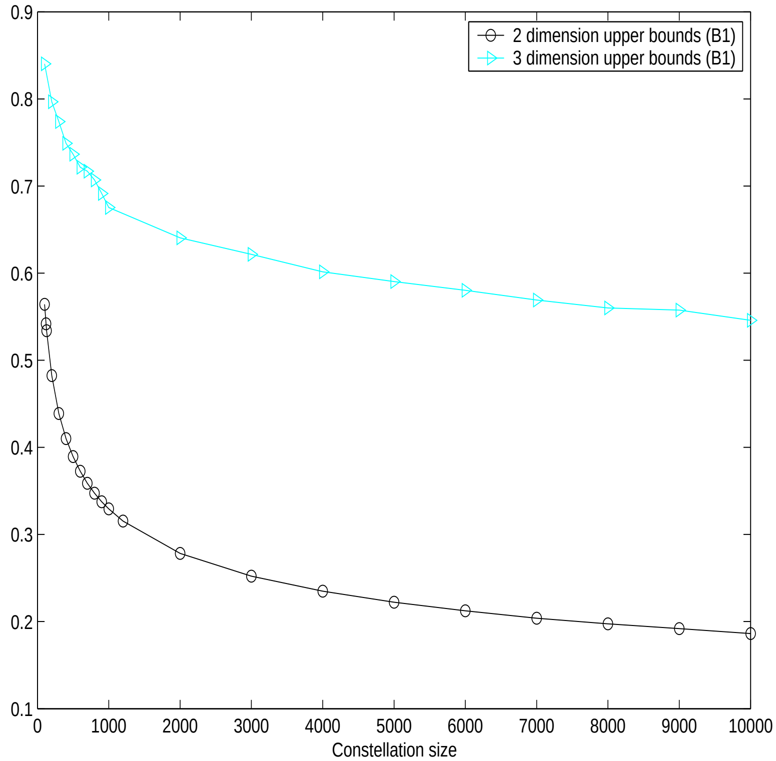

Figure 2: Upper bounds for 2 and 3 dimensional constellations

existing bounds when $n = 2$ and $m \geq 100$. When n is large the exact computation of r_0^E is rather involved and hence it is also computationally difficult to compute the bounds (B1) and (B2). Nonetheless it is our belief that the resulting upper bounds (B1) and (B2) become fairly tight as soon as m is sufficiently large.

It was pointed out that the resulted upper bounds also apply for the diversity product, although the bounds seem to be less tight in this situation. The future work may involve the derivation of a tighter upper bound analysis for the diversity product of unitary constellations using differential geometric means.

Acknowledgments

We started this research while spending a month at the Institute Mittag-Leffler in Stockholm in May 2003. The hospitality and the financial support of the Institute Mittag-Leffler are greatly acknowledged. We are also grateful for stimulating discussions we had on this subject with Professor Uwe Helmke and Professor Xuebin Liang.

References

[1] S. M. Alamouti. A simple transmitter diversity scheme for wireless communications. *IEEE J. Selected Areas of Commun.*, pages 1451–1458, October 1998.

[2] K. E. Atkinson. *An Introduction to Numerical Analysis*. J. Wiley, New York, 1978.

[3] J. H. Conway and N. J. A. Sloane. *Sphere packings, lattices and groups.* Springer-Verlag, New York, second edition, 1993. With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.

[4] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353 (electronic), 1999.

[5] R. Goodman and N. R. Wallach. *Representations and invariants of the classical groups.* Cambridge University Press, Cambridge, U.K. ; New York, NY, USA, 1998.

[6] G. Han and J. Rosenthal. A website of unitary space time constellations with large diversity. http://www.nd.edu/~eecoding/space-time/.

[7] G. Han and J. Rosenthal. Geometrical and numerical design of structured unitary space time constellations, December 2003. E-print math.OC/0312170.

[8] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces.* Graduate studies in mathematics. Providence, R.I. : American Mathematical Society, 2001.

[9] B. Hochwald and W. Sweldens. Differential unitary space-time modulation. *IEEE Trans. Comm.*, pages 2041–2052, December 2000.

[10] B. M. Hochwald and T. L. Marzetta. Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading. *IEEE Trans. Inform. Theory*, 46(2):543–564, 2000.

[11] X.-B. Liang and X.-G. Xia. Unitary signal constellations for differential space-time modulation with two transmit antennas: Parametric codes, optimal designs and bounds. *IEEE Trans. Inform. Theory*, 48(8):2291–2322, August 2002.

[12] L. A. Santaló. *Integral geometry and geometric probability*, volume 1 of *Encyclopedia of mathematics and its applications ; v. 1 Section, Probability.* Addison-Wesley Pub. Co., 1976.

[13] A. Shokrollahi, B. Hassibi, B. M. Hochwald, and W. Sweldens. Representation theory for high-rate multiple-antenna code design. *IEEE Trans. Inform. Theory*, 47(6):2335–2367, 2001.