Golden Space–Time Block-Coded Modulation

Laura Luzzi, Ghaya Rekaya-Ben Othman, *Member, IEEE*, Jean-Claude Belfiore, *Member, IEEE*, and Emanuele Viterbo, *Senior Member, IEEE*

***Abstract*—In this paper, block-coded modulation is used to design a 2×2 multiple-input multiple-output (MIMO) space–time code for slow fading channels. The Golden Code is chosen as the inner code; the scheme is based on a set partitioning of the Golden Code using two-sided ideals whose norm is a power of two. In this case, a lower bound for the minimum determinant is given by the minimum Hamming distance. The description of the ring structure of the quotients suggests further optimization in order to improve the overall distribution of determinants. Simulation results show that the proposed schemes achieve a significant gain over the uncoded Golden Code.**

***Index Terms*—Coding gain, Golden Code, Reed–Solomon codes, space–time block codes.**

I. Introduction

THE wide diffusion of wireless communications has led to a growing demand for high-capacity, highly reliable transmission schemes over fading channels. The use of multiple transmit and receive antennas can greatly improve performance because it increases the *diversity order* of the system, defined as the number of independent transmit–receive paths. In order to exploit fully the available diversity, a new class of code designs, called Space–Time Block Codes, was developed in [1]. In the *coherent*, *block fading* model, where the channel coefficients are supposed to be known at the receiver, and remain constant for a time block, the fundamental criteria for code design are

— the *rank criterion*, stating that the difference of two distinct codewords or "space–time blocks" must be a full-rank matrix,

— the *determinant criterion*, stating that its minimum determinant ought to be maximized [1], [2].

Codes meeting these two criteria can be constructed using tools from algebraic number theory. In particular, by choosing a subset of a *division algebra* over a number field as the code, one ensures that all the nonzero codewords are invertible. If, furthermore, this subset is contained in an *order* of the algebra, the minimum determinant over all nonzero codewords will be bounded from below and will not vanish when the size of the constellation grows to infinity [3]–[5].

The construction of codes from cyclic division algebras was first introduced in [6]. In the 2×2 multiple-input multiple-output (MIMO) case, Belfiore *et al.* [7] designed the *Golden Code* $\mathcal{G}$, a full-rate, full-rank, and information-lossless code satisfying the nonvanishing determinant condition. The $n \times n$ MIMO codes that achieve these properties were called *Perfect Codes* in [8] and also studied in [9].

In this paper, we focus on the slow block-fading channel, where the fading coefficients are assumed to be constant for a certain number of time blocks T.[1] Traditionally, the design of space–time codes has focused either on the case of short block lengths, where the quasi-static interval T is approximately the same as the number of transmit antennas M [6], [11], [8], [9], [5], [12], or on the case $T \gg M$ [1], [13], [12]. Here we consider the case of moderate block lengths, that falls in between these two categories.

Even though fading hinders transmission with respect to the AWGN case, fast fading is actually beneficial because the transmission paths at different times can be regarded as independent. On the contrary, with slow fading the ergodicity assumption must be dropped, leading to a performance loss. This loss can be compensated using *coded modulation*: in a general setting, a full-rank space–time block code is used as an *inner code* to guarantee full diversity, and is combined with an *outer code* which improves the minimum determinant.

We will take as inner code the Golden Code: we focus on the problem of designing a *block code* $\{\mathbf{X} = (X_1, \ldots, X_n)\}$, where each component X_i is a Golden codeword. We will assume that the quasi-static interval has length $T = 2n$. In order to increase the minimum determinant, one can consider the ideals of $\mathcal{G}$. In [14] and [13], a *set partitioning* of the Golden Code is described; it is based on a chain of left ideals $\mathcal{G}_k = \mathcal{G}B^k$, such that the minimum determinant in $\mathcal{G}_k$ is 2^k times that of $\mathcal{G}$.

Choosing the components X_i independently in $\mathcal{G}_k$, one obtains a very simple block code. For small sizes of the signal constellation these subcodes already yield a performance gain with respect to the "uncoded" Golden Code (that is, with respect to choosing $X_i \in \mathcal{G}$ independently). However, the gain is cancelled out asymptotically by the loss of rate as the size of the signal set grows to infinity, since an energy increase is required to maintain the same spectral efficiency, or bit-rate per channel use. A better performance is achieved when the X_i are not chosen in an independent fashion. In [13], two encoders are combined: a trellis encoder whose output belongs to the quotient

Manuscript received December 17, 2007; revised: August 18, 2008. Current version published February 11, 2009. The material in this paper was presented in part at the 2008 IEEE Information Theory Workshop, Porto, Portugal, May 2008.

L. Luzzi, G. R.-B. Othman, and J.-C. Belfiore are with TELECOM ParisTech, 75013 Paris, France (e-mail: luzzi@telecom-paristech.fr; rekaya@telecom-paristech.fr; belfiore@telecom-paristech.fr).

E. Viterbo is with DEIS–Università Della Calabria, 87036 Rende (CS), Italy (e-mail: viterbo@deis.unical.it).

Communicated by B. S. Rajan, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2008.2009846

[1] This kind of behavior might be caused by large obstructions between transmitter and receiver. The model is realistic if T is smaller than the coherence time of the channel; for most practical applications, it has been estimated [10] that the coherence time is greater than 0.01 s, so that $T < 100$ is a legitimate assumption.

$\mathcal{G}_k/\mathcal{G}_{k+1}$, and a lattice encoder for $\mathcal{G}_{k+1}$ (trellis coded modulation).

The global minimum determinant for the block code is

$$\Delta_{\min} = \min_{\mathbf{X}\neq 0} \det \left(\sum_{i=1}^{n} X_i X_i^H \right).$$

This expression is difficult to handle because its "mixed terms" are Frobenius norms of products in $\mathcal{G}$. The codes described in [13] are designed to maximize the approximate parameter $\Delta'_{\min} = \min_{\mathbf{X}\neq 0} \sum_{i=1}^{n} \det\left(X_i X_i^H\right)$ and so *a priori* they might be suboptimal; we will consider the mixed terms and so obtain a tighter bound for $\Delta_{\min}$.

A rough estimate of the coding gain for the block code comes from its minimum "Hamming distance", that is, the minimum number of nonzero components. To increase the Hamming weight, we will take as our outer code an error correcting code over the quotient of $\mathcal{G}$ by one of its ideals.

The paper is organized as follows. After an Overview of the proposed encoder (Section II), we recall the algebraic construction of the Golden Code and its properties in Section III. In Section IV, we describe the general setting for Golden block codes and the coding gain estimates; in Section V, we study the two-sided ideals of $\mathcal{G}$ that are suitable for binary partitioning. In Sections VI and VII, we introduce the repetition codes and the Reed–Solomon block codes over $\mathcal{G}$ and discuss their performance obtained through simulations. The interested reader can find in the Appendix the main definitions and theorems concerning quaternion algebras that are cited in the paper.

II. Overview

The general structure of the encoder is the following (see Fig. 1).

a) The binary information message is divided into two parts. A first data block of $4Nk$ bits, where $0 < k < n$, is encoded into a block of length $4Nn$ by a linear binary code with generator matrix $G \in \mathbb{F}_2^{(4Nn\times 4Nk)}$. The second, of length $4Nn\varepsilon_0$ bits, is left uncoded. This block is optional, that is $\varepsilon_0 \in \{0,1\}$.

b) The two binary sequences are "mixed up" and rearranged into n vectors of $4N(1+\varepsilon_0)$ bits each. Each of the n binary vectors is then used to modulate four signals a, b, c, d belonging a QAM constellation of size $2^{(1+\varepsilon_0)N}$. (We only considered the case of square constellations.)

c) Finally, each of the n quadruples (a,b,c,d) is used to encode a Golden Codeword

$$X = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha(a+b\theta) & \alpha(c+d\theta) \\ \bar{\alpha} i(c+d\bar{\theta}) & \bar{\alpha}(a+b\bar{\theta}) \end{bmatrix} \tag{1}$$

where $\theta = \frac{1+\sqrt{5}}{2}$, $\bar{\theta} = 1-\theta$.

The Golden Code is full-rate and transmits two symbols per channel use; each symbol carries $(1+\varepsilon_0)N$ bits. The total information rate of the binary code is $\frac{k+\varepsilon_0 n}{n(1+\varepsilon_0)}$. The rate of the block code will be $\frac{2(k+\varepsilon_0 n)N}{n}$ bpcu.

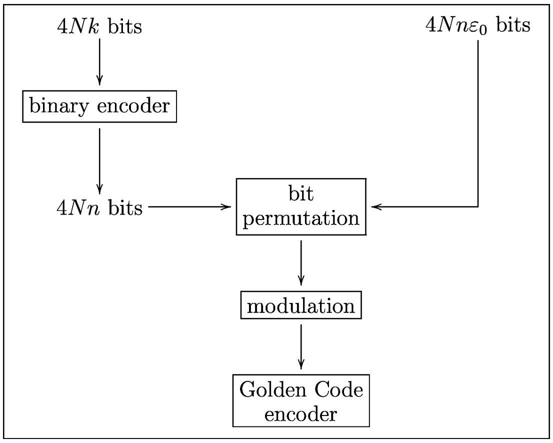

Fig. 1. The general structure of the encoder.

The encoder will be based on a partition of the Golden Code $\mathcal{G}$ of cardinality 2^{4N}:

$$\mathcal{G} = \bigcup_{i=1}^{2^{4N}} (m\mathcal{G} + c_i), \quad c_i \in \mathcal{G}.$$

Each subset is a coset of $m\mathcal{G}$, where $m = (1+i)^l \in \mathbb{Z}[i]$ is a scalar factor.

The permutation of the bits in Step b) and the labelling of the constellation are chosen in such a way that for each of the n vectors, the coded bits select one of the 2^{4N} cosets of $\mathcal{G}$ and the uncoded bits specify a point in the coset.

More precisely, we consider a partition of the QAM constellation into 2^N subsets of size $2^{\varepsilon_0 N}$ (if $\varepsilon_0 = 0$, the subsets consist of a single point). Each subset is a scaled version of a $2^{\varepsilon_0 N}$-QAM constellation. For instance, a 16-QAM constellation can be partitioned into four scaled 4-QAM constellations, as shown in Fig. 8.

Two examples of block-coded modulation will be described in detail, namely, the following.

- A "repetition code" over the cosets of $(1+i)\mathcal{G}$. In this case the scaling factor m is $1+i$, the block length n is 2, $k=1$, $N=1$, $\varepsilon_0 = 1$, so that 4-QAM constellations are used, with a partitioning into two BPSK constellations.
- A "Golden Reed–Solomon" scheme, for which the binary code is derived from an $(n,k,d_{\min})$ error-correcting code over the finite field $\mathbb{F}_{2^{4N}}$, simply by representing the elements of the finite field as polynomials of degree $4N-1$ with binary coefficients. In particular, we consider two cases for a scaling factor $m=2$:
  - —In the first case we use 4-QAM constellations, and the uncoded part is empty ($N=2, \varepsilon_0 = 0$); the partition consists of individual points.
  - —In the second, we consider 16-QAM constellations partitioned into four 4-QAM constellations ($N=2, \varepsilon_0=1$).

III. The Golden Code

Since we are interested in the partitioning of the Golden Code, we begin by recalling its algebraic construction. For the sake of simplicity, definitions and theorem statements are collected in Appendix B.

The Golden Code $\mathcal{G}$, introduced in [7], is a full rate, full-diversity, information lossless and DMT achieving code for two transmit and two or more receive antennas. This code is constructed using the cyclic division algebra $\mathcal{A} = (\mathbb{Q}(i,\theta)/\mathbb{Q}(i),\sigma,\gamma)$ over the number field $\mathbb{Q}(i,\theta)$, where $\theta = \frac{\sqrt{5}+1}{2}$ is the golden number. The set $\mathcal{A}$ is the $\mathbb{Q}(i,\theta)$-vector space $\mathbb{Q}(i,\theta) \oplus \mathbb{Q}(i,\theta)j$, where j is such that $j^2 = \gamma \in \mathbb{Q}(i)^*$, $xj = j\bar{x}\ \forall\, x \in \mathbb{Q}(i,\theta)$.

Here we denote by σ the canonical conjugacy sending an element $x = a + b\theta \in \mathbb{Q}(i,\theta)$ to $\bar{x} = a + b\bar{\theta}$, where

$$\bar{\theta} = 1 - \theta = \frac{1-\sqrt{5}}{2}, \quad \theta\bar{\theta} = -1.$$

As its degree over its center $\mathbb{Q}(i)$ is 4, $\mathcal{A}$ is also called a *quaternion algebra*. If we choose $\gamma = i$, γ is not a norm in $\mathbb{Q}(i,\theta)/\mathbb{Q}(i)$, and this implies that $\mathcal{A}$ is a division algebra (see Theorem 7 in Appendix B) [7].

$\mathbb{Q}(i,\theta)$ is a splitting field for $\mathcal{A}$ (see Theorem 8 in Appendix B), and so $\mathcal{A}$ is isomorphic to a subalgebra of $\mathcal{M}_2(\mathbb{Q}(i,\theta))$. The inclusion is given by

$$x \mapsto \begin{pmatrix} x & 0 \\ 0 & \bar{x} \end{pmatrix} \ \forall\, x \in \mathbb{Q}(i,\theta), \qquad j \mapsto \begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}. \tag{2}$$

That is, every element $X \in \mathcal{A}$ admits a matrix representation

$$X = \begin{bmatrix} x_1 & x_2 \\ i\bar{x}_2 & \bar{x}_1 \end{bmatrix}, \quad x_1, x_2 \in \mathbb{Q}(i,\theta). \tag{3}$$

The Golden Code $\mathcal{G}$ is a subring of $\mathcal{A}$ having two additional properties: the minimum determinant

$$\delta = \min_{X \neq X', X, X' \in \mathcal{G}} |\det(X - X')|^2$$

should be strictly bounded away from 0, and moreover the code is information lossless.

For the first condition, if one requires that the matrix elements of X belong to the ring of integers $\mathbb{Z}[i,\theta]$ of $\mathbb{Q}(i,\theta)$, then X belongs to the $\mathbb{Z}[i]$-*order*

$$\mathcal{O} = \left\{ \begin{bmatrix} x_1 & x_2 \\ i\bar{x}_2 & \bar{x}_1 \end{bmatrix}, x_1, x_2 \in \mathbb{Z}[i,\theta] \right\} \tag{4}$$

Since $x \in \mathbb{Z}[i,\theta]$ implies that the reduced norm $N(x) = x\bar{x}$ belongs to $\mathbb{Z}[i]$, we have $\det(X) \in \mathbb{Z}[i]$, so $|\det(X)| \geq 1$ for every $X \in \mathcal{O} \setminus \{0\}$.

Each codeword of $\mathcal{O}$ carries two symbols $x_1 = a + b\theta$, $x_2 = c + d\theta$ in $\mathbb{Z}[i,\theta]$, or equivalently four information symbols $(a,b,c,d) \in \mathbb{Z}[i]^4$: the code is *full-rate*.

In order to have an information lossless code, a right principal ideal of $\mathcal{O}$ of the form $\alpha\mathcal{O}$ was used, where $\alpha = 1 + i\bar{\theta}$: its matrix representation is

$$A = \begin{bmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{bmatrix} \in \mathcal{O}. \tag{5}$$

The Golden Code is defined as $\mathcal{G} = \frac{1}{\sqrt{5}}\alpha\mathcal{O}$. Every codeword in $\mathcal{G}$ is of the form $X = \frac{1}{\sqrt{5}}AW$, with $W \in \mathcal{O}$: see (1).

Remark 1: We have seen that $\forall\, W \in \mathcal{O} \setminus \{0\}$, $|\det(W)| \geq 1$. Consequently, $\forall\, X \in \mathcal{G} \setminus \{0\}$, $|\det(X)|^2 \geq \delta = \frac{1}{5}$.

In fact, if $X = \frac{A}{\sqrt{5}}W$, $|\det(X)| = \frac{|N(\alpha)|}{5}|\det(W)| = |\frac{\det(W)}{\sqrt{5}}|$, since $|N(\alpha)| = |2+i| = \sqrt{5}$.

The code $\mathcal{G}$ has *cubic shaping*: it is isometric to the cubic lattice $\mathbb{Z}[i]^4$ (and so it is information lossless). In fact, if we consider the linear mapping $\phi : \mathcal{A} \to \mathbb{C}^4$ that vectorizes matrices

$$\phi\left(\begin{bmatrix} a & c \\ b & d \end{bmatrix}\right) = (a,b,c,d) \in \mathbb{C}^4$$

then $\phi(\mathcal{G}) = R\mathbb{Z}[i]^4$, where R is the unitary matrix

$$R = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha & -\bar{\alpha}i & 0 & 0 \\ 0 & 0 & \bar{\alpha}i & \alpha \\ 0 & 0 & \alpha & -\bar{\alpha}i \\ \bar{\alpha} & -\alpha i & 0 & 0 \end{bmatrix}. \tag{6}$$

Even though $\mathcal{G}$ is defined (up to a scaling constant) as a right ideal, it is easy to see that actually it is a *two-sided ideal*: if $w = w_1 + w_2 j \in \mathcal{O}$, $w_1, w_2 \in \mathbb{Z}[i,\theta]$

$$\alpha(w_1 + w_2 j) = w_1\alpha + w_2 j\bar{\alpha} = (w_1 + i\theta w_2 j)\alpha$$

observing that $\alpha i\theta = i\theta + 1 = \bar{\alpha}$. But

$$\xi : w_1 + w_2 j \mapsto w_1 + i\theta w_2 j \tag{7}$$

is an homomorphism of $\mathbb{Z}[i]$-modules that maps $\mathcal{O}$ into itself bijectively, therefore $\alpha\mathcal{O} = \mathcal{O}\alpha$.

Finally, $\sqrt{5}\mathcal{G}$ is an *integral ideal* because it is contained in $\mathcal{O}$.

Remark 2: For the sake of simplicity, in this section we have described the Golden Code as an infinite code. However in a practical transmission scheme, one considers a finite subset of $\mathcal{G}$, by choosing the information symbols a, b, c, d in a QAM constellation carved from $\mathbb{Z}[i]$.

IV. Golden Block Codes

A. System Model

We consider a slow block-fading channel, where the channel coefficients remain constant during the transmission of n codewords. The transmitted signal $\mathbf{X} = (X_1, \ldots, X_n)$ will be a vector of Golden codewords in a block code $\mathcal{S} \subset \mathcal{G}^n$. The received signal is given by

$$\mathbf{Y} = H\mathbf{X} + \mathbf{W}, \quad \mathbf{X}, \mathbf{Y}, \mathbf{W} \in \mathbb{C}^{2 \times 2n} \tag{8}$$

where the entries of $H \in \mathbb{C}^{2\times 2}$ are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variance per real dimension equal to $\frac{1}{2}$, and $\mathbf{W}$ is the complex Gaussian noise with i.i.d. entries of zero mean and variance N_0. We consider the coherent case, where the channel matrix H is known at the receiver.

The pairwise error probability is bounded by [1]

$$P(\mathbf{X} \mapsto \mathbf{X}') \leq \frac{1}{\left(\sqrt{\Delta_{\min} \frac{E_{\mathcal{S}}}{N_0}}\right)^4} \tag{9}$$

In the above formula, $E_{\mathcal{S}}$ is the average energy per symbol of $\mathcal{S}$ and

$$\Delta_{\min} = \min_{\mathbf{X} \in \mathcal{S} \setminus \{0\}} \quad |\det(\mathbf{X}\mathbf{X}^H)|.$$

In order to minimize the PEP for a given SNR, we should maximize $\Delta_{\min}$. We will show that

$$|\det(\mathbf{X}\mathbf{X}^H)| \geq (w_H(\mathbf{X}))^2 \delta$$

where $w_H(\mathbf{X})$ is the number of nonzero codewords in $(X_1, \dots, X_n)$ (a sort of "Hamming weight"), and $\delta = \frac{1}{5}$ is the minimum square determinant of the Golden Code.

If we simply choose $X_1, \dots, X_n$ independently in the Golden Code, the code performance will be poor compared to the fast block fading model. We call this scheme the "uncoded Golden Code": in this case $\Delta_{\min} = \delta$, for any length n. To compare the error probability of a block code with that of the uncoded Golden Code of equal length n with the same data rate, we can employ the asymptotic coding gain defined in [13]

$$\gamma_{\text{as}} = \frac{\sqrt{\Delta_{\min}}/E_{\mathcal{S}}}{\sqrt{\Delta_{\min,U}}/E_{\mathcal{S},U}} \tag{10}$$

where $\Delta_{\min}$, $\Delta_{\min,U}$ and $E_{\mathcal{S}}$, $E_{\mathcal{S},U}$ are the minimum determinants and average constellation energies of the block code and the uncoded case, respectively.

In all the cases that we considered, the theoretical gain γ_{as} turned out to be smaller than the actual gain evidenced by computer simulations. This is not surprising, since γ_{as} is only a comparison of the dominant terms in the pairwise error probability.

B. Estimates of the Frobenius Norm

First of all, we give a more explicit expression for $\det(\mathbf{X}\mathbf{X}^H)$.

We define the quaternionic conjugacy in the algebra $\mathcal{A}$

$$X = \begin{bmatrix} x_1 & x_2 \\ i\bar{x}_2 & \bar{x}_1 \end{bmatrix} \mapsto \widetilde{X} = \begin{bmatrix} \bar{x}_1 & -x_2 \\ -i\bar{x}_2 & x_1 \end{bmatrix}.$$

Observe that $\forall\, X \in \mathcal{A}$,

$$\widetilde{X}X = \det(X)\mathbb{1} \tag{11}$$

$$\widetilde{X} + X = (x_1 + \bar{x}_1)\mathbb{1} = \operatorname{tr}(X)\mathbb{1} \tag{12}$$

$$\det(X) = \det(\widetilde{X}) \tag{13}$$

where $\mathbb{1}$ denotes the identity matrix.

Recall that the *Frobenius norm* of a matrix $M = (m_{i,j})$ is

$$||M||_F = \sqrt{\sum_{i,j} |m_{i,j}|^2}.$$

Then the following formula holds.

Lemma 1: $\forall\, \mathbf{X} = (X_1, \dots, X_n) \in \mathcal{A}^n$

$$\begin{aligned}\det(\mathbf{X}\mathbf{X}^H) &= \det\left(\sum_{i=1}^n X_i X_i^H\right) \\ &= |\det(X_1)|^2 + \cdots + |\det(X_n)|^2 + \sum_{j>i} ||\widetilde{X}_j X_i||_F^2.\end{aligned} \tag{14}$$

The proof can be found in Appendix A.

We also state some simple properties of the quaternionic conjugate and of the Frobenius norm that will be useful in the sequel.

Remark 3:

a) If $W \in \mathcal{O}$, $||w||_F^2 \in \mathbb{Z}$.

b) Let X, Y be two 2×2 complex-valued matrices. Then

$$\begin{aligned} ||X||_F^2 &\geq 2|\det(X)|, \\ ||\widetilde{X}Y||_F^2 &\geq 2|\det(X)||\det(Y)|. \end{aligned} \tag{15}$$

In particular $\forall\, W \in \mathcal{O} \setminus \{0\}$

$$||W||_F^2 \geq 2|\det(W)| \geq 2. \tag{16}$$

c) If $X_1, X_2 \in \mathcal{G} \setminus \{0\}$,

$$||\widetilde{X}_2 X_1||_F^2 \geq \frac{2}{5} = 2\delta. \tag{17}$$

From (15), it follows that the determinant is bounded from below by the squared Hamming weight.

Lemma 2: Let $\mathbf{X} = (X_1, \dots, X_n) \in \mathcal{G}^n$. Then

$$\det(\mathbf{X}\mathbf{X}^H) \geq \left(\sum_{i=1}^n |\det(X_i)|\right)^2 \geq (w_H(\mathbf{X}))^2 \delta$$

where $w_H(\mathbf{X}) = \#\{i \in \{1, \dots, n\} | X_i \neq 0\}$ is the Hamming weight of the block $\mathbf{X}$.

V. Two-Sided Ideals of $\mathcal{G}$

The choice of a good block code of length n will be based on a partition chain of ideals of the Golden Code. We would like to obtain a binary partition, which is simpler to use for coding and fully compatible with the choice of a QAM constellation: we must then use ideals whose index is a power of 2, that is, whose norm is a power of $1 + i$.

A similar construction appears in [13] and employs one-sided ideals.

We prefer to choose two-sided ideals; this is a necessary and sufficient condition for the quotient group to be also a ring (see [15, Th. 2.7]). Moreover, two-sided ideals are also invariant under involution (see Theorem 15 in Appendix C and the following Remark). This will be useful for code optimization in Section VI.

In this section we describe the structure of the two-sided ideals of $\mathcal{G}$ whose norm is a power of $1 + i$. Unfortunately, it turns out that the only two-sided ideals with this property are the trivial ones. We then study the corresponding quotient rings, which are rings of matrices over nonintegral rings.

For these constructions, we will need some notions from noncommutative algebra (see Appendix C), relating the existence of two-sided ideals to the ramification of primes over the base field.

As we have seen in Section III, $\mathcal{O} = \mathbb{Z}[i,\theta] \oplus \mathbb{Z}[i,\theta]j$ is a $\mathbb{Z}[i]$-order of $\mathcal{A}$, and $\overline{\mathcal{G}} = \sqrt{5}\mathcal{G} = \alpha\mathcal{O}$ is a two-sided principal ideal of $\mathcal{O}$.

$\sqrt{5}\mathcal{G}$ is also a prime ideal since $\sqrt{5}\mathcal{G} \cap \mathbb{Z}[i] = (2+i)$ is a prime ideal of $\mathbb{Z}[i]$ (see Theorem 13 in Appendix C).

Observe that the prime ideals $(2+i)$ and $(2-i)$ of $\mathbb{Z}[i]$ are both ramified in $\mathcal{A}$: in fact

$$(2+i) = (\alpha)^2, \text{ and } (2-i) = (\alpha')^2, \text{ where } \alpha' = 1 - i\bar{\theta}.$$

(Remark that $\alpha = i\theta\bar{\alpha}$, $\alpha' = -i\bar{\theta}\bar{\alpha}'$).

It has been shown in [3] and [5], Section V, that $\mathcal{O}$ is a maximal order and its reduced discriminant is $5\mathbb{Z}[i]$: consequently, from Proposition 11 we learn that $(2+i)$ and $(2-i)$ are the only ramified primes in $\mathcal{A}$.

Then Theorem 15 implies that the prime two-sided ideals of $\mathcal{O}$ are either of the form $p\mathcal{O}$, where p is prime in $\mathbb{Z}[i]$, or belong to $\{\alpha\mathcal{O}, \alpha'\mathcal{O}\}$.

It follows that the only two-sided ideals of $\mathcal{G}$ whose norm is a power of $1+i$ are the trivial ideals of the form $(1+i)^l\mathcal{G}$.

A. The Quotient Ring $\overline{\mathcal{G}}/(1+i)\overline{\mathcal{G}}$

In the sequel, we will denote by $\overline{\mathcal{G}}$ the integral ideal $\sqrt{5}\mathcal{G}$.

Consider the prime ideal $(1+i)\mathcal{O}$. $\overline{\mathcal{G}}$ and $(1+i)\mathcal{O}$ are *coprime* ideals, that is $\overline{\mathcal{G}} + (1+i)\mathcal{O} = \mathcal{O}$; as a consequence,

$$\overline{\mathcal{G}} \cap (1+i)\mathcal{O} = \overline{\mathcal{G}}(1+i)\mathcal{O} = (1+i)\overline{\mathcal{G}}.$$

Recall the following basic result:

Theorem 3: (Third isomorphism theorem for rings). Let I and J be ideals in a ring R. Then $\frac{I}{I \cap J} \cong \frac{I+J}{J}$.

Taking $I = \overline{\mathcal{G}}$ and $J = (1+i)\mathcal{O}$, we get

$$\frac{\overline{\mathcal{G}}}{(1+i)\overline{\mathcal{G}}} \cong \frac{\mathcal{O}}{(1+i)\mathcal{O}}. \tag{18}$$

If $\pi_{\overline{\mathcal{G}}} : \overline{\mathcal{G}} \longrightarrow \overline{\mathcal{G}}/(1+i)\overline{\mathcal{G}}$ and $\pi_{\mathcal{O}} : \mathcal{O} \longrightarrow \mathcal{O}/(1+i)\mathcal{O}$ are the canonical projections on the quotient, the ring isomorphism in (18) is simply given by $\pi_{\overline{\mathcal{G}}}(g) \mapsto \pi_{\mathcal{O}}(g)$.

Theorem 13 implies that $\mathcal{O}/(1+i)\mathcal{O}$ is a simple algebra over $\mathbb{Z}[i]/(1+i) \cong \mathbb{F}_2$. We denote the image of $x \in \mathcal{O}$ through $\pi_{\mathcal{O}}$ with $[x]$.

Lemma 4: $\mathcal{O}/(1+i)\mathcal{O}$ is isomorphic to the ring $\mathcal{M}_2(\mathbb{F}_2)$ of 2×2 matrices over $\mathbb{F}_2$.

Proof: We use the well-known lemma [15]:

Lemma 5: Let R be a ring with identity, I a proper ideal of R, M a free R-module with basis X and $\pi : M \longrightarrow M/IM$ the canonical projection. Then M/IM is a free R/I-module with basis $\pi(X)$ and $|\pi(X)| = |X|$.

We know that $\mathcal{O}/(1+i)\mathcal{O}$ is a $\mathbb{Z}[i]$-module; the lemma implies that it is also a free $\mathbb{Z}[i]/(1+i)$-module, that is a vector space over $\mathbb{F}_2$, whose basis is $\{[1], [\theta], [j], [\theta j]\}$.

We define an homomorphism of $\mathbb{F}_2$-vector spaces $\psi : \mathcal{O}/(1+i)\mathcal{O} \longrightarrow \mathcal{M}_2(\mathbb{F}_2)$ by specifying the images of the basis

$$\psi([1]) = \mathbb{1}, \quad \psi([\theta]) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$
$$\psi([j]) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \psi([\theta j]) = \psi([\theta])\psi([j]).$$

It is one-to-one since $\psi([1]), \psi([\theta]), \psi([j]), \psi([\theta j])$ are linearly independent. To prove that ψ is also a ring homomorphism, it is sufficient to verify that $\psi(w_i w_j) = \psi(w_i)\psi(w_j)$ for all pairs of basis vectors w_i, w_j. □

Recall that as a $\mathbb{Z}[i]$-lattice, $\overline{\mathcal{G}}$ is isometric to $\sqrt{5}\mathbb{Z}[i]^4$, and a canonical basis is given by $\{\alpha, \alpha\theta, \alpha j, \alpha\theta j\}$. The corresponding elements $\psi([\alpha]), \psi([\alpha\theta]), \psi([\alpha j]), \psi[\alpha\theta j])$ of $\mathcal{M}_2(\mathbb{F}_2)$ are

$$\mathbf{e}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$
$$\mathbf{e}_3 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{e}_4 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{19}$$

It is easy to check that the only invertible elements in $\mathcal{M}_2(\mathbb{F}_2)$ are

$$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_1 + \mathbf{e}_2 = \mathbb{1}, \quad \mathbf{e}_3 + \mathbf{e}_4 = \varphi(j).$$

Observe that the lifts to $\mathcal{G}$ of non-invertible elements have a higher determinant.

Remark 4: If $M \in \mathcal{M}_2(\mathbb{F}_2) \setminus \{0\}$ is noninvertible

$$\min_{X \in \mathcal{G}, \quad \pi_{\overline{\mathcal{G}}}(\sqrt{5}X) = M} |\det(X)|^2 \geq 2\delta.$$

Proof: $\pi_{\overline{\mathcal{G}}}(X)$ is noninvertible in $\overline{\mathcal{G}}/(1+i)\overline{\mathcal{G}}$ if and only if its determinant is noninvertible in $\mathbb{Z}[i]/(1+i)$, that is, $\det(X) = \widetilde{X}X \in (1+i) \setminus \{0\}$. (If $M \neq 0$, $\det(X) \neq 0$, since $\mathcal{A}$ is a division ring.)

Then $|\det(\widetilde{X}X)| = |\det(X)|^2 \geq 2\delta$. □

B. The Quotient Ring $\overline{\mathcal{G}}/2\overline{\mathcal{G}}$

Again, $\overline{\mathcal{G}}$ and $2\mathcal{O}$ are coprime and so $\overline{\mathcal{G}} + 2\mathcal{O} = \mathcal{O}$, $\overline{\mathcal{G}} \cap \mathcal{O} = 2\overline{\mathcal{G}}$; from the third isomorphism theorem for rings, $\frac{\overline{\mathcal{G}}}{2\overline{\mathcal{G}}} \cong \frac{\mathcal{O}}{2\mathcal{O}}$.

Lemma 6: $\mathcal{O}/2\mathcal{O}$ is isomorphic to the ring $\mathcal{M}_2(\mathbb{F}_2[i])$ of 2×2 matrices over the ring $\mathbb{F}_2[i]$.

Proof: First of all, Lemma 5 implies that $\mathcal{O}/2\mathcal{O}$ is a free $\mathbb{Z}[i]/2$-module, that is a free $\mathbb{F}_2[i]$-module, of dimension 4. As in the previous case, we can construct an explicit homomorphism of $\mathbb{F}_2[i]$-modules $\phi : \mathcal{O}/2\mathcal{O} \rightarrow \mathcal{M}_2(\mathbb{F}_2[i])$:

$$\phi([1]) = \mathbb{1}, \quad \phi([\theta]) = \begin{pmatrix} 1+i & 1 \\ i & i \end{pmatrix}$$
$$\phi([j]) = \begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}, \quad \phi([\theta j]) = \phi([\theta])\phi([j]).$$

One can easily check that ϕ is bijective (the images of the basis elements being linearly independent) and that it is a ring homomorphism. □

To find an explicit isomorphism between $\overline{\mathcal{G}}/2\overline{\mathcal{G}}$ and $\mathcal{M}_2(\mathbb{F}_2)$, consider the following diagram, where $\pi_{\overline{\mathcal{G}}} : \overline{\mathcal{G}} \rightarrow \overline{\mathcal{G}}/2\overline{\mathcal{G}}$ is the projection on the quotient, φ is given by the third isomorphism theorem for rings, and $\phi : \mathcal{O}/2\mathcal{O} \rightarrow \mathcal{M}_2(\mathbb{F}_2[i])$ is the mapping defined in Lemma 6

$$\overline{\mathcal{G}} \xrightarrow{\pi_{\overline{\mathcal{G}}}} \overline{\mathcal{G}}/2\overline{\mathcal{G}} \xrightarrow{\varphi} \mathcal{O}/2\mathcal{O} \xrightarrow{\phi} \mathcal{M}_2(\mathbb{F}_2[i]).$$

The basis $\{\alpha, \alpha\theta, \alpha j, \alpha\theta j\}$ of $\overline{\mathcal{G}}$ as a $\mathbb{Z}[i]$-module is also a basis of $\overline{\mathcal{G}}/2\overline{\mathcal{G}}$ as an $\mathbb{F}_2[i]$-module. The isomorphism φ is simply the composition of the inclusion $\overline{\mathcal{G}} \hookrightarrow \mathcal{O}$ and the quotient mod $2\mathcal{O}$. We can compute the images through ϕ of the basis vectors: observing that

$$\alpha = 1 + i - i\theta, \quad \alpha\theta = \theta - i$$
$$\alpha j = (1 + i - i\theta)j, \quad \alpha\theta j = (\theta - i)j$$

we get

$$\phi(\alpha) = \begin{pmatrix} 0 & i \\ 1 & i \end{pmatrix}, \quad \phi(\alpha\theta) = \begin{pmatrix} 1 & 1 \\ i & 0 \end{pmatrix} \tag{20}$$

$$\phi(\alpha j) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \phi(\alpha\theta j) = \begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}. \tag{21}$$

Also in this case, the lifts X of non-invertible elements of $\mathcal{M}_2(\mathbb{F}_2[i])$ in $\mathcal{G}$ will have non-invertible determinant, that is $|\det(X)|^2 \geq 2$.

C. The Encoder

The codes that we consider follow the general outline of Forney's *coset codes* [16], taking advantage of the decomposition $\mathcal{G} = [\mathcal{G}/I] + I$, where I is $(1+i)\mathcal{G}$ or $2\mathcal{G}$, and $[\mathcal{G}/I]$ denotes a set of coset leaders.

—An $(n, k, d_{\min})$ linear code over $(\mathbb{F}_2)^4$ or $(\mathbb{F}_2)^8$ operates on part of the information data, and these coded bits are used to select $(C_1, \ldots, C_n) \in (\mathcal{G}/I)^n$.

—The remaining information bits are left uncoded and used to select $(Z_1, \ldots, Z_n) \in I^n$.

—The corresponding block codeword is $\mathbf{X} = (c_1 + Z_1, \ldots, c_n + Z_n) \in \mathcal{G}^n$, where c_i is the coset leader of C_i.

For a coset code, $\Delta_{\min}$ is bounded by the minimum determinant of I and the minimum distance $d_{\min}$ of the linear code

$$\Delta_{\min} \geq \min\left(\min_{X \in I \setminus \{0\}} |\det(X)|^2, d_{\min}^2 \delta \right). \tag{22}$$

In fact, if $(c_1, \ldots, c_n) = \mathbf{0}$, then $\mathbf{X} \in I^n$, and for $\mathbf{X} \neq \mathbf{0}$, $\det(\mathbf{X}\mathbf{X}^H) \geq \min_{X \in I \setminus \{0\}} |\det(X)|^2$. If on the contrary $(c_1, \ldots, c_n) \neq \mathbf{0}$, there are at least $d_{\min}$ components of $\mathbf{X}$ which do not belong to I, and consequently are nonzero, and $\det(\mathbf{X}\mathbf{X}^H) \geq \delta w_H(\mathbf{X}) \geq \delta d_{\min}^2$.

So the performance of a coset code will be always limited by the minimum determinant of I, except if the code on I^n is the zero code.

If I is simply $(1+i)\mathcal{G}$ or $2\mathcal{G}$, the set of possible coordinates (a, b, c, d) for the coset leaders of I in $\mathcal{G}$ *coincides* with the $(\text{BPSK})^4$ and $(\text{4-QAM})^4$ constellations respectively. This makes it much easier to implement coset codes with high Hamming distance.

VI. The Repetition Code

Here we consider the case where $I = (1+i)\mathcal{G}$, and the linear code is simply the repetition code of length 2 over $\mathcal{G}/I$. Recall that the quotient group $\mathcal{G}/I$ is a ring because I is a two-sided ideal of $\mathcal{G}$.

If $\pi : \mathcal{G} \rightarrow \overline{\mathcal{G}}/(1+i)\overline{\mathcal{G}}$ is the projection on the quotient ring $\left(\pi(X) = \pi_{\overline{\mathcal{G}}}\left(\sqrt{5}X\right)\right)$, we define

$$\mathcal{C} = \left\{ \mathbf{X} = (X_1, X_2) \in \mathcal{G}^2 | \pi(X_1) = \pi(X_2) \right\}.$$

A. The Minimum Determinant

Recall that as we have seen in Lemma 1

$$\det(\mathbf{X}\mathbf{X}^H) = |\det(X_1)|^2 + |\det(X_2)|^2 + \|\widetilde{X}_2 X_1\|_F^2.$$

With the code $\mathcal{C}$, we have $\Delta_{\min} = 4\delta$. In fact if $(X_1, 0)$ (respectively, $(0, X_2)$) is a codeword of Hamming weight 1, clearly $\pi(X_1) = 0$ and $\det(\mathbf{X}\mathbf{X}^H) = |\det(X_1)|^2$ is greater than the minimum square determinant in $(1+i)\mathcal{G}$, which is 4δ. If on the contrary $\pi(X_1) = \pi(X_2) \neq 0$,

$$\det(\mathbf{X}\mathbf{X}^H) \geq (|\det(X_1)| + |\det(X_2)|)^2 \geq 4\delta$$

because of (15).

By choosing any bijection h of the quotient ring $\overline{\mathcal{G}}/(1+i)\overline{\mathcal{G}}$ in itself, one obtains a simple variation of the repetition scheme

$$\mathcal{C}_h = \left\{ \mathbf{X} = (X_1, X_2) \in \mathcal{G}^2 | \pi(X_2) = h(\pi(X_1)) \right\}.$$

Remark 5: A suitable choice of h can slightly improve performance. In the case of the repetition code, suppose that $\pi(X_1) = \pi(X_2) = C_i$.

—If C_i is invertible in $\mathcal{M}_2(\mathbb{F}_2)$, then $\widetilde{C}_i C_i = \det(C_i)\mathbb{1} = \mathbb{1} = \mathbf{e}_1 + \mathbf{e}_2$ in the basis (19), and so the minimum determinant of a codeword $\widetilde{X}_2 X_1 \in \pi^{-1}(\widetilde{C}_i C_i)$ is also 1, and the minimum of $\|\widetilde{X}_2 X_1\|_F^2$ is 2δ. Thus $\det(\mathbf{X}\mathbf{X}^H) \geq (1 + 1 + 2)\delta = 4\delta$.

—If on the other side C_i corresponds to a non-invertible, nonzero element in $\mathcal{M}_2(\mathbb{F}_2)$, then (see Remark 4)

$$\min_{X \in \pi^{-1}(C_i)} |\det(X)| \geq \sqrt{2\delta}$$

and $\det(\mathbf{X}\mathbf{X}^H) \geq (|\det(X_1)| + |\det(X_2)|)^2 \geq \left(2\sqrt{2\delta}\right)^2 = 8\delta$.

This remark suggests that it might be more convenient to consider a group homomorphism $h : \mathcal{M}_2(\mathbb{F}_2) \rightarrow \mathcal{M}_2(\mathbb{F}_2)$ which maps invertible elements into non-invertible elements, raising the minimum determinant to 6δ if C_i invertible, $h(C_i)$ non-invertible: $\|\widetilde{X}_2 X_1\|_F^2 \geq 2\sqrt{2}\delta$, but $\|\widetilde{X}_2 X_1\|_F^2 \in \delta\mathbb{Z}$ (see Remark 3) and so $\|\widetilde{X}_2 X_1\|_F^2 \geq 3\delta$, and $\det(\mathbf{X}\mathbf{X}^H) \geq (1 + 2 + 3)\delta = 6\delta$.

Such a function h is not difficult to define, and in the case of 4-QAM modulation, an exhaustive search on the finite lattice shows that the distribution of determinants for $\mathcal{C}_{\bar{h}}$ is indeed better.[2]

B. The Encoder

Only 4 bits are needed to select an element of $\overline{\mathcal{G}}/(1+i)\overline{\mathcal{G}} \cong \mathcal{M}_2(\mathbb{F}_2)$, while the number of bits needed to select an element in the ideal depends on the chosen modulation scheme. Using 4-QAM constellations, the two choices of an element in $(1+i)\mathcal{G}$ require 4 bits each: in total, each codeword carries 12 information bits, yielding a spectral efficiency of 3 bpcu.

Suppose that $(b_1,\ldots,b_{12})$ is the binary input.

— $(b_1,\ldots,b_4)$ are used to select the matrix $b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3 + b_4\mathbf{e}_4 \in \mathcal{M}_2(\mathbb{F}_2)$ in the basis (19). The corresponding element of $[\mathcal{G}/(1+i)\mathcal{G}]$ is $C = [b_1\alpha + b_2\alpha\theta + b_3\alpha j + b_4\alpha\theta j]$.

— $(b_5,\ldots,b_{12})$ are used to select two codewords in $(1+i)\mathcal{G}$: $X_1 = (1+i)(b_5\alpha + b_6\alpha\theta + b_7\alpha j + b_8\alpha\theta j)$, $X_2 = (1+i)(b_9\alpha + b_{10}\alpha\theta + b_{11}\alpha j + b_{12}\alpha\theta j)$.

— The final block codeword is $(C + X_1, h(C) + X_2)$.

C. Asymptotic Coding Gain

Since the minimum determinant does not change, the asymptotic coding gain estimate is the same for all choices of h.

We compare these schemes with the uncoded Golden Code at 3 bpcu, using 4-QAM constellations for the symbols a, c and BPSK constellations for the symbols b, d in each Golden codeword (see (1)). The average energy per symbol is $E_{\mathcal{S}} = 0.5(0.5 + 0.25) = 0.375$, and

$$\gamma_{\text{as}} = \frac{\sqrt{\Delta_{\min}}/E_{\mathcal{S}}}{\sqrt{\Delta_{\min,U}}/E_{\mathcal{S},U}} = \frac{2/0.5}{1/0.375} = 1.5$$

This computation gives a theoretical gain of at least $10\log_{10}(1.5)$ dB $= 1.7$ dB.

D. Simulation Results

Fig. 2 shows the performance of the codes $\mathcal{C}_{\text{Id}}$ and $\mathcal{C}_{\bar{h}}$, which gain 2.4 and 2.9 dB, respectively, over the uncoded scheme at 3 bpcu at the frame error rate of 10^{-3}, supposing that the channel is constant for two time blocks.

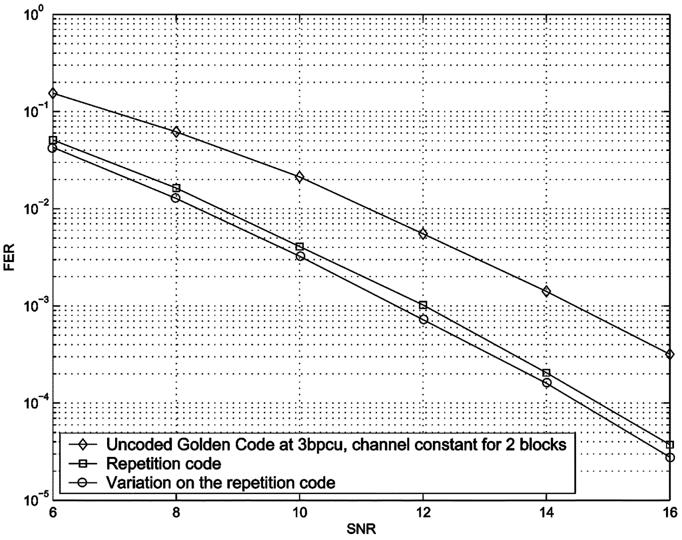

Fig. 2. Performance of the repetition code $\mathcal{C}_{\text{Id}}$ and of the variation $\mathcal{C}_{\bar{h}}$ at 3 bpcu compared with the uncoded Golden Code scheme with the same spectral efficiency. The channel is supposed to be constant for two time blocks.

VII. Golden Reed–Solomon Codes

The repetition code has the advantage of simplicity, but clearly its performance is limited by the fact that the minimum Hamming distance is only 1. To increase the Hamming distance, we need to use a more sophisticated error-correcting code.

As we have seen in the previous sections, in addition to the minimum Hamming distance, the multiplicative structure and the minimum number of noninvertible components also have a significant influence on the coding gain of a block code design. Ideally, in order to keep track of these parameters, one ought to employ error-correcting codes on $\mathcal{M}_2(\mathbb{F}_2[i])$. However, at present very little is known about codes over noncommutative rings; we choose shortened Reed–Solomon codes instead because they are maximum distance separable and their implementation is very simple; we will restrict our attention to the additive structure, defining a group isomorphism between $\overline{\mathcal{G}}/2\overline{\mathcal{G}}$ and the finite field $\mathbb{F}_{256}$.

A. The 4-QAM Case

Using 4-QAM constellations to modulate each of the four symbols a, b, c, d in a Golden codeword (1), we obtain a total of 256 codewords, one in each coset of $2\mathcal{G}$.

We consider an $(n, k, d_{\min})$ Reed–Solomon code over $\mathbb{F}_{256}$. Each quadruple (a, b, c, d) of 4-QAM signals carries 8 bits or one byte; each block of n Golden codewords will carry n bytes, corresponding to k information bytes.

The encoding procedure involves several steps.

a) Reed–Solomon encoding:
Each information byte can be seen as a binary polynomial of degree < 8, that is, an element of the Galois Field $\mathbb{F}_{256}$. An information message of k bytes, seen as a vector $\mathbf{U} = (U_1,\ldots,U_k) \in \mathbb{F}_{256}^k$, is encoded into a codeword $\mathbf{V} = (V_1,\ldots,V_n) \in \mathbb{F}_{256}^n$ using the $\text{RS}(n,k,d_{\min})$ shortened code $\mathcal{C}$. For our purposes, it is much better to use a *systematic* version of the code that preserves the first k bits of the input.

b) From the Galois field $\mathbb{F}_{256}$ to the matrix ring $\mathcal{M}_2(\mathbb{F}_2[i])$:
We can represent the elements of $\mathcal{M}_2(\mathbb{F}_2[i])$ as bytes, simply by vectorizing each matrix and separating real and imaginary parts. Since we are only working with the additive structure, we can identify $\mathbb{F}_{256}$ and $\mathcal{M}_2(\mathbb{F}_2[i])$, which are both $\mathbb{F}_2$-vector spaces of dimension 8. According to our simulation results, it seems that the choice of the linear

[2] In fact, if we define $\bar{h}(\mathbf{e}_1) = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_4$, $\bar{h}(\mathbf{e}_2) = \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$, $\bar{h}(\mathbf{e}_3) = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $\bar{h}(\mathbf{e}_4) = \mathbf{e}_1 + \mathbf{e}_3 + \mathbf{e}_4$ with respect to the basis (19), we have

$$\sum_{\mathbf{X}\in\mathcal{C}} q^{Det(\mathbf{X}\mathbf{X}^H)} = 1 + 66q^4 + 120q^8 + 48q^{10} + 202q^{16} + \cdots$$

$$\sum_{\mathbf{X}\in\mathcal{C}_{\bar{h}}} q^{Det(\mathbf{X}\mathbf{X}^H)} = 1 + 24q^4 + 61q^8 + 24q^9 + 8q^{11} + 74q^{12} + \cdots.$$

identification has very little influence on the code performance.

c) From the matrix ring $\mathcal{M}_2(\mathbb{F}_2[i])$ to the quotient ring $\overline{\mathcal{G}}/2\overline{\mathcal{G}}$:
For this step, we make use of the isomorphism of $\mathbb{F}_2[i]$-modules $(\varphi \circ \phi)^{-1} : \mathcal{M}_2(\mathbb{F}_2[i]) \to \overline{\mathcal{G}}/2\overline{\mathcal{G}}$ described in Section V-B that relates the coordinates with respect to the bases $\mathcal{B}_{\overline{\mathcal{G}}} = \{\alpha, \alpha\theta, \alpha j, \alpha\theta j\}$ and (20). Let $(a,b,c,d) \in \mathbb{Z}_2[i]^4$ be the coordinates of a codeword in the basis $\mathcal{B}_{\overline{\mathcal{G}}}$.

d) Golden Code encoding:
For each of the n vector components, the symbols $a,b,c,d \in \mathbb{Z}_2[i]$ correspond to four 4-QAM signals, and can be encoded into a Golden codeword of the form (1). Thus we have obtained a Golden block $\mathbf{X} = (X_1, X_2, \ldots, X_n) = \xi(\mathbf{V})$, where $\xi : \mathbb{F}_{256}^n \to \mathcal{G}^n$ is injective.

B. Decoding

ML decoding consists in the search for the minimum of the Euclidean distance

$$\sum_{i=1}^{n} \|HX_i - Y_i\|^2$$

over all the images $\mathbf{X} = \xi(\mathbf{V}')$ of Reed–Solomon codewords.

One can first compute and store in memory the Euclidean distances

$$d(i,j) = \left\| HX^{(j)} - Y_i \right\|^2 \tag{23}$$

for every component $i = 1, \ldots, n$ of the received vector $\mathbf{Y}$ and for all the Golden codewords $X^{(j)}$, $j = 0, \ldots, 255$ that can be obtained from a quadruple $U^{(j)}$ of 4-QAM symbols.

The search for the minimum can be carried out using the Viterbi algorithm or a tree search algorithm.

Stack Decoding: For our computer simulations, we have chosen to use a stack decoding algorithm. If the code is based on an $(n, k, d_{\min})$ Reed–Solomon code with systematic generator matrix, the $(256)^k$ codewords are the possible paths in a full tree with height k and 256 outgoing branches per node.

The decoder will store in a stack a certain number of triples $(s, \mathbf{u}, d_{\mathbf{u}})$, where $\mathbf{u}$ is an incomplete path of length s in the tree, and $d_{\mathbf{u}}$ is its distance from the initial segment $(Y_1, \ldots, Y_s)$ of $\mathbf{Y}$.

An upper bound T for the minimum distance of the received point to the lattice of Golden-RS codewords will be used as a "cost function" for the stack.

a) Sorting of distances: Before the search, for each component i, the distances $\{d(i,j)\}_{j=0,..,255}$ of (23) are sorted in increasing order: let

$$d(i, j_1(i)), d(i, j_2(i)), \ldots, d(i, j_{256}(i))$$

be the resulting sequence.

b) First step: At the beginning, the initial segments of length 1 are inserted into a previously empty stack: the triples

$$(1, j_1(0), d(0, j_1(0))), \ldots, (1, j_{256}(0), d(0, j_{256}(0)))$$

are entered in decreasing order with respect to the distance, discarding those whose distances are greater than T.

c) Intermediate steps: At each iteration of the algorithm, the triple $(s, \mathbf{u} = (j^{(1)}, \ldots, j^{(s)}), d_{\mathbf{u}})$ currently at the top of the stack is examined.

- If $s < k$, its "children" nodes

$$(s, (\mathbf{u}, r) = (j^{(1)}, \ldots, j^{(s)}, r), d_{(\mathbf{u},r)})$$
$$\text{for } r = j_1(s+1), j_2(s+1), \ldots, j_{256}(s+1)$$

are generated, updating the corresponding Euclidean distances

$$d_{(\mathbf{u},r)} = d_{\mathbf{u}} + d(s+1, r).$$

The "parent" node is deleted from the stack and the children are inserted in the stack and sorted with respect to distance, or discarded if the distance is greater than T.

(Knowing the minimum distances component-wise, one can require a stronger condition without losing optimality, namely, $d_{(u,r)} + \sum_{t=s+1}^{n} d(t, j_1(t)) < T$).

- If $s = k$, generate the Reed–Solomon codeword $\mathbf{v} = (v_1, \ldots, v_n) = G\mathbf{u}$ and store $(n, \mathbf{v}, d_{\mathbf{v}})$ in the stack (recall that $\mathbf{u}$ is an initial segment of $\mathbf{v}$), where

$$d_{\mathbf{v}} = d_{\mathbf{u}} + \sum_{t=k+1}^{n} d(t, v_t).$$

- If $s = n$, the search terminates and the initial segment of length k of $\mathbf{u}$ is the decoded message.

d) Choice of the cost function T: A simple bound for the decoder may be the distance from the received signal of the (unique) Golden-RS codeword corresponding to the "closest choice" $\left(U^{(j_0(1))}, \ldots, U^{(j_0(k))}\right)$ for the first k components. Any subset of k components may be used as well to improve the minimum provided that the corresponding lines in the Reed–Solomon generator matrix are linearly independent.

C. Simulation Results

In the 4-QAM case, the spectral efficiency of the Golden Reed–Solomon codes is given by

$$\frac{8k \text{ bits}}{2n \text{ channel uses}} = \frac{4k}{n} \text{ bpcu}.$$

From Lemma 2, it follows that using an $(n, k, d_{\min})$ Reed–Solomon code, $\Delta_{\min} \geq \delta d_{\min}^2$.

If $k = \frac{n}{2}$, the spectral efficiency is 2 bpcu. Comparing the 4-QAM, $(n, k, d_{\min})$ Golden-RS design ($E_{\mathcal{S}} = 0.5$) with the uncoded Golden Code using BPSK ($E_{\mathcal{S},U} = 0.25$), the asymptotic coding gain is

$$\gamma_{\text{as}} = \frac{\sqrt{\Delta_{\min}}/E_{\mathcal{S}}}{\sqrt{\Delta_{\min,U}}/E_{\mathcal{S},U}} = \frac{d_{\min}/0.5}{1/0.25} = \frac{d_{\min}}{2}. \tag{24}$$

Figs. 3 and 4 show the performance comparisons of the Golden-RS codes $(4, 2, 3)$ and $(6, 3, 4)$ with the corresponding uncoded schemes at the spectral efficiency of 2 bpcu.

Assuming the channel to be constant for 4 blocks and 6 blocks, respectively, the Golden-RS codes outperform the uncoded scheme by 6.1 dB and 7.0 dB.

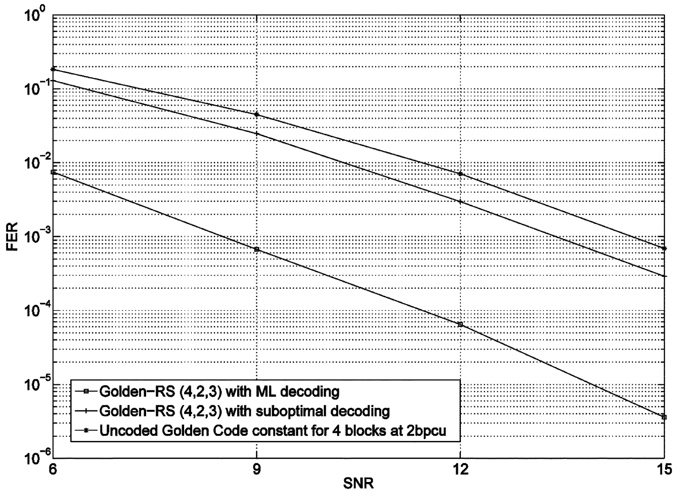

Fig. 3. Comparison between suboptimal decoding and ML decoding for the $\mathrm{RS}(4,2,3)$ code at 2 bpcu. The first method achieves a gain of only 1.1 dB over the uncoded case, compared to the 6.1 dB of the second.

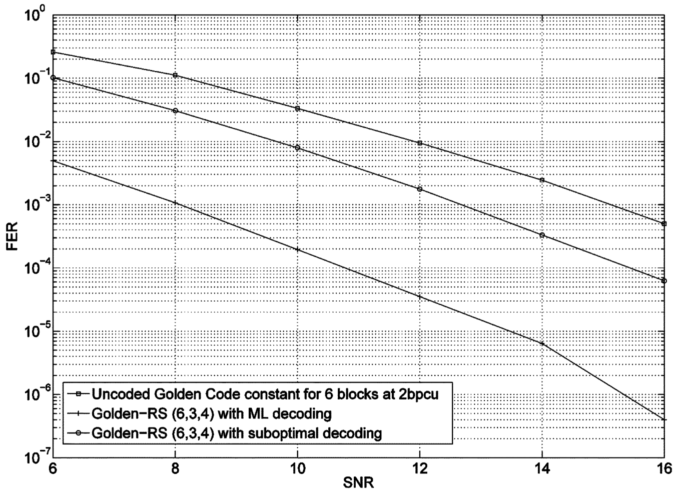

Fig. 4. Comparison between suboptimal decoding and ML decoding for the $\mathrm{RS}(6,3,4)$ code at 2 bpcu. The first method achieves a gain of 2.4 dB over the uncoded case, compared to the 7.0 dB of ML decoding.

The gain for the $(4,2,3)$ code is unexpectedly high compared with the theoretical coding gain (24) for $d=3$, that is $10\log_{10}\left(\frac{3}{2}\right)$ dB $=1.7$ dB. The rough estimate (24) is based on the worst possible occurrence, that of a codeword of Hamming weight 3 in which all three nonzero components correspond to invertible elements in the quotient.

However, we can verify empirically that in the 4-QAM case and with our choice of the $(4,2,3)$ code, this event does not take place and in fact the actual value for $\Delta_{\min}$ found by computer search is 18, giving an estimate for the gain of 3.2 dB, a little closer to the observed value.

This favorable behavior might be due to the fact that the chosen constellation contains only one point in each coset, so that the codewords of Hamming distance 3 are few.

Also for the $(6,3,4)$ code, the actual gain (7.0 dB) is higher than the theoretical gain ($10\log_{10}2$ dB $=3.0$ dB based solely on the minimum Hamming distance; 5.3 dB using the true value of $\Delta_{\min}$, that is 46.)

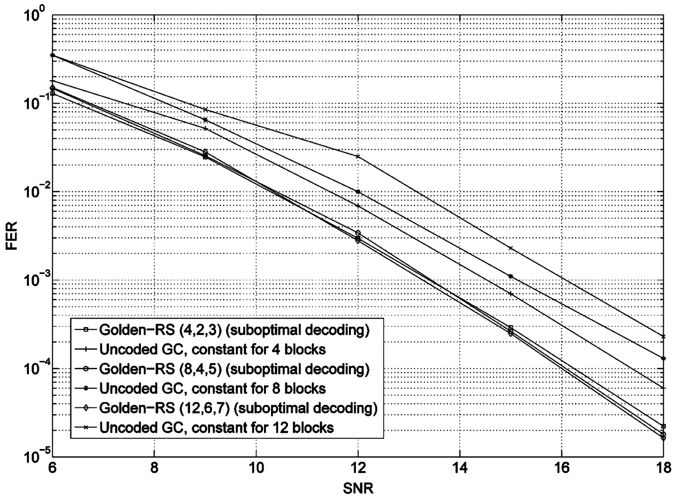

Fig. 5. Performance of $(4,2,3)$, $(8,4,5)$, and $(12,6,7)$ Golden Reed–Solomon codes with suboptimal decoding at 2 bpcu compared to the uncoded Golden Code scheme with the same spectral efficiency.

D. Suboptimal Decoding

One can replace maximum likelihood (ML) decoding with n separate Sphere Decoders on each of the n components of $\mathbf{Y}$. The signal is then demodulated, and mapped to a vector $(\hat{V}_1,\ldots,\hat{V}_n)$ in $\mathbb{F}_{256}^n$ using the inverse mappings of Steps c) and b) in Section VII-A. The received sequence $(\hat{V}_1,\ldots,\hat{V}_n)$ does not necessarily belong to the RS code, so a final step of RS decoding is needed. This "hard" decoding has the advantage of speed and allows to use longer Reed–Solomon codes with high minimum distance. However it is highly suboptimal; performance simulations show that with this method the coding gain is almost entirely cancelled out (see Fig. 3).

Suboptimal decoding also provides a good initial bound of the distance of the received point from the lattice, which can be used as a cost function for the stack decoder described in Section VII-B.

- **2 bpcu:** Fig. 5 shows the performance comparison of the Golden-RS codes with suboptimal decoding with the uncoded scheme at the spectral efficiency of 2 bpcu.

  Assuming the channel to be constant for 4, 8, and 12 blocks, respectively, the $(4,2,3)$, $(8,4,5)$ and $(12,6,7)$ Golden-RS codes outperform the uncoded scheme at the same spectral efficiency by 1.1, 1.7, and 2.8 dB at the FER of 10^{-3}.

  The Golden-RS schemes seem to be more robust on slow fading channels; in fact the performances of the Golden-$\mathrm{RS}(n,k,d_{\min})$ codes on a channel which is constant for n blocks remain almost unchanged (the variation is less than 0.2 dB) when n varies between 4 and 12, while the uncoded Golden Code has a loss of almost 1.5 dB.
- **3 bpcu:** Assuming the channel to be constant for 8, 16, and 24 blocks, respectively, the $(8,6,3)$, $(16,12,5)$ and $(24,18,7)$ Golden-RS codes gain 1.5, 2.2, and 2.8 dB over the uncoded scheme at the FER of 10^{-3} (see Fig. 6).

  Similarly to the previous case, the Golden-$\mathrm{RS}(n,k,d_{\min})$ codes lose less than 0.3 dB when n varies between 8 and 24, while the Golden Code has a loss of 1.1 dB.

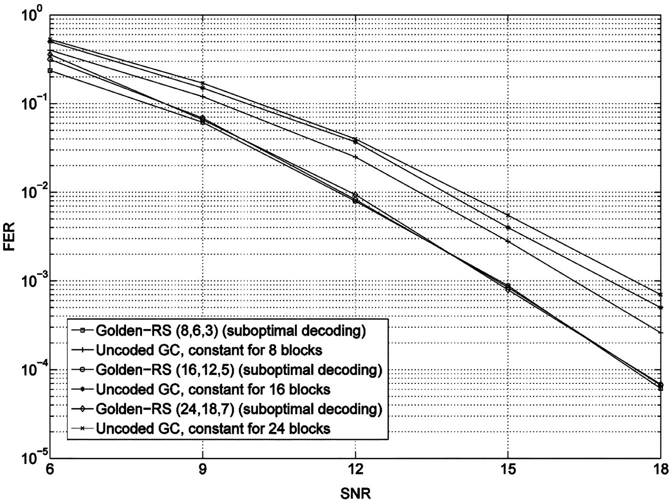

Fig. 6. Performance of $(8,6,3)$, $(16,12,5)$, and $(24,18,7)$ Golden Reed–Solomon codes with suboptimal decoding at 3 bpcu compared to the uncoded Golden Code scheme with the same spectral efficiency.

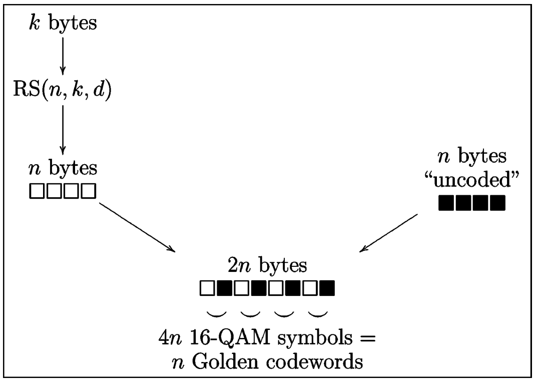

Fig. 7. The output of the Reed–Solomon code and the uncoded bits are "mingled" before modulation.

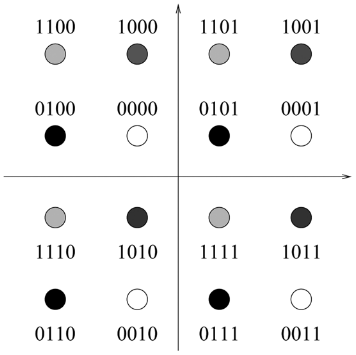

Fig. 8. The labelling of the 16-QAM constellation used for performance simulations. The first and second bit identify one of the four cosets of $2\mathbb{Z}[i]$ in $\mathbb{Z}[i]$ (drawn in different shades of gray); the third and fourth bit identify one of the four points in the coset. We remark that this type of labeling cannot be a Gray mapping.

E. The 16-QAM Case

Using 16-QAM modulation for each symbol a, b, c, d in a Golden codeword, there are 2^{16} available Golden codewords, or 256 words for each of the 256 cosets of $2\mathcal{G}$ in $\mathcal{G}$.

As in the 4-QAM case, we consider coset codes where the outer code is an $(n,k,d_{\min})$ Reed–Solomon code $\mathcal{C}$ on the quotient $\mathcal{G}/2\mathcal{G}$. Intuitively, the minimum distance of the Reed–Solomon code "protects" the cosets from being decoded wrongly; if this choice is correct, the estimate for the right point in the coset is protected by the minimum determinant in $2\mathcal{G}$.

The total information bits transmitted are $8k+8n$; they will be encoded into $8n+8n=16n$ bits.

— The code $\mathcal{C}$ outputs $8n$ bits, which are used to encode the first two bits of $4n$ 16-QAM constellations, that is the bits which identify one of the four cosets of $2\mathbb{Z}[i]$ in $\mathbb{Z}[i]$; each byte corresponds to a different coset configuration of (a,b,c,d) (see Fig. 8).

— the other $8n$ bits, left uncoded, are used to choose the last two bits of each 16-QAM signal.

In total, we have $4n$ 16-QAM symbols, that is a vector of n Golden codewords $\mathbf{X}=(X_1,\ldots,X_n)$. The resulting spectral efficiency is

$$\frac{8(k+n)\text{ bits}}{2n\text{ channel uses}}=\frac{4(k+n)}{n}\text{ bpcu.}$$

In this case, the coding gain depends on the minimum determinant of the ideal in addition to the minimum Hamming distance in the quotient: we have seen in (22) that

$$\begin{aligned}\Delta_{\min}&\geq\min\left(\min_{X\in 2\mathcal{G}\setminus\{0\}}|\det(X)|^2,d_{\min}^2\delta\right)\\&=\min\left(16\delta,d_{\min}^2\delta\right).\end{aligned}$$

With an error-correcting code of rate $k=\frac{n}{2}$, the spectral efficiency is 6 bpcu.

— If $d_{\min}\geq 4$, $\gamma_{\text{as}}=\frac{4/2.5}{1/1.5}=2.4$, leading to an approximate gain of 3.8 dB. Thus, it does not seem worthwhile to use long codes with a high minimum distance with this scheme.

— If $d_{\min}=3$, $\gamma_{\text{as}}=\frac{3/2.5}{1/1.5}=1.8$, making for a gain of 2.5 dB.

F. Decoding

The ML decoding procedure for the 16-QAM case requires only a slight modification with respect the 4-QAM case illustrated in Section VII-B. In the first phase, for each component $i=1,\ldots,n$ and for each coset leader W_j, $j=0,\ldots,255$, we find the closest point in that coset to the received component Y_i, that is

$$\hat{X}_{i,j}=\operatorname*{argmin}_{X\in 2\mathcal{G}}\|Y_i-H(X+W_j)\|^2.$$

Computing HX and HW_j separately allows to perform only 512 products instead of 256^2. The second phase can be performed as in the 4-QAM case, and the search is limited to the "closest points" $\hat{X}_{i,j}+W_j$ determined in the previous phase, i.e.,

$$\hat{\mathbf{X}}=\operatorname*{argmin}_{(\hat{X}_{1,j_1}+W_{j_1},\ldots,\hat{X}_{n,j_n}+W_{j_n})}\sum_{i=1}^{n}\|H(\hat{X}_{i,j_i}+W_{j_i})-Y_i\|^2$$

over all the images $(W_{j_1},\ldots,W_{j_n})$ of Reed–Solomon codewords.

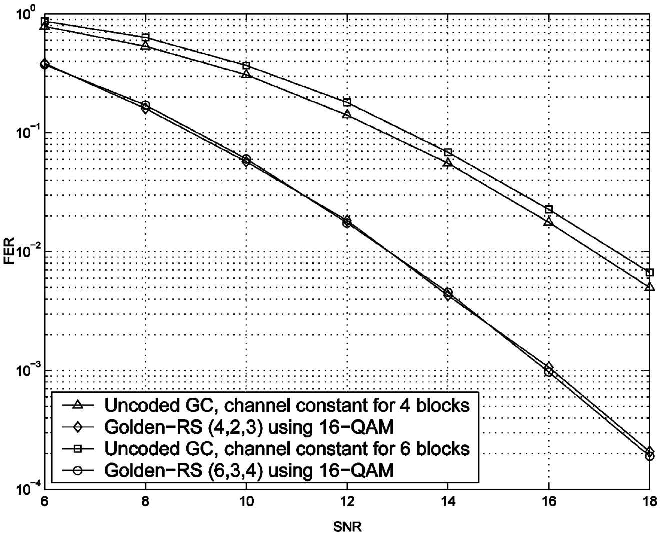

Fig. 9. Performance of the $(4,2,3)$ and $(6,3,4)$ Golden Reed–Solomon codes with ML decoding at 6 bpcu compared to the uncoded schemes with the same spectral efficiency.

G. Simulation Results

In the 16-QAM case, the $(4,2,3)$ and $(6,3,4)$ Golden Reed–Solomon codes achieve a gain of 3.9 and 4.3 dB, respectively, over the uncoded scheme at 6 bpcu at the frame error rate of 10^{-2}, supposing that the channel is constant for 4 and 6 time blocks (see Fig. 9).

VIII. Conclusions and Perspectives

In this paper, we have presented Golden-RS codes, a coded modulation scheme for 2×2 slow-fading MIMO channels, where the inner code is the Golden Code.

We consider a simple binary partitioning based on a two-sided ideal of the Golden Code, whose set of coset leaders coincides with a QAM symbol constellation. With a Reed–Solomon code as the outer code in order to increase the minimum Hamming distance among the codewords, we obtain a significant performance gain with respect to the uncoded case.

Future work will deal with exploiting the ring structure of the quotient to improve the overall distribution of determinants, instead of focusing only on the minimum determinant. Our coded modulation approach could also be used to improve the performance of the 4×4 and 6×6 Perfect Codes in [8] in the slow fading case.

Appendix A
Proofs

We report here some of the proofs for the results stated in the main part of the paper.

Proof of Lemma 1: For all $i = 1, \dots, n$, let $Q_i = X_i X_i^H$: then

$$\begin{aligned}\det\left(X_1X_1^H + \cdots + X_nX_n^H\right)\mathbb{1} &= \det(Q_1 + \cdots + Q_n)\mathbb{1}\\ &= (\widetilde{Q}_1 + \cdots + \widetilde{Q}_n)(Q_1 + \cdots + Q_n)\mathbb{1}\\ &= \sum_{i,j=1}^{n} \widetilde{Q}_i Q_j = \sum_{i=1}^{n} \det(Q_i)\mathbb{1} + \sum_{i \neq j} \widetilde{Q}_i Q_j.\end{aligned}$$

We need to show that $\widetilde{Q}_iQ_j + \widetilde{Q}_jQ_i = \|\widetilde{X}_jX_i\|_F^2\mathbb{1}$.

But $\|X\|_F^2 = \operatorname{tr}(XX^H)$, and therefore $\|\widetilde{X}_jX_i\|_F^2 = \operatorname{tr}\left(\widetilde{X}_jX_iX_i^H\widetilde{X}_j^H\right)$, and

$$\begin{aligned}\widetilde{Q}_jQ_i &= \widetilde{X}_j^H\widetilde{X}_jX_iX_i^H, \quad \widetilde{Q}_iQ_j = \widetilde{\widetilde{Q}_jQ_i}\\ &\Rightarrow \widetilde{Q}_iQ_j + \widetilde{Q}_jQ_i = \operatorname{tr}(\widetilde{Q}_iQ_j)\mathbb{1}\\ &= \operatorname{tr}\left(\widetilde{X}_jX_iX_i^H\widetilde{X}_j^H\right)\mathbb{1},\end{aligned}$$

recalling that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. □

Proof of Remark 3:

a) Let

$$W = \begin{bmatrix} w_1 & w_2 \\ i\overline{w}_2 & \overline{w}_1 \end{bmatrix}$$

where $w_1, w_2 \in \mathbb{Z}[i,\theta]$. Then $\|W\|_F^2 = |w_1|^2 + |\overline{w}_1|^2 + |w_2|^2 + |\overline{w}_2|^2$. But $w_1 = a + b\theta + i(c + d\theta)$ for some $a, b, c, d \in \mathbb{Z}$, and

$$\begin{aligned}&|w_1|^2 + |\overline{w}_1|^2\\ &\quad = (a + b\theta)^2 + (c + d\theta)^2 + (a + b\bar{\theta})^2 + (c + d\bar{\theta})^2\\ &\quad = 2a^2 + 3b^2 + 2ab + 2c^2 + 3d^2 + 2cd \in \mathbb{Z}.\end{aligned}$$

The same is true for $|w_2|^2 + |\overline{w}_2|^2$.

b) If $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$\begin{aligned}\|X\| &= |a|^2 + |b|^2 + |c|^2 + |d|^2 \geq 2(|ad| + |bc|)\\ &\geq 2|ad - bc| = 2|\det(X)|\end{aligned}$$

and

$$\|\widetilde{X}Y\|_F^2 \geq 2|\det(\widetilde{X}Y)| = 2|\det(X)\det(Y)|.$$

c) Let $X_1 = \frac{1}{\sqrt{5}}AW_1, X_2 = \frac{1}{\sqrt{5}}AW_2, W_1, W_2 \in \mathcal{O}$. Then

$$\begin{aligned}\|\widetilde{X}_2X_1\|_F^2 &= \frac{1}{25}\|\widetilde{W}_2\widetilde{A}AW_1\|_F^2\\ &= \frac{|N(\alpha)|^2}{25}\|\widetilde{W_2}W_1\|_F^2 = \frac{1}{5}\|\widetilde{W_2}W_1\|_F^2 \geq \frac{2}{5}\end{aligned}$$

since $W = \widetilde{W_2}W_1$ belongs to $\mathcal{O}$. □

Appendix B
Quaternion Algebras

This section summarizes some basic facts about quaternion algebras that are used in the paper. Our main references are the books of Vignéras [17] and Reiner [18].

Definition 1 (Quaternion Algebras): Let K be a field. A *quaternion algebra* $\mathcal{H}$ of center K is a central simple algebra of dimension 4 over K, such that there exists a separable quadratic extension L of K, and an element $\gamma \in K^*$, with

$$\mathcal{H} = L \oplus Le, \quad e^2 = \gamma, \quad ex = \sigma(x)e \quad \forall\, x \in L$$

where σ is the non-trivial K-automorphism of L. L is called a *maximal subfield* of $\mathcal{H}$. $\mathcal{H}$ will be denoted by the triple $(L/K, \sigma, \gamma)$.

Quaternion algebras are a special case of *cyclic algebras*. To obtain a representation of $\mathcal{H}$ as a K-module, consider a primitive element i such that $L = K(i)$, and let $j = e$, $k = ij = j\sigma(i)$. Then

$$\mathcal{H} = \{a + bi + cj + dk | a, b, c, d \in K\}. \qquad (25)$$

The following theorem, which follows from a more general result about cyclic algebras ([18, Corollary 30.7]) gives a sufficient condition for a quaternion algebra to be a division ring.

Theorem 7: Let $\mathcal{H} = (L/K, \sigma, \gamma)$ be a quaternion algebra. If γ is not a reduced norm of any element of L, then $\mathcal{H}$ is a skew field.

Definition 2 (Splitting Fields): Let $\mathcal{H}$ be a central simple K-algebra. An extension field E of K *splits* $\mathcal{H}$, or is a *splitting field* for $\mathcal{H}$, if

$$E \otimes_K \mathcal{H} \cong M_r(E).$$

In the case of division algebras, every maximal subfield is a splitting field ([18, Th. 7.15]):

Theorem 8: Let $\mathcal{D}$ be a skew field with center K, with finite degree over K. Then every maximal subfield E of $\mathcal{D}$ contains K, and is a splitting field for $\mathcal{D}$.

In the following paragraphs we will always consider a Dedekind domain R, its quotient field K, and a quaternion algebra $\mathcal{H}$ over K.

Definition 3 (Lattices and Orders): A *full R-lattice* or *ideal* in $\mathcal{H}$ is a finitely generated R-submodule I in $\mathcal{H}$ such that $KI = \mathcal{H}$, where

$$KI = \left\{ \sum_{i=1}^{n} k_i x_i \middle| k_i \in K, x_i \in I, n \in \mathbb{N} \right\}.$$

An *R-order* Θ in $\mathcal{H}$ is a full R-lattice which is also a subring of $\mathcal{H}$ with the same unity element. A *maximal R-order* is an order which is not properly contained in any other order of $\mathcal{H}$.

The following proposition is a consequence of [18, Th. 10.3].

Proposition 9: A subring of $\mathcal{H}$ containing a basis for $\mathcal{H}$ over K is an order if and only if all its elements are integral over R.

Remark 6: The notion of order is a generalization of the notion of the ring of integers for commutative extensions. However, in the noncommutative case the set of elements which are integral over the base field might not be a ring.

Definition 4 (Properties of Ideals): Given an ideal I of $\mathcal{H}$, we can define the *left order* and the *right order* of I as follows:

$$\Theta_l(I) = \{x \in \mathcal{H} | Ix \subset I\},$$
$$\Theta_r(I) = \{x \in \mathcal{H} | xI \subset I\}$$

$\Theta_l(I)$ and $\Theta_r(I)$ are orders. I is called

- *two-sided* if $\Theta_l(I) = \Theta_r(I)$,
- *normal* if $\Theta_l(I)$ and $\Theta_r(I)$ are maximal,
- *integral* if $I \subset \Theta_l(I)$, $I \subset \Theta_r(I)$,
- *principal* if $I = \Theta_l(I)x = x\Theta_r(I)$ for some $x \in \mathcal{H}$

The *inverse* of I is the fractional ideal $I^{-1} = \{x \in \mathcal{H} | IxI \subset I\}$.

The *norm* $N(I)$ of an ideal I is the set of reduced norms of its elements, and it is an ideal of R. If $I = \Theta x$ is principal, $N(I) = RN(x)$.

Appendix C
Ideals, Valuations and Maximal Orders

Definition 5 (Valuations and Local Fields): A *valuation* v of K is a positive real function of K such that $\forall\, k, h \in K$,

a) $v(k) = 0 \Leftrightarrow k = 0$,
b) $v(kh) = v(k)v(h)$,
c) $v(k + h) \leq v(k) + v(h)$.

v is *non-Archimedean* if $v(k + h) \leq \max(v(k), v(h))\ \forall\, k, h \in K$; it is *discrete* if $v(K^*)$ is an infinite cyclic group.

K can be endowed with a topology induced by v in the following way: a neighborhood basis of a point k is given by the sets

$$U_\varepsilon(k) = \{h \in K | v(h - k) < \varepsilon\}$$

K will be called *complete* if it is complete with respect to this topology.

If v is non-Archimedean, the set

$$R_v = \{k \in K | v(k) \leq 1\}$$

is a local ring, called the *valuation ring* of v. The quotient R_v/P_v, where P_v is the unique maximal ideal of R_v, is called the field of residues of K.

K is a *local field* if it is complete with respect to a discrete valuation v and if R_v/P_v is finite.

Definition 6 (Places): A *place* v of K is an immersion $i_v : K \to K_v$ into a local field K_v. If v is non-archimedean, we say that it is a *finite place*; otherwise, that it is an *infinite place*.

The finite places of K arise from discrete P-adic valuations of K, where P ranges over the maximal ideals in the ring of integers R of K. (Recall that the ring of integers in a number field is always a Dedekind domain, and so the maximal ideals coincide with the prime ideals).

In the case of infinite places P, the P-adic completion can be $\mathbb{R}$ (*real primes*) or $\mathbb{C}$ (*complex primes*).

The notion of ramification for quaternion algebras is a generalization of the notion of ramification for field extensions.

Definition 7 (Ramified Places): Let $\mathcal{H}$ be a quaternion algebra over K, and P a place of K.

Consider the K-module $\mathcal{H}_P = \mathcal{H} \otimes_K K_P$; $\mathcal{H}_P$ is isomorphic to a matrix algebra $M_r(D)$ over a skew field D of center K_P and index m_P over K_P; m_P is called the *local index* of $\mathcal{H}$ at P. We say that P is *ramified* in $\mathcal{H}$ if $m_P > 1$.

Complex primes are never ramified [18].

Given a maximal order Θ, the set $\mathrm{Ram}(\mathcal{H})$ of ramified places of $\mathcal{H}$ is related to a particular two-sided ideal of Θ:

Definition 8 (Different and Discriminant): Let Θ be an order. The set

$$\Theta^* = \{x \in \mathcal{H} | tr(x\Theta) \subset R\}$$

is a two-sided ideal, called the *dual* of Θ. Its inverse $\mathfrak{D} = (\Theta^*)^{-1}$ is a two-sided integral ideal, called the *different* of Θ. If $\{w_1, \dots, w_4\}$ is a basis of Θ as a free R-module,

$$(n(\mathfrak{D}))^2 = R \det(\mathrm{tr}(w_i w_j)).$$

The ideal $n(\mathfrak{D})$ of R is called the *reduced discriminant* of Θ and is denoted by $d(\Theta)$.

Proposition 10: If Θ, Θ' are two orders and $\Theta' \subsetneq \Theta$, then $d(\Theta') \subsetneq d(\Theta)$.

For the following results see [17, Corollaire 5.3 and p. 86], respectively.

Proposition 11: Let $\mathcal{H}$ be a quaternion algebra unramified at infinity.

A necessary and sufficient condition for an order Θ to be maximal is that

$$d(\Theta) = \prod_{P \in \mathrm{Ram}(\mathcal{H}) \setminus \infty} P.$$

Definition 9 (Prime Ideals): Let Θ be an order, $\mathfrak{P}$ a two-sided ideal of Θ. $\mathfrak{P}$ is *prime* if it is nonzero and $\forall\, I, J$ integer two-sided ideals of Θ, $IJ \subset \mathfrak{P} \Rightarrow I \subset \mathfrak{P}$ or $J \subset \mathfrak{P}$.

The following theorems are a consequence of more general results for central simple algebras ([18, Ths. 22.3, 22.4, 22.10, 25.7]).

Theorem 12: The two-sided ideals of an order Θ form a free group generated by the prime ideals.

Theorem 13: Let Θ be a maximal order in a quaternion algebra $\mathcal{H}$. Then the prime ideals of Θ coincide with the maximal two-sided ideals of Θ, and there is a one-to-one correspondence between the prime ideals $\mathfrak{P}$ in $\mathcal{H}$ and the prime ideals P of R, given by $P = R \cap \mathfrak{P}$.

Moreover, $\Theta/\mathfrak{P}$ is a simple algebra over the finite field R/P.

Theorem 14: Let Θ be a maximal order in $\mathcal{H}$. For each place P of K, let m_P be the local index of $\mathcal{H}$ at P, and let $\mathfrak{P}$ be the prime ideal of Θ corresponding to P (see Theorem 13). Then $m_P > 1$ only for a finite number of places P, and

$$\begin{aligned} P\Theta &= \mathfrak{P}^{m_P}, \\ \mathfrak{D} &= \prod_{P \in \mathrm{Ram}(\mathcal{H})} \mathfrak{P}^{m_P - 1} \end{aligned} \tag{26}$$

Theorem 15: The two-sided ideals of a maximal order Θ form a commutative group with respect to multiplication, which is generated by the ideals of R and the ideals of reduced norm P, where P varies over the prime ideals of R that are ramified in $\mathcal{H}$.

Remark 7: For any prime ideal $\mathfrak{P}$ of Θ, let $\widetilde{\mathfrak{P}} = \{\widetilde{x} | x \in \mathfrak{P}\}$. Since $N(\mathfrak{P}) = N(\widetilde{\mathfrak{P}})$, $\mathfrak{P}$ and $\widetilde{\mathfrak{P}}$ divide the same prime $P\Theta$.

If $\mathfrak{P}$ is prime, from (26) we have $\mathfrak{P} = \widetilde{\mathfrak{P}}$. Since two-sided ideals can be decomposed into ideals of R and prime two-sided ideals, they are invariant under involution.

References

[1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, 1998.

[2] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE VTC '96*, pp. 136–140.

[3] C. Hollanti and J. Lahtonen, "A new tool: Constructing STBCs from maximal orders in central simple algebras," in *Proc. Inf. Theory Workshop*, 2006, pp. 322–326.

[4] C. Hollanti, J. Lahtonen, and H. F. Lu, "Maximal orders in the design of dense space-time lattice codes," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4493–4510, Oct. 2008.

[5] C. Hollanti, J. Lahtonen, K. Ranto, and R. Vehkalahti, "On the densest MIMO lattices from cyclic division algebras," *IEEE Trans. Inf. Theory*, Dec. 2006, submitted for publication.

[6] B. A. Sethuraman, B. S. Rajan, and V. Shashidar, "Full-diversity, high-rate space-time block codes from division algebras," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2596–2616, 2003.

[7] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The Golden Code: A 2 × 2 full-rate space-time code with non-vanishing determinants," *IEEE Trans. Inf. Theory*, vol. 51, 2005.

[8] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space-time blocks codes," *IEEE Trans. Inf. Theory*, vol. 52, 2006.

[9] P. Elia, B. A. Sethuraman, and P. V. Kumar, "Perfect space-time codes with minimum and non-minimum delay for any number of transmit antennas," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3853–3868, 2007.

[10] S. Benedetto and E. Biglieri, *Principles of Digital Transmission With Wireless Applications*. New York: Kluwer, 1999.

[11] P. Elia, K. R. Kumar, S. A. Pawar, P. V. Kumar, and H.-F. Lu, "Explicit, minimum delay space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3869–3884, 2006.

[12] K. R. Kumar and G. Caire, "Space-time codes from structured lattices," *IEEE Trans. Inf. Theory*, Apr. 2008, submitted for publication.

[13] Y. Hong, E. Viterbo, and J.-C. Belfiore, "Golden space-time Trellis coded modulation," *IEEE Trans. Inf. Theory*, vol. 53, 2007.

[14] D. Champion, J.-C. Belfiore, G. Rekaya, and E. Viterbo, "Partitionning the Golden Code: A framework to the design of space-time coded modulation," presented at the Canadian Workshop on Information Theory, 2005.

[15] T. W. Hungerford, *Algebra*. New York: Springer-Verlag, 1974.

[16] G. D. Forney, "Coset codes- Part I: Introduction and geometrical classification," *IEEE Trans. Inf. Theory*, vol. 34, 1988.

[17] M.-F. Vignéras, "Arithmétique des algèbres de quaternions," in *Lecture Notes in Mathematics*. New York: Springer Verlag, 1980.

[18] I. Reiner, *Maximal Orders*. Oxford, U.K.: Clarendon Press, 2003.

Laura Luzzi was born in Pavia, Italy, in 1980. She received the degree (Laurea) in mathematics from the University of Pisa, Italy, in 2003 and the Ph.D. degree in mathematics for technology and industrial applications from Scuola Normale Superiore, Pisa, Italy, in 2007.

She is now a Postdoctoral Fellow at the Department of Communications and Electronics at TELECOM ParisTech, Paris, France. Her research interests include algebraic space–time coding and decoding and continued fractions.

Ghaya Rekaya-Ben Othman (M'08) was born in Tunis, Tunisia, in 1977. She received the degree in electrical engineering from ENIT, Tunisia, in 2000 and the Ph.D. degree from the Ecole Nationale Supérieure des Télécommunications (ENST) Paris, France, in 2004.

Since 2005, she has been with the Department of Communications and Electronics, TELECOM ParisTech, Paris, France, as Assistant Professor. Her research interests are in space–time coding and lattice reduction and decoding.

Jean-Claude Belfiore (M'90) received the Diplôme d'ingénieur (Eng. degree) from Supelec in 1985, the doctorat (Ph.D. degree) from Ecole Nationale Supérieure des Télécommunications (ENST Paris) in 1989 and the "Habilitationà diriger des Recherches" (HdR) from Université Pierre et Marie Curie

(UPMC) in 2001. He was enrolled in Alcatel in 1985 and then in 1989 in ENST, Paris where he is now a full professor.

He is carrying out research at the "Laboratoire de Traitement et Communication de l'Information" (LTCI), joint research laboratory between ENST and the "Centre National de la Recherche Scientifique" (CNRS), UMR 5141, where he is in charge of research activities in the fields of wireless communications, coding for wireless networks and space–time coding. He has made pioneering contributions on signal design for wireless communication systems, space–time coding, cooperative and multiuser communications.

Emanuele Viterbo (M'95–SM'05) was born in Torino, Italy, in 1966. He received the degree (Laurea) in electrical engineering in 1989 and the Ph.D. degree in electrical engineering in 1995, both from the Politecnico di Torino, Torino, Italy.

From 1990 to 1992, he was with the European Patent Office, The Hague, The Netherlands, as a patent examiner in the field of dynamic recording and error-control coding. Between 1995 and 1997, he held a postdoctoral position in the Dipartimento di Elettronica of the Politecnico di Torino in Communications Techniques over Fading Channels. He became Associate Professor at Politecnico di Torino, Dipartimento di Elettronica in 2005, and since November 2006, he has been a Full Professor in Dipartimento di Elettronica, Informatica e Sistemistica (DEIS) at Università della Calabria, Italy. In 1993, he was Visiting Researcher in the Communication Department of DLR, Oberpfaffenhofen, Germany. In 1994 and 1995, he was Visiting the École Nationale Supérieure des Télécomm.(E.N.S.T.), Paris. In 1998, he was Visiting Researcher in the Information Sciences Research Center of AT&T Research, Florham Park, NJ. In 2003, he was Visiting Researcher at the Mathematics Department of EPFL, Lausanne, Switzerland. In 2004, he was Visiting Researcher at the Telecommunications Department of UniCamp, Campinas, Brazil. In 2005 and 2006, he was Visiting Researcher at the ITR of UniSA, Adelaide, Australia.

Dr. Viterbo was awarded a NATO Advanced Fellowship in 1997 from the Italian National Research Council. His main research interests are in lattice codes for the Gaussian and fading channels, algebraic coding theory, algebraic space–time coding, digital terrestrial television broadcasting, and digital magnetic recording. He is Associate Editor of IEEE TRANSACTIONS ON INFORMATION THEORY, *European Transactions on Telecommunications* and *Journal of Communications and Networks*.