

A Receiver-centric Approach to Interference Management: Fairness and Outage Optimization

Mehdi Ashraphijuo, Ali Tajer, *Senior Member, IEEE*, Chen Gong, and Xiaodong Wang, *Fellow, IEEE*

Abstract—Effective interference management in the multiuser interference channel necessitates that the users form their transmission and interference management decisions in coordination, and adapt them to the state of the channel. Establishing such coordination, often facilitated through information exchange, is prohibitive in fast-varying channels, especially when the network size grows. This paper focuses on the multiuser Gaussian interference channel and offers a receiver-centric approach to interference management. In this approach, the transmitters deploy rate-splitting and superposition coding to generate their messages according to independent Gaussian codebooks. The receivers can freely decode any arbitrary set of interfering messages along with their designated messages in any desired joint or ordered fashion, and treat the rest of the interferers as Gaussian noise. The proposed receiver-centric interference management approach is applied to two class of problems (outage optimization and fairness-constrained rate allocation), and constructive proofs are provided to establish the following properties for the proposed approach: 1) the optimal set of codebooks to be decoded by each receiver is a local decision made by each receiver based on its local channel state information (CSI); 2) the *globally* optimal transmission rates are related to *locally* optimal rates computed by the receivers based on their local information, which implies that the transmitters do not require explicit knowledge of the CSI and can determine their rates via limited feedback from the receivers; and 3) obtaining the optimal interference management strategy at each receiver has controlled complexity.

Index Terms—Distributed interference management, fairness, group decoding, outage, superposition coding.

I. INTRODUCTION

A. Motivation

INVESTIGATING multiuser interference channels has a rich history, spanning from Shannon's seminal work in [1] and the best known achievable rate-region derived in [2], to the their limits in the high signal-to-noise-ratio (SNR) regimes [3] and [4]. Motivated by the ambitious spectral

Manuscript received June 23, 2015; revised March 3, 2016; accepted July 9, 2016. Date of publication July 27, 2016; date of current version September 13, 2016. This work was supported by the U. S. National Science Foundation under Grant ECCS-1455228 and Grant CCF-1526215.

M. Ashraphijuo and X. Wang are with the Electrical Engineering Department, Columbia University, New York, NY 10027 USA (e-mail: mehdi@ee.columbia.edu; wangx@ee.columbia.edu).

A. Tajer is with the Electrical, Computer, and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: tajer@ecse.rpi.edu).

C. Gong is with the Department of Electronic and Information Science, University of Science and Technology of China, Hefei 230026, China (e-mail: cgong821@ustc.edu.cn).

Communicated by S.-Y. Chung, Associate Editor for Shannon Theory.

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

Digital Object Identifier 10.1109/TIT.2016.2594219

efficiency goals and universal frequency reuse in cellular networks, wireless networks are growing to be increasingly interference-limited. Therefore, understanding the fundamental limits of operation in the multiuser interference channel, which captures the essence of interference-limited networks, is of paramount importance for designing wireless networks. The effectiveness of interference management strategies in the multiuser interference channel strongly depends on the availability of the channel state information (CSI) at the transmitter as well as the receiver sides. While the receivers can acquire the CSI through training sessions, acquiring it by the transmitters, often facilitated via feedback, consumes excessive communication resources, especially when the size of network grows.

Driven by the challenges associated with acquiring the CSI at the transmitters (CSIT), and based on the availability extent of the CSIT there exist different research directions on interference management. Specifically, the existing directions can be grouped under those assuming *perfect and instantaneous* CSIT, *perfect and delayed* CSIT, and *imperfect* CSIT. When the CSIT is assumed to be available perfectly and instantaneously, in one significant approach the capacity region of a two-user Gaussian interference channel (GIC) within 1-bit accuracy is approximated in [5]. In another important approach, the notion of interference alignment, introduced in [3] and [4] characterizes the achievable degrees of freedom in the multiuser interference channel. The capacity region of the K -user GIC when the users are constrained to use point-to-point random codes is established in [6] and [7]. This result is generalized to the case that encoding is restricted to random code ensembles with time sharing and superposition coding in [8]. While providing imperative insights about the fundamental limits of multiuser interference-limited networks, these recent developments on interference channels strongly hinge on the availability of the CSIT. In a second direction, motivated by the fact that acquiring the perfect CSIT is not always feasible in real time, and pioneered by the study in [9], there has been extensive recent research on analyzing the effects of delayed perfect CSIT on interference management (c.f. [10]–[13]). Finally, in the third direction the effects of partial (imperfect) CSIT on interference management and the achievable rate are studied (c.f. [14]–[18]).

Meeting such CSIT requirements consumes communication resources and can become prohibitive in large networks. This motivates investigating and analyzing interference management when the CSI is only available at the receivers. Despite

its significance, the studies on the interference channels when the transmitters do not have explicit access to the CSI is not well-investigated [19] and [20].

B. Contributions

This paper focuses on the K -user GIC and assumes no explicit availability of the CSIT (except for limited functions of the CSI) and also assumes availability of only *local* CSI at the receivers (CSIR). That is, each receiver knows only the states of its incoming channels from different transmitters but not those of the channels to other receivers. The transmitters employ rate splitting, in which the message of each transmitter is generated by superimposing messages drawn from multiple independent Gaussian codebooks, where the power levels allocated to the codebooks are pre-specified. The attendant interference management scheme used by the receivers under this setting consists in each receiver dynamically partitioning the interfering codebooks into two disjoint sets, where one set is decoded along with the intended codebooks while the other set is discarded as Gaussian noise.

Under this setting, this paper proposes a receiver-centric approach to interference management and analyzes the notion of *constrained partial group decoding (CPGD)* which aims to perform rate optimization over the rate region achievable under the aforementioned rate splitting and interference management strategies such that some notions of *fairness* in rate allocation and *controlled complexity* in decoding the codebooks are satisfied. The important observation is that having only local CSIR and no explicit CSIT, in conjunction with reporting some functions of the CSI from the receivers to the transmitters suffice to identify the optimal¹ codebook rates and the associated interference management strategies. We provide a constructive proof to demonstrate this observation, which consists in the following main components to establish the optimality guarantees for rate allocation, while recognizing the fairness and complexity constraints.

- 1) The message of each transmitter is generated by superimposing messages generated by M random codebooks. M remains a constant as a design parameter and does not depend on the CSI. Allocation of the power to different codebooks in each transmitter is pre-specified. While the focus of the paper is on equal power allocation, all the analyses can be readily generalized to accommodate pre-specified non-equal power allocation. Hence, the structure of the encoding strategy adopted is independent of the CSI. On the other hand, the decoding dynamics is not fixed and dynamically adapts to the CSI. Specifically, the receivers can freely select any arbitrary set of interferers to decode in any desired joint or ordered fashion.
- 2) Based on the definition of the rate optimization problem over the network we define an individual *local* problem corresponding to each transmitter-receiver pair. Such local problems are formulated and solved by their respective receivers and based on the local CSIR available at the receivers. The outcome of the process

at each receiver includes *local* interference management decisions (i.e., what codebooks should be decoded by each receiver) and a set of rates computed for *all* transmitters.

- 3) The globally optimal transmission rates are related to the *locally* optimal ones computed by the receivers through a known function. This implies that for identifying the optimal rates for the transmitters, the information required about the channel states is entirely embedded in the local rates computed by the receivers. Hence, the receivers refrain from feeding back the CSI to the transmitters, and instead only report their locally computed rates, the information content of which is substantially smaller than that of the full CSI.
- 4) The interference management strategies obtained *locally* turn out to be also the *globally* optimal strategies. The important implication is that no coordination among the users is necessary for identifying the set of codebooks to be decoded by each receiver, and that remains a local decision.
- 5) Also, each receiver employs a successive *group* decoding strategy in which at each stage the number of codebooks that the receiver affords to decode jointly via maximum likelihood (ML) decoding is controlled not to exceed a pre-specified level. Such decoding complexity constraint can be relaxed by selecting the threshold to be sufficiently large.

The main properties of this receiver-centric approach to interference management are summarized below.

- 1) **Distributed implementation:** Being amenable to distributed implementation follows from the observed properties that the optimal codebook rates are related to the locally computed ones through simple functions, and that the optimal decoding strategy at each receiver remains a local decision.
- 2) **Search complexity:** The search complexity for identifying the best set of codebooks to be decoded at each receiver grows exponentially with the number of users and the number of codebooks per user. By deploying a successive decoding approach and leveraging the matroid structure of the achievable rate regions, we show that this search complexity can be reduced to polynomial complexity. Similar search complexity problem when there exists only one codebook per transmitter is studied for point-to-point communication under the presence of undesired interferers in [21] and for multiuser interference channel in [22].
- 3) **Decoding complexity:** The complexity of jointly decoding the optimal set of codebooks at each receiver grows with the number of users and codebooks per user. The proposed approach can control the number of codebooks can be decoded via ML decoding, based on the decoding complexity that each receiver affords. This is an extension of the conventional successive interference cancellation (SIC) decoding such that at each decoding stage, instead of one codebook, a subset of codebooks are jointly decoded. The significance of such decoders is that they span a broad spectrum of decoding strategies,

¹We remark that optimality refers to the maximum achievable rates under the specified rate splitting, superposition coding, power allocation, and interference management strategies.

ranging from the low-complexity SIC decoder to the high-complexity ML decoder. The decoding complexity aspect of CPGD can be considered as an extension of the successive group decoder (SGD) initially introduced in [23] for the uncoded Gaussian code-division multiple-access (CDMA) channel, which is studied extensively over the fading multiple-access channel (MAC) in [25]. The practical merits of the proposed approach when multiple codebooks per user are deployed and practical rates are designed for each codebook are studied in [25]. The remainder of this paper is organized as follows. In Section II we present the system model for the GIC and describe constrained partial group decoders. In Section III outage minimization, symmetric fair rate allocation, and max-min fair rate allocation problems are formalized, which are treated in sections IV, V, and VI, respectively. Simulation results and concluding remarks are provided in sections VII and VIII, respectively. The detailed proofs are relegated to the appendices.

II. PRELIMINARIES

A. Channel Model

Consider a slow-fading fully connected K -user GIC consisting of K transmitters each intending to communicate with one designated receiver. Denote the fading channel from the j^{th} transmitter to the i^{th} receiver by $h_{i,j} \in \mathbb{C}$. By denoting the input of the j^{th} transmitter to the channel during the n^{th} symbol interval by $X_j[n]$, the output of the channel at the i^{th} receiver is

$$Y_i[n] = \sum_{j=1}^K h_{i,j} X_j[n] + Z_i[n], \quad \forall n \in \mathbb{N}, \quad (1)$$

where $Z_i[n]$ accounts for the noise at the i^{th} receiver during the n^{th} transmission interval. Channel inputs are statistically independent and channel noise values $Z_i[n]$ are statistically independent of the channel inputs and temporally uncorrelated with distribution $\mathcal{N}_{\mathbb{C}}(0, 1)$.² Furthermore, channel inputs are subject to the power constraints $\mathbb{E}[|X_i[n]|^2] \leq P_i$ for all $i \in \{1, \dots, K\}$. We define the channel vector \mathbf{h}_i as the vector of incoming channels to the i^{th} receiver, i.e.,

$$\mathbf{h}_i \triangleq [h_{i,1}, \dots, h_{i,K}], \quad \forall i \in \{1, \dots, K\}, \quad (2)$$

and define the channel matrix \mathbf{H} by concatenating vectors $\{\mathbf{h}_i\}_{i=1}^K$ as its rows, i.e., $\mathbf{H} \triangleq [\mathbf{h}_1^T, \dots, \mathbf{h}_K^T]^T$.

B. Rate Splitting

The ultimate goal of each receiver is to effectively decode the messages transmitted by its respective transmitter while suppressing the disruptive effects of the interfering messages transmitted by other transmitters. For effective interference management each receiver *may* or *may not* benefit from decoding the messages of the interferers depending on their strengths. The optimal interference management strategies in

² $\mathcal{N}_{\mathbb{C}}(a, b)$ denotes symmetric complex Gaussian with mean a and variance b .

the extreme cases of weak and strong interference are well-studied, where in the former case the best strategy is to treat the interferers as Gaussian noise [26] and [27], whereas the optimal strategy in the latter is to fully decode the interferers [2], [28], and [29]. In the more general settings in which the strength of the message and interference signals can vary arbitrarily, however, a universally optimal interference management strategy is unknown. Nevertheless, one effective approach inspired by the Han-Koabayashi interference management scheme is to provide the receivers with the freedom to dynamically decide which interfering messages to decode along with their desired ones. Hence, each receiver must be able to identify the *optimal*³ set of the interfering signals to be decoded. Moreover, since a receiver is not ultimately interested in the messages of the interfering transmitters, it is beneficial to allow the receivers to decode the interfering transmitters only *partially*. For this purpose, the message of each transmitter is split into multiple smaller messages, each drawn from an *independent* Gaussian codebook.

Let us define M as the number of codebooks used by each transmitter and denote the set of codebooks of the i^{th} transmitter by $\tilde{\mathcal{X}}_i \triangleq \{\mathcal{X}_{i,1}, \dots, \mathcal{X}_{i,M}\}$ where $\mathcal{X}_{i,m}$ is a Gaussian codebook with rate $R_{i,m}$ per channel use. Increasing the number of codebooks M allows the receivers to approximate the best power allocation with a higher accuracy, where in the asymptote of large values of M (i.e., $M \rightarrow \infty$) the setting becomes equivalent to performing optimal power allocation across codebooks, which is viable at the expense of higher complexity in identifying the optimal set of decodable codebooks at each receiver. Hence, increasing M provides the receivers with the freedom to decide what fraction of the interfering messages to decode and to treat what fraction as noise. In order to formalize this, define $X_{i,m}[n]$ as the unit-power input from codebook $\mathcal{X}_{i,m}$ to the channel during the n^{th} channel use. Therefore, we have

$$X_i[n] = \sqrt{\frac{P_i}{M}} \sum_{m=1}^M X_{i,m}[n], \quad (3)$$

which satisfies the power constraint of P_i for all $i \in \{1, \dots, K\}$. It is noteworthy that the power of different codebooks in each transmitter is pre-specified to be P_i/M . While the focus of the paper is on equal power allocation to all transmitters for simplicity, all the analyses can be readily generalized to accommodate pre-specified non-equal power allocation. Subsequently, the rate of the i^{th} transmitter is

$$R_i = \sum_{m=1}^M R_{i,m}, \quad \text{for all } i \in \{1, \dots, K\}. \quad (4)$$

We define the rate matrix \mathbf{R} such that $[\mathbf{R}]_{i,m} = R_{i,m}$ and use the pair (i, m) to denote the index of codebook $\mathcal{X}_{i,m}$ and rate $R_{i,m}$. Furthermore, we define the set \mathcal{K} as the set of such index pairs, i.e.,

$$\mathcal{K} \triangleq \{(i, m) \mid i \in \{1, \dots, K\} \text{ and } m \in \{1, \dots, M\}\}. \quad (5)$$

³Definition of the *optimal* subset of the transmitted signals to be decoded by each receiver depends on the objective sought to be optimized for the interference channel and will be defined rigorously in the subsequent sections.

For any set $\mathcal{U} \subseteq \mathcal{K}$, we define the $K \times M$ indicator matrix $\mathbb{1}_{\mathcal{U}}$ such that $\forall (i, m) \in \mathcal{K}$

$$[\mathbb{1}_{\mathcal{U}}]_{i,m} = \begin{cases} 1 & \text{if } (i, m) \in \mathcal{U} \\ 0 & \text{if } (i, m) \notin \mathcal{U}. \end{cases}$$

Finally, we represent the Hadamard product of matrices \mathcal{U} and \mathcal{V} by $\mathcal{U} \circ \mathcal{V}$, and denote the ℓ_1 -norm of matrix \mathbf{R} by

$$\|\mathbf{R}\| = \sum_{i=1}^K \sum_{m=1}^M [\mathbf{R}]_{i,m}.$$

C. Constrained Partial Group Decoding

Motivated by the premise that a receiver *might* benefit from decoding a subset of the messages of the interferers, each receiver partitions the set of all codebooks $\{\bar{\mathcal{X}}_1 \cup \bar{\mathcal{X}}_2 \cup \dots \cup \bar{\mathcal{X}}_K\}$ into a set of *decodable* codebooks that it will decode, and a set of *non-decodable* codebooks that it will suppress by treating them as noise. Optimal partitioning of the codebooks hinges on the utility function that one seeks to optimize for the network, which will be defined in Section III, as well as the instantaneous realization of the channel. As a result, designing the optimal transmission and receiving strategies requires addressing the following issues.

- **Coordination:** When the CSI is revealed to all the transmitters and receivers, the rate allocation problem and the attendant interference management strategies can be solved by all parties. Due to the communication overhead incurred by revealing the CSI, the important question is how much coordination (information exchange) across the network is necessary in order to determine the optimal rate allocation at the transmitter sites and interference management at the receiver sites.
- **Search complexity:** Equipping the receivers with the freedom to dynamically identify the optimal decodable set of codebooks is viable at expense of two types of complexities. One is the *search* complexity, that is the complexity due to dynamically identifying the optimal decodable set for each receiver. This complexity can be readily shown to be growing *exponentially* with the number of users and the number of codebooks per user.
- **Decoding complexity:** The other type of complexity is pertinent to jointly decoding multiple codebooks by each receiver.

The notion of CPGD aims to address the aforementioned issues in a unified framework. The remainder of this section is focused on providing a few definitions, which are instrumental to circumventing both types of complexities and also formalizing the structure of coordination.

In order to control search complexity the CPGD, which is a generalization of successive group decoders originally proposed for multiple access channels [23] and [30], provides a successive search approach to replace the exhaustive search, and will be shown to break the exponential complexity in MK to polynomial complexity in MK . In order to formalize the search process, define $\mathcal{Q}_i \subseteq \mathcal{K}$ as the set of the indices of the decodable codebooks by the i^{th} receiver (i.e., the set of codebooks to be decoded by the i^{th} receiver).

Subsequently, $\mathcal{K} \setminus \mathcal{Q}_i$ contains the indices of the codebooks to be treated as noise by the i^{th} receiver. The CPGD at the i^{th} receiver further partitions \mathcal{Q}_i into p_i disjoint and ordered sets, for some $p_i \in \mathbb{N}$, as

$$\bar{\mathcal{Q}}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}, \quad (6)$$

and successively decodes all the codebooks included in \mathcal{Q}_i . More specifically, corresponding to the ordered partitions $\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}$ the i^{th} receiver performs a p_i -stage successive decoding procedure, in which in stage $k \in \{1, \dots, p_i\}$ it jointly decodes the messages in \mathcal{Q}_i^k while treating those in $\mathcal{K} \setminus \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{k-1}\}$ as Gaussian noise.

In order to control the decoding complexity, we define μ_i as number of codebooks that the i^{th} receiver affords to decode *jointly*. Hence, by defining \mathcal{K} as the set of all non-empty subsets of \mathcal{K} , corresponding to each receiver we define the bounding function

$$f_i : \mathcal{K} \rightarrow \{0, 1\},$$

such that for any $\mathcal{U} \in \mathcal{K}$, $f_i(\mathcal{U}) = 1$ indicates that the i^{th} receiver affords to jointly decode the codebooks in \mathcal{U} , whereas $f_i(\mathcal{U}) = 0$ means that these codebooks cannot be decoded jointly. In order to capture the decoding complexity that each receiver can afford, we set the bounding function f_i as the set size control function

$$\forall \mathcal{U} \in \mathcal{K} : f_i(\mathcal{U}) = \begin{cases} 1 & \text{if } |\mathcal{U}| \leq \mu_i \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

for the pre-specified positive integers $\{\mu_1, \dots, \mu_K\}$. It is noteworthy that μ_i 's can increase arbitrarily and this constraint can be lifted by setting $\mu_i = +\infty$. Based on these definitions, we state $\mathcal{Q}_i \triangleq \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ is a *valid* ordered partition of the codebooks if the following conditions are satisfied.

- 1) $f(\mathcal{Q}_i^k) = 1$ for all $k \in \{1, \dots, p_i\}$, i.e., the number of codebooks to be decoded jointly by receiver i at each step does not exceed μ_i .
- 2) $\forall m \in \{1, \dots, M\} : (i, m) \in \mathcal{Q}_i$, i.e., all the codebooks of the i^{th} transmitter will be decoded by the i^{th} receiver.
- 3) $\exists m \in \{1, \dots, M\}$ such that $(i, m) \in \mathcal{Q}_i^{p_i}$ as otherwise, $\mathcal{Q}_i^{p_i}$ can be combined with $\mathcal{K} \setminus \mathcal{Q}_i$ and be treated as noise.

We define $\bar{\mathcal{Q}}_i$ as the set of all valid ordered set of partitions $\bar{\mathcal{Q}}_i$ for the i^{th} receiver.

D. Outage Event

For any two *disjoint* sets $\mathcal{U}, \mathcal{V} \subseteq \mathcal{K}$, and a given channel realization \mathbf{H} we define the achievable rate region $\mathcal{C}_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V})$ as the set of all rate matrices \mathbf{R} that the i^{th} receiver supports when it jointly decodes the codebooks in \mathcal{U} while treating those in \mathcal{V} as Gaussian noise. The achievable rate region $\mathcal{C}_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V})$ can be characterized as

$$\begin{aligned} \mathcal{C}_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}) \\ = \left\{ \mathbf{R} \in \mathbb{R}_+^{K \times M} : \|\mathbf{R} \circ \mathbb{1}_{\mathcal{D}}\| \leq \mathcal{R}_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}), \forall \mathcal{D} \subseteq \mathcal{U} \right\}, \end{aligned} \quad (8)$$

where we have defined

$$\mathcal{R}_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}) \triangleq \log \left(1 + \frac{\sum_{(j,m) \in \mathcal{D}} P_j |h_{i,j}|^2}{M + \sum_{(j,m) \in \mathcal{V}} P_j |h_{i,j}|^2} \right). \quad (9)$$

Also, for given disjoint sets $\mathcal{D}, \mathcal{V} \subseteq \mathcal{K}$, and a given rate matrix \mathbf{R} we define

$$\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}) \triangleq \mathcal{R}_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}) - \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\|, \quad (10)$$

which measures the gap between the boundary of the achievable rate region and the instantaneous sum-rate of the codebooks in \mathcal{D} . The following two properties can be readily verified for function $\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})$:

1) **Chain rule:** For any disjoint sets \mathcal{U}, \mathcal{V} , and set \mathcal{D} :

$$\begin{aligned} \Delta_i(\mathbf{h}_i, \mathcal{U} \cup \mathcal{V}, \mathcal{D}, \mathbf{R}) \\ = \Delta_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V} \cup \mathcal{D}, \mathbf{R}) + \Delta_i(\mathbf{h}_i, \mathcal{V}, \mathcal{D}, \mathbf{R}). \end{aligned} \quad (11)$$

2) **Subset:** For any set $\mathcal{D} \subseteq \mathcal{B}$

$$\Delta_i(\mathbf{h}_i, \mathcal{U}, \mathcal{D}, \mathbf{R}) \geq \Delta_i(\mathbf{h}_i, \mathcal{U}, \mathcal{B}, \mathbf{R}). \quad (12)$$

Based on the definition in (10), next for any two disjoint sets $\mathcal{U}, \mathcal{V} \subseteq \mathcal{K}$ we define

$$\varepsilon_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \triangleq \min_{\mathcal{D} \subseteq \mathcal{U}, \mathcal{D} \neq \emptyset} \Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}), \quad (13)$$

which identifies the subset of \mathcal{U} that has the smallest sum-rate gap with the one given by the boundary of the achievable rate region. It can be readily verified that

$$\mathbf{R} \in \mathcal{C}_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}) \Leftrightarrow \varepsilon_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \geq 0. \quad (14)$$

Hence, for a given valid ordered partition $\mathcal{Q}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ the i^{th} receiver is in *outage* if $\exists k \in \{1, \dots, p_i\}$ such that

$$\mathbf{R} \notin \mathcal{C}(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^k \mathcal{Q}_i^j). \quad (15)$$

An outage event can be equivalently characterized by further defining $\bar{\varepsilon}_i(\mathbf{h}_i, \mathcal{Q}_i, \mathbf{R})$ for a given set of ordered partitions as

$$\bar{\varepsilon}_i(\mathbf{h}_i, \mathcal{Q}_i, \mathbf{R}) \triangleq \min_{k \in \{1, \dots, p_i\}} \varepsilon_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^k \mathcal{Q}_i^j, \mathbf{R}), \quad (16)$$

which indicates that the i^{th} receiver is in outage if and only if $\bar{\varepsilon}_i(\mathbf{h}_i, \mathcal{Q}_i, \mathbf{R}) < 0$. When the i^{th} receiver is not in outage, it can deploy a p_i -stage successive procedure by using the maximum likelihood rule. The contribution of the decoded codebooks to the received signal is removed and the received signal is consequently updated to

$$Y_i[n] \leftarrow Y_i[n] - \sum_{(j,m) \in \mathcal{Q}_i^k} \sqrt{\frac{P_j}{M}} h_{i,j} X_{j,m}[n] \quad (17)$$

for further processing in the subsequent stages.

III. PROBLEM FORMULATIONS

A. Fixed Rate Mode

The specific structure of the partial group decoders, and in particular delineating partitions $\{\mathcal{Q}_i, \mathcal{K} \setminus \mathcal{Q}_i\}$ and $\bar{\mathcal{Q}}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ for each receiver $i \in \{1, \dots, K\}$ depends on the utility function to be optimized for the network. In this paper we consider two classes of problems. First we assume that the rates and power of all codebooks are fixed and invariant to channel variations. The objective in this class of problems reduces to identifying the best CPGD at each receiver such that the likelihood of an outage event for the network is minimized.

Specifically, for a given channel realization \mathbf{H} and given rates \mathbf{R} , the objective is to identify a valid partition $\mathcal{Q}_i^* = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ that maximizes $\bar{\varepsilon}_i(\mathbf{h}_i, \mathcal{Q}_i, \mathbf{R})$. Hence, the fixed-rate outage minimization problem can be cast as

$$\mathcal{Q}_i^* = \arg \max_{\mathcal{Q}_i \in \bar{\mathcal{Q}}_i} \bar{\varepsilon}_i(\mathbf{h}_i, \mathcal{Q}_i, \mathbf{R}), \quad (18)$$

where $\bar{\varepsilon}_i(\mathbf{h}_i, \mathcal{Q}_i, \mathbf{R})$ is defined in (16). Having a combinatorial nature, a crude exhaustive search can identify the the optimal partitioning \mathcal{Q}_i^* for the i^{th} receiver. In Section IV we discuss the complexity of such exhaustive search and how CPGDs can be leveraged to reduce the complexity.

B. Rate Adaptation Mode

In the second class the objective is to concurrently design the rates of the codebooks and the CPGDs to be used by different receivers. The rates in this class are dynamically updated based on channel variations such that some notion of fairness in rate allocation to different users is satisfied. In order to formalize this assume that at some time instance T_n , for $n \in \mathbb{N}$, the K -user interference channel is in some state, which we denote by S_n . This state is influenced by the amount of resources (power and spectrum) available to the transmitters as well as the fading status of the wireless channels at time instance T_n . In order to account for channel variations, we use $h_{i,j}^n$ to denote the state of channel $h_{i,j}$ at time instance T_n . Also assume that at time instance T_n the users are operating at some *decodable* rate matrix \mathbf{R}^n , i.e., the i^{th} transmitter-receiver link sustains the rates $\mathbf{R}_i^n \triangleq [[\mathbf{R}^n]_{i,1}, \dots, [\mathbf{R}^n]_{i,M}]$. The channel remains in the same state for some duration, and due to some variation in the available resources or channel statics, at time instance T_{n+1} changes to the state S_{n+1} . Due to such change the rates \mathbf{R}^n may remain decodable if $\forall i$, the corresponding codebooks of \mathbf{R}_i^n are all decodable by the i^{th} receiver, and will not be decodable if $\exists i$ such that the corresponding codebooks of \mathbf{R}_i^n are not decodable by the i^{th} receiver. Rate adaptation seeks updating \mathbf{R}^n and obtaining a new set of *decodable* rates \mathbf{R}^{n+1} such that the following three conditions are satisfied.

- 1) Some notion of fairness is maintained, i.e., no user sacrifices its rate in favor of the other users.
- 2) \mathbf{R}^{n+1} is optimal in the sense that it cannot be increased without violating the fairness constraints.
- 3) We assume that each receiver has perfect and instantaneous access to the *local* CSI, i.e. \mathbf{h}_i is known *only* to the i^{th} receiver, and the transmitters have no CSI. Hence, rate updates are accomplished in a distributed

way such that each user updates its rate only based on its *local* information about network dynamics and with some limited information exchange with other users.

1) *Symmetric Fairness*: In the symmetric fairness model, we consider updating the rates of different codebooks based on a pre-determined set of priorities for them. More specifically, we are interested in finding the largest possible $x \in \mathbb{R}$ such that after updating the rate matrix as $\mathbf{R}^{n+1} = \mathbf{R}^n + x \cdot \mathbf{T}$ for a given $K \times M$ matrix \mathbf{T} with $[\mathbf{T}]_{i,m} \triangleq t_{i,j} \geq 0, \forall i \in \{1, \dots, K\}, \forall m \in \{1, \dots, M\}$, the updated rate matrix \mathbf{R}^{n+1} remains decodable. Matrix \mathbf{T} can capture different notions of fairness. For instance, setting $\mathbf{T} = \mathbf{1}_{K \times M}$ provides all codebooks with identical rate changes, or setting $\mathbf{T} = \mathbf{R}^n$ leads to scaling all the rates identically. We call x the *symmetric rate adaptation factor* and its optimal value is determined by solving

$$\beta^*(n) = \begin{cases} \max & x \\ \text{s.t.} & \mathbf{R}^{n+1} = \mathbf{R}^n + x \cdot \mathbf{T} \text{ is decodable} \\ & \mathcal{Q}_i \in \mathcal{Q}_i \quad \forall i \in \{1, \dots, K\}. \end{cases} \quad (19)$$

When the rate matrix \mathbf{R}^n is decodable under the new network, state $\beta^*(n)$ is expected to be non-negative and the rates can possibly be incremented beyond \mathbf{R}^n . On the other hand, when the rate matrix \mathbf{R}^n is not decodable after the change in the network, state $\beta^*(n)$ will be negative and the rates should be decremented in order to avoid outage. The objective of symmetric-fair rate allocation problem is to determine the *unique* rate matrix \mathbf{R}^{n+1} that satisfies the constraints in (19).

2) *Max-Min Fairness*: In this model, by denoting the rate variation of the m^{th} codebook of the i^{th} user by $[\mathbf{r}]_{i,m} \triangleq r_{i,m}$, our objective is to maximize $\min_{\{i,m\}} \frac{r_{i,m}}{t_{i,m}}$ such that after updating $\mathbf{R}^{n+1} = \mathbf{R}^n + \mathbf{r}$ for some given \mathbf{T} , \mathbf{R}^{n+1} is decodable. By defining the *max-min rate adaptation factor* y , the max-min rate allocation problem yields

$$\gamma^*(n) = \begin{cases} \max & \min_{\{i,m\}} \frac{r_{i,m}}{t_{i,m}} \\ \text{s.t.} & \mathbf{R}^{n+1} = \mathbf{R}^n + \mathbf{r} \text{ is decodable} \\ & \mathcal{Q}_i \in \mathcal{Q}_i \quad \forall i \in \{1, \dots, K\}. \end{cases} \quad (20)$$

Similar to the symmetric fairness model, if \mathbf{R}^n is decodable, then $\gamma^*(n)$ is non-negative and otherwise it is negative. One major difference between the max-min rate adaptation factor $\gamma^*(n)$ and its symmetric counterpart $\beta^*(n)$ is that unlike the symmetric case, which admits a unique rate matrix \mathbf{R}^{n+1} , there exists potentially a set of distinct rate matrices \mathbf{R}^{n+1} that satisfy the constraints given in (20) by having equal smallest normalized rate increment value $r_{i,m}/t_{i,m}$. We define the max-min rate allocation strategy as the one resulting in a *pareto-optimal* solution of (20).

Remark 1: It is noteworthy that the rate allocation problems formalized in (19) and (20) aim to optimize rate allocations on the *codebooks* level. These problems subsume the problems of rate allocation on *users* level, which correspond to the settings in which the elements within each row of \mathbf{T} are identical. This can be readily proved by contradiction. In the symmetric fairness problem, for instance, if user-level rate allocation problem leads to distinct optimal rates $\mathbf{R}^n(i, m)$ and $\mathbf{R}^n(i, m')$ for codebooks $\mathcal{X}_{i,m}$ and $\mathcal{X}_{i,m'}$, respectively, then

by symmetry in the setting (both codebooks experience the same channels), rates $\mathbf{R}^n(i, m')$ and $\mathbf{R}^n(i, m)$ are also optimal choices for codebooks $\mathcal{X}_{i,m}$ and $\mathcal{X}_{i,m'}$, respectively. By a time-sharing argument, as a result, $\frac{1}{2}(\mathbf{R}^n(i, m) + \mathbf{R}^n(i, m'))$ and $\frac{1}{2}(\mathbf{R}^n(i, m) + \mathbf{R}^n(i, m'))$ are also optimal choices for codebooks $\mathcal{X}_{i,m}$ and $\mathcal{X}_{i,m'}$, which contradicts the assumption that optimal rates being distinct.

IV. FIXED RATE MODE

In the fixed rate mode the rates of all the codebooks and transmitters are pre-specified and these rates might fall inside or outside the achievable rate region. The task of each receiver is to identify a decodable set of codebooks, decoding which leads to minimizing the likelihood of an outage event. Since there is no rate optimization involved, the tasks of the receivers reduces to only identifying their optimal decodable sets of codebooks.

In the first step, in order to motivate the need for a computationally efficient procedure for partitioning the codebooks at each receiver, we compute the complexity of the exhaustive search process. By computing the cardinality of \mathcal{Q}_i that is the set of all valid ordered partitions of to be deployed by the i^{th} receiver we can assess the scaling rate of $|\mathcal{Q}_i|$ in terms of M and K .

Lemma 1: The size of the set of valid partitions \mathcal{Q}_i scales exponentially with the number of users K or the number of codebooks per user M . ■

Proof: See Appendix A.

Next, for given channel and rate matrices \mathbf{H} and \mathbf{R} , we propose a successive procedure which can determine the optimal partitioning at each receiver without requiring any coordination among the receivers, and has *polynomial* complexity in KM . Specifically this proposed procedure (Algorithm 1) has at most $(K-1)M$ stages each of which bearing polynomial complexity of an order not exceeding KM .

In order to establish the tools for evaluating whether there exists a valid partitioning \mathcal{Q}_i such that each outage can be avoided at the i^{th} receiver, for any two *disjoint* subsets of the messages $\mathcal{U}, \mathcal{V} \subseteq \mathcal{K}$, and for any receiver $i \in \{1, \dots, K\}$ we define the outage indicator variable

$$\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \triangleq \begin{cases} \max & x \\ \text{s.t.} & \mathbf{R} + x \cdot \mathbf{1}_{K \times M} \in \mathcal{C}_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}), \end{cases} \quad (21)$$

where $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ is the maximum rate change that we can achieve by jointly decoding the codebooks listed in \mathcal{U} and treating those listed in \mathcal{V} as noise. This immediately establishes the following necessary and sufficient condition:

$$\begin{aligned} \alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) < 0 \Leftrightarrow & \text{decoding } \mathcal{U} \text{ while} \\ & \text{treating } \mathcal{V} \text{ as noise leads to outage at receiver } i. \end{aligned} \quad (22)$$

Therefore, for a valid partitioning $\mathcal{Q}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ of the codebooks at the i^{th} receiver, at stage $k \in \{1, \dots, p_i\}$ the decoder experiences outage if and only if

$$\alpha_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}) < 0. \quad (23)$$

Hence, given \mathcal{Q}_i , outage at the i^{th} receiver can be avoided if and only if

$$\min_{k \in \{1, \dots, p_i\}} \{\alpha_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R})\} \geq 0. \quad (24)$$

Hence, in order to avoid outage at the i^{th} receiver, it suffices to identify one valid partition $\mathcal{Q}_i \in \mathcal{Q}_i$ corresponding to which the condition in (24) holds. As a result, a necessary and sufficient condition for avoiding outage at the i^{th} receiver is $\alpha_i^* \geq 0$, where we have defined

$$\alpha_i^* \triangleq \max_{\mathcal{Q}_i \in \mathcal{Q}_i} \left\{ \min_k \{\alpha_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R})\} \right\}, \quad (25)$$

with the corresponding optimal partitioning \mathcal{Q}_i^* given by

$$\mathcal{Q}_i^* \triangleq \arg \max_{\mathcal{Q}_i \in \mathcal{Q}_i} \left\{ \min_k \{\alpha_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R})\} \right\}. \quad (26)$$

Verifying the condition $\alpha_i^* \geq 0$ involves computing $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ and searching for an optimal partition \mathcal{Q}_i^* . Throughout the rest of this section we provide systematic approaches for performing these tasks with controlled complexity. Specifically we show that:

- 1) Function $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ can be computed in polynomial time.
- 2) The optimal \mathcal{Q}_i can be identified via a successive approach that consists of at most MK successions each with polynomial complexity in MK .

A. Computing $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$

Computing $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ has a central role in determining whether the i^{th} receiver is experiencing outage. We provide the following lemma, which establishes a connection between $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ and $\Delta_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ defined in (10), and provides a systematic approach for solving $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$. The result of this lemma can be proved by using the properties of *polymatroids* and using the techniques developed in [31].

Lemma 2: For a submodular function $f_{\mathcal{V}} : 2^{\mathcal{U}} \rightarrow \mathbb{R}_+$ on the ground set $\mathcal{U} \subseteq \mathcal{K}$ define the polymatroid

$$P_{f_{\mathcal{V}}}(\mathcal{U}) \triangleq \left\{ \mathbf{R} \in \mathbb{R}_+^{K \times M} : \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\| \leq f_{\mathcal{V}}(\mathcal{D}), \forall \mathcal{D} \subseteq \mathcal{U} \right\}.$$

For function

$$\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) = \begin{cases} \max & x \\ \text{s.t.} & \mathbf{R} + x \cdot \mathbf{1}_{K \times M} \in P_{f_{\mathcal{V}}}(\mathcal{U}) \end{cases}, \quad (27)$$

we have

$$\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{f_{\mathcal{V}}(\mathcal{D}) - \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\|}{|\mathcal{D}|}.$$

Proof: See Appendix B. \blacksquare

Since the region $\mathcal{R}_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V})$ can be readily shown to be a polymatroid of the form $P_{f_{\mathcal{V}}}(A)$ defined in Lemma 2 with the submodular function $f_{\mathcal{V}}(\mathcal{D}) = \mathcal{R}_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V})$, by applying Lemma 2, the solution to (21) can be found as

$$\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}. \quad (28)$$

Therefore, this lemma establishes that solving (21) reduces to solving a combinatorial optimization of a submodular function over a polymatroid, which is feasible in a *polynomial* time [21] and [32].

B. Outage Minimization

As shown in Lemma 1, a naive exhaustive search for optimal partition \mathcal{Q}_i has a complexity that scales exponentially with KM . We, instead, propose an efficient successive procedure, which finds \mathcal{Q}_i^* with a complexity that is polynomial in KM . We start by briefly explaining the steps involved in this procedure and then provide the details followed by their optimality properties. The procedure starts by including all the codebooks as candidates to be decoded by the i^{th} receiver, which forms a multiple access channel between the collection of all codebooks and the i^{th} receiver. The capacity region of this multiple access channel is characterized by $(2^{|\mathcal{K}|} - 1)$ inequalities of the form

$$\forall \mathcal{D} \subseteq \mathcal{K}, \quad \mathcal{D} \neq \emptyset : \quad \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\| \leq \mathcal{R}_k(\mathbf{h}_i, \mathcal{D}, \emptyset). \quad (29)$$

Hence, the value of function $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ for $\mathcal{U} = \mathcal{K}$ and $\mathcal{V} = \emptyset$, by its definition, yields the smallest per-codebook gap between the two sides of inequality in (29), i.e.,

$$\alpha_i^1 \triangleq \alpha_i(\mathbf{h}_i, \mathcal{K}, \emptyset, \mathbf{R}) = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{K}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \emptyset, \mathbf{R})}{|\mathcal{D}|}, \quad (30)$$

and the bottleneck set of codebooks corresponding to this smallest gap is

$$\mathcal{V}_1 \triangleq \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{K}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \emptyset, \mathbf{R})}{|\mathcal{D}|}. \quad (31)$$

We will prove that $\alpha_i^1 \geq 0$ is a sufficient condition for the i^{th} user not being in outage. On the other hand if $\alpha_i^1 < 0$ we update sets \mathcal{V} and \mathcal{U} based on the structure and cardinality of set \mathcal{V}_1 and parameter α_i^1 and compute a second parameter α_i^2 , which itself can serve as an alternative sufficient condition for verifying whether the i^{th} user is in outage. The procedure continues by successively refining \mathcal{V} and \mathcal{U} , and generating the sequence of parameters $\{\alpha_i^1, \alpha_i^2, \alpha_i^3, \dots\}$, which collectively establish *necessary* and *sufficient* conditions for checking whether the i^{th} receiver is in outage. The detailed steps of this successive procedure are presented in Algorithm 1, and the optimality properties are demonstrated in the following lemmas and theorem.

The following lemma demonstrates the evolution of the sequence $\{\alpha_i^k\}_{k=1}^q$ corresponding to the i^{th} receiver.

Lemma 3: For the sequence $\{\alpha_i^k\}_{k=1}^q$ computed by Algorithm 1 we have $\alpha_i^1 \leq \alpha_i^2 \leq \dots \leq \alpha_i^q$. \blacksquare

Proof: See Appendix C.

By leveraging Lemma 3, we provide the following lemma, which is instrumental to proving the optimality of Algorithm 1.

Lemma 4: If $\alpha_i^* > \alpha_i^q$, then for the sets \mathcal{V}_i^q and \mathcal{G}_i yielded by Algorithm 1 we have

$$\mathcal{V}_i^q \subseteq \mathcal{Q}_i \subseteq \mathcal{G}_i.$$

Proof: See Appendix D. \blacksquare

Lemmas 3 and 4 establish the optimality of Algorithm 1 in determining whether the i^{th} receiver is in outage.

Theorem 1: Algorithm 1 determines whether the i^{th} receiver is in outage for given channel and rate matrices \mathbf{H} and \mathbf{R} . Also, a necessary and sufficient condition for avoiding outage at the i^{th} receiver is that $\alpha_i^q \geq 0$.

Algorithm 1 Determining Outage

```

1: Initialize  $\mathcal{G} = \mathcal{K}$ ,  $\mathcal{V} = \emptyset$ , and  $k = 1$ .
2: repeat
3:   Find  $\alpha_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
4:   Find  $\mathcal{V}_i^k = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
5:   if  $(i, m) \notin \mathcal{V}_i^k$  for all  $m \in \{1, \dots, M\}$ 
6:      $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{V} \leftarrow \mathcal{V} \cup \mathcal{V}_i^k$  and  $k \leftarrow k + 1$ .
7:   end if
8: until  $\exists m \in \{1, \dots, M\}$  such that  $(i, m) \in \mathcal{V}_i^k$ .
9: set  $p = k$ 
10: repeat
11:   Find  $\alpha_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
12:   Find  $\mathcal{V}_i^k = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
13:   if  $(i, m) \notin \mathcal{V}_i^k$  for all  $m \in \{1, \dots, M\}$ 
14:      $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{V} \leftarrow \mathcal{V} \cup \mathcal{V}_i^k$  and  $k \leftarrow k + 1$ .
15:   end if
16: until  $\exists m \in \{1, \dots, M\}$  such that  $(i, m) \in \mathcal{V}_i^k$ .
17:  $q = k$ 
18: if  $\alpha_i^q < 0$  declare outage
19: if  $\alpha_i^q \geq 0$  declare decodability
20: set  $\mathcal{G}_i = \mathcal{G}$ 

```

Proof: See Appendix E. ■

It is noteworthy that Algorithm 1 not only determines whether a receiver is in outage, but also measures the gap value α_i^* between the actual sum-rate based on \mathbf{R} and the achievable one based on the achievable rate region. The latter, while necessary for proving the optimality of Algorithm 1, is redundant when the objective is to determine whether the i^{th} receiver is in outage. Motivated by this observation, we modify Algorithm 1 and propose Algorithm 2 as an alternative with a lower computational complexity, which focuses on determining whether the i^{th} receiver is in outage. Algorithm 2 is designed by taking into account that

$$\alpha_i^1 \leq \alpha_i^2 \leq \dots \leq \alpha_i^q = \alpha_i^*. \quad (32)$$

This relationship implies throughout the successions in Algorithm 1 when we encounter $\alpha_i^k \geq 0$ for some $k \in \{1, \dots, q-1\}$, it is a *sufficient* condition for concluding that $\alpha_i^* \geq 0$, and hence the i^{th} receiver is not in outage. Therefore, in Algorithm 2 the process terminates as soon as a non-negative element in the sequence $\{\alpha_i^k\}_{k=0}^{q-1}$ is identified.

V. SYMMETRIC FAIRNESS

The goal of fairness-constrained rate allocation is to dynamically adjust and update the rates of different users (and codebooks) when the network undergoes a change in the channel states, while in parallel some fairness in rate adjustments is ensured. This section focuses on rate adjustment under symmetric fairness as formalized in (19), where the goal is to update the rate matrix \mathbf{R}^n to $\mathbf{R}^{n+1} = \mathbf{R}^n + x \cdot \mathbf{T}$ for a given matrix \mathbf{T} such that the updated rate matrix \mathbf{R}^{n+1} remains decodable. In this section, we offer the procedure for identifying the optimal rate updates as well as the attendant optimal interference management strategy at each receiver.

Algorithm 2 Determining Outage

```

1: Initialize  $\mathcal{G} = \mathcal{K}$ ,  $\mathcal{V} = \emptyset$ , and  $k = 1$ .
2: repeat
3:   Find  $\alpha_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
4:   Find  $\mathcal{V}_i^k = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
5:   if  $\alpha_i^k \geq 0$ 
6:     declare decodability
7:     stop
8:   end if
9:   if  $(i, m) \notin \mathcal{V}_i^k$  for all  $m \in \{1, \dots, M\}$ 
10:     $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{V} \leftarrow \mathcal{V} \cup \mathcal{V}_i^k$  and  $k \leftarrow k + 1$ .
11:   end if
12: until  $\exists m \in \{1, \dots, M\}$  such that  $(i, m) \in \mathcal{V}_i^k$ .
13: repeat
14:   Find  $\alpha_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
15:   Find  $\mathcal{V}_i^k = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$ .
16:   if  $\alpha_i^k \geq 0$ 
17:     declare decodability
18:     stop
19:   end if
20:   if  $(i, m) \notin \mathcal{V}_i^k$  for all  $m \in \{1, \dots, M\}$ 
21:      $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{V} \leftarrow \mathcal{V} \cup \mathcal{V}_i^k$  and  $k \leftarrow k + 1$ .
22:   end if
23: until  $\exists m \in \{1, \dots, M\}$  such that  $(i, m) \in \mathcal{V}_i^k$ .
24: declare outage

```

We propose the CPGD procedure for solving the symmetric fair rate adaptation problem formulated in (19). The notable structure of this proposed procedure is that it breaks the rate adaption problem into K *local* problems each specialized for one transmitter-receiver pair. These local problems are solved by the receivers in parallel, after which the receivers perform a round of information exchange, which facilitates providing the optimal solution to all the users. By noting that for achieving a network-wide optimal interference management strategy different users cannot operate autonomously, the important findings of the proposed constructive procedure are the following:

- 1) Identifying the best partitions $\{\mathcal{Q}_1, \dots, \mathcal{Q}_K\}$ are purely local decisions such that determining \mathcal{Q}_i can be carried out by the i^{th} receivers based on its limited information about the network. Recall that we assume that the i^{th} receiver knows only its incoming channels \mathbf{h}_i defined in (2). Hence, there is no coordination necessary among the users for establishing the decoding procedure.
- 2) Each receiver provides a local solution for the rate adaptation factor. The optimal solution is shown to be a function of these local solutions. Hence, for determining the optimal rates at the transmitters the receivers need to feed their local solutions back to transmitters.

A. Local Interference Management

For a given rate matrix \mathbf{R}^n , the fairness constraint embedded in \mathbf{T} , any two *disjoint* subsets of the codebooks $\mathcal{U}, \mathcal{V} \subseteq \mathcal{K}$, and for any receiver $i \in \{1, \dots, K\}$ we define a rate change

factor as

$$\begin{aligned} \beta_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T}) \\ \triangleq \begin{cases} \max & x \\ \text{s.t.} & \mathbf{R}^n + x \cdot \mathbf{T} \in \mathcal{C}_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}), \end{cases} \end{aligned} \quad (33)$$

where $\beta_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T})$ is the maximum rate change that we can achieve by jointly decoding \mathcal{U} and treating \mathcal{V} as noise. Given the definition in (33), the maximum rate change factor that the i^{th} user can afford for a valid partitioning $\bar{\mathcal{Q}}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ of the codebooks at the k^{th} stage, for $k \in \{1, \dots, p_i\}$ is

$$\beta_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T}). \quad (34)$$

Therefore, given partitions $\bar{\mathcal{Q}}_i$, the maximum rate change factor corresponding to which the i^{th} receiver remains decodable throughout all decoding stages is

$$\min_{k \in \{1, \dots, p_i\}} \{\beta_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\}. \quad (35)$$

Finally, an optimal valid partition of the codebooks at the i^{th} receiver can be obtained by maximizing the rate factor change, i.e.,

$$\beta_i^*(n) \triangleq \max_{\bar{\mathcal{Q}}_i \in \mathcal{Q}_i} \left\{ \min_k \{\beta_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\} \right\}, \quad (36)$$

with the corresponding optimal partitioning \mathcal{Q}_i^* given by

$$\mathcal{Q}_i^* \triangleq \arg \max_{\bar{\mathcal{Q}}_i \in \mathcal{Q}_i} \left\{ \min_k \{\beta_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\} \right\}. \quad (37)$$

The remaining part of this section will be devoted to solving (36) with controlled complexity constraints given in (7). Solving (36) involves two levels of complexity. One stems from the complexity of searching for the *best* choice of partition $\bar{\mathcal{Q}}_i$ that solves (36), and the other one is related to the complexity of solving (33) for given sets \mathcal{U} and \mathcal{V} . We discuss each of these levels of complexity and our approaches for mitigating them separately.

By following the same line of arguments as in Section IV, we use the properties of *polymatroids* for solving $\beta_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T})$ for given disjoint sets \mathcal{U} and \mathcal{V} . We provide the following lemma, which is instrumental for solving the optimization problem in (33).

Lemma 5: For a submodular function $f_{\mathcal{V}} : 2^{\mathcal{U}} \rightarrow \mathbb{R}_+$ on the ground set $\mathcal{U} \subseteq \mathcal{K}$ define the polymatroid

$$P_{f_{\mathcal{V}}}(\mathcal{U}) \triangleq \left\{ \mathbf{R} \in \mathbb{R}_+^{K \times M} : \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\| \leq f_{\mathcal{V}}(\mathcal{D}), \forall \mathcal{D} \subseteq \mathcal{U} \right\}.$$

For a given matrix $\mathbf{T} \succeq 0$,

$$\beta_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T}) = \begin{cases} \max & x \\ \text{s.t.} & \mathbf{R}^n + x \cdot \mathbf{T} \in P_{f_{\mathcal{V}}}(\mathcal{U}) \end{cases}$$

is equal to

$$\beta_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T}) = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{f_{\mathcal{V}}(\mathcal{D}) - \|\mathbf{R}^n \circ \mathbf{1}_{\mathcal{D}}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}.$$

Proof: See Appendix F. \blacksquare

In order to apply the above lemma on the problem at hand, note that the region $\mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V})$ can be readily shown to be a polymatroid of the from $P_{f_{\mathcal{V}}}(\mathcal{A})$ defined in Lemma 5 with the submodular function $f_{\mathcal{V}}(\mathcal{D}) = \mathcal{R}_k(\mathbf{h}_k^{n+1}, \mathcal{D}, \mathcal{V})$ given in (9). Hence, by applying Lemma 5, the solution of (33) can be found as

$$\begin{aligned} \beta_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T}) \\ = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{\mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}) - \|\mathbf{R}^n \circ \mathbf{1}_{\mathcal{D}}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}. \end{aligned} \quad (38)$$

Therefore, solving (33) reduces to solving a combinatorial optimization of a submodular function over a polymatroid, which is feasible in a polynomial time [21] and [32].

By using the result of Lemma 5, we start by briefly explaining the steps involved in this procedure, and then provide the details followed by their optimality properties. Without loss of generality, assume that the network state changes are in favor of increasing the rates of the i^{th} user beyond its currently operating rates, i.e., $\beta_i^*(n) \geq 0$. Corresponding to each valid ordered partition $\bar{\mathcal{Q}}_i$, the i^{th} receiver observes a multiple access channel that needs to decode the messages included in $\{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ successively and its noise level at the k^{th} stage is shaped by the interference induced by the messages included in $\mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j$. At the k^{th} stage, the capacity region of the associated multiple access channel corresponding is characterized by $2^{|\mathcal{Q}_i^k|} - 1$ inequalities of the following form. For all non-empty sets $\mathcal{D} \subseteq \mathcal{Q}_i^k$ we have

$$\|\mathbf{R}^{n+1} \circ \mathbf{1}_{\mathcal{D}}\| \leq \mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j). \quad (39)$$

Based on these inequalities, the rate of the messages in \mathcal{D} can be incremented as much as this inequality is not violated. Hence, the highest rate increments occur when

$$\|\mathbf{R}^{n+1} \circ \mathbf{1}_{\mathcal{D}}\| = \mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j),$$

or equivalently

$$\begin{aligned} \|\mathbf{R}^n + x \cdot \mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\| &= \|\mathbf{R}^n \circ \mathbf{1}_{\mathcal{D}}\| + x \cdot \|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\| \\ &= \mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j), \end{aligned} \quad (40)$$

which provides that the maximum rate change factor corresponding to set \mathcal{D} is

$$x = \frac{\mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j) - \|\mathbf{R}^n \circ \mathbf{1}_{\mathcal{D}}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}. \quad (41)$$

Now, among all possible choices for \mathcal{D} , the one that yields the smallest rate change factor x given in (41) will constitute the bottleneck set of messages in the sense that these users determine what the maximum rate change factor is. The procedure that identifies this bottleneck set for the i^{th} receiver is initialized by including *all* messages as candidates for being decoded by the i^{th} receiver, which forms a multiple access channel between the collection of all codebooks and the i^{th} receiver. The capacity region of this multiple access channel is characterized by $2^{|\mathcal{K}|} - 1$ inequalities of the form

$$\forall \mathcal{D} \subseteq \mathcal{K}, \mathcal{D} \neq \emptyset : \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\| \leq \mathcal{R}_i(\mathbf{h}_i, \mathcal{D}, \emptyset). \quad (42)$$

Hence, the value of function $\beta_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ for $\mathcal{U} = \mathcal{K}$ and $\mathcal{V} = \emptyset$, by its definition, yields the smallest *normalized* per-codebook gap between the two sides of inequality in (42), i.e.,

$$\begin{aligned}\beta_i^1 &\triangleq \beta_i(\mathbf{h}_i^{n+1}, \mathcal{K}, \emptyset, \mathbf{R}^n, \mathbf{T}) \\ &= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{K}} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \emptyset, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|},\end{aligned}\quad (43)$$

and the bottleneck set of codebooks corresponding to this smallest gap is

$$\mathcal{V}_1 \triangleq \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{K}} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \emptyset, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}.\quad (44)$$

Based on the structure of \mathcal{V}_1 and value of β_i^1 we update sets \mathcal{V} and \mathcal{U} and compute a second parameter β_i^2 . The procedure continues by successively refining \mathcal{V} via eliminating the codebooks that are not deemed beneficial if decoded, and hence are treated as noise. Specifically, this procedure successively identifies the group of codebooks which can be safely treated as Gaussian noise and partitions those to be decoded such that in each decoding step the number of codebooks to be decoded jointly does not exceed μ_i , thus limiting the decoding complexity. Furthermore, the procedure generates a sequence of parameters $\{\beta_i^1, \beta_i^2, \beta_i^3 \dots\}$, which collectively are sufficient for computing the *locally* optimal rate adaptation factor $\beta_i^*(n)$ for the i^{th} receiver, defined in (36). This procedure has at most $(K-1)M$ steps, where each step solves a problem of the form in (43), which as discussed earlier has a polynomial complexity in KM . The steps of this successive message elimination procedure are formalized in Algorithm 3. The optimality of this algorithm is demonstrated in the subsequent lemmas and theorem.

Lemma 6: For the sequence $\{\beta_i^k\}_{k=1}^q$ computed by Algorithm 1 we have $\beta_i^1 \leq \beta_i^2 \leq \dots \leq \beta_i^q$.

Proof: See Appendix G. \blacksquare

By leveraging Lemma 6 we provide the following lemma, which is instrumental to proving that Algorithm 3 yields the optimal local rate adaptation factor.

Lemma 7: If $\beta_i^*(n) > \beta_i^q$ then for the sets \mathcal{V}_i^q and \mathcal{G}_i yielded by Algorithm 3 we have

$$\mathcal{V}_i^q \subseteq \bar{\mathcal{Q}}_i^* \subseteq \mathcal{G}_i.$$

Proof: See Appendix H. \blacksquare

Algorithm 3 provides a constructive approach for determining the optimal local rate allocation factor $\beta_i^*(n)$. This constructive approach computes the optimal local rate adaptation factor $\beta_i^*(n)$ for the i^{th} receiver and generates partitions

$$\bar{\mathcal{Q}}_i^* = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\},\quad (45)$$

which achieve $\beta_i^*(n)$. Algorithm 3 also provides the steps involved in successive decoding by successively finding the set of messages that constitute the decoding bottleneck through exhibiting the smallest normalized gap between the mutual information function and the operating rates as formalized in (43). The optimality of Algorithm 3 is established in Theorem 2.

Algorithm 3 Symmetric Fair Rate Adaptation: Local Calculations

```

1: Initialize  $\mathcal{G} = \mathcal{K}$ ,  $\mathcal{V} = \emptyset$ , and  $k = 1$ .
2: repeat
3:   Find  $\beta_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
4:   Find  $\mathcal{V}_i^k = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
5:   if  $(i, m) \in \mathcal{V}_i^k$  for all  $m \in \{1, \dots, M\}$ 
6:      $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{V} \leftarrow \mathcal{V} \cup \mathcal{V}_i^k$  and  $k \leftarrow k + 1$ .
7:   end if
8: until  $\exists m \in \{1, \dots, M\}$  such that  $(i, m) \in \mathcal{V}_i^k$ .
9:  $p = k$ .
10: repeat
11:   Find  $\beta_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
12:   Find  $\mathcal{V}_i^k = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
13:   if  $(i, m) \in \mathcal{V}_i^k$  for all  $m \in \{1, \dots, M\}$ 
14:      $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{V} \leftarrow \mathcal{V} \cup \mathcal{V}_i^k$  and  $k \leftarrow k + 1$ .
15:   end if
16: until  $\exists m \in \{1, \dots, M\}$  such that  $(i, m) \in \mathcal{V}_i^k$ .
17:  $q = k$ .
18: set  $\mathcal{G}_i = \mathcal{G}$ .
19: repeat
20:   Find  $\beta_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
21:   Find  $\mathcal{Q}_i^{k-q+1} = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
22:   Set  $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{Q}_i^{k-q+1}$  and  $k \leftarrow k + 1$ .
23: until  $\nexists m \in \{1, \dots, M\}: (i, m) \in \mathcal{G}$ .
24: Output  $p_i = k - q$  and  $\{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ .

```

Theorem 2: Algorithm 3 yields the local optimal rate adaptation factor and the attendant optimal set of valid partitions $\bar{\mathcal{Q}}_i^*$, i.e.,

$$\beta_i^*(n) = \beta_i^q,$$

and

$$\bar{\mathcal{Q}}_i^* = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}.$$

Proof: See Appendix I. \blacksquare

B. Coordination

In Algorithm 3, each user acts autonomously and determines a *local* optimal rate change factor and the associated rate update policy based on its local information. However, distinct users do not necessarily prescribe identical rate updates. We show that for reaching a consensus among all users about the optimal rate change policy it is sufficient to have them exchange their local and independent computation results. Specifically, as it will be proven in Theorem 3, the i^{th} user has to report the value of $\beta_i^*(n)$ (one scalar) to all the transmitters. The globally optimal rate change factor $\beta^*(n)$ defined in (19) is then found as the minimum of all $\{\beta_i^*(n)\}_{i=1}^K$. Consequently, the new rate matrix \mathbf{R}^{n+1} is

$$\mathbf{R}^{n+1} = \mathbf{R}^n + \min_i \beta_i^*(n) \cdot \mathbf{T}.\quad (46)$$

Algorithm 4 Symmetric Fair Rate Adaptation - User Coordination

```

1: Input  $\mathbf{R}^n$ 
2: for  $i = 1, \dots, K$  do
3: Use Algorithm 2 to determine  $\beta_i^*(n)$  and  $\bar{\mathcal{Q}}_i^*$ .
4: end for
5: Update  $\mathbf{R}^{n+1} \leftarrow \mathbf{R}^n + \min_i \{\beta_i^*(n)\} \cdot \mathbf{T}$ 
6: Output  $\mathbf{R}^{n+1}$  and  $\{\bar{\mathcal{Q}}_i^*\}_{i=1}^K$ 

```

In Algorithm 4, first the optimal rate adaptation factor for each user is determined (through $\{\beta_i^*(n)\}_{i=1}^K$). Then the smallest such rate adaptation factor among all users (i.e., $\min_i \beta_i^*(n)$) is chosen to be the global rate adaptation factor, i.e., $\beta^*(n)_i = \min_i \{\beta_i^*(n)\}$.

The optimality of this algorithm is demonstrated in the following theorem.

Theorem 3: The updated rate matrix yielded by (46) satisfies $\mathbf{R}^{n+1} \succeq \tilde{\mathbf{R}}$ where $\tilde{\mathbf{R}}$ is any decodable rate matrix such that $\tilde{\mathbf{R}} = \mathbf{R}^n + \tilde{x} \cdot \mathbf{T}$ for some $\tilde{x} \in \mathbb{R}$.

Proof: If there exists a decodable rate-matrix $\tilde{\mathbf{R}}$ as defined above such that $\mathbf{R}^{n+1} \prec \tilde{\mathbf{R}}$, then we have

$$\tilde{x} > \beta^*(n) = \min_i \beta_i^*(n).$$

By defining $\ell \triangleq \arg \min_i \beta_i^*(n)$ we conclude $\tilde{x} > \beta_\ell^*$. Since $\tilde{\mathbf{R}}$ is decodable, then $\mathbf{R}^n + \tilde{x} \cdot \mathbf{T}$ is decodable by the ℓ^{th} receiver and $\tilde{x} > \beta_\ell^*(n)$, which contradicts the optimality of $\beta_\ell^*(n)$ as the optimal rate change factor for the ℓ^{th} user (Theorem 2). \blacksquare

VI. MAX-MIN FAIRNESS

This section focuses on rate adjustment under max-min fairness as formalized in (20), where the goal is to update the rate matrix \mathbf{R}^n to $\mathbf{R}^{n+1} = \mathbf{R}^n + \mathbf{r}$ such that the max-min fairness constraint in (20) is satisfied and the updated rate matrix \mathbf{R}^{n+1} remains decodable. In this section, we offer the procedure for identifying the optimal rate updates as well as the attendant optimal interference management strategy at each receiver. Similar to symmetric fairness, we propose the CPGD procedure for solving the max-min fair rate adaptation problem. Similarly to the symmetric fairness setting, the notable structure of this proposed procedure is that it breaks the rate adaption problem into K local problems, each specialized for one transmitter-receiver pair. These local problems are solved by the receivers in parallel, after which the receivers perform information exchange, which facilitates providing the optimal solution to all the users.

A. Local Interference Management

For a given rate matrix \mathbf{R}^n , the fairness constraint embedded in \mathbf{T} , any two disjoint subsets of the codebooks $\mathcal{U}, \mathcal{V} \subseteq \mathcal{K}$, and for any receiver $i \in \{1, \dots, K\}$ we define a rate change

factor as

$$\begin{aligned} \gamma_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T}) \\ \triangleq \begin{cases} \max \min_{\{i,m\}} \frac{r_{i,m}}{t_{i,m}} \\ \text{s.t. } \mathbf{R}^n + x \cdot \mathbf{r} \in \mathcal{C}_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}), \end{cases} \end{aligned} \quad (47)$$

where $\gamma_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T})$ is the maximum fair rate change that we can achieve by jointly decoding \mathcal{U} and treating \mathcal{V} as noise. Given the definition in (47), the maximum rate change factor that the i^{th} user can afford for a valid partitioning $\bar{\mathcal{Q}}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ of the codebooks at the k^{th} stage, for $k \in \{1, \dots, p_i\}$ is

$$\gamma_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T}). \quad (48)$$

Therefore, given partitions $\bar{\mathcal{Q}}_i$, the maximum rate change factor corresponding to which the i^{th} receiver remains decodable throughout all decoding stages is

$$\min_{k \in \{1, \dots, p_i\}} \{\gamma_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\}. \quad (49)$$

As a result, an optimal valid partition of the codebooks at the i^{th} receiver can be obtained by maximizing the rate factor change, i.e.,

$$\gamma_i^*(n) \triangleq \max_{\bar{\mathcal{Q}}_i \in \mathcal{Q}_i} \left\{ \min_k \{\gamma_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\} \right\}, \quad (50)$$

with the corresponding optimal partitioning \mathcal{Q}_i^* given by

$$\mathcal{Q}_i^* \triangleq \arg \max_{\bar{\mathcal{Q}}_i \in \mathcal{Q}_i} \left\{ \min_k \{\gamma_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\} \right\}. \quad (51)$$

The following lemma provides the underlying connection between the solution of the max-min fair rate change problem in (47) and the symmetric fair rate change problem in (33).

Lemma 8: For the i^{th} user and any two disjoint subsets \mathcal{U}, \mathcal{V} of \mathcal{K} , we have

$$\gamma_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T}) = \beta_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V}, \mathbf{R}^n, \mathbf{T}).$$

Proof: See Appendix J. \blacksquare

Based on Lemma 8 and the definition of $\gamma_i^*(n)$ in (50) we have

$$\gamma_i^*(n) \triangleq \max_{\bar{\mathcal{Q}}_i \in \mathcal{Q}_i} \left\{ \min_k \{\beta_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\} \right\}, \quad (52)$$

where $\mathcal{G} \subseteq \mathcal{K}$, $(i, m) \in \mathcal{G}$, $\forall m \in \{1, \dots, M\}$. By solving (52), the i^{th} receiver identifies a rate change policy. In particular, the i^{th} receiver recommends a rate update for each user such that the i^{th} receiver can successfully decode its designated user and max-min fairness is sustained, i.e., $r_{j,m}^k \triangleq \gamma_i^* t_{j,m}$. Algorithm 5 is a computationally efficient procedure with polynomial complexity in KM for finding the set of rate increments $\mathbf{r}^i = \{r_1^i, \dots, r_K^i\}$ by receiver i , where $\mathbf{r}_j^i \triangleq [r_{j,1}^i, \dots, r_{j,M}^i]$. Note that deploying an exhaustive search costs a complexity which scales as $O(3^{KM})$.

Theorem 4: If the i^{th} user is decodable under the rate \mathbf{R}^n , then it is also decodable under the rate matrix

Algorithm 5 Max-Min Rate Adaptation: Local Calculations

```

1: Initialize  $\mathcal{G} = \mathcal{K}$  and  $\mathcal{V} = \emptyset$ ,  $k = 1$  and  $\ell = 1$ .
2: repeat
3:   Find  $\beta_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
4:   Find  $\mathcal{V}_i^k = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
5:   if  $\exists m, (i, m) \in \mathcal{V}_i^k$ 
6:     Find  $\beta_i^k = \min_{\substack{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G} \\ |\mathcal{D}| \leq \mu_i}} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
7:     Find  $\mathcal{V}_i^k = \arg \min_{\substack{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G} \\ |\mathcal{D}| \leq \mu_i}} \frac{\Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n)}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}$ .
8:   end if
9:   if  $\exists m, (i, m) \in \mathcal{V}_i^k$ 
10:     $r_{j,m}^i = \beta_i^k t_{j,m}$  for all  $(j, m) \in \mathcal{V}_i^k$ ,
11:     $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{Q}_i^\ell = \mathcal{V}_i^k$ , and  $\ell \leftarrow \ell + 1$ .
12:   else
13:     $r_{j,m}^i = +\infty$  for all  $(j, m) \in \mathcal{V}_i^k$ ,
14:     $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{V}_i^k$ ,  $\mathcal{V} \leftarrow \mathcal{V} \cup \mathcal{V}_i^k$ .
15:   end if
16:    $k \leftarrow k + 1$ .
17: until  $\mathcal{G} = \emptyset$ 
18: Output  $\{r_j^i\}_{j=1}^K$  and  $\{\mathcal{Q}_i^j\}_{j=1}^{p_i}$ .

```

$\mathbf{R}^{n+1} = \mathbf{R}^n + \mathbf{r}^i$, where $[r_1^i, r_2^i, \dots, r_K^i]$ is yielded by Algorithm 5. Furthermore,

$$\min_{(\ell, m) \in \mathcal{K}} \frac{r_{\ell, m}^i}{t_{\ell, m}} \geq \min_{(\ell, m) \in \mathcal{K}} \frac{\tilde{r}_{\ell, m}^i}{t_{\ell, m}},$$

where $[\tilde{r}_1^i, \dots, \tilde{r}_K^i]$ is any other arbitrary rate update matrix for which the i^{th} user is decodable under the rates $\mathbf{R}^n + [\tilde{r}_1^i, \dots, \tilde{r}_K^i]$. \blacksquare

Proof: See Appendix K. \blacksquare

Therefore, according to Theorem 4, for each specific user i Algorithm 5 identifies the optimal rate changes for all users with the constraint that the i^{th} user remains decodable at its designated receiver. In the following subsection we show how the local rate changes computed by different users should be processed in order to find the globally optimal fair rate allocation.

B. Coordination

The core idea is that each user i solves the problem (52) and computes $\gamma_i^*(n)$ independently of the rest without any information exchange. These local solutions are then combined with minimal information exchange. Specifically, each user computes a rate change for its codebooks as well as other users' codebooks, which means that each user receives K rate change suggestions corresponding to each codebook one computed by itself and $(K-1)$ ones by the others. Next, each user selects the smallest rate change suggested for each of its codebooks. The steps for such rate change are formalized in Algorithm 6. The optimal properties of this algorithm are enumerated in Theorem 5.

Theorem 5: The distributed (iterative) max-min fair rate change algorithm (Algorithm 6) has the following properties

Algorithm 6 Max-Min Fair Rate Adaptation - User Coordination

```

1: Input  $\mathbf{R}^{(0)} = \mathbf{R}^n$  and  $q = 0$ .
2: repeat
3:   for  $i = 1, \dots, K$  do
4:     Run Algorithm 5 to determine  $\mathbf{r}^i$  and  $\bar{\mathcal{Q}}_i^*$ .
5:   end for
6:   Update  $q \leftarrow q + 1$ ,  $R_{i,m}^{(q)} = R_{i,m}^{\min} + \min_{1 \leq \ell \leq K} \{r_{i,m}^\ell\}$  for all  $i \in \{1, \dots, K\}$  and  $m \in \{1, \dots, M\}$ , and  $\mathbf{R}^{\min} \leftarrow \mathbf{R}^{(q)}$ .
7: until  $\mathbf{R}^{(q)}$  converges.
8: Output  $\mathbf{R}^{n+1} = \mathbf{R}^{(q)}$  and  $\{\bar{\mathcal{Q}}_i^*\}_{i=1}^K$ 

```

- 1) It is monotonic in the sense that $\mathbf{R}^{(q+1)} \succeq \mathbf{R}^{(q)}$, and it is convergent.
- 2) The matrix $\mathbf{R}^{(q)}$ is max-min optimal, i.e., for any other arbitrary decodable rate matrix $\tilde{\mathbf{R}} \succeq \mathbf{R}^{\min}$, we have

$$\min_{(i, m) \in \mathcal{K}} \frac{R_{i,m}^{(q)} - R_{i,m}^{\min}}{t_{i,m}} \geq \min_{(i, m) \in \mathcal{K}} \frac{\tilde{R}_{i,m} - R_{i,m}^{\min}}{t_{i,m}}.$$

- 3) The rate change yielded by Algorithm 6 is also pareto-optimal.

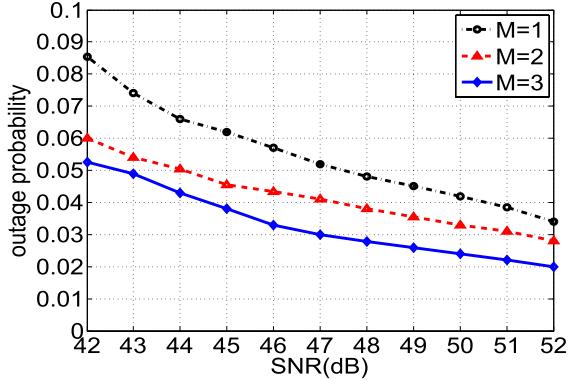
Proof: See Appendix L. \blacksquare

The second property in the theorem above, establishes that Algorithm 6 solves (20) and computes $\gamma^*(n)$. The pareto-optimality property rules out the possibility of unilaterally increasing the rate of one user. In other words, it states that the solution provided by Algorithm 6 has the property that any increase in the rate of one user must be penalized by a decrease in the rate of some other user(s).

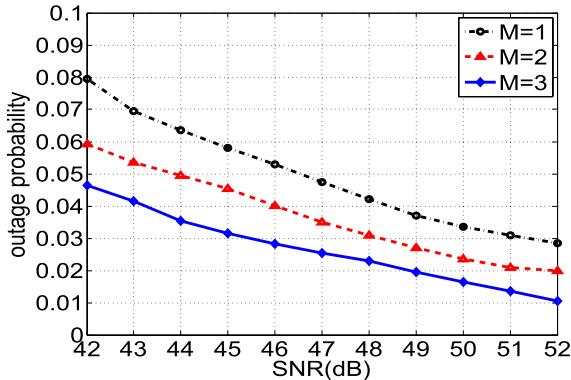
Remark 2: The solution to the max-min problem differs from that to the symmetric problem several aspects. Specifically, in the max-min fairness setting the optimal rate adaptation factor gives a set of possible rates as opposed to the symmetric case in which the rate choice is unique. Hence, over all possible rate matrices, we look for the one that is pareto-optimal (i.e., we cannot improve any rate without penalizing others). Also, the algorithms for finding the pareto-optimal rate are iterative such that in each iteration we find a rate matrix that outperforms the previous one and the procedure continues until convergence. Therefore, there are consecutive rounds of local computations and coordination (as opposed to symmetric case where the problem is solved in one shot).

VII. SIMULATION RESULTS

In this section, we provide simulation results to assess the performance of different CPGD algorithms and rate allocation schemes. Throughout all simulations, we consider a CPGD with three single-antenna transmitter-receiver pairs ($K = 3$). Each transmitter divides its message into M parts (we consider $M = 1, 2$, and 3) and all the partial messages of all the users are considered to have equal priority, i.e., $\mathbf{T} = \mathbf{1}_{M \times K}$. We assume that the channels are all quasi-static i.i.d. with



(a)



(b)

Fig. 1. The outage probability versus SNR for the fixed-rate setting. (a) With decoder size of $\mu_i = 2$. (b) With no decoder size constraint ($\mu_i = +\infty$).

Rayleigh fading distribution. In all simulations, 500 channel realizations are used.

First, we investigate the fixed rate mode and assume that transmitters' rate is $\frac{1}{2}$ bits per channel use and they have identical power constraints. Therefore, the set of rates for $M = 1$ is equal to

$$\mathbf{R} = \frac{1}{2} \cdot \mathbf{1}_{3 \times 1}, \quad (53)$$

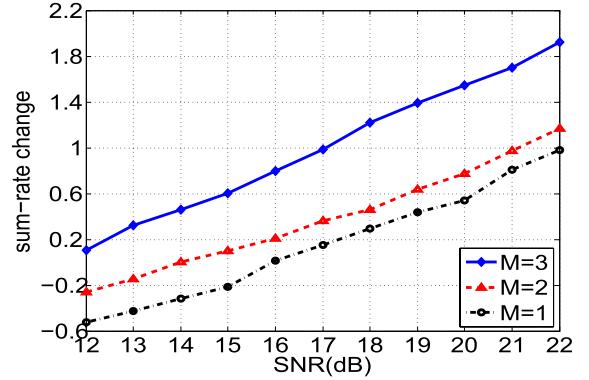
for $M = 2$ is equal to

$$\mathbf{R} = \frac{1}{4} \cdot \mathbf{1}_{3 \times 2}, \quad (54)$$

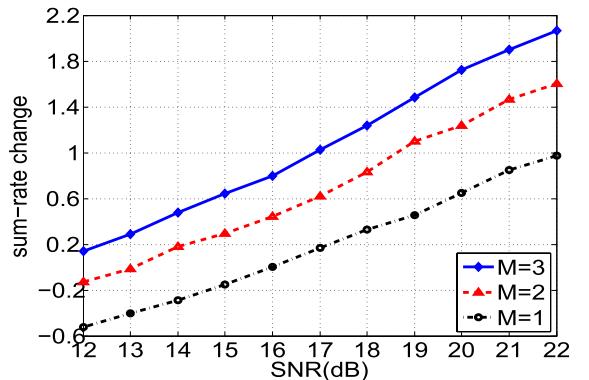
and for $M = 3$ is equal to

$$\mathbf{R} = \frac{1}{6} \cdot \mathbf{1}_{3 \times 3}. \quad (55)$$

In Figure 1, by using Algorithm 1 the outage probability (probability of at least one of the users being in outage) over a range of SNR is demonstrated. It is observed that splitting the messages into several sub-messages via CPGD even for $M = 2$ provides a substantial improvement in the outage probability. Also, comparing Figure 1(a) and Figure 1(b) demonstrates that imposing the decoding complexity constraint $\mu_i = 2$ penalizes the outage probability only slightly, while offering significant gain in decoding complexity.



(a)

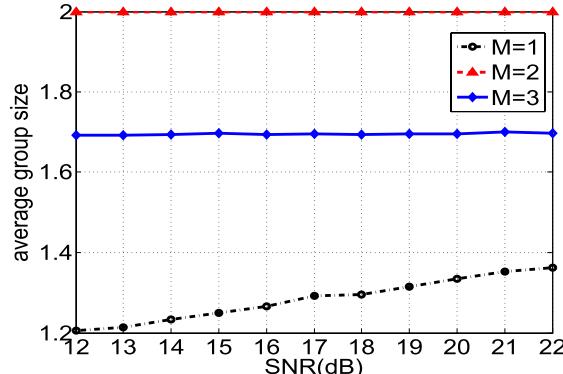


(b)

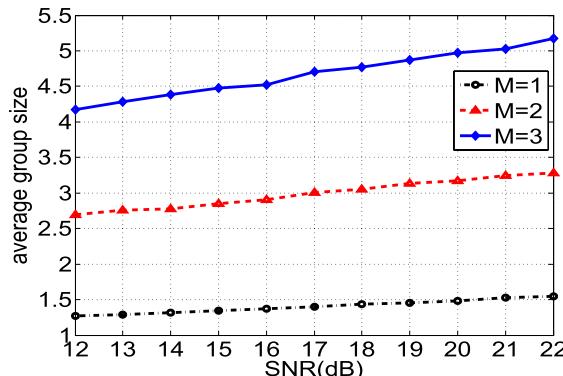
Fig. 2. The sum-rate versus SNR for the symmetric fairness setting. (a) With decoder size of $\mu_i = 2$. (b) With no decoder size constraint ($\mu_i = +\infty$).

Figures 2 and 3 focus on the symmetric fairness setting, where the initial rates are set according to the rates given in (53), (54) and (55) for $M = 1, 2$, and 3 , respectively. In Figure 2, by using Algorithm 3 and Algorithm 4, we compare sum-rate change, which is the product of the rate adaption factor and the number of sub-messages per user, over a range of SNR for different values of M . It is observed that splitting the messages into several sub-messages via CPGD even with $M = 2$ provides an improvement in the sum-rate. Also, comparing Figure 2(a) and Figure 2(b) demonstrates that imposing the decoding complexity constraint $\mu_i = 2$ penalizes the sum-rate change only slightly, while offering significant gain in decoding complexity. In both Figures 2(a) and 2(b), it is observed that for $\text{SNR}=12$ and $M = 1$ the sum-rate change is negative, which means that we need to decrease the rate of the transmitter to sustain reliable communication, which after increasing the number of each user's sub-messages to $M = 3$, the sum-rate change becomes positive, which means that the transmitters' messages are decodable and even the rates can be further increased. This observation confirms the effectiveness of rate splitting and superposition coding.

In Figure 3, we compare the average group size, i.e., $|\bar{\mathcal{Q}}_i| = ((\sum_{i=1}^K \sum_{k=1}^{p_i} |\mathcal{Q}_i^k|) / (\sum_{i=1}^K p_i))$, over a range of SNR for different values of M for the symmetric fairness setting.



(a)

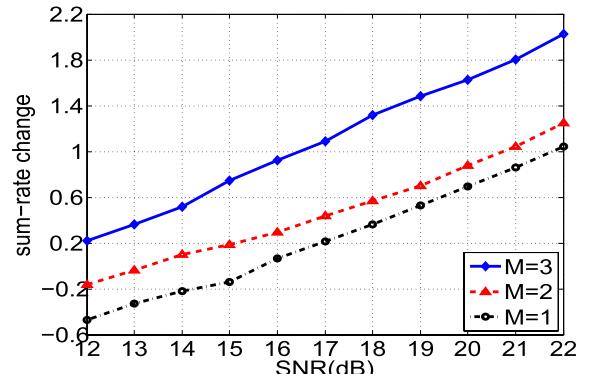


(b)

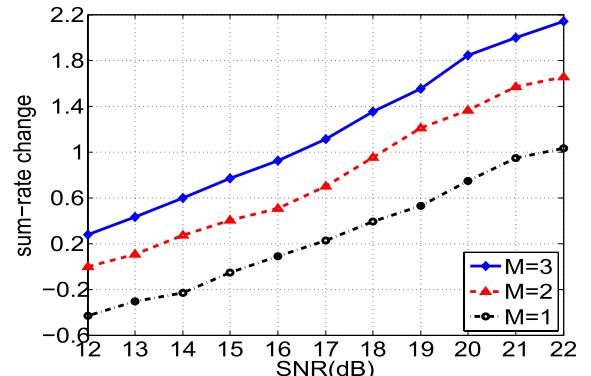
Fig. 3. The average group size versus SNR for the symmetric fairness setting. (a) With decoder size of $\mu_i = 2$. (b) With no decoder size constraint ($\mu_i = +\infty$).

An interesting observation is that splitting the messages into more sub-messages does not necessarily increase $|\bar{\mathcal{Q}}_i|$. In Figure 3(b), which presents the results for $\mu = +\infty$, there is a direct relationship between $|\bar{\mathcal{Q}}_i|$ and M as expected. However, an interesting behaviour is observed in Figure 3(a), which presents the results for $\mu_i = 2$, and that is the average group size for the case of $M = 2$ is larger than that for the case of $M = 3$. Comparing figures 2 and 3 indicates that imposing the constraint $\mu_i = 2$ does not penalize the rate significantly, while it decreases the decoding complexity significantly.

In Figure 4, the max-min fairness setting is evaluated, where the initial rates are set according to (53), (54), and (55) for $M = 1, 2$, and 3 , respectively. By using Algorithm 5 and Algorithm 6, we compare the sum of rate changes of all the sub-messages versus varying SNR for different values of M . It is observed that breaking the messages into several sub-messages via CPGD even with $M = 2$ provides an improvement in the achieved rate. We have the same observation here as in the case of symmetric fairness case and comparing Figure 4(a) and Figure 4(b) demonstrates that imposing the decoding complexity constraint $\mu_i = 2$ penalizes the sum-rate change only slightly, while offering significant gain in decoding complexity.



(a)



(b)

Fig. 4. The sum-rate versus SNR for the max-min fairness setting. (a) With decoder size of $\mu_i = 2$. (b) With no decoder size constraint ($\mu_i = +\infty$).

As a representative of the structure of CPGD for symmetric fairness model, next we consider the special case where $\forall i \in \{1, \dots, K\}$ we have set $\mu_i = 1$. For user i , each of the sequential decoding stages of the elements of a \mathcal{Q}_i^k for all $k \in \{1, \dots, p_i\}$ can achieve one edge of its corresponding achievable rate region. If the direction of \mathbf{T} (an image of \mathbf{T} corresponding to the elements of \mathcal{Q}_i^k) and the direction of one of the \mathcal{Q}_i^k 's achievable rate region edges are the same, the sequential decoding corresponding to that edge of the achievable rate region is optimum. In Figure 5 it is observed that the achievable rates via sequential decoding for two symmetric models. Figure 5(a) presents a setting in which we have 3 codebooks ($M = 3$) and aim to decode codebooks 1 and 2. By decoding them successively, we can achieve various corners of the achievable rate region. If we show the coordination of these $2! = 2$ corners shown in Figure 5(a) with (a_1, b_1) and (a_2, b_2) , with the constraint of $\mu_i = 1$ we can achieve the union of the two-dimensional intervals created by these two points, i.e., $\bigcup_{i=1}^2 ([0, a_i] \times [0, b_i])$. Figure 5(b) demonstrates the case that we have 4 codebooks and aim to decode codebooks 1, 2, and 3. By decoding these codebooks successively, we can achieve various corners of the achievable rate region. If we denote the coordinations of these $3! = 6$ corners shown in Figure 5(b) by $(a_1, b_1, c_1), \dots, (a_6, b_6, c_6)$, we can achieve the union of the three-dimensional regions created by these six points, i.e., $\bigcup_{i=1}^6 ([0, a_i] \times [0, b_i] \times [0, c_i])$.

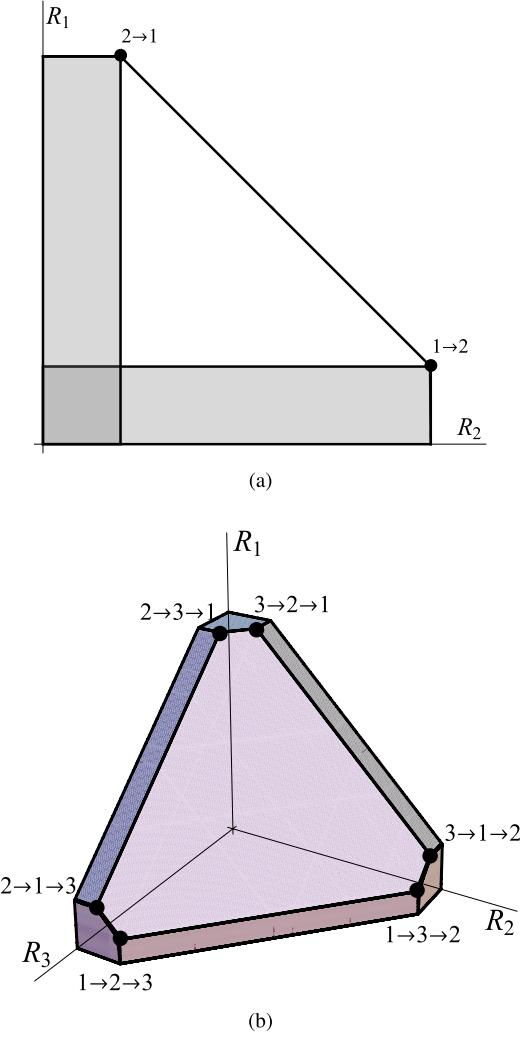


Fig. 5. Achievable rate regions with sequential decoding with $\mu_i = 1$. (a) With two signals of power P being decoded and one signal of power P being treated as noise. (b) With three signals of power P being decoded and one signal of power P being treated as noise.

VIII. CONCLUSIONS

Motivated by the challenges associated with acquiring the channel state information at the transmitters sites in the multi-user interference channel, we have proposed and analyzed constrained partial group decoders as an effective approach for interference management without requiring the channel state information at the transmitter and imposing only limited coordination among the users. This a receiver-centric approach to interference management, which in contrast to the majority of the existing art on interference management, relies strongly on processing at the receiver sites. The approach is discussed for certain fairness-constrained rate allocation problems where it is shown that optimal rate allocation is possible when 1) the transmitters do not have any channel state information, 2) the receivers have only local channel state information, 3) interference management decisions at each receiver are formed only based on the local information available to each receiver, and 4) the decoding complexity each receiver is controlled at a desired level.

APPENDIX A PROOF OF LEMMA 1

Each partition $\{\mathcal{K}, \mathcal{K} \setminus \bar{\mathcal{Q}}_i\}$ is of the form $\bar{\mathcal{Q}}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$, where

- 1) $p_i \geq 1$,
- 2) $\exists m \in \{1, \dots, M\}$ such that $(i, m) \in \mathcal{Q}_i^{p_i}$,
- 3) and $(i, m) \in \mathcal{Q}_i$ for all $m \in \{1, \dots, M\}$.

First, consider a particular choice of $\mathcal{Q}_i^{p_i}$ of size q , where $1 \leq q \leq \mu_i$ and define d as

$$d \triangleq |\{m \mid (k, m) \in \mathcal{Q}_i^{p_i}\}|, \quad (56)$$

where clearly $1 \leq d \leq \min\{q, M\}$ and therefore there exist

$$\sum_{d=1}^{\min\{q, M\}} \binom{(K-1)M}{q-d} \binom{M}{d}$$

such choices. For each such choice, there exist $\binom{(K-1)M-(q-d)}{s-(M-d)}$ ways to select s messages from the remaining messages where $(M-d)$ number of them are the main messages and the remaining $(s-(M-d))$ are chosen from the interfering codebooks to be decoded in other stages. Therefore, given a selection of q messages in $\mathcal{Q}_i^{p_i}$ and s included in sets $\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i-1}$, the number of possibilities to partition the aforementioned s messages using valid ordered partitions is given by

$$L_{(s, \mu_i)} = \sum_{\{b_s \in Z_+\}_{s=1}^{\mu_i}, \sum_{t=1}^{\mu_i} tb_t = s} \frac{(\sum_{t=1}^{\mu_i} b_t)! s!}{b_1! \dots b_{\mu_i}! (1!)^{b_1} \dots (\mu_i!)^{b_{\mu_i}}}. \quad (57)$$

Therefore, by defining $T_{K, M}$ as the cardinality of \mathcal{Q}_i when we have K user pairs each with M codebooks, we have

$$T_{K, M} = \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M\}} \binom{(K-1)M}{q-d} \binom{M}{d} \right) \times \left[\sum_{s=M-d}^{MK-q} \binom{(K-1)M-(q-d)}{s-(M-d)} L_{(s, \mu_i)} \right]. \quad (58)$$

Moreover, we can also derive the following difference equation, which establishes a connection between $T_{K+1, M}$ and $T_{K, M}$:

$$\begin{aligned} T_{K+1, M} &= \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M\}} \binom{KM}{q-d} \binom{M}{d} \right) \\ &\quad \times \left[\sum_{s=M-d}^{M(K+1)-q} \binom{KM-(q-d)}{s-(M-d)} L_{(s, \mu_i)} \right] \\ &= \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M\}} \binom{KM-1}{q-d-1} \binom{M}{d} \right) \\ &\quad \times \left[\sum_{s=M-d}^{M(K+1)-q} \binom{KM-1-(q-d)}{s-(M-d)} L_{(s, \mu_i)} \right] \end{aligned}$$

$$\begin{aligned}
& + \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M\}} \binom{KM-1}{q-d} \binom{M}{d} \right) \\
& \times \left[\sum_{s=M-d}^{M(K+1)-q} \binom{KM-1-(q-d)}{s-(M-d)} L_{(s, \mu_i)} \right] \\
& + \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M\}} \binom{KM}{q-d} \binom{M}{d} \right) \\
& \times \left[\sum_{s=M-d}^{M(K+1)-q-1} \binom{KM-1-(q-d)}{s-(M-d)} L_{(s+1, \mu_i)} \right] \\
& \geq 2T_{K,M} + \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M\}} \binom{KM-1}{q-d-1} \binom{M}{r} \right) \\
& \times \left[\sum_{s=M-d}^{M(K+1)-q} \binom{KM-1-(q-d)}{s-(M-d)} L_{(s, \mu_i)} \right] \\
& \geq 2T_{K,M}, \tag{59}
\end{aligned}$$

which shows that $T_{K,M}$ grows exponentially in K . Similarly we can find a relationship between $T_{K,M+1}$ and $T_{K,M}$ as follows.

$$\begin{aligned}
& T_{K,M+1} \\
& = \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M+1\}} \binom{(K-1)(M+1)}{q-d} \binom{M+1}{d} \right) \\
& \times \left[\sum_{s=M+1-d}^{(M+1)K-q} \binom{(K-1)(M+1)-(q-r)}{s-(M+1-d)} L_{(s, \mu_i)} \right] \\
& = \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M+1\}} \binom{(K-1)(M+1)-1}{q-d-1} \binom{M+1}{d} \right) \\
& \times \left[\sum_{s=M+1-d}^{(M+1)K-q} \binom{(K-1)(M+1)-1-(q-d)}{s-(M+1-d)} L_{(s, \mu_i)} \right] \\
& + \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M+1\}} \binom{(K-1)(M+1)-1}{q-d} \binom{M+1}{d} \right) \\
& \times \left[\sum_{s=M+1-d}^{(M+1)K-q} \binom{(K-1)(M+1)-1-(q-d)}{s-(M+1-d)} L_{(s, \mu_i)} \right] \\
& + \sum_{q=1}^{\mu_i} \left(\sum_{r=1}^{\min\{q, M+1\}} \binom{(K-1)(M+1)}{q-d} \binom{M+1}{d} \right) \\
& \times \left[\sum_{s=M+1-d}^{(M+1)K-q-1} \binom{(K-1)(M+1)-1-(q-d)}{s-(M+1-d)} L_{(s+1, \mu_i)} \right] \\
& \geq 2T_{K,M} \\
& + \sum_{q=1}^{\mu_i} \left(\sum_{d=1}^{\min\{q, M+1\}} \binom{(K-1)(M+1)-1}{q-d-1} \binom{M+1}{d} \right)
\end{aligned}$$

$$\begin{aligned}
& \times \left[\sum_{s=M+1-d}^{(M+1)K-q} \binom{(K-1)(M+1)-1-(q-d)}{s-(M+1-d)} L_{(s, \mu_i)} \right] \\
& \geq 2T_{K,M}, \tag{60}
\end{aligned}$$

which shows that $T_{K,M}$ grows exponentially in M .

APPENDIX B PROOF OF LEMMA 2

Since $P_{f_V}(\mathcal{U})$ is a polymatroid and

$$(\mathbf{R} + \alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \cdot \mathbf{1}_{K \times M}) \in P_{f_V}(\mathcal{U}),$$

then the condition that for every non-empty set $\mathcal{D} \subseteq \mathcal{U}$

$$\|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\| + \alpha_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}) |\mathcal{D}| \leq f_V(\mathcal{D}), \tag{61}$$

indicates that

$$\begin{aligned}
& \alpha_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}) \\
& \leq \alpha'_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}) \\
& \triangleq \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{f_V(\mathcal{D}) - \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\|}{|\mathcal{D}|}. \tag{62}
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
& \|\mathbf{R} \circ \mathbf{1}_{\mathcal{U}}\| + \alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) |\mathcal{U}| \\
& \leq \|\mathbf{R} \circ \mathbf{1}_{\mathcal{U}}\| + \frac{f_V(\mathcal{U}) - \|\mathbf{R} \circ \mathbf{1}_{\mathcal{U}}\|}{|\mathcal{U}|} |\mathcal{U}| \\
& = f_V(\mathcal{U}), \tag{63}
\end{aligned}$$

which implies that $(\mathbf{R} + \alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \cdot \mathbf{1}_{K \times M})$ falls within the polymatroid and, therefore, the upper bound $\alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ is achievable. As a result, $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) = \alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$.

APPENDIX C PROOF OF LEMMA 3

First we show that $\alpha_i^1 \leq \dots \leq \alpha_i^p$, where p is defined in line 9 of Algorithm 1. For $k \in \{1, \dots, p-1\}$ we have

$$\alpha_i^k = \alpha_i(\mathbf{h}_i, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R}) \tag{64}$$

$$\stackrel{(a)}{=} \frac{\Delta_k(\mathbf{h}_i, \mathcal{V}_k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^k|} \tag{65}$$

$$\stackrel{(b)}{=} \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j} \frac{\Delta_k(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^k|} \tag{66}$$

$$\stackrel{(c)}{\leq} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_i^k \cup \mathcal{V}_i^{k+1}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^{k+1}| + |\mathcal{V}_i^k|} \tag{67}$$

$$\stackrel{(d)}{=} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_{k+1}, \bigcup_{j=1}^k \mathcal{V}_i^j, \mathbf{R}) + \Delta_i(\mathbf{h}_i, \mathcal{V}_i^k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^{k+1}| + |\mathcal{V}_i^k|} \tag{68}$$

$$\stackrel{(e)}{=} \frac{\alpha_i^{k+1} |\mathcal{V}_{k+1}| + \alpha_i^k |\mathcal{V}_k|}{|\mathcal{V}_i^{k+1}| + |\mathcal{V}_i^k|}, \tag{69}$$

where (a) is based on definition of \mathcal{V}_i^k in line 4 of Algorithm 1, (b) is based on definition of \mathcal{V}_i^k in line 4 of Algorithm 1, (c) is based on optimality of \mathcal{V}_i^k , (d) is according to the chain rule in (11), and (e) is based on the definitions of α_i^{k+1} and α_i^k .

The inequality between (64) and (69) provides

$$\alpha_i^k (|\mathcal{V}_i^{k+1}| + |\mathcal{V}_i^k|) \leq \alpha_i^{k+1} |\mathcal{V}_i^{k+1}| + \alpha_i^k |\mathcal{V}_i^k|, \quad (70)$$

or equivalently

$$\alpha_i^k \leq \alpha_i^{k+1}, \quad \forall k \in \{1, \dots, p-1\}. \quad (71)$$

Next we show that $\alpha_i^p \leq \alpha_i^{p+1}$ based on their definitions. Specifically,

$$\alpha_i^p = \alpha_i(\mathbf{h}_i, \mathcal{K} \setminus \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \mathbf{R}) \quad (72)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{p-1} \mathcal{V}_i^j} \frac{\Delta_k(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_p|} \quad (73)$$

$$\stackrel{(a)}{\leq} \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_k(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_p|} \quad (74)$$

$$= \alpha_i^{p+1}, \quad (75)$$

where (a) is due to the additional constraint. Finally, by following a similar line of argument as in (64)-(69), for $k \in \{p, \dots, q-1\}$ we have

$$\alpha_i^k = \alpha_i(\mathbf{h}_i, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R}) \quad (76)$$

$$\stackrel{(a)}{=} \frac{\Delta_k(\mathbf{h}_i, \mathcal{V}_i^k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^k|} \quad (77)$$

$$\stackrel{(b)}{=} \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_k(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^k|} \quad (78)$$

$$\stackrel{(c)}{\leq} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_k \cup \mathcal{V}_{k+1}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^{k+1}| + |\mathcal{V}_i^k|} \quad (79)$$

$$\stackrel{(d)}{=} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_{k+1}, \bigcup_{j=1}^k \mathcal{V}_i^j, \mathbf{R}) + \Delta_i(\mathbf{h}_i, \mathcal{V}_k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{|\mathcal{V}_i^{k+1}| + |\mathcal{V}_i^k|} \quad (80)$$

$$\stackrel{(e)}{=} \frac{\alpha_i^{k+1} |\mathcal{V}_i^{k+1}| + \alpha_i^k |\mathcal{V}_i^k|}{|\mathcal{V}_i^{k+1}| + |\mathcal{V}_i^k|}, \quad (81)$$

where (a) is based on definition of \mathcal{V}_i^k in line 12 of Algorithm 1, (b) is based on definition of \mathcal{V}_i^k in line 12 of Algorithm 1, (c) is based on optimality of \mathcal{V}_i^k , (d) is according to the chain rule in (11), and (e) is based on the definitions of α_i^{k+1} and α_i^k , which establishes that

$$\alpha_i^k \leq \alpha_i^{k+1}, \quad \forall k \in \{p, \dots, q-1\}. \quad (82)$$

Hence, (71), (75), and (82) collectively establish that $\alpha_i^1 \leq \alpha_i^2 \leq \dots \leq \alpha_i^q$.

APPENDIX D PROOF OF LEMMA 4

1) $\mathcal{Q}_i \subseteq \mathcal{G}_i$: Note that $\mathcal{G}_i = \mathcal{K} \setminus \bigcup_{k=1}^{q-1} \mathcal{V}_i^k$. In order to show that $\mathcal{Q}_i \subseteq \mathcal{G}_i$ we equivalently show that $\mathcal{Q}_i \cap \mathcal{V}_i^k = \emptyset$ for $k \in \{1, \dots, q-1\}$. By contradiction, let us assume that \mathcal{Q}_i has non-empty intersection with some of the sets $\{\mathcal{V}_i^1, \dots, \mathcal{V}_i^{q-1}\}$ and denote by j the smallest value such that $\mathcal{Q}_i \cap \mathcal{V}_i^j \neq \emptyset$,

while for $k \in \{1, \dots, j-1\}$, we have $\mathcal{Q}_i \cap \mathcal{V}_i^j = \emptyset$. By using the expansion

$$\mathcal{V}_i^j = \{\mathcal{Q}_i \cap \mathcal{V}_i^j\} \cup \{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j\}, \quad (83)$$

and the properties of Δ_i defined in (11)-(12) along with the definitions of α_i^j and \mathcal{V}_i^j (Algorithm 3) we get

$$\alpha_i^j |\mathcal{V}_i^j| \stackrel{(a)}{=} \Delta_i(\mathbf{h}_i, \mathcal{V}_i^j, \bigcup_{k=1}^{j-1} \mathcal{V}_k, \mathbf{R}) \quad (84)$$

$$\stackrel{(b)}{=} \Delta_i(\mathbf{h}_i, \{\mathcal{Q}_i \cap \mathcal{V}_i^j\} \cup \{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j\}, \bigcup_{k=1}^{j-1} \mathcal{V}_k, \mathbf{R}) \quad (85)$$

$$\stackrel{(c)}{=} \Delta_i(\mathbf{h}_i, \{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j, \bigcup_{k=1}^{j-1} \mathcal{V}_i^j, \mathbf{R}) + \Delta_i(\mathbf{h}_i, \mathcal{Q}_i \cap \mathcal{V}_i^j, \{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j\} \cup \{\bigcup_{k=1}^{j-1} \mathcal{V}_i^j\}, \mathbf{R}) \quad (86)$$

$$\stackrel{(d)}{\geq} \Delta_i(\mathbf{h}_i, \underbrace{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j}_{\subseteq \mathcal{V}_i^j}, \bigcup_{i=1}^{j-1} \mathcal{V}_i, \mathbf{R}^n) + \Delta_i(\mathbf{h}_i, \underbrace{\mathcal{Q}_i \cap \mathcal{V}_i^j}_{\subseteq \mathcal{Q}_i}, \mathcal{K} \setminus \mathcal{Q}_i \mathbf{R}), \quad (87)$$

where (a) holds based on definition of \mathcal{V}_i^k in line 4 of Algorithm 1, (b) is based on (83), (c) holds according to the chain rule in (11), and (d) follows from $\{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j\} \cup \{\bigcup_{k=1}^{j-1} \mathcal{V}_i^j\} \subseteq \mathcal{K} \setminus \mathcal{Q}_i$. Next, note that for any $\mathcal{D} \subseteq \mathcal{V}_i^j$ we have

$$\alpha_i^j = \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_i^j, \bigcup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R})}{|\mathcal{V}_i^j|} \leq \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \bigcup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R})}{|\mathcal{D}|}, \quad (88)$$

or equivalently

$$\Delta_i(\mathbf{h}_i, \mathcal{D}, \bigcup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R}) \geq \alpha_i^j |\mathcal{D}|.$$

Now, by setting $\mathcal{D} = \{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j$ we get

$$\Delta_i(\mathbf{h}_i, \{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j, \bigcup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R}) \geq \alpha_i^j |\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j|. \quad (89)$$

By following the same line of argument we can also show that

$$\Delta_i(\mathbf{h}_i, \mathcal{Q}_i \cap \mathcal{V}_i^j, \mathcal{K} \setminus \mathcal{Q}_i, \mathbf{R}) \geq \alpha_i^* |\mathcal{Q}_i \cap \mathcal{V}_i^j|. \quad (90)$$

Combining (87)-(90) provides that

$$\alpha_i^j |\mathcal{V}_i^j| \geq \alpha_i^j |\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j| + \alpha_i^* |\mathcal{Q}_i \cap \mathcal{V}_i^j| \quad (91)$$

$$> \alpha_i^j \left(|\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j| + |\mathcal{Q}_i \cap \mathcal{V}_i^j| \right) \quad (92)$$

$$= \alpha_i^j |\mathcal{V}_i^j|, \quad (93)$$

where (92) follows from $\alpha_i^* > \alpha_i^j$. Comparing (91) and (93) indicates a contradiction. Hence, for all $j \in \{1, \dots, q-1\}$, $\mathcal{Q}_k \cap \mathcal{V}_i^j = \emptyset$ and as a result $\mathcal{Q}_i \subseteq \mathcal{G}_i$.

2) $\mathcal{V}_i^q \subseteq \mathcal{Q}_i$: If $\mathcal{V}_i^q \not\subseteq \mathcal{Q}_i$, then by noting that $\mathcal{V}_i^q \subseteq \mathcal{G}_i$ we conclude $\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\} \neq \emptyset$. By expanding $\mathcal{V}_i^q = \{\mathcal{V}_i^q \cap \mathcal{Q}_i\} \cup \{\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}\}$ and following the same line of argument as in the previous case we have

$$\alpha_i^q |\mathcal{V}_i^q| = \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q, \bigcup_{k=1}^{q-1} \mathcal{V}_i^k, \mathbf{R}) \quad (94)$$

$$\stackrel{(a)}{=} \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q \cap \mathcal{Q}_i, \{\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}\} \cup \{\cup_{k=1}^{q-1} \mathcal{V}_i^k\}, \mathbf{R}) \\ + \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}, \cup_{k=1}^{q-1} \mathcal{V}_i^k, \mathbf{R}^n) \quad (95)$$

$$\stackrel{(b)}{\geq} \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q \cap \mathcal{Q}_i, \mathcal{K} \setminus \mathcal{Q}_i, \mathbf{R}) + \alpha_i^q |\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}| \quad (96)$$

$$\stackrel{(c)}{\geq} \alpha_i^* |\mathcal{V}_i^q \cap \mathcal{Q}_i| + \alpha_i^q |\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}| \quad (97)$$

$$\stackrel{(d)}{>} \alpha_i^q |\mathcal{V}_i^q|, \quad (98)$$

where (a) is according to the chain rule in (11), and (b) follows from $\{\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}\} \cup \{\cup_{k=1}^{q-1} \mathcal{V}_i^k\} \subset \mathcal{K} \setminus \mathcal{Q}_i$, (c) follows from $\mathcal{V}_i^q \cap \mathcal{Q}_i \subseteq \mathcal{Q}_i$, and (d) follows from $\alpha_i^* > \alpha_i^q$. Comparing (94) and (98) establishes a contradiction, and as a result $\mathcal{V}_i^q \subseteq \mathcal{Q}_i$.

APPENDIX E PROOF OF THEOREM 1

We show that for the output of Algorithm 1 we have $\alpha_i^* = \alpha_i^q$, which in conjunction with the condition for outage provided in (25), establishes the desired result.

1) $\alpha_i^* \leq \alpha_i^q$:

We provide the proof by contradiction. Suppose that we have $\alpha_i^* > \alpha_i^q$. Then, according to Lemma 4 we have $\mathcal{V}_i^q \subseteq \mathcal{Q}_i \subseteq \mathcal{G}_i$. Let us define $\mathcal{F}_i \triangleq \mathcal{G}_i \setminus \mathcal{Q}_i$. Therefore

$$\mathcal{Q}_i = \mathcal{G}_i \setminus \mathcal{F}_i, \quad \mathcal{V}_i^q \subseteq \mathcal{G}_i \setminus \mathcal{F}_i, \text{ and } \mathcal{K} \setminus \mathcal{Q}_i = \{\mathcal{K} \setminus \mathcal{G}_i\} \cup \mathcal{F}_i. \quad (99)$$

Consequently, we get

$$\alpha_i^q = \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_i^q, \cup_{k=1}^{q-1} \mathcal{V}_i^k, \mathbf{R})}{|\mathcal{V}_i^q|} \quad (100)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{K} \setminus \mathcal{G}_i, \mathbf{R})}{|\mathcal{D}|} \quad (101)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}_i \setminus \mathcal{F}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{K} \setminus \mathcal{G}_i, \mathbf{R})}{|\mathcal{D}|} \quad (102)$$

$$\geq \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}_i \setminus \mathcal{F}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \{\mathcal{K} \setminus \mathcal{G}_i\} \cup \mathcal{F}_i, \mathbf{R})}{|\mathcal{D}|} \quad (103)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{Q}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{K} \setminus \mathcal{Q}_i, \mathbf{R})}{|\mathcal{D}|} \quad (104)$$

$$= \alpha_i^*, \quad (105)$$

where (102) holds since $\mathcal{V}_i^q \subseteq \mathcal{G}_i \setminus \mathcal{F}_i$ and therefore minimizing over the sets \mathcal{G}_i or $\mathcal{G}_i \setminus \mathcal{F}_i$ yields the same result. Equation (103) follows from (12), and (104) follows from the definitions provided in (99). Comparing (100) and (105) shows that $\alpha_i^* \leq \alpha_i^q$, which contradicts the initial assumption that $\alpha_i^* > \alpha_i^q$. Hence, we have $\alpha_i^* \leq \alpha_i^q$.

2) $\alpha_i^* \geq \alpha_i^q$:

Next we show, constructively, that there exist a set of partitions $\{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ which achieves α_i^q , which in conjunction with $\alpha_i^* \leq \alpha_i^q$, establishes that $\alpha_i^* = \alpha_i^q$. Specifically, we show that given the outputs of Algorithm 1, i.e., \mathcal{G}_i and q , the sets $\{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ yielded by Algorithm 1 offer a valid successive decoding order corresponding to which the optimal rate gap factor is α_i^q .

Algorithm 7 Partitioning $\bar{\mathcal{Q}}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$

1: Initialize $\mathcal{G} = \mathcal{Q}_i$, $\mathcal{V} = \mathcal{K} \setminus \mathcal{Q}_i$ and $k = q$.
2: **repeat**
3: Find $\beta_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$.
4: Find $\mathcal{Q}_i^{k-q+1} = \arg \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|}$.
5: Set $\mathcal{G} \leftarrow \mathcal{G} \setminus \mathcal{Q}_i^{k-q+1}$ and $i \leftarrow i + 1$.
6: **until** $\nexists m \in \{1, \dots, M\}: (i, m) \in \mathcal{G}$.
7: Output $p_i = k - q$ and $\{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$.

To this end, we show that

$$\alpha_i^q \leq \dots \leq \alpha_i^{q+p_i-1}, \quad (106)$$

which can be readily verified by noting that for $k \in \{q, \dots, q+p_i-1\}$ we have

$$\alpha_i^k = \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{Q}_i \setminus \cup_{j=1}^{k-q} \mathcal{Q}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|} \quad (107)$$

$$\leq \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{Q}_i \setminus \cup_{j=1}^{k-q+1} \mathcal{Q}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{|\mathcal{D}|} \quad (108)$$

where $q \leq i \leq q+p_i-2$, and (107) holds due to the fact that $\mathcal{Q}_i \setminus \cup_{j=1}^{k-q+1} \mathcal{Q}_i^j$ is a subset of $\mathcal{Q}_i \setminus \cup_{j=1}^{k-q} \mathcal{Q}_i^j$. By recalling the definition of α_i^* from (25) we have

$$\alpha_i^* \triangleq \max_{\mathcal{Q}_i \in \mathcal{Q}_i} \left\{ \min_k \{\alpha_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \cup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R})\} \right\} \quad (109)$$

$$\geq \left\{ \min_k \{\alpha_i(\mathbf{h}_i, \bar{\mathcal{Q}}_i^k, \mathcal{K} \setminus \cup_{j=1}^{k-1} \bar{\mathcal{Q}}_i^j, \mathbf{R})\} \right\} \quad (110)$$

$$= \min_{k \in \{q, \dots, q+p_i\}} \alpha_i^k \quad (111)$$

$$= \alpha_i^q, \quad (112)$$

where the last step holds by invoking the order in (106). Equations (105) and (112) prove that $\alpha_i^* = \alpha_i^q$.

APPENDIX F PROOF OF LEMMA 5

Since $P_{f_{\mathcal{V}}}(\mathcal{U})$ is a polymatroid and $\mathbf{R} + \alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \cdot \mathbf{1}_{K \times M} \in P_{f_{\mathcal{V}}}(\mathcal{U})$, then the inequality $\mathcal{D} \subseteq \mathcal{U}$ for the nonempty set \mathcal{D}

$$\|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\| + \alpha_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}) \|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\| \leq f_{\mathcal{V}}(\mathcal{D}) \quad (113)$$

indicates that

$$\alpha_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}) \leq \alpha'_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R}) \triangleq \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{f_{\mathcal{V}}(\mathcal{D}) - \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}, \quad (114)$$

which means that $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ is bounded above by

$$\min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{f_{\mathcal{V}}(\mathcal{D}) - \|\mathbf{R} \circ \mathbf{1}_{\mathcal{D}}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{D}}\|}. \quad (115)$$

On the other hand,

$$\|\mathbf{R} \circ \mathbf{1}_{\mathcal{U}}\| + \alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \|\mathbf{T} \circ \mathbf{1}_{\mathcal{U}}\|$$

$$\leq \|\mathbf{R} \circ \mathbf{1}_{\mathcal{U}}\| + \frac{f_{\mathcal{V}}(\mathcal{U}) - \|\mathbf{R} \circ \mathbf{1}_{\mathcal{U}}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{U}}\|} \|\mathbf{T} \circ \mathbf{1}_{\mathcal{U}}\| \\ = f_{\mathcal{V}}(\mathcal{U}) \quad (116)$$

which shows that $(\mathbf{R} + \alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) \cdot \mathbf{1}_{K \times M})$ falls within the polymatroid and therefore the upperbound $\alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$ is achievable, or $\alpha_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R}) = \alpha'_i(\mathbf{h}_i, \mathcal{U}, \mathcal{V}, \mathbf{R})$.

APPENDIX G PROOF OF LEMMA 6

First we show that $\beta_i^1 \leq \dots \leq \beta_i^p$. For $k \in \{1, \dots, p-1\}$ we have

$$\beta_i^k = \beta_i(\mathbf{h}_i, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R}) \quad (117)$$

$$\stackrel{(a)}{=} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (118)$$

$$\stackrel{(b)}{=} \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (119)$$

$$\stackrel{(c)}{\leq} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_i^k \cup \mathcal{V}_i^{k+1}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (120)$$

$$\stackrel{(d)}{=} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_{k+1}, \bigcup_{j=1}^k \mathcal{V}_i^j, \mathbf{R}) + \Delta_i(\mathbf{h}_i, \mathcal{V}_i^k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (121)$$

$$\stackrel{(e)}{=} \frac{\beta_i^{k+1} \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_{k+1}}\| + \beta_i^k \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|}, \quad (122)$$

where (a) is based on definition of \mathcal{V}_i^k in line 4 of Algorithm 3, (b) is based on definition of \mathcal{V}_i^k in line 4, (c) is based on optimality of \mathcal{V}_i^k , (d) is according to the chain rule in (11), and (e) is based on the definitions of β_i^{k+1} and β_i^k . The inequality between (117) and (122) provides

$$\beta_i^k \left(\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\| \right) \\ \leq \beta_i^{k+1} \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \beta_i^k \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|, \quad (123)$$

or equivalently

$$\beta_i^k \leq \beta_i^{k+1}, \quad \forall k \in \{1, \dots, p-1\}. \quad (124)$$

Next we show that $\beta_i^p \leq \beta_i^{p+1}$ based on their definitions. Specifically,

$$\beta_i^p = \beta_i(\mathbf{h}_i, \mathcal{K} \setminus \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \mathbf{R}) \quad (125)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{p-1} \mathcal{V}_i^j} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_p}\|} \quad (126)$$

$$\stackrel{(a)}{\leq} \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{p-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_p}\|} \quad (127)$$

$$= \beta_i^{p+1}, \quad (128)$$

where (a) is due to the additional constraint. Finally, by following a similar line of argument as in (117)-(122), for $k \in \{p, \dots, q-1\}$ we have

$$\beta_i^k = \beta_i(\mathbf{h}_i, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R}) \quad (129)$$

$$\stackrel{(a)}{=} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_i^k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (130)$$

$$\stackrel{(b)}{=} \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \in \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (131)$$

$$\stackrel{(c)}{\leq} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_k \cup \mathcal{V}_{k+1}, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (132)$$

$$\stackrel{(d)}{=} \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_{k+1}, \bigcup_{j=1}^k \mathcal{V}_i^j, \mathbf{R}) + \Delta_i(\mathbf{h}_i, \mathcal{V}_k, \bigcup_{j=1}^{k-1} \mathcal{V}_i^j, \mathbf{R})}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|} \quad (133)$$

$$\stackrel{(e)}{=} \frac{\beta_i^{k+1} \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_{k+1}}\| + \beta_i^k \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|}{\|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^{k+1}}\| + \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^k}\|}, \quad (134)$$

where (a) is based on definition of \mathcal{V}_i^k in line 12 of Algorithm 3, (b) is based on definition of \mathcal{V}_i^k in line 12 of Algorithm 3, (c) is based on optimality of \mathcal{V}_i^k , (d) is according to the chain rule in (11), and (e) is based on the definitions of β_i^{k+1} and β_i^k , which establishes that

$$\beta_i^k \leq \beta_i^{k+1}, \quad \forall k \in \{p, \dots, q-1\}. \quad (135)$$

Hence, (124), (128), and (135) collectively establish that $\beta_i^1 \leq \beta_i^2 \leq \dots \leq \beta_i^q$.

APPENDIX H PROOF OF LEMMA 7

1) $\mathcal{Q}_i \subseteq \mathcal{G}_i$: Note that $\mathcal{G}_i = \mathcal{K} \setminus \bigcup_{k=1}^{q-1} \mathcal{V}_i^k$. In order to show that $\mathcal{Q}_i \subseteq \mathcal{G}_i$ we equivalently show that $\mathcal{Q}_i \cap \mathcal{V}_i^k = \emptyset$ for $k \in \{1, \dots, q-1\}$. By contradiction, let us assume that \mathcal{Q}_i has non-empty intersection with some of the sets $\{\mathcal{V}_i^1, \dots, \mathcal{V}_i^{q-1}\}$ and denote by j the smallest value such that $\mathcal{Q}_i \cap \mathcal{V}_i^j \neq \emptyset$, while for $k \in \{1, \dots, j-1\}$, we have $\mathcal{Q}_i \cap \mathcal{V}_i^j = \emptyset$. By using the expansion

$$\mathcal{V}_i^j = \{\mathcal{Q}_i \cap \mathcal{V}_i^j\} \cup \{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j\}, \quad (136)$$

and the properties of Δ_i defined in (11)-(12) along with the definitions of β_i^j and \mathcal{V}_i^j (Algorithm 3) we get

$$\begin{aligned} & \beta_i^j \|\mathbf{T} \circ \mathbf{1}_{\mathcal{V}_i^j}\| \\ & \stackrel{(a)}{=} \Delta_i(\mathbf{h}_i, \mathcal{V}_i^j, \bigcup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R}) \\ & \stackrel{(b)}{=} \Delta_i(\mathbf{h}_i, \{\mathcal{Q}_i \cap \mathcal{V}_i^j\} \cup \{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j\}, \bigcup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R}) \\ & \stackrel{(c)}{=} \Delta_i(\mathbf{h}_i, \{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j, \bigcup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R}) \\ & \quad + \Delta_i(\mathbf{h}_i, \mathcal{Q}_i \cap \mathcal{V}_i^j, \{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j\} \cup \{\bigcup_{k=1}^{j-1} \mathcal{V}_i^k\}, \mathbf{R}) \\ & \stackrel{(d)}{\geq} \Delta_i(\mathbf{h}_i, \underbrace{\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j}_{\subseteq \mathcal{V}_i^j}, \bigcup_{i=1}^{j-1} \mathcal{V}_i^j, \mathbf{R}^n) \\ & \quad + \Delta_i(\mathbf{h}_i, \underbrace{\mathcal{Q}_i \cap \mathcal{V}_i^j}_{\subseteq \mathcal{Q}_i}, \mathcal{K} \setminus \mathcal{Q}_i, \mathbf{R}), \end{aligned} \quad (137)$$

where (a) is based on definition of \mathcal{V}_i^k in line 4 in Algorithm 3, (b) is based on (136), (c) follows chain rule in (11), and (d)

follows from $\{\mathcal{K} \setminus \mathcal{Q}_k\} \cap \mathcal{V}_i^j\} \cup \{\cup_{i=1}^{j-1} \mathcal{V}_i^j\} \subseteq \mathcal{K} \setminus \mathcal{Q}_k$. Next, note that for any $\mathcal{D} \subseteq \mathcal{V}_i^j$ we have

$$\beta_i^j = \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_i^j, \cup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^j}\|} \leq \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \cup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|}, \quad (138)$$

or equivalently

$$\Delta_i(\mathbf{h}_i, \mathcal{D}, \cup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R}) \geq \beta_i^j \|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|.$$

Now, by setting $\mathcal{D} = \{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j$ we get

$$\Delta_i(\mathbf{h}_i, \{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j, \cup_{k=1}^{j-1} \mathcal{V}_i^k, \mathbf{R}) \geq \beta_i^j |\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j|. \quad (139)$$

By following the same line of argument we can also show that

$$\Delta_i(\mathbf{h}_i, \mathcal{Q}_i \cap \mathcal{V}_i^j, \mathcal{K} \setminus \mathcal{Q}_i, \mathbf{R}) \geq \beta_i^* |\mathcal{Q}_i \cap \mathcal{V}_i^j|. \quad (140)$$

Combining (137)-(140) provides that

$$\beta_i^j \|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^j}\| \geq \beta_i^j |\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j| + \beta_i^* |\mathcal{Q}_i \cap \mathcal{V}_i^j| \quad (141)$$

$$> \beta_i^j (|\{\mathcal{K} \setminus \mathcal{Q}_i\} \cap \mathcal{V}_i^j| + |\mathcal{Q}_i \cap \mathcal{V}_i^j|) \quad (142)$$

$$= \beta_i^j \|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^j}\|, \quad (143)$$

where (142) follows from $\beta_i^* > \beta_i^j$. Comparing (141) and (143) indicates a contradiction. Hence, for all $j \in \{1, \dots, q-1\}$, $\mathcal{Q}_k \cap \mathcal{V}_i^j = \emptyset$ and as a result $\mathcal{Q}_i \subseteq \mathcal{G}_i$.

2) $\mathcal{V}_i^q \subseteq \mathcal{Q}_i$:

If $\mathcal{V}_i^q \not\subseteq \mathcal{Q}_i$, then by noting that $\mathcal{V}_i^q \subseteq \mathcal{G}_i$ we conclude $\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\} \neq \emptyset$. By expanding $\mathcal{V}_i^q = \{\mathcal{V}_i^q \cap \mathcal{Q}_i\} \cup \{\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}\}$ and following the same line of argument as in the previous case we have

$$\begin{aligned} & \beta_i^q \|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^q}\| \\ &= \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q, \cup_{k=1}^{q-1} \mathcal{V}_i^k, \mathbf{R}) \end{aligned} \quad (144)$$

$$\begin{aligned} &= \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q \cap \mathcal{Q}_i, \{\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}\} \cup \{\cup_{k=1}^{q-1} \mathcal{V}_i^k\}, \mathbf{R}) \\ &+ \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}, \cup_{k=1}^{q-1} \mathcal{V}_i^k, \mathbf{R}) \end{aligned} \quad (145)$$

$$\geq \Delta_i(\mathbf{h}_i, \mathcal{V}_i^q \cap \mathcal{Q}_i, \mathcal{K} \setminus \mathcal{Q}_i, \mathbf{R}) + \beta_i^q |\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}| \quad (146)$$

$$\geq \beta_i^* \|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^q} \| \cap \mathcal{Q}_i | + \beta_i^q \|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^q} \| \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}| \quad (147)$$

$$> \beta_i^q \|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^q}\|, \quad (148)$$

where (146) follows from $\{\mathcal{V}_i^q \cap \{\mathcal{G}_i \setminus \mathcal{Q}_i\}\} \cup \{\cup_{k=1}^{q-1} \mathcal{V}_i^k\} \subseteq \mathcal{K} \setminus \mathcal{Q}_i$, (147) follows from $\mathcal{V}_i^q \cap \mathcal{Q}_i \subseteq \mathcal{Q}_i$, and (148) follows from $\beta_i^* > \beta_i^q$. Comparing (144) and (148) establishes a contradiction, and as a result $\mathcal{V}_i^q \subseteq \mathcal{Q}_i$.

APPENDIX I PROOF OF THEOREM 2

We show that for the output of Algorithm 3 we have $\beta_i^* = \beta_i^q$, which in conjunction with the definition of optimal local rate adaptation factor in (19), establishes the desired result.

1) $\beta_i^* \leq \beta_i^q$:

We provide the proof by contradiction. Suppose that we have $\beta_i^* > \beta_i^q$. Then, according to Lemma 7 we have $\mathcal{V}_i^q \subseteq \mathcal{Q}_i \subseteq \mathcal{G}_i$. Let us define $\mathcal{F}_i \triangleq \mathcal{G}_i \setminus \mathcal{Q}_i$. Therefore

$$\mathcal{Q}_i = \mathcal{G}_i \setminus \mathcal{F}_i, \quad \mathcal{V}_i^q \subseteq \mathcal{G}_i \setminus \mathcal{F}_i, \quad \text{and} \quad \mathcal{K} \setminus \mathcal{Q}_i = \{\mathcal{K} \setminus \mathcal{G}_i\} \cup \mathcal{F}_i. \quad (149)$$

Consequently, we get

$$\beta_i^q = \frac{\Delta_i(\mathbf{h}_i, \mathcal{V}_i^q, \cup_{k=1}^{q-1} \mathcal{V}_i^k, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{V}_i^q}\|} \quad (150)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{K} \setminus \mathcal{G}_i, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|} \quad (151)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}_i \setminus \mathcal{F}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{K} \setminus \mathcal{G}_i, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|} \quad (152)$$

$$\geq \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{G}_i \setminus \mathcal{F}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \{\mathcal{K} \setminus \mathcal{G}_i\} \cup \mathcal{F}_i, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|} \quad (153)$$

$$= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{Q}_i, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{K} \setminus \mathcal{Q}_i, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|} \quad (154)$$

$$= \beta_i^*, \quad (155)$$

where (152) holds since $\mathcal{V}_i^q \subseteq \mathcal{G}_i \setminus \mathcal{F}_i$ and therefore minimizing over the sets \mathcal{G}_i or $\mathcal{G}_i \setminus \mathcal{F}_i$ yields the same result. Equation (153) follows from (12) and (154) follows from the definitions provided in (149). Comparing (150) and (155) shows that $\beta_i^* \leq \beta_i^q$, which contradicts the initial assumption that $\beta_i^* > \beta_i^q$. Hence, we have $\beta_i^* \leq \beta_i^q$.

2) $\beta_i^* \geq \beta_i^q$:

Next we show that the set of partitions $\{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ yielded by Algorithm 3 achieves β_i^q , which in conjunction with $\beta_i^* \leq \beta_i^q$, establishes that $\beta_i^* = \beta_i^q$. To this end, we show that

$$\beta_i^q \leq \dots \leq \beta_i^{q+p_i-1}, \quad (156)$$

which can be readily verified by noting that for $k \in \{q, \dots, q+p_i-1\}$ we have

$$\begin{aligned} \beta_i^k &= \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{Q}_i \setminus \cup_{j=1}^{k-q} \mathcal{Q}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|} \\ &\leq \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{Q}_i \setminus \cup_{j=1}^{k-q+1} \mathcal{Q}_i^j, |\mathcal{D}| \leq \mu_i} \frac{\Delta_i(\mathbf{h}_i, \mathcal{D}, \mathcal{V}, \mathbf{R})}{\|\mathbf{T} \circ \mathbb{1}_{\mathcal{D}}\|} \quad (157) \\ &= \beta_i^{k+1}, \quad (158) \end{aligned}$$

where $q \leq i \leq q+p_i-2$, and (157) holds due to the fact that $\mathcal{Q}_i \setminus \cup_{j=1}^{k-q+1} \mathcal{Q}_i^j$ is a subset of $\mathcal{Q}_i \setminus \cup_{j=1}^{k-q} \mathcal{Q}_i^j$. By recalling the definition of β_i^* from (25) we have

$$\beta_i^* \triangleq \max_{\mathcal{Q}_i \in \mathcal{Q}_i} \left\{ \min_k \{\beta_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \cup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R})\} \right\} \quad (159)$$

$$\geq \left\{ \min_k \{\beta_i(\mathbf{h}_i, \mathcal{Q}_i^k, \mathcal{K} \setminus \cup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R})\} \right\} \quad (160)$$

$$= \min_{k \in \{q, \dots, q+p_i\}} \beta_i^k \quad (161)$$

$$= \beta_i^q, \quad (162)$$

where the last step holds by invoking the order in (156). Equations (155) and (162) prove that $\beta_i^* = \beta_i^q$.

APPENDIX J PROOF OF LEMMA 8

Define $\lambda \triangleq \min_{\{i,m\}} \frac{r_{i,m}}{t_{i,m}}$. The constraint that $\mathbf{R}^n + \mathbf{r}$ should be decodable requires that

$$\lambda \sum_{(i,m) \in \mathcal{D}} t_{i,m} \leq \sum_{(i,m) \in \mathcal{D}} r_{i,m} \leq \Delta_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}, \mathbf{R}^n),$$

$\forall \emptyset \neq \mathcal{D} \subseteq \mathcal{U}$, or equivalently,

$$\lambda \leq \lambda' \triangleq \min_{\mathcal{D} \neq \emptyset, \mathcal{D} \subseteq \mathcal{U}} \frac{\mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{D}, \mathcal{V}) - \|\mathbf{R}^n \circ \mathbf{1}_D\|}{\|\mathbf{T} \circ \mathbf{1}_D\|}.$$

Also, it can be readily verified that $\mathbf{R}_{\mathcal{U}}^n + \lambda' \cdot \mathbf{t}_{\mathcal{U}}$ is in the achievable rate region $\mathcal{R}_i(\mathbf{h}_i^{n+1}, \mathcal{U}, \mathcal{V})$ and, as a result, $\max \min_{\{i,m\}} \frac{r_{i,m}}{t_{i,m}} = \lambda'$. By invoking (43) the proof is complete.

APPENDIX K PROOF OF THEOREM 4

Assume that Algorithm 5 partitions the set \mathcal{K} to the disjoint sets $\{\mathcal{V}_i^1, \dots, \mathcal{V}_i^p\}$ with corresponding parameters $\{\beta_i^1, \dots, \beta_i^p\}$ such that $d \leq p-1$ is the largest number that $\forall m \in \{1, \dots, M\}$ we have $(i, m) \in \bigcup_{s=d}^{p-1} \mathcal{V}_i^{s+1}$. By taking the same approach as in Lemma 6 we can readily show that

$$\beta_i^1 \leq \dots \leq \beta_i^{s+1} \leq \dots \leq \beta_i^p. \quad (163)$$

As proposed by the algorithm, $r_{j,m}^i = +\infty$ for $(j, m) \in \bigcup_{\ell=1}^d \mathcal{V}_i^\ell$ and for $(i, m) \in \mathcal{V}_i^\ell$, where $\ell \geq d+1$, $\frac{r_{j,m}^i}{t_{j,m}} = \beta_i^\ell$. Therefore,

$$\min_{(i,m) \in \mathcal{K}} \frac{r_{j,m}^i}{t_{j,m}} = \min\{+\infty, \beta_i^{d+1}, \dots, \beta_i^p\} = \beta_i^{d+1}. \quad (164)$$

Now consider any arbitrary partitioning $\bar{\mathcal{Q}}_i = \{\mathcal{Q}_i^1, \dots, \mathcal{Q}_i^{p_i}\}$ that supports the rate increments $\{\tilde{r}_{j,m}^i\}$ and satisfies the max-min optimality. Based on the definition in (47) we have

$$\min_{(i,m) \in \mathcal{K}} \frac{\tilde{r}_{j,m}^i}{t_{j,m}} = \min_k \{\gamma_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\}, \quad (165)$$

and by invoking the result of Lemma 8 we have

$$\min_{(i,m) \in \mathcal{K}} \frac{\tilde{r}_{j,m}^i}{t_{j,m}} = \min_k \{\beta_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\}. \quad (166)$$

Recall that according to Theorem 2, $\min_k \{\beta_i(\mathbf{h}_i^{n+1}, \mathcal{Q}_i^k, \mathcal{K} \setminus \bigcup_{j=1}^{k-1} \mathcal{Q}_i^j, \mathbf{R}^n, \mathbf{T})\}$ is maximized by deploying Algorithm 3 and its respective maximum value is β_i^{d+1} , which is the same in Algorithm 3 and Algorithm 5 due to their similarities in constructing the sets $\{\mathcal{V}_i^k\}$ as well as computing parameters $\{\beta_i^k\}$. Therefore,

$$\min_{(i,m) \in \mathcal{K}} \frac{\tilde{r}_{j,m}^i}{t_{j,m}} \leq \beta_i^{d+1}. \quad (167)$$

Equations (164) and (167) together establish the desired result.

APPENDIX L PROOF OF THEOREM 5

- 1) Since \mathbf{R}^{\min} is decodable, as an straightforward application of Theorem 4 we find that $\mathbf{R}^{(1)}$ is also decodable and $\mathbf{R}^{(1)} \succeq \mathbf{R}^{\min}$. In general, at the $(q+1)^{th}$ iteration for finding the rate matrix $\mathbf{R}^{(q+1)}$ we have set $\mathbf{R}^{\min} = \mathbf{R}^{(q)}$ where again by using Theorem 4 we conclude that $\mathbf{R}^{(q+1)}$ is decodable and $\mathbf{R}^{(q+1)} \succeq \mathbf{R}^{(q)}$. Finally, as the set of rate matrices $\{\mathbf{R}^{(q+1)}\}$ is monotonically increasing and the set of decodable rate matrices is bounded, the convergence is guaranteed.
- 2) By invoking $\mathbf{R}^{(q)} \succeq \dots \succeq \mathbf{R}^{(1)}$ from the first part we get

$$\min_{(i,m) \in \mathcal{K}} \frac{R_{i,m}^{(q)} - R_{i,m}^{\min}}{t_{i,m}} \geq \min_{(i,m) \in \mathcal{K}} \frac{R_{i,m}^{(1)} - R_{i,m}^{\min}}{t_{i,m}}. \quad (168)$$

Now, for the given rate matrix $\tilde{\mathbf{R}}$ let us define $\tilde{r}_{i,m}^j \triangleq \tilde{R}_{i,m} - R_{i,m}^{\min}$ for $j = 1, \dots, K$. By noting that $R_{i,m}^{(1)} = R_{i,m}^{\min} + \min_{1 \leq j \leq K} \{r_{i,m}^j\}$, where $\{r_{i,m}^j\}$ are the rate recommendations made after the first iteration, we get

$$\begin{aligned} & \min_{(i,m) \in \mathcal{K}} \frac{R_{i,m}^{(1)} - R_{i,m}^{\min}}{t_{i,m}} \\ &= \min_{(i,m) \in \mathcal{K}} \frac{\min_{1 \leq j \leq K} \{r_{i,m}^j\}}{t_{i,m}} \\ &= \min_{1 \leq j \leq K} \min_{(i,m) \in \mathcal{K}} \frac{r_{i,m}^j}{t_{i,m}} \\ &\geq \min_{1 \leq j \leq K} \min_{(i,m) \in \mathcal{K}} \frac{\tilde{r}_{i,m}^j}{t_{i,m}} \end{aligned} \quad (169)$$

$$\begin{aligned} &= \min_{(i,m) \in \mathcal{K}} \min_{1 \leq j \leq K} \frac{\tilde{r}_{i,m}^j}{t_{i,m}} \\ &= \min_{(i,m) \in \mathcal{K}} \min_{1 \leq j \leq K} \frac{R_{i,m} - R_{i,m}^{\min}}{t_{i,m}} \\ &= \min_{(i,m) \in \mathcal{K}} \frac{R_{i,m} - R_{i,m}^{\min}}{t_{i,m}}, \end{aligned} \quad (170)$$

where (169) holds due to Theorem 4. By putting together (168) and (170) the desired result is established.

- 3) We denote the output of Algorithm 6 by $\mathbf{R}^* = \lim_{q \rightarrow \infty} \mathbf{R}^{(q)}$ and show that for this rate allocation, any increase in the rate of any user will incur a decrease in the rate of some other user and therefore, \mathbf{R}^* is the pareto-optimal solution. For this purpose, we investigate the possibility of increasing the rate of a specific codebook while keeping the others' unchanged. Without loss of generality we examine whether the rate matrix $\tilde{\mathbf{R}} = \mathbf{R}^* + \varepsilon \mathbf{1}_{\mathcal{U}}$ where $\mathcal{U} = \{1, 1\}$ is decodable for some $\varepsilon > 0$.

At each iteration, each specific codebook of each user receives rate increment suggestions by all other users, among which the user with the lowest rate increment suggestion identifies the rate increment for that specific codebook. At the final iteration, let us assume that

the lowest rate increment recommendation for the first codebook of user 1 is made by the i^{th} user, i.e., $r_{1,1}^i = \min_j r_{1,1}^j = 0$. Also, let $\{\mathcal{V}_i^1, \dots, \mathcal{V}_i^m, \mathcal{V}_i^{m+1}, \dots, \mathcal{V}_i^p\}$ denote the sets found for the i^{th} user in the last iteration of Algorithm 6, using \mathbf{R}^* as the minimum rate matrix and denote their respective values by $\{\beta_i^1, \dots, \beta_i^d, \beta_i^{d+1}, \dots, \beta_i^p\}$. Suppose $d \leq p-1$ is the largest number that $\forall m \in \{1, \dots, M\}$ we have $(i, m) \in \bigcup_{s=d}^{p-1} \mathcal{V}_i^{s+1}$ and since all the codebooks of the i^{th} user must be decodable, we must have $\beta_i^{d+1} \geq 0$. Also recall that $\beta_i^1 \leq \beta_i^2 \leq \dots \leq \beta_i^p$. Based on this observation we can deduce the following properties for the sets $\{\mathcal{V}_i^\ell\}$ and $\{\beta_i^\ell\}$:

- a) $\beta_i^{d+2} > 0$: Clearly when $\beta_i^{d+1} > 0$ we must have $\beta_i^{d+2} > 0$. Now suppose $\beta_i^{d+1} = 0$ so that $\beta_i^{d+2} \geq 0$. Assume $\beta_i^{d+2} = 0$. Then, since $\Delta_i(\mathbf{h}_i, \mathcal{V}_i^{d+2}, (\bigcup_{j=1}^{d+1} \mathcal{V}_i^j), \mathbf{R}^*) = \Delta_i(\mathbf{h}_i, \mathcal{V}_i^{d+1}, (\bigcup_{j=1}^d \mathcal{V}_i^j), \mathbf{R}^*) = 0$ it can be shown that $\Delta_i(\mathbf{h}_i, \mathcal{V}_i^{d+2}, (\bigcup_{j=1}^d \mathcal{V}_i^j), \mathbf{R}^*) = 0$. This is a contradiction since it implies that in Algorithm 6, line 4, we could have chosen \mathcal{V}_i^{d+2} instead of \mathcal{V}_i^{d+1} . Thus, $\beta_i^{d+2} > \beta_i^{d+1} \geq 0$.
- b) $(1, 1) \in \mathcal{V}_i^{d+1}$: First, $(1, 1) \notin \mathcal{V}_i^j$ for $j \leq d$ since otherwise the i^{th} user would recommend $r_{1,1}^i = +\infty$ which is a contradiction. On the other hand, if $\exists m : (i, m) \in \mathcal{V}_i$ for $j \geq d+2$ as the i^{th} user would recommend the rate increment $\beta_i^j t_{1,1} > 0$ which is also a contradiction.
- c) $\beta^{d+1} = 0$: Since $(1, 1) \in \mathcal{V}_i^{d+1}$, due to \mathbf{R}^* being the convergence point, β^{d+1} cannot be greater than zero as otherwise it leads to a higher rate for the first codebook of the 1^{st} user. By taking into account that $\beta^{d+1} \geq 0$ we get $\beta^{d+1} = 0$.

As argued in Theorem 4, we have $\beta^1 \leq \dots \leq \beta^p$. Now, define $c \in \{1, \dots, d\}$ such that $\beta^1 \leq \dots \leq \beta^c < 0$ and $\beta^{c+1} \leq \dots \leq \beta^{d+1} = 0$ and construct the sets

$$\begin{aligned} \mathcal{D}^- &\triangleq \mathcal{V}_i^1 \cup \dots \cup \mathcal{V}_i^c, \\ \mathcal{D}^0 &\triangleq \mathcal{V}_i^{c+1} \cup \dots \cup \mathcal{V}_i^{d+1}, \\ \mathcal{D}^+ &\triangleq \mathcal{V}_i^{d+2} \cup \dots \cup \mathcal{V}_i^p. \end{aligned}$$

Recall that $\tilde{\mathbf{R}} \geq \mathbf{R}^*$. Consequently, it follows that no codebook with index in \mathcal{D}^- can be decoded at receiver i , under the rate assignment $\tilde{\mathbf{R}}$. Thus, the codebooks in \mathcal{D}^- must be treated as Gaussian interferers. Next since the rates of the signals in \mathcal{D}^+ remain unaltered, these codebooks are decodable using the partition $\{\bigcup_{j=d+2}^p \mathcal{V}_i^j, \mathcal{K} \setminus \bigcup_{j=d+2}^p \mathcal{V}_i^j\}$ under the rate assignment $\tilde{\mathbf{R}}$. Thus, without loss of optimality, we can assume that these users have been perfectly decoded and expurgated. In the following, for simplicity we define the partitioning operator $\cup_k [X]_k$ for any set X .

Let us focus on any arbitrary partitioning of sets $\{\tilde{\mathcal{G}}, \mathcal{D}^- \cup \mathcal{D}^0 \setminus \tilde{\mathcal{G}}\}$ and $\cup_k [\tilde{\mathcal{G}}]_k$ such that $(i, m) \in \tilde{\mathcal{G}}$ for all $m \in \{1, \dots, M\}$. Our objective is to show that at

least one of the codebooks of user i is not decodable under the rate assignment $\tilde{\mathbf{R}}$ using any such partition. First consider the case $(1, 1) \notin \tilde{\mathcal{G}}$. For all the codebooks of user i to be decodable we must have

$$\begin{aligned} &\min_k \left\{ \Delta_i(\mathbf{h}_i, [\tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup \mathcal{D}^0 \setminus \bigcup_{j=1}^k [\tilde{\mathcal{G}}]_j, \tilde{\mathbf{R}}) \right\} \\ &= \min_k \left\{ \Delta_i(\mathbf{h}_i, [\tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup \mathcal{D}^0 \setminus \bigcup_{j=1}^k [\tilde{\mathcal{G}}]_j, \mathbf{R}^*) \right\} \geq 0. \end{aligned}$$

Using the fact that $\beta^{c+1} = 0$, we can conclude that for any partitioning $\cup_k [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_k$ we have

$$\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_j, \mathbf{R}^*) \right\} \geq 0.$$

However, since for any partitioning $\cup_k [\mathcal{D}^0]_k$ we have

$$\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0]_j, \mathbf{R}^*) \right\} = 0,$$

we must have that both

$$\min_k \left\{ \Delta_i(\mathbf{h}_i, [\tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup \mathcal{D}^0 \setminus \bigcup_{j=1}^k [\tilde{\mathcal{G}}]_j, \mathbf{R}^*) \right\}$$

and

$$\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_j, \mathbf{R}^*) \right\}$$

are equal to zero. Again using the fact that we have $\beta^{c+1} = 0$, it is concluded that for any partitioning $\cup_k [\mathcal{D}^0 \setminus \tilde{\mathcal{G}} \cup (\bigcup_{j=c+1}^d \mathcal{V}_i^j)]_k$ we have

$$\begin{aligned} &\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0 \setminus \tilde{\mathcal{G}} \cup (\bigcup_{j=c+1}^d \mathcal{V}_i^j)]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0 \setminus \tilde{\mathcal{G}} \cup (\bigcup_{j=c+1}^d \mathcal{V}_i^j)]_j, \mathbf{R}^*) \right\} \\ &\geq 0. \end{aligned} \tag{171}$$

However, since for any partitioning $\cup_k [\bigcup_{j=c+1}^d \mathcal{V}_i^j]_k$ we have

$$\begin{aligned} &\min_k \left\{ \Delta_i(\mathbf{h}_i, [\bigcup_{j=c+1}^d \mathcal{V}_i^j]_k, \mathcal{D}^- \cup_{j>k} [\bigcup_{j=c+1}^d \mathcal{V}_i^j]_j, \mathbf{R}^*) \right\} \\ &= 0, \end{aligned} \tag{172}$$

and

$$\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_j, \mathbf{R}^*) \right\} = 0, \tag{173}$$

we immediately have

$$\begin{aligned} &\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0 \setminus \tilde{\mathcal{G}} \cup (\bigcup_{j=c+1}^d \mathcal{V}_i^j)]_k, (\bigcup_{j=c+1}^d \mathcal{V}_i^j) \right. \\ &\quad \left. \cup \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0 \setminus \tilde{\mathcal{G}} \cup (\bigcup_{j=c+1}^d \mathcal{V}_i^j)]_j, \mathbf{R}^*) \right\} \\ &= 0. \end{aligned} \tag{174}$$

This yields us the desired contradiction since the partitioned set $\cup_k [\mathcal{D}^0 \setminus \tilde{\mathcal{G}} \cup (\bigcup_{j=c+1}^d \mathcal{V}_i^j)]_k$ does not contain (i, m) for all $m \in \{1, \dots, M\}$ but was not selected instead of \mathcal{V}_i^{d+1} in line 4 of Algorithm 5. Consequently, we conclude that the codebooks of user i is not decodable using the partitioning $\{\tilde{\mathcal{G}}, \mathcal{D}^- \cup \mathcal{D}^0 \setminus \tilde{\mathcal{G}}\}$ and $\cup_k [\tilde{\mathcal{G}}]_k$ under rate assignment \mathbf{R}^* and hence under rate assignment $\tilde{\mathbf{R}}$.

Finally, we need to rule out partitions $\{\tilde{\mathcal{G}}, \mathcal{D}^- \cup \mathcal{D}^0 \setminus \tilde{\mathcal{G}}\}$ and $\cup_k [\tilde{\mathcal{G}}]_k$ such that $(1, 1) \in \tilde{\mathcal{G}}$ and $(i, m) \in \tilde{\mathcal{G}}$ for all

$m \in \{1, \dots, M\}$. For user i to be decodable, we must have that

$$\min_k \left\{ \Delta_i(\mathbf{h}_i, [\tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup \mathcal{D}^0 \setminus \bigcup_{j=1}^k [\tilde{\mathcal{G}}]_j, \tilde{\mathbf{R}}) \right\} \geq 0.$$

Using the facts that $\beta^{c+1} = 0$ and $(1, 1) \notin \mathcal{D}^0 \setminus \tilde{\mathcal{G}}$, we conclude that

$$\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0 \setminus \tilde{\mathcal{G}}]_j, \mathbf{R}^*) \right\} \geq 0.$$

These facts collectively provide that $\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0]_j, \mathbf{R}^*) \right\} \geq 0$. However, this is a contradiction since $\min_k \left\{ \Delta_i(\mathbf{h}_i, [\mathcal{D}^0]_k, \mathcal{D}^- \cup_{j>k} [\mathcal{D}^0]_j, \mathbf{R}^*) \right\} = 0$ and $\tilde{\mathbf{R}} \geq \mathbf{R}^*$ with $\tilde{R}_{1,1} > R_{1,1}^*$.

REFERENCES

- [1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math., Statist. Probab.*, Sep. 1961, pp. 611–644.
- [2] 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.
- [3] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K -user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [4] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.
- [5] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
- [6] A. S. Motahari and A. K. Khandani, "To decode the interference or to consider it as noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1274–1283, Mar. 2011.
- [7] F. Baccelli, A. El Gamal, and D. N. C. Tse, "Interference networks with point-to-point codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2582–2596, May 2011.
- [8] B. Bandemer, A. E. Gamal, and Y.-H. Kim, "Optimal achievable rates for interference networks with random codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6536–6549, Dec. 2015.
- [9] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.
- [10] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, Jan. 2013.
- [11] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of K -user SISO interference and X channels with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6542–6561, Oct. 2013.
- [12] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6093–6130, Oct. 2014.
- [13] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4396–4417, Jul. 2012.
- [14] Q. Li, K. H. Li, and K. C. Teh, "Diversity-multiplexing tradeoff of fading interference channels with source cooperation and partial CSIT," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2781–2804, May 2011.
- [15] T. T. Kim and M. Skoglund, "Diversity-multiplexing tradeoff in MIMO channels with partial CSIT," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2743–2759, Aug. 2007.
- [16] F. Brah, L. Vandendorpe, and J. Louveaux, "Constrained resource allocation in OFDMA downlink systems with partial CSIT," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, May 2008, pp. 4144–4148.
- [17] R. K. Farsani, "Capacity bounds for wireless ergodic fading broadcast channels with partial CSIT," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 927–931.
- [18] M. Kobayashi, M. Debbah, and J.-C. Belfiore, "Outage efficient strategies for network MIMO with partial CSIT," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jul. 2009, pp. 249–253.
- [19] S. A. Jafar, "Exploiting channel correlations—Simple interference alignment schemes with no CSIT," in *Proc. IEEE Global Telecommun. Conf.*, Miami, FL, USA, Dec. 2010, pp. 1–5.
- [20] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5354–5374, Aug. 2012.
- [21] A. S. Motahari and A. K. Khandani, "M-user Gaussian interference channels: To decode the interference or to consider it as noise," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 2196–2200.
- [22] A. Tajer, N. Prasad, and X. Wang, "Beamforming and rate allocation in MISO cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 362–377, Jan. 2010.
- [23] M. K. Varanasi, "Group detection for synchronous Gaussian code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1083–1096, Jul. 1995.
- [24] N. Prasad, G. Yue, X. Wang, and M. Varanasi, "Optimal successive group decoders for MIMO multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4821–4826, Nov. 2008.
- [25] C. Gong, A. Tajer, and X. Wang, "Interference channel with constrained partial group decoding," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3059–3071, Nov. 2011.
- [26] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.
- [27] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.
- [28] A. Carleial, "A case where interference does not reduce capacity (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 21, no. 5, pp. 569–570, Sep. 1975.
- [29] S. Sridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Capacity of symmetric K -user Gaussian very strong interference channels," in *Proc. IEEE Global Telecommun. Conf.*, New Orleans, LA, USA, Nov. 2008, pp. 1–5.
- [30] N. Prasad and X. Wang, "Outage minimization and rate allocation for the multiuser Gaussian interference channels with successive group decoding," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5540–5557, Dec. 2009.
- [31] M. A. Maddah-Ali, H. Mahdavi-Doost, and A. K. Khandani, "Optimal order of decoding for max-min fairness in K -user memoryless interference channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 2866–2870.
- [32] A. Schrijver, "A combinatorial algorithm minimizing submodular functions in strongly polynomial time," *J. Combinat. Theory*, vol. 80, no. 2, pp. 346–355, Nov. 2000.
- [33] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

Mehdi Ashraphijuo received the B.Sc. and M.Sc. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2010 and 2012, respectively. He is currently working towards the Ph.D. degree in the Department of Electrical Engineering, Columbia University, New York, NY. His research interests lie in the general areas of network information theory and statistical signal processing with applications in wireless communication. His current research focuses primarily on developing fundamental principles for communication network design, with emphasis on developing interference management and understanding the role of feedback and user cooperation. He was the winner of Qualcomm Innovation Fellowship in 2014.

Ali Tajer (S'05–M'10–SM'15) is an Assistant Professor of Electrical, Computer, and Systems Engineering at Rensselaer Polytechnic Institute. During 2007–2010 he was with Columbia University where he received the M.A. degree in Statistics and Ph.D. degree in Electrical Engineering, and during 2010–2012 he was with Princeton University as a Postdoctoral Research Associate. His research interests include mathematical statistics and network information theory, with applications in wireless communications and power grids. Dr. Tajer Serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Editor for the IEEE TRANSACTIONS ON SMART GRID, and as the Guest Editor-in-Chief for the IEEE TRANSACTIONS ON SMART GRID, Special Issue on Theory of Complex Systems with Applications to Smart Grid Operations. He is a senior member of the IEEE and received the NSF CAREER award in 2016.

Chen Gong received the B.S. degree in electrical engineering and mathematics (minor) from Shanghai Jiaotong University, Shanghai, China in 2005, the M.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2008, and the Ph.D degree from Columbia University, New York City, NY, USA, in 2012. He was a Senior Systems Engineer at Qualcomm Research, San Diego, CA, USA, between April 2012 and December 2013. Now, he is on the faculty of the University of Science and Technology of China (USTC). His research interests falls into the general area of wireless communications, optical wireless communications, and signal processing. He was selected by the “Young 1000 Talent Program” of China Government in 2014.

Xiaodong Wang (S'98–M'98–SM'04–F'08) received the Ph.D degree in Electrical Engineering from Princeton University. He is a Professor of Electrical Engineering at Columbia University in New York. Dr. Wang's research interests fall in the general areas of computing, signal processing and communications, and has published extensively in these areas. Among his publications is a book entitled *Wireless Communication Systems: Advanced Techniques for Signal Reception*, published by Prentice Hall in 2003. His current research interests include wireless communications, statistical signal processing, and genomic signal processing. Dr. Wang received the 1999 NSF CAREER Award, the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award, and the 2011 IEEE Communication Society Award for Outstanding Paper on New Communication Topics. He has served as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON INFORMATION THEORY. He is a Fellow of the IEEE and listed as an ISI Highly-cited Author