Fast lattice decodability of space-time block codes

Grégory Berhuy, Nadya Markin, B.A. Sethuraman
Institut Fourier, 100 rue des maths, BP 74, 38402 St Martin d'Hères cedex, France
Division of Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371
Department of Mathematics, California State University Northridge, Northridge, CA 91330, USA.
gregory.berhuy@ujf-grenoble.fr, nmarkin@ntu.edu.sg, al.sethuraman@csun.edu

***Abstract*—We study fast decodability for general rate space-time block codes. We derive bounds on the number of generators that can be in mutually orthogonal groups and use these to show that that the ML decoding complexity of a full-rate $n \times n$ space-time block code is unfortunately quite high: at least of the order of $|S|^{n^2+1}$, where S is the effective real constellation. We also show that g-group decodability is not possible for full-rate codes.**

***Index Terms*—Fast-decodability, space-time block codes, skew-Hermitian matrices.**

I. Introduction

The transmission of a space-time block code consisting of elements of $n \times n$ matrices in $\mathcal{M}_{n\times n}(\mathbb{C})$ is modelled by considering points on a lattice inside $\mathbb{R}^{2n^2}$. When $2l$ is the real dimension of the lattice, the complexity of Maximum Likelihood (ML) decoding by brute force checking of all $2l$-tuples is $|S|^{2l}$, where S is the effective real alphabet in use. Efforts to improve on this decoding complexity have initiated a search for codes with reduced decoding complexity [1], which have been a subject of research since, for instance [2]–[9]. Roughly speaking, the improvement on the decoding complexity is achieved when for given values in a (possibly empty) subset of information symbols, the complement can be partitioned into groups which are decoded independently of one another. The authors of [3] give a sufficient condition, which we refer to as *mutual orthogonality*, on the generating matrices of the code, which guarantees the existence of such a partitioning into independently decodable symbols regardless of the channel matrix H. In this paper we prove that this mutual orthogonality condition is not only sufficient but indeed necessary, in the sense that it must be satisfied if the partitions of the information symbols are to be decoded independently for any matrix H (see also Remark 1 ahead). This leads us to formulate a new definition of fast decodability which is intrinsic to the code as it involves solely the (mutual orthogonality) properties of the generating matrices of the code, and does not depend on the channel matrix H. We then prove the equivalence of this new definition to the current definition [3], which does depend implicitly on the channel matrix H.

The necessity of the mutual orthogonality criterion for fast decodability allows us to analyze lower bounds on ML decoding complexity of full-rate ($l = n^2$, l as above) space-time block codes (Section IV). Using only the elementary fact that the dimension of the $\mathbb{R}$-space of skew-Hermitian matrices is n^2, we find that the lower bound on ML decoding complexity is of the order $|S|^{n^2+1}$, which is unfortunately quite high. Further, we show that in fact g-group decodability (where the information symbols are completely partitioned into proper independently decodable groups) is not possible for full-rate codes.

This paper is organized as follows. We give the system model and the background on decoding in Section II. We prove the necessity of the mutual orthogonality condition for fast-decodability in Section III, and then propose a new definition of fast (lattice) decodability independent of the channel matrix; we also demonstrate its equivalence to the current definition. Finally, we establish our results on lower bounds on ML decoding complexity and impossibility of g-group decodability for full-rate codes in Section IV

II. System Model and Maximum Likelihood Decoding

We consider transmission over a quasi-static Rayleigh fading channel with perfect channel state information at the receiver. We assume that the number of receive antennas and the number of transmit antennas are the same, namely n.

The received codeword Y is given by

$$Y = HX + N \tag{1}$$

where the codeword X, the channel matrix H and the noise matrix N are all $n \times n$ complex valued matrices, i.e., elements of $\mathcal{M}_{n\times n}(\mathbb{C})$, with the assumption that the entries of H are i.i.d. circularly symmetric complex Gaussian with zero mean and variance 1, and the entries of N are i.i.d. complex Gaussian with zero mean and variance N_0.

Each codeword X is represented as $X = \sum_{i=1}^{2l} s_i A_i$, in terms of generating matrices A_i weighted by information symbols s_i which arise from the effective real alphabet S obtained by splitting each complex symbol into its real and imaginary parts. The generating matrices A_i, also referred to as a basis of the code, are fixed $\mathbb{R}$-linearly independent complex valued matrices. We assume that the generating matrices are all invertible, which is not a significant constraint from the viewpoint of diversity. Maximum-likelihood (ML) decoding amounts to finding the information symbols $s_1, \ldots, s_{2l}$ that result in a codeword $X = \sum_{i=1}^{2l} s_i A_i$ which minimizes the squared Frobenius norm

$$||Y - HX||_F^2. \tag{2}$$

The matrices appearing in Equation 1 are first converted to vectors in real space via the map

$$\mathrm{Vec}_{\mathbb{R}} : \mathcal{M}_{n\times n} \to \mathbb{C}^{n^2} \to \mathbb{R}^{2n^2}$$

which first stacks the columns of a matrix to obtain a vector in $\mathbb{C}^{n^2}$ and then separates each complex entry into its real and imaginary parts to obtain a vector in $\mathbb{R}^{2n^2}$. We have

$$\mathrm{Vec}_{\mathbb{R}}(Y) = \sum_{j=1}^{2l} s_j \mathrm{Vec}_{\mathbb{R}}(HA_j) + \mathrm{Vec}_{\mathbb{R}}(N).$$

Given a channel matrix H, we define the matrix $T = T(H) \in \mathcal{M}_{2n^2\times 2l}(\mathbb{R})$ to be the real matrix whose j-th column is $\mathrm{Vec}_{\mathbb{R}}(HA_j)$. Then we have

$$\mathrm{Vec}_{\mathbb{R}}(\sum_{j=1}^{2l} s_j HA_j) = T\begin{pmatrix} s_1 \\ \vdots \\ s_{2l} \end{pmatrix}.$$

We can view T as the basis matrix for the $2l$-dimensional lattice in $\mathbb{R}^{2n^2}$ from which points are transmitted. Letting $\mathbf{s}$ denote the transpose of $(s_1, \ldots, s_{2l})$, the decoding problem now becomes to find an information vector $\mathbf{s}$ which minimizes the Euclidean distance

$$|\mathrm{Vec}_{\mathbb{R}}(Y) - T\mathbf{s}| \tag{3}$$

of vectors in $\mathbb{R}^{2n^2}$.

III. Fast Lattice Decodability

When a code has no special structure, the number of computations needed in order to minimize the distance in Equation (3) above is $|S|^{2l}$. Several authors [1], [3] studied improved lattice decodability of space-time block codes by considering a QR decomposition of the transmitted lattice matrix T in Equation 3 above, and rewriting Equation 3 as

$$Q^*\mathrm{Vec}_{\mathbb{R}}(Y) = R\cdot\mathbf{s} + Q^*\mathrm{Vec}_{\mathbb{R}}(N). \tag{4}$$

Since Q^* is unitary, the new noise vector $Q^*\mathrm{Vec}_{\mathbb{R}}(N)$ is still i.i.d. real Gaussian, so the maximum likelihood estimate for $\mathbf{s}$ is given by minimizing $|Q^*\mathrm{Vec}_{\mathbb{R}}(Y) - R\cdot\mathbf{s}|$. The location of zero blocks in the R matrix indicates which symbols s_i can be decoded independently of one another. Producing a fast decodable code was accomplished [2], [3] by choosing the basis matrices A_i satisfying a mutual orthogonality condition (see Definition 1 below), which suffices to guarantee that the matrix R has a certain zero block structure, regardless of the channel matrix H on which T and hence R depend. We start by showing the necessity of the aforementioned mutual orthogonality condition on A_i. This leads us to an alternative definition of fast decodability which depends only on mutual orthogonality properties of the generating matrices of the code, and does not depend on channel matrix H.

Definition 1. *We say that two complex matrices, A, B are* mutually orthogonal *if $AB^* + BA^* = 0$.*

The reason for the choice of this term is because, as we show in the following theorem, two basis matrices A_i and A_j satisfy $A_iA_j^* + A_jA_i^* = 0$ if and only if the i-th and j-th columns of $T(H)$ for any H are mutually orthogonal as vectors in $\mathbb{R}^{2n^2}$. Although our proof is new, see Remark 1 ahead.

Theorem 1. *The i-th and j-th columns of $T = T(H)$ are orthogonal as vectors in $\mathbb{R}^{2n^2}$ for all channel matrices H if and only if the basis matrices A_i satisfy $A_iA_j^* + A_jA_i^* = 0$.*

Proof. It is elementary that the dot product $\mathrm{Vec}_{\mathbb{R}}(A)\cdot\mathrm{Vec}_{\mathbb{R}}(B)$ equals $\mathrm{Re}\,(\mathrm{Tr}(AB^*))$ for any matrices A, B. It follows that orthogonality of the i-th and j-th columns of T is equivalent to the condition $\mathrm{Re}\,(\mathrm{Tr}((HA_i)(HA_j)^*)) = 0$. Also, note that $\mathrm{Tr}((HA_i)(HA_j)^*) = \mathrm{Tr}(HA_iA_j^*H^*) = \mathrm{Tr}((A_iA_j^*)(H^*H))$, where the second equality is because $\mathrm{Tr}(XY) = \mathrm{Tr}(YX)$ for two matrices X and Y.

Assume that $A_iA_j^* + A_jA_i^* = 0$ for $i \neq j$. Then $A_iA_j^*$ is skew Hermitian, while H^*H is of course Hermitian. If M is skew Hermitian and N is Hermitian, then note that $(MN)^* = N^*M^* = -NM$. Since for any matrix X we have $\mathrm{Re}(\mathrm{Tr}(X)) = \mathrm{Re}(\mathrm{Tr}(X^*))$, we find that for $X = MN$, $\mathrm{Re}(\mathrm{Tr}(MN)) = \mathrm{Re}(\mathrm{Tr}((MN)^*)) = \mathrm{Re}(\mathrm{Tr}(-NM)) = -\mathrm{Re}(\mathrm{Tr}(NM)) = -\mathrm{Re}(\mathrm{Tr}(MN))$. It follows that $\mathrm{Re}(\mathrm{Tr}(MN)) = 0$. In particular, for $M = A_iA_j^*$ and $N = H^*H$, we find $0 = \mathrm{Re}\left(\mathrm{Tr}((A_iA_j^*)(H^*H))\right) = \mathrm{Re}\left(\mathrm{Tr}((HA_i)(A_j^*H^*))\right) = \mathrm{Re}\,(\mathrm{Tr}((HA_i)(HA_j)^*))$.

Now assume that the trace condition holds. We write this as $\mathrm{Re}\left(\mathrm{Tr}((A_iA_j^*)(H^*H))\right) = 0$ for all matrices H. Write M for $A_iA_j^*$. We wish to show that M is skew Hermitian. The matrix $E_{k,k}$ that has 1 in the (k,k) slot and zeros elsewhere satisfies $E_{k,k}^*E_{k,k} = E_{k,k}$. Choosing $H = E_{k,k}$, we find that the matrix $MH^*H = ME_{k,k}$ will have the k-th column of M in the k-th column, and zeros elsewhere. The trace condition now shows that the (k,k) element of M is purely imaginary. We next need to show that $m_{l,k} = -\overline{m_{k,l}}$ for $k \neq l$, where we have written $m_{i,j}$ for the (i,j)-th entry of M. Computing directly, we find the following relations hold (where $E_{i,j}$ has 1 in the (i,j) slot and zeros everywhere else):

$$\begin{aligned} E_{k,k} + E_{k,l} + E_{l,k} + E_{l,l} &= (E_{k,k} + E_{l,k})\cdot(E_{k,k} + E_{k,l}) \\ E_{k,k} - \imath E_{k,l} + \imath E_{l,k} + E_{l,l} &= (E_{k,k} + \imath E_{l,k})\cdot(E_{k,k} - \imath E_{k,l}) \end{aligned}$$

Thus, each of the matrices on the left sides of the two equations above can be written as H^*H, where H is the second factor on the right. Again computing directly, we find that $M\cdot(E_{k,k}+E_{k,l}+E_{l,k}+E_{l,l})$ has $m_{k,k}+m_{k,l}$ in the (k,k) slot and $m_{l,k}+m_{l,l}$ in the (l,l) slot, and zeros elsewhere in the diagonal. Hence, $\mathrm{Re}(\mathrm{Tr}(M\cdot(E_{k,k}+E_{k,l}+E_{l,k}+E_{l,l}))) = \mathrm{Re}(m_{k,k}+m_{k,l}+m_{l,k}+m_{l,l})$. Since we have already seen that the diagonal elements of M are purely imaginary, we find $\mathrm{Re}(m_{k,l} + m_{l,k}) = 0$. Similarly, we find $\mathrm{Re}(\mathrm{Tr}(M\cdot(E_{k,k} - \imath E_{k,l} + \imath E_{l,k} + E_{l,l}))) = \mathrm{Re}(m_{k,k} + \imath m_{k,l} - \imath m_{l,k} + m_{l,l})$. Once again, because the diagonal elements of M are purely imaginary, we find $\mathrm{Im}(m_{k,l} - m_{l,k}) = 0$. These two together show that $m_{l,k} = -\overline{m_{k,l}}$ for $k \neq l$. Together with the fact that

the diagonal elements of M are purely imaginary, we find $M = A_iA_j^*$ is skew Hermitian, as desired.

□

Remark 1. As mentioned in Section I, the sufficiency of the condition $A_iA_j^* + A_jA_i^* = 0$ for orthogonality of the columns of T, and hence for fast decodability, was already considered before ([7, Theorem 2], [2, Theorem 1]). What is new here is the necessity of the condition. It is the consequences of the necessity that enables us to analyze lower bounds on fast decodability in the section ahead. We should point out however, that we noticed after we proved our results that the authors of [2] also mention the necessity of this condition. However, they do not give a proof of the necessity in that paper. Tracking this further, we discovered that the authors of [10] have actually provided a proof of this result. Their proof is by an explicit computation. Indeed, they write down the entries of $T(H)$, blockwise, in terms of the matrices H and A_i, and compute $T(H)^*T(H)$. From the derived block structure of $T(H)^*T(H)$ they read off the necessity of the mutual orthogonality. This is of course very different from our approach.

The theorem above allows us to rephrase orthogonality properties among columns of the matrix $T(H)$ in terms of mutual orthogonality of generating matrices of the code, independently of H. Thanks to that, we can now define fast decodability using only the properties of generating matrices of the code, independent of the channel matrix H.

Definition 2. *[cf. [3, Definition 4]] We will say that the space-time block code defined by the matrices $X = \sum_{i=1}^{2l} s_iA_i$ admits fast (lattice) decodability if for $g \geq 2$ there exist disjoint subsets $\Gamma_1, \ldots, \Gamma_g, \Gamma_{g+1}$, with Γ_{g+1} possibly empty, of cardinalities $n_1, \ldots, n_g, n_{g+1}$ respectively, whose union is $\{1, \ldots, 2l\}$, such that for all $u \in \Gamma_i$ and $v \in \Gamma_j$ $(1 \leq i < j \leq g)$, the generating matrices A_u, A_v are mutually orthogonal.*

Remark 2. Given a code that admits fast (lattice) decodability, we can define a permutation $\pi : \{1, \ldots, 2l\} \to \Gamma_1 \cup \ldots \cup \Gamma_g \cup \Gamma_{g+1}$, which sends the first n_1 elements $\{1, \ldots, n_1\}$ to Γ_1, the next n_2 elements $\{n_1+1, \ldots, n_1+n_2\}$ to Γ_2 and so on, where $n_i = |\Gamma_i|$ for $i = 1, \ldots, g+1$. Given such a permutation π, we write T_π for the matrix whose i-th column is the $\pi(i)$-th column of T, namely, $\mathrm{Vec}_{\mathbb{R}}(HA_{\pi(i)})$; note that $T_\pi = T_\pi(H)$ depends on H.

Current definition of fast decodability given in [3, Definition 4] depends on the structure of the matrix R arising from the QR decomposition of T, which invokes the channel matrix H, and implicitly requires that matrix R has the block structure given below in (5). We show below that the two definitions are equivalent:

Proposition 1. *The space-time block code $X = \sum_{i=1}^{2l} s_iA_i$ admits fast (lattice) decodability as per Definition 2 if and only if for all channel matrices H, there exists a permutation π of the index set $\{1, \ldots, 2l\}$, integers $g \geq 2$, $n_i \geq 1$ $(i = 1, \ldots, g)$, and $n_{g+1} \geq 0$, with $n_1 + \cdots + n_{g+1} = 2l$, such that the matrix R obtained by doing a QR decomposition on $T_\pi = T_\pi(H)$ by doing a Gram-Schmidt orthogonalization in the order first column, then second column, and so on, has the special block form below:*

$$\begin{pmatrix} B_1 & & & & N_1 \\ & B_2 & & & N_2 \\ & & \ddots & & N_3 \\ & & & B_g & N_g \\ & & & & N_{g+1} \end{pmatrix} \tag{5}$$

for some matrices $B_1, \ldots, B_g$, and $N_1, \ldots, N_{g+1}$. Here, all empty spaces are filled by zeros, the B_i are of size $n_i \times n_i$ and N_{g+1} is of size $n_{g+1} \times n_{g+1}$.

Before we prove this, we remark in more detail why previous authors have been interested in the special form of R above: On applying the permutation π to Equation (3), we get $\mathrm{Vec}_{\mathbb{R}}(Y) = T_\pi \cdot \mathbf{s}_\pi + \mathrm{Vec}_{\mathbb{R}}(N)$, and then, as in the beginning of this section, premultiplying by Q^* we find $Q^*\mathrm{Vec}_{\mathbb{R}}(Y) = R \cdot \mathbf{s}_\pi + Q^*\mathrm{Vec}_{\mathbb{R}}(N)$. It is clear from the block structure of the matrix R that after fixing the values of the last n_{g+1} variables in $\mathbf{s}_\pi$, the remaining variables can be decoded in g parallel steps, the i-th step involving n_i variables. The worst-case decoding complexity for this system is then of the order $|S|^{n_{g+1}+\max n_i}$. This is in contrast to the complexity of $|S|^{2l}$ if the matrix R has no special structure.

Proof. If X is fast decodable as per Definition 2, then as described in Remark 2, the subsets $\Gamma_1, \ldots, \Gamma_g, \Gamma_{g+1}$ provide a permutation π of $\{1, \ldots, 2l\}$, and integers $g \geq 2$, $n_1, \ldots, n_g, n_{g+1}$ with the properties described.

Definition 2 together with Theorem 1 tell us that every column of T_π indexed by elements of $\pi^{-1}(\Gamma_i)$ is orthogonal to every column indexed by $\pi^{-1}(\Gamma_j)$ $(1 \leq i < j \leq g)$. It follows immediately that on applying a QR decomposition to T_π in the order first column, then second column, etc., that the R matrix, which results from the Gram-Schmidt orthogonalizations of the columns of T_π in this order, will have the property that the columns indexed by $\pi^{-1}(\Gamma_i)$ will be perpendicular to those indexed by $\pi^{-1}(\Gamma_j)$. (This can be seen easily from how the Gram-Schmidt process works, but this can also be checked from the explicit form of the matrix R obtained from this Gram-Schmidt orthogonalization, described for instance in [1, Section III] or [7, Section VI].)

As for the other direction, assume that for all channel matrices H there is a permutation π of $\{1, \ldots, 2l\}$ and integers $g \geq 2$, and n_i $(i = 1, \ldots, g+1)$ with $n_1 + \cdots + n_{g+1} = 2l$, such that $T_\pi = QR$, where Q is unitary and R has the form as in Equation (5) above. Define the sets Γ_i in terms of the integers n_i so that $\Gamma_1 = \pi(\{1, \ldots, n_1\})$ is the image under π of the first n_1 elements $\{1, \ldots, n_1\}$, Γ_2 is the image of the next n_2 elements, and so on. It is clear from the block form of R that if $\pi(u) \in \Gamma_i$ and $\pi(v) \in \Gamma_j$ $(1 \leq i < j \leq 2l)$, then u-th and v-th columns of R are orthogonal as vectors in $\mathbb{R}^{2n^2}$. Since Q is unitary, the same holds for the matrix T_π. Equivalently,

the $\pi(u)$-th and $\pi(v)$-th columns of T are orthogonal for all H. Thus, by Theorem 1, $A_\pi(u)$ and $A_\pi(v)$ are mutually orthogonal, so X is fast decodable as per Definition 2. □

We summarize what we have shown in the next corollary:

Corollary 2. *The following are equivalent for disjoint subsets $\Gamma_i, \Gamma_j \subset \{1, \ldots, 2l\}$:*

- *for all $u \in \Gamma_i$ and $v \in \Gamma_j$ $A_u A_v^* + A_v A_u^* = 0$.*
- *for all $u \in \Gamma_i$ and $v \in \Gamma_j$, the u-th and v-th columns of $T = T(H)$ are orthogonal as real vectors for any H.*
- *there exists a permutation on the index set $\{1, \ldots, 2l\}$ so that such that the matrix R arising as in the statement of Proposition 1 has a zero block in the entries $(\pi^{-1}(\Gamma_i), \pi^{-1}(\Gamma_j))$ and $(\pi^{-1}(\Gamma_j), \pi^{-1}(\Gamma_i))$.*

Corollary 3. *The Definition 2 is equivalent to one given in [3, Definition 4].*

Definition 3. *We say that the code $X = \sum_{i=1}^{2l} s_i A_i$ is g-group decodable ($g \geq 2$) if it is fast decodable for this g (Definition 2) and if the set Γ_{g+1} in Definition 2 is empty, so the matrix R of Proposition 1 has a block-diagonal form.*

IV. Bounds on decoding complexity for full-rate codes

In this section, we will analyze the mutual orthogonality condition $A_i A_j^* + A_j A_i^* = 0$ of Theorem 1 and show that for full-rate codes, the best possible decoding complexity is not better than $|S|^{n^2+1}$, and that g-group decoding is in fact not possible for full-rate codes.

Theorem 2. *There can be at most $n^2 - 1$ $\mathbb{R}$-linearly independent matrices in $\mathcal{M}_{n\times n}(\mathbb{C})$ that are both skew-Hermitian and pairwise mutually orthogonal.*

Proof. For, suppose to the contrary that $A_1, \ldots, A_{n^2}$ were $\mathbb{R}$-linearly independent, skew-Hermitian, and mutually orthogonal. The matrix $\imath I_n$ is skew-Hermitian. Suppose first that one of these A_i, say A_1, is an $\mathbb{R}$-multiple of $\imath I_n$. This is already a contradiction, since $A_1 A_2^*$ is skew-Hermitian by the mutual orthogonality condition, but $A_1 A_2^*$ is a real multiple of $\imath A_2^*$ and is therefore Hermitian. Now suppose that no A_i is an $\mathbb{R}$-multiple of $\imath I_n$. The matrix $\imath I_n$, being skew-Hermitian, can be written as a linear combination of these matrices A_i since they form a basis for the skew-Hermitian matrices, so $\imath I_n = \sum a_j A_j$ for real a_j. Now A_1 is not a real multiple of $\imath I_n$ by assumption. Consider $\imath A_1^*$. This is Hermitian. On the other hand, $(\sum a_j A_j) A_1^* = a_1 A_1 A_1^* + (\sum a_j A_j) A_1^*$, where this second sum runs from $j = 2$ onwards. But for $j = 2$ onwards, $A_j A_1^*$ is skew-Hermitian by the mutual orthogonality condition, while both $\imath A_1^*$ and $a_1 A_1 A_1^*$ are Hermitian. For this to happen, $(\sum a_j A_j) A_1^*$, where the sum is over $j \geq 2$, must be zero, and $\imath A_1^*$ must equal $a_1 A_1 A_1^*$. On canceling A_1^* (recall our assumption that the basis matrices are invertible), we find that A_1 is a multiple of $\imath I_n$, contradiction. □

Example 1. In the 2×2 matrices $M_2(\mathbb{C})$ over the complex numbers $\mathbb{C}$, consider the three matrices $A_1 = \begin{pmatrix} \imath & 0 \\ 0 & -\imath \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and $A_3 = \begin{pmatrix} 0 & -\imath \\ -\imath & 0 \end{pmatrix}$. These three matrices are $\mathbb{R}$-linearly independent, skew-Hermitian, and pairwise mutually orthogonal matrices. Together with the identity matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, they form a $\mathbb{C}$-basis for $M_2(\mathbb{C})$, and as can be checked, no $\mathbb{C}$-linear combination of I, A_1, A_2, and A_3 is both skew-Hermitian and mutually orthogonal to A_1, A_2, and A_3. Thus, the $2^2 - 1$ matrices A_1, A_2, and A_3 exemplify the contention of this theorem.

We get a quick corollary:

Corollary 4. *For a space-time block code generated by invertible $n \times n$ matrices, the maximum number of groups g in the notation of Definition 2 is n^2.*

Proof. If the number of groups is more than n^2, then we can find $n^2 + 1$ matrices that are $\mathbb{R}$-linearly independent and mutually orthogonal. Note that if matrices X and Y are mutually orthogonal, then so are MX and MY for any matrix M, as can be easily seen. Thus, if A is one of the $n^2 + 1$ matrices, and C and D are any two of the remaining n^2 matrices, then taking $M = A^{-1}$, we find I_n and $A^{-1}C$ are mutually orthogonal, which is to say that $A^{-1}C$ is skew-Hermitian. Moreover, $A^{-1}C$ and $A^{-1}D$ are mutually orthogonal. Thus, we find n^2 skew-Hermitian and mutually orthogonal $\mathbb{R}$-linearly independent matrices. But this contradicts Theorem 2. □

Theorem 3. *If any $g - 1$ of the groups $\Gamma_1, \ldots, \Gamma_g$ from Definition 2 together have at least n^2 matrices in them, then the remaining group can only have one matrix in it and those $g - 1$ groups must then have exactly n^2 elements in them.*

Proof. Say the last $g - 1$ groups, for simplicity, have at least n^2 matrices, and suppose that the first group Γ_1 has at least two elements, call them A and B. By multiplying throughout by A^{-1}, we can assume that the two elements are I_n and B, without destroying the mutual orthogonality, as in the proof of Corollary 4 above. Note that after multiplying by A^{-1}, the matrices in the remaining groups all become skew-Hermitian, also as in the proof above. Because there are at least n^2 skew-Hermitian ($\mathbb{R}$-linearly independent) matrices, we find that there must be exactly n^2 of them because the dimension of the skew-Hermitian matrices is n^2. Call these n^2 matrices $C_1, \ldots, C_{n^2}$. We must have $\imath I_n$ in the linear span of these C_i because $\imath I_n$ is also skew-Hermitian. Thus, $\imath I_n = \sum a_i C_i$. Now multiply on the right by B^*, where B is as above. Each of the products $C_i B^*$ is skew-Hermitian because of the mutual orthogonality condition that requires $C_i B^* + B C_i^* = 0$. Thus, $\imath B^*$ is also skew-Hermitian. It follows from this that B^* is Hermitian, i.e., B is Hermitian. But now, we consider $C_i B^*$ for any i. The mutual orthogonality condition says that this is skew-Hermitian, so it equals $-(BC_i^*)$, and since C_i^* is skew-Hermitian, this equals BC_i. On the other hand, we just saw that B is Hermitian, so $C_i B^* = C_i B$. Thus, B commutes with all C_i, i.e, with all skew-Hermitian matrices. But this means B commutes with all the Hermitian matrices as well, because every Hermitian matrix is of the form $\imath$ times a skew-

Hermitian matrix. Since every matrix is a sum of a Hermitian and a skew-Hermitian matrix, B commutes with all matrices, and is Hermitian, so it must be real scalar. But this violates the $\mathbb{R}$-linear independence of I_n and B. □

Corollary 5. *If, as in the notation of Definition 2, $n_i \geq 2$ for any i, then the total number of matrices in the g groups is at most $n^2 + n_i - 1$. In particular, if $k = \min n_i \geq 2$, then the total number is at most $n^2 - 1 + k$.*

Proof. Since the i-th group has size $n_i \geq 2$, the remaining groups must have less than n^2 matrices in them, or else, the lemma above will be violated. It follows that there at most $n^2 + n_i - 1$ matrices in the g groups. □

We now come to our main result on decoding complexity.

Theorem 4. *The decoding complexity of a full-rate space-time block code $X = \sum_{i=1}^{2n^2} s_i A_i$ is not better than $|S|^{n^2+1}$, where $|S|$ is the size of the effective real constellation.*

Proof. Consider the basis matrices A_i: if there are at least two mutually orthogonal groups, then, then code is fast decodable, and Proposition 1 tells us the R matrix that comes from $T = T(H)$ will have the form as in Equation (5). Consider the integers $n_i = |\Gamma_i|$. If any $n_i \geq 2$, then by Corollary 5, the total number of matrices in the g groups is at most $n^2 + n_i - 1$. Thus, the subset Γ_{g+1} will be of size at least $(n^2 - n_i + 1)$. Having conditioned the last group of symbols, decoding the first g groups of symbols has a decoding complexity at least $|S|^{n_i}$, so we find that the decoding complexity must be at least $|S|^{n^2-n_i+1} \cdot |S|^{n_i} = |S|^{n^2+1}$. If on the other hand all $n_i = 1$, then we have g groups of size 1 each. By Corollary 4, $g \leq n^2$, so Γ_{g+1} is of size at least n^2. Thus the decoding complexity is at least $|S|^{n^2} \cdot |S| = |S|^{n^2+1}$. □

Example 2. Silver Code: This 2×2 code for four complex signal elements s_1, s_2, s_3, s_4 is given by $X(s_1, s_2) + TX(z_1, z_2)$, where for any a and b, $X(a, b) = \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix}$, and $T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The signal elements s_3 and s_4 are related to z_1 and z_2 by $(z_1, z_2)^T = M(s_3, s_4)^T$, where $M = \frac{1}{\sqrt{7}} \begin{pmatrix} 1+\imath & -1+2\imath \\ 1+2\imath & 1-\imath \end{pmatrix}$. This code has a decoding complexity of at most $|S|^5$ (see [3] for instance). This example thus shows that our bound n^2+1 is strict. Moreover, Theorem 4 shows that the Silver code cannot have a lower lattice decoding complexity than the known $|S|^5$.

Theorem 5. *It is not possible to arrange for the full-rate space-time block code $X = \sum_{i=1}^{2n^2} s_i A_i$ to have g-group decodability for any g.*

Proof. Assume the code is g-group decodable. First suppose some $n_i \geq 2$. Then by Corollary 5, the total number of matrices in the g groups, namely $2n^2$, is bounded above by $n^2 + n_i - 1$. Thus, $n_i \geq n^2 + 1$. But this violates Theorem 3, since the total number in any $g - 1$ groups can at most be n^2. Thus, all groups must have just one matrix each. But by Corollary 4, the total number of groups is at most n^2, so we find $2n^2 \leq n^2$, a contradiction. Hence g-group decodability is not possible for full-rate codes. □

V. Conclusion

We first give a proof in this paper of the necessity of the mutual orthogonality condition $A_i A_j^* + A_j A_i^* = 0$ of generators of a space-time block code for fast decoding, which allows us to propose a more intrinsic definition of fast decodability based on the generators alone. We then use this necessary orthogonality to provide strict lower bounds on the decoding complexity of space-time block codes. Indeed we prove that full-rate linear codes subject only to the constraint that their generators are invertible will unfortunately suffer from high ML decoding complexity. We also show that full-rate codes cannot be g-group decodable. This work is part of results to appear in a full length paper, which contains sharper bounds on the maximum number of groups g in codes admitting fast lattice decodability.

Acknowledgements: N. Markin was supported by the Singapore National Research Foundation under Research Grant NRF-RF2009-07. B.A. Sethuraman was supported by a U.S. National Science Foundation grant CCF-1318260. G. Berhuy and B.A. Sethuraman wish to thank Prof. Frederique Oggier and Nanyang Technological University, Singapore, for hosting their visit during which the ideas for this paper germinated.

References

[1] E. Biglieri, Y. Hong and E. Viterbo, "On fast-decodable space-time block codes", *IEEE Trans. Inform. Theory*, vol. 55, no. 2, Feb 2009.

[2] T.P. Ren, Y.L. Guan, C. Yuen, and R.J. Shen, "Fast-group-decodable space-time block code," Proceedings IEEE Workshop (ITW 2010), 2010.

[3] G.R. Jithamithra, B.S Rajan , Minimizing the Complexity of Fast Sphere Decoding of STBCs, *IEEE Transactions on Wireless Communications*, vol 12, no. 12, 2013.

[4] R. Vehkalahti, C. Hollanti, F. Oggier, "Fast-Decodable Asymmetric Space-Time Codes from Division Algebras," *IEEE Transactions on Information Theory*, vol. 58, no. 4, April 2012.

[5] N. Markin, F. Oggier, "Iterated Space-Time Code Constructions From Cyclic Algebras," Information Theory, IEEE Transactions on, vol.59, no.9, pp.5966–5979, Sept. 2013.

[6] L. Luzzi, F. Oggier, "A family of fast-decodable MIDO codes from crossed-product algebras over $\mathbb{Q}$," *Proc. IEEE Int. Symp. Inform. Theory*, St Petersburg, July 2011.

[7] K. P. Srinath, B. S. Rajan, "Low ML-decoding complexity, large coding gain, full-diversity STBCs for 2×2 and 4×2 MIMO systems," *IEEE J. on Special Topics in Signal Processing: managing complexity in multi-user MIMO systems*, 2010.

[8] K. P. Srinath, B. S. Rajan, "Generalized Silver Codes," *IEEE Trans. Inform. Theory*, vol. 57, no. 9, Sep 2011.

[9] L. P. Natarajan, B. S. Rajan, "Asymptotically-Good, Multigroup Decodable Space-Time Block Codes," *IEEE Transactions on Wireless Communications*, vol. 12, no 10, pp. 5035-5047, 2013.

[10] Chau Yuen, Yong Liang Guan, Tjeng Thiang Tjhung, "On the Search for High-Rate Quasi-Orthogonal SpaceTime Block Code," *Int. J. Wireless Inf. Network*, vol. 13, pp. 329-340, Oct. 2006.

[11] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, "Full-diversity, high-rate space-time block codes from division algebras," *IEEE Trans. on Information Theory*, vol. 49, no. 10, pp. 2596-2616, Oct. 2003.

[12] G. Berhuy and F. Oggier, An introduction to central simple algebras and their applications to wireless communication, Mathematical Surveys and Monographs, Amer. Math. Soc., vol. 191, 2013.