Optimal Factorization in Lattice-Reduction-Aided and Integer-Forcing Linear Equalization

Sebastian Stern and Robert F.H. Fischer
Institute of Communications Engineering, Ulm University, Ulm, Germany
Email: {sebastian.stern,robert.fischer}@uni-ulm.de

Abstract—**Lattice-reduction-aided (LRA) equalization is a very interesting multi-user equalization technique as it enables a low-complexity full-diversity detection. To this end, the multiple-input/multiple-output channel is factorized into a reduced variant and a unimodular integer matrix. Inspired by the closely related finite-field processing strategy of integer-forcing (IF) equalization, this factorization task has recently been relaxed to non-unimodular integer matrices. In this paper, the claim of a significant performance gain induced by the IF philosophy is revisited. For that purpose, lattice-basis-reduction approaches are reviewed; the optimal one with respect to channel equalization is identified. A fair comparison between unimodular LRA and IF strategy is given, complemented by detailed numerical results.**

I. Introduction

Multiple-input/multiple-output (MIMO) transmission has become one of the most important principles in modern communication systems. In the past decade, research has especially been focused on multi-user communication, where several users simultaneously transmit to or receive from one central multi-antenna instance. Restricting to uplink transmission, this is known as *MIMO multiple-access channel*.

Efficient strategies had to be found to handle the multi-user interference. Possible solutions were either a simple linear channel equalization or the technique of decision-feedback equalization using the principle of successive interference cancellation. However, both approaches were not really convincing as they could not exploit the MIMO channel's diversity.

In order to achieve full diversity [13], *lattice-reduction-aided* (LRA) equalization has been proposed [16], [14], where the MIMO channel is interpreted as the generator matrix of a lattice. The channel matrix is factorized into a reduced version thereof (*reduced basis* of the lattice) and a unimodular integer part which describes the change of basis. To this end, well-known lattice-basis-reduction strategies could be applied, especially the Lenstra-Lenstra-Lovász (LLL) algorithm. In the sequel, advanced approaches like Hermite-Korkine-Zolotareff (HKZ) and Minkowski (MK) reduction were studied, e.g., in [18]. Evaluations have most often been performed from a mathematical point of view, e.g., the orthogonality defect of the reduced channel has been assessed. This, however, may not necessarily be the optimal way in terms of communications.

Some time ago, the mathematically-driven lattice basis reduction has been queried by the *integer-forcing* (IF) equalization strategy [17]. The IF receiver employs the basic structure of the LRA receiver, but performs the interference cancellation over a finite field. Thereby, the integer part does not have to describe a change of basis any more. It only has to consist of linearly independent lattice points, optimally determined by algorithms solving the *successive-minima problem* [4], [5]. Recently, this relaxation to non-unimodular (full-rank) integer matrices has been generalized [5]: even the (conventional) LRA receiver is able to handle any full-rank integer matrix; a finite-field processing (including its restrictions on the signal constellation) is not required. Certainly, the question arises if this novel factorization approach can improve the performance.

Even though some comparisons between conventional (unimodular) lattice basis reduction and (non-unimodular) IF factorization philosophy have been given in literature, e.g., in [10], [4], [5], a fair comparison on equal terms is still an open point—especially for the high-dimensional case. Hence, in this paper, we first derive the optimal factorization strategy for conventional lattice reduction and compare it to state-of-the-art reduction schemes. Via this, a fair comparison of the lattice-reduction and IF philosophy is given, supported by extensive numerical results. This particularly includes a clarification of the gain directly induced by dropping the unimodularity.

The paper is structured as follows: In Sec. II, basic properties of lattices and lattice-basis-reduction schemes are explained. Sec. III starts with the channel model and the structure of the LRA and IF receiver. Following this, an optimal factorization is derived and/or reviewed for both approaches. Numerical simulations and a detailed comparison are provided in Sec. IV. A summary and conclusions are given in Sec. V.

II. Lattices and Related Problems

In this section, important properties of lattices as well as the principle of lattice basis reduction and related algorithms are reviewed. Since a complex-valued transmission is considered in the following, we restrict to complex-valued lattices and related definitions. More precisely, the signal points (constellation) are assumed to be drawn from the *Gaussian integers* $\mathbb{G} = \mathbb{Z} + \mathrm{j}\,\mathbb{Z}$, i.e., the integer lattice in the complex plane.

A complex-valued lattice $\boldsymbol{\Lambda}(\boldsymbol{G})$ is defined by

$$\boldsymbol{\Lambda}(\boldsymbol{G}) = \left\{ \boldsymbol{G}\,[u_1, \ldots, u_K]^{\mathsf{T}} = \sum_{k=1}^{K} u_k\, \boldsymbol{g}_k \mid u_k \in \mathbb{G} \right\} , \quad (1)$$

where $\boldsymbol{G} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K] \in \mathbb{C}^{N \times K}$ is its generator matrix which consists of $K \in \mathbb{N}$ linearly independent basis vectors $\boldsymbol{g}_k \in \mathbb{C}^N$, $N \geq K$, $N \in \mathbb{N}$ (N-dimensional lattice of rank K).

This work has been supported by Deutsche Forschungsgemeinschaft (DFG) within the framework COIN under grant Fi 982/4-3.

A. Gram-Schmidt Orthogonalization

Any matrix $\boldsymbol{G}$ can be decomposed into the form $\boldsymbol{G} = \boldsymbol{G}^\circ \boldsymbol{R}$, where $\boldsymbol{G}^\circ = [\boldsymbol{g}_1^\circ, \ldots, \boldsymbol{g}_K^\circ]$ is the Gram-Schmidt orthogonalization of $\boldsymbol{G}$ with orthogonal columns $\boldsymbol{g}_1^\circ, \ldots, \boldsymbol{g}_K^\circ$. Thereby, $\boldsymbol{R} \in \mathbb{C}^{K\times K}$ is upper triangular with unit main diagonal.

The vectors of $\boldsymbol{G}^\circ$ can be obtained successively for index $k = 1, \ldots, K$ via the equation $\boldsymbol{g}_k^\circ = \boldsymbol{g}_k - \sum_{l=1}^{k-1} r_{l,k}\, \boldsymbol{g}_l^\circ$, where the coefficients of $\boldsymbol{R}$ are given by

$$r_{l,k} = \frac{(\boldsymbol{g}_l^\circ)^{\mathsf{H}} \boldsymbol{g}_k^\circ}{\|\boldsymbol{g}_l^\circ\|_2^2} \;, \quad l = 1, \ldots, k \; . \tag{2}$$

B. Minkowski's Successive Minima and Related Problems

The k^{th} *successive minimum* of $\boldsymbol{\Lambda}(\boldsymbol{G})$, $k = 1, \ldots, K$, is denoted as $\rho_k(\boldsymbol{\Lambda}(\boldsymbol{G}))$. It is defined as

$$\rho_k(\boldsymbol{\Lambda}(\boldsymbol{G})) = \inf\left\{ r_k \mid \dim\left(\operatorname{span}\left(\boldsymbol{\Lambda}(\boldsymbol{G}) \cap \boldsymbol{B}_N(r_k)\right)\right) = k \right\} , \tag{3}$$

where $\boldsymbol{B}_N(r)$ is the N-dimensional ball (over $\mathbb{C}$) with radius r centered at the origin [3], [8], [4]. Concisely said, r_k has to be chosen as the smallest radius for which $\boldsymbol{B}_N(r_k)$ contains k linearly independent lattice vectors.

1) Shortest Independent Vector Problem (SIVP): For a given complex-valued lattice $\boldsymbol{\Lambda}(\boldsymbol{G})$ of rank K, a set of K linearly independent lattice vectors $\mathcal{G} = \{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_K\}$ has to be obtained, such that $\|\boldsymbol{\lambda}_k\|_2 \leq \rho_K(\boldsymbol{\Lambda}(\boldsymbol{G}))$, $k = 1, \ldots, K$. Briefly speaking, the lattice vector(s) in the set $\mathcal{G}$ with largest Euclidean norm has (have) to be as short as possible, i.e.,

$$\max_{k=1,\ldots,K} \|\boldsymbol{\lambda}_k\|_2 = \rho_K(\boldsymbol{\Lambda}(\boldsymbol{G})) \; . \tag{4}$$

The norms of all shorter vectors do not matter in this case.

2) Successive Minima Problem (SMP): Given a lattice $\boldsymbol{\Lambda}(\boldsymbol{G})$ of rank K, a set of K linearly independent lattice vectors $\mathcal{G} = \{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_K\}$ has to be found, such that

$$\|\boldsymbol{\lambda}_k\|_2 = \rho_k(\boldsymbol{\Lambda}(\boldsymbol{G})) \;, \quad k = 1, \ldots, K \; . \tag{5}$$

Hence, all lattice vectors in the set $\mathcal{G}$ have to be as short as possible. Naturally, in that case, the SIVP is solved, too.

C. Lattice Basis Reduction

Regarding the problem of lattice basis reduction, a set of K linearly independent lattice vectors $\mathcal{G} = \{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_K\}$ of a (complex-valued) lattice $\boldsymbol{\Lambda}(\boldsymbol{G})$ of rank K has to be found as well. However, here these vectors additionally have to form a (reduced) basis of the lattice. As a consequence, the lattice can equivalently be defined by the reduced generator matrix $\boldsymbol{G}_{\mathrm{r}} = [\boldsymbol{g}_{\mathrm{r},1}, \ldots, \boldsymbol{g}_{\mathrm{r},K}] = [\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_K]$, i.e., $\boldsymbol{\Lambda}(\boldsymbol{G}) = \boldsymbol{\Lambda}(\boldsymbol{G}_{\mathrm{r}})$.

The change of basis is then described by

$$\boldsymbol{G} = \boldsymbol{G}_{\mathrm{r}} \boldsymbol{U} \;, \tag{6}$$

where $\boldsymbol{U} \in \mathbb{G}^{K\times K}$ is unimodular ($|\det(\boldsymbol{U})| = 1$), ensuring the existence of an integer inverse $\boldsymbol{U}^{-1} \in \mathbb{G}^{K\times K}$. In dependency of the reduction criterion/algorithm, the orthogonality defect of $\boldsymbol{G}_{\mathrm{r}}$ and/or the Euclidean norm of its basis vectors may be reduced in comparison to the original basis $\boldsymbol{G}$. In the following, the most important reduction criteria are listed.

1) Lenstra-Lenstra-Lovász (LLL) Reduction: A generator matrix $\boldsymbol{G} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K] \in \mathbb{C}^{N\times K}$ with Gram-Schmidt orthogonal basis $\boldsymbol{G}^\circ = [\boldsymbol{g}_1^\circ, \ldots, \boldsymbol{g}_K^\circ]$ and upper triangular matrix $\boldsymbol{R}$ according to (2) is called (C)LLL-reduced [6], if

i) for $1 \leq l < k \leq K$, it is *size-reduced* according to

$$|\mathrm{Re}\{r_{l,k}\}| \leq 0.5 \quad \text{and} \quad |\mathrm{Im}\{r_{l,k}\}| \leq 0.5 \;, \tag{7}$$

ii) for $k = 2, \ldots, K$ and a parameter $0.5 < \delta \leq 1$,

$$\|\boldsymbol{g}_k^\circ\|_2^2 \geq (\delta - |r_{k-1,k}|^2) \|\boldsymbol{g}_{k-1}^\circ\|_2^2 \; . \tag{8}$$

The parameter δ controls the trade-off between "strength" of the LLL reduction and computational complexity. The case $\delta = 1$ is denoted as *optimal LLL reduction* [2]; in all other cases the reduction is suboptimal. Usually, $\delta = 0.75$ is chosen.

2) Hermite-Korkine-Zolotareff (HKZ) Reduction: A generator matrix $\boldsymbol{G} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K] \in \mathbb{C}^{N\times K}$ with Gram-Schmidt orthogonal basis $\boldsymbol{G}^\circ = [\boldsymbol{g}_1^\circ, \ldots, \boldsymbol{g}_K^\circ]$ and upper triangular $\boldsymbol{R}$ according to (2) is called (C)HKZ-reduced [8], [7], if

i) $\boldsymbol{R}$ is size-reduced according to (7),

ii) for $k = 1, \ldots, K$, the columns of $\boldsymbol{G}^\circ$ fulfill

$$\|\boldsymbol{g}_k^\circ\|_2 = \rho_1(\boldsymbol{\Lambda}(\boldsymbol{G}^{(k)})) \; . \tag{9}$$

Thereby, $\boldsymbol{\Lambda}(\boldsymbol{G}^{(k)})$ is the sublattice of rank $K - k + 1$ and dimension N, which is defined by the generator matrix[1] $\boldsymbol{G}^{(k)} = [0, \ldots, 0, \boldsymbol{g}_k^\circ, \ldots, \boldsymbol{g}_K^\circ]\boldsymbol{R}$. Hence, the k^{th} column of the Gram-Schmidt orthogonal basis has to be a shortest (non-zero) vector in $\boldsymbol{\Lambda}(\boldsymbol{G}^{(k)})$ with norm $\rho_1(\boldsymbol{\Lambda}(\boldsymbol{G}^{(k)}))$.

3) Minkowski (MK) Reduction: We call a generator matrix $\boldsymbol{G} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K] \in \mathbb{C}^{N\times K}$ (C)MK-reduced [9], [18], if $\forall \boldsymbol{G}' = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{k-1}, \boldsymbol{g}'_k, \ldots, \boldsymbol{g}'_K]$ with $\boldsymbol{\Lambda}(\boldsymbol{G}') = \boldsymbol{\Lambda}(\boldsymbol{G})$,

$$\|\boldsymbol{g}_k\|_2 \leq \|\boldsymbol{g}'_k\|_2 \;, \quad k = 1, \ldots, K \; . \tag{10}$$

In words, $\boldsymbol{G}$ is Minkowski-reduced if for $k = 1, \ldots, K$ the basis vector $\boldsymbol{g}_k$ has the minimum norm among all possible lattice points $\boldsymbol{g}'_k$ for which the set $\{\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_{k-1}, \boldsymbol{g}'_k\}$ can be extended to a basis of the lattice $\boldsymbol{\Lambda}(\boldsymbol{G})$. Hence, in contrast to the SMP (5) where only the K shortest *independent* lattice vectors have to be found, now the K shortest vectors have to be obtained that *additionally form a basis of the lattice*.

D. Algorithms for Lattice Basis Reduction and the SMP/SIVP

Algorithms for lattice basis reduction are known for quite some time and have partially been extended to the complex-valued case: The complex LLL reduction can, e.g., be performed with the CLLL algorithm in [6] and the complex HKZ reduction with the one in [7]. In [18], efficient algorithms for HKZ and MK reduction are given.[2] Recently, efficient strategies for solving the (C)SMP were proposed in [4], [5].

Only the suboptimal (C)LLL reduction ($\delta < 1$) has polynomial complexity.[3] To calculate an HKZ- or MK-reduced basis or to solve the SMP/SIVP, a solution to the NP-hard *shortest vector problem* has to be found once or several times.

[1] $\boldsymbol{\Lambda}(\boldsymbol{G}^{(k)})$ is the orthogonal projection of $\boldsymbol{\Lambda}(\boldsymbol{G})$ onto the orthogonal complement of $\{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{k-1}\}$.

[2] The reduction algorithms in [18] are described for the real-valued case, but they can easily be adapted to complex-valued lattices.

[3] For $\delta = 1$, at least the convergence of the LLL algorithm is ensured [2].

III. Optimal Factorization in LRA and IF Receiver-Side Linear Equalization

In the following, we introduce the channel model and review factorization approaches for conventional LRA, IF and an advanced LRA equalization strategy. The optimal lattice basis reduction is derived and the differences to IF are discussed.

A. MIMO Multiple-Access Channel

We consider a discrete-time MIMO multiple-access channel with K single-antenna non-cooperating transmitters (users) and a joint receiver with $N \geq K$ antennas (complex-baseband domain). The channel is described by the system equation

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{11}$$

and the respective system model is depicted in Fig. 1.

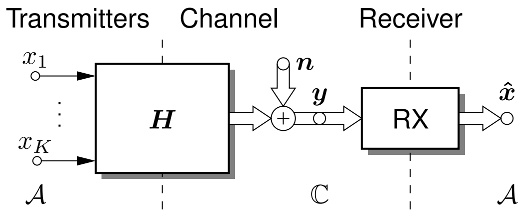

Fig. 1. System model for multi-user uplink transmission over the MIMO multiple-access channel: non-cooperating transmitters and joint receiver (RX).

In each time step, the transmitters radiate their data symbols x_k, $k = 1, \ldots, K$, in vector notation $\boldsymbol{x} = [x_1, \ldots, x_K]^{\mathsf{T}}$. The symbols are drawn from a zero-mean signal constellation $\mathcal{A}$ with variance σ_a^2, i.e., $\mathrm{E}\{|x_k|^2\} = \sigma_x^2 = \sigma_a^2$. The constellation has to form a subset of $\mathbb{G}$ (i.e., QAM constellations are suited).

The coefficients of the $N \times K$ channel matrix $\boldsymbol{H}$ are assumed to be i.i.d. zero-mean unit-variance complex Gaussian. Additive zero-mean white Gaussian noise with variance σ_n^2 is present at each receiving antenna. The (independent) noise components are combined into the vector $\boldsymbol{n} = [n_1, \ldots, n_N]^{\mathsf{T}}$.

Finally, on the basis of the incoming disturbed receive symbols $\boldsymbol{y} = [y_1, \ldots, y_N]^{\mathsf{T}}$, a joint receiver-side processing is used to obtain estimated data symbols $\hat{\boldsymbol{x}} = [\hat{x}_1, \ldots, \hat{x}_K]^{\mathsf{T}}$.

B. Lattice-Reduction-Aided Linear Receiver

The LRA receiver structure allows a full-diversity MIMO equalization. To this end, a channel factorization of the form

$$\boldsymbol{H} = \boldsymbol{H}_{\mathrm{r}}\boldsymbol{Z} \tag{12}$$

is performed, where $\boldsymbol{H}_{\mathrm{r}}$ denotes the reduced channel matrix and $\boldsymbol{Z}$ an integer matrix with elements drawn from $\mathbb{G}$. The LRA channel equalization is realized as illustrated in Fig. 2.

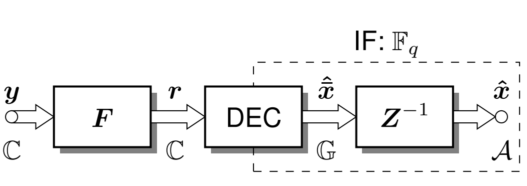

Fig. 2. Block diagram of the LRA linear receiver. The dashed framed part indicates the finite-field processing performed in the IF receiver instead.

First, the reduced channel $\boldsymbol{H}_{\mathrm{r}}$ is equalized via the $K \times N$ equalization matrix $\boldsymbol{F}$ (*non-integer equalization*). Second, channel decoding (DEC) is performed on the linearly equalized symbols $\boldsymbol{r} \in \mathbb{C}^K$, resulting in decoded symbols $\hat{\hat{\boldsymbol{x}}} \in \mathbb{G}^K$. In case of uncoded transmission, channel decoding is a simple quantization to $\mathbb{G}$. An *integer equalization* via the inverse integer matrix $\boldsymbol{Z}^{-1}$ finally results in the vector of estimated data symbols $\hat{\boldsymbol{x}} \in \mathcal{A}^K$.

C. Optimal Factorization for Conventional LRA Equalization

Classically, e.g., in [16], [14], [18], the channel factorization is directly realized according to (12). In this case, $\boldsymbol{\Lambda}(\boldsymbol{H})$ is considered, i.e., the lattice spanned by the channel matrix. The lattice basis reduction (6) is calculated for the generator matrix $\boldsymbol{G} = \boldsymbol{H}$, and hence $\boldsymbol{G}_{\mathrm{r}} = \boldsymbol{H}_{\mathrm{r}}$ and $\boldsymbol{U} = \boldsymbol{Z}$. The zero-forcing (ZF) non-integer linear equalization matrix reads $\boldsymbol{F} = \boldsymbol{H}_{\mathrm{r}}^{-1}$, or generally $\boldsymbol{F} = \boldsymbol{H}_{\mathrm{r}}^{+}$ if $\boldsymbol{H}$ is non-square.[4] In principle, any lattice-reduction approach can be employed. Most often, the suboptimal (C)LLL reduction is used to obtain a reduced basis. To assess/predict the transmission performance for different reduction algorithms, performance criteria have been given in the literature, e.g., in [18]: since MK- and HKZ-reduced bases $\boldsymbol{G}_{\mathrm{r}} = \boldsymbol{H}_{\mathrm{r}}$ have a lower average *orthogonality defect* or *proximity factor* when compared with LLL-reduced ones, the related transmissions achieve a lower bit-error rate (BER).

In [15], an extension of (12) to minimum mean-square error (MMSE) linear non-integer equalization has been proposed. Thereby, the $(N+K) \times K$ *augmented* channel matrix $\boldsymbol{\mathcal{H}} = \left[\begin{smallmatrix}\boldsymbol{H} \\ \sqrt{\zeta}\boldsymbol{I}\end{smallmatrix}\right]$ spans the lattice to be reduced, where $\boldsymbol{I}$ is the identity matrix and $\zeta = \sigma_n^2/\sigma_x^2$. Then, the performance may be assessed by evaluating properties of the *augmented reduced* channel matrix $\boldsymbol{\mathcal{H}}_{\mathrm{r}}$ as the reduced basis, whose pseudoinverse $\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{H}}_{\mathrm{r}}^{+}$ yields[5] the MMSE linear equalization matrix $\boldsymbol{F}$.

However, in terms of error-rate performance, the above approaches are suboptimal since substitutional performance criteria are used. Actually, the aim has to be to minimize the noise enhancement (ZF criterion) or the mean-square error (MMSE criterion), which are determined by the squared row norms of $\boldsymbol{F}$ or $\boldsymbol{\mathcal{F}}$, respectively. Restricting to the MMSE case, this aim can be achieved by the factorization [5]

$$(\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}} = (\boldsymbol{\mathcal{H}}_{\mathrm{r}}^{+})^{\mathsf{H}}\boldsymbol{Z}^{-\mathsf{H}} = \boldsymbol{\mathcal{F}}^{\mathsf{H}}\boldsymbol{Z}^{-\mathsf{H}} \,, \tag{13}$$

i.e., a lattice basis reduction with respect to the $(N+K) \times K$ generator matrix[6] $\boldsymbol{G} = (\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}}$, where $\boldsymbol{U} = \boldsymbol{Z}^{-\mathsf{H}}$. Then, the columns of $\boldsymbol{G}_{\mathrm{r}} = \boldsymbol{\mathcal{F}}^{\mathsf{H}}$, i.e., the rows of $\boldsymbol{\mathcal{F}} = [\boldsymbol{f}_1^{\mathsf{H}}, \ldots, \boldsymbol{f}_K^{\mathsf{H}}]^{\mathsf{H}}$, are directly affected. Substitutional performance criteria are not needed any more; the (squared) norms $\|\boldsymbol{f}_k^{\mathsf{H}}\|_2^2$ can directly be minimized. Noteworthy, the respective ZF-based approach $(\boldsymbol{H}^{+})^{\mathsf{H}} = \boldsymbol{F}^{\mathsf{H}}\boldsymbol{Z}^{-\mathsf{H}}$ has first been proposed in [13] and is obtained from (13) for $\zeta = 0$. We hence restrict to augmented matrices to cover both equalization criteria below.

[4] $\boldsymbol{A}^{+} = (\boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathsf{H}}$ denotes the left pseudoinverse of a matrix $\boldsymbol{A}$. For the inverse of the Hermitian $\boldsymbol{A}^{\mathsf{H}}$, we write $(\boldsymbol{A}^{\mathsf{H}})^{-1} = (\boldsymbol{A}^{-1})^{\mathsf{H}} = \boldsymbol{A}^{-\mathsf{H}}$.

[5] More precisely, $\boldsymbol{F}$ is the $K \times N$ left part of $\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{H}}_{\mathrm{r}}^{+}$.

[6] The lattice $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$ is the *dual lattice* [1] of $\boldsymbol{\Lambda}(\boldsymbol{\mathcal{H}})$. Thus, (13) is the dual factorization task of $\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{H}}_{\mathrm{r}}\boldsymbol{Z}$, cf. Table I in [5].

In LRA (receiver-side) equalization, the (error) performance is dominated by the maximum row norm of $\boldsymbol{\mathcal{F}}$, i.e., $\max_k \|\boldsymbol{f}_k^{\mathsf{H}}\|_2^2 \rightarrow \min$ has to be achieved. Utilizing the factorization approach (13), in terms of lattice basis reduction we hence have to solve the *shortest basis problem* (SBP)

$$\boldsymbol{Z}^{\mathsf{H}} = \underset{\substack{\boldsymbol{Z}^{\mathsf{H}} \in \mathbb{G}^{K\times K}, \\ |\det(\boldsymbol{Z}^{\mathsf{H}})|=1}}{\operatorname{argmin}} \max_{k=1,\dots,K} \left\|(\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}} \boldsymbol{z}_k\right\|_2^2 , \tag{14}$$

where $\boldsymbol{Z}^{\mathsf{H}} = [\boldsymbol{z}_1, \dots, \boldsymbol{z}_K]$. The question remains which reduction criterion should be used to solve the SBP for $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$.

Obviously, since LLL- or HKZ-reduced bases are defined by their Gram-Schmidt orthogonalization (cf. (7), (8), and (9)), the maximum norm of their basis vectors is generally not as short as possible. In contrast, an MK-reduced basis is directly defined by the length of its basis vectors. According to (10), the reduced basis consists of the K shortest lattice vectors that form a basis of the lattice. Consequently, not only the maximum row norm of $\boldsymbol{\mathcal{F}}$, but all of them are as short as possible. Restricting to lattice basis reduction, an MK reduction according to (13) thus results in an *optimal integer matrix* $\boldsymbol{Z}$. To the best knowledge of the authors this statement has not yet been given in the literature so far, even though respective numerical simulations have been performed in [10], [4] for comparisons of the LRA and IF receive strategy.

D. Integer-Forcing and Advanced LRA Linear Receiver

The unimodularity of $\boldsymbol{Z}$ has been queried by the concept of *integer-forcing* linear equalization [17]. The IF receiver employs the structure of the conventional LRA receiver—with the difference that channel (de)coding and integer equalization (via $\boldsymbol{Z}^{-1}$) are performed in a joint finite-field modulo arithmetic (dashed framed part in Fig. 2). To this end, *algebraic signal constellations* have to be applied that are isomorphic to finite fields $\mathbb{F}_q$, q a prime or a squared prime, cf. [11]. Since in a (finite) field an inverse of a matrix always exists as long as it has full rank, the unimodularity constraint on $\boldsymbol{Z}$ is weakened to $\operatorname{rank}(\boldsymbol{Z}) = K$. A detailed comparison of the LRA and IF receiver philosophy can be found in [5].

Actually, the strategy of dropping the unimodularity is not limited to a finite-field processing but can be generalized to the conventional LRA receiver [5]: If $\boldsymbol{Z}$ can be any full-rank integer matrix, then $|\det(\boldsymbol{Z})| \geq 1$. As a consequence, $\boldsymbol{Z}$ may not describe a change of basis any more, but at least a sublattice of $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$ is always described since $\det(\boldsymbol{Z}) \in \mathbb{G}$. An ordinary decoding/quantization to $\mathbb{G}$ can still be applied. After the equalization via $\boldsymbol{Z}^{-1}$ (which may have non-integer coefficients now), again the original lattice $\mathbb{G}$ is present as $\det(\boldsymbol{Z}^{-1}) = (\det(\boldsymbol{Z}))^{-1}$. A restriction to algebraic signal constellations is not necessary; the conventional LRA receiver can handle any full-rank matrix indeed. In the following, the relaxation to full-rank matrices for the LRA receiver structure is called *advanced LRA equalization*.

E. Optimal Factorization for IF and Advanced LRA Receiver

In principle, conventional LRA and IF or advanced LRA equalization share the same optimal factorization task (13).

TABLE I
OVERVIEW ON FACTORIZATION APPROACHES (LRA MMSE LIN. EQU.).

Property	classical LRA	optimal LRA	IF / adv. LRA
Lattice	$\boldsymbol{\Lambda}(\boldsymbol{\mathcal{H}})$	$\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$	$\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$
Constraint on $\boldsymbol{Z}$	$\lvert\det(\boldsymbol{Z})\rvert=1$	$\lvert\det(\boldsymbol{Z})\rvert=1$	$\operatorname{rank}(\boldsymbol{Z})=K$
Lattice Problem	SBP	SBP	SIVP
Algorithm	Usually (C)LLL	(C)MK	(C)SMP

The only difference is the relaxation to a non-unimodular $\boldsymbol{Z}$ in the IF or advanced LRA case,[7] weakening (14) to

$$\boldsymbol{Z}^{\mathsf{H}} = \underset{\substack{\boldsymbol{Z}^{\mathsf{H}} \in \mathbb{G}^{K\times K}, \\ \operatorname{rank}(\boldsymbol{Z}^{\mathsf{H}})=K}}{\operatorname{argmin}} \max_{k=1,\dots,K} \left\|(\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}} \boldsymbol{z}_k\right\|_2^2 . \tag{15}$$

Thus, the *shortest independent vector problem* as defined in (4) has to be solved for $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$. More specifically, $\max_k \|\boldsymbol{f}_k^{\mathsf{H}}\|_2 = \rho_K((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$ should be obtained.

It is quite evident that (15) is *optimally* solved if—as a stricter demand—the integer vectors of $\boldsymbol{Z}^{\mathsf{H}}$ yield the K successive minima of $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$. Then, similar to the MK reduction, all related lattice vectors are as short as possible. Since they, in contrast, do not necessarily form a basis of the lattice any more, a potential gain in performance is enabled.

In summary, a fair comparison is only achieved if IF or advanced LRA and optimal instead of classical LRA equalization are contrasted, i.e., the MK-reduced basis vs. the solution to the SMP for $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$, cf. Table I. In the literature [10], [4], some respective simulation results can be found which indicate a similar performance. Nevertheless, the exact role of the unimodularity has so far rather remained unclear.

IV. NUMERICAL RESULTS AND COMPARISON

In the following, we present numerical results to asses all mentioned lattice-reduction approaches and the solution to the SMP. The results are averaged over all users and more than one million channel realizations, each with a large number of transmit symbols and noise samples. We restrict to the optimal factorization task (13) unless otherwise specified.

A. Transmission Performance

We first asses the transmission performance. In order to focus on the channel equalization, we consider the (uncoded) LRA receiver structure as any constellation in $\mathbb{G}$ can be used and both the unimodular and non-unimodular case are covered.

In Fig. 3, the BER of 16QAM transmission is illustrated over the signal-to-noise ratio (SNR) in dB for the high-diversity case $K = N = 8$ applying both the ZF and MMSE equalization criterion. The SNR is expressed as energy per bit over the noise power density $E_{\mathrm{b,TX}}/N_0 = \sigma_x^2/(\sigma_n^2 \log_2(16))$. First, we see that the MMSE criterion generally results in a gain of more than $1\,\mathrm{dB}$ when compared with the ZF approach. In case of lattice basis reduction, the suboptimal LLL reduction ($\delta = 0.75$) shows the poorest performance.

[7]In the literature on IF receivers, $\boldsymbol{L}^{\mathsf{H}} = (\boldsymbol{H}^{\mathsf{H}}\boldsymbol{H} + \zeta\boldsymbol{I})^{-1/2}$ is most often factorized instead, e.g., in [17]. This approach is equivalent to (13), cf. [5].

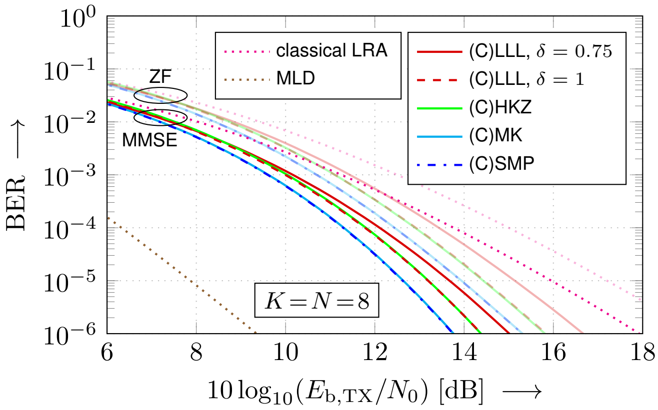

Fig. 3. BER of 16QAM transmission in dependency of $E_{\mathrm{b,TX}}/N_0$ in dB for both ZF and MMSE LRA linear equalization. Variation of the factorization algorithm; $\mathbf{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$; $K = N = 8$. The results of classical LRA equalization according to Table I (LLL, $\delta = 0.75$) and MLD are given.

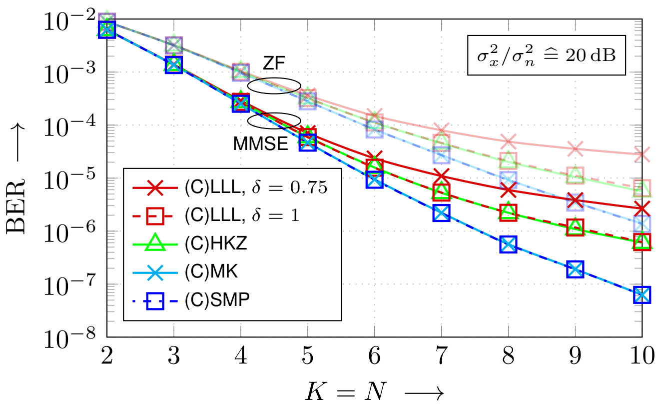

Fig. 4. BER of 16QAM transmission in dependency of $K = N$ for both ZF and MMSE criterion and $\sigma_x^2/\sigma_n^2 \widehat{=} 20\,\mathrm{dB}$, i.e., $E_{\mathrm{b,TX}}/N_0 \approx 14\,\mathrm{dB}$. Variation of the factorization algorithm for the LRA receiver; $\mathbf{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$.

Choosing $\delta = 1$ instead results in a considerable gain of about $0.5\,\mathrm{dB}$ (high-SNR regime); the performance is nearly the same as for HKZ reduction. Calculating the optimal basis via MK reduction results in about $0.5\,\mathrm{dB}$ additional gain. In contrast, dropping the unimodularity by solving the SMP does not show any further noticeable advantage. Both strategies (MMSE criterion) perform about $4\,\mathrm{dB}$ worse than a maximum-likelihood detection (MLD) via the sphere decoder [1]. The classical LRA equalization approach according to Table I (LLL, $\delta = 0.75$) significantly decreases the performance: the factorization of the (augmented) channel matrix instead of the inverse one induces a loss of more than $2\,\mathrm{dB}$.

In Fig. 4, the BER is depicted in dependency of the dimensions $K = N$ for an SNR of $\sigma_x^2/\sigma_n^2 \widehat{=} 20\,\mathrm{dB}$, i.e., $E_{\mathrm{b,TX}}/N_0 \approx 14\,\mathrm{dB}$ (16QAM transmission). As above, the superiority of the MMSE criterion is visible. For the case $K = N \leq 4$, all approaches almost perform the same, cf. [5]. When increasing the dimensions, the optimal LLL as well as the HKZ reduction show a gain when compared with the standard LLL one; the difference goes up to nearly one decade in BER ($K = N = 10$). Applying the MK reduction, an additional gain of about one decade is visible. In contrast, calculating the solution to the SMP again does not result in any significant further gain even for the high-dimensional case when $K = N = 10$.

TABLE II
PERCENTAGE OF CHANNELS WHERE MINKOWSKI BASIS SOLVES SMP.

SNR$\backslash K = N$	2	4	6	8	10
$\sigma_x^2/\sigma_n^2 \widehat{=} 15\,\mathrm{dB}$	100 %	99.0 %	95.7 %	90.3 %	83.8 %
$\sigma_x^2/\sigma_n^2 \widehat{=} 20\,\mathrm{dB}$	100 %	99.0 %	95.6 %	89.8 %	82.3 %
$\sigma_x^2/\sigma_n^2 \to \infty$	100 %	99.0 %	95.5 %	89.4 %	81.5 %

TABLE III
PERCENTAGE OF CHANNELS WHERE MINKOWSKI BASIS SOLVES SIVP.

SNR$\backslash K = N$	2	4	6	8	10
$\sigma_x^2/\sigma_n^2 \widehat{=} 15\,\mathrm{dB}$	100 %	99.2 %	97.0 %	94.0 %	90.6 %
$\sigma_x^2/\sigma_n^2 \widehat{=} 20\,\mathrm{dB}$	100 %	99.2 %	97.0 %	93.5 %	89.3 %
$\sigma_x^2/\sigma_n^2 \to \infty$	100 %	99.2 %	96.9 %	93.2 %	88.5 %

B. *Optimality of Minkowski Lattice Basis Reduction*

Obviously, the gain in performance induced by dropping the unimodularity is—at least in realistic scenarios—negligible. One reason for the minor impact is shown in Table II: There, the percentages of the channel realizations are listed for which the (complex) MK reduction also solves the SMP (5), i.e., where the K shortest independent lattice vectors form a basis of the lattice. We consider the mid- and high-SNR regime as well as the ZF approach ($\sigma_x^2/\sigma_n^2 \to \infty$). For the complex case and $K = N = 2$, an MK-reduced basis is always a solution to the SMP.[8] For $K = N = 4$, this is still the case for $99\,\%$ of the channels. Going up to $K = N = 10$, we still have more than $80\,\%$ in dependency of the actual SNR.[9]

However, still more relevant with respect to the performance is Table III, where the percentages of the channels are listed for which the (complex) MK reduction solves the SIVP (4). Now, $\max_k \|\boldsymbol{f}_k^{\mathsf{H}}\|_2^2 = \rho_K^2$ is sufficient since only the maximum norm actually determines the performance (cf. Sec. III).[10] For the case $K = N = 2$, the SIVP is consistently solved by the MK reduction, too. Considering $K = N = 4$, for additional $0.2\,\%$ of the channels we have the same maximum norm, but one or several others are increased when compared with the optimal non-unimodular solution. For $K = N = 10$, about $7\,\%$ of the channel realizations even fall into this category.

C. *Distribution of Maximum Squared Row Norm*

Though the unimodularity constraint still allows an optimal performance for the vast majority of channel realizations, the

[8]For real-valued lattices the K (squared) norms of MK-reduced basis vectors are bounded by $\rho_k^2 \leq \|\boldsymbol{g}_k\|_2^2 \leq \max\{1, (5/4)^{K-4}\}\,\rho_k^2$ [9]. For $K \leq 4$, they hence always yield the successive minima. Since an $N \times K$ complex lattice has an equivalent real-valued $2N \times 2K$ representation [4], for the complex case $K = N = 2$, the optimal matrix $\boldsymbol{Z}$ is always unimodular.

[9]For higher SNRs (including the ZF solution), the percentage is a little bit lower since a larger number of channels is badly conditioned, cf. [5].

[10]The channels in Table II are a subset of the ones in Table III, since not only a difference in the maximum norm but in any of the K norms is counted.

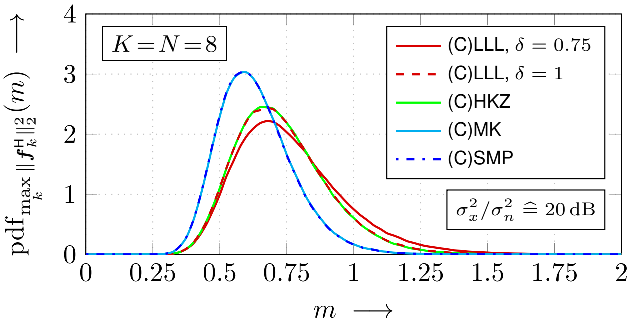

Fig. 5. Probability density function of the maximum squared norm of the non-integer equalization matrix for the LRA receiver; $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$. Variation of the factorization algorithm; $K = N = 8$; MMSE criterion ($\sigma_x^2/\sigma_n^2 \mathrel{\widehat{=}} 20\,\mathrm{dB}$).

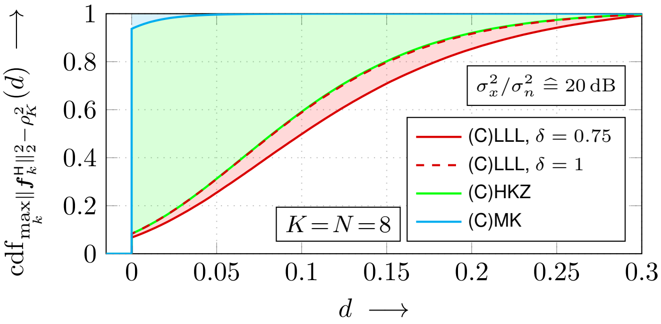

Fig. 6. Cumulative distribution function of the difference between the maximum squared norm of the non-integer equalization matrix of several lattice-reduction approaches and the squared K^{th} successive minimum (solution to (C)SMP/SIVP). $\boldsymbol{\Lambda}((\boldsymbol{\mathcal{H}}^{+})^{\mathsf{H}})$; $K = N = 8$; MMSE ($\sigma_x^2/\sigma_n^2 \mathrel{\widehat{=}} 20\,\mathrm{dB}$).

question of a potential gain for all the other channels remains. To clarify this, we first have a look at Fig. 5, where the probability density functions (pdfs) of $\max_k \|\boldsymbol{f}_k^{\mathsf{H}}\|_2^2$ are shown. The relations between these pdfs are also reflected in the relations of the BER curves in Fig. 4 (MMSE, $\sigma_x^2/\sigma_n^2 \mathrel{\widehat{=}} 20\,\mathrm{dB}$, $K = N = 8$). Applying suboptimal LLL reduction, high magnitudes ($\max_k \|\boldsymbol{f}_k^{\mathsf{H}}\|_2^2 > 1$) are more likely than for optimal LLL or HKZ reduction. For MK reduction and the solution to the SMP/SIVP ($\max_k \|\boldsymbol{f}_k^{\mathsf{H}}\|_2^2 = \rho_K^2$) high magnitudes barely appear; the squared norm is lowered on average. However, again hardly any difference between both is present—even though at least for some channels ρ_K^2 should be lower.

The reason for that finally becomes clear in Fig. 6, where the cumulative distribution function (cdf) of the difference between $\max_k \|\boldsymbol{f}_k^{\mathsf{H}}\|_2^2$ of each reduced basis and ρ_K^2 (SMP/SIVP) is illustrated (cf. [4, Fig. 6]; parameters as in Fig. 5). Considering LLL and HKZ reduction, for roughly $10\,\%$ of the channels both quantities are the same; for all others the difference is mostly located in the range up to 0.3. In contrast, an MK basis is optimal for nearly $94\,\%$ of the channels (cf. Table III). However—most important—in almost all other cases the difference is smaller than 0.05, i.e., negligible. For that reason, the pdfs in Fig. 5 are nearly the same, leading to an almost identical transmission performance.

V. Summary and Conclusions

We have reviewed and compared factorization approaches for LRA and IF linear receivers. Restricting to lattice basis reduction, we have derived the optimal factorization criterion which is solved by Minkowski reduction. The possibility of dropping the unimodularity constraint by solving the successive minima problem has been discussed. In theory, for special channel matrices, this may result in an infinite gain if $K \to \infty$, cf. the example in [17]. Nevertheless, numerical results have revealed that—even for the high-dimensional case—the gain is negligible in practice: the Minkowski reduction is most often already optimal; if differences occur they are barely relevant.

Further gains in performance may be achieved by utilizing the complex hexagonal lattice, also known as *Eisenstein integers*, cf. [11], [12]. In addition, an adaption to LRA decision-feedback equalization still has to be investigated, as well as to the dual scenario of the MIMO broadcast channel, cf. [12].

References

[1] E. Agrell, T. Eriksson, A. Vardy, K. Zeger. Closest Point Search in Lattices. *IEEE Trans. Inf. Theory* **48** (8), pp. 2201–2214, Aug. 2002.

[2] A. Akhavi. The Optimal LLL Algorithm is Still Polynomial in Fixed Dimension. *Theoretical Computer Science* **297**, pp. 3–23, Mar. 2003.

[3] J.W.S. Cassels. *An Introduction to the Geometry of Numbers*. Springer Berlin/Heidelberg, Reprint of the 1971 Edition, 1997.

[4] L. Ding, K. Kansanen, Y. Wang, J. Zhang. Exact SMP Algorithms for Integer-Forcing Linear MIMO Receivers. *IEEE Trans. Wireless Commun.* **14** (12), pp. 6955–6966, Dec. 2015.

[5] R.F.H. Fischer, M. Cyran, S. Stern. Factorization Approaches in Lattice-Reduction-Aided and Integer-Forcing Equalization. Proc. *Int. Zurich Seminar on Communications*, Zurich, Switzerland, Mar. 2016.

[6] Y.H. Gan, C. Ling, W.H. Mow. Complex Lattice Reduction Algorithm for Low-Complexity Full-Diversity MIMO Detection. *IEEE Trans. Signal Process.* **57** (7), pp. 2701–2710, July 2009.

[7] H. Jiang, S. Du. Complex Korkine-Zolotareff Reduction Algorithm for Full-Diversity MIMO Detection. *IEEE Commun. Letters* **17** (2), pp. 381–384, Feb. 2013.

[8] J.C. Lagarias, H.W. Lenstra, C.P. Schnorr. Korkin-Zolotarev Bases and Successive Minima of a Lattice and its Reciprocal Lattice. *Combinatorica* **10** (4), pp. 333–348, 1990.

[9] H. Minkowski. Über die positiven quadratischen Formen und über kettenbruchähnliche Algorithmen. *Journal für die reine und angewandte Mathematik* **107**, pp. 278–297, 1891.

[10] A. Sakzad, J. Harshan, E. Viterbo. Integer-Forcing MIMO Linear Receivers Based on Lattice Reduction. *IEEE Trans. Wireless Commun.* **12** (10), pp. 4905–4915, Oct. 2013.

[11] S. Stern, R.F.H. Fischer. Lattice-Reduction-Aided Preequalization over Algebraic Signal Constellations. Proc. *Int. Conf. on Signal Processing and Communication Systems*, Cairns, Australia, Dec. 2015.

[12] S. Stern and R.F.H. Fischer. Advanced Factorization Strategies for Lattice-Reduction-Aided Preequalization. Proc. *IEEE Int. Symposium on Information Theory*, pp. 1471–1475, Barcelona, Spain, July 2016.

[13] M. Taherzadeh, A. Mobasher, A.K. Khandani. LLL Reduction Achieves the Receive Diversity in MIMO Decoding. *IEEE Trans. Inf. Theory* **53** (12), pp. 4801–4805, Dec. 2007.

[14] C. Windpassinger, R.F.H. Fischer. Low-Complexity Near-Maximum-Likelihood Detection and Precoding for MIMO Systems using Lattice Reduction. Proc. *IEEE Information Theory Workshop*, Paris, France, pp. 345–348, Mar. 2003.

[15] D. Wübben, R. Böhnke, V. Kühn, K.D. Kammeyer. Near-Maximum-Likelihood Detection of MIMO Systems using MMSE-Based Lattice Reduction. Proc. *IEEE Int. Conference on Communications*, pp. 798–802, Paris, France, June 2004.

[16] H. Yao, G.W. Wornell. Lattice-Reduction-Aided Detectors for MIMO Communication Systems. Proc. *IEEE Global Telecommunications Conference*, Taipei, Taiwan, Nov. 2002.

[17] J. Zhan, B. Nazer, U. Erez, M. Gastpar. Integer-Forcing Linear Receivers. *IEEE Trans. Inf. Theory* **60** (12), pp. 7661–7685, Dec. 2014.

[18] W. Zhang, S. Qiao, Y. Wei. HKZ and Minkowski Reduction Algorithms for Lattice-Reduction-Aided MIMO Detection. *IEEE Trans. Signal Process.* **60** (11), pp. 5963–5976, Nov. 2012.