Acknowledgment

The authors would like to thank Prof. Sandeep Pradhan, University of Michigan, Ann Arbor, for insightful discussions and feedback. We would also like to thanks the reviewers for their insightful comments and suggestions.

References

[1] *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information: Video (MPEG-2), 2nd Edition*, , 2000, ISO/IEC JTC 1/SC 2913818-2.
[2] *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information—Part 3: Audio (MP3), 2nd Edition*, , 1998, ISO/IEC JTC 1/SC 2913818-3.
[3] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 1–10, Jan. 1976.
[4] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Contr. Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
[5] T. M. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 1, pp. 2–14, Jan. 1972.
[6] T. J. Goblick, Jr, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 4, pp. 558–567, Oct. 1965.
[7] T. Berger and D. W. Tufts, "Optimum pulse amplitude modulation part I: Transmitter-receiver design and bounds from information theory," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 2, pp. 196–208, Apr. 1967.
[8] S. Shamai (Shitz), S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 564–579, Mar. 1998.
[9] U. Mittal and N. Phamdo, "Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1082–1102, May 2002.
[10] Z. Reznic, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3778–3788, Aug. 2006.
[11] V. M. Prabhakaran, R. Puri, and K. Ramchandran, "Hybrid analog-digital strategies for source-channel broadcast," in *Proc. 43rd Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Sep. 2005.
[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[13] J. Ziv, "On universal quantization," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 344–347, May 1985.
[14] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Trans. Inf. Theory,*, vol. 38, no. 2, pp. 428–436, Mar. 1992.
[15] W. H. R. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 269–275, Mar. 1991.
[16] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.
[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[18] [Online]. Available: http://www.mosek.com
[19] N. Merhav and S. Shamai (Shitz), "On joint source–channel coding for the Wyner–Ziv source and the Gel'fand–Pinsker channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 11, pp. 2844–2855, Nov. 2003.
[20] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.
[21] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder-ii: General sources," *Inf. Contr.*, vol. 38, pp. 60–80, Jul. 1978.

Optimal Normalized Diversity Product of 2 × 2 Lattice-Based Diagonal Space–Time Codes From QAM Signal Constellations

Haiquan Wang and Xiang-Gen Xia, *Senior Member, IEEE*

***Abstract*—In this correspondence, we prove that the optimal normalized diversity product of 2×2 lattice-based diagonal space–time block codes with Gaussian integer (or QAM) signal constellations, i.e., $\mathbb{Z}[\mathbf{i}]$, and any generating matrices of complex entries (not necessarily algebraic extensions of $\mathbb{Z}[\mathbf{i}]$ as commonly used) is $1/\sqrt{3}$. This result implies that 2×2 lattice-based diagonal space–time block codes with Gaussian integer signal constellations and generating matrices of entries from quadratic algebraic extensions of $\mathbb{Z}[\mathbf{i}]$ have already reached the optimal normalized diversity product.**

***Index Terms*—Algebraic extension, Gaussian integers, geometry of numbers, lattice-based space–time block codes, normalized diversity product.**

I. Introduction

Lattice-based diagonal space–time block codes (L-DSTBC) can be used in single-input–single-output (SISO) wireless systems to achieve signal space diversity, see for example [1]–[3] and can also be used in multiple-input–multiple-output (MIMO) wireless systems, see for example [4] and in the meantime they can be treated as the base of layered/threaded full-rate space–time codes [5]–[7], [10]. The basic idea of L-DSTBC is as follows. Let n be the number of transmit antennas in MIMO systems or the block size in SISO systems. Let $s_1, \ldots, s_n$ be n (complex) information symbols located on a two-dimensional real lattice $\mathcal{L}$ called *base lattice*, such as the Gaussian integer ring $\mathbb{Z}[\mathbf{i}]$ or the Eisenstein integer ring $\mathbb{Z}[\mathbf{j}]$. These n information symbols are linearly transformed into another n complex values $x_1, \ldots, x_n$: $(x_1, \ldots, x_n)^T = G(s_1, \ldots, s_n)^T$ with an $n \times n$ complex entry matrix G such that

$$\min_{(s_1,\ldots,s_n)\neq \mathbf{0}} |x_1 \cdots x_n| > 0. \tag{1}$$

Then, these n transformed complex values are put onto the diagonal to form an $n \times n$ L-DSTBC and the property (1) ensures that this L-DSTBC has full diversity. To construct a full diversity L-DSTBC as above depends on how to choose a base lattice $\mathcal{L}$ where information symbols are located and how to choose a linear transformation matrix G. Due to the existing algebraic number theory, there are many ways to construct such G and $\mathcal{L}$ and they are based on algebraic extension approach, see for example [1]–[4], [8], [9]. Since there are many designs of G and $\mathcal{L}$ for full diversity L-DSTBC, the question then is which one is optimal in the sense that which one has the smallest mean transmission signal power when its diversity product (or determinant distance) is fixed. By using the packing theory, the following normalized diversity product has been introduced in [9] to design an L-DSTBC:

$$\xi(G, \mathcal{L}) \triangleq \frac{\min_{(s_1,\ldots,s_n)\neq \mathbf{0}} |x_1 \cdots x_n|}{|\det(G)| \cdot |\mathcal{L}|^{n/2}} \tag{2}$$

Manuscript received December 7, 2006; revised January 22, 2007. Their work was supported in part by the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-05-1-0161 and the National Science Foundation under Grant CCR-0325180.

The authors are with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: hwang@ee.udel.edu; xxia@ee.udel.edu).

Communicated by S. McLaughlin, Associate Editor for Coding Techniques.
Digital Object Identifier 10.1109/TIT.2008.917688

where $|\mathcal{L}|$ denotes the absolute value of the determinant of the 2×2 generating matrix of the two-dimensional real base lattice $\mathcal{L}$. The rule to design an L-DSTBC is to design G and $\mathcal{L}$ such that the above normalized diversity product $\xi(G, \mathcal{L})$ is as large as possible. When G and $\mathcal{L}$ are over cyclotomic fields/rings, optimal L-DSTBC have been obtained in [9], [10].

In [11], it is shown that, when the base lattice $\mathcal{L} = \mathbb{Z}[\mathbf{i}]$, the Gaussian integer ring (or QAM, i.e., square lattice), and the entries of the 2×2 generating matrix G are the roots of a quadratic polynomial over $\mathbb{Z}[\mathbf{i}]$, the largest normalized diversity product $\xi(G_2, \mathbb{Z}[\mathbf{i}])$ of 2×2 L-DSTBC is $1/\sqrt{3}$, and furthermore if the entries of the 2×2 generating matrix G are arbitrary complex numbers, then the largest normalized diversity product $\xi(G, \mathbb{Z}[\mathbf{i}])$ of 2×2 L-DSTBC is upper bounded by $1/\sqrt{2}$ and this upper bound is not reachable.

In [11], it is also shown that, when the base lattice $\mathcal{L} = \mathbb{Z}[\mathbf{j}]$, the Eisenstein integer ring (or equal-literal triangular lattice), and the entries of the 2×2 generating matrix G are the roots of a quadratic polynomial over $\mathbb{Z}[\mathbf{j}]$, the largest normalized diversity product $\xi(G_2, \mathbb{Z}[\mathbf{j}])$ is $2/(13^{1/4}\sqrt{3})$.

In this correspondence, we show that the largest normalized diversity product $\xi(G, \mathbb{Z}[\mathbf{i}])$ of 2×2 L-DSTBC when the entries of 2×2 generating matrix G are general complex numbers (not necessarily algebraic extensions) is also upper bounded by $1/\sqrt{3}$. This implies that 2×2 L-DSTBC, when base lattice $\mathcal{L} = \mathbb{Z}[\mathbf{i}]$ and entries of generating matrix G are quadratic algebraic extensions of $\mathbb{Z}[\mathbf{i}]$, can already reach the optimal normalized diversity product among all possible 2×2 generating matrices G of any complex entries. This result also shows that the L-DSTBC obtained in [9]

$$D_{2,2} = \left\{ \begin{pmatrix} s_1 + \exp(\mathbf{i}\pi/6)s_2 & 0 \\ 0 & s_1 + \exp(\mathbf{i}5\pi/6)s_2 \end{pmatrix} : \; s_1, s_2 \in \mathbb{Z}[\mathbf{i}] \right\} \tag{3}$$

has the optimal normalized diversity $1/\sqrt{3}$ as long as the information symbols s_1 and s_2 are from a QAM constellation, i.e., $\mathbb{Z}[\mathbf{i}]$. In this code $D_{2,2}$, the 2×2 generating matrix G_2 is

$$G_2 = \begin{pmatrix} 1 & \exp(\mathbf{i}\pi/6) \\ 1 & \exp(\mathbf{i}5\pi/6) \end{pmatrix}$$

and $\exp(\mathbf{i}\pi/6)$ and $\exp(\mathbf{i}5\pi/6)$ are the two roots of quadratic polynomial $x^2 - \mathbf{i}x - 1$.

II. New Upper Bound of Normalized Diversity Product of 2×2 L-DSTBC

In what follows, we always consider $\mathbb{Z}[\mathbf{i}]$ as the base lattice, i.e., $\mathcal{L} = \mathbb{Z}[\mathbf{i}]$, where $\mathbf{i} = \sqrt{-1}$. Thus, $|\mathcal{L}| = 1$. We only consider 2×2 L-DSTBC, i.e., $n = 2$. For notational convenience, a 2×2 generating matrix G is denoted as

$$G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{4}$$

where a, b, c, d are complex numbers and $\det(G) = ad - bc \neq 0$. We are interested in the following L-DSTBC generated from G:

$$\mathcal{C}(a,b,c,d) \triangleq \begin{pmatrix} ax + by & 0 \\ 0 & cx + dy \end{pmatrix}, \; x, y \in \mathbb{Z}[\mathbf{i}]. \tag{5}$$

Its normalized diversity product becomes

$$\xi(a,b,c,d) \triangleq \xi(G, \mathbb{Z}[\mathbf{i}]) = \frac{d_{\min}(a,b,c,d)}{|ad - bc|} \tag{6}$$

where

$$d_{\min}(a,b,c,d) \triangleq \min_{x,y \in \mathbb{Z}[\mathbf{i}], (x,y) \neq (0,0)} |ax + by| \cdot |cx + dy|. \tag{7}$$

The main result is as follows.

Theorem 1: The largest possible normalized diversity product of 2×2 lattice based diagonal space–time block codes when information symbols are taken from Gaussian integer ring $\mathbb{Z}[\mathbf{i}]$ is $1/\sqrt{3}$, i.e.

$$\max_{a,b,c,d \in \mathbb{C}} \xi(a,b,c,d) = 1/\sqrt{3} \tag{8}$$

where $\mathbb{C}$ denotes the field of complex numbers.

Proof: By a result from [9] that the normalized diversity product of code $D_{2,2}$ in (3) is $1/\sqrt{3}$, the inequality $\max_{a,b,c,d \in \mathbb{C}} \xi(a,b,c,d) \geq 1/\sqrt{3}$ is proved. The next is to prove the upper bound side, i.e., we need to prove

$$\max_{a,b,c,d \in \mathbb{C}} \xi(a,b,c,d) \leq 1/\sqrt{3}. \tag{9}$$

To do so, we first list two lemmas and their proofs are straightforward and therefore omitted.

Lemma 1: For any two nonzero complex numbers λ_1, λ_2, we have

$$\xi(\lambda_1 a, \lambda_1 b, \lambda_2 c, \lambda_2 d) = \frac{d_{\min}(\lambda_1 a, \lambda_1 b, \lambda_2 c, \lambda_2 d)}{|(\lambda_1 a)(\lambda_2 d) - (\lambda_1 b)(\lambda_2 c)|} = \frac{d_{\min}(a,b,c,d)}{|ad - bc|} = \xi(a,b,c,d).$$

This lemma is the scaling invariance of the normalized diversity product as also mentioned in [9].

Lemma 2: Let T be an unimodular 2×2 matrix over $\mathbb{Z}[\mathbf{i}]$, i.e.

$$T = \begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \end{pmatrix}$$

where $p_1, p_2, q_1, q_2 \in \mathbb{Z}[\mathbf{i}]$ and $|\det(T)| = 1$. Then, the transformation

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix}$$

on information symbols x, y does not change $d_{\min}(a,b,c,d)$ or $\xi(a,b,c,d)$ of the code $\mathcal{C}(a,b,c,d)$.

We now apply these two lemmas to simplify the following optimization problem $\mathcal{P}$:

$$\max_{a,b,c,d \in \mathbb{C}} \xi(a,b,c,d). \tag{10}$$

First, we may assume $ac \neq 0$. In fact, if $a = 0$, then the code becomes

$$\mathcal{C}(0,b,c,d) = \begin{pmatrix} by & 0 \\ 0 & cx + dy \end{pmatrix}.$$

In this case, letting $y = 0$ implies $d_{\min}(0,b,c,d) = 0$, and therefore $\xi(0,b,c,d) = 0$, which can not be the maximum. The case of $c = 0$ is similar.

Similarly, we may also assume $bd \neq 0$. Hence, it is enough to consider the optimization problem $\mathcal{P}$ under conditions $ac \neq 0$ and $bd \neq 0$.

We next convert the form of the optimization problem $\mathcal{P}$ into a simpler form. Assume that the minimum of $|ax + by||cx + dy|$ is achieved at point $(x, y) = (p_0, q_0)$ with the minimal value d_0, i.e.,

$$\min_{(x,y) \neq (0,0), x,y \in \mathbb{Z}[\mathbf{i}]} |ax + by||cx + dy| = d_0 = |ap_0 + bq_0||cp_0 + dq_0|. \tag{11}$$

We claim that p_0 and q_0 are coprime over $\mathbb{Z}[\mathbf{i}]$, i.e., $|\gcd(p_0, q_0)| = 1$. In fact, let $p = \gcd(p_0, q_0)$. Then, $|p| \geq 1$ and p_0/p and q_0/p are Gaussian integers. Dividing $|p|^2$ at the two sides of (11), we obtain another product form

$$\frac{d_0}{|p|^2} = |a\frac{p_0}{p} + b\frac{q_0}{p}||c\frac{p_0}{p} + d\frac{q_0}{p}|.$$

Since d_0 is the minimum among all the above product forms, we have $|p| \le 1$. Hence, $|p| = 1$.

Since p_0 and q_0 are coprime over $\mathbb{Z}[\mathbf{i}]$, by a basic result, see for example pg. 13 in [12], there are Gaussian integers p_1 and q_1 such that the following matrix:

$$T_1 = \begin{pmatrix} p_0 & p_1 \\ q_0 & q_1 \end{pmatrix}$$

is an unimodular matrix over $\mathbb{Z}[\mathbf{i}]$. Using the following transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} = T_1 \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{12}$$

a code $\mathcal{C}(a,b,c,d)$ can be changed into code $\mathcal{C}(a',b',c',d')$, where

$$\begin{aligned} a' &= ap_0 + bq_0; \quad & b' &= ap_1 + bq_1; \\ c' &= cp_0 + dq_0; \quad & d' &= cp_1 + dq_1. \end{aligned}$$

Since T_1 is unimodular, by Lemma 2, the new code $\mathcal{C}(a',b',c',d')$ is equivalent to the old one in terms of the normalized diversity product. Noticing that $|a'c'| = |(ap_0 + bq_0)(cp_0 + dq_0)| = d_0$, we know that in this new code, point $(x,y) = (1,0)$ achieves the minimum.

Since $|a'c'| = d_0 \ne 0$, we set

$$\lambda_1 \triangleq \frac{1}{a'} = \frac{1}{ap_0 + bq_0}, \quad \lambda_2 \triangleq \frac{1}{c'} = \frac{1}{cp_0 + dq_0}.$$

By Lemma 1, code $\mathcal{C}(\lambda_1 a', \lambda_1 b', \lambda_2 c', \lambda_2 d')$ and code $\mathcal{C}(a',b',c',d')$ have the same normalized diversity product while $\mathcal{C}(\lambda_1 a', \lambda_1 b', \lambda_2 c', \lambda_2 d')$ has the following form:

$$\mathcal{C}(1,t_1,1,t_2) = \begin{pmatrix} x + t_1 y & 0 \\ 0 & x + t_2 y \end{pmatrix} \tag{13}$$

where t_1 and t_2 are complex numbers. Furthermore

$$\min_{(x,y)\ne(0,0),x,y\in\mathbb{Z}[\mathbf{i}]} |x + t_1 y||x + t_2 y| = 1. \tag{14}$$

In the following, we always assume that a code, i.e., $\mathcal{C}(a,b,c,d) = \mathcal{C}(1,t_1,1,t_2)$, has the properties (13) and (14) without loss of generality. Note that the determinant absolute value of the 2×2 generating matrix G of this code is $|t_1 - t_2|$ and therefore the normalized diversity product is $1/|t_1 - t_2|$. Then, we can change the optimization problem $\mathcal{P}$ into the following equivalent one denoted by $\mathcal{P}_1$:

To find the minimum of $|t_1 - t_2|$ among all nonzero complex numbers t_1 and t_2, subject to

$$|x + t_1 y||x + t_2 y| \ge 1 \tag{15}$$

for all $(x,y) \ne (0,0)$ and $x,y \in \mathbb{Z}[\mathbf{i}]$.

In the following, we prove that the minimum of the above optimization problem $\mathcal{P}_1$ is greater than or equal to $\sqrt{3}$, which then implies Theorem 1. Before going to the proof, we need the following lemma.

Lemma 3: Let $\mathbb{P}$ be a convex quadrangle on the plane with four vertices $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_4$ as shown in Fig. 1, where convex means that the set of the points inside $\mathbb{P}$ is a convex set. Let p_i be the angle corresponding to the vertex $\mathbf{P}_i$ for $i = 1,2,3,4$, as shown in Fig. 1. Denote the length of the segment from $\mathbf{P}_i$ to $\mathbf{P_j}$ as l_{ij}, where $1 \le i,j \le 4$. If $l_{13} = 1$, $l_{12}l_{14} \ge 1$ and $l_{32}l_{34} \ge 1$, then $l_{24} \ge \sqrt{3}$.

This lemma is about some basic plane geometry. Its proof is in Appendix.

Assume that $k \in \mathbb{Z}[\mathbf{i}]$. Then, the following transformation to the information symbols x and y

$$\begin{pmatrix} 1 & -k \\ 0 & 1 \end{pmatrix}$$

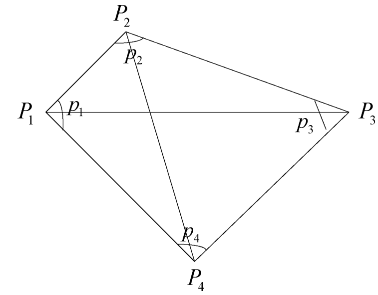

Fig. 1. Quadrangle 1.

does not change the normalized diversity product according to Lemma 2, since the transformation is unimodular. By absorbing this matrix into G, it maintains the properties (13) and (14) while parameters (t_1, t_2) are changed into $(t_1 - k, t_2 - k)$, which does not change the above optimization problem $\mathcal{P}_1$. Therefore, if we take $k \in \mathbb{Z}[\mathbf{i}]$ such that it is the nearest point to t' from the left-below side of the complex plane, where $t' = t_1$ if $\mathrm{Re}(t_1) > \mathrm{Re}(t_2)$ and $|\mathrm{Re}(t_1 - t_2)| > 1$, or $t' = t_2$ if $\mathrm{Re}(t_2) > \mathrm{Re}(t_1)$ and $|\mathrm{Re}(t_1 - t_2)| > 1$; when $|\mathrm{Re}(t_1 - t_2)| \le 1$, we let $t' \in \{t_1, t_2\}$ such that it has the greater imaginary part. Thus, we have $0 \le \mathrm{Re}(t' - k) \le 1$ and $0 \le \mathrm{Im}(t' - k) \le 1$. In what follows, without loss of generality, we assume $t' = t_1$. If we use the polar coordinate to denote t_1 and t_2 as

$$t_1 = r_1 \exp(\mathbf{i}\alpha) \text{ and } t_2 = r_2 \exp(\mathbf{i}\beta)$$

then, we can assume

$$0 \le r_1 \cos(\alpha) \le 1, \quad 0 \le r_1 \sin(\alpha) \le 1 \quad \text{and} \quad 0 \le \alpha \le \pi/2.$$

For β, by the above selection of t' and the fact $|t_1 - t_2| \ge \sqrt{2}$ that comes from the result $\max_{a,b,c,d\in\mathbb{C}} \xi(a,b,c,d) < 1/\sqrt{2}$ obtained in [11], we have $-\pi \le \beta \le 0$ or $\pi/2 \le \beta \le \pi$. If $\pi/2 \le \beta \le \pi$, we consider $(t_1', t_2') \triangleq (-\mathbf{i}t_2, -\mathbf{i}t_1)$. Clearly, the code using the numbers (t_1', t_2') has the same normalized diversity product as the one using (t_1, t_2), but the angle of t_2' is in the interval $[-\pi/2, 0]$ and the angle of t_1' is in the interval $[0, \pi/2]$. Therefore, we can assume $-\pi \le \beta \le 0$. In summary, in the following, we always assume:

$$0 \le r_1 \cos(\alpha) \le 1, \quad 0 \le r_1 \sin(\alpha) \le 1$$

and

$$0 \le \alpha \le \pi/2, -\pi \le \beta \le 0. \tag{16}$$

In condition (15), if we let $(x,y) = (0,1)$, we obtain $|t_1||t_2| \ge 1$ and if we let $(x,y) = (1,-1)$, we obtain $|1 - t_1||1 - t_2| \ge 1$. Therefore

$$r_1 r_2 \ge 1 \quad \text{and} \quad |1 - r_1 \exp(\mathbf{i}\alpha)||1 - r_2 \exp(\mathbf{i}\beta)| \ge 1. \tag{17}$$

Define

$$\begin{aligned} f(t_1,t_2) &\triangleq |t_1 - t_2|^2 = |r_1 \exp(\mathbf{i}\alpha) - r_2 \exp(\mathbf{i}\beta)|^2 \\ &= r_1^2 + r_2^2 - 2r_1 r_2 \cos(\alpha - \beta). \end{aligned}$$

We next prove that the minimum of the above function f is greater than or equal to 3 under condition (15). Let us assume that there are two points t_1 and t_2 such that $f(t_1,t_2) < 3$, and we will derive a contradiction.

Since $f(t_1,t_2) < 3$, i.e., $r_1^2 + r_2^2 - 2r_1 r_2 \cos(\alpha - \beta) < 3$, we have

$$\begin{aligned} 3 &> 2r_1 r_2 - 2r_1 r_2 \cos(\alpha - \beta) \\ &= 2r_1 r_2 (1 - \cos(\alpha - \beta)) \ge 2(1 - \cos(\alpha - \beta)). \end{aligned}$$

Therefore, $\cos(\alpha - \beta) > -1/2$, i.e., $\alpha - \beta < 120^0$ or $\alpha - \beta > 240^0$.

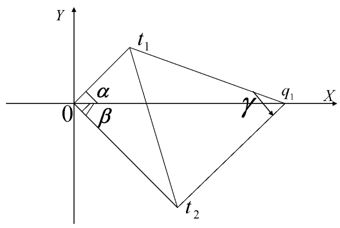

Fig. 2. Quadrangle 2.

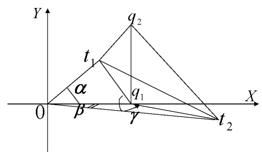

Fig. 3. Quadrangle 3.

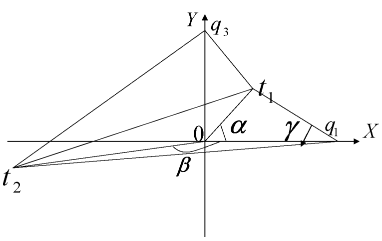

Fig. 4. Quadrangle 4.

Similarly, if we denote the angle from vector $t_1 - q_1$ to vector $t_2 - q_1$ on the complex plane according to the inverse clock-wise direction by γ as shown in Figs. 2–4, where $q_1 = 1$, then, by using triangle $\Delta t_1 q_1 t_2$ on the complex plane, we have

$$|t_1 - q_1|^2 + |t_2 - q_1|^2 - 2|t_1 - q_1| \cdot |t_2 - q_1| \cos(\gamma) = |t_1 - t_2|^2 < 3.$$

Hence, $2|t_1 - q_1| \cdot |t_2 - q_1|(1 - \cos(\gamma)) < 3$. Noticing that $|t_1 - q_1| = |r_1 \exp(\mathbf{i}\alpha) - 1|$ and $|t_2 - q_1| = |r_2 \exp(\mathbf{i}\beta) - 1|$ and by condition (17), we also have $\cos(\gamma) > -1/2$, and therefore, $\gamma < 120^0$ or $\gamma > 240^0$.

Consider the case when $0 \leq \alpha - \beta < 120^0$ and $0 \leq \gamma < 120^0$. In this case, because point t_1 is above the X-axis and point t_2 is below the X-axis, four complex numbers (or points) $0, t_1, q_1, t_2$ form a convex quadrangle (see Fig. 2) on the complex plane. Furthermore, in this quadrangle, the product of the lengthes of two segments from 0 to t_1 and from 0 to t_2 is $r_1 r_2$, which is greater than or equal to 1; and the product of two segments from q_1 to t_1 and from q_1 to t_2 is $|(1 - t_1)(1 - t_2)|$, which is also greater than or equal to 1 from (17). Thus, by Lemma 3, the length of the segment from t_1 to t_2 is greater than or equal to $\sqrt{3}$, i.e., $|t_1 - t_2| \geq \sqrt{3}$, which contradicts with the assumption $f(t_1, t_2) < 3$.

Consider the case when $0 \leq \alpha - \beta < 120^0$ and $\gamma > 240^0$. Because $-\pi \leq \beta \leq 0^0$, point t_2 on the complex plane is below the X-axis. On the other hand, by condition (16), point t_1 on the complex plane is on the left side of the line $X = 1$. Thus, in this case, four points q_1, t_1, q_2, t_2 on the complex plane form a convex quadrangle (see Fig. 3), where $q_2 = 1 + \mathbf{i}$. Let $(x, y) = (1 + \mathbf{i}, -1)$ in (15), we get $|q_2 - t_1| \cdot |q_2 - t_2| \geq 1$. Thus, combining conditions $|q_1 - t_1| \cdot |q_1 - t_2| \geq 1$ and $|q_1 - q_2| = 1$, again by Lemma 3, we also have $|t_1 - t_2| \geq \sqrt{3}$, which contradicts with the assumption $f(t_1, t_2) < 3$.

Consider the case when $\alpha - \beta > 240^0$ and $0^0 \leq \gamma < 120^0$. Because point t_2 on the complex plane is below the X-axis and point t_1 is on the right side of the Y-axis, four points $0, t_1, q_3, t_2$ form a convex quadrangle (see Fig. 4), where $q_3 = \mathbf{i}$. Let $(x, y) = (\mathbf{i}, -1)$ in (15), we get $|q_3 - t_1| \cdot |q_3 - t_2| \geq 1$. Thus, by Lemma 3, we also have the contradiction.

The last case is when $240^0 \leq \alpha - \beta < 270^0$ and $240^0 \leq \gamma$, which is impossible because $\alpha - \beta + \gamma \geq 480^0 > 360^0$.

Summarizing the above cases, we have proved the theorem. $\mathbf{q.e.d.}$

III. Conclusion

In this correspondence, we have proved that $1/\sqrt{3}$ is the optimal normalized diversity product of lattice based 2×2 diagonal space time block codes (L-DSTBC) with generating matrices of complex entries and information symbols in $\mathbb{Z}[\mathbf{i}]$, i.e., a QAM constellation. This result implies that for Gaussian integer information symbols, i.e., QAM signal constellations, the optimal normalized diversity product of 2×2 L-DSTBC can be reached when their generating matrices are over quadratic algebraic extensions of Gaussian integers.

Appendix
Proof of Lemma 3

We use the notations given in Fig. 1. Assume that $l_{24} < \sqrt{3}$. We will get a contradiction. Before we go to the proof, we cite a fact from a book [14, p. 66]:

Lemma 4: For any four points $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$, the following identity holds:

$$|\mathbf{AC}|^2 \cdot |\mathbf{BD}|^2 = |\mathbf{AB}|^2 \cdot |\mathbf{CD}|^2 + |\mathbf{AD}|^2 \cdot |\mathbf{BC}|^2 - 2|\mathbf{AB}| \cdot |\mathbf{BC}| \cdot |\mathbf{CD}| \cdot |\mathbf{DA}| \cdot \cos(\angle ABC + \angle CDA).$$

Applying this lemma to four points $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3, \mathbf{P}_4$, we get

$$\cos(p_2 + p_4) = \frac{l_{12}^2 l_{34}^2 + l_{14}^2 l_{23}^2 - l_{13}^2 l_{24}^2}{2 l_{12} l_{23} l_{34} l_{41}}.$$

Because $l_{13} = 1, l_{12} l_{14} \geq 1$ and $l_{23} l_{34} \geq 1, l_{24} < \sqrt{3}$, we have

$$\cos(p_2 + p_4) > \frac{2 l_{12} l_{23} l_{34} l_{41} - 3}{2 l_{12} l_{23} l_{34} l_{41}} = 1 - \frac{3}{2 l_{12} l_{23} l_{34} l_{41}} \geq 1 - \frac{3}{2} = -\frac{1}{2}.$$

Therefore

$$p_2 + p_4 < 120^0 \quad \text{or} \quad p_2 + p_4 > 240^0. \tag{18}$$

Because $p_1 + p_3 = 360^0 - (p_2 + p_4)$, we also have

$$p_1 + p_3 > 240^0 \quad \text{or} \quad p_1 + p_3 < 120^0. \tag{19}$$

On the other hand, on the triangle $\Delta \mathbf{P}_1 \mathbf{P}_2 \mathbf{P}_4$, because $l_{24}^2 = l_{12}^2 + l_{14}^2 - 2 l_{12} l_{14} \cos(p_1) < 3$, i.e.

$$\cos(p_1) > \frac{l_{12}^2 + l_{14}^2 - 3}{2 l_{12} l_{14}} \geq \frac{2 l_{12} l_{14} - 3}{2 l_{12} l_{14}} \geq 1 - \frac{3}{2 l_{12} l_{14}} \geq 1 - \frac{3}{2} = -\frac{1}{2}$$

where the last inequality is from the assumption $l_{12} l_{14} \geq 1$. Therefore, $p_1 < 120^0$ or $p_1 > 240^0$. Because $\mathbb{P}$ is convex, we have $p_1 < 120^0$. Similarly, using triangle $\Delta \mathbf{P}_3 \mathbf{P}_2 \mathbf{P}_4$, we also have $p_3 < 120^0$. Thus, from (18) and (19), we obtain

$$p_1 + p_3 < 120^0 \quad \text{and} \quad p_2 + p_4 > 240^0. \tag{20}$$

From $p_1+p_3 < 120^0$, we know that $p_1 < 60^0$ or $p_3 < 60^0$. Without loss of generality, we can assume

$$p_1 < 60^0. \tag{21}$$

Also from $p_2 + p_4 > 240^0$, we know that $p_2 > 120^0$ or $p_4 > 120^0$. Without loss of generality, we can also assume

$$p_2 > 120^0. \tag{22}$$

From the convexity of quadrangle $\mathbb{P}$, we have $p_2 \leq 180^0$. Hence, $\cos(p_2) < -1/2$. Therefore, on the triangle $\Delta \mathbf{P}_1\mathbf{P}_2\mathbf{P}_3$

$$1 = l_{13}^2 = l_{12}^2 + l_{32}^2 - 2l_{12}l_{32}\cos(p_2) > l_{12}^2 + l_{32}^2.$$

The above inequality implies $l_{12} < 1$ and $l_{32} < 1$. By conditions $l_{12}l_{14} \geq 1$ and $l_{32}l_{34} \geq 1$, we have

$$l_{14} > 1, \quad \text{and} \quad l_{34} > 1.$$

Thus $l_{14}l_{34} > 1$.

On the triangle $\Delta \mathbf{P}_1\mathbf{P}_4\mathbf{P}_3$, we have $1 = l_{13}^2 = l_{14}^2 + l_{34}^2 - 2l_{14}l_{34}\cos(p_4)$. So,

$$\begin{aligned}\cos(p_4) &= \frac{l_{14}^2 + l_{34}^2 - 1}{2l_{14}l_{34}} \geq \frac{2l_{14}l_{34} - 1}{2l_{14}l_{34}} \\ &\geq 1 - \frac{1}{2l_{14}l_{34}} > 1 - \frac{1}{2} = \frac{1}{2}\end{aligned} \tag{23}$$

which implies $p_4 < 60^0$. Thus, from (20), we have $p_2 > 240^0 - p_4 > 240^0 - 60^0 = 180^0$, which contradicts with the fact that quadrangle $\mathbb{P}$ is convex. We have thus proved the lemma. **q.e.d.**

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their careful reading and useful comments for improving the presentation of this paper and for pointing out some errors in the early versions of this paper.

REFERENCES

[1] K. Boulle and J.-C. Belfiore, "Modulation schemes designed for the Rayleigh fading channel," in *Proc. CISS'92*, Princeton, NJ, Mar. 1992.
[2] X. Giraud, E. Boutillon, and J.-C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Trans. Inf. Theory*, vol. 43, pp. 938–952, May 1997.
[3] J. Boutros and E. Viterbo, "Signal space diversity: A power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. Inf. Theory*, vol. 44, pp. 1453–1467, Jul. 1998.
[4] M. O. Damen, K. A. Meraim, and J.-C. Belfiore, "Diagonal algebraic space-time block codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 628–636, Mar. 2002.
[5] M. O. Damen, A. Tewfik, and J.-C. Belfiore, "A construction of a space-time code based on number theory," *IEEE Trans. Inf. Theory*, vol. 48, pp. 753–760, Mar. 2002.
[6] H. El Gamal and A. R. Hammons Jr., "A new approach to layered space-time code and signal processing," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2335–2367, Sep. 2001.
[7] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1097–1119, May 2003.
[8] B. A. Sethuraman, B. S. Rajan, and V. Shashidahar, "Full-diversity, high-rate space-time block codes from division algebra," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2596–2616, Oct. 2003.
[9] G. Wang, H. Liao, H. Wang, and X.-G. Xia, "Systematic and optimal cyclotomic lattice and diagonal space-time block code designs," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3348–3360, Dec. 2004.
[10] G. Wang and X.-G. Xia, "On optimal multi-layer cyclotomic space-time code designs," *IEEE Trans. Inf. Theory*, vol. 51, pp. 1102–1135, Mar. 2005.
[11] H. Liao, H. Wang, and X.-G. Xia, "Some designs and normalized diversity product upper bounds for lattice based diagonal and full rate space-time block codes," *IEEE Trans. Inform. Theory*, Sep. 2004, submitted for publication.
[12] M. Newman, *Integral Matrices*. New York: Academic, 1972.
[13] J. W. S. Cassels, *An Introduction to the Geometry of Numbers*. Berlin, Germany: Springer-Verlag, 1959.
[14] R. A. Johnson, *Modern Geometry*. Cambridge, MA: Houghton Mifflin, 1929.

Maximum Entropy for Sums of Symmetric and Bounded Random Variables: A Short Derivation

Yaming Yu

***Abstract*—Let $X_1, \ldots, X_n$ be n independent, symmetric, random variables on the interval $[-1, 1]$. Ordentlich (2006) showed that the differential entropy of $S_n = \sum_{i=1}^n X_i$ is maximized when X_i, $i = 1, \ldots, n-1$ are symmetric Bernoulli random variables and X_n is $\mathrm{uniform}(-1, 1)$. We give a short derivation of this result via an alternative proof of a key lemma of Ordentlich (2006).**

***Index Terms*—Differential entropy, maximum entropy.**

I. INTRODUCTION

Given n independent, symmetric, random variables $X_1, \ldots, X_n$ on the interval $[-1, 1]$, Ordentlich (2006) has the following result.

Theorem 1.1: Let $Z_1, \ldots, Z_{n-1}$ be independent and identically distributed (i.i.d.) random variables taking on 1 and -1 with equal probability. Let U be independent of $Z_1, \ldots, Z_{n-1}$ and uniformly distributed on $[-1, 1]$. Then

$$h\left(\sum_{i=1}^n X_i\right) \leq h\left(U + \sum_{i=1}^{n-1} Z_i\right)$$

where $h(S) = -\int_{-\infty}^{\infty} f(s)\log_2 f(s)\,ds$ is the differential entropy for a continuous random variable S with density $f(s)$.

In other words, the entropy of $\sum_{i=1}^n X_i$ is maximized when $X_1, \ldots, X_{n-1}$ are symmetric Bernoulli random variables and X_n is $\mathrm{uniform}(-1, 1)$. For the information-theoretic background on this problem, see Ordentlich [1].

Theorem 1.1 is a consequence of the following key lemma.

Lemma 1.1: If $Z_1, \ldots, Z_n$ are i.i.d. Bernoulli random variables taking on values 1 and -1 with equal probability, and if constants $a_1, \ldots, a_n$ satisfy $0 \leq a_i \leq 1$, then

$$\begin{aligned}&\Pr\left(\sum_{i=1}^n Z_i a_i \in [-n+2j,\ n-2j]\right) \\ &\qquad \geq \Pr\left(\sum_{i=1}^{n-1} Z_i \in [-n+2j,\ n-2j]\right)\end{aligned} \tag{1}$$

where j is any integer such that $n - 2j > 0$.

Manuscript received September 14, 2007; revised December 12, 2007.
The author is with the Department of Statistics, University of California, Irvine, CA 92697-1250 USA (e-mail: yamingy@uci.edu).
Communicated by W. Szpankowski, Associate Editor for Source Coding.
Digital Object Identifier 10.1109/TIT.2008.917660