Relay-Aided NOMA in Uplink Cellular Networks

Wonjae Shin, Heecheol Yang, Mojtaba Vaezi, Jungwoo Lee, and H. Vincent Poor

***Abstract*—A new relay-aided non-orthogonal multiple access (NOMA) technique is proposed for multi-cell uplink cellular networks in which each cell supports K single-antenna users by its respective base station (BS) equipped with $N(\ll K)$ antennas. Cooperative relaying transmission is used to accommodate more than one user per orthogonal resource block in the context of interference-limited cellular networks. With the proposed relay-aided NOMA, an Alamouti structure of the desired symbol can further be generated at each BS free of interference, which gives rise to diversity gain of two. The proposed scheme does not require any channel state information (CSI) at users. Further, only limited CSI is required at relays and BSs, which can greatly reduce the control overhead.**

***Index Terms*—Non-orthogonal multiple access (NOMA), relay.**

I. Introduction

Non-orthogonal multiple access (NOMA) is a promising technology for fifth generation wireless networks in which massive numbers of users must be supported with low latency [1]. Existing NOMA techniques mostly exploit code or power domains to allow multiple users to share time/frequency resources in a single cell setting [2]. NOMA can also be applied to interference-limited cellular networks by jointly managing inter-cell and intra-cell interference under the premise of perfect channel knowledge at all nodes [3]–[5]. This can however lead to large training and/or feedback overhead. To avoid these issues, coordinated multipoint (CoMP) transmission can be applied to NOMA schemes [6]. Nevertheless, CoMP imposes a rather heavy burden on backhual due to data sharing among neighboring cells with an extremely low latency.

The role of cooperative relaying has recently been investigated for multi-user interference networks [7]–[9]. Although the use of relay(s) cannot yield a degrees-of-freedom (DoF) gain, it is still useful in practice because the channel state information at transmitters (CSIT) requirement can be relaxed to implement interference alignment (IA) algorithms [9]. In the context of single-cell NOMA, relays have been investigated for various purposes, e.g., user cooperation [10]–[12], reliability [13]–[17], and ergodic sum-capacity [18], [19]. However, there is no previous work considering the use of cooperative relays for interference-limited NOMA systems as a means of increasing the number of concurrently served users. This motivates us to explore the benefits of relays in cellular networks where inter-cell/intra-cell interference severely limits the number of served users.

Our main contribution is to introduce a relay-aided NOMA technique for uplink cellular networks in which the number of antennas at each base station (BS) is limited by the number of users in each cell. We also derive feasibility conditions for the proposed scheme as a function of the relay configuration. Unlike many existing NOMA schemes, we shed light on a robust form of NOMA which does not rely on channel gain disparity among multiplexed users.

Throughout this paper, we use $\mathbf{A}^T$, $\bar{\mathbf{A}}$, and $\mathbf{A}^\dagger$ to indicate the transpose, conjugate, and conjugate transpose of a matrix $\mathbf{A}$, respectively. $\mathbf{0}_{x\times y}$ is used for a zero matrix with size $x \times y$; $\mathbf{I}_n$ represents an identity matrix of size n. In addition, $\mathbb{E}[\cdot]$ and $\mathrm{rk}(\cdot)$ represent the expectation of a random quantity and the rank of the matrix, respectively. The cardinality of a set $\mathcal{A}$ is denoted by $|\mathcal{A}|$, and $\mathcal{B}\setminus\mathcal{A}$ represents the relative complement of the set $\mathcal{A}$ in the set $\mathcal{B}$. $\|\mathbf{x}\|$ is the Euclidean norm of the vector $\mathbf{x}$. $\mathcal{G}$ and $\mathcal{L}$ denote the sets $\{1, 2, \ldots, G\}$ and $\{1, 2, \ldots, L\}$, respectively.

II. System Model

We consider a G cell uplink cellular network, in which each cell consists of K users with L shared relays ($G, K \geq 2$). The relays employ the amplify-and-forward protocol to relay their input signals, and they operate in half-duplex mode. It is assumed that each user and BS is equipped with single and $N(= K/2)$ antennas, respectively[1]; the ℓth relay has M_ℓ antennas where $M_\ell \leq K/2$.[2] When K is odd, a two time-slot symbol extension is additionally required in line with [7].

Let $x^{[k,i]}(t)$ and $\mathbf{x}_{\mathsf{R}}^{\ell}(t)$ denote symbols transmitted at time slot t by user $[k, i]$ and the ℓth relay, respectively, where $[k, i]$ refers to the kth user in cell i. The average transmit power at each node is bounded by P. $\mathbf{h}_{[k,i]}^{\mathsf{DS},j}(t)$ and $\mathbf{h}_{[k,i]}^{\mathsf{RS},\ell}(t)$ are the channel vectors from user $[k, i]$ to BS j and the ℓth relay, respectively; $\mathbf{H}_{[j,\ell]}^{\mathsf{DR}}(t)$ denotes channel matrix from the ℓth relay to BS j. We assume that all the channel coefficients are independent and identically distributed (i.i.d.) complex Gaussian random variables with unit variance. $\mathbf{n}^{[i]}(t)$ and $\mathbf{n}_{\mathsf{R}}^{[\ell]}(t)$ are noise vectors at BS i and the ℓth relay, respectively, all of whose entries are i.i.d. complex Gaussian random variables with zero mean and variance σ^2. At each time slot $t\ (\geq 1)$, the received signals at the BS in cell i

[1]Note that the overloading factor (i.e., the number of multiplexed users per orthogonal resource) is two in multi-cell NOMA systems.

[2]The relays cost less and have smaller form factors than the BS, so the number of RF chains at the relays is limited by that at the BSs.

Manuscript received August 21, 2017; accepted September 11, 2017. Date of publication September 21, 2017; date of current version October 27, 2017. The work of W. Shin, H. Yang, and J. Lee was supported by the Basic Science Research Program (NRF-2017R1A2B2007102) through the National Research Foundation of Korea. The work of H. V. Poor was supported by the U.S. National Science Foundation under Grant CCF-1420575 and Grant ECCS-1647198. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alessio Zappone. *(Corresponding author: Jungwoo Lee.)*

W. Shin, M. Vaezi, and H. V. Poor are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: wonjae.shin@princeton.edu; mvaezi@princeton.edu; poor@princeton.edu).

H. Yang and J. Lee are with the Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, South Korea (e-mail: hee2070@snu.ac.kr; junglee@snu.ac.kr).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2017.2755049

and the ℓth relay are, respectively, given by

$$\mathbf{y}^{[i]}(t) = \sum_{k=1}^{K} \mathbf{h}_{[k,i]}^{\mathsf{DS},i}(t)x^{[k,i]}(t) + \sum_{\substack{j=1\\ j\neq i}}^{G}\sum_{k=1}^{K} \mathbf{h}_{[k,j]}^{\mathsf{DS},i}(t)x^{[k,j]}(t) + \sum_{\ell=1}^{L} \mathbf{H}_{i,\ell}^{\mathsf{DR}}(t)\mathbf{x}_{\mathsf{R}}^{\ell}(t) + \mathbf{n}^{[i]}(t), \quad i \in \mathcal{G} \tag{1}$$

$$\mathbf{y}_{\mathsf{R}}^{[\ell]}(t) = \sum_{j=1}^{G}\sum_{k=1}^{K} \mathbf{h}_{[k,j]}^{\mathsf{RS},\ell}(t)x^{[k,j]}(t) + \mathbf{n}_{\mathsf{R}}^{[\ell]}(t), \quad \ell \in \mathcal{L}. \tag{2}$$

Due to the causality along with the half-duplex constraint, $\mathbf{x}_{\mathsf{R}}^{\ell}(t)$ depends only on the past received signals $\mathbf{y}_{\mathsf{R}}^{[\ell]}(t'), \forall t' < t$; thereby, $\mathbf{x}_{\mathsf{R}}^{\ell}(1) = 0$ for $\forall \ell$. We assume that users do not have any CSIT, whereas each relay and BS has the global CSI and local CSI, respectively.[3]

The per-cell-DoF is defined as the prelog factor of the achievable sum rate in each cell. The individual DoF of user $[k, i]$ and the per-cell DoF are expressed as

$$d^{[k,i]} \triangleq \lim_{\mathsf{SNR}\to\infty} \frac{R^{[k,i]}(\mathsf{SNR})}{\log(\mathsf{SNR})} \text{ and } \mathrm{DoF}_{\mathsf{cell}} = \frac{1}{G}\sum_{\forall k,i} d^{[k,i]} \tag{3}$$

respectively, where the SNR is given by $\frac{P}{\sigma^2}$ and $R^{[k,i]}(\mathsf{SNR})$ denotes the achievable rate of user $[k, i]$. The diversity gain of user $[k, i]$'s message is defined as $r \triangleq -\lim_{\mathsf{SNR}\to\infty} \frac{\log P_e(\mathsf{SNR})}{\log(\mathsf{SNR})}$, where $P_e(\mathsf{SNR})$ denotes the error rate measured by its respective BS for a given SNR.

III. Relay-Aided NOMA for Cellular Networks

A. Motivating Example

In this section, we present the key concept of relay-aided NOMA for cellular networks. For ease of exposition, we first consider a two-cell uplink cellular network with four users in each cell and four relays, in which each BS and relay has two antennas and each user has single antenna ($K = 4, G = N = M_\ell = 2, L = 4, \forall \ell$). A generalization of this method will be presented in the following section.

In this setting, we show that a total of four data symbols can be delivered to each BS over $G + 1 (= 3)$ time slots (i.e., $4/3$ per-cell DoF) while a diversity gain of two is achievable for each symbol. Note that relays are exploited to achieve the maximum DoF that can be achieved for the case when there are no relays and we have full CSI through *asymptotic IA* [20], along with additional diversity gains. To this end, we design the transmitted signal at user $[k, i]$ with no beamforming as

$$x^{[k,i]}(t) = \begin{cases} s^{[k,i]}, & \text{if } t = i \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $s^{[k,i]}$ is the information symbol of user $[k, i]$.

Each relay cannot transmit and receive signals simultaneously due to the half-duplex constraint so that it keeps silent for the first G time slots. Based on the past received signals at the relays, the following signal is sent at the ℓth relay:

$$x_{\mathsf{R}}^{\ell}(t) = \begin{cases} \sum_{i=1}^{G} \mathbf{F}_i^{[\ell]}\mathbf{y}_{\mathsf{R}}^{[\ell]}(i), & \text{if } t = G+1, \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where $\mathbf{F}_i^{[\ell]}$ is the beamforming matrix at relay ℓ associated with the signals of users in cell i, whose size is $M_\ell \times M_\ell$.

The signal received at BS 1 can be expressed as

$$\mathbf{y}^{[1]}(1) = \mathbf{h}_{[1,1]}^{\mathsf{DS},1}(1)s^{[1,1]} + \mathbf{h}_{[2,1]}^{\mathsf{DS},1}(1)s^{[2,1]} + \mathbf{h}_{[3,1]}^{\mathsf{DS},1}(1)s^{[3,1]} + \mathbf{h}_{[4,1]}^{\mathsf{DS},1}(1)s^{[4,1]} + \mathbf{n}^{[1]}(1), \tag{6}$$

$$\mathbf{y}^{[1]}(2) = \mathbf{h}_{[1,2]}^{\mathsf{DS},1}(2)s^{[1,2]} + \mathbf{h}_{[2,2]}^{\mathsf{DS},1}(2)s^{[2,2]} + \mathbf{h}_{[3,2]}^{\mathsf{DS},1}(2)s^{[3,2]} + \mathbf{h}_{[4,2]}^{\mathsf{DS},1}(2)s^{[4,2]} + \mathbf{n}^{[1]}(2), \tag{7}$$

$$\mathbf{y}^{[1]}(3) = \sum_{\ell=1}^{4} \mathbf{H}_{1,\ell}^{\mathsf{DR}}(3)\mathbf{F}_1^{[\ell]}\left[\sum_{k=1}^{4}\mathbf{h}_{[k,1]}^{\mathsf{RS},\ell}(1)s^{[k,1]}\right] + \sum_{\ell=1}^{4} \mathbf{H}_{1,\ell}^{\mathsf{DR}}(3)\mathbf{F}_2^{[\ell]}\left[\sum_{k=1}^{4}\mathbf{h}_{[k,2]}^{\mathsf{RS},\ell}(2)s^{[k,2]}\right] + \widehat{\mathbf{n}}^{[i]}(3) \tag{8}$$

where $\widehat{\mathbf{n}}^{[i]}(t) = \sum_{\ell=1}^{4}\mathbf{H}_{i,\ell}^{\mathsf{DR}}(t)\sum_{j=1}^{G}\mathbf{F}_j^{[\ell]}\mathbf{n}_{\mathsf{R}}^{[\ell]}(j) + \mathbf{n}^{[i]}(t)$ is the equivalent noise in the second hop of time slot t.

Let us choose the relay beamforming matrix $\mathbf{F}_j^{[\ell]}$ so that the set of inter-cell interference (ICI) signals are aligned while making effective channel for all desired symbols orthogonal (e.g., an Alamouti structure [21], [22]) as follows:

1) IA for interfering symbols (cell 2 → cell 1):

$$c_{1,2}^{[1]}\sum_{\ell=1}^{4}\mathbf{H}_{1,\ell}^{\mathsf{DR}}(3)\mathbf{F}_2^{[\ell]}\mathbf{h}_{[k,2]}^{\mathsf{RS},\ell}(2) = \mathbf{h}_{[k,2]}^{\mathsf{DS},1} \quad \forall k \tag{9}$$

2) Signal shaping for desired symbols (cell 1 → cell 1):

$$c_{1,1}^{[1]}\sum_{\ell=1}^{4}\mathbf{H}_{1,\ell}^{\mathsf{DR}}(3)\mathbf{F}_1^{[\ell]}\mathbf{h}_{[k,1]}^{\mathsf{RS},\ell}(1) = (-1)^{\bar{k}}\bar{\mathbf{h}}_{[\bar{k},1]}^{\mathsf{DS},1} \quad \forall k \tag{10}$$

where $\bar{k} = k + (-1)^{k+1}$, and $c_{k,j}^{[i]}$ is an arbitrary complex value which can be chosen independently of SNR while satisfying the power constraint, thereby not affecting the DoF and diversity gains.

Provided that the beamforming matrix exists, we can further convert the set of received signals into a vector form as[4]

$$\begin{aligned}
&\left[\mathbf{y}^{[1]}(1)^T,\ \mathbf{y}^{[1]}(2)^T,\ \mathbf{y}^{[1]}(3)^T\operatorname{diag}(\mathbf{c}_1^{[1]})^{-1}\right]^T \\
&= \left[\begin{array}{cc|cc}
\mathbf{h}_{[1,1]}^{\mathsf{DS},1}(1) & \mathbf{h}_{[2,1]}^{\mathsf{DS},1}(1) & \mathbf{h}_{[3,1]}^{\mathsf{DS},1}(1) & \mathbf{h}_{[4,1]}^{\mathsf{DS},1}(1) \\
\mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} \\
\bar{\mathbf{h}}_{[2,1]}^{\mathsf{DS},1}(1) & -\bar{\mathbf{h}}_{[1,1]}^{\mathsf{DS},1}(1) & \bar{\mathbf{h}}_{[4,1]}^{\mathsf{DS},1}(1) & -\bar{\mathbf{h}}_{[3,1]}^{\mathsf{DS},1}(1)
\end{array}\right]\mathbf{s}^{[1]} \\
&+ \left[\begin{array}{cc|cc}
\mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} \\
\mathbf{h}_{[1,2]}^{\mathsf{DS},1}(2) & \mathbf{h}_{[2,2]}^{\mathsf{DS},1}(2) & \mathbf{h}_{[3,2]}^{\mathsf{DS},1}(2) & \mathbf{h}_{[4,2]}^{\mathsf{DS},1}(2) \\
\mathbf{h}_{[1,2]}^{\mathsf{DS},1}(2) & \mathbf{h}_{[2,2]}^{\mathsf{DS},1}(2) & \mathbf{h}_{[3,2]}^{\mathsf{DS},1}(2) & \mathbf{h}_{[4,2]}^{\mathsf{DS},1}(2)
\end{array}\right]\mathbf{s}^{[2]} \\
&+ \left[\mathbf{n}^{[1]}(1)^T,\ \mathbf{n}^{[1]}(2)^T,\ \widehat{\mathbf{n}}^{[1]}(3)^T\operatorname{diag}(\mathbf{c}_1^{[1]})^{-1}\right]^T
\end{aligned} \tag{11}$$

where $\mathbf{s}^{[i]} \triangleq \left[s^{[1,i]},\ s^{[2,i]}, \ldots, s^{[K,i]}\right]^T$ and $\mathbf{c}_1^{[1]} \triangleq [c_{1,1}^{[1]},\ c_{1,2}^{[1]}]^T$.

[3] A low-cost user terminal has severe limitations on signaling overhead in contrast with (fixed) multi-antenna relays, which makes this assumption reasonable and also relevant in practice.

[4] The beamforming matrix should be designed not only for BS 1 but also BS 2 at the same time; the feasibility of such beamforming design will be demonstrated in Section IV in greater detail.

Since the equivalent channels for inter-cell interference are aligned, the aligned interference can be efficiently canceled by

$$\left[\mathbf{y}^{[1]}(1)^T,\ \mathbf{y}^{[1]}(3)^T\operatorname{diag}(\mathbf{c}_1^{[1]})^{-1}-\mathbf{y}^{[1]}(2)^T\right]^T \triangleq \hat{\mathbf{y}}^{[1]}$$
$$=\begin{bmatrix}\mathbf{h}^{\mathsf{DS},1}_{[1,1]}(1) & \mathbf{h}^{\mathsf{DS},1}_{[2,1]}(1) & \mathbf{h}^{\mathsf{DS},1}_{[3,1]}(1) & \mathbf{h}^{\mathsf{DS},1}_{[4,1]}(1)\\ \bar{\mathbf{h}}^{\mathsf{DS},1}_{[2,1]}(1) & -\bar{\mathbf{h}}^{\mathsf{DS},1}_{[1,1]}(1) & \bar{\mathbf{h}}^{\mathsf{DS},1}_{[4,1]}(1) & -\bar{\mathbf{h}}^{\mathsf{DS},1}_{[3,1]}(1)\end{bmatrix}\mathbf{s}^{[1]}$$
$$+\underbrace{\left[\mathbf{n}^{[1]}(1)^T,\ \hat{\mathbf{n}}^{[1]}(3)^T\operatorname{diag}(\mathbf{c}_1^{[1]})^{-1}-\mathbf{n}^{[1]}(2)^T\right]^T}_{=\hat{\mathbf{n}}^{[1]}}. \quad (12)$$

Furthermore, we introduce a new matrix $\widehat{\mathbf{H}}^{\mathsf{DS},1}_{i,j}$ having the Alamouti structure defined as

$$\widehat{\mathbf{H}}^{\mathsf{DS},1}_{i,j} \triangleq \begin{bmatrix} h_a & h_b \\ h_b^* & -h_a^*\end{bmatrix} \quad (13)$$

where h_a and h_b are the ith component of $\mathbf{h}^{\mathsf{DS},1}_{[2\times j-1,1]}(1)$ and $\mathbf{h}^{\mathsf{DS},1}_{[2\times j,1]}(1)$, respectively. By performing elementary row operations in (12), we have

$$\mathbf{P}_{\Pi}\times\hat{\mathbf{y}}^{[1]}=\begin{bmatrix}\hat{\mathbf{y}}^{[1]}_{\Pi,1}\\ \hat{\mathbf{y}}^{[1]}_{\Pi,2}\end{bmatrix}=\begin{bmatrix}\widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,1} & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,2}\\ \widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,1} & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,2}\end{bmatrix}\mathbf{s}^{[1]}+\hat{\mathbf{n}}^{[1]}_{\Pi} \quad (14)$$

where $\mathbf{P}_{\Pi}$ is a permutation matrix, which is

$$\mathbf{P}_{\Pi}=\left[\mathbf{e}^T_{\Pi(1)}\ \mathbf{e}^T_{\Pi(2)}\ \cdots\ \mathbf{e}^T_{\Pi(2N)}\right]^T \quad (15)$$

with $\mathbf{e}_j$ denoting a row vector of length $2N$ with 1 in the jth position and 0 in every other position; $\Pi(i)=\lceil\frac{i}{2}\rceil$ for odd i and $\Pi(i)=N+\frac{i}{2}$ for even i; and $\hat{\mathbf{n}}^{[1]}_{\Pi}=\mathbf{P}_{\Pi}\hat{\mathbf{n}}^{[1]}$.

To extract individual symbols with diversity gain at BS 1, we apply a two-step decoding procedure to cancel the interuser interference (IUI) within a cell while maintaining the orthogonality of effective channels for desired symbols as follows:

Step 1) IUI cancellation

Let us first focus on how to resolve $s_1^{[1]}$ and $s_2^{[1]}$ from $\hat{\mathbf{y}}^{[1]}_{\Pi,1}$ and $\hat{\mathbf{y}}^{[1]}_{\Pi,2}$, all of which are linear combinations of all the symbols sent by users in cell 1. We can completely remove the interfering symbols $s_3^{[1]}$ and $s_4^{[1]}$ by the following operation:

$$\hat{\mathbf{y}}^{[1]'}=\frac{\widehat{\mathbf{H}}^{\mathsf{DS},1\dagger}_{1,2}}{\|\widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,2}\|^2}\hat{\mathbf{y}}^{[1]}_{\Pi,1}-\frac{\widehat{\mathbf{H}}^{\mathsf{DS},1\dagger}_{2,2}}{\|\widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,2}\|^2}\hat{\mathbf{y}}^{[1]}_{\Pi,2}$$
$$=\left(\frac{\widehat{\mathbf{H}}^{\mathsf{DS},1\dagger}_{1,2}\widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,1}}{\|\widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,2}\|^2}-\frac{\widehat{\mathbf{H}}^{\mathsf{DS},1\dagger}_{2,2}\widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,1}}{\|\widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,2}\|^2}\right)\begin{bmatrix}s_1^{[1]}\\ s_2^{[1]}\end{bmatrix}+\hat{\mathbf{n}}^{[1]'}_{\Pi}$$
$$=\widehat{\mathbf{G}}^{\mathsf{DS},1}\left[s_1^{[1]}\ s_2^{[1]}\right]^T+\hat{\mathbf{n}}^{[1]'}_{\Pi} \quad (16)$$

where $\widehat{\mathbf{G}}^{\mathsf{DS},1}$ still has the Alamouti structure (i.e., orthogonal) due to the completeness of the Alamouti matrix under addition and multiplication [21].

Step 2) Decoupling two symbols by Maximal ratio combining (MRC)

MRC is applied to further achieve the diversity gain, i.e.,

$$\begin{bmatrix}\hat{y}_1^{[1]'}\\ \hat{y}_2^{[1]'}\end{bmatrix}=\widehat{\mathbf{G}}^{\mathsf{DS},1\dagger}\hat{\mathbf{y}}^{[1]'}=(|\hat{\mathbf{g}}^{\mathsf{DS},1}_{1,1}|^2+|\hat{\mathbf{g}}^{\mathsf{DS},1}_{1,2}|^2)\begin{bmatrix}s_1^{[1]}\\ s_2^{[1]}\end{bmatrix}+\hat{\mathbf{n}}^{[1]''}_{\Pi}$$

where $\hat{\mathbf{g}}^{\mathsf{DS},1}_{m,n}$ denotes the (m,n)th element of $\widehat{\mathbf{G}}^{\mathsf{DS},1}$; and $\hat{\mathbf{n}}^{[1]'}_{\Pi}$ and $\hat{\mathbf{n}}^{[1]''}_{\Pi}$ are the equivalent noise terms after applying the linear operations in step 1 and step 2, respectively.

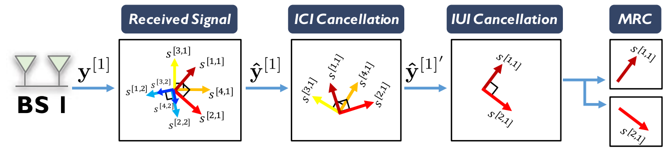

Fig. 1. Illustration of receiver operations for relay-aided NOMA.

As a result, $s_1^{[1]}$ and $s_2^{[1]}$ can be decoded by

$$s_k^{[1]}=\arg\max_{s_k^{[1]}\in\mathcal{S}}\left|\hat{y}_k^{[1]'}-\left(|\hat{\mathbf{g}}^{\mathsf{DS},1}_{1,1}|^2+|\hat{\mathbf{g}}^{\mathsf{DS},1}_{1,2}|^2\right)s_k^{[1]}\right|,\ k\in\{1,2\}$$

where $\mathcal{S}$ denotes the set of constellation points.

By symmetry, a similar argument can be used to demonstrate that each user achieves $1/3$ DoF in both cell 1 and cell 2 with a diversity gain of two, and thus $4/3$ per-cell DoF can be achieved with the diversity benefit. Note that the decoding procedures require knowledge of only incoming channel vectors associated with itself (i.e., local CSI), not global CSI. The procedure at BS 1 are illustrated in Fig. 1.

B. Extension to General Uplink Cellular Networks

With the understanding of the special case in the previous section, let us generalize relay-aided NOMA to arbitrary system configurations. Following the transmission strategies as shown in (4) and (5), the received signals at BS i are given by

$$\mathbf{y}^{[i]}(j)=\sum_{k=1}^{K}\mathbf{h}^{\mathsf{DS},i}_{[k,j]}(j)s^{[k,j]}+\mathbf{n}^{[i]}(j),\ j\in\{1,\dots,G\}$$
$$\mathbf{y}^{[i]}(G+1)=\sum_{\ell\in\mathcal{L}}\mathbf{H}^{\mathsf{DR}}_{i,\ell}(G+1)\sum_{j=1}^{G}\mathbf{F}_j^{[\ell]}\left[\sum_{k=1}^{K}\mathbf{h}^{\mathsf{RS},\ell}_{[k,j]}(j)s^{[k,j]}\right]$$
$$+\hat{\mathbf{n}}^{[i]}(G+1) \quad (17)$$

where $\hat{\mathbf{n}}^{[i]}(G+1)=\sum_{\ell\in\mathcal{L}}\mathbf{h}^{\mathsf{DR}}_{i,\ell}(G+1)\sum_{j=1}^{G}\mathbf{F}_j^{[\ell]}\mathbf{n}_R^{[\ell]}(j)+\mathbf{n}^{[i]}(G+1)$ is the equivalent noise in the second hop of the time slot. To simplify notation, we write the channel matrices as $\mathbf{H}^{\mathsf{RS}}_{\ell,j}(j)\triangleq\left[\mathbf{h}^{\mathsf{RS},\ell}_{[1,j]}(j)\cdots\mathbf{h}^{\mathsf{RS},\ell}_{[K,j]}(j)\right]$ and $\mathbf{H}^{\mathsf{DS}}_{i,j}(j)\triangleq\left[\mathbf{h}^{\mathsf{DS},i}_{[1,j]}(j)\cdots\mathbf{h}^{\mathsf{DS},i}_{[K,j]}(j)\right]$. We further drop time index of $\mathbf{H}^{\mathsf{DR}}_{i,\ell}(t)$ for simplicity.

Similar to the previous example, we design relay beamforming matrices such that two types of conditions are satisfied. Using a property of the Kronecker product [25], i.e., $\operatorname{vec}(\mathbf{AXB})=(\mathbf{B}^T\otimes\mathbf{A})\operatorname{vec}(\mathbf{X})$ where $\otimes$ denotes the Kronecker product between two matrices, we have

1) IA for interfering symbols (cell $j\to$ cell i):

$$\sum_{\ell\in\mathcal{L}}\mathbf{H}^{\mathsf{DR}}_{i,\ell}\mathbf{F}_j^{[\ell]}\mathbf{H}^{\mathsf{RS}}_{\ell,j}(j)=\operatorname{diag}(\mathbf{c}_j^{[i]})\mathbf{H}^{\mathsf{DS}}_{i,j}(j)\ \ \forall i,j(\neq i)$$
$$\Rightarrow\left[\mathbf{H}^{\mathsf{RS}}_{1,j}(j)^T\otimes\mathbf{H}^{\mathsf{DR}}_{i,1},\ \mathbf{H}^{\mathsf{RS}}_{2,j}(j)^T\otimes\mathbf{H}^{\mathsf{DR}}_{i,2},\cdots,\mathbf{H}^{\mathsf{RS}}_{L,j}(j)^T\otimes\mathbf{H}^{\mathsf{DR}}_{i,L}\right]$$
$$\times\begin{bmatrix}\operatorname{vec}(\mathbf{F}_j^{[1]})\\ \operatorname{vec}(\mathbf{F}_j^{[2]})\\ \vdots\\ \operatorname{vec}(\mathbf{F}_j^{[L]})\end{bmatrix}=\left(\mathbf{H}^{\mathsf{DS}}_{i,j}(j)^T\otimes\mathbf{I}_K\right)\operatorname{vec}\left(\operatorname{diag}(\mathbf{c}_j^{[i]})\right).$$

2) Signal shaping for desired symbols (cell $i\to$ cell i):

$$\sum_{\ell\in\mathcal{L}}\mathbf{H}^{\mathsf{DR}}_{i,\ell}\mathbf{F}_i^{[\ell]}\mathbf{H}^{\mathsf{RS}}_{\ell,i}(i)=\operatorname{diag}(\mathbf{c}_i^{[i]})\hat{\mathbf{H}}^{\mathsf{DS}}_{i,i}(i)\quad\forall i \quad (18)$$

$$\Rightarrow \left[\mathbf{H}^{\mathsf{RS}}_{1,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{i,1},\ \mathbf{H}^{\mathsf{RS}}_{2,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{i,2} \cdots\ \mathbf{H}^{\mathsf{RS}}_{L,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{i,L}\right] \times \begin{bmatrix} \mathrm{vec}(\mathbf{F}^{[1]}_i) \\ \mathrm{vec}(\mathbf{F}^{[2]}_i) \\ \vdots \\ \mathrm{vec}(\mathbf{F}^{[L]}_i) \end{bmatrix} = \left(\hat{\mathbf{H}}^{\mathsf{DS}}_{i,i}(i)^T \otimes \mathbf{I}_K\right) \mathrm{vec}\left(\mathrm{diag}(\mathbf{c}^{[i]}_i)\right)$$

where $\hat{\mathbf{H}}^{\mathsf{DS}}_{i,i}(i) = \bar{\mathbf{H}}^{\mathsf{DS}}_{i,i}(i) \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Now, consider a unified system of equations by aggregating all signal shaping and IA conditions for every receiver to jointly construct the relay beamforming matrices, which is given by

$$\underbrace{\begin{bmatrix} \mathbf{H}^{\mathsf{RS}}_{1,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{i,1} & \cdots & \mathbf{H}^{\mathsf{RS}}_{L,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{i,L} \\ \mathbf{H}^{\mathsf{RS}}_{1,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{\bar{i}(1),1} & \cdots & \mathbf{H}^{\mathsf{RS}}_{L,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{\bar{i}(1),L} \\ \vdots & & \vdots \\ \mathbf{H}^{\mathsf{RS}}_{1,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{\bar{i}(G-1),1} & \cdots & \mathbf{H}^{\mathsf{RS}}_{L,i}(i)^T \otimes \mathbf{H}^{\mathsf{DR}}_{\bar{i}(G-1),L} \end{bmatrix}}_{GNK \times \sum_{\ell\in\mathcal{L}} M_\ell^2}$$

$$\times \begin{bmatrix} \mathrm{vec}(\mathbf{F}^{[1]}_i) \\ \mathrm{vec}(\mathbf{F}^{[2]}_i) \\ \vdots \\ \mathrm{vec}(\mathbf{F}^{[L]}_i) \end{bmatrix} = \begin{bmatrix} \left(\hat{\mathbf{H}}^{\mathsf{DS}}_{i,i}(i)^T \otimes \mathbf{I}_K\right) \mathrm{vec}\left(\mathrm{diag}(\mathbf{c}^{[i]}_i)\right) \\ \left(\mathbf{H}^{\mathsf{DS}}_{\bar{i}(1),i}(i)^T \otimes \mathbf{I}_K\right) \mathrm{vec}\left(\mathrm{diag}(\mathbf{c}^{[\bar{i}(1)]}_i)\right) \\ \vdots \\ \left(\mathbf{H}^{\mathsf{DS}}_{\bar{i}(G-1),i}(i)^T \otimes \mathbf{I}_K\right) \mathrm{vec}\left(\mathrm{diag}(\mathbf{c}^{[\bar{i}(G-1)]}_i)\right) \end{bmatrix}$$

$$\Rightarrow \mathbf{H}_i \times \mathbf{f}_i = \mathbf{c}_i\ \forall i \in \{1,2,\ldots,G\} \tag{19}$$

where $\bar{i}(j) = \mathrm{mod}(i+j-1, G) + 1$ and $j \in \{1,\ldots,G-1\}$.

Suppose there exists $\mathbf{F}^{[\ell]}_i, \forall \ell$ such that (19) is satisfied due to there being enough relay antennas.[5] We then focus on the decoding process at BS 1. Provided that (19) has a solution, the aligned inter-cell interference can be suppressed by

$$\widehat{\mathbf{y}}^{[1]} = \left[\mathbf{y}^{[1]}(1)^T,\ \mathbf{y}^{[1]}(G+1)^T - \textstyle\sum_{j=2}^{G} \mathbf{y}^{[1]}(j)^T \mathrm{diag}(\mathbf{c}^{[1]}_j)\right]^T$$

which can be recast into

$$\underbrace{\begin{bmatrix} \widehat{\mathbf{y}}^{[1]}_{\Pi,1} \\ \widehat{\mathbf{y}}^{[1]}_{\Pi,2} \\ \vdots \\ \widehat{\mathbf{y}}^{[1]}_{\Pi,\frac{K}{2}} \end{bmatrix}}_{=\mathbf{P}_\Pi \times \widehat{\mathbf{y}}^{[1]}} = \underbrace{\begin{bmatrix} \widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,1} & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,2} & \cdots & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{1,\frac{K}{2}} \\ \widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,1} & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,2} & \ddots & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{2,\frac{K}{2}} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\mathbf{H}}^{\mathsf{DS},1}_{\frac{K}{2},1} & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{\frac{K}{2},2} & \cdots & \widehat{\mathbf{H}}^{\mathsf{DS},1}_{\frac{K}{2},\frac{K}{2}} \end{bmatrix}}_{\triangleq \mathcal{H}_1} \mathbf{s}^{[1]} + \widehat{\mathbf{n}}^{[1]}_\Pi$$

where $\mathbf{P}_\Pi$ follows the previous definition as in (15).

In order to extract the kth user's symbol, the aligned intracell interference except for one neighboring user, the $(k+(-1)^{k+1})$th user can be further removed by the generalized zero-forcing approach as in [23] as follows:

$$\left(\mathbf{I}_{2N} - \mathsf{Proj}_{\mathcal{H}_1^{(k')}}\right) \begin{bmatrix} \widehat{\mathbf{y}}^{[1]}_{\Pi,1} \\ \widehat{\mathbf{y}}^{[1]}_{\Pi,2} \\ \vdots \\ \widehat{\mathbf{y}}^{[1]}_{\Pi,\frac{K}{2}} \end{bmatrix} = \widehat{\mathbf{G}}^{\mathsf{DS},k'} \left[s^{[1]}_{2k'-1}\ s^{[1]}_{2k'}\right]^T + \widehat{\mathbf{n}}^{[1]'}_\Pi$$

where $k' = \lceil \frac{k}{2} \rceil$ and $\widehat{\mathbf{G}}^{\mathsf{DS},k'} \triangleq (\mathbf{I}_{2N} - \mathsf{Proj}_{\mathcal{H}_1^{(k')}})\mathcal{H}_1$; $\mathsf{Proj}_{\mathcal{H}_1^{(k)}}$ denotes a projection matrix that projects into the subspace spanned by the columns of $\mathcal{H}_1$ except the kth column. Note that due to the completeness of Alamouti matrices under addition and multiplication, the equivalent channel matrix, $\widehat{\mathbf{G}}^{\mathsf{DS},k'}$, has an Alamouti structure. This allows us to decouple two neighboring users' symbols by simply applying the MRC.

With these decoding steps, all the K users' symbols over $G+1$ time slots with a diversity gain of two can be delivered at BS 1 with only $K/2$ antennas at the BS. By symmetry, the other $G-1$ BSs can also achieve the same performance gains as BS 1, which results in $\frac{K}{G+1}$ per-cell DoF. Notice that if users have no channel knowledge, the number of served users in the network without relays is limited by the number of BS antennas [24], thereby achieving $\frac{K}{2G}$ per-cell DoF. That is, relay-aided NOMA approximately doubles per-cell DoF for large G.

Remark 1: For two-cell cellular networks having K users per cell and $\frac{K}{2}$ antennas at each BS, relay-aided NOMA can achieve the maximum per-cell DoF of $\frac{K}{3}$ that can be achieved by asymptotic IA even with full CSIT using no relays [20], along with the diversity gain. Note that this does not require infinite symbol extension over time/frequency or full CSI at nodes, which facilitates IA with the help of relays in practice.

IV. Analysis of a Feasibility Condition

We establish a feasibility condition on the numbers of relay antennas needed to realize NOMA in interference-limited cellular networks. The feasibility of relay-aided NOMA relies only on the solvability of the sets of linear (19), which depends on the rank of $\mathbf{H}_i$ as a function of the relay configuration. To prove the feasibility condition, we first introduce the following lemma.

Lemma 1: Suppose $\mathbf{A}$ and $\mathbf{B}_i$ are, respectively, as $m \times n$ and $p \times n$ random matrices whose entries are i.i.d. complex Gaussian. Then, we have almost surely

$$\mathrm{rk}\left(\begin{bmatrix} \mathbf{A} \otimes \mathbf{B}_1 \\ \vdots \\ \mathbf{A} \otimes \mathbf{B}_G \end{bmatrix}\right) = n^2 \tag{20}$$

where $\min\{m,p\} \geq n$ and $G \geq 2$.

Proof: Using the properties of the Kronecker product, the left-hand side of (20) can be rewritten as

$$\mathrm{rk}\,(\mathbf{A}) \cdot \mathrm{rk}\left(\left[\mathbf{B}_1^T, \cdots, \mathbf{B}_G^T\right]\right) = \min\{m,n\} \cdot \min\{Gp, n\} \overset{(a)}{=} n^2$$

where (a) follows from the assumption of this lemma.

Theorem 1: For a G-cell uplink cellular network with K single-antenna users and each BS equipped with $K/2$ antennas, relay-aided NOMA achieves $\frac{K}{G+1}$ per-cell DoF and a diversity gain of two when $\sum_{\ell\in\mathcal{L}} M_\ell^2 \geq \frac{1}{2}GK^2$.

Proof: Suppose $\mathbf{H}_i = \left[\mathbf{H}^{[1]}_i, \mathbf{H}^{[2]}_i, \ldots, \mathbf{H}^{[L]}_i\right]$, where $\mathbf{H}^{[\ell]}_i$ is the ℓth submatrix whose size is $GK^2/2 \times M_\ell^2$. According to Lemma 1, the rank of the submatrix is almost surely the number of columns it has, which implies a full *column* rank matrix. Considering the independence across submatrices,[6] it can be argued that the matrix $\mathbf{H}_i$ is a full *row* rank matrix as long as the number of columns is greater than or equal to the number of rows of a set of submatrices, $\sum_{\ell\in\mathcal{L}} M_\ell^2 \geq \frac{1}{2}GK^2$.

[5]Feasibility conditions will be provided in the following section.

[6]All the elements of the submatrix $\mathbf{H}^{[\ell]}_i$ are drawn independently from those of the other submatrices $\mathbf{H}^{[j]}_i, j \in \mathcal{L}\backslash\{\ell\}$.

REFERENCES

[1] V. W. Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, *Key Technologies for 5G Wireless Systems.* Cambridge U.K.: Cambridge Univ. Press, 2017.

[2] L. Dai, B. Wang, Y. Yuan, S. Han, C.-L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, 2015.

[3] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, "Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 176–183, 2017.

[4] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, "Coordinated beamforming for multi-cell MIMO-NOMA," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 84–87, 2017.

[5] W. Shin, M. Vaezi, J. Lee, and H. V. Poor, "On the number of users served in MIMO-NOMA cellular networks," in *Proc. Int. Symp. Wireless Commun. Syst.*, Poznan, Poland, 2016, pp. 638–642.

[6] J. Choi, "Non-orthogonal multiple access in downlink coordinated two-point systems," *IEEE Commun. Lett.*, vol. 18, no. 2, pp. 313–316, 2014.

[7] V. R. Cadambe and S. Jafar, "Interference alignment and the degrees of freedom of the K-User interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, 2008.

[8] V. R. Cadambe and S. Jafar, "Interference alignment and the degrees of freedom of wireless X networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3893–3908, 2009.

[9] V. R. Cadambe and S. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation, and full duplex operation," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2334–2344, 2009.

[10] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, 2015.

[11] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, 2016.

[12] Z. Zhang, Z. Ma, M. Xiao, Z. Ding, and P. Fan, "Full-duplex device-to-device-aided cooperative nonorthogonal multiple access," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4467–4471, 2017.

[13] Z. Ding, H. Dai, and H. V. Poor, "Relay selection for cooperative NOMA," *IEEE Wireless Commun. Lett.*, vol. 5, no. 4, pp. 416–419, 2016.

[14] J.-B. Kim and I.-H. Lee, "Non-orthogonal multiple access in coordinated direct and relay transmission," *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 2037–2040, 2015.

[15] C. Zhong and Z. Zhang, "Non-orthogonal multiple access with cooperative full-duplex relaying," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2478–2481, 2016.

[16] X. Yue, Y. Liu, S. Kang, and A. Nallanathan, "Performance analysis of NOMA with fixed gain relaying over Nakagami-m fading channels," *IEEE Access*, vol. 5, pp. 5445–5454, 2017.

[17] X. Liang, Y. Wu, D. W. K. Ng, Y. Zuo, S. Jin, and H. Zhu, "Outage performance for cooperative NOMA transmission with an AF relay," *IEEE Commun. Lett.*, vol. PP, no. 99, pp. 1–1, 2017.

[18] M. F. Kader, M. B. Shahab, and S. Y. Shin, "Exploiting non-orthogonal multiple access in cooperative relay sharing," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1159–1162, 2017.

[19] C. Zhong and Z. Zhang, "Non-orthogonal multiple access with cooperative full-duplex relaying," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2478–2481, 2016.

[20] G. Sridharan and W. Yu, "Degrees of freedom of MIMO cellular networks: decomposition and linear beamforming design," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3339–3364, 2015.

[21] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, 1998.

[22] L. Li and H. Jafarkhani, "Maximum-rate transmission with improved diversity gain for interference networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5313–5330, 2013.

[23] L. Li and H. Jafarkhani, "Multi-user downlink system design with bright transmitters and blind receivers," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 4074–4084, 2012.

[24] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5354–5374, 2012.

[25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.