The Diagonal Reduction Algorithm Using Fast Givens

Wen Zhang, Sanzheng Qiao, and Yimin Wei

Abstract Recently, a new lattice basis reduction notion, called diagonal reduction, was proposed for lattice-reduction-aided detection (LRAD) of multiinput multioutput (MIMO) systems. In this paper, we improve the efficiency of the diagonal reduction algorithm by employing the fast Givens transformations. The technique of the fast Givens is applicable to a family of LLL-type lattice reduction methods to improve efficiency. Also, in this paper, we investigate dual diagonal reduction and derive an upper bound of the proximity factors for a family of dual reduction aided successive interference cancelation (SIC) decoding. Our upper bound not only extends an existing bound for dual LLL reduction to a family of dual reduction methods, but also improves the existing bound.

Keywords Lattice reduction, diagonal reduction, fast Givens, dual reduction, MIMO detection.

1 Introduction

Lattice basis reduction plays an important role in the detection of wireless multiple-input multiple-output (MIMO) systems. For detection problems of lattice type, the

Wen Zhang is supported by 973 Program Project under grant 2010CB327900.
School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao, 266000, P. R. China. e-mail: zhangwen9801@gmail.com

Sanzheng Qiao is partially supported by Shanghai Key Laboratory of Contemporary Applied Mathematics.
Department of Computing and Software, McMaster University, Hamilton, ON, L8S 4K1, Canada. e-mail: qiao@cas.mcmaster.ca

Yimin Wei is supported by the National Natural Science Foundation of China under grant 11271084.
School of Mathematical Sciences and Shanghai Key Laboratory of Contemporary Applied Mathematics, Fudan University, Shanghai, 200433, P.R. China. e-mail: ymwei@fudan.edu.cn

optimal maximum-likelihood (ML) decoding can be modeled as the closest vector problem (CVP) [1, 16], which has been proved to be NP-hard [2]. Although many algorithms, like the sphere decoding algorithm [14, 5], can solve CVP exactly, the complexity of these algorithms increases exponentially with the number of transmit antennas [1, 5, 6]. Thus, such optimal solvers are infeasible for real-time systems, where timing is critical. To satisfy the time constraint, many sub-optimal solvers with polynomial complexity, like the successive interference cancelation (SIC) decoding, have been proposed [3, 12]. However, the sub-optimal detectors may suffer from heavy performance loss at a low signal-to-noise ratio (SNR). It has been found that lattice reduction, used as an efficient preprocessor, has the potential to achieve high performance for sub-optimal decoding algorithms. Recently, many reduction algorithms, such as the Lenstra-Lenstra-Lovász (LLL) algorithm [7], effective LLL algorithm [10], partial LLL algorithm [11, 17], and diagonal reduction algorithm [19], have been proposed for SIC decoding. It is proved in [9, 18] that SIC decoding aided by the above reduction notions can achieve the same receive diversity order as the infinite lattice decoding (ILD).

Of all the aforementioned lattice reduction algorithms, the diagonal reduction algorithm is the most efficient one. From our observation [19], the total computation of the diagonal reduction is dominated by the computation of the Givens rotations. Thus, in this paper, we propose to improve the efficiency of the diagonal reduction by replacing the Givens rotation with the more efficient and mathematically equivalent fast Givens transformation [4, Page 218]. The improvement is achieved by substantially reducing the number of multiplication operations required, because two entries of the 2-by-2 fast Givens matrix equal 1. Moreover, the fast Givens technique is general in that it can be incorporated into all the LLL-type lattice reduction methods to enhance performance.

Also, we investigate the basis reduction for dual lattices. In [9], the LLL and effective LLL algorithms for dual lattices are presented. In this paper, we investigate the diagonal reduction for dual lattices and prove that the dual basis of a diagonal reduced basis is also diagonal reduced. In addition, we derive an upper bound for the proximity factors of a family of dual LLL-type reduction aided SIC decoding. Our upper bound not only extends an existing bound for LLL reduction in [9] to a family of reduction methods, but also improves the existing one.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the systems model and review the diagonal reduction algorithm. The new algorithm using the fast Givens is given in Section 3. Section 4 presents the diagonal reduction for dual lattices and our new upper bound for the proximity factors. In Section 5, we demonstrate our simulation results.

Notations: $\mathbf{B}^{\mathrm{T}}$, $\mathbf{B}^{\dagger}$, and $\det(\mathbf{B})$ denote the transpose, the Moore-Penrose inverse, and the determinant of a matrix $\mathbf{B}$ respectively, $\Re(z)$ and $\Im(z)$ the real and imaginary parts of a complex number z, $\lfloor a \rceil$ the integer nearest to a real number a.

2 Lattice Basis Reduction

2.1 System Model

Consider a MIMO system consisting of n_T transmit antennas and m_T receive antennas. The relationship between the $n_T \times 1$ transmitted signal vector $\mathbf{x}$ and the $m_T \times 1$ received signal vector $\mathbf{y}$ is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{H}$, $\mathbf{y}, \mathbf{n}$ represent the channel matrix, the received and additive noise signals, respectively. In general, the entries of both $\mathbf{H}$ and $\mathbf{n}$ are assumed to be complex-valued independently and identically distributed (i.i.d.) Gaussian variables. Treating the real and imaginary parts of (1) separately, an equivalent real-valued system of doubled size can be obtained:

$$\mathbf{y} = \mathbf{B}\mathbf{x} + \mathbf{n}, \tag{2}$$

where

$$\mathbf{y} = \begin{bmatrix} \Re(\mathbf{y}) \\ \Im(\mathbf{y}) \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} \Re(\mathbf{n}) \\ \Im(\mathbf{n}) \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \Re(\mathbf{H}) & -\Im(\mathbf{H}) \\ \Im(\mathbf{H}) & \Re(\mathbf{H}) \end{bmatrix}.$$

Given a MIMO system modeled as (2), the optimum ML decoding is equivalent to the following CVP:

$$\min_{\mathbf{x} \in \mathscr{A}} \|\mathbf{y} - \mathbf{B}\mathbf{x}\|_2. \tag{3}$$

where the constellation $\mathscr{A}$ is of lattice type. Unfortunately, CVP has been proved to be NP-hard [2], and all existing algorithms for solving (3) has an exponential complexity with the lattice dimension n [5, 6]. Recently, lattice-reduction-aided SIC decoding turned out to be extremely promising, since its bit-error-rate (BER) performance can effectively approximate the ML decoding with a complexity of only $O(n^3)$ operations [15, 9].

2.2 Diagonal Reduction Algorithm

In this section, we first introduce some concepts of lattices and the SIC decoding, then we describe the diagonal reduction method [19].

Given a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$ $(n \leq m)$ of full column rank, then a *lattice* generated by $\mathbf{B}$ is defined by $L(\mathbf{B}) = \{\mathbf{Bz} : \mathbf{z} \in \mathbb{Z}^n\}$. The columns of $\mathbf{B}$ form a *basis* for the lattice $L(\mathbf{B})$. An integer matrix $\mathbf{Z} \in \mathbb{Z}^{n \times n}$ is called *unimodular* if $|\det(\mathbf{Z})| = 1$. The columns of a matrix $\mathbf{B}'$ can form a basis for $L(\mathbf{B})$ if and only if there exists a unimodular matrix $\mathbf{Z}$ such that $\mathbf{B}' = \mathbf{BZ}$. The *volume* of $L(\mathbf{B})$ is defined as $\mathrm{vol}(L(\mathbf{B})) = \sqrt{\det(\mathbf{B}^{\mathrm{T}}\mathbf{B})}$, which is independent of the choice of basis. Let $\lambda(L)$ be the Euclidean length of the shortest nonzero vector in a lattice L, then it is well known that $\lambda(L)/\mathrm{vol}(L)^{1/n}$ is upper bounded over all n-dimension lattices L, and the *Hermite's constant* γ_n is defined as the supremum of $\lambda(L)^2/\mathrm{vol}(L)^{2/n}$ over all n-dimention lattices. Finding the exact value of γ_n is very difficult. The exact value of γ_n is only known for $1 \leq n \leq 8$ and $n = 24$ [13, Page 33]. For an arbitrary dimension n, an upper bound of the Hermite's constant is given in [13, Page 35]:

$$\gamma_n \leq 1 + \frac{n}{4}, \quad \text{for all } n \geq 1. \tag{4}$$

A *lattice reduction algorithm* finds a unimodular matrix $\mathbf{Z}$ for a given $\mathbf{B}$ such that the columns of $\mathbf{BZ}$ are reasonably short. Lattice reduction has now become a powerful tool for enhancing the performance of sub-optimal MIMO detectors, since it can significantly improve the orthogonality of the channel matrix.

Given a lattice generator matrix $\mathbf{B} \in \mathbb{R}^{m\times n}$ and its QR decomposition $\mathbf{B} = \mathbf{QR}$, where $\mathbf{Q} \in \mathbb{R}^{m\times n}$ has orthonormal columns and $\mathbf{R} \in \mathbb{R}^{n\times n}$ is upper triangular. From [8, 17], the efficiency of sphere decoding and the performance of SIC decoding is determined by the arrangement of the diagonal elements of $\mathbf{R}$. Based on this fact, various reduction notions, such as the LLL reduction [7], effective LLL reduction [10], partial LLL reduction [11, 17], and diagonal reduction [19], have been proposed. Among all the aforementioned reduction notions, the diagonal reduction is the weakest, consequently, the least computationally demanding.

Definition 1 (Diagonal reduction [19]). A basis matrix $\mathbf{B} \in \mathbb{R}^{m\times n}$ is said to be diagonal reduced with the parameter ω $(1/4 < \omega < 1)$, if the entries $r_{i,j}$ of the upper triangular factor $\mathbf{R}$ in its QR decomposition $\mathbf{B} = \mathbf{QR}$ satisfy

$$(r_{k-1,k} - \mu_k r_{k-1,k-1})^2 + r_{k,k}^2 \geq \omega r_{k-1,k-1}^2, \tag{5}$$

for all $1 < k \leq n$, where $\mu_k = \lfloor r_{k-1,k}/r_{k-1,k-1} \rceil$.

From the above definition, diagonal reduction only imposes one simple constraint on the diagonal entries of $\mathbf{R}$. However, it is proved in [19] that diagonal-reduction-aided SIC decoding has identical performance as LLL-reduction-aided SIC decoding. A generic implementation of diagonal reduction can be found in Figure 1.

3 Diagonal Reduction Using Fast Givens

From Figure 1, the computational cost of the diagonal reduction algorithm includes two parts: the size-reduction (lines 7-8) and the Givens rotation (lines 11-13). The simulation results in [19] indicate that the overall complexity of the algorithm is dominated by the Givens rotations as the lattice dimension n increases. Thus, we propose the use of the fast Givens transformation in place of the Givens rotations to speed up the diagonal reduction algorithm.

Like the Givens rotation, the fast Given can be used to introduce zeros into selected positions. Specifically, given a lattice generator matrix $\mathbf{B}$, the fast Givens transformation is based on the following decomposition:

```
**Input:** $\mathbf{Q} \in \mathbb{R}^{m\times n}$, $\mathbf{R} \in \mathbb{R}^{n\times n}$, $\omega$
**Output:** Updated $\mathbf{Q}$ and updated $\mathbf{R}$ that is diagonal reduced with the parameter $\omega$ and a unimodular $\mathbf{Z}$ that reduces $\mathbf{R}$
Initialization $\mathbf{Z} \leftarrow \mathbf{I}_n$
$k \leftarrow 2$
while $k \le n$ do
    $\mu_k \leftarrow \lfloor \mathbf{R}(k-1,k)/\mathbf{R}(k-1,k-1) \rceil$
    if the condition (5) is not satisfied then
        if $\mu_k \neq 0$ then
            $\mathbf{R}(1:k-1,k) \leftarrow \mathbf{R}(1:k-1,k) - \mu_k \mathbf{R}(1:k-1,k-1)$
            $\mathbf{Z}(:,k) \leftarrow \mathbf{Z}(:,k) - \mu_k \mathbf{Z}(:,k-1)$
        end if
        swap columns $k-1$ and $k$ in $\mathbf{R}$ and $\mathbf{Z}$
        find a Givens rotation $\mathbf{G}$ to restore the upper triangular structure of $\mathbf{R}$
        $\mathbf{R}(k-1:k,k-1:n) \leftarrow \mathbf{G}\mathbf{R}(k-1:k,k-1:n)$
        $\mathbf{Q}(:,k-1:k) \leftarrow \mathbf{Q}(:,k-1:k)\mathbf{G}^{\mathrm{T}}$
        $k \leftarrow \max(k-1,2)$
    else
        $k \leftarrow k+1$
    end if
end while
```

Fig. 1 Diagonal reduction algorithm (DR) [19]

$$\mathbf{B} = \mathbf{F}\mathbf{D}^{-1}\mathbf{R}, \tag{6}$$

where $\mathbf{D} = \mathrm{diag}(d_i)$ is a positive diagonal matrix, $\mathbf{F}\mathbf{D}^{-1/2}$ represents the orthogonal factor in the QR decomposition of $\mathbf{B}$, and $\mathbf{D}^{-1/2}\mathbf{R}$ represents the upper triangular factor.

How can the fast Givens introduce zeros? In the 2-by-2 case, given $\mathbf{x} = [x_1, x_2]^{\mathrm{T}}$ and the corresponding diagonal elements $d_1, d_2 > 0$, we first compute

$$\alpha = -x_1/x_2, \quad \beta = -\alpha d_2/d_1, \quad \text{and } \gamma = -\alpha\beta.$$

When $\gamma \le 1$, we have the type 1 fast Givens:

$$\mathbf{F} = \begin{bmatrix} \beta & 1 \\ 1 & \alpha \end{bmatrix} \tag{7}$$

and update d_1 and d_2:

$$\widehat{d_1} \leftarrow (1+\gamma)d_2 \quad \text{and} \quad \widehat{d_2} \leftarrow (1+\gamma)d_1. \tag{8}$$

When $\gamma > 1$, setting

$$\alpha \leftarrow 1/\alpha, \quad \beta \leftarrow 1/\beta, \quad \text{and } \gamma \leftarrow 1/\gamma,$$

we have the type 2 fast Givens:

$$\mathbf{F} = \begin{bmatrix} 1 & \beta \\ \alpha & 1 \end{bmatrix} \tag{9}$$

and update d_1 and d_2:

$$\widehat{d_1} \leftarrow (1+\gamma)d_1 \quad \text{and} \quad \widehat{d_2} \leftarrow (1+\gamma)d_2. \tag{10}$$

Then it can be verified that

$$\mathbf{F} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \times \\ 0 \end{bmatrix}$$

and

$$\begin{bmatrix} \widehat{d_1} & 0 \\ 0 & \widehat{d_2} \end{bmatrix}^{-1/2} \mathbf{F} \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}^{1/2}$$

is orthogonal.

In our fast Givens based diagonal reduction algorithm, all the transformations are based on the decomposition (6). In the beginning, we compute the QR decomposition $\mathbf{B} = \mathbf{QR}$ and set $\mathbf{F} = \mathbf{Q}$ and $\mathbf{D} = \mathbf{I}_n$. Thus, in this case, the size-reduction in each iteration is the same as lines 7-8 of Figure 1. But the diagonal reduction condition (5) becomes

$$d_{k-1}^{-1}(r_{k-1,k} - \mu_k r_{k-1,k-1})^2 + d_k^{-1} r_{k,k}^2 \geq \omega d_{k-1}^{-1} r_{k-1,k-1}^2, \tag{11}$$

for $1 < k \leq n$. The diagonal reduction algorithm using fast Givens (DRFG) is summarized in Figure 2.

In comparison with the original diagonal reduction algorithm, DRFG saves a substantial number of multiplication operations, since two entries of the 2-by-2 fast Givens matrix are equal to 1. However, DRFG introduces overhead, such as the computations in line 14 and line 20. Our simulation results presented in Section 5 show that overall DRFG is more efficient than DR.

4 Dual Diagonal Reduction

In this section, after a brief introduction to dual basis, we first investigate diagonal reduction of dual bases and prove that if a primal basis is diagonal reduced, then its dual basis is also diagonal reduced. Then we derive an upper bound of proximity factor of SIC decoding, which not only improves an existing bound for the dual LLL reduction in [9], but also extends it to the family of dual LLL-type reductions.

```
Input: $\mathbf{Q} \in \mathbb{R}^{m\times n}$, $\mathbf{R} \in \mathbb{R}^{n\times n}$, $\omega$
Output: Updated $\mathbf{Q}$ and updated $\mathbf{R}$ that is diagonal reduced with the parameter $\omega$ and a unimodular $\mathbf{Z}$ that reduces $\mathbf{R}$
$\mathbf{F} \leftarrow \mathbf{Q}$, $\mathbf{D} \leftarrow \mathbf{I}_n$, $\mathbf{Z} \leftarrow \mathbf{I}_n$
$k \leftarrow 2$
while $k \le n$ do
    $\mu_k \leftarrow \lfloor \mathbf{R}(k-1,k)/\mathbf{R}(k-1,k-1) \rceil$
    if the condition (11) is not satisfied then
        if $\mu_k \neq 0$ then
            $\mathbf{R}(1:k-1,k) \leftarrow \mathbf{R}(1:k-1,k) - \mu_k \mathbf{R}(1:k-1,k-1)$
            $\mathbf{Z}(:,k) \leftarrow \mathbf{Z}(:,k) - \mu_k \mathbf{Z}(:,k-1)$
        end if
        swap columns $k-1$ and $k$ in $\mathbf{R}$ and $\mathbf{Z}$
        construct fast Givens matrix $\mathbf{F}$ of type (7) or type (9) such that $\mathbf{F}\begin{bmatrix} r_{k-1,k-1} \\ r_{k,k-1} \end{bmatrix} = \begin{bmatrix} \times \\ 0 \end{bmatrix}$
        $\mathbf{R}(k-1:k,k-1:n) \leftarrow \mathbf{F}\mathbf{R}(k-1:k,k-1:n)$
        $\mathbf{F}(:,k-1:k) \leftarrow \mathbf{F}(:,k-1:k)\mathbf{F}^{\mathrm{T}}$
        using (8) or (10) to update $d_{k-1}$ and $d_k$
        $k \leftarrow \max(k-1,2)$
    else
        $k \leftarrow k+1$
    end if
end while
$\mathbf{Q} \leftarrow \mathbf{F}\mathbf{D}^{-1/2}$, $\mathbf{R} \leftarrow \mathbf{D}^{-1/2}\mathbf{R}$
```

Fig. 2 Diagonal reduction algorithm using fast Givens (DRFG)

4.1 Dual Lattice Reduction

Let L be an n-dimensional lattice in $\mathbb{R}^m$, then the dual lattice L^* of L is defined as the set

$$L^* = \{\mathbf{u} \mid \langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{Z}, \text{for all } \mathbf{v} \in L\}, \tag{12}$$

where $\langle \mathbf{u}, \mathbf{v} \rangle$ is the inner product of $\mathbf{u}$ and $\mathbf{v}$. Suppose that $\mathbf{B}$ is a primal basis matrix of L, then it is obvious that the columns of $\mathbf{B}^{\dagger\mathrm{T}}$ form a basis for its dual lattice L^*. In this paper, we adopt the definition of the dual basis $\mathbf{B}^* \triangleq \mathbf{B}^{\dagger\mathrm{T}}\mathbf{J}$ [9], where

$$\mathbf{J} \triangleq \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \cdots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}$$

A dual lattice is closely related to its corresponding primal lattice. For instance, we have $L^{**} = L$ and $\det(L^*) = 1/\det(L)$.

Given a primal basis matrix $\mathbf{B}$, then the dual basis reduction is to perform a lattice reduction algorithm on its dual basis $\mathbf{B}^*$. Like the primal basis reduction, dual basis reduction can also return a well reduced basis of the primal lattice. Suppose that $\mathbf{Z}^*$ is the unimodular matrix that reduces the dual basis $\mathbf{B}^*$, then the corresponding reduced primal basis is given by

$$\mathbf{B}' = (\mathbf{B}^{\dagger\mathrm{T}}\mathbf{J}\mathbf{Z}^*)^{\dagger\mathrm{T}}\mathbf{J} = \mathbf{B}\mathbf{J}(\mathbf{Z}^*)^{\dagger\mathrm{T}}\mathbf{J},$$

where $\mathbf{J}(\mathbf{Z}^*)^{\dagger\mathrm{T}}\mathbf{J}$ is the unimodular matrix associated with the primal lattice.

To study the reduction properties of diagonal reduction on dual lattices, the following result is essential.

Lemma 1. *Let* $\mathbf{B} = \mathbf{Q}\mathbf{R}$ *and* $\mathbf{B}^* = \mathbf{Q}^*\mathbf{R}^*$ *be the QR decompositions of the primal basis* $\mathbf{B}$ *and its dual basis* $\mathbf{B}^*$*, respectively. Then*

$$\mathbf{Q}^* = \mathbf{Q}\mathbf{J}, \quad \mathbf{R}^* = \mathbf{J}\mathbf{R}^{-\mathrm{T}}\mathbf{J}. \tag{13}$$

Proof: It is easy to verify that $\mathbf{B}^\dagger = \mathbf{R}^{-1}\mathbf{Q}^{\mathrm{T}}$. Thus, we have

$$\begin{aligned}\mathbf{B}^* &= \mathbf{B}^{\dagger\mathrm{T}}\mathbf{J} = (\mathbf{R}^{-1}\mathbf{Q}^{\mathrm{T}})^{\mathrm{T}}\mathbf{J} = \mathbf{Q}\mathbf{R}^{-\mathrm{T}}\mathbf{J} \\ &= (\mathbf{Q}\mathbf{J})\cdot(\mathbf{J}\mathbf{R}^{-\mathrm{T}}\mathbf{J}).\end{aligned} \tag{14}$$

Obviously, $\mathbf{Q}\mathbf{J}$ has orthonormal columns and $\mathbf{J}\mathbf{R}^{-\mathrm{T}}\mathbf{J}$ is an upper triangular matrix, thus the proof is completed.

Based on the above lemma, we can obtain the following result.

Proposition 1. *If the lattice basis matrix* $\mathbf{B}$ *is diagonal reduced, then its dual basis* $\mathbf{B}^*$ *is also diagonal reduced.*

Proof: Let $\mathbf{R} = [r_{i,j}]$ and $\mathbf{R}^*$ be the upper triangular factors of $\mathbf{B}$ and $\mathbf{B}^*$, respectively. Then from Lemma 1,

$$\begin{aligned}\mathbf{R}^* &= \mathbf{J}\mathbf{R}^{-\mathrm{T}}\mathbf{J} \\ &= \begin{bmatrix} \frac{1}{r_{n,n}} & -\frac{r_{n-1,n}}{r_{n-1,n-1}r_{n,n}} & \times & & \times \\ & \frac{1}{r_{n-1,n-1}} & -\frac{r_{n-2,n-1}}{r_{n-2,n-2}r_{n-1,n-1}} & & \times \\ & & \frac{1}{r_{n-2,n-2}} & \ddots & \vdots \\ & & & \ddots & -\frac{r_{1,2}}{r_{1,1}r_{2,2}} \\ & & & & \frac{1}{r_{1,1}} \end{bmatrix}\end{aligned} \tag{15}$$

Since $\mathbf{B}$ is diagonal reduced, we then have

$$\left(r_{k-1,k} - \left\lfloor \frac{r_{k-1,k}}{r_{k-1,k-1}} \right\rceil \cdot r_{k-1,k-1}\right)^2 + r_{k,k}^2 \geq \omega r_{k-1,k-1}^2, \tag{16}$$

for all $1 < k \leq n$. Multiplying the both sides of (16) with $\frac{1}{(r_{k-1,k-1}r_{k,k})^2}$, we obtain

$$\left(\frac{r_{k-1,k}}{r_{k-1,k-1}r_{k,k}} - \left\lfloor \frac{r_{k-1,k}}{r_{k-1,k-1}} \right\rceil \cdot \frac{1}{r_{k,k}}\right)^2 + \left(\frac{1}{r_{k-1,k-1}}\right)^2 \geq \omega \left(\frac{1}{r_{k,k}}\right)^2,$$

which implies that $\mathbf{R}^*$ is also diagonal reduced.

4.2 Proximity Factor

To characterize the performance gap between sub-optimal decoding and ILD, a proximity factor was defined in [8] and further discussed in [9, 18]. Given a lattice generator matrix $\mathbf{B} = [\mathbf{b}_1, ..., \mathbf{b}_n]$, denote ϕ_i the acute angle between $\mathbf{b}_i$ and the linear space spanned by the previous $i-1$ basis vectors, then the proximity factor of SIC decoding is defined as:

$$\rho_i \triangleq \sup_{\mathbf{B}\in\mathscr{B}_{\text{Red}}} \frac{\lambda^2(L(\mathbf{B}))}{\|\mathbf{b}_i\|_2^2 \sin^2\phi_i}, \tag{17}$$

where the supremum is taken over the set $\mathscr{B}_{\text{Red}}$ of bases satisfying a certain reduction notion for any n-dim lattice L. We further define $\rho \triangleq \max_i\{\rho_i\}$. From [9], the average error probability of SIC decoding can be bounded as

$$P_{e,SIC}(\text{SNR}) \leq \sum_{i=1}^{n} P_{e,ILD}\left(\frac{\text{SNR}}{\rho_i}\right) \leq nP_{e,ILD}\left(\frac{\text{SNR}}{\rho}\right)$$

for arbitrary SNR.

Denote ρ_{LLL}, ρ_{DLLL}, ρ_{DR}, and ρ_{DDR} the proximity factors of SIC decoding aided by LLL reduction, dual LLL reduction, diagonal reduction, and dual diagonal reduction, respectively. An upper bound of ρ_{LLL} is given in [18]:

$$\rho_{LLL} \leq \gamma_n \cdot \beta^{\frac{n-1}{2}} \leq \left(1+\frac{n}{4}\right)\beta^{\frac{n-1}{2}}, \tag{18}$$

where $\beta = 1/(\omega - 1/4) \geq 4/3$. Following the argument in [19], it is easy to prove that

$$\rho_{i,DR} = \rho_{i,LLL} \leq \gamma_i \cdot \beta^{\frac{i-1}{2}}. \tag{19}$$

Thus,

$$\rho_{DR} = \rho_{n,DR} \leq \gamma_n \cdot \beta^{\frac{n-1}{2}}. \tag{20}$$

For dual reduction, an upper bound of ρ_{DLLL} is given in [9]:

$$\rho_{DLLL} \leq \beta^{n-1}. \tag{21}$$

In the following, we improve the upper bound (21). From (19) and Proposition 1, we can obtain that

$$\rho_{i,DDR} = \sup_{\mathbf{B}^*\in\mathscr{B}_{DR}} \frac{\lambda^2(L(\mathbf{B}))}{r_{i,i}^2} = \sup_{\mathbf{B}\in\mathscr{B}_{DR}} \frac{\lambda^2(L(\mathbf{B}))}{r_{i,i}^2} \leq \gamma_i \cdot \beta^{\frac{i-1}{2}}. \tag{22}$$

Thus,

$$\rho_{DDR} = \rho_{n,DDR} \leq \gamma_n \cdot \beta^{\frac{n-1}{2}}. \tag{23}$$

Following the above argument, it is easy to prove that the proximity factors of SIC decoding aided by all dual LLL-type reduction notions, such as dual LLL reduction, dual effective LLL reduction, and dual partial LLL reduction, can be upper bounded by the right-hand side of (23). Comparing (23) with (20), SIC decoding aided by primal and dual diagonal reductions are expected to have the same performance. This shall be confirmed by the simulation results presented in Section 5.

5 Simulation Results

In this section, we present our simulation results on comparing the efficiency of the proposed algorithm DRFG with the original algorithm DR. All experiments were performed on matrices with random entries, drawn from an i.i.d. zero-mean, unit variance Gaussian distribution. Without loss of generality, all testing matrices were set to square matrices. For each size, we generated 1000 random matrices and took an average. The parameter ω in the reduction algorithms was set to 0.99.

Although the new algorithm DRFG is expected to be faster than the original algorithm DR, the computations in line 14 and line 20 of Figure 2 introduce overhead. To compare the overall complexity of the algorithms, we experimented on the floating-point operations (flops)[1] carried out by the algorithms. Figure 3 depicts our results on the average numbers of flops performed by the reduction algorithms. The figure shows that in both cases of primal and dual lattice reduction, DRFG is more efficient than DR, and the performance gap between them widens quickly as the dimension increases. This indicates that the overhead introduced by the fast Givens is insignificant. Also note that the DR (DRFG) algorithm is slightly faster than its dual counter part dual DR (dual DRFG) algorithm. This is due to the additional computation, for instance, the calculation of $\mathbf{B}^{\dagger}$, required by the dual reduction.

We also investigated the reduction quality of different reduction algorithms measured by the BER performance of the SIC decoding. Specifically, using a 64-QAM constellation, Figure 4 depicts the simulated BER curves of lattice-reduction-aided SIC over an 8×8 uncoded MIMO fading channel. We have found that the SIC aided by the four diagonal reduction algorithms have identical BER performance with that aided by the LLL algorithm. This is consistent with the theoretical analysis presented in Section 4.2.

Acknowledgements We would like to thank Professor Lihong Zhi and referees for their useful comments.

[1] Flop count: addition/multiplication/division/max/rounding, 1 flop.

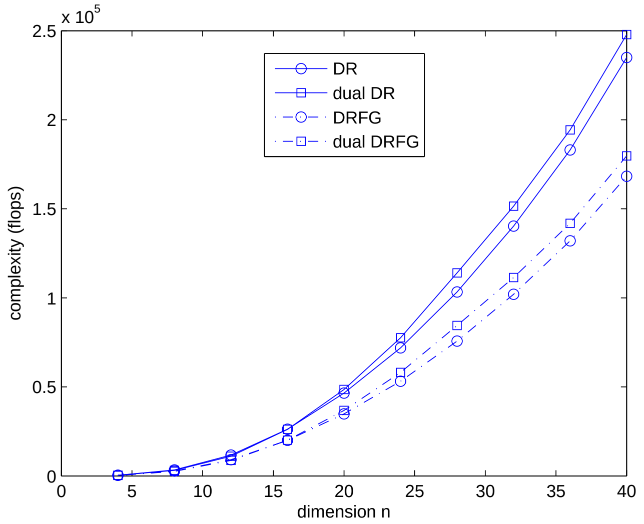

Fig. 3 The average complexity (in flops) of the diagonal reduction (DR), dual diagonal reduction (dual DR), DRFG, and dual DRFG algorithms.

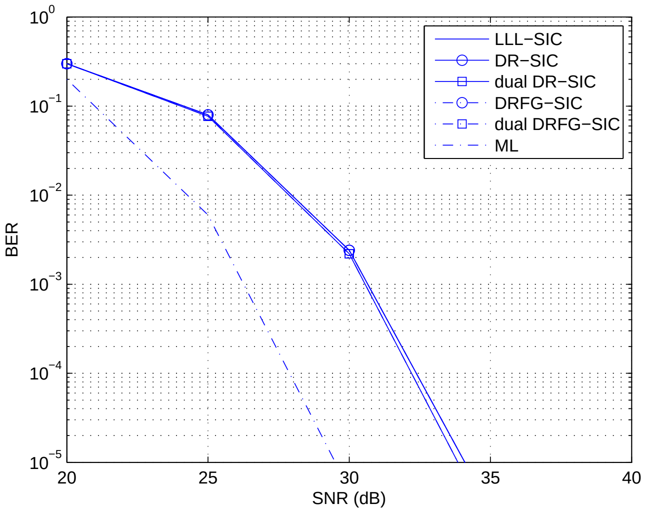

Fig. 4 Simulated BER of SIC aided by the LLL, DR, dual DR, DRFG, and the dual DRFG for 64-QAM over an 8 × 8 uncoded MIMO fading channel.

References

1. Agrell, E., Eriksson, T., Vardy, A., Zeger, K.: Closest point search in lattices. IEEE Trans. Inf. Theory. **48**, 2201-2214 (2002)

2. Arora, S., Babai, L., Stern, J.: The hardness of approximate optima in lattices, codes, and systems of linear equations. J. Comput. Syst. Sci. **54**, 317-331 (1997)
3. Babai, L.: On Lovász's lattice reduction and the nearest lattice point problem. Combinatorica. **6**, 1-13 (1986)
4. Golub, G. H., Van Loan, C. F.: Matrix Computations. The Johns Hopkins University Press, Baltimore, MD (1996)
5. Hassibi, B., Vikalo, H.: On the sphere-decoding algorithm I: Expected complexity. IEEE Trans. Signal Process. **53**, 2806-2818 (2005)
6. Jaldén, J., Ottersen, B.: On the complexity of sphere decoding in digital communications. IEEE Trans. Signal Process. **53**, 1474-1484 (2005)
7. Lenstra, A. K., Lenstra, H. W., Lovász, L.: Factorizing polynomials with rational coefficients. **261**, 515-534 (1982)
8. Ling, C.: Towards characterizing the performance of approximate lattice decoding. Proc. Int. Symp. Turbo Codes/Int. Conf. Source Channel Coding '06. Munich, Germany (2006)
9. Ling, C.: On the proximity factors of lattice reduction-aided decoding. IEEE Trans. Signal Process. **59**, 2795-2808 (2011)
10. Ling, C., Howgrave-Graham, N.: Effective LLL reduction for lattice decoding. Proc. IEEE Int. Symp. Inf. Theory (ISIT). Nice, France (2007)
11. Ling, C., Mow, W. H., Gan, L.: Dual-lattice ordering and partial lattice reduction for SIC-based MIMO detection. IEEE J. Sel. Topics Signal Process. **3** 975-985 (2009)
12. Mow, W. H.: Universal lattice decoding: Principle and recent advances. Wireless Commun. Mobile Comput., Special Issue on Coding and Its Appl. Wireless CDMA Syst. **3**, 553-569 (2003)
13. Nguyen, P. Q., Vallée, B.: The LLL Algorithm: Survey and Applications. Springer-Verlag, Berlin, Germany (2009)
14. Schnorr, C. P., Euchner, M.: Lattice basis reduction: Improved practical algorithms and solving subset sum problems. Math. Programming. **66**, 181-199 (1994)
15. Taherzadeh, M., Mobasher, A., Khandani, A. K.: LLL reduction achieves the receive diversity in MIMO decoding. IEEE Trans. Inf. Theory, **53** 4801-4805 (2007)
16. Wübben, D., Seethaler, D., Jaldén, J., Marz, G.: Lattice reduction: a survey with applications in wireless communications. IEEE Signal Process. Mag. **28** 70-91 (2011)
17. Xie, X., Chang, X. W., Borno, M. A.: Partial LLL reduction. Proc. IEEE GLOBECOM (2011)
18. Zhang, W., Qiao, S., Wei, Y.: HKZ and Minkowski reduction algorithms for lattice-reduction-aided MIMO detection. IEEE Trans. Signal Process. **60** 5963-5976 (2012)
19. Zhang, W., Qiao, S., Wei, Y.: A diagonal lattice reduction algorithm for MIMO detection. IEEE Signal Process. Lett. **19** 311-314 (2012)