Rate-Splitting Multiple Access to Mitigate the Curse of Mobility in (Massive) MIMO Networks

Onur Dizdar, *Member, IEEE,* Yijie Mao, *Member, IEEE,* and Bruno Clerckx, *Senior Member, IEEE*

***Abstract*—Rate-Splitting Multiple Access (RSMA) is a robust multiple access scheme for downlink multi-antenna wireless networks. RSMA relies on multi-antenna Rate-Splitting (RS) at the transmitter and Successive Interference Cancellation (SIC) at the receivers. In this work, we study the performance of RSMA under the important setup of imperfect Channel State Information at the Transmitter (CSIT) originating from user mobility and latency/delay (between CSI acquisition and data transmission) in the network. We derive a lower bound on the ergodic sum-rate of RSMA for an arbitrary number of transmit antennas, number of users, user speed and transmit power. Then, we study the power allocation between common and private streams and obtain a closed-form solution for optimal power allocation that maximizes the obtained lower bound. The proposed power allocation greatly reduces precoder design complexity for RSMA. By Link-Level Simulations (LLS), we demonstrate that RSMA with the proposed power allocation is robust to the degrading effects of user mobility and has significantly higher performance compared to conventional multi-user (massive) Multiple-Input Multiple-Output (MIMO) strategies. The work has important practical significance as results demonstrate that, in contrast to conventional multi-user (massive) MIMO whose performance collapse under mobility, RSMA can maintain reliable multi-user connectivity in mobile deployments.**

***Index Terms*—Rate-splitting, multi-antenna broadcast channel, outdated CSIT, ergodic sum-rate.**

I. Introduction

An essential requirement for the next generation communication systems is to meet the demand for exponentially increasing data rate and accessibility. A key technology expected to meet such requirement is small-scale and large-scale (massive) Multi-User Multiple-Input Multiple-Output (MU-MIMO). Aside having been studied in a vast amount of literature, MU-MIMO concept has also been used in real life wireless communications standards, with 4G and 5G being among the most notable ones. The performance achieved by beamforming and interference management methods in MU-MIMO heavily relies on the Channel State Information (CSI) accuracy. However, acquiring accurate CSI at the Transmitter (CSIT) is a major challenge in real life scenarios. CSIT is commonly subject to errors, which is known as imperfect CSIT. Imperfect CSIT is a common problem in wireless communication networks and causes severe multi-user interference in multi-antenna Broadcast Channel (BC).

A major cause of imperfect CSIT in modern wireless communication networks is the CSI acquisition/feedback latency from mobile users, which results in outdated CSIT due to the variations in the wireless medium under mobility. Indeed, one of the earliest problems reported in field trials for MU-MIMO communications in 5G is outdated CSIT due to mobility. The processing delay in the system easily becomes larger than the coherence time of the channel even when the users are moving with a speed of 30km/h, making the CSIT inaccurate for multi-user beamforming [1]. Moreover, mobility scenarios are expected to become increasingly important in next generation wireless communication systems. In addition to the broadband data services, use cases such as Vehicle-to-Everything (V2X), very high speed trains and transportation, and very low earth orbit satellite systems, are especially prone to the disruptive effects of mobility. Therefore, maintaining high-rate multi-user connectivity and communications in mobile networks becomes a crucial and fundamental problem in future communication systems [2]–[4]. Therefore, it is of paramount importance to seek strategies that can mitigate the detrimental effects induced by user mobility and latency in multi-user multi-antenna communications.

The most common method of achieving MU-MIMO communications is to use multi-user linear precoding techniques to separate users' streams at the transmitter. At the receivers, multi-user interference is treated as noise during signal detection. Such strategy is often named as Space Division Multiple Access (SDMA). A major drawback of SDMA is the performance degradation due to the abovementioned problem of imperfect CSIT. Studies investigating both theoretical (e.g., Degrees-of-Freedom (DoF) [5]) and practical (e.g., throughput [1]) metrics demonstrate the dependence of SDMA performance on the CSIT quality. Therefore, maintaining multi-user connectivity and performance improvements under mobility conditions remain a crucial problem for SDMA-assisted Multi-User (Massive) MIMO.

A method to tackle the issues of latency and mobility in SDMA-assisted Multi-User (Massive) MIMO is to perform switching between Single-User (SU) and Multi-User (MU) transmission modes [6], [7]. Such an approach restricts transmission to a small number of users (even to the SU case in the extreme case) so as to reduce the interference level and prevent a sum-rate loss in the network. Consequently, MU communications cannot be maintained unless interference is reduced to certain levels. Another strategy is to play at the channel acquisition level and conduct channel prediction and reconstruction at the transmitter [8]–[12]. However, the complexity of channel prediction algorithms becomes quickly overwhelming with an increasing number of transmit and receive antennas and subcarriers in an Orthogonal Frequency

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London, UK. e-mail: ({o.dizdar,y.mao16,b.clerckx}@imperial.ac.uk).
This research was partially funded by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under grant EP/R511547/1.

Division Multiplexing (OFDM) MU-MIMO systems.

In this work, we depart from conventional precoding techniques and take a different approach that is inherently robust to imperfect CSIT and multi-user interference to tackle the mobility problem. We consider Rate-Splitting Multiple Access (RSMA), which is a multiple access technique for multi-antenna Broadcast Channel (BC) that relies on linearly precoded Rate-Splitting (RS) at the transmitter and Successive Interference Cancellation (SIC) at the receivers [5], [13]. RSMA is based on splitting the user messages into common and private parts at the transmitter. The common parts of the split messages of each user are combined and encoded into a single common message to be decoded by all users. The single common message and the private parts of the user messages are encoded separately into common and private streams, precoded and superposed in a non-orthogonal manner. Thanks to the message split and the presence of common and private streams, RSMA manages interference by partially decoding the interference and treating the remaining interference as noise by means of SIC receivers.

Originally introduced for the Single-Input Single-Output (SISO) Interference Channel (IC) [14], [15], RS has been adapted to multi-antenna BC due to recent information theoretic studies showing that RS achieves the entire DoF region of K-user Multiple-Input Single-Output (MISO) BC with partial CSIT [16], [17], and outperforms Dirty Paper Coding (DPC) [18]. The performance gain of RSMA comes from its ability to manage the interference in a flexible manner (namely bridging the extremes of treating multi-user interference as noise, as in SDMA, and fully decoding the multi-user interference, as in Non-Orthogonal Multiple Access (NOMA)) by splitting the messages of the users and performing SIC at the receivers. Being a multiple access scheme designed for BC with multiple transmit (and receive) antennas, RSMA is naturally a promising candidate for operating under interference in multi-antenna networks. Indeed, numerous studies have demonstrated that RSMA bridges and outperforms existing multiple access schemes, i.e., SDMA, NOMA, Orthogonal Multiple Access (OMA) and multicasting [13], [17], [19]. In its simplest form, 1-layer RSMA relies on a single common stream and requires a single SIC at each receiver [13].

In the context of RSMA use in practical systems, 1-layer RSMA with low-complexity precoders for the common and private streams is considered. A problem arising in such system is how to allocate the total transmit power between the common and private streams in an optimal manner for an improved sum-rate performance under imperfect CSIT.

A. Related Work

An analysis of sum-rate for 1-layer RSMA for the two-user MISO BC is studied in [20] with ZF precoders for the private streams and perfect CSIT. The obtained sum-rate expression is used to derive a power allocation method to distribute power between the common and private streams for the considered scenario.

The authors investigate power allocation between the common and private streams of 1-layer RSMA for the two-user MISO BC in [21] under channel quantization error and ZF precoders for the private streams. The channel estimation error is bounded (i.e, isotropically distributed) and the interference from the private stream of the unintended user is neglected, which is reasonable for a scenario with two users and a low level of channel quantization error. Equal power allocation is considered among the private streams of the two users. The obtained power allocation method is optimal in the sense that it minimizes the ergodic sum-rate difference of SDMA with perfect CSIT and RSMA with imperfect CSIT.

Another work on the power allocation between the common and private streams is [22], which considers a massive MIMO setup with imperfect CSIT and unbounded channel estimation error (i.e, Complex Gaussian distributed). The sum-rate analysis and obtained closed-form power allocation algorithm are based on the asymptotic case of a large number of transmit antennas and number of users.

Precoder design and power allocation based on second-order channel statistics are investigated for RSMA with hybrid precoding for Millimeter Wave (mmWave) in [23]. In the low Signal to Noise (SNR) regime, the proposed power allocation assigns all transmit power to the private streams. On the contrary, in the high SNR regime, only a portion of the total transmit power is assigned to the private streams such that the private streams are not interference limited, *i.e.*, the noise power remains larger or equal to the interference left between the private streams.

The studies in the literature so far have investigated specific cases in terms of the number of transmit antennas and/or users, and/or resort to optimization frameworks for RSMA [13], [17] that do not lend themselves to practical implementations. Therefore, obtaining a power allocation algorithm under imperfect CSIT for an arbitrary number of transmit antennas and users (and not necessarily infinitely many) still remains an open problem. Furthermore, there is a lack of consideration for the imperfections in the channel estimate quality originating from user mobility and latency in the network. The severity of CSIT quality degradation under user mobility causes non-negligible interference terms in the rate expressions, making it more difficult to analytically design practical power allocation algorithms.

B. Contributions

The contributions of this paper can be listed as follows:

- RSMA has been shown to outperform existing multiple access schemes under imperfect CSIT resulting from imperfect channel estimation and quantized feedback. In this work, we design RSMA specifically for CSIT imperfections caused by user mobility in Multi-User (Massive) MIMO. This is the first paper that shows that RSMA can be used to maintain reliable multi-user connectivity and mitigate the curse of mobility in Multi-User MIMO networks.
- We propose a system model where a multi-antenna transmitter employing RSMA serves multiple single-antenna users with mobility. We consider an arbitrary number of transmit antennas, number of users, total transmit

power and user speed in the system. We consider low-complexity transmit precoders for RSMA for the common and private streams. The splitting of the user rates is controlled by a power allocation coefficient that distributes the total transmit power between common and private precoders.

- Taking user mobility into consideration, we derive a lower bound for the approximated ergodic sum-rate of 1-layer RSMA for an arbitrary number of transmit antennas, number of users, amount of total transmit power and user mobility in a tractable form. The analytical expressions show that the obtained lower bound is also a function of the power allocation coefficient that distributes the total transmit power between common and private precoders. We demonstrate by simulations that the obtained lower bound captures the behaviour of the ergodic sum-rate for different sets of system parameters and different values of the power allocation coefficient.
- Using the derived lower bound, we obtain a closed-form solution for the power allocation coefficient that distributes the total transmit power between the common stream and private streams and propose a power allocation algorithm based on the obtained closed-form solution. The obtained closed-form solution is optimal in the sense that it maximizes the lower bound for the approximated ergodic sum-rate. The proposed power allocation method is valid for any number of transmit antennas, number of users, total transmit power and user speed.
- We perform simulations to analyse the performance of RSMA with the proposed power allocation method. We use Monte-Carlo and Link-Level Simulations (LLS) based on finite constellations and finite length polar codes to investigate the sum-rate and throughput performance. Furthermore, we compare the performance of RSMA and SDMA to highlight the significant sum-rate gain offered by RSMA. The results demonstrate that, in contrast to SDMA and conventional Multi-User (massive) MIMO whose performance collapse under mobility, RSMA can maintain multi-user connectivity in mobile multi-antenna systems. The results also show that the benefits are visible in a wide range of deployments, from small MIMO regime to large-scale (massive) MIMO regime, with a small and a large number of users.

The rest of the paper is organized is as follows. Section II gives the system model. We derive the lower bound for the ergodic sum-rate of 1-layer RSMA with ZF precoding in Section III. We obtain a closed-form expression and perform an asymptotic analysis for the power allocation algorithm for RSMA in Section IV. Section V gives the simulation results on the performance of RSMA with the proposed power allocation algorithm. Section VI concludes the paper.

Notation: Vectors are denoted by bold lowercase letters. The operations $|.|$ and $||.||$ denote the cardinality of a set or absolute value of a scalar and l_2-norm of a vector, respectively. $\mathbf{a}^T$ and $\mathbf{a}^H$ denote the transpose and Hermitian transpose of a vector $\mathbf{a}$, respectively. $\mathcal{CN}(0,\sigma^2)$ denotes the Circularly Symmetric Complex Gaussian distribution with zero mean and variance σ^2. $\mathbf{I}$ denotes the identity matrix. $\lfloor . \rceil$ denotes the round operation. $\mathrm{Gamma}(D,\theta)$ represents the Gamma distribution with the probability density function $f(x) = \frac{1}{\Gamma(D)\theta^D} x^{D-1} e^{-\frac{x}{\theta}}$. For any complex x with a positive real part, $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the gamma function and $\Gamma'(x)$ is the derivative of $\Gamma(x)$ with respect to x.

II. System Model

We consider the MISO BC setting consisting of one transmitter with n_t transmit antennas and K single-antenna users indexed by $\mathcal{K} = \{1,\ldots,K\}$, with $n_t \geq K$. We consider RSMA for multi-user transmission. Fig. 1 shows the transmission model of 1-layer RSMA with K users. RSMA relies on splitting the user messages at the transmitter and partially decoding the interference and partially treating the remaining interference as noise at the receivers.

The message W_k, which is intended for user-k, is split into one common part $W_{c,k}$ and one private part $W_{p,k}$, $\forall k \in \mathcal{K}$. The split messages are assumed to be independent. The common parts of the messages, $W_{c,k}$, $k \in \mathcal{K}$, are combined into the single common message W_c. The common message W_c and the K private messages $W_{p,k}, k \in \mathcal{K}$ are independently encoded into streams s_c, s_k, $k \in \mathcal{K}$, respectively. Linear precoding is applied to all streams. The transmit signal is expressed as

$$\mathbf{x} = \sqrt{P(1-t)}\mathbf{p}_c s_c + \sqrt{\frac{Pt}{K}} \sum_{k\in\mathcal{K}} \mathbf{p}_k s_k. \tag{1}$$

The precoders $\mathbf{p}_c, \mathbf{p}_k \in \mathbb{C}^{n_t}$ satisfy $||\mathbf{p}_c||^2 = 1$ and $||\mathbf{p}_k||^2 = 1$, $\forall k \in \mathcal{K}$. The transmitted streams satisfy $\mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\right\} = \mathbf{I}$, where $\mathbf{s} = [s_c, s_1, \ldots, s_k]$. The power allocation coefficient t determines the distribution of power among the common precoder and the private precoders. Note that $t = 1$ turns off the common stream and the strategy in (1) simply boils down to conventional SDMA with uniform power allocation.

The signal received by user-k is written as

$$y_k = \mathbf{h}_k^H[m]\mathbf{x} + n_k, \quad k \in \mathcal{K}, \tag{2}$$

where $\mathbf{h}_k[m] \in \mathbb{C}^{n_t}$ is the channel vector between the transmitter and user-k at time instant m and $n_k \sim \mathcal{CN}(0,1)$ is the Additive White Gaussian Noise (AWGN) component for user-k.

The flexible treatment[1] of interference is enabled by the joint use of message split at the transmitter and SIC at the receivers. The SIC processing at an arbitrary receiver (*e.g.*, user-k, $k = \{1,2,\ldots,K\}$) can be explained in the following 3 steps.

1) First, the interference from all users is partially decoded by decoding the common message estimate $\widehat{W}_c$ while treating the private streams of all (K) users as noise. Since the common message contains the common part of the message of each user (including the common part of the message of user-k), each user partially decodes the

[1]Readers are referred to [5], [13] for more discussion on rate-splitting.

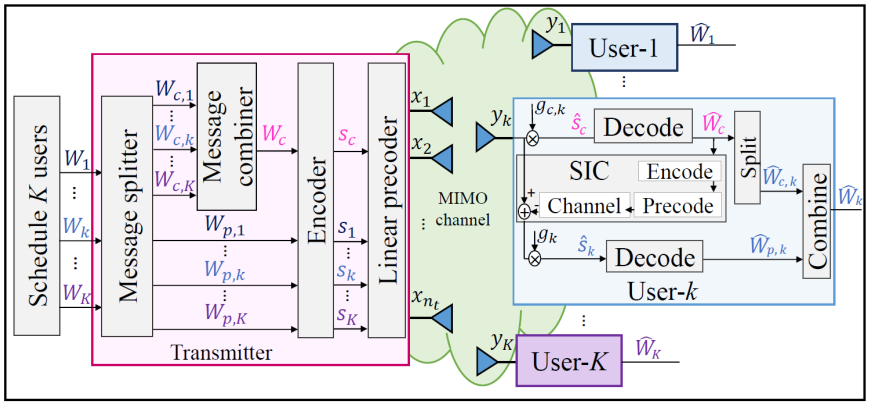


Fig. 1: Transmission model of K-user 1-layer RSMA.

message of another user (and therefore partially decodes interference).

2) Then, the common stream is reconstructed using the common message estimate $\widehat{W}_c$ and subtracted from the received signal by SIC to remove the decoded interference from the received signal.
3) Finally, the remaining (private) part of the message of user-k (private part) is decoded by treating the interference from the private streams of all other $K-1$ users as noise.

The overall decoded message of user-k, $\widehat{W}_k$, is reconstructed by combining the decoded common part, $\widehat{W}_{c,k}$, and the decoded private part, $\widehat{W}_{p,k}$.

The Signal-to-Interference-plus-Noise Ratio (SINR) for the common stream at user-k is

$$\gamma_{c,k} = \frac{P(1-t)|\mathbf{h}_k^H[m]\mathbf{p}_c|^2}{1+\frac{Pt}{K}\sum_{j\in\mathcal{K}}|\mathbf{h}_k^H[m]\mathbf{p}_i|^2}. \tag{3}$$

The SINR for the private stream at user-k is given by

$$\gamma_k = \frac{\frac{Pt}{K}|\mathbf{h}_k^H[m]\mathbf{p}_k|^2}{1+\frac{Pt}{K}\sum_{j\in\mathcal{K},j\neq k}|\mathbf{h}_k^H[m]\mathbf{p}_i|^2}. \tag{4}$$

The ergodic rates of the common and private streams are expressed as

$$R_c(t)=\mathbb{E}\left\{\log_2\left(1+\min_{k\in\mathcal{K}}\left[\frac{P(1-t)|\mathbf{h}_k^H[m]\mathbf{p}_c|^2}{1+\frac{Pt}{K}\sum_{j\in\mathcal{K}}|\mathbf{h}_k^H[m]\mathbf{p}_j|^2}\right]\right)\right\},$$
$$R_k(t)=\mathbb{E}\left\{\log_2\left(1+\frac{\frac{Pt}{K}|\mathbf{h}_k^H[m]\mathbf{p}_k|^2}{1+\frac{Pt}{K}\sum_{j\in\mathcal{K},j\neq k}|\mathbf{h}_k^H[m]\mathbf{p}_j|^2}\right)\right\}, \tag{5}$$

where the expectations are defined over the user channels $\mathbf{h}_k$. The ergodic sum-rate for the power allocation coefficient t is given by

$$R_{\text{RSMA}}(t) = R_c(t) + \sum_{k=1}^{K} R_k(t). \tag{6}$$

We consider the scenarios in which imperfect CSIT occurs due to user mobility and delay in CSI reports from the users. We adopt the channel model in [24], where the channel coefficients at time instant m are expressed as

$$\mathbf{h}_k[m] = \sqrt{\epsilon^2}\mathbf{h}_k[m-1] + \sqrt{1-\epsilon^2}\mathbf{e}_k[m]. \tag{7}$$

In the expression, $\mathbf{h}_k[m]$ denotes the spatially uncorrelated Rayleigh flat fading channel with i.i.d. entries distributed according to $\mathcal{CN}(0,1)$[2], $\epsilon = J_0(2\pi f_D T)$ denotes the time correlation coefficient obeying the Jakes' model [25] and $\mathbf{e}_k[m]$ has i.i.d. entries distributed according to $\mathcal{CN}(0,1)$. The parameter $f_D = vf_c/c$ is the maximum Doppler frequency for given carrier frequency f_c and user speed v (c is the speed of light) and T is the channel instantiation interval. In words, $\mathbf{h}_k[m]$ is the instantaneous channel vector observed at time instant m. However, the transmitter only has knowledge of the channel vector observed at time instant $m-1$, $\mathbf{h}_k[m-1]$, due to the latency in CSI reports from the users (T models the CSI report delay in seconds). Consequently, the transmitter uses the CSIT $\mathbf{h}_k[m-1]$ at time instant m to calculate the precoders. We consider ZF precoders with equal power allocation for the private streams[3], so that $|\mathbf{h}_j^H[m-1]\mathbf{p}_k| = 0$, $\forall j \in \mathcal{K}\backslash k$. The common precoder is assumed to be a random beamformer independent of $\mathbf{h}_k[m-1]$, $\mathbf{e}_k[m]$, and $\mathbf{p}_k$, $\forall k \in \mathcal{K}$, in order to be able to obtain tractable expressions for the rate of the common stream[4].

III. Lower Bound for the Ergodic Sum-Rate

Our aim is to find a power allocation method that maximizes the ergodic sum-rate of RSMA in (6) under imperfect CSIT due to mobility. As explained in Section II, we consider the RSMA scheme with ZF beamformers for the private streams with a portion Pt of the total power distributed equally among the users. The rest of the power, $P(1-t)$, is allocated to the common stream.

We focus on deriving a lower bound for the sum-rate of 1-layer RSMA scheme in an effort to obtain a closed-form solution for the optimum power allocation coefficient t_{opt}. We obtain the lower bound for an approximation of the ergodic sum-rate, which will be referred as the Approximated Ergodic Sum-Rate (AESR) for the rest of the paper, to obtain

[2]The assumption of identical second order statistics for user channels is for the brevity of the derivations and to obtain a closed-form solution for the optimal power allocation coefficient t. The derivations for the lower bound on the ergodic sum-rate can be extended to user channels with different second-order statistics following the same steps in Section III.

[3]Aside from being extremely common in literature (for example, see [21]–[23], [26] and the references therein), equal power allocation among precoders in MU-MIMO is employed in practical systems like 4G and 5G (for example, see [27], [28]).

[4]In Section V, we show that RSMA achieves significant sum-rate gain even with random beamformers for common streams.

tractable expressions and a closed-form solution for the power allocation coefficient. As a result, the obtained lower bound is for the AESR

$$\widetilde{R}_{\mathrm{RSMA}}(t) = \widetilde{R}_c(t) + \sum_{k=1}^{K} \widetilde{R}_k(t), \tag{8}$$

where $\widetilde{R}_c(t)$ and $\widetilde{R}_k(t)$ are Approximated Ergodic Rates (AERs), which are approximations for $R_c(t)$ and $R_k(t)$, respectively. The approximations are explained in detail in the following subsections. We demonstrate the tightness of the obtained lower bound by simulations in Section V.

We perform the derivation in two parts: the derivation of a lower bound for the sum of the AERs of the private streams, $\sum_{k=1}^{K} \widetilde{R}_k(t)$, and the derivation of a lower bound for the AER of the common stream, $\widetilde{R}_c(t)$.

A. Lower Bound for $\sum_{k=1}^{K} \widetilde{R}_k(t)$

We assume that all users experience channels with the same second order statistics. Therefore, we first find a lower bound for the ergodic rate of user-k, $R_k(t)$, and generalize it to K users to obtain the lower bound for $\sum_{k=1}^{K} R_k(t)$ [21].

We rewrite the expression for $R_k(t)$ in (5) as

$$\begin{aligned} R_k(t) =& \mathbb{E}\left\{\log_2\left(1+\frac{Pt}{K}\left(|\mathbf{h}_k^H[m]\mathbf{p}_k|^2+\sum_{j\in\mathcal{K},j\neq k}|\mathbf{h}_k^H[m]\mathbf{p}_j|^2\right)\right)\right\} \\ &- \mathbb{E}\left\{\log_2\left(1+\frac{Pt}{K}\sum_{j\in\mathcal{K},j\neq k}|\mathbf{h}_k^H[m]\mathbf{p}_j|^2\right)\right\}. \end{aligned} \tag{9}$$

Expression (9) consists of the terms $|\mathbf{h}_k^H[m]\mathbf{p}_j|^2$, for $k, j \in \mathcal{K}$. For tractability, we approximate the terms as

$$|\mathbf{h}_k^H[m]\mathbf{p}_k|^2 \approx \epsilon^2|\mathbf{h}_k^H[m-1]\mathbf{p}_k|^2+(1-\epsilon^2)|\mathbf{e}_k^H[m]\mathbf{p}_k|^2, \tag{10}$$

$$\begin{aligned} |\mathbf{h}_k^H[m]\mathbf{p}_j|^2 &\approx \epsilon^2|\mathbf{h}_k^H[m-1]\mathbf{p}_j|^2+(1-\epsilon^2)|\mathbf{e}_k^H[m]\mathbf{p}_j|^2 \\ &= (1-\epsilon^2)|\mathbf{e}_k^H[m]\mathbf{p}_j|^2. \end{aligned} \tag{11}$$

The approximations in (10) and (11) are due to the neglected terms containing both $\mathbf{h}_k^H[m-1]$ and $\mathbf{e}_k[m]$ [6], [7], [26]. Observing the expressions (10) and (11), one needs the distributions of the random variables (r.v.s) $|\mathbf{h}_k^H[m-1]\mathbf{p}_k|^2$ and $|\mathbf{e}_j^H[m]\mathbf{p}_k|^2$ for $j, k \in \mathcal{K}$.

With ZF beamforming, the precoder for a specific user is chosen to be orthogonal to the $K-1$ channels of the other users. Due to the isotropic nature of independent and identically distributed (i.i.d) Rayleigh fading, the orthogonality reduces the degrees of freedom at the transmitter to $n_t - K + 1$. Therefore, for ZF precoders with equal power allocation, the term $|\sqrt{2}\mathbf{h}_k^H[m-1]\mathbf{p}_k|^2$ has a chi-squared distribution with a degree of freedom $2(n_t - K+1)$, i.e., $|\sqrt{2}\mathbf{h}_k^H[m-1]\mathbf{p}_k|^2 \sim \chi^2_{2(n_t-K+1)}$. Equivalently, $|\mathbf{h}_k^H[m-1]\mathbf{p}_k|^2 \sim \mathrm{Gamma}(n_t-K+1,1)$ [26], [29], [30].

The ZF precoder $\mathbf{p}_k$ is isotropically distributed and independent of the Gaussian error $\mathbf{e}_j^H[m]$, $\forall j,k \in \mathcal{K}$, leading to an exponentially distributed r.v. with unit mean, or equivalently $|\mathbf{e}_j^H[m]\mathbf{p}_k|^2 \sim \mathrm{Gamma}(1,1)$. Assuming each interference term is independent, the term $\sum_{j\in\mathcal{K},j\neq k}|\mathbf{e}_k^H[m]\mathbf{p}_j|^2$ can be approximated by a random variable (r.v.) with the distribution $\mathrm{Gamma}(K-1,1)$ [6], [26].

We use the r.v.s and the distributions discussed above to obtain a lower bound for $\sum_{k=1}^{K}\widetilde{R}_k(t)$. We start the derivations from the first term in the expression (9). Lemma 1 gives an approximate distribution for the r.v. $\sum_{j\in\mathcal{K}}|\mathbf{h}_k^H[m]\mathbf{p}_j|^2$.

Lemma 1: The random variable $X = \sum_{j\in\mathcal{K}}|\mathbf{h}_k^H[m]\mathbf{p}_j|^2$ can be approximated by a random variable $\widetilde{X}$ with the distribution $\mathrm{Gamma}(\widehat{D},\widehat{\theta})$, where

$$\begin{aligned} \widehat{D} &= \frac{\left[\epsilon^2(n_t+1)+(1-2\epsilon^2)K\right]^2}{\epsilon^4(n_t+1)+(1-2\epsilon^2)K}, \\ \widehat{\theta} &= \frac{\epsilon^4(n_t+1)+(1-2\epsilon^2)K}{\epsilon^2(n_t+1)+(1-2\epsilon^2)K}. \end{aligned} \tag{12}$$

Proof: See Appendix A. ∎

We write an approximation $\widetilde{R}_k(t)$ for $R_k(t)$ using Lemma 1 and (11) with (9) as

$$\begin{aligned} R_k(t) &\approx \widetilde{R}_k(t) \\ &= \mathbb{E}\left\{\log_2\left(1+\frac{Pt}{K}\widetilde{X}\right)\right\} - \mathbb{E}\left\{\log_2\left(1+\frac{Pt}{K}\widetilde{Z}\right)\right\}, \end{aligned} \tag{13}$$

with $\widetilde{Z} = (1-\epsilon^2)\sum_{j\in\mathcal{K},j\neq k}|\mathbf{e}_k^H[m]\mathbf{p}_j|^2$. Proposition 1 gives a lower bound for $\widetilde{R}_k(t)$.

Proposition 1: The AER $\widetilde{R}_k(t)$ is lower bounded by

$$\widetilde{R}_k(t) \geq \log_2\left(1+\frac{P}{K}e^{\mu}t\right) - \log_2\left(1+\frac{(K-1)(1-\epsilon^2)P}{K}t\right), \tag{14}$$

where $\mu = \ln(\widehat{\theta}) + \Gamma'(\widehat{D})/\Gamma(\widehat{D})$ and the variables $\widehat{D}$ and $\widehat{\theta}$ are given in (12).

Proof: We start by writing a lower bound for $\widetilde{R}_k(t)$ as

$$\widetilde{R}_k(t) \geq \mathbb{E}\left\{\log_2\left(1+\frac{Pt}{K}\widetilde{X}\right)\right\} - \log_2\left(1+\frac{Pt}{K}\mathbb{E}\{\widetilde{Z}\}\right) \tag{15a}$$

$$\geq \log_2\left(1+\frac{Pt}{K}e^{\mathbb{E}\{\ln(\widetilde{X})\}}\right) - \log_2\left(1+\frac{Pt}{K}\mathbb{E}\{\widetilde{Z}\}\right). \tag{15b}$$

The expressions (15a) and (15b) follow from the Jensen's inequality and the fact that $\log_2(1+ax)$ is a concave function of x for any $ax > 0$ and $\log_2(1+ae^x)$ is a convex function of x for any $a > 0$ [21], respectively.

The mean of $\ln(\widetilde{X})$ can be shown to be

$$\mu \triangleq \mathbb{E}\{\ln(\widetilde{X})\} = \ln(\widehat{\theta}) + \Gamma'(\widehat{D})/\Gamma(\widehat{D}). \tag{16}$$

As a result,

$$\log_2\left(1+\frac{Pt}{K}e^{\mathbb{E}\{\ln\widetilde{X}\}}\right) = \log_2\left(1+\frac{Pt}{K}e^{\mu}\right).$$

For the second term in (13), we can directly write $\mathbb{E}\{\widetilde{Z}\} = (1-\epsilon^2)(K-1)$ as

$$\sum_{j\in\mathcal{K},j\neq k}|\mathbf{e}_k^H[m]\mathbf{p}_j|^2 \sim \mathrm{Gamma}(K-1,1).$$

Therefore,

$$\log_2\left(1+\frac{Pt}{K}\mathbb{E}\{\widetilde{Z}\}\right) = \log_2\left(1+\frac{P(1-\epsilon^2)(K-1)}{K}t\right),$$

which completes the proof. ■

We can generalize the lower bound given in Proposition 1 to K users with independent channels having identical second order statistics as

$$\sum_{k=1}^{K} \widetilde{R}_k(t) \geq K \log_2 \left(1 + \frac{P}{K} e^{\mu} t\right) - K \log_2 \left(1 + \frac{(K-1)(1-\epsilon^2)P}{K} t\right). \quad (17)$$

B. Lower Bound for $\widetilde{R}_c(t)$

Observing the expression for $R_c(t)$ in (5), we define the r.v.s

$$Y_k = \frac{|\mathbf{h}_k^H[m]\mathbf{p}_c|^2}{1 + \frac{Pt}{K}\sum_{j\in\mathcal{K}} |\mathbf{h}_k^H[m]\mathbf{p}_j|^2}, \quad (18)$$

and

$$Y = \min_{k\in\mathcal{K}} Y_k. \quad (19)$$

For tractability of the expressions, we approximate the r.v. Y by a r.v. $\widetilde{Y}$, so that

$$\widetilde{R}_c(t) = \log_2 \left(1 + P(1-t)\widetilde{Y}\right). \quad (20)$$

The following lemma gives the distribution of $\widetilde{Y}$.

Lemma 2: The r.v. Y can be approximated by a r.v. $\widetilde{Y}$ with the CDF

$$F_{\widetilde{Y}}(y) \approx 1 - \frac{e^{-Ky}}{\left(y\frac{P\widehat{\theta}t}{K} + 1\right)^{\widehat{D}K}}. \quad (21)$$

Proof: See Appendix B. ■

Before we start to derive the lower bound, we give another useful result in Lemma 3.

Lemma 3: The expected value of the natural logarithm of the r.v. $\widetilde{Y}$ is given by

$$\mathbb{E}\{\ln(\widetilde{Y})\} = -\gamma - \ln(K) - e^{\frac{K^2}{P\widehat{\theta}t}} \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right), \quad (22)$$

where $\gamma \approx 0.577$ is the Euler-Mascheroni constant and

$$E_v(x) = \int_1^{\infty} \frac{e^{-tx}}{t^v} dt, \quad x > 0,\ v \in \mathbb{R} \quad (23)$$

is the generalized exponential-integral function [31], [32].

Proof: The expected value of the function $g(x)$ of the random variable x can be calculated from the CDF $F_X(x)$ as

$$\mathbb{E}\{g(x)\} = \int_{-\infty}^{\infty} g(x)\frac{dF_X(x)}{dx} = -\int_{-\infty}^{\infty} g(x)\frac{d(1-F_X(x))}{dx}. \quad (24)$$

Integrating by parts gives

$$\mathbb{E}\{g(x)\} = -g(x)(1-F_X(x))\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty}(1-F_X(x))\frac{d(g(x))}{dx}. \quad (25)$$

We use the property in (25) and the CDF for $\ln(\widetilde{Y})$ given in (21) to obtain $\mathbb{E}\{\ln(\widetilde{Y})\}$ as

$$\mathbb{E}\{\ln(\widetilde{Y})\} = -\ln(y)\frac{e^{-Ky}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}}\Big|_0^{\infty} + \int_0^{\infty} \frac{e^{-Ky}}{y\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}} dy. \quad (26)$$

We assume that the power allocated to the private streams is always non-zero, so that $\frac{P\widehat{\theta}t}{K} > 0$. Under the assumption, we perform a change of variables in the second part of the expression by $z = \left(y\frac{P\widehat{\theta}t}{K} + 1\right)$, so that,

$$\begin{aligned}
\mathbb{E}\{\ln(\widetilde{Y})\} &= -\ln(y)\frac{e^{-Ky}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}}\Big|_0^{\infty} + \int_1^{\infty} \frac{e^{-K(z-1)\frac{K}{P\widehat{\theta}t}}}{z^{\widehat{D}K}\frac{z-1}{\left(\frac{P\widehat{\theta}t}{K}\right)}} \frac{dz}{\left(\frac{P\widehat{\theta}t}{K}\right)} \\
&= -\ln(y)\frac{e^{-Ky}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}}\Big|_0^{\infty} + e^{\frac{K^2}{P\widehat{\theta}t}} \int_1^{\infty} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}z}}{z^{\widehat{D}K}(z-1)} dz \\
&= -\ln(y)\frac{e^{-Ky}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}}\Big|_0^{\infty} + e^{\frac{K^2}{P\widehat{\theta}t}}\Theta(\widehat{D}K). \quad (27)
\end{aligned}$$

The integral $\Theta(\widehat{D}K)$ is not tractable. Our aim is to transform it into a form that is expressed in terms of known integrals. We use the expansion $\frac{1}{z(z-1)} = \frac{1}{z-1} - \frac{1}{z}$ to write

$$\begin{aligned}
\Theta(\widehat{D}K) &= \int_1^{\infty} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}z}}{z^{\widehat{D}K}(z-1)} dz \\
&= \int_1^{\infty} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}z}}{z^{(\widehat{D}K-1)}z(z-1)} dz \\
&= \int_1^{\infty} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}z}}{z^{(\widehat{D}K-1)}}\left(\frac{1}{z-1} - \frac{1}{z}\right) dz \\
&= \int_1^{\infty} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}z}}{z^{(\widehat{D}K-1)}(z-1)} dz - \int_1^{\infty} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}z}}{z^{\widehat{D}K}} dz \\
&= \Theta(\widehat{D}K-1) - \mathrm{E}_{\widehat{D}K}\left(\frac{K^2}{P\widehat{\theta}t}\right). \quad (28)
\end{aligned}$$

In order to define a tractable recursive expression to solve (28), we need $\widehat{D}K = \frac{\left[\epsilon^2(n_t+1)+(1-2\epsilon^2)K\right]^2}{\epsilon^4(n_t+1)/K+(1-2\epsilon^2)}$ to be a positive integer. Since the multiplication is not guaranteed to give such a result, we replace $\widehat{D}K$ with $\left\lfloor \widehat{D}K \right\rceil$. We rewrite the expression in (28) as

$$\begin{aligned}
\Theta\left(\left\lfloor \widehat{D}K \right\rceil\right) &= \Theta\left(\left\lfloor \widehat{D}K \right\rceil - 1\right) - \mathrm{E}_{\lfloor \widehat{D}K \rceil}\left(\frac{K^2}{P\widehat{\theta}t}\right) \\
&= \Theta(0) - \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right). \quad (29)
\end{aligned}$$

One needs a closed-form solution for $\Theta(0) = \int_1^{\infty} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}z}}{z-1} dz$ in order to obtain a closed-form expression for (29). For this

purpose, we perform a change of variables $u = \frac{K^2}{P\widehat{\theta}t}z - \frac{K^2}{P\widehat{\theta}t}$, so that

$$\Theta(0) = e^{-\frac{K^2}{P\widehat{\theta}t}} \int_0^\infty \frac{e^{-u}}{u} du = e^{-\frac{K^2}{P\widehat{\theta}t}} \int_0^\infty e^{-u} \frac{d(\ln(u))}{du}. \quad (30)$$

Next, we perform integration by parts to obtain

$$\Theta(0) = e^{-\frac{K^2}{P\widehat{\theta}t}} e^{-u}\ln(u)\Big|_0^\infty + e^{-\frac{K^2}{P\widehat{\theta}t}} \int_0^\infty e^{-u}\ln(u)du \quad (31a)$$

$$= e^{-\frac{K^2}{P\widehat{\theta}t}} e^{-u}\ln(u)\Big|_0^\infty - e^{-\frac{K^2}{P\widehat{\theta}t}}\gamma \quad (31b)$$

$$= e^{-\frac{K^2}{P\widehat{\theta}t}} \left(e^{-Ky}\ln(y)\Big|_0^\infty - \ln(K) - \gamma \right). \quad (31c)$$

The expression (31c) follows from the fact that $u = Ky$. We solve the recursive expression in (29) by substituting (31c) to get

$$\Theta\left(\left\lfloor \widehat{D}K \right\rceil\right) = e^{-\frac{K^2}{P\widehat{\theta}t}} \left(e^{-Ky}\ln(y)\Big|_0^\infty - \ln(K) - \gamma \right) - \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right). \quad (32)$$

Finally, we substitute the solution in (32) into (27) to get

$$\begin{aligned}
&\mathbb{E}\{\ln(\widetilde{Y})\} \\
&= -\ln(y)\frac{e^{-Ky}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}}\Big|_0^\infty + \ln(y)e^{-Ky}\Big|_0^\infty - \ln(K) - \gamma \\
&\qquad - e^{\frac{K^2}{P\widehat{\theta}t}} \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right) \\
&= \ln(y)e^{-Ky}\left(1 - \frac{1}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}}\right)\Big|_0^\infty - \ln(K) - \gamma \\
&\qquad - e^{\frac{K^2}{P\widehat{\theta}t}} \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right) \\
&= -\gamma - \ln(K) - e^{\frac{K^2}{P\widehat{\theta}t}} \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right). \quad \blacksquare
\end{aligned}$$

Proposition 2: The AER $\widetilde{R}_c(t)$ is lower bounded by

$$\widetilde{R}_c(t) \geq \log_2\left(1 + P(1-t)e^{\beta}\right), \quad (33)$$

where $\beta \triangleq -\gamma - \ln(K) - e^{\frac{K^2}{P\widehat{\theta}t}} \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right)$.

Proof: The proof is straightforward using the Jensen's inequality as in (15b) and the result in Lemma 3. ■

The results in Propositions 1 and 2 are combined to obtain a lower bound for an approximation of the sum-rate of the 1-layer RSMA with imperfect CSIT due to mobility.

Proposition 3: The ergodic sum-rate $R_{\mathrm{RSMA}}(t)$ of the 1-layer RSMA with imperfect CSIT due to mobility can be approximated by $\widetilde{R}_{\mathrm{RSMA}}(t)$, which is lower bounded by

$$\begin{aligned}
&\widetilde{R}_{\mathrm{RSMA}}(t) \\
&\geq K\log_2\left(1 + \frac{P}{K}e^{\mu}t\right) - K\log_2\left(1 + \frac{(K-1)(1-\epsilon^2)P}{K}t\right) \\
&\qquad + \log_2\left(1 + P(1-t)e^{\beta}\right). \quad (34)
\end{aligned}$$

We note here that the lower bound on the ergodic sum-rate can be extended to scenarios including user channels with different second-order statistics following the same steps in this section.

IV. Closed-Form Solution for the Optimal Power Allocation

We are interested in the value of the power allocation coefficient t, which maximizes the lower bound in (34). Such a coefficient can be found by an exhaustive search over the expression (34). However, obtaining the optimum t by an exhaustive search is not suitable for use in practical systems. In this section, we aim to find a closed-form solution for the optimal power allocation coefficient, which we denote as t_{opt}.

A. Derivation

We assume the asymptotic scenario where $Pt \to \infty$ in an effort to obtain the closed-form solution. With such assumption, we obtain the following result.

Lemma 4: As $Pt \to \infty$, the term $\beta = -\gamma - \ln(K) - e^{\frac{K^2}{P\widehat{\theta}t}} \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right)$ can be approximated as

$$\beta \approx -\gamma + \ln\left(\frac{K}{P\widehat{\theta}\left(\left\lfloor \widehat{D}K \right\rceil - 1\right)t}\right) - \frac{1}{2(\lfloor \widehat{D}K \rfloor - 1)}. \quad (35)$$

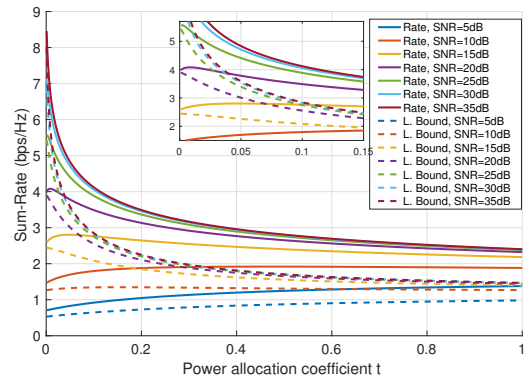
Proof: See Appendix C. ■

Based on the assumption $Pt \to \infty$ and Lemma 4, we can rewrite the lower bound expression (34) as

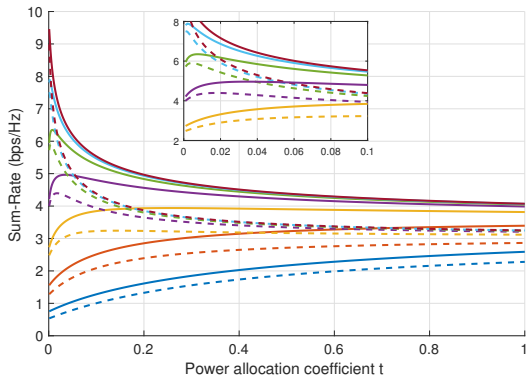
$$\begin{aligned}
&\widetilde{R}_{\mathrm{RSMA}}(t) \\
&\geq K\log_2\left(1 + \frac{Pe^{\mu}}{K}t\right) - K\log_2\left(1 + \frac{(K-1)(1-\epsilon^2)P}{K}t\right) \\
&\qquad + \log_2\left(1 + P(1-t)e^{\beta}\right) \\
&\approx K\log_2\left(\frac{Pe^{\mu}}{K}t\right) - K\log_2\left(1 + \frac{(K-1)(1-\epsilon^2)P}{K}t\right) \\
&\qquad + \log_2\left(1 + \frac{(1-t)}{t}\frac{K}{\hat{\theta}(\left\lfloor \widehat{D}K \right\rceil - 1)}e^{-\gamma - \frac{1}{2(\lfloor \widehat{D}K \rceil - 1)}}\right) \\
&= -K\log_2\left(\frac{1}{\tau t} + \frac{\omega}{\tau}\right) + \log_2\left((1-\rho) + \frac{\rho}{t}\right). \quad (36)
\end{aligned}$$

with the terms

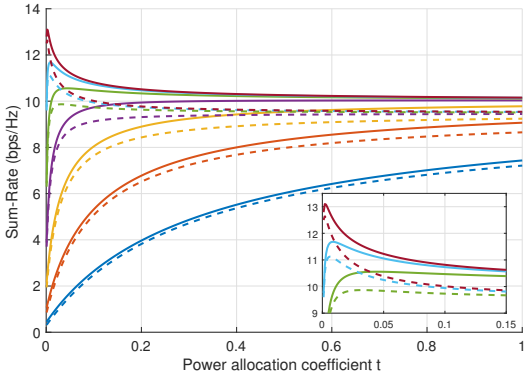
$$\tau \triangleq \frac{Pe^{\mu}}{K}, \qquad \omega \triangleq \frac{(K-1)(1-\epsilon^2)P}{K},$$
$$\rho \triangleq \frac{K}{\widehat{\theta}(\left\lfloor \widehat{D}K \right\rceil - 1)}e^{-\gamma - \frac{1}{2(\lfloor \widehat{D}K \rceil - 1)}}. \quad (37)$$



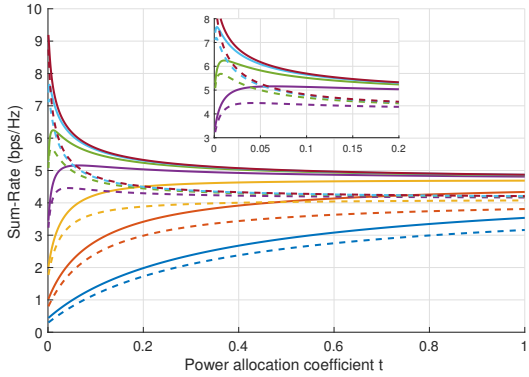
(a) $K = 4$, $n_t = 4$, $\epsilon = 0.5$.



(b) $K = 4$, $n_t = 8$, $\epsilon = 0.5$.



(c) $K = 8$, $n_t = 32$, $\epsilon = 0.5$.



(d) $K = 8$, $n_t = 32$, $\epsilon = 0.3$.

Fig. 2: Sum-rate and lower bound.

Taking the derivate of (36) with respect to t and equating the resulting expression to zero yields

$$t_{opt} = \frac{\rho(K-1)}{\rho(\omega+K)-K}. \tag{38}$$

In order to obtain a feasible value for t_{opt}, i.e., $t_{opt} \in (0, 1]$, we set the condition $\rho(\omega+K)-K > \rho(K-1)$. The resulting power allocation algorithm is expressed as

$$t_{opt} = \begin{cases} \frac{\rho(K-1)}{\rho(\omega+K)-K}, & \text{if } \rho(\omega+1)/K > 1 \\ 1 \quad , & \text{otherwise.} \end{cases} \tag{39}$$

In the next subsection, we analyze the expression (39) for extreme cases of CSIT quality (i.e., $\epsilon = 0$ and $\epsilon = 1$) in order to get an insight on the behaviour of the proposed power allocation method.

B. Asymptotic Analysis

We analyze the values of the power allocation coefficient for the asymptotic values of 0 and 1 for $\epsilon \in [0, 1]$. As the first step in the analyses, we check whether the condition in (39) is feasible (i.e., can be satisfied for at least one combination of K, n_t and P) in the considered asymptotic case.

1) $\epsilon \to 0$: We substitute $\epsilon = 0$ in (12) and (37) to obtain the following: $\hat{D} = K$, $\hat{\theta} = 1$, $\omega = \frac{P(K-1)}{K}$, and $\rho = \frac{K}{(K^2-1)} e^{-\gamma - \frac{1}{2(K^2-1)}}$. Then, the condition in (39) becomes

$$1 < \frac{P(1-\frac{1}{K})+1}{K^2-1} e^{-\gamma-\frac{1}{2(K^2-1)}}. \tag{40}$$

The expression (40) is difficult to analyze in its current form. We divert our analysis to the scenarios with the smallest possible number ($K = 2$) and a large ($K-1 \approx K$) number of users in an effort to obtain an understanding for the coefficient behaviour at extreme values.

For $K = 2$, the power allocation algorithm becomes

$$t_{opt} = \begin{cases} \frac{2}{P-8.62}, & \text{if } 10.62 < P \\ 1 \quad , & \text{otherwise.} \end{cases} \tag{41}$$

It is seen from (41) that the resulting power allocation algorithm is dependent only on the transmit power level. The optimal power allocation coefficient $t_{opt} \to 0$ as $P \to \infty$.

For large values of K, i.e., $K-1 \approx K$, the power allocation condition is approximated as

$$1.78 < \frac{P+1}{K^2} e^{-\frac{1}{2K^2}} \approx \frac{P+1}{K^2}.$$

Then, the power allocation algorithm becomes

$$t_{opt} = \begin{cases} \frac{1}{\frac{P}{K}+1-1.78K}, & \text{if } 1.78 < \frac{P+1}{K^2} \\ 1 \quad , & \text{otherwise.} \end{cases} \tag{42}$$

The algorithm in (42) is similar to that in (41) in the sense that both rely on the transmit power level and the convergence behaviour of the optimal coefficient, i.e., $t_{opt} \to 0$ as $P \to \infty$.

The resulting algorithms indicate that at asymptotically low CSIT quality, more transmit power is assigned to the common stream as the transmit power increases. The threshold for power allocation to the common stream depends on the

transmit power and the number of users. Such behaviour is justified since assigning more power (and equally distributing it) to the private streams does not improve the sum-rate in the interference-limited regime.

2) $\epsilon \rightarrow 1$: We substitute $\epsilon = 1$ in (12) and (37) to obtain the following: $\hat{D} = n_t + 1 - K$, $\hat{\theta} = 1$, $\omega = 0$, and $\rho = \frac{K}{(n_t+1-K)K-1} e^{-\gamma - \frac{1}{2((n_t+1-K)K-1)}}$. Then, the condition in (39) becomes

$$1 < \frac{1}{(n_t+1-K)K-1} e^{-\gamma - \frac{1}{2((n_t+1-K)K-1)}}. \tag{43}$$

As $e^{-\gamma - \frac{1}{2((n_t+1-K)K-1)}} < 1$, $K > 1$, and $n_t \geq K$, the condition is not satisfied for any values of K or n_t. Therefore, the optimal power allocation coefficient becomes $t_{opt} = 1$, which indicates that no power is allocated to the common stream when there is perfect CSIT. Such behaviour is justified since multi-user interference can be cancelled out perfectly by ZF precoding under perfect CSIT.

V. Simulation Results

In this section, we demonstrate the behaviour of the lower bound compared to the ergodic sum-rate and the performance of 1-layer RSMA with the proposed power allocation algorithm by simulations. In the first part of the analysis, we perform Monte Carlo simulations over different channel realizations to investigate the sum-rate performance. We compare the sum-rate performance achieved by using the power allocation coefficient obtained from the expression (39) and the coefficient that maximizes the lower bound (34), which is found by exhaustive search. By such comparison, we aim to show that the proposed power allocation algorithm can achieve near-optimal performance. Furthermore, we compare the performance of SDMA and RSMA with the abovementioned power allocation methods to show that RSMA achieves significant performance gain over SDMA.

In the second part of the analysis, we investigate the throughput performance by Link-Level Simulations (LLS) using the proposed power allocation algorithm with a practical transceiver architecture and a realistic 3rd Generation Partnership Project (3GPP) channel model. Moreover, we compare the performance of SDMA and RSMA with the proposed power allocation algorithm to verify the performance gain achieved by RSMA in a realistic setup.

A. Sum-Rate Performance

We start our analysis by investigating the behaviour of the derived lower bound in order to get an insight on the success of the obtained power allocation algorithm. Fig. 2a-2d show the ergodic sum-rate with Gaussian inputs obtained by expression (6) and Monte Carlo simulations, and the lower bound calculated from expression (34), both as functions of t for different values of ϵ, n_t, K, and P. We define SNR as $10\log_{10}(P)$, owing to the unit noise variance assumption in the system model (without loss of generality). We note that $t = 1$ corresponds to the SDMA scheme, *i.e.,* no power is allocated to the common stream and each message W_k is directly encoded into a private stream.

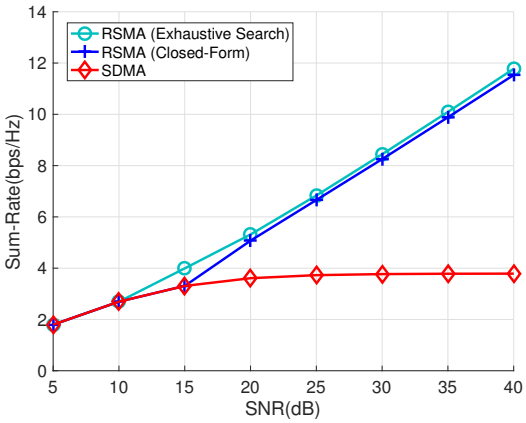
An immediate observation from the figures is the improved sum-rate performance of RSMA ($0 \leq t \leq 1$) compared to SDMA ($t = 1$), *i.e.*, conventional MU-MIMO precoding. The optimal value of t that maximizes the sum-rate of RSMA varies with the parameters n_t, K, P, ϵ and the maximal points at values $t < 1$ prove that message splitting is required to maximize the sum-rate. A general trend observed in all figures is that as the total transmit power increases, the increase in the sum-rate of SDMA saturates as a result of the multi-user interference becoming more dominant than noise (the system enters the interference-limited regime). Under such operating conditions, the optimal strategy to maximize the sum-rate is to allocate a certain portion of the total transmit power to the common stream. The figures show that the optimal portion varies with the system parameters and the CSIT quality. One can also note here that RSMA achieves significant performance gain over SDMA even with random beamformers employed for the common streams.

Another observation from Fig. 2a-2d is that the obtained lower bound captures the behaviour of the ergodic sum-rate with the varying parameters successfully. The results indicate that the power allocation coefficient t that maximizes the obtained lower bound is almost identical to the coefficient that maximizes the actual ergodic sum-rate. Consequently, one can expect that the proposed low-complexity power allocation algorithm in (39) achieves a near-optimal performance.

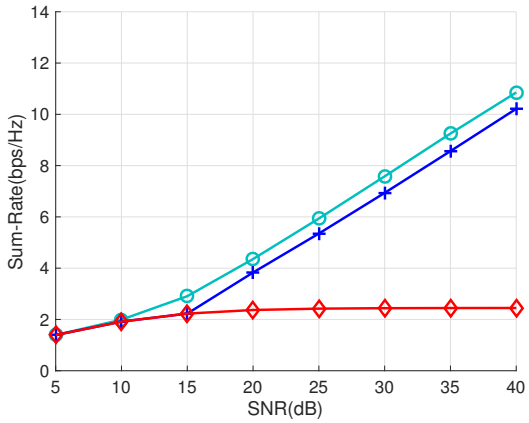
Next, we investigate the performance of 1-layer RSMA with the proposed power allocation algorithm and compare its performance with that of SDMA. We consider the ergodic sum-rate for Gaussian inputs. The performance of RSMA is studied by using power allocation coefficients obtained using two separate methods. For the first method, we perform an exhaustive search over the power allocation coefficient t to find the value that maximizes the ergodic sum-rate expression given in (6). For the exhaustive search procedure, we consider all values of t in the interval $(0, 1]$ with a granularity of 0.001. For the second method, we use the coefficient t_{opt} given by the closed-form solution in (39). We analyse the performance with the user speed values of 20km/h and 30km/h, a CSI acquisition delay of 2.5ms and an operating frequency of 3.5GHz. For an enhanced performance for the common stream, we consider the leftmost eigenvector (the eigenvector corresponding to the maximum eigenvalue) of the CSIT matrix $\mathbf{H}[m-1] = [\mathbf{h}_1[m-1], \mathbf{h}_1[m-1], \ldots, \mathbf{h}_K[m-1]]$ as the common precoder. Fig. 3a-3d show the ergodic sum-rates achieved by SDMA and RSMA with the two power allocation methods for different values of K, n_t, and user speed.

As seen from the figures, the exhaustive search and the closed-form power allocation methods achieve similar sum-rate performance for all of the considered cases. The results demonstrate that the closed-form method provides a near-optimal solution for the ergodic sum-rate maximization problem with a low-complexity algorithm. An important conclusion from the figures is that RSMA achieves a non-saturating performance while the sum-rate of SDMA saturates in the interference limited SINR regime, as also expected from the results in Fig. 2a-2d.

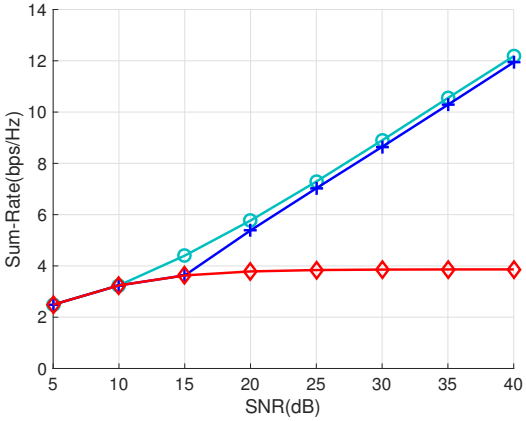
Finally, we compare the behaviours of the closed-form



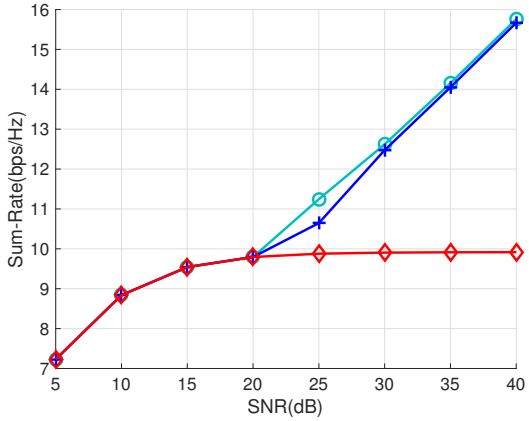
(a) $K = 4$, $n_t = 4$, Speed=20km/h.



(b) $K = 4$, $n_t = 4$, Speed=30km/h.

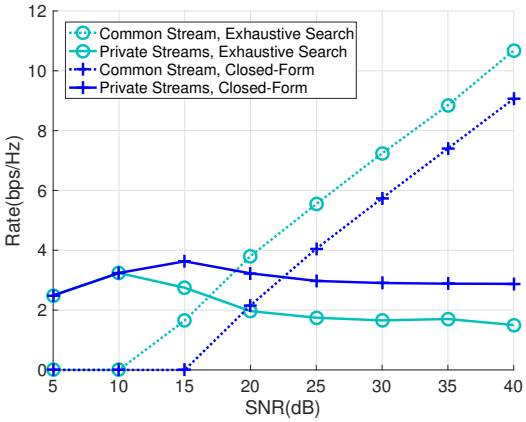


(c) $K = 4$, $n_t = 8$, Speed=30km/h.

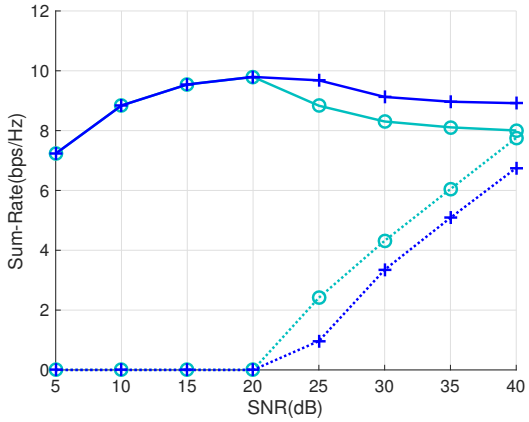


(d) $K = 8$, $n_t = 32$, user speed=30km/h.

Fig. 3: Sum-rate vs. SNR.



(a) $K = 4$, $n_t = 8$, user speed=30km/h.



(b) $K = 8$, $n_t = 32$, user speed=30km/h.

Fig. 4: Stream rates vs. user speed.

and exhaustive search power allocation methods in different settings. For this purpose, we plot the rates allocated to the common and private streams separately. Fig. 4a-4b show the rate allocations achieved by exhaustive search over the obtained lower bound and the closed-form solution for two different scenarios. The figures show that both methods behave similarly when distributing the power among the common and private streams. The portion of the total power allocated to the common stream increases with increasing SNR as the system moves into the interference-limited regime for both methods. Such behaviour allows RSMA to achieve a non-saturating sum-rate performance as opposed to SDMA, whose sum-rate saturates with SNR as observed from Fig. 3a-3d. The results imply that the proposed closed-form method achieves a near-optimal power allocation with a practical algorithm.

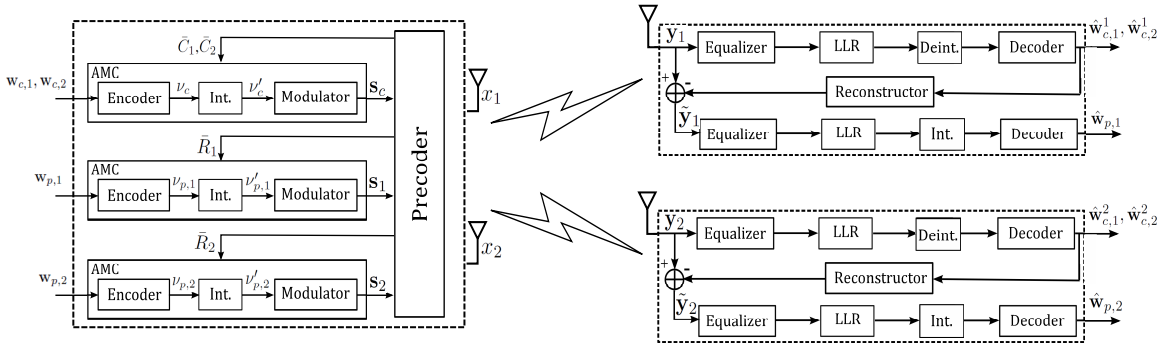
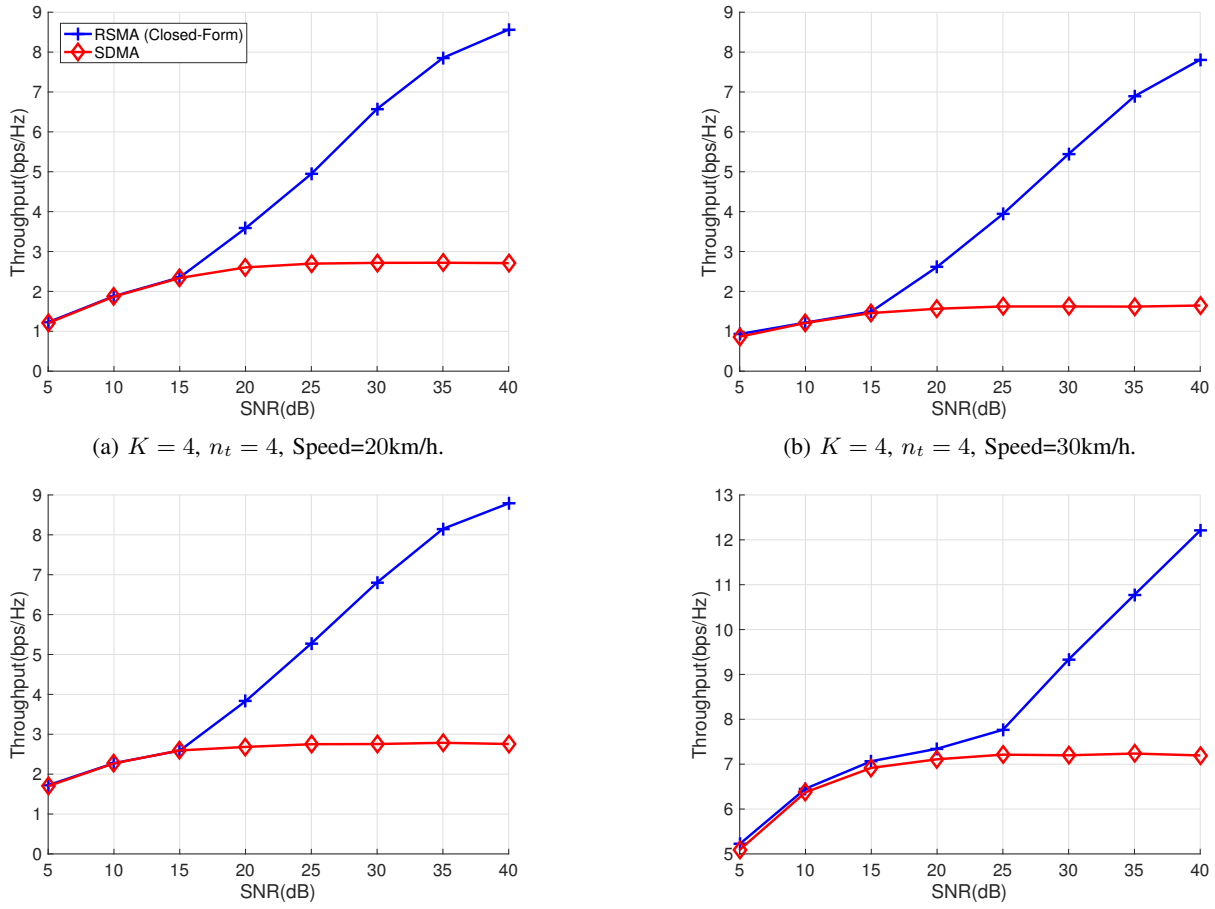


Fig. 5: Transmitter and receiver structures for 1-layer RSMA [34].



(a) $K = 4$, $n_t = 4$, Speed=20km/h.

(b) $K = 4$, $n_t = 4$, Speed=30km/h.

(c) $K = 4$, $n_t = 8$, Speed=30km/h.

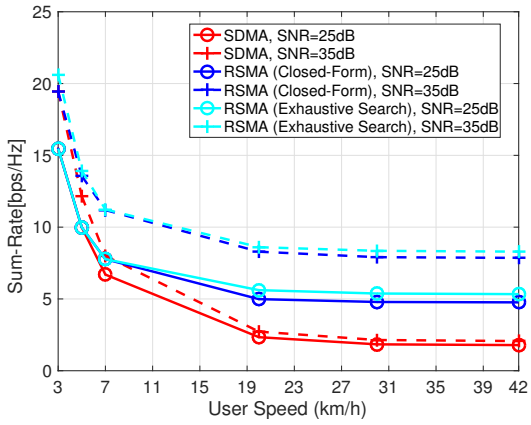
(d) $K = 8$, $n_t = 32$, Speed=30km/h.

Fig. 6: Throughput vs. SNR.

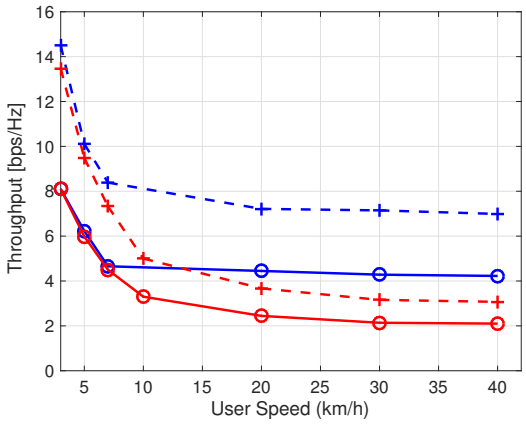
B. Link-Level Simulations

In this section, we investigate the throughput performance of SDMA and RSMA with the proposed power allocation algorithm by LLS. We use a transceiver architecture for RSMA with finite alphabet modulation schemes, finite-length channel coding, and an Adaptive Modulation and Coding (AMC) algorithm. We employ the modulation schemes 4-QAM, 16-QAM, 64-QAM and 256-QAM and polar coding [33]. The AMC algorithm selects a suitable modulation-coding rate pair for data transmission depending on the calculated transmit rates. The transmit rate calculations for AMC algorithm are performed assuming the instantaneous channel $\mathbf{h}_k[m], \forall k \in \mathcal{K}$ are known in the AMC module (only for the purpose of modulation-coding scheme selection). For completeness, the overall architecture for two transmit antennas and two single-antenna users is presented in Fig. 5, which is modified in this work to simulate a system with K (single-antenna and multi-antenna) users. The reader is referred to [34] for more details on the architecture and the LLS platform.

Let $S^{(l)}$ denote the number of channel uses in the l-th Monte-Carlo realization and $D_{s,k}^{(l)}$ denote the number of successfully recovered information bits by user-k in the common stream (excluding the common part of the message intended for the other user) and its private stream. Then, we calculate the throughput as

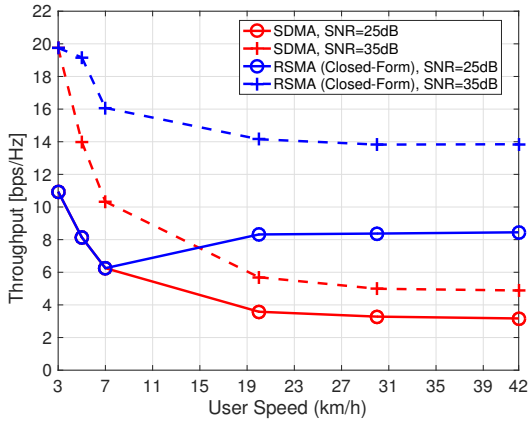


(a) Sum-rate vs. user speed.

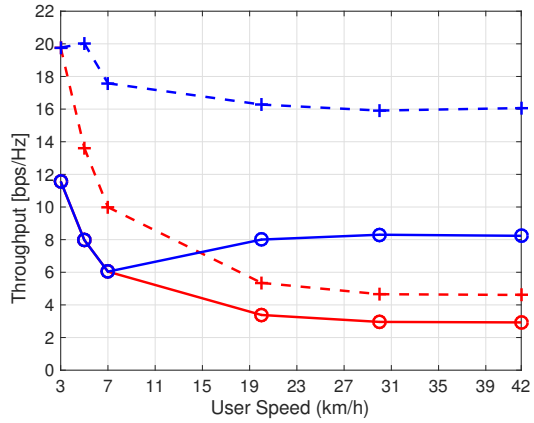


(b) Throughput vs. user speed.

Fig. 7: Sum-rate and throughput vs. user speed, OFDM waveform, 3GPP channel model.



(a) Throughput vs. user speed, $Q_c = 2$, $Q_p = 2$.



(b) Throughput vs. user speed, $Q_c = 3$, $Q_p = 3$.

Fig. 8: Sum-rate and throughput vs. user speed, OFDM waveform, 3GPP channel model, multiple receive antennas.

$$\text{Throughput[bps/Hz]} = \frac{\sum_l (D_{s,1}^{(l)} + D_{s,2}^{(l)})}{\sum_l S^{(l)}}. \quad (44)$$

Fig. 6a-6d show the throughput performance of RSMA and SDMA for the settings used in Fig. 3a-3d. The results demonstrate that the throughput performance follow the exact same trends of the sum-rate curves. RSMA can achieve the performance enhancements demonstrated by the ergodic sum-rate analysis in terms of throughput with practical transmitter and receiver designs.

We take our performance analysis in practical scenarios one step further and study the performance of SDMA and RSMA with the proposed power allocation algorithm in more realistic channel models. We use the Quadriga Channel Generator [35] to generate channels according to the 3GPP Urban Macro-Cell channel model [36]. The channels of each user have a delay spread of 300ns and 23 clusters, with each cluster consisting of 20 rays.

We use Cyclic-Prefix (CP)-OFDM waveform for transmission. The OFDM subcarrier spacing is set as 30kHz. Data is carried over 256 subcarriers. The CP length is 10μs. We assume the system is operating at 3.5GHz and the CSI acquisition delay is 10ms. We calculate and apply precoders separately for each subcarrier. The power allocation coefficient t_{opt} is constant on each subcarrier as the parameters n_t, K, P and ϵ do not vary with subcarrier index.

Fig. 7a and 7b show the sum-rate and throughput performance of RSMA and SDMA with respect to user speed. We analyse the scenarios with $n_t = 32$, $K = 8$ and average received SNR levels of 25dB and 35dB in the simulations.

First, we compare the performance of the two power allocation algorithms from Fig. 7a. The power allocation coefficients obtained from the closed-form solution (39) can achieve a performance close to the coefficients obtained by the exhaustive search for all user speed and SNR values, strengthening our conclusion from the previous results that the proposed power allocation algorithm can achieve near-optimal performance.

Fig. 7a and 7b clearly demonstrate the robustness of 1-layer RSMA with the proposed power allocation algorithm under user mobility. The sum-rate and achieved throughput of SDMA drops significantly with increasing user speed. Furthermore, the performance of SDMA saturates with increasing SNR, implying that the system operates in the interference-limited regime. On the other hand, RSMA limits the performance degradation starting at low user speed and achieves a robust performance as the speed increases. It can be observed that the sum-rate and achieved throughput of RSMA increases

with increasing SNR, meaning that RSMA can operate in the interference-limited regime without a saturating throughput.

Although we consider single-antenna receivers in the derivations in this work, we extend our LLS analysis to a system model in which each receiver has n_r receive antennas. We consider the transmission of a vector of Q_c common streams to be decoded by all users and K vectors of Q_p private streams (*i.e.,* the total number of streams is $Q_c + KQ_p$). The precoders for the private streams are obtained by the Block-Diagonalization (BD) method to suppress the multi-user interference [37]. We employ the same power allocation algorithm in (39) to distribute power among the precoders for the common and private streams. The power allocated to the common precoders is distributed equally among the common streams and the power allocated to the private precoders is distributed equally among the private streams of all users.

The receivers employ Minimum-Mean Square Error (MMSE) equalization to detect the intended common stream by suppressing the interference from the other common streams and all private streams. The common streams are decoded in parallel without any interference cancellation in-between. After all common streams are decoded, the overall common signal is reconstructed and interference cancellation is performed using the reconstructed common signal. Finally, the receivers apply an identical procedure to detect and decode the private streams.

Fig. 8a and 8b demonstrate the throughput results for $n_t = 64$, $n_r = 4$, $K = 8$ and different values of Q_c and Q_p. The throughput curve of RSMA is observed to have a non-monotonic behaviour since the power allocation algorithm is sub-optimal for systems with multiple receive antennas, thus resulting in loss of throughput for specific user speed values. The figures clearly show that RSMA with the proposed power allocation algorithm can perform significantly better than SDMA also with multi-antenna receivers. Future works include the refinement of the lower bound derivations and the closed-form solution for the power allocation algorithm for systems with multi-antenna receivers.

VI. Conclusion

In this work, we derive a lower bound for the approximate ergodic sum-rate of RSMA under user mobility for arbitrary number of transmit antennas, number of users, user speed and transmit power. Using this lower bound, we find a closed-form solution for the optimal power allocation coefficient that distributes the transmit power between the common stream and private streams. We perform simulations with realistic channel models to demonstrate the performance of RSMA with the obtained power allocation coefficient. Theoretical bounds and link-level simulation results clearly show that RSMA with the proposed power allocation scheme is robust to the degrading effects of user mobility and has significantly higher performance than that of SDMA. The key takeaway message is that in mobile MIMO networks, to maintain reliable multi-user connectivity, the transmission strategy should depart from conventional SDMA and Multi-User (massive) MIMO architecture that treats residual interference as noise, and adopts RSMA that is inherently robust to imperfect CSIT and mobility thanks to its ability to partially decode and partially treat interference as noise.

Appendix A
Proof of Lemma 1

Using the expressions (10) and (11), we can approximate the r.v. X with the r.v. $\widehat{X}$, such that

$$\widehat{X} = \epsilon^2 |\mathbf{h}_k^H[m-1]\mathbf{p}_k|^2 + (1-\epsilon^2)\sum_{k\in\mathcal{K}} |\mathbf{e}_k^H[m]\mathbf{p}_j|^2. \quad (45)$$

As explained in Section III-A, the terms $|\mathbf{h}_k^H[m-1]\mathbf{p}_k|^2$ and $|\mathbf{e}_k^H[m]\mathbf{p}_j|^2$ are r.v.s with Gamma distributions. Although the Gamma-distributed r.v.s in (45) are correlated and have different scales, one can approximate $\widehat{X}$ with another Gamma-distributed random variable $\widetilde{X}$ by matching the first two moments [26]. Specifically, a random variable $Z = \sum_i A_i$, where $A_i \sim \mathrm{Gamma}(D_i, \theta_i)$ can be approximated with a random variable $\widetilde{Z} \sim \mathrm{Gamma}(\widehat{D}, \widehat{\theta})$, where

$$\widehat{D} = \frac{(\sum_i D_i\theta_i)^2}{\sum_i D_i\theta_i^2}, \qquad \widehat{\theta} = \frac{\sum_i D_i\theta_i^2}{\sum_i D_i\theta_i}. \quad (46)$$

Therefore, we approximate $\widehat{X}$ by $\widetilde{X} \sim \mathrm{Gamma}(\widehat{D}, \widehat{\theta})$, where

$$\widehat{D} = \frac{\left[\epsilon^2(n_t+1) + (1-2\epsilon^2)K\right]^2}{\epsilon^4(n_t+1) + (1-2\epsilon^2)K},$$
$$\widehat{\theta} = \frac{\epsilon^4(n_t+1) + (1-2\epsilon^2)K}{\epsilon^2(n_t+1) + (1-2\epsilon^2)K}.$$

Appendix B
Proof of Lemma 2

We start the proof by finding the CDF of the r.v. Y_k.

$$Y_k = \frac{Q_{k1}}{1 + \frac{P\widehat{\theta}t}{K} Q_{k2}}, \quad (47)$$

with $Q_{k1} = |\mathbf{h}_k^H[m]\mathbf{p}_c|^2$ and $Q_{k2} = \frac{1}{\widehat{\theta}}\sum_{j\in\mathcal{K}} |\mathbf{h}_k^H[m]\mathbf{p}_j|^2$.

It is assumed that the precoder for the common stream, $\mathbf{p}_c$, to be independent of $\mathbf{h}_k[m]$, $\forall k \in \mathcal{K}$ and isotropically distributed on the unit sphere [21]. As discussed in Section III-A, these assumptions result in $|\mathbf{h}_k^H[m]\mathbf{p}_c|^2 \sim \mathrm{Gamma}(1,1)$.

Lemma 1 states that the r.v. $X = \sum_{j\in\mathcal{K}} |\mathbf{h}_k^H[m]\mathbf{p}_j|^2$ can be approximated by another r.v. $\widetilde{X} \sim \mathrm{Gamma}(\widehat{D}, \widehat{\theta})$ with $\widehat{D}$ and $\widehat{\theta}$ given in (12). Using this result, we can approximate Q_{k2} with a r.v. $\widetilde{Q}_{k2} \sim \mathrm{Gamma}(\widehat{D}, 1)$.

For tractability, we assume independence among the terms in the numerator and the denominator. Then, we can approximate Y_k by $\widetilde{Y}_k$ with the CDF

$$\begin{aligned} F_{\widetilde{Y}_k}(y) &= \int_0^\infty P\left(Q_{k1} < y\left(1 + \frac{P\widehat{\theta}t}{K}q\right)\right) f_{\widetilde{Q}_{k2}}(q)dq \\ &= \int_0^\infty \left(1 - e^{-y\left(1+\frac{P\widehat{\theta}t}{K}q\right)}\right) \frac{e^{-q}q^{\widehat{D}-1}}{\Gamma(\widehat{D})}dq \\ &= \int_0^\infty \frac{e^{-q}q^{\widehat{D}-1}}{\Gamma(\widehat{D})}dq - \frac{e^{-y}}{\Gamma(\widehat{D})}\int_0^\infty e^{-q\left(y\frac{P\widehat{\theta}t}{K}+1\right)} q^{\widehat{D}-1}dq \\ &= 1 - \frac{e^{-y}}{\Gamma(\widehat{D})}\int_0^\infty e^{-q\left(y\frac{P\widehat{\theta}t}{K}+1\right)} q^{\widehat{D}-1}dq. \end{aligned} \quad (48)$$

We perform a change of variables by $z = q\left(y\frac{P\widehat{\theta}t}{K}+1\right)$, so that we can rewrite (48) as

$$\begin{aligned}
&F_{\widetilde{Y}_k}(y)\\
&= 1 - \frac{e^{-y}}{\Gamma(\widehat{D})}\int_0^\infty e^{-z}\frac{z^{\widehat{D}-1}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{(\widehat{D}-1)}}\frac{dz}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)},\\
&= 1 - \frac{e^{-y}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}}}\int_0^\infty \frac{e^{-z}z^{\widehat{D}-1}}{\Gamma(\widehat{D})}dz\\
&= 1 - \frac{e^{-y}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}}}, \qquad (49)
\end{aligned}$$

with the support set $[0,\infty)$.

Using the assumption that Y_k are independent [21], we can approximate Y by $\widetilde{Y}$ with the CDF $F_{\widetilde{Y}}(y) = (1-(1-F_{\widetilde{Y}_k}(y))^K)$, which yields

$$\begin{aligned}
F_{\widetilde{Y}}(y) &= 1 - \left(1 - 1 + \frac{e^{-y}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}}}\right)^K\\
&= 1 - \frac{e^{-Ky}}{\left(y\frac{P\widehat{\theta}t}{K}+1\right)^{\widehat{D}K}}. \qquad (50)
\end{aligned}$$

Appendix C
Proof of Lemma 4

We focus on the term $e^{\frac{K^2}{P\widehat{\theta}t}}\sum_{m=1}^{\lfloor\widehat{D}K\rceil}\mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right)$ since it is the only term in β that is a function of P and t. In order to approximate this term under the assumption $Pt\to\infty$, we make use of a recurrence relation given in [32] as

$$m\mathrm{E}_{m+1}(x) = e^{-x} - x\mathrm{E}_m(x). \qquad (51)$$

Starting with the relation (51), one can write

$$\mathrm{E}_m(x) = \frac{e^{-x}}{m-1} + \frac{e^{-x}}{(m-1)!}\sum_{k=1}^{m-2}(-x)^k + \frac{(-x)^{m-1}}{(m-1)!}\mathrm{E}_1(x), \qquad (52)$$

for any integer $m>1$.

We use (52) to expand $e^{\frac{K^2}{P\widehat{\theta}t}}\sum_{m=1}^{\lfloor\widehat{D}K\rceil}\mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right)$. The resulting expression is given in (53d). The term $\phi(K,P,t,\widehat{D},\widehat{\theta})$ in (53d) is defined as

$$\begin{aligned}
\phi(K,P,t,\widehat{D},\widehat{\theta}) &\triangleq e^{\frac{K^2}{P\widehat{\theta}t}}\mathrm{E}_1\left(\frac{K^2}{P\widehat{\theta}t}\right)\left(\frac{\Gamma\left(\left\lfloor\widehat{D}K\right\rceil-1,-\frac{K^2}{P\widehat{\theta}t}\right)}{e^{\frac{K^2}{P\widehat{\theta}t}}\Gamma\left(\left\lfloor\widehat{D}K\right\rceil-1\right)}\right)\\
&- \frac{\frac{K^2}{P\widehat{\theta}t}}{\left(1+\frac{K^2}{P\widehat{\theta}t}\right)}\left(\frac{\lfloor\left(\left\lfloor\widehat{D}K\right\rceil-1\right)!e\rfloor}{\left(\left\lfloor\widehat{D}K\right\rceil-1\right)!}-2\right)\\
&- \frac{1}{\left(1+\frac{K^2}{P\widehat{\theta}t}\right)}\left(\frac{\Gamma\left(\left\lfloor\widehat{D}K\right\rceil-1,-\frac{K^2}{P\widehat{\theta}t}\right)}{e^{\frac{K^2}{P\widehat{\theta}t}}\Gamma\left(\left\lfloor\widehat{D}K\right\rceil-1\right)}-1+\frac{K^2}{P\widehat{\theta}t}\right). \qquad (54)
\end{aligned}$$

The approximation in (53b) follows from $\sum_{k=1}^{n}\frac{1}{k}\approx\ln n+\gamma+\frac{1}{2n}$ for the finite harmonic series, applied on the second term in (53a), such that

$$\begin{aligned}
e^{\frac{K^2}{P\widehat{\theta}t}}\sum_{m=2}^{\lfloor\widehat{D}K\rceil}\frac{e^{-\frac{K^2}{P\widehat{\theta}t}}}{m-1} &= \sum_{m'=1}^{\lfloor\widehat{D}K\rceil-1}\frac{1}{m'}\\
&= \gamma+\ln\left(\left\lfloor\widehat{D}K\right\rceil-1\right)+\frac{1}{2\left(\left\lfloor\widehat{D}K\right\rceil-1\right)}.
\end{aligned}$$

The expression (53c) follows from the solution for the alternating geometric series, i.e.,

$$\sum_{k=1}^{m-2}\left(-\frac{K^2}{P\widehat{\theta}t}\right)^k = \frac{-\frac{K^2}{P\widehat{\theta}t}-\left(-\frac{K^2}{P\widehat{\theta}t}\right)^m}{\left(1+\frac{K^2}{P\widehat{\theta}t}\right)}. \qquad (55)$$

The second and third terms in (54) follow from the properties

$$\sum_{k=0}^{n}\frac{1}{k!} = \frac{\lfloor n!e\rfloor}{n!}, \quad \sum_{k=0}^{n}\frac{(-x)^k}{k!} = \frac{\Gamma(n,-x)}{e^x\Gamma(n)}, \qquad (56)$$

applied to the last two terms in (53c).

By the assumptions $Pt\to\infty$ and $e^x\mathrm{E}_1(x)\approx-\gamma-\ln(x)$ as $x\to0$, we approximate (53d) as

$$\begin{aligned}
&e^{\frac{K^2}{P\widehat{\theta}t}}\sum_{m=1}^{\lfloor\widehat{D}K\rceil}\mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right)\\
&\quad\approx\ln\left(\left\lfloor\widehat{D}K\right\rceil-1\right)+\frac{1}{2\left(\left\lfloor\widehat{D}K\right\rceil-1\right)}-\ln\left(\frac{K^2}{P\widehat{\theta}t}\right). \qquad (57)
\end{aligned}$$

Finally, substituting (57) into the expression for β, we get (35).

References

[1] W. Wang et al., "Field trial on TDD massive MIMO system with polar code," *Proc. IEEE 28th Annu. Intern. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Montreal, QC, 2017, pp. 1–6.

[2] F. Tariq, M. Khandaker, K.-K. Wong, M. Imran, M. Bennis, M. Debbah, "A Speculative Study on 6G," arXiv, 1902.06700, 2019.

[3] E. Calvanese Strinati et al., "6G: the next frontier: from holographic messaging to artificial intelligence using subterahertz and visible light communication," *IEEE Vehic. Tech. Mag.*, vol. 14, no. 3, pp. 42–50, Sept. 2019.

[4] O. Dizdar, Y. Mao, W. Han, and B. Clerckx, "Rate-splitting multiple access: a new frontier for the PHY layer of 6G," *Proc. 2020 IEEE 92nd Vehic. Tech. Conf. (VTC)* (Accepted), 2020.

[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] J. Zhang, R.W. Heath, M. Kountouris, and J. G. Andrews, "Mode switching for the multi-antenna broadcast channel based on delay and channel quantization," *EURASIP J. Adv. Signal Process.*, vol. 2009, pp. 1—29, Feb. 2009.

[7] J. Zhang, M. Kountouris, J. G. Andrews and R. W. Heath, "Multi-mode transmission for the MIMO broadcast channel with imperfect channel state information," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 803–814, March 2011.

[8] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *J. Commun. Networks*, vol. 15, no. 4, pp. 338–351, Aug. 2013.

[9] A. K. Papazafeiropoulos and T. Ratnarajah, "Deterministic equivalent performance analysis of time-varying massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5795–5809, Oct. 2015.

[10] C. Kong, C. Zhong, A. K. Papazafeiropoulos, M. Matthaiou and Z. Zhang, "Sum-rate and power scaling of massive MIMO systems with channel aging," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4879–4893, Dec. 2015.

$$e^{\frac{K^2}{P\widehat{\theta}t}} \sum_{m=1}^{\lfloor \widehat{D}K \rceil} \mathrm{E}_m\left(\frac{K^2}{P\widehat{\theta}t}\right)$$

$$= e^{\frac{K^2}{P\widehat{\theta}t}} \left(\mathrm{E}_1\left(\frac{K^2}{P\widehat{\theta}t}\right) + \sum_{m=2}^{\lfloor \widehat{D}K \rceil} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}}}{m-1} + \sum_{m=3}^{\lfloor \widehat{D}K \rceil} \frac{e^{-\frac{K^2}{P\widehat{\theta}t}}}{(m-1)!} \sum_{k=1}^{m-2} \left(-\frac{K^2}{P\widehat{\theta}t}\right)^k + \mathrm{E}_1\left(\frac{K^2}{P\widehat{\theta}t}\right) \sum_{m=2}^{\lfloor \widehat{D}K \rceil} \frac{(-\frac{K^2}{P\widehat{\theta}t})^{m-1}}{(m-1)!} \right) \tag{53a}$$

$$\approx \gamma + \ln\left(\left\lfloor \widehat{D}K \right\rceil - 1\right) + \frac{1}{2\left(\left\lfloor \widehat{D}K \right\rceil - 1\right)} + \sum_{m'=2}^{\lfloor \widehat{D}K \rceil - 1} \frac{1}{m'!} \sum_{k=1}^{m'-1} \left(-\frac{K^2}{P\widehat{\theta}t}\right)^k + e^{\frac{K^2}{P\widehat{\theta}t}} \mathrm{E}_1\left(\frac{K^2}{P\widehat{\theta}t}\right) \left(1 + \sum_{m'=1}^{\lfloor \widehat{D}K \rceil - 1} \frac{\left(-\frac{K^2}{P\widehat{\theta}t}\right)^{m'}}{m'!} \right) \tag{53b}$$

$$= \gamma + \ln\left(\left\lfloor \widehat{D}K \right\rceil - 1\right) + \frac{1}{2\left(\left\lfloor \widehat{D}K \right\rceil - 1\right)} + \sum_{m'=2}^{\lfloor \widehat{D}K \rceil - 1} \frac{-\frac{K^2}{P\widehat{\theta}t} - \left(-\frac{K^2}{P\widehat{\theta}t}\right)^{m'}}{m'!\left(1+\frac{K^2}{P\widehat{\theta}t}\right)} + e^{\frac{K^2}{P\widehat{\theta}t}} \mathrm{E}_1\left(\frac{K^2}{P\widehat{\theta}t}\right) \left(1 + \sum_{m'=1}^{\lfloor \widehat{D}K \rceil - 1} \frac{\left(-\frac{K^2}{P\widehat{\theta}t}\right)^{m'}}{m'!} \right) \tag{53c}$$

$$= \gamma + \ln\left(\left\lfloor \widehat{D}K \right\rceil - 1\right) + \frac{1}{2\left(\left\lfloor \widehat{D}K \right\rceil - 1\right)} + \phi(K, P, t, \widehat{D}, \widehat{\theta}). \tag{53d}$$

[11] S. Kashyap, C. Mollén, E. Björnson and E. G. Larsson, "Performance analysis of (TDD) massive MIMO with Kalman channel prediction," *Proc. 2017 IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, New Orleans, LA, 2017, pp. 3554–3558.

[12] H. Yin, H. Wang, Y. Liu and D. Gesbert, "Addressing the curse of mobility in massive MIMO with prony-based angular-delay domain channel predictions," *IEEE J. Sel. Areas Commun.* , Early Access, 2020.

[13] 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. 2018, no. 1, p. 133, 2018.

[14] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.

[15] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.

[16] E. Piovano, B. Clerckx, "Optimal DoF region of the K-user MISO BC with partial CSIT," *IEEE Commun. Lett.*, vol. 21, no.11, pp. 2368-–2371, Nov. 2017.

[17] 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.

[18] Y. Mao and B. Clerckx, "Beyond dirty paper coding for multi-antenna broadcast channel with partial CSIT: a rate-splitting approach," *IEEE Trans. Commun.* , Early Access, 2020.

[19] 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.

[20] 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, March 2020.

[21] 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, Sept. 2015.

[22] 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, July 2016.

[23] M. Dai and B. Clerckx, "Multiuser millimeter wave beamforming strategies with quantized and statistical CSIT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7025-7038, Nov. 2017.

[24] T. Kim, D. J. Love and B. Clerckx, "MIMO systems with limited rate differential feedback in slowly varying channels," *IEEE Trans. Commun.*, vol. 59, no. 4, pp. 1175–1189, April 2011.

[25] J. G. Proakis, "Digital Communications," 4th edition. McGraw Hill, 2000.

[26] D. Jaramillo-Ramírez, M. Kountouris and E. Hardouin, "Coordinated multi-point transmission with imperfect CSI and other-cell interference," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 1882–1896, April 2015.

[27] C. Lim, T. Yoo, B. Clerckx, B. Lee and B. Shim, "Recent trend of multiuser MIMO in LTE-advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 127-135, March 2013.

[28] G. Liu et al., "3-D-MIMO With Massive Antennas Paves the Way to 5G Enhanced Mobile Broadband: From System Design to Field Trials," *IEEE Journ. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1222-1233, June 2017.

[29] N. Jindal, J. G. Andrews and S. Weber, "Rethinking MIMO for wireless networks: linear throughput increases with multiple receive antennas," *Proc. 2009 IEEE Intern. Conf. Commun. (ICC)*, Dresden, 2009, pp. 1–6.

[30] N. Jindal, J. G. Andrews and S. Weber, "Multi-antenna communication in ad hoc networks: achieving MIMO gains with SIMO transmission," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 529–540, Feb. 2011.

[31] M. S. Milgram, "The generalized integro-exponential function," *Math. Comput.*, vol. 44, no. 170, 1985, pp. 443–458.

[32] Chiccoli, C., Lorenzutta, S., Maino, G., "Recent results for generalized exponential integrals," *Comp. & Math. App.*, vol. 19, issue 5, 1990, pp. 21–29.

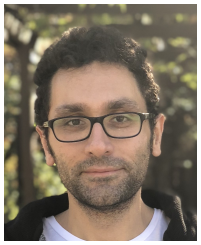
[33] E. Arıkan, "Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[34] O. Dizdar, Y. Mao, W. Han, and B. Clerckx, "Rate-splitting multiple access for downlink multi-antenna communications: physical layer design and link-level simulations," *Proc. IEEE 31st Annu. Intern. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, London, UK, 2020, pp. 1–6.

[35] S. Jaeckel, L. Raschkowski, K. Börner, L. Thiele, F. Burkhardt and E. Eberlein, "QuaDRiGa - quasi deterministic radio channel generator, user manual and documentation," Fraunhofer Heinrich Hertz Institute, Tech. Rep. v2.2.0, 2019.

[36] 3GPP TR 36.873 v12.7.0, "Study on 3D channel model for LTE," Tech. Rep., 2018.

[37] Q. H. Spencer, A. L. Swindlehurst and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Proc.*, vol. 52, no. 2, pp. 461-471, Feb. 2004.



Onur Dizdar Onur Dizdar received the B.Sc. and M.Sc. degrees in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey, in 2008 and 2011. He received his Ph.D. in Electrical and Electronics Engineering from Bilkent University, Ankara, Turkey, in 2017. He also worked as a communications system design engineer in ASELSAN, Turkey from 2008 to 2019. He is currently a postdoctoral research associate in the Communications and Signal Processing Group at Imperial College London. His research interests include wireless communications, error-correcting codes and decoding algorithms, rate-splitting and non-orthogonal multiple access and signal processing.



Yijie Mao Yijie Mao received the B.Eng. degree from Beijing University of Posts and Telecommunications, and the B.Eng. (Hons.) degree from Queen Mary University of London (London, United Kingdom) in 2014. She received the Ph.D. degree in the Electrical and Electronic Engineering Department from the University of Hong Kong (Hong Kong, China) in 2018. She was a Postdoctoral Research Fellow at the University of Hong Kong from 2018 to 2019. She is currently a postdoctoral research associate with the Communications and Signal Processing Group (CSP), Department of the Electrical and Electronic Engineering at Imperial College London. Her research interests include Multiple Input Multiple Output (MIMO) communication networks, rate-splitting and non-orthogonal multiple access for 5G and beyond.



Bruno Clerckx Bruno Clerckx is a (Full) Professor, the Head of the Wireless Communications and Signal Processing Lab, and the Deputy Head of the Communications and Signal Processing Group, within the Electrical and Electronic Engineering Department, Imperial College London, London, U.K. He received the M.S. and Ph.D. degrees in Electrical Engineering from the Université Catholique de Louvain, Louvain-la-Neuve, Belgium, in 2000 and 2005, respectively. From 2006 to 2011, he was with Samsung Electronics, Suwon, South Korea, where he actively contributed to 4G (3GPP LTE/LTE-A and IEEE 802.16m) and acted as the Rapporteur for the 3GPP Coordinated Multi-Point (CoMP) Study Item. Since 2011, he has been with Imperial College London, first as a Lecturer from 2011 to 2015, Senior Lecturer from 2015 to 2017, Reader from 2017 to 2020, and now as a Full Professor. From 2014 to 2016, he also was an Associate Professor with Korea University, Seoul, South Korea. He also held various long or short-term visiting research appointments at Stanford University, EURECOM, National University of Singapore, The University of Hong Kong, Princeton University, The University of Edinburgh, The University of New South Wales, and Tsinghua University. He has authored two books on "MIMO Wireless Communications" and "MIMO Wireless Networks", 200 peer-reviewed international research papers, and 150 standards contributions, and is the inventor of 80 issued or pending patents among which 15 have been adopted in the specifications of 4G standards and are used by billions of devices worldwide. His research area is communication theory and signal processing for wireless networks. He has been a TPC member, a symposium chair, or a TPC chair of many symposia on communication theory, signal processing for communication and wireless communication for several leading international IEEE conferences. He was an Elected Member of the IEEE Signal Processing Society SPCOM Technical Committee. He served as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He has also been a (lead) guest editor for special issues of the EURASIP Journal on Wireless Communications and Networking, IEEE ACCESS, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, and the PROCEEDINGS OF THE IEEE. He was an Editor for the 3GPP LTE-Advanced Standard Technical Report on CoMP. He is an IEEE Communications Society Distinguished Lecturer 2021-2022.