

“© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

A Linear Network Coding Approach for Uplink Distributed MIMO Systems: Protocol and Outage Behavior

Tao Yang, *Member, IEEE*, Qifu Tyler Sun, *Member, IEEE*, Jian A. Zhang, *Senior Member, IEEE* and Jinhong Yuan, *Senior Member, IEEE*

Abstract—A distributed MIMO system consists of M users served by L distributed base stations (BSs), where the BSs are connected to a central unit (CU) via L independent rate-constrained backhaul (BH) links. In this paper we consider the design of an uplink distributed MIMO system where the channel state information is not available at the transmitters (users), under the BH rate-constraint. We propose a new linear network coding (LNC) based protocol. In this protocol, each BS attempts to forward N pre-assigned linear functions of the M users' messages, where N is determined by the BH rate-constraint. The CU recovers all users' messages by solving the linear functions, provide that the linear functions are full rank. The coefficients of the pre-assigned linear functions are carefully designed to minimize the probability of rank deficiency of the distributed MIMO system. Then we analyze the outage probability (OP) performance of the proposed scheme. We analytically show that as long as the BH rate is greater than the individual data rate of one user, the OP of the proposed scheme decays like $\frac{1}{\text{SNR}^L}$ at high SNR. As the BH rate-constraint approaches M times the data rate of one user, the performance of the proposed scheme is $\frac{10}{L} \log_{10}(L!)$ dB away from that of the full MIMO scenario at high SNR. We also develop a structured way to efficiently construct the coefficients of pre-assigned linear functions that yield the optimized OP performance. Numerical results show that the proposed scheme has significantly improved OP and frame error rate performance over existing schemes.

I. INTRODUCTION

In a distributed multiple-input multiple-output (D-MIMO) system, a number of mobile users are served by a number of distributed base stations. These base stations are connected to a central unit via independent backhaul links. By carefully designing the functionality of the base stations, a D-MIMO system can better leverage multi-user or inter-cell interference and improve the reception quality, leading to enhanced spectral efficiency, system capacity and energy efficiency [1]. A D-MIMO system has potentials to significantly improve the performance of current and future wireless communication networks, such as coordinated multi-point (CoMP) [2] in LTE-advanced cellular networks, mobile over Fiber to the x (FTTx) networks based on a multi-source multi-relay model [3], and local area networks with WiFi coordination. Moreover, D-MIMO provides an platform for incorporating new technologies such as cloud radio access [4], [5], [6], [7] and cloud storage [8].

A. State of the Art

The design of D-MIMO systems under rate-constrained backhaul has received extensive interests [9], [10]. An approach based on quantize re-map and forward (QMF) was proposed in

[11]. This approach was extended in [12] to a general network, referred to as noisy network coding. In the QMF approach, the base stations perform vector quantization of their received signals at some rate satisfying the backhaul rate-constraint. It is shown that QMF or noisy network coding achieves a rate region within a bounded gap from the cut-set outer bound. Yet, from an implementation viewpoint, QMF and noisy network coding may not be practical due to the very high complexity of vector quantization.

Recently, Nazer and Gastpar proposed a compute-and-forward (CF) strategy [10] that can be employed in the uplink D-MIMO system. This strategy is of particular interest in a practical scenario where the channel state information (CSI) is not available at the transmitters (mobile users), and the base stations cannot share the CSI. In CF, each base station attempts to decode and forward some finite-field linear combinations of a subset of all users' messages, by exploiting the structural property of conceptual nested lattice codes. The information theoretic performance of CF was examined for the uplink D-MIMO system in the case of the overly simplistic Wyner model [13], rather than a practical fading channel model. Recently, we proposed a simplified CF based scheme, referred to as linear physical-layer network coding, aiming at the practical implementations of CF in fading channels [14]. Two major impairments to CF in an uplink D-MIMO system are the *non-integer penalty* and the high probability of *rank deficiency*. The non-integer penalty is due to the severe quantization noise incurred by the mismatch between the channel observations (with non-integer channel coefficients) and the integer superposition of lattice codewords [10]. The high probability of *rank deficiency* is due to that the CF coefficients are distributively determined by the relays. In a realistic wireless channel with fading and shadowing, a high probability of full rank CF coefficients is not guaranteed, in contrast to the overly simplified Wyner model case [13]. These two impairments are particularly significant when the transmitters lack the CSI.

Another approach for the uplink D-MIMO system is quantize-and-forward (QF) (or compress-and-forward), where the base stations are referred to as oblivious base stations [1], [15], [16]. In QF, each base station performs scalar quantization of the signal, where the number of quantized bits are determined by the rate-constraint of the backhaul link. The CU performs multi-user decoding based on the quantized signals, while combating the quantization noise. In a practical scenario where the backhaul links have non-trivial rate-constraints, the performance of QF may be poor due to the severe quantization noise.

From the literature, the ultimate performance of an uplink D-

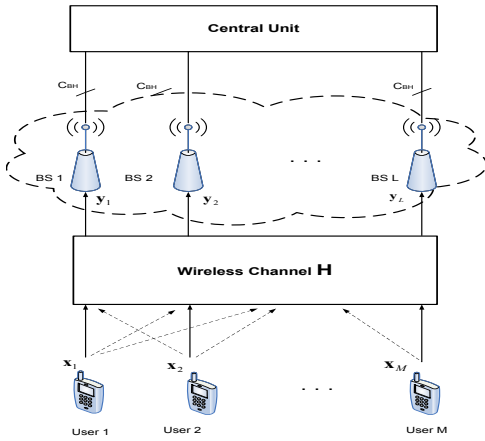


Fig. 1. Block Diagram of an M-user L-BS uplink distributed MIMO system with rate-constrained backhaul links.

MIMO system remains unknown. For a practical scenario where the mobile users do not have the CSI, and the base stations do not share their CSI, the design of a high-performance D-MIMO scheme becomes even more challenging. It remains difficult to quantify the benefits of D-MIMO and evaluate its performance gap to the full MIMO system.

B. Contributions

In this paper, we aim to design a high-performance uplink D-MIMO system with moderate-to-stringent backhaul rate-constraint, e.g., the backhaul rate-constraints and the data rates are of the same order of magnitude. We focus on the scenario where the transmitters do not have CSI and the base stations do not share the CSI. The main contributions of this paper are summarized as follows: 1) We propose a new linear network coding based protocol. In the proposed protocol, each of the L base stations attempts to forward N pre-assigned linear network coding functions of the M users' messages, satisfying the rate-constraint of the backhaul link. The central unit recovers all M users' messages by solving the linear functions, provide that the linear functions are full rank. The coefficients of the pre-assigned linear functions are carefully designed to minimize the probability of rank deficiency, leading to improved performance over CF. 2) We analyze the information rate outage probability of the proposed scheme. We analytically show that for a system of $M \geq L$, as long as the backhaul rate-constraint is greater than the individual data rate of one user, the outage probability of the proposed scheme decays like $\frac{1}{\text{SNR}^L}$ at a high SNR. As the backhaul rate-constraint approaches M times of the data rate of one user, the performance of the proposed scheme is $L/10\log_{10}L!$ dB away from that of the full MIMO scenario at a high SNR. 3) We also develop a structured way to efficiently construct the pre-assigned linear functions that yields the optimized outage probability of the proposed scheme. Numerical results show that the proposed scheme has significantly improved outage probability and frame error rate performance over existing CF and QF schemes, e.g., more than 3 dB improvement in a $M = L = 2$ setup.

II. SYSTEM MODEL

Consider an uplink D-MIMO system with M single-antenna users, L single-antenna base stations (BSs) and a common central

unit (CU). The BSs are connected to the CU via L independent rate-constrained backhaul (BH) link. We consider that the BSs are not mutually connected, and they do not share data or channel state information (CSI). We consider an *open-loop* uplink system, i.e., there is no feedback from CU to BSs, and from BSs to the users. Thus, the users do not have the CSI.

The block diagram of the system is depicted in Fig. 1. The system consists of two layers: *air-interface* layer and *backhaul* (BH) layer. At the air-interface layer, M users transmit their messages towards the BSs using the same frequency band and time-slot. At BH layer, each BS processes its received signal and forwards it to the CU. Upon collecting the signals from the L BSs, the CU recovers all M users' messages.

A. Air-Interface Layer

User m has a length- k_m message sequence denoted by \mathbf{b}_m , $m \in \{1, 2, \dots, M\}$. We consider that each user's message is drawn independently and uniformly over the binary field, i.e., $\mathbf{b}_m \in \{0, 1\}^{k_m}$, $m \in \{1, 2, \dots, M\}$. The m th user's encoder maps \mathbf{b}_m into a length- n coded-and-modulated complex-valued symbol sequence $\mathbf{x}_m = [x_m[1], \dots, x_m[n]]$. Denote the encoding functions by

$$\mathbf{x}_m = \varepsilon_m(\mathbf{b}_m), m \in \{1, 2, \dots, M\}. \quad (1)$$

The information rate of user m is $R_m = \frac{k_m}{n}$ bits/channel-use. The power constraints are

$$E \left(\sum_{t=1}^n |x_m[t]|^2 \right) \leq P, m \in \{1, 2, \dots, M\}, \quad (2)$$

where we consider that all users have the same power constraint P . Our result in this paper can be easily extended to cases where the users have different power constraints.

We consider a block fading channel, i.e. the channel coefficients remain constant over a block of n symbols and vary over blocks. In one block, the signal received by the l th BS is

$$\mathbf{y}_l = \sum_{m=1}^M h_{lm} \mathbf{x}_m + \mathbf{z}_l, l \in \{1, 2, \dots, L\}, \quad (3)$$

where h_{lm} denotes the complex-valued flat-fading channel coefficient from the m th user to the l th BS, and \mathbf{z}_l is a length- n complex-valued additive white Gaussian noise (AWGN) sequence of zero mean and variance $\frac{N_0}{2}$ per real-dimension. The signal-to-noise ratio (SNR) is defined as $\rho \triangleq \frac{P}{N_0}$. In this paper, we consider perfect synchronization among users at the relay. We note that our new scheme to be proposed in Section IV also applies to the asynchronous scenario. As we will see later in Section IV, each relay will carry out joint multi-user decoding w.r.t. all users' messages. In the case with asynchronous access, joint multi-user decoding such as iterative soft cancellation and decoding can be carried out, as in the asynchronous code-division multiple-access (CDMA) or interleave-division multiple-access (IDMA) [17]. For our proposed scheme, the performance of the multi-user decoding in the asynchronous case will be the same as that in the synchronous case, as the interference structure and per-user signal-to-interference plus noise ratio (SINR) are the same for the synchronous and asynchronous cases.

The l th BS processes its received signal y_l . The reconstructed signal is generally written as

$$f_l(y_l), l \in \{1, 2, \dots, L\}. \quad (4)$$

Remark 1 (Degree of channel state information): In this paper, we consider that the channel state information (CSI) is not available at the transmitters of mobile users. We assume that the channel coefficients h_{lm} , $m \in \{1, 2, \dots, M\}$, are perfectly known by the l th BS but not known by the l' -th BS for all $l' \neq l$.

B. Backhaul Layer

Recall that the l th BS is connected to the CU via the l th rate-constrained BH link, $l \in \{1, 2, \dots, L\}$. For more detailed discussion on the rate-constrained BH link, please refer to [9], [18]. In this paper, we denote by \bar{C}_l^{BH} the maximum rate per-block that can be supported by the l th BH link.

The signal $f_l(y_l)$ generated by the l th BS is forwarded to the CU via the l th BH link. We refer to $f_l(y_l)$ as the BH function of the l th BS. Due to the rate-constraint of the BH link, the BH function must satisfy

$$H(f_l(y_l)) \leq \bar{C}_l^{BH}, l \in \{1, 2, \dots, L\} \quad (5)$$

where $H(\cdot)$ denotes the entropy function. We assume error free transmission of the reconstructed signal (4) over the BH links if the rate constraint (5) is met.

Denote by $\mathbf{u}_l = f_l(y_l)$, $l \in \{1, 2, \dots, L\}$. Upon collecting $\mathbf{u}_1, \dots, \mathbf{u}_L$, the CU attempts to recover all M users' messages $\mathbf{b}_1, \dots, \mathbf{b}_M$. This decoding process is represented by

$$[\hat{\mathbf{b}}_1, \dots, \hat{\mathbf{b}}_M] = \text{dec}(\mathbf{u}_1, \dots, \mathbf{u}_L). \quad (6)$$

An decoding error is declared if $\hat{\mathbf{b}}_m \neq \mathbf{b}_m$ for any $m \in \{1, \dots, M\}$. This finishes one block of transmission.

Remark 2: In this paper, we focus on that all BH links have the same rate-constraint per-block, i.e., $\bar{C}_1^{BH} = \dots = \bar{C}_L^{BH} = \bar{C}^{BH}$. Our result can be generalized to cases where the BH links have different rate-constraints.

III. PROBLEM STATEMENT

A. Definition of Achievable Symmetric-Rate

In an open-loop uplink system considered in this paper, adaptive rate selection cannot be performed at the users, since CSI is not available at the users and there is no feedback from BSs to users. We focus on the symmetric-rate scenario where all users have the same per-user data rate, i.e., $R_1 = \dots = R_M = R$, or equivalently $k_1 = \dots = k_M = k$. We will discuss the non-symmetric rate scenarios later in Section IV. For notation convenience, the normalized BH rate-constraint C^{BH} is defined as

$$C^{BH} = \frac{\bar{C}^{BH}}{k}. \quad (7)$$

For simplicity, C^{BH} will be referred to as the BH rate-constraint hereafter.

Let \mathbf{H} be a matrix which consists of all the channel coefficients, i.e., the $\{l, m\}$ th entry of \mathbf{H} is given by

$$\mathbf{H}_{\{l, m\}} = h_{lm}. \quad (8)$$

Each row of \mathbf{H} consists of the channel coefficients $\mathbf{h}_l^T = [h_{l1}, \dots, h_{lM}]$ w.r.t. the l th BS. We consider that the entries of \mathbf{H} are i.i.d. and follow Rayleigh distribution.

For a deterministic \mathbf{H} and a given BH rate-constraint C^{BH} , a (R, n, P_e) code consists of the following:

1. M encoding functions $\varepsilon_1(\cdot), \dots, \varepsilon_M(\cdot)$ of rate R .
2. L BH functions $f_1(\cdot), \dots, f_L(\cdot)$ that satisfy the BH rate-constraint C^{BH} (as specified in (5)).
3. A decoding function $\text{dec}(\cdot)$ at the CU (as specified in (6)).
4. An error probability P_e , given by

$$P_e = \Pr\{\text{dec}(\mathbf{u}_1, \dots, \mathbf{u}_L) \neq [\mathbf{b}_1, \dots, \mathbf{b}_M]\}. \quad (9)$$

A symmetric rate R is said to be *achievable* if, for any $\epsilon > 0$, there exists a (R, n, P_e) code such that $P_e \leq \epsilon$ as $n \rightarrow \infty$. We emphasize that R is dependent on \mathbf{H} and C^{BH} .

B. Performance Metric: Information Rate Outage Probability

Let the target symmetric rate of each user be denoted by R_0 . For the block fading channel considered in this paper, R_0 may or may not be achievable, depending on the realization of \mathbf{H} as well as C^{BH} . In a channel realization that the achievable symmetric rate R of a scheme is smaller than the target information rate R_0 , an outage event is declared. The information rate *outage probability* (OP) is defined as

$$P_{\text{out}}(R_0, C^{BH}) \triangleq \Pr\{R < R_0\}. \quad (10)$$

We will use $P_{\text{out}}(R_0, C^{BH})$ as the performance metric of the uplink D-MIMO system. In the literature, the OP is often employed as the performance metric of a system without CSI at the transmitter side. The OP provides a lower bound on the frame error rate of a practical coding and modulation scheme.

Following the convention, a scheme is said to achieve a diversity order of d as $\rho \rightarrow \infty$ if

$$P_{\text{out}}(R_0, C^{BH}) \propto \rho^{-d}. \quad (11)$$

Here $a \propto b$ denotes that a is proportional to b . Note that achieving the maximal diversity order is necessary for achieving the optimal OP at a high SNR.

C. Stringent-to-moderate BH rate-constraint

The OP performance of the D-MIMO system is dependent on 1) the target data rate R_0 and 2) the BH rate-constraint C^{BH} . In this paper, we consider the case with "stringent-to-moderate" BH rate-constraint, where R_0 and C^{BH} have the *same order of magnitude*. In particular, we will focus on $C^{BH} = NR_0$, where $N \in \{1, \dots, M\}$. For $C^{BH} \gg R_0$, e.g., $kC^{BH} > \max_{l \in \{1, \dots, L\}} H(y_l)^1$, the BH capacity is sufficient for the BSs to forward their entire received signals y_1, y_2, \dots, y_L to the CU via the BH links, and the CU can be viewed as a receiver with L co-located antennas. This setup is called the *full MIMO* case in this paper². For $C^{BH} \ll R_0$, the performance of the system will be bottle-necked by C^{BH} which is not the interest in this paper.

¹Note that $H(y_l) \geq H(x_l) = kR_0$. Here, kR_0 is the total information rate per-block.

²More precisely, this case is referred to as the MIMO multiple-access setup. For simplicity, we just call it the full MIMO setup.

IV. PROPOSED NEW LINEAR NETWORK CODING BASED PROTOCOL

In this section, we propose a new linear network coding based protocol for the D-MIMO system.

A. Air-interface Layer

Let the encoding functions $\varepsilon_1(\cdot), \dots, \varepsilon_M(\cdot)$ be globally known to all BSs. In the air-interface layer, the M users transmit simultaneously. Given the received signals in (3), each BS attempts to decode $\mathbf{b}_1, \dots, \mathbf{b}_M$, e.g., using multi-user decoding³. In general, each BS may *successfully* (or correctly) *decode* a subset of the M users' message, depending on R_0 and the fading channel realization. We note that since the BSs are not mutually connected, joint processing among the BSs cannot be performed.

B. Backhaul Layer

In this part, we will present the proposed protocol for the system with $C^{BH} = R_0$ (or $N = 1$), i.e., the normalized BH capacity is the same as the target information rate. After that, we will extend our proposed scheme to the cases with $C^{BH} = NR_0, N = 2, \dots, M$.

1) *Pre-assigned LNC Function*: Denote by \mathbf{w}_m the q -ary conversion of \mathbf{b}_m , $m \in \{1, 2, \dots, M\}$. The q -ary conversion of a binary vector $\mathbf{b} = [b_0, \dots]^T$ is invertible and given by $\mathbf{w} = [w_0, \dots]^T : \sum_{i=0,1,\dots} w_i q^i = \sum_{j=0,1,\dots} b_j 2^j$, where the entries

of \mathbf{w} belong to $\{0, \dots, q-1\}$.

In the proposed protocol, a *pre-assigned linear combination* of the M users' messages is given by

$$\alpha_{l,1} \otimes \mathbf{w}_1 \oplus \alpha_{l,2} \otimes \mathbf{w}_2 \oplus \dots \oplus \alpha_{l,M} \otimes \mathbf{w}_M = \bigoplus_{m=1}^M \alpha_{l,m} \otimes \mathbf{w}_m \quad (12)$$

where $\alpha_{l,1}, \dots, \alpha_{l,M} \in F_q$, $l \in \{1, 2, \dots, L\}$, and F_q denotes the Galois field of size q . Here, " \otimes " represents the finite field multiplication of a vector by a scalar, and " \oplus " represents the finite field addition between two vectors. In particular, when q is a prime, " \otimes " and " \oplus " respectively denote the modulo- q multiplication and addition. We refer to (12) as a pre-assigned LNC function and $\alpha_{l,1}, \dots, \alpha_{l,M}$ as the *pre-assigned LNC coefficients* for the l th BS. Let the pre-assigned LNC coefficients be denoted by a vector:

$$\boldsymbol{\alpha}_l = [\alpha_{l,1}, \dots, \alpha_{l,M}]^T, l \in \{1, 2, \dots, L\},$$

which is known to BS l . Note that BS l does not have to know $\boldsymbol{\alpha}_{l'}$ for $l' \neq l$. The pre-assigned LNC coefficients for all the L BSs are denoted by the following matrix

$$\mathbf{A} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_L]^T \quad (13)$$

which is referred to as a *pre-assigned LNC coefficient matrix*.

³In this section, we do not consider CF or physical-layer network coding methods, which suffer from severe rank deficiency problem. Later, we will show that our proposed scheme outperforms the CF in terms of outage probability in Rayleigh fading channel.

2) *Active LNC Function for $C^{BH} = R_0$ ($N = 1$)*: In the ideal case where the l th BS can successfully decode all M users' messages, the exact pre-assigned LNC function in (12) can be constructed indeed. In general, for each fading channel realization, the l th BS may successfully decode only a subset of $\mathbf{w}_1, \dots, \mathbf{w}_M$. Let this subset be denoted by \mathcal{S}_l . Then, in the proposed scheme, BS l forwards the following LNC function⁴:

$$\mathbf{u}_l = \left(\bigoplus_{m=1, m \in \mathcal{S}_l}^M \beta_{l,m} \otimes \mathbf{w}_m \right) \oplus \left(\bigoplus_{m=1, m \in \mathcal{S}_l^c}^M \beta_{l,m} \otimes \mathbf{w}_m \right) \quad (14)$$

where

$$\beta_{l,m} = \begin{cases} \alpha_{l,m} & \text{if } m \in \mathcal{S}_l \\ 0 & \text{if } m \in \mathcal{S}_l^c \end{cases} \quad (15)$$

and \mathcal{S}_l^c denotes the complementary set of \mathcal{S}_l . We refer to \mathbf{u}_l as an *active LNC function*, and $\beta_{l,1}, \dots, \beta_{l,M}$ as the *active LNC coefficients* for the l th BS. Note that \mathbf{u}_l has the same rate as the target symmetric data rate. Thus, $H(\mathbf{u}_l) = kR_0, l \in \{1, 2, \dots, L\}$. For $N = 1$, it is clear that (14) satisfies the BH rate-constraint in (5).

Remark 3: In our proposed scheme, BS l is notified with the pre-assigned LNC coefficients in an off-line manner. These coefficients are carefully designed and notified to the users before data transmission. During the data transmission, BS l 's operation will always be based on this pre-assigned $\alpha_{l,1}, \dots, \alpha_{l,M}$, for all fading realizations. Yet, for different fading channel realizations, BS l can successfully decode different subsets of all users' message, which yields different active LNC coefficients $\beta_{l,1}, \dots, \beta_{l,M}$.

Remark 4: In each block, BS l will first advise the CU its active LNC coefficients. If the pre-assigned matrix \mathbf{A} is known to the CU, by forwarding only the indices of zeros element in $[\beta_{l,1}, \dots, \beta_{l,M}]$, the overhead is at most M bits per-block per BS. Such an overhead is negligible if M is much smaller than the block length k .

3) *Active LNC Function for $C^{BH} = N \cdot R_0, N \geq 2$* : In this case, each BS can forward up to N LNC functions of the users' messages while satisfying the BH rate-constraint. We propose to use the following LNC functions at BS l :

$$\begin{aligned} \mathbf{u}_l^{(1)} &= \bigoplus_{m=1, m \in \mathcal{S}_l}^M \beta_{l,m}^{(1)} \otimes \mathbf{w}_m \oplus \bigoplus_{m=1, m \in \mathcal{S}_l^c}^M \beta_{l,m}^{(1)} \otimes \mathbf{w}_m \\ &\vdots \\ \mathbf{u}_l^{(N)} &= \bigoplus_{m=1, m \in \mathcal{S}_l}^M \beta_{l,m}^{(N)} \otimes \mathbf{w}_m \oplus \bigoplus_{m=1, m \in \mathcal{S}_l^c}^M \beta_{l,m}^{(N)} \otimes \mathbf{w}_m, \end{aligned} \quad (16)$$

where the active LNC coefficients are

$$\beta_{l,m}^{(i)} = \begin{cases} \alpha_{\text{mod}(l+i-2, L)+1, m} & \text{if } m \in \mathcal{S}_l \\ 0 & \text{if } m \in \mathcal{S}_l^c \end{cases}, i = 1, \dots, N. \quad (17)$$

Note that the N LNC coefficient vectors of any BS can be obtained from a single \mathbf{A} . For each BS, its N active LNC coefficient vectors are some of N consecutive rows of \mathbf{A} in a cyclic-shift manner. Denote the N active LNC coefficient vectors

⁴Note that the second term in (14) is equal to zero and could be removed from the expression. Here, the second term in (14) is explicitly shown in order to highlight that the active network coding coefficients for those users' messages that are not successfully decoded are set to zero.

of BS l by

$$\beta_l^{(1)} = [\beta_{l,1}^{(1)}, \dots, \beta_{l,M}^{(1)}]^T, \dots, \beta_l^{(N)} = [\beta_{l,1}^{(N)}, \dots, \beta_{l,M}^{(N)}]^T. \quad (18)$$

The active LNC coefficient vectors for all the L BSs are denoted by

$$\mathbf{B} = [\beta_1^{(1)}, \dots, \beta_1^{(N)}, \beta_2^{(1)}, \dots, \beta_L^{(N)}]^T \quad (19)$$

which is referred to as the *active LNC coefficient matrix*. We refer to each zero element in \mathbf{B} as an *erasure*.

In each block, given the received signals in (3), each BS attempts to decode $\mathbf{b}_1, \dots, \mathbf{b}_M$ using joint multi-user decoding, such as joint typicality decoding [16] or iterative soft interference cancellation and decoding [22]. In general, each BS may successfully decode a subset of the M users' message, and the zero terms in the active LNC coefficients β_l are determined. Then, BS l will advise the CU its active LNC coefficients β_l .

C. Processing at the Central Unit

Given the $N \cdot L$ functions $\mathbf{u}_1^{(1)}, \dots, \mathbf{u}_1^{(N)}, \mathbf{u}_2^{(1)}, \dots, \mathbf{u}_L^{(N)}$ and \mathbf{B} , the CU attempts to recover all users' messages $\mathbf{w}_1, \dots, \mathbf{w}_M$ by solving the following linear equations in F_q :

$$\mathbf{B} \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_M \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^{(1)} \\ \vdots \\ \mathbf{u}_L^{(N)} \end{bmatrix}. \quad (20)$$

We note that the CU can successfully recover all $\mathbf{w}_1, \dots, \mathbf{w}_M$ if and only if \mathbf{B} has a full rank of M in F_q . An error happens if \mathbf{B} contains some patterns of erasures (zero elements) that lead to rank deficiency of \mathbf{B} .

Example 1: Consider $M = L = 2$ and $N = 1$. Let $q = 3$ and $\mathbf{A} = [1 \ 1; 1 \ 2]$. Here, \mathbf{A} has full rank in F_q where F_q is an integer field of $\{0, 1, 2\}$. Suppose that BS 1 can successfully decode both users' messages, the active LNC coefficients will be $[\beta_{1,1}, \beta_{1,2}] = [1, 1]$. The LNC function delivered to the CU will be $\mathbf{w}_1 \oplus \mathbf{w}_2$. Suppose that BS 2 cannot successfully decode the first user's message but can successfully decode the second user's message, the active LNC coefficients will be $[\beta_{2,1}, \beta_{2,2}] = [0, 2]$. Then, the LNC function delivered to the CU will be \mathbf{w}_2 . The active LNC coefficient matrix is $\mathbf{B} = [1 \ 1; 0 \ 2]$ which has full rank and the CU can recover both \mathbf{w}_1 and \mathbf{w}_2 . Here a single erasure will not affect the recovery of the two users' messages.

D. Comments on the Proposed Scheme

The choice of the pre-assigned LNC matrix \mathbf{A} is critical. Intuitively, a well-designed \mathbf{A} should be able to accommodate as many erasures (i.e., the zero elements in \mathbf{B}) as possible while still having a full rank in the finite field. This will lead to improved OP performance relative to that with an arbitrarily chosen \mathbf{A} , as we will show in the next section. The details on the design of \mathbf{A} will be presented in Section VI.

In the proposed LNC based protocol, the BSs exhibit the feature of "cloud base stations". The users are not aware of which BSs are actually serving them, while each BS is not aware of the operations of other BSs. This leads to several advantages in practice, e.g., there is no need to design soft handover and the cell planning can be simplified. For LTE cellular networks, in

particular, there is no need to change the system architecture in order to employ the proposed protocol. The proposed protocol also shares some features of the cloud storage using network coding [8]. The erasures are due to that a BS cannot successfully decode some of the users' messages, which are largely decided by the random realization of fading channels.

The proposed scheme differs from compute-and-forward in that the LNC coefficients used by the BSs in our scheme are pre-assigned (optimized) in advance. By carefully designing the pre-assigned LNC matrix \mathbf{A} (see next two sections), the probability of rank deficiency at the CU can be minimized. In contrast, in compute-and-forward, each BS selects the best N LNC coefficients, but this frequently leads to rank deficiency at the CU. We note that the method proposed in [19] for designing the CF coefficient required feedback from the CU to the BSs, which cannot be implemented in the open loop uplink system considered in this paper.

We would like to note that our proposed scheme can be extended to non-symmetric rate scenarios. Specifically, at the relay we can partition the message sequence of a high-rate user into several segments, where the length of each segment is identical to the length of the message sequence of a low-rate user. Next, the multiple segments of a high-rate user are associated with multiple LNC coefficient vectors. Then, the LNC operations at the relay and the decoding at the destination can be straightforwardly applied.

We note that our proposed scheme applies to the case where the number of users changes. Specifically, when a user leaves the network, its pre-assigned LNC coefficients will be released. When a user joins the network, it will be allocated with some pre-assigned LNC coefficients that meet the LNC design requirement (Cond. 1 on Page 17). Meanwhile, the allocation of pre-assigned LNC coefficients for other users will remain unchanged. In this way, only the pre-assigned LNC coefficients for the users that join or leave the network will be computed. The resultant overhead is not significant when most users keep being active in the network.

V. OUTAGE BEHAVIOR IN RAYLEIGH FADING CHANNELS

In this section, we study the OP behavior of the proposed LNC based protocol for the uplink D-MIMO system in Rayleigh fading channels. Throughout this section, we will focus on the case of ⁵ $M \leq L$. Our scheme can be generalized to the case of $L < M$ while $N > 1$ is required.

A. Main Asymptotic Results

In this section, we focus on the outage probability for the case that all channels are i.i.d. and follow Rayleigh distribution, i.e., $h_{l,m} \sim \mathcal{CN}(0, 1)$. The analysis can be extended to the case with different fading channel distributions. In general, deriving the exact OP for a specific SNR is very difficult. Here, we present some main results on the OP of our proposed protocol in the high SNR regime. Our results will be useful for evaluating the outage behavior of the proposed scheme, as well as its performance gap to the full MIMO setup.

⁵In a practical system e.g. the 4G cellular system, when a single-antenna BS need to serve multiple users, the BS will allocate different frequency bands for them, e.g. using OFDM). This is the main reason why we do not consider the case of $L < M$.

Theorem 1: For $M \leq L$, as long as $C^{BH} \geq R_0$ (or $N \geq 1$), there exists a pre-assigned LNC matrix \mathbf{A} such that, as $\rho \rightarrow \infty$, the OP of the proposed scheme satisfies

$$P_{out}(R_0, C^{BH}) \propto \frac{1}{\rho^L}. \quad (21)$$

It is noteworthy that achieving the maximal diversity order is a necessary condition for a scheme to be OP optimal at a relatively high SNR. It can be shown that in a conventional scheme where BS l only recovers and forwards \mathbf{w}_j , $j \in \{1, \dots, M\}$, the OP is proportional to $\frac{1}{\rho}$. This suggests that, as long as $C^{BH} \geq R_0$, our proposed scheme outperforms the conventional scheme at a sufficiently high SNR.

Theorem 2: For $M \leq L$, as long as $C^{BH} \geq MR_0$ (or $N \geq M$), the OP of the proposed scheme becomes

$$P_{out}(R_0, C^{BH}) \xrightarrow{\rho \rightarrow \infty} M \left(\frac{2^{R_0} - 1}{\rho} \right)^L. \quad (22)$$

Using Theorem 2, we can evaluate the performance gap to the full MIMO setup as follows.

Corollary 1: For $M \leq L$, it can be shown that the OP of the full MIMO setup is given by

$$P_{out}^{MIMO}(R_0) \xrightarrow{\rho \rightarrow \infty} \frac{M}{L!} \left(\frac{2^{R_0} - 1}{\rho} \right)^L. \quad (23)$$

Comparing (22) and (23), our proposed scheme is $\frac{10}{L} \log_{10}(L!)$ dB away from the performance of the full MIMO setup as long as $C^{BH} \geq MR_0$ at a high SNR. The gap is 1.505 dB, 2.59dB and 3.45 dB for $L = 2, 3, 4$ respectively. In Section VII, we will show by numerical results that the above asymptotic results hold when the SNR is greater than about 15 dB (or 25 dB) for D-MIMO systems where $R_0 = 1$ (or $R_0 = 2$).

We note that a well-designed matrix \mathbf{A} is necessary to achieve the above results.

B. Outage Probability Analysis (Proof of Main Results)

We now analyze the OP of the proposed scheme and prove the main results.

1) Preliminaries:

Definition 1 (Completely decodable region): For a given target symmetric rate R_0 , the achievable channel parameter region $\mathcal{A}(\mathbf{w}_1, \dots, \mathbf{w}_M)$ is defined as the collection of all (h_{l1}, \dots, h_{lM}) such that all of $\mathbf{w}_1, \dots, \mathbf{w}_M$ can be successfully decoded by BS l . We refer to such a region as the *completely decodable (CD) region*.

Lemma 1: For a given R_0 , at BS l , $\mathcal{A}(\mathbf{w}_1, \dots, \mathbf{w}_M)$ is given by the collection of all (h_{l1}, \dots, h_{lM}) s.t.

$$\begin{aligned} |h_{lm}|^2 &\geq \frac{2^{R_0} - 1}{\rho}, \forall m \in \{1, \dots, M\} \\ |h_{lm}|^2 + |h_{lm'}|^2 &\geq \frac{2^{2R_0} - 1}{\rho}, \forall m, m' \in \{1, \dots, M\}, m \neq m' \\ &\vdots \\ \sum_{m=1}^M |h_{lm}|^2 &\geq \frac{2^{MR_0} - 1}{\rho}. \end{aligned} \quad (24)$$

Proof: This follows from the well-known capacity region of the M -user multiple-access channel (MAC). The proof was

based on superposition encoding and joint typicality decoding [20]. ■

An example of the CD region is illustrated in Fig. 2 for the $M = 2$ case.

Remark 5: From the literature, it is known that there exist spatially coupled codes that exhibit vanishing error probability (as $n \rightarrow \infty$) as long as (h_{l1}, \dots, h_{lM}) belong to the CD region specified above [21].

In conventional M -user MAC, the receiver is required to decode all users' messages. In contrast, in the uplink D-MIMO system, it is not necessary that every BS must successfully decode all users' messages. As long as the L BSs collectively forward a number of linear message functions that have rank M , the CU can recover all users' messages. The following definition and lemma will be used to reflect this feature.

Definition 2 (Partially decodable region): For a given set of user-message index $\mathcal{S}_l \subseteq \{\mathbf{w}_1, \dots, \mathbf{w}_M\}$ w.r.t. BS l , the partially decodable (PD) region, denoted by $\mathcal{A}(\mathcal{S}_l)$, is defined as the collection of all channel parameters (h_{l1}, \dots, h_{lM}) such that the messages $\mathbf{w}_m, \mathbf{w}_m \in \mathcal{S}_l$, can be successfully decoded while $\mathbf{w}_m, \mathbf{w}_m \in \mathcal{S}_l^c$, cannot be successfully decoded.

Remark 6: When \mathcal{S}_l collects all M users' message indices, i.e., $\mathcal{S}_l = \{\mathbf{w}_1, \dots, \mathbf{w}_M\}$, the PD region $\mathcal{A}(\mathcal{S}_l)$ becomes the CD region. When \mathcal{S}_l is an empty set \emptyset , $\mathcal{A}(\emptyset)$ specifies the region that no user's messages can be successfully decoded. We refer to $\mathcal{A}(\emptyset)$ as a *non-decodable region*. In general, there are 2^M non-overlapping PD regions for BS l , including the CD region and the non-decodable region.

Lemma 2: For a given R_0 , the PD region $\mathcal{A}(\mathcal{S}_l)$ is given by the collection of all (h_{l1}, \dots, h_{lM}) such that

$$\begin{aligned} |h_{lm}|^2 &\geq \frac{2^{R_0} - 1}{\rho} \left(\rho \sum_{\bar{m} \in \mathcal{S}_l^c} |h_{l\bar{m}}|^2 + 1 \right), \forall m \in \mathcal{S}_l \\ |h_{lm}|^2 + |h_{lm'}|^2 &\geq \frac{2^{2R_0} - 1}{\rho} \left(\rho \sum_{\bar{m} \in \mathcal{S}_l^c} |h_{l\bar{m}}|^2 + 1 \right), \forall m, m' \in \mathcal{S}_l, m' \neq m \\ &\vdots \\ \sum_{m \in \mathcal{S}_l} |h_{lm}|^2 &\geq \frac{2^{|\mathcal{S}_l|R_0} - 1}{\rho}, \\ |h_{lm}|^2 &< \frac{2^{R_0} - 1}{2^{2R_0}\rho} \left(\rho \sum_{\bar{m} \in \mathcal{S}_l^c} |h_{l\bar{m}}|^2 + 1 \right), \forall m \in \mathcal{S}_l^c \\ |h_{lm}|^2 + |h_{lm'}|^2 &< \frac{2^{2R_0} - 1}{2^{2R_0}\rho} \left(\rho \sum_{\bar{m} \in \mathcal{S}_l^c} |h_{l\bar{m}}|^2 + 1 \right), \forall m, m' \in \mathcal{S}_l^c, m' \neq m \\ &\vdots \\ \sum_{\bar{m} \in \mathcal{S}_l^c} |h_{l\bar{m}}|^2 &< \frac{2^{|\mathcal{S}_l^c|R_0} - 1}{\rho}. \end{aligned} \quad (25)$$

Proof: It can be shown that if the inequalities in (25) are satisfied, for all users belonging to \mathcal{S}_l , their messages can be successfully decoded while treating all the rest users' signals as noise. Also, it can be shown that if the inequalities in (26) are satisfied, for all users belonging to \mathcal{S}_l^c , none of their messages can be successfully decoded. Combining these two parts, we obtain the PD region for a given \mathcal{S}_l . ■

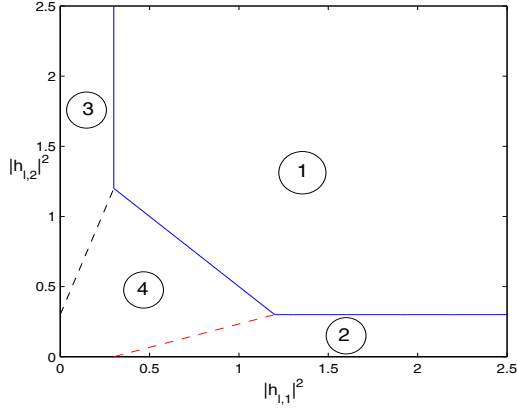


Fig. 2. Illustration of CD and PD regions of a $M = L = 2$ D-MIMO system. The SNR is 10 dB and $R_0 = 2$. Area 1 depicts the CD region $\mathcal{A}(S_l = \{\mathbf{w}_1, \mathbf{w}_2\})$. Areas 2 and 3 respectively depict the PD regions $\mathcal{A}(S_l = \mathbf{w}_1)$ (or $\mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2)$) and $\mathcal{A}(S_l = \mathbf{w}_2)$ (or $\mathcal{A}(\tilde{\mathbf{w}}_1, \mathbf{w}_2)$). Area 4 depicts the non-decodable region $\mathcal{A}(S_l = \emptyset)$ (or $\mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2)$).

Example 2: Consider $M = 2$. As long as $|h_{l1}|^2 > \frac{2^{R_0}-1}{\rho}(\rho|h_{l2}|^2 + 1)$, BS l can successfully decode \mathbf{w}_1 by treating user 2's signal as noise, and the interference from user 1 to user 2 can be removed. In addition, if $|h_{l2}|^2 < \frac{2^{R_0}-1}{\rho}$, BS l cannot successfully decode \mathbf{w}_2 . Lemma 2 generalize this to the M user case.

In Fig. 2, we depict the CD region $\mathcal{A}(S_l = \{\mathbf{w}_1, \mathbf{w}_2\})$ and the PD regions $\mathcal{A}(S_l = \mathbf{w}_1)$ and $\mathcal{A}(S_l = \mathbf{w}_2)$ w.r.t. BS l , for a $M = 2$ system. Here, the SNR is 10 dB and the symmetric rate is $R_0 = 2$. The horizontal axis denotes $|h_{l1}|^2$ and the vertical axis denotes $|h_{l2}|^2$. The whole area is divided into four disjoint areas. In particular, area 1 depicts the CD region $\mathcal{A}(S_l = \{\mathbf{w}_1, \mathbf{w}_2\})$ which is specified in (24). For a channel realization where $|h_{l1}|^2$ and $|h_{l2}|^2$ fall into this region, both \mathbf{w}_1 and \mathbf{w}_2 can be successfully decoded. Area 2 depicts the PD region $\mathcal{A}(S_l = \mathbf{w}_1)$. For a channel realization where $|h_{l1}|^2$ and $|h_{l2}|^2$ falls into this area, \mathbf{w}_1 can be successfully decoded but \mathbf{w}_2 cannot. Similarly, area 3 depicts $\mathcal{A}(S_l = \mathbf{w}_2)$. Area 4 depicts the non-decodable region $\mathcal{A}(S_l = \emptyset)$ where neither \mathbf{w}_1 nor \mathbf{w}_2 can be successfully decoded.

2) *Outage Events at the CU:* Recall that in the proposed LNC based protocol, each BS forwards N LNC functions to the CU via its BH link. Then, the CU attempts to recover all M users' messages based on the $L \cdot N$ LNC functions. It is clear that the positions of erasures (zero elements) in the active LNC coefficients are determined by the indices of the users that cannot be successfully decoded at BS l . To be specific, for BS l , the active LNC coefficient vector $\beta_l = [\beta_{l,1}, \dots, \beta_{l,M}]^T$ is determined by the PD region that the channel realization \mathbf{h}_l falls into. By considering all L BSs, the outage event at the CU is given below.

Lemma 3: An outage event at the CU happens if $\mathbf{h}_1, \dots, \mathbf{h}_L$ belong to the L PD regions such that the corresponding active LNC coefficient matrix \mathbf{B} does not have a full rank of M in F_q .

To help with explaining the outage event, let us consider an example with $M = L = 2$. Consider the CD and PD regions as depicted in Fig. 2 for both BS 1 and BS 2. For a certain realization of (h_{11}, h_{12}) , BS 1 is subject to one of the four disjoint regions. Similarly, for a certain realization of (h_{21}, h_{22}) ,

BS2 is subject to one of the four disjoint regions. Then, the erasures in the active LNC matrix \mathbf{B} is also determined. If \mathbf{B} turns out to be full rank, the CU can recover both users' messages. Otherwise, an outage event happens. The outage events for the $M = L = 2$ case are described below. For notation simplicity, let $\mathcal{A}(S_l = \mathbf{w}_1)$ and $\mathcal{A}(S_l = \mathbf{w}_2)$ be respectively denoted by $\mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2)$ and $\mathcal{A}(\tilde{\mathbf{w}}_1, \mathbf{w}_2)$.

Corollary 2: For $M = L = 2$, an outage happens if $(h_{11}, h_{12}, h_{21}, h_{22}) \in \mathcal{R}_{out}$ where:

A) For $C^{BH} = R_0$ (or $N = 1$),

$$\mathcal{R}_{out} = \{ \mathbf{H}: (h_{11}, h_{12}) \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \cap (h_{21}, h_{22}) \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \\ \text{or } (h_{11}, h_{12}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \mathbf{w}_2) \cap (h_{21}, h_{22}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \mathbf{w}_2) \\ \text{or } (h_{11}, h_{12}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) \cup (h_{21}, h_{22}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) \}. \quad (27)$$

B) For $C^{BH} = 2R_0$ (or $N = 2$),

$$\mathcal{R}_{out} = \{ \mathbf{H}: (h_{11}, h_{12}) \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \cap (h_{21}, h_{22}) \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \\ \text{or } (h_{11}, h_{12}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \mathbf{w}_2) \cap (h_{21}, h_{22}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \mathbf{w}_2) \\ \text{or } (h_{11}, h_{12}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) \cap (h_{21}, h_{22}) \in \mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) \}. \quad (28)$$

Note that (27) and (28) differ in their last lines, i.e., “ \cup ” in (27) and “ \cap ” in (28).

3) *Asymptotic OP Results (Proof of Main Results):* As the size of the D-MIMO system, i.e., M and L , become large, there will be a large number of events that will lead to outages at the CU. This makes it very difficult to characterize the exact OP of the proposed protocol for general M, L and N , even at the infinite SNR regime. The results presented in Theorem 1 and Theorem 2 (in Section IV. A) showed the diversity order as well as the asymptotic performance gap to the full MIMO setup, for general M and L of $M \leq L$. The remainder of this section is devoted to the proofs of these results.

Recall the PD-regions in Definition 2. The event that user m 'th message cannot be successfully decoded at BS l is given by $h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l)$. Here, $h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l)$ specifies the event that h_{lm} falls into the union of the PD-regions $\mathcal{A}(S_l)$ where S_l does not contain the m th user's message \mathbf{w}_m . As $\rho \rightarrow \infty$, the probability of such an event can be upper-bounded as follows.

Proposition 1: As $\rho \rightarrow \infty$ and for $R_0 < \infty$,

$$\Pr \left\{ h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l) \right\} \leq \frac{2^{MR_0} - (M-1)2^{R_0} + M - 2}{\rho} \\ \propto \frac{1}{\rho}. \quad (29)$$

Proof: It can be shown that the following region

$$|h_{lm}|^2 < \frac{2^{MR_0} - 1 - (M-1)(2^{R_0} - 1)}{\rho}, \forall m = 1, \dots, M, \quad (30)$$

is strictly contained in the CD region. This follows from the fact that the CD region is convex and thus any linear combination of the corner points of the CD region is inside the CD region. Geometrically, the region specified in (30) is an M -dimension cubic that is strictly inside the pentagon of the CD region. Then, based on the above relaxation, we have

$$\Pr \left\{ h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l) \right\} \leq \Pr \left\{ |h_{lm}|^2 < \frac{2^{MR_0} - 1 - (M-1)(2^{R_0} - 1)}{\rho} \right\} \\ = \frac{2^{MR_0} - (M-1)2^{R_0} + M - 2}{\rho}.$$

We now prove Theorem 1.

Proof of Theorem 1: For $C^{BH} = R_0$ (or $N = 1$), the active LNC matrix \mathbf{B} is of size L by M . Note that each event of $h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l)$ contributes to a zero element in \mathbf{B} . Let τ be the number of zero elements in the active LNC matrix \mathbf{B} . From the above proposition, as $\rho \rightarrow \infty$, we have

$$\Pr[\tau = T] \leq \binom{M \cdot L}{T} \left(\frac{2^{MR_0} - (M-1)2^{R_0} + M - 2}{\rho} \right)^T \quad (31)$$

From Prop. 3 (on page 19), there exists a pre-assigned LNC matrix \mathbf{A} that is robust to $L-1$ erasures. As long as $\tau \leq L-1$, the CU can successfully recover all M users' messages. Thus the OP is upper-bounded by

$$\begin{aligned} P_{out}(R_0, C^{BH}) &\leq \sum_{T=L}^{M \cdot L} \binom{M \cdot L}{T} \left(\frac{2^{MR_0} - (M-1)2^{R_0} + M - 2}{\rho} \right)^T \\ &= \binom{M \cdot L}{L} \left(\frac{2^{MR_0} - (M-1)2^{R_0} + M - 2}{\rho} \right)^L \\ &\quad + o\left(\frac{1}{\rho^L}\right) \propto \frac{1}{\rho^L} \end{aligned} \quad (32)$$

In addition, since increasing N will not reduce the slope of $P_{out}(R_0, C^{BH})$, we have $P_{out}(R_0, C^{BH}) \propto \frac{1}{\rho^L}$ for $N > 1$. This completes the proof. ■

We next consider $N = M$ and prove the result of Theorem 2. The following lemma will be used.

Lemma 4: As $\rho \rightarrow \infty$, for $R_0 < \infty$, we have

$$\Pr \left\{ h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l) \right\} \approx \frac{2^{R_0} - 1}{\rho} + o\left(\frac{1}{\rho}\right) \quad (33)$$

for $m = 1, \dots, M, l = 1, \dots, L$. The proof of Lemma 4 is given in the appendix.

Proof of Theorem 2: As long as the pre-assigned LNC matrix satisfies Proposition 2, for $N = M$, the CU cannot correctly recover user m ' message only if there is no BS that can successfully decode user m 's message. The probability of such event is given by

$$\Pr \left\{ \bigcap_{l=1}^L h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l) \right\} \approx \left(\frac{2^{R_0} - 1}{\rho} \right)^L + o\left(\frac{1}{\rho^L}\right) \quad (34)$$

where we have used the result of Lemma 4. As $\rho \rightarrow \infty$, for $R_0 < \infty$ and $N = M$, the OP is equal to the probability that the CU cannot correctly recover at least one of the M users' messages, which is given by

$$\begin{aligned} P_{out}(R_0, C^{BH}) &= \sum_{m=1}^M \Pr \left\{ \bigcap_{l=1}^L h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l) \right\} \\ &\approx M \left(\frac{2^{R_0} - 1}{\rho} \right)^L + o\left(\frac{1}{\rho^L}\right). \end{aligned} \quad (35)$$

This completes the proof. ■

4) *Impact of BH Rate-Constraint:* To end this section, we briefly discuss the impact of the BH rate-constraint (i.e. N) on the OP performance, where we focus on a small system. For $M = L = 2$, we can derive an exact closed-form expression for

the OP at a high SNR for $N = 1$, as presented below and proved in Appendix:

Theorem 3: For $M = L = 2$, as $\rho \rightarrow \infty$, the OP with $C^{BH} = R_0$ ($N = 1$) is given by

$$P_{out}(R_0, C^{BH} = R_0) \approx \frac{4(2^{R_0} - 1)^2 + 2(2^{R_0} - 1)^3 + (2^{R_0} - 1)^4}{\rho^2}. \quad (36)$$

Remark 7: From Theorem 2, for $M = L = 2$ and $N = 2$, we have

$$P_{out}(R_0, C^{BH} = 2R_0) \approx 2 \left(\frac{2^{R_0} - 1}{\rho} \right)^2. \quad (37)$$

Comparing $P_{out}(R_0, C^{BH} = 2R_0)$ and $P_{out}(R_0, C^{BH} = R_0)$ in (36), it is clear that a larger N gives rise to a reduced OP. Furthermore, as R_0 increases, the improvement due to a larger C^{BH} becomes greater. We note that this behavior holds when M and L become large. (See Section VI.)

VI. DESIGN OF PRE-ASSIGNED LNC MATRIX \mathbf{A}

In this section, we first present the existence of a pre-assigned LNC matrix \mathbf{A} that yields the OP results in the previous section. Then, we study how to construct a desired matrix \mathbf{A} in an efficient manner.

A. Existence of Pre-assigned Matrix \mathbf{A}

As a preliminary, we first describe a condition. As long as this condition is met, the OP results shown in the previous section is guaranteed.

Condition 1 (Robust-to- Γ -erasures condition): A full rank matrix is said to be *robust to Γ erasures* if after any $\tau \leq \Gamma$ entries are set to be 0, the resultant matrix is still of full rank.

As an example, the 5×2 matrix $\begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 2 \end{bmatrix}^T$ over GF(3) is robust to 3 erasures but not to 4 erasures, because if we set all 4 nonzero entries to be zero, then the matrix is not full rank anymore. In order to have a 5×2 matrix robust to 4 erasures, we need a larger finite field than GF(3). For instance, $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 \end{bmatrix}^T$ is a matrix over GF(5) robust to 4 erasures. More generally, when $M = 2$, an $L \times M$ matrix \mathbf{A} which is robust to $L-1$ erasures can be constructed over any GF(q), $q \geq 4$, by selecting the first L rows in $\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & \dots \\ 1 & a & b & 1 & a & b & 1 & \dots \end{bmatrix}^T$, where a, b are arbitrary two distinct elements in GF(q) other than 0 and 1. It remains to discuss the cases when $M \geq 3$. As exemplifying matrices, one can check that

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 12 & 12 \\ 1 & 12 & 1 & 12 \\ 1 & 12 & 12 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 12 & 12 & 2 \\ 1 & 12 & 1 & 12 & 3 \end{bmatrix}^T$$

over GF(13) are respectively robust to 3 and 4 erasures. In the remaining part of this subsection, we shall show, by rather technical but standard arguments in the network coding literature (See [18] for example), that there always exists an $L \times M$ matrix robust to $L-1$ erasures over a sufficiently large finite field.

1) Case $M = L$:

Proposition 2: When $q > \sum_{n=2}^{M-1} \binom{M^2-1}{n}$, there exists an $M \times M$ matrix \mathbf{A} over $\text{GF}(q)$ that is robust to $M - 1$ erasures, i.e., after any $n \leq M - 1$ entries of \mathbf{A} are set to be 0, the resultant matrix is of full rank M .

Remark 8: In the proposed scheme, each zero element in the active LNC matrix \mathbf{B} corresponds to a decoding failure of a user's message at a specific BS. For a matrix \mathbf{A} that satisfies Cond.1, only when there are M zero elements in \mathbf{B} , \mathbf{B} will have rank deficiency and an outage event happens. In the proof of Theo. 1, we showed that the probability of M zero elements is proportional to $\frac{1}{\rho}$ as $\rho \rightarrow \infty$.

Proof: Let X be a set of indeterminates $\{x_{j,k}\}_{1 \leq j,k \leq M}$, and $A(X)$ be the $M \times M$ matrix with $(j,k)^{\text{th}}$ entry being $x_{j,k}$. Further let Ψ be the collection of all subsets of X with $0 \leq n < M$ elements, and $A(X, \psi)$, where $\psi \in \Psi$, be the matrix obtained from $A(X)$ by setting the $(j,k)^{\text{th}}$ entry to be 0 whenever $x_{j,k} \in \psi$. Then, $\det(A(X, \psi))$, to be denoted by $f_\psi(X)$, is a nonzero polynomial in indeterminates in X over $\text{GF}(q)$. In order to show that there exists a matrix \mathbf{A} over $\text{GF}(q)$ subject to Cond. 1., it is equivalent to show that there exists an assignment of $a_{j,k} \in \text{GF}(q)$ to $x_{j,k} \in X$ such that the evaluation $f_\psi((a_{j,k})_{1 \leq j,k \leq M})$ is nonzero for all $\psi \in \Psi$. Denote by Ψ' (resp. Ψ'') the collection of all such members in Ψ that each $\psi \in \Psi'$ (resp. $\psi \in \Psi''$) contains $M - 1$ indeterminates which all appear in a same row (resp. in a same column) in $A(X)$. Note that $\Psi' \subsetneq \Psi$ due to $M \geq 3$. Moreover, since $\prod_{\psi \in \Psi'} f_\psi(X) = \prod_{\psi \in \Psi''} f_\psi(X)$, it suffices to show that there exists an assignment of $a_{j,k} \in \text{GF}(q)$ to $x_{j,k} \in X$ such that the evaluation $\prod_{\psi \in \Psi \setminus \Psi'} f_\psi((a_{j,k})_{1 \leq j,k \leq M})$ is nonzero.

Consider an arbitrary indeterminate $x_{j,k}$. Note that there are totally $\sum_{n=0}^{M-1} \binom{M^2}{n} - M^2$ members in $\Psi \setminus \Psi'$. Since the degree of $x_{j,k}$ in $f_\psi(X)$ is at most one for every $\psi \in \Psi$, the degree of $x_{j,k}$ in $\prod_{\psi \in \Psi \setminus \Psi'} f_\psi(X)$ is upper bounded by $\sum_{n=0}^{M-1} \binom{M^2}{n} - M^2$. By a more dedicate counting, we can further show that the degree of $x_{j,k}$ in $\prod_{\psi \in \Psi \setminus \Psi'} f_\psi(X)$ is at most $\sum_{n=0}^{M-1} \binom{M^2}{n} - M^2 - \left(\sum_{n=0}^{M-2} \binom{M^2-1}{n} - M + 1 \right) - (M - 1) = \sum_{n=0}^{M-1} \binom{M^2}{n} - \sum_{n=0}^{M-2} \binom{M^2-1}{n} - M^2 = \sum_{n=2}^{M-1} \binom{M^2-1}{n}$, where the last equality is due to the fact that $\binom{M}{n} = \binom{M}{n-1} + \binom{M-1}{n-1}$ for any $M, n > 1$. By applying the lemma below with $f(x_1, \dots, x_n) = \prod_{\psi \in \Psi \setminus \Psi'} f_\psi(X)$, we conclude that if $q > \sum_{n=2}^{M-1} \binom{M^2-1}{n}$, there exist $a_{j,k} \in \text{GF}(q)$ for $1 \leq j, k \leq M$ such that $\prod_{\psi \in \Psi \setminus \Psi'} f_\psi((a_{j,k})_{1 \leq j,k \leq M}) \neq 0$, which guarantees $f_\psi((a_{j,k})_{1 \leq j,k \leq M}) \neq 0$ for all $\psi \in \Psi$. The proposition is proved. ■

Lemma 5: Consider a polynomial $f(x_1, \dots, x_n)$ over $\text{GF}(q)$. Denote by d_j the highest degree of x_j in $f(x_1, \dots, x_n)$. If $q > \max\{d_1, \dots, d_n\}$, then there exist $a_1, \dots, a_n \in \text{GF}(q)$ such that $f(a_1, \dots, a_n) \neq 0$.

Proof: This can be proved by an inductive argument on n , which is obviously true for $n = 1$. ■

Remark 9: The technique adopted to prove Prop. 2 is to show the existence of a nonzero evaluation of a properly designed multi-variable polynomial over $\text{GF}(q)$. It is well known that there is a solution over $\text{GF}(q)$ for a multicast network if q is larger than the number of receivers in the network (See [22]). However, the gap between the number of receivers and the minimum field size q_{\min} for the existence of an LNC solution can be extremely large and the characterization of q_{\min} remains open. Similarly, in

Prop. 2, the existence of an $M \times M$ square matrix \mathbf{A} over $\text{GF}(q)$ subject to Cond.1 is only guaranteed when $q > \sum_{n=2}^{M-1} \binom{M^2-1}{n}$. For instance, for $M = 3$, we need $q > 28$ and $q >$ respectively in Prop. 2. However, the minimum field size by computer search subject to Cond.1 is $q = 8$ with matrix $\begin{bmatrix} 1 & 1 & 1 \\ 1 & \alpha & \alpha^3 \\ 1 & \alpha^5 & \alpha \end{bmatrix}$, where α is a primitive element in $\text{GF}(8)$.

2) Case $M \leq L$:

Proposition 3: When $q > M \sum_{n=2}^{M-1} \binom{M^2-1}{n}$, there is an $L \times M$ matrix \mathbf{A} over $\text{GF}(q)$ which is robust to $L - 1$ erasures, i.e., after any $n \leq L - 1$ entries of \mathbf{A} are set to be 0, the resultant matrix is of full rank M .

Proof: Let X denote a set of indeterminates $\{x_{j,k}\}_{1 \leq j \leq L, 1 \leq k \leq M}$, and $A(X)$ denote the $L \times M$ matrix with $(j,k)^{\text{th}}$ entry being $x_{j,k}$. Further let Ψ be the collection of all subsets of X with $n < L$ elements, and $A(X, \psi)$, where $\psi \in \Psi$, be the matrix obtained from $A(X)$ by setting $x_{j,k} = 0$ whenever $x_{j,k} \in \psi$. Corresponding to each $A(X, \psi)$, define a collection $\mathcal{A}(\psi)$ of $L - M \times M$ submatrices $A_1(X, \psi), \dots, A_L(X, \psi)$ of $A(X, \psi)$ in such a way that each $A_n(X, \psi)$ consists of the M rows in $A(X, \psi)$ indexed by $\{n, \dots, n + M - 1\}$, where the indices are modulo L if they are larger than L .

Observe that for any $\psi \in \Psi$, there is at least one matrix in $\mathcal{A}(\psi)$ that contains at most $M - 1$ zeros. This is because the total number of zero entries in all matrices in $\mathcal{A}(\psi)$ is equal to $M|\psi| \leq M(L - 1)$, whereas there are only L matrices in $\mathcal{A}(\psi)$, so that it is impossible for every matrix in $\mathcal{A}(\psi)$ to contain no less than M zeros. For every $1 \leq n \leq L$, denote by $X_n \subseteq X$ the set of indeterminates $\{x_{j,k}\}_{n \leq j \leq n + M - 1, 1 \leq k \leq M}$, where the index j is modulo L when it is larger than L , and denote by $\Psi_n \subseteq \Psi$ the collection of all subsets of X_n with $M' < M$ elements. Then, for each $\psi \in \Psi$, at least one $A_n(X, \psi) \in \mathcal{A}(\psi)$ belongs to $\{A_n(X, \psi) : \psi \in \Psi_n\}$. It remains to show that when $q > M \sum_{n=2}^{M-1} \binom{M^2-1}{n}$, there is an assignment of $a_{j,k} \in \text{GF}(q)$ to $x_{j,k} \in X$ s.t. every matrix in $\bigcup_{1 \leq n \leq L} \{A_n(X, \psi) : \psi \in \Psi_n\}$ is of full rank M . Same as in the proof of Prop. 2, specific to an arbitrary n , we can establish such a polynomial, denoted by $f_n(X_n)$, in indeterminates in X_n that (i) the degree of every indeterminate is upper bounded by $\sum_{n=2}^{M-1} \binom{M^2-1}{n}$; (ii) based on an assignment of $a_{j,k} \in \text{GF}(q)$ to $x_{j,k} \in X_n$, every matrix in $\{A_n(X, \psi) : \psi \in \Psi_n\}$ is of full rank M iff the evaluation of this polynomial is nonzero. Then, it is equivalent to show that there is an assignment of $a_{j,k} \in \text{GF}(q)$ to $x_{j,k} \in X$ such that the evaluation of the polynomial $\prod_{1 \leq n \leq L} f_n(X_n)$ is nonzero. Since every indeterminate only appears in at most M matrices in $\mathcal{A}(\psi)$, the degree of every indeterminate in $\prod_{1 \leq n \leq L} f_n(X_n)$ is upper bounded by $M \sum_{n=2}^{M-1} \binom{M^2-1}{n}$. When q is larger than this value, a desired assignment is feasible according to Lemma 5. ■

B. Construction of Pre-assigned Matrix \mathbf{A}

The previous subsection proves the existence of a pre-assigned matrix \mathbf{A} subject to robust-to- $(L-1)$ -erasure condition. Recall that matrix \mathbf{A} is pre-assigned in the proposed scheme. Once \mathbf{A} is determined, it will remain unchanged for different fading channel realizations. Therefore, from a practical perspective, the complexity for obtaining matrix \mathbf{A} will not be a major concern. Yet, it will be of theoretical interests to investigate how to efficiently construct such a matrix as follows. First, we shall

reduce the construction of an $L \times M$ matrix robust to $L - 1$ erasures to the construction of an $(L - 1) \times (M - 1)$ matrix.

Proposition 4: Let \mathbf{A} be an $L \times M$ matrix over $\text{GF}(q)$ satisfying robust-to- $(L-1)$ -erasure condition. Every entry in \mathbf{A} is nonzero and there exists a matrix over $\text{GF}(q)$ satisfying robust-to- $(L-1)$ -erasure condition with all entries in the first row and first column equal to 1.

Proof: Denote by $a_{j,k}$ the $(j,k)^{\text{th}}$ entry in \mathbf{A} . Consider an arbitrary $a_{j,k}$. Denote by A' the matrix obtained from \mathbf{A} via setting $a_{j',k} = 0$ for all $j' \neq j$. Since \mathbf{A} satisfies condition robust-to- $(L-1)$ -erasure condition, A' has full rank M . If $a_{j,k}$ is equal to 0, then the k^{th} column in A' is a zero column and thus $\text{rank}(A') < M$, a contradiction. Therefore, $a_{j,k} \neq 0$. Next, we can obtain another matrix A'' from \mathbf{A} via sequentially multiplying the j^{th} row by $a_{j,1}^{-1}$ for all $1 \leq j \leq L$, and then sequentially multiplying the k^{th} column in the new matrix by $a_{1,1}^{-1}a_{1,k}^{-1}$ for all $2 \leq k \leq M$. All entries in the first row and first column in A'' are then equal to 1. Since the rank of an arbitrary matrix over a field keeps the same after any row or column in the matrix is multiplied by a nonzero value in the field, matrix A'' satisfies robust-to- $(L-1)$ -erasure condition too. ■

Recall that the size of a finite field must be a power of a prime integer. Assume that q is a prime greater than $M \sum_{n=2}^{M-1} \binom{M^2-1}{n}$, so that the existence of a matrix robust to $L - 1$ erasures is guaranteed by Prop. 3.

Assume that q is a prime power greater than $M \sum_{n=2}^{M-1} \binom{M^2-1}{n}$. If we need to find a matrix over $\text{GF}(q)$ s.t. robust-to- $(L-1)$ -erasure condition by exhaustive search, the total number of matrices over $\text{GF}(q)$ we need to check is in the order $O(q^{LM})$. If we assume that the computational complexity is $O(LM^2)$ to check whether an $L \times M$ matrix has full rank (by Gaussian elimination for example,) the computational complexity to find a matrix subject to robust-to- $(L-1)$ -erasure condition over $\text{GF}(q)$ by exhaustive search is $O(LM^2q^{LM})$.

We next establish a more structured way to construct a matrix subject to robust-to- $(L-1)$ -erasure condition. A matrix is called a *mixed matrix* over $\text{GF}(q)$ if every entry in it contains either a value in $\text{GF}(q)$ or an indeterminate. Let \mathcal{A} denote a collection of such $M \times M$ square mixed matrices over $\text{GF}(q)$ that every indeterminate appears only at most once in every matrix, and let X denote the set of indeterminates appearing in \mathcal{A} . A *max-rank matrix completion* of \mathcal{A} is a procedure to assign values in $\text{GF}(q)$ to indeterminates in X such that the rank of every matrix in \mathcal{A} is preserved. When $q > |\mathcal{A}|$, an algorithm is designed in [23] for max-rank completion of \mathcal{A} with computational complexity $O(|\mathcal{A}|(M^3 \log M + |X|M^2))$.

Now, in order to construct an $M \times M$ matrix over $\text{GF}(q)$ subject to robust-to- $(L-1)$ -erasure condition, it suffices to construct a matrix over $\text{GF}(q)$ with all entries in the first row and the first column equal to 1 as a consequence of Proposition 4. Let \mathbf{A} denote a mixed matrix such that (i) the entries in the first row and the first column are equal to 1; (ii) the $(j,k)^{\text{th}}$ entry is an indeterminate $x_{j,k}$ for all $2 \leq j, k \leq M$. Further denote by X the set of ordered pairs $\{(j,k) : 1 \leq j, k \leq M\}$, by Ψ the collection of all subsets of X with $n < M$ elements, and by $A(\psi)$ where $\psi \in \Psi$ the matrix obtained from \mathbf{A} via setting the $(j,k)^{\text{th}}$ entry to be 0 whenever $(j,k) \in \psi$. Then, a max-rank matrix completion procedure on $\mathcal{A} = \{A(\psi) : \psi \in \Psi\}$ is able to construct a matrix subject to

robust-to- $(L-1)$ -erasure condition. Specific to an entry indexed by (j,k) , consider two members in Ψ : $\psi' = \{(j',k) : j' \neq j\}$, $\psi'' = \{(j,k') : k' \neq k\}$. Since $\det(A(\psi')) = \det(A(\psi''))$, a max-rank matrix completion of $\{\mathbf{A}, A(\psi')\}$ is equivalent to a max-rank matrix completion of $\{\mathbf{A}, A(\psi'')\}$. As a result, it suffices to conduct a max-rank matrix completion procedure on $\mathcal{A} = \{A(\psi) : \psi \in \Psi \setminus \Psi'\}$ to find a matrix subject to robust-to- $(L-1)$ -erasure condition, where $\Psi' = \{\{(j_1,k), \dots, (j_{M-1},k)\} : 1 \leq j_1 < \dots < j_{M-1} \leq M, 1 \leq k \leq M\}$. Then, based on the algorithm proposed in [23], when $q > |\mathcal{A}| = \sum_{n=0}^{M-1} \binom{M^2}{n} - M^2 = \sum_{n=1}^{M-1} \binom{M^2}{n}$, an $M \times M$ matrix \mathbf{A} subject to robust-to- $(L-1)$ -erasure condition can be constructed with computational complexity $O(\sum_{n=1}^{M-1} \binom{M^2}{n} (M^3 \log M + M^4)) = O(M^4 \sum_{n=1}^{M-1} \binom{M^2}{n})$. As an extension, we can also adopt the max-rank matrix completion approach to construct an $L \times M$ matrix subject to robust-to- $(L-1)$ -erasure condition. Let \mathbf{A} denote an $L \times M$ mixed matrix s.t. (i) the entries in the first row and the first column are equal to 1; (ii) the $(j,k)^{\text{th}}$ entry is an indeterminate $x_{j,k}$ for all $2 \leq j \leq L$ and $2 \leq k \leq M$. Based on \mathbf{A} , define L $M \times M$ submatrices $\mathbf{A}_1, \dots, \mathbf{A}_L$ of \mathbf{A} in such a way that each \mathbf{A}_n , $1 \leq n \leq L$ consists of the M rows in \mathbf{A} indexed by $\{n, \dots, n + M - 1\}$, where the indices are modulo L if they are larger than L . As explained in the proof of Prop. 3, after any $L - 1$ entries in \mathbf{A} are set to 0, there is at least one submatrix \mathbf{A}_n among $\mathbf{A}_1, \dots, \mathbf{A}_L$ containing at most M zeros. Thus, a max-rank matrix completion procedure on the family $\mathcal{A} = \{\mathbf{A}_n(\psi) : \psi \in \Psi, 1 \leq n \leq L\}$ of mixed matrices, where Ψ and $\mathbf{A}_n(\psi)$ are defined in the same way as in the previous paragraph, is sufficient generate an $L \times M$ matrix satisfying robust-to- $(L-1)$ -erasure condition. By a similar argument as in the previous paragraph, we can further show that when $q' > L \sum_{n=1}^{M-1} \binom{M^2}{n}$, the algorithm proposed in [23] can construct an $L \times M$ matrix satisfying robust-to- $(L-1)$ -erasure condition over $\text{GF}(q')$ with the computational complexity $O(L \sum_{n=1}^{M-1} \binom{M^2}{n} (M^3 \log M + (L-1)(M-1)M^2)) = O(L^2 M^3 \sum_{n=1}^{M-1} \binom{M^2}{n})$.

VII. NUMERICAL RESULTS

In this section, we present numerical results of the uplink D-MIMO system. Here we first present the numerical results of OPs studied in the preceding sections. They serve as upper bounds on the frame error rate (FER) of practical schemes with state-of-the-art error-control codes and digital modulations. We will compare the OP of the proposed scheme with those of existing schemes in the literature. After that, we show the FER performance of a practical turbo coded scheme with multi-user decoding at the BSs. The pre-assigned LNC coefficient matrix \mathbf{A} used in the simulations are shown in Table I.

A. Outage Probability

We first consider an $M = L = 2$ system, where the per-user information rate is $R_0 = 1$ bit/channel-use. In Fig. 3, the red curve (in solid line and with circles) is the OP of a baseline scheme where BS 1 (or BS 2) only decodes and forwards its user 1's (or user 2's) message⁶. The green curve

⁶In the baseline system, this assignment is fixed. In the model under consideration, the cloud base stations do not share channel state information. Therefore, the base stations cannot perform adaptive user selection according to the instantaneous SNR.

No. of users and BSs	$M = L = 2$	$M = L = 3$	$M = L = 4$
Size of F_q	$q = 3$	$q = 8$	$q = 13$
Pre-assigned matrix \mathbf{A}	$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$	$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \gamma & \gamma^3 \\ 1 & \gamma^5 & \gamma \end{bmatrix}$	$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 12 & 12 \\ 1 & 12 & 1 & 12 \\ 1 & 12 & 12 & 1 \end{bmatrix}$

TABLE I
PRE-ASSIGNED LNC COEFFICIENT MATRIX \mathbf{A} . HERE, γ IS A PRIMITIVE ELEMENT OF THE FIELD (FOR $M = L = 3$).

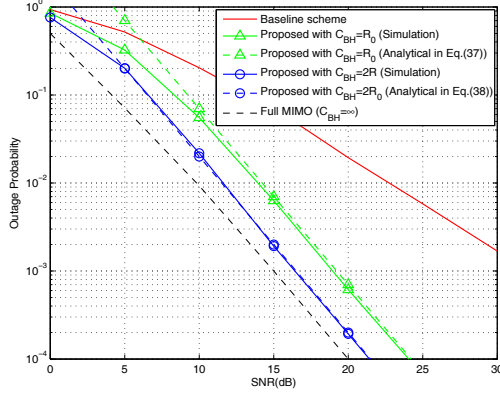


Fig. 3. OP performance of $M = L = 2$ and $R_0 = 1$.

(in solid line and with triangles) is the OP of the proposed LNC based scheme, where $C_{BH} = R_0$ ($N = 1$). We observe that, with our designed matrix \mathbf{A} , the proposed scheme yields dramatic performance improvement over the baseline scheme. In particular, our proposed scheme achieves a diversity order of two, which agrees with our analysis in Section V. We also plot the derived closed-form result in Eq. (37), which matches with the simulated results at SNR greater than 15 dB. We note that, if an arbitrary \mathbf{A} (rather than that described in Section VI) is used, the OP of the LNC based scheme may be much worse than that shown in this figure due to rank deficiency. In Fig. 3, we also plot the OP of the proposed scheme, where $C^{BH} = 2R_0$ ($N = 2$). At a medium-to-high SNR, a 2.6 dB improvement relative to that of $C^{BH} = R_0$ is observed. The simulated curve matches with our derived close-form result in Eq. (37) at SNR greater than 15 dB. We also observe that the OP performance is about 1.5 dB away from that of the full MIMO setup, and this agrees with our analysis presented in Section V. A.

We evaluate the performance for $R_0 = 2$ as shown in Fig. 4, which lead to similar observations. We observe that the performance improvement with $C^{BH} = 2R_0$ over that with $C^{BH} = R_0$ is about 4.3 dB, which is greater than the 2.6 dB in the $R_0 = 1$ case in Fig. 3. This shows that as the data rate R_0 increases, increasing N will lead to a greater performance improvement. This agrees with our discussion at the end of Section V.

In Fig. 5, we compare the OP performance of the proposed scheme with the compute-and-forward (CF) scheme [10] and the quantize-and-forward (QF) scheme [24]. For the CF scheme, each BS locally selects the N sets of coefficient vectors that yield the highest computation rates which are greater than R_0 [10]. For the QF scheme, each BS quantizes its received signal into NR_0 bits. The variance of the quantization noise was shown in [24].

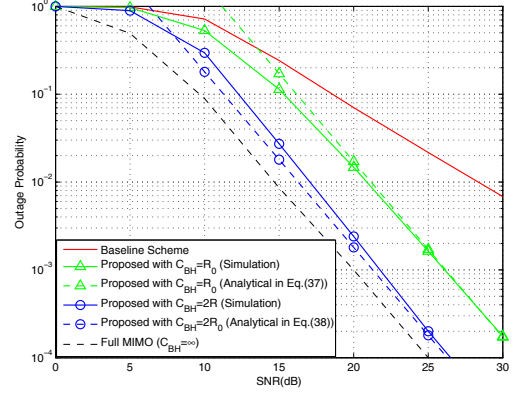


Fig. 4. OP performance of $M = L = 2$ and $R_0 = 2$.

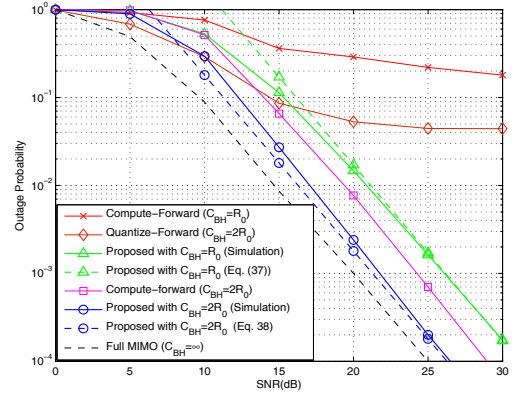


Fig. 5. Comparison of OP performance of various schemes, where $M = L = 2$ and $R_0 = 2$.

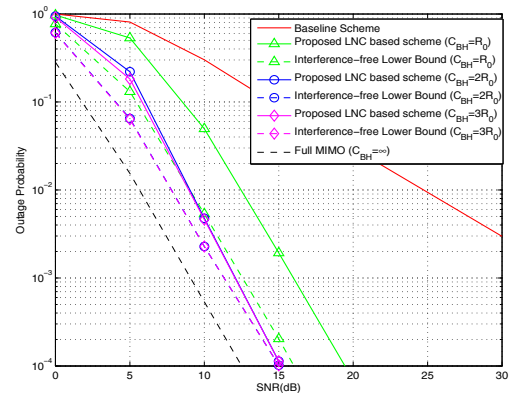


Fig. 6. OP performance for $M = L = 3$, $R_0 = 1$.

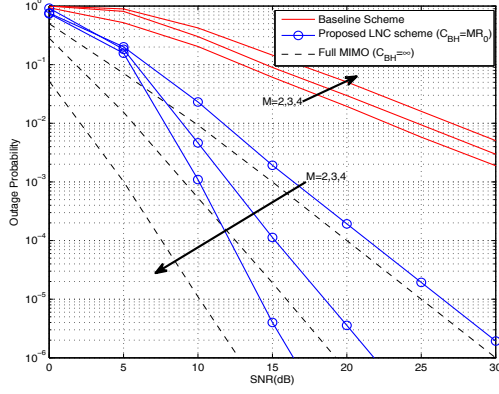


Fig. 7. OP performance for various $M, M = L$. $R_0 = 1$.

We observe that the proposed LNC based scheme outperforms both the CF scheme and the QF scheme. For $N = 1$, for the CF scheme, there is a high probability that the coefficient vectors separately selected by the BSs are not full rank. In contrast, the proposed scheme can dramatically reduce the probability of the rank deficiency by carefully designing the pre-assigned LNC matrix \mathbf{A} . For $N = 2$, the rank deficiency problem of the CF scheme is alleviated, but its OP performance is still inferior to that of the proposed scheme. Moreover, it can be seen that the QF scheme is subject to a severe error floor, which is due to the severe quantization noise. Note that we can derive the closed-form result on the OP of our proposed scheme. At the moment, a closed-form result cannot be done for QF and CF.

In Fig. 6, we consider an $M = L = 3$ system, where the per-user information rate is $R_0 = 1$. The BH capacities under consideration are $C^{BH} = R_0$ and $C^{BH} = 2R_0$. It is clear that proposed scheme yields dramatic performance improvement over the baseline scheme. In particular, our proposed scheme achieves a diversity order of three, which agrees with our analysis in Section V.

In Fig. 7, we plot the OPs for various numbers of M and L where $R_0 = 1$. Here, we consider $M = L = N$. It is clear that as M increases, the OP of the D-MIMO system with the proposed LNC based approach improves. In contrast, the performance of the baseline scheme (where BS l only decodes user l 's message) becomes worse as M increases, due to that the interference becomes more severe. At a SNR higher than 15 dB, we observe that our proposed scheme is about 1.6 dB, 2.6 dB, 3.7 dB away from the OP of the full MIMO case, for $M = 2, 3, 4$, respectively. This agrees with our results in Theorem 2 and Corollary 1.

B. Frame Error Rates of Turbo-Coded and Modulated Systems

In Fig. 8, we consider an $M = L = 2$ system, where every user employs a rate-1/2 turbo code with generator polynomials $[37, 21]_8$ and QPSK. The information rate is $R_0 = 1$ bit/channel-use. In our simulation, each block consists of 256 message bits. Each BS carries out multi-user decoding via iterative a posteriori probability (APP) detection and decoding [25], [26]. Note that each BS is not aware of other BSs' processing. For $C^{BH} = R_0$ ($N = 1$), the simulated FER performance of the practical coded and modulated system is about 1.5 dB away from our derived theoretical OP (at a medium-to-high SNR regime). We conjecture

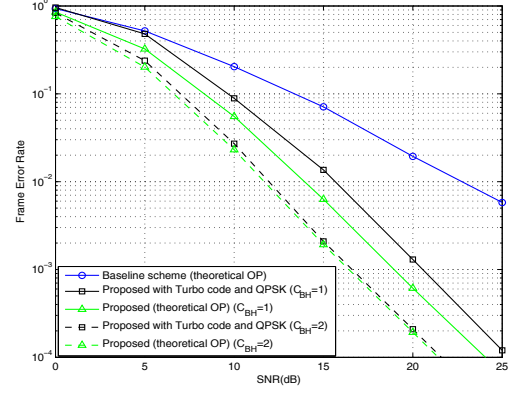


Fig. 8. FER performance of $M = L = 2$. Each user employs a rate-half turbo code and QPSK.

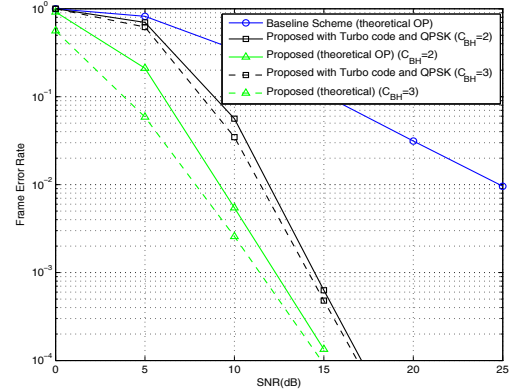


Fig. 9. FER performance of $M = L = 3$. Each user employs a rate-half turbo code and QPSK.

that this gap is primarily due to that the simulated coding and modulation scheme may not be able to achieve the full rate-region of the MAC. From the literature, it is known that universal codes, such as spatially coupled LDPC codes, can be employed to achieve the whole rate-region of the multiple-access channel [21], and yield further improved FER performance. Interestingly, for $C^{BH} = 2R_0$, the simulated FER performance of the practical coded and modulated system is only about 0.2 dB away from our derived theoretical OP, and the gap becomes almost unnoticeable. These FER results further confirm our derived OP results.

In Fig. 9, we consider an $M = L = 3$ system, where again each user employs the rate-1/2 turbo code and QPSK. We observe that, the simulated FER performance of the proposed scheme is about 2.0 dB away from the theoretical OP, for both $C^{BH} = 2R_0$ and $C^{BH} = 3R_0$. This increased gap between the FER and the theoretical OP, relative to the $M = L = 2$ case, is primarily due to that the simulated coding and modulation scheme may not be able to achieve the full rate-region of the MAC, and the gap becomes greater as the number of users M increases. We expect that spatially coupled LDPC codes can be designed to considerably reduce this gap [21]. This will be considered in our future work.

VIII. CONCLUSIONS

We proposed a LNC approach for uplink D-MIMO systems where the transmitters do not have CSI. By carefully designing the pre-assigned LNC matrix, the proposed scheme exhibited significantly improved outage probability performance over the base-line single-cell processing scheme, the compute-and-forward scheme and the quantize-and-forward scheme. We showed that as long as the BH rate-constraint is not smaller than the data rate, i.e., $C_{BH} \geq R_0$, the proposed scheme achieves a full diversity order of L determined by the number of distributed BSs. As $C_{BH} \geq M \cdot R_0$, the scheme is only $\frac{10}{L} \log_{10}(L!)$ dB away from the performance of the full MIMO setup at a high SNR. In addition, the performance improves as the size of the D-MIMO system scales up. We also showed that the frame error rate performance of a practical turbo coded system is reasonably close to our derived theoretical OP.

The work in this paper can be enriched from several aspects. For example, it is an open question whether the performance gap (of $\frac{10}{L} \log_{10}(L!)$ dB when $C_{BH} \geq M \cdot R_0$) to the full MIMO case can be further reduced. Also, the smallest field size q w.r.t. the LNC matrix that guarantees this high-performance is still unknown. In addition, the design of the LNC based approach for the multi-antenna setup remains open. These interesting open problems may deserve further investigations.

APPENDIX

Proof of Lemma 4: Define the following region:

$$\mathcal{D} = \{ (h_{l1}, \dots, h_{lM}) : h_{lm} \geq \frac{2^{R_0} - 1}{\rho}, \forall m = 1, \dots, M. \quad (38)$$

Clearly, $\mathcal{A}(\mathbf{w}_1, \dots, \mathbf{w}_M) \subseteq \mathcal{D}$, i.e., the CD region is contained in the region \mathcal{D} .

Also, define another region as:

$$\mathcal{G} \triangleq \mathcal{A}^c(\mathbf{w}_1, \dots, \mathbf{w}_M) \cap \mathcal{D} \quad (39)$$

which is the part of \mathcal{D} not overlapped with the CD region $\mathcal{A}(\mathbf{w}_1, \dots, \mathbf{w}_M)$. It can be shown that $\mathcal{G} \subseteq \mathcal{A}(\emptyset)$.

We next show that as $\rho \rightarrow \infty$, \mathcal{G} approaches an empty set. We prove this by contradiction.

Consider $R_0 < \infty$. Suppose that \mathcal{G} is non-empty. According to the definition of \mathcal{G} in the above, none of the M users can successfully decode its messages if $(h_{l1}, \dots, h_{lM}) \in \mathcal{G}$. This suggests that

$$|h_{lm}|^2 < \frac{2^{R_0} - 1}{\rho} \left(1 + (M-1) \frac{2^{R_0} - 1}{\rho} \right) \xrightarrow{\rho \rightarrow \infty} \frac{2^{R_0} - 1}{\rho} \quad (40)$$

for $(h_{l1}, \dots, h_{lM}) \in \mathcal{G}$, since otherwise user m 's message can be successfully decoded by treating the other $(M-1)$ users' messages as noise. This leads to absurdity from the definition of \mathcal{D} and $\mathcal{G} \subseteq \mathcal{D}$. Therefore \mathcal{G} is an empty set as $\rho \rightarrow \infty$. The above suggests that $\mathcal{A}(\mathbf{w}_1, \dots, \mathbf{w}_M) \rightarrow \mathcal{D}$ as $\rho \rightarrow \infty$. Then we have

$$\begin{aligned} \Pr \left\{ h_{lm} \in \bigcup_{S_l: \mathbf{w}_m \notin S_l} \mathcal{A}(S_l) \right\} &\approx \left(|h_{lm}|^2 < \frac{2^{R_0} - 1}{\rho} \right) \\ &= \frac{2^{R_0} - 1}{\rho} + o\left(\frac{1}{\rho}\right). \end{aligned}$$

Proof of Theorem 3: As $\rho \rightarrow \infty$, the probability that (h_{l1}, h_{l2}) belongs to $\mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2)$ is given by

$$\begin{aligned} &\Pr \{ h_{l1}, h_{l2} \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \} \\ &= \frac{2^R - 1}{\rho} \left(1 - \frac{1}{2} \frac{(2^R - 1)^2}{\rho} - \frac{2^R - 1}{\rho} \right) + o\left(\frac{1}{\rho}\right) \\ &= \frac{2^R - 1}{\rho} + o\left(\frac{1}{\rho}\right), l \in \{1, 2\}. \end{aligned} \quad (41)$$

This follows from integrating over the PD-region $\mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2)$ over the pdf of a Rayleigh distribution. As $\rho \rightarrow \infty$, $\Pr \{ h_{l1}, h_{l2} \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \}$ is just the area of $\mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2)$, which leads to (41). Similarly,

$$\Pr \{ h_{l1}, h_{l2} \in \mathcal{A}(\tilde{\mathbf{w}}_1, \mathbf{w}_2) \} = \frac{2^R - 1}{\rho} + o\left(\frac{1}{\rho}\right), l \in \{1, 2\}. \quad (42)$$

As $\rho \rightarrow \infty$, the probability that (h_{l1}, h_{l2}) belongs to the non-decodable region $\mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2)$ is given by

$$\begin{aligned} &\Pr \{ h_{l1}, h_{l2} \in \mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) \} \\ &= \frac{(2^R - 1)^2 + (2^R - 1)^3 + \frac{1}{2} (2^R - 1)^4}{\rho^2} + o\left(\frac{1}{\rho^2}\right) \end{aligned} \quad (43)$$

Consider the event specified in the second line of (27). The probability of this event is

$$\begin{aligned} P^{(1)} &= \Pr \{ h_{l1}, h_{l2} \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \} \Pr \{ h_{21}, h_{22} \in \mathcal{A}(\mathbf{w}_1, \tilde{\mathbf{w}}_2) \} \\ &= \frac{(2^R - 1)^2}{\rho^2} + o\left(\frac{1}{\rho^2}\right) \end{aligned}$$

where we have used Eq. (41). Similarly, the probability w.r.t. the event w.r.t. the third line of (27) is given by

$$P^{(2)} = \frac{(2^R - 1)^2}{\rho^2} + o\left(\frac{1}{\rho^2}\right). \quad (44)$$

Moreover, it can be shown that the probability of the event specified in the fourth line of (27) is given by

$$\begin{aligned} P^{(3)} &= \Pr \{ h_{l1}, h_{l2} \in \mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) \} + \Pr \{ h_{21}, h_{22} \in \mathcal{A}(\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) \} \\ &= \frac{2(2^R - 1)^2 + (2^R - 1)^3 + \frac{1}{2} (2^R - 1)^4}{\rho^2} + o\left(\frac{1}{\rho^2}\right) \end{aligned}$$

where we have used Eq. (43). Finally, considering all error events, we have

$$\begin{aligned} P_{out}(R_0, C^{BH} = R_0) &\leq P^{(1)} + P^{(2)} + P^{(3)} \\ &= \frac{4(2^R - 1)^2 + 2(2^R - 1)^3 + (2^R - 1)^4}{\rho^2} + o\left(\frac{1}{\rho^2}\right) \end{aligned}$$

and this completes the proof. \blacksquare

REFERENCES

- [1] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multi-cell mimo cooperative networks: a new look at interference," *IEEE Jour. Sel. Area. Comm.*, vol. 28, no. 9, pp. 1390–1408, Dec. 2010.
- [2] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzaresse, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in lte-advanced: deployment scenarios and operational challenges," *IEEE Communications Magazine*, vol. 50, no. 2, pp. 148–155, 2012.
- [3] J. Zhang, I. Collings, C. Chen, L. Roullet, L. Luo, S. Ho, and J. Yuan, "Evolving small-cell communications towards mobile-over-FTTx networks," *IEEE Communications Magazine*, vol. 51, no. 12, pp. 92–101, 2013.

- [4] M. Hadzialic, B. Dosenovic, M. Dzaferagic, and J. Musovic, "Cloud-ran: Innovative radio access network architecture," *ELMAR, 2013 55th International Symposium*, pp. 115–120, 2013.
- [5] C. Mobile, "C-ran: the road towards green ran," *White Paper, ver. 2.5, China Mobile Research Institute*, Oct. 2011.
- [6] C. Bernardos, A. D. Domenico, J. Ortin, P. Rost, and D. Wübben, "Challenges of designing jointly the backhaul and radio access network in a cloud-based mobile network," in *Future Network & Mobile Summit 2013, Lisbon, Portugal*, Jul 2013.
- [7] Y. Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," *IEEE Jour. Selected Area in Comm.*, vol. 32, no. 6, pp. 1295–1307, June 2014.
- [8] H. C. Chen, Y. Hu, P. P. Lee, and Y. Tang, "Nccloud: A network-coding-based storage system in a cloud-of-clouds," *IEEE Trans. Comput.*, vol. 63, no. 1, pp. 31–44, Jan. 2014.
- [9] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Distributed MIMO in multi-cell wireless systems via finite-capacity links," in *Proc. ISCCSP*, 2008.
- [10] 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.
- [11] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
- [12] S. Lim, Y. H. Kim, A. E. Gamal, and S. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May. 2011.
- [13] B. Nazer, A. Sanderovich, M. Gastpar, and S. Shamai, "Structured superposition for backhaul constrained cellular uplink," *IEEE Intern. Symp. Inform. Theory*, Jun. 2009.
- [14] T. Yang and I. B. Collings, "On the optimal design and performance of linear physical-layer network coding for fading two-way relay channels," *IEEE Trans. Wireless Comm. (early-access available)*, 2014.
- [15] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Local base station cooperation via finite-capacity links for the uplink of linear cellular networks," *IEEE Trans. Inform. Theory*, vol. 55, no. 1, pp. 190–204, Jan. 2009.
- [16] R. Zakhour and S. Hanly, "Base station cooperation on the downlink: Large system analysis," *IEEE Trans. Inform. Theory*, vol. 58, no. 4, pp. 2079–2106, Apr. 2012.
- [17] T. Yang, J. Yuan, and Z. Shi, "Rate optimization for idma systems with iterative joint multi-user decoding," *IEEE Trans. Wireless Comm.*, vol. 8, no. 3, pp. 1148–1153, Mar. 2009.
- [18] P. Marsch, "Coordinated multi-point under a constrained backhaul and imperfect channel knowledge," a *PhD thesis of TU Dresden*, 2010.
- [19] L. Wei and W. Chen, "Compute-and-forward network coding design over multi-source multi-relay channels," *IEEE Trans. Wireless Comm.*, vol. 11, no. 9, pp. 3348–3357, 2012.
- [20] T. M. Cover and J. A. Thomas, "Elements of information theory," *John Wiley & Sons, Inc.*, 1991.
- [21] A. Yedla, P. S. Nguyen, H. D. Pfister, and K. Narayanan, "Universal codes for the gaussian MAC via spatial coupling," available at arxiv.org/pdf/1110.0252.
- [22] S.-Y. R. Li, Q. T. Sun, and Z. Shao, "Linear network coding: theory and algorithms," *Proc. IEEE*, vol. 99, no. 3, pp. 372–387, Mar. 2011.
- [23] N. J. A. Harvey, D. R. Karger, and K. Murota, "Deterministic network coding by matrix completion," *Annual ACM-SIAM Symposium on Discrete Algorithm*, 2005.
- [24] A. Sanderovich, S. Shamai, and Y. Steinberg, "Distributed MIMO receiver - achievable rates and upper bounds," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4419–4438, Oct. 2009.
- [25] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Comm.*, vol. 51, no. 3, pp. 389–400, Mar. 2003.
- [26] T. Yang, J. Yuan, and Z. Shi, "Jointly gaussian approximation and multi-stage LLR combining in the iterative receiver for MIMO-BICM systems," *IEEE Trans. Wireless Comm.*, vol. 7, no. 12, pp. 5250–5256, Dec. 2008.



Tao Yang (S'07, M'10) received B.Sc. degree in electronic engineering in 2003 from Beijing University of Aeronautics and Astronautics (Beihang University), Beijing, China. He received Master by research and Ph.D. degrees in electrical engineering from the University of New South Wales, Sydney, Australia, in 2006 and 2010, respectively. He was an OCE postdoc research fellow in the Wireless and Networking Technologies Laboratory (WNTL) at Commonwealth Scientific and Industrial Research Organization (CSIRO), Sydney, Australia. He is now with School of electrical engineering and telecommunications, the University of New South Wales, Australia. He is holding an Australian Research Council (ARC) Discovery Early Career Research Award fellowship. His research expertise and interests include physical-layer network coding, multi-user and MIMO communications, cooperative communications, error-control coding and iterative signal processing. He has published more than 30 research articles in IEEE journals and conferences. He served as the TPC members of IEEE ICC and WCNC. He was the recipient of Australian Postgraduate Award (APA), NICTA research project award (NRPA) and Supplementary Engineering Award (SEA) from the University of New South Wales, and Publication Award from CSIRO ICT centre.