Received 8 June 2021; accepted 23 June 2021. Date of publication 30 June 2021; date of current version 13 July 2021.
Digital Object Identifier 10.1109/OJCOMS.2021.3092815

Rate Splitting With Finite Constellations: The Benefits of Interference Exploitation vs Suppression

ABDELHAMID SALEM [1] (Member, IEEE), CHRISTOS MASOUROS [1] (Senior Member, IEEE), AND BRUNO CLERCKX [2] (Senior Member, IEEE)

[1] Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, U.K.
[2] Electrical and Electronic Engineering Department, Imperial College London, London SW7 2AZ, U.K.

CORRESPONDING AUTHOR: A. SALEM (e-mail: a.salem@ucl.ac.uk)

This work was supported in part by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/R007934/1, Grant EP/N015312/1, and Grant EP/R511547/1. Parts of this paper were presented at the WCNC 2019 [1].

ABSTRACT Rate-Splitting (RS) has been shown recently to be a powerful approach for the design of non-orthogonal transmission, multiple access, and interference management strategies in multi-user multi-antenna systems. RS, through the split of messages into common and private parts, relies on the transmission of common streams decoded by all users, and private streams decoded only by its intended user. This enables RS to bridge the extreme of fully decode interference and fully treat interference as noise. In this paper, we depart from Gaussian signaling and study RS under finite input alphabet for multi-user multi-antenna system and propose a constructive interference (CI) exploitation approach to further enhance the sum-rate achieved by RS. To that end, new analytical expressions for the ergodic sum-rate are derived for two precoding techniques of the private messages, namely, 1) a traditional interference suppression zero-forcing (ZF) precoding approach, 2) a closed-form CI precoding approach. Our analysis is presented for perfect channel state information at the transmitter (CSIT), and is extended to imperfect CSIT knowledge. A novel power allocation strategy, specifically suited for the finite alphabet setup, is derived and shown to lead to superior performance for RS over conventional linear precoding not relying on RS (NoRS). The results in this work validate the significant sum-rate gain of RS with CI over the conventional RS with ZF and NoRS.

INDEX TERMS Rate splitting, zero forcing, constructive interference, phase-shift keying signaling.

I. INTRODUCTION

The recent years have witnessed the widespread application of multi-user multiple-input single-output (MU-MISO) systems, due to their reliability and high spectral efficiency [2], [3]. However, in practical communication networks, the advantages of MU-MISO systems are often impacted by interference [2], [3]. Consequently, a considerable amount of researches has focused on improving the performance of MU-MISO systems [3], [4]. In this regard, Rate-Splitting (RS) approach was recently proposed and investigated in different scenarios to enhance the performance of MU-MISO systems [5]–[9]. RS scheme splits the users' messages into a common message and private messages, and superimposes the common message on top of the private messages. Using Successive Interference Cancellation (SIC) at the receivers, the common message is first decoded by all the users, and each private message is then decoded only by its intended user. By adjusting the message split and the power allocated to the common and private messages, RS has the ability to better handle the multiuser interference. RS has been studied in multiuser multi-antenna setups with both perfect and imperfect CSIT. In [7], the authors analyzed the sum-rate gain achieved by RS over conventional multi-user linear precoding without using RS (NoRS),[1] in a two-user multi-antenna broadcast channel with imperfect CSIT, and considered that the common message is transmitted via a space and space-time design.

1. Hereafter, all the conventional techniques that do not split messages are denoted by NoRS to contrast with RS.

In [6]–[9], again considering imperfect CSIT, the authors leveraged convex optimization to optimize the precoders of the common and private messages to maximize the ergodic sum-rate and the max-min rate, respectively, and again showed the superiority of RS over NoRS. In [10], RS was designed and its performance analyzed for massive multiple-input multiple-output (MIMO) systems with imperfect CSIT and shown to outperform the conventional NoRS approach. In [11], RS was designed for a multi-antenna multi-cell system with imperfect CSIT, and showed the superiority in a Degrees-of-Freedom sense over NoRS. The benefits of RS have also been highlighted in multiuser multi-antenna system in [12]–[15], and the performance gains were highlighted over both NoRS and power-domain Non-Orthogonal Multiple Access (NOMA) techniques. Particularly, NOMA (and conventional SDMA/multi-user MISO) is shown to be a particular subset of RS, and is therefore outperformed by RS [12]–[15].

Another line of research has recently proposed constructive interference (CI) precoding techniques to enhance the performance of downlink MU-MISO systems [16]–[20]. In contrast to the conventional interference mitigation techniques, where the knowledge of the interference is used to cancel it, the main idea of the CI is to use the interference to improve the system performance. The CI precoding technique exploits interference that can be known to the transmitter to increase the useful signal received power [16]–[18]. That is, with the knowledge of the CSI and users' data symbols, the interference can be classified as constructive and destructive. The interference signal is considered to be constructive to the transmitted signal if it pushes/moves the received symbols away from the decision thresholds of the constellation towards the direction of the desired symbol. Accordingly, the transmit precoding can be designed such that the resulting interference is constructive to the desired symbol. The concept of CI has been extensively studied in literature. This line of work has been introduced in [16], where the CI precoding scheme for the downlink of PSK-based MIMO systems has been proposed. In this work, it was shown that the effective signal to interference-plus-noise ratio (SINR) can be enhanced without the need to increase the transmitted signal power at the base station (BS). In [17], an optimization-based precoder in the form of pre-scaling has been designed for the first time using the concept of CI. Thereof, [18] proposed transmit beamforming schemes for the MU-MISO downlink that minimize the transmit power for generic PSK signals. In [21], a transmission algorithm that exploits the constructive multiuser interference was proposed. In [22], [23], CI precoding scheme was applied in wireless power transfer scenario in order to minimize the transmit power. Further works in [24], [25] applied the CI concept to massive MIMO and cognitive radio networks. Several block-level and symbol-level precoding techniques for MU-MISO systems have been presented in [26]–[28]. Very recently, the authors in [29] derived closed-form precoding expression for CI exploitation in the MU-MISO downlink. The closed-form precoder in this work has for the first time made the application of CI exploitation practical, and has further paved the way for the development of communication theoretic analysis of the benefits of CI [30]–[32]. Based on this precoding formula, a new closed form expression of the ergodic sum-rate for the CI precoding without using RS has been presented in [33]. The authors in [34] provided a tutorial on CI exploitation schemes from the perspective of precoding design. While the above literature has addressed traditional downlink transmission, the application of the CI concept to RS approaches remains an open problem, due to the finite constellation input that CI requires.

Accordingly, in this paper, we provide the first attempt to combine those two lines of research on RS and CI, and employ the CI precoding technique to further enhance the sum-rate achieved by RS scheme in MU-MISO systems under finite input alphabet.[2] In this regard and in order to provide fair comparison, new analytical expressions for the ergodic sum-rate are derived for two precoding techniques of the private messages, namely, 1) a traditional interference suppression zero-forcing (ZF) precoding approach, 2) a closed-form CI precoding approach. Thus we provide a comparison between two techniques to deal with the interference, suppression and exploitation.[3] Our analysis is presented for perfect channel state information at the BS (CSIT), and extended to imperfect CSIT. Additionally, since RS subsumes NoRS as a special case whenever no power is allocated to the common stream, the conventional transmission NoRS is also studied in this paper. Furthermore, a power allocation scheme that can achieve superiority of RS over the NoRS in finite alphabet systems is proposed and investigated.

For clarity we list the major contributions of this work as follows.

1) First, new analytical expressions for the ergodic sum-rate are derived for RS based on finite constellations with ZF and CI precoding schemes for the private messages. Both perfect CSIT and imperfect CSIT are considered. This contrasts with the existing literature that either study NoRS based on finite constellation with ZF/CI precoding, or RS based on Gaussian inputs. This is the first paper that a) studies RS with finite constellations, b) combines RS with CI precoding.
2) Second, a novel power allocation algorithm is introduced to optimize the resulting sum-rate in the finite alphabet scenario.
3) Third, Monte-Carlo simulations are provided to confirm the analysis, and the impact of the different system parameters on the achievable sum-rate are examined and investigated.

2. We note that, while traditional analysis focus on Gaussian signaling, the study of finite constellation signaling is of particular importance, since finite constellations are applied in practice.

3. Conventional transmission (without precoding) provides lower-bound of the ZF and CI techniques [33].

The results in this work show clearly that the sum-rate of RS with CI outperforms the sum rate of RS with ZF and NoRS (with either ZF or CI) transmission techniques.

Notations: h, $\mathbf{h}$, and $\mathbf{H}$ denote a scalar, a vector and a matrix, respectively. $(\cdot)^H$, $(\cdot)^T$ and diag(.) denote conjugate transposition, transposition and diagonal of a matrix, respectively. $\mathcal{E}[.]$ denotes average operation. $[\mathbf{h}]_k$ denotes the k^{th} element in $\mathbf{h}$, $|.|$ denotes the absolute value, and $\|.\|^2$ denotes the second norm. $\mathbb{C}^{K\times N}$ represents an $K \times N$ matrix, and $\mathbf{I}$ denotes the identity matrix.

II. SYSTEM MODEL

We consider a MU-MISO system, in which an $N-$antennas BS node communicates with K-single antenna users in a downlink scenario using the RS strategy. In this system the channels are assumed to be independent identically distributed (i.i.d) Rayleigh fading channels. The channel matrix between the BS and the K users is denoted by $\mathbf{H} \in \mathbb{C}^{K\times N}$, which can be written as $\mathbf{H} = \mathbf{D}^{1/2}\mathrm{G}$ where $\mathrm{G} \in \mathbb{C}^{K\times N}$ contains i.i.d $\mathcal{CN}(0,1)$ elements represent small-scale fading coefficients and $\mathbf{D} \in \mathbb{C}^{K\times K}$ is a diagonal matrix represents the path-loss attenuation with $[\mathbf{D}]_{kk} = d_k^{-m}$, where d_k is the distance between the BS and the k^{th} user, and m is the path loss exponent.[4]

Therefore, the BS transmits K independent messages $Q_{\tilde{t},1}, \ldots, Q_{\tilde{t},K}$ uniformly drawn from the sets $\mathcal{Q}_{\tilde{t},1}, \ldots, \mathcal{Q}_{\tilde{t},K}$, and intended for users $1, \ldots, K$ respectively. In RS, each user message is split into a common part and a private part, i.e., $Q_{\tilde{t},k} = \{Q_{c,k}, Q_{p,k}\}$[5] with $Q_{c,k} \in \mathcal{Q}_{c,k}$, $Q_{p,k} \in \mathcal{Q}_{p,k}$, and $\mathcal{Q}_{c,k} \times \mathcal{Q}_{p,k} = \mathcal{Q}_{\tilde{t},k}$. The common message is composed by packing the common parts such that $Q_c = \{Q_{c,1}, \ldots, Q_{c,K}\} \in \mathcal{Q}_{c,1} \times \cdots \times \mathcal{Q}_{c,K}$. The resulting $K+1$ messages are encoded into the independent data streams $x_c, x_1, \ldots, ., x_K$, where x_c and x_k represent the encoded common and private symbols [6]. The $K+1$ symbols are grouped in a signal vector $\mathbf{x} = [x_c, x_1, \ldots, ., x_K]^T \in \mathbb{C}^{K+1}$, where $\mathcal{E}\{\mathbf{x}\mathbf{x}^H\} = \mathrm{I}$. Then, the K private streams are mapped to the BS antenna array through MU-precoders, (e.g., ZF/ CI), while the common part is precoded in a multicast fashion.[6] Accordingly, the symbols are mapped to the BS antennas by a linear precoding matrix defined as $\mathbf{W} = [\mathbf{w}_c, \mathbf{w}_1, \ldots, \mathbf{w}_K]$ where $\mathbf{w}_c \in \mathbb{C}^N$ denotes the common precoder and $\mathbf{w}_k \in \mathbb{C}^N$ is the k^{th} private precoder. Therefore, the transmitted signal can be mathematically expressed by

$$\mathbf{s} = \mathbf{W}\mathbf{x} = \sqrt{P_c}\mathbf{w}_c x_c + \sum_{k=1}^{K} \sqrt{P_p}\mathbf{w}_k x_k \tag{1}$$

where $\mathbf{W} = [\mathbf{w}_c, \mathbf{w}_1, \ldots, \mathbf{w}_K]$, $\mathbf{w}_c$ denotes the common precoder of the common message and $\mathbf{w}_k$ is the k^{th} private precoder. In addition, $P_c = (1-t)P$ and $P_p = \frac{tP}{K}$ are the power allocated to the common message and the power allocated to the private message, respectively, while P is the total power and $t \in (0, 1]$ is the fraction of the total power allocated to the private parts[7] [5]–[7]. Conventional multi-user linear precoding without RS, NoRS, is a particular instance of the RS strategy and is obtained by turning of the common message and allocating all transmit power exclusively to the privates messages. The received signal at the k^{th} user in this system can be written as

$$y_k = \sqrt{P_c}\mathbf{h}_k\mathbf{w}_c x_c + \sum_{k=1}^{K}\sqrt{P_p}\mathbf{h}_k\mathbf{w}_k x_k + n_k \tag{2}$$

where $\mathbf{h}_k^T$ is the $K \times 1$ channel vector from the BS to user k, n_k is the additive wight Gaussian noise (AWGN) at the k^{th} user, $n_k \sim \mathcal{CN}(0, \sigma_k^2)$.[8] At the user side, the common symbol is decoded firstly by treating the interference from the private messages as noise, and then each user decodes its own message after canceling the common message using SIC technique. It is assumed that, the signals are transmitted with strong codes (infinite length) at an ergodic rate. Hence the rates are achievable and errors do not occur.[9] Therefore, after removing the contribution from the common message, the received signal at the k^{th} user in this system can be written as

$$y_k^p = \mathbf{h}_k\mathbf{W}^p\mathbf{x}^p + n_k = \sum_{k=1}^{K}\sqrt{P_p}\mathbf{h}_k\mathbf{w}_k x_k + n_k \tag{3}$$

where $\mathbf{x}^p = [x_1, \ldots, , x_K]^T$ and $\mathbf{W}^p = [\mathbf{w}_1, \ldots, \mathbf{w}_K]$. The sum rate in this scenario can be expressed by $R = R^c + \sum_{[k=1]}^{K} R_k^p$, where R^c is the rate for the common part, $R^c = \min(R_1^c, R_2^c, .., R_k^c, .., R_K^c)$, R_k^c is the rate for the common message at user k, and R_k^p is the rate for the private part at the k^{th} user. In this work, both perfect CSIT and imperfect CSIT are considered, and delay-tolerant transmission is assumed. Hence the channel coding can be achieved over a long sequence of channel states. Therefore, transmitting the common and the k^{th} private messages at ergodic rates $\mathcal{E}\{R_k^c\}$ and $\mathcal{E}\{R_k^p\}$, respectively, guarantees successful decoding by the k^{th} user [9]. Hence, to guarantee the common message, x_c, is successfully decoded and then canceled by the users, it should be sent at an ergodic rate not exceeding $\min_j(\mathcal{E}\{R_j^c\})_{j=1}^K$. Finally, the ergodic sum rate can be

4. The value of the path-loss exponent, m, depends on the propagation environment. The actual value of m can be calculated by fitting the model to the empirical data [35].

5. The subscript $\tilde{t}$ here denotes total, which is explained that $Q_{\tilde{t},k}$ is composed of two parts. The subscripts c and p are used for common part and private parts, respectively.

6. For more details about the architecture of MU-MISO system with RS we refer the reader to [5] where the basic concepts of RS scheme was discussed in detail.

7. We assume a uniform power allocation among all the private symbols, similarly to other works on RS [7], [10]. Although this assumption does not produce the optimal performance, it allows us to find tractable results. This assumption is commonly used in practice, e.g., LTE and LTE-A.

8. The additive noise may arise from amplifiers and electronic components at the receiver and it is characterized as thermal noise. This type of noise is characterized statistically as a Gaussian noise process with zero mean and variance σ_n^2 [36].

9. Considerations with finite block lengths is beyond the scope of the paper.

expressed by

$$\mathcal{E}\{R\} = \min_{j}\left(\mathcal{E}\left\{R_j^c\right\}\right)_{j=1}^{K} + \sum_{[k=1]}^{K} \mathcal{E}\left\{R_k^p\right\} \tag{4}$$

Next, the ergodic sum-rate has been derived in two different scenarios based on the availability of the CSIT.

III. ERGODIC SUM RATE ANALYSIS UNDER PERFECT CSIT

In this scenario, the BS has perfect knowledge of the CSI, and the precoding matrices have been designed based on this perfect knowledge. Therefore, in this section two precoding techniques are considered. In the first one, we use weighted maximum ratio transmission (MRT) for the common message and ZF for the private messages, and in the second one we use weighted MRT for the common message and CI for the private messages.

A. RS: MRT/ZF

In this case we implement weighted MRT technique for common signal and ZF precoding for the private messages. The analysis provided in this section can be applied to any finite alphabet input such as phase shift-keying (PSK), and quadrature amplitude modulation (QAM). Therefore, the precoding for the common and the private messages can be written, respectively, as

$$\mathbf{w}_c = \sum_{i=1}^{K} \beta_c \mathbf{h}_i^H \quad \text{and} \quad \mathbf{W}_{ZF}^p = \beta_p\, \mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H\right)^{-1}, \tag{5}$$

where β_c and β_p are the scaling factors to meet the transmit power constraint at the transmitter, which can be expressed as $\beta_c = \frac{1}{\sqrt{\| \sum_{i=1}^{K} \mathbf{h}_i^H \|^2}}$ and $\beta_p = \sqrt{\frac{1}{\mathbf{x}^H(\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{x}}}$ [10], [29]. For simplicity the normalization constants β_c and β_p are designed to ensure that the long-term total transmit power at the BS is constrained. The long-term power constraint is well justified in the literature, for instance [37]. Accordingly, the scaling factors can be written as [4], [29], $\beta_c = \frac{1}{\sqrt{\mathcal{E}\{\| \sum_{i=1}^{K} \mathbf{h}_i^H \|^2\}}}$ and $\beta_p = \frac{1}{\sqrt{\mathcal{E}\{\mathbf{x}^H(\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{x}\}}}$, respectively. Since $\| \sum_{i=1}^{K} \mathbf{h}_i^H \|^2$ and $\frac{1}{\mathbf{s}^H(\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{s}}$ both have Gamma distribution [16], we can find that, $\beta_c = \frac{1}{\sqrt{N \sum_{[i=1]}^{K} \varpi_k}}$ and $\beta_p = \sqrt{\frac{\Gamma(2-K+N)}{(K(N-K)!)}}$ where $\varpi_k = d_k^{-m}$ [38].

i) Ergodic rate for the common part: the ergodic rate for the common part at user k can be written as in (6), shown at the bottom of the page, where $\mathbf{W} = [\mathbf{w}_c, \mathbf{W}_{ZF}^p]$, $\mathbf{x}_{m,i} = \mathrm{x}_m - \mathbf{x}_i$, $\mathbf{x}_m$ and $\mathbf{x}_i$ contain N symbols, which are taken from the equiprobable constellation set with cardinality M.[10] Similar to the Gaussian input assumption case, (6) reveals that the achievable rate suffering from the interference caused by other signals. The first term in (6), φ, contains all the received signals at user k, while the second term, ψ, contains only the interference signals.

Proof: The proof of the above follows known derivations from the finite constellation rate analysis literature [3], [39], [40], and due to the paper length limitation, the proof of (6) has been omitted in this paper. ∎

By using Jensen inequality, an accurate approximation of the ergodic rate for the common part at user k can be obtained following a common approach in [3] and [33]. Firstly, the first term in (6), φ, can be written as

$$\varphi = \mathcal{E}_{\mathbf{h},n_k}\left\{\log_2 \sum_{i=1}^{M^N} e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}\right\}$$
$$\leq \log_2 \sum_{i=1}^{M^N} \mathcal{E}_{\mathbf{h},n_k}\left\{e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}\right\}. \tag{7}$$

Since the noise n_k has Gaussian distribution, the average over the noise can be derived as

$$\mathcal{E}_{n_k}\left\{e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}\right\} = \frac{1}{\pi\sigma^2}\int_{n_k} e^{-\frac{\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2 + |n_k|^2}{\sigma_k^2}}\, dn_k. \tag{8}$$

By using the integrals of exponential function in [41], the average over the noise can be obtained as

$$\mathcal{E}_n\left\{e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}\right\} = e^{-\frac{\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i}\right|^2}{2\sigma_k^2}}. \tag{9}$$

Now, we can write φ as

$$\varphi = \log_2 \sum_{i=1}^{M^N} \mathcal{E}_{\mathbf{h}}\left\{e^{-\frac{\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i}\right|^2}{2\sigma_k^2}}\right\} \tag{10}$$
$$= \log_2 \sum_{i=1}^{M^N} \mathcal{E}_{\mathbf{h}}\left\{e^{-\frac{\left|\sqrt{P_c}\mathbf{h}_k \mathbf{w}_c [\mathbf{x}_{m,i}]_1 + \sqrt{P_p}\mathbf{h}_k \mathbf{W}_{ZF}^p \mathbf{x}_{m,i}^p\right|^2}{2\sigma_k^2}}\right\} \tag{11}$$
$$= \log_2 \sum_{i=1}^{M^N} \mathcal{E}_{\mathbf{h}}\left\{e^{-\frac{P\left|\sqrt{(1-t)}\beta_c\left(\sum_{i=1}^{K}\mathbf{h}_k\mathbf{h}_i^H\right)[\mathbf{x}_{m,i}]_1 + \sqrt{t}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}}\right\}. \tag{12}$$

10. Each input $\mathbf{x}_i$ consists of symbols taken from the M-PSK constellation.

$$\mathcal{E}\left\{R_k^c\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N} \underbrace{\mathcal{E}_{\mathbf{h},n_k} \log_2 \sum_{i=1}^{M^N} e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}}_{\varphi} + \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}} \underbrace{\mathcal{E}_{\mathbf{h},n_k} \log_2 \sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}_{ZF}^p \mathbf{x}_{m,i}^p + n_k\right|^2}{\sigma_k^2}}}_{\psi} \tag{6}$$

Since the term $Y = (\sum_{i=1}^{K} \mathbf{h}_k \mathbf{h}_i^H)$ has Gamma distribution, i.e., $Y \sim \Gamma(\upsilon, \theta)$, the average can be derived as

$$\varphi = \log_2 \sum_{i=1}^{M^N} \int_0^{\infty} e^{-\frac{P\left|\sqrt{(1-t)}\beta_c y\left[\mathbf{x}_{m,i}\right]_1 + \sqrt{t}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}} \frac{y^{\upsilon-1} e^{\frac{-y}{\theta}}}{\Gamma(\upsilon)\theta^{\upsilon}} dy. \tag{13}$$

Applying Gaussian Quadrature rule, the average can be obtained by

$$\varphi = \log_2 \sum_{i=1}^{M^N} \sum_{r=1}^{n} \frac{(y_r)^{\upsilon-1} \mathrm{H}_r}{\Gamma(\upsilon)} e^{-\frac{P\left|\sqrt{(1-t)}\beta_c \theta y_r\left[\mathbf{x}_{m,i}\right]_1 + \sqrt{t}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}} \tag{14}$$

where y_r and H_r are the r^{th} zero and the weighting factor of the Laguerre polynomials, respectively [42]. Similarly, for the second term ψ, using Jensen inequality we can write

$$\psi = \log_2 \sum_{t=1}^{M^{N-1}} \mathcal{E}_{n_k} \left\{ e^{-\frac{\left|\sqrt{tP}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k + n_k\right|^2}{\sigma_k^2}} \right\}. \tag{15}$$

Using the integrals of exponential function in [41], the average over the noise can be obtained as

$$\psi = \log_2 \sum_{t=1}^{M^{N-1}} e^{-\frac{\left|\sqrt{tP}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}}. \tag{16}$$

Substituting (14) and (16) into (6), we can get an accurate approximation of the ergodic rate for the common part at user k. The accuracy of above derivation is verified in the following by its close match with simulated performance in Figs. 1, 2, and 3 of our results section.

ii) Ergodic rate for the private part: the ergodic rate for the private message at the k^{th} user, under finite alphabet signaling using ZF precoding technique can be written as

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}} \sum_{m=1}^{M^{N-1}} \underbrace{\mathcal{E}\left\{ \log_2 \sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}_{ZF}^p \mathbf{x}_{m,i}^p\right|^2}{2\sigma_k^2}} \right\}}_{\psi}. \tag{17}$$

By using Jensen inequality, and following similar steps as in the previous sub-section we can find the average of the term ψ in (17) as in (16). Consequently, the ergodic rate for the k^{th} private message using ZF can be expressed as,

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}} \sum_{m=1}^{M^{N-1}} \log_2 \sum_{t=1}^{M^{N-1}} e^{-\frac{\left|\sqrt{tP}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}}. \tag{18}$$

B. RS: MRT/CI

In this scenario weighted MRT technique is used for common message and CI precoding for the private messages. As the interference exploitation scheme is modulation dependent, in this section CI precoder with PSK and QAM input signals are considered.

1) CI PRECODING WITH PSK INPUT ALPHABET

The precoder for the common and the private messages in this case can be written, respectively, as

$$\mathbf{w}_c = \sum_{i=1}^{K} \beta_c \mathbf{h}_i^H \quad \text{and}$$
$$\mathbf{W}_{CI}^p = \frac{1}{K} \beta_p \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} \mathrm{diag}\left\{\mathbf{V}^{-1}\mathbf{u}\right\}, \tag{19}$$

where $\beta_c = \frac{1}{\sqrt{\|\sum_{i=1}^{K} \mathbf{h}_i^H\|^2}}$ and $\beta_p = \frac{1}{\sqrt{\mathbf{u}^H \mathbf{V}^{-1} \mathbf{u}}}$ are the scaling factor to meet the transmit power constraint at the transmitter, while $\mathbf{V} = \mathrm{diag}(\mathbf{x}^{pH})(\mathbf{H}\mathbf{H}^H)^{-1}\mathrm{diag}(\mathbf{x}^p)$ and $\mathbf{1}^T\mathbf{u} = 1$[11] [10], [29]. For simplicity the normalization constants β_c and β_p are designed to ensure that the long-term total transmit power at the BS is constrained, so it can be written as [4], [29] $\beta_c = \frac{1}{\sqrt{\mathcal{E}\{\|\sum_{i=1}^{K} \mathbf{h}_i^H\|^2\}}}$ and $\beta_p = \frac{1}{\sqrt{\mathcal{E}\{\mathbf{u}^H \mathbf{V}^{-1} \mathbf{u}\}}}$. Since $\|\sum_{i=1}^{K} \mathbf{h}_i^H\|^2$ and $(\mathbf{H}\mathbf{H}^H)$ has Gamma and Wishart distributions respectively, we can find that, $\beta_c = \frac{1}{\sqrt{N\sum_{i=1}^{K}\varpi_k}}$ and $\beta_p = \frac{1}{\sqrt{\mathbf{u}^H \mathrm{diag}(\mathbf{x}^H)^{-1} N\Sigma(\mathrm{diag}(\mathbf{x}))^{-1}\mathbf{u}}}$, where $\Sigma = \mathrm{D}$ [38]. From (4), we now need to calculate the ergodic rate for the common and private messages as follows.

i) Ergodic rate for the common part: the ergodic rate for the common part at user k under PSK signaling can be written as

$$\mathcal{E}\left\{R_k^c\right\} = \log_2 M - \frac{1}{M^N} \sum_{m=1}^{M^N} \underbrace{\mathcal{E}_{\mathbf{h},n_k} \log_2 \sum_{i=1}^{M^N} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}\mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}}_{\varphi}$$
$$+ \frac{1}{M^{N-1}} \sum_{m=1}^{M^{N-1}} \underbrace{\mathcal{E}_{\mathbf{h},n_k} \log_2 \sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}_{CI}^p \mathbf{x}_{m,i}^p + n_k\right|^2}{\sigma_k^2}}}_{\psi}, \tag{20}$$

where $\mathbf{W} = [\mathbf{w}_c, \mathbf{W}_{CI}^p]$. By invoking Jensen inequality, an accurate approximation of the ergodic rate for the common part at user k can be derived as follows. The first term in (20), φ, can be expressed by

$$\varphi = \mathcal{E}_{\mathbf{h},n_k} \left\{ \log_2 \sum_{i=1}^{M^N} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}\mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}} \right\}$$

11. Please, note that CI precoding expression is conditional on $\mathbf{u}$, thus all the derived expressions in this work for CI are conditional on$\mathbf{u}$. The values of this vector should satisfy the condition $\mathbf{1}^T\mathbf{u} = 1$. The impact of this vector on the system performance has been investigated in [31], [33]. Although the elements of $\mathbf{u}$ should be selected to maximize the performance, using any values larger than zero and satisfy $\mathbf{1}^T\mathbf{u} = 1$ can show the superiority of the CI over the conventional techniques.

$$\leq \log_2 \sum_{i=1}^{M^N} \mathcal{E}_{\mathbf{h},n_k} \left\{ e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}} \right\} \tag{21}$$

Since the noise, n_k, has Gaussian distribution, the average over the noise can be derived using the integrals of exponential function in [41] as

$$\mathcal{E}_n \left\{ e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}} \right\} = e^{-\frac{\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i}\right|^2}{2\sigma_k^2}}. \tag{22}$$

Now, the average over the channel can be derived as

$$\mathcal{E}_{\mathbf{h}} \left\{ e^{-\frac{\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i}\right|^2}{2\sigma_k^2}} \right\} = \mathcal{E}_{\mathbf{h}} \left\{ e^{-\frac{\left|\sqrt{P_c}\mathbf{h}_k \mathbf{w}_c \left[\mathbf{x}_{m,i}\right]_1 + \sqrt{P_p}\mathbf{h}_k \mathbf{W}_{CI}^p \mathbf{x}_{m,i}^p\right|^2}{2\sigma_k^2}} \right\} \tag{23}$$

which can be written as in (24), shown at the bottom of the page, where $\mathbf{a}_k$ is a $1 \times K$ vector all the elements of this vector are zeros except the k^{th} element is one. Therefore, the first term φ can be expressed as in (25) shown at the bottom of the page, where $\mathbf{A} = (\mathbf{H}\mathbf{H}^H)$, $\xi = \frac{\beta_p}{K}\mathbf{a}_k(\mathrm{diag}(\mathbf{x}^{pH}))^{-1}(\Sigma)(\mathrm{diag}(\mathbf{x}^p))^{-1}\mathbf{u}$. Now, we can simplify the last expression in (25) to

$$\varphi = \log_2 \sum_{i=1}^{M^N} \mathcal{E}_{\mathbf{h}} \left\{ e^{-\frac{P|\xi|^2 \|\mathbf{A}_k\|^2 \Psi_{m,i}}{2\sigma_k^2}} \right\}. \tag{26}$$

The term $\Psi_{m,i}$ in (25) can be reduced to $\Psi_{m,i} = N(|(\sqrt{1-t}[\mathbf{x}_{m,i}]_1) + (\sqrt{t}\varpi_k^{-1}[\mathbf{x}_{m,i}^p]_k)|^2)$. Let $z_k = \|\mathbf{A}_k\|^2$ which has Gamma distribution, i.e., $z_k \sim \Gamma(\upsilon_k, \theta_k)$, with $\upsilon_k = N(N+1)$ degrees of freedom, therefore the average over $\mathbf{h}$ is the moment generating function (MGF) of the term, $\frac{P|\xi|^2\|\mathbf{A}_k\|^2\Psi_{m,i}}{2\sigma_k^2}$ which can be found easily as

$$\varphi = \log_2 \sum_{i=1}^{M^N} \left(1 + \frac{P|\xi|^2 \theta_k \Psi_{m,i}}{2\sigma_k^2} \right)^{-\upsilon_k}. \tag{27}$$

For the second term, ψ, similarly using Jensen inequality we can write

$$\begin{aligned} \psi &= \mathcal{E}_{\mathbf{h},n_k} \left\{ \log_2 \sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}_{CI}^p \mathbf{x}_{m,i}^p + n_k\right|^2}{\sigma_k^2}} \right\} \\ &\leq \log_2 \sum_{i=1}^{M^{N-1}} \mathcal{E}_{\mathbf{h},n_k} \left\{ e^{\frac{-\left|\mathbf{h}_k \mathbf{W}_{CI}^p \mathbf{x}_{m,i}^p + n_k\right|^2}{\sigma_k^2}} \right\}. \end{aligned} \tag{28}$$

Since n_k has Gaussian distribution, we can write ψ as

$$\begin{aligned} \psi &= \log_2 \sum_{i=1}^{M^{N-1}} \mathcal{E}_{\mathbf{h}} \left\{ e^{-\frac{\left|\left(\frac{\sqrt{P_p}\beta_p}{K} \boldsymbol{a}(\mathbf{H}\mathbf{H}^H)\boldsymbol{b}\left[\mathbf{x}_{m,i}^p\right]_k\right)\right|^2}{2\sigma_k^2}} \right\} \\ &= \log_2 \sum_{i=1}^{M^{N-1}} \mathcal{E}_{\mathbf{h}} \left\{ e^{-\frac{\left|c_k Y \left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}} \right\} \end{aligned} \tag{29}$$

where $a = \mathbf{a}_k(\mathrm{diag}(\mathbf{x}^{pH}))^{-1}$, $b = (\mathrm{diag}(\mathbf{x}^p))^{-1}\mathbf{u}$, $c_k = \frac{\sqrt{P_p}\beta_p \mathbf{a}\Sigma \boldsymbol{b}}{K}$ and $Y = \frac{\boldsymbol{a}(\mathbf{H}\mathbf{H}^H)\boldsymbol{b}}{\mathbf{a}\Sigma\boldsymbol{b}}$. It was shown that, Y has a Gamma distribution [38]. Therefore we can write

$$\psi = \log_2 \sum_{i=1}^{M^{N-1}} \int_0^\infty e^{-\frac{\left|c Y \left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}} \frac{e^{-Ky}(Ky)^{N-K} K}{(N-K)!} dy, \tag{30}$$

which can be obtained as

$$\psi = \log_2 \sum_{i=1}^{M^{N-1}} \Lambda_{k_{m,i}} \tag{31}$$

where $\Lambda_{k_{m,i}}$ is expressed in (32), shown at the bottom of the next page where ${}_1\mathrm{F}_1$ is the Hypergeometric function.

Substituting (27) and (31) into (20), we can get an accurate approximation of the ergodic rate for the common part at user k.

$$\mathcal{E}_{\mathbf{h}} \left\{ e^{-\frac{\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i}\right|^2}{2\sigma_k^2}} \right\} = \mathcal{E}_{\mathbf{h}} \left\{ e^{-\frac{\left|\sqrt{P_c}\beta_c\left(\sum_{i=1}^{K}\mathbf{h}_k\mathbf{h}_i^H\right)\left[\mathbf{x}_{m,i}\right]_1 + \left(\frac{\sqrt{P_p}\beta_p}{K}\mathbf{a}_k\left(\mathrm{diag}\left(\mathbf{x}^{pH}\right)\right)^{-1}\left(\mathbf{H}\mathbf{H}^H\right)(\mathrm{diag}(\mathbf{x}^p))^{-1}\mathbf{u}\left[\mathbf{x}_{m,i}^p\right]_k\right)\right|^2}{2\sigma_k^2}} \right\} \tag{24}$$

$$\varphi = \log_2 \sum_{i=1}^{M^N} \mathcal{E}_{\mathbf{h}} e^{-\frac{P|\xi|^2\|\mathbf{A}_k\|^2 \overbrace{\left| \frac{\sqrt{(1-t)}\beta_c\left(\sum_{i=1}^K \mathbf{h}_k\mathbf{h}_i^H\right)\left[\mathbf{x}_{m,i}\right]_1}{|\xi|\|\mathbf{A}_k\|} + \frac{\sqrt{t}\beta_p\mathbf{a}_k\left(\mathrm{diag}\left(\mathbf{x}^{pH}\right)\right)^{-1}\mathbf{A}(\mathrm{diag}(\mathbf{x}^p))^{-1}\mathbf{u}\left[\mathbf{x}_{m,i}^p\right]_k}{|\xi|K\|\mathbf{A}_k\|} \right|^2}^{\Psi_{m,i}}}{2\sigma_k^2}} \tag{25}$$

ii) Ergodic rate for the private part: the ergodic rate for the private part at user k under PSK signaling, using CI precoding technique can be written as

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}} \sum_{m=1}^{M^{N-1}} \underbrace{\mathcal{E}\left\{\log_2 \sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}_{CI}^p \mathbf{x}_{m,i}^p\right|^2}{2\sigma_k^2}}\right\}}_{\psi}. \tag{33}$$

By using Jensen inequality, and following similar steps as in the previous sub-section we can find the average of the term ψ in (33) as in (31). Consequently, the ergodic rate for the k^{th} private message using CI can be expressed as [3], [39], [40],

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}} \sum_{m=1}^{M^{N-1}} \log_2 \sum_{i=1}^{M^{N-1}} \Lambda_{k_{m,i}}. \tag{34}$$

2) CI PRECODING WITH QAM INPUT ALPHABET

In this case, the precoder for the private messages can be written as [43]

$$\mathbf{W}_{CI}^p = \frac{1}{K}\beta_p \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} \mathbf{U}\mathrm{diag}\left\{\mathbf{F}^{-1}\tilde{\mathbf{V}}^{-1}\mathbf{u}_1\right\}\mathbf{x}^p \hat{\mathbf{x}}^{pT}, \tag{35}$$

where $\beta_p = \frac{1}{\sqrt{\mathcal{E}\{\mathbf{u}_1^H \tilde{\mathbf{V}}^{-1}\mathbf{u}_1\}}}$ is the scaling factor to meet the transmit power constraint at the transmitter, $\mathbf{F}$ is $2K \times 2K$ transformation matrix, $\tilde{\mathbf{V}} = \mathbf{F}\mathbf{V}\mathbf{F}^T$, $\mathbf{V} = \mathrm{diag}(\mathbf{x}^{pH})\mathbf{U}^H(\mathbf{H}\mathbf{H}^H)^{-1}\mathbf{U}\mathrm{diag}(\mathbf{x}^p)$, $\hat{\mathbf{x}}^p = [\frac{1}{x_1^p}, \ldots, ., \frac{1}{x_K^p}]^T$ and $\mathbf{1}^T\mathbf{u}_1 = 1$ [43].

i) Ergodic rate for the common part: by substituting (35) into (20) and following similar steps as in PSK scenario, the average of the first and second terms in (20), φ and ψ, can be found, respectively, as

$$\begin{aligned}\varphi &= \log_2 \sum_{i=1}^{M^N} \int_0^\infty e^{-\frac{Pz\Psi_{m,i}}{2\sigma_k^2}} \frac{e^{-\frac{\beta}{z}}(z)^{-\alpha-1}\beta^\alpha}{\Gamma(\alpha)} dz \\ &= \log_2 \sum_{i=1}^{M^N}\sum_{n=1}^{\mathcal{N}} \mathrm{H}_n e^{-z_n\left(\frac{P\Psi_{m,i}}{2\sigma_k^2}-1\right)} \frac{e^{-\frac{\beta}{z_n}}(z_n)^{-\alpha-1}\beta^\alpha}{\Gamma(\alpha)},\end{aligned} \tag{36}$$

and

$$\psi = \log_2 \sum_{i=1}^{M^{N-1}} \Xi_{k_{m,i}} \tag{37}$$

where z_n and H_n are the n^{th} zero and the weighting factor of the Laguerre polynomials, respectively [41] and $\Xi_{k_{m,i}}$ is given in (38) shown at the bottom of the page.

Proof: Due to the paper length limitation and in order to avoid the repetition, the proofs of (36) and (37) have been omitted in this paper. ■

ii) Ergodic rate for the private part: by substituting (35) in (33), and following similar steps as in the previous sub-section we can find the average of the term ψ in (33) as in (37). Consequently, the ergodic rate for the k^{th} private message using CI can be expressed as [3], [39], [40],

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}} \sum_{m=1}^{M^{N-1}} \log_2 \sum_{i=1}^{M^{N-1}} \Xi_{k_{m,i}}. \tag{39}$$

$$\begin{aligned}\Lambda_{k_{m,i}} = &\left(\left(\frac{2^{\left(\frac{1}{2}(N-K-1)\right)} K^{(N-K+1)} \left|c_k\left[\mathbf{x}_{m,i}^p\right]_k\right|^{-2+K-N}}{(N-K)!}\right)\left(\left(\frac{1}{\sigma_k^2}\right)^{\frac{1}{2}(K-N-2)}\right)\right) \\ &\times \left(\sqrt{\frac{1}{\sigma_k^2}}\left(\left|c\left[\mathbf{x}_{m,i}^p\right]_k\right|\right)\Gamma\left(\frac{1}{2}(N-K+1)\right){}_1\mathrm{F}_1\left(\frac{1}{2}(N-K+1), \frac{1}{2}, \frac{K^2\sigma_k^2}{2\left|c_k\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}\right)\right. \\ &\left. - \sqrt{2}K\,\Gamma\left(\frac{1}{2}(N-K+2)\right){}_1\mathrm{F}_1\left(\frac{1}{2}(N-K+2), \frac{3}{2}, \frac{K^2\sigma_k^2}{2\left|c_k\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}\right)\right)\end{aligned} \tag{32}$$

$$\begin{aligned}\Xi_{k_{m,i}} = &\left(\left(\frac{\tilde{K}^{1-K+N}}{2c^2(N-K)!}\right)\left(\left(\frac{c^2}{\sigma_k^2}\right)^{\frac{1}{2}(K-N)}\right)\sigma_k^2\right) \times \left(\sqrt{\frac{c}{\sigma_k^2}}\Gamma\left(\frac{1}{2}(N-K+1)\right){}_1\mathrm{F}_1\left(\frac{1}{2}(N-K+1), \frac{1}{2}, \frac{K^2\sigma_k^2}{4c}\right)\right. \\ &\left. - K\,\Gamma\left(\frac{1}{2}(N-K+2)\right){}_1\mathrm{F}_1\left(\frac{1}{2}(N-K+2), \frac{3}{2}, \frac{K^2\sigma_k^2}{4c}\right)\right)\end{aligned} \tag{38}$$

C. CONVENTIONAL TRANSMISSION WITHOUT RATE SPLITTING (NORS)

The ergodic rate at the k^{th} user in conventional transmission without RS is expressed by

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N} \underbrace{\mathcal{E}\left\{\log_2 \sum\nolimits_{i=1}^{M^N} e^{\frac{-\left|\mathbf{h}_k \mathbf{W} \mathbf{x}_{m,i}\right|^2}{2\sigma_k^2}}\right\}}_{\psi}. \tag{40}$$

In ZF scenario the precoding matrix $\mathbf{W}$ is given in (5), and then the expectation in (40) can be derived using Jensen inequality as in (16). Consequently, the ergodic rate for the k^{th} user without RS using ZF can be expressed as

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N} \log_2 \sum_{t=1}^{M^N} e^{-\frac{\left|\sqrt{tP}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}}. \tag{41}$$

On the other hand, in CI with PSK signals the precoding matrix $\mathbf{W}$ is given in (19), and the expectation in (40) can be derived using Jensen inequality as in (31). Consequently, the ergodic rate for the k^{th} user without RS using CI can be expressed as

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N} \log_2 \sum_{i=1}^{M^N} \Lambda_{k_{m,i}}. \tag{42}$$

In case CI with QAM signals, the expectation in (40) can be derived using Jensen inequality as in (37). Consequently, the ergodic rate for the k^{th} user without RS can be expressed as

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N} \log_2 \sum_{i=1}^{M^N} \Xi_{k_{m,i}}. \tag{43}$$

Please note that, in case the users' locations are randomly distributed, the ergodic sum-rate with respect to each user location can be calculated easily by averaging the derived sum-rate expression over all possible user locations.

IV. ERGODIC SUM RATE ANALYSIS UNDER IMPERFECT CSI

Now, we consider the case when the channel matrix H is imperfectly known to the BS. Considering minimum mean square error (MMSE) channel estimator, which is widely investigated in the literature [8], [44], [45]. The current channels in terms of the estimated channels, and the estimation error can be written as, $\mathrm{H} = \hat{\mathrm{H}} + \mathbf{E}$, where $\hat{\mathrm{H}}$ is the estimated channel matrix, $\mathbf{E}$ is the estimation error matrix. By the property of MMSE estimation the two matrices $\hat{\mathrm{H}}$, and $\mathbf{E}$ are uncorrelated and distributed as $\hat{\mathrm{H}} \sim \mathcal{CN}(0, \hat{\boldsymbol{D}})$ and $\mathbf{E} \sim \mathcal{CN}(0, \boldsymbol{D} - \hat{\boldsymbol{D}})$, where $\hat{\boldsymbol{D}}$ is a diagonal matrix with $[\hat{\mathbf{D}}]_{kk} = \hat{\sigma}_k^2 = \frac{p_u \varpi_k^2}{p_u \varpi_k + 1}$ and $[\boldsymbol{D} - \hat{\mathbf{D}}]_{kk} = \hat{\sigma}_{ek}^2 = \frac{\varpi_k}{p_u \varpi_k + 1}$ [8], [44], [45], while $p_u = \tau p_p$, τ is number of symbols used for channel training and p_p is the transmit power for each pilot symbol. Consequently, the received signal can be written now as,

$$\hat{y}_k = \sqrt{P_c}\hat{\mathbf{h}}_k \hat{\mathbf{w}}_c x_c - \sqrt{P_c}\boldsymbol{e}_k \hat{\mathbf{w}}_c x_c + \sum_{i=1}^{K}\sqrt{P_p}\hat{\mathbf{h}}_k \hat{\mathbf{w}}_i^p x_i - \sum_{i=1}^{K}\sqrt{P_p}\boldsymbol{e}_k \hat{\mathbf{w}}_i^p x_i + n_k. \tag{44}$$

A. RS: MRT/ZF

In this case the precoding for the common and the private messages based on the estimated channels can be written, respectively, as

$$\hat{\mathbf{w}}_c = \sum_{i=1}^{K}\beta_c \hat{\mathbf{h}}_i^H \quad \text{and} \quad \hat{\mathbf{W}}^p = \beta_p \hat{\mathbf{H}}^H\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)^{-1}. \tag{45}$$

Therefore, the received signal is given by

$$\hat{y}_k = \sqrt{P_c}\beta_c N \sum_{i=1}^{K}\frac{1}{N}\hat{\mathbf{h}}_k \hat{\mathbf{h}}_i^H x_c - \sqrt{P_c}\beta_c N \sum_{i=1}^{K}\frac{1}{N}\boldsymbol{e}_k \hat{\mathbf{h}}_i^H x_c + \sqrt{P_p}\sum_{i=1}^{K}\hat{\mathbf{h}}_k \hat{\mathbf{w}}_i^p x_i - \sqrt{P_p}N\sum_{i=1}^{K}\frac{1}{N}\boldsymbol{e}_k \hat{\mathbf{w}}_i^p x_i + n_k. \tag{46}$$

i) Ergodic rate for the common part: the ergodic rate for the common part at user k under finite alphabet signaling in imperfect CSI scenario can be written as in (47) shown at the bottom of the page.

As one can see from (47), the ergodic rate is hard to further simplify, since the expectations involve several random

$$\mathcal{E}\left\{R_k^c\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N} \underbrace{\mathcal{E}_{\hat{\mathbf{h}},\boldsymbol{e},n_k}\left\{\log_2 \sum_{i=1}^{M^N} e^{\frac{-\left|\hat{\mathbf{h}}_k \hat{\mathbf{W}}\mathbf{x}_{m,i} + \boldsymbol{e}_k \hat{\mathbf{W}}\mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}\right\}}_{\varphi} + \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}} \underbrace{\mathcal{E}_{\hat{\mathbf{h}},\boldsymbol{e},n_k}\left\{\log_2 \sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\hat{\mathbf{h}}_k \hat{\mathbf{W}}^p \mathbf{x}_{m,i}^p + \boldsymbol{e}_k \hat{\mathbf{W}}\mathbf{x}_{m,i} + n_k\right|^2}{\sigma_k^2}}\right\}}_{\psi} \tag{47}$$

variables. However, an approximation based on large number of antennas at the BS can be derived.

Analysis for large N: in this case we analyze the ergodic rate when the number of BS antennas is large ($N \gg K$), driven by the increasing research interest in MU-MISO systems with a large number of BS antennas.

Lemma 1: Let $\boldsymbol{a} = [a_1 \cdots ..a_n]^T$ and $\boldsymbol{b} = [b_1 \cdots ..b_n]^T$ be $n \times 1$ independent vectors contain i.i.d entries with zero-mean and variances $\mathcal{E}\{|a_i|^2\} = \sigma_a^2$ and $\mathcal{E}\{|b_i|^2\} = \sigma_b^2$. Therefore, following the law of large numbers, we can get [45]

$$\frac{1}{n}\boldsymbol{a}^H\boldsymbol{a} \overset{\text{a.s}}{\to} \sigma_a^2, \frac{1}{n}\boldsymbol{b}^H\boldsymbol{b} \overset{\text{a.s}}{\to} \sigma_b^2, \frac{1}{n}\boldsymbol{a}^H\boldsymbol{b} \overset{\text{a.s}}{\to} 0,$$
$$\text{and } \frac{1}{\sqrt{n}}\boldsymbol{a}^H\boldsymbol{b} \overset{\text{d}}{\to} \mathcal{CN}\left(0,\sigma_a^2\sigma_b^2\right) \quad (48)$$

where $\overset{\text{a.s}}{\to}$ and $\overset{\text{d}}{\to}$denote almost-sure and distribution convergence, respectively.

It is well known that by deploying very large number of antennas at the BS, the small-scale fading can be averaged out. Therefore, we now can elaborate more on analyzing the impact of large-scale fading on the system performance. Using the facts in Lemma 1, the rate for the common part at user k can be written as

$$\begin{aligned}\mathcal{E}\left\{R_k^c\right\} &= N\log_2 M \\ &- \frac{1}{M^N}\sum_{m=1}^{M^N} \\ &\times \underbrace{\mathcal{E}_{d_k,n_k}\left\{\log_2\sum_{i=1}^{M^N} e^{\frac{-\left|\sqrt{P_c}\beta_c N\hat{\sigma}_k^2 x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\}}_{\varphi} \\ &+ \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}} \\ &\times \underbrace{\mathcal{E}_{d_k,n_k}\left\{\log_2\sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\sqrt{P_p}\beta_p x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\}}_{\psi}. \end{aligned} \quad (49)$$

By using Jensen inequality, the first term in (49), φ, can be expressed by

$$\begin{aligned}\varphi &= \mathcal{E}_{d_k,n_k}\left\{\log_2\sum_{i=1}^{M^N} e^{\frac{-\left|\sqrt{P_c}\beta_c N\hat{\sigma}_k^2 x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\} \\ &\le \log_2\sum_{i=1}^{M^N}\mathcal{E}_{d_k,n_k}\left\{e^{\frac{-\left|\sqrt{P_c}\beta_c N\hat{\sigma}_k^2 x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\}. \end{aligned} \quad (50)$$

Since the noise n_k has Gaussian distribution, the average over the noise using the integrals of exponential function can be derived as

$$\begin{aligned}&\mathcal{E}_n\left\{e^{\frac{-\left|\sqrt{P_c}\beta_c N\hat{\sigma}_k^2 x_{m,i}^c+\sqrt{P_p}\beta_p N\hat{\sigma}_k^2 x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\} \\ &= e^{-\frac{\left|\sqrt{P_c}\beta_c N\hat{\sigma}_k^2 x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}. \end{aligned} \quad (51)$$

For analytical convenience, in this section we assume that the cell shape is approximated by a circle of radius R, and the users are uniformly distributed in the cell [46]. Hence, the PDF of the users at radius r relative to the BS is [46] $f_d(r) = \frac{2(r-R_0)}{(R-R_0)^2}$, $R_0 \le r \le R$, where R_0 is the closest distance between a user and the BS. Therefore, the average over d_k can be derived as

$$\begin{aligned}&\mathcal{E}_{d_k}\left\{e^{-\frac{\left|\sqrt{P_c}\beta_c N\hat{\sigma}_k^2 x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}\right\} \\ &= \int_{R_0}^{R} e^{-\frac{\left|\sqrt{P_c}\beta_c N\hat{\sigma}_k^2 x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}\frac{2(r-R_0)}{(R-R_0)^2}dr \end{aligned} \quad (52)$$

which can be found using Gaussian Quadrature rules as in (53) shown at the bottom of the page. For the second term ψ, using Jensen inequality the average can be derived as

$$\psi = \log_2\sum_{t=1}^{M^{N-1}} e^{-\frac{\left|\sqrt{tP}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}. \quad (54)$$

ii) Ergodic rate for the private part: the ergodic rate for the private message at the k^{th} user, under finite alphabet

$$\begin{aligned}&\int_{R_0}^{R} e^{-\frac{\left|\sqrt{P_c}\beta_c N\left(\frac{\tau p_p \varpi_k^2}{\tau p_p \varpi_k+1}\right)x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}\frac{2(r-R_0)}{(R-R_0)^2}dr \\ &= \sum_{j=1}^{C}\mathrm{H}_j\, e^{-\frac{\left|\sqrt{P_c}\beta_c N\left(\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\right)x_{m,i}^c+\sqrt{P_p}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}\frac{2\left(\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2} \end{aligned} \quad (53)$$

signaling using ZF precoding technique can be written as

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}} \underbrace{\mathcal{E}\left\{\log_2 \sum\nolimits_{i=1}^{M^{N-1}} e^{\frac{-\left|\mathbf{h}_k \mathbf{W}_{ZF}^p \mathbf{x}_{m,i}^p\right|^2}{2\sigma_k^2}}\right\}}_{\psi}. \tag{55}$$

By using Jensen inequality, and following similar steps as in the previous section, we can find the average of ψ as in (54). Consequently, the ergodic rate for the k^{th} private message using ZF can be expressed as

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}} \log_2 \sum_{t=1}^{M^{N-1}} e^{-\frac{\left|\sqrt{tP}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}. \tag{56}$$

B. RS: MRT/CI

In this scenario weighted MRT technique is used for common message and CI precoding for the private messages.

1) CI PRECODING WITH PSK INPUT ALPHABET

In this scenario, the precoder for the common and the private messages based on the estimated channels can be written, respectively, as

$$\begin{aligned}\hat{\mathbf{w}}_c &= \sum_{i=1}^{K}\beta_c \hat{\mathbf{h}}_i^H \quad \text{and } \hat{\mathbf{W}}^p \\ &= \frac{1}{K}\beta_p\, \hat{\mathbf{H}}^H\left(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right)^{-1}\mathrm{diag}\left\{\hat{\mathbf{V}}^{-1}\mathbf{u}\right\}.\end{aligned} \tag{57}$$

The received signal at user k can be now written as

$$\begin{aligned}\hat{y}_k &= \sqrt{P_c}\beta_c\sum_{i=1}^{K}\hat{\mathbf{h}}_k\hat{\mathbf{h}}_i^H x_c - \sqrt{P_c}\beta_c\sum_{i=1}^{K}\boldsymbol{e}_k\hat{\mathbf{h}}_i^H x_c \\ &+ \sqrt{P_p}\sum_{i=1}^{K}\hat{\mathbf{h}}_k\hat{\mathbf{w}}_i^p x_i - \sqrt{P_p}\sum_{i=1}^{K}\boldsymbol{e}_k\hat{\mathbf{w}}_i^p x_i + n_k.\end{aligned} \tag{58}$$

i) Ergodic rate for the common part: the ergodic rate for the common part at user k under PSK signaling in imperfect CSIT scenario, can be written as in (59) shown at the bottom of the page.

Analysis for large N: using the facts in Lemma 1, (59) becomes as in (60) at the bottom of the page. By invoking Jensen inequality, the first term in (60), φ, can be expressed as in (61) at the bottom of the next page. Since the noise n_k has Gaussian distribution, using the integrals of exponential function, we can find

$$\mathcal{E}_n\left\{e^{\frac{-\left|\sqrt{P_c}\beta_c N\left(\frac{p u\varpi_k^2}{p u\varpi_k+1}\right)x_{m,i}^c+\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p u\varpi_k^2}{p u\varpi_k+1}\right)x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\}$$

$$\begin{aligned}\mathcal{E}\left\{R_k^c\right\} = \log_2 M &- \frac{1}{M^N}\sum_{m=1}^{M^N}\underbrace{\mathcal{E}_{\hat{\mathbf{h}},\boldsymbol{e},n_k}\left\{\log_2\sum_{i=1}^{M^N} e^{\frac{-\left|\hat{\mathbf{h}}_k\hat{\mathbf{W}}\mathbf{x}_{m,i}+\boldsymbol{e}_k\hat{\mathbf{W}}\mathbf{x}_{m,i}+n_k\right|^2}{\sigma_k^2}}\right\}}_{\varphi} \\ &+\frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}}\underbrace{\mathcal{E}_{\hat{\mathbf{h}},\boldsymbol{e},n_k}\left\{\log_2\sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\hat{\mathbf{h}}_k\hat{\mathbf{W}}^p\mathbf{x}_{m,i}^p+\boldsymbol{e}_k\hat{\mathbf{W}}\mathbf{x}_{m,i}+n_k\right|^2}{\sigma_k^2}}\right\}}_{\psi}\end{aligned} \tag{59}$$

$$\begin{aligned}\mathcal{E}\left\{R_k^c\right\} = \log_2 M &- \frac{1}{M^N}\sum_{m=1}^{M^N}\underbrace{\mathcal{E}_{d_k,n_k}\left\{\log_2\sum_{i=1}^{M^N} e^{\frac{-\left|\sqrt{P_c}\beta_c N\left(\frac{p u\varpi_k^2}{p u\varpi_k+1}\right)x_{m,i}^c+\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p u\varpi_k^2}{p u\varpi_k+1}\right)x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\}}_{\varphi} \\ &+\frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}}\underbrace{\mathcal{E}_{d_k,n_k}\left\{\log_2\sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p u\varpi_k^2}{p u\varpi_k+1}\right)x_{m,i}^p+n_k\right|^2}{\sigma_k^2}}\right\}}_{\psi}\end{aligned} \tag{60}$$

$$= e^{-\frac{\left|\sqrt{P_c}\beta_c N\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^c_{m,i}+\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^p_{m,i}\right|^2}{2\sigma_k^2}}. \quad (62)$$

Now, to find the average over the user location, similar to the ZF scenario, we assume that the cell shape is approximated by a circle of radius R and the users are uniformly distributed in the cell [46]. Therefore, we can find the average over d_k using Gaussian Quadrature rules as,

$$\mathcal{E}_{d_k}\left\{e^{-\frac{\left|\frac{p_u}{p_u d_k^m+d_k^{2m}}\right|^2\zeta}{2\sigma_k^2}}\right\} = \int_{R_0}^{R} e^{-\frac{\left|\frac{p_u}{p_u r^m+r^{2m}}\right|^2\zeta}{2\sigma_k^2}}\frac{2(r-R_0)}{(R-R_0)^2}dr \quad (63)$$

$$= \sum_{j=1}^{C}\mathrm{H}_j\, e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\right|^2\zeta}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2} \quad (64)$$

where $\zeta = |\sqrt{P_c}\beta_c N x^c_{m,i} + \sqrt{P_p}N\frac{1}{K}\beta_p u_k x^p_{m,i}|^2$and r_j and H_j are the j^{th} zero and the weighting factors of the Laguerre polynomials, respectively [42]. For the second term, ψ, similarly using Jensen inequality we can write

$$\psi = \mathcal{E}_{d_k,n_k}\log_2\sum_{i=1}^{M^{N-1}} e^{\frac{-\left|\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^p_{m,i}+n_k\right|^2}{\sigma_k^2}}$$

$$\leq \log_2\sum_{i=1}^{M^{N-1}}\mathcal{E}_{d_k,n_k} e^{\frac{-\left|\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^p_{m,i}+n_k\right|^2}{\sigma_k^2}}. \quad (65)$$

Since n_k has Gaussian distribution, we can get

$$\psi = \log_2\sum_{i=1}^{M^{N-1}}\mathcal{E}_{d_k}\left\{e^{-\frac{\left|\frac{p_u}{p_u d_k^m+d_k^{2m}}\right|^2\vartheta}{2\sigma_k^2}}\right\} \quad (66)$$

where $\vartheta = |\sqrt{P_p}N\frac{1}{K}\beta_p u_k x^p_{m,i}|^2$. The average in (66) can be obtained as in (63) and (64), which is given by

$$\mathcal{E}_{d_k}\left\{e^{-\frac{\left|\frac{p_u}{p_u d_k^m+d_k^{2m}}\right|^2\vartheta}{2\sigma_k^2}}\right\} = \int_{R_0}^{R} e^{-\frac{\left|\frac{p_u}{p_u r^m+r^{2m}}\right|^2\vartheta}{2\sigma_k^2}}\frac{2(r-R_0)}{(R-R_0)^2}dr \quad (67)$$

$$= \sum_{j=1}^{C}\mathrm{H}_j\, e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\right|^2\vartheta}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2}. \quad (68)$$

ii) Ergodic rate for the private part: the ergodic rate for the private part at user k under PSK signaling, using CI precoding technique can be written as

$$\mathcal{E}\{R_k^p\} = \log_2 M - \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}} \underbrace{\mathcal{E}_{\mathbf{h},n_k}\left\{\log_2\sum\nolimits_{i=1}^{M^{N-1}} e^{\frac{-\left|\mathbf{h}_k\mathbf{W}^p_{CI}\mathbf{x}^p_{m,i}\right|^2}{2\sigma_k^2}}\right\}}_{\psi}. \quad (69)$$

By using Jensen inequality, and following similar steps as in the previous sub-section we can find the average of ψ in (69) as in (68). Consequently, the ergodic rate for the k^{th} private message using CI can be expressed as

$$\mathcal{E}\{R_k^p\} = \log_2 M - \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}}\log_2\sum_{i=1}^{M^{N-1}}\sum_{j=1}^{C}\mathrm{H}_j \times e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\right|^2\vartheta}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2}. \quad (70)$$

$$\varphi = \mathcal{E}_{d_k,n_k}\left\{\log_2\sum_{i=1}^{M^N} e^{\frac{-\left|\sqrt{P_c}\beta_c N\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^c_{m,i}+\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^p_{m,i}+n_k\right|^2}{\sigma_k^2}}\right\}$$

$$\leq \log_2\sum_{i=1}^{M^N}\mathcal{E}_{d_k,n_k}\left\{e^{\frac{-\left|\sqrt{P_c}\beta_c N\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^c_{m,i}+\sqrt{P_p}N\frac{1}{K}\beta_p u_k\left(\frac{p_u\varpi_k^2}{p_u\varpi_k+1}\right)x^p_{m,i}+n_k\right|^2}{\sigma_k^2}}\right\} \quad (61)$$

2) CI PRECODING WITH QAM INPUT ALPHABET

In this scenario, the precoder for the private messages based on the estimated channels can be written as [43]

$$\mathbf{W}_{CI}^{p} = \frac{1}{K}\beta_p\, \hat{\mathbf{H}}^H\Big(\hat{\mathbf{H}}\hat{\mathbf{H}}^H\Big)^{-1}\mathbf{U}\mathrm{diag}\Big\{\mathbf{F}^{-1}\tilde{\mathbf{V}}^{-1}\mathbf{u}_1\Big\}\mathbf{x}^p\hat{\mathbf{x}}^{pT}. \tag{71}$$

i) Ergodic rate for the common part: by substituting (71) into (59) and following similar steps as in PSK scenario, the average of the first and second terms in (59), φ and ψ, can be found, respectively, as

$$\varphi = \sum_{j=1}^{C}\mathrm{H}_j\, e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\zeta\right|^2}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}\,r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2} \tag{72}$$

$$\psi = \sum_{j=1}^{C}\mathrm{H}_j\, e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\vartheta\right|^2}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}\,r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2}. \tag{73}$$

where $\zeta = |\sqrt{P_c}\beta_c N x_{m,i}^c + \sqrt{P_p}N\frac{1}{K}\beta_p u_k[\tilde{x}_{m,i}^p]_k|^2$, $\vartheta = |\sqrt{P_p}N\frac{1}{K}\beta_p u_k[\tilde{x}_{m,i}^p]_k|^2$, r_j and H_j are the j^{th} zero and the weighting factors of the Laguerre polynomials, respectively [42].

Proof: Due to the paper length limitation and in order to avoid the repetition, the proofs of (72) and (73) have been omitted. ∎

ii) Ergodic rate for the private part: by substituting (71) in (69), and following similar steps as in the previous subsection we can find the average of the term ψ in (69) as in (73). Consequently, the ergodic rate for the k^{th} private message using CI can be expressed as [3], [39], [40],

$$\mathcal{E}\left\{R_k^p\right\} = \log_2 M - \frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}}\log_2\sum_{i=1}^{M^{N-1}}\sum_{j=1}^{C}\mathrm{H}_j \times e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\vartheta\right|^2}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}\,r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2}. \tag{74}$$

C. CONVENTIONAL TRANSMISSION NORS

The rate at the k^{th} user in conventional transmission without RS is expressed by

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N}\underbrace{\mathcal{E}\left\{\log_2\sum_{i=1}^{M^N}e^{\frac{-\left|\mathbf{h}_k\mathbf{W}\mathbf{x}_{m,i}\right|^2}{2\sigma_k^2}}\right\}}_{\psi}. \tag{75}$$

For sake of comparison with using RS technique in this scenario, we study approximation of the ergodic user rate based on large number of antennas. In ZF scenario the precoding matrix is given in (45), and then the expectation in (75) can be derived using Jensen inequality as in (54). Consequently, the ergodic rate for the k^{th} user without RS using ZF can be expressed as

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N}\log_2\sum_{t=1}^{M^N}e^{-\frac{\left|\sqrt{t P}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}}. \tag{76}$$

In case CI with PSK signals the precoding matrix is given in (57), and the expectation in (75) can be derived using Jensen inequality as in (66) and (68). Consequently, the ergodic rate for the k^{th} user without RS is

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N}\log_2\sum_{i=1}^{M^N}\sum_{j=1}^{C}\mathrm{H}_j \times e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\vartheta\right|^2}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}\,r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2}. \tag{77}$$

In case CI with QAM, the expectation in (75) can be derived using Jensen inequality as in (74). Consequently, the ergodic rate for the k^{th} user without RS is

$$\mathcal{E}\left\{R_k^{NoRS}\right\} = \log_2 M - \frac{1}{M^N}\sum_{m=1}^{M^N}\log_2\sum_{i=1}^{M^N}\sum_{j=1}^{C}\mathrm{H}_j \times e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^m+\left(\frac{R-R_0}{2}r_j+\frac{R+R_0}{2}\right)^{2m}}\vartheta\right|^2}{2\sigma_k^2}} \times \frac{2\left(\left(\frac{R-R_0}{2}\,r_j+\frac{R+R_0}{2}\right)-R_0\right)}{(R-R_0)^2}. \tag{78}$$

From the rate expressions presented in this Section for MRT/ZF and MRT/CI, we can observe the following. Firstly, it is clear that the imperfect estimation of the channels has a baneful impact on the achievable sum-rates. The analysis in this section clearly reveals that the channel estimation errors deteriorate the achievable sum-rates of the considered precoding schemes. Furthermore, when number of BS antennas, N, is very large, the small-scaling fading is averaged out due to the hardening effect of the large scale antenna system, and the ergodic rate becomes dependent essentially on the distribution of the users location in the network. Therefore, the ergodic sum-rates in this case are limited by the variances of the estimated channels, the estimation errors, and hence by the large-scale fading. Actually, the large scale fading has influence on the estimated channels and plays important role in the achievable rates in this scenario. Moreover, from these expressions we can also notice that, in communication systems with finite alphabet signals, the rate increases as the transmit SNR increases and at high SNR the rate for both perfect and imperfect CSI saturate at $\log_2 M$, thus the rate loss is negligible at high SNR. On the other hand, at very low SNR the ergodic rates in the cases of both perfect and imperfect CSI approach zero, and the rate loss is also negligible at low SNR values. It is also obvious that, the achievable sum-rate increases with the number of the BS antennas N and number of users K. In addition, we would like to mention that, all the derived expressions are very accurate, explicit and in closed-form, thus effective power allocation techniques to enhance the sum-rates or provide fairness among the users can be developed from the derived expressions provided in this paper.

V. RATE MAXIMIZATION THROUGH RS POWER ALLOCATION

In this section, we formulate a power allocation problem for maximizing the ergodic sum-rate of the RS transmission schemes described in the previous sections. The optimal value of t can be obtained by solving the following problem

$$\max_{0\le t\le 1}\left(\min_j\left(\mathcal{E}\left\{R_j^c\right\}\right)_{j=1}^{K}+\sum_{k=1}^{K}\mathcal{E}\left\{R_k^p\right\}\right) \tag{79}$$

where the ergodic rates for the common part $\mathcal{E}\{R_j^c\}$ and the private part $\mathcal{E}\{R_k^p\}$ for ZF and CI have been derived in (Section III-A), (Section III-B), (Section (IV-A)) and (Section (IV-B)). It is worth noting that the availability of perfect CSIT enables the BS to maximize the instantaneous sum-rate by adapting the power split among the common and private messages based on the channel status. Consequently, following [9], the maximization in (79) can be moved inside the expectation and the optimum solution can be found for each channel state. In case the BS has imperfect CSIT, the BS can not evaluate the instantaneous rates, but it can access the average rates which are the expected rates for a given channel estimate. Hence, maximizing the ergodic sum-rate under imperfect CSIT can be achieved for each estimated channel [9]. For simplicity, we consider ergodic sum-rate maximization problem in the two scenarios. On one hand, the analytical optimization for the case of finite constellation signaling using the derived formulas above becomes intractable. On the other hand, the optimal t can be obtained by a simple one dimensional search over $0 \le t \le 1$. Hence, the optimal t can be found by using line search methods such as golden section technique. The overall steps of golden section method to obtain the optimal t is stated in Algorithm 1 [4]. Moreover, in order to reduce the complexity, two sub-optimal solutions can be considered in finite alphabet scenarios, as follows.

Algorithm 1 Golden Section Method

Initialize $\varrho = 0$, $\zeta = 1$, and $\lambda = \frac{-1+\sqrt{5}}{2}$.
Repeat
Update $t_1 = \varrho + (1-\lambda)\zeta$ and $t_2 = \zeta + (1-\lambda)\varrho$.
Obtain $R(t_1)$ and $R(t_2)$ from (4).
If $R(t_1) > R(t_2)$, set $\varrho = t_1$. Else set $\zeta = t_2$.
Until $|\varrho - \zeta|$ converges.
Find $t^* = (\varrho + \zeta)/2$.

- In the first solution, we allocate a fraction t of the total power for the private messages to achieve the same sum-rate as the conventional techniques with full power. Then, the remaining power can be allocated for the common message, as considered in [10]. The sum-rate payoff of the RS scheme over the NoRS can be determined by

$$\Delta R = \mathcal{E}\{R_c\} + \sum_{k=1}^{K}\left(\mathcal{E}\{R_k^p\} - \mathcal{E}\left\{R_k^{NoRS}\right\}\right) \tag{80}$$

Consequently, the ratio t that achieves the superiority can be obtained by satisfying the equality, $\mathcal{E}\{R_k^p\} = \mathcal{E}\{R_k^{NoRS}\}$. The necessary condition is $\mathcal{E}\{R_k^p\} \le \mathcal{E}\{R_k^{NoRS}\}$. The equality holds when the power splitting ratio t achieves $\mathcal{E}\{R_k^p\} \approx \mathcal{E}\{R_k^{NoRS}\} = r_k^{NoRS}$.

- In the first case perfect CSIT MRT/ZF we can write

$$\frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}}\log_2\sum_{i=1}^{M^{N-1}} e^{-\frac{\left|\sqrt{tP}\beta_p\left[\mathbf{x}_{m,i}^p\right]_k\right|^2}{2\sigma_k^2}} \approx X - r_k^{NoRS} \tag{81}$$

where $X = \log_2 M$. The optimal t in this case is the value that can satisfy (81), we can find this value numerically by changing t from 0 to 1.

- In the second case perfect CSIT MRT/CI we can write[12]

$$\frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}}\log_2\sum_{i=1}^{M^{N-1}}\Lambda_{k_{m,i}} \approx X - r_k^{NoRS} \tag{82}$$

The optimal t is the value that can achieve (82), which can be obtained by changing t from 0 to 1.

- In the third case imperfect CSIT MRT/ZF we can write

$$\frac{1}{M^{N-1}}\sum_{m=1}^{M^{N-1}}\log_2\sum_{t=1}^{M^{N-1}} e^{-\frac{\left|\sqrt{P_p}\beta_p x_{m,i}^p\right|^2}{2\sigma_k^2}} \approx X - r_k^{NoRS}. \tag{83}$$

12. Due to the paper length limitation, we use the sum-rate expressions for CI with PSK signals.

The optimal t in this case is the value that can satisfy (83), which can be obtained numerically by changing t from 0 to 1.

- In the last case imperfect CSIT MRT/CI the expression can be written as in (84) at the bottom of the page, where $\Xi = \frac{2((\frac{R-R_0}{2} r_j + \frac{R+R_0}{2}) - R_0)}{(R-R_0)^2}$. Therefore the optimal t in this case is the value that can achieve (84), which can be obtained easily by changing t from 0 to 1.

- In the second solution, since the achievable data rate in the finite alphabet systems saturates at maximum predefined value ($R_m = (K+1)\log_2 M$), here at high SNR the optimal value of t is the value that achieves the maximum rate with less transmit power P, as in the following expression

$$(K+1)\log_2 M = \min_j \Big(\mathcal{E}\Big\{R_j^c\Big\}\Big)_{j=1}^{K} + \sum_{k=1}^{K} \mathcal{E}\{R_k^p\} \tag{85}$$

Therefore, the optimum value of t at high SNR is the value that satisfies (85) with minimum power P. For simplicity, in this solution t can be obtained by substituting $r_k^{NoRS} = \log_2 M$ in the all above equations (81), (82), (83) and (84).

VI. NUMERICAL RESULTS

In this section, we present numerical results of the analytical expressions derived in this work. Monte-Carlo simulations are conducted where the channel coefficients are randomly generated. The path loss exponent is chosen to be $m = 2.7$, and $p_u = 5\,\text{W}$. Assuming the users have same noise power, σ^2, and the total transmission power is p, the transmit signal to noise ratio (SNR) is defined as SNR $= \frac{p}{\sigma^2}$. The optimal value of t is obtained in this section by using Algorithm 1, and the elements of vector $\mathbf{u}$ has been chosen to be, $u_k = \frac{1}{K}$.

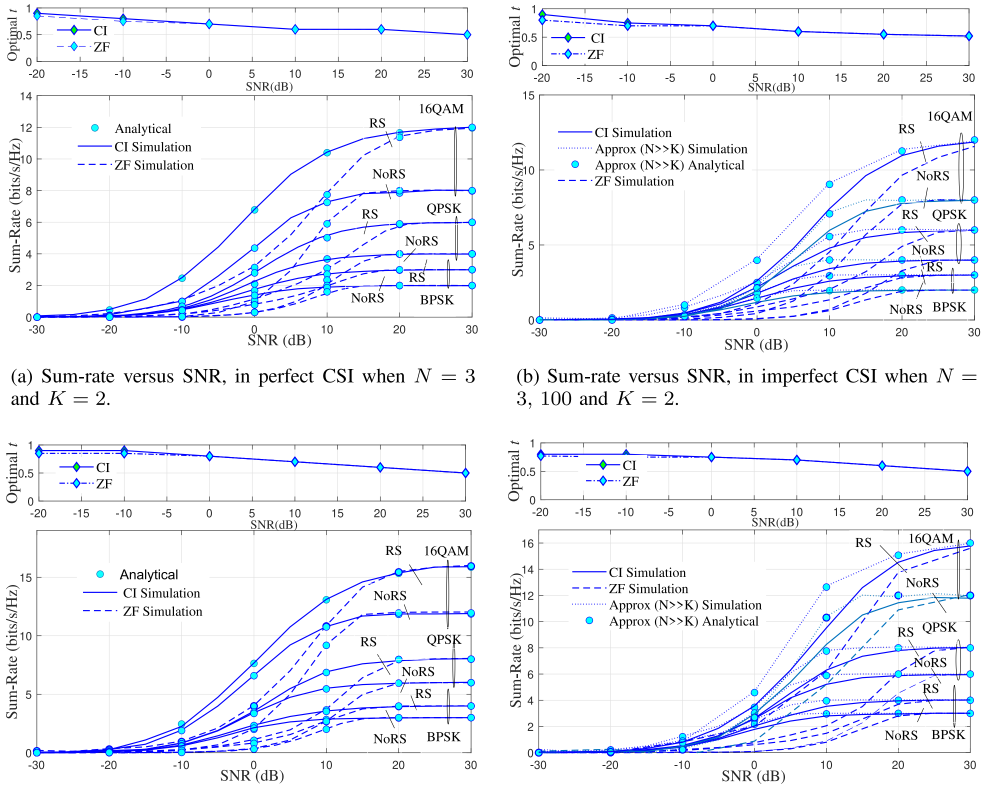
Firstly, in Fig. 1, we illustrate the sum-rate for the RS and NoRS using MRT-CI and MRT-ZF in perfect CSIT scenario and imperfect CSIT scenario, respectively, subject to BPSK, QPSK and 16 QAM when the distances are normalized to unit value. Fig. 1(a) and Fig. 1(c) present the sum-rate in perfect CSIT scenario when $N = 3$, $K = 2$ and $N = 4$, $K = 3$, respectively, while Fig. 1(b) and Fig. 1(d) show the sum-rate in imperfect CSIT scenario when $N = 3$, $K = 2$ and $N = 4$, $K = 3$, and $N = 100$ for the approximation results. The good agreement between the analytical and simulated results confirms the validity of the analysis introduced in this paper. Several observations can be extracted from these figures. Firstly, it is clear that the sum rate saturates at a certain SNR value, owing to the finite constellation. Secondly, the RS scheme enhances the sum-rate of the considered system and tackles the sum-rate saturation occurred in the communication systems with PSK signaling. In addition, it is evident that the CI precoding techniques outperforms the ZF technique in the all considered scenarios for a wide SNR range with an up to 10dB gain in the SNR for a given sum rate. Furthermore, as anticipated the system performance degrade notably in the imperfect CSIT scenario. In addition, we can observe that when the number of BS antennas is high $N \gg K$, the ZF achieves the same performance as the CI; ZF precoding can be considered as a special case of the CI precoding technique [29].

To elaborate more on the impact of the number of BS antennas and number of users on the system performance. In Fig. 2 we plot the sum-rate versus the SNR for the considered transmission schemes with BPSK, and QPSK, when the distances are normalized to unit value. Fig. 2(a) present the sum-rate in perfect CIST scenario when $N = 8$, and $K = 4$ and Fig. 2(b) shows the sum-rate in imperfect CSIT scenario, when $N = 8$, and $K = 4$ and $N = 100$ for the approximation results. looking closer at the results in this figure, it is clear that increasing the number of users K and/or the number of antennas N results in enhancing the achievable sum-rate in all the considered scenarios. From these results, it is clear that increasing number of users K and/or the number of antennas N results in enhancing the achievable sum-rate, and reducing the gap performance between the CI and ZF in all the considered scenarios.

Furthermore, in Fig. 3 we show the sum-rate when the users are uniformly distributed inside a circle area with a radius of 40m and the BS is located at the center of this area, for two system set up $N = 3$, $K = 2$ and $N = 4$, $K = 3$ and $N = 100$ for the approximation results. The well agreement between the analytical and simulated results confirms the analysis provided in this paper. Comparing the results in this figure with that in Fig. 1, one can notice that, in general, increasing the distance always degrades the achievable sum rates. In addition, when the distance between the BS and the users increases the rate saturation occurs at high SNR values, due to larger path-loss. It is also clear that, the superiority of RS with CI over RS with ZF and NoRS does not depend on the users' locations.

Moreover, from the results presented in Figs. 1–3 we can notice that, the optimal value of the power fraction t at low SNR values is close to one, which means that splitting the messages and transmitting a common message is not beneficial in this SNR range. In this case only the private messages are transmitted and the RS degenerates to NoRS. This is because the users are experienced similar SNR. On the other hand at high SNR values the optimal value of t is less than one, $t < 1$, which indicates that the common message is transmitted with the remaining power beyond the

$$\frac{1}{M^{N-1}} \sum_{m=1}^{M^{N-1}} \sum_{j=1}^{C} \mathrm{H}_j \,\Xi \times e^{-\frac{\left|\frac{p_u}{p_u\left(\frac{R-R_0}{2} r_j + \frac{R+R_0}{2}\right)^m + \left(\frac{R-R_0}{2} r_j + \frac{R+R_0}{2}\right)^{2m}}\right|^2 \left|\sqrt{P_p} N \frac{1}{K} \beta_p u_k x_{m,i}^p\right|^2}{2\sigma_k^2}} \approx X - r_k^{NoRS} \tag{84}$$



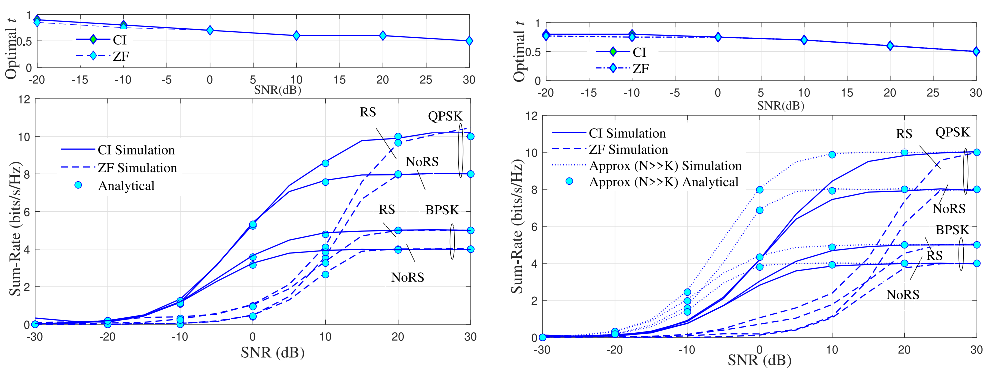
(a) Sum-rate versus SNR, in perfect CSI when $N = 3$ and $K = 2$.

(b) Sum-rate versus SNR, in imperfect CSI when $N = 3, 100$ and $K = 2$.

(c) Sum-rate versus SNR, in perfect CSI when $N = 4$ and $K = 3$.

(d) Sum-rate versus SNR, in imperfect CSI when $N = 4, 100$ and $K = 3$.

FIGURE 1. **Sum-rate versus SNR for RS and NoRS with different types of input in perfect and imperfect CSI, when the distances are normalized to unit value.**



(a) Sum-rate versus SNR, in perfect CSI when $N = 8$ and $K = 4$.

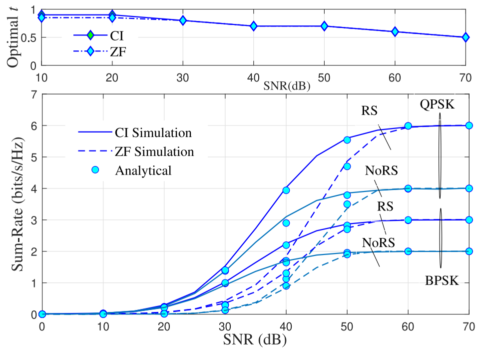
(b) Sum-rate versus SNR, in imperfect CSI when $N = 8, 100$ and $K = 4$.

FIGURE 2. **Sum-rate versus SNR for RS and NoRS with different types of input in perfect and imperfect CSI, when the distances are normalized to unit value.**

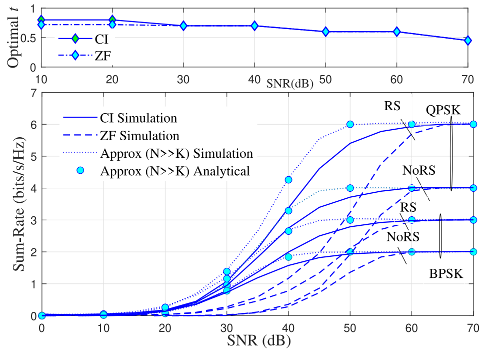
saturation of the private message transmission. Therefore, the optimal value of t decreases with increasing the SNR values. This observation is elaborated in the next figures.

In order to clearly illustrate the impact of the power fraction t on the system performance, we plot in Fig. 4 the sum-rate versus SNR for various values of t with the CI precoding under QPSK, when $N = 3$, $K = 2$, $d_1 = 1$m and $d_2 = 5$m. Interestingly enough, it is noted that at low SNR values, SNR ≤ 12 dB, the sum-rate degrades as t becomes small, and the optimal t in this range is approximately close to 1. In addition, at high SNR values, SNR ≥ 12 dB, the sum-rate degrades as the value of t increases, till the sum-rate reaches the achievable rate in case NoRS when $t = 1$.

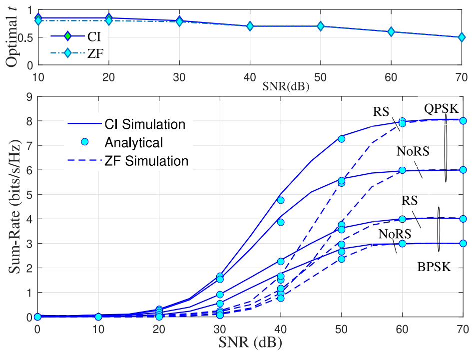
Finally, in order to compare the sub-optimal solutions of the power allocation with the optimal algorithm, in Fig. 5 we plot the sum-rate versus the transmit SNR for the optimal power allocation scheme and the sum-optimal schemes, when $N = 3$, $K = 2$, $d_1 = 1$m and $d_2 = 5$m, under BPSK signals. Notable and as anticipated, the first sub-optimal solution gives similar results to Algorithm 1 (optimal) in low SNR values, and lower sum-rate in high SNR values. On contrary,



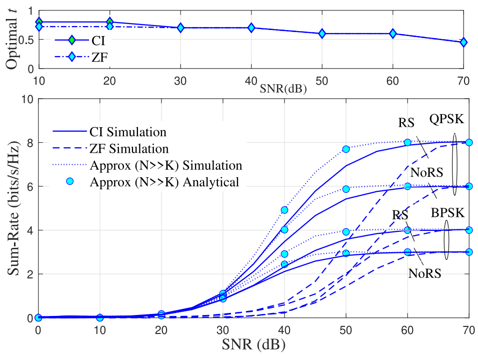
(a) Sum-rate versus SNR in perfect CSIT, when $N = 3$ and $K = 2$.



(b) Sum-rate versus SNR in imperfect CSIT, when $N = 3, 100$ and $K = 2$.



(c) Sum-rate versus SNR in perfect CSIT, when $N = 4$ and $K = 3$.



(d) Sum-rate versus SNR in imperfect CSIT, when $N = 4, 100$ and $K = 3$.

FIGURE 3. Sum-rate versus SNR with different types of input when the users are randomly distributed.

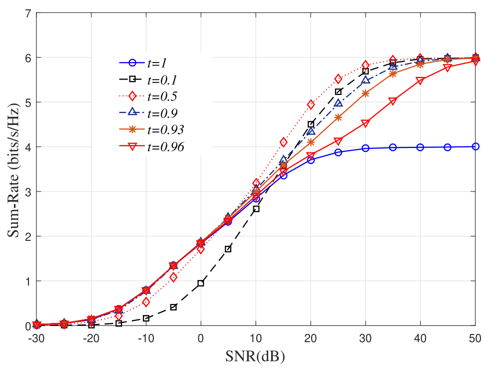


FIGURE 4. Sum-rate for RS with CI and QPSK modulation versus SNR for various values of *t*.

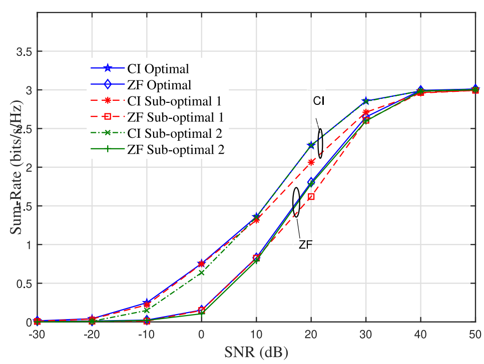


FIGURE 5. Sum-rate for RS with BPSK versus SNR for various power allocation schemes.

the second sub-optimal solution which is proposed based on high SNR values, produces similar results to Algorithm 1 (optimal) in high SNR regime, and smaller sum-rate at low SNRs.

VII. CONCLUSION

In this paper we employed the CI precoding technique to enhance the sum-rate performed by RS scheme in MU-MISO systems with finite alphabet inputs. New analytical expressions for the ergodic sum-rate have been derived for ZF precoding technique and CI precoding technique in RS and NoRS scenarios. Furthermore, a power allocation scheme that achieves superiority of RS over NoRS in the presence of finite constellation was proposed. The results presented in this work demonstrated that RS with CI has greater sum-rate than RS with ZF and NoRS transmission techniques. In addition, increasing the number of BS antennas and/ or the number of users enhances the achievable sum-rate.

REFERENCES

[1] A. Salem and C. Masouros, "Rate splitting approach under PSK signaling using constructive interference precoding technique," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Marrakesh, Morocco, 2019, pp. 1–6.

[2] M. S. John G. Proakis, *Digital Communications*, 5th ed. New York, NY, USA: McGraw-Hill, 2008.

[3] Y. Wu, C. Xiao, X. Gao, J. D. Matyjas, and Z. Ding, "Linear precoder design for MIMO interference channels with finite-alphabet signaling," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3766–3780, Sep. 2013.

[4] A. Salem, L. Musavian, and K. A. Hamdi, "Wireless power transfer in distributed antenna systems," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 737–747, Jan. 2019.

[5] B. Clerckx, H. Joudeh, C. Hao, M. Dai, and B. Rassouli, "Rate splitting for MIMO wireless networks: A promising PHY-layer strategy for LTE evolution," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 98–105, May 2016.

[6] H. Joudeh and B. Clerckx, "Robust transmission in downlink multiuser MISO systems: A rate-splitting approach," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6227–6242, Dec. 2016.

[7] C. Hao, Y. Wu, and B. Clerckx, "Rate analysis of two-receiver MISO broadcast channel with finite rate feedback: A rate-splitting approach," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3232–3246, Sep. 2015.

[8] A. Papazafeiropoulos and T. Ratnarajah, "Rate-splitting robustness in multi-pair massive MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5623–5636, Aug. 2018.

[9] H. Joudeh and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A rate-splitting approach," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4847–4861, Nov. 2016.

[10] M. Dai, B. Clerckx, D. Gesbert, and G. Caire, "A rate splitting strategy for massive MIMO with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4611–4624, Jul. 2016.

[11] C. Hao and B. Clerckx, "MISO networks with imperfect CSIT: A topological rate-splitting approach," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2164–2179, May 2017.

[12] B. Clerckx, Y. Mao, R. Schober, and H. V. Poor, "Rate-splitting unifying SDMA, OMA, NOMA, and multicasting in MISO broadcast channel: A simple two-user rate analysis," *IEEE Wireless Commun. Lett.*, vol. 9, no. 3, pp. 349–353, Mar. 2020.

[13] B. Clerckx *et al.*, "Is noma efficient in multi-antenna networks? A critical look at next generation multiple access techniques," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 1310–1343, 2021.

[14] H. Joudeh and B. Clerckx, "Rate-splitting for max–min fair multigroup multicast beamforming in overloaded systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7276–7289, Nov. 2017.

[15] Y. Mao, B. Clerckx, and V. O. K. Li, "Rate-splitting multiple access for downlink communication systems: Bridging, generalizing, and outperforming SDMA and NOMA," *EURASIP J. Wireless Commun. Netw.*, vol. 133, May 2018, doi: 10.1186/s13638-018-1104-7.

[16] C. Masouros and E. Alsusa, "Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1396–1404, Mar. 2009.

[17] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector perturbation based on symbol scaling for limited feedback MISO downlinks," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 562–571, Feb. 2014.

[18] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3628–3640, Jul. 2015.

[19] C. Masouros and E. Alsusa, "Two-stage transmitter precoding based on data-driven code-hopping and partial zero forcing beamforming for MC-CDMA communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3634–3645, Jul. 2009.

[20] E. Alsusa and C. Masouros, "Adaptive code allocation for interference management on the downlink of DS-CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2420–2424, Jul. 2008.

[21] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[22] S. Timotheou, G. Zheng, C. Masouros, and I. Krikidis, "Exploiting constructive interference for simultaneous wireless information and power transfer in multiuser downlink systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1772–1784, May 2016.

[23] M. R. A. Khandaker, C. Masouros, and K.-K. Wong, "Constructive interference based secure precoding: A new dimension in physical layer security," *IEEE Trans. Inf. Forensics Security*, vol. 13, pp. 2256–2268, 2018.

[24] P. V. Amadori and C. Masouros, "Large scale antenna selection and precoding for interference exploitation," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4529–4542, Oct. 2017.

[25] F. A. Khan, C. Masouros, and T. Ratnarajah, "Interference-driven linear precoding in multiuser MISO downlink cognitive radio network," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2531–2543, Jul. 2012.

[26] M. Alodeh *et al.*, "Symbol-level and multicast precoding for multiuser multiantenna downlink: A state-of-the-art, classification, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1733–1757, 3rd Quart., 2018.

[27] A. Garcia-Rodriguez and C. Masouros, "Power-efficient tomlinson-harashima precoding for the downlink of multi-user MISO systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1884–1896, Jun. 2014.

[28] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Maximizing energy efficiency in the vector precoded MU-MISO downlink by selective perturbation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4974–4984, Sep. 2014.

[29] A. Li and C. Masouros, "Interference exploitation precoding made practical: Optimal closed-form solutions for PSK modulations," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7661–7676, Nov. 2018.

[30] A. Salem and C. Masouros, "On the finite constellation sum rates for ZF and CI precoding," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Marrakesh, Morocco, Apr. 2019, pp. 1–6.

[31] A. Salem and C. Masouros, "Error probability analysis and power allocation for interference exploitation over Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, early access, Apr. 8, 2021, doi: 10.1109/TWC.2021.3069803.

[32] A. Salem, C. Masouros, and K.-K. Wong, "On the secrecy performance of interference exploitation with PSK: A non-Gaussian signaling analysis," *IEEE Trans. Wireless Commun.*, early access, May 25, 2021, doi: 10.1109/TWC.2021.3080686.

[33] A. Salem, C. Masouros, and K.-K. Wong, "Sum rate and fairness analysis for the MU-MIMO downlink under PSK signalling: Interference suppression vs exploitation," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6085–6098, Sep. 2019.

[34] A. Li *et al.*, "A tutorial on interference exploitation via symbol-level precoding: Overview, state-of-the-art and future directions," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 796–839, 2nd Quart., 2020.

[35] Y.-W. P. Hong, W.-J. Huang, and C.-C. J. Kuo, *Cooperative Communications and Networking Technologies and System Design*. New York, NY, USA: Springer, 2010.

[36] J. Proakis, *Digital Communications*, 3rd ed. New York, NY, USA: McGraw-Hill, 1995.

[37] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1273–1289, May 2001.

[38] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York, NY, USA: Wiley, 1982.

[39] W. Wu, K. Wang, W. Zeng, Z. Ding, and C. Xiao, "Cooperative multi-cell MIMO downlink precoding with finite-alphabet inputs," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 766–779, Mar. 2015.

[40] Y. Wu, M. Wang, C. Xiao, Z. Ding, and X. Gao, "Linear precoding for MIMO broadcast channels with finite-alphabet constraints," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2906–2920, Aug. 2012.

[41] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Washington, DC, USA: U.S. Dept. Commerce, 1972.

[42] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. Washington, DC, USA: U.S. Dept. Commerce, 1972.

[43] A. Li, C. Masouros, Y. Li, B. Vucetic, and A. L. Swindlehurst, "Interference exploitation precoding for multi-level modulations: Closed-form solutions," *IEEE Trans. Commun.*, vol. 69 no. 1, pp. 291–308, Jan. 2021.

[44] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, May 2006.

[45] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[46] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1047–1066, Jul. 1999.