Full Diversity Unitary Precoded Integer-Forcing

Amin Sakzad and Emanuele Viterbo[†]

Abstract

We consider a point-to-point flat-fading MIMO channel with channel state information known both at transmitter and receiver. At the transmitter side, a lattice coding scheme is employed at each antenna to map information symbols to independent lattice codewords drawn from the same codebook. Each lattice codeword is then multiplied by a unitary precoding matrix $\mathbf{P}$ and sent through the channel. At the receiver side, an integer-forcing (IF) linear receiver is employed. We denote this scheme as *unitary precoded integer-forcing* (UPIF). We show that UPIF can achieve full-diversity under a constraint based on the shortest vector of a lattice generated by the precoding matrix $\mathbf{P}$. This constraint and a simpler version of that provide design criteria for two types of full-diversity UPIF. Type I uses a unitary precoder that adapts at each channel realization. Type II uses a unitary precoder, which remains fixed for all channel realizations. We then verify our results by computer simulations in 2×2, and 4×4 MIMO using different QAM constellations. We finally show that the proposed Type II UPIF outperform the MIMO precoding X-codes at high data rates.

Index Terms

MIMO, Integer-Forcing, unitary precoding, lattice codes, full-diversity.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) channels with multiple antennas appeared in the early 1990's as the key technology to achieve high spectral efficiencies in wireless channels. These

channels are known to obtain higher capacity gains in comparison to single antenna point-to-point channels. For a MIMO channel matrix $\mathbf{H}_{n_t \times n_r}$ with n_t transmit antennas and n_r receive antennas, a *diversity gain* of at most $n_t n_r$ is achievable. If an encoder/decoder pair gives the maximum possible diversity $n_t n_r$, then the entire system is said to achieve *full-diversity*. Let n_s denote the number of symbols transmitted per channel use. If $n_s = \min\{n_r, n_t\}$, then the encoding scheme is said to have *full-rate*.

For MIMO channels with channel state information at the receiver (CSIR), carefully designed space-time block-codes (STBC) can provide full-diversity [1]. Algebraic full-diversity full-rate STBCs with large coding gain were later designed in [2], [3]. The design criteria for these codes are known as the *rank* and the *non-vanishing determinant* criteria [1], [3]. In order to take full advantage from these codes, an optimal maximum likelihood (ML) detector such as a sphere decoder [4], with high computational complexity should be implemented at the receiver side. Other well-known linear receivers [5], including zero-forcing (ZF) and minimum mean square estimation (MMSE) linear receivers, trade-off error performance for reduced computational complexity. Specific families of full-diversity STBCs were investigated and constructed in [6]–[11] for use with linear receivers (ZF and MMSE). In these schemes, full-diversity is obtained at the cost of limited rate.

Precoding schemes that assume CSIR, channel state information at the transmitter (CSIT), and an ML decoder usually make use of a singular value decomposition (SVD) to convert the MIMO channel into parallel subchannels. Then appropriate precoders employed over these parallel channels to minimize the error probability of the system over different constellation sizes. Some examples are the well-known full-diversity E-$d_{\min}$ precoders for 4-QAM [12], and MIMO precoding X- and Y- codes [13], which are known to provide a diversity order of $d_{n_t,n_r,n_s} = \left(n_r - \frac{n_s}{2} + 1\right)\left(n_t - \frac{n_s}{2} + 1\right)$, and the full-diversity precoders introduced and investigated in [14]. Note that d_{n_t,n_r,n_s} shows a trade-off between full-rate and full-diversity. In one hand, if $n_t = n_r$ and we use a full-rate MIMO precoding X- or Y- code, then the diversity order is $d_{n_t,n_r,n_s} = \left(\frac{n_t}{2} + 1\right)^2 < n_t^2$ (full-diversity). On the other hand, if we want to ensure full-diversity for MIMO precoding X- or Y- codes, we need $n_s = 2$, which means the system has very limited rate. For a detailed comparison between the ML-decoding complexity of the above three schemes, we refer

the reader to Table III of [14].

On the other hand, in the presence of CIST and having a linear receiver such as MMSE linear receiver at the destination, optimal precoder designs have been extensively studied ([15]–[18]). The design criteria for these schemes include maximizing the rate or minimizing the error performance. The MIMO linear precoding techniques such as regularized ZF [19] are alternative approaches that can provide full-diversity in conjunction with linear receivers, see [20] and references therein. The decoding complexity of such schemes is relatively low in comparison to ML-decoding algorithms but incur some error performance degradation.

In this paper, assuming CSIT and CSIR we propose a full-diversity unitary precoding scheme for $n \times n$ MIMO ($n_r = n_t = n$), where the receiver is equipped with the recently proposed integer-forcing (IF) linear receiver [21]. In IF framework, the transmitter employs a lattice coding scheme, and sends independent lattice codewords from the same codebook, across different transmit antennas (or *layers*). At each receive antenna, we recover an integer linear combination of the lattice codewords depending on the entries of the channel matrix $\mathbf{H}$. As all the lattice codewords are drawn from the same lattice codebook, the integer linear combinations are themselves guaranteed to be lattice points. To further estimate the original transmitted lattice codewords, a system of linear equations with integer coefficient matrix $\mathbf{A}$ should be solved at the receiver. In fact, the IF linear receiver exploits the linearity of lattices to first eliminate the noise and then hareness the interference between transmitted data across different layers. The key step underpinning IF is the selection of $\mathbf{A}$ to approximate the channel matrix $\mathbf{H}$. Hence, a matrix $\mathbf{B}$ is needed such that $\mathbf{A}$ is full rank and $\mathbf{BH} \approx \mathbf{A}$, with minimum quantization error at high signal-to-noise ratio (SNR) values. The problem of finding $\mathbf{A}$ based on $\mathbf{H}$ using lattice reduction algorithms is addressed in [22], [23]. The integer-forcing linear receiver is also shown to provide a full receive diversity order $n_r = n$ and a full multiplexing gain in [21], [22].

Using perfect STBCs like the Golden code [25], [3], the authors of [26] have shown that under an IF decoder, a constant gap to capacity is attainable. A perfect STBC code is a linear dispersion space-time code over a Quadrature-Amplitude Modulation (QAM) constellation, which is full-rate, satisfies non-vanishing determinant property, and its generator matrix is a unitary matrix. In this paper, we aim at relaxing the first two constraints of perfect STBCs and keeping the last one

only. We analyze the diversity order of a lattice space-time encoding scheme, which is precoded by a unitary matrix and uses an IF linear receiver. First, singular value decomposition (SVD) precoding is used to diagonalize the channel, and then an optimal precoder matrix is designed. Our optimization criterion is based on maximizing the minimum distance of the input lattice, similar to the approach of [27]. We note that differently from [27], our precoder matrices are unitary and our decoder is the IF linear receiver and not the ML receiver. Related works focus on a different design criterion based on maximizing the mutual information ([28], [29], [30], and [31] and references therein).

From the receiver point of view, using IF linear receiver instead of ML decoders in slow-fading channels is advantageous in terms of complexity. For the IF receiver, the lattice reduction algorithm is performed only at the beginning of each quasi-static interval for which the channel $\mathbf{H}$ is assumed to remain constant over multiple codeword transmissions. Subsequently, for each codeword transmission, only a system of linear equations has to be solved to recover an estimate of that codeword. On the other hand, for the ML decoding, a sphere decoder should be used for each codeword within a quasi-static channel interval.

We summarize the contributions of this paper as follows:

- **Unitary precoded integer-forcing (UPIF)** – Considering the availability of CSIT, the SVD decomposition $\mathbf{W\Sigma V}^h$ for $\mathbf{H}_{n \times n}$ is first performed. Using $\mathbf{V}$ at the transmitter and $\mathbf{W}$ at the receiver, the channel is reduced to a diagonal channel matrix $\mathbf{\Sigma}$. An additional unitary precoding matrix $\mathbf{P}$, is then employed as $\mathbf{VP}$. At the destination, the receiver is equipped with an IF linear receiver.

- **Full-diversity constraint** – We find an upper bound on the probability of error of UPIF (Theorem 1) and show analytically that if $\mathbf{P}$ satisfies a specific constraint (Theorem 2), the full diversity (n^2) can be achieved by our UPIF. The exact constraint is hard to verify, since it is related to finding the minimum distance of a lattice depending on $\mathbf{\Sigma}$ and $\mathbf{P}$. Theorem 3 provides a simpler constraint, based on the non-vanishing *minimum product distance* of the lattice generated by $\mathbf{P}$ and guarantees the full-diversity of UPIF.

- **Precoder design criteria** – The constraints on $\mathbf{P}$ provide the design criteria for full-diversity-achieving precoders for IF MIMO linear receivers. We define two types of full-

diversity UPIF schemes: (*i*) the precoder matrix $\mathbf{P}$ depending on the channel matrix singular values (diagonal elements of $\mathbf{\Sigma}$) is adapted at each quasi-static channel interval (*Type I UPIF*), (*ii*) a fixed unitary precoder $\mathbf{P}$ (independent of $\mathbf{\Sigma}$) is used for all channel realizations (*Type II UPIF*). For Type I UPIF, finding an analytical optimization technique for $\mathbf{P}$ remains an open problem. For Type II UPIF we propose the use of full-diversity lattice generator matrices with non-vanishing minimum product distance [32]-[37].

- **Simulation results** – Monte-Carlo simulations are presented in support the derived theoretical results. For example, it is shown that a 2×2 full-diversity algebraic rotation matrix given in [37] in a Type II UPIF scheme with a 64-QAM constellation, can perform as close as 1dB away from ML decoder for the same code and 2dB better than MIMO precoding X-codes [13].

Notation. The superscripts $(\cdot)^T$ and $(\cdot)^h$ denote transposition and Hermitian transposition, respectively. Let $\mathcal{G}$ and $\mathcal{G}'$ be a group and its subgroup respectively, then $\mathcal{G}/\mathcal{G}'$ denotes the quotient group. The sets $\mathbb{Z}$, $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{Z}[i]$ denote the ring of rational integers, the field of complex numbers, the field of real numbers, and the ring of Gaussian integers, respectively, where $i^2 = -1$. Real and imaginary parts of a complex number z are denoted by $\Re(z)$ and $\Im(z)$, and $|z|$ and $\arg(z)$ are the modulus and the unique phase, respectively. The notation $\|\mathbf{v}\|$ stands for the Euclidean norm of a vector $\mathbf{v} \in \mathbb{C}^n$. A $k \times k$ matrix $\mathbf{X} = [\mathbf{x}_1^T, \ldots, \mathbf{x}_k^T]^T$ is formed by stacking the $k-$dimensional row vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$, and $\mathbf{I}_k$ denotes the $k \times k$ identity matrix. Finally, let $\mathbf{v}$ be a vector, then its j-th entry is represented by $[\mathbf{v}]_j$.

II. System Model

We first review the notion of complex lattices [38] which are essential for the rest of the paper.

Definition 1: Given a set of d-dimensional vectors $\{\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_k\}$ in $\mathbb{C}^d$, a k-dimensional *complex lattice* Λ is the set of points

$$\left\{ \sum_{m=1}^{k} z_m \boldsymbol{\ell}_m : z_m \in \mathbb{Z}[i], \ \boldsymbol{\ell}_m \in \mathbb{C}^d \right\}.$$

More briefly we can write $\Lambda = \left\{ \mathbf{z}\mathbf{L} \colon \mathbf{z} \in \mathbb{Z}[i]^k \right\}$, where $\mathbf{L} = [\boldsymbol{\ell}_1^T, \ldots, \boldsymbol{\ell}_k^T]^T$ represents a *basis* of the lattice Λ. In other words, every point $\mathbf{x} \in \Lambda$ can be represented as a Gaussian integer linear combination of basis vectors. For any point $\mathbf{x} \in \Lambda$ the *Voronoi cell* $\mathcal{V}(\mathbf{x})$ is

$$\left\{ \mathbf{v} = \sum_{m=1}^{k} \alpha_m \boldsymbol{\ell}_m \colon \|\mathbf{v} - \mathbf{x}\| \leq \|\mathbf{v} - \mathbf{y}\|, \ \forall \mathbf{y} \in \Lambda, \ \alpha_m \in \mathbb{C} \right\}.$$

The Voronoi cell for the origin is denoted by $\mathcal{V}$. Let $\mathbf{L}$ be a matrix with $\boldsymbol{\ell}_m$ as its rows, then $\mathbf{L}$ is called the *generator matrix* of the lattice $\Lambda_{\mathbf{L}}$. Throughout the paper, we only consider full rank lattices where $d = k$. For example $\mathbb{Z}[i]^d$ is a lattice with the identity matrix as generator matrix. Let $\mathcal{S} \subseteq \mathbb{C}^n$, then the linear span of $\mathcal{S}$ is defined as:

$$\mathrm{span}(\mathcal{S}) \triangleq \left\{ \sum_{p=1}^{q} c_p \mathbf{v}_p \colon q \in \mathbb{N}, \ \mathbf{v}_p \in \mathcal{S}, \ c_p \in \mathbb{C} \right\}.$$

Let

$$\mathcal{B}_r(\mathbf{0}) = \left\{ \mathbf{x} \in \mathbb{C}^d \colon \|\mathbf{x}\| \leq r \right\},$$

be the d-dimensional ball of radius r centered at $\mathbf{0}$ and $\dim(\mathcal{V})$ denotes the dimension of a subspace $\mathcal{V}$ of $\mathbb{C}^d$. For a d-dimensional lattice $\Lambda_{\mathbf{L}}$, we define the m-th successive minima, for $1 \leq m \leq d$ as

$$\epsilon_m(\Lambda_{\mathbf{L}}) \triangleq \inf \left\{ r \colon \dim\left(\mathrm{span}\left(\Lambda_{\mathbf{L}} \cap \mathcal{B}_r(\mathbf{0})\right)\right) \geq m \right\}. \tag{1}$$

The m-th successive minima of $\Lambda_{\mathbf{L}}$ is the infimum radius r such that there are m independent vectors of $\Lambda_{\mathbf{L}}$ in $\mathcal{B}_r(\mathbf{0})$.

A d-dimensional lattice $\Lambda_{\mathbf{L}}$ is called *full-diversity* if for all distinct $\mathbf{x}, \mathbf{y} \in \Lambda_{\mathbf{L}}$, $[\mathbf{x}]_m \neq [\mathbf{y}]_m$ for all $1 \leq m \leq d$. The *minimum product distance* of a full-diversity lattice $\Lambda_{\mathbf{L}}$ is denoted by $d_{p,\min}(\Lambda_{\mathbf{L}})$ ([35]):

$$d_{p,\min}(\Lambda_{\mathbf{L}}) \triangleq \min_{\mathbf{x} \in \Lambda_{\mathbf{L}} \setminus \{\mathbf{0}\}} \prod_m \left| [\mathbf{x}]_m \right|. \tag{2}$$

A *lattice code* $\mathcal{C} \subseteq \Lambda$ is a finite set of points of Λ. One way of constructing lattice codes is through sublattices (nested lattices). A subset $\Lambda' \subseteq \Lambda$ is called a *sublattice* if Λ' is a lattice itself. Given a sublattice Λ', we define the quotient Λ/Λ' as a lattice code. The fine lattice Λ is called

the *coding lattice* and the coarse lattice Λ' is usually referred as the *shaping lattice*. The lattice code $\mathcal{C} = \Lambda/\Lambda'$ corresponds to a finite constellation of lattice points carved from the lattice Λ, within a fixed Voronoi cell of Λ'. In other words, let $\mathcal{V}'$ be the Voronoi cell of origin for lattice Λ', then $\mathcal{C} = \Lambda \cap \mathcal{V}'$. Therefore, the shape of such constellation is governed by $\mathcal{V}'$. For example, the Voronoi region of the 4-dimensional checkerboard lattice

$$\Lambda' = D_4 \triangleq \left\{ (x_1, \cdots, x_4) \in \mathbb{Z}^4 \colon \sum_{m=1}^{4} x_i = \text{even} \right\},$$

has exactly 24 faces, with the relevant vectors being the 24 minimum Euclidian lattice points [39]. This lattice is a sublattice of 4-dimensional integer lattice, *i.e.* $D_4 \subseteq \mathbb{Z}^4$ and hence so is the scaled lattice

$$8D_4 \triangleq \{8\mathbf{x} \colon \mathbf{x} \in D_4\} \subseteq \mathbb{Z}^4.$$

The lattice code $\Lambda/\Lambda' = \mathbb{Z}^4/8D_4$ includes all 4-dimensional integer points inside the Voronoi region of the lattice $8D_4$.

In general, a common choice for the sublattice Λ' of Λ is the scaled lattice $g\Lambda$ for some integer $g \in \mathbb{Z}$ as in [38]. As another simpler example, for $\Lambda = \mathbb{Z}[i]^d$ and $\Lambda' = g\mathbb{Z}[i]^d$, we have $\Lambda/\Lambda' = \mathbb{Z}_g[i]$ for which $\mathbb{Z}_g$ is the ring of integers modulo g. For $g = 2$ the cubic shaped lattice code $\mathbb{Z}_2[i]$ is simply a 4-QAM constellation. For the rest of this paper we use $\Lambda = \mathbb{Z}[i]^d$ and choose g to be a power of 2.

We consider a quasi-static flat-fading $n \times n$ MIMO channel, where the channel state information is available at both transmitter and receiver. The channel matrix is denoted by $\bar{\mathbf{H}} \in \mathbb{C}^{n \times n}$, where the entries of $\bar{\mathbf{H}}$ are i.i.d. complex Gaussian random variables $\sim \mathcal{CN}(0, 1)$. We use an n-layer lattice coding scheme, where the information transmitted across different antennas are independent. For $1 \leq m \leq n$, the m-th layer is equipped with a lattice encoder $E \colon \mathcal{R}^k \to \mathbb{C}^n$ which maps a message $\bar{\mathbf{s}}_m \in \mathcal{R}^k$ over the ring $\mathcal{R}$ into a lattice codeword $\bar{\mathbf{x}}_m \in \Lambda/\Lambda' \subset \mathbb{C}^n$ in the complex space. The matrix $\bar{\mathbf{X}} = [\bar{\mathbf{x}}_1^T, \ldots, \bar{\mathbf{x}}_n^T]^T$ is actually a space-time codeword, where its rows are all lattice codewords. This lattice space-time codeword will then be precoded using a unitary matrix. The precoding matrix $\bar{\mathbf{U}}$ can be derived using the components of the channel matrix $\bar{\mathbf{H}}$ and another unitary matrix $\bar{\mathbf{P}}$, to be optimized later. Together with $\bar{\mathbf{U}}$, the matrix $\bar{\mathbf{X}}$

forms a space-time codeword $\bar{\mathbf{U}}\bar{\mathbf{X}}$ to be sent through the channel.

Let $\bar{\mathbf{H}} = \bar{\mathbf{W}}\bar{\mathbf{\Sigma}}\bar{\mathbf{V}}^h$ be the singular value decomposition (SVD) of the channel matrix where $\bar{\mathbf{W}}, \bar{\mathbf{V}} \in \mathbb{C}^{n \times n}$ are two unitary matrices and $\bar{\mathbf{\Sigma}}$ is a diagonal matrix given by $\bar{\mathbf{\Sigma}} = \mathrm{diag}(\bar{\sigma}_1, \ldots, \bar{\sigma}_n)$ with $\bar{\sigma}_1 \geq \cdots \geq \bar{\sigma}_n$ all in $\mathbb{R}$. A unitary precoder matrix

$$\bar{\mathbf{U}} = \bar{\mathbf{V}}\bar{\mathbf{P}} \tag{3}$$

is then employed at the transmitter where $\bar{\mathbf{P}} \in \mathbb{C}^{n \times n}$ is a unitary matrix that needs to be optimized. If $\bar{\mathbf{X}} = [\bar{\mathbf{x}}_1^T, \ldots, \bar{\mathbf{x}}_n^T]^T$ denotes the matrix of transmitted vectors, the received signal $\bar{\mathbf{Y}}$ is given by

$$\bar{\mathbf{Y}} = \sqrt{\rho} \cdot \bar{\mathbf{H}}\bar{\mathbf{U}}\bar{\mathbf{X}} + \bar{\mathbf{Z}}, \tag{4}$$

where $\rho = \frac{\mathrm{SNR}}{n}$ denotes the average signal-to-noise ratio at each receive antenna and the entries of $\bar{\mathbf{Z}}$ are i.i.d. distributed as $\mathcal{CN}(0, 1)$. Let $\bar{\mathbf{M}}$ be the generator matrix of a complex lattice $\Lambda_{\bar{\mathbf{M}}}$, then by using the standard conversion from complex lattice to the equivalent real one, we have the real lattice generator matrix as:

$$\mathbf{M} = \begin{pmatrix} \Re(\bar{\mathbf{M}}) & \Im(\bar{\mathbf{M}}) \\ -\Im(\bar{\mathbf{M}}) & \Re(\bar{\mathbf{M}}) \end{pmatrix}.$$

With the above transformation, (4) can be written as

$$\mathbf{Y} = \sqrt{\rho} \cdot \mathbf{H}\mathbf{U}\mathbf{X} + \mathbf{Z}, \tag{5}$$

where all the matrices $\mathbf{Y}, \mathbf{H}, \mathbf{U}, \mathbf{X}$, and $\mathbf{Z}$ are in $\mathbb{R}^{2n \times 2n}$. A suitable block diagram is then as Fig. 1. Upon receiving $\mathbf{Y}$ at the destination, we multiply it by $\mathbf{W}^h$ to get $\mathbf{Y}' \triangleq \mathbf{W}^h\mathbf{Y}$. Substituting $\mathbf{U}$ from (3) into (5) the channel can be modeled as:

$$\mathbf{Y}' = \sqrt{\rho} \cdot \mathbf{\Sigma}\mathbf{P}\mathbf{X} + \mathbf{Z}', \tag{6}$$

where $\mathbf{Z}' = \mathbf{W}^h\mathbf{Z}$. Note that $\mathbf{Z}'$ continues to be an i.i.d. complex Gaussian matrix with entries $\sim \mathcal{CN}(0, 1)$ because $\mathbf{W}$ is unitary.

At this point, the receiver employs integer-forcing. That is a linear receiver architecture that
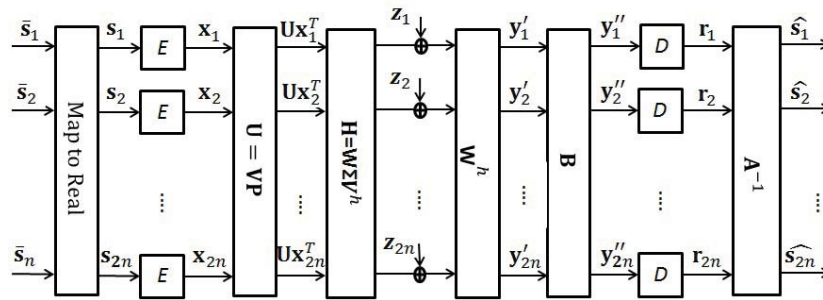
Fig. 1. Real Block diagram of unitary precoded integer-forcing.

creates an effective integer-valued channel matrix. The entries of this effective channel matrix are used to recover integer combinations of the codewords. This is in contrast to the other linear receivers and ML-based decoders, which attempt to recover the transmitted codewords directly. A feature of integer-forcing linear receiver is in using the same lattice codes at both the transmitter as well as the receiver. In this framework, the sender uses a layered transmission scheme, and transmits independent lattice codewords simultaneously across the layers. At the receiver side, each layer is allowed to decode an integer linear combination of transmitted lattice codewords. Since any integer linear combination of lattice points is another lattice point, the decoded point will be another lattice point. Assuming that the effective channel matrix is full rank, these integer combinations can then be solved for the original lattice codewords.

The goal of integer-forcing linear receiver is to project ΣP (by left multiplying it with a receiver filtering matrix $\mathbf{B}$) onto a non-singular integer matrix $\mathbf{A}$. In order to uniquely recover the information symbols, the matrix $\mathbf{A}$ must be invertible over the ring $\mathcal{R}$. Thus, we have

$$\mathbf{Y}'' = \mathbf{B}\mathbf{Y}' = \sqrt{\rho} \cdot \mathbf{B}\Sigma\mathbf{P}\mathbf{X} + \mathbf{B}\mathbf{Z}'. \tag{7}$$

For the IF receiver [21] formulation, the model is given by

$$
\begin{aligned}
\mathbf{Y}'' &= \sqrt{\rho} \cdot \mathbf{A}\mathbf{X} + \underbrace{\sqrt{\rho} \cdot (\mathbf{B}\Sigma\mathbf{P} - \mathbf{A})\mathbf{X}}_{\text{quantization noise term}} + \mathbf{B}\mathbf{Z}' \\
&\triangleq \sqrt{\rho} \cdot \mathbf{A}\mathbf{X} + \mathbf{E} ,
\end{aligned} \tag{8}
$$

where $\sqrt{\rho} \cdot \mathbf{AX}$ is the desired signal component, and the effective noise is

$$\sqrt{\rho} \cdot (\mathbf{B\Sigma P} - \mathbf{A})\mathbf{X} + \mathbf{BZ}'. \tag{9}$$

We further denote the effective noise term along the m-th layer by $\mathbf{e}_m$ and by stacking these row vectors, we get $\mathbf{E}$. The average energy of the effective noise along the m-th row of $\mathbf{Y}''$ is defined as

$$G(\mathbf{a}_m, \mathbf{b}_m) \triangleq \rho \|\mathbf{b}_m \mathbf{\Sigma P} - \mathbf{a}_m\|^2 + \|\mathbf{b}_m\|^2, \tag{10}$$

where $\mathbf{a}_m$ and $\mathbf{b}_m$ denote the m-th row of $\mathbf{A}$ and $\mathbf{B}$, respectively. Note that in order to increase the m-th layer effective signal-to-noise ratio [21]

$$\mathrm{SNR_{eff}} = \frac{\rho}{G(\mathbf{a}_m, \mathbf{b}_m)} \tag{11}$$

the term $G(\mathbf{a}_m, \mathbf{b}_m)$ has to be minimized, for each m, by appropriately selecting the matrices $\mathbf{A}$ and $\mathbf{B}$. We refer to the above signal model as *unitary precoded integer-forcing*. Note that the signal strength in the numerator of (11) is simply ρ (instead of $\rho\|\mathbf{a}_m\|^2$) due to the fact that $\mathbf{a}_m^T \mathbf{x}$ is itself a codeword, which is decodable up to an effective noise variance of ρ. Specifically, for lattice codes, it is often easier to think of each codeword as surrounded by a ball of radius $\sqrt{n\rho}$. If the noise falls within this ball, then decoding is successful. The received signal-strength viewpoint is more appropriate for i.i.d. random coding methods, where a joint typicality analysis can be applied.

A. Decoding Complexity Advantage of UPIF

We highlight the decoding complexity advantage of UPIF in comparison to ML decoder techniques such as sphere decoder [4] in slow-fading MIMO channels. In these channels, the phases and the amplitudes of the channel coefficients can be assumed approximately constant over a long period. One such period is called a *quasi-static channel interval* or simply an interval. Let the channel coefficients remain unchanged for a time period t, which is much larger than the codeword length $2n$, so that many codewords can be transmitted in this interval. For example, let $m = \frac{t}{2n} \in \mathbb{Z}$ be the number of feasible codeword transmissions within one such quasi-static

channel interval. In IF, the matrices $\mathbf{A}$ and $\mathbf{B}$ are computed once, based on $\mathbf{H}$ and estimated ρ in the first codeword transmission. In the subsequent $m-1$ codeword transmissions within the same interval, only a system of linear equations, with integer coefficients $\mathbf{A}$, will be solved to recover the original information symbols at the receiver. Referring to Table II of [22], the overall complexity of using IF in an interval is dominated by the complexity of finding the matrix $\mathbf{A}$ as the m systems of $2n$ linear equations with $2n$ unknown variables can be solved using Gauss-Jordan elimination process with computational complexity of order $m \times (2n)^3$. This is in contrast to ML decoding algorithms in slow-fading MIMO channels. In these scenarios, an algorithm such as sphere decoder should be employed for each and all m codeword transmissions inside one quasi-static channel interval. Therefore, the overall computational complexity for an interval is m times the complexity of the sphere-decoder algorithm [4]. This makes IF linear receiver computationally more efficient than ML decoding algorithms in slow-fading channels.

III. DIVERSITY ANALYSIS

We first recall the definition of diversity and a weak version of Woodbury identity [40], which are used in the rest of this Section.

Definition 2: In an $2n \times 2n$ MIMO system and at a high signal-to-noise ration ρ, if the average probability P_e, that a transmitted symbol vector is wrongly decoded, is approximated by $(c.\rho)^{-\delta}$, then δ is called the *diversity order* (or *diversity gain*) and c is called the *coding gain*. For a MIMO system with precoding, if $\delta = (2n)^2$, then, we say that the precoder achieves full-diversity order.

Lemma 1 (Woodbury Identity): Let $\mathbf{M}_1$ and $\mathbf{M}_2$ be two $d \times d$ invertible matrices, then the following identity holds:

$$\mathbf{I}_d - \mathbf{M}_1 \left(\mathbf{I}_d + \mathbf{M}_2\mathbf{M}_1\right)^{-1} \mathbf{M}_2 = \left(\mathbf{I}_d + \mathbf{M}_1\mathbf{M}_2\right)^{-1}.$$

We now proceed to analyze the diversity order of UPIF. The effective noise in (9) is not Gaussian distributed due to the quantization noise term. Note that matrix $\left(\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h\mathbf{\Sigma}\right)^{-1}$ is a positive definite matrix as it is a diagonal matrix with positive entries, hence it admits a Cholesky

decomposition

$$\left(\mathbf{I}_{2n} + \rho \cdot \boldsymbol{\Sigma}^h \boldsymbol{\Sigma}\right)^{-1} = \mathbf{L}\mathbf{L}^h. \tag{12}$$

The optimum value of $\mathbf{b}_m$ that minimizes (10) given $\mathbf{a}_m$ is [21]

$$\mathbf{b}_m = \rho \cdot \mathbf{a}_m \boldsymbol{\Sigma} \mathbf{P}^h \left(\mathbf{I}_{2n} + \rho \cdot \boldsymbol{\Sigma} \mathbf{P} \left(\boldsymbol{\Sigma} \mathbf{P}\right)^h\right)^{-1}. \tag{13}$$

By defining

$$\mathbf{S} \triangleq \mathbf{I}_{2n} + \rho \cdot \boldsymbol{\Sigma} \mathbf{P} \left(\boldsymbol{\Sigma} \mathbf{P}\right)^h,$$

we re-write (13) as

$$\mathbf{b}_m = \rho \cdot \mathbf{a}_m \boldsymbol{\Sigma} \mathbf{P}^h \mathbf{S}^{-1}. \tag{14}$$

Let

$$\mathbf{L}_p \triangleq \mathbf{P}^h \mathbf{L}, \tag{15}$$

then the average energy of the effective noise term along the m-th layer is given by

$$
\begin{aligned}
G(\mathbf{a}_m, \mathbf{b}_m) \;\; &= \; \rho \|\mathbf{b}_m \boldsymbol{\Sigma} \mathbf{P} - \mathbf{a}_m\|^2 + \|\mathbf{b}_m\|^2 \\
&= \; \rho \cdot \mathbf{a}_m (\mathbf{I}_{2n} - (\boldsymbol{\Sigma} \mathbf{P})^h \, \mathbf{S}^{-1} \boldsymbol{\Sigma} \mathbf{P}) \mathbf{a}_m^h \tag{16} \\
&= \; \rho \cdot \mathbf{a}_m \left(\mathbf{I}_{2n} + \rho \cdot (\boldsymbol{\Sigma} \mathbf{P})^h \, \boldsymbol{\Sigma} \mathbf{P}\right)^{-1} \mathbf{a}_m^h \tag{17} \\
&= \; \rho \cdot \mathbf{a}_m \mathbf{P}^h \left(\mathbf{I}_{2n} + \rho \cdot \boldsymbol{\Sigma}^h \boldsymbol{\Sigma}\right)^{-1} \mathbf{P} \mathbf{a}_m^h \\
&= \; \rho \cdot \mathbf{a}_m \mathbf{P}^h \mathbf{L}\mathbf{L}^h \mathbf{P} \mathbf{a}_m^h \tag{18} \\
&= \; \rho \cdot \mathbf{a}_m \mathbf{L}_p \mathbf{L}_p^h \mathbf{a}_m^h, \tag{19}
\end{aligned}
$$

where (16) holds because of (14) and equation (9) in [22], (17) is true because of the Woodbury identity in Lemma 1 for

$$d = 2n, \;\; \mathbf{M}_1 = \sqrt{\rho} \cdot (\boldsymbol{\Sigma} \mathbf{P})^h, \;\; \mathbf{M}_2 = \sqrt{\rho} \cdot \boldsymbol{\Sigma} \mathbf{P},$$

and (18) and (19) are obtained based on (12) and (15), respectively. We denote the probability of error for decoding the m-th layer in the integer lattice $\mathbb{Z}^{2n}$ by $P_e(m, \boldsymbol{\Sigma} \mathbf{P}, \mathbb{Z}^{2n})$. It follows

that

$$P_e\left(m, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right) = \Pr\left(|\mathbf{e}_m| \geq \frac{\sqrt{\rho}}{2}\right), \tag{20}$$

where $\mathbf{e}_m$ is the m-th row of $\mathbf{E}$ defined in (8).

The following Theorem provides an upper bound on $P_e\left(m, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right)$.

Theorem 1 (Upper Bound on Probability of Error): For all $1 \leq m \leq 2n$, the term $P_e\left(m, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right)$ as in (20), is upper bounded as

$$P_e\left(m, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right) \leq \exp\left(-c\epsilon_{2n-m+1}^2(\Lambda_{\mathbf{L}_p^{-1}})\right), \tag{21}$$

where c is a constant independent of ρ and $\epsilon_{2n-m+1}^2(\Lambda_{\mathbf{L}_p^{-1}})$ is the $(2n - m + 1)$-th successive minima of the lattice with generator matrix $\mathbf{L}_p^{-1}$ defined in (19).

Proof: The proof is given in Appendix A. Note that in the proof we assume the use of dithering which makes $\mathbf{x}_m$ a random variable distributed uniformly within the shaping region.

∎

We are interested in $P_e\left(2n, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right)$ which is an upper bound for $P_e\left(m, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right)$, for all $1 \leq m \leq 2n$. We define the error probability for unitary precoded integer forcing over $\mathbb{Z}^{2n}$ as

$$P_e\left(\mathbf{\Sigma P}, \mathbb{Z}^{2n}\right) \triangleq P_e\left(2n, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right). \tag{22}$$

Based on (21), we have

$$P_e(\mathbf{\Sigma P}, \mathbb{Z}) \leq \exp\left(-c\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}})\right). \tag{23}$$

Let the average probability

$$P_e = \mathbb{E}_{\mathbf{H}}\left(P_e(\mathbf{\Sigma P}, \mathbb{Z})\right),$$

where the expectation is taken over all channel matrices $\mathbf{H}$.

Theorem 2: Let the precoding matrix $\mathbf{P}$ be such that $[\mathbf{Pv}]_1 \neq 0$, where $\mathbf{v} \in \mathbb{Z}^{2n}$ is the vector satisfying $\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}}) = \|\mathbf{L}_p^{-1}\mathbf{v}\|^2$, then the unitary precoded integer-forcing explained through (5) to (10) achieves full-diversity $(2n)^2$.

Proof: The proof is given in Appendix B.

∎

Based on the proof of the above theorem, to achieve full-diversity, all we need is to select a unitary precoder matrix $\mathbf{P}$ such that if $\mathbf{v} \in \mathbb{Z}^{2n}$ be the vector, which arises to the minimum distance of the lattice $\Lambda_{\mathbf{L}_p^{-1}}$, then it satisfies $[\mathbf{P}\mathbf{v}]_1 \neq 0$. However, enforcing this condition seems to be unpractical because it involves computation of $d_{\min}\left(\Lambda_{\mathbf{L}_p^{-1}}\right)$ which is an NP-hard problem. For that reason, we choose to work with a stronger condition that is $d_{p,\min}(\Lambda_{\mathbf{P}}) \neq 0$. It is clear that if $d_{p,\min}(\Lambda_{\mathbf{P}}) \neq 0$ then $[\mathbf{P}\mathbf{v}]_1 \neq 0$ for every vector $\mathbf{0} \neq \mathbf{v} \in \mathbb{Z}^{2n}$.

Theorem 3: Let the precoding matrix $\mathbf{P}$ be such that $d_{p,\min}(\Lambda_{\mathbf{P}}) \neq 0$, then the achievable diversity of the unitary precoded integer-forcing explained through (5) to (10) is $(2n)^2$.

According to the above theorem, if lattice codes along with a non-zero minimum product distance precoder are used at the transmitter of a MIMO channel and integer-forcing is employed at the receiver, full (transmit and receive) diversity can be achieved. We next provide two different design criteria for obtaining unitary precoders suitable for integer-forcing linear receivers.

IV. Optimal Design of Full-Diversity Unitary Precoders

Based on the above two theorems, we next design classes of full-diversity unitary precoders for integer-forcing linear receivers. We first propose a method of designing optimal precoders based on Theorem 2 and then we provide our optimal precoders based on Theorem 3.

A. Design of Type I UPIF

Based on (23) and Theorem 2, the optimal Type I UPIF is as follows:

$$\mathbf{P}_{1,\text{opt}} = \arg \max_{\mathbf{P} \in \mathcal{O}_{2n}} \min_{\substack{\mathbf{v} \in \mathbb{Z}^{2n} \setminus \{\mathbf{0}\} \\ [\mathbf{P}\mathbf{v}]_1 \neq 0}} \|\mathbf{L}^{-1}\mathbf{P}\mathbf{v}\|^2, \tag{24}$$

where $\mathcal{O}_{2n}$ is the orthogonal group of all $2n \times 2n$ orthogonal matrices with matrix multiplication operation. In other words, we should design a precoder matrix $\mathbf{P}$ such that the minimum distance of the lattice $\Lambda_{\mathbf{L}_p^{-1}}$ is maximized, while $[\mathbf{P}\mathbf{v}]_1 \neq 0$ for $\mathbf{v} \in \mathbb{Z}^{2n}$ and $\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}}) = \|\mathbf{L}_p^{-1}\mathbf{v}\|^2$.

Remark 1: If we relax the constraint on $\mathbf{P} \in \mathcal{O}_{2n}$ to be $\text{Tr}\left(\mathbf{P}\mathbf{P}^T\right) \leq \rho_0$, where $\text{Tr}(.)$ denotes the trace of a matrix, and remove the constraint $[\mathbf{P}\mathbf{v}]_1 \neq 0$, we get a more general optimization problem, which is the main subject of study in [42]. The optimal solution for this problem is derived to be the hexagonal lattice $\mathcal{A}_2$ [43] for 2×2 real MIMO channels and Schlafli lattice
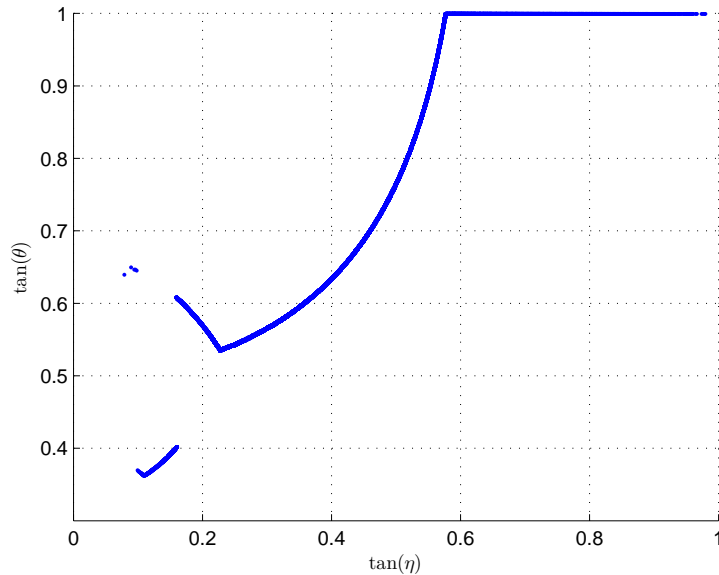
Fig. 2. The variation of $\tan\theta$ based on the variation of $\tan\eta$ in a 2×2 complex MIMO Channel using real Type I UPIF.

$\mathcal{D}_4$ [43] for 2×2 complex MIMO channels. However, the generator matrix of these lattices are not unitary. In this paper, we focus on solving the optimization problem in (24), which is slightly different from that of [42].

The problem stated in (24) seems to be a hard combinatorial problem, which we could not solve analytically. However, a systematic approach of finding $\mathbf{P}_{1,\text{opt}}$ is to perform an exhaustive search in $\mathcal{O}_{2n}$. A convenient parametrization of $\mathcal{O}_{2n}$ using $(2n)^2$ parameters is presented in [44]. Such a parametrization uses the fact that an orthogonal transformation $\mathbf{P}$ can be composed from elementary orthogonal transformations in two-dimensional sub-spaces. The induced matrices using the algorithm given in [44] will be uniformly distributed with respect to Haar measure.

For the 2×2 MIMO system, since we imposed the orthogonality constraint on $\mathbf{P}$, we parameterize it using a single angle θ as follows:

$$\mathbf{P}(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{25}$$

It can be easily seen by the symmetry of integers that it suffices to consider $\theta \in [0, \pi/4]$ for

the maximization in (24). Hence, we numerically search for

$$\mathbf{P}_{1,\text{opt}}^{(\mathbb{R})} = \arg \max_{\mathbf{P}(\theta) \in \mathcal{O}_2} \min_{[\mathbf{P}(\theta)\mathbf{v}]_1 \neq 0} \|\mathbf{L}^{-1}\mathbf{P}(\theta)\mathbf{v}\|^2, \tag{26}$$

where $\mathcal{O}_2$ is the group of all 2×2 orthogonal matrices with matrix multiplication operation. This exhaustive search can be done by discretizing $\theta \in [0, \pi/4]$ using fine steps of 0.001 radians and consequently finding the minimum distance of $\Lambda_{\mathbf{L}_p^{-1}}$ using Gauss reduction algorithm [45]. The 2×2 real precoding matrix, which results in the highest minimum distance of the corresponding lattice $\Lambda_{\mathbf{L}_p^{-1}}$ is considered as a 2×2 real approximately optimal percoder matrix $\mathbf{P}_{1,\text{opt}}^{(\mathbb{R})}$ to be employed for both real and complex 2×2 UPIF. Using (18), the matrix $\mathbf{L}^{-1}$ is given by

$$\mathbf{L}^{-1} = \begin{pmatrix} \sqrt{1 + \rho\sigma_1^2} & 0 \\ 0 & \sqrt{1 + \rho\sigma_2^2} \end{pmatrix} = \xi \begin{pmatrix} \cos\eta & 0 \\ 0 & \sin\eta \end{pmatrix},$$

where $\xi = \sqrt{\left(\sqrt{1 + \rho\sigma_1^2}\right)^2 + \left(\sqrt{1 + \rho\sigma_2^2}\right)^2} = \sqrt{2 + \rho(\sigma_1^2 + \sigma_1^2)}$ and

$$\eta = \tan^{-1}\left(\frac{\sqrt{1 + \rho\sigma_2^2}}{\sqrt{1 + \rho\sigma_1^2}}\right).$$

It follows that $\mathbf{L}_p^{-1}$ is equal to

$$\mathbf{L}^{-1}\mathbf{P}(\theta) = \xi \begin{pmatrix} \cos\eta & 0 \\ 0 & \sin\eta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

In Fig. 2, we plot the variation of $\tan\theta$ based on $\tan\eta$ for 10^5 samples of 2×2 complex MIMO systems employing real 2×2 precoders $\mathbf{P}(\theta)$. In particular,

$$\tan\eta = \frac{\sqrt{1 + \rho\sigma_2^2}}{\sqrt{1 + \rho\sigma_1^2}} = \frac{1}{\beta_{\mathbf{L}_p^{-1}}},$$

where $\beta_{\mathbf{L}_p^{-1}}$ is the condition number of $\mathbf{L}_p^{-1}$. For large values of ρ, $\tan\eta \approx \frac{\sigma_2}{\sigma_1} = \frac{1}{\beta_{\mathbf{H}}}$, where $\beta_{\mathbf{H}}$ is the condition number of the channel $\mathbf{H}$. The generator matrix of the lattice $\Lambda_{\mathbf{L}_p^{-1}}$ can be considered now as the equivalent channel matrix. It can be seen that if the condition number is less than $\sqrt{3}$, then the best precoder obtains by putting $\theta = \pi/4$. In Fig. 3, we plot the variation
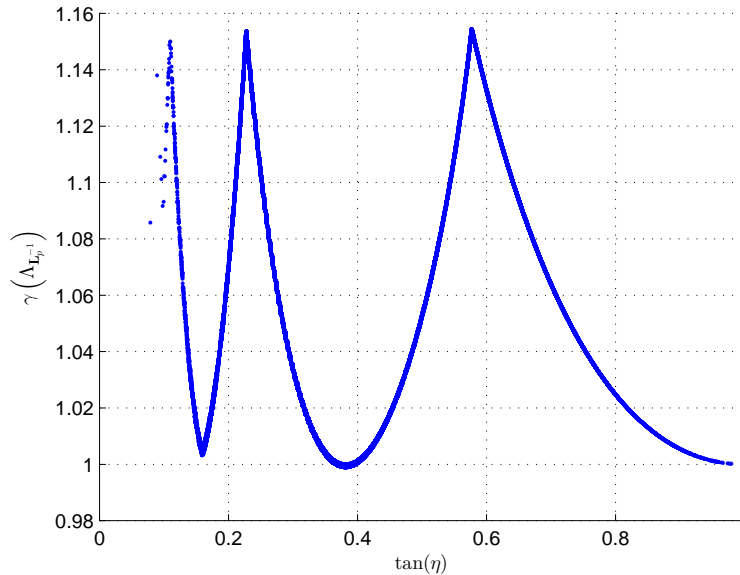
Fig. 3. The variation of $\gamma(\Lambda_{\mathbf{L}_p^{-1}})$ based on the variation of $\tan\eta$ in a 2×2 complex MIMO Channel using real Type I UPIF.

of the coding gain of the lattice $\Lambda_{\mathbf{L}_p^{-1}}$ versus $\tan\eta$. The coding gain formula is:

$$\gamma(\Lambda_{\mathbf{L}_p^{-1}}) = \frac{\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}})}{\det\left(\mathbf{L}_p^{-1}\right)^{\frac{2}{2n}}}. \tag{27}$$

The coding gain measures the increase in density of $\Lambda_{\mathbf{L}_p^{-1}}$ over the integer lattice $\mathbb{Z}^{2n}$ with $\gamma\left(\mathbb{Z}^{2n}\right) = 1$. Based on Fig. 3, we can safely say that the equivalent precoded channel matrix $\mathbf{L}_p^{-1}$ is at least perform as good as an integer 2×2 channel matrix. In other words, the precoder matrix will change the bad channels to equivalent channels as good as an integer one. Fig. 4 shows the histogram of $\gamma(\Lambda_{\mathbf{L}_p^{-1}})$ for the same 10^5 channel samples. There is a probability mass at around 1, which states that unitary matrices has cleared all the bad channels with $\gamma\left(\Lambda_{\mathbf{H}}\right) < 1$ to convert them to an equivalent channel matrix $\mathbf{L}_p^{-1}$ with $\gamma(\Lambda_{\mathbf{L}_p^{-1}}) \geq 1$. Note also that the maximum attainable coding gain (Hermite's constant [43]) for 2-dimensional lattices is $\frac{2}{\sqrt{3}}$. As it is clear from Figs. 3 and 4, this quantity has been achieved by few lattices in our simulations too.
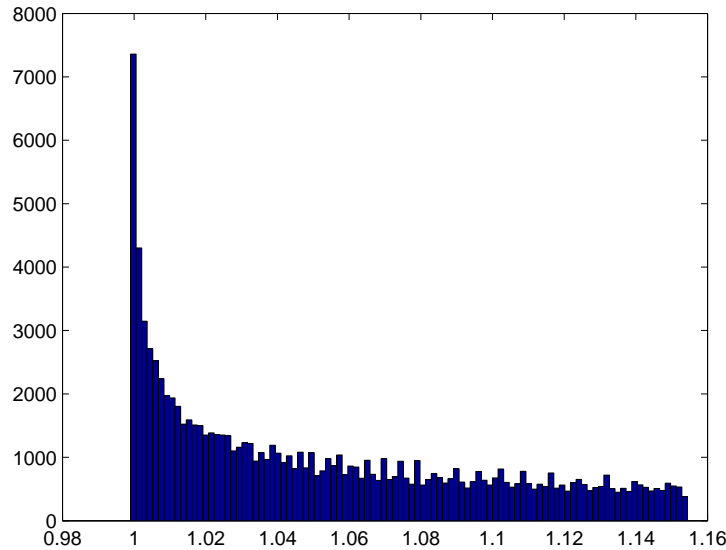
Fig. 4. The histogram of $\gamma(\Lambda_{\mathbf{L}_p^{-1}})$ in a 2×2 complex MIMO Channel using real Type I UPIF.

B. Design of Type II UPIF

In previous Subsection, we introduced $\mathbf{P}_{1,\text{opt}}$ as a suitable precoder matrix for UPIF. Equations (25) and (26) show that Type I UPIF precoder matrices $\mathbf{P}_{1,\text{opt}}$ and $\mathbf{P}_{1,\text{opt}}^{(\mathbb{R})}$ are adapted based on $\mathbf{L}^{-1}$ and consequently $\mathbf{H}$ and ρ. In this Subsection, however we introduce Type II UPIF, where the optimal precoder matrix $\mathbf{P}_{2,\text{opt}}$ is fixed and it does not change by varying the channel matrix $\mathbf{H}$.

By further expanding $\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}})$ in (23), we get:

$$
\begin{aligned}
\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}}) &= \|\mathbf{L}_p^{-1}\mathbf{v}\|^2 \\
&= \mathbf{v}^h \mathbf{L}_p^{-h} \mathbf{L}_p^{-1} \mathbf{v} = \mathbf{v}^h \left(\mathbf{L}_p \mathbf{L}_p^h\right)^{-1} \mathbf{v} \\
&= \mathbf{v}^h \mathbf{P}^h \left(\left(\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h \mathbf{\Sigma}\right)^{-1}\right)^{-1} \mathbf{P}\mathbf{v} \qquad (28) \\
&= \mathbf{v}^h \mathbf{P}^h \left(\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h \mathbf{\Sigma}\right) \mathbf{P}\mathbf{v}, \qquad (29)
\end{aligned}
$$

where (28) follows from (19). Since $\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h \mathbf{\Sigma}$ is a positive definite matrix, we can perform a Cholesky decomposition to obtain $\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h \mathbf{\Sigma} = \mathbf{D}\mathbf{D}^h$. With this, (29) can be written as

$\|\mathbf{DPv}\|^2$. Hence, we get

$$
\begin{aligned}
\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}}) &= \|\mathbf{DPv}\|^2 \\
&= \sum_{m=1}^{2n} [\mathbf{D}]_{m,m}^2 [\mathbf{Pv}]_m^2 \\
&\geq 2n \left(\prod_{m=1}^{2n} [\mathbf{D}]_{m,m}^2 \prod_{m=1}^{2n} [\mathbf{Pv}]_m^2 \right)^{\frac{1}{2n}} \quad (30) \\
&= 2n \left(\prod_{m=1}^{2n} \left(1 + \rho \sigma_m^2(\mathbf{H}) \right) \prod_{m=1}^{2n} [\mathbf{Pv}]_m^2 \right)^{\frac{1}{2n}} \\
&\geq 2n \prod_{m=1}^{2n} \left(\rho \sigma_m^2(\mathbf{H}) \right)^{\frac{1}{2n}} \left(d_{p,\min}^2 (\Lambda_{\mathbf{P}}) \right)^{\frac{1}{2n}} \quad (31) \\
&= 2n\rho \left(\det(\mathbf{H}) \right)^{\frac{1}{n}} d_{p,\min}^{\frac{1}{n}} (\Lambda_{\mathbf{P}}), \quad (32)
\end{aligned}
$$

where (30) follows from Cauchy–Schwarz inequality, and (31) is true because $1 + \rho \sigma_m^2(\mathbf{H}) >$ $\rho \sigma_m^2(\mathbf{H})$, for all $1 \leq m \leq 2n$ and $d_{p,\min}^2 (\Lambda_{\mathbf{P}}) \leq \prod_{m=1}^{2n} [\mathbf{Pu}]_m^2$ for all $\mathbf{u} \in \mathbb{Z}^{2n}$.

Based on (32) and since almost surely $\det(\mathbf{H}) \neq 0$, the optimal Type II UPIF can be obtained by solving the following optimization problem:

$$
\mathbf{P}_{2,\text{opt}} = \arg \max_{\mathbf{P} \in \mathcal{O}_{2n}} d_{p\,\min}^{\frac{1}{n}} (\Lambda_{\mathbf{P}}), \quad (33)
$$

The solution for the maximization in (33) is provided in [32]–[36] using algebraic number theoretic lattices. A list of full-diversity algebraic rotations is available in [37]. All the lattices provided there satisfy non-vanishing minimum product distance criterion too. Hence, thanks to [37], the optimal Type II UPIF along with their corresponding $d_{p,\min}$ are available. We now proceed to verify this result by conducting simulations.

V. SIMULATION RESULTS

The UPIF scheme and MIMO precoding X-codes and Y-codes [13] share similar properties, which make them suitable for comparison: (*i*) both schemes use SVD decomposition technique to transform the channel matrix into a diagonal one, (*ii*) the precoder matrices in both systems must be unitary/orthogonal matrices, (*iii*) both the detectors at the receiver side, *i.e.* lattice reduction
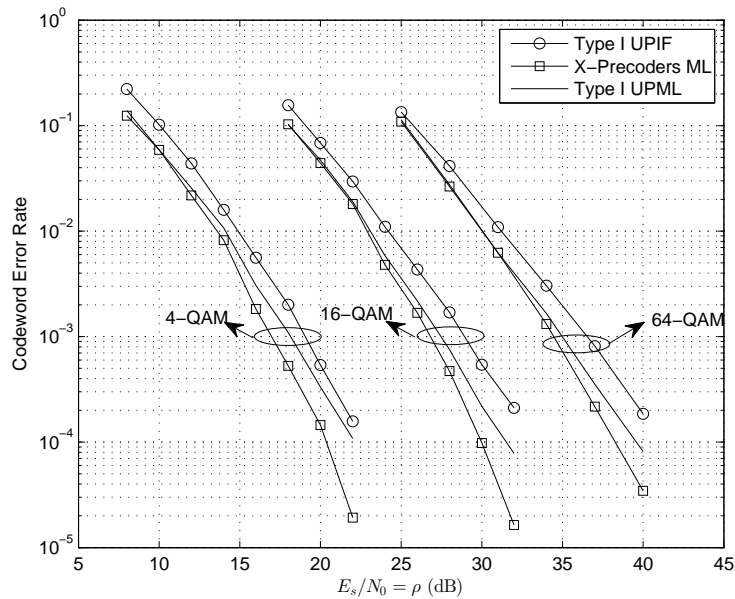
Fig. 5. Type I UPIF in comparison with, X-Precoders decoded with sphere decoding algorithm, and Type II UPML in a 2×2 complex MIMO Channel.

based IF linear receiver [22] and a combination of two 2-dimensional ML decoders, provide full receive diversity in 2×2 MIMO.

For comparison we show the performance of UPIF precoding when the IF receiver is replaced by an ML decoder. We will denote this with the acronym UPML. In the following subsections, we simulate and compare our proposed Type I/II UPIF decoded using lattice reduction based IF receiver, with Type I/II UPML, and MIMO X-Precoders/X-Codes.

A. Type I UPIF versus X-Precoders

We show simulations for a 2×2 MIMO UPIF channel over $4/16/64$-QAM constellations. The matrices $\mathbf{A}$ and $\mathbf{B}$ for IF linear receiver were found using Algorithm2 presented in [22]. We have chosen our Type I UPIF by running an exhaustive search based on the design criterion given in (26). For reference purposes, we also presented MIMO X-precoder [13] scheme decoded under sphere decoding algorithm [4]. The step size for our brute force search for both the cases (Type I UPIF and X-precoders) is 0.001 radians. Fig. 5 shows the codeword error rate [22] curves versus signal-to-noise ratio ρ of the different schemes. Since the performance of UPIF

decreases by growing ρ parallel to ML decoded curve, full-diversity of UPIF is guaranteed. In particular, at codeword error rate 10^{-3} over $4/16/64$-QAM constellations, the gaps of 2.2dB, 2.15dB, and 2.1dB between Type I UPIF and X-precoder are observed. This shows a decrease in coding loss by increasing the constellation size. The 2×2 X-precoders outperform the other schemes because of two main reasons: (*i*) these precoders are designed for the particular lattice constellation (for example, 4-QAM, 16-QAM, or 64-QAM) while Type I UPIF are designed for the (infinite) lattice $\mathbb{Z}^{2n}$: (*ii*) since we are only able to simulate Type I UPIF over a 2×2 MIMO channel, the diversity gain of Type I UPIF over X-precoders cannot be captured. This will be more visible in the next subsection, where Type II UPIF outperform X-codes in MIMO channels with larger number of antennas employed at larger constellation sizes.

B. Type II UPIF versus X-Codes

We have conducted simulations for 2×2 and 4×4 MIMO UPIF channel over $4/16/64$-QAM constellations. For 2×2 MIMO channels, the matrices $\mathbf{A}$ and $\mathbf{B}$ for IF linear receiver were found using Algorithm2 presented in [22] while for larger MIMOs we have employed Algorithm1. We have chosen our Type II UPIFs based on the design criterion given in (33). The 2×2 and 4×4 full-diversity algebraic rotation [37] matrices which has been used for simulations have the following minimum product distances ($d_{p,\min}^{\frac{1}{n}}$) 0.668740, 0.438993, and 0.289520. For reference purposes, we also presented MIMO X-codes [13] decoded under sphere decoding algorithm [4]. For 4-QAM constellation the optimal angle for a X-Code is 26.6 degrees while for 16-QAM it is equal to 15 degrees [13]. We also found the optimal angle for 64-QAM to be equal to 8 degrees. In Figs. 6 and 7, we plot the codeword error performance of Type II UPIF versus MIMO X-codes and Type II unitary precoders decoded using a maximum likelihood decoder (Type II UPML) such as sphere decoding algorithm versus signal-to-noise ratio $E_s/N_0 = \rho$. As it is clear from Figs. 6 and 7, the Type II UPIF outperform MIMO X-codes for larger constellation sizes in higher MIMO dimensions. The main reason is that the Type II UPIF precoder matrices are designed to achieve full-diversity and the diversity gain shows itself when the number of antennas increases. In particular, the diversity order [13] of the optimal MIMO X-codes (in terms of diversity not the error performance) is known to be $d_{n_t,n_r,n_s} = \left(n_r - \frac{n_s}{2} + 1\right)\left(n_t - \frac{n_s}{2} + 1\right),$
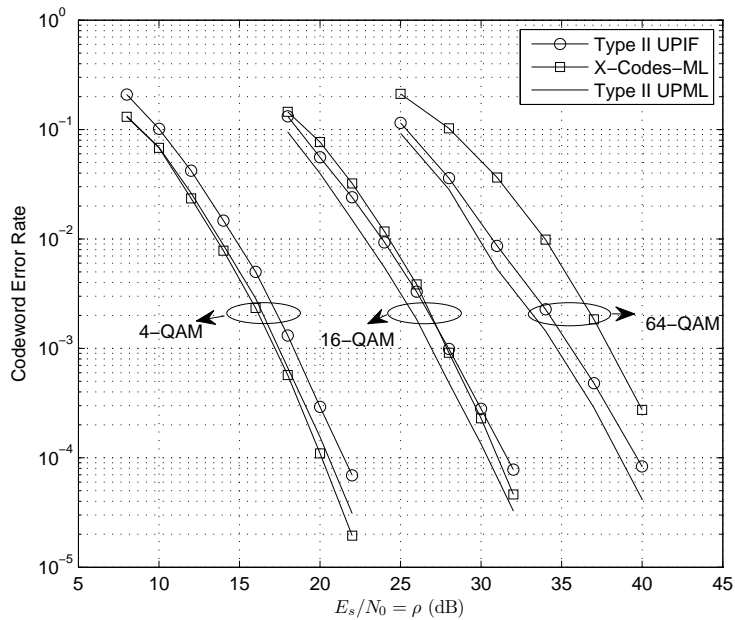
Fig. 6. Type II UPIF in comparison with X-Codes and Type II UPML in a 2×2 complex MIMO Channel.

where n_r and n_t denote the number of receive and transmit antennas, respectively, and n_s denotes the number of symbols. In our setting, for 2×2 and 4×4 MIMO X-codes, we have $d_{2,2,2} = 4$ and $d_{4,4,4} = 9$ in comparison with full-diversity orders 4 and 16 for Type II UPIF, respectively.

C. Type I UPIF versus Type II UPIF

In this subsection, we compare Type I and Type II UPIF together. Fig. 8 shows the codeword error performance of Type I and Type II UPIF in comparison to their corresponding UPML. It is obvious that the Type II UPIF and UPML outperform Type I UPIF and UPML precoders by no more than 0.5dB and 1dB, respectively.

The sub-optimality of unitary precoder matrices obtained using quantized exhaustive search algorithm is the main reason for having Type II UPIF performing better than Type I UPIF. Furthermore, we note that the design criteria given in (26) and (33) are different and we cannot say which one can perform better than the other by just looking at these equations. Hence, considering the simplicity of Type II UPIF in comparison to Type I and their better error performance, we suggest using of Type II UPIF for practical purposes.
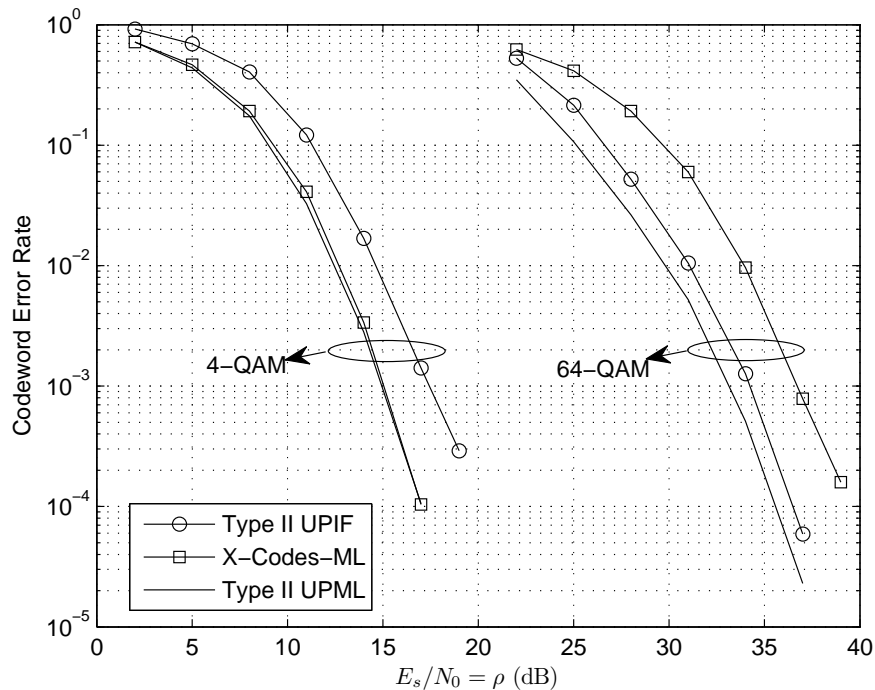
Fig. 7. Type II UPIF in comparison with X-Codes and Type II UPML in a 4×4 complex MIMO Channel.

VI. SUMMARY AND DIRECTIONS FOR FUTURE WORK

A unitary precoding scheme (UPIF) has been introduced for transmission over a flat-fading MIMO channel with both CSIT and CSIR, and an IF linear receiver. The diversity gains of the proposed approach have been analyzed both theoretically and numerically. Two different design criteria for two types of unitary precoders, Type I UPIF and Type II UPIF, were given. For both cases, an SVD is performed first to transform the channel to a diagonal one. Type I UPIF includes precoder matrices adapting based on the channel matrix singular values, while Type II UPIF consists of precoding matrices, which remain fixed by changing the channel matrix.

Designing full-diversity unitary precoders or more generally space-time block codes for IF receivers without CSIT would be of interest. Another direction is to let the transmitter have access to limited feedback over a delay-free link from the IF receiver [46], [47]. In this case, designing a suitable codebook of unitary precoding matrices attaining higher rates and obtaining higher coding gains in error performance seems to be a promising research topic.
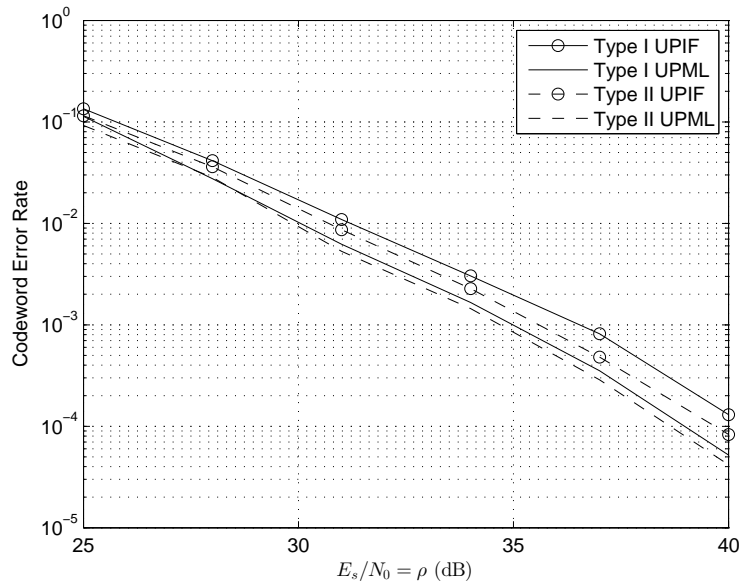
Fig. 8. Type I versus Type II UPIF and UPML schemes in a 2×2 complex MIMO Channel.

APPENDIX A

PROOF OF THEOREM 1

For the proof of this we follow the lines of the proof of Theorem 7 in [41]. Since the minimum Euclidean distance of $\mathbb{Z}$ is unity, an error is declared if $\mathbf{e}_m \geq \frac{\sqrt{\rho}}{2}$. Therefore, we have

$$
\begin{aligned}
P_e\left(m, \mathbf{\Sigma P}, \mathbb{Z}^{2n}\right) &= \operatorname{Pr}\left(|\mathbf{e}_m| \geq \frac{\sqrt{\rho}}{2}\right) \\
&= 2\operatorname{Pr}\left(\mathbf{e}_m \geq \frac{\sqrt{\rho}}{2}\right) \quad (34) \\
&\leq 2\min_{t>0} \frac{\mathbb{E}(\exp(t\mathbf{e}_m))}{\exp\left(\frac{\sqrt{\rho}t}{2}\right)} \quad (35) \\
&= 2\min_{t>0} \frac{\mathbb{E}\left(\exp\left(t\sqrt{\rho} \cdot \langle \mathbf{b}_m\mathbf{\Sigma P} - \mathbf{a}_m, \mathbf{x}_m \rangle + t \cdot \langle \mathbf{b}_m, \mathbf{z}'_m \rangle\right)\right)}{\exp\left(\frac{\sqrt{\rho}t}{2}\right)} \\
&= 2\min_{t>0} \frac{\mathbb{E}(\exp(t\sqrt{\rho} \cdot \langle \mathbf{b}_m\mathbf{\Sigma P} - \mathbf{a}_m, \mathbf{x}_m \rangle)\mathbb{E}\left(\exp\left(t \cdot \langle \mathbf{b}_m, \mathbf{z}'_m \rangle\right)\right)}{\exp\left(\frac{\sqrt{\rho}t}{2}\right)} \quad (36)
\end{aligned}
$$

where (34) follows from the symmetry of effective noise around zero, (35) follows from Chernoff's bound and (36) is true due to the independence of $\mathbf{x}_m$ and $\mathbf{z}'_m$. We continue to use

moment generating functions of uniform and Gaussian distributions to further upper bound the two exponential terms in (36). Since the components of $\mathbf{z}'_m$ are distributed based on a Gaussian distribution with zero mean and unit variance, we have

$$\mathbb{E}\left(\exp\left(t \cdot \langle \mathbf{b}_m, \mathbf{z}'_m \rangle\right)\right) \leq \exp\left(\frac{t^2 \|\mathbf{b}_m\|^2}{2}\right), \tag{37}$$

for the second term in (36). We now investigate its first term. Let $\mathbf{q}_m \triangleq t\sqrt{\rho} \cdot (\mathbf{b}_m \boldsymbol{\Sigma} \mathbf{P} - \mathbf{a}_m)$. The entries of $\mathbf{x}_m \in \mathbb{Z}^{2n}$ are uniformly distributed over the hypercube $\left[-t\sqrt{\rho}\|\mathbf{q}_m\|, t\sqrt{\rho}\|\mathbf{q}_m\|\right]^{2n}$, therefore we have

$$
\begin{aligned}
\mathbb{E}\left(\exp(t\sqrt{\rho} \cdot \langle \mathbf{q}_m, \mathbf{x}_m \rangle)\right) &= \mathbb{E}\left(\exp\left(t\sqrt{\rho} \cdot \sum_{j=1}^{2n} [\mathbf{q}_m]_j [\mathbf{x}_m]_j\right)\right) \\
&= \prod_{j=1}^{2n} \mathbb{E}\left(\exp\left(t\sqrt{\rho} \cdot [\mathbf{q}_m]_j [\mathbf{x}_m]_j\right)\right) \tag{38} \\
&\leq \prod_{j=1}^{2n} \frac{\sinh\left(t\sqrt{\rho}|[\mathbf{q}_m]_j [\mathbf{x}_m]_j|\right)}{t\sqrt{\rho}|[\mathbf{q}_m]_j [\mathbf{x}_m]_j|} \tag{39} \\
&\leq \prod_{j=1}^{2n} \exp\left(\frac{t^2 \rho |[\mathbf{q}_m]_j|^2}{6}\right) \tag{40} \\
&\leq \exp\left(\frac{t^2 \rho \|\mathbf{q}_m\|^2}{2}\right), \tag{41}
\end{aligned}
$$

where (38) follows from the independence of components of $\mathbf{x}_m$, (39) follows from the moment generating function of a uniform distribution, and (40) follows from the inequality

$$\frac{\sinh(y)}{y} \leq \exp\left(\frac{y^2}{6}\right),$$

for every $y \in \mathbb{R}$. Combining (41) and (37) in (36), we get

$$
\begin{aligned}
P_e(m, \boldsymbol{\Sigma} \mathbf{P}, \mathbb{Z}) &\leq 2 \min_{t>0} \frac{\exp\left(\frac{t^2 \rho \|\mathbf{b}_m \boldsymbol{\Sigma} \mathbf{P} - \mathbf{a}_m\|^2}{2}\right) \exp\left(\frac{t^2 \|\mathbf{b}_m\|^2}{2}\right)}{\exp\left(\frac{\sqrt{\rho} t}{2}\right)} \\
&= 2 \min_{t>0} \frac{\exp\left(\frac{t^2 G(\mathbf{a}_m, \mathbf{b}_m)}{2}\right)}{\exp\left(\frac{\sqrt{\rho} t}{2}\right)} \tag{42} \\
&= 2 \exp\left(\frac{-\rho}{4 G(\mathbf{a}_m, \mathbf{b}_m)}\right), \tag{43}
\end{aligned}
$$

where (43) obtained by optimizing t, that is $t = \frac{\sqrt{\rho}}{2G(\mathbf{a}_m, \mathbf{b}_m)}$.

Now since $G(\mathbf{a}_m, \mathbf{b}_m)$ needs to be minimized, $\mathbf{a}_m$ and $\mathbf{b}_m$ should be chosen appropriately as in [21] and [22], thus we get

$$\frac{G(\mathbf{a}_m, \mathbf{b}_m)}{\rho} = \epsilon_m^2(\Lambda_{\mathbf{L}_p^h}),$$

where $\epsilon_m^2(\Lambda_{\mathbf{L}_p^h})$ denotes the m-th successive minimum of the lattice

$$\Lambda_{\mathbf{L}_p^h} = \left\{ \mathbf{L}^h \mathbf{P} \mathbf{d} \mid \forall \mathbf{d} \in \mathbb{Z}^{2n} \right\}.$$

Here $\mathbf{L}_p^{-1}$ is a generator of the dual lattice $\Lambda_{\mathbf{L}_p^h}^*$. Thus we have the relation (see Lemma 4 in [21])

$$\epsilon_m^2(\Lambda_{\mathbf{L}_p^h}) \leq \frac{(2n)^3 + (3n)^2}{\epsilon_{2n-m+1}^2(\Lambda_{\mathbf{L}_p^h}^*)} = \frac{(2n)^3 + (3n)^2}{\epsilon_{2n-m+1}^2(\Lambda_{\mathbf{L}_p^{-1}})}, \tag{44}$$

where $\epsilon_{2n-m+1}^2(\Lambda_{\mathbf{L}_p^h}^*)$ is the $(2n-m+1)$-th successive minima of the lattice $\Lambda_{\mathbf{L}_p^h}^*$. Therefore, we have

$$\frac{\rho}{G(\mathbf{a}_m, \mathbf{b}_m)} \geq \frac{\epsilon_{2n-m+1}^2(\Lambda_{\mathbf{L}_p^{-1}})}{c_0}, \tag{45}$$

where $c_0 = (2n)^3 + (3n)^2$. Using inequality of (45) in (43), the probability of error for decoding the m-th layer is upper bounded as

$$P_e\left(m, \mathbf{\Sigma}\mathbf{P}, \mathbb{Z}^{2n}\right) \leq \exp\left(-c\epsilon_{2n-m+1}^2(\Lambda_{\mathbf{L}_p^{-1}})\right), \tag{46}$$

where $c = \frac{1}{4c_0}$. This completes the proof.

APPENDIX B

PROOF OF THEOREM 2

Toward the proof of this theorem, we make use of the following two lemmas. The first lemma gives the expression for the first-order expansion of the marginal probability density function (PDF) of eigenvalues of $\mathbf{H}$.

Lemma 2: Let the entries of the matrix $\mathbf{H}$ be i.i.d. complex Gaussian random variables with zero mean and unit variance. The first-order expansion of the marginal PDF of the k-th largest

eigenvalue λ_k of the complex central Wishart matrix $\mathbf{H}\mathbf{H}^h$ is given by

$$F_{\lambda_k}(\lambda_k) = q_k \lambda_k^{d_k} + o\left(\lambda_k^{d_k}\right),$$

as $\lambda_k \to 0$, with $d_k = (2n - k + 1)^2 - 1$ and q_k being positive constants.

Lemma 3: For a scalar channel modeled by $y = \sqrt{\text{SNR}}\beta x + n$, where $n \sim \mathcal{N}(0, 1)$, and $\mathbb{E}[|x|^2] = 1$ and $\alpha = \beta^2$ is a nonnegative random variable whose probability density function (PDF) $F_\alpha(\alpha)$ is such that as

$$F_\alpha(\alpha) = q\alpha^t + o(\alpha^t), \tag{47}$$

the average word error probability (WEP) P_e, which is given by

$$P_e = \mathbb{E}\left(P_e(\alpha)\right) = \int_0^\infty \exp\left(-k\alpha\text{SNR}\right) F_\alpha(\alpha) d\alpha,$$

is such that as $\text{SNR} \to \infty$

$$P_e = \frac{2^t q \Gamma(t + \frac{3}{2})}{\sqrt{\pi}(t + 1)}(k\text{SNR})^{-(t+1)} + o\left(\text{SNR}^{-(t+1)}\right).$$

Now we proceed to give the proof of Theorem 2. Let $\mathbf{v}$ be the Gaussian vector for which

$$
\begin{aligned}
\epsilon_1^2(\Lambda_{\mathbf{L}_p^{-1}}) &= \|\mathbf{L}_p^{-1}\mathbf{v}\|^2 \\
&= \mathbf{v}^h \mathbf{L}_p^{-h}\mathbf{L}_p^{-1}\mathbf{v} = \mathbf{v}^h \left(\mathbf{L}_p\mathbf{L}_p^h\right)^{-1}\mathbf{v} \\
&= \mathbf{v}^h\mathbf{P}^h\left(\left(\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h\mathbf{\Sigma}\right)^{-1}\right)^{-1}\mathbf{P}\mathbf{v} \qquad (48) \\
&= \mathbf{v}^h\mathbf{P}^h\left(\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h\mathbf{\Sigma}\right)\mathbf{P}\mathbf{v} \\
&\triangleq \|\mathbf{D}\mathbf{P}\mathbf{v}\|^2 \geq [\mathbf{D}]_{1,1}^2 [\mathbf{P}\mathbf{v}]_1^2 \qquad (49) \\
&= (1 + \rho\sigma_1^2(\mathbf{H})) [\mathbf{P}\mathbf{v}]_1^2 \geq \rho\sigma_1^2(\mathbf{H}) [\mathbf{P}\mathbf{v}]_1^2 \qquad (50) \\
&\geq \sqrt{\rho\lambda_1(\mathbf{H})} [\mathbf{P}\mathbf{v}]_1^2 \qquad (51)
\end{aligned}
$$

where (48) follows from (19), (49) follows from Cholesky decomposition of $\mathbf{I}_{2n} + \rho \cdot \mathbf{\Sigma}^h\mathbf{\Sigma}$ and the fact that $\mathbf{D}$ is a diagonal matrix, (50) is true because $1 + \rho\sigma_1^2(\mathbf{H}) > \rho\sigma_1^2(\mathbf{H})$, and (51) follows from the fact that $\sigma_1^2(\mathbf{H}) = \lambda_1(\mathbf{H})$.

Since the quantity $[\mathbf{P}\mathbf{v}]_1$ is not zero. Using Lemmas 2, 3 and (51), we have

$$P_e \triangleq \mathbb{E}_{\mathbf{H}}\left(P_e(\boldsymbol{\Sigma}\mathbf{P}, \mathbb{Z}^{2n})\right) \leq t_0 \rho^{-(2n)^2} + o\left(\rho^{-(2n)^2}\right),$$

where t_0 is a function of n, q_1 defined in (47), and $[\mathbf{P}\mathbf{v}]_1 \neq 0$.

REFERENCES

[1] V. Tarokh, N. Seshadri, and A.R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. on Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

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

[3] F. Oggier, G. Rekaya, J.C. Belfiore, and E. Viterbo, "Perfect space-time block-codes," *IEEE Trans. on Inform. Theory,* vol. 52, no. 9, pp. 3885–3902, Sep. 2006.

[4] E. Viterbo and J.J. Boutros, "A Universal Lattice Decoder for Fading Channels," *IEEE Trans. on Inform. Theory,* vol. 45, no. 5, pp. 1639–1642, July. 1999.

[5] K. Kumar, G. Caire, and A. Moustakas, "Asymptotic performance of linear receivers in MIMO fading channels," *IEEE Trans. on Inform. Theory*, vol. 55, no. 10, pp. 4398–4418, Oct. 2009.

[6] J.K. Zhang, J. Liu, and K.M. Wong, "Linear Toeplitz space time block codes," in Proc. of *IEEE ISIT-05*, Adelaide, Australia, Sep. 2005, pp. 1942–1946.

[7] J. Liu, J.K. Zhang, and K.M. Wong, "On the design of minimum BER linear space-time block codes for MIMO systems equipped with MMSE receivers," *IEEE Trans. on Signal Process.,* vol. 54, no. 8, pp. 3147-3158, Aug. 2006.

[8] G. Susinder Rajan and B.S. Rajan, "MMSE Optimal Algebraic Space-Time Codes," *IEEE Trans. on Wireless Communications,* vol. 07, no. 7, pp. 2468–2472, Jul. 2008.

[9] Y. Shang and X.-G. Xia, "Space time block codes achieving full diversity with linear receivers," *IEEE Trans. on Inform. Theory,* vol. 54, no. 10, pp. 4528–4547, Oct. 2008.

[10] W. Zhang, T. Xu, and X.-G. Xia, "Two designs of space-time block codes achieving full diversity with partial interference cancellation group decoding," *IEEE Trans. on Inform. Theory,* vol. 58, no. 2, pp. 747–764, Oct. 2012.

[11] H. Wang, X.-G. Xia, Q. Yin, and B. Li, "A family of space-time block codes achieving full diversity with linear receivers," *IEEE Trans. on Communications,* vol. 57, no. 12, pp. 3607–3617, Dec. 2009.

[12] B. Vrigneau, J. Letessier, P. Rostaing, L. Collin, and G. Burel, "Extension of the MIMO precoder based on the minimum Euclidean distance: A cross-form matrix," *IEEE J. Sel. Topics Signal Process.,* vol. 2, no. 2, pp. 135–146, Apr. 2008.

[13] S.K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "MIMO Precoding With X- and Y-Codes," *IEEE Trans. on Inform. Theory,* vol. 57, no. 6, pp. 3542–3566, June 2011.

[14] K.P. Srinath, B.S. Rajan, "A Low ML-Decoding Complexity, Full-Diversity, Full-Rate MIMO Precoder," *IEEE Trans. on Signal Process.,* vol. 59, no. 11, pp. 5485–5498, Nov. 2011.

[15] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.,* vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[16] A. Scaglione, P. Stoica, S. Barbarossa, G.B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. on Signal Process.,* vol. 50, no. 5, pp. 1051–1064, May 2002.

[17] P. Rostaing, O. Berder, G. Burel, and L. Collin, "Minimum BER diagonal precoder for MIMO digital transmission," *Signal Process.,* vol. 82, no. 10, pp. 1477–1480, Oct. 2002.

[18] D.P. Palomar, J.M. Cioffi, and M.A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE. Trans. on Signal Process.,* vol. 51, no. 9, pp. 2381–2401, Sep. 2003.

[19] C. Peel, B. Hochwald, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Trans. Commun.,* vol. 53, no. 1, pp. 195–202, Jan. 2005.

[20] A.H. Mehana and A. Nosratinia, "Diversity of MIMO Linear Precoding," *IEEE Trans. on Inform. Theory,* vol. 60, no. 2, pp. 1019–1038, Feb. 2014.

[21] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, "Integer-forcing linear receivers," *To appear in IEEE Trans. on Inform. Theory.* Available online at arXiv:1003.5966v3.

[22] A. Sakzad, J. Harshan, and E. Viterbo, "Integer-forcing MIMO linear receivers based on lattice reduction," *IEEE Trans. on Wireless Communications,* vol. 12, no. 10, pp. 4905–4915, Nov. 2013.

[23] A. Sakzad, J. Harshan, and E. Viterbo, "On complex LLL algorithm for integer forcing linear receivers," in the Proc. of in Proc. *IEEE Australian Communication Theory Workshop 2013,* Adelaide, Australia, Jan. 2013, pp. 13–18.

[24] A. Sakzad and E. Viterbo, "Unitary Precoding for Integer-Forcing MIMO Linear Receivers," *To appear in Proc. of IEEE Information Theory Workshop (ITW) 2014.*

[25] J.C.Belfiore, G. Rekaya, and E. Viterbo, "The golden code: a 22 full-rate space-time code with nonvanishing determinants," *IEEE Trans. on Inform. Theory,* vol. 51 no. 4, pp. 1432–1436, April 2005.

[26] O. Ordentlich and U. Erez, "Precoded integer-forcing universally achieves the MIMO capacity to within a constant gap," submitted to *IEEE Trans. on Inform. Theory,* 2013. Also available online at arXiv:1301.6393v2.

[27] S. Bergman and B. Ottersten, "Lattice-based linear pre-coding for MIMO channels with transmitter CSI," *IEEE Trans. on Signal Process.,* vol. 56, no. 7, pp. 29002–2914, July 2008.

[28] A. Lozano, A. Tulino, and S. Verdu, "Optimum power allocation for parallel Gaussian channels with arbitrary in-put distributions," *IEEE Trans. on Inform. Theory,* vol. 52 no. 7, pp. 3033–3051, July 2006.

[29] F. Perez-Cruz, M.R. Rodrigues, and S. Verdu, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Trans. on Inform. Theory,* vol. 56, no. 3, pp. 1070–1084, Mar. 2010.

[30] E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Trans. on Telecommunications,* vol. 10 no. 6, pp. 585–595, 1999.

[31] M. Payaro and D.P. Palomar, "Hessian and Concavity of Mutual Information, Differential Entropy, and Entropy Power in Linear Vector Gaussian Channels," *IEEE Trans. on Inform. Theory,* vol. 55, no. 8, pp. 3613–3628, Aug. 2009.

[32] F. Oggier and E. Viterbo, "Algebraic number theory and code design for Rayleigh fading channels," in *Foundations and Trends in Communications and Information Theory,* vol. 1, pp. 333–415, 2004.

[33] E. Bayer-Fluckiger, F. Oggier, and E. Viterbo, "New Algebraic Constructions of Rotated $\mathbb{Z}^n$ Lattice Constellations for the Rayleigh Fading Channel," *IEEE Trans. on Inform. Theory,* vol. 50, no. 4, pp. 702–714, Apr. 2004.

[34] F. Oggier and E. Bayer-Fluckiger, "Best rotated cubic lattice constellations for the Rayleigh fading channel," in Proc. of *IEEE ISIT-03,* Yokohama, Japan, Jul. 2003, pp. 37.

[35] J.J. Boutros and E. Viterbo "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. on Inform. Theory,* vol. 44, no. 4, pp. 1453–1467, July. 1998.

[36] X. Giraud, E. Boutillon, and J.C. Belfiore, "Algebraic tools to build modulation schemes for fading channels," *IEEE Trans. on Inform. Theory,* vol. 43, no. 3, pp. 938–952, May 1997.

[37] List of full diversity algebraic rotations, available online at: http://www.ecse.monash.edu.au/staff/eviterbo/rotations/rotations.html.

[38] J. Conway and N. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. on Inform. Theory,* vol. 29, no. 6, pp. 820-824, Nov. 1983.

[39] J.H. Conway and N.J.A. Sloane, "Sphere Packing, Lattices and Groups," 3rd ed., New York, Springer-Verlag, 1998.

[40] M.A. Woodbury, "Inverting modified matrices," *Statistical Research Group, Memo. Rep., Princeton University, Princeton, N. J.,* no. 42 pp. 65.0X, 1950.

[41] C. Feng, D. Silva, and F.R. Kschichang, "An algebraic approach to physical-layer network coding," *IEEE Trans. on Inform. Theory,* vol., 59, no. 11, pp. 7576–7596, Nov. 2013.

[42] D. Kapetanovic, H.V. Cheng, W.H. Mow, and F. Rusek, "Optimal Two-Dimensional Lattices for Precoding of Linear Channels," *IEEE Trans. on Wireless Communications,* vol. 12, no. 5, pp. 2104–2113, May 2013.

[43] J.H. Conway and N.J.A. Sloane, "Sphere Packing, Lattices and Groups," 3rd ed., New York, Springer-Verlag, 1998.

[44] K. Zyczkowskit and M. Kus "Random unitary matrices," *Journal of Phys. A: Math. Gen.,* vol. 27, pp. 4235–4245, 1994.

[45] B. Vallee and A. Vera, "Lattice reduction in two dimensions: analyses under realistic probabilistic models," In*Proceedings of the 13th Conference on Analysis of Algorithms, AofA 07. DMTCS Proceedings,* 2007.

[46] A. Sakzad, E. Viterbo, J.J. Boutros, and Y. Hong, "Phase Precoding for the Compute-and-Forward Protocol," *submitted to IEEE Trans. on Inform. Theory.* Available online at: http://arxiv.org/abs/1404.4157.

[47] A. Sakzad, E. Viterbo, J.J. Boutros, and Y. Hong, "Phase precoded compute-and-forward with partial feedback," *Proceedings of 2014 IEEE International Symposium on Information Theory (ISIT),* pp. 2117–2121, June. 29 2014-July 4 2014.