This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787



# Link Adaptation for Rate Splitting Systems with Partial CSIT

Carlos Mosquera, Felipe Gómez-Cuba

* atlanTTic Research Center, Universidade de Vigo, Vigo, Spain
Email: mosquera@gts.uvigo.es, gomezcuba@gts.uvigo.es

***Abstract*—Rate Splitting (RS) has emerged as a flexible multiple-access scheme for Multi-User Multiple-Input Multiple-Output (MU-MIMO) systems, generalizing more conventional approaches such as linear precoding or Non-Orthogonal Multiple Access (NOMA), and providing additional robustness against imperfect channel state information at the transmitter. Actual achievable rates cannot be pre-calculated when facing imperfect and/or outdated channel state information at the transmitter (CSIT), so practical real-time mechanisms in the form of link adaptation (LA) are required. The available CSIT is used to design the RS precoders, which can accommodate some degree of robustness, whereas the physical layer feedback reports the outcome of the decoding of the codewords, in the form of acknowledgement (ACK) or negative ACK (NACK), providing the additional meta-robustness needed when the statistical model of CSIT errors is not accurate, and/or the precoders cannot fully track the channel dynamics. We characterize the outage rate region of RS and analyze the convergence of the complete LA-RS system under partial CSIT (non-accurate and non-permanent), providing relevant insights for addressing the adaptation of private and common rates, designing robust precoders with throughput as relevant metric, and handling the time-variant CSIT quality.**

***Index Terms*—Rate Splitting, multiuser MIMO, outer loop link adaptation, outage rate region.**

## I. Introduction

Wireless Multi-User Multiple-Input Multiple-Output (MU-MIMO) is a well established technology to serve several users simultaneously from a multiple-antenna base station (BS) on the same time and frequency resources. Most incarnations require some sort of channel knowledge at the BS. For example, linear precoding is an efficient scheme to avoid interference at the transmitter side, with the receiver treating the residual interference as noise. Most recently, Rate Splitting (RS) has gained traction as a promising enhancement over classic linear MU-MIMO schemes [1]. Practical RS requires linear precoding at the transmitter and Successive Interference Cancellation (SIC) at the receivers. The user messages are split into common and private parts, in such a way that the common parts, encoded together, are decoded by all receivers. After that, the common message interference is completely removed, and each private message is decoded by the corresponding user, following a SIC process.

In the RS scheme, the transmitter needs to compute private and common precoding vectors, and set the transmission rate for the selection of the encoder and symbol mapping. As a design rule for any MU-MIMO scheme, the robustness of the operation against the inaccuracies of the available channel knowledge is to be prioritized. Due to its greater practical challenges, it becomes instrumental to address the case of imperfect Channel State Information at the Transmitter (CSIT). CSIT can be hampered by the errors in the estimation of the channel response, limited channel quality indication feedback, and/or time variability which is not fully tracked.

Regardless of the cause of CSIT inaccuracy, practical RS implementations require some form of Link Adaptation (LA) to adjust the transmission rates, which must be chosen carefully to keep decoding errors within admissible limits. It is key to characterize the Block Error Rate (BLER) associated with each Modulation and Coding Scheme (MCS) that LA can select. Even under perfect CSIT, the BLER will not be zero due to the use of finite length codes or suboptimal detection and decoding, among other factors. In addition, imperfect CSIT precludes the knowledge of the actual instantaneous achievable rates. In scenarios with a fully accurate statistical model of the CSIT error, the achievable rates vs. outage can be pre-calculated for the different receivers, performing *open loop* rate selection for the desired outage probability. Nonetheless, when it comes to the imperfect statistical knowledge of CSIT, it is more robust to rely on statistical outage metrics and some sort of feedback control mechanism.

### A. Prior work

*1) Rate Splitting:* RS has been shown to provide merits in multi-user settings under perfect and imperfect CSIT. With perfect CSIT, RS generalizes Space Division Multiple Access (SDMA) based on linear precoding, power-domain Non-

Orthogonal Multiple Access (NOMA), Orthogonal Multiple Access (OMA) and multicasting, which are but special instances [2], [3]. RS is also especially suited for those cases with imperfect CSIT, since the SIC operation at the receivers attenuates the additional interference caused by CSIT uncertainty. This fact is confirmed by the optimality of RS in the Degrees of Freedom (DoF) [4], [5] and generalized DoF (GDoF) [6] frameworks, under which the channel estimation error decays with increasing Signal-to-Noise Ratio (SNR) at a rate of $\mathcal{O}(\text{SNR}^{-\alpha})$ with $\alpha \geq 0$. In short, DoF and GDoF are metrics which serve to approximate the capacity at very large SNR, useful for benchmarking purposes in the absence of actual capacity expressions. In general, models for the CSI uncertainty, either stochastic or deterministic, are used to optimize metrics such as transmit power [7], max-min fairness with a bounded uncertainty set in the form of a ball [4], [8], the average sum-rate under a stochastic CSIT model [9], or the ergodic sum-rate in a cloud radio access network with a statistical CSIT description [10]. [11] considers random vector quantization for finite feedback, whereas RS precoders for the multigroup multicast case are designed in [12], [13]. In some extreme settings, such as [14], phase information in the CSIT may be absent all-together, which poses additional hurdles for the transmission, and space-time coding is required along with RS.

Several prior works have addressed the computational cost of obtaining the RS precoders, which can be substantial especially if they have to be recomputed for every encoding block under a permanent renewal of the channel estimate. The sum-rate maximization problem is non-convex, and different methods can be found in the literature such as Weighted MMSE (WMMSE) [9], Successive Convex Approximation (SCA) [15], [16], or Semidefinite Relaxation (SDR) [17]. In the former case, the WMMSE approach, introduced in [18], allows to transform the sum-rate maximization problem into a convex problem which is solved by Alternating Optimization. A previous sample average step is needed to account for statistical uncertainty in the CSIT values, rephrasing the associated stochastic problem into a deterministic counterpart [9]. These methods, despite lacking proven convergence, usually find solutions close to the global optimal as discussed recently in [19] after extensive numerical experiments. However, the high cost associated to the recalculation of the RS precoders is stimulating the search for alternatives which can alleviate this complex task. For example, [20] and [21] propose to use precoders with lower computational burden, such as Minimum Mean Square Error (MMSE) or Zero Forcing (ZF) for the private symbols, which are recomputed for each codeword, and put the effort on the allocation of power to private and common messages; this could follow an adaptive strategy [20] or closed-form expressions after some simplifications [21]. In any case, the permanent feedback of the channel estimate is required. Note that a careful power allocation is instrumental to achieve a competitive performance with different types of non-orthogonal signalling schemes, such as power-domain NOMA [22], in addition to RS.

It is in this context that our paper proposes a broader model of LA-RS for MU-MIMO links. By leveraging LA, meta-robustness can be provided for scenarios with imperfect CSIT of *unknown statistics*. Moreover, our scheme also supports additional error in tracking channel variation over time, reducing the CSIT feedback and precoding recalculation periodicity.

*2) Link Adaptation:* The characterization of the outage probability and/or the Signal-to-Interference-plus-Noise Ratio (SINR) distribution at the receivers of MU-MIMO systems is not an easy task. The optimization of the outage probability can be addressed to maximize, for example, the sum throughput[1], provided that all the required channel statistical information is available, similarly to the single user case [23]. When defining the outage capacity region, it is necessary to determine very precisely which states are declared in outage, given the existence of a common message to be decoded by all terminals, absent in seminal works dealing with fading multi-user channels such as [24]. For those cases for which the operation conditions allow it, worst case or ergodic rates are metrics which have been proposed with RS precoding, see [4] or [9].

An estimation-induced outage capacity is defined in [25], where the outage capacity is refined through the acquired knowledge of the channel. In a similar line of thought, CSI is obtained by minimum mean square error (MMSE) estimation in [26], so that the unknown channel is Gaussian distributed, with a mean value equal to the currently available channel estimate; a back-off margin based on the nominal statistical CSIT description is then applied to guarantee the prescribed outage probability.

The recent work [27] performs an outage analysis of transmit beamforming for Multi-User Multiple-Input Single-Output (MU-MISO) with outdated CSI; in this case, linear Zero-Forcing precoders are employed, and the relevance of a back-off margin[2] to keep the outage probability under control is stressed. According to [27], this margin should be adaptively chosen to match the user Doppler. Thus, the use of LA algorithms reveals as instrumental to guarantee practical operation performance indicators. Some occasional references to its use in combination with SIC detection have been made in the literature [28], although no specific discussion has been found. In the case of RS, and for the purpose of MCS selection, [29]

[1]The throughput, explicitely defined in Section VI, is the effective received bit rate, i.e., successfully decoded.

[2]Scaling factor is the chosen denomination in [27] for the SINR back-off term.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787



assumes perfect CSIT, whereas [30] uses back-off margins to maximize the throughput for a given maximum admissible BLER, which are obtained by simulations.

Specific link adaptation algorithms have been studied extensively, mainly for single user systems, as in [31], [32] or [33]. among others. In the case of multiuser systems, for which analysis might become easily intractable, fewer results are available, in several instances resorting to machine learning. This is the case, for example, of [34], which considers link adaptation for multicast, whereas [35] applies data-driven link adaptation for multiuser MIMO, in both works making use of Support Vector Machines.

### B. Contributions

We study a MU-MIMO LA-RS system with imperfect CSIT feedback. To the best of the authors' knowledge, the interplay of LA and RS subsystems has not been studied in prior work. The channel is assumed to follow a block-fading model, which remains constant for the duration of a codeword. We will assume perfect operation at the receiver, with no degradation coming from the channel state information at the receiver (CSIR) and SIC, to focus on the precoding and rate selection decisions at the transmitter, under some unavoidable channel uncertainty. For some channel use blocks, a channel snapshot is available at the transmitter, after being reported by the receivers or estimated locally by the transmitter if TDD (Time Division Duplexing) is in place. Thus, the actual channel lies within an uncertainty region centered at the imperfect CSIT. This uncertainty region spans both the observation error and the slow time-drift of the channel since the last block in which CSIT was observed.

We study the problem of robust RS precoding and LA schemes against the statistical CSIT error model. Both sum-rate and outage probability are considered for design purposes, due to the uncertainty of the channel. In our model, the RS precoders are computed only when the channel estimate is updated, which may happen only at certain time instants under a low-rate feedback signalling channel. Moreover, SINR back-off margins are applied for LA rate allocation purposes, thus provisioning for the channel time variations and protecting against errors. Additionally, a specific challenge of the LA-RS operation needs to be dealt with, due to the fact that the receiver is intended to decode two messages with two rate selections, but only a single bit of protocol feedback is available for LA purposes [3]. We analyze the convergence and correct configuration of the LA-RS scheme in order to address this ambiguity, achieving the best sum-rate vs. error probability trade-off across both RS messages.

[3]This would apply for both SIC and joint decoding.

Our scheme enforces a given outage probability as the maximum admissible probability of decoding errors of the codewords at the different receivers. We will declare a decoding error at a given user if either the common or the private messages cannot be decoded. In terms of throughput computation, the common message will contribute as the proportional rate to those successful users, which means that all users contribute equally to the common message. LA, as a key feature for fading wireless channels, will be used to choose the transmission rate able to satisfy a prescribed error metric, and in such a way that a cost-effective RS precoding strategy can be used without resorting to a permanent update of the channel estimate. Note that additional signaling is needed to embed the ACK bit from every terminal at each frame, used to target a prescribed BLER which is not necessarily optimal in terms of throughput.

To summarize, the focus of this paper resides on the design of a LA scheme suited to RS precoding, with the following contributions:

i) Instantaneous rate adaptation operating at the physical layer is presented for RS systems, which avoids the need for an accurate knowledge of the statistical distribution of the CSIT uncertainty, and can endure additional CSIT errors due to the time-variation of the channel. Operation of well-established single-user link-adaptation algorithms is extended to a novel setting making user of RS and SIC.
ii) An outage characterization of RS is presented, along with its interplay with the convergence analysis of LA. Well-established conclusions of rate splitting schemes in ergodic settings are seen to hold when instantaneous rates and outage probabilities are considered.
iii) A simulation study is performed to illustrate the robustness and cost effectiveness of the proposed embodiment for static and mobile channels.

### C. Organization and notation

The rest of the paper is organized as follows: in Section II, we present the channel and CSIT models, together with the design of RS precoders. Next, in Section III, we study the framework for outage characterization of multi-user systems with RS precoding. Subsequently, Section IV addresses the embedding of LA into an RS-based system, with a discussion on convergence and ensuing implications included in Section V. We perform a numerical evaluation in Section VI, before presenting the conclusions and discussing future research lines in Section VII.

As notation convention, upper (lower) boldface letters denote matrices (vectors). When capital letters must be used for vectors, the notation $\vec{X}$ will be used instead. $(\cdot)^t$, $(\cdot)^H$ and $\mathbf{I}_N$

denote transpose, Hermitian transpose and the $N \times N$ identity matrix, respectively. $[\mathbf{A}]_{m,n}$ is the $m$th row, $n$th column entry of $\mathbf{A}$, and $\mathbf{1}$ is the all-ones vector. The operators $\|\cdot\|_q$, Prob$\{\cdot\}$, $\mathbb{E}[\cdot]$ and $\mathrm{tr}\{\cdot\}$ represent the $q$-norm, the statistical probability, the expectation, and the matrix trace, respectively. At some points, $\bar{x} \triangleq \mathbb{E}[x]$ will be also used for the expected value. $\mathcal{CN}(0, \sigma^2)$ denotes the Circularly Symmetric Complex Gaussian distribution of zero mean and variance $\sigma^2$, $\sim$ stands for "distributed as", and i.i.d. refers to independent and identically distributed. The probability density function of random variable $X$ conditioned on random variable $Y$ reads as $f(x|y)$. The function $f(x)$ is said to be $\mathcal{O}(g(x))$, with $\mathcal{O}(\cdot)$ the Landau symbol, if $|f(x)| \leq b|g(x)|$ for some constant $b$ and all values of $x$.

For reference purposes, a notation list has been included in Table I.

| Parameter | Description |
|---|---|
| $N$ | Number of antennas at the BS |
| $K$ | Number of single-antenna users |
| $P_t, \sigma_n^2$ | Transmit and noise power |
| $\mathbf{H}[b]$ | Channel response at time (block) index $b$ |
| $\hat{\mathbf{H}}[b]$ | CSIT at time index $b$ |
| $\tilde{\mathbf{H}}[b]$ | CSIT error at time index $b$ |
| $L_F$ | Frame length |
| $E_k[b]$ | Acknowledgement bit of user $k$ at time index $b$ |
| $\mathbf{e}[b]$ | Vector of users ACKs at time index $b$ |
| $s_c$, $\{s_k\}_{k=1}^K$ | Common, private symbols |
| $p_c$, $\{p_k\}_{k=1}^K$ | Common, private precoders |
| $\gamma_{c,k}, \gamma_{p,k}$ | Common, private SINR at $k$th user |
| $\hat{\gamma}_{c,k}, \hat{\gamma}_{p,k}$ | Estimate of $k$th user common and private SINR |
| $\dot{R}(\gamma)$ | Achievable rate for SNR $\gamma$ |
| $R_c, R_{p,k}$ | Common and private transmission rates |
| $F^C(x)$ | Complementary Cumulative Distribution Function |
| $m_{c,k}, m_{p,k}$ | Margins for common and private SINR of $k$th user |
| $\mathbf{m}[b]$ | Vector of private margins at time index $b$ |
| $P_{out,k}$ | Outage probability of user $k$ |

TABLE I: Notation summary.

## II. System Model

### A. Discrete Equivalent Channel Model

We consider a time-varying frequency-flat Multiple-Input Single-Output Broadcast Channel (MISO BC) with $N$ antennas at the BS and $K$ single-antenna users[4], as depicted on the top right region of Fig. 1, and $N \geq K$. The signal received by the $k$-th user at a given channel use reads as

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, k = 1, \ldots, K \quad (1)$$

where $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ is the channel vector from the transmitter to the $k$th user, $\mathbf{x} \in \mathbb{C}^{N \times 1}$ is the vector of transmit samples, and $n_k$ is the Additive White Gaussian Noise at receiver $k$, with $n_k \sim \mathcal{C}N(0, \sigma_n^2)$; the noise variance is assumed to be the same for all users. In compact form, we model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, with $\mathbf{y} \in \mathbb{C}^{K \times 1}$, $\mathbf{H} \in \mathbb{C}^{K \times N}$, $\mathbf{x} \in \mathbb{C}^{K \times 1}$, with transmit power given by $\mathbb{E}\left[\mathrm{tr}\{\mathbf{x}\mathbf{x}^H\}\right] = P_t$, and $\mathbf{n} \in \mathbb{C}^{K \times 1}$ such that $\mathbb{E}\left[\mathbf{n}\mathbf{n}^H\right] = \sigma_n^2 \mathbf{I}_K$. The transmit SNR is given by $P_t/\sigma_n^2$.

We assume a slow block-fading model, in which the channel remains constant at least for the duration of a codeword. When the block time index $b$ is of relevance, we will use the notation $\mathbf{H}[b]$ for the corresponding channel response; otherwise, the index will be dropped for simplicity, as in Fig. 1. For generality purposes, an arbitrary statistical dependence $f(\mathbf{H}[b+1]|\mathbf{H}[b])$ between successive channel observations $\mathbf{H}[b]$ and $\mathbf{H}[b+1]$ can be expected.

### B. CSIT Feedback Model

We investigate a limited feedback scenario, under which the channel entries are estimated and reported only for some blocks, so that delay and error can be expected. Therefore, the bit rate cannot be matched to the instantaneous fading of each block, as in a fast-fading regime [38]. Every several blocks, the RS precoding vectors are computed, in response to the latest CSIT, and possibly taking into account the fading statistics for the upcoming channel realizations. For the sake of clarity as well as tractability, we will restrict our analysis to the case of a regular CSIT feedback being received once every $L_F$ blocks, which form a frame. This framing structure is depicted in Fig. 2. Nonetheless, our model could readily be extended to scenarios with irregular, unreliable or random inter-feedback intervals, which we leave for future work.

The RS precoders to be used for the next upcoming $L_F$ transmissions can be obtained under different schemes, some of them described in Section II-C. In addition, and for each block, the LA subsystem detailed in Section IV is in charge of selecting the transmission rates. For this, we assume that each block is decoded independently, and LA seeks to enforce a prescribed maximum probability of outage. Although joint decoding of several consecutive frames could be considered as an extension of our work, this could increase the application delay, so many practical systems perform decoding by blocks. Thus, each block is a separate codeword that is decoded independently, with success or failure outcome reported with one bit of "protocol feedback".

In every CSIT feedback transmission, we assume a stochastic model of imperfect CSIT in which the transmitter receives an "extracted channel estimate" provided by the receiver at

[4]We consider single-antenna receivers for simplicity. Multi-antenna terminals have been also explored in several RS works, see, e.g., [29], [36] or [37].

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787



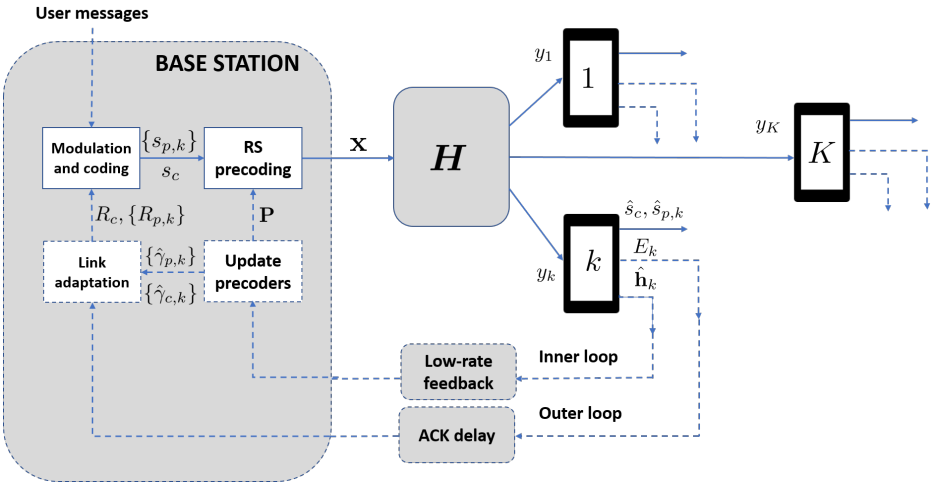



Fig. 1: Diagram of the RS channel, information exchanges, feedback and LA procedure. Dashed lines represent exchanges at block (codeword) level.

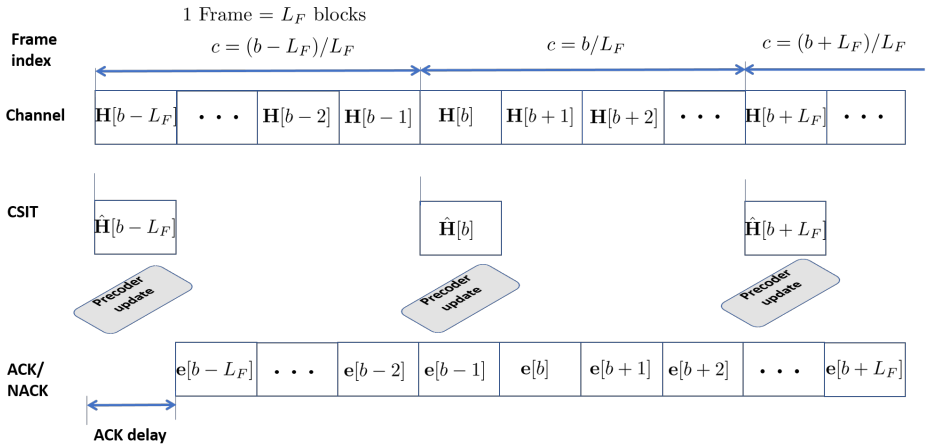



Fig. 2: Frame structure, including channel blocks, CSIT and ACK/NACK from the perspective of the BS. In this example, the decoding ACK/NACK are received with one block delay. Precoders which are designed based on $\hat{\mathbf{H}}[cL_F]$ start operating at the block $cL_F + 1$.

a given time instant[5]. For example, the transmitter would get $\hat{\mathbf{H}}[cL_F], c = 0, 1, \ldots$ in the case of Fig. 2. Thus, the true channel at block $b$ is written as

$$\mathbf{H}[b] = \hat{\mathbf{H}}[b - \ell] + \tilde{\mathbf{H}}[b], \tag{2}$$

where the offset $\ell = b \mod L_F$ represents the age of the latest received CSIT feedback at block $b$. This channel contains randomly distributed elements according to a distribution $f(\tilde{\mathbf{H}}[b])$ independent of $\hat{\mathbf{H}}[b - \ell]$. A Gaussian channel error model is commonly used in the literature $[\tilde{\mathbf{H}}[b]]_{m,n} \sim \mathcal{CN}(0, \sigma_e^2)$; see, e.g., [7] and references therein. Moreover, the variance of the CSIT error is assumed to be known in the vast majority of studies; for example a CSIT quality factor $\alpha$ such that $\sigma_e^2 = \beta \cdot \text{SNR}^{-\alpha}$ is assumed in [4]. Only in cases when $f(\tilde{\mathbf{H}}[b])$ can be statistically characterized, the transmission rates at the MCS subsystem can be selected without any explicit decoding report from the receivers. Otherwise, feedback-based LA is necessary even for the case when there is feedback in every single block, with frame duration $L_F = 1$.

Due to the reasons above, in this paper we assume that the covariance matrix of the CSIT may be difficult to obtain. Our scheme only requires the general assumption that both mean and covariance are slowly-varying channel statistics [39]. In order to operate in settings with unknown $f(\tilde{\mathbf{H}}[b])$, the design of a scheme with meta-robustness is one of the main contribu-

[5]This would be the case for Frequency Division Duplex (FDD) systems. In case of Time Division Duplex (TDD), channel estimation for the downlink can take place at the BS.

tion of our work. In exchange, we must consider carefully the LA subsystem, which acts as the "uncertainty safety net" by adjusting the transmission rates to satisfy a given performance metric over several successive blocks that share a common latest feedback, i.e., blocks $\mathbf{H}[b], \mathbf{H}[b+1], \ldots, \mathbf{H}[b+L_F-1]$ for known $\hat{\mathbf{H}}[b]$.

As we discussed in Sec. I-A, a drawback of the works that address the highly non-convex extraction of optimal precoders, given CSIT $\hat{\mathbf{H}}[b]$, is their excessive computational complexity; these include the WMMSE algorithm [9] or the use of Successive Convex Approximation (SCA) [40], in both cases to maximize the weighted sum-rate. Some works can also be found which make use of faster-computation suboptimal precoders such as ZF [2], regularized ZF or block diagonalization [21], which can be computed within one block period, after assuming feedback in every block. However, there may be scenarios where the suboptimal precoder performance is unacceptable [40].

Thus, the two-fold innovation of our system model is the study of Rate Splitting permitting communications with unknown statistical characterization of the CSIT error in the case $L_F = 1$, and also the extension to the case $L_F > 1$, which reduces the precoding CSIT feedback periodicity, so that some additional complexity on the calculation of the RS precoders can be afforded if needed.

A key to support more infrequent feedback is that, in addition to channel estimation and quantization errors, the additive imperfect CSIT model (2) may also characterize other forms of uncertainty. For example, the CSIT $\hat{\mathbf{H}}[b]$ may correspond to the Line of Sight (LoS) component of a Rician fading channel with known LoS propagation but unknown multipath scattering, or known information about spatial correlations between channels of adjacent users, or outdated CSIT in a temporally-correlated channel [41].

Robustness in the design of RS precoders for CSIT model such as (2) can be provided by methods such as WMMSE, usually under the convention that the statistical description is accurately known. The embedding of LA as in Fig. 1 provides additional robustness, since the variance of $\tilde{\mathbf{H}}[b+\ell]$ in (2) may not be accurately known, and it might be even time variant. It is important to remark again that the notation (2) is fully general as it admits any arbitrary distribution for the changes in the channel since the last feedback, denoted $f(\tilde{\mathbf{H}}[b+\ell]) \equiv f(\mathbf{H}[b+\ell]|\hat{\mathbf{H}}[b])$. In the numerical results in Section VI we will showcase two well-known channel models:

1) **Static Rice channel.** In the first example we consider a LOS channel where the LOS component is invariant and known, whereas a random i.i.d. Rician multipath scattering affects each block. Therefore $L_F = \infty$, $\hat{\mathbf{H}}[b] = \mathbf{H}_{\text{LOS}}\ \forall b$, and the actual channels follow a distribution $f(\mathbf{H}[b]|\hat{\mathbf{H}}[b]) \sim \mathcal{CN}(\mathbf{H}_{\text{LOS}}, \sigma_e^2 \mathbf{I}_{NK})$ that is independent in each block.
2) **Mobile Rayleigh channel.** The second case corresponds to a first-order Markov renewal for spatially uncorrelated Rayleigh flat fading channel, with $\mathbf{H}[b] = \rho \mathbf{H}[b-1] + \sqrt{1-\rho^2}\mathbf{G}[b]$, where $\mathbf{G}[b] \sim \mathcal{CN}(0, \mathbf{I}_{NK})$ is a Gaussian i.i.d. "channel update" random variable and $\rho \in [0,1]$. While for $L_F = 1$ the channel update does not need to be reflected in the error distribution, in the case $L_F > 1$ we would have that different blocks respond to slightly different CSIT error distributions. More concisely, the combination of CSIT error and channel evolution in time can be modeled as $f(\mathbf{H}[b]|\hat{\mathbf{H}}[b-\ell]) \sim \mathcal{CN}(\rho^\ell \hat{\mathbf{H}}[b], (\rho^{2\ell}\sigma_e^2 + 1 - \rho^{2\ell})\mathbf{I}_{NK})$, for $b = cL_F + \ell$, with integer $c$ and $0 \leq \ell < L_F$.

These two examples illustrate how the CSIT error variance in (2) may be unknown or time-variant in some settings.

### C. Rate Splitting: precoder design

In the RS scheme, the transmitted symbol vector $\mathbf{x}$ is constructed as a linear combination of $K+1$ message sequences as

$$\mathbf{x} = \mathbf{p}_c s_c + \sum_{k=1}^{K} \mathbf{p}_k s_k \tag{3}$$

where $s_c$ is the *common symbol*, which will be first decoded by all receivers, and $\{s_k\}_{k=1}^K$ are the private symbols, which are each decoded only by their respective receiver. We assume that $\mathbb{E}\left[\mathbf{s}\mathbf{s}^H\right] = \mathbf{I}_{K+1}$, with $\mathbf{s} = [s_c, s_1, \ldots, s_K]^T$. With this, the rate towards each receiver user is split by the transmitter into private and common messages, with the latter being concatenated into a common group message to be decoded by all users and removed prior to the decoding of the private streams. The linear combination is achieved by means of the common and private precoding vectors $\mathbf{p}_c$ and $\{\mathbf{p}_k\}_{k=1}^K$, respectively. For convenience, we can write $\mathbf{x} = \mathbf{P}\mathbf{s}$ where $\mathbf{P} = [\mathbf{p}_c|\mathbf{p}_1|\ldots|\mathbf{p}_K]$, and thus the transmit power is expressed as a function of the precoders: $\mathbb{E}\left[\text{tr}\{\mathbf{x}\mathbf{x}^H\}\right] = \text{tr}\{\mathbf{P}\mathbf{P}^H\} = \|\mathbf{p}_c\|_2^2 + \sum_{k=1}^K \|\mathbf{p}_k\|_2^2 = P_t$.

For our model, we will distinguish two SINR values per user $k$, $\gamma_{c,k}$ and $\gamma_{p,k}$, associated with the common and private streams, respectively. The common stream is decoded first, surrounded by interference from all private streams, including that of $k$. The receiver knows $\mathbf{h}_k$ perfectly, and thus the common SINR reads as

$$\gamma_{c,k} = \frac{|\mathbf{h}_k^H \mathbf{p}_c|^2}{\sum_{j=1}^K |\mathbf{h}_k^H \mathbf{p}_j|^2 + \sigma_n^2}. \tag{4}$$

Next, each user can remove the interference of the term $\mathbf{h}_k^H \mathbf{p}_c s_c$ from $y_k$, and thus the private stream is decoded with SINR

$$\gamma_{p,k} = \frac{|\mathbf{h}_k^H \mathbf{p}_k|^2}{\sum_{j \neq k} |\mathbf{h}_k^H \mathbf{p}_j|^2 + \sigma_n^2}. \quad (5)$$

The design of the RS precoders can follow conventional MU linear precoding schemes such as ZF, regularized ZF (RZF), or block diagonalization [40]; also, numerical methods to maximize the weighted sum-rate have been used [4], [9], [42], to address this highly non-convex problem. Imperfect CSIT can cause an important performance degradation, so robustness has been embedded in several designs, mainly in the iterative methods which try to find close-to-optimal solutions such as WMMSE [4] or SCA [40]. In this paper, we do not propose a new design scheme for RS precoders, but rather focus on the interplay between the inner and the outer loop in Fig. 1, i.e., the relation between the design of the precoders and the satisfaction of given error metrics in the presence of imperfect channel models. For simulation purposes, we will adopt two particular schemes featuring two different approaches: (i) a simpler method which avoids inter-user interference for the private messages by means of ZF precoding, see, e.g., [2] or [20], labeled as RS-ZF; and (ii) a more robust scheme, WMMSE [9], which maximizes the average sum-rate at the cost of much higher complexity. This maximization is as good as the statistical model of the CSIT error in (2), which may be imperfect.

More in detail, the selected design methods operate as follows:

i) **RS-ZF**. The private precoders are designed to cancel out the multiuser interference of the private messages: $[\mathbf{p}_1, \ldots, \mathbf{p}_K] = \hat{\mathbf{H}}(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)^{-1}$. The common precoder is chosen as the first column of matrix $\mathbf{V}$ in the SVD of $\hat{\mathbf{H}} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$ [9], [20], [21]. The transmit power is divided between common and private precoders so that the sum rate is maximized (this can be solved analytically for the $N = K = 2$ case, see [2]).

ii) **WMMSE**. The expectation of the achievable rates $\dot{R}(\gamma)$, detailed in Def. 1, and conditioned on $\hat{\mathbf{H}}$, is written as

$$\bar{R}_{c,k} \triangleq \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}[\dot{R}(\gamma_{c,k})|\hat{\mathbf{H}}],\ \bar{R}_{p,k} \triangleq \mathbb{E}_{\mathbf{H}|\hat{\mathbf{H}}}[\dot{R}(\gamma_{p,k})|\hat{\mathbf{H}}]. \quad (6)$$

The RS precoding design is now based on the maximization of the expected achievable rate:

$$\max_{\mathbf{P}} \bar{R}_c + \sum_{k=1}^{K} \bar{R}_{p,k} \quad \text{s.t. } \bar{R}_{c,k} \geq \bar{R}_c,\ \mathrm{tr}\{\mathbf{P}\mathbf{P}^H\} \leq P_t \quad (7)$$

where the precoder vectors are packed together in $\mathbf{P}$, and the maximum transmit power is $P_t$. This stochastic problem is first formulated as a deterministic version using the Sampling Average Approximation, and then solved using the WMMSE approach by means of the Alternating Optimization algorithm, as discussed in [4][6].

The RS precoder design is performed once per frame, every $L_F$ blocks, when a new channel estimate $\hat{\mathbf{H}}[b]$ is extracted. Actual performance data in the form of packet acknowledgements will be used to support the selection of transmission rates, as exposed in Section IV, since bit rates can only be matched to the channel statistics, which may not even fully known.

## III. Rate Splitting: Outage Analysis

A formal framework for the outage characterization of the downlink RS multiple user case under partial CSIT is presented in this section. This is a prerequisite to understand the operation of adaptive algorithms aimed to operate with uncertainty at the transmitter leading to some unavoidable decoding outage. Most of the RS literature adopts an ergodic approach, with only a few mentions to an outage formulation such as [43], which is sidestepped, for example, by means of lower bounds on the ergodic rates. Thus, we extend the characterization of achievable outage rate regions from more conventional MIMO multiuser channels [44] to RS systems.

**Definition 1.** *We define the* ***achievable rate***

$$\dot{R}(\gamma) = \log_2(1 + \gamma) \quad (8)$$

*as the capacity of an equivalent AWGN channel where the SNR is $\gamma$. The achievable rate is widely used in the literature to design linear precoders with symbol decisions which treat the interference as noise. This is because $\dot{R}(\gamma)$ lower bounds the practical rates achievable under interference with any decoder in which the SINR is $\gamma$. The justification is two-fold: first, the worst noise distribution of known covariance is Gaussian [45]; and second, recent coding techniques such as turbo, LDPC and polar codes can almost achieve the mutual information of any given "virtual" AWGN channel of such form [38].*

Note that we assume perfect CSIR, with the achievable rates $\dot{R}_{c,k}(\gamma_{c,k})$ and $\dot{R}_{p,k}(\gamma_{p,k})$ evaluated as a function of $\mathbf{h}_k$. We adopt an outage-based link error model where the receiver is guaranteed to decode with a negligibly small probability of error provided a decent channel coding is used, and as long as the actual chosen transmit rates, $R_{c,k}$ and $R_{p,k}$, are selected below the achievable rate: $R_{c,k} < \dot{R}_{c,k}$ and $R_{p,k} < \dot{R}_{p,k}$. We assume codewords long enough as to neglect decoding

[6]Sampling Average Approximation performs a deterministic maximization of the rate averaged across the sampled uncertaintity region, as an approximation of the rate expectation, which cannot be handled analytically. Both values can differ as the knowledge of the channel statistics becomes inaccurate.

errors if the operation takes place within the admissible theoretical achievable rates. In case this assumption does not hold, imperfect decoding and SIC needs to be addressed, since successful decoding cannot be guaranteed even though the transmission rate is properly determined. The addition of these effects into the design of RS systems is a promising line of work as frequently noted in recent references, see, e.g. [46].

In the presented setting, due to the lack of perfect CSIT, the transmitter cannot choose the optimal rates nor the optimal precoding vectors easily. The selection of transmit rates is addressed by the LA outer loop. The corresponding precoding vectors are designed with the support of an inner loop (see Fig. 1), which leaves, as a result, a region of achievable rates for the LA to choose from.

Unless perfect CSI is available at the transmitter at all times, there is a non-zero probability of outage, i.e, of transmitting a bit rate which cannot be supported. Thus, given a fixed set of precoders, we define an outage achievable rate region, or outage region for brevity, as the set of rate vectors $\vec{R} \in \mathcal{R}(\epsilon)$ which can be sustained with a probability $1-\epsilon$, where we define $\vec{R} \triangleq (R_c, R_{p,1}, \ldots, R_{p,K})^T$.

**Definition 2.** *Let $\epsilon$ denote the outage probability. Then $\vec{R} \in \mathcal{R}(\epsilon)$ if*

$$\text{Prob}\{\dot{R}(\gamma_{c,k}) > R_c, \dot{R}(\gamma_{p,k}) > R_{p,k}\} \geq 1-\epsilon, \quad k = 1, \ldots, K. \quad (9)$$

*From $R = \sum_k R_{p,k} + R_c$, we can define the outage achievable sum-rate $R \leq R_{\text{out}}(\epsilon)$ as the maximum sum-rate which is achievable within the outage region, that is,*

$$R_{\text{out}}(\epsilon) = \sup_{\vec{R} \in \mathcal{R}(\epsilon)} \|\vec{R}\|_1. \quad (10)$$

*For the RS operation, we assume that the common and private messages are decoded independently, and SIC decoding is declared successful if both common and private messages are decoded without errors. Thus, we can define an individual outage region for each receiver $\mathcal{R}_k(\epsilon) = \{(R_{c,k}, R_{p,k})\}$, given by the pair of rates such that*

$$\text{Prob}\{\dot{R}(\gamma_{c,k}) > R_{c,k}, \dot{R}(\gamma_{p,k}) > R_{p,k}\} \geq 1-\epsilon, \ k = 1, \ldots, K. \quad (11)$$

*With this, the full outage region of the MU-MIMO system* (9) *is immediately constructed from the Cartesian product of the individual regions $\mathcal{R}_k(\epsilon)$.*

Since we assume that the user channels $\mathbf{h}_k$ are i.i.d., possibly except for different received SNR, we will focus on the case when the same outage probability is enforced for all user terminals. This permits to better understand the dynamics in which different terminals could be in outage at different time instants, producing only partial system outages, and thus studying the marginal axes of the outage region corresponding to individual terminals separately. With this, no optimization of common rate allocation is performed, since the common message payload is equally divided across users. For a discussion on common rate allocation for the maximization of a certain utility function, see the recent survey [3]. More refined extensions, to account for different outage probabilities or to set a common avoidance of outage, could follow the corresponding definitions in [44], therein developed for the MISO interference channel.

Note that the expected and outage rates are related through Markov's bound

$$R_{\text{out}}(1 - \text{Prob}\{R \leq R_{\text{out}}\}) \leq \mathbb{E}[R], \quad (12)$$

so that the maximization of the expected rate, as described in Sec. II-C, pulls-up indirectly the outage achievable rates through their upper bound.

For given precoders and maximum transmit power $P_t$, the BS computes an estimate of the SINR levels that the different codewords will experience at the user terminals at time index $b$, using the channel estimates in place of the true values:

$$\begin{aligned} \hat{\gamma}_{c,k}[b] &= \frac{|\hat{\mathbf{h}}_k[b]^H \mathbf{p}_c|^2}{\sum_{j=1}^K |\hat{\mathbf{h}}_k[b]^H \mathbf{p}_j|^2 + \sigma_n^2} \\ \hat{\gamma}_{p,k}[b] &= \frac{|\hat{\mathbf{h}}_k[b]^H \mathbf{p}_k|^2}{\sum_{j \neq k} |\hat{\mathbf{h}}_k[b]^H \mathbf{p}_j|^2 + \sigma_n^2}, \end{aligned} \quad (13)$$

for $k = 1, \ldots, K$.

Since there is a correspondence between the random variable $\mathbf{H}[b]|\hat{\mathbf{H}}[b-\ell]$ and the pair of random variables $(\gamma_{c,k}[b]|, \gamma_{p,k}[b])|\hat{\mathbf{H}}[b-\ell]$, the outage probability of each user $k$ is evaluated with the complementary CDF (CCDF) of the pair of SINRs

$$P_{out,k}[b] = 1 - F^C_{(\gamma_{c,k}[b], \gamma_{p,k}[b])}\left(2^{R_c} - 1, 2^{R_{p,k}} - 1\right), \quad (14)$$

where it is important to note that the two random variables are dependent, and the joint CCDF takes the form $F^C_{(\gamma_{c,k}, \gamma_{p,k})}(x, y) = (1 - F_{\gamma_{c,k}}(x))(1 - F_{\gamma_{p,k}|\gamma_{c,k}}(y))$.

**Example 1.** *Figure 3 shows the individual outage regions (cf. Definition 2) in a two user case, for a SIC probability of outage $\epsilon = 0.05$ per user, SNR = 15 dB. The channel error model follows* (2)*, with the CSIT given by $\hat{\mathbf{h}}_1 = [1, 1]^H/\sqrt{2}$ and $\hat{\mathbf{h}}_2 = 0.7[1, e^{j\pi/2}]^H/\sqrt{2}$. The CSIT errors are generated by an i.i.d. random perturbation with Gaussian distribution of variance $\sigma_e^2 = SNR^{-1}$. The RS precoders were computed with the WMMSE method to maximize the average sum-rate* (7)*.*

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787



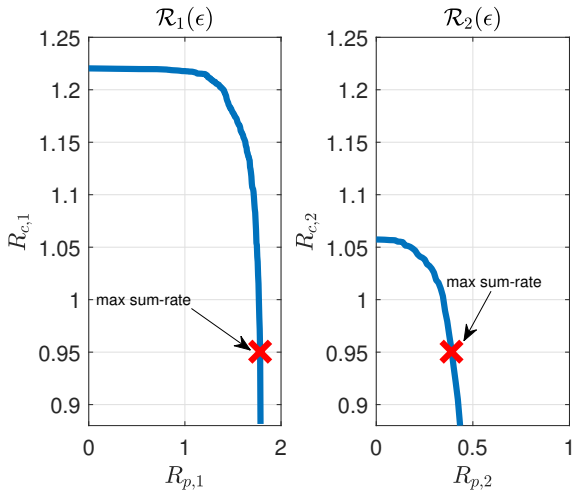



Fig. 3: Individual outage regions for a two user case, outage probability $\epsilon = 0.05$, SNR = 15 dB, CSIT error variance $\sigma_e^2 = \text{SNR}^{-1}$. The maximum sum-rate $R_{\text{out}}(\epsilon)$ corresponds to the two red cross points.

We reiterate that even though this example has assumed accurate knowledge of the channel uncertainty statistical model, so the outage regions and the optimum operation point can be easily depicted, practical CSIT errors do not necessarily follow known statistics: uncertainty margins may be difficult to estimate, or the model may not hold long enough in time. This brings up the need for an empirical adjustment of the operation margins, which can be carried out by a LA algorithm, as described in the next section. Finally, let us remark that this way of computing the outage rate, based on imperfect CSIT, finds relevant connections with previous works such as [25] and [26].

## IV. Link Adaptation

We assume that in each block the BS must choose the common message and private user rates such that outages, as defined in Def. 2, are kept within prescribed bounds. Due to the inaccuracy in the estimation of the SINR values (4)-(5), the rates are selected with some safety multiplicative margins applied on the SINR estimates. We employ the operation margins $\{m_{c,k}[b], m_{p,k}[b]\}_{k=1}^K \in \mathbb{R}^{2K}$ in such a way that, for each block $b$, the BS selects the transmission rates as those that would be achievable after applying the corresponding scaling on the estimated SINR:

$$\begin{aligned} R_c[b] &= \min_k \dot{R}(\hat{\gamma}_{c,k}[b] \cdot m_{c,k}[b]) \\ R_{p,k}[b] &= \dot{R}(\hat{\gamma}_{p,k}[b] \cdot m_{p,k}[b]). \end{aligned} \tag{15}$$

Alternatively, a single common margin could be adapted, such that $R_c[b] = \dot{R}(\min_k\{\hat{\gamma}_{c,k}[b]\} \cdot m_c[b])$. However, frequent spurious adjustments of this margin would take place, since a single bit reports the outcome of the SIC operation of each user. As a result, individual common margins are employed, in the understanding that the common rate will be set by the most restrictive user at each channel use, according to the link adaptation procedure. Under this framework, the specific task of the LA scheme will be the iterative update of the parameters $\{m_{c,k}[b], m_{p,k}[b]\}_{k=1}^K$ based on the ongoing observation of packet delivery successes and failures, reported by a single ACK bit of protocol feedback. Here, we note that the protocol feedback admits possibly different periodicity for the ACK and CSIT reporting, as depicted in Fig. 2.

For LA purposes, it is more convenient to write the margin implementation in dB, so that the estimated SINRs (13) are mapped, after being shifted, as follows for block $b$:

$$\begin{aligned} R_c[b] &= \min_k \Pi(\hat{\gamma}_{c,k}[b] + m_{c,k}[b] \text{ [dB]}) \\ R_{p,k}[b] &= \Pi(\hat{\gamma}_{p,k}[b] + m_{p,k}[b] \text{ [dB]}) \end{aligned} \tag{16}$$

where $\Pi(\cdot)$ can be any appropriate function between the "corrected SINR" in dB and the selected rate. We note that a negative margin will be needed when the true channel SINR is overestimated. In practical systems the set of MCS available to the transmitter is discrete, and the function $\Pi(\cdot)$ can map SNR intervals to discrete rates. Nevertheless, in our analysis we will not restrict ourselves to any specific MCS family, and we rather use the achievable rate (cf. Def. 1) for mapping purposes in (16). We remark that the use of more practical encoding strategies can simply be absorbed into an additional slight degradation with respect to the Shannon achievable rate, which can be taken care of by the protection margins.

As discussed previously, margins can simply be chosen for an optimal operation point, for example in the sum-rate sense, by inverting (14) if a full statistical characterization of the CSIT error is available. The proposed embodiment in this paper accommodates also the sheer lack of a CSIT error model, with the support of an LA feedback loop. Nevertheless, these two cases are extremes of a range of possibilities. Partial statistical knowledge of the CSIT error is relevant even when practical modeling errors need to be considered, since it can be exploited by an LA algorithm as prior information to speed up the convergence of the margins. These back-off margins should not be more conservative than strictly needed, so that they are adaptively tuned based on the outcome of the decoding operations, reported by the users.

Link adaptation, as a central feature of the wireless physical layer, gives rise to different implementations. Arguably, one of the most popular ones is the so called Outer Loop LA (OLLA) scheme [47]. OLLA addresses error sources due to outdated feedback, interference variability or inaccuracies in the channel report from the other end, with the goal of maintaining a specified BLER target, by exploiting the error count

reported by the corresponding receiver. The superposition of common and private messages, along with SIC detection at the receivers, generates new questions as to the relation between the margins for both types of messages. To the authors' knowledge, this has not been touched upon in the literature, even though one of the early works on NOMA [28] already pointed out the need for an optimized LA scheme.

Thus, the application of the OLLA scheme will serve to adjust at the BS the back-off margins of all users, with possibly different values for the common and private messages. We assume that only one feedback bit is available per user per block, $E_k[b]$, signaling the decoding error of either common or private messages by $k$th user at block $b$. That is, $E_k[b] = 0$ indicates that *both* common and private decoding have been successful, whereas $E_k[b] = 1$ reports an error in either one. With this, margins are updated by following a stochastic gradient descent [32]:

$$\begin{aligned} m_{c,k}[b+1] &= m_{c,k}[b] - \mu_c(E_k[b-b_0] - \epsilon) \text{ [dB]} \\ m_{p,k}[b+1] &= m_{p,k}[b] - \mu_p(E_k[b-b_0] - \epsilon) \text{ [dB]} \end{aligned} \tag{17}$$

where $b_0$ denotes the potential presence of a protocol ACK feedback delay, and $\epsilon$ is the target BLER. The notation exhibits constant stepsizes across users and time, although some additional flexibility could be embedded if needed. The top left region in Fig. 1 details the rate adaptation and precoding adaptation blocks at the transmitter. The OLLA receives one bit of feedback for every transmitted block, and adjusts the SINR margins in order to alter the transmitter rates, whereas the inner loop of CSIT feedback is actually updated more slowly, only once per frame, every $L_F$ blocks.

For the purpose of completeness, we should mention that to speed up the convergence of the basis OLLA scheme, some refinements have been proposed in the literature, such as *sequential hypothesis testing* [48] or, more recently, reinforcement learning under different Bayesian flavors [33].

## V. Convergence Analysis

The convergence properties of OLLA and tuning of stepsize for single-user systems are discussed in several references, see, e.g., [31] or [32]. Suffice to say that convergence in average is usually guaranteed for basic OLLA and some of its variants. In the proposed setting, different OLLA loops are deployed for the different messages and users, entangled through the use of common messages and multiuser interference. The framing and periodic update of precoders, as exposed in Fig. 2, complicates even further the convergence analysis of LA-RS, that we will perform for the average behaviour in the general case, with some insights of practical relevance for particular cases. We will make use of a few assumptions to simplify the study:

(A1) The ACK delay is set to $b_0 = 0$ in the recursions (17).
(A2) At a given point, channel symmetry will be invoked, so that all channels are i.i.d, with the same SINR distribution statistics.
(A3) For notation purposes, and without loss of generality, it will be convenient to consider that all margins are initially set to 0 dB.
(A4) The stepsizes for the adaptation of the private and common messages will be the same for all the users, $\mu_p$ and $\mu_c$, respectively.

As a consequence of the use of a single feedback bit $E_k[b]$ for adaptation purposes, the evolution of the common and private margins is bonded at all times, so that the two terms in (17) can be easily seen to be linearly related as

$$m_{c,k}[b] = \frac{\mu_c}{\mu_p} m_{p,k}[b] + m_{c,k}[0] - \frac{\mu_c}{\mu_p} m_{p,k}[0]. \tag{18}$$

More concisely, under assumption (A3), the second and third terms of (18) are zero. Thus, despite the fact that the private messages are decoded independently, the private margins of different users, $m_{p,k}[b]$ and $m_{p,k'}[b]$, display a dependency in their evolution caused by the common message errors and the above linear connection. Thus, the analysis that follows provides a sufficient convergence condition which depends on $\mu_p$, with the ratio $\frac{\mu_c}{\mu_p}$ bearing upon the achieved sum-rate.

For convenience, we put together the private margins and errors within the following vectors of size $K$: $\mathbf{m}[b] \triangleq (m_{p,1}[b], \ldots, m_{p,K}[b])^t$, and $\mathbf{e}[b] \triangleq (E_1[b], \ldots, E_K[b])$. From (18), we note that at any given time instant a similar vector of common margins $\mathbf{m}_c[b] \triangleq (m_{c,1}[b], \ldots, m_{c,K}[b])^t$ is a scaled version of $\mathbf{m}[b]$: $\mathbf{m}_c[b] = \frac{\mu_c}{\mu_p}\mathbf{m}[b]$.

We combine the update of all private margins in (17), and write the expected value conditioned on $\mathbf{m}[b]$ as

$$\mathbb{E}\left[\mathbf{m}[b+1]|\mathbf{m}[b]\right] = \mathbf{m}[b] - \mu_p(\mathbb{E}\left[\mathbf{e}[b]|\mathbf{m}[b]\right] - \epsilon). \tag{19}$$

The evaluation of the expected value of the errors contained in $\mathbf{e}[b]$ is highly complex, given their dependency on the CSIT, CSIT error, periodic recalculation of precoders, and relative position within the frame. If we take the expectation with respect to $\mathbf{m}[b]$, we get to the following recursion for the expected margin vector:

$$\mathbb{E}\left[\mathbf{m}[b+1]\right] = \mathbb{E}\left[\mathbf{m}[b]\right] - \mu_p(\mathbb{E}_{\mathbf{m}[b]}\mathbb{E}\left[\mathbf{e}[b]|\mathbf{m}[b]\right] - \epsilon\mathbf{1}), \tag{20}$$

which can be also written as

$$\overline{\mathbf{m}}[b+1] = \overline{\mathbf{m}}[b] - \mu_p\overline{\vec{P}}_{out}[b] + \mu_p\epsilon\mathbf{1}, \tag{21}$$

with $\overline{\mathbf{m}}[b] \triangleq \mathbb{E}\left[\mathbf{m}[b]\right]$, and $\overline{\vec{P}}_{out}[b]$ the vector of average **instantaneous** decoding error probabilities at instant $b$:

$$\overline{\vec{P}}_{out}[b] = \left[\overline{P}_{out,1}[b], \ldots, \overline{P}_{out,K}[b]\right] \tag{22}$$

where, combining (9) and (16), the outage probability for user $k = 1, \ldots, K$, can be expressed as

$$P_{out,k}[b] = 1 - \text{Prob}\{\gamma_{c,k}[b] > \hat{\gamma}_{c,k}[b], \gamma_{p,k}[b] > \hat{\gamma}_{p,k}[b]\}. \tag{23}$$

After substituting (18) and the average of (23), we get

$$\overline{P}_{out,k} = 1 - \mathbb{E}\left[F^{C}_{(\gamma_{c,k},\gamma_{p,k})}\left(\frac{\mu_c}{\mu_p}\min_{k'}\hat{\gamma}_{c,k'}m_{c,k'}, \hat{\gamma}_{p,k'}m_{p,k}\right)\right] \tag{24}$$

with the CCDF $F^{C}_{(\gamma_{c,k},\gamma_{p,k})}(\cdot,\cdot)$ defined as in (14), and where we have dropped the block index $[b]$ for notation simplicity.

More in detail, we address next two specific instances in line with the CSIT feedback models presented in Section II-B.

### A. Static Rice channel - i.i.d. CSIT

We focus on the cases where $f(\tilde{\mathbf{H}}[b]|\hat{\mathbf{H}}[b-\ell])$ are i.i.d. $\forall \ell$. An example is the static LOS Rice channel, where we have that $\hat{\mathbf{H}}[b] = \mathbf{H}_{LOS}$, and $L_F = \infty$. Banach's convergence theorem [49] can be applied to the mapping (21), expressed as $\overline{\mathbf{m}}[b+1] = T(\overline{\mathbf{m}}[b])$. For the channel model at hand, it can be readily seen that this mapping is time-invariant, so the recursion will converge to a fixed point such that $T(\overline{\mathbf{m}}^{\star}) = \overline{\mathbf{m}}^{\star}$ if the mapping is contractive.

**Definition 3.** *We say $T(\cdot)$ is a **contraction mapping** if $\exists \kappa \in [0,1)$ such that $d(T(\mathbf{m}^{(1)}), T(\mathbf{m}^{(2)})) \leq \kappa d(\mathbf{m}^{(1)}, \mathbf{m}^{(2)})$, with $d(\mathbf{x}_1, \mathbf{x}_2)$ any p-norm distance, under which $(\mathbb{R}^K, d)$ is a complete metric space.*

Even for this simple channel model, the study of the contractiveness of the update mapping becomes quite involved with respect to previous single-user single-antenna works such as [32], due to the coupling of the common message across users. A sufficient condition for the contractiveness of $T(\cdot)$ is that the $p$-norm of the Jacobian of the transformation is bounded, i.e., $\|J_T(\mathbf{m})\|_p < 1$, with $\mathbf{m}$ defined in a convex set, and $T(\mathbf{m})$ containing continuous partial derivatives. From (20)-(21), the Jacobian matrix can be expressed, for $\mathbf{m} = [m_{p,1}, \ldots, m_{p,K}]^t$, as $J_T(\mathbf{m}) = \mathbf{I}_K - \mu_p \boldsymbol{\Upsilon} \triangleq$

$$\begin{pmatrix} 1-\mu_p\Upsilon_{1,1} & -\mu_p\Upsilon_{1,2} & \cdots & -\mu_p\Upsilon_{1,K} \\ -\mu_p\Upsilon_{2,1} & 1-\mu_p\Upsilon_{2,2} & \cdots & -\mu_p\Upsilon_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ -\mu_p\Upsilon_{K,1} & -\mu_p\Upsilon_{K,2} & \cdots & 1-\mu_p\Upsilon_{K,K} \end{pmatrix} \tag{25}$$

where

$$\Upsilon_{k,k'} = \frac{\partial \overline{P}_{out,k}[b]}{\partial m_{p,k'}[b]} \tag{26}$$

are non-negative values of difficult evaluation[7]. As we saw in (24), the selection of the common rate, and therefore the outage probability of the common message of user $k$, depends on $\frac{\mu_c}{\mu_p}\min_{k'}\hat{\gamma}_{c,k'}m_{c,k'}$. Due to this, in general the partial derivatives for the off-diagonal elements of the Jacobian (25) are non-zero.

The admissible values for the stepsizes which grant convergence in mean would be obtained from $\|J_T(\mathbf{m})\|_p < 1$ in the set of margins of interest. In the case of the spectral norm, or 2-norm, the condition for convergence is written as

$$\|\mathbf{I}_K - \mu_p\boldsymbol{\Upsilon}\|_2 = \max_i \sigma_i(\mathbf{I}_K - \mu_p\boldsymbol{\Upsilon}) < 1 \tag{27}$$

where $\sigma_i(\mathbf{X})$ denotes the $i$th singular value of $\mathbf{X}$.

For illustration purposes, let us resort to assumption (A2), and consider the fully symmetric case for which all channels are statistically identical. From (24), only two different values need to be considered for the entries in (25), namely, the diagonal values $\Upsilon_d \triangleq \Upsilon_{1,1} = \cdots = \Upsilon_{K,K}$, and the off-diagonal values $\Upsilon_x = \Upsilon_{1,2} = \Upsilon_{2,1} = \cdots = \Upsilon_{K,K-1}$. In such a case, the eigenvalues of $\boldsymbol{\Upsilon}$ will determine the convergence, since

$$\max_i \sigma_i(\mathbf{I}_K - \mu_p\boldsymbol{\Upsilon}) = \max_i |1 - \mu_p\lambda_i(\boldsymbol{\Upsilon})| \tag{28}$$

with $\lambda_i(\boldsymbol{\Upsilon})$ the $i$th eigenvalue of $\boldsymbol{\Upsilon}$. Given the highly structured matrix $\boldsymbol{\Upsilon}$ for this case, it is easy to see that its eigenvalues are equal to $\Upsilon_d + (K-1)\Upsilon_x$ and $\Upsilon_d - \Upsilon_x$, the latter with multiplicity $K-1$. With this, the spectral norm in (28) reads as $\max_i \sigma_i(\mathbf{I}_K - \mu_p\boldsymbol{\Upsilon}) =$

$$\begin{cases} |1-\mu_p(\Upsilon_d - \Upsilon_x)|, & \mu_p \leq \frac{1}{\Upsilon_d - \Upsilon_x + K\Upsilon_x/2} \\ |1-\mu_p(\Upsilon_d + (K-1)\Upsilon_x)|, & \mu_p \geq \frac{1}{\Upsilon_d - \Upsilon_x + K\Upsilon_x/2}. \end{cases} \tag{29}$$

From this, we obtain the range of stepsizes for which the spectral norm is upper bounded by 1, given by $|1-\mu_p(\Upsilon_d - \Upsilon_x) - \mu_p K\Upsilon_x| < 1$ or, equivalently,

$$0 < \mu_p < \frac{2}{\Upsilon_d + (K-1)\Upsilon_x}. \tag{30}$$

If we particularize for the single-user case, $K = 1$, then we obtain $0 < \mu_p < 2/\Upsilon_d$, which is the well-known upper bound for the average convergence of OLLA [47]. This would also apply if $K$ users operate in the absence of a common message, as there would be no inter-user dependency and $\Upsilon_x = 0$. In other words, the analysis shows how the common message decoding in RS has some negative impact on the LA convergence.

### B. Mobile Rayleigh channel - periodical CSIT

The analysis can be extended for cases where the subsequences $f(\tilde{\mathbf{H}}[b]|\hat{\mathbf{H}}[b-\ell])$ and $f(\tilde{\mathbf{H}}[b+L_F]|\hat{\mathbf{H}}[b+L_F-\ell])$

[7]Please note that the error probability grows with margins.

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787



are "periodically" i.i.d., even if the CSIT errors have different distributions for different values of $\ell$. This is the case of the Markov renewal process described in Section II-B, that gives rise to a spatially uncorrelated Rayleigh flat fading channel, which is estimated once every $L_F$ frames. Given the periodic update of the precoders in Fig. 2, the expression (21) is now a time-variant transformation, which requires a different approach. Thus, we will be interested in analyzing the average periodic evolution of the margins subsequence $\mathbf{s}[c] \triangleq \mathbb{E}\left[\mathbf{m}[cL_F]\right], c = 0, 1, \ldots$ Note that this periodic mapping is time-invariant if the statistical description of the true channel $[\mathbf{H}[cL_F], \ldots, \mathbf{H}[(c+1)L_F - 1]$ and the CSIT error $[\tilde{\mathbf{H}}[cL_F], \ldots, \tilde{\mathbf{H}}[(c+1)L_F - 1]$ are independent of the index frame $c$. However, the corresponding mapping updating the average margins from one block to the next within a frame would evolve with the relative position $\ell$ within the frame, since the statistics of the CSIT error $\tilde{\mathbf{H}}[cL_F + \ell]$ change with $\ell$ in general. Following the same reasoning as before, if the mapping is contractive, it can be readily seen that at the fixed point such that $\mathbf{s}^*[c+1] = \mathbf{s}^*[c]$, then $0 = -\mu_p \sum_{\ell=1}^{L_F} \overline{\overline{P}}_{out}[cL_F+\ell] + \mu_p \epsilon L_F$, from which the mean outage probability within each frame converges to the target value $\epsilon$ for all users. Note that the average error probability can take different values according to the relative position within the frame, even after convergence. In fact, it displays a periodical behavior with periodicity $L_F$ in general, as illustrated later by the numerical results. It can be shown that the Jacobian of $\mathbf{s}[c+1]$ would be of the same form as (25), but with its terms defined now as $\Upsilon_{k,k'} = \frac{\partial \sum_{\ell=0}^{L_F-1} \overline{P}_{out,k}[cL_F+\ell]}{\partial s_{k'}[c]}$. Thus, a similar discussion of convergence would ensue.

**Corollary 1.** *If the mapping* $\mathbf{s}[c+1] = T(\mathbf{s}[c])$*, result of sampling both sides of* (21) *once per frame, is contractive, then the mean error probabilities averaged within the different CSIT frames converge to the desired value:*

$$\lim_{c\to\infty} \frac{1}{L_F} \sum_{\ell=0}^{L_F-1} \overline{\overline{P}}_{out}[cL_F + \ell] = \epsilon \mathbf{1}. \tag{31}$$

### C. Sum-rate maximization vs. $\frac{\mu_c}{\mu_p}$

Note that the derived conditions for convergence apply on the stepsize for the private margins $\mu_p$ only. The role of the relative value of the ratio $\frac{\mu_c}{\mu_p}$ is twofold: first, it modifies the value of the Jacobian terms $\Upsilon_{k,k'}$ through (24) and (26). In addition, the ratio $\frac{\mu_c}{\mu_p}$ will have an impact on the asymptotic margins. As an illustrative case, let us consider the trajectories in Fig. 4, which depict the evolution for a single time realization of the spectral efficiency with respect to the outage region and the optimal sum-rate pairs, for a channel with constant mean $\hat{\mathbf{H}}[b]$, and independent errors across blocks according to model (2).

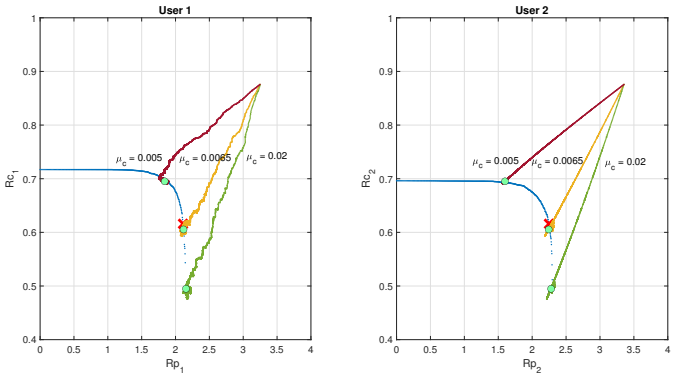



Fig. 4: Trajectory of selected rates with respect to the individual outage regions of two users. $\hat{\mathbf{h}}_1 = [1, 1]^H/\sqrt{2}$, $\hat{\mathbf{h}}_2 = [1, e^{j\pi/2}]^H/\sqrt{2}$. $\epsilon = 0.05$, SNR = 18 dB, CSIT error variance $\sigma_e^2 = \text{SNR}^{-1}$. The OLLA scheme uses the same stepsizes for both users, $\mu_p = 0.01$. All margins are initially set to 0 dB. The red cross denotes the maximum sum-rate point, whereas the green circles correspond to the rates after convergence of the corresponding LA scheme.

As mentioned above, we should not expect to handle a genie-aided statistical model of the channel, which grants us a relation between the CSIT and the transmit rates for a given outage probability, otherwise we would be able to apply well-defined pre-computed margins, thus avoiding the need to learn from the empirical decoding at the receivers. Nevertheless, the insights obtained from the discussion of this example can be applied for practical LA. Convergence takes place in all three cases in Fig. 4 to the prescribed probability of outage. However, the maximum sum-rate operation point is only achieved with the right $\mu_c/\mu_p$ value, which requires full knowledge of the CSIT model; otherwise, the stationary point will drift away from the optimal one. RS precoders have been designed with WMMSE, by assuming the knowledge of the CSIT statistical model, although the conclusions would hold regardless of the scheme used to obtain the precoders; what would change is the outage region and the asymptotic selected rates.

To explore the role of the ratio $\frac{\mu_c}{\mu_p}$ further, let us consider the spectral efficiency which can be achieved for different $\mu_c/\mu_p$ values, depicted for a channel with constant mean in Figure 5, with the RS precoder computed again with the WMMSE method. There is an optimum ratio of step-sizes $\mu_c/\mu_p$, in the case of the figure around 0.5. In other words, the convergence point depends on the relative evolution speed of both private and common margins for a given user, and the contribution of each to the BLER target can be optimized. In fact, it can be readily seen from Definition 2 that the outage probability is bounded as

$$\max\{\epsilon_c, \epsilon_p\} < \epsilon < 2\max\{\epsilon_c, \epsilon_p\}, \tag{32}$$

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787



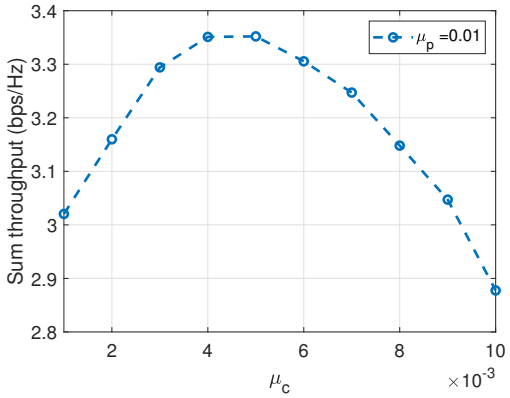



Fig. 5: Sum throughput for a two user case, the channel mean is invariant with time: $\hat{\mathbf{h}}_1 = [1,1]^H/\sqrt{2}$, $\hat{\mathbf{h}}_2 = [1, e^{j\pi/4}]^H/\sqrt{2}$, $\epsilon = 0.05$, SNR = 15 dB, CSIT error variance $\sigma_e^2 = \text{SNR}^{-1}$. The OLLA scheme uses the same stepsizes for both users.

with $\epsilon_c$ and $\epsilon_p$ the outage probabilities of common and private messages, respectively. Among all different combinations of $\epsilon_c$ and $\epsilon_p$ giving rise to the same BLER $\epsilon$, a metric such as the sum rate can be maximized. This problem is equivalent to optimizing the applied margins, as stated next.

**Proposition 1.** *If the nominal margins $\{m^\star_{c,k}, m^\star_{p,k}\}$ are to be achieved for a given outage probability $\epsilon$, the corresponding stepsizes of the OLLA algorithm in* (17) *should satisfy, for user $k$, the following:*

$$\frac{\mu_c}{\mu_p} = \frac{m^\star_{c,k}}{m^\star_{p,k}} \tag{33}$$

*for initial margins $m_{p,k}[0] = m_{c,k}[0] = 0$ dB, $k = 1, \ldots, K$.*

The proof readily follows from (18).

## VI. Results and Discussion

We have tested the embedding of link adaptation along with RS in two scenarios, fixed and mobile. RS precoders were designed with robustness constraints (RS-WMMSE), or by assuming perfect CSIT (RS-ZF). For benchmarking purposes, ZF precoding (no common messages and no SIC) is also tested at some instances. In all cases, the corresponding precoders were designed to maximize the sum-rate, and the state-of-the-art OLLA algorithm was used for link adaptation purposes, to enforce an outage probability $\epsilon$ of $0.05$ or $0.1$. We assumed a perfectly known channel at the receiver, with ideal SIC, and error-free feedback channel for the ACK/NACK reporting, which is performed on a block basis. Unless otherwise stated, CSIT is not perfect, with an error model which is only partially known, and may be time variant. Throughput is computed as $\frac{1}{B}\sum_{b=1}^{B} SE[b]$, with

$$SE[b] = \sum_{k=1}^{K}(1 - E_k[b])\left(R_{p,k}[b] + \frac{1}{K}R_c[b]\right).$$

$B$ is the number of blocks used for averaging, also used for BLER estimation purposes, and the rates $R_c[b], R_{p,k}[b]$ are defined in (16).

### A. Static Rice channel

First, we consider a channel with invariant mean, chosen as in [2], with line-of-sight CSIT and Gaussian uncertainty: $\hat{\mathbf{h}}_1 = 1/\sqrt{2}[1,1]^H$ and $\hat{\mathbf{h}}_2 = 1/\sqrt{2}[1, e^{j2\pi/5}]$. The CSIT error is described by $[\tilde{\mathbf{H}}[b]]_{m,n} \sim \mathcal{CN}(0, \text{SNR}^{-\alpha}), 0 \leq m, n \leq 2$. It is usual to operate with $0 \leq \alpha \leq 1$ [3], with $\alpha = 0$ used for finite precision CSIT, and $\alpha = 1$ corresponding to perfect CSIT in the DoF sense. Stepsizes were chosen to maximize the sum throughput at SNR = 15 dB, and $\alpha_{csit} = 1$ (WMMSE case). We want to analyze the impact of miss-modeling the error when using a robust design method such as WMMSE, and also the performance of RS-ZF, which does not take into account the CSIT errors. Fig. 6 shows the throughput for both methods in the presence of inaccurate statistical description of the channel uncertainty. As upper bound the genie-aided performance is included, which assumes the perfect knowledge of the achievable rate at each time instant and for all messages. As expected, WMMSE outperforms RS-ZF thanks to its robust design to accommodate the channel uncertainty. Interestingly, more robust designs (lower $\alpha_{csit}$) yield a higher throughput when applying WMMSE, which is designed to maximize the average sum-rate over a wider uncertainty region, which in turn helps to increase the outage rate.

### B. Mobile Rayleigh channel

The model for imperfect CSIT in this second case corresponds to a time-varying channel, as described in [21]:

$$\mathbf{h}_k[b] = \rho \mathbf{h}_k[b-1] + \sqrt{1-\rho^2}\mathbf{g}_k[b] \tag{34}$$

with the entries of $\mathbf{h}_k[b]$ and $\mathbf{g}_k[b]$ i.i.d. following $\mathcal{CN}(0,1)$, and $\rho$ the time correlation coefficient according to the Jakes's model, $\rho = J_0(2\pi f_D T)$, for Doppler frequency $f_D$ and time spacing $T$. This simple recursion is a good touchstone to evaluate LA performance in the presence of outdated CSIT under mobility [50]. For simplicity, we assume that all users channels are i.i.d, such that the correlation coefficient $\rho$ is the same. The precoders are recomputed once per frame, when a new channel reading is reported to the transmitter, as in (2), with the CSIT updating period $L_F = 100$. Again, the

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787



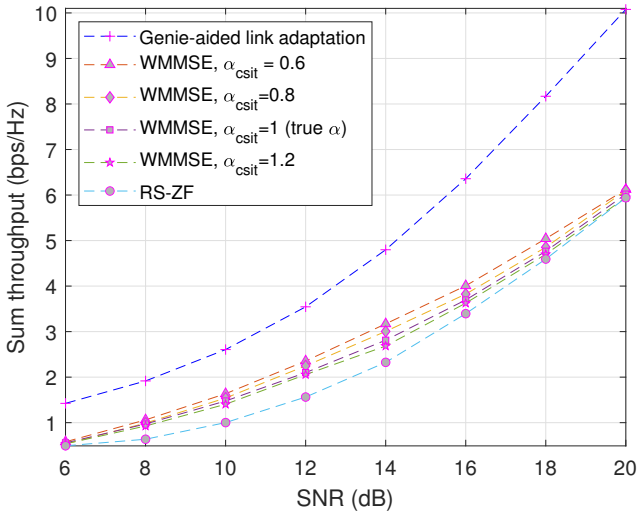
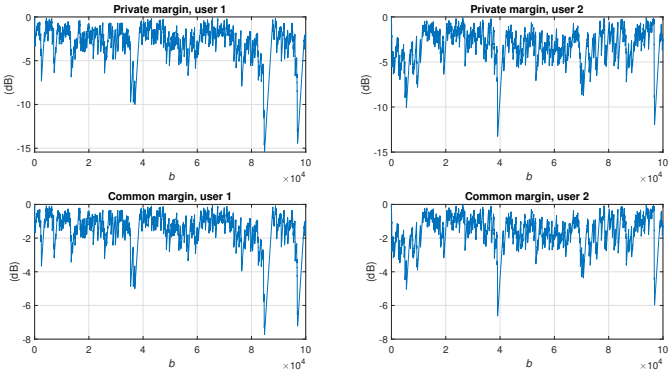



Fig. 6: Sum throughput, static Rice channel, two user case. $\hat{\mathbf{h}}_1 = [1,1]^H/\sqrt{2}$, $\hat{\mathbf{h}}_2 = [1, e^{j2\pi/5}]^H/\sqrt{2}$, $\epsilon = 0.05$. The number of simulated blocks is $10^5$, $B = 10^4$, one realization. Same stepsizes for both users, optimized for SNR=15 dB, and $\alpha_{csit} = 1$ in the WMMSE case. The ideal (genie-aided) link adaptation makes use of the RS-ZF precoder.

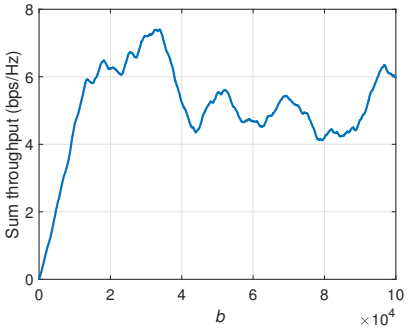
role of LA is to keep the error within prescribed bounds, to complement the design of robust RS precoders. Results are shown for RS-ZF precoders and one realization in Fig. 7, with the following settings: SNR = 15 dB, $\rho = 0.9999$, $\sigma_e^2 = 0$, $N = 2$, $K = 2$, $\epsilon = 0.05$, $\mu_p = 0.1$, $\mu_c = 0.05$ (same stepsizes for both users), ACK/NACK delay $b_0 = 1$. No convergence can be expected to fixed back-off margins for a single realization, since the genie-aided values, provided all the information is available, change with every CSIT reporting and the corresponding precoder update. Note the fluctuation of the BLER of both users around the target $\epsilon = 0.05$ in Fig. 7c. In order to gauge the improvement of RS-ZF over ZF linear precoding, well established in the literature, and how it extends to practical settings with unknown achievable rates due to inaccurate tracking of the channel at the transmitter, we plot the average throughput after 1000 realizations in Fig. 8, this time with $N = K = 4$, $\epsilon = 0.1$, $\mu_p = 1$, $\mu_c = 0.5$ (same stepsizes for both users), and averaging period $B = 100$. Both ideal ($\sigma_e^2 = 0$) and non-ideal ($\sigma_e^2 = \text{SNR}^{-1}$) CSIT are showcased. Again, the total spectral efficiency of the genie-aided case, which does not suffer from outage since the maximum achievable rate is assumed to be known for all messages, is included for benchmarking purposes. The remaining parameters are the same as before. The following observations can be made:

(i) RS outperforms ZF precoding, for both perfect and non-perfect CSIT, also in combination with LA. This extends previous results in the literature which applied to ergodic settings.

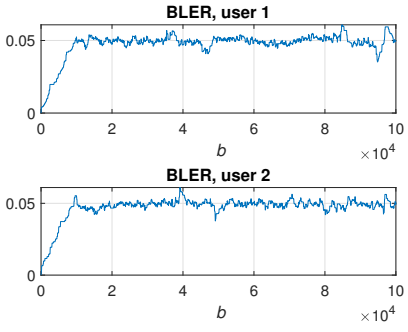
(ii) The OLLA algorithm converges in mean to a periodic pattern, as anticipated in the discussion in Section V. At the beginning of each frame, when precoders are recomputed based on a new (ideal) CSIT, margins increase in average (back-off decreases). As the time passes, the CSIT gets more outdated, so that the margins need to compensate eventually for the channel errors and decrease on average till the next frame starts over.




(a)




(b)



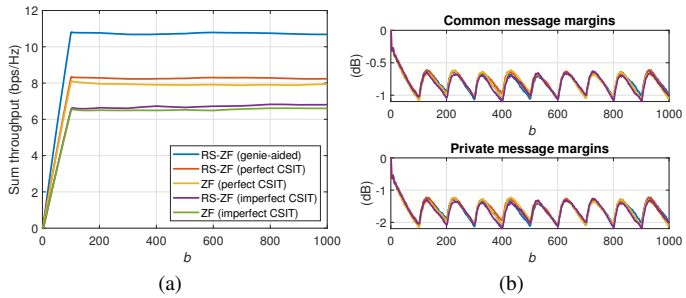

(c)

Fig. 7: Mobile Rayleigh channel, SNR = 15 dB, $N = K = 2$, a single time realization, RS-ZF precoder. (a) Margins. (b) Throughput. (c) BLER.

Finally, Fig. 9 compares the average sum throughput performance of RS-ZF and ZF achieved at convergence, for a wide

This article has been accepted for publication in IEEE Journal on Selected Areas in Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSAC.2023.3240787






(a) (b)

Fig. 8: Mobile Rayleigh channel, SNR=15 dB, 1000 realizations, $N = K = 4$, $L_F = 100$. (a) Time evolution of sum throughput under perfect CSIT ($\sigma_e^2 = 0$) and imperfect CSIT ($\sigma_e^2 = \text{SNR}^{-1}$), with the genie-aided version as reference. (b) Average margins for the four users, perfect CSIT.

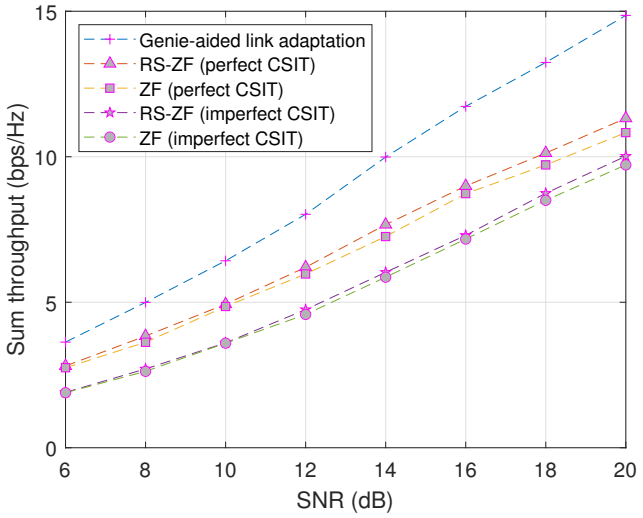



Fig. 9: Sum throughput, mobile Rayleigh channel, $10^4$ realizations, $N = K = 4$, $L_F = 100$, $\epsilon = 0.1$, $\rho = 0.9999$, $\mu_c = 0.5$, $\mu_p = 1$. Perfect CSIT ($\sigma_e^2 = 0$) and imperfect CSIT ($\sigma_e^2 = \text{SNR}^{-1}$). The ideal (genie-aided) link adaptation makes use of the RS-ZF precoder.

range of SNR values. With the ideal LA as an upper reference, it can be seen that RS-ZF outperforms ZF for all SNR values, both with perfect and imperfect CSIT, as anticipated in Fig. 8a for SNR=15 dB. The gap with respect to the genie-aided performance is caused by the outdated CSIT, only reported once every $L_F$ blocks, and the CSIT error in the case of imperfect CSIT.

## VII. Conclusions

This work has considered the embedding of link adaptation into the operation of a MIMO broadcast channel which makes use of Rate Splitting precoding. Regardless of the selected design method for RS precoders, the transmitter has to face channels with uncertain achievable rates, due to fading and imperfect CSIT. Thus, link adaptation is a key instrument at the physical layer to guarantee an error performance when the CSIT is imperfect. It also provides meta-robustness when the statistical models of the channel dynamics and CSIT errors are not accurate enough, which precludes the off-line computation of safety margins which are not overly conservative. From the convergence analysis, it was concluded that the adaptation speed of the back-off private and common margins determines the stationary point; additionally, the periodic update of CSIT and recalculation of precoders gives rise to a periodic behavior of the margins to enforce the prescribed average error probability. Further efforts can be put into the practical application of the exposed ideas, considering other link adaptation schemes, possibly based on Bayesian reasoning. And although SIC detection was assumed throughout the paper, joint decoding could be implemented while keeping the same signaling feedback, and improving the outage rate regions. The trade-offs among frame spacing, CSI quality and link adaptation schemes can be also further explored.

## VIII. Acknowledgements

The authors wish to thank Dr. Yijie (Lina) Mao, from ShanghaiTech University, and Prof. Bruno Clerckx, from Imperial College London, for sharing the code to compute the WMMSE RS precoders, which is available at http://www.ee.ic.ac.uk/bruno.clerckx/Research.html.

This work has been supported by the Agencia Estatal de Investigación (Spain) and the European Regional Development Fund (ERDF) through the projects RODIN (PID2019-105717RB-C21) and CASTRO (PID2021-122483OA-I00).

## References

[1] B. Clerckx, Y. Mao, R. Schober, E. A. Jorswieck, D. J. Love, J. Yuan, L. Hanzo, G. Y. Li, E. G. Larsson, and G. Caire, "Is NOMA Efficient in Multi-Antenna Networks? A Critical Look at Next Generation Multiple Access Techniques," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1310–1343, 2021.

[2] 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 Communications Letters*, vol. 9, no. 3, pp. 349–353, 2020.

[3] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-Splitting Multiple Access: Fundamentals, Survey, and Future Research Trends," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2022.

[4] 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, 2016.

[5] E. Piovano and B. Clerckx, "Optimal DoF region of the $K$-user MISO BC with partial CSIT," *IEEE Communications Letters*, vol. 21, no. 11, pp. 2368–2371, 2017.

[6] A. G. Davoodi, B. Yuan, and S. A. Jafar, "GDoF region of the MISO BC: Bridging the gap between finite precision and perfect CSIT," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7208–7217, 2018.

[7] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5690–5705, 2014.

[8] L. Yin and B. Clerckx, "Rate-Splitting Multiple Access for Multigroup Multicast and Multibeam Satellite Systems," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 976–990, 2021.

[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] A. A. Ahmad, Y. Mao, A. Sezgin, and B. Clerckx, "Rate Splitting Multiple Access in C-RAN: A Scalable and Robust Design," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5727–5743, 2021.

[11] G. Lu, L. Li, H. Tian, and F. Qian, "MMSE-Based Precoding for Rate Splitting Systems With Finite Feedback," *IEEE Communications Letters*, vol. 22, no. 3, pp. 642–645, 2018.

[12] A. Z. Yalcin, M. Yuksel, and B. Clerckx, "Rate splitting for multi-group multicasting with a common message," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 12 281–12 285, 2020.

[13] H. Chen, D. Mi, T. Wang, Z. Chu, Y. Xu, D. He, and P. Xiao, "Rate-splitting for multicarrier multigroup multicast: Precoder design and error performance," *IEEE Transactions on Broadcasting*, vol. 67, no. 3, pp. 619–630, 2021.

[14] C. Mosquera, N. Noels, T. Ramírez, M. Caus, and A. Pastore, "Space-Time Rate Splitting for the MISO BC With Magnitude CSIT," *IEEE Transactions on Communications*, vol. 69, no. 7, pp. 4417–4432, 2021.

[15] J. Zhang, J. Zhang, Y. Zhou, H. Ji, J. Sun, and N. Al-Dhahir, "Energy and spectral efficiency tradeoff via rate splitting and common beamforming coordination in multicell networks," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7719–7731, 2020.

[16] Z. Yang, M. Chen, W. Saad, and M. Shikh-Bahaei, "Optimization of Rate Allocation and Power Control for Rate Splitting Multiple Access (RSMA)," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5988–6002, 2021.

[17] X. Su, L. Li, H. Yin, and P. Zhang, "Robust Power- and Rate-Splitting-Based Transceiver Design in $K$ -User MISO SWIPT Interference Channel Under Imperfect CSIT," *IEEE Communications Letters*, vol. 23, no. 3, pp. 514–517, 2019.

[18] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, 2008.

[19] B. Matthiesen, Y. Mao, A. Dekorsy, P. Popovski, and B. Clerckx, "Globally optimal spectrum-and energy-efficient beamforming for rate splitting multiple access," *arXiv preprint arXiv:2204.00273*, 2022.

[20] A. R. Flores and R. C. De Lamare, "Robust and Adaptive Power Allocation Techniques for Rate Splitting based MU-MIMO systems," *IEEE Transactions on Communications*, pp. 1–1, 2022.

[21] O. Dizdar, Y. Mao, and B. Clerckx, "Rate-splitting multiple access to mitigate the curse of mobility in (massive) MIMO networks," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6765–6780, 2021.

[22] J. Choi, "Power allocation for max-sum rate and max-min rate proportional fairness in NOMA," *IEEE Communications Letters*, vol. 20, no. 10, pp. 2055–2058, 2016.

[23] P. Wu and N. Jindal, "Coding versus ARQ in fading channels: How reliable should the PHY be?" *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3363–3374, 2011.

[24] L. Li, N. Jindal, and A. Goldsmith, "Outage capacities and optimal power allocation for fading multiple-access channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1326–1347, 2005.

[25] P. Piantanida, G. Matz, and P. Duhamel, "Outage behavior of discrete memoryless channels under channel estimation errors," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4221–4239, 2009.

[26] R. Fritzsche, P. Rost, and G. P. Fettweis, "Robust rate adaptation and proportional fair scheduling with imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4417–4427, 2015.

[27] G. Liu, E. Bala, L. Li, and L. J. Cimini, "Outage Analysis for MISO Zero-Forcing Precoding With Outdated CSI," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9152–9156, 2020.

[28] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2013, pp. 611–615.

[29] A. Mishra, Y. Mao, O. Dizdar, and B. Clerckx, "Rate-Splitting Multiple Access for Downlink Multiuser MIMO: Precoder Optimization and PHY-Layer Design," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 874–890, 2022.

[30] 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," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, 2020, pp. 1–6.

[31] T. Cui, F. Lu, V. Sethuraman, A. Goteti, S. P. Rao, and P. Subrahmanya, "Throughput optimization in high speed downlink packet access (HS-DPA)," *IEEE Transactions on Wireless Communications*, vol. 10, no. 2, pp. 474–483, 2011.

[32] F. Blánquez-Casado, G. Gómez, M. d. C. Aguayo-Torres, and J. T. Entrambasaguas, "eOLLA: an enhanced outer loop link adaptation for cellular networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, pp. 1–16, 2016.

[33] V. Saxena, H. Tullberg, and J. Jaldén, "Reinforcement learning for efficient and tuning-free link adaptation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 768–780, 2022.

[34] S. Yun, C. Caramanis, and R. W. Heath Jr, "Distributed link adaptation for multicast traffic in MIMO-OFDM systems," *Physical Communication*, vol. 4, no. 4, pp. 286–295, 2011.

[35] A. Rico-Alvariño and R. W. Heath, "Learning-based adaptive transmission for limited feedback multiuser MIMO-OFDM," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3806–3820, 2014.

[36] C. Kaulich, M. Joham, and W. Utschick, "Rate-Splitting for the Weighted Sum Rate Maximization under Minimum Rate Constraints in the MIMO BC," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2021, pp. 1–6.

[37] A. R. Flores, R. C. De Lamare, and B. Clerckx, "Tomlinson-Harashima Precoded Rate-Splitting With Stream Combiners for MU-MIMO Systems," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3833–3845, 2021.

[38] A. Lozano and N. Jindal, "Are yesterday's information-theoretic fading models and performance metrics adequate for the analysis of today's wireless systems?" *IEEE Communications Magazine*, vol. 50, no. 11, pp. 210–217, 2012.

[39] J. Wang and D. P. Palomar, "Worst-Case Robust MIMO Transmission With Imperfect Channel Knowledge," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3086–3100, 2009.

[40] A. Krishnamoorthy and R. Schober, "Downlink MIMO-RSMA with Successive Null-Space Precoding," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2022.

[41] M. Bashar, Y. Lejosne, D. Slock, and Y. Yuan-Wu, "MIMO broadcast channels with Gaussian CSIT and application to location based CSIT," in *2014 Information Theory and Applications Workshop (ITA)*. IEEE, 2014, pp. 1–7.

[42] Y. Mao, B. Clerckx, and V. O. Li, "Rate-splitting multiple access for downlink communication systems: bridging, generalizing, and outperforming SDMA and NOMA," *EURASIP J Wirel Commun Netw*, vol. 2018, no. 1, p. 133, 2018.

[43] Z. Li, C. Ye, Y. Cui, S. Yang, and S. Shamai, "Rate splitting for multi-antenna downlink: Precoder design and practical implementation," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1910–1924, 2020.

[44] J. Lindblom, E. Karipidis, and E. G. Larsson, "Achievable outage rate regions for the MISO interference channel," *IEEE Wireless Communications Letters*, vol. 2, no. 4, pp. 439–442, 2013.

[45] S. Diggavi and T. Cover, "The worst additive noise under a covariance constraint," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 3072–3081, 2001.

[46] J. Park, J. Choi, N. Lee, W. Shin, and H. Vincent Poor, "Rate-Splitting Multiple Access for Downlink MIMO: A Generalized Power Iteration Approach," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2022.

[47] A. Rico-Alvariño, J. Arnau, and C. Mosquera, "Link adaptation in mobile satellite links: schemes for different degrees of CSI knowledge," *International Journal of Satellite Communications and Networking*, vol. 34, no. 5, pp. 679–694, 2016.

[48] R. A. Delgado, K. Lau, R. Middleton, R. S. Karlsson, T. Wigren, and Y. Sun, "Fast convergence outer loop link adaptation with infrequent updates in steady state," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2017, pp. 1–5.

[49] S. Banach, "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales," *Fund. math*, vol. 3, no. 1, pp. 133–181, 1922.

[50] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 844–855, 1999.