Recovery of a Lattice Generator Matrix from its Gram Matrix for Feedback and Precoding in MIMO

Francisco A. Monteiro, *Student Member, IEEE*, and Ian J. Wassell

Abstract— **Either in communication or in control applications, multiple-input multiple-output systems often assume the knowledge of a matrix that relates the input and output vectors. For discrete inputs, this linear transformation generates a multidimensional lattice. The same lattice may be described by an infinite number of generator matrixes, even if the rotated versions of a lattice are not considered. While obtaining the Gram matrix from a given generator matrix is a trivial operation, the converse is not obvious for non-square matrixes and is a research topic in algorithmic number theory. This paper proposes a method to execute such a conversion and applies it in a novel MIMO system implementation where some of the complexity is taken from the receiver to the transmitter. Additionally, given the symmetry of the Gram matrix, the number of elements required in the feedback channel is nearly halved.**

I. Introduction

Any multiple-input multiple-output (MIMO) system is traditionally described by a generator matrix. In the wireless (and recently also in wired [1],[2]) communications systems context, the matrix storing the fading coefficients between transmit and receive antennas is known as the *channel matrix,* however in other contexts which operate with vectorial spaces, the matrix receives other names. Considering that the inputs are restricted to a set of discrete inputs isomorphic to $\mathbb{Z}$, these systems can be framed in the general theory of lattices.

The regularity of a lattice lends itself to the representation of problems where different signals are interpreted as a point in a multidimensional space. They appear in many areas of signal processing such as quantization[3][4][5] or image processing [6]. Recently, lattices have also become a central tool in cryptography [7] [8]; they are also used in numerical integration (i.e., *quadrature*) of multi-dimensional functions constituting *lattice rules* [9][10], and have a long history in the fields of geometry of numbers, algorithmic number theory [11], multidimensional sphere packing (important in coding theory) [12] and also in integer programming [13].

The communication theory community has recently seen topics that were thought to be distinct (such as the multiple access channel, the broadcast channel, precoding, space-time coding, MIMO spatial multiplexing, and even OFDM) unified from a lattice perspective as a general equalization problem (e.g., [14]). Advances in lattice theory are therefore of great interest for MIMO engineering.

There are several ways of describing a lattice (e.g., via modular equations [15] or trellis structures [16]), however, the two most popular ones in engineering applications are *i)* the generator matrix and *ii)* the Gram matrix. The computation of the latter given the former is trivial. The reverse is not, and an efficient algorithm for this conversion remains an open problem in the theory of lattices.

It should be noticed that an efficient algorithm for this reverse operation can allow a lattice to be described using only about half the number of elements usually required when the dimensionality of the space is sufficiently high, provided that the Gram matrix is always symmetric. For example, in MIMO communications with channel state information at the transmitter (CSIT), this means that about half the number of coefficients would need to be sent to the transmitter (T_x) when compared with that when using traditional feedback [17]. Using the traditional example in [17] , while in a single-input single-output configuration (with BPSK modulation) the channel state information is conveyed by one coefficient only, in a 4×4 antenna system one has 16 complex variables describing the channel, or equivalently 32 real coefficients, that need to be periodically fedback to the transmitter. In fact, the number of coefficients to be fedback is the product of the number of antennas at the transmitter, at the receiver (R_x), the delay spread and, in multi-user environments, also proportional to the product with the number of users.

This paper shows how an approximate solution to an open problem in algorithmic number theory may lead to a more efficient CSIT mechanism. The paper proposes an algorithm to obtain a close approximation for a generator matrix given a Gram matrix of a lattice. The algorithm is based on an exact technique recently proposed by Lenstra [11] (an historical figure in the fields of algorithms for lattices). This paper uses the proposed algorithm as a constituent block in a novel strategy for closed-loop MIMO communication.

II. Lattice Basics

A lattice is a discrete subgroup (of maximal rank) in a Euclidean space and can be defined in a number of ways, as listed in Section I. We summarise here the most common ones.

A. *The generator matrix*

A n-dimensional lattice Λ may be defined by

$$\Lambda = \left\{ \mathbf{y} \in \mathbb{R}^n : \mathbf{y} = \sum_{i=1}^{n} \mathbf{h}_i x_i = \mathbf{H} \cdot \mathbf{x},\, x_i \in \mathbb{Z} \right\} \qquad (1)$$

Manuscript received November 23, 2009. This work was supported in part by a scholarship from the Foundation for Science and Technology, Portugal.

Francisco A. Monteiro is with the Department of Engineering and with The Computer Laboratory, University of Cambridge, UK, and also with the Telecommunications Institute at the Lisbon University Institute, Portugal (e-mail: fatbnm2@cam.ac.uk).

Ian J. Wassell is with The Computer Laboratory, University of Cambridge, UK (e-mail: ijw24@cam.ac.uk).

Revised version of the paper in the ISCCSP 2010 proceedings: a QR is moved into the receiver (Fig.3) and the implications from that.

where $\mathbf{y}$ are the points of the lattice, $\mathbf{h}_i$ are the *generating vectors* where each corresponds to the i^{th} column of the $n \times n$ *generator matrix* $\mathbf{H}$ (considering full-rank lattices only). Each integer x_i is an element of the column vectors $\mathbf{x}$. Thus, a lattice defined as in (1) is the span of the column space of the generator matrix $\mathbf{H}$, when we restrict the input to integers. Note that the prevalent notation in MIMO literature considers column vectors while in channel coding or in other fields in mathematics lattices are traditionally represented by the span of the row space of a generator matrix. Notice that rows and columns of a given $\mathbf{H}$ span different lattices.

Any generator matrix $\mathbf{H}$ can be transformed into one representing an *equivalent lattice* defined by

$$\mathbf{H}_{eq} = \mathbf{Q} \cdot \mathbf{H} \cdot \mathbf{M} \tag{2}$$

where $\mathbf{Q}$ is an *unitary matrix* (with real elements and $\det(\mathbf{Q}) = 1$), and where $\mathbf{M}$ is a *unimodular matrix* (with all elements integers and $\det(\mathbf{M}) = 1$).

With this in mind, the unitary transformation $\mathbf{Q}$ performs a rigid rotation of the lattice structure (i.e., of its generating vectors), while the unimodular matrix replaces a set of generating vectors by a different set that still generates the same lattice. Essentially, $\mathbf{M}$ finds an equivalent basis for the same structure and $\mathbf{Q}$ rotates the entire structure in space.

One of the hardest lattice problems is the *lattice distinguishing problem*, i.e., to discover if two lattices are “the same”. Once $\mathbf{Q}$ or $\mathbf{M}$ are fixed, answering the question becomes trivial. But when both transformations are unknown the question is difficult to solve in some particular problems [18],[19]. This decision problem has a simple solution when the lattices are *rational lattices* (when all entries are in $\mathbb{Q}$). In that case two lattices are equivalent if and only if their generating matrixes have the same Hermite Normal Form [13],[20]. However, the real lattices that arise in communication problems lead to numerical problems given the large numerators and denominators in the fractions representing fading coefficients. In that case the best approach is to use the fact that the QR decomposition is unique up to signs in the main diagonal.

B. *The Gram matrix*

Gram matrix is obtained from

$$\mathbf{G} = \mathbf{H}^H \mathbf{H}, \tag{3}$$

where H is the Hermitian operator (conjugate and transpose). The elements of $\mathbf{G}$ correspond to all the possible inner products $\mathbf{h}_i \cdot \mathbf{h}_j$ between all generating vectors and thus is unique to a lattice subject to unitary transformations $\mathbf{Q}$ (albeit not unique for unimodular transformations $\mathbf{M}$). It should be noticed that, by construction, $\mathbf{G}$ is always symmetric (because the inner product is commutative) and a *definite positive matrix*. This second property can be verified from the squared Euclidean norm of $\mathbf{y}$ (using the Hermitian operator):

$$\left\|\mathbf{y}\right\|^2 = \mathbf{y}^H \cdot \mathbf{y} = \left(\mathbf{Hx}\right)^H \cdot \left(\mathbf{Hx}\right) = \mathbf{x}^H \mathbf{H}^H \mathbf{H} \mathbf{x} = \mathbf{x}^H \mathbf{G} \mathbf{x}. \tag{4}$$

Consequently, one can state that $\mathbf{G}$ induces a quadratic form and is definite positive because $\left\|\mathbf{y}\right\| > 0$ for any $\mathbf{x} \neq 0$. This permits us to say that $\mathbf{G}$ always has a $\mathbf{LDL}^T$ decomposition [21][20][22].

Obtaining a valid $\mathbf{H}$ from $\mathbf{G}$ is not simple. $\mathbf{G}$ defines an abstract lattice, however, two versions of a lattice will have the same Gram matrix and in general, for a given $\mathbf{G}$, obtaining a possible $\mathbf{H}$ is named the *Gram matrix factorization problem*. When $\mathbf{H}$ is square, the Cholesky decomposition offers a good solution as it applies to symmetric definite positive matrices [21]. For the general $m \times n$ case, obtaining a basis from a specified Gram matrix had no available method in the literature until recently [11].

C. *The volume of lattices*

Full rank lattices are specified by a full-rank generator matrix and the volume of a lattice (e.g., [8]) is given by

$$\mathrm{Vol}(\Lambda) = \det\left(\mathbf{H}\right). \tag{5}$$

When $\mathbf{H}$ is rank-deficient (which is the case when $\mathbf{H}$ is non-square), the volume is

$$\mathrm{Vol}\left(\Lambda\right) = \sqrt{\det\left(\mathbf{H}^H \mathbf{H}\right)} = \sqrt{\det\left(\mathbf{G}\right)}. \tag{6}$$

III. MIMO MODELS

For a traditional complex representation with N_T inputs and N_R antennas as outputs, the received signal is

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n}, \text{ with} \tag{7}$$

$\mathbf{y} = \left[y_1, y_2, \ldots, y_{N_R}\right]^T \in \mathbb{C}^{N_R \times 1}$ $\mathbf{x} = \left[x_1, x_2, \ldots, x_{N_T}\right]^T \in \mathbb{C}^{N_T \times 1}$ and $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$, where each entry $h_{i,j}$ is a zero-mean circularly symmetric Gaussian random variable with unitary variance. The noise vector is $\mathbf{n} = \left[n_1, n_2, \ldots, n_{N_R}\right]^T \in \mathbb{C}^{N_R \times 1}$ with independent circularly symmetric Gaussian random variables, each one with a certain variance σ_n^2.

In the remainder of the paper we will resort to the equivalent real model of complex lattices:

$$\mathbf{y}_r = \mathbf{H}_r \mathbf{x}_r + \mathbf{n}_r \Leftrightarrow \begin{bmatrix} \Re\mathbf{y} \\ \Im\mathbf{y} \end{bmatrix} = \begin{bmatrix} \Re\mathbf{H} & -\Im\mathbf{H} \\ \Im\mathbf{H} & \Re\mathbf{H} \end{bmatrix} \begin{bmatrix} \Re\mathbf{x} \\ \Im\mathbf{x} \end{bmatrix} + \begin{bmatrix} \Re\mathbf{n} \\ \Im\mathbf{n} \end{bmatrix} \tag{8}$$

IV. THE MATRIX CONVERSION METHOD

Given a rational Gram matrix $\mathbf{G}$ it is possible to diagonalise the quadratic form as

$$\mathbf{G} = \mathbf{L} \cdot \mathbf{D} \cdot \mathbf{L}^T, \tag{9}$$

where $\mathbf{L}$ is a rational $n \times n$ lower triangular matrix with ones in the diagonal (i.e., is a unit matrix) and $\mathbf{D}$ is a $n \times n$ diagonal matrix with rational diagonal entries $d_{jj} > 0$.

We propose to expand these $d_{jj} \in \mathbb{Q}$ into a sum of a fixed number of squares of *R* rational numbers, that is,

$$\tilde{d}_{jj} = z_{j,1}^2 + z_{j,2}^2 + \cdots z_{j,R}^2 . \qquad (10)$$

An exact expansion of these d_{jj} can be accomplished by applying a naive greedy algorithm as proposed for the first time by Lenstra in [11]. Imposing an exact expansion for each d_{jj} often leads to a large number of terms in the sum (10) and for that reason Lenstra also proposed the use of a randomized algorithm given in [23], which assures the bound $R \leq 4$.

This paper proposes to replace an exact conversion from $\mathbf{G}$ to $\mathbf{H}$ by an approximate conversion (leading to a $\tilde{\mathbf{H}}$) using a fixed complexity algorithm that can be applied in a real time communication system. Based on the results in [23], we use a simple greedy algorithm for the expansion and truncate the number of terms to $R \leq 4$ leading to a *truncated Lenstra algorithm*. One way of achieving this is by using *R* equal terms in (10). One may notice that when *R*=1, the algorithm resorts to an approximated Cholesky decomposition.

One starts by constructing a *tall matrix* $\mathbf{B}$ with $R \cdot n$ rows and *n* columns. For the case with *R*=4 terms for each of the d_{jj}, $\mathbf{B}$ has the form

$$\mathbf{B} = \begin{bmatrix} z_{1,1} & z_{1,2} & z_{1,3} & z_{1,4} & & 0 & 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & 0 & & z_{j,1} & z_{j,2} & z_{j,3} & z_{j,4} & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & & z_{n,3} & z_{n,4} \end{bmatrix}^T \qquad (11)$$

where each row has one and only one non-zero entry. Now, one can re-construct the diagonal matrix $\mathbf{D}$ from $\mathbf{B}$ using

$$\tilde{\mathbf{D}} = \mathbf{B}^T \cdot \mathbf{B} =$$

We have made this matrix multiplication explicit to emphasise how this ensures that each d_{jj} is a sum of squares as defined in (10). Finally, the approximated generator matrix can be seen to be

$$\tilde{\mathbf{H}} = \mathbf{B} \cdot \mathbf{L}^T , \qquad (12)$$

because

$$\begin{aligned} \tilde{\mathbf{G}} &= \mathbf{L} \cdot \tilde{\mathbf{D}} \cdot \mathbf{L}^T = \mathbf{L} \cdot \mathbf{B}^T \cdot \mathbf{B} \cdot \mathbf{L}^T \\ &= \left(\mathbf{B} \cdot \mathbf{L}^T\right)^T \cdot \mathbf{B} \cdot \mathbf{L}^T = \tilde{\mathbf{H}}^T \cdot \tilde{\mathbf{H}}, \end{aligned} \qquad (13)$$

which verifies (3).

The complexity of this technique is cubic in the dimension *n*, due to the $\mathbf{LDL}^T$ steps, to which we should add $\mathcal{O}\left(n\right)$ for the rational approximation steps. The overall complexity is $\mathcal{O}\left(n^3 + n\right)$, however, as it will be shown later in Table 1, as the $\mathbf{LDL}^T$ decomposition will be reused, only the *n* term will add to the complexity of the technique to be presented.

It should be noticed that, unlike Cholesky decomposition, this technique is applicable to both square and non-square matrixes, allowing us to retrieve a rectangular $\mathbf{H}$ from $\mathbf{G}$.

V. CLOSED LOOP TECHNIQUE

Using traditional singular value decomposition (SVD) [1][24][25],

$$\begin{aligned} \mathbf{y} &= \underbrace{\mathbf{USV}^T}_{\mathbf{H}} \mathbf{x} + \mathbf{n} \\ \underbrace{\mathbf{U}^T\mathbf{y}}_{\mathbf{y}_d} &= \mathbf{U}^T\mathbf{USV}^T \left(\mathbf{Vx}\right) + \mathbf{U}^T\mathbf{n} = \mathbf{S} \cdot \mathbf{x} + \mathbf{U}^T\mathbf{n} \end{aligned} \qquad (14)$$

where $\mathbf{S}$ is a diagonal matrix and $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices (i.e., norm-preserving rotations, and therefore $\mathbf{U}^T$ preserves the statistics of the noise term). A SVD-based scheme implies one SVD decomposition (requiring at least $\mathcal{O}(6n^3)$ flops [22]) at the R_x in addition to matrix multiplications both at R_x and T_x (each multiplication requiring $\mathcal{O}(n^3)$ flops). This technique achieves capacity through *water filling* power allocation (according to the singular values) [26].

In [27] it was shown that the $\mathbf{LDL}^T$ decomposition achieves better performance than standard SVD, while being slightly less complex. Most importantly, that new approach takes advantage of having a precoding matrix with $(n^2 - n)/2$ zero elements. In fact it requires a in the precoder instead of the unitary $\mathbf{V}$ in (14). That lower triangular matrix is fedback from the R_x to the T_x, saving bandwidth in the feedback channel.

This paper proposes a technique that achieves the same performance as [27] while removing most of the complexity from the receiver side to the transmitter side, which is an important feature in scenarios where the transmitter is a base station and the receiver terminal should be made as simple as possible, i.e., by avoiding expensive processing.

Remark: as in [27], this paper is not assessing the capacity-achieving regime and thus, for simplicity, uniform power is allocated to the transmit antennas.

Both this proposal and [27] have a pre-processing stage at R_x consisting of the generation of a definite positive matrix **G,** the Gram matrix of the lattice defined by the columns of $\mathbf{H}$. This Gram matrix is formed by the left multiplication (3)

In this proposal it is imperative CSIT so that T_x can construct the precoding matrix $\mathbf{P}$. This paper shows that his can be achieved with the feedback of a lower triangular matrix only. Given the symmetry of $\mathbf{G}$, the T_x only needs to receive $(n^2 + n)/2$ coefficients and from them is able to reconstruct the entire matrix, achieving the same bandwidth savings seen in [27] for the feedback channel. After reconstructing the entire $\mathbf{G}$ (by symmetry), the T_x can use the truncated Lenstra algorithm described in Section IV to obtain an equivalent generator matrix (Section II) for the lattice. This matrix, $\tilde{\mathbf{H}}$, is not the same as $\mathbf{H}$ but rather an equivalent generator matrix for the lattice, holding the same Gram matrix. However, it is possible to obtain from them the same and unique generator matrix resulting from QR decompositions remembering that a QR decomposition is unique when imposing the positiveness of elements in the main diagonal. Thus

$$\mathbf{H} = \mathbf{QR} \text{ and } \tilde{\mathbf{H}} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}} \tag{15}$$

lead to $\mathbf{R}$ and $\tilde{\mathbf{R}}$, which would be the same matrix (up signs in main diagonal) if there was no distortion associated with $\tilde{\mathbf{H}}$ in the truncated Lenstra algorithm. A central aspect is that the same Gram matrix is also obtainable from $\mathbf{R}$ alone, $\tilde{\mathbf{R}}$ alone or even a mixture of both, given their closeness:

$$\mathbf{G} = \mathbf{R}^T\mathbf{R} \cong \tilde{\mathbf{R}}^T\tilde{\mathbf{R}} \cong \mathbf{R}^T\tilde{\mathbf{R}}\,. \tag{16}$$

As indicated in Section II.B and in Section IV, $\mathbf{G}$ would have a $\mathbf{LDL}^T$ decomposition, which can be calculated not only given $\mathbf{H}$ or $\tilde{\mathbf{H}}$ (in both cases applying (3)) but also given $\mathbf{R}$, or $\tilde{\mathbf{R}}$, or both. However, this will not be the matrix to be $\mathbf{LDL}^T$ decomposed.

Taking advantage of these facts the mode in (7) and be changed into

$$\mathbf{y} = \mathbf{Q}\,\mathbf{R}\,\mathbf{x} + \mathbf{n} \tag{17}$$

and then, applying a precoding matrix $\mathbf{P} = \tilde{\mathbf{R}}^T(\tilde{\mathbf{L}}^T)^{-1}\tilde{\mathbf{D}}^{-1/2}$,

$$\mathbf{y}' = \mathbf{Q}\underbrace{\mathbf{R}\,(\tilde{\mathbf{R}}^T}_{\mathbf{G}'}(\tilde{\mathbf{L}}^T)^{-1}\tilde{\mathbf{D}}^{-1/2})\,\mathbf{x} + \mathbf{n}\,. \tag{18}$$

The matrixes used in (18) will be presented and justified on the fallowing. First, notice that this made a matrix $\mathbf{G}' = \mathbf{R}\tilde{\mathbf{R}}^T$ to appear (a permutation matrix may be needed together with $\tilde{\mathbf{R}}$ to have a unique QR), which, despite not being the Gram matrix of the underlying lattice, it is the (approximate) Gram matrix of the lattice spanned by the row lattice (as indicated in Section II). This matrix $\mathbf{G}'$ also has an $\mathbf{LDL}^T$ decomposition and thus (18) can equivalently be written as

$$\mathbf{y}' = \mathbf{Q}\underbrace{\tilde{\mathbf{L}}\,\tilde{\mathbf{D}}\,\tilde{\mathbf{L}}^T}_{\mathbf{G}'}(\tilde{\mathbf{L}}^T)^{-1}\,\tilde{\mathbf{D}}^{-1/2}\,\mathbf{x} + \mathbf{n}\,. \tag{19}$$

Finally, after the detection filter at the receiver, the entire chain becomes

$$\underbrace{\left(\tilde{\mathbf{L}}^{-1}\mathbf{Q}^{-1}\right)\mathbf{y}'}_{\mathbf{y}'_d} = \left(\tilde{\mathbf{L}}^{-1}\mathbf{Q}^{-1}\right)\mathbf{Q}\,\tilde{\mathbf{L}}\,\tilde{\mathbf{D}}\,\tilde{\mathbf{L}}^T\,(\tilde{\mathbf{L}}^T)^{-1}\,\tilde{\mathbf{D}}^{-1/2}\,\mathbf{x} + \left(\tilde{\mathbf{L}}^{-1}\mathbf{Q}^{-1}\right)\mathbf{n}$$
$$\mathbf{y}'_d = \tilde{\mathbf{D}}\tilde{\mathbf{D}}^{-1/2}\cdot\mathbf{x} + \left(\tilde{\mathbf{L}}^{-1}\mathbf{Q}^T\right)\mathbf{n} = \mathbf{D}_{eq}\cdot\mathbf{x} + \underbrace{\left(\tilde{\mathbf{L}}^{-1}\mathbf{Q}^T\right)\mathbf{n}}_{n_d} \tag{20}$$

It should be noticed that, as $\mathbf{Q}$ is orthogonal (unitary if considering complex models), then $\mathbf{Q}^{-1} = \mathbf{Q}^T$, which further simplifies the computations at the receiver. The $\mathrm{R_x}$ will then just have to apply the filtering $\mathbf{F} = \tilde{\mathbf{L}}^{-1}\mathbf{Q}^T$ to the incoming precoded signal, i.e., the unavoidable filtering multiplications that are present in all detectors. Besides that, the $\mathrm{R_x}$ only needs to compute $\mathbf{G}$ and (given its symmetry) send back to the $\mathrm{T_x}$ only the lower or the upper parts of $\mathbf{G}$, which will be denoted by $\mathbf{G}_{1/2}$. Moreover, both $\tilde{\mathbf{L}}^{-1}$ and $\mathbf{Q}^T$ are computed and sent from the $\mathrm{T_x}$ to the $\mathrm{R_x}$.

The resulting transmission chain (20) can be interpreted in two ways: *i)* algebraically it corresponds to a set of independent transmission channels and *ii)* geometrically it corresponds to a communication problem over a rectangular lattice. Thus, it is convenient to think of a lattice as the result of a linear transformation of the cubic lattice $\mathbb{Z}^n$.

The diagonal matrix $\mathbf{D}_{eq}$ corresponds to a set of orthogonal generating vectors which span a rectangular lattice (this lattice would even have the simplest trellis representation one may have for lattices, e.g., [16]). The performance increase can be geometrically interpreted from this insight. A rectangular lattice is obtained from a deformation of a cubic lattice $\mathbb{Z}^n$ by stretching each dimension according to each d_{jj}. Its decision regions are rectangles and thus even a zero-forcing detector experiences no performance penalty.

The right multiplication of the channel matrix by $\tilde{\mathbf{R}}^T$ in (17) changes the power at the transmitter. The geometric interpretation is also useful on this matter. The "row lattice" $\mathcal{L}(\mathbf{H}^T)$ has volume

$$\mathrm{Vol}\left(\mathcal{L}(\tilde{\mathbf{R}}^T) =\right)\sqrt{\det\left(\tilde{\mathbf{R}}\tilde{\mathbf{R}}^T\right)}\,. \tag{21}$$

At the same time, because $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{L}}^T$ are unit matrixes,

$$\mathrm{Vol}\left(\mathcal{L}(\tilde{\mathbf{D}})\right) = \det\left(\tilde{\mathbf{D}}\right) = \det\left(\tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T\right)$$
$$= \det\left(\tilde{\mathbf{G}}'\right) = \det\left(\tilde{\mathbf{R}}\tilde{\mathbf{R}}^T\right). \tag{22}$$

Subsequently, one also needs the insertion of a diagonal scaling $\mathbf{D}^{-1/2}$ at the precoding stage so that the volume of both lattices underlying the transmission scheme becomes the same.

Figure 1 depicts the overall transmission scheme that is proposed while in Figure 2 and in Figure 3 one can observe in detail the processing required respectively at the $\mathrm{T_x}$ and at the $\mathrm{R_x}$ as well as the fluxes of CSI between both of them.

At the $\mathrm{R_x}$ it is important to highlight that there are two parallel processing occurring at different stages and each one associated with a different fading block: *i)* obtain $\mathbf{G}$ that will be sent back to $\mathrm{T_x}$ in the form of a triangular matrix and *ii)* construct the receive filter from a received strictly upper triangular matrix and $\mathbf{Q}$. In fact, $\tilde{\mathbf{L}}^{-1}$ is not only a lower triangular but also a unit lower triangular (ones in the diagonal). This saves the transmission of the diagonal and thus only the $(n^2-n)/2$ coefficients of $\tilde{\mathbf{L}}^{-1}$ are needed to be forwarded to the $\mathrm{R_x}$. These coefficients are denoted by $\tilde{\mathbf{L}}_u^{-1}$. The process is summarised in Algorithm 1. (The channel is assumed to remain unchanged between adjacent symbols as it is common in the slow fading assumption.)

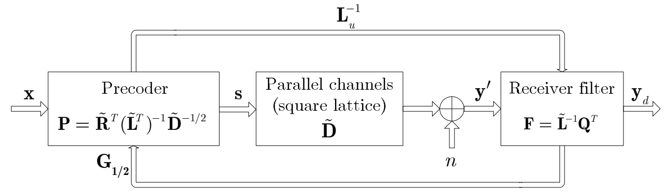

Figure 1: Proposed closed loop transmission scheme.

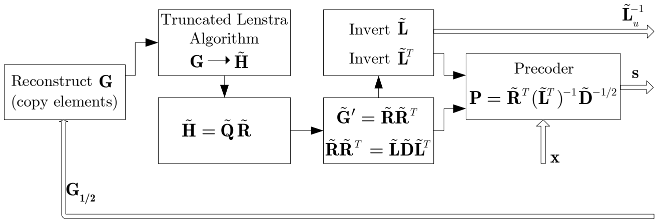

Figure 2: Processing at the transmitter.

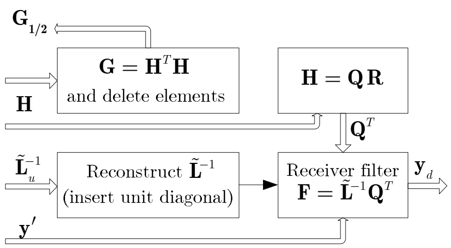

Figure 3: Processing at the receiver.

ALGORITHM 1: CLOSED LOOP TECHNIQUE

1: Channel estimation at R_x : $\mathbf{H}$

2: Gram matrix of the lattice at R_x : $\mathbf{G} = \mathbf{H}^T\mathbf{H}$

3: Lower triangular matrix is feedback: $\mathbf{G}_{1/2}$ ($R_x \rightarrow T_x$)

4: Gram matrix is reconstructed at T_x:

$$\mathbf{G} = \mathbf{G}_{1/2} + \left(\mathbf{G}_{1/2}\right)^T - \mathrm{diag}\left(\mathbf{G}_{1/2}\right)$$

5: Obtain approximate $\tilde{\mathbf{H}}$ from $\mathbf{G}$ using the algorithm in Section IV (which encompasses a $\mathbf{LDL}^T$ for $\mathbf{G}$)

6: Compute QR decomposition of $\tilde{\mathbf{H}}$: $\tilde{\mathbf{H}} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}$

7: Compute the Gram matrix of the "row lattice" at T_x: $\tilde{\mathbf{G}}' = \tilde{\mathbf{R}}\tilde{\mathbf{R}}^T$

8: Decomposition of $\mathbf{G}'$ at T_x: $\tilde{\mathbf{G}}' = \tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T$

9: Precoding at T_x: $\mathbf{P} = \tilde{\mathbf{R}}^T(\tilde{\mathbf{L}}^T)^{-1}\tilde{\mathbf{D}}^{-1/2}$; $\mathbf{s}$ is sent

(note: for a non-squared $\tilde{\mathbf{R}}$ the non-zero rows must be deleted)

10: Strictly lower triangular matrix is sent: $\tilde{\mathbf{L}}_u^{-1}$ ($T_x \rightarrow R_x$)

11: Compute QR decomposition of $\mathbf{H}$: $\mathbf{H} = \mathbf{QR}$

(note: can be computed at the same time as steps 4-10)

12: $\tilde{\mathbf{L}}^{-1}$ is reconstructed at R_x: $\tilde{\mathbf{L}}^{-1} = \tilde{\mathbf{L}}_u^{-1} + \mathbf{I}_{n\times n}$

13: Receiver filter $\mathbf{F} = \tilde{\mathbf{L}}^{-1}\mathbf{Q}^T$ multiplies the received chain and the received vector becomes $\mathbf{y}_d = \tilde{\mathbf{D}}^{1/2}\mathbf{x} + \mathbf{n}_d$

The number of flops required by the $\mathbf{LDL}^T$ decomposition is $\mathcal{O}(n^3/3)$, which is half of the number of flops needed in Gaussian elimination, the number of flops of QR decomposition is $\mathcal{O}(2n^3)$ and for the standard matrix multiplication one has $\mathcal{O}(n^3)$ [21][22][28](though there are more efficient algorithms for matrix multiplication). Table 1 contains a comparison of the proposed technique with SVD and with [27] in terms of the number of flops and number of coefficients flowing in both the uplink and downlink. The number of operations in Table 1 is presented in a way that shows the contribution of each individual processing stage to the total number of operations of the R_x or T_x (matrix multiplications are counted as only one $\mathcal{O}(n^3)$ though). The complexity at the receiver comes from a QR decomposition and two matrix multiplications: one to initially obtain $\mathbf{G}$ (similar to [27]) and then the unavoidable filtering multiplication by $\mathbf{F}$. One should remember that this last multiplication is common to all types of receivers in both closed or open-loop configurations.

TABLE 1

	SVD	[27]	Proposal
# flops at R_x	$\mathcal{O}\left(4n^3 + n^3\right)$	$\mathcal{O}\left(n^3 + n^3/3 + n^3\right)$	$\mathcal{O}\left(2n^3 + n^3\right)$
# flops at T_x	$\mathcal{O}\left(n^3\right)$	$\mathcal{O}\left(n^3\right)$	$\mathcal{O}\left(n + 2n^3 + n^3/3 + n^3\right)$
Coefficients in feedback	n^2	$\left(n^2 + n\right)/2$	$\left(n^2 + n\right)/2$
Coefficients in downlink	–	–	$\left(n^2 - n\right)/2$
Total of coefficients	n^2	$\left(n^2 + n\right)/2$	n^2

VI. ASSESSMENT OF THE APPROXIMATION

To access the approximation one first computes the error matrix of the Gram matrix involved (i.e., the Gram matrix associated with the "row lattice", as indicated in Section V)

$$\mathbf{E} = \mathbf{G}' - \tilde{\mathbf{G}}' \tag{23}$$

and one applies to it the squared *Frobenius matrix norm* [22]

$$\left\|\mathbf{E}\right\|_F^2 = \sum_{i,j}\left|e_{i,j}\right|^2 = \mathrm{Trace}\left(\mathbf{E}^H\mathbf{E}\right) \tag{24}$$

as the evaluation metric.

Figure 4 shows the distribution of this error for three example cases having the number of real dimensions most common in MIMO wireless communications (and with variance 0.5 per real component).

Notice that despite the Gram matrix of the "row lattice" and the one of "column lattice" being different, they hold the same distribution because $\mathbf{H}$ and $\mathbf{H}^H$ exhibit the same statistics and consequently they are interchangeable in (3).

For a N_T =4, N_R =4 configuration (i.e., n=8 dimensions) under a Rayleigh fading channel and using 16-QAM modulation, it was verified that with the $\mathbf{LDL}^T$ decomposition proposed in this paper the error shown in Figure 4 leads to a negligible performance penalty in terms of symbol error rate (SER) in respect to the results presented in [27] for the same configuration when using the same minimum mean squared error (MMSE) receiver.

VII. CONCLUSIONS

This paper shows how to reconstruct (with very low distortion and fixed complexity) a generator matrix of a lattice from one given Gram matrix of the same lattice for non-square matrixes. This opens new possibilities in several problems of engineering and computer science that rely on lattice geometry as it breaks through the restriction that the Cholesky decompositions imposes (valid for squared matrixes). Subsequently, the paper shows how the algorithm devised in Section IV can be central to a technique for channel diagonalisation of MIMO systems. With this technique: *i)* $\mathbf{LDL}^T$ decomposition takes place at the transmit side; *ii)* the number of elements to be fedback to T_x is $(n^2+n)/2$, as in [27]; *iii)* the filtering matrix at R_x is build from a unit lower triangular and an orthogonal matrix, which further reduces the complexity of the filtering matrix multiplication at R_x. The extra cost to bear is a QR decomposition at the R_x. However, a QR module would have to exist at R_x if typical open-loop spatial multiplexing schemes are also to be supported. For large number of antennas, the presented closed loop architecture (i.e., with CSIT) for MIMO communications nearly halves the number of coefficients traditionally needed to represent the channel.

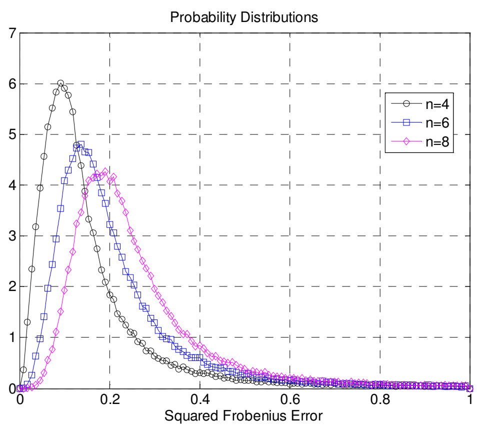

Figure 4: Probability distribution of the squared Frobenius norm of the error matrix for $\tilde{\mathbf{G}}$ (or $\tilde{\mathbf{G}}'$).

REFERENCES

[1] Fernando Pérez-Cruz, Miguel R. Rodrigues, and Sergio Verdú, "Optimal precoding for digital subscriber lines," in *Proceedings of ICC'08 - The IEEE Inter. Conf. on Communications*, Beijing, May 2008, pp. 1200-1204.

[2] Per Ödlig, Thomas Magesacher, Per Ola Börjesson Stefan öst, Miguel Berg, and Enrique Areizaga, "The forth generation broadband concept," *IEEE Communications Magazine*, vol. 43, no. 1, pp. 63-69, January 2009.

[3] M. Vedat Eyuboglu and G. David Forney, "Lattice and trellis quantization with lattice- and trellis-bounded codebooks - high-rate theory for memoryless sources," *IEEE Transactions on Information theory*, vol. 39, no. 1, pp. 46-58, January 1993.

[4] Erik Agrell and Thomas Eriksson, "Optimization of lattices for quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 5, pp. 1814-1828, September 1998.

[5] Robert M. Gray and David L. Neuhoff, "Quantization," *IEEE Transactions on Information theory*, vol. 44, no. 6, pp. 2325-2383, October 1998.

[6] R. Cortelazzo and G. Manduchi, "On the determinant of all sublattices of preassigned index and its applications to multidimensional sampling," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 3, no. 4, pp. 318-320, August 1993.

[7] Daniele Micciancio and Oded Regev, "Lattice-based cryptography," in *Post-Quantum Cryptography*, Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen, Eds. Berlin, Germany: Springer, 2009, pp. 146-191.

[8] Daniele Micciancio and Shafi Goldwasser, *Complexity of Lattice Problems - A Cryptographic Perspective*. Norwell, Massachusetts, USA: Kluwer Academic Publishers, 2002.

[9] Ian H. Sloan and Stephen Joe, *Lattice Methods for Multiple Integration*. Oxford, UK: Oxford University Press, 1994.

[10] J. N. Lyness, "An introduction to lattice rules and their generator matrices," *IMA Journal of Numerical Analysis - Oxford University Press*, vol. 9, no. 3, pp. 405-419, July 1989.

[11] Hendrik W. Lenstra, "Lattices," in *Algorithmic Number Theory*, J. P. Buhler and P. Stevenhagen, Eds. Cambridge, UK: Cambridge University Press, 2008, pp. 127-181.

[12] John H. Conway and Neil J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed. New York, New York, USA: Springer, 1999.

[13] Alexandre Schrijver, *Theory of Linear and Integer Programming*. Chischester, UK: John Wiley & Sons, 1986, ch. 4, 5, 6.

[14] Wei Zhang, Xiaoli Ma, B. Gestner, and D. Anderson, "Designing low-complexity equalizers for wireless systems," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 56-62, January 2009.

[15] A. Paz and C. P. Schnorr, "Approximating integer lattices by lattices with cycle factor groups," in *Proceedings of 14th Int Conf on Automata, Languages and Programming*, Karlsruhe, Germany, LNCS 267, July 1987, pp. 386-393.

[16] G. David Forney, "Density / length profiles and trellis complexity of lattices," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1753-1772, November 1994.

[17] David J. Love, Robert W. Heath, Wiroonsak Santipach, and Michael L. Honig, "What is the value of limited feedback for MIMO channels," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54-59, October 2004.

[18] Michael Szydlo, "Hypercubic lattice reduction and analysis of GGH and NTRU signatures," in *Proceedings of Eurocrypt 2003*, Warsaw, Poland, LNCS 2656, Germany, 2003, pp. 433-448.

[19] Erik Agrell, Alexander Vardy Thomas Eriksoon, and Kenneth Zeger, "Closest point in lattices," *IEEE Transactions on Information Theory*, vol. 48, no. 8, pp. 2201-2214, August 2002.

[20] Robert Piziak and P. L. Odell, *Matrix Theory - From Generalized Inverses to Jordan Form*. Boca Raton, FL: Chapman & Hall - CRC, 2007.

[21] G. W. Stewart, *Introduction to Matrix Computations*. London, UK: Academic Press, 1973.

[22] Gene H. Golub and Charles F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, Maryland, USA: The Hohns Hopkins University Press, 1996.

[23] M. O. Rabin and J. O. Shallit, "Randomized algorithms in number theory," *Communications on Pure and Applied Mathematics*, vol. 39:Suplement, pp. S239-S256, 1986.

[24] Jürgen Lindner, "Transmission over MIMO channels: on diversiy and spacial multiplexing," in *Proceedings of ISCTA´07 - The Inter. Symposium on Communication Theory and Applications*, Ambleside, July 2007.

[25] Kenichi Kobayashi, Tomoaki Ohtsuki, and Toshinobu Kaneko, "MIMO Systems in the Presence of Feedback Delay," *IEICE Transactions on Communications*, vol. E91-B, no. 3, pp. 829-836, March 2008.

[26] Ezio Biglieri and Giorgio Taricco, *Transmission and Reception with Multiple Antennas: Theorectical Foundations*. Hanover, Massachusetts, USA: now Publishers, 2004.

[27] Che-Chen Chou, Hsi-Chei Chen, and Jen-Ming Wu, "A low complexity channel decomposition and feedback strategy for MIMO precoder design," in *Proc. of ICASSP'09 - The IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, April 2009, pp. 2705-2707.

[28] Ran Raz, "On the complexity of matrix product," in *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, Montreal, Canada, May 2002, pp. 144-151.