IEEE Access
The journal for rapid open access publishing

Received September 18, 2015, accepted November 19, 2015, date of publication December 7, 2015, date of current version January 6, 2016.

Digital Object Identifier 10.1109/ACCESS.2015.2506261

Fairness-Aware Non-Orthogonal Multi-User Access With Discrete Hierarchical Modulation for 5G Cellular Relay Networks

MEGUMI KANEKO, (Member, IEEE), HIROFUMI YAMAURA, YOUHEI KAJITA, KAZUNORI HAYASHI, (Member, IEEE), AND HIDEAKI SAKAI, (Life Fellow, IEEE)
Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan

Corresponding author: M. Kaneko (meg@i.kyoto-u.ac.jp)

This work was supported in part by the Japan Society for the Promotion of Science within the Grants-in-Aid for Scientific Research through the Ministry of Education, Science, Sports, and Culture of Japan under Grant 23760334, Grant 26820143, Grant 15H2252, and Grant 15K06064, and in part by the Telecommunications Advancement Foundation.

ABSTRACT Non-orthogonal multiple access based on superposition coding (SC) for 5G cellular systems without relays is gaining increasing interest from both academia and industries. Since relay stations will be an integral part of future cellular networks, we propose evolved non-orthogonal multi-access schemes for both direct and relayed users. Our schedulers are built upon SC-relaying schemes with practical discrete hierarchical modulations (HMs), where the messages of two selected users are superposed into different HM layers, each layer being allocated an optimized amount of power and bearing a message flow to be decoded through either a direct or relayed link. As opposed to conventional schedulers that allocate orthogonal resources to each user in wireless relaying systems, our non-orthogonal schedulers allow a pair of selected users to simultaneously share their allocated resource unit. Moreover, unlike the SC-relaying schemes in the literature based on Gaussian codebooks, the proposed schemes are designed and analyzed under the practical constraints of discrete HMs. In spite of the complexity of the power optimization under discrete HMs, we provide a simple and near-optimal power allocation method for sum-rate maximization and proportional fairness. The simulation results show that the conventional orthogonal schedulers are outperformed by the proposed schedulers in terms of sum rate and fairness, even under the practical assumption of HMs.

INDEX TERMS Cellular relay system, radio resource allocation, superposition coding, hierarchical modulation, non-orthogonal multiple access.

I. INTRODUCTION

Many recent research works have demonstrated the efficiency of relaying techniques for wireless communications such as Multi-Hop (MH) and Cooperative Diversity (CD) transmissions [1], [2], providing various advantages such as rate improvement and coverage extension with only low deployment costs [3], [4]. For the single-user relay channel, [5] has proposed a relaying scheme based on Superposition Coding (SC) where two superposed messages in the modulation domain are sent through the direct and relayed paths and recovered by Successive Interference Cancelation (SIC) at the destination, improving the achievable rates of MH and CD.

A crucial aspect of cellular relay systems concerns the design of efficient scheduling algorithms as surveyed in [6] which can be a very complex problem given the large number of possible sender to destination paths. Most existing algorithms for relayed systems as in [6]–[8] allocate an orthogonal resource, which can be in time, frequency, code or space, to each user within a cell or sector, i.e., each Base Station (BS) or Relay Station (RS) serves only a unique user per resource unit. However, the suboptimality of orthogonal allocation is an established fact, as higher spectral efficiency may be achieved with non-orthogonal allocation where multiple users are simultaneously served by one access point over each resource unit [9], [10]. This is the principle of SC with SIC, the capacity-achieving scheme in the Gaussian Broadcast Channel (GBC) [10], where the messages to multiple users are superposed with an appropriate power ratio [11]. The principle of Non-Orthogonal Multiple Access (NOMA) based on SC is gaining more and more interests not only in the research community but also in the industry, such as NTT Docomo which is actively developing the NOMA concept based on SC for the Future Radio Access in the 2020s, for 5G cellular systems without relays [12], [13]. Given the generalized deployment of relays for next generation systems, it can be naturally foreseen that NOMA based on SC for relay-aided systems will be the next major milestone. Thus, non-orthogonal SC schedulers for cellular relay systems were designed in [14]–[17], built upon a generic scheme designed for the two-user Relay Broadcast Channel (RBC) of Fig. 1 in which two users, MS_1 and MS_2, are simultaneously served by a BS with the help of a RS. In particular, the scheme proposed in [14] superposes the two users' messages into three SC layers, exploiting one high quality direct link (e.g., BS-MS_1 in Fig. 1) and the two relayed links (BS-RS-MS_1 and BS-RS-MS_2). In that scheme, power ratios are optimized over the three SC layers to maximize the sum rate [15] or Proportional Fairness (PF) criterion [16]. It was shown that these schedulers outperform conventional non-orthogonal schedulers in terms of sum-rate, outage probability and user fairness. However, all these non-orthogonal schedulers and their underlying SC relaying schemes have been designed under the assumption of Gaussian codebooks. As practical wireless relay systems make use of discrete modulations, the schemes and power allocation optimization in [14]–[17] are not applicable. Thus, feasible SC schemes are strongly required to be designed taking into account these practical constraints.

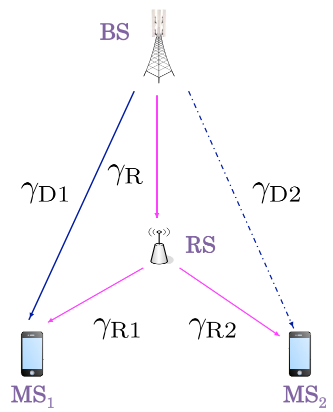

FIGURE 1. Two-user relay broadcast channel.

The aim of this work is to design SC relaying schemes and non-orthogonal schedulers making use of discrete Hierarchical Modulations (HM)s [18]–[20], which is necessary for their implementation into next generation relay-aided cellular systems. HM consists of non-uniformly spaced constellation points, providing different levels of error protection to messages superposed in the same symbol. Given the discrete nature of HMs, the whole problem setup is totally different compared to the previous schemes designed under the Gaussian codebook assumptions. This is why new scheduling design and power allocation analysis for this evolved problem setup are required.

A. RELATED WORKS ON HM-BASED RELAYING AND/OR SCHEDULING

Several works have considered the design of HM schemes for various relaying and scheduling scenarios. In [21], a coded cooperation scheme based on HM has been proposed where two users cooperate through their direct link. The problem of multi-user scheduling with HM schemes has been considered in [11]–[13] and [22], but in a downlink system without any relays. HM in the single-user relay channel was proposed in [23], with two types of HM levels, whereas [24], [25] designed HM schemes with optimized power allocation among the HM layers. For the same channel, [26] proposed an HM-based cooperative scheme for Symbol Error Rate (SER) minimization, while optimal Signal-to-Noise Ratio (SNR) threshold switching between two types of HMs was designed in [27]. Then, [28] applied the two-layer HM scheme for the two-user RBC, which is shown to outperform the scheme without relaying in terms of SER. However, no power optimization is performed. Finally, other HM-based schemes can be found for interference mitigation purposes as in [29], or for a single source/destination channel aided by multiple relays [30]. Note also that the NOMA concept by NTT Docomo [12] which is currently attracting huge research interests has only dealt with SC in cellular systems without relaying so far.

B. OUR NEW CONTRIBUTIONS

In this work, we first propose a generic two-user SC scheme based on discrete HMs, the *Three-Layer Two-User HM (3L2UHM)* scheme, that simultaneously serves two users whose messages are superposed into three discrete HM layers with specific power allocation ratios. Two different objective functions are considered in the analysis of power allocation optimization: Sum-Rate (SR) and General Proportional Fairness (GPF) maximization. Next, each of these schemes are integrated into the proposed non-orthogonal schedulers for the Downlink (DL) of a multi-user cellular relay system, namely the *Two-User Sum-Rate Maximizing (TU-SRM)* scheduler which selects the sum-rate maximizing users per scheduling frame, and the *Two-User GPF Maximizing (TU-GPFM)* scheduler that serves the user pair whose GPF metric is maximal. The proposed non-orthogonal multi-user access schemes built upon these generic three-layer HM schemes provide an additional non-orthogonal option for the scheduler compared to previous orthogonal access schemes. A low-complexity version of the *TU-GPFM* scheduler is also designed. Thus, the main advantages of our schemes and new contributions can be given as follows:

- As opposed to the previous works limited to two HM levels for relay channels, our scheme comprises three HM levels with optimized power ratios, allowing to further performance enhancement by taking full advantage of the three best link flows. Although a maximum of four link flows could be exploited in a two-user RBC, a four-level HM would lead to a very complex scheme with marginal performance gain.[1] Therefore, the proposed relaying scheme based on three-level HMs allows the best compromise between performance and feasibility.

- Power optimization among the three HM levels is performed for two different metrics: sum-rate and PF. The simulation results show that, for the two-user RBC, the proposed generic schemes outperform the rate and GPF performance compared to conventional relaying schemes.

- Unlike the previous works on HM-based relaying which mainly considered a one or two user system, our system-level simulations show the effectiveness of the proposed non-orthogonal schedulers both in terms of sum-rate and GPF, compared to the conventional orthogonal schedulers which allocate each resource unit to a unique user, while the proposed low-complexity algorithm achieves an excellent trade-off between the involved performance metrics.

- Most importantly, we have confirmed that even under the more stringent constraints of discrete HMs, non-orthogonal schedulers outperform traditional orthogonal schedulers in terms of various system level metrics, similarly to the main conclusions of previous works [15], [17] that assumed Gaussian codebooks. Thus, our results open up new perspectives towards the integration of non-orthogonal multiple access into next-generation cellular relay systems.

The sequel of the paper is organized as follows. The system model is introduced in Section II, then the reference schemes in Section III. The proposed *3L2UHM* scheme and power allocation analysis are described in Section IV, then the proposed schedulers in Section V. Numerical results are presented in Section VI. Finally, conclusions are given in Section VII.

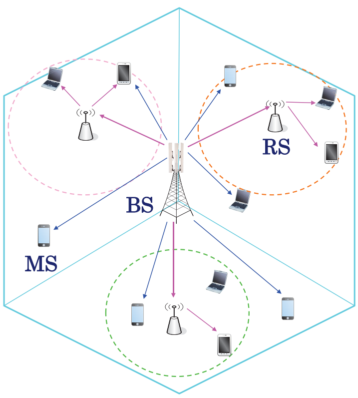

FIGURE 2. Cellular multi-user relay system.

II. SYSTEM MODEL

We consider the DL multi-user cellular relay system in Fig. 2, where each cell contains three sectors with one RS each. Scheduling is performed independently for each sector, assuming orthogonalized resources in frequency or time among sectors. BS, RS and MS are each equipped with a single antenna. The half-duplex RS works under the Decode-and-Forward (DF) protocol and SIC detector. The *direct* and *relayed* links refer to the BS-MS and BS-RS-MS links, respectively. In each scheduling frame composed of the two steps below, one or two users are served through the BS, the RS, or both depending on the relaying scheme. The two user system scheduled in each frame by the proposed non-orthogonal schedulers constitute the TU-RBC in Fig. 1.

In Step 1, the BS transmits a vector of N symbols $\boldsymbol{x} = [x(1), \ldots, x(N)]^{\mathrm{T}}$. The received signals at the RS and MS_i, $i = 1, 2$ are given by

$$\begin{aligned} y_{\mathrm{R}} &= h_{\mathrm{R}}\boldsymbol{x} + z_{\mathrm{R}}, \\ y_{\mathrm{D}i} &= h_{\mathrm{D}i}\boldsymbol{x} + z_{\mathrm{D}i}, \end{aligned} \tag{1}$$

respectively.

In Step 2, the RS transmits a vector of N_{R} symbols $\boldsymbol{x}_{\mathrm{R}} = [x_{\mathrm{R}}(1), \ldots, x_{\mathrm{R}}(N_{\mathrm{R}})]$. The received signal at MS_i, $i = 1, 2$ is given by

$$y_{\mathrm{R}i} = h_{\mathrm{R}i}\boldsymbol{x}_{\mathrm{R}} + z_{\mathrm{R}i}. \tag{2}$$

In (1)-(2), $h_I, I \in \{\mathrm{R}, \mathrm{D}i, \mathrm{R}i\}, i = 1, 2$, denote the complex channel coefficients of links BS-RS, BS-MS_i, RS-MS_i, while z_I are vectors of a circular-symmetric complex Additive White Gaussian Noise (AWGN) whose elements have mean zero and variance σ^2. The instantaneous link SNRs are defined as

$$\gamma_I = \frac{|h_I|^2}{\sigma^2}, \quad I \in \{\mathrm{R}, \mathrm{D}i, \mathrm{R}i\}, \quad i = 1, 2. \tag{3}$$

TABLE 1. Discrete modulation set.

Modulation	BPSK	QPSK	8-PAM	16-QAM	64-QAM
Modulation Level m	1	2	3	4	5
Rate $r(m)$ [bits/symbol]	1	2	3	4	6

All channel coefficients $h_I, I \in \{\mathrm{R}, \mathrm{D}i, \mathrm{R}i\}$, are assumed to be constant during each transmission time frame and to be known by the BS and the RS before transmission. We assume transmitted symbols to have mean zero $\mathrm{E}[x(n)] = \mathrm{E}[x_{\mathrm{R}}(n_{\mathrm{R}})] = 0$, and an average power of $\mathrm{E}[|x(n)|^2] = \mathrm{E}[|x_{\mathrm{R}}(n_{\mathrm{R}})|^2] = 1$. The bandwidth of transmitted signals is assumed to be normalized to 1. We consider the discrete modulations: BPSK, QPSK, 8-PAM, 16-QAM and 64-QAM as in Table 1 where $m = 1, \ldots, 5$ is the modulation level with rate $r(m)$.[2] We also define the two-layer Hierarchical QAM (HQAM): $[\mathrm{BPSK}]^2$-HQAM in Fig. 3, $[\mathrm{QPSK}]^2$-HQAM in Fig. 4 and similarly, $[8\text{-PAM}]^2$-HQAM as explained in Table 2 where $\bar{m} = 1, 2, 3$ is the HM level.

[1]Note that it was also proved in [17] that using the three best links, and hence three-layer SC scheme was sum-rate optimal.

[2]Note that we have considered the discrete modulations and HMs in Tables 1 and 2 without loss of generality, i.e., the proposed schemes may be applied to different sets of modulations.

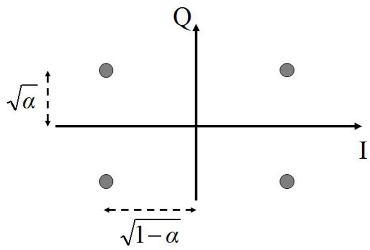

FIGURE 3. **$[\text{BPSK}]^2$-HQAM constellation.**

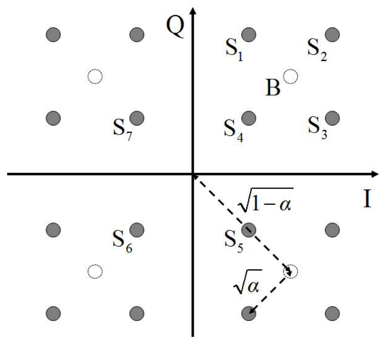

FIGURE 4. **$[\text{QPSK}]^2$-HQAM constellation.**

TABLE 2. **Discrete two-layer HM set.**

HM	$[\text{BPSK}]^2$-HQAM	$[\text{QPSK}]^2$-HQAM	$[8-\text{PAM}]^2$-HQAM
HM Level $\bar{m}$	1	2	3
Modulation of first layer	BPSK	QPSK	8-PAM
Modulation of second layer	$\pi/2$-shift BPSK	QPSK	$\pi/2$-shift 8-PAM

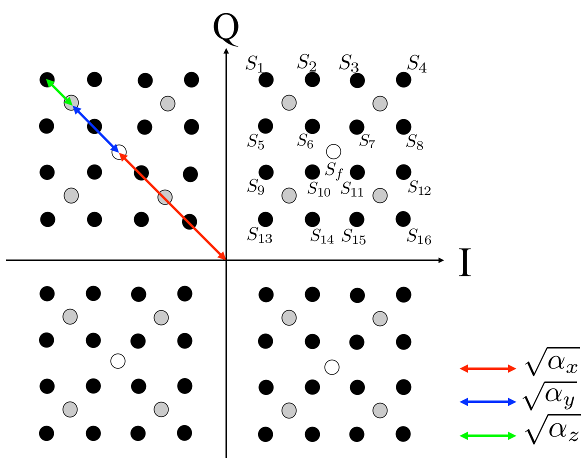

FIGURE 5. **$[\text{QPSK}]^3$-HQAM constellation.**

Moreover, we introduce $[\text{QPSK}]^3$-HQAM which is a three-layer HQAM created by superimposing three QPSK symbols, as shown in Fig. 5 where three QPSK symbols have power α_x, α_y and $\alpha_z = 1-\alpha_x-\alpha_y$ in HM layers 1, 2 and 3, respectively. We assume the following conditions for α_x and α_y so that each constellation point stays within the same quadrant as its corresponding QPSK symbol in the first layer,

$$\begin{cases} \frac{1}{2}(1-\alpha_x) + \frac{1}{2}\sqrt{\alpha_x(2-3\alpha_x)} < \alpha_y < 1-\alpha_x, \\ \qquad\qquad\qquad\qquad \text{for } \frac{1}{2} < \alpha_x < \frac{2}{3}, \\ \frac{1}{2}(1-\alpha_x) < \alpha_y < 1-\alpha_x, \\ \qquad\qquad\qquad\qquad \text{for } \frac{2}{3} \leq \alpha_x < 1. \end{cases} \tag{4}$$

Thus in Fig. 5, the hierarchical constellation points S_i $(i = 1, \ldots, 16)$ generated from the QPSK symbol S_f in the first layer stay within the first quadrant. For $\alpha_x = 16/21$ and $\alpha_y = 4/21$, $[\text{QPSK}]^3$-HQAM has a square constellation, i.e., the same one as for 64-QAM.

III. REFERENCE TRANSMISSION SCHEMES AND SCHEDULERS

We describe the reference SU and TU transmission schemes used in reference schedulers.

A. REFERENCE SU AND TU TRANSMISSION SCHEMES

1) DIRECT TRANSMISSION (DT) SCHEME

During both steps 1 and 2 described in Section II, BS transmits to an MS_i directly without any help from the RS. The rate is defined as the number of correct bits received by MS per unit symbol time, where the whole message is discarded if at least one bit in the message is not decoded correctly. The rate of DT scheme is obtained as

$$R_{\text{DT}} = r(m)\,(1 - P_1(m, \gamma_{\text{D}i}))^N,$$

where $P_1(m, \gamma)$ is the Symbol Error Rate (SER) for modulation level m and SNR γ. Defining $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-t^2)\,dt$, the SER is given by [31]

$$P_1(m, \gamma) = \begin{cases} \frac{1}{2}\text{erfc}\left(\sqrt{\gamma}\right), & m = 1, \\ 1 - \left(1 - \frac{1}{2}\text{erfc}\left(\sqrt{\frac{\gamma}{2}}\right)\right)^2, & m = 2, \\ \frac{7}{8}\text{erfc}\left(\sqrt{\frac{\gamma}{21}}\right), & m = 3, \\ 1 - \left(1 - \frac{3}{4}\text{erfc}\left(\sqrt{\frac{\gamma}{10}}\right)\right)^2, & m = 4, \\ 1 - \left(1 - \frac{7}{8}\text{erfc}\left(\sqrt{\frac{\gamma}{42}}\right)\right)^2, & m = 5. \end{cases} \tag{5}$$

2) MULTI-HOP (MH) SCHEME

Only the relayed signal is considered at MS_i, not the direct one. After RS decodes the received signal from the BS in Step 1, it forwards in Step 2 the remodulated signal given SNR $\gamma_{\text{R}i}$. BS selects modulation level m_1 in Step 1, while RS uses modulation level m_2 in Step 2. In Step 2, the RS forwards $Nr(m_1)$ bits to the MS with the rate $r(m_2)$. Thus, the duration of Step 2 is $\frac{Nr(m_1)}{r(m_2)}$. The rate of MH scheme is

$$R_{\text{MH}} = \frac{Nr(m_1)}{N + \frac{Nr(m_1)}{r(m_2)}} (1 - P_1(m_1, \gamma_{\text{R}}))^N \times (1 - P_1(m_2, \gamma_{\text{R}i}))^{\frac{Nr(m_1)}{r(m_2)}}.$$

3) COOPERATIVE DIVERSITY (CD) SCHEME

In the CD scheme of [32], different modulation rates can be used at the BS and the RS. In Step 1, the signal transmitted by the BS with modulation level m_1 is received by both RS and MS$_i$. At RS, the received signal is decoded and retransmitted to MS$_i$ in Step 2 using the modulation adapted to SNR $\gamma_{\mathrm{R}i}$. If RS uses the same modulation as the BS, MS$_i$ performs MRC of the signal received from BS in Step 1 with the one received from RS in Step 2 and decodes the combined signal. Otherwise, RS uses modulation level $m_2 \neq m_1$ and MS$_i$ decodes his message by combining the received signals from the BS and the RS by using Log-Likelihood Ratio (LLR) combining for each bit.

When $m_1 = m_2$, the duration of Step 1 is equal to that of Step 2. Thus, the rate is

$$R_{\mathrm{CD}} = \frac{r(m_1)}{2}\left(1 - P_1(m_1, \gamma_{\mathrm{R}})\right)^N \left(1 - P_1(m_1, \gamma_{\mathrm{D}i} + \gamma_{\mathrm{R}i})\right)^N.$$

When $m_1 \neq m_2$, the duration of Step 2 is $\frac{Nr(m_1)}{r(m_2)}$. In this case, the SNR of the combined signal is no longer given by the sum of the SNRs of the direct and the relayed signals. To solve this problem, we use the approximated Bit Error Rate (BER) expression $\mathrm{BER}(\gamma) = a\exp\left(-\frac{b\gamma}{k(m)}\right)$ in [33] and [34], where a and b are the fitting parameters for the BER of BPSK and $k(m)$ is a specific parameter for modulation level m which is equal to one in the case of BPSK. We obtain $k(m) = 1, 2, 21, 10, 42$ for $m = 1, 2, 3, 4, 5$, respectively, where $a = 0.268$ and $b = 1.0358$ were determined as least square solutions in logarithmic scale, respectively. It was found that the expression $\mathrm{BER} = a\exp\left(-b\left(\frac{\gamma_{\mathrm{D}i}}{k(m_1)} + \frac{\gamma_{\mathrm{R}i}}{k(m_2)}\right)\right)$ gave a very good approximation of the BER of the combined signal at the MS, giving the rate

$$R_{\mathrm{CD}} = \frac{Nr(m_1)}{N + \frac{Nr(m_1)}{r(m_2)}}\left(1 - P_1(m_1, \gamma_{\mathrm{R}})\right)^N \times \left[1 - a\exp\left(-b\left\{\frac{\gamma_{\mathrm{D}i}}{k(m_1)} + \frac{\gamma_{\mathrm{R}i}}{k(m_2)}\right\}\right)\right]^{Nr(m_1)}.$$

4) SINGLE-USER SUPERPOSITION CODING (SUSC) SCHEME

In the scheme of [24] and [25], in Step 1, BS generates $\boldsymbol{x} = \sqrt{1-\alpha}\boldsymbol{x}_{\mathrm{b}} + \sqrt{\alpha}\boldsymbol{x}_{\mathrm{s}}$ using [BPSK]^2-HQAM, [QPSK]^2-HQAM or [8-PAM]^2-HQAM, where $\alpha \in (0, 1)$ is the power allocation parameter between the basic and the superposed symbols $\boldsymbol{x}_{\mathrm{b}}$ and $\boldsymbol{x}_{\mathrm{s}}$ intended for MS. RS decodes $\boldsymbol{x}_{\mathrm{b}}$ from the received signal, treating $\sqrt{\alpha}\boldsymbol{x}_{\mathrm{s}}$ as noise. Subtracting $\sqrt{1-\alpha}\boldsymbol{x}_{\mathrm{b}}$ from the received signal, RS decodes $\boldsymbol{x}_{\mathrm{s}}$, while MS keeps the received signal from the BS in memory. In Step 2, RS transmits $\boldsymbol{x}_{\mathrm{R}}$, the remodulated signal of $\boldsymbol{x}_{\mathrm{s}}$ correctly decoded in Step 1, using the modulation in Table 1 adapted to SNR γ_{R1}. MS decodes $\boldsymbol{x}_{\mathrm{R}}$ ($\boldsymbol{x}_{\mathrm{s}}$) from the signal received from RS and cancels its contribution from the signal kept in Step 1, then decodes $\boldsymbol{x}_{\mathrm{b}}$. BS selects HM level $\bar{m}_1$ in Step 1, while RS uses level m_2 in Step 2. The rate of SUSC scheme is obtained as

$$\begin{aligned} R_{\mathrm{SUSC}} = &\frac{r(\bar{m}_1)r(m_2)}{r(\bar{m}_1) + r(m_2)}\{1 - P_2(\bar{m}_1, \gamma_{\mathrm{R}}, 1-\alpha, \alpha)\}^N \\ &\times \{1 - P_1(\bar{m}_1, \gamma_{\mathrm{R}}\alpha)\}^N \{1 - P_1(m_2, \gamma_{\mathrm{R1}})\}^{\frac{Nr(\bar{m}_1)}{r(m_2)}} \\ &\times \left[1 + \{1 - P_1(\bar{m}_1, \gamma_{\mathrm{D1}}(1-\alpha))\}^N\right]. \end{aligned}$$

In (6), $P_2(\bar{m}, \gamma, \alpha_1, \alpha_2)$ is the SER of the symbol in the first layer of two-layer HM with SNR γ, and α_1, α_2 are the power ratios allocated to the symbols in the first and second layers, respectively. The SER $P_2(\bar{m}, \gamma, \alpha_1, \alpha_2)$ is given by

$$P_2(\bar{m}, \gamma, \alpha_1, \alpha_2) = \begin{cases} \frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma\alpha_1}\right), & \text{for } \bar{m} = 1, \\ 1 - \left\{1 - \frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma\alpha_1}{2}} - \sqrt{\frac{\gamma\alpha_2}{2}}\right) - \frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma\alpha_1}{2}} + \sqrt{\frac{\gamma\alpha_2}{2}}\right)\right\}^2, & \text{for } \bar{m} = 2, \\ \frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma\alpha_1}{21}}\right), & \text{for } \bar{m} = 3. \end{cases} \tag{6}$$

5) TWO-LAYER TWO-USER SUPERPOSITION CODING (2L2USC) SCHEME

An initial two-user SC based scheme was proposed in [35] for the TU-RBC as in Fig. 1 assuming Gaussian codebooks. Here, we introduce the reference 2L2USC scheme with discrete HM based on the scheme in [35]. It simply follows SUSC by setting $\boldsymbol{x}_{\mathrm{b}} = \boldsymbol{x}_1$ destined to MS$_1$ and $\boldsymbol{x}_{\mathrm{s}} = \boldsymbol{x}_2$ destined to MS$_2$. Thus, the sum rate of 2L2USC scheme is

$$\begin{aligned} R_{\mathrm{2L2USC}} = &\frac{r(\bar{m}_1)r(m_2)}{r(\bar{m}_1) + r(m_2)}\{1 - P_2(\bar{m}_1, \gamma_{\mathrm{R}}, 1-\alpha, \alpha)\}^N \\ &\times \{1 - P_1(\bar{m}_1, \gamma_{\mathrm{R}}\alpha)\}^N \{1 - P_1(m_2, \gamma_{\mathrm{R2}})\}^{\frac{Nr(\bar{m}_1)}{r(m_2)}} \\ &\times \left[1 + \{1 - P_1(\bar{m}_1, \gamma_{\mathrm{D1}}(1-\alpha))\}^N\right]. \end{aligned}$$

B. REFERENCE SINGLE-USER SCHEDULERS

In the DL multi-user cellular relay system, conventional scheduling algorithms allocate orthogonal channels (typically, different time slots/frequency channels) to different users, where the user with the corresponding reference single user transmission scheme of Section III-A that achieves the best target metric is served in each frame. We consider sum-rate maximization and proportional fairness in the schedulers described below.

1) **Single User Sum-Rate Maximizing (*SU-SRM*) Scheduler:** For each user k, ($k = 1, \ldots, K$), the scheduler selects the SU transmission scheme among DT, MH, CD and SUSC that achieves the highest rate. Let R_k be the rate of user k in the current frame, with the selected best SU scheme. Then, the user k^* with the highest rate is selected to be scheduled in the current frame, $k^* = \arg\max_k R_k$.

2) **SU Normalized Proportional Fairness Maximizing (*SU-NPFM*) Scheduler:** as shown in [36], the PF scheduler is defined as the scheduler $\mathcal{S}$ that maximizes the GPF metric

$$\Gamma_{\mathcal{S}} = \sum_{k=1}^{K} \log R_k^{(\mathcal{S})}, \tag{7}$$

where $R_k^{(\mathcal{S})}$ denotes user k's achievable rate with scheduler $\mathcal{S}$.

In the SU case, one conventional PF scheduler [37] is the one that selects the user maximizing the Normalized PF (NPF) metric

$$\rho_k = \frac{R_k}{\bar{R}_k}, \tag{8}$$

where $\bar{R}_k$ is user k's long-term average rate given by its average SNR.

Therefore, the Reference *SU-NPFM* scheduler works as follows: after selecting the best SU transmission scheme for each user as in *SU-SRM* scheduler, it chooses the user k^* with the highest NPF metric in each frame, i.e., $k^* = \arg\max_k \rho_k$.

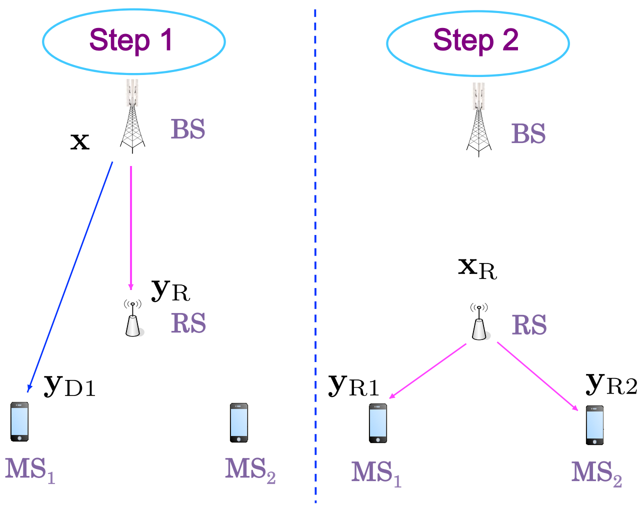

FIGURE 6. Steps of the proposed scheme.

IV. PROPOSED RELAYING SCHEMES WITH HM

A. DESCRIPTION OF THE STEPS

We describe the proposed *Three-Layer Two-User HM (3L2UHM)* for the TU-RBC in Fig. 6.

Step 1: Using $[\mathrm{QPSK}]^3$-HQAM, BS generates the three-layer hierarchically modulated signal

$$\boldsymbol{x} = \sqrt{\alpha_{\mathrm{b1}}}\boldsymbol{x}_{\mathrm{b1}} + \sqrt{\alpha_{\mathrm{s1}}}\boldsymbol{x}_{\mathrm{s1}} + \sqrt{\alpha_2}\boldsymbol{x}_2, \tag{9}$$

where $\alpha_{\mathrm{b1}}, \alpha_{\mathrm{s1}}, \alpha_2 \in [0, 1]$ denote the power allocation parameters of HM layers 1, 2 and 3, respectively, summing up to one, $\alpha_{\mathrm{b1}} + \alpha_{\mathrm{s1}} + \alpha_2 = 1$. $\boldsymbol{x}_{\mathrm{b1}}$ and $\boldsymbol{x}_{\mathrm{s1}}$ are destined to MS_1 and $\boldsymbol{x}_2$ to MS_2. The RS receives

$$\boldsymbol{y}_{\mathrm{R}} = h_{\mathrm{R}}(\sqrt{\alpha_{\mathrm{b1}}}\boldsymbol{x}_{\mathrm{b1}} + \sqrt{\alpha_{\mathrm{s1}}}\boldsymbol{x}_{\mathrm{s1}} + \sqrt{\alpha_2}\boldsymbol{x}_2) + \boldsymbol{z}_{\mathrm{R}}. \tag{10}$$

The proposed scheme may be applied whenever the link SNRs satisfy $\gamma_{\mathrm{R}} \geq \gamma_{\mathrm{R}i} \geq \gamma_{\mathrm{D}i}$, $i = 1, 2$ or, $\gamma_{\mathrm{R}} \geq \gamma_{\mathrm{R}i} \geq \gamma_{\mathrm{D}i}$ and $\gamma_{\mathrm{R}} \geq \gamma_{\mathrm{D}j} \geq \gamma_{\mathrm{R}j}$ for $i \neq j$. Therefore in the analysis, the following link SNR ordering will be assumed without loss of generality for the proposed scheme

$$\gamma_{\mathrm{D2}} \leq \gamma_{\mathrm{D1}} \leq \gamma_{\mathrm{R1}} \leq \gamma_{\mathrm{R2}} \leq \gamma_{\mathrm{R}}, \tag{11}$$

as the analysis applies in all cases by adapting the SIC decoding orders at each receiver as follows. That is, as explained in [17] and the SC scheme in [9] for broadcast channel, we allocate more power to the messages with weaker link qualities for improving user fairness. Since at the final decoding stage, $\boldsymbol{x}_{\mathrm{b1}}, \boldsymbol{x}_{\mathrm{s1}}, \boldsymbol{x}_2$ are message flows over the links with SNRs satisfying $\gamma_{\mathrm{D1}} \leq \gamma_{\mathrm{R1}} \leq \gamma_{\mathrm{R2}}$, this implies the power ratio constraint $\alpha_{\mathrm{b1}} \geq \alpha_{\mathrm{s1}} \geq \alpha_2$ for ensuring higher power to the messages on weaker links. Thus, the optimal SIC decoding order at RS is $\boldsymbol{x}_{\mathrm{b1}} \rightarrow \boldsymbol{x}_{\mathrm{s1}} \rightarrow \boldsymbol{x}_2$, since the message with highest power should be successively decoded and canceled, following the SIC principle. Thus, RS decodes $\boldsymbol{x}_{\mathrm{b1}}$ from $\boldsymbol{y}_{\mathrm{R}}$, treating the contributions of $\boldsymbol{x}_{\mathrm{s1}}$ and $\boldsymbol{x}_2$ as noise. Subtracting $\sqrt{\alpha_{\mathrm{b1}}}\boldsymbol{x}_{\mathrm{b1}}$ from $\boldsymbol{y}_{\mathrm{R}}$, we obtain $\boldsymbol{y}'_{\mathrm{R}} = \boldsymbol{y}_{\mathrm{R}} - h_{\mathrm{R}}\sqrt{\alpha_{\mathrm{b1}}}\boldsymbol{x}_{\mathrm{b1}} = h_{\mathrm{R}}(\sqrt{\alpha_{\mathrm{s1}}}\boldsymbol{x}_{\mathrm{s1}} + \sqrt{\alpha_2}\boldsymbol{x}_2) + \boldsymbol{z}_{\mathrm{R}}$, from which the RS decodes $\boldsymbol{x}_{\mathrm{s1}}$, treating $h_{\mathrm{R}}\sqrt{\alpha_2}\boldsymbol{x}_2$ as noise, giving $\boldsymbol{y}''_{\mathrm{R}} = \boldsymbol{y}'_{\mathrm{R}} - h_{\mathrm{R}}\sqrt{\alpha_{\mathrm{s1}}}\boldsymbol{x}_{\mathrm{s1}} = h_{\mathrm{R}}\sqrt{\alpha_2}\boldsymbol{x}_2 + \boldsymbol{z}_{\mathrm{R}}$. The RS finally decodes $\boldsymbol{x}_2$ from $\boldsymbol{y}''_{\mathrm{R}}$. On the other hand, MS_1 receives

$$\boldsymbol{y}_{\mathrm{D1}} = h_{\mathrm{D1}}(\sqrt{\alpha_{\mathrm{b1}}}\boldsymbol{x}_{\mathrm{b1}} + \sqrt{\alpha_{\mathrm{s1}}}\boldsymbol{x}_{\mathrm{s1}} + \sqrt{\alpha_2}\boldsymbol{x}_2) + \boldsymbol{z}_{\mathrm{D1}}, \tag{12}$$

and keeps it in memory. As the link SNR γ_{D2} is the worst, MS_2 ignores its received signal.

Step 2: The RS transmits the two-layer hierarchically modulated signal

$$\boldsymbol{x}_{\mathrm{R}} = \sqrt{1-\beta}\boldsymbol{x}_{\mathrm{R1}} + \sqrt{\beta}\boldsymbol{x}_{\mathrm{R2}}, \tag{13}$$

created with $[\mathrm{BPSK}]^2$-HQAM, $[\mathrm{QPSK}]^2$-HQAM or $[8\text{-PAM}]^2$-HQAM. $\boldsymbol{x}_{\mathrm{R1}}$ and $\boldsymbol{x}_{\mathrm{R2}}$ are the remodulated signals of $\boldsymbol{x}_{\mathrm{s1}}$ and $\boldsymbol{x}_2$, respectively. $\beta \in (0, 1)$ is the power allocation parameter. The received signal at MS_2 is

$$\boldsymbol{y}_{\mathrm{R2}} = h_{\mathrm{R2}}(\sqrt{1-\beta}\boldsymbol{x}_{\mathrm{R1}} + \sqrt{\beta}\boldsymbol{x}_{\mathrm{R2}}) + \boldsymbol{z}_{\mathrm{R2}}. \tag{14}$$

Since $\gamma_{\mathrm{R1}} \leq \gamma_{\mathrm{R2}}$, MS_2 first decodes $\boldsymbol{x}_{\mathrm{R1}}$ then $\boldsymbol{x}_{\mathrm{R2}}$ through SIC. From $\boldsymbol{y}_{\mathrm{R2}}$, MS_2 decodes $\boldsymbol{x}_{\mathrm{R1}}$ and hence $\boldsymbol{x}_{\mathrm{s1}}$. Canceling the component of $\boldsymbol{x}_{\mathrm{R1}}$ from $\boldsymbol{y}_{\mathrm{R2}}$, MS_2 obtains $\boldsymbol{y}'_{\mathrm{R2}} = \boldsymbol{y}_{\mathrm{R2}} - h_{\mathrm{R2}}\sqrt{1-\beta}\boldsymbol{x}_{\mathrm{R1}} = h_{\mathrm{R2}}\sqrt{\beta}\boldsymbol{x}_{\mathrm{R2}} + \boldsymbol{z}_{\mathrm{R2}}$, from which it decodes $\boldsymbol{x}_{\mathrm{R2}}$ ($\boldsymbol{x}_2$). Meanwhile, MS_1 receives

$$\boldsymbol{y}_{\mathrm{R1}} = h_{\mathrm{R1}}(\sqrt{1-\beta}\boldsymbol{x}_{\mathrm{R1}} + \sqrt{\beta}\boldsymbol{x}_{\mathrm{R2}}) + \boldsymbol{z}_{\mathrm{R1}}. \tag{15}$$

MS_1 decodes $\boldsymbol{x}_{\mathrm{R1}}$ from $\boldsymbol{y}_{\mathrm{R1}}$, treating $h_{\mathrm{R1}}\sqrt{\beta}\boldsymbol{x}_{\mathrm{R2}}$ as noise. Thus, MS_1 obtains $\boldsymbol{x}_{\mathrm{s1}}$ and cancels $\boldsymbol{x}_{\mathrm{s1}}$ from $\boldsymbol{y}_{\mathrm{D1}}$ kept in memory in Step 1, getting $\boldsymbol{y}'_{\mathrm{D1}} = \boldsymbol{y}_{\mathrm{D1}} - h_{\mathrm{D1}}\sqrt{\alpha_{\mathrm{s1}}}\boldsymbol{x}_{\mathrm{s1}} = h_{\mathrm{D1}}\sqrt{\alpha_{\mathrm{b1}}}\boldsymbol{x}_{\mathrm{b1}} + \boldsymbol{z}_{\mathrm{D1}}$, from which MS_1 finally decodes $\boldsymbol{x}_{\mathrm{b1}}$.

B. EXPRESSIONS FOR SUM RATE AND GPF METRIC

In this section, we derive the sum rate of the proposed *3L2UHM* scheme defined as the sum of the expected number of correctly received bits at MS_1 and MS_2 per unit symbol time. It is assumed that if there is at least one bit error in a decoded message, the whole message is discarded. Let $\bar{\boldsymbol{x}}_i^{(\mathrm{R})}$ ($i \in \{\mathrm{b1}, \mathrm{s1}, 2\}$) be the message $\boldsymbol{x}_i$ decoded at the RS. We define $Q_{\mathrm{b1}}^{(\mathrm{R})}$ as the probability of correct decoding of message $\boldsymbol{x}_{\mathrm{b1}}$ by RS, and $Q_{\mathrm{s1}}^{(\mathrm{R})}$ as the probability of correct decoding of message $\boldsymbol{x}_{\mathrm{s1}}$ given the correct decoding of $\boldsymbol{x}_{\mathrm{b1}}$ by RS, where we have denoted the conditional probability of event A given B as $\Pr(A \mid B)$. Similarly, $Q_2^{(\mathrm{R})}$ is defined as the probability of correct decoding of $\boldsymbol{x}_2$ given the correct decoding of $\boldsymbol{x}_{\mathrm{b1}}$ and $\boldsymbol{x}_{\mathrm{s1}}$ by RS, namely

$$\begin{aligned}
Q_{\mathrm{b1}}^{(\mathrm{R})} &= \Pr(\bar{\boldsymbol{x}}_{\mathrm{b1}}^{(\mathrm{R})} = \boldsymbol{x}_{\mathrm{b1}}),\\
Q_{\mathrm{s1}}^{(\mathrm{R})} &= \Pr(\bar{\boldsymbol{x}}_{\mathrm{s1}}^{(\mathrm{R})} = \boldsymbol{x}_{\mathrm{s1}} \mid \bar{\boldsymbol{x}}_{\mathrm{b1}}^{(\mathrm{R})} = \boldsymbol{x}_{\mathrm{b1}}),\\
Q_2^{(\mathrm{R})} &= \Pr(\bar{\boldsymbol{x}}_2^{(\mathrm{R})} = \boldsymbol{x}_2 \mid \bar{\boldsymbol{x}}_{\mathrm{s1}}^{(\mathrm{R})} = \boldsymbol{x}_{\mathrm{s1}}, \bar{\boldsymbol{x}}_{\mathrm{b1}}^{(\mathrm{R})} = \boldsymbol{x}_{\mathrm{b1}}).
\end{aligned}$$

Let $\bar{\boldsymbol{x}}_i^{(\mathrm{MS}_j)}$ ($i \in \{\mathrm{b1}, \mathrm{s1}, 2\}, j \in \{1, 2\}$) be the message $\boldsymbol{x}_i$ decoded at MS_j. We also define probabilities $Q_{\mathrm{R1}}^{(\mathrm{MS}_2)}$, $Q_{\mathrm{R2}}^{(\mathrm{MS}_2)}$, $Q_{\mathrm{R1}}^{(\mathrm{MS}_1)}$ and $Q_{\mathrm{b1}}^{(\mathrm{MS}_1)}$ of correct decoding of the involved messages at MS_1, MS_2 similarly as above, giving

$$\begin{aligned}
Q_{\mathrm{R1}}^{(\mathrm{MS}_2)} &= \Pr(\bar{\boldsymbol{x}}_{\mathrm{R1}}^{(\mathrm{MS}_2)} = \boldsymbol{x}_{\mathrm{R1}} \mid \bar{\boldsymbol{x}}^{(\mathrm{R})} = \boldsymbol{x}),\\
Q_{\mathrm{R2}}^{(\mathrm{MS}_2)} &= \Pr(\bar{\boldsymbol{x}}_{\mathrm{R2}}^{(\mathrm{MS}_2)} = \boldsymbol{x}_{\mathrm{R2}} \mid \bar{\boldsymbol{x}}_{\mathrm{R1}}^{(\mathrm{MS}_2)} = \boldsymbol{x}_{\mathrm{R1}}, \bar{\boldsymbol{x}}^{(\mathrm{R})} = \boldsymbol{x}),\\
Q_{\mathrm{R1}}^{(\mathrm{MS}_1)} &= \Pr(\bar{\boldsymbol{x}}_{\mathrm{R1}}^{(\mathrm{MS}_1)} = \boldsymbol{x}_{\mathrm{R1}} \mid \bar{\boldsymbol{x}}^{(\mathrm{R})} = \boldsymbol{x}),\\
Q_{\mathrm{b1}}^{(\mathrm{MS}_1)} &= \Pr(\bar{\boldsymbol{x}}_{\mathrm{b1}}^{(\mathrm{MS}_1)} = \boldsymbol{x}_{\mathrm{b1}} \mid \bar{\boldsymbol{x}}_{\mathrm{R1}}^{(\mathrm{MS}_1)} = \boldsymbol{x}_{\mathrm{R1}}, \bar{\boldsymbol{x}}^{(\mathrm{R})} = \boldsymbol{x}),
\end{aligned}$$

where $\bar{\boldsymbol{x}}^{(\mathrm{R})} = \boldsymbol{x}$ represents the correct decoding of all messages $\boldsymbol{x}_{\mathrm{b1}}$, $\boldsymbol{x}_{\mathrm{s1}}$ and $\boldsymbol{x}_2$ at RS in Step 1. Since BS uses $[\mathrm{QPSK}]^3$-HQAM in Step 1, each message $\boldsymbol{x}_i$ ($i \in \{\mathrm{b1}, \mathrm{s1}, 2\}$) is composed of $2N$ bits where N is the number of symbols per message block. Message $\boldsymbol{x}_{\mathrm{s1}}$ is correctly decoded at MS_1 if all messages $\boldsymbol{x}_{\mathrm{b1}}$, $\boldsymbol{x}_{\mathrm{s1}}$ and $\boldsymbol{x}_2$ were correctly decoded at the RS, and if $\boldsymbol{x}_{\mathrm{s1}}$ was correctly decoded at MS_1. Thus, the expected number of bits of $\boldsymbol{x}_{\mathrm{s1}}$ that MS_1 correctly decodes is $2NQ_{\mathrm{b1}}^{(\mathrm{R})}Q_{\mathrm{s1}}^{(\mathrm{R})}Q_2^{(\mathrm{R})}Q_{\mathrm{R1}}^{(\mathrm{MS}_1)}$. Similarly, the expected number of bits of $\boldsymbol{x}_{\mathrm{b1}}$ that MS_1 correctly decodes is $2NQ_{\mathrm{b1}}^{(\mathrm{R})}Q_{\mathrm{s1}}^{(\mathrm{R})}Q_2^{(\mathrm{R})}Q_{\mathrm{R1}}^{(\mathrm{MS}_1)}Q_{\mathrm{b1}}^{(\mathrm{MS}_1)}$, and the expected number of bits of $\boldsymbol{x}_2$ that MS_2 correctly decodes is $2NQ_{\mathrm{b1}}^{(\mathrm{R})}Q_{\mathrm{s1}}^{(\mathrm{R})}Q_2^{(\mathrm{R})}Q_{\mathrm{R1}}^{(\mathrm{MS}_2)}Q_{\mathrm{R2}}^{(\mathrm{MS}_2)}$. Since BS transmits N symbols in Step 1, and RS N_{R} symbols in Step 2, the total duration is $N + N_{\mathrm{R}}$ symbol times. Thus, the achievable rates of MS_1 and MS_2 can be written as

$$R_{\mathrm{MS}_1} = \frac{2N}{N+N_R} Q_{\mathrm{b1}}^{(\mathrm{R})}Q_{\mathrm{s1}}^{(\mathrm{R})}Q_2^{(\mathrm{R})}Q_{\mathrm{R1}}^{(\mathrm{MS}_1)}(1 + Q_{\mathrm{b1}}^{(\mathrm{MS}_1)}), \tag{16}$$

$$R_{\mathrm{MS}_2} = \frac{2N}{N+N_R} Q_{\mathrm{b1}}^{(\mathrm{R})}Q_{\mathrm{s1}}^{(\mathrm{R})}Q_2^{(\mathrm{R})}Q_{\mathrm{R1}}^{(\mathrm{MS}_2)}Q_{\mathrm{R2}}^{(\mathrm{MS}_2)}, \tag{17}$$

and hence the overall sum-rate as

$$\begin{aligned}
R_{\mathrm{3L2UHM}} = \frac{2N}{N+N_{\mathrm{R}}} & Q_{\mathrm{b1}}^{(\mathrm{R})}Q_{\mathrm{s1}}^{(\mathrm{R})}Q_2^{(\mathrm{R})}\\
&\times \left\{Q_{\mathrm{R1}}^{(\mathrm{MS}_1)}(1+Q_{\mathrm{b1}}^{(\mathrm{MS}_1)})+Q_{\mathrm{R1}}^{(\mathrm{MS}_2)}Q_{\mathrm{R2}}^{(\mathrm{MS}_2)}\right\}.
\end{aligned} \tag{18}$$

On the other hand, the GPF metric defined in (7) is given by

$$\Gamma_{\mathrm{3L2UHM}} = \log R_{\mathrm{MS}_1} + \log R_{\mathrm{MS}_2} = \log R_{\mathrm{MS}_1}R_{\mathrm{MS}_2}, \tag{19}$$

so maximizing Γ_{3L2UHM} is equivalent to maximizing the user rates' product $R_{\mathrm{MS}_1}R_{\mathrm{MS}_2}$,

$$\begin{aligned}
R_{\mathrm{MS}_1}R_{\mathrm{MS}_2} = &\left\{\frac{2N}{N+N_R} Q_{\mathrm{b1}}^{(\mathrm{R})}Q_{\mathrm{s1}}^{(\mathrm{R})}Q_2^{(\mathrm{R})}\right\}^2\\
&\times Q_{\mathrm{R1}}^{(\mathrm{MS}_1)}(1 + Q_{\mathrm{b1}}^{(\mathrm{MS}_1)})Q_{\mathrm{R1}}^{(\mathrm{MS}_2)}Q_{\mathrm{R2}}^{(\mathrm{MS}_2)}.
\end{aligned} \tag{20}$$

Next, we derive the expressions of each term in R_{3L2UHM} and Γ_{3L2UHM}. First, $Q_{\mathrm{b1}}^{(\mathrm{R})}$ is given by

$$Q_{\mathrm{b1}}^{(\mathrm{R})} = \{1 - P_3(\gamma_{\mathrm{R}}, \alpha_{\mathrm{b1}}, \alpha_{\mathrm{s1}})\}^N, \tag{21}$$

where P_3 is the SER of the symbol in the first layer of $[\mathrm{QPSK}]^3$-HQAM, determined as

$$\begin{aligned}
&P_3(\gamma, \alpha_{\mathrm{b1}}, \alpha_{\mathrm{s1}})\\
&= 1 - \left\{1 - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma\alpha_2}{2}}\right)\right.\\
&\qquad -\frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma\alpha_2}{2}}\right)\\
&\qquad -\frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma\alpha_2}{2}}\right)\\
&\qquad \left.-\frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma\alpha_2}{2}}\right)\right\}^2.
\end{aligned}$$

Next, after canceling $\boldsymbol{x}_{\mathrm{b1}}$ from the received signal $\boldsymbol{y}_R$ in Step 1, the resulting signal $\boldsymbol{y}'_{\mathrm{R}}$ is regarded as a $[\mathrm{QPSK}]^2$-HQAM symbol with AWGN. Thus, $Q_{\mathrm{s1}}^{(\mathrm{R})}$ is obtained as

$$Q_{\mathrm{s1}}^{(\mathrm{R})} = \{1 - P_2(2, \gamma_{\mathrm{R}}, \alpha_{\mathrm{s1}}, \alpha_2)\}^N, \tag{22}$$

where SER $P_2(\bar{m}, \gamma, \alpha_1, \alpha_2)$ was given in (6).

Then, $\boldsymbol{y}''_{\mathrm{R}}$ obtained by canceling $\boldsymbol{x}_{\mathrm{s1}}$ from $\boldsymbol{y}'_{\mathrm{R}}$ is a QPSK symbol with AWGN, giving

$$Q_2^{(\mathrm{R})} = \{1 - P_1(2, \gamma_{\mathrm{R}}\alpha_2)\}^N, \tag{23}$$

where SER $P_1(m, \gamma)$ was given in (5).

In Step 2, if RS uses HM level $\bar{m}_2$, $Q_{\mathrm{R1}}^{(\mathrm{MS}_2)}$, $Q_{\mathrm{R2}}^{(\mathrm{MS}_2)}$, $Q_{\mathrm{R1}}^{(\mathrm{MS}_1)}$ and $Q_{\mathrm{b1}}^{(\mathrm{MS}_1)}$ can be written as

$$\begin{aligned}
Q_{\mathrm{R1}}^{(\mathrm{MS}_2)} &= \{1 - P_2(\bar{m}_2, \gamma_{\mathrm{R2}}, \beta, 1-\beta)\}^{N_{\mathrm{R}}},\\
Q_{\mathrm{R2}}^{(\mathrm{MS}_2)} &= \{1 - P_1(\bar{m}_2, \gamma_{\mathrm{R2}}\beta)\}^{N_{\mathrm{R}}},\\
Q_{\mathrm{R1}}^{(\mathrm{MS}_1)} &= \{1 - P_2(\bar{m}_2, \gamma_{\mathrm{R1}}, \beta, 1-\beta)\}^{N_{\mathrm{R}}},\\
Q_{\mathrm{b1}}^{(\mathrm{MS}_1)} &= \{1 - P_2(2, \gamma_{\mathrm{D1}}, \alpha_{\mathrm{b1}}, \alpha_2)\}^{N}.
\end{aligned} \tag{24}$$

Finally, the analytical expressions of R_{3L2UHM} in (18) and Γ_{3L2UHM} in (19) are obtained by substituting these probabilities by their expressions in (21), (22), (23) and (24).

C. POWER ALLOCATION OPTIMIZATION FOR SUM-RATE MAXIMIZATION

We perform the power allocation optimization for maximizing the sum-rate (18). To achieve the global solution, all parameters α_{b1}, α_{s1} and β should be jointly optimized. However, this problem is very difficult due to the complexity and non-convexity of the sum rate expression, i.e., optimality can only be guaranteed by exhaustive search over the three dimensions, which is very time-consuming. Instead, we take a suboptimal approach where we separate the optimization in two steps, first on β and then on $(\alpha_{b1}, \alpha_{s1})$, as $(\alpha_{b1}, \alpha_{s1})$ determine the power allocation in the first step, and β that of the second step. This approach will be compared to the optimal solution obtained by exhaustive search.

1) OPTIMIZING α_{b1} AND α_{s1} FOR GIVEN β

To determine the values of α_{b1} and α_{s1} that maximize R_{3L2UHM} for given β, we differentiate the sum-rate expression in (18) with respect to α_{b1} and α_{s1}. Denoting $I \in \{b1, s1\}$, we get

$$\frac{\partial R_{3L2UHM}}{\partial \alpha_I} = \left[\frac{\partial Q_{b1}^{(R)}}{\partial \alpha_I} Q_{s1}^{(R)} Q_2^{(R)} + Q_{b1}^{(R)} \frac{\partial Q_{s1}^{(R)}}{\partial \alpha_I} Q_2^{(R)} + Q_{b1}^{(R)} Q_{s1}^{(R)} \frac{\partial Q_2^{(R)}}{\partial \alpha_I}\right] \times \left\{C_1\left(1+Q_{b1}^{(MS_1)}\right) + C_2\right\} + C_1 Q_{b1}^{(R)} Q_{s1}^{(R)} Q_2^{(R)} \frac{\partial Q_{b1}^{(MS_1)}}{\partial \alpha_I}, \tag{25}$$

where the constants $C_1 = \frac{2N}{N+N_R} Q_{R1}^{(MS_1)}$ and $C_2 = \frac{2N}{N+N_R} Q_{R1}^{(MS_2)} Q_{R2}^{(MS_2)}$ are only functions of β. In (25), the partial derivatives of $Q_{b1}^{(R)}$, $Q_{s1}^{(R)}$, $\frac{\partial Q_2^{(R)}}{\partial \alpha_I}$ and $\frac{\partial Q_{b1}^{(MS_1)}}{\partial \alpha_I}$ with respect to α_I are given by

$$\begin{aligned}
\frac{\partial Q_{b1}^{(R)}}{\partial \alpha_I} &= N\{1 - P_3(\gamma_R, \alpha_{b1}, \alpha_{s1})\}^{N-1} \times \frac{\partial P_3}{\partial \alpha_I}(\gamma_R, \alpha_{b1}, \alpha_{s1}),\\
\frac{\partial Q_{s1}^{(R)}}{\partial \alpha_I} &= N\{1 - P_2(2, \gamma_R, \alpha_{s1}, \alpha_2)\}^{N-1} \times \frac{\partial P_2}{\partial \alpha_I}(2, \gamma_R, \alpha_{s1}, \alpha_2),\\
\frac{\partial Q_2^{(R)}}{\partial \alpha_I} &= N\{1 - P_1(2, \gamma_R \alpha_2)\}^{N-1} \times \frac{\partial P_1}{\partial \alpha_I}(2, \gamma_R \alpha_2),\\
\frac{\partial Q_{b1}^{(MS_1)}}{\partial \alpha_I} &= N\{1 - P_2(2, \gamma_{D1}, \alpha_{b1}, \alpha_2)\}^{N-1} \times \frac{\partial P_2}{\partial \alpha_I}(2, \gamma_{D1}, \alpha_{b1}, \alpha_2),
\end{aligned} \tag{26}$$

where the expressions of $\frac{\partial P_3}{\partial \alpha_I}(\gamma_R, \alpha_{b1}, \alpha_{s1})$, $\frac{\partial P_2}{\partial \alpha_I}(2, \gamma_R, \alpha_{s1}, \alpha_2)$, $\frac{\partial P_1}{\partial \alpha_{b1}}(2, \gamma_R \alpha_2) = \frac{\partial P_1}{\partial \alpha_{s1}}(2, \gamma_R \alpha_2)$ and $\frac{\partial P_2}{\partial \alpha_I}(2, \gamma_{D1}, \alpha_{b1}, \alpha_2)$ are given in Appendix A due to their extensiveness.

To optimize α_{b1} and α_{s1}, we should solve the equations $\frac{\partial R_{3L2UHM}}{\partial \alpha_{b1}} = 0$ and $\frac{\partial R_{3L2UHM}}{\partial \alpha_{s1}} = 0$ but these are still difficult problems as shown in Appendix A. Instead, we will make use of the hill climbing method in Algorithm 1 to determine α_{b1} and α_{s1}, where $\frac{\partial R_{3L2UHM}}{\partial \alpha_{b1}}(a, b)$ means that $\frac{\partial R_{3L2UHM}}{\partial \alpha_{b1}}$ is evaluated at the point $(\alpha_{s1} = a, \alpha_{b1} = b)$, and similarly for $\frac{\partial R_{3L2UHM}}{\partial \alpha_{s1}}(a, b)$. The function $sgn(x)$ in Algorithm 1 is given by $sgn(x) = 1$ if $x > 0$, $sgn(x) = -1$ if $x < 0$ and else $sgn(x) = 0$. The initial power allocation parameters $(\alpha_{b1}^{(0)}, \alpha_{s1}^{(0)})$ are updated iteratively in the positive direction of the gradient of R_{3L2UHM}, improving the achievable sum rate.

Algorithm 1 Hill Climbing Algorithm

Require: Initial values $(\alpha_{b1}^{(0)}, \alpha_{s1}^{(0)})$ and step size δ
Ensure: Power allocation parameters $(\alpha_{b1}^{*}, \alpha_{s1}^{*})$

1: $(\alpha_{b1}^{*}, \alpha_{s1}^{*}) \leftarrow (\alpha_{b1}^{(0)}, \alpha_{s1}^{(0)})$
2: **repeat**
3: $\quad (a, b) \leftarrow (\alpha_{b1}^{*}, \alpha_{s1}^{*})$
4: $\quad a' \leftarrow a + \delta \cdot sgn\left(\frac{\partial R_{3L2UHM}}{\partial \alpha_{b1}}(a, b)\right)$
5: $\quad b' \leftarrow b + \delta \cdot sgn\left(\frac{\partial R_{3L2UHM}}{\partial \alpha_{s1}}(a, b)\right)$
6: $\quad (\alpha_{b1}^{*}, \alpha_{s1}^{*}) = \arg\max_{(x,y)\in\mathcal{A}} R_{3L2UHM}(x, y)$, where $\mathcal{A} = \{(a, b), (a, b'), (a', b), (a', b')\}$
7: **until** $(\alpha_{b1}^{*}, \alpha_{s1}^{*}) = (a, b)$

2) OPTIMIZING β FOR GIVEN α_{b1} AND α_{s1}

Next, we optimize β for given α_{b1} and α_{s1}. By differentiating (18) with respect to β, we get

$$\frac{\partial R_{3L2UHM}}{\partial \beta} = C_3 C_4 \frac{\partial Q_{R1}^{(MS_1)}}{\partial \beta} + C_3 \frac{\partial Q_{R1}^{(MS_1)}}{\partial \beta} Q_{R1}^{(MS_2)} + C_3 Q_{R1}^{(MS_1)} \frac{\partial Q_{R1}^{(MS_2)}}{\partial \beta}, \tag{27}$$

where $C_3 = \frac{2N}{N+N_R} Q_{b1}^{(R)} Q_{s1}^{(R)} Q_2^{(R)}$ and $C_4 = 1 + Q_{b1}^{(MS_1)}$ are only functions of α_{b1} and α_{s1}. The expression of (27) and hence the analysis will differ given the two-layer HM level used in Step 2. In the sequel, we will detail the case for [BPSK]^2-HQAM, while the analysis for the case using [QPSK]^2-HQAM will be found in Appendix B.

Differentiating R_{3L2UHM}, we get

$$\begin{aligned}
\frac{\partial R_{3L2UHM}}{\partial \beta} &= C_3 C_4 N_R \left\{1 - \frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R1}(1-\beta)}\right)\right\}^{N_R - 1} \times \left(-\frac{1}{2}\right)\sqrt{\frac{\gamma_{R1}}{\pi}}\sqrt{\frac{1}{1-\beta}}\exp(-\gamma_{R1}(1-\beta))\\
&+ C_3 N_R \sqrt{\frac{\gamma_{R2}}{\pi}} \left\{1 - \frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}(1-\beta)}\right)\right\}^{N_R - 1} \left(\frac{-1}{2}\right) \times \sqrt{\frac{1}{1-\beta}}\exp(-\gamma_{R2}(1-\beta)) \left\{1 - \frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}\beta}\right)\right\}^{N_R}\\
&+ C_3 N_R \left\{1 - \frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}(1-\beta)}\right)\right\}^{N_R} \times \sqrt{\frac{\gamma_{R2}}{\pi}} \left\{1 - \frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}\beta}\right)\right\}^{N_R - 1} \frac{1}{2}\sqrt{\frac{1}{\beta}}\exp(-\gamma_{R2}\beta).
\end{aligned} \tag{28}$$

The solution to $\frac{\partial R_{3L2UHM}}{\partial \beta} = 0$ gives a necessary condition for an optimal β. We prove that solutions for $\frac{\partial R_{3L2UHM}}{\partial \beta} = 0$ always exist in $0 < \beta < 1$. From (28), we obtain

$$\begin{aligned}
&\lim_{\beta\to 0^+} \frac{\partial R_{3L2UHM}}{\partial \beta} \\
&= -\frac{C_3 C_4 N_R}{2}\sqrt{\frac{\gamma_{R1}}{\pi}}\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R1}}\right)\right\}^{N_R-1} \\
&\quad \times\exp(-\gamma_{R1}) + \frac{C_3 N_R}{2^{N_R}}\sqrt{\frac{\gamma_{R2}}{\pi}}\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}}\right)\right\}^{N_R-1} \\
&\quad \times\left[\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}}\right)\right\}\left\{\lim_{\beta\to 0^+}\sqrt{\frac{1}{\beta}}\right\} - \frac{1}{2}\exp(-\gamma_{R2})\right], \\
&\lim_{\beta\to 1^-} \frac{\partial R_{3L2UHM}}{\partial \beta} = -\frac{C_3 C_4 N_R}{2}\sqrt{\frac{\gamma_{R1}}{\pi}}\left(\frac{1}{2}\right)^{N_R-1} \\
&\quad \times\left\{\lim_{\beta\to 1^-}\sqrt{\frac{1}{1-\beta}}\right\} \\
&\quad -\frac{C_3 N_R}{2^{N_R}}\sqrt{\frac{\gamma_{R2}}{\pi}}\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}}\right)\right\}^{N_R-1} \\
&\quad \times\left[\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\gamma_{R2}}\right)\right\}\left\{\lim_{\beta\to 1^-}\sqrt{\frac{1}{1-\beta}}\right\} - \frac{1}{2}\exp(-\gamma_{R2})\right]
\end{aligned}$$

where $\lim_{x\to c^+} f(x)$ and $\lim_{x\to c^-} f(x)$ represent the right-hand and left-hand limits of a function $f(x)$ as x approaches c. It can be seen that $\lim_{\beta\to 0^+}\frac{\partial R_{3L2UHM}}{\partial \beta} = +\infty > 0$ and $\lim_{\beta\to 1^-}\frac{\partial R_{3L2UHM}}{\partial \beta} = -\infty < 0$, thus the intermediate-value theorem guarantees the existence of at least a solution in the considered range. Although it is difficult to prove the uniqueness of the solution in (0, 1), the value β^* obtained by solving $\frac{\partial R_{3L2UHM}}{\partial \beta} = 0$ by standard numerical methods such as bisection method will be used in the proposed scheme.

3) OPTIMIZING α_{b1}, α_{s1} AND β ITERATIVELY

We adopt an iterative procedure for optimizing α_{b1}, α_{s1} and β. We initially set $\alpha_{b1}^{(0)} = 0.76$ and $\alpha_{s1}^{(0)} = 0.19$, for which the corresponding $[\mathrm{QPSK}]^3$-HQAM coincides with a uniform square 64-QAM constellation. Then, we first optimize β by determining β^* numerically as in Subsection IV-C.2. Next, fixing $\beta = \beta^*$ we optimize α_{b1} and α_{s1} using the procedure in Subsection IV-C.1, and so on. The results in Section VI will show that despite the non-convexity of the rate/GPF expressions, these initial values guarantee a near-optimal performance for a large range of SNRs. The steps are summarized in Algorithm 2.

Although it will not be developed here due to lack of space, the power ratios that maximize the GPF metric in (19) may be derived similarly. Thus, the proposed scheme with the power ratios derived for maximizing the sum-rate in (25) will be referred as the **Three-Layer Two-User HM Sum-Rate (*3L2UHM-SR*)** scheme, while the proposed scheme derived for maximizing the GPF in (19) will be referred as the **Three-Layer Two-User HM-GPF (*3L2UHM-GPF*)** scheme.

Algorithm 2 Proposed Power Allocation Optimization Algorithm

Require: Initial values $\alpha_{b1}^{(0)} = 0.76$, $\alpha_{s1}^{(0)} = 0.19$ and iteration number L

Ensure: Power allocation parameters α_{b1}^*, α_{s1}^* and β^*

1: Obtaining $\beta^{(0)}$ for given $\alpha_{b1}^{(0)}$ and $\alpha_{s1}^{(0)}$ by using the method in Subsection IV-C.2
2: **for** $i \leftarrow 1$ **to** L **do**
3: Obtaining $\alpha_{b1}^{(i)}$ and $\alpha_{s1}^{(i)}$ for given $\beta^{(i-1)}$ as in Subsection IV-C.1
4: Obtaining $\beta^{(i)}$ for given $\alpha_{b1}^{(i)}$ and $\alpha_{s1}^{(i)}$ as in Subsection IV-C.2
5: **end for**
6: $\alpha_{b1}^* \leftarrow \alpha_{b1}^{(L)}$, $\alpha_{s1}^* \leftarrow \alpha_{s1}^{(L)}$ and $\beta^* \leftarrow \beta^{(L)}$

V. PROPOSED MULTI-USER SCHEDULING ALGORITHMS

We now focus on the DL transmissions in the cellular relay system in Fig. 2, where each cell contains three sectors with one RS each, and propose sum-rate and fairness enhancing algorithms.

A. SUM-RATE MAXIMIZING SCHEDULER

We first focus on sum-rate maximization and propose a scheduler based on the Proposed *3L2UHM-SR* scheme where the scheme achieving the best rate among the *3L2UHM-SR* scheme and the reference SU schemes is chosen. Thus, one or two users may be allocated in each frame. It will be referred as the Proposed **Two User Sum-Rate Maximizing (*TU-SRM*)** scheduler.

1. For each user k ($k = 1, \ldots, K$), the scheduler selects the SU scheme achieving the highest rate among DT, MH, CD and SUSC schemes in Section III. Let R_k^{SU} be the achievable rate of the user k with its best SU scheme.
2. The scheduler selects the user k^* with the highest rate, i.e., $k^* = \arg\max_k R_k^{\mathrm{SU}}$.
3. For every user pair (i, j) where $\gamma_{Di} \geq \gamma_{Dj}$, the scheduler applies the Proposed *3L2UHM-SR* scheme to this user pair where i, j correspond to users MS_1, MS_2 in the scheme, respectively. The SIC decoding order at the RS is adapted given the SNR ordering of γ_{Di}, γ_{Ri} and γ_{Rj}, as explained in Section IV. Let $R_{i,j}^{\mathrm{TU}}$ given by (18) be the sum rate of user pair (i, j) served by the *3L2UHM-SR* scheme.
4. The scheduler obtains the user pair (i^*, j^*) with the best sum rate $R_{i^*,j^*}^{\mathrm{TU}}$,
$$(i^*, j^*) = \arg\max_{(i,j)} R_{i,j}^{\mathrm{TU}}. \tag{29}$$
5. If $R_{k^*}^{\mathrm{SU}} \leq R_{i^*,j^*}^{\mathrm{TU}}$, the scheduler allocates the current frame to the user pair (i^*, j^*) simultaneously using the Proposed *3L2UHM-SR* scheme. Otherwise, the user k^* only is served by its corresponding best-rate SU scheme.

B. PROPORTIONAL FAIR SCHEDULERS

To improve fairness, we propose the scheduler based on the *3L2UHM-GPF* scheme where the algorithm chooses the scheme giving the best GPF metric among the *3L2UHM-GPF* scheme and the SU schemes in Section III. It is referred as the Proposed **Two User GPF Maximizing (*TU-GPFM*)** scheduler.

1. For each user k ($k = 1, \ldots, K$), the scheduler selects the SU scheme achieving the highest rate among DT, MH, CD and SUSC schemes in Section III.
2. The scheduler selects the user k^* with the best NPF in (8), i.e., $k^* = \arg\max_k \rho_k^{\mathrm{SU}}$, and let $\Gamma_{k^*}^{\mathrm{SU}} = \log R_{k^*}$.
3. For every user pair (i, j) where $\gamma_{\mathrm{D}i} \geq \gamma_{\mathrm{D}j}$, the scheduler applies the *3L2UHM-GPF* where users i, j correspond to users MS_1 and MS_2 in the scheme, respectively. Let $\Gamma_{i,j}^{\mathrm{TU}}$ be the GPF of the user pair (i, j) served by *3L2UHM-GPF*.
4. The scheduler determines the pair (i^*, j^*) with the best GPF $\Gamma_{i^*,j^*}^{\mathrm{TU}}$,
$$(i^*, j^*) = \arg\max_{(i,j)} \Gamma_{i,j}^{\mathrm{TU}}. \qquad (30)$$
5. If $\Gamma_{k^*}^{\mathrm{SU}} \leq \Gamma_{i^*,j^*}^{\mathrm{TU}}$, the scheduler allocates the current frame to the pair (i^*, j^*) using the *3L2UHM-GPF*. Otherwise, k^* only is served by its corresponding SU scheme.

One drawback of the Proposed *TU-GPFM* scheduler is its high computational complexity due to exhaustive search over all user pairs, so we propose the low-complexity **Two User Normalized PF Maximizing (*TU-NPFM*)** scheduler:

1. For each user k ($k = 1, \ldots, K$), the scheduler selects the SU scheme achieving the highest rate among DT, MH, CD and SUSC schemes. Let ρ_k^{SU} be user k's NPF metric and Γ_k^{SU} its GPF metric.
2. The scheduler selects the user pair (i^*, j^*) with the best and the second best NPF metric $\rho_{i^*}^{\mathrm{SU}}$ and $\rho_{j^*}^{\mathrm{SU}}$, and applies the *3L2UHM-GPF* scheme. Their GPF metric is denoted $\Gamma_{i^*,j^*}^{\mathrm{TU}}$.
3. If $\Gamma_{k^*}^{\mathrm{SU}} \leq \Gamma_{i^*,j^*}^{\mathrm{TU}}$, the scheduler allocates the current frame to the pair (i^*, j^*) using the *3L2UHM-GPF*. Otherwise, k^* only is served by its SU scheme.

VI. NUMERICAL RESULTS

First, the proposed *3L2UHM-SR* and *3L2UHM-GPF* schemes will be evaluated assuming the TU-RBC setting. Next, the proposed schedulers in Section V will be compared to the reference SU schedulers in Section III-B in the multi-user cellular relay system.

A. TWO-USER RBC ENVIRONMENT

Fig. 7 illustrates the sum rate performance of the schemes for the given parameters and SNR values. In the TU-RBC, the reference *SU-SRM* scheme simply selects the single best-rate user for each set of SNRs. Here, it serves MS_2 by MH for $\gamma_{\mathrm{D1}} \leq 11$ dB and MS_1 by DT for $\gamma_{\mathrm{D1}} > 11$ dB. For the proposed scheme, we compare five possible power allocation methods:

- *3L2UHM-SR (cst)*: We fix $(\alpha_{\mathrm{b1}}, \alpha_{s1}) = (0.76, 0.19)$, for which the constellation of $[\mathrm{QPSK}]^3$-HQAM corresponds to a square 64-QAM constellation. Similarly, for having a square constellation, we set $\beta = 0.2$ for $[\mathrm{QPSK}]^2$-HQAM in Step 2.
- *3L2UHM-SR (opt.β)*: $(\alpha_{\mathrm{b1}}, \alpha_{s1}) = (0.76, 0.19)$ and optimal β obtained by exhaustive search.
- *3L2UHM-SR (opt.α)*: with optimal $\alpha_{\mathrm{b1}}, \alpha_{s1}$ obtained by exhaustive search and $\beta = 0.2$.
- *3L2UHM-SR (exh)*: with optimal power allocation parameters obtained by exhaustive search.
- *3L2UHM-SR (num)*: with α_{b1}^*, α_{s1}^* and β^* obtained numerically based on the iterative algorithm in Subsection IV-C.3, setting $L = 1$.

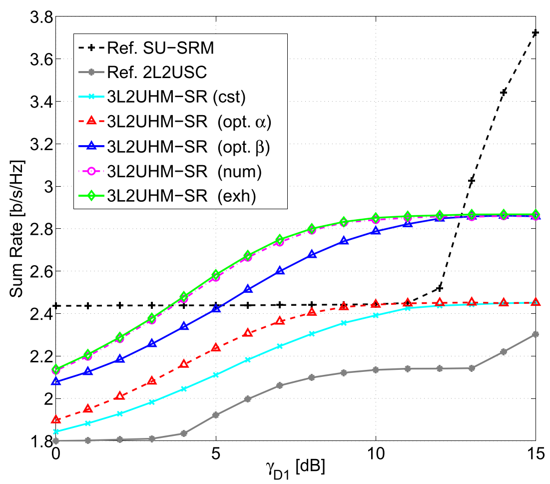

FIGURE 7. Sum rate performance of the proposed and reference schemes with $N = 12$, $\gamma_{\mathrm{R}} = 40$ dB, $\gamma_{\mathrm{R1}} = 10$ dB, $\gamma_{\mathrm{R2}} = 20$ dB.

Fig. 7 shows that the proposed *3L2UHM-SR (num)* closely approaches *3L2UHM-SR (exh)*, the optimal sum rate given by exhaustive search, showing the validity of our scheme. For these SNRs, optimizing β in *3L2UHM-SR (opt.β)* has a greater impact on the sum rate compared to optimizing α_{b1} and α_{s1} in *3L2UHM-SR (opt.α)*. This is due to the low value of γ_{R1}, so that β should be optimized to reduce the decoding errors for $\boldsymbol{x}_{\mathrm{s1}}$. The large gains of *3L2UHM-SR (num)* over *3L2UHM-SR (cst)* illustrate the necessity of the proposed power optimization. Moreover, the proposed scheme always outperforms the reference *2L2USC*, as *3L2UHM-SR (num)* enables to make efficient use of the high quality relayed link to MS_1, thanks to the three-layered HM signal from BS. Compared to *SU-SRM*, our scheme achieves a higher rate in the region $3.7 \leq \gamma_{\mathrm{D1}} \leq 12.7$ dB. Thus, the best strategy here is to use MH to MS_2 for $0 < \gamma_{\mathrm{D1}} < 3.7$ dB, the proposed *3L2UHM-SR (num)* for $3.7 \leq \gamma_{\mathrm{D1}} \leq 12.7$ dB, then DIR to MS_1 for $12.7\ dB < \gamma_{\mathrm{D1}}$ with high modulation levels (16-QAM and 64-QAM) given the high direct link quality.

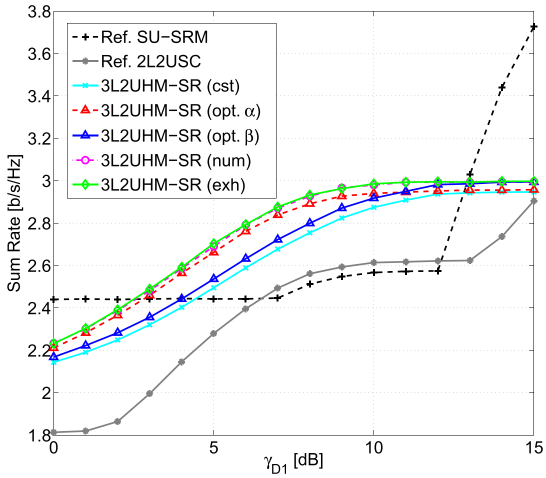

FIGURE 8. Sum rate of the proposed and reference schemes with $N = 12$, $\gamma_R = 40$ dB, $\gamma_{R1} = 20$ dB, $\gamma_{R2} = 15$ dB.

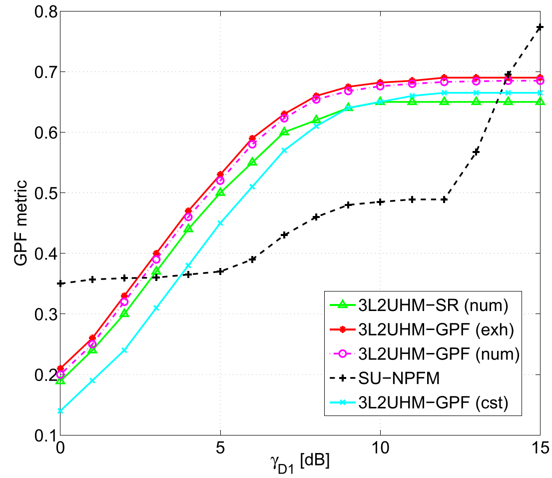

FIGURE 10. GPF metric for proposed and reference schemes with PF with $N = 12$, $\gamma_R = 40$ dB, $\gamma_{R1} = 20$ dB, $\gamma_{R2} = 15$ dB.

Fig. 8 shows the sum rate performance for a different set of SNRs. Again, *3L2UHM-SR (num)* closely matches the optimal performance *3L2UHM-SR (exh)*. In this case, since both γ_{R1} and γ_{R2} have higher values, we get a larger impact by optimizing α_{b1} and α_{s1} than β, as shown by *3L2UHM-SR (opt.α)* and *3L2UHM-SR (opt.β)*. Here, the reference *SU-SRM* serves MS_1 by CD for $\gamma_{D1} < 12.9$ dB, then by DT for 12.9 dB $< \gamma_{D1}$. Moreover, *3L2UHM-SR (num)* outperforms *SU-SRM* over $3.5 \leq \gamma_{D1} \leq 12.9$ dB, so the best strategy is CD to MS_1 for $0 < \gamma_{D1} < 3.5$ dB, *3L2UHM-SR (num)* for $3.5 \leq \gamma_{D1} \leq 12.9$ dB, and then DT to MS_1 for $\gamma_{D1} > 12.9$ dB.

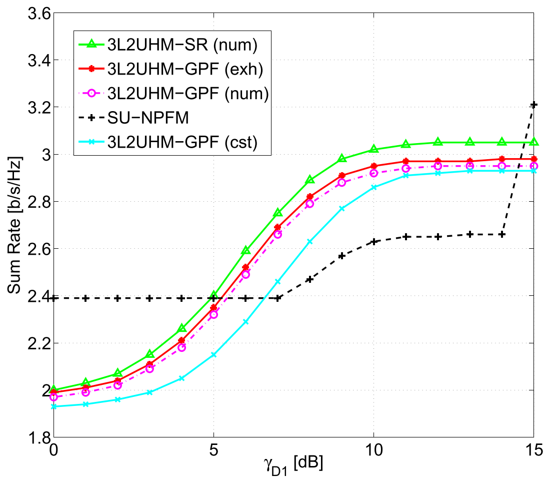

FIGURE 9. Sum rate of proposed and reference schemes with PF with $N = 12$, $\gamma_R = 40$ dB, $\gamma_{R1} = 20$ dB, $\gamma_{R2} = 15$ dB.

Next, the Proposed *3L2UHM-GPF* scheme is evaluated for $N = 12$, $\gamma_R = 40$ dB, $\gamma_{R1} = 20$ dB, $\gamma_{R2} = 15$ dB in Figs. 9 and 10 in terms of sum-rate and GPF metric, respectively. The benchmark *SU-NPFM* scheme selects the user with the best NPF metric among the two users, for each set of SNRs. Proposed *3L2UHM-GPF (num)* achieves the sum rate and GPF very close to the optimal ones given by *3L2UHM-GPF (exh)*. Compared to the proposed scheme designed for sum-rate maximization *3L2UHM-SR (num)*, *3L2UHM-GPF (num)* achieves a slightly lower sum-rate but larger GPF metric, showing the effectiveness of the proposed GPF maximizing power allocation. Moreover, *3L2UHM-GPF (num)* significantly outperforms that with constant power ratios *3L2UHM-GPF (cst)*, both in terms of sum-rate and GPF metric. For GPF, *3L2UHM-GPF (num)* outperforms *SU-NPFM* over $2.7 \leq \gamma_{D1} \leq 13.9$ dB. Thus, for GPF maximization, one should use the CD scheme to MS_1 for $0 < \gamma_{D1} < 2.7$ dB, *3L2UHM-GPF (num)* for $3.5 \leq \gamma_{D1} \leq 13.9$ dB, then the DT scheme to MS_1 for 13.9 dB $< \gamma_{D1}$.

These simulations have confirmed the validity of the proposed method in Algorithm 2, as it achieves near-optimal performance over a large range of SNRs despite the complexity and non-convexity of the sum-rate/GPF expressions. In addition, the proposed method considerably reduces the required computational complexity compared to exhaustive search, as a near-optimal performance with only $L = 1$ iteration can be achieved using standard numerical methods. Moreover, as we assume only static or low mobility scenarios, the power optimization only needs to be performed when the channels vary significantly, i.e., typically every few frames.

B. MULTI-USER SCHEDULING

In this section, we evaluate the different algorithms in a multi-user cellular environment. The simulation parameters are set as follows. The radius of the cell is fixed to $D = 1000$ m, and the BS-RS distance to $D_R = 600$ m. The fixed RS is deployed so that a high quality BS-RS link in Line-Of-Sight is guaranteed as recommended in [3], hence it is modeled as an AWGN channel with an average SNR of 40 dBs. All the other channels undergo Rayleigh fading. The average SNR $\bar{\gamma}_k$ for MS_k is given by $\bar{\gamma}_k = P(\frac{D}{d_k})^{\mu}$ where d_k is the distance of MS_k to the BS or RS, and $\mu = 3$ the path loss exponent, which is a typical value for outdoor environments [38]. BS transmit power P is given such that a received SNR level of 5 dBs at the cell edge is ensured, and the RS transmit power P_R is equal to P.

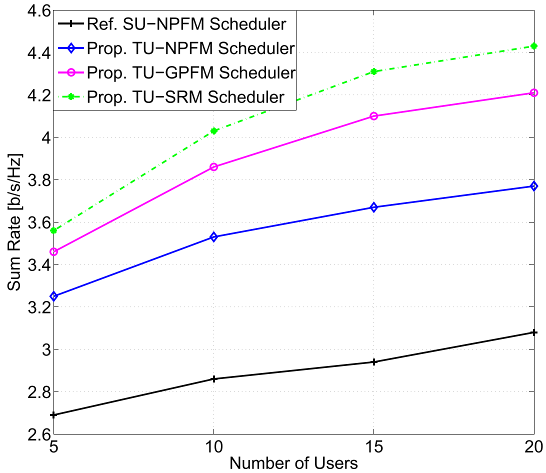

FIGURE 11. Sum rate performance of proposed and reference schedulers.

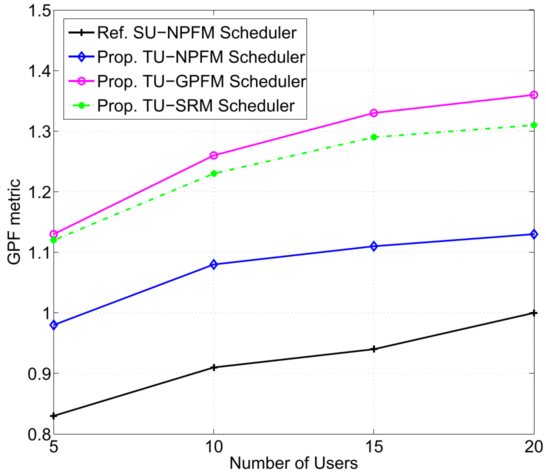

FIGURE 12. GPF performance of proposed and reference schedulers.

In Figs. 11 and 12, we compare the sum rate and GPF metric of the proposed and reference schedulers against the number of users K per sector, respectively. First of all, we observe that all the proposed algorithms outperform the reference *SU-NPFM* scheduler of Section III-B that allocates a single user per scheduling frame. This is because the proposed schedulers enable to make use of the best combination of high quality direct and relayed links that may belong to two different users. Moreover, the Proposed *3L2UHM-SR* and *3L2UHM-GPF* schemes underlying those schedulers can take full advantage of the available links thanks to the optimized power ratios among the three HM layers. As expected, the proposed *TU-SRM* scheduler achieves a higher sum-rate than the proposed *TU-GPFM* scheduler, and vice versa concerning the GPF performance. Nevertheless, both schedulers largely outperform the reference *SU-NPFM* scheduler for both metrics. Note that, even if only sum-rate maximization is aimed by *TU-SRM* scheduler, it also greatly improves the fairness level as well, as shown in Fig. 12. This is because in the Proposed *3L2UHM-SR* scheme, any user having either a good direct or relayed link gets a higher chance to be selected as one user of the allocated pair, while only users experiencing high quality over both direct and relayed links are favored by reference SU schedulers.

Moreover, despite its suboptimality, the proposed *TU-NPFM* scheduler still outperforms the reference *SU-NPFM* scheduler both in terms of sum-rate and GPF metric. This performance gain is achieved with much lower computational complexity compared to the Proposed *TU-GPFM* scheduler, which is quite promising for practical implementation.

VII. CONCLUSION

We have proposed SC-based relaying schemes and non-orthogonal schedulers designed under the practical constraints of discrete HMs. In the proposed schemes, the messages destined to two users are superposed into three HM layers with given power ratios that are tuned to the qualities of the corresponding three links, i.e., the best direct and the two relayed links. We have first derived the sum-rate and GPF analytical expressions, taking into account the decoding errors for all messages at each step. Given the intractability of the joint power optimization problem, we have instead proposed a sub-optimal power allocation based on iterative optimization and numerical methods. This gave rise to the *3L2UHM-SR* scheme for sum-rate maximization and the *3L2UHM-GPF* scheme for GPF maximization. Next, these schemes were integrated into two schedulers for the DL of a multi-user cellular relay system, the *TU-SRM* and *TU-GPFM* schedulers for sum-rate and GPF enhancement, respectively. A low-complexity scheduler based on PF, *TU-NPFM*, was also proposed. The simulation results have shown that both *3L2UHM-SR* and *3L2UHM-GPF* schemes with the proposed power allocation closely approached the optimal performance. Moreover, these schemes improved the rate and GPF metrics compared to conventional relaying schemes over a large SNR region. Finally, in the multi-user setting, the proposed schedulers outperformed conventional orthogonal schedulers both in terms of sum-rate and fairness. More importantly, we could confirm that the main conclusions of previous works such as [15] and [17] assuming Gaussian channels were also true under the more stringent constraints of discrete HMs, namely the superiority of non-orthogonal schedulers over conventional orthogonal schedulers for various system level metrics. These conclusions open up new perspectives for the design and integration of non-orthogonal multiple access into next-generation cellular relay systems.

In the future work, channel coding will be considered in conjunction of the discrete HM levels in the proposed schemes, and the impact of adaptive discrete HM and coding on the different schemes and schedulers will be evaluated.

APPENDIX A

Below are given all the partial derivatives' expressions involved in (25)-(26) for sum-rate optimization in IV-C.

$$
\begin{aligned}
&\frac{\partial P_3}{\partial \alpha_{\mathrm{b1}}}(\gamma_{\mathrm{R}}, \alpha_{\mathrm{b1}}, \alpha_{\mathrm{s1}}) \\
&= \frac{1}{4}\sqrt{\frac{\gamma_{\mathrm{R}}}{2\pi}} \\
&\quad \times \Bigg\{ \left(\sqrt{\frac{1}{\alpha_{\mathrm{b1}}}} + \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\qquad + \left(\sqrt{\frac{1}{\alpha_{\mathrm{b1}}}} - \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\qquad + \left(\sqrt{\frac{1}{\alpha_{\mathrm{b1}}}} + \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\qquad + \left(\sqrt{\frac{1}{\alpha_{\mathrm{b1}}}} - \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \Bigg\} \\
&\quad \times \Bigg\{ 1 - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \\
&\qquad - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \\
&\qquad - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \\
&\qquad - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial P_3}{\partial \alpha_{\mathrm{s1}}}(\gamma_{\mathrm{R}}, \alpha_{\mathrm{b1}}, \alpha_{\mathrm{s1}}) \\
&= \frac{1}{4}\sqrt{\frac{\gamma_{\mathrm{R}}}{2\pi}} \\
&\quad \times \Bigg\{ \left(\sqrt{\frac{1}{\alpha_{\mathrm{s1}}}} + \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\qquad + \left(-\sqrt{\frac{1}{\alpha_{\mathrm{s1}}}} - \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\qquad + \left(\sqrt{\frac{1}{\alpha_{\mathrm{s1}}}} + \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\qquad + \left(\sqrt{\frac{1}{\alpha_{\mathrm{s1}}}} - \sqrt{\frac{1}{\alpha_2}}\right) \\
&\qquad \times \exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \Bigg\} \\
&\quad \times \Bigg\{ 1 - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \\
&\qquad - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \\
&\qquad - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \\
&\qquad - \frac{1}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{b1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \Bigg\},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial P_1}{\partial \alpha_{\mathrm{b1}}}(2, \gamma_{\mathrm{R}}\alpha_2) \\
&= \frac{\partial P_1}{\partial \alpha_{\mathrm{s1}}}(2, \gamma_{\mathrm{R}}\alpha_2) \\
&= -\sqrt{\frac{\gamma_{\mathrm{R}}}{2\pi\alpha_2}}\exp\left(-\frac{\gamma_{\mathrm{R}}\alpha_2}{2}\right)\left\{1 - \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)\right\},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial P_2}{\partial \alpha_{\mathrm{b1}}}(2, \gamma_{\mathrm{R}}, \alpha_{\mathrm{s1}}, \alpha_2) \\
&= \frac{1}{2}\sqrt{\frac{\gamma_{\mathrm{R}}}{\pi}} \times \Bigg\{ \sqrt{\frac{1}{2\alpha_2}}\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\qquad\qquad - \sqrt{\frac{1}{2\alpha_2}}\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \Bigg\} \\
&\quad \times \left\{ 1 - \frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \right. \\
&\qquad \left. \times - \frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \right\},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial P_2}{\partial \alpha_{\mathrm{s1}}}(2, \gamma_{\mathrm{R}}, \alpha_{\mathrm{s1}}, \alpha_2) \\
&= \frac{1}{2}\sqrt{\frac{\gamma_{\mathrm{R}}}{\pi}} \\
&\quad \times \Bigg\{ \left(\sqrt{\frac{1}{2\alpha_{\mathrm{s1}}}} + \sqrt{\frac{1}{2\alpha_2}}\right)\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \\
&\quad + \left(\sqrt{\frac{1}{2\alpha_{\mathrm{s1}}}} - \sqrt{\frac{1}{2\alpha_2}}\right)\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right)^2\right) \Bigg\} \\
&\quad \times \left\{ 1 - \frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} - \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \right. \\
&\qquad \left. \times - \frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}}\alpha_{\mathrm{s1}}}{2}} + \sqrt{\frac{\gamma_{\mathrm{R}}\alpha_2}{2}}\right) \right\},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial P_2}{\partial \alpha_{\mathrm{b1}}}(2,\gamma_{\mathrm{D1}},\alpha_{\mathrm{b1}},\alpha_2)\\
&= \frac{1}{2}\sqrt{\frac{\gamma_{\mathrm{D1}}}{2\pi}}\left\{1-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}-\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)\right.\\
&\qquad\qquad \left.-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}+\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)\right\}\\
&\quad\times\left\{\left(\sqrt{\frac{1}{\alpha_{\mathrm{b1}}}}+\sqrt{\frac{1}{\alpha_2}}\right)\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}-\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)^2\right)\right.\\
&\quad\left.+\left(\sqrt{\frac{1}{\alpha_{\mathrm{b1}}}}-\sqrt{\frac{1}{\alpha_2}}\right)\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}+\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)^2\right)\right\},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial P_2}{\partial \alpha_{\mathrm{s1}}}(2,\gamma_{\mathrm{D1}},\alpha_{\mathrm{b1}},\alpha_2)\\
&= \frac{1}{2}\sqrt{\frac{\gamma_{\mathrm{D1}}}{2\pi}}\left\{1-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}-\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)\right.\\
&\qquad\qquad \left.-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}+\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)\right\}\\
&\quad\times\left\{\sqrt{\frac{1}{\alpha_2}}\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}-\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)^2\right)\right.\\
&\qquad\left.-\sqrt{\frac{1}{\alpha_2}}\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_{\mathrm{b1}}}{2}}+\sqrt{\frac{\gamma_{\mathrm{D1}}\alpha_2}{2}}\right)^2\right)\right\}.
\end{aligned}
$$

APPENDIX B

We provide the analysis for optimizing β given α_{b1}, α_{s1} in (18) as in Section IV-C.2 for HM level $[\mathrm{QPSK}]^2$-HQAM and $[8\text{-PAM}]^2$-HQAM.

A. $[\mathrm{QPSK}]^2$-HQAM IN STEP 2

With $[\mathrm{QPSK}]^2$-HQAM in Step 2, we have $0<\beta<1/2$,

$$
\frac{\partial R_{\mathrm{3L2UHM}}}{\partial \beta} = F_1(\beta)+F_2(\beta), \quad \text{where}
$$

$$
\begin{aligned}
F_1(\beta) &= -\frac{C_1C_2N_{\mathrm{R}}}{2}\sqrt{\frac{\gamma_{\mathrm{R1}}}{\pi}}\\
&\quad\times\left\{1-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R1}}(1-\beta)}{2}}-\sqrt{\frac{\gamma_{\mathrm{R1}}\beta}{2}}\right)\right.\\
&\qquad\left.-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R1}}(1-\beta)}{2}}+\sqrt{\frac{\gamma_{\mathrm{R1}}\beta}{2}}\right)\right\}^{2N_{\mathrm{R}}-1}\\
&\quad\times\left\{\left(\sqrt{\frac{1}{2(1-\beta)}}+\sqrt{\frac{1}{2\beta}}\right)\right.\\
&\qquad\times\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R1}}(1-\beta)}{2}}-\sqrt{\frac{\gamma_{\mathrm{R1}}\beta}{2}}\right)^2\right)\\
&\quad+\left(\sqrt{\frac{1}{2(1-\beta)}}-\sqrt{\frac{1}{2\beta}}\right)\\
&\qquad\left.\times\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R1}}(1-\beta)}{2}}+\sqrt{\frac{\gamma_{\mathrm{R1}}\beta}{2}}\right)^2\right)\right\},
\end{aligned}
$$

$$
\begin{aligned}
F_2(\beta) &= -\frac{C_1N_{\mathrm{R}}}{2}\sqrt{\frac{\gamma_{\mathrm{R2}}}{\pi}}\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)\right\}^{2N_{\mathrm{R}}}\\
&\quad\times\left\{1-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}(1-\beta)}{2}}-\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)\right.\\
&\qquad\left.-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}(1-\beta)}{2}}+\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)\right\}^{2N_{\mathrm{R}}-1}\\
&\quad\times\left\{\left(\sqrt{\frac{1}{2(1-\beta)}}+\sqrt{\frac{1}{2\beta}}\right)\right.\\
&\qquad\times\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R2}}(1-\beta)}{2}}-\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)^2\right)\\
&\quad+\left(\sqrt{\frac{1}{2(1-\beta)}}-\sqrt{\frac{1}{2\beta}}\right)\\
&\qquad\left.\times\exp\left(-\left(\sqrt{\frac{\gamma_{\mathrm{R2}}(1-\beta)}{2}}+\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)^2\right)\right\}\\
&\quad+C_1N_{\mathrm{R}}\sqrt{\frac{\gamma_{\mathrm{R2}}}{\pi}}\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)\right\}^{2N_{\mathrm{R}}-1}\\
&\quad\times\sqrt{\frac{1}{2\beta}}\exp\left(-\frac{\gamma_{\mathrm{R2}}\beta}{2}\right)\\
&\quad\times\left\{1-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}(1-\beta)}{2}}-\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)\right.\\
&\qquad\left.-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}(1-\beta)}{2}}+\sqrt{\frac{\gamma_{\mathrm{R2}}\beta}{2}}\right)\right\}^{2N_{\mathrm{R}}}.
\end{aligned}
$$

From which we obtain

$$
\begin{aligned}
&\lim_{\beta\to 0^+}\frac{\partial R_{\mathrm{3L2UHM}}}{\partial \beta}\\
&= -\frac{C_1C_2N_{\mathrm{R}}}{2\sqrt{2}}\sqrt{\frac{\gamma_{\mathrm{R1}}}{\pi}}\exp\left(-\frac{\gamma_{\mathrm{R1}}}{2}\right)\\
&\quad\times\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R1}}}{2}}\right)\right\}^{2N_{\mathrm{R}}-1}\\
&\quad+C_1N_{\mathrm{R}}\sqrt{\frac{\gamma_{\mathrm{R2}}}{\pi}}\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}}{2}}\right)\right\}^{2N_{\mathrm{R}}-1}\left(\frac{1}{2}\right)^{2N_{\mathrm{R}}-1}\\
&\quad\times\left[\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R2}}}{2}}\right)\right\}\left\{\lim_{\beta\to 0^+}\sqrt{\frac{1}{2\beta}}\right\}\right.\\
&\qquad\left.-\frac{1}{4\sqrt{2}}\exp\left(-\frac{\gamma_{\mathrm{R2}}}{2}\right)\right],
\end{aligned}
$$

$$
\begin{aligned}
&\lim_{\beta\to \frac{1}{2}^-}\frac{\partial R_{\mathrm{3L2UHM}}}{\partial \beta}\\
&= -C_1C_2N_{\mathrm{R}}\sqrt{\frac{\gamma_{\mathrm{R1}}}{\pi}}\\
&\quad\times\left\{\frac{3}{4}-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\gamma_{\mathrm{R1}}}\right)\right\}^{2N_{\mathrm{R}}-1}
\end{aligned}
$$

$$-C_1N_{\mathrm{R}}\sqrt{\frac{\gamma_{\mathrm{R}2}}{\pi}}\left\{\frac{3}{4}-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\gamma_{\mathrm{R}2}}\right)\right\}^{2N_{\mathrm{R}}-1}\times\left\{1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}}{4}}\right)\right\}^{2N_{\mathrm{R}}-1}\times\left[1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}}{4}}\right)-\left\{\frac{3}{4}-\frac{1}{4}\mathrm{erfc}\left(\sqrt{\gamma_{\mathrm{R}2}}\right)\right\}\times\exp\left(-\frac{\gamma_{\mathrm{R}2}}{4}\right)\right].$$

For analyzing these limits, we define the following function of x,

$$f(x)=1-\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{x}{4}}\right)-\left\{\frac{3}{4}-\frac{1}{4}\mathrm{erfc}\left(\sqrt{x}\right)\right\}\exp\left(-\frac{x}{4}\right).$$

Differentiating $f(x)$ with respect to x, we have

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x}=\frac{1}{4}\exp\left(-\frac{x}{4}\right)\times\left[\sqrt{\frac{1}{\pi x}}(1-\exp(-x))+\frac{3}{4}-\frac{1}{4}\mathrm{erfc}\left(\sqrt{x}\right)\right].$$

Since $\frac{\mathrm{d}f(x)}{\mathrm{d}x}>0$ for $x>0$ and $f(0)=0$, $f(\gamma_{\mathrm{R}2})$ is always positive. Thus, $\lim_{\beta\to\frac{1}{2}^-}\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}<0$. Moreover, we see that $\lim_{\beta\to0^+}\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=+\infty>0$. Therefore, the intermediate-value theorem guarantees the existence of at least a solution in the considered range. Again, the proposed scheme chooses the value β^* obtained by solving $\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=0$ numerically.

B. [8-PAM]^2-HQAM IN STEP 2

In this case, we have $0<\beta<1$ and

$$\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=C_1C_2N_{\mathrm{R}}\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}1}(1-\beta)}{21}}\right)\right\}^{N_{\mathrm{R}}-1}\times\left(-\frac{7}{8}\right)\sqrt{\frac{\gamma_{\mathrm{R}1}}{\pi}}\sqrt{\frac{1}{21(1-\beta)}}\exp\left(-\frac{\gamma_{\mathrm{R}1}(1-\beta)}{21}\right)$$
$$+C_1N_{\mathrm{R}}\sqrt{\frac{\gamma_{\mathrm{R}2}}{\pi}}\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}(1-\beta)}{21}}\right)\right\}^{N_{\mathrm{R}}-1}\times\left(\frac{-7}{8}\right)\sqrt{\frac{1}{21(1-\beta)}}\exp\left(-\frac{\gamma_{\mathrm{R}2}(1-\beta)}{21}\right)\times\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}\beta}{21}}\right)\right\}^{N_{\mathrm{R}}}$$
$$+C_1N_{\mathrm{R}}\sqrt{\frac{\gamma_{\mathrm{R}2}}{\pi}}\left\{1-\frac{7}{8}\mathrm{erfc}\left(\frac{\sqrt{\gamma_{\mathrm{R}2}(1-\beta)}}{21}\right)\right\}^{N_{\mathrm{R}}}\times\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}\beta}{21}}\right)\right\}^{N_{\mathrm{R}}-1}\frac{7}{8}\sqrt{\frac{1}{21\beta}}\exp\left(-\frac{\gamma_{\mathrm{R}2}\beta}{21}\right).$$

We obtain

$$\lim_{\beta\to0^+}\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=-\frac{7C_1C_2N_{\mathrm{R}}}{8}\sqrt{\frac{\gamma_{\mathrm{R}1}}{21\pi}}\exp\left(-\frac{\gamma_{\mathrm{R}1}}{21}\right)\times\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}1}}{21}}\right)\right\}^{N_{\mathrm{R}}-1}$$
$$+\frac{7C_1N_{\mathrm{R}}}{8}\sqrt{\frac{\gamma_{\mathrm{R}2}}{21\pi}}\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}}{21}}\right)\right\}^{N_{\mathrm{R}}-1}\left(\frac{1}{8}\right)^{N_{\mathrm{R}}-1}\times\left[\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}}{21}}\right)\right\}\left\{\lim_{\beta\to0^+}\sqrt{\frac{1}{\beta}}\right\}-\frac{1}{8}\exp\left(-\frac{\gamma_{\mathrm{R}2}}{21}\right)\right],$$

$$\lim_{\beta\to1^-}\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=-\frac{7C_1C_2N_{\mathrm{R}}}{8}\sqrt{\frac{\gamma_{\mathrm{R}1}}{21\pi}}\left(\frac{1}{8}\right)^{N_{\mathrm{R}}-1}\left\{\lim_{\beta\to1^-}\sqrt{\frac{1}{1-\beta}}\right\}$$
$$-\frac{7C_1N_{\mathrm{R}}}{8}\sqrt{\frac{\gamma_{\mathrm{R}2}}{21\pi}}\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}}{21}}\right)\right\}^{N_{\mathrm{R}}-1}\left(\frac{1}{8}\right)^{N_{\mathrm{R}}-1}\times\left[\left\{1-\frac{7}{8}\mathrm{erfc}\left(\sqrt{\frac{\gamma_{\mathrm{R}2}}{21}}\right)\right\}\left\{\lim_{\beta\to1^-}\sqrt{\frac{1}{1-\beta}}\right\}\frac{1}{8}\exp\left(-\frac{\gamma_{\mathrm{R}2}}{21}\right)\right].$$

We have $\lim_{\beta\to0^+}\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=+\infty>0$, $\lim_{\beta\to1^-}\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=-\infty<0$, so the intermediate-value theorem guarantees the existence of a solution in the considered range. Similarly to the other cases, we use the value β^* obtained by solving $\frac{\partial R_{3\mathrm{L}2\mathrm{UHM}}}{\partial\beta}=0$ numerically.

REFERENCES

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[2] M. Kaneko, K. Hayashi, P. Popovski, K. Ikeda, H. Sakai, and R. Prasad, "Amplify-and-forward cooperative diversity schemes for multi-carrier systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1845–1850, May 2008.

[3] R. Pabst *et al.*, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Wireless Commun. Mag.*, vol. 42, no. 9, pp. 80–89, Sep. 2004.

[4] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I. System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[5] P. Popovski and E. de Carvalho, "Improving the rates in wireless relay systems through superposition coding," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4831–4836, Dec. 2008.

[6] M. Salem *et al.*, "An overview of radio resource management in relay-enhanced OFDMA-based networks," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 3, pp. 422–438, Third Quarter 2010.

[7] M. Kaneko, P. Popovski, and K. Hayashi, "Throughput-guaranteed resource-allocation algorithms for relay-aided cellular OFDMA system," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1951–1964, May 2009.

[8] M. Salem *et al.*, "Fairness-aware joint routing and scheduling in OFDMA-based cellular fixed relay networks," in *Proc. IEEE ICC*, Dresden, Germany, Jun. 2009, pp. 1–6.

[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[10] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 2006.

[11] J. Hossain, M.-S. Alouini, and V. K. Bhargava, "Rate adaptive hierarchical modulation-assisted two-user opportunistic scheduling," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2076–2085, Jun. 2008.

[12] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. 77th IEEE VTC-Spring*, Dresden, Germany, Jun. 2013, pp. 1–5.

[13] N. Otao, Y. Kishiyama, and K. Higuchi, "Performance of non-orthogonal access with SIC in cellular downlink using proportional fair-based resource allocation," in *Proc. IEEE ISWCS*, Paris, France, Aug. 2012, pp. 476–480.

[14] M. Kaneko, K. Hayashi, and H. Sakai, "Sum rate maximizing superposition coding scheme for a two-user wireless relay system," *IEEE Commun. Lett.*, vol. 15, no. 4, pp. 428–430, Apr. 2011.

[15] M. Kaneko, K. Hayashi, P. Popovski, and H. Sakai, "Fairness-aware superposition coded scheduling for a multi-user cooperative cellular system," *IEICE Trans. Commun.*, vol. E94-B, no. 12, pp. 3272–3279, Dec. 2011.

[16] M. Kaneko, K. Hayashi, P. Popovski, and H. Sakai, "Proportional fair scheduling with superposition coding in a cellular cooperative relay system," *Ann. Telecommun.*, vol. 68, no. 9, pp. 525–537, Oct. 2013.

[17] M. Kaneko, K. Hayashi, and H. Sakai, "Superposition coding based user combining schemes for non-orthogonal scheduling in a wireless relay system," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3232–3243, Jun. 2014.

[18] P. K. Vitthaladevuni and M.-S. Alouini, "BER computation of 4/M-QAM hierarchical constellations," *IEEE Trans. Broadcast.*, vol. 47, no. 3, pp. 228–239, Sep. 2001.

[19] P. K. Vitthaladevuni and M.-S. Alouini, "Errata for 'BER computation of 4/M-QAM hierarchical constellations,'" *IEEE Trans. Broadcast.*, vol. 49, no. 4, p. 408, Dec. 2003.

[20] P. K. Vitthaladevuni and M.-S. Alouini, "A recursive algorithm for the exact BER computation of generalized hierarchical QAM constellations," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 297–307, Jan. 2003.

[21] K. Zheng, L. Wang, and W. Wang, "Performance analysis of coded cooperation with hierarchical modulation," in *Proc. IEEE ICC*, Beijing, China, May 2008, pp. 4978–4982.

[22] X. Wang and L. Cai, "Proportional fair scheduling in hierarchical modulation aided wireless networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1584–1593, Apr. 2013.

[23] C. Hausl and J. Hagenauer, "Relay communication with hierarchical modulation," *IEEE Commun. Lett.*, vol. 11, no. 1, pp. 64–66, Jan. 2007.

[24] H. Yamaura, M. Kaneko, K. Hayashi, and H. Sakai, "Adaptive hierarchical modulation and power allocation for superposition-coded relaying," *EURASIP J. Wireless Commun. Netw.*, vol. 2013, p. 233, Dec. 2013. DOI: 10.1186/1687-1499-2013-233

[25] H. Yamaura, M. Kaneko, K. Hayashi, and H. Sakai, "Superposition coding scheme with discrete adaptive modulation for wireless relay systems," in *Proc. IEEE VTC-Fall*, San Francisco, LA, USA, Sep. 2011, pp. 1–5.

[26] C. Hucher and P. Sadeghi, "Hierarchical modulation-based cooperative scheme: Minimizing the symbol error probability," in *Proc. IEEE 17th ICT*, Doha, Qatar, Apr. 2010, pp. 309–315.

[27] H. X. Nguyen, H. H. Nguyen, and T. Le-Ngoc, "Signal transmission with unequal error protection in wireless relay networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2166–2178, Jun. 2010.

[28] R. J. Whang, H. Liu, and E.-K. Hong, "Multiuser cooperative relay communication employing hierarchical modulation," in *Proc. IEEE 71st VTC-Spring*, Taipei, Taiwan, May 2010, pp. 1–5.

[29] H. Son and S. Lee, "Hierarchical modulation based cooperative relaying over a multi-cell OFDMA network," *Wireless Netw.*, vol. 19, no. 5, pp. 577–590, Jul. 2013.

[30] T. V. Nguyen, P. C. Cosman, and L. B. Milstein, "Optimized receiver design for decode-and-forward relays using hierarchical modulation," in *Proc. IEEE Asilomar CSSC*, Pacific Grove, CA, USA, Nov. 2013, pp. 1535–1539.

[31] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 1995.

[32] K. Kimura, M. Nakada, T. Obara, and F. Adachi, "Single-carrier cooperative DF relay using adaptive modulation," in *Proc. IEEE APWCS*, Singapore, Aug. 2011, pp. 22–23.

[33] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels: A Unified Approach to Performance Analysis*, 1st ed. New York, NY, USA: Wiley, 2000.

[34] Y. Zhang, Y. Ma, and R. Tafazolli, "Modulation-adaptive cooperation schemes for wireless networks," in *Proc. IEEE VTC-Spring*, Singapore, May 2008, pp. 1320–1324.

[35] P. Popovski, E. de Carvalho, K. Sivanesan, E.-T. Lim, H.-K. Choi, and Y.-K. Cho, "Method and apparatus for transmitting and receiving data using multi-user superposition coding in a wireless relay system," U.S. Patent 2008 0 227 388 A1, Sep. 18, 2008.

[36] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, Mar. 1998.

[37] J.-G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 766–778, Mar. 2007.

[38] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2001.

MEGUMI KANEKO (M'05) received the B.S. and M.Sc. degrees in communication engineering from the Institut National des Télécommunications, France, in 2003 and 2004, respectively, the M.Sc. degree from Aalborg University, Denmark, and the Ph.D. degree from Aalborg University, in 2007. In 2007, she was a Visiting Researcher with Kyoto University, Kyoto, Japan, and a JSPS Post-Doctoral Fellow from 2008 to 2010. She is currently an Assistant Professor with the Department of Systems Science, Graduate School of Informatics, Kyoto University. Her research interests include wireless communication, cross-layer protocol design, and communication theory. She received the 2009 Ericsson Young Scientist Award, the IEEE Globecom 2009 Best Paper Award, the 2011 Funai Young Researcher's Award, the WPMC 2011 Best Paper Award, and the 2012 Telecom System Technology Award.

HIROFUMI YAMAURA received the B.E. degree in applied mathematics and physics and the M.E. degree in systems science from the Department of Systems Science, Graduate School of Informatics, Kyoto University, Japan, in 2011 and 2013, respectively. His research interests include signal processing and radio resource allocation in wireless communication systems. He received the IEEE VTC2011-Fall Young Researcher's Encouragement Award from the IEEE VTS Japan Chapter.

YOUHEI KAJITA received the B.E. degree from Osaka Prefecture University, in 2012, and the M.E. degree in systems science from the Department of Systems Science, Graduate School of Informatics, Kyoto University, Japan, in 2014. His research interests include signal processing and radio resource allocation in wireless communication systems.

KAZUNORI HAYASHI (M'97) received the B.E., M.E., and Ph.D. degrees in communication engineering from Osaka University, Osaka, Japan, in 1997, 1999, and 2002, respectively. Since 2002, he has been with the Department of Systems Science, Graduate School of Informatics, Kyoto University, where he is currently an Associate Professor. His research interests include statistical signal processing for communication systems. He is a member of IEICE and ISCIE. He received the ICF Research Award from the KDDI Foundation in 2008, the IEEE Globecom 2009 Best Paper Award, the IEICE Communications Society Best Paper Award in 2010, the WPMC'11 Best Paper Award, the Telecommunications Advancement Foundation Award in 2012, and the IEICE Communications Society Best Tutorial Paper Award in 2013.

HIDEAKI SAKAI (LF'14) received the B.E. and D.E. degrees in applied mathematics and physics from Kyoto University, Kyoto, Japan, in 1972 and 1981, respectively. He spent six months with Stanford University as a Visiting Scholar from 1987 to 1988. He has been a Professor with the Department of Systems Science, Graduate School of Informatics, Kyoto University, where he is currently a Professor Emeritus. His research interests are in the areas of adaptive and statistical signal processing. He served twice as an Associate Editor of the IEEE Transactions on Signal Processing.