IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 27, 2019, accepted June 16, 2019, date of publication June 20, 2019, date of current version July 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924020

Joint Integer-Forcing Precoder Design for MIMO Multiuser Relay System

HUA JIANG[1], (Member, IEEE), JIAN ZHAO[1], LIZHEN SHEN[2], HAO CHENG[1], AND GUOQING LIU[1]
[1]School of Physical and Mathematical Sciences, Nanjing Tech University, Nanjing 211816, China
[2]School of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, Nanjing 211816, China

Corresponding author: Hua Jiang (quantumcat@163.com)

This work was supported in part by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant 17KJB510024, and in part by the National Natural Science Foundation of China under Grant 61501224 and Grant 61702259.

ABSTRACT In this paper, we study a relay-aided multiuser amplify-and-forward relaying system where a multi-antenna base station (BS) transmits downlink to multiple single-antenna mobile stations (MSs) via a multi-antenna relay station (RS). We propose a novel precoder design scheme by using the joint integer-forcing method to achieve high rate with the algebraic structure of lattice codes. Inspired by the concept of decoding the linear equations in compute-and-forward relay systems, the linear precoding matrix is generalized as an integer matrix that is approximate to the channel matrix, thereby mitigating the multiuser interference (MUI) at the MSs. The power allocation optimization problem at the BS is taken into account to improve the performance of the scheme. The simulation results show the outstanding performance of the proposed scheme.

INDEX TERMS Multiuser, integer-forcing, multiple-input multiple-output (MIMO), compute-and-forward, linear precoding.

I. INTRODUCTION

Cooperative data transmission has received much research attention in recent years. By equipping these networks with one or more relays, cooperative communication can enhance the coverage, reliability, power efficiency, and capacity of wireless networks [1]–[5]. According to the re-transmission mode of the relay, various cooperative transmission protocols can be used, such as the amplify-and-forward (AF) and decode-and-forward (DF) protocols. Recently, the compute-and-forward (CF) protocol has been developed as a new relaying technique that can offer significantly higher rates compared with the other relaying strategies [6]–[13]. The underlying principle of CF is that relays decode the integer linear equations of transmitted messages from the source nodes. When given enough linear independent equations, the destination can solve for the original messages by using a matrix of full-rank integer coefficients [6]. Based on the idea of CF, a linear receiver structure for multiple-input multiple-output (MIMO) system called integer-forcing receiver is proposed in [14]–[16] to recover different integer combinations of lattice codewords. Given the algebraic structure of lattice codes, any integer combination of lattice codewords is still considered a codeword, and such property can help detect the original message. With its favorable approximation of the channel real coefficients by finding the optimal integer coefficients, the CF scheme can offer higher rates with low complexity compared with the other protocols [17]–[19].

In a practical MIMO multiuser relay system, due to the small size and cost constraints, MSs are generally equipped with one antenna. Cooperation among the MSs is assumed to be unavailable. Therefore, the CF technique faces challenges related to the inherent multiuser interference (MUI) in the system. However, the MUI can be nullified by an efficient precoder in a traditional MIMO system [20]–[25]. When the channel state information is available at the transmitter, precoding design becomes an attractive approach for avoiding the MUI at the MSs. Furthermore, precoder design for power allocation can optimize the performance of MIMO multiuser relay systems. Many studies have investigated precoding designs for AF MIMO relay systems that are combined with power allocation optimization, such as the joint zero-forcing and joint MMSE precoder designs [3], [4]. In [2], the multiuser precoding strategy is implemented under

The associate editor coordinating the review of this manuscript and approving it for publication was Liang Yang.

the AF protocol by jointly designing Tomlison-Harashima precoding (THP) at the MIMO BS and linear signal processing at the MIMO RS. While these methods are simple to implement, they suffer a performance penalty under the poor channel conditions.

In this paper, we study a MIMO multiuser relay network and design a new precoding strategy jointly at the BS and RS to avoid MUI and to maximize the total rate. We use an integer-forcing (IF) approach inspired by the CF framework to achieve a promising gain. Different from the MIMO scenario, the multiuser relay system contains the RS retransmission progress and the corresponding transmission model is much complicated. The message precoding for RS is incorporated in the BS precoding design. The precoding approach can be described as a concatenation of three stages. In the first stage, the BS performs message precoding at the message level in the finite field. The integer matrices corresponding to the channel matrices between the BS and RS and between the RS and MSs are pre-inverted to utilize the precoder. In the second stage, the auxiliary messages are mapped to the codewords in the nested lattice. In the last stage, after performing quantization and the modulo operation on the received signal, the RS applies linear beamforming with the AF process to cancel out the MUI. Benefiting from multi-level precoding, the MSs can recover the original message as intended while avoiding MUI.

This paper is organized as follows. Section II introduces the system model and briefly describes multiuser relaying. Section III proposes a novel precoding strategy. Section IV performs some numerical simulations to validate the usefulness of the proposed method. Section V presents some concluding remarks.

II. PRELIMINARIES

We regard the nested code and the relay MIMO channel as a serially concatenated scheme, with a nested lattice encoder, precoded channel matrices and message precoding. In this section, we review the MIMO relay architecture and the lattice structure, including the nested lattice encoding at the BS and CF (lattice) decoding at the RS and MSs.

A. SYSTEM MODEL

As shown in Fig. 1, we consider an AF MIMO multiuser relay system that comprises an M-antenna BS and an N-antenna RS, with $M \geq N$. K MSs are selected to be served via

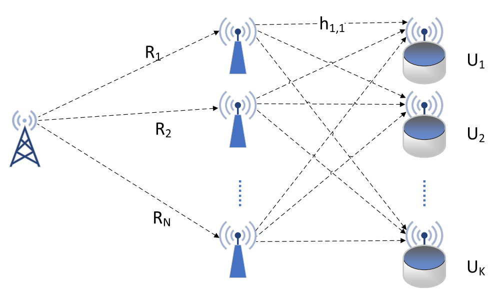


FIGURE 1. Multiuser broadcast system with a fixed multi-antenna RS.

the fixed half-duplex RS. To facilitate the system utilization, we assume $K = N$. In such case, the channel between BS and RS is a point-to-point MIMO link, whereas the RS-MS channel is a multiple-input single-output broadcast channel (MISO-BC). The direct connections between the BS and MSs are reasonably be neglected. The MIMO channel matrices involved are always complex valued. Therefore, the complex-valued systems have to be decoupled into their real and imaginary parts. For example, $\mathrm{y} = \mathrm{Hx} + \mathrm{n}$ is decoupled as,

$$\begin{bmatrix} Re(\mathrm{y}) \\ Im(\mathrm{y}) \end{bmatrix} = \begin{bmatrix} Re(\mathrm{H}) & -Im(\mathrm{H}) \\ Im(\mathrm{H}) & Re(\mathrm{H}) \end{bmatrix} \begin{bmatrix} Re(\mathrm{x}) \\ Im(\mathrm{x}) \end{bmatrix} + \begin{bmatrix} Re(\mathrm{n}) \\ Im(\mathrm{n}) \end{bmatrix}. \quad (1)$$

For the purpose of analysis, we consider the effective real-valued systems. Note that the re-formulation will double the effective lattice dimension and that the corresponding complexity also increases.

Supposing in every transmission period with length L channel uses at the BS, each transmit antenna sends a data vector of length L to the RS. $\mathbf{x}_m,\ m = 1, \cdots, M$ denote the precoded signal vector and are transformed to $\mathbf{X} \in \mathbb{R}^{M \times L}$ in a row-wise manner. The relay works in the half-duplex mode, and the BS and RS transmissions are synchronized with the typical two-hop mode. In the first hop, the BS transmits the symbol vectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M$ to the RS, and the received signal in the RS can be expressed as

$$\mathbf{Y}_{RB} = \mathbf{H}_{RB}\mathbf{X} + \mathbf{N}_{RB}, \quad (2)$$

where $\mathbf{H}_{RB}$ denotes the channel matrix between BS and RS where each element is assumed to be an independent and identically distributed (i.i.d) zero-mean unit-variance Gaussian random variable, while $\mathbf{N}_{RB}$ is an additive Gaussian noise where each term is modeled as an i.i.d. zero-mean Gaussian random variable with unit variance.

The signal $\mathbf{Y}_{RB}$ is a combination of the same set of source signals. The relay needs to apply quantization and modulo operations on the received signal. We consider the AF processing at the RS. By assuming that the re-encoded signal is $\overline{\mathbf{X}}_R$, the transmit signal in the second hop is given by

$$\mathbf{X}_R = \mathbf{W}_R\overline{\mathbf{X}}_R. \quad (3)$$

Let $\mathbf{h}_{RM,k}$ denote the MISO channel between the RS and kth MS. The received signal with respect to the RS transmission at the kth MS can be expressed as

$$\mathbf{y}_{MR,k} = \mathbf{h}_{MR,k}\mathbf{X}_R + \mathbf{n}_{MR,k} \quad (4)$$

where $\mathbf{n}_{MR,k}$ is the noise at the kth MS receiver that is assumed to be an i.i.d. zero-mean Gaussian random variable with unit variance. We stack the vector $\mathbf{h}_{RM,k}$ row-wise to form the $K \times N$ matrix $\mathbf{H}_{MR}$.

B. LATTICE AND NESTED LATTICE CODES

Lattices are exploited to develop strong source and channel codes for many communications scenarios. Signal processing applications where lattice structure has been successfully used include lattice reduction, nest lattice codes [8], [26]–[28]. Lattice reduction is referred to finding

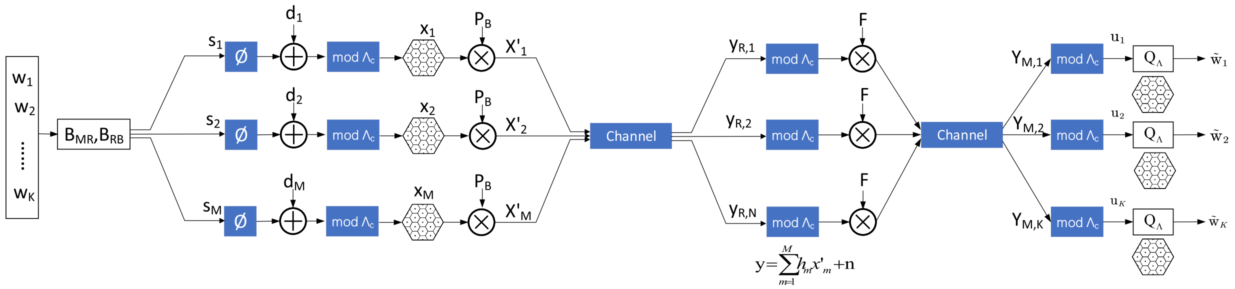


FIGURE 2. **The BS and RS operations for the MIMO multiuser relay network.**

better representations of a given lattice using algorithms like Lenstra, Lenstra, Lovász (LLL) reduction [29]. On the other hand, nested lattice codes are concerned with lattice encoding and decoding for its good statistical and algebraic properties for constellation shaping.

Nested lattice codes are important components of the CF framework. To facilitate the further work, We initially review some basic concepts related to these codes. An L-dimensional lattice Λ with a basis set $\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_K \in \mathbb{R}^L$ represents a set of all possible vectors or points and is given by

$$\Lambda = \{\mathbf{s} = \mathbf{G}\mathbf{w} : \mathbf{w} \in \mathbb{Z}^K\}, \tag{5}$$

where $\mathbf{G} = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_K]$ is a lattice-generating matrix of Λ. We only consider full rank lattices where $L = K$. The fundamental Voronoi region of Λ is defined as $\mathcal{V} = \{\mathbf{x} \in \mathbb{R}^L : \mathcal{Q}(\mathbf{x}) = 0\}$ and comprises all points in $\mathbb{R}^L$ closest to the zero vector. Meanwhile, $\mathcal{Q}(\mathbf{x})$ denotes the nearest neighbor quantizer of $\mathbf{x}$ on Λ. The modulo-Λ is defined as

$$\mathbf{x} \bmod \Lambda = \mathbf{x} - \mathcal{Q}(\mathbf{x}). \tag{6}$$

If two lattices Λ_1 and Λ_2 are available, then $\Lambda_1 \subseteq \Lambda_2$ holds when Λ_1 is nested in Λ_2. The nested lattice code set $\mathcal{C}$ is formed by taking all the elements of Λ_2 that lie in the fundamental Voronoi region of Λ_1, that is, $\mathcal{C} = \Lambda_2 \bigcap \mathcal{V}_{\Lambda_1}$.

III. JOINT-INTEGER-FORCING-BASED PRECODING

The Integer-forcing detection is a natural extension of CF, which directly recovers a linear combination of interfering data streams without noise amplification. This concept can be used for precoding in the MIMO broadcast channel [15]. For a multiuser MIMO relay downlink, the interferences among the MSs are always originated from the transmitters, and in principle, a precoder can be well designed to essentially mitigate its effects and provide performance improvement. We focus on the formulation of an integer-forcing-based precoding for multiuser relay systems. Our strategy involves message precoding, nested lattice encoding, linear beamforming and power allocation, all of which are performed at the BS/RS transmitters together with CF decoding at the RS/MS receivers.

A. PRECODING STRUCTURE

The proposed integer-forcing-based precoding scheme for multiuser relay systems is illustrated in Fig. 2. We first build a nest lattice code together with message precoding. Let $\mathbf{w}_i$ denote the information vectors to be transmitted at the ith antenna of the BS. $\mathbf{w}_i$ is drawn independently and uniformly over a prime-size finite field $\mathbb{F}_p^L = \{0, 1, \cdots, p-1\}^L$, where p is prime. A *fine* lattice, also called a coding lattice, is denoted by $\Lambda = \mathbf{G}\mathbb{Z}^L$, where $\mathbf{G} \in \mathbb{R}^{M \times L}$ is a full-rank generator matrix. A *coarse* lattice, also called a shaping lattice, is denoted by $\Lambda_c = p\Lambda$. The corresponding nested lattice can be represented as $\mathcal{C} = \Lambda \bigcap \mathcal{V}_{\Lambda_c}$. Let ϕ be a bijective map between the set of message $\mathbf{w}$ and the elements of the nested lattice code $\phi(\mathbf{w})$. The function $[\phi(\mathbf{w}) = \mathbf{G}\mathbf{w}] \bmod \Lambda$ is a such map that satisfies the one-to-one condition. Specifically, if we take $\mathbf{G} = \mathbf{I}$, we recover Λ to a $\mathbb{Z}^L$ lattice.

We assume two non-singular integer matrices $\mathbf{A}_{RB} \in \mathbb{Z}^{N \times N}$ and $\mathbf{A}_{MR} \in \mathbb{Z}^{K \times K}$ to approximate the channels $\mathbf{H}_{RB}$ and $\mathbf{H}_{MR}$, respectively. Let $\tilde{\mathbf{A}}_{RB}$ and $\tilde{\mathbf{A}}_{MR}$ be the $p-$modulo versions of $\mathbf{A}_{RB}$ and $\mathbf{A}_{MR}$, respectively, i.e., $\tilde{\mathbf{A}}_{RB} = [\mathbf{A}_{RB}] \bmod p$ and $\tilde{\mathbf{A}}_{MR} = [\mathbf{A}_{MR}] \bmod p$. $\tilde{\mathbf{A}}_{RB}$ and $\tilde{\mathbf{A}}_{MR}$ are full rank. That is, $\mathbf{A}_{RB}$ and $\mathbf{A}_{MR}$ are invertible modulo p, i.e., $[\det(\mathbf{A}_{RB})] \bmod p \neq 0$ and $[\det(\mathbf{A}_{MR})] \bmod p \neq 0$. This sufficient condition given for the inverses of $\tilde{\mathbf{A}}_{RB}$ and $\tilde{\mathbf{A}}_{MR}$ taking over $\mathbb{Z}_p$ is accessible. The inverses of $\tilde{\mathbf{A}}_{RB}$ and $\tilde{\mathbf{A}}_{MR}$ can be calculated as

$$\mathbf{B}_{RB} = \frac{[\tilde{\mathbf{A}}_{RB}^*] \bmod p}{[\det(\tilde{\mathbf{A}}_{RB})] \bmod p} \tag{7}$$

and

$$\mathbf{B}_{MR} = \frac{[\tilde{\mathbf{A}}_{MR}^*] \bmod p}{[\det(\tilde{\mathbf{A}}_{MR})] \bmod p}, \tag{8}$$

where $\tilde{\mathbf{A}}_{RB}^*$ and $\tilde{\mathbf{A}}_{MR}^*$ are the adjoint matrices of $\tilde{\mathbf{A}}_{RB}$ and $\tilde{\mathbf{A}}_{MR}$, respectively. Let the original messages $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K$ comprise a $K \times L$ matrix $\mathbf{W}$ in a row-wise way, i.e., $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K]^T$. The original precoded data vector at the BS can be expressed,

$$\mathbf{S} = \mathbf{B}_{MR} \bigotimes \mathbf{W}, \tag{9}$$

where $\bigotimes$ denotes multiplication over the finite field. $\mathbf{S}$ is then precoded with $\mathbf{B}_{RB}$, thereby leading to the precoded messages

$\mathbf{s}'_1, \cdots, \mathbf{s}'_K$, which form the rows of the matrix;

$$\mathbf{S}' = \mathbf{B}_{RB} \bigotimes \mathbf{S}. \tag{10}$$

Afterward, each precoded message $\mathbf{s}'_i$ is mapped to the nested lattice code $\mathbf{t}_i = \phi(\mathbf{s}'_i) \in \mathcal{C}$. A dither vector that is uniformly distributed in the Voronoi region $\mathcal{V}$ of Λ_c should be considered to meet the following power constraint:

$$P_i = \frac{1}{n} E\left[\|\mathbf{t}_i\|^2\right] \leq P_{\Lambda_c}, \tag{11}$$

where P_{Λ_c} is the mean power of the codewords distributed over $\mathcal{V}_{\Lambda_c}$, and $\mathbf{x}_i = [\mathbf{t}_i - \mathbf{d}_i] \bmod \Lambda_c$ is uniformly distributed over Λ_c. We assume that the signal is power normalized by a normalization factor P_{Λ_c}. Note that $\mathbf{d}_i$ is known at the BS for dither elimination and signal recovery.

Provided that the transmit power consumed at the BS in one transmission period is denoted by ρ_B^2, a beamforming matrix $\mathbf{P}_B$ is imposed on $\mathbf{X}$ and the transmitted signal is computed as

$$\mathbf{X}' = \frac{\rho_B}{c_B}\mathbf{P}_B\mathbf{X}^T. \tag{12}$$

c_B is then provided to enforce the power constraint:

$$c_B = \sqrt{\mathrm{tr}(\mathbf{P}_B^T\mathbf{P}_B)} \tag{13}$$

According to the conclusion in [14], in an integer-forcing detection, the sum rate performance is optimal when the equalization matrix is given as $\mathbf{A}\mathbf{H}^T((1/\mathrm{SNR})\mathbf{I} + \mathbf{H}^T\mathbf{H})^{-1}$. We consider the power allocation and express zero-forcing beamforming matrix $\mathbf{P}_{RB}$ as

$$\mathbf{P}_B = \mathbf{H}_{RB}^T(\mathbf{H}_{RB}^T\mathbf{H}_{RB})^{-1}\mathbf{A}_{RB}. \tag{14}$$

By using this precoder, the received signal at the RS is given by

$$\begin{aligned} \mathbf{Y} &= \mathbf{H}_{RB}\mathbf{X}' + \mathbf{N}_{RB} \\ &= \frac{\rho_B}{c_B}\mathbf{H}_{RB}\mathbf{P}_B\mathbf{X}^T + \mathbf{N}_{RB} \\ &= \frac{\rho_B}{c_B}\mathbf{A}_{RB}\mathbf{X}^T + \mathbf{N}_{RB}. \end{aligned} \tag{15}$$

The RS takes advantage of an automatic gain control (AGC) operation to adjust the analog signal level to the dynamic range of an analog-to-digital converter (ADC), thereby yielding

$$\mathbf{Y}' = \mathbf{A}_{RB}\mathbf{X}^T + \frac{c_B}{\rho_B}\mathbf{N}_{RB}. \tag{16}$$

After removing the dithering signals, the effective received signal at ith RS antenna can be written as

$$\begin{aligned} \mathbf{y}'_i &= \mathbf{a}_{RB,i}\mathbf{T}^T + \frac{c_B}{\rho_B}\mathbf{n}_{RB,i} \\ &= \sum_{j=1}^{N} a_{RB,i,j}\mathbf{t}_j + \frac{c_B}{\rho_B}\mathbf{n}_{RB,i}, \end{aligned} \tag{17}$$

where $\mathbf{T} = [\mathbf{t}_1^T, \cdots, \mathbf{t}_N^T]$.

$\{\mathbf{y}'_i\}$ is practically a combination of the original signal $\mathbf{t}_i$, thereby producing correlations among different antennas. Therefore, forwarding $\{\mathbf{y}'_i\}$ directly from the RS to the MSs results in information redundancy. To address this problem, we feed each $\mathbf{y}_i$ to the modulo function with lattice Λ, which yielding,

$$\mathbf{v}_i = \mathbf{y}'_i \bmod \Lambda_c. \tag{18}$$

$\mathbf{v}_i$ can be equivalently written as

$$\mathbf{v}_i = [\mathbf{G}\mathbf{s}_i + \frac{c_B}{\rho_B}\mathbf{n}_{RB,i}] \bmod \Lambda_c. \tag{19}$$

Proof: We review the modulo theory that $[\mathbf{a}+\mathbf{b}] \bmod \Lambda = [\mathbf{a} \bmod \Lambda + \mathbf{b}] \bmod \Lambda$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^L$ holds [6]. In this case, $\mathbf{v}_i$ can be rewritten as

$$\begin{aligned} \mathbf{v}_i &= \mathbf{y}'_i \bmod \Lambda_c \\ &= [\sum_{j=1}^{N} a_{RB,i,j}\mathbf{t}_j + \frac{c_B}{\rho_B}\mathbf{n}_{RB,i}] \bmod \Lambda_c \\ &= \left[[\sum_{j=1}^{N} a_{RB,i,j}\mathbf{t}_j] \bmod \Lambda_c + \frac{c_B}{\rho_B}\mathbf{n}_{RB,i} \right] \bmod \Lambda_c \end{aligned} \tag{20}$$

Recall that since $\lambda_c = p\Lambda$ and $\mathbf{G}$ is the generator matrix of Λ, $p\mathbf{G}\mathbb{Z}^n \in \Lambda_c$, we have

$$\begin{aligned} &\left[\sum_{j=1}^{N} a_{RB,i,j}\mathbf{t}_j\right] \bmod \Lambda_c \\ &= \left[\sum_{j=1}^{N}(a_{RB,i,j}\mathbf{G}[\sum_{l=1}^{N} b_{RB,j,l}\mathbf{s}_l] \bmod p)\right] \bmod \Lambda_c \\ &= \left[\sum_{j=1}^{N}(a_{RB,i,j}[\mathbf{G}\sum_{j=1}^{N} b_{RB,j,l}\mathbf{s}_l - \mathbf{G}\mathcal{Q}_p(\sum_{l=1}^{N} b_{RB,j,l}\mathbf{s}_l)])\right] \bmod \Lambda_c \\ &\overset{(a)}{=} \left[\sum_{j=1}^{N}(a_{RB,i,j}[\mathbf{G}\sum_{l=1}^{N} b_{RB,j,l}\mathbf{s}_l])\right] \bmod \Lambda_c \\ &= \left[\sum_{l=1}^{N}\mathbf{G}\mathbf{s}_l\sum_{j=1}^{N} b_{RB,j,l}a_{i,j}\right] \bmod \Lambda_c \\ &= \left[\sum_{l=1}^{N}\mathbf{G}\mathbf{s}_l\sum_{j=1}^{N} b_{RB,j,l}a_{RB,i,j} - \mathbf{G}\mathcal{Q}_p(\mathbf{s}_l\sum_{j=1}^{N} a_{RB,i,j}b_{RB,j,l})\right] \bmod \Lambda_c \\ &\overset{(b)}{=} \left[\sum_{l=1}^{N}\mathbf{G}\mathbf{s}_l[\sum_{j=1}^{N} b_{RB,j,l}[a_{RB,i,j}] \bmod p] \bmod p\right] \bmod \Lambda_c \\ &\overset{(c)}{=} [\mathbf{G}\mathbf{s}_i] \bmod \Lambda_c, \end{aligned} \tag{21}$$

where $\mathcal{Q}_p(\cdot)$ denotes the quantization over $p\mathbb{Z}$. (a) and (b) follow because $\mathbf{G}\mathcal{Q}_p(\cdot)$ is an element of Λ_c, whereas (c) follows because $\mathbf{B}_{RB}[\mathbf{A}_{RB}] \bmod p = \mathbf{I}$.

Given that $\mathbf{v}_i$ is enforced in the Voronoi region $\mathcal{V}$ of Λ_c, the corresponding power normalization factor is incorporated to meet the power constraint. After performing the simple modulo operation, a similar ZF filter with a suitable power allocation is imposed on $\mathbf{V}$, where $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_K]$. We assume that this precoding matrix at the BS is

$$\mathbf{F} = \mathbf{H}_{MR}^T(\mathbf{H}_{MR}^T\mathbf{H}_{MR})^{-1}\mathbf{A}_{MR}. \tag{22}$$

c_R is provided to enforce the power constraint:

$$c_R = \sqrt{\mathrm{tr}(\mathbf{F}^T\mathbf{F})}. \tag{23}$$

The transmit power consumed at the RS is assumed to be ρ_R^2. By using this precoder, the received signal at the i the antenna of the RS is given by

$$\begin{aligned}\mathbf{Y}_{MR} &= \frac{\rho_R}{c_R}\mathbf{H}_{MR}\mathbf{F}\mathbf{V} + \mathbf{N}_{MR} \\ &= \frac{\rho_R}{c_R}\mathbf{A}_{MR}\mathbf{V} + \mathbf{N}_{MR}\end{aligned} \tag{24}$$

To compensate for the effect of amplification by a factor of ρ_R/c_R, the received signal at the ith MS is divided by c_R/ρ_R via the automatic gain control (AGC) circuit and can be expressed as

$$\mathbf{y}'_{MR,i} = \mathbf{a}_{MR,i}\mathbf{V} + \frac{c_R}{\rho_R}\mathbf{n}_{MR,i}. \tag{25}$$

We feed $\mathbf{y}'_{MR,i}$ to the modulo function with respect to the lattice Λ_c. Based on Eq. (21), we have

$$\begin{aligned}\mathbf{u}_i &= \mathbf{y}'_{MR,i} \bmod \Lambda_c \\ &= [\mathbf{a}_{MR,i}\mathbf{V} + \frac{c_R}{\rho_R}\mathbf{n}_{MR,i}] \bmod \Lambda_c \\ &= [[\sum_{j=1}^{K} a_{MR,i,j}\mathbf{G}\mathbf{s}_j] \bmod \Lambda_c + \frac{c_B}{\rho_B}\sum_{j=1}^{K} a_{MR,i,j}\mathbf{n}_{RB,j} \\ &\quad + \frac{c_R}{\rho_R}\mathbf{n}_{MR,i}] \bmod \Lambda_c \qquad (26) \\ &= [\mathbf{G}\mathbf{w}_i + \frac{c_B}{\rho_B}\sum_{j=1}^{K} a_{MR,i,j}\mathbf{n}_{RB,j} + \frac{c_R}{\rho_R}\mathbf{n}_{MR,i}] \bmod \Lambda_c.\end{aligned}$$

The full rank of matrix $\mathbf{G}$ provides sufficient conditions for recovering original messages via the quantization of $\mathbf{u}_i$, i.e.,

$$\tilde{\mathbf{w}}_i = \mathbf{G}^{-1}\mathcal{Q}_\Lambda(\mathbf{u}_i) \tag{27}$$

where $\mathcal{Q}_\Lambda$ denotes the quantization over lattice Λ. The choice of the lattice generator matrix $\mathbf{G}$ is irrelevant for achievable rates. If we take $\mathbf{G} = \mathbf{I}$, then an integer-lattice is involved and can construct a regular constellation (e.g., PAM and QAM).

B. POWER ALLOCATION AND IF INTEGER MATRIX SEARCH

According to (26), the ith MS equivalently processes detection over an independent modulo-Λ_c channel without interference from the other MS. With the channel input-output relationship, the achievable rate of the ith MS is given by

$$\begin{aligned}R_i &= \log_2^+\left(\frac{c_B^2}{\rho_B^2}\sum_{j=1}^{N}|a_{MR,i,j}|^2 + \frac{c_R^2}{\rho_R^2}\right)^{-1} \\ &= \log_2^+\left(N\frac{c_B^2}{\rho_B^2}\|\mathbf{a}_{MR,i}\|^2 + \frac{c_R^2}{\rho_R^2}\right)^{-1}.\end{aligned} \tag{28}$$

The sum rate achievable by the precoding scheme is then given by

$$R_s = \sum_{i=1}^{N}\log_2^+\left(N\frac{c_B^2}{\rho_B^2}\|\mathbf{a}_{MR,i}\|^2 + \frac{c_R^2}{\rho_R^2}\right)^{-1}. \tag{29}$$

Let P_{total} represent the total transmit power consumed for the whole system at each period. P_{total} is computed as the sum of transmit power ρ_B^2 at the BS and transmit power ρ_R^2 at the Rs, i.e., $P_{total} = \rho_B^2 + \rho_R^2$. Our objective is to maximize the sum rate while keeping the total power P_{total} fixed. Provided that the integer matrices $\mathbf{A}_{RB}$ and $\mathbf{A}_{MR}$ are fixed, the power allocation between BS and RS can be formulated as

$$\begin{aligned}\min_{\rho_B,\rho_R} \quad & \sum_{i=1}^{N}\log_2^+\left(N\frac{c_B^2}{\rho_B^2}\|\mathbf{a}_{MR,i}\|^2 + \frac{c_R^2}{\rho_R^2}\right) \\ \text{subject to} \quad & \rho_B^2 + \rho_R^2 = P_{total}\end{aligned} \tag{30}$$

The global optimal solution of the problem can be simply obtained by using standard convex optimization approaches that typically fast algorithms and a low computational complexity. Eq. (30) can be converted to a convex unconstrained optimization problem with a single variable. Assuming $x = \rho_B^2$ and

$$u_i(x) = N\frac{c_B^2}{x}\|a_{MR,i}\|^2 + \frac{c_R^2}{P_{total} - x}, \tag{31}$$

we have

$$f(x) = \sum_{i=1}^{N}\log_2^+ u_i(x). \tag{32}$$

The unconstrained minimization problem $\min f(x)$ can be solved efficiently by using the Newton method [30]. We first solve the equation

$$\begin{aligned}f(x)' + f(x)''\Delta x &= 0 \\ \Delta x &= -\frac{f'(x)}{f''(x)}\end{aligned} \tag{33}$$

where $f(x)'$ is the derivative of $f(x)$ and $f(x)''$ is the second derivative of $f(x)$, i.e.,

$$\begin{aligned}f'(x) &= \sum_{i=1}^{N}\frac{u_i'(x)}{u_i(x)} \\ f''(x) &= \sum_{i=1}^{N}\frac{u_i''(x)u_i(x) - u_i'^2(x)}{u_i^2(x)},\end{aligned} \tag{34}$$

where

$$\begin{aligned}u_i'(x) &= \frac{c_R^2}{(P_{total} - x)^2} - N\frac{c_B^2}{x^2}\|a_{MR,i}\|^2 \\ u_i''(x) &= -\frac{2c_R^2}{(P_{total} - x)^3} + 2N\frac{c_B^2}{x^3}\|a_{MR,i}\|^2,\end{aligned} \tag{35}$$

Given the initial guess x_0, the iterative scheme can be generalized as

$$x_{n+1} = x_n - f'(x_n)/f''(x_n), \quad n = 0, 1, \cdots \tag{36}$$

The sequence $x_1, x_2, \cdots$ converges to a point x^*, thereby satisfying $f'(x^*)$. As n increases, the approximation becomes more accurate.

Eq. (29) shows that while the transmit powers ρ_B^2 and ρ_R^2 are fixed, the minimization of the achievable sum rate is determined by the power enhancement c_B^2 and c_R^2. By appropriately choosing the integer matrices $\mathbf{A}_{RB}$ and $\mathbf{A}_{MR}$ subject to the full-rank constraint, we aim to minimize c_B^2 and c_R^2. This objective problem can be transformed into

$$\mathbf{A}_{RB} = \arg\min \sum_{i=1}^{N} \|\mathbf{H}_{RB}^{\dagger}\mathbf{a}_{RB,i}\|^2 \quad \text{subject to} \quad \mathrm{Rank}(\mathbf{A}_{RB}) = N \tag{37}$$

and

$$\mathbf{A}_{MR} = \arg\min \sum_{i=1}^{K} \|\mathbf{H}_{MR}^{\dagger}\mathbf{a}_{MR,i}\|^2 \quad \text{subject to} \quad \mathrm{Rank}(\mathbf{A}_{MR}) = N. \tag{38}$$

Finding the optimal integer matrix is known to be a short vector problem in the CF architecture. An exhaustive search guarantees the optimal integer matrix every time but introduces significant computational complexity. To overcome this problem, some alternative algorithms have been proposed including the greedy algorithm and sphere coding algorithm (for more details, refer to [14], [31], [32]).

IV. NUMERICAL RESULTS

In this section, we present the numerical results for the sum-rate and bit error rate (BER) performance of the proposed schemes for the integer-forcing precoding design. We average over 10000 randomly generated channel realizations. The channel matrix is generated randomly from one burst to the next with i.i.d elements $H_{i,j} \sim \mathcal{CN}(0, 1)$. The total power is assumed to be fixed and the power allocation between BS and RS is optimized by using the newton method. In the initiation setup, half of the total power is allocated to the BS while the other half is allocated to the RS. In other words, x_0 is initialized to $P_{total}/2$. With this appropriate initiation, the iteration algorithm demonstrates a rapid convergence.

A. SUM RATE RESULTS

Figure 3 illustrates the achievable sum rate of the proposed integer-forcing precoding scheme with respect to total power. The number of transmit antennas at the BS and RS is $M = N = 4$, and $K = 4$ MSs are available in the system. For comparison, the sum capacity obtained by lattice reduction (LR) precoding and THP SVD scheme [2] is also shown. The transmitted symbols for LR precoding are modulated with 4-QAM. At high SNR regime, the THP SVD scheme suffers from a loss of multiplexing gain. The performance difference between the integer-forcing precoding and the LLL precoding is up to 1 dB at a sum rate of 10 bits/s/Hz. The difference between the slopes of the integer-forcing precoding and LLL precoding curves increases along with SNR

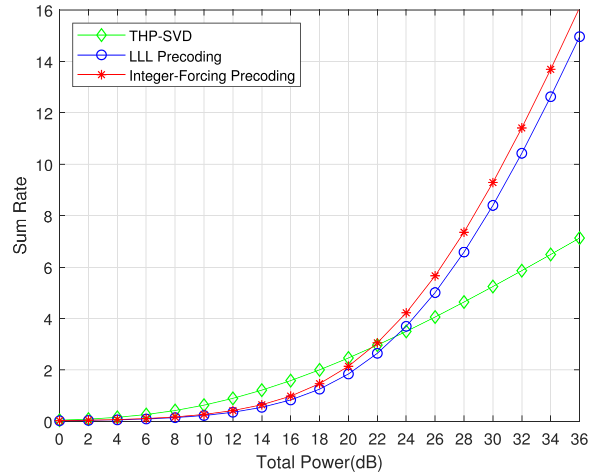


FIGURE 3. Average sum rate for different precoding strategies.

because the integer matrix $\mathbf{U}$ in LR must be unimodular, which is a strong condition. This constraint would keep the LR precoder from finding the optimal integer matrix $\mathbf{U}$. In comparison, the integer-forcing method only requires a non-singular propriety for the integer matrix $\mathbf{A}$. This relaxation in the constraint offers a better integer matrix $\mathbf{A}$ and explains the weakness of the LR-aid precoder and the benefits of the integer-forcing precoder for the multiuser relay system.

B. BER RESULTS

To examine the BER performance, we adopt the points drawing from a 2-dimensional integer lattice as codewords. With reference to the precoding structure shown in Section III, we use $M = N = K = 4$ and $\Lambda = \mathbb{Z}$. The transmitted symbols are modulated with 4-QAM. The corresponding finite ring is $\mathbb{Z}_2 = \{0, 1\}$, the finite constellation is $\mathcal{S} = \{0, 1, i, 1+i\}$, and $\mathcal{S}$ is the set of coset representatives of $\mathbb{Z}[i]/2\mathbb{Z}[i]$. The symbols of $\mathcal{S}$ are transformed into scaled and shifted versions with a reduced average transmit power. By using Eq. (1), the complex symbols are unfolded to real vectors whose elements are in $\{0, 1\}$. In this way the modulo lattice operation has been well implemented with respect to precoding and decoding.

Figure 4 shows the BER performance of the proposed integer-forcing precoding schemes. For comparison, the LLL-aided precoding and THP SVD schemes is also investigated. The integer-forcing precoding method performs much better than the LLL precoding method, yielding a clear advantage and evidently demonstrates a better performance by over 1 dB at a BER of 10^{-4}. In addition, it is worth noting that LR is helpful to collect the high diversity. When SNR is large, the THP SVD scheme exhibits poor BER performance. This is because the distribution of singular value is highly unequal and the worst subchannel (smallest singular value) is the dominant contributor to the BER.

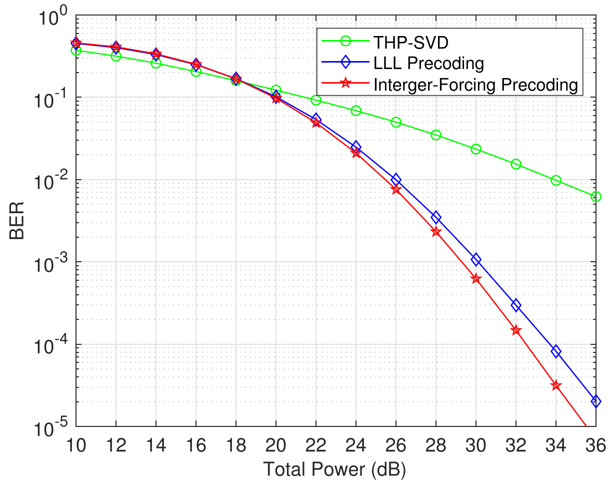


FIGURE 4. BER performance of precoded MIMO relay systems.

V. CONCLUSIONS

In this paper, we propose a new and efficient integer-forcing-based precoding scheme for MIMO multiuser systems. To benefit from the high computation rate of the CF scheme, we design the precoded nested lattice codes utilized in the transmitter. An upside of this proposed approach is that the channels from the BS to RS and from the RS to MSs are all involved in the integer-forcing process. The total power is allocated optimally between the BS and RS to maximize the sum rate performance. We analyze the differences between the proposed integer-forcing based precoding scheme and the LLL reduced based precoding scheme and find that the former outperforms the latter in terms of sum rate and bit error.

REFERENCES

[1] N. Guo, J. Ge, Q. Bu, and C. Zhang, "Multi-user cooperative non-orthogonal multiple access scheme with hybrid full/half-duplex user-assisted relaying," *IEEE Access*, vol. 7, pp. 39207–39226, 2019.

[2] C.-B. Chae, T. Tang, R. W. Heath, Jr., and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 727–738, Feb. 2008.

[3] U. Phuyal, S. C. Jha, and V. K. Bhargava, "Joint zero-forcing based precoder design for QoS-aware power allocation in MIMO cooperative cellular network," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 350–358, Feb. 2012.

[4] F.-S. Tseng, W.-R. Wu, and J.-Y. Wu, "Joint source/relay precoder design in nonregenerative cooperative systems using an MMSE criterion," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4928–4933, Oct. 2009.

[5] Y. Huang, F. Al-Qahtani, C. Zhong, Q. Wu, J. Wang, and H. Alnuweiri, "Performance analysis of multiuser multiple antenna relaying networks with co-channel interference and feedback delay," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 59–73, Jan. 2014.

[6] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[7] Y. Tan and X. Yuan, "Compute-compress-and-forward: Exploiting asymmetry of wireless relay networks," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 511–524, Jan. 2016.

[8] Y. C. Huang, K. R. Narayanan, and P. C. Wang, "Lattices over algebraic integers with an application to compute-and-forward," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6863–6877, Oct. 2018.

[9] J. Zhu and M. Gastpar, "Gaussian multiple access via compute-and-forward," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2678–2695, May 2017.

[10] M. Nokleby and B. Aazhang, "Cooperative compute-and-forward," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 14–27, Jan. 2016.

[11] C. Guo, A. N. Aljalai, C. Feng, L. Zhao, V. C. M. Leung, and R. K. Ward, "Compute-and-forward for uplink non-orthogonal multiple access," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 986–989, Dec. 2018.

[12] H. Cheng, X. Yuan, and Y. Tan, "Generalized compute-compress-and-forward," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 462–481, Jan. 2019.

[13] S. Ashrafi, C. Feng, and S. Roy, "Compute-and-forward for random-access: The case of multiple access points," *IEEE Trans. Commun.*, vol. 66, no. 8, pp. 3434–3443, Aug. 2018.

[14] L. Wei and W. Chen, "Integer-forcing linear receiver design with slowest descent method," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2788–2796, Jun. 2013.

[15] D. Silva, G. Pivaro, G. Fraidenraich, and B. Aazhang, "On integer-forcing precoding for the Gaussian MIMO broadcast channel," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4476–4488, Jul. 2017.

[16] S. M. Azimi-Abarghouyi, M. Hejazi, B. Makki, M. Nasiri-Kenari, and T. Svensson, "Integer-forcing message recovering in interference channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4124–4135, May 2018.

[17] W. Liu and C. Ling, "Efficient integer coefficient search for compute-and-forward," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8039–8050, Dec. 2016.

[18] S. Sahraei and M. Gastpar, "Polynomially solvable instances of the shortest and closest vector problems with applications to compute-and-forward," *IEEE Trans. Inf. Theory*, vol. 63, no. 12, pp. 7780–7792, Dec. 2017.

[19] H. Cheng, X. Yuan, and Y. Tan, "Compute-compress-and-forward: New results," in *Proc. IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.

[20] A. Sakzad and V. Emanuele, "Full diversity unitary precoded integer-forcing," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4316–4327, Aug. 2015.

[21] A. S. Guerreiro, G. Fraidenraich, and L. L. Mendes, "Approximate sum rate for integer-forcing receiver with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11767–11775, Dec. 2018.

[22] T. Yang, "Distributed MIMO broadcasting: Reverse compute-and-forward and signal-space alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 581–593, Jan. 2017.

[23] Y. Chen, Y. Song, and N. Devroye, "The capacity region of the L-user Gaussian inverse compute-and-forward problem," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 6953–6968, Dec. 2016.

[24] H. Cheng, X. Yuan, and T. Yang, "Generalized signal-space alignment based physical-layer network coding for distributed MIMO systems," *IEEE Access*, vol. 7, pp. 48430–48444, 2019.

[25] M. M. Molu, K. Cumanan, and A. Burr, "Low-complexity compute-and-forward techniques for multisource multirelay networks," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 926–929, May 2016.

[26] M. M. Molu, K. Cumanan, M. Bashar, and A. Burr, "On convolutional lattice codes and lattice decoding using trellis structure," *IEEE Access*, vol. 4, pp. 9702–9715, 2016.

[27] S. Kim and J. Chun, "Capacity and performance of lattice reduction aided linear processing with lattice encoding and decoding in limited feedback systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1567–1577, Oct. 2008.

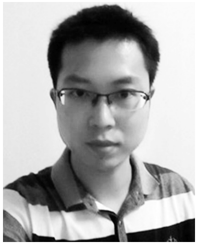
[28] A. Sakzad, J. Harshan, and V. Emanuele, "Integer-forcing MIMO linear receivers based on lattice reduction," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 4905–4915, Oct. 2013.

[29] A. K. Lenstra, H. W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, no. 4, pp. 515–534, 1982.

[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: University Press, 2004.

[31] Q. Huang and A. Burr, "Low complexity coefficient selection algorithms for compute-and-forward," *IEEE Access*, vol. 5, pp. 19182–19193, 2017.

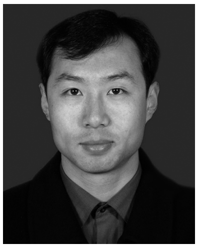
[32] Y.-L. Chuang, C.-I. Fan, and Y.-F.Tseng, "An efficient algorithm for the shortest vector problem," *IEEE Access*, vol. 6, pp. 61478–61487, 2018.



HUA JIANG was born in Nantong, China. He received the Ph.D. degree in electrical engineering from Nanjing University, in 2013. He is currently a Lecturer with the School of Physical and Mathematical Sciences, Nanjing Tech University. His research interests include cognitive radio, OFDM, and cooperative networks for wireless communications.

JIAN ZHAO received the M.E.E. degree in computer science and technology and the Ph.D. degree from the College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University, Nanjing, China, in 2011 and 2016, respectively. He was a Visiting Scholar under the supervision of Prof. H. Jiang with Northwestern University, Evanston, IL, USA, from 2017 to 2018. He is currently an Associate Professor with Nanjing Tech University. His current research interests include in the fields of digital signal processing, biological sequences analysis of bioinformatics, and mathematical modeling.

LIZHEN SHEN was born in Nanjing, China. He is currently pursuing the Ph.D. degree with the College of Biotechnology and Pharmaceutical Engineering, Nanjing Tech University. His research interests include signal processing and wireless communications.



HAO CHENG was born in Laiyang, China. He received the M.S. degree from Nanjing Tech University, in 2004. He is currently an Associate Professor with Nanjing Tech University. His research interests include signal processing and scientific computation.

GUOQING LIU received the Ph.D. degree in mathematics from Nanjing University, in 2001. He is currently a Professor with the College of Sciences, Nanjing Tech University, China. His research interests include statistical signal processing, signal representation, and blind source separation.