On Non-Orthogonal Multiple Access With Finite-Alphabet Inputs in Z-Channels

Zheng Dong, He Chen, *Member, IEEE*, Jian-Kang Zhang, *Senior Member, IEEE*,
and Lei Huang, *Senior Member, IEEE*

Abstract—This paper focuses on the design of non-orthogonal multiple access in a classical two-transmitter two-receiver Z-channel, wherein one transmitter sends information to its intended receiver from the direct link while the other transmitter sends information to both receivers from the direct and cross links. Unlike most existing designs using (continuous) Gaussian input distribution, we consider the practical finite-alphabet (i.e., discrete) inputs by assuming that the widely used quadrature amplitude modulation constellations are adopted by both transmitters. To balance the error performance of two receivers, we apply the max–min fairness design criterion in this paper. More specifically, we propose to jointly optimize the scaling factors at both transmitters, which control the minimum Euclidean distance of transmitting constellations, to maximize the smaller minimum Euclidean distance of two resulting constellations at the receivers, subject to an individual average power constraint at each transmitter. The formulated problem is a *mixed continuous-discrete* optimization problem and is thus intractable in general. By resorting to the Farey sequence, we manage to attain the closed-form expression for the optimal solution to the formulated problem. This is achieved by dividing the overall feasible region of the original optimization problem into a finite number of sub-intervals and deriving the optimal solution in each sub-interval. Through carefully observing the structure of the optimal solutions in all sub-intervals, we obtain compact and closed-form expressions for the optimal solutions to the original problem in three possible scenarios defined by the relative strength of the cross link. Simulation studies are provided to validate our analysis and demonstrate the merits of the proposed design over existing orthogonal or non-orthogonal schemes.

Index Terms—Non-orthogonal multiple access (NOMA), Z-channel, finite-alphabet inputs, quadrature amplitude modulation, max-min fairness, Farey sequence.

Z. Dong is with Shenzhen University, Shenzhen 518060, China, and also with McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: dongz3@mcmaster.ca).

H. Chen is with The University of Sydney, Sydney, NSW 2006, Australia (e-mail: he.chen@sydney.edu.au).

J.-K. Zhang is with McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: jkzhang@mail.ece.mcmaster.ca).

L. Huang is with Shenzhen University, Shenzhen 518060, China (e-mail: lhuang@szu.edu.cn).

I. INTRODUCTION

MULTIPLE access technologies have been playing an important role in determining the performance of each generation of mobile communication systems. Based on how the resources are allocated to users, multiple access technologies can generally be categorized into two types: orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA) [1, Ch. 6]. The current generation of cellular networks, known as 4G, and all previous generations have primarily adopted the OMA technologies, which include frequency-division multiple access (FDMA) for 1G, time-division multiple access (TDMA) for 2G, code-division multiple access (CDMA) for 3G, and orthogonal frequency-division multiple access (OFDMA) for 4G [2]. In these OMA schemes, the resource is partitioned into orthogonal blocks in time/frequency/code domain, and each resource block is then assigned to one single user exclusively. In this sense, there is no inter-user interference in OMA, leading to low-complexity receiver and scheduling algorithms. Moreover, after the resource allocation, the multiple-user problem is divided into several point-to-point problems such that the well-established single-user encoder/decoder techniques can be directly applied. However, early information-theoretic studies showed that compared with NOMA, OMA has lower spectral efficiency as it normally cannot achieve the multi-user channel capacity region [3]. Besides, OMA is not scalable as the total number of orthogonal resources and their granularity strictly limit the maximum number of served users.

Different from OMA, NOMA exploits the power domain to multiplex multiple users together such that they can be served in the same time/frequency/code resources [2]–[6]. As such, with proper multi-user detection techniques to deal with the inter-user interference at the receiver side (e.g., successive interference cancellation (SIC) [7]), NOMA is capable of achieving improved spectral efficiency and serving much more users simultaneously. In fact, the uplink and downlink versions of NOMA, well-known as multiple access channel (MAC) and broadcast channel (BC) respectively, have been intensively investigated for several decades in the information theory community, see, e.g., [8]–[11]. However, due to the high complexity of interference cancellation, these studies mainly lied in the theoretical aspects and their results were not implemented in practical communication systems. With the fast advances of hardware, the implementation of NOMA with interference cancellation becomes more affordable and feasible. Actually, NOMA has been regarded as a key enabling

technology to meet the unprecedented requirements of 5G wireless networks due to its significant network throughput gain and great potential to support massive connectivity, low latency and user fairness [2], [5], [7], [12]–[17]. Furthermore, a two-user downlink scenario of NOMA, termed multiuser superposition transmission (MUST), has been incorporated in the 3rd Generation Partnership Project (3GPP) Long Term Evolution-Advanced (LTE-A) [18], [19].

Most conventional information-theoretic and recent studies on NOMA adopted the assumption of Gaussian input distribution [5], [8], [9], [20]–[25]. Although the designs with Gaussian signaling can approach most of the known capacity inner bounds, such as in [8]–[11], their direct implementation in practical communication systems may lead to significant performance loss [26]. Moreover, Gaussian signaling will require unaffordable encoding and decoding efforts, which could lead to extremely high hardware cost, huge storage capability, high computational complexity, and long delay. Therefore, Gaussian inputs could arguably be infeasible for current hardware and it acts mostly as the theoretical benchmark. The inputs of practical wireless systems are actually drawn from finite constellations, such as phase shift keying (PSK) modulation or quadrature amplitude modulation (QAM), which are essentially different from the continuous Gaussian inputs. When it comes to a NOMA system with finite input constellations, the key design challenge is to guarantee that each user's codeword can be uniquely decoded from their sum signal at the receiver side [27]–[29]. For the two-user MAC with finite-alphabet inputs, a constellation rotation (CR) scheme and a constellation power allocation (CPA) scheme were proposed in [30] and [31] to construct an unambiguous sum constellation at the receiver, respectively. This is achieved by strategically introducing certain angle of rotation between the input constellations in the CR scheme and appropriately controlling the transmit power of each user in the CPA scheme. The results in [30] and [31] have been extended to various multiple-antenna scenarios, see, e.g., [32] and references therein. The aforementioned NOMA designs were primarily for the PSK modulations by utilizing its circular structure. The studies on NOMA with QAM, another practical modulation scheme that has been widely adopted in cellular systems due to its higher spectral efficiency, are quite limited. Very recently, the mutual information were used as the performance metric to optimize NOMA systems with QAM in [33] and [34]. However, the optimal NOMA designs in [33] and [34] were achieved by numerical approaches with high computational complexities.

We also notice that the existing studies on NOMA mainly focused on MAC and BC, due to their wide applications in centralized systems like cellular networks. Also there are some initial efforts considering the NOMA design for the interference channel (IC) [3]. With recent advances in non-centralized networks (e.g., wireless sensor networks and ad hoc networks), Z-channel (ZC) was proposed in [35] and attracted considerable attention in the past decade [36]–[42]. As a special case of the classical two-user IC [43], a two-user ZC consists of two transmitters and two receivers, wherein one transmitter sends information to its intended receiver from

the direct link without causing interference to the unintended receiver, while the other transmitter sends information to both receivers from the direct and cross links. In this sense, the ZC is a general channel model that includes MAC and BC as special cases. The ZC is also closely related to the Z-interference channel (ZIC) [44]–[47], wherein each transmitter transmits information to its corresponding receiver only from the direct link and the received signal from the cross link carries no desired information, and thus is treated as interference at the receiver side. It is also worth emphasizing that there are two messages transmitted from the two direct links in ZIC, while three messages are sent via the two direct links and the cross link in ZC. Both ZC and ZIC are proper models for the multi-cell downlink transmission, where one user is located near to the cell edge and thus can receive signals from both transmitters, while the other user is near the cell center and suffers from no interference. Another example corresponds to the two-user IC, where one cross link is blocked by obstacles with large pathloss such as tall buildings or thick walls, while the other user is still exposed to interference [45], [48]. Despite of the existing great efforts on ZC, to our best knowledge, the design of NOMA with practical finite-alphabet inputs in ZC is still an *open problem* in the literature.

Motivated by this gap, in this paper we concentrate on the practical design of NOMA with QAM in a two-transmitter two-receiver ZC. It is worth emphasizing that the design of NOMA with QAM is much more challenging than that for PSK modulation. This is mainly because the unambiguity of sum QAM constellations is much harder to maintain since its signal points are distributed more evenly and there is a higher probability that more than one signal points coincide or close to each other on the sum constellation. The main contributions of this paper can be summarized as follows:

1) We, for the first time, develop a practical NOMA framework with QAM and max-min fairness in ZC. In our framework, we optimize the scaling factors of both transmitters, which adjust the minimum Euclidean distance of the transmitting constellations, to maximize the smaller minimum Euclidean distance among the resulting constellations at both receivers subject to an individual power constraint on each transmitter. Through our design, the average error performance of both transmitters in the considered ZC can be minimized with good user fairness, which is fundamentally different from the existing designs that mainly focused on the channel capacity maximization.

2) The formulated optimization problem is shown as a mixed continuous-discrete optimization problem, which is challenging to solve in general. By carefully observing the features of the formulated problem, we realize that the Farey sequence (also known as Farey series) [49] can be applied to resolve the problem. More specifically, by taking the advantage of Farey sequence associated with the finite-alphabet, we strategically partition the entire feasible region of the original optimization problem into a finite number of sub-intervals and attain the closed-form solution in each sub-interval. Then, by a
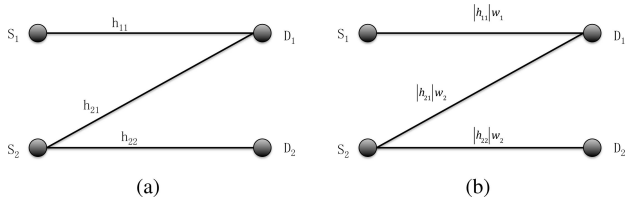
Fig. 1. (a) Two-user complex Gaussian ZC. (b) Two-user real Gaussian ZC.

careful observation on the structure of the solutions in all sub-intervals, the overall solution is obtained in a compact closed-form for three complementary scenarios divided by the relative strength of the cross link.

3) We verify the correctness of the analytical results by conducting simulations in both deterministic and random fading channels. Simulation results show that the sum constellation at the receiver side is still a regular QAM constellation with a larger size for most scenarios, but could be a hierarchical QAM with two-resolution [50], e.g., when the cross link is very strong relative to the direct link. We adopt the bit error rate (BER) as the performance metric to compare the proposed NOMA design with the existing OMA and NOMA schemes under random fading channels. The comparison illustrates that our scheme can achieve a significant lower BER performance than the benchmark schemes, which validates the effectiveness of our design.

II. System Model of Complex Gaussian ZC with QAM Constellations

We consider a two-user complex Gaussian ZC consisting of two transmitters S_1 and S_2, and two receivers D_1 and D_2, as depicted in Fig. 1(a). We consider that each node is equipped with a single antenna and works in a half-duplex mode. As per the ZC, only the cross link between S_2 and D_1 is assumed to be available. Moreover, S_1 sends one unicasting message x_1 to D_1, while S_2 transmits one multicasting message x_2 to both D_1 and D_2. S_1 and S_2 transmit their messages simultaneously using the same frequency band. The equivalent complex baseband signals observed at D_1 and D_2 can be given, respectively, by the following equations:

$$z_1 = h_{11}x_1 + h_{21}x_2 + \xi_1, \tag{1a}$$

$$z_2 = h_{22}x_2 + \xi_2, \tag{1b}$$

where $h_{k\ell} \in \mathbb{C}$, $k, \ell \in \{1, 2\}$ denotes the complex channel coefficients from transmitter S_k to receiver D_ℓ. Hereafter, we call h_{11} and h_{22} the direct links, while h_{21} is referred to as the cross link. In line with [35]–[39], [41], [42], and [45], all the channel links are assumed to be known perfectly at all the terminals. The additive noise processes $\xi_1, \xi_2 \sim \mathcal{CN}(0, 2\sigma^2)$ are independent and identically distributed (i.i.d.) over time and are assumed to be circularly symmetric complex Gaussian (CSCG). Note that the case with different noise levels at receivers can be incorporated into our model by scaling operations. We suppose that QAM constellations are used by both transmitters since it is more spectrally

efficient than other frequently-used modulation schemes such as phase-shift keying (PSK), and is also relatively easy to implement [51, Ch. 5.3.3]. A predefined average transmitted power constraint is imposed to both transmitters,[1] i.e., $\mathbb{E}[|x_1|^2] \leq P_1$ and $\mathbb{E}[|x_2|^2] \leq P_2$. In this paper, the system signal-to-noise ratio (SNR) is defined by $\rho = \frac{1}{2\sigma^2}$.

For analytical simplicity, we decompose the considered complex Gaussian ZC given in (1) into two parallel real scalar Gaussian ZCs [35], [36], [45], which are called the in-phase and quadrature components, respectively. We note that this method was commonly used in the study of IC and the real IC was studied directly in [48] and [52]–[54]. Actually, designing two-dimensional QAM constellations is an extremely challenging problem even for two-user MAC, see e.g., [30], [31], and references therein. In this paper, instead of designing the two-dimensional QAM constellations directly, we propose a practical design that decomposes the complex Gaussian ZC into two parallel real scalar Gaussian ZCs such that we can split the two-dimensional QAM constellation into two one-dimensional PAM constellations. In fact, although we use complex baseband representation in (1), the actual modulated and demodulated signals are all real since the oscillator at the transmitter can only generate real sinusoids rather than complex exponentials, and the channel just introduces amplitude and phase distortion to the transmitted signals [51, Ch. 2.2]. By this means, the original two-dimensional QAM constellation can be split into two one-dimensional pulse amplitude modulation (PAM) constellations for both the in-phase and quadrature components. Mathematically, for the complex Gaussian ZC described in (1), the in-phase and quadrature components can be attained by rotating x_1 and x_2 according to the instantaneous channel coefficients to compensate for the phase offset, and then taking the real and imaginary parts, respectively. First of all, we note that (1) is equivalent to

$$z_1 = |h_{11}| \exp(j \arg(h_{11}))x_1 + |h_{21}| \exp \\ \times (j \arg(h_{21}))x_2 + \xi_1$$

$$\exp\left(j \arg\left(\frac{h_{21}}{h_{22}}\right)\right)z_2 = |h_{22}| \exp(j \arg(h_{21}))x_2 \\ + \exp\left(j \arg\left(\frac{h_{21}}{h_{22}}\right)\right)\xi_2. \tag{2}$$

Now, we set

$$y_1 = \text{Re}(z_1), \quad y_2 = \text{Re}\left(\exp\left(j \arg\left(\frac{h_{21}}{h_{22}}\right)\right)z_2\right), \tag{3a}$$

$$w_1 s_1 = \text{Re}\left(\exp(j \arg(h_{11}))x_1\right), \tag{3b}$$

$$w_2 s_2 = \text{Re}\left(\exp(j \arg(h_{21}))x_2\right), \tag{3c}$$

$$n_1 = \text{Re}(\xi_1), \quad n_2 = \text{Re}\left(\exp\left(j \arg\left(\frac{h_{21}}{h_{22}}\right)\right)\xi_2\right); \tag{3d}$$

$$y_1' = \text{Im}(z_1), \quad y_2' = \text{Im}\left(\exp\left(j \arg\left(\frac{h_{21}}{h_{22}}\right)\right)z_2\right), \tag{3e}$$

$$w_1' s_1' = \text{Im}\left(\exp(j \arg(h_{11}))x_1\right), \tag{3f}$$

$$w_2' s_2' = \text{Im}\left(\exp(j \arg(h_{21}))x_2\right), \tag{3g}$$

[1]Our design can also be generalized to the case with peak power constraint straightforwardly.

$$n'_1 = \text{Im}(\xi_1), \quad n'_2 = \text{Im}\big(\exp(j \arg\big(\tfrac{h_{21}}{h_{22}}\big))\xi_2\big), \quad (3h)$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ are the real and imaginary parts of the complex number, respectively. We also assume that $s_1 \in \mathcal{A}_{M_1}$, $s'_1 \in \mathcal{A}_{M'_1}$, sent by S_1, and $s_2 \in \mathcal{A}_{M_2}$, $s'_2 \in \mathcal{A}_{M'_2}$, transmitted by S_2, are the information-bearing symbols, which are drawn from standard PAM constellation with equal probability, in which $\mathcal{A}_M \triangleq \{\pm (2k-1)\}_{k=1}^{M/2}$ is a M-ary PAM constellation set. Moreover, the scaling factors w_1, w_2, w'_1, and w'_2 are real positive scalars that determine the minimum Euclidean distance of the corresponding PAM constellation set.

From (3), we obtain $\exp\big(j \arg(h_{11})\big)x_1 = (w_1 s_1 + w'_1 s'_1 j) \in Q_1$ and $\exp\big(j \arg(h_{21})\big)x_2 = (w_2 s_2 + w'_2 s'_2 j) \in Q_2$, where

$$Q_1 \triangleq \{\pm w_1(2k-1) \pm w'_1(2\ell-1)j :$$
$$k = 1, \ldots, M_1/2, \ell = 1, \ldots, M'_1/2\}, \quad (4a)$$
$$Q_2 \triangleq \{\pm w_2(2k-1) \pm w'_2(2\ell-1)j :$$
$$k = 1, \ldots, M_2/2, \ell = 1, \ldots, M'_2/2\}, \quad (4b)$$

are $M_1 M'_1$- and $M_2 M'_2$-ary QAM constellations, respectively. If $w_1 = w'_1$ and $w_2 = w'_2$, we call Q_1 and Q_2 symmetric QAM constellations. Otherwise, we are using unsymmetric signaling [45], [48], [55]. In addition, $n_1, n_2, n'_1, n'_2 \sim \mathcal{N}(0, \sigma^2)$ are i.i.d. real additive white Gaussian since the complex noise terms are assumed to be CSCG. Then, the in-phase and quadrature sub-channels of (1), as illustrated in Fig. 1(b), can be reformulated by

$$y_1 = |h_{11}|w_1 s_1 + |h_{21}|w_2 s_2 + n_1,$$
$$y_2 = |h_{22}|w_2 s_2 + n_2, \quad (5a)$$
$$y'_1 = |h_{11}|w'_1 s'_1 + |h_{21}|w'_2 s'_2 + n'_1,$$
$$y'_2 = |h_{22}|w'_2 s'_2 + n'_2. \quad (5b)$$

The transmitted signals over both subchannels should still be subject to average power constraints, i.e., $\mathbb{E}[w_1^2|s_1|^2] \le p_1$, $\mathbb{E}[w_2^2|s_2|^2] \le p_2$, $\mathbb{E}[w_1'^2|s'_1|^2] \le p'_1$, $\mathbb{E}[w_2'^2|s'_2|^2] \le p'_2$ such that $p_1 + p'_1 = P_1$ and $p_2 + p'_2 = P_2$. The following power allocation among the in-phase and quadrature components is normally performed to balance the minimum Euclidean distance of the two PAM constellations [51, Ch.6.1.4], i.e.,

$$p_1 = \frac{(M_1^2 - 1)P_1}{M_1^2 + M_1'^2 - 2}, \quad p_2 = \frac{(M_2^2 - 1)P_2}{M_2^2 + M_2'^2 - 2}, \quad (6a)$$
$$p'_1 = \frac{(M_1'^2 - 1)P_1}{M_1^2 + M_1'^2 - 2}, \quad p'_2 = \frac{(M_2'^2 - 1)P_2}{M_2^2 + M_2'^2 - 2}. \quad (6b)$$

It can be observed that, if square-QAM constellations are used at both transmitters with $M_1 = M'_1$ and $M_2 = M'_2$, we have $p_1 = p'_1 = P_1/2$ and $p_2 = p'_2 = P_2/2$.

An important problem for the considered ZC is that for any given QAM constellation sizes of both messages, how to optimize the values of scaling coefficients w_1, w_2, w'_1 and w'_2 to minimize the average error probability at both receivers, subject to the individual average power constraint at both transmitters. By leveraging the decomposable property of the complex Gaussian ZC and the symmetry of the two subchannels, we can simply focus on the design for one of

the two real Gaussian ZCs with PAM constellation sets, which will be elaborated in the next section.[2]

III. THE CONSTELLATION DESIGN FOR THE REAL GAUSSIAN ZC

In this section, we consider the constellation design problem, i.e., finding the optimal values of w_1 and w_2 for the in-phase real Gaussian ZC characterized by (5a). The optimal solution to the quadrature component can be obtained in a similar fashion and hence omitted for brevity. In particular, if $M_1 = M'_1$ and $M_2 = M'_2$, then the two sub-channels are identical. It is worth noting that, similar design for BC or MAC can be included as a special case of our proposed design for the considered ZC.

A. Problem Formulation

As the first effort towards the design of NOMA with finite-alphabet inputs in ZC, in this paper we concentrate on the case that $M_1 = M_2 = M$ and $M'_1 = M'_2 = M'$. As a result, we have $s_1, s_2 \in \mathcal{A}_M = \{\pm (2k-1)\}_{k=1}^{M/2}$. As $\mathbb{E}[w_1^2|s_1|^2] \le p_1$ and $\mathbb{E}[w_2^2|s_2|^2] \le p_2$, we thus have $0 < w_1 \le \sqrt{\frac{3 \, p_1}{M^2-1}}$ and $0 < w_2 \le \sqrt{\frac{3 \, p_2}{M^2-1}}$.

In our scheme, the transmitted signal from S_1 and S_2 are superimposed together at D_1, which is inherently a non-orthogonal transmission. In line with [36], we use a *joint decoding*[3] at the receiver D_1 since the error performance is dominated by the minimum Euclidean of the resulting sum-constellation at D_1. We assume that each receiver uses a coherent maximum-likelihood (ML) detector to estimate the transmitted signals in a symbol-by-symbol fashion.[4] For receivers D_1 and D_2, the estimated signals can be expressed as

$$(\hat{s}_1, \hat{s}_2) = \arg \min_{(s_1, s_2)} \big| y_1 - (|h_{11}|w_1 s_1 + |h_{21}|w_2 s_2) \big|,$$
$$\hat{s}'_2 = \arg \min_{s_2} \big| y_2 - |h_{22}|w_2 s_2 \big|.$$

By applying the nearest neighbor approximation method [51, Ch. 6.1.4] at high SNRs for the ML receiver, the average error rate is dominated by the minimum Euclidean distance of the received constellation points owing to the exponential decaying of the Gaussian distribution. To balance the error performance of both receivers, in this paper, we aim to devise the optimal value of w_1 and w_2 by applying the max-min fairness criterion on the minimum Euclidean

---

[2]It should be pointed out that this design is a practical but not necessarily optimal approach, which has been widely adopted in practice [36], [45], [48], [52]–[57].

[3]We note that, for ZIC, the joint decoder used by D_1 may not necessarily be the most efficient one. Instead, we should use a treat-interference-as-noise (TIN) receiver when the channel gain of the cross link is very low and use a successive-interference-cancellation (SIC) receiver when the channel gain of cross link is very strong compared with the direct link [36], [43], [44]. In general, a joint decoder can be used when the cross link is moderately strong [36], [43], which will result in a similar design as our case. However, how to extend our design to ZIC with finite-alphabet input is still an open problem and has been left as a future work.

[4]Since we are doing a symbol-by-symbol detection, the decoding complexity is $O(M^2)$ with M being the PAM constellation size of S_1 and S_2, respectively. Although we can use the message passing algorithm (MPA) [58] to further decrease the decoding complexity, however, our method is feasible.

distance of the received signal constellation points among both receivers.

The Euclidean distance between two received signals $y_1(s_1, s_2)$ and $y_1(\tilde{s}_1, \tilde{s}_2)$ at D_1 and that between $y_2(s_2)$ and $y_2(\tilde{s}_2)$ at D_2 for the transmitted signal vectors (s_1, s_2) and $(\tilde{s}_1, \tilde{s}_2)$ at S_1 and S_2 in the noise-free case are given, respectively, by

$$|y_1(s_1, s_2) - y_1(\tilde{s}_1, \tilde{s}_2)| = \big||h_{11}|w_1(s_1 - \tilde{s}_1)$$
$$- |h_{21}|w_2(\tilde{s}_2 - s_2)\big|, \quad (7a)$$
$$|y_2(s_2) - y_2(\tilde{s}_2)| = |h_{22}|w_2|s_2 - \tilde{s}_2|. \quad (7b)$$

Note that $s_1, \tilde{s}_1, s_2$ and $\tilde{s}_2$ are all odd numbers, and we thus can let $s_1 - \tilde{s}_1 = 2n$ and $\tilde{s}_2 - s_2 = 2m$, in which $m, n \in \mathbb{Z}_{M-1}$ with $\mathbb{Z}_{M-1} \triangleq \{0, \pm 1, \ldots, \pm M - 1\}$ denoting the set containing all the possible differences. Similarly, we also define $\mathbb{Z}_{M-1}^2 \triangleq \{(a, b) : a, b \in \mathbb{Z}_{M-1}\}$, $\mathbb{N}_{M-1} \triangleq \{0, 1, \cdots, M - 1\}$ and $\mathbb{N}_{M-1}^2 \triangleq \{(a, b) : a, b \in \mathbb{N}_{M-1}\}$. From the above definition, $(s_1, s_2) \neq (\tilde{s}_1, \tilde{s}_2)$ is equivalent to $(m, n) \neq (0, 0)$. Here, by $(m, n) \neq (0, 0)$, we mean that $m \neq 0$ or $n \neq 0$. To proceed, we define

$$d_1(m, n) = \frac{1}{2}|y_1(s_1, s_2) - y_1(\tilde{s}_1, \tilde{s}_2)|$$
$$= \big||h_{11}|w_1 \, n - |h_{21}|w_2 \, m\big|,$$
$$\text{for } (m, n) \in \mathbb{Z}_{M-1}^2 \setminus \{(0, 0)\},$$
$$d_2(m) = \frac{1}{2}|y_2(s_2) - y_2(\tilde{s}_2)|$$
$$= |h_{22}|w_2|m|, \quad \text{for } m \in \mathbb{Z}_{M-1} \setminus \{0\}.$$

We are now ready to formally formulate the following max-min optimization problem:

Problem 1 (Optimal Design of NOMA in Real Scalar ZC With PAM Constellation): Find the optimal value of (w_1^*, w_2^*) subject to the individual average power constraint such that the minimum value of the minimum Euclidean distances of the received signal constellation points over both received signals is maximized, i.e.,

$$(w_1^*, w_2^*) = \arg \max_{(w_1, w_2)} \min \Big\{ \underbrace{\min_{(m,n) \in \mathbb{Z}_{M-1}^2 \setminus \{(0,0)\}} d_1(m, n)}_{T_1},$$
$$\underbrace{\min_{m \in \mathbb{Z}_{M-1} \setminus \{0\}} d_2(m)}_{T_2} \Big\} \quad (9a)$$

$$\text{s.t. } 0 < w_1 \le \sqrt{\frac{3p_1}{M^2 - 1}}$$
$$\text{and } 0 < w_2 \le \sqrt{\frac{3p_2}{M^2 - 1}}. \quad (9b)$$

∎

Note that the inner optimization problem of finding the minimum Euclidean distances is a discrete one, while the outer optimization problem on (w_1, w_2) is a continuous problem. In other words, Problem 1 is a *mixed continuous-discrete* optimization problem and it is in general hard to solve. To the best of our knowledge, only numerical solutions to such kind of problems are available in the literature, see e.g., [30]–[32] and references therein.

To optimally and systematically solve this problem, we now develop a novel framework based on the *Farey sequence* (also known as Farey series) [49], which can divide the entire feasible region of (w_1, w_2) into a finite number of mutually exclusive sub-intervals. Then for each sub-interval, the formulated optimization problem can be solved optimally with a closed-form solution, and subsequently the overall maximum value of Problem 1 can be attained by taking the maximum value of the objective function among all the sub-intervals.

For the inner optimization problem of T_2 given in (9a), it can be observed that

$$\min_{m \in \mathbb{Z}_{M-1} \setminus \{0\}} d_2(m) = \min_{m \in \mathbb{Z}_{M-1} \setminus \{0\}} |h_{22}|w_2|m|$$
$$= |h_{22}|w_2, \text{ with } m = 1.$$

However, for T_1, we have

$$\min_{(m,n) \in \mathbb{Z}_{M-1}^2 \setminus \{(0,0)\}} d_1(m, n)$$
$$= \min_{(m,n) \in \mathbb{Z}_{M-1}^2 \setminus \{(0,0)\}} \big||h_{11}|w_1 n - |h_{21}|w_2 m\big|. \quad (10)$$

We should point out that the closed-form solution to the optimal (m, n) is not trivial, since the solution depends on the values of $|h_{11}|$ and $|h_{22}|$, which can span the whole positive real axis. Moreover, the value of w_1 and w_2 can not be determined beforehand. Actually, the problem in (10) is essentially equivalent to finding a real rational number with finite order to approximate a real irrational number as closely as possible. This naturally leads us to resorting to the Farey sequence, which particularly plays a critical role in solving such kind of problems [49]. In the subsequent section, we will introduce the definition and some important properties of Farey sequence.

B. Farey Sequence

The Farey sequence characterizes the relationship between two positive integers and the formal definition is given as follows:

Definition 1 (Farey Sequence [49]): The Farey sequence $\mathfrak{F}_K$ of order K is the ascending sequence of irreducible fractions between 0 and 1 whose denominators are less than or equal to K. ∎

By the definition, $\mathfrak{F}_K = \big(\frac{b_k}{a_k}\big)_{k=1}^{|\mathfrak{F}_K|}$ is a sequence of fractions $\frac{b_k}{a_k}$ such that $0 \le b_k \le a_k \le K$ and $\langle a_k, b_k \rangle = 1$ arranged in an increasing order, where $\langle a, b \rangle$ denotes the largest common divider of non-negative integers a, b. $|\mathfrak{F}_K| = 1 + \sum_{m=1}^K \varphi(m)$ is the cardinality of $\mathfrak{F}_K$ with $\varphi(\cdot)$ being the Euler's totient function [49]. Some examples of Farey sequences are given as follows:

Example 1: $\mathfrak{F}_5$ is the ordered sequence

$$\Big(\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}\Big).$$

It can be observed that each Farey sequence begins with number 0 (fraction $\frac{0}{1}$) and ends with 1 (fraction $\frac{1}{1}$). The series of breakpoints after $\frac{1}{1}$ is the reciprocal version of the Farey sequence. We call the Farey number sequence together with

its reciprocal version as the *extended Farey sequence* which is formally defined as follows:

Definition 2 (Extended Farey Sequence): The extended Farey sequence $\mathfrak{S}_K$ of order K is the sequence of ascending irreducible fractions, where the maximum value of the numerator and denominator do not exceed K. ∎

From the definition, we have $\mathfrak{S}_K = \left(\frac{b_k}{a_k}\right)_{k=1}^{|\mathfrak{S}_K|}$ with $\langle a_k, b_k \rangle = 1$ and $|\mathfrak{S}_K| = 1 + 2\sum_{m=1}^{K}\varphi(m)$. We have the following example:

Example 2: $\mathfrak{S}_5$ is the sequence

$$\left(\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1},\right.$$
$$\left.\frac{5}{4}, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, \frac{2}{1}, \frac{5}{2}, \frac{3}{1}, \frac{4}{1}, \frac{5}{1}, \frac{1}{0}\right).$$

It can be observed that the extended Farey sequence starts with number 0 (fraction $\frac{0}{1}$) and end with ∞ (fraction $\frac{1}{0}$).

The positive real axis can be divided by the extended Farey sequence $\mathfrak{S}_K$ into a *finite number* (i.e., $|\mathfrak{S}_K| - 1 = 2\sum_{m=1}^{K}\varphi(m)$) of intervals. In this paper, we call the fractions consisting of adjacent terms in the extended Farey sequence as a *Farey pair*, and the interval between the Farey pair is referred to as a *Farey interval*. We then have the Farey interval set formally defined as follows:

Definition 3 (Farey Interval Set): A Farey interval set $\mathcal{S}_K$ of order K is the set containing all the Farey intervals generated by the Farey pair of the extended Farey sequence $\mathfrak{S}_K$. ∎

By the definition, we have $\mathcal{S}_K = \left\{\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)\right\}_{k=1}^{|\mathcal{S}_K|}$, where $|\mathcal{S}_K| = |\mathfrak{S}_K| - 1 = 2\sum_{m=1}^{K}\varphi(m)$. Note that, with a slight abuse of notations, $\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$ denotes the interval between end nodes $\frac{b_k}{a_k}$ and $\frac{b_{k+1}}{a_{k+1}}$ rather than a vector, and this will be clear from the context.

Example 3: The Farey interval set $\mathcal{S}_5$ is the set given by

$$\left\{\left(\frac{0}{1}, \frac{1}{5}\right), \left(\frac{1}{5}, \frac{1}{4}\right), \left(\frac{1}{4}, \frac{1}{3}\right), \ldots, \left(\frac{3}{1}, \frac{4}{1}\right), \left(\frac{4}{1}, \frac{5}{1}\right), \left(\frac{5}{1}, \frac{1}{0}\right)\right\}.$$

The Farey interval set can be further divided into two subsets $\mathcal{U}_K^L = \left\{\left(\frac{b_k'}{a_k'}, \frac{b_{k+1}'}{a_{k+1}'}\right) \in \mathcal{S}_K : b_k' + b_{k+1}' \geq L\right\}$ and $\mathcal{V}_K^L = \left\{\left(\frac{b_k''}{a_k''}, \frac{b_{k+1}''}{a_{k+1}''}\right) \in \mathcal{S}_K : b_k'' + b_{k+1}'' < L\right\}$ for $L = 1, 2, \ldots, 2K$ such that $\mathcal{S}_K = \mathcal{U}_K^L \cup \mathcal{V}_K^L$ and $\mathcal{U}_K^L \cap \mathcal{V}_K^L = \varnothing$. In particular, $\mathcal{U}_K^1 = \mathcal{S}_K$ and $\mathcal{V}_K^1 = \varnothing$ while $\mathcal{U}_K^{2K} = \varnothing$ and $\mathcal{V}_K^{2K} = \mathcal{S}_K$.

Example 4: For the Farey interval set $\mathcal{S}_5$, we have

$$\mathcal{U}_5^4 = \left\{\left(\frac{1}{2}, \frac{3}{5}\right), \left(\frac{3}{5}, \frac{2}{3}\right), \left(\frac{2}{3}, \frac{3}{4}\right), \left(\frac{3}{4}, \frac{4}{5}\right), \left(\frac{4}{5}, \frac{1}{1}\right),\right.$$
$$\left(\frac{1}{1}, \frac{5}{4}\right), \left(\frac{5}{4}, \frac{4}{3}\right), \left(\frac{4}{3}, \frac{3}{2}\right), \left(\frac{3}{2}, \frac{5}{3}\right), \left(\frac{5}{3}, \frac{2}{1}\right),$$
$$\left.\left(\frac{2}{1}, \frac{5}{2}\right), \left(\frac{5}{2}, \frac{3}{1}\right), \left(\frac{3}{1}, \frac{4}{1}\right), \left(\frac{4}{1}, \frac{5}{1}\right), \left(\frac{5}{1}, \frac{1}{0}\right)\right\},$$

and

$$\mathcal{V}_5^4 = \left\{\left(\frac{0}{1}, \frac{1}{5}\right), \left(\frac{1}{5}, \frac{1}{4}\right), \left(\frac{1}{4}, \frac{1}{3}\right), \left(\frac{1}{3}, \frac{2}{5}\right), \left(\frac{2}{5}, \frac{1}{2}\right)\right\}.$$

We now review some elementary properties of Farey sequences [49] which are also true for extended Farey sequences.

Property 1: If $\frac{n_1}{m_1}$ and $\frac{n_2}{m_2}$ are two adjacent terms (Farey pairs) of $\mathfrak{S}_K$ with $K \geq 1$, such that $\frac{n_1}{m_1} < \frac{n_2}{m_2}$, then,
1) $m_1 n_2 - m_2 n_1 = 1$.
2) $\frac{n_1+n_2}{m_1+m_2} \in \left(\frac{n_1}{m_1}, \frac{n_2}{m_2}\right)$, $\frac{m_1+m_2}{n_1+n_2} \in \left(\frac{m_2}{n_2}, \frac{m_1}{n_1}\right)$. ∎

Property 2: If $\frac{n_1}{m_1}$, $\frac{n_2}{m_2}$ and $\frac{n_3}{m_3}$ are three consecutive terms of $\mathfrak{S}_K$ with $K \geq 1$ such that $\frac{n_1}{m_1} < \frac{n_2}{m_2} < \frac{n_3}{m_3}$, then, $\frac{n_2}{m_2} = \frac{n_1+n_3}{m_1+m_3}$. ∎

Property 3: Given $K \geq 2$, we assume $\frac{n_1}{m_1}, \frac{n_2}{m_2}, \frac{n_3}{m_3}, \frac{n_4}{m_4} \in \mathfrak{S}_K$ and $\frac{n_1}{m_1} < \frac{n_2}{m_2} < \frac{n_3}{m_3} < \frac{n_4}{m_4}$. If $\frac{n_2}{m_2}$ and $\frac{n_3}{m_3}$ form one Farey pair, then $\frac{n_1+n_3}{m_1+m_3} \leq \frac{n_2}{m_2}$ and $\frac{n_3}{m_3} \leq \frac{n_2+n_4}{m_2+m_4}$. ∎

C. The Minimum Euclidean Distance of the Received Signal Constellation Points

We are now ready to solve the problem in (10) to find the constellation point pairs (m, n) that have the minimum Euclidean distance. To that end, we first have the following preliminary propositions.

Proposition 1: Let $\mathbb{F}_K^2 = \{(m, n) : \frac{n}{m} \in \mathfrak{S}_K\}$, where $\mathfrak{S}_K$ is the extended Farey number sequence of order K, and then we have

$$\min_{(m,n)\in\mathbb{Z}_K^2 \setminus \{(0,0)\}} d_1(m, n) = \min_{(m,n)\in\mathbb{F}_K^2} d_1(m, n).$$

∎

The proof is given in Appendix-A.

Proposition 2: Consider the Farey interval $\left(\frac{n_1}{m_1}, \frac{n_2}{m_2}\right) \in \mathcal{S}_K$, with $K \geq 1$, and $\frac{n_1}{m_1} < \frac{n_2}{m_2}$. Then, for $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in \left(\frac{n_1}{m_1}, \frac{n_2}{m_2}\right)$ and $d_1(m, n) = ||h_{11}|w_1 n - |h_{21}|w_2 m|$, we have

1) If $\frac{|h_{21}|w_2}{|h_{11}|w_1} = \frac{n_1+n_2}{m_1+m_2}$, then $d_1(m_1, n_1) = d_1(m_2, n_2)$;
2) If $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in \left(\frac{n_1}{m_1}, \frac{n_1+n_2}{m_1+m_2}\right)$, then $d_1(m_1, n_1) < d_1(m_2, n_2)$;
3) If $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in \left(\frac{n_1+n_2}{m_1+m_2}, \frac{n_2}{m_2}\right)$, then $d_1(m_1, n_1) > d_1(m_2, n_2)$. ∎

The proof can be found in Appendix-B.

Proposition 3: Consider $\frac{n_1}{m_1}, \frac{n_2}{m_2}, \frac{n_3}{m_3}, \frac{n_4}{m_4} \in \mathfrak{S}_K$, with $K \geq 2$, such that $\frac{n_1}{m_1} < \frac{n_2}{m_2} < \frac{n_3}{m_3} < \frac{n_4}{m_4}$ where $\frac{n_2}{m_2}, \frac{n_3}{m_3}$ form one Farey pair,

1) If $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in \left(\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}\right)$, then

$$\min_{(m,n)\in\mathbb{F}_{M-1}^2} d_1(m, n) = d_1(m_2, n_2)$$
$$= |h_{21}|w_2 m_2 - |h_{11}|w_1 n_2.$$

2) If $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in \left(\frac{n_2+n_3}{m_2+m_3}, \frac{n_3}{m_3}\right)$, then

$$\min_{(m,n)\in\mathbb{F}_{M-1}^2} d_1(m, n) = d_1(m_3, n_3)$$
$$= |h_{11}|w_1 n_3 - |h_{21}|w_2 m_3.$$

∎

The proof is provided in Appendix-C.

Proposition 4: For the Farey interval set $\mathcal{S}_{M-1} = \left\{\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)\right\}_{k=1}^{|\mathcal{S}_{M-1}|}$, if $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, then $\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|} \leq \frac{a_k+a_{k+1}}{b_k+b_{k+1}} \leq \frac{a_k}{b_k} - \frac{|h_{22}|}{b_k|h_{21}|}$. ∎

The proof is given in Appendix-D.

D. Optimal Solution to Problem 1 for $\frac{|h_{21}|w_2}{|h_{11}|w_1}$ in Certain Farey Interval

In this section, we solve Problem 1 by restricting $\frac{|h_{21}|w_2}{|h_{11}|w_1}$ into a certain Farey interval where a closed-form solution is attainable. We consider the Farey interval set $\mathcal{S}_{M-1}$ given by $\mathcal{S}_{M-1} = \{(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}})\}_{k=1}^{|\mathcal{S}_{M-1}|}$ where $|\mathcal{S}_{M-1}| = 2\sum_{m=1}^{M-1}\varphi(m)$. Now we consider the case $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in (\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}})$, $k = 1, 2, \ldots, |\mathcal{S}_{M-1}|$ and we aim to find the optimal $(w_1^*(k), w_2^*(k))$ such that

$$g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \max_{(w_1, w_2)} \min\left\{\min_{(m,n)\in\mathbb{F}_{M-1}^2} d_1(m, n),\right.$$

$$\left.\min_{m\in\mathbb{Z}\setminus\{0\}} d_2(m)\right\} \quad (11a)$$

$$\text{s.t.} \quad \frac{b_k}{a_k} < \frac{|h_{21}|w_2}{|h_{11}|w_1} \leq \frac{b_{k+1}}{a_{k+1}},$$

$$0 < w_1 \leq \sqrt{\frac{3p_1}{M^2-1}}$$

$$\text{and } 0 < w_2 \leq \sqrt{\frac{3p_2}{M^2-1}}. \quad (11b)$$

By applying the propositions in last subsections, we manage to attain the following lemma in terms of the optimal solution to problem (11).

Lemma 1: The optimal solution to (11) is given as follows.

1) If $\frac{|h_{21}|}{|h_{22}|} \leq b_k + b_{k+1}$, the following statements are true:

  a) If $\frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}(a_k+a_{k+1})}{\sqrt{p_1}(b_k+b_{k+1})}$, then $g(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}) = \frac{|h_{21}|}{b_k+b_{k+1}}\sqrt{\frac{3p_2}{M^2-1}}$ and $(w_1^*(k), w_2^*(k)) = \left(\frac{(a_k+a_{k+1})|h_{21}|}{(b_k+b_{k+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right)$.

  b) If $\frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}(a_k+a_{k+1})}{\sqrt{p_1}(b_k+b_{k+1})}$, then $g(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}) = \frac{|h_{11}|}{a_k+a_{k+1}}\sqrt{\frac{3p_1}{M^2-1}}$ and $(w_1^*(k), w_2^*(k)) = \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{(b_k+b_{k+1})|h_{11}|}{(a_k+a_{k+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2-1}}\right)$.

2) If $\frac{|h_{21}|}{|h_{22}|} > b_k + b_{k+1}$, we have the following results:

  a) If $\frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}}{\sqrt{p_1}}(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|})$, then $g(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}) = |h_{22}|\sqrt{\frac{3p_2}{M^2-1}}$ and $(w_1^*(k), w_2^*(k)) = \left(\frac{a_{k+1}|h_{21}|+|h_{22}|}{b_{k+1}|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right)$.

  b) If $\frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}}{\sqrt{p_1}}(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|})$, then $g(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}) = \frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}|+|h_{22}|}\sqrt{\frac{3p_1}{M^2-1}}$ and $(w_1^*(k), w_2^*(k)) = \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|+|h_{22}|}\sqrt{\frac{3p_1}{M^2-1}}\right)$.

The proof of Lemma 1 can be found in Appendix-E.

Remark 1: We have the following insights from the above lemma,

1) It can be observed from Lemma 1 that at least one transmitter should employ the maximum allowable power, since otherwise we could scale up both transmitted power without violating the power constraint such that the minimum Euclidean distance is enlarged.

2) We can see from Lemma 1 that the optimal value of the objective function together with $(w_1^*(k), w_2^*(k))$ substantially depend on the relative strength of the

channel coefficients. Inspired by [36], we divide the considered real Gaussian ZC into three scenarios:

  a) Gaussian ZC with *weak cross link*, i.e., $\frac{|h_{21}|}{|h_{22}|} \in (0, 1]$;
  b) Gaussian ZC with *strong cross link*, i.e., $\frac{|h_{21}|}{|h_{22}|} \in (1, 2M)$;
  c) Gaussian ZC with *very strong cross link*, i.e., $\frac{|h_{21}|}{|h_{22}|} \in [2M, \infty)$.

Then, a compact closed-form expression for Problem 1 can be established for three complimentary scenarios, which constitutes main contents of the subsequent subsection. ∎

E. The Optimal NOMA Design With PAM Constellation for the Gaussian ZC

Now we are ready to give the closed-form optimal solution of (w_1^*, w_2^*) to Problem 1 which maximizes the minimum Euclidean distance over all the Farey intervals for the aforementioned three scenarios.

1) Scenario 1 (ZC With Weak Cross Link): For this case, we have $\frac{|h_{21}|}{|h_{22}|} \in (0, 1]$. Consider the Farey interval set $\mathcal{S}_{M-1} = \{(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}})\}_{k=1}^{|\mathcal{S}_{M-1}|}$. By Property 1, we have $\frac{a_{k+1}+a_{k+2}}{b_{k+1}+b_{k+2}} < \frac{a_{k+1}}{b_{k+1}} < \frac{a_k+a_{k+1}}{b_k+b_{k+1}}$, and therefore the positive real axis can be divided into the $|\mathcal{S}_{M-1}| + 1$ intervals in an increasing order:

$$\left\{\left(0, \frac{a_{|\mathcal{S}_{M-1}|} + a_{|\mathcal{S}_{M-1}|+1}}{b_{|\mathcal{S}_{M-1}|} + b_{|\mathcal{S}_{M-1}|+1}}\right), \left(\frac{a_{|\mathcal{S}_{M-1}|} + a_{|\mathcal{S}_{M-1}|+1}}{b_{|\mathcal{S}_{M-1}|} + b_{|\mathcal{S}_{M-1}|+1}},\right.\right.$$
$$\left.\left.\frac{a_{|\mathcal{S}_{M-1}|-1} + a_{|\mathcal{S}_{M-1}|}}{b_{|\mathcal{S}_{M-1}|-1} + b_{|\mathcal{S}_{M-1}|}}\right), \cdots, \left(\frac{a_1 + a_2}{b_1 + b_2}, \infty\right)\right\}.$$

In particular, we have $\frac{a_1+a_2}{b_1+b_2} = M$ and $\frac{a_{|\mathcal{S}_{M-1}|}+a_{|\mathcal{S}_{M-1}|+1}}{b_{|\mathcal{S}_{M-1}|}+b_{|\mathcal{S}_{M-1}|+1}} = \frac{1}{M}$.

Theorem 1 (Gaussian ZC with weak cross link): Suppose that $\frac{|h_{21}|}{|h_{22}|} \in (0, 1]$. Then, the optimal power scaling factors to Problem 1 are explicitly determined as follows:

1) If $\frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}}{M\sqrt{p_1}}$, then, we have $(w_1^*, w_2^*) = \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{M|h_{11}|}{|h_{21}|}\sqrt{\frac{3p_1}{M^2-1}}\right)$;

2) If $\frac{|h_{11}|}{|h_{21}|} \geq \frac{M\sqrt{p_2}}{\sqrt{p_1}}$, then, we have $(w_1^*, w_2^*) = \left(\frac{M|h_{22}|}{|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right)$;

3) Let $\frac{|h_{11}|}{|h_{21}|} \in \left(\frac{\sqrt{p_2}(a_{\ell_1+1}+a_{\ell_1+2})}{\sqrt{p_1}(b_{\ell_1+1}+b_{\ell_1+2})}, \frac{\sqrt{p_2}(a_{\ell_1}+a_{\ell_1+1})}{\sqrt{p_1}(b_{\ell_1}+b_{\ell_1+1})}\right)$ for some $\ell_1 = 1, \cdots, |\mathcal{S}_{M-1}| - 1$. If we let $\tilde{\ell}_a = \arg\min_k\{(a_1 + a_2), \cdots, (a_{\ell_1} + a_{\ell_1+1})\}$ and $\tilde{\ell}_b = \arg\min_k\{(b_{\ell_1+1} + b_{\ell_1+2}), \cdots, (b_{|\mathcal{S}_{M-1}|} + b_{|\mathcal{S}_{M-1}|+1})\}$, then we have

$$(w_1^*, w_2^*)$$

$$= \begin{cases} \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{(b_{\tilde{\ell}_a} + b_{\tilde{\ell}_a+1})|h_{11}|}{(a_{\tilde{\ell}_a} + a_{\tilde{\ell}_a+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}(a_{\tilde{\ell}_a} + a_{\tilde{\ell}_a+1})}{\sqrt{p_1}(b_{\tilde{\ell}_b} + b_{\tilde{\ell}_b+1})}; \\ \left(\frac{(a_{\tilde{\ell}_b} + a_{\tilde{\ell}_b+1})|h_{21}|}{(b_{\tilde{\ell}_b} + b_{\tilde{\ell}_b+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}(a_{\tilde{\ell}_a} + a_{\tilde{\ell}_a+1})}{\sqrt{p_1}(b_{\tilde{\ell}_b} + b_{\tilde{\ell}_b+1})}. \end{cases}$$

∎

The proof of Theorem 1 is provided in Appendix-F.

2) Scenario 2 (ZC With Strong Cross Link): In this case, $\frac{|h_{21}|}{|h_{22}|} \in (1, 2M)$. We suppose that $L - 1 < \frac{|h_{21}|}{|h_{22}|} \leq L$. Then, the optimal solution to Problem 1 can be obtained by considering the following two cases: the Farey interval set $\mathcal{U}_{M-1}^L = \{(\frac{b_k'}{a_k'}, \frac{b_{k+1}'}{a_{k+1}'}) \in \mathcal{S}_{M-1} : b_k' + b_{k+1}' \geq L\}$ and $\mathcal{V}_{M-1}^L = \{(\frac{b_k''}{a_k''}, \frac{b_{k+1}''}{a_{k+1}''}) \in \mathcal{S}_{M-1} : b_k'' + b_{k+1}'' < L\}$. The whole discussions on them can be summarized as the following theorem.

Theorem 2 (ZC With Strong Cross Link): The optimal solution to Problem 1 in this case can be obtained by finding the maximum value of the objective functions in the following two cases which are explicitly attained as follows:

1) Let $(\frac{b_k'}{a_k'}, \frac{b_{k+1}'}{a_{k+1}'}) \in \mathcal{U}_{M-1}^L$. Then, the following three statements are true.

    a) If $\frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}(a'_{|\mathcal{U}_{M-1}^L|} + a'_{|\mathcal{U}_{M-1}^L|+1})}{\sqrt{p_1}(b'_{|\mathcal{U}_{M-1}^L|} + b'_{|\mathcal{U}_{M-1}^L|+1})}$, then we have
$$(w_{u_1}^*, w_{u_2}^*) = \left(\sqrt{\frac{3 p_1}{M^2-1}}, \frac{(b'_{\tilde{\ell}_c} + b'_{\tilde{\ell}_c+1})|h_{11}|}{(a'_{\tilde{\ell}_c} + a'_{\tilde{\ell}_c+1})|h_{21}|}\sqrt{\frac{3 p_1}{M^2-1}}\right),$$
where $\tilde{\ell}_c = \arg\min_k\{(a_1' + a_2'), \ldots, (a'_{|\mathcal{U}_{M-1}^L|} + a'_{|\mathcal{U}_{M-1}^L|+1})\}$.

    b) If $\frac{|h_{11}|}{|h_{21}|} \geq \frac{M\sqrt{p_2}}{\sqrt{p_1}}$, then we have $(w_{u_1}^*, w_{u_2}^*) = \left(\frac{M|h_{21}|}{|h_{11}|}\sqrt{\frac{3 p_2}{M^2-1}}, \sqrt{\frac{3 p_2}{M^2-1}}\right)$.

    c) Suppose that $\frac{|h_{11}|}{|h_{21}|} \in \left(\frac{\sqrt{p_2}(a'_{\ell_2+1} + a'_{\ell_2+2})}{\sqrt{p_1}(b'_{\ell_2+1} + b'_{\ell_2+2})}, \frac{\sqrt{p_2}(a'_{\ell_2} + a'_{\ell_2+1})}{\sqrt{p_1}(b'_{\ell_2} + b'_{\ell_2+1})}\right)$ for some $\ell_2' = 1, \cdots, |\mathcal{U}_{M-1}^L|$. If we let $\tilde{\ell}_d = \arg\min_k\{(a_1 + a_2), \ldots, (a_{\ell_2} + a_{\ell_2+1})\}$ and $\tilde{\ell}_e = \arg\min_k\{(b_{\ell_2+1} + b_{\ell_2+2}), \ldots, (b_{|\mathcal{U}_{M-1}^L|} + b_{|\mathcal{U}_{M-1}^L|+1})\}$, then we have
$$(w_{u_1}^*, w_{u_2}^*)$$
$$= \begin{cases} \left(\sqrt{\frac{3 p_1}{M^2-1}}, \frac{(b'_{\tilde{\ell}_d} + b'_{\tilde{\ell}_d+1})|h_{11}|}{(a'_{\tilde{\ell}_d} + a'_{\tilde{\ell}_d+1})|h_{21}|}\sqrt{\frac{3 p_1}{M^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}(a'_{\tilde{\ell}_d} + a'_{\tilde{\ell}_d+1})}{\sqrt{p_1}(b'_{\tilde{\ell}_e} + b'_{\tilde{\ell}_e+1})}; \\ \left(\frac{(a'_{\tilde{\ell}_e} + a'_{\tilde{\ell}_e+1})|h_{21}|}{(b'_{\tilde{\ell}_e} + b'_{\tilde{\ell}_e+1})|h_{11}|}\sqrt{\frac{3 p_2}{M^2-1}}, \sqrt{\frac{3 p_2}{M^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}(a'_{\tilde{\ell}_d} + a'_{\tilde{\ell}_d+1})}{\sqrt{p_1}(b'_{\tilde{\ell}_e} + b'_{\tilde{\ell}_e+1})}. \end{cases}$$

2) Let $(\frac{b_{k-1}''}{a_{k-1}''}, \frac{b_k''}{a_k''}) \in \mathcal{V}_{M-1}^L$. Then, the following two statements are true.

    a) If $\frac{|h_{11}|}{|h_{22}|} \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}$, then we have $(w_{v_1}^*, w_{v_2}^*) = \left(\sqrt{\frac{3 p_1}{M^2-1}}, \frac{b''_{|\mathcal{V}_{M-1}|+1}|h_{11}||h_{22}|}{a''_{|\mathcal{V}_{M-1}|+1}|h_{21}|+|h_{22}|}\sqrt{\frac{3 p_1}{M^2-1}}\right)$.

    b) If $\frac{|h_{11}|\sqrt{p_1}}{|h_{21}|\sqrt{p_2}} \in \left(\frac{a''_{\ell_3+1}}{b''_{\ell_3+1}} + \frac{|h_{22}|}{b''_{\ell_3+1}|h_{21}|}, \frac{a''_{\ell_3}}{b''_{\ell_3}} + \frac{|h_{22}|}{b''_{\ell_3}|h_{21}|}\right)$ for some $\ell_3 = 1, \cdots, |\mathcal{V}_{M-1}|$, then we

have
$$(w_{v_1}^*, w_{v_2}^*)$$
$$= \begin{cases} \left(\sqrt{\frac{3 p_1}{M^2-1}}, \frac{b''_{\ell_3}|h_{11}|}{a''_{\ell_3}|h_{21}|+|h_{22}|}\sqrt{\frac{3 p_1}{M^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a''_{\ell_3}}{b''_{\ell_3}} + \frac{|h_{22}|}{b''_{\ell_3}|h_{21}|}\right); \\ \left(\frac{a''_{\ell_3+1}|h_{21}|+|h_{22}|}{b''_{\ell_3+1}|h_{11}|}\sqrt{\frac{3 p_2}{M_2^2-1}}, \sqrt{\frac{3 p_2}{M_2^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{\ell_3''}}{b''_{\ell_3}} + \frac{|h_{22}|}{b''_{\ell_3}|h_{21}|}\right). \end{cases}$$

∎

The proof of Theorem 2 is very similar to that of Theorem 1 and the following Theorem 3 and thus, is omitted due to space limitation.

3) Scenario 3 (ZC With Very Strong Cross Link): In this case, $\frac{|h_{21}|}{|h_{22}|} \in [2M, \infty)$. Likewise, we consider the Farey interval set $\mathcal{S}_{M-1} = \{(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}})\}_{k=1}^{|\mathcal{S}_{M-1}|}$. Note that $\frac{|h_{21}|}{|h_{22}|} \geq 2M > b_k + b_{k+1}$ for $k = 1, \ldots, |\mathcal{S}_{M-1}|$. Then, by Property 1 and Proposition 4, we have $\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|} \leq \frac{a_k + a_{k+1}}{b_k + b_{k+1}} < \frac{a_k}{b_k} < \frac{a_k}{b_k} + \frac{|h_{22}|}{b_k|h_{21}|}$. As a result, the positive real axis can be divided into the following intervals in increasing order:
$$\left\{\left(0, \frac{a_{|\mathcal{S}_{M-1}|+1}}{b_{|\mathcal{S}_{M-1}|+1}} + \frac{|h_{22}|}{b_{|\mathcal{S}_{M-1}|+1}|h_{21}|}\right), \right.$$
$$\left. \ldots, \left(\frac{a_2}{b_2} + \frac{|h_{22}|}{b_2|h_{21}|}, \frac{a_1}{b_1} + \frac{|h_{22}|}{b_1|h_{21}|}\right)\right\},$$
where $\frac{a_1}{b_1} + \frac{|h_{22}|}{b_1|h_{21}|} = \infty$.

Theorem 3 (Gaussian ZC With Very Strong Cross Link): Let $\frac{|h_{21}|}{|h_{22}|} \in [2M, \infty)$. Then, the optimal solution to Problem 1 is given below:

1) If $\frac{|h_{11}|}{|h_{22}|} \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}$, then we have $(w_1^*, w_2^*) = \left(\sqrt{\frac{3 p_1}{M^2-1}}, \frac{|h_{11}|}{|h_{22}|}\sqrt{\frac{3 p_1}{M^2-1}}\right)$.

2) If $\frac{|h_{11}|\sqrt{p_1}}{|h_{21}|\sqrt{p_2}} \in \left(\frac{a_{\ell_4+1}}{b_{\ell_4+1}} + \frac{|h_{22}|}{b_{\ell_4+1}|h_{21}|}, \frac{a_{\ell_4}}{b_{\ell_4}} + \frac{|h_{22}|}{b_{\ell_4}|h_{21}|}\right)$ for some $\ell_4 = 1, \ldots, |\mathcal{S}_{M-1}|$, then we have
$$(w_1^*, w_2^*) = \begin{cases} \left(\sqrt{\frac{3 p_1}{M^2-1}}, \frac{b_{\ell_4}|h_{11}|}{a_{\ell_4}|h_{21}|+|h_{22}|}\sqrt{\frac{3 p_1}{M^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{\ell_4}}{b_{\ell_4}} + \frac{|h_{22}|}{b_{\ell_4}|h_{21}|}\right); \\ \left(\frac{a_{\ell_4+1}|h_{21}|+|h_{22}|}{b_{\ell_4+1}|h_{11}|}\sqrt{\frac{3 p_2}{M^2-1}}, \sqrt{\frac{3 p_2}{M^2-1}}\right), \\ \qquad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{\ell_4}}{b_{\ell_4}} + \frac{|h_{22}|}{b_{\ell_4}|h_{21}|}\right). \end{cases}$$

∎

The proof of Theorem 3 is provided in Appendix-G.

IV. SIMULATION RESULTS AND DISCUSSIONS

In this section, computer simulations are carried out to demonstrate the effectiveness of our proposed NOMA design under different channel configurations. More precisely, we compare our proposed NOMA design with CR based NOMA [30], time-division multiple access (TDMA) and
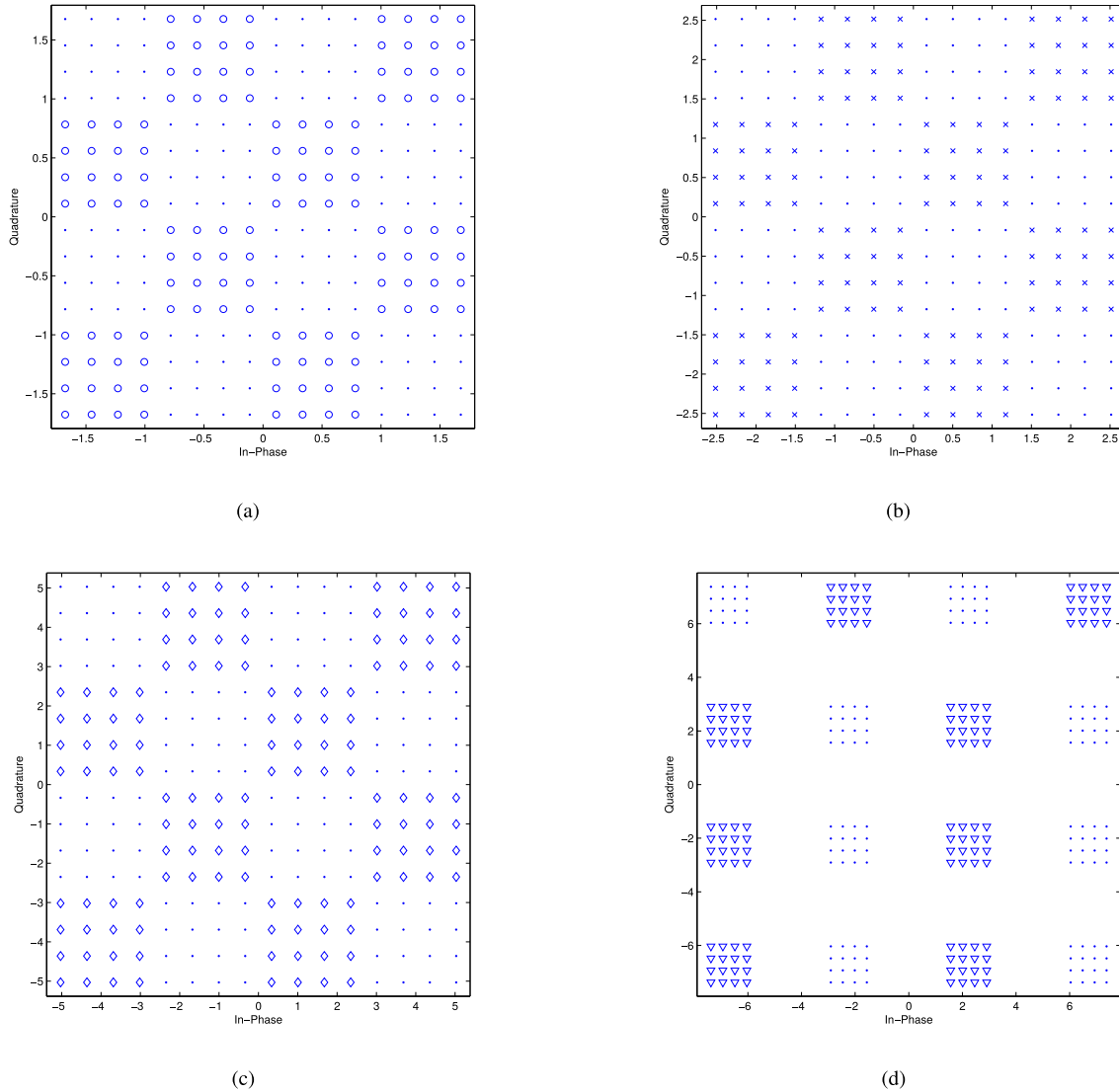
(a)

(b)

(c)

(d)

Fig. 2. Optimal resulting constellations at receiver D_1 for various cases: (a) Weak cross link with $h_{11} = 1$, $h_{21} = 1/2$ and $h_{22} = 1$. (b) Strong cross link with $h_{11} = 1$, $h_{21} = 3/2$ and $h_{22} = 1$. (c) Strong cross link with $h_{11} = 1$, $h_{21} = 3$ and $h_{22} = 1$ and (d) Very strong cross link with $h_{11} = 1$, $h_{21} = 5$ and $h_{22} = 1/2$.

frequency-division multiple access (FDMA) approaches. Without loss of generality, we set $P_1 = P_2 = 1$. For simplicity, in the simulations, we assume that the same square-QAM constellation is adopted by both users, i.e., $M = M'$ and according to (6), we have $p_1 = p'_1 = P_1/2 = 1/2$ and $p_2 = p'_2 = P_2/2 = 1/2$.

A. The Resulting Optimal Sum Constellation at Receiver D_1 For Several Deterministic Channels

Here, we consider several deterministic channels corresponding to the three scenarios of *weak cross link*, *strong cross link* and *very strong cross link*. We assume that 16-QAM constellations are used by both users with $M = M' = 4$. We discuss the results of these cases one by one as follows:

1) Weak Cross Link: In this case, we assume that $h_{11} = 1$, $h_{21} = 1/2$ and $h_{22} = 1$. Based on the derived expressions for the optimal solution provided in the previous section,

we can readily obtain that $(w_1^*, w_2^*) = (0.4472, 0.2236)$. The corresponding received constellation at D_1 is plotted in Fig. 2(a). It can be observed that the sum-constellation at D_1 is a regular 256-QAM generated by the superposition of two 16-QAM. Hereafter, we call the signal constellation with smaller minimum Euclidean distance as the satellite constellation. By observing that $w_1^* h_{11} = 0.4472$ and $w_2^* h_{21} = 0.1118$ (i.e., $w_2^* h_{21} = \frac{1}{4} w_1^* h_{11}$), we can deduce that the constellation used by S_2 forms the satellite constellation of the sum constellation at D_1.

2) Strong Cross Link: We investigate two channel realizations for this scenario. For the first realization, we let $h_{11} = 1$, $h_{21} = 3/2$ and $h_{22} = 1$. We then have $(w_1^*, w_2^*) = (0.1677, 0.4472)$ and the resulting sum-constellation at D_1 is also regular, as illustrated in Fig. 2(b). Since $w_1^* h_{11} = 0.1677$ and $w_2^* h_{21} = 0.6708$ (i.e., $w_2^* h_{21} = 4 w_1^* h_{11}$) in this case, the constellation used by S_1 forms the satellite constellation at D_1. For the second realization, we set $h_{11} = 1$, $h_{21} = 3$
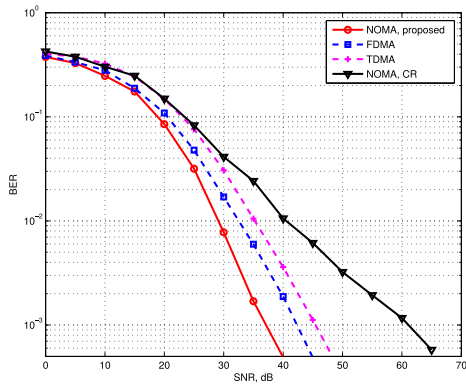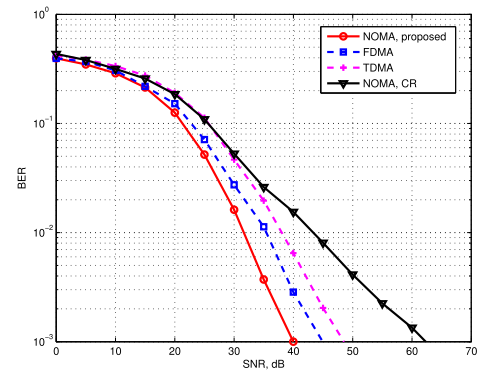
(a)
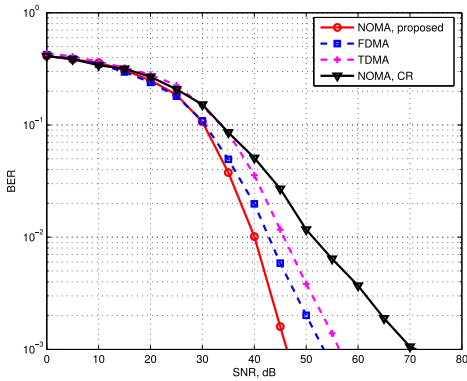
(b)

Fig. 3. Comparison of the proposed NOMA, CR based NOMA, TDMA and FDMA with $(\delta_{11}^2, \delta_{21}^2, \delta_{22}^2) = (1, 1, 1)$: (a) 16-QAM is used for proposed NOMA while 16-PSK is used for CR based NOMA. (b) 64-QAM is used for proposed NOMA while 64-PSK is used for CR based NOMA.

(a)

(b)

Fig. 4. Comparison of the proposed NOMA, CR based NOMA, TDMA and FDMA with $(\delta_{11}^2, \delta_{21}^2, \delta_{22}^2) = (1, 1/4, 1)$: (a) 16-QAM is used for proposed NOMA while 16-PSK is used for CR based NOMA. (b) 64-QAM is used for proposed NOMA while 64-PSK is used for CR based NOMA.

and $h_{22} = 1$, leading to $(w_1^*, w_2^*) = (0.3354, 0.4472)$. The resulting constellation plotted in Fig. 2(c) is also uniform as in the previous two scenarios. We have $w_1^* h_{11} = 0.3354$ and $w_2^* h_{21} = 1.3416$ (i.e., $w_2^* h_{21} = 4 \ w_1^* h_{11}$). Thus, the constellation used by the transmitter S_1 forms the satellite constellation at D_1.
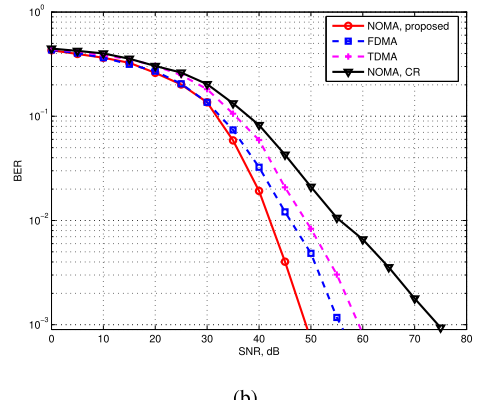
3) Very Strong Cross Link: In this case, we suppose that $h_{11} = 1$, $h_{21} = 5$ and $h_{22} = 1/2$, generating $(w_1^*, w_2^*) = (0.2236, 0.4472)$. The obtained constellation at D_1 is shown in Fig. 2(d). In this case, we have $w_1^* h_{11} = 0.2236$ and $w_2^* h_{21} = 2.236$ (i.e., $w_2^* h_{21} = 10 \ w_1^* h_{11}$). However, it can be observed that $w_2^* h_{22} = 0.2236$, i.e., $w_1^* h_{11} = w_2^* h_{22}$.

B. Average Error Performance Comparison in Rayleigh Fading Channels

We now compare the average BER of the proposed NOMA scheme with that of three existing methods, including TDMA, FDMA, and NOMA with CR [30] methods, over Rayleigh fading channels with $h_{11} \sim \mathcal{CN}(0, \delta_{11}^2)$, $h_{21} \sim \mathcal{CN}(0, \delta_{21}^2)$ and $h_{22} \sim \mathcal{CN}(0, \delta_{22}^2)$. Recall that we use error performance (i.e., BER) as the design criterion for the NOMA in ZCs with finite-alphabet inputs using *fixed transmission rate* (i.e., fixed constellation size). However, we are unable

to compare the error performance of the considered system using finite-alphabet inputs with that of Gaussian inputs. This is because for Gaussian input, it is intractable to evaluate the BER for uncoded system since its input signal is continuous. Moreover, the BER for coded system with Gaussian input is hard to simulate due to the huge storage capacity requirement for the large codebook and the high computational complexity [59, Ch. 9].

For TDMA, we assume that both users transmit alternatively by using half of the total time slots and thus no interference occurs at the destination side. More importantly, the individual instantaneous power constraints on both users S_1 and S_2 *remain unchanged*. On the other hand, for FDMA, each user uses only half of the available bandwidth. Due to the orthogonality between different frequency band, there is also no interference occurring at the destination side. Note that, in FDMA, the bandwidth occupied by each user is halved and the noise arises at the receiver is assumed to be white Gaussian. Therefore, the variance of the noise is also halved. In addition, for the CR based NOMA as proposed in [30], we let each user transmit at the maximum allowable power by using constellations $\{\exp(\frac{j2\pi k}{N})\}_{k=0}^{N-1}$ and $\{\exp(\frac{j2\pi \ell + j\pi}{N})\}_{\ell=0}^{N-1}$ for user S_1 and S_2, respectively.

In Fig. 3, we consider that the variances of all channels are the same, i.e., $(\delta_{11}^2, \delta_{21}^2, \delta_{22}^2) = (1, 1, 1)$, and the average BER
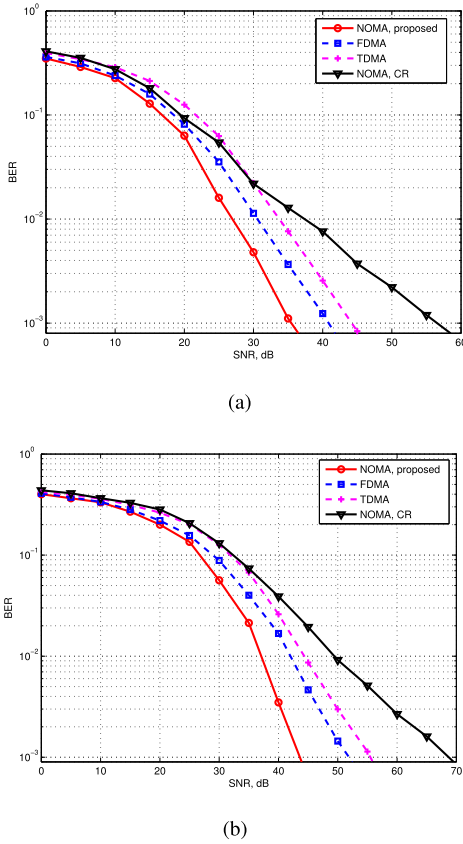
Fig. 5. Comparison of the proposed NOMA, CR based NOMA, TDMA and FDMA with $(\delta_{11}^2, \delta_{21}^2, \delta_{22}^2) = (1, 4, 1)$: (a) 16-QAM is used for proposed NOMA while 16-PSK is used for CR based NOMA. (b) 64-QAM is used for proposed NOMA while 64-PSK is used for CR based NOMA.

over both receivers of all the methods are plotted against the SNR $\rho = \frac{1}{2\sigma^2}$. In Fig. 3(a), 16-QAM is used for the proposed NOMA scheme while 16-PSK is employed by the NOMA with CR. Since only half of the total time slots or total bandwidths are available for each transmitter, to maintain the same data rate for each user in each block compared with NOMA methods, we should increase the constellation size by using 256-QAM in both TDMA and FDMA. It can be observed that, our proposed method has significant BER gain over TDMA and FDMA methods, which confirms the effectiveness of the NOMA scheme. From the simulation results, we also find that FDMA has a smaller BER compared with TDMA. This is because the variance of the effective noise is smaller than that of TDMA. As the rotation based method uses the PSK constellation, which is not spectrally efficient, it has the worst BER performance. Then, in Fig. 3(b), the average BER of all the cases are plotted against the SNR where 64-QAM is used by each user for the proposed NOMA and 64-PSK is used by CR based NOMA while 4096-QAM constellations are used by TDMA and FDMA methods. We also simulate another two cases with unequal channel variances. Specifically, the average BER of both receivers for all the considered methods is plotted in Fig. 4, wherein the variances of three channels are set as $(\delta_{11}^2, \delta_{21}^2, \delta_{22}^2) = (1, 1/4, 1)$. The case with the variances of three channels $(\delta_{11}^2, \delta_{21}^2, \delta_{22}^2) = (1, 4, 1)$ is provided in Fig. 5. Similar observations can be seen as the previous case with equal channel variances, which further

verifies the effectiveness of our proposed NOMA design.

V. CONCLUSIONS

In this paper, we developed a novel and practical design framework for the non-orthogonal multiple access (NOMA) in a classical two-transmitter two-receiver Z-channel with widely-used quadrature amplitude modulation (QAM) and max-min user fairness. Specifically, we formulated a max-min optimization problem to jointly optimize the scaling factors at both transmitters to maximize the smaller minimum Euclidean distance among the two resulting signal constellations at both receivers, subject to the individual average power constraint at each transmitter. The formulated mixed continuous-discrete problem was successfully resolved in compact closed-form by strategically applying the Farey sequences and their unique properties. Simulation results verified the correctness of our analytical derivations and showed that the proposed NOMA design significantly outperforms the existing orthogonal multiple access and NOMA schemes, especially at high signal-to-noise ratio. Furthermore, the performance gap of the proposed scheme over its existing counterparts can be further enlarged when the size of constellations used at the transmitter side becomes larger.

APPENDIX

A. Proof of Proposition 1

We first divide the feasible region $\mathbb{Z}_K^2 \setminus \{(0, 0)\}$ into four subsets given by $\mathbb{S}_{K,1}^2 = \{(m, n) : m, n \in \mathbb{N}_K, (m, n) \neq (0, 0)\}$, $\mathbb{S}_{K,2}^2 = \{(-m, n) : m, n \in \mathbb{N}_K, (m, n) \neq (0, 0)\}$, $\mathbb{S}_{K,3}^2 = \{(m, -n) : m, n \in \mathbb{N}_K, (m, n) \neq (0, 0)\}$, $\mathbb{S}_{K,4}^2 = \{(-m, -n) : m, n \in \mathbb{N}_K, (m, n) \neq (0, 0)\}$. Then clearly we have $\mathbb{Z}_K^2 \setminus \{(0, 0)\} = \cup_{k=1}^4 \mathbb{S}_{K,k}^2$. For $(-m, n) \in \mathbb{S}_{K,2}^2$ with $m, n \geq 0$, we can always find $(m, n) \in \mathbb{S}_{K,1}^2$ such that $d_1(m, n) = ||h_{11}|w_1 \, n - |h_{21}|w_2 \, m| \leq ||h_{11}|w_1 \, n - |h_{21}|w_2(-m)| = d_1(-m, n)$. Hence, $\min_{(m,n) \in \mathbb{S}_{K,1}^2} d_1(m, n) \leq \min_{(m,n) \in \mathbb{S}_{K,2}^2} d_1(m, n)$. By a similar argument on $\mathbb{S}_{K,3}^2$ and $\mathbb{S}_{K,4}^2$, it follows that $\min_{(m,n) \in \mathbb{S}_{K,1}^2} d_1(m, n) = \min_{(m,n) \in \mathbb{Z}_K^2 \setminus \{(0,0)\}} d_1(m, n)$.

Then, we further divide $\mathbb{S}_{K,1}^2$ into $\mathbb{F}_K^2$ and $\mathbb{S}_{K,1}^2 \setminus \mathbb{F}_K^2$. In what follows, we will show that $\min_{(m,n) \in \mathbb{F}_K^2} d_1(m, n) = \min_{(m,n) \in \mathbb{S}_{K,1}^2} d_1(m, n)$. This can be proved by contradiction. Suppose that $\min_{(m,n) \in \mathbb{S}_1^2 \setminus \mathbb{F}_K^2} d_1(m, n) = \min_{(m,n) \in \mathbb{S}_{K,1}^2} d_1(m, n)$, where the minimum is achieved by $(m^*, n^*) \in \mathbb{S}_{K,1}^2 \setminus \mathbb{F}_K^2$ such that $\langle m^*, n^* \rangle = \ell > 1$ by the definition of Farey sequences. Then, we can find $(\frac{m^*}{\ell}, \frac{n^*}{\ell}) \in \mathbb{F}_K^2$ such that $d_1(\frac{m^*}{\ell}, \frac{n^*}{\ell}) = \frac{1}{\ell} d_1(m^*, n^*) < d_1(m^*, n^*)$, which contradicts the assumption. This completes the proof. □

B. Proof of Proposition 2

Recall that $d_1(m, n) = ||h_{11}|w_1 \, n - |h_{21}|w_2 \, m|$. Therefore, for $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in \left(\frac{n_1}{m_1}, \frac{n_2}{m_2}\right)$, we have: $d_1(m_1, n_1) - d_1(m_2, n_2) = (m_1 + m_2)|h_{11}|w_1 \left(\frac{|h_{21}|w_2}{|h_{11}|w_1} - \frac{n_1 + n_2}{m_1 + m_2}\right)$. The results presented in the proposition can be readily obtained, and we complete the proof. □

C. Proof of of Proposition 3

We first consider the case $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in (\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3})$. By using Proposition 2, we have $d_1(m_2, n_2) < d_1(m_3, n_3)$. Then, with the help of Property 1 and 3, we have $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in (\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}) \subset (\frac{n_2}{m_2}, \frac{n_3}{m_3}) \subseteq (\frac{n_2}{m_2}, \frac{n_2+n_4}{m_2+m_4})$ and using Property 2 again, we have $d_1(m_2, n_2) \leq d_1(m_4, n_4)$. On the other hand, by Property 1 and 3, we have $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in (\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}) \subset (\frac{n_2}{m_2}, \frac{n_3}{m_3}) \subseteq (\frac{n_1+n_3}{m_1+m_3}, \frac{n_3}{m_3})$ and using Property 2 again, we have $d_1(m_3, n_3) \leq d_1(m_1, n_1)$.

As $\frac{n_1}{m_1}$ and $\frac{n_4}{m_4}$ are randomly picked entry in $\mathfrak{S}_K$, or equivalently (m_1, n_1) and (m_4, n_4) are randomly picked in $\mathbb{F}_{M-1}^2 \setminus \{(m_2, n_2), (m_3, n_3)\}$, this proves that for $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in (\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3})$, $\min_{(m,n)\in\mathbb{F}_{M-1}^2} d_1(m, n) = d_1(m_2, n_2) = |h_{21}|w_2 \, m_2 - |h_{11}|w_1 \, n_2$. The other case can also be proved by a similar argument and hence omitted. We completes the proof. $\square$

D. Proof of Proposition 4

First of all, we calculate the following difference $\frac{a_k+a_{k+1}}{b_k+b_{k+1}} - \frac{a_{k+1}}{b_{k+1}} - \frac{|h_{22}|}{b_{k+1}|h_{21}|} = \frac{|h_{21}|-|h_{22}|(b_k+b_{k+1})}{(b_k+b_{k+1})b_{k+1}|h_{21}|}$. Hence, if $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, then $\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|} \leq \frac{a_k+a_{k+1}}{b_k+b_{k+1}}$. Similarly, we have $\frac{a_k}{b_k} - \frac{|h_{22}|}{b_k|h_{21}|} - \frac{a_k+a_{k+1}}{b_k+b_{k+1}} = \frac{|h_{21}|-|h_{22}|(b_k+b_{k+1})}{b_k \, (b_k+b_{k+1})|h_{21}|}$. As a result, if $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, then $\frac{a_k+a_{k+1}}{b_k+b_{k+1}} \leq \frac{a_k}{b_k} - \frac{|h_{22}|}{b_k|h_{21}|}$. This completes the proof. $\square$

E. Proof of Lemma 1

According to proposition 3 and notice that $(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}) = (\frac{b_k}{a_k}, \frac{b_k+b_{k+1}}{a_k+a_{k+1}}) \cup (\frac{b_k+b_{k+1}}{a_k+a_{k+1}}, \frac{b_{k+1}}{a_{k+1}})$, problem in (11) can be further divided into the following two sub-problems and the overall solution is the maximum value of the two problems:

Problem 2 (Sub-Problem 1): We aim to solve the following optimization problem:

$$g_1\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \max_{w_1, w_2} \ \min \ \{|h_{11}|w_1 b_{k+1} - |h_{21}|w_2 a_{k+1},$$

$$|h_{22}|w_2\} \quad (12a)$$

$$\text{s.t.} \ \frac{b_k+b_{k+1}}{a_k+a_{k+1}} \leq \frac{|h_{21}|w_2}{|h_{11}|w_1} \leq \frac{b_{k+1}}{a_{k+1}}, \quad (12b)$$

$$0 < w_1 \leq \sqrt{\frac{3p_1}{M^2-1}},$$

$$0 < w_2 \leq \sqrt{\frac{3p_2}{M^2-1}}. \quad (12c)$$

$\blacksquare$

Problem 3 (Sub-Problem 2): The optimization problem is stated as follows:

$$g_2\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \max_{w_1, w_2} \min\{|h_{21}|w_2 a_k - |h_{11}|w_1 b_k, |h_{22}|w_2\}$$

$$(13a)$$

$$\text{s.t.} \ \frac{b_k}{a_k} \leq \frac{|h_{21}|w_2}{|h_{11}|w_1} < \frac{b_k+b_{k+1}}{a_k+a_{k+1}}, \quad (13b)$$

$$0 < w_1 \leq \sqrt{\frac{3p_1}{M^2-1}},$$

$$0 < w_2 \leq \sqrt{\frac{3p_2}{M^2-1}}. \quad (13c)$$

$\blacksquare$

We first consider Sub-problem 1, which can be divided into the following two case:

1) Case 1: Receiver 1 has smaller Euclidean distance.

$$g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$

$$= \max_{w_1, w_2} \ b_{k+1}|h_{11}|w_1 - a_{k+1}|h_{21}|w_2$$

$$\text{s.t.} \ b_{k+1}|h_{11}|w_1 - a_{k+1}|h_{21}|w_2 \leq |h_{22}|w_2, \quad (14a)$$

$$\frac{b_k+b_{k+1}}{a_k+a_{k+1}} \leq \frac{|h_{21}|w_2}{|h_{11}|w_1} \leq \frac{b_{k+1}}{a_{k+1}}, \quad (14b)$$

$$0 < w_1 \leq \sqrt{\frac{3p_1}{M^2-1}}, \quad 0 < w_2 \leq \sqrt{\frac{3p_2}{M^2-1}}. \quad (14c)$$

2) Case 2: Receiver 2 has smaller Euclidean distance.

$$g_{12}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$

$$= \max_{w_1, w_2} \ |h_{22}|w_2$$

$$\text{s.t.} \ |h_{22}|w_2 < b_{k+1}|h_{11}|w_1 - a_{k+1}|h_{21}|w_2, \quad (15a)$$

$$\frac{b_k+b_{k+1}}{a_k+a_{k+1}} \leq \frac{|h_{21}|w_2}{|h_{11}|w_1} < \frac{b_{k+1}}{a_{k+1}}, \quad (15b)$$

$$0 < w_1 \leq \sqrt{\frac{3p_1}{M^2-1}}, \quad 0 < w_2 \leq \sqrt{\frac{3p_2}{M^2-1}}. \quad (15c)$$

For **Case-1 of Sub-problem 1**, constraint (14a) is equivalent to $\frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|+|h_{22}|}w_1 \leq w_2$ and constraint (14b) means $\frac{(b_k+b_{k+1})|h_{11}|}{(a_k+a_{k+1})|h_{21}|}w_1 \leq w_2 \leq \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|}w_1$. Also we notice that if $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$ then $\frac{(b_k+b_{k+1})|h_{11}|}{(a_k+a_{k+1})|h_{21}|}w_1 \leq \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|+|h_{22}|}w_1$. Hence, the optimization problem can be further divided into two cases:

1) If $\frac{|h_{21}|}{|h_{22}|} \leq b_k + b_{k+1}$, then

$$g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$

$$= \max_{w_1, w_2} \ b_{k+1}|h_{11}|w_1 - a_{k+1}|h_{21}|w_2$$

$$\text{s.t.} \ \frac{(b_k+b_{k+1})|h_{11}|}{(a_k+a_{k+1})|h_{21}|}w_1 \leq w_2 \leq \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|}, \quad (16a)$$

$$0 < w_1 \leq \sqrt{\frac{3p_1}{M^2-1}}, \quad 0 < w_2 \leq \sqrt{\frac{3p_2}{M^2-1}}. \quad (16b)$$

We let $w_2 = \frac{(b_k+b_{k+1})|h_{11}|}{(a_k+a_{k+1})|h_{21}|}w_1$, then the objective function is $\frac{|h_{11}|w_1}{a_k+a_{k+1}}$. In this case, (16b) is equivalent to $w_1 \leq \frac{(a_k+a_{k+1})|h_{21}|}{(b_k+b_{k+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}$ and $w_1 \leq \sqrt{\frac{3p_1}{M^2-1}}$ and hence, we have $g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \min\left\{\frac{|h_{11}|}{a_k+a_{k+1}}\frac{(a_k+a_{k+1})|h_{21}|}{(b_k+b_{k+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \frac{|h_{11}|}{a_k+a_{k+1}}\sqrt{\frac{3p_1}{M^2-1}}\right\}$.

As a consequence, we have

$$
g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)
$$

$$
= \begin{cases}
\frac{|h_{21}|}{b_k + b_{k+1}}\sqrt{\frac{3p_2}{M^2 - 1}}, \text{ where} \\
(w_1, w_2) \\
\quad = \left(\frac{(a_k + a_{k+1})|h_{21}|}{(b_k + b_{k+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2 - 1}}, \sqrt{\frac{3p_2}{M^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}; \\
\frac{|h_{11}|}{a_k + a_{k+1}}\sqrt{\frac{3p_1}{M^2 - 1}}, \text{ where} \\
(w_1, w_2) \\
\quad = \left(\sqrt{\frac{3p_1}{M^2 - 1}}, \frac{(b_k + b_{k+1})|h_{11}|}{(a_k + a_{k+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})},
\end{cases}
$$

2) If $\frac{|h_{21}|}{|h_{22}|} > b_k + b_{k+1}$, then

$$
g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)
$$
$$
= \max_{w_1, w_2} b_{k+1}|h_{11}|w_1 - a_{k+1}|h_{21}|w_2
$$
$$
\text{s.t. } \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}w_1 \leq w_2 \leq \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|}w_1, \tag{18a}
$$
$$
0 < w_1 \leq \sqrt{\frac{3p_1}{M^2 - 1}}, \quad 0 < w_2 \leq \sqrt{\frac{3p_2}{M^2 - 1}}. \tag{18b}
$$

We first notice that $\frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}w_1 < \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|}w_1$ and hence the problem is always feasible. By letting $w_2 = \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}w_1$, the objective function can be written by $\frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}| + |h_{22}|}w_1$. In this case, the constraints in (18b) are equivalent to $w_1 \leq \sqrt{\frac{3p_1}{M^2-1}}$ and $w_1 \leq \frac{a_{k+1}|h_{21}| + |h_{22}|}{b_{k+1}|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}$. Therefore, the solution is

$$
g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)
$$

$$
= \begin{cases}
\frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}| + |h_{22}|}\sqrt{\frac{3p_1}{M^2 - 1}}, \text{ where} \\
(w_1, w_2) = \left(\sqrt{\frac{3p_1}{M_1^2 - 1}}, \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}\sqrt{\frac{3p_1}{M^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right); \\
|h_{22}|\sqrt{\frac{3p_2}{M_2^2 - 1}}, \text{ where} \\
(w_1, w_2) = \left(\frac{a_{k+1}|h_{21}| + |h_{22}|}{b_{k+1}|h_{11}|}\sqrt{\frac{3p_2}{M^2 - 1}}, \sqrt{\frac{3p_2}{M^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} > \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right).
\end{cases}
$$

For **Case-2 of Sub-problem 1**, constraint (15a) is equivalent to $w_2 < \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}w_1$ and constraint (15b) implies

$\frac{(b_k + b_{k+1})|h_{11}|}{(a_k + a_{k+1})|h_{21}|}w_1 \leq w_2 \leq \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|}w_1$. By noticing that $\frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|} < \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}|}$, the problem is equivalent to

$$
g_{12}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)
$$
$$
= \max_{w_1, w_2} |h_{22}|w_2
$$
$$
\text{s.t. } \frac{(b_k + b_{k+1})|h_{11}|}{(a_k + a_{k+1})|h_{21}|}w_1 \leq w_2 \leq \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}w_1, \tag{20a}
$$
$$
0 < w_1 \leq \sqrt{\frac{3p_1}{M^2 - 1}}, \quad 0 < w_2 \leq \sqrt{\frac{3p_2}{M^2 - 1}}. \tag{20b}
$$

Constraint (20a) is feasible if $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$. In this case, the solution is

$$
g_{12}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)
$$

$$
= \begin{cases}
\frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}| + |h_{22}|}\sqrt{\frac{3p_1}{M^2 - 1}}, \text{ where} \\
(w_1, w_2) = \left(\sqrt{\frac{3p_1}{M^2 - 1}}, \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}\sqrt{\frac{3p_1}{M^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right); \\
|h_{22}|\sqrt{\frac{3p_2}{M_2^2 - 1}}, \text{ where} \\
(w_1, w_2) = \left(\frac{a_{k+1}|h_{21}| + |h_{22}|}{b_{k+1}|h_{11}|}\sqrt{\frac{3p_2}{M_2^2 - 1}}, \sqrt{\frac{3p_2}{M_2^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} > \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right).
\end{cases}
$$

The solution of Sub-problem 2 can be attained in a similar fashion as Sub-problem 1, and hence is omitted. Then, for the subinterval division $\frac{|h_{21}|w_2}{|h_{11}|w_1} \in \left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$, **the solution to both Sub-problems can be summarized as follows:**

1) Scenario 1, $\frac{b_k + b_{k+1}}{a_k + a_{k+1}} \leq \frac{|h_{21}|w_2}{|h_{11}|w_1} < \frac{b_{k+1}}{a_{k+1}}$:

 a) If $\frac{|h_{21}|}{|h_{22}|} \leq b_k + b_{k+1}$

$$
g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)
$$

$$
= \begin{cases}
\frac{|h_{21}|}{b_k + b_{k+1}}\sqrt{\frac{3p_2}{M^2 - 1}}, \text{ where} \\
(w_1, w_2) \\
\quad = \left(\frac{(a_k + a_{k+1})|h_{21}|}{(b_k + b_{k+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2 - 1}}, \sqrt{\frac{3p_2}{M^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}; \\
\frac{|h_{11}|}{a_k + a_{k+1}}\sqrt{\frac{3p_1}{M^2 - 1}}, \text{ where} \\
(w_1, w_2) \\
\quad = \left(\sqrt{\frac{3p_1}{M^2 - 1}}, \frac{(b_k + b_{k+1})|h_{11}|}{(a_k + a_{k+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2 - 1}}\right), \\
\qquad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})},
\end{cases}
$$

b) If $\frac{|h_{21}|}{|h_{22}|} > b_k + b_{k+1}$

$$g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$

$$= \begin{cases} \frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}| + |h_{22}|}\sqrt{\frac{3p_1}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{b_{k+1}|h_{11}|}{a_{k+1}|h_{21}| + |h_{22}|}\sqrt{\frac{3p_1}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right); \\[2mm] |h_{22}|\sqrt{\frac{3p_2}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\frac{a_{k+1}|h_{21}| + |h_{22}|}{b_{k+1}|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} > \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right). \end{cases}$$

2) Scenario 2, $\frac{b_k}{a_k} \leq \frac{|h_{21}|w_2}{|h_{11}|w_1} \leq \frac{b_k + b_{k+1}}{a_k + a_{k+1}}$:

 a) Case 1: Receiver 1 has smaller Euclidean distance.

 i) If $\frac{|h_{21}|}{|h_{22}|} \leq b_k + b_{k+1}$, the solution is

$$g_{21}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$

$$= \begin{cases} \frac{|h_{21}|}{b_k + b_{k+1}}\sqrt{\frac{3p_2}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\frac{(a_k + a_{k+1})|h_{21}|}{(b_k + b_{k+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}; \\[2mm] \frac{|h_{11}|}{a_k + a_{k+1}}\sqrt{\frac{3p_1}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{(b_k + b_{k+1})|h_{11}|}{(a_k + a_{k+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}. \end{cases}$$

ii) If $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$,

$$g_{21}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$

$$= \begin{cases} |h_{22}|\sqrt{\frac{3p_2}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\frac{|h_{21}|a_k - |h_{22}|}{b_k|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_k}{b_k} - \frac{|h_{22}|}{|h_{21}|b_k}\right); \\[2mm] \frac{b_k|h_{11}||h_{22}|}{|h_{21}|a_k - |h_{22}|}\sqrt{\frac{3p_1}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{b_k|h_{11}|}{|h_{21}|a_k - |h_{22}|}\sqrt{\frac{3p_1}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} < \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_k}{b_k} - \frac{|h_{22}|}{|h_{21}|b_k}\right). \end{cases}$$

 b) Case 2: Receiver 2 has smaller Euclidean distance.

 i) If $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, the solution is

$$g_{22}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$

$$= \begin{cases} \frac{(b_k + b_{k+1})|h_{11}||h_{22}|}{(a_k + a_{k+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\sqrt{\frac{3p_1}{M^2-1}}, \frac{(b_k + b_{k+1})|h_{11}|}{(a_k + a_{k+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}; \\[2mm] |h_{22}|\sqrt{\frac{3p_2}{M^2-1}}, & \text{where} \\[2mm] (w_1, w_2) \\[2mm] = \left(\frac{(a_k + a_{k+1})|h_{21}|}{(b_k + b_{k+1})|h_{11}|}\sqrt{\frac{3p_2}{M^2-1}}, \sqrt{\frac{3p_2}{M^2-1}}\right), \\[2mm] \quad \text{if } \frac{|h_{11}|}{|h_{21}|} > \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}. \end{cases}$$

Now, we aim to combine $g_{21}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$ and $g_{22}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$. By Proposition 4, for $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, we have $\frac{a_k}{b_k} - \frac{|h_{22}|}{|h_{21}|b_k} \geq \frac{a_k + a_{k+1}}{b_k + b_{k+1}}$. Also, for $\frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_k}{b_k} - \frac{|h_{22}|}{b_k|h_{21}|}\right)$, we have $\sqrt{\frac{3p_2}{M^2-1}} \geq \frac{b_k|h_{11}|}{|h_{21}|a_k - |h_{22}|}\sqrt{\frac{3p_1}{M^2-1}}$. In addition, for $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, we have $\frac{(b_k + b_{k+1})}{(a_k + a_{k+1})|h_{21}|} \geq \frac{b_k}{|h_{21}|a_k - |h_{22}|}$. Hence, we have $g_{22}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) \geq g_{21}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$. We notice that for $\frac{|h_{21}|}{|h_{22}|} \leq b_k + b_{k+1}$, $g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = g_{21}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$. Then for $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, we combine $g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$ and $g_{22}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$. By Proposition 4, for $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, we have $\frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right) \leq \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}$. Also, for $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, we attain $\frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}| + |h_{22}|}\sqrt{\frac{3p_1}{M^2-1}} \geq \frac{(b_k + b_{k+1})|h_{11}||h_{22}|}{(a_k + a_{k+1})|h_{21}|}\sqrt{\frac{3p_1}{M^2-1}}$. In addition, for $\frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}(a_k + a_{k+1})}{\sqrt{p_1}(b_k + b_{k+1})}$, we have $|h_{22}|\sqrt{\frac{3p_2}{M^2-1}} \geq$

$\frac{(b_k+b_{k+1})|h_{11}||h_{22}|}{(a_k+a_{k+1})|h_{21}|}\sqrt{\frac{3\ p_1}{M^2-1}}$. In conclusion, for $\frac{|h_{21}|}{|h_{22}|} \geq b_k + b_{k+1}$, $g_{11}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) \geq g_{22}\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$.

With the above discussion, we have the result in Lemma 1 and we complete the proof. □

F. Proof of Theorem 1

For the weak cross link, we have $\frac{|h_{21}|}{|h_{22}|} \leq 1 \leq b_k + b_{k+1}$ for $k = 1, \ldots, |\mathcal{S}_{M-1}|$.

1) If $\frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}}{M\sqrt{p_1}}$, we have $\frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}(a_k+a_{k+1})}{\sqrt{p_1}(b_k+b_{k+1})}$, and then by Lemma 1, we attain $g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \frac{|h_{11}|}{a_k+a_{k+1}}\sqrt{\frac{3\ p_1}{M^2-1}}$ for $k = 1, \ldots, |\mathcal{S}_{M-1}|$. Also, note that $|\mathcal{S}_{M-1}| = \arg\min_k \{(a_1 + a_2), \ldots, (a_{|\mathcal{S}_{M-1}|} + a_{|\mathcal{S}_{M-1}|+1})\}$, then $\max\left\{\frac{|h_{11}|}{a_1+a_2}\sqrt{\frac{3\ p_1}{M^2-1}}, \cdots, \frac{|h_{11}|}{a_{|\mathcal{S}_{M-1}|}+a_{|\mathcal{S}_{M-1}|+1}}\sqrt{\frac{3\ p_1}{M^2-1}}\right\} = |h_{11}|\sqrt{\frac{3\ p_1}{M^2-1}}$, with $(w_1^*, w_2^*) = \left(\sqrt{\frac{3\ p_1}{M^2-1}}, \frac{M|h_{11}|}{|h_{21}|}\sqrt{\frac{3\ p_1}{M^2-1}}\right)$.

2) If $\frac{|h_{11}|}{|h_{21}|} \geq \frac{M\sqrt{p_2}}{\sqrt{p_1}}$, we have $\frac{|h_{11}|}{|h_{21}|} \geq \frac{(a_k+a_{k+1})\sqrt{p_2}}{(b_k+b_{k+1})\sqrt{p_1}}$ and then, by using Lemma 1, we have $g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \frac{|h_{21}|}{b_k+b_{k+1}}\sqrt{\frac{3\ p_2}{M^2-1}}$ for $k = 1, \ldots, |\mathcal{S}_{M-1}|$. Also, note that $1 = \arg\min_k \{(b_1 + b_2), \ldots, (b_{|\mathcal{S}_{M-1}|} + b_{|\mathcal{S}_{M-1}|+1})\}$. Then,

$$\max\left\{\frac{|h_{21}|}{b_1+b_2}\sqrt{\frac{3\ p_2}{M^2-1}}, \cdots, \frac{|h_{21}|}{b_{|\mathcal{S}_{M-1}|}+b_{|\mathcal{S}_{M-1}|+1}}\sqrt{\frac{3\ p_2}{M^2-1}}\right\}$$
$$= |h_{21}|\sqrt{\frac{3\ p_2}{M^2-1}},$$

with $(w_1^*, w_2^*) = \left(\frac{M|h_{21}|}{|h_{11}|}\sqrt{\frac{3\ p_2}{M^2-1}}, \sqrt{\frac{3\ p_2}{M^2-1}}\right)$.

3) If $\frac{|h_{11}|}{|h_{21}|} \in \left(\frac{\sqrt{p_2}(a_{\ell_1+1}+a_{\ell_1+2})}{\sqrt{p_1}(b_{\ell_1+1}+b_{\ell_1+2})}, \frac{\sqrt{p_2}(a_{\ell_1}+a_{\ell_1+1})}{\sqrt{p_1}(b_{\ell_1}+b_{\ell_1+1})}\right)$, for $\ell_1 = 1, \ldots, |\mathcal{S}_{M-1}| - 1$, then, with the help of Lemma 1, we have

$$g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$
$$= \begin{cases} \frac{|h_{11}|}{a_k+a_{k+1}}\sqrt{\frac{3\ p_1}{M^2-1}} & \text{for } k = 1, \ldots, \ell_1, \\ \frac{|h_{21}|}{b_k+b_{k+1}}\sqrt{\frac{3\ p_2}{M^2-1}} & \text{for } k = \ell_1+1, \ldots, |\mathcal{S}_{M-1}|. \end{cases}$$

Note that, $\tilde{\ell}_a = \arg\min_k\{(a_1 + a_2), \ldots, (a_{\ell_1} + a_{\ell_1+1})\}$ and $\tilde{\ell}_b = \arg\min_k\{(b_{\ell_1+1} + b_{\ell_1+2}), \ldots, (b_{|\mathcal{S}_{M-1}|} + b_{|\mathcal{S}_{M-1}|+1})\}$, hence we have

$$\max\left\{\frac{|h_{11}|}{a_1+a_2}\sqrt{\frac{3\ p_1}{M^2-1}}, \cdots, \frac{|h_{11}|}{a_{\ell_1}+a_{\ell_1+1}}\sqrt{\frac{3\ p_1}{M^2-1}}, \frac{|h_{21}|}{b_{\ell_1+1}+b_{\ell_1+2}}\sqrt{\frac{3\ p_2}{M^2-1}}, \cdots, \frac{|h_{21}|}{b_{|\mathcal{S}_{M-1}|}+b_{|\mathcal{S}_{M-1}|+1}}\sqrt{\frac{3\ p_2}{M^2-1}}\right\}$$

$$= \max\left\{\frac{|h_{11}|}{a_{\tilde{\ell}_a}+a_{\tilde{\ell}_a+1}}\sqrt{\frac{3\ p_1}{M^2-1}}, \frac{|h_{21}|}{b_{\tilde{\ell}_b}+b_{\tilde{\ell}_b+1}}\sqrt{\frac{3\ p_2}{M^2-1}}\right\}.$$

Therefore, if $\frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}(a_{\tilde{\ell}_a}+a_{\tilde{\ell}_a+1})}{\sqrt{p_1}(b_{\tilde{\ell}_b}+b_{\tilde{\ell}_b+1})}$, we have $\frac{|h_{11}|}{a_{\tilde{\ell}_a}+a_{\tilde{\ell}_a+1}}\sqrt{\frac{3\ p_1}{M^2-1}} \geq \frac{|h_{21}|}{b_{\tilde{\ell}_b}+b_{\tilde{\ell}_b+1}}\sqrt{\frac{3\ p_2}{M^2-1}}$ and hence, $(w_1^*, w_2^*) = \left(\sqrt{\frac{3\ p_1}{M^2-1}}, \frac{(b_{\tilde{\ell}_a}+b_{\tilde{\ell}_a+1})|h_{11}|}{(a_{\tilde{\ell}_a}+a_{\tilde{\ell}_a+1})|h_{21}|}\sqrt{\frac{3\ p_1}{M-1}}\right)$ and else we have $(w_1^*, w_2^*) = \left(\frac{(a_{\tilde{\ell}_b}+a_{\tilde{\ell}_b+1})|h_{21}|}{(b_{\tilde{\ell}_b}+b_{\tilde{\ell}_b+1})|h_{11}|}\sqrt{\frac{3\ p_2}{M^2-1}}, \sqrt{\frac{3\ p_2}{M^2-1}}\right)$.

This completes the proof. □

G. Proof of Theorem 3

1) If $|h_{11}| \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}|h_{22}|$, we have $\frac{|h_{11}|}{|h_{21}|} \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}\frac{|h_{22}|}{|h_{21}|} = \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{|\mathcal{S}_{M-1}|+1}}{b_{|\mathcal{S}_{M-1}|+1}} + \frac{|h_{22}|}{b_{|\mathcal{S}_{M-1}|+1}|h_{21}|}\right) \leq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{k+1}}{b_{k+1}} + \frac{|h_{22}|}{b_{k+1}|h_{21}|}\right)$ for $k = 1, \ldots, |\mathcal{S}_{M-1}|$. According to Lemma 1, we have $g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}$, for $k = 1, \ldots, |\mathcal{S}_{M-1}|$. We also have

$$\max\left\{\frac{b_2|h_{11}||h_{22}|}{a_2|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}, \cdots, \frac{b_{|\mathcal{S}_{M-1}|+1}|h_{11}||h_{22}|}{a_{|\mathcal{S}_{M-1}|+1}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}\right\}$$
$$= |h_{11}|\sqrt{\frac{3\ p_1}{M^2-1}},$$

with $(w_1^*, w_2^*) = (\sqrt{\frac{3\ p_1}{M^2-1}}, \frac{|h_{11}|}{|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}})$.

2) If $\frac{|h_{11}|}{|h_{21}|} \in \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{\ell_4+1}}{b_{\ell_4+1}} + \frac{|h_{22}|}{b_{\ell_4+1}|h_{21}|}, \frac{a_{\ell_4}}{b_{\ell_4}} + \frac{|h_{22}|}{b_{\ell_4}|h_{21}|}\right)$ for $\ell_4 = 1, \ldots, |\mathcal{S}_{M-1}|$.

Then, according to Lemma 1, we have $g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}$, for $k = 1, \ldots, \ell_4-1$ and $g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = |h_{22}|\sqrt{\frac{3\ p_2}{M^2-1}}$ for $k = \ell_4, \ldots, |\mathcal{S}_{M-1}|$.

$$g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right)$$
$$= \begin{cases} \frac{b_{k+1}|h_{11}||h_{22}|}{a_{k+1}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}} & \text{for } k = 1, \ldots, \ell_4-1, \\ |h_{22}|\sqrt{\frac{3\ p_2}{M^2-1}} & \text{for } k = \ell_4, \ldots, |\mathcal{S}_{M-1}|. \end{cases}$$

Then, we have,

$$\max\left\{\frac{b_2|h_{11}||h_{22}|}{a_2|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}, \cdots, \frac{b_{\ell_4}|h_{11}||h_{22}|}{a_{\ell_4}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}, |h_{22}|\sqrt{\frac{3\ p_2}{M^2-1}}\right\}$$
$$= \max\left\{\frac{b_{\ell_4}|h_{11}||h_{22}|}{a_{\ell_4}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}, |h_{22}|\sqrt{\frac{3\ p_2}{M^2-1}}\right\}.$$
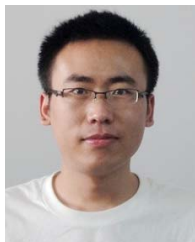
As a result, if $\frac{|h_{11}|}{|h_{21}|} \geq \frac{\sqrt{p_2}}{\sqrt{p_1}}\left(\frac{a_{\ell_4}}{b_{\ell_4}} + \frac{|h_{22}|}{b_{\ell_4}|h_{21}|}\right)$, then $\frac{b_{\ell_4}|h_{11}||h_{22}|}{a_{\ell_4}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}} \geq |h_{22}|\sqrt{\frac{3\ p_2}{M^2-1}}$ and $(w_1^*, w_2^*) = \left(\sqrt{\frac{3\ p_1}{M^2-1}}, \frac{b_{\ell_4}|h_{11}|}{a_{\ell_4}|h_{21}|+|h_{22}|}\sqrt{\frac{3\ p_1}{M^2-1}}\right)$. Else, we can attain $(w_1^*, w_2^*) = \left(\frac{a_{\ell_4+1}|h_{21}|+|h_{22}|}{b_{\ell_4+1}|h_{11}|}\sqrt{\frac{3\ p_2}{M^2-1}}, \sqrt{\frac{3\ p_2}{M^2-1}}\right)$. This completes the proof. □

REFERENCES

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

[2] L. Dai, B. Wang, Y. Yuan, S. Han, C. L. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, Sep. 2015.

[3] W. Shin, M. Vaezi, B. Lee, D. J. Love, J. Lee, and H. V. Poor, "Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges," *CoRR*, Jun. 2017. [Online]. Available: http://arxiv.org/abs/1611.01607

[4] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[5] Z. Ding *et al.*, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.

[6] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[7] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, London, U.K., Sep. 2013, pp. 611–615.

[8] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.

[9] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[10] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 1, pp. 2–10, Jan. 1974.

[11] R. S. Cheng and S. Verdu, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 773–785, May 1993.

[12] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, "System-level performance of downlink NOMA for future LTE enhancements," in *Proc. Globecom Workshops (GC Wkshps)*, Dec. 2013, pp. 66–70.

[13] 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.

[14] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, "5G network capacity: Key elements and technologies," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 71–78, Mar. 2014.

[15] P. Wang, J. Xiao, and L. P, "Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems," *IEEE Veh. Technol. Mag.*, vol. 1, no. 3, pp. 4–11, Sep. 2006.

[16] Z. Wei, J. Yuan, D. W. K. Ng, M. Elkashlan, and Z. Ding, "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks," *CoRR*, Jun. 2017. [Online]. Available: http://arxiv.org/abs/1609.01856

[17] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[18] *Study on Downlink Multiuser Superposition Transmission (MUST) for LTE (Release 13)*, document TR 36.859 V13.0.0, 3GPP, Dec. 2015.

[19] H. Lee, S. Kim, and J. H. Lim, "Multiuser superposition transmission (MUST) for LTE-A systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[20] Q. Sun, S. Han, C. L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405–408, Aug. 2015.

[21] Y. Liu, G. Pan, H. Zhang, and M. Song, "On the capacity comparison between MIMO-NOMA and MIMO-OMA," *IEEE Access*, vol. 4, pp. 2123–2129, May 2016.

[22] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.

[23] Z. Ding, H. Dai, and H. V. Poor, "Relay selection for cooperative NOMA," *IEEE Wireless Commun. Lett.*, vol. 5, no. 4, pp. 416–419, Aug. 2016.

[24] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.

[25] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Nonorthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10152–10157, Dec. 2016.

[26] A. Lozano, A. M. Tulino, and S. Verdu, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033–3051, Jul. 2006.

[27] T. Kasami and S. Lin, "Coding for a multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 22, no. 2, pp. 129–137, Mar. 1976.

[28] T. Kasami and S. Lin, "Bounds on the achievable rates of block coding for a memoryless multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 2, pp. 187–197, Mar. 1978.

[29] R. Ahlswede and V. B. Balakirsky, "Construction of uniquely decodable codes for the two-user binary adder channel," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 326–330, Jan. 1999.

[30] J. Harshan and B. Rajan, "On two-user Gaussian multiple access channels with finite input constellations," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1299–1327, Mar. 2011.

[31] J. Harshan and B. S. Rajan, "A novel power allocation scheme for two-user GMAC with finite input constellations," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 818–827, Feb. 2013.

[32] Y. Wu, C. K. Wen, C. Xiao, X. Gao, and R. Schober, "Linear precoding for the MIMO multiple access channel with finite alphabet inputs and statistical CSI," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 983–997, Feb. 2015.

[33] J. Choi, "On the power allocation for a practical multiuser superposition scheme in noma systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 438–441, Mar. 2016.

[34] S. L. Shieh and Y. C. Huang, "A simple scheme for realizing the promised gains of downlink nonorthogonal multiple access," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1624–1635, Apr. 2016.

[35] S. Vishwanath, N. Jindal, and A. Goldsmith, "The 'Z' channel," in *Proc. IEEE Global Commu. Conf. (GLOBECOM)*, vol. 3. Dec. 2003, pp. 1726–1730.

[36] H.-F. Chong, M. Motani, and H. K. Garg, "Capacity theorems for the 'Z' channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1348–1365, Apr. 2007.

[37] N. Liu and S. Ulukus, "On the capacity region of the Gaussian Z-channel," in *Proc. IEEE Global Commu. Conf. (GLOBECOM)*, vol. 1. Nov. 2004, pp. 415–419.

[38] V. R. Cadambe, S. A. Jafar, and S. Vishwanath, "The capacity region of a class of deterministic Z channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2009, pp. 2634–2638.

[39] S. Salehkalaibar and M. R. Aref, "On the capacity region of the degraded Z channel," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Aug. 2010, pp. 1–5.

[40] J. Lee, K. Kim, S. C. Cho, and S. H. Lee, "Linear beamforming with one-way cooperation for MISO interference channels," in *Proc. IEEE 77th Veh. Tech. Conf. (VTC Spring)*, Jun. 2013, pp. 1–5.

[41] R. Duan, Y. Liang, A. Khisti, and S. Shamai (Shitz), "State-dependent Gaussian Z-channel with mismatched side-information and interference," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Seville, Spain, Sep. 2013, pp. 1–5.

[42] N. Liu and W. Kang, "The capacity region of a class of Z channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6144–6153, Oct. 2014.

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

[44] M. Costa, "On the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 31, no. 5, pp. 607–615, Sep. 1985.

[45] E. Kurniawan and S. Sun, "Improper Gaussian signaling scheme for the Z-interference channel," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3912–3923, Jul. 2015.

[46] R. Kolte, A. Özgür, and H. Permuter, "The capacity region of a class of deterministic state-dependent Z-interference channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2014, pp. 656–660.

[47] M. Vaezi and H. V. Poor, "Simplified Han–Kobayashi region for one-sided and mixed Gaussian interference channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[48] S. Lagen, A. Agustin, and J. Vidal, "On the superiority of improper Gaussian signaling in wireless interference MIMO scenarios," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3350–3368, Aug. 2016.

[49] G. Hardy and E. Wright, *An Introduction to the Theory of Numbers*, 4th ed. Oxford, U.K.: Oxford Univ. Press, 1975.

[50] K. Ramchandran, A. Ortega, K. M. Uz, and M. Vetterli, "Multiresolution broadcast for digital HDTV using joint source/channel coding," *IEEE J. Sel. Areas Commun.*, vol. 11, no. 1, pp. 6–23, Jan. 1993.

[51] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[52] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[53] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[54] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4566–4592, Sep. 2010.

[55] V. R. Cadambe, S. A. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling—Settling the Høst–Madsen–Nosratinia conjecture," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4552–4565, Sep. 2010.

[56] V. R. Cadambe, S. A. Jafar, and S. Shamai (Shitz), "Interference alignment on the deterministic channel and application to fully connected Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 269–274, Jan. 2009.

[57] R. Prasad, S. Bhashyam, and A. Chockalingam, "On the Gaussian many-to-one X channel," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 244–259, Jan. 2016.

[58] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, Apr. 2008.

[59] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

Jian-Kang Zhang (SM'09) received the B.S. degree in information science (mathematics) from Shaanxi Normal University, Xi'an, China, the M.S. degree in information and computational science (mathematics) from Northwest University, Xi'an, and the Ph.D. degree in electrical engineering from Xidian University, Xi'an, in 1983, 1988, and 1999, respectively.

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. He holds research positions at McMaster University and Harvard University. His research interests are in the general area of signal processing, digital communication, signal detection and estimation, wavelet, and time-frequency analysis, mainly emphasizing mathematics-based new technology innovation and exploration for variety of signal processing and practical applications, and specifically, number theory, matrix theory, and linear algebra-based various kinds of signal processing. His current research focuses on transceiver designs for multiuser communication systems, coherent and noncoherent space-time signal, and receiver designs for MIMO and cooperative relay communications. He is the Co-Author of the paper that received the IEEE Signal Processing Society Best Young Author Award in 2008. He has served as an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS. He is currently serving as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the *Journal of Electrical and Computer Engineering*.

Zheng Dong received the B.Sc. and M.Eng. degrees from the School of Information Science and Engineering, Shandong University, Jinan, China, in 2009 and 2012, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, in 2016. He received the Outstanding Bachelor Thesis of Shandong University and the Outstanding Master Thesis of Shandong Province. Currently, he is a Post-Doctoral Research Fellow with the College of Information Engineering, Shenzhen University, China. He also holds a research position at McMaster University. His research interests include signal processing and information theory.

He (Henry) Chen (S'10–M'16) received the Ph.D. degree in electrical engineering (telecommunications) from The University of Sydney, Sydney, Australia, in 2015. He is currently a Post-Doctoral Research Fellow at the School of Electrical and Information Engineering, The University of Sydney. His current research interests are in the field of wireless communications, with a particular focus on ultra-reliable low latency communications, nonorthogonal multiple access, millimeter wave wireless communications, and wireless energy harvesting and transfer.
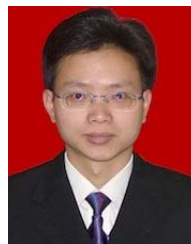
He received the Outstanding Bachelor Thesis of Shandong University, the Outstanding Master Thesis of Shandong Province, and the Chinese Government Award for Outstanding Self-Financed Students Abroad. He served as a Guest Editor of the 2016 special issue of millimeter wave wireless communications and networks in MDPI Sensors, and an Organizer and a Co-Chair of the GLOBECOM 2016 Workshop on wireless energy harvesting communication networks. He has been serving as a Regular Reviewer for over ten prestigious IEEE journals as well as the TPC Member of several IEEE flagship conferences, such as GLOBECOM and ICC.

Lei Huang (M'07–SM'14) was born in Guangdong, China. He received the B.Sc., M.Sc., and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2000, 2003, and 2005, respectively. From 2005 to 2006, he was a Research Associate with the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA. From 2009 to 2010, he was a Research Fellow with the Department of Electronic Engineering, City University of Hong Kong, and a Research Associate with the Department of Electronic Engineering, The Chinese University of Hong Kong. From 2011 to 2014, he was a Professor with the Department of Electronic and Information Engineering, Harbin Institute of Technology Shenzhen Graduate School. Since 2014, he has been with the College of Information Engineering, Shenzhen University, where he is currently a Distinguished Professor and the Director of the Shenzhen Key Lab of Advanced Navigation Technology. His research interests include array signal processing, statistical signal processing, sparse signal processing, and their applications in radar, navigation, and wireless communications. He is currently serving as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the DIGITAL SIGNAL PROCESSING. Additionally, he is an elected member of the Sensor Array and Multichannel Technical Committee of the IEEE Signal Processing Society.