Uplink Non-Orthogonal Multiple Access With Finite-Alphabet Inputs

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

***Abstract*— This paper focuses on the non-orthogonal multiple access (NOMA) design for a classical two-user multiple access channel (MAC) with finite-alphabet inputs. In contrast to most of the existing NOMA designs using continuous Gaussian input distributions, we consider practical quadrature amplitude modulation (QAM) constellations at both transmitters, the sizes of which are assumed to be not necessarily identical. We propose maximizing the minimum Euclidean distance of the received sum constellation with a maximum likelihood (ML) detector by adjusting the scaling factors (i.e., instantaneous transmitted powers and phases) of both users. The formulated problem is a *mixed continuous–discrete* optimization problem, which is nontrivial to resolve in general. By carefully observing the structure of the objective function, we define a new type of Farey sequence, termed *punched Farey sequence* to tackle the formulated problem. Based on this, we manage to achieve a closed-form optimal solution to the original problem by first dividing the entire feasible region into a finite number of Farey intervals and then taking the maximum over all possible intervals. The resulting sum constellation is proved to be a regular QAM constellation of a larger size, and hence, a simple quantization receiver can be implemented as the ML detector for the demodulation. Moreover, the superiority of NOMA over time-division multiple access in terms of minimum Euclidean distance is rigorously proved. We subsequently address how to extend our design framework intended for the two-user MAC to systems with multiple users and multiple antennas. Finally, simulation results are provided to verify our theoretical analysis and demonstrate the merits of the proposed NOMA over existing orthogonal and non-orthogonal designs.**

***Index Terms*— Non-orthogonal multiple access (NOMA), finite-alphabet inputs, multiple access channel (MAC), quadrature amplitude modulation (QAM), Farey sequence.**

I. Introduction

NON-ORTHOGONAL multiple access (NOMA) has recently emerged as a key enabling radio access technology to meet the unprecedented requirements of forthcoming fifth generation (5G) networks, due to its inherent advantages of high spectral efficiency, massive connectivity, and low transmission latency [1]–[3]. The concept of NOMA has multiple variants, such as power-domain NOMA, sparse code multiple access, pattern division multiple access, low density spreading, and lattice partition multiple access [3]. In this paper, we mainly consider the power-domain NOMA. The basic principle of power-domain NOMA is to serve more than one user with distinct channel conditions simultaneously in the same orthogonal resource block along the time, frequency, or code axes. This can be achieved by applying the superposition coding (SC) at the transmitter as well as multiuser detector (e.g., successive interference cancelation (SIC)) at the receiver side to distinguish the co-channel users. As such, NOMA is fundamentally different from conventional orthogonal multiple access (OMA) methods primarily used in the previous generations of mobile systems, where each user is allocated to one dedicated orthogonal radio resource block exclusively.

Although the OMA schemes have been widely used in the past several decades, they generally cannot achieve the whole multiuser capacity region and thus tend to have a lower spectral efficiency than NOMA approaches [1], [3]–[5]. For example, in OMA, a resource block allocated to a user with a poor channel condition cannot be reused by another user with a much stronger channel state. Apart from that, OMA is in general not scalable. This is because the amount of resource blocks as well as the granularity of user scheduling strictly limit the number of users that can be supported at the same time. On the contrary, by breaking the orthogonality of the radio resource allocation, NOMA has been shown to be able to provide better user fairness and improve physical layer security in addition to the advantages mentioned above [1], [3].

A. Related Work

Despite the fact that the deployment of NOMA as a new radio access technology in next-generation mobile systems is

Manuscript received January 10, 2018; revised April 24, 2018; accepted June 9, 2018. Date of publication June 28, 2018; date of current version September 10, 2018. The work of Z. Dong and L. Huang was supported in part by the National Natural Science Foundation of China under Grants U1713217 and U1501253, in part by the Natural Science Foundation of Guangdong Province under Grant 2015A030311030, and in part by the Foundation of Shenzhen City under Grant ZDSYS201507081625213. The work of H. Chen and B. Vucetic was supported by the Australian Research Council under Grant FL160100032. The work of J.-K. Zhang was supported by the Natural Sciences and Engineering Research Council of Canada. The associate editor coordinating the review of this paper and approving it for publication was M. R. Nakhai. *(Corresponding author: Lei Huang.)*

Z. Dong is with the College of Information Engineering, Shenzhen University, Shenzhen 518060, China, and also with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: dongz3@mcmaster.ca).

H. Chen and B. Vucetic are with the School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: he.chen@sydney.edu.au; branka.vucetic@sydney.edu.au).

J.-K. Zhang is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: jkzhang@mail.ece.mcmaster.ca).

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

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

Digital Object Identifier 10.1109/TWC.2018.2849413

relatively new, the performance of NOMA has been studied extensively in the information theory society for various channel topologies such as broadcast channel (BC) [6], multiple access channel (MAC) [7], and interference channel (IC) [8]. However, these results concentrated mainly on the study of the channel capacity region with the assumption of unlimited encoding/decoding complexity, and therefore lie mostly in the theoretical aspects due to their extremely high implementation cost. Thanks to the rapid progress of the radio frequency (RF) chain and the processing capability of mobile devices in the past decades, the implementation of NOMA is becoming more and more feasible and thus has drawn tremendous attention from both academia and industry very recently [3]. More specifically, by taking practical constraints on user fairness and/or radio resource management into consideration, NOMA has been investigated in various wireless systems, such as cognitive radio [9], cooperative communications [10], cellular uplink [11], cellular downlink [12]–[16], short-packet communications [17], and multi-cell networks [18], [19]. In fact, a two-user downlink scenario of NOMA, known as multiuser superposition transmission (MUST), has already been incorporated in the 3rd Generation Partnership Project (3GPP) Long Term Evolution-Advanced (LTE-A) [20], [21].

We note that, up to now, the vast majority of existing NOMA designs assumed the use of Gaussian input signals [6], [8]–[13], [18], [22]–[24]. Although the Gaussian input is of great significance both theoretically and practically, its implementation in reality will require huge storage capacity, unaffordable computational complexity and extremely long decoding delay [4, Ch. 9]. More importantly, the actual transmitted signals in real communication systems are drawn from finite-alphabet constellations, such as pulse amplitude modulation (PAM), quadrature amplitude modulation (QAM), and phase-shift keying (PSK) [25, Ch. 5]. Applying the results derived from the Gaussian inputs to the signals with finite-alphabet inputs can lead to significant performance loss [26]. In this sense, Gaussian input serves mostly as the theoretical benchmark.

Motivated by the above facts, the NOMA design with finite-alphabet inputs is of utmost importance and has attracted considerable efforts, see e.g., [19], [27]–[31], and references therein. The main principle[1] of these efforts is to ensure that the signal originated from each user can be uniquely decoded from the received sum-signal at the receiver side. By using mutual information as a performance measure, Harshan and Rajan [27], [28] considered the NOMA design in an ideal two-user Gaussian MAC with finite-input constellations under individual power constraint on each user. Specifically, NOMA was realized by strategically introducing certain constellation rotations (CR) to the adopted PSK signals in [27] or using proper power control in [28]. However, only numerical solutions to the optimal NOMA designs were provided in [27] and [28]. Moreover, linear precoders were considered for the MIMO MAC in [29], where the expression of the weighted sum-rate was asymptotic and the optimal solution was also numerical. Besides, the downlink NOMA system with discrete input distributions was studied in [30], where the solution is intuitive based on the deterministic approximation of the actual fading channel. The discrete input alphabets were also considered for a two-user interference channel to evaluate the capacity inner bound in [31]. In other words, all NOMA designs provided in [27]–[31] used mutual information as the performance measure, where the solutions were numerical and limited insights on the relationship between the sum-constellation and each user's constellation can thus be drawn from the obtained solutions.

[1]Note that the principle was originally proposed in the seminal work [32], [33], wherein the finite-length codeword design problem in the binary domain were considered from an information-theoretical perspective.

B. Motivation and Contributions

Inspired by the aforementioned work, in this paper we target a *closed-form* NOMA design for a classical two-user Gaussian MAC with finite-alphabet inputs and an optimal maximum likelihood (ML) detector at the receiver, where the two users are allowed to transmit simultaneously in the same frequency band. Finding the capacity bound of a Gaussian MAC with Gaussian inputs and adaptive power control has always been a classic problem, see e.g., [7], [23], [24]; the optimal power control scheme for the Gaussian MAC with finite-alphabet inputs, however, is still an open problem and only numerical solutions are available [21], [27], [28], [34]. To fill this gap, in this paper we, for the first time, investigate the optimal power control problem for the two-user Gaussian MAC with *finite square QAM constellations* that maximizes the minimum Euclidean distance of the received signals with the maximum likelihood (ML) detector. In this paper, we mainly concentrate on the error performance, which is determined by the geometry of the received sum-constellation [27], [28]. The error performance is a complicated function of the geometry of the sum-constellation, which makes our optimization problem challenging to solve. Note that QAM signaling is more spectrally efficient than other commonly-used constellations such as PSK signaling. Nevertheless, the NOMA design with QAM is more challenging than that with PSK since in QAM both the amplitude and the phase of the modulated signal vary, while in PSK only the phase is different, and thus the unambiguity of the sum-constellation at the receiver side is much more difficult to maintain. The main contributions of this paper can be summarized as follows:

1) We develop a practical NOMA design for the classical two-user complex Gaussian MAC, where the two users are allowed to adopt not necessarily the same QAM constellations. In our design framework, we aim to maximize the minimum Euclidian distance of the received sum-constellation at the receiver side, which dominates the error performance of the considered system, by adjusting the transmit power and phase of each user. To this end, we first decompose the complex MAC design problem into two real MAC design problems by strategically rotating the phase of the input signals at the two users. Nevertheless, the decomposed problems are still non-trivial due to their mixed continuous-and-discrete feature.

2) To address this challenging problem, we define a new type of Farey sequence, termed *punched Farey sequence*, which is essential for our NOMA design with not necessarily the same QAM constellations. This concept is even mathematically new to the best of our knowledge [35]. We identify and rigourously prove several important properties of the punched Farey sequence in parallel to the conventional Farey sequence. Based on the punched Farey sequence and its important properties, we manage to resolve the above decomposed problem for each channel branch by providing a neat closed-form optimal solution, which reveals that the optimal sum-constellation is a regular QAM constellation of a larger size. Due to this nice structure of the sum-constellation, a simple quantization decoder can be employed to implement the ML detector.
3) Based on the obtained closed-form solution, we prove the superiority of this NOMA design over the time-division multiple access (TDMA) approach in terms of the minimum Euclidean distance at the receiver for arbitrary given channel realization and rate allocation. Actually, this is a surprising result since the new NOMA method can achieve a better error performance than TDMA in a high SNR regime even if there is no near-far effect. We subsequently address how to extend our design framework intended for the two-user MAC to systems with multiple users and multiple antennas. Finally, simulation results are provided to verify our theoretical analysis and demonstrate the merits of the proposed NOMA over existing orthogonal and non-orthogonal designs.

II. Two-User Gaussian Multiple-Access Channel

We consider a two-user Gaussian MAC given by

$$z = h_1 x_1 + h_2 x_2 + \xi, \tag{1}$$

where z is the received signal at the base station (BS), h_k denotes the complex channel coefficient between the transmitter S_k and BS for $k = 1, 2$, and ξ is the additive zero-mean, circularly symmetric complex Gaussian (CSCG) noise with variance $2\sigma^2$, i.e., $\xi \sim \mathcal{CN}(0, 2\sigma^2)$. We assume that perfect channel state information (CSI) is available to all the nodes and symbol synchronization is maintained at BS. The transmitted symbols x_k are superimposed at the receiver in a NOMA manner which are chosen randomly, independently and equally likely from the (finite) square QAM constellation $\mathcal{Q}_k$, and are subject to individual average power constraint P_k, i.e., $\mathbb{E}[|x_k|^2] \le P_k$ for $k = 1, 2$.

Although we use a complex baseband representation in (1), the modulated and demodulated signals are real since the oscillator at the transmitter can only generate real sinusoids rather than complex exponentials, and the channel then introduces amplitude and phase distortion to the transmitted signals [25]. As such, we follow [28] to decompose the considered complex Gaussian MAC given in (1) into two parallel real-scalar Gaussian MACs, which are called the in-phase and quadrature components, respectively [25]. This means that the original two-dimensional QAM constellation can be split into two one-dimensional PAM constellations to be transmitted via the in-phase and quadrature branches. Besides, since the in-phase and quadrature components of the sum-constellation are separable, they can be decoded independently at the receiver, thereby reducing the decoding complexity. Mathematically, we notice that (1) is equivalent to

$$z = |h_1| x_1 \exp(j \arg(h_1)) + |h_2| x_2 \exp(j \arg(h_2)) + \xi. \tag{2}$$

To simplify the subsequent expressions, we let $y = \mathrm{Re}(z)$, $y' = \mathrm{Im}(z)$, $w_1 s_1 = \mathrm{Re}(x_1 \exp(j \arg(h_1)))$, $w_1' s_1' = \mathrm{Im}(x_1 \exp(j \arg(h_1)))$, $w_2 s_2 = \mathrm{Re}(x_2 \exp(j \arg(h_2)))$, $w_2' s_2' = \mathrm{Im}(x_2 \exp(j \arg(h_2)))$, $n = \mathrm{Re}(\xi)$ and $n' = \mathrm{Im}(\xi)$, where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ are the real and imaginary parts of the complex number, respectively. Besides, w_1, w_2, w_1', and w_2' are the real non-negative scalars determining the minimum Euclidean distance of the actual transmitted PAM constellation sets, which are referred to as the *weighting coefficients* throughout this paper. Now, the in-phase and quadrature branches of (1) can be reformulated by

$$y = |h_1| w_1 s_1 + |h_2| w_2 s_2 + n, \tag{3a}$$

$$y' = |h_1| w_1' s_1' + |h_2| w_2' s_2' + n', \tag{3b}$$

where $n, n' \sim \mathcal{N}(0, \sigma^2)$ are independent and identically distributed (i.i.d.) real additive white Gaussian components since the complex noise term ξ is assumed to be CSCG noise.

Without loss of generality, we assume that $x_1 \exp(j \arg(h_1)) \in \mathcal{Q}_1$ and $x_2 \exp(j \arg(h_2)) \in \mathcal{Q}_2$, where $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are M_1^2- and M_2^2-ary square QAM constellations (M_1 and M_2 are both no less than 2 but not necessarily equal to each other), respectively, given by $\mathcal{Q}_1 \triangleq \{\pm w_1(2k-1) \pm w_1'(2\ell-1)j : k, \ell = 1, \ldots, M_1/2\}$ and $\mathcal{Q}_2 \triangleq \{\pm w_2(2k-1) \pm w_2'(2\ell-1)j : k, \ell = 1, \ldots, M_2/2\}$. As a result, the information-bearing symbols $s_1, s_1' \in \mathcal{A}_{M_1} = \{\pm(2k-1)\}_{k=1}^{M_1/2}$, sent by S_1, and $s_2, s_2' \in \mathcal{A}_{M_2} = \{\pm(2k-1)\}_{k=1}^{M_2/2}$, transmitted by S_2, are drawn from the standard PAM constellations with equal probability. We consider that an equal power allocation between two branches is performed to balance the minimum Euclidean distance of the two PAM constellations [25, Ch. 6.1.4] and the transmitted signals over both subchannels should still be subject to average power constraints, i.e., $\mathbb{E}[w_1^2 |s_1|^2] \le P_1/2$, $\mathbb{E}[w_1'^2 |s_1'|^2] \le P_1/2$, $\mathbb{E}[w_2^2 |s_2|^2] \le P_2/2$, and $\mathbb{E}[w_2'^2 |s_2'|^2] \le P_2/2$.

An important problem for the considered MAC is, 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 the receiver, subject to the individual average power constraints at both transmitters. As the in-phase and quadrature subchannels are symmetric, if the same algorithm is applied to both branches, we will expect to have $w_1 = w_1'$ and $w_2 = w_2'$, and we call $\mathcal{Q}_1$ and $\mathcal{Q}_2$ the symmetric square QAM constellations. It is worth mentioning that our framework can be readily extended to un-symmetric signaling [36], [37], i.e., un-equal power allocation between the two branches. By leveraging the decomposable property of the complex Gaussian MAC and the symmetry of the two subchannels, we can simply focus on the design for one of the two real-scalar Gaussian MACs with PAM constellation sets, which will be elaborated in next section.[2]

III. The Weighting Coefficients Design for the Real-Scalar Gaussian MAC

In this section, we consider the constellation design problem, i.e., finding the optimal weighting coefficients w_1 and w_2, for the in-phase real-scalar Gaussian MAC. As the two subchannels are symmetric, the optimal solution to the quadrature component can be obtained in exactly the same way and hence is omitted for brevity.

A. Problem Formulation

Recall that $\mathbb{E}[w_1^2|s_1|^2] \le P_1/2$, $\mathbb{E}[w_2^2|s_2|^2] \le P_2/2$, and hence $0 < w_1 \le \sqrt{\frac{3\ P_1}{2(M_1^2-1)}}$, $0 < w_2 \le \sqrt{\frac{3\ P_2}{2(M_2^2-1)}}$. For notation simplicity, we set $|\tilde{h}_1| = \sqrt{\frac{3\ P_1}{2(M_1^2-1)}}|h_1|$, $|\tilde{h}_2| = \sqrt{\frac{3\ P_2}{2(M_2^2-1)}}|h_2|$ and

$$\tilde{w}_1 = \sqrt{\frac{2(M_1^2-1)}{3P_1}}w_1, \quad \tilde{w}_2 = \sqrt{\frac{2(M_2^2-1)}{3P_2}}w_2, \tag{4}$$

where $0 < \tilde{w}_1 \le 1$ and $0 < \tilde{w}_2 \le 1$. The received signal in (3a) can thus be re-written as

$$y = |\tilde{h}_1|\tilde{w}_1 s_1 + |\tilde{h}_2|\tilde{w}_2 s_2 + n, \tag{5}$$

where $s_1 \in \mathcal{A}_{M_1} = \{\pm(2k-1)\}_{k=1}^{M_1/2}$ and $s_2 \in \mathcal{A}_{M_2} = \{\pm(2k-1)\}_{k=1}^{M_2/2}$.

We note that both SIC and joint decoding are popular decoding techniques for non-orthogonal transmissions [4], [5]. In this paper, instead of using a SIC receiver as in downlink NOMA, we assume that a coherent joint maximum-likelihood (ML) detector is used at BS to estimate the transmitted signals in a symbol-by-symbol fashion. This is mainly because that the receiver wants to decode the information from both users, which makes joint decoding more appropriate to apply. Furthermore, it has been shown that SIC decoder can only achieve the corner point of the pentagonal capacity region of the MAC in the uplink without a further time-sharing operation, while the joint decoder can achieve any point in the capacity region without the need of time-sharing [5, Ch. 4.5]. Another important advantage of joint detection relative to SIC is that there is no error propagation. This is crucial for the considered symbol-by-symbol detector at the BS, where SIC is prone to the error propagation. On the other hand, it is worth mentioning that compared to the SIC receiver, a potential drawback of a joint decoder can be its relatively higher decoding complexity. Since we perform a symbol-by-symbol detection, the decoding complexity is at most $\mathcal{O}(M_1\ M_2)$ with M_1 and M_2 being the PAM constellation size of s_1 and s_2, respectively. We later show that a simple quantization receiver can be used to implement the joint ML detection with a low complexity of $\mathcal{O}(1)$. Mathematically, the estimated signals can be expressed as

$$(\hat{s}_1, \hat{s}_2) = \arg \min_{(s_1,s_2)} \left|y - (|\tilde{h}_1|\tilde{w}_1 s_1 + |\tilde{h}_2|\tilde{w}_2 s_2)\right|. \tag{6}$$

By applying the nearest neighbor approximation method [25, Ch. 6.1.4] at high SNRs for 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. As such, in this paper, we aim to devise the optimal value of $(\tilde{w}_1, \tilde{w}_2)$ (or equivalently constellations $\mathcal{Q}_1$ and $\mathcal{Q}_2$) to maximize the minimum Euclidean distance of constellation points of the received signal. The Euclidean distance between the two received signals $y(s_1, s_2)$ and $y(\tilde{s}_1, \tilde{s}_2)$ at the receiver for (s_1, s_2) and $(\tilde{s}_1, \tilde{s}_2)$ in the noise-free case is given by

$$|y(s_1,s_2) - y(\tilde{s}_1,\tilde{s}_2)| = \left||\tilde{h}_1|\tilde{w}_1(s_1-\tilde{s}_1) - |\tilde{h}_2|\tilde{w}_2(\tilde{s}_2-s_2)\right|. \tag{7}$$

Note that s_1, $\tilde{s}_1$, s_2 and $\tilde{s}_2$ are all odd numbers, and thus we can let $s_1 - \tilde{s}_1 = 2\ n$ and $\tilde{s}_2 - s_2 = 2\ m$, in which $n \in \mathbb{Z}_{M_1-1}$ and $m \in \mathbb{Z}_{M_2-1}$ with $\mathbb{Z}_N \triangleq \{0, \pm 1, \cdots, \pm N\}$ denoting the set containing all the possible differences. Similarly, we also define $\mathbb{Z}^2_{(M_1-1,M_2-1)} \triangleq \{(a,b) : a \in \mathbb{Z}_{M_1-1}, b \in \mathbb{Z}_{M_2-1}\}$, and $\mathbb{N}^2_{(M_1-1,M_2-1)} \triangleq \{(a,b) : a \in \mathbb{N}_{M_1-1}, b \in \mathbb{N}_{M_2-1}\}$ where $\mathbb{N}_N \triangleq \{0, 1, \cdots, N\}$. From the definitions above, $(s_1, s_2) \ne (\tilde{s}_1, \tilde{s}_2)$ is equivalent to $(m, n) \ne (0, 0)$ (i.e., $m \ne 0$ or $n \ne 0$). To proceed, we define

$$\begin{aligned} d(m,n) &= \frac{1}{2}|y(s_1,s_2) - y(\tilde{s}_1,\tilde{s}_2)| \\ &= \left||\tilde{h}_1|\tilde{w}_1\ n - |\tilde{h}_2|\tilde{w}_2\ m\right|, \\ &\quad (m,n) \in \mathbb{Z}^2_{(M_1-1,M_2-1)} \setminus \{(0,0)\}, \end{aligned} \tag{8}$$

where $\mathcal{A} \setminus \mathcal{B} \triangleq \{x \in \mathcal{A} \text{ and } x \notin \mathcal{B}\}$. We are at a point to formally formulate the following max-min optimization problem,

Problem 1 (Power Control of NOMA in Real-Scalar MAC With PAM Constellation): Find the optimal value of $(\tilde{w}_1^*, \tilde{w}_2^*)$ subject to the individual average power constraint such that the minimum Euclidean distance d^* of the received signal constellation points is maximized, i.e.,

$$(\tilde{w}_1^*, \tilde{w}_2^*) = \arg \max_{(\tilde{w}_1,\tilde{w}_2)} \min_{(m,n)\in\mathbb{Z}^2_{(M_1-1,M_2-1)}\setminus\{(0,0)\}} d(m,n) \tag{9a}$$

$$\text{s.t. } 0 < \tilde{w}_1 \le 1 \text{ and } 0 < \tilde{w}_2 \le 1. \tag{9b}$$

Note that the inner optimization variable of finding the minimum Euclidean distances is discrete, while the outer one $(\tilde{w}_1, \tilde{w}_2)$ is continuous. 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 open literature [21], [27], [28], [34]. To optimally and systematically solve this problem, we now develop a design framework based on the *Farey sequence* [35], in which the entire feasible region of $(\tilde{w}_1, \tilde{w}_2)$ is divided into a finite

[2]It should be pointed out that designing two PAM constellations for both subchannels separately is a practical but not necessarily optimal approach. In fact, this approach has been widely adopted in literature, such as in [36]–[38]. How to design a two-dimensional complex constellation directly for the Gaussian MAC has been left as a future work.

number of mutually exclusive sub-regions. Then, for each sub-region, 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 possible sub-regions. We first consider the inner optimization problem in (9) given by:

Problem 2 (Finding Differential Pairs With the Minimum Euclidean Distance):

$$\min_{(m,n)\in\mathbb{Z}^2_{(M_1-1,M_2-1)}\backslash\{(0,0)\}} d(m,n) = \min_{(m,n)\in\mathbb{Z}^2_{(M_1-1,M_2-1)}\backslash\{(0,0)\}} \left||\tilde{h}_1|\tilde{w}_1 n - |\tilde{h}_2|\tilde{w}_2 m\right|. \quad (10)$$

We should point out that finding the closed-form solution to the optimal (m,n) for (10) is not trivial since the solution depends on the values of $|\tilde{h}_1|$ and $|\tilde{h}_2|$, which can span the whole positive real axis. Moreover, the values of $\tilde{w}_1$ and $\tilde{w}_2$ will be optimized later and cannot be determined beforehand. It is worth mentioning here that a similar optimization problem was formulated and resolved for a Gaussian Z channel in [19]. In [19], we resorted to the existing Farey sequence to solve the formulated problem. However, due to the inherent symmetric structure between numerators and denominators of the conventional Farey sequence, our results presented in [19] refers only to the case where both transmitters need to use exactly *identical* constellation size (i.e., the same transmission rate) and thus cannot be applied to the problem in this paper with M_1 and M_2 not necessarily the same. Motivated by this, in this paper we define a new type of Farey sequence, termed punched Farey sequence. In the subsequent section, we will introduce the definition and some important properties of the original Farey sequence and the developed punched 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 [35]): The Farey sequence $\mathfrak{F}_K$ is the ascending sequence of irreducible fractions between 0 and 1 whose denominators are less than or equal to K. ∎

By the definition, $\mathfrak{F}_K = \left(\frac{b_k}{a_k}\right)_{k=1}^{|\mathfrak{F}_K|}$ 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. In addition, $|\mathfrak{F}_K| = 1 + \sum_{m=1}^{K}\varphi(m)$ is the cardinality of $\mathfrak{F}_K$ with $\varphi(\cdot)$ being the Euler's totient function [35]. An example of Farey sequence is given as follows:

Example 1: $\mathfrak{F}_5$ is the ordered sequence $\left(\frac{0}{1},\frac{1}{5},\frac{1}{4},\frac{1}{3},\frac{2}{5},\frac{1}{2},\frac{3}{5},\frac{2}{3},\frac{3}{4},\frac{4}{5},\frac{1}{1}\right)$.

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

It can be observed that the extended Farey sequence starts with number 0 (fraction $\frac{0}{1}$) and end with ∞ (fraction $\frac{1}{0}$). We now propose a new definition called *Punched Farey sequence* in number theory as follows.

Definition 3 (Punched Farey Sequence): The punched (extended) Farey sequence $\mathfrak{P}_K^L$ is the ascending sequence of irreducible fractions whose denominators are no greater than K and numerators are no greater than L. ∎

Example 3: $\mathfrak{P}_5^2$ is the ordered sequence $\left(\frac{0}{1},\frac{1}{5},\frac{1}{4},\frac{1}{3},\frac{2}{5},\frac{1}{2},\frac{2}{3},\frac{1}{1},\frac{2}{1},\frac{1}{0}\right)$.

From Definition 3, when $L = K$, $\mathfrak{P}_K^K$ degenerates into Farey sequence $\mathfrak{F}_K$, i.e., $\mathfrak{P}_K^K = \mathfrak{F}_K$. We can also observe that each punched Farey sequence begins with number 0 (fraction $\frac{0}{1}$) and ends with ∞ (fraction $\frac{1}{0}$).

We now develop some elementary properties of the punched Farey sequence in line with Farey sequences [35]. It is worth pointing out that, although for some properties, we can find the counterparts in conventional Farey sequences, the extension to the punched Farey sequences is non-trivial and the following results are new.

Property 1: If $\frac{n_1}{m_1}$ and $\frac{n_2}{m_2}$ are two adjacent terms (called Farey pairs) in $\mathfrak{P}_K^L$ ($\min\{K,L\} \ge 2$) such that $\frac{n_1}{m_1} < \frac{n_2}{m_2}$, then, 1) $\frac{n_1+n_2}{m_1+m_2} \in \left(\frac{n_1}{m_1},\frac{n_2}{m_2}\right)$, $\frac{m_1+m_2}{n_1+n_2} \in \left(\frac{m_2}{n_2},\frac{m_1}{n_1}\right)$; 2) $m_1\ n_2 - m_2\ n_1 = 1$; 3) If $n_1+n_2 \le L$, then $m_1+m_2 > K$ and if $m_1+m_2 \le K$, then $n_1+n_2 > L$; 4) $n_1+n_2 \ge 1$ where the equality is attained if and only if $\frac{n_1}{m_1} = \frac{0}{1}$ and $\frac{n_2}{m_2} = \frac{1}{K}$. Likewise, $m_1+m_2 \ge 1$ where the equality is attained if and only if $\frac{n_1}{m_1} = \frac{L}{1}$ and $\frac{n_2}{m_2} = \frac{1}{0}$. ∎

The proof is given in Appendix-A.

Property 2: If $\frac{n_1}{m_1}$, $\frac{n_2}{m_2}$ and $\frac{n_3}{m_3}$ are three consecutive terms in $\mathfrak{P}_K^L$ with $\min\{K,L\} \ge 2$ such that $\frac{n_1}{m_1} < \frac{n_2}{m_2} < \frac{n_3}{m_3}$, then $\frac{n_2}{m_2} = \frac{n_1+n_3}{m_1+m_3}$. ∎

The proof is provided in Appendix-B.

Property 3: Consider $\frac{n_1}{m_1},\frac{n_2}{m_2},\frac{n_3}{m_3},\frac{n_4}{m_4} \in \mathfrak{P}_K^L$ with $\min\{K,L\} \ge 3$, such that $\frac{n_1}{m_1} < \frac{n_2}{m_2} < \frac{n_3}{m_3} < \frac{n_4}{m_4}$ where $\frac{n_2}{m_2},\frac{n_3}{m_3}$ are successive in $\mathfrak{P}_K^L$, then $\frac{n_1+n_3}{m_1+m_3} \le \frac{n_2}{m_2}$ and $\frac{n_3}{m_3} \le \frac{n_2+n_4}{m_2+m_4}$. ∎

The proof is provided in Appendix-C.

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

We are now ready to solve Problem 2 to find the differential pairs (m,n) having the minimum Euclidean distance. To this end, we first introduce the following preliminary propositions.

Proposition 1: Let $\mathbb{F}^2_{(M_1-1,M_2-1)} = \{(m,n) : \frac{n}{m} \in \mathfrak{P}^{M_1-1}_{M_2-1}\}$, and then $\min_{(m,n)\in\mathbb{Z}^2_{(M_1-1,M_2-1)}\setminus\{(0,0)\}} d(m,n) = \min_{(m,n)\in\mathbb{F}^2_{(M_1-1,M_2-1)}} d(m,n)$. ■

The proof is similar to [19, Appendix A] and hence is omitted for brevity.

Proposition 2: Let $\frac{n_1}{m_1}$ and $\frac{n_2}{m_2}$ be two terms of $\mathfrak{P}^{M_1-1}_{M_2-1}$ such that $\frac{n_1}{m_1} < \frac{n_2}{m_2}$. Then, for $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in (\frac{n_1}{m_1}, \frac{n_2}{m_2})$ and $d(m,n) = \big||\tilde{h}_1|\tilde{w}_1\ n - |\tilde{h}_2|\tilde{w}_2\ m\big|$, we have 1) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} = \frac{n_1+n_2}{m_1+m_2}$, then $d(m_1,n_1) = d(m_2,n_2)$; 2) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \big(\frac{n_1}{m_1}, \frac{n_1+n_2}{m_1+m_2}\big)$, then $d(m_1,n_1) < d(m_2,n_2)$; 3) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \big(\frac{n_1+n_2}{m_1+m_2}, \frac{n_2}{m_2}\big)$, then $d(m_2,n_2) < d(m_1,n_1)$. ■

The proof can be found in Appendix-D.

We now give the solution to Problem 2 in the following proposition:

Proposition 3: For any $\frac{n_1}{m_1}, \frac{n_2}{m_2}, \frac{n_3}{m_3}, \frac{n_4}{m_4} \in \mathfrak{P}^{M_1-1}_{M_2-1}$ with $|\mathfrak{P}^{M_1-1}_{M_2-1}| \geq 4$, such that $\frac{n_1}{m_1} < \frac{n_2}{m_2} < \frac{n_3}{m_3} < \frac{n_4}{m_4}$, and $\frac{n_2}{m_2}, \frac{n_3}{m_3}$ are successive in $\mathfrak{P}^{M_1-1}_{M_2-1}$, we have 1) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \big(\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}\big)$, then $\min_{(m,n)\in\mathbb{F}^2_{(M_1-1,M_2-1)}} d(m,n) = d(m_2,n_2) = |\tilde{h}_2|\tilde{w}_2\ m_2 - |\tilde{h}_1|\tilde{w}_1\ n_2$; 2) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \big(\frac{n_2+n_3}{m_2+m_3}, \frac{n_3}{m_3}\big)$, then $\min_{(m,n)\in\mathbb{F}^2_{(M_1-1,M_2-1)}} d(m,n) = d(m_3,n_3) = |\tilde{h}_1|\tilde{w}_1\ n_3 - |\tilde{h}_2|\tilde{w}_2\ m_3$. ■

The proof is given in Appendix-E.

D. Closed-Form Optimal Solution to Problem 1

With the optimal solution to the inner optimization problem (i.e., Problem 2) given in Proposition 3, we now find the optimal solution to the outer optimization on $(\tilde{w}_1,\tilde{w}_2)$ of Problem 1 in a closed form. To facilitate our analysis, we denote the entire feasible region of Problem 1 in terms of $(\tilde{w}_1,\tilde{w}_2)$ as $\mathcal{U} = \big\{(\tilde{w}_1,\tilde{w}_2) : 0 < \tilde{w}_1 \leq 1, 0 < \tilde{w}_2 \leq 1\big\}$. From the above discussion, we know that the positive real axis can be divided into $C-1$ mutually exclusive sub-intervals by the punched Farey sequence $\mathfrak{P}^{M_1-1}_{M_2-1} = \big(\frac{b_1}{a_1}, \frac{b_2}{a_2}, \cdots, \frac{b_C}{a_C}\big)$, where $C = |\mathfrak{P}^{M_1-1}_{M_2-1}|$. By denoting $\mathcal{A}_k = \big\{(\tilde{w}_1,\tilde{w}_2) : 0 < \tilde{w}_1 \leq 1, 0 < \tilde{w}_2 \leq 1, \frac{b_k}{a_k} < \frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \leq \frac{b_{k+1}}{a_{k+1}}\big\}$, $k = 1,2,\ldots,C-1$ where $\mathcal{U} = \bigcup_{k=1}^{C-1}\mathcal{A}_k$, we can solve Problem 1 by restricting $(\tilde{w}_1,\tilde{w}_2) \in \mathcal{A}_k$. More specifically, we aim to find the optimal $(\tilde{w}_1^*(k),\tilde{w}_2^*(k))$ such that

$$g\Big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\Big) = \max_{(\tilde{w}_1,\tilde{w}_2)} \min_{(m,n)\in\mathbb{F}^2_{(M_1-1,M_2-1)}} d(m,n) \quad \text{s.t. } \frac{b_k}{a_k} < \frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \leq \frac{b_{k+1}}{a_{k+1}},\ 0<\tilde{w}_1\leq 1 \text{ and } 0<\tilde{w}_2\leq 1. \tag{11}$$

By applying the propositions in last subsections, we obtained the following lemma related to the optimal solution to problem (11).

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

$$g\Big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\Big) = \begin{cases} \frac{|\tilde{h}_2|}{b_k+b_{k+1}}, \quad \text{with} \\ (\tilde{w}_1^*(k),\tilde{w}_2^*(k)) = \big(\frac{|\tilde{h}_2|(a_k+a_{k+1})}{|\tilde{h}_1|(b_k+b_{k+1})}, 1\big), \\ \text{if } \frac{|\tilde{h}_2|}{|\tilde{h}_1|} \leq \frac{b_k+b_{k+1}}{a_k+a_{k+1}}; \\ \frac{|\tilde{h}_1|}{a_k+a_{k+1}}, \quad \text{with} \\ (\tilde{w}_1^*(k),\tilde{w}_2^*(k)) = \big(1, \frac{|\tilde{h}_1|(b_k+b_{k+1})}{|\tilde{h}_2|(a_k+a_{k+1})}\big), \\ \text{if } \frac{|\tilde{h}_2|}{|\tilde{h}_1|} > \frac{b_k+b_{k+1}}{a_k+a_{k+1}}. \end{cases}$$

■

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

Now, we are ready to present the closed-form optimal solution to Problem 1 in terms of (w_1^*,w_2^*) instead of $(\tilde{w}_1^*,\tilde{w}_2^*)$ defined in (4) for clarity, which maximizes the minimum Euclidean distance of the sum-constellation, denoted by d_{noma}, over the entire feasible region. The main idea is to find the optimal solutions to problem (11) for all the possible sub-intervals $\mathcal{A}_k$, $k = 1,2,\ldots,C-1$, and then taking the one that maximizes the objective function of problem (9).

Theorem 1 (Closed-Form Optimal Weighting Coefficients): The optimal solution to Problem 1 in terms of (w_1^*,w_2^*) is given by:

$$(w_1^*,w_2^*) = \begin{cases} \Big(\sqrt{\frac{3P_2M_2^2}{2(M_2^2-1)}}\frac{|h_2|}{|h_1|}, \sqrt{\frac{3P_2}{2(M_2^2-1)}}\Big), \\ \quad \text{if } \frac{|h_2|}{|h_1|} \leq \sqrt{\frac{P_1(M_2^2-1)}{P_2M_2^2(M_1^2-1)}}; \\ \Big(\sqrt{\frac{3P_1}{2(M_1^2-1)}}, \sqrt{\frac{3P_1}{2M_2^2(M_1^2-1)}}\frac{|h_1|}{|h_2|}\Big), \\ \quad \text{if } \sqrt{\frac{P_1(M_2^2-1)}{P_2M_2^2(M_1^2-1)}} < \frac{|h_2|}{|h_1|} \leq \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2M_2^2(M_1^2-1)}}; \\ \Big(\sqrt{\frac{3P_2}{2M_1^2(M_2^2-1)}}\frac{|h_2|}{|h_1|}, \sqrt{\frac{3P_2}{2(M_2^2-1)}}\Big), \\ \quad \text{if } \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2M_2^2(M_1^2-1)}} < \frac{|h_2|}{|h_1|} \leq \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2(M_1^2-1)}}; \\ \Big(\sqrt{\frac{3P_1}{2(M_1^2-1)}}, \sqrt{\frac{3P_1M_1^2}{2(M_1^2-1)}}\frac{|h_1|}{|h_2|}\Big), \\ \quad \text{if } \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2(M_1^2-1)}} < \frac{|h_2|}{|h_1|}. \end{cases} \tag{12}$$

The resulting minimum Euclidean distance d_{noma} in each case is:

$$d_{\mathrm{noma}} = \begin{cases} \sqrt{\frac{3P_2}{2(M_2^2-1)}}|h_2|, & \text{if } \frac{|h_2|}{|h_1|} \le \sqrt{\frac{P_1(M_2^2-1)}{P_2M_2^2(M_1^2-1)}}; \\ \sqrt{\frac{3P_1}{2M_2^2(M_1^2-1)}}|h_1|, & \text{if } \sqrt{\frac{P_1(M_2^2-1)}{P_2M_2^2(M_1^2-1)}} < \frac{|h_2|}{|h_1|} \le \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2M_2^2(M_1^2-1)}}; \\ \sqrt{\frac{3P_2}{2M_1^2(M_2^2-1)}}|h_2|, & \text{if } \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2M_2^2(M_1^2-1)}} < \frac{|h_2|}{|h_1|} \le \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2(M_1^2-1)}}; \\ \sqrt{\frac{3P_1}{2(M_1^2-1)}}|h_1|, & \text{if } \sqrt{\frac{P_1M_1^2(M_2^2-1)}{P_2(M_1^2-1)}} < \frac{|h_2|}{|h_1|}. \end{cases} \tag{13}$$

■

The proof is provided in Appendix-G.

Remark 1: We note that for each transmission block, the complexity of solving Problem 1 using the exhaustive search approach can be as high as $\mathcal{O}(L_1\ L_2\ M_1\ M_2)$, where L_1 and L_2 are the discretization levels for the feasible ranges of $\tilde{w}_1$ and $\tilde{w}_2$, respectively. In contrast, our proposed solution in (12) can only involve some simple algebraic operations. Furthermore, the number of these operations is constant irrelevant to the values of L_1, L_2, M_1, and M_2. Hence the computational complexity of the proposed method in (12) is $\mathcal{O}(1)$. ■

We have the following remark regarding the choice of constellation size M_1, M_2.

Remark 2: In order to attain the results in Theorem 1 with the aid of Farey sequence, we assume that $\min\{M_1, M_2\} \ge 2$. However, it can be verified that for $M_1 = 1, M_2 \ge 2$ or $M_1 \ge 2, M_2 = 1$, although (12) is no longer true, (13) still holds. In fact, if $M_1 = 1, M_2 \ge 2$, we have $(w_1^*, w_2^*) = (0, \sqrt{\frac{3\ P_2}{2(M_2^2-1)}})$. Else if $M_1 \ge 2, M_2 = 1$, we have $(w_1^*, w_2^*) = (\sqrt{\frac{3\ P_1}{2(M_1^2-1)}}, 0)$. That is, by assuming $M_k = 1$, $k = 1, 2$, i.e., no information is transmitted by user S_k, we should let it keep silent, and thus all the channel resources are allocated to the other user exclusively, who should transmit at its maximum allowable power. ■

Remark 3: Note that our design has already taken the user fairness into consideration. Specifically, we maximize the minimum Euclidean distance of both users at the receiver side, which is essentially a max-min design. We can further optimize the values of M_1 and M_2 to improve the user fairness, e.g., we can simply let $M_1 = M_2$ to achieve uncoded transmission rate fairness. ■

Remark 4: The final neat results shown in Theorem 1 can be applied to arbitrary but known channel condition. In practice, to implement the proposed scheme in fading channels, channel training should be performed at the beginning of each transmission block to obtain the instantaneous values of h_1 and h_2. Once the channel coefficients are known, the optimal weighting coefficients that maximize the minimum Euclidean distance of the received signal can be obtained immediately by applying the results given in Theorem 1. It is also worth mentioning that the assumption of the availability of CSI at the receiver side is practical in 3G and beyond, where it can be obtained by using training pilot [25]. In practice, Problem 1 can be solved by applying Theorem 1 at BS which knows the global CSI and then feeds back the weighting coefficients to all users. ■

Remark 5: Although the proposed NOMA design given in Theorem 1 with finite-alphabet inputs is targeted for uncoded system, it can be readily applied to coded systems with hard-decision based detectors. By using the maximum-likelihood (ML) detector to extract the transmitted signals in a symbol-by-symbol fashion in (6), we get an estimated value of $(\hat{s}_1, \hat{s}_2)$ for each component, or equivalently $(\hat{x}_1, \hat{x}_2)$ for the original complex channel. Then, the considered two-user MAC is split into two virtual point-to-point digital channels with QAM constellations, where independent error-correcting channel codes can be naturally employed to each digital channel. In practical systems, soft-decision based multiuser detectors (e.g., bit log-likelihood ratio (LLR) based decoder) are used more widely than the hard-decision based ones due to the superior performance. As shown in [39] and [40], when bit interleaved coded modulation (BICM) is employed with bit LLR based soft-decision decoder, for given channel coding (gain), the minimum Euclidean distance of the received sum-constellation plays a dominant role in the system bit error rate (BER). In this sense, the adopted minimum-Euclidean-distance-maximization principle is also a reasonable design criterion in many systems with soft-decision based decoders. ■

We also have the following corollary about the optimal solution described in Theorem 1:

Corollary 1: The sum-constellation at the receiver is a *standard* $M_1^2\ M_2^2$-QAM constellation with the minimum Euclidean distance d_{noma} affected by the instantaneous channel realizations as given in (13). ■

The proof is provided in Appendix-H.

Due to this nice structure of the sum-constellation, the ML decoder reduces to a simple quantizer for the complex constellation with complexity of $\mathcal{O}(1)$ [14], where the detection can be performed for the in-phase and quadrature components separately since they are separable. It is worth mentioning that if $\frac{|h_2|}{|h_1|} \le \sqrt{\frac{P_1\ M_1^2(M_2^2-1)}{P_2\ M_2^2(M_1^2-1)}}$, we have $\frac{|h_1|w_1^*}{|h_2|w_2^*} = M_2$, i.e., the constellation of S_2 will have a smaller Euclidean distance than that of S_1 at the receiver side; Otherwise if $\frac{|h_2|}{|h_1|} > \sqrt{\frac{P_1\ M_1^2(M_2^2-1)}{P_2\ M_2^2(M_1^2-1)}}$, we attain $\frac{|h_1|w_1^*}{|h_2|w_2^*} = \frac{1}{M_1}$, i.e., the constellation of S_1 will have a smaller Euclidean distance than that of S_2.

E. The Superiority of NOMA Over TDMA

It is significant to conduct comparisons between NOMA and OMA, such as in [41]. For the sake of fairness, we compare the minimum Euclidean distance of the proposed NOMA and that of TDMA under the same channel realization for the real-scalar in-phase component of the Gaussian MAC where both methods are using PAM constellations. In general, for TDMA, the overall available frame is partitioned *uniformly* into orthogonal time slots of the same length for the ease of symbol synchronization. Specifically, for a two-user TDMA, we assume that each user has half of the total available time slots and therefore, they should employ M_1^2- and M_2^2-ary PAM constellations, respectively, to maintain the same transmission rate. In this comparison, we also assume that the channel state of both users remains unchanged (i.e., quasi-static) during the two consecutive time slots.

For TDMA, the minimum Euclidean distance for users S_1 and S_2 are $d_{\text{oma},1} = \sqrt{\frac{3\ P_1}{2(M_1^4-1)}}|h_1|$ and $d_{\text{oma},2} = \sqrt{\frac{3\ P_2}{2(M_2^4-1)}}|h_2|$, respectively. Now, we denote the minimum Euclidean distance among the two users as:

$$\begin{aligned} d_{\text{oma}} &= \min\ \{d_{\text{oma},1}, d_{\text{oma},2}\} \\ &= \min\ \left\{\sqrt{\frac{3P_1}{2(M_1^4-1)}}|h_1|,\ \sqrt{\frac{3P_2}{2(M_2^4-1)}}|h_2|\right\}. \end{aligned} \tag{14}$$

We then have the following corollary regarding the resulting minimum Euclidean distance of both schemes:

Corollary 2: The minimum Euclidean distance of the proposed NOMA, d_{noma} given in (13), is strictly larger than that of the TDMA scheme, d_{oma} given in (14), with equal time-slot allocation. That is, $d_{\text{noma}} > d_{\text{oma}}$ holds for arbitrary given channel realizations h_1, h_2 and constellation sizes M_1, M_2.■

The proof is provided in Appendix-I. From Corollary 2, since $d_{\text{noma}} > d_{\text{oma}}$, it is expected that NOMA outperforms TDMA in terms of error performance, especially in moderate and high SNR regions as can be confirmed by numerical results.

IV. Extensions to Multi-User/Multi-Antenna Systems

A. Extension to Single-Antenna Multi-User Multiple Access Channel

In this subsection, we show how to extend our design framework for two single-antenna users described in (1) to systems with multiple single-antenna users. More specifically, we consider a Gaussian MAC, where N single-antenna users transmit to a single-antenna receiver simultaneously, given by

$$z = \sum_{k=1}^{N} h_k x_k + \xi, \tag{15}$$

where h_k denotes the complex channel coefficients between the transmitter S_k and BS, $x_k \exp(j\arg(h_k))$ is drawn uniformly from the M_k^2-ary square QAM constellation $\mathcal{Q}_k$ such that $\mathcal{Q}_k \triangleq \{\pm w_k(2k-1) \pm w'_k(2\ell-1)j : k, \ell = 1, \ldots, M_k/2\}$, which is subject to an individual average power constraint P_k, i.e., $\mathbb{E}[|x_k|^2] \leq P_k$ for $k = 1, 2, \ldots, N$. In line with (5), the in-phase component of (15) can be formulated by

$$y = \sum_{k=1}^{N} |\tilde{h}_k| \tilde{w}_k s_k + n, \tag{16}$$

where $0 < \tilde{w}_k = \sqrt{\frac{2(M_k^2-1)}{3\ P_k}} w_k \leq 1$ is the normalized weighting coefficients, $|\tilde{h}_k| = \sqrt{\frac{3\ P_k}{2(M_k^2-1)}}|h_k|$ is the normalized channel coefficients, and $s_k \in \mathcal{A}_{M_k} = \{\pm(2k-1)\}_{k=1}^{M_k/2}$ is the information bearing signal drawn from M_k-ary PAM constellation, for $k = 1, 2, \ldots, N$.

We now show how to solve the weighting coefficient design problem iteratively with the aid of Theorem 1. For $N \geq 3$, we have

$$y = \underbrace{\sum_{k=1}^{N-2} |\tilde{h}_k| \tilde{w}_k s_k}_{\mathrm{T}_1} + \underbrace{|\tilde{h}_{N-1}| \tilde{w}_{N-1} s_{N-1} + |\tilde{h}_N| \tilde{w}_N s_N}_{\mathrm{T}_2} + n.$$

Without loss of generality, we consider the maximization of the minimum Euclidean distance of $\mathrm{T}_2 = |\tilde{h}_{N-1}|\tilde{w}_{N-1}s_{N-1} + |\tilde{h}_N|\tilde{w}_N s_N$ by temporarily ignoring the term T_1. We assume that $\max\{\tilde{w}_{N-1}, \tilde{w}_N\} \leq \tilde{w}'_{N-1}$, where $0 < \tilde{w}'_{N-1} \leq 1$ is the scaling factor we introduced for T_2 relative to T_1. With the help of Theorem 1, when the minimum Euclidean distance of T_2 is maximized, we have $\mathrm{T}_2 = |\tilde{h}'_{N-1}|\tilde{w}'_{N-1}s'_{N-1}$, where s'_{N-1} is a standard $M_{N-1}M_N$-ary PAM constellation, i.e., $s'_{N-1} \in \mathcal{A}'_{M_{N-1}M_N} = \{\pm(2k-1)\}_{k=1}^{M_{N-1}M_N/2}$ such that

1) If $\frac{|\tilde{h}_N|}{|\tilde{h}_{N-1}|} \leq \frac{1}{M_N}$, then $|\tilde{h}'_{N-1}| = |\tilde{h}_N|$, $s'_{N-1} = M_N s_{N-1} + s_N$, and $(\tilde{w}^*_{N-1}, \tilde{w}^*_N) = \left(M_N \frac{|\tilde{h}_N|}{|\tilde{h}_{N-1}|}\tilde{w}'_{N-1}, \tilde{w}'_{N-1}\right)$;
2) If $\frac{1}{M_N} < \frac{|\tilde{h}_N|}{|\tilde{h}_{N-1}|} \leq \frac{M_{N-1}}{M_N}$, then $|\tilde{h}'_{N-1}| = \frac{|\tilde{h}_{N-1}|}{M_N}$, $s'_{N-1} = M_N s_{N-1} + s_N$, and $(\tilde{w}^*_{N-1}, \tilde{w}^*_N) = \left(\tilde{w}'_{N-1}, \frac{|\tilde{h}_{N-1}|}{M_N|\tilde{h}_N|}\tilde{w}'_{N-1}\right)$;
3) If $\frac{M_{N-1}}{M_N} < \frac{|\tilde{h}_N|}{|\tilde{h}_{N-1}|} \leq M_{N-1}$, then $|\tilde{h}'_{N-1}| = \frac{|\tilde{h}_N|}{M_{N-1}}$, $s'_{N-1} = s_{N-1} + M_{N-1}s_N$, and $(\tilde{w}^*_{N-1}, \tilde{w}^*_N) = \left(\frac{|\tilde{h}_N|}{M_{N-1}|\tilde{h}_{N-1}|}\tilde{w}'_{N-1}, \tilde{w}'_{N-1}\right)$;
4) If $M_{N-1} < \frac{|\tilde{h}_N|}{|\tilde{h}_{N-1}|}$, then $|\tilde{h}'_{N-1}| = |\tilde{h}_{N-1}|$, $s'_{N-1} = s_{N-1} + M_{N-1}s_N$ and $(\tilde{w}^*_{N-1}, \tilde{w}^*_N) = \left(\tilde{w}'_{N-1}, M_{N-1}\frac{|\tilde{h}_{N-1}|}{|\tilde{h}_N|}\tilde{w}'_{N-1}\right)$.

Now, (16) can be reformulated by

$$\begin{aligned} y &= \sum_{k=1}^{N-2} |\tilde{h}_k|\tilde{w}_k s_k + |\tilde{h}'_{N-1}|\tilde{w}'_{N-1}s'_{N-1} + n \\ &= \sum_{k=1}^{N-1} |\tilde{h}'_k|\tilde{w}'_k s'_k + n, \end{aligned} \tag{17}$$

where $|\tilde{h}'_k| = |\tilde{h}_k|$, $s_k = s'_k$, and $\tilde{w}'_k = \tilde{w}_k$ for $k = 1, 2, \ldots, N-2$, $0 < \tilde{w}'_k \leq 1$ for $k = 1, 2, \ldots, N-1$. In this way, the original N-user NOMA design problem is converted into a $(N-1)$-user NOMA design problem. By repeating this procedure for $(N-1)$ times, the weighting coefficients can be obtained for the N-user NOMA system. Note that, the ordering of the users can affect the performance, and there are totally $N!/2$ possible ordering methods and the optimal ordering of users can be obtained by an exhaustive search.

B. Extension to Multi-Antenna Multi-User Multiple Access Channel

Now, we extend our design to networks with multiple antennas. In line with [42] and [43], we consider a Gaussian MAC consisting of N users who transmit simultaneously to BS equipped with K antennas. We assume that the k-th user is equipped with L_k antennas for $k = 1, 2, \ldots, N$. Then, the received signal at BS is given by

$$\mathbf{z} = \sum_{k=1}^{N} \mathbf{H}_k \mathbf{p}_k x_k + \boldsymbol{\xi}$$

where x_k is the information bearing symbol, $\mathbf{H}_k \in \mathbb{C}^{K \times L_k}$ is the channel coefficient between user S_k and BS, $\mathbf{p}_k$ is the precoding vector of S_k for $k = 1, 2, \ldots, N$. Also, $\boldsymbol{\xi} \sim \mathcal{CN}(\mathbf{0}_{K\times K}, 2\sigma^2 \mathbf{I}_{K\times K})$ is the additive noise vector arising at the receiver side.

Then, we assume that the receiver uses a vector $\mathbf{g}$ to combine the received signal and we have

$$\hat{z} = \mathbf{g}^H \mathbf{z} = \sum_{k=1}^{N} \mathbf{g}^H \mathbf{H}_k \mathbf{p}_k x_k + \mathbf{g}^H \boldsymbol{\xi} = \sum_{k=1}^{N} \hat{h}_k x_k + \hat{\xi},$$

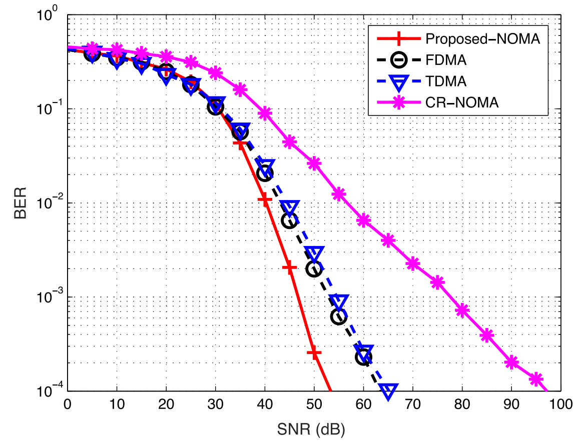
where $\hat{h}_k = \mathbf{g}^H \mathbf{H}_k \mathbf{p}_k$ are the equivalent complex scalar channels between S_k and BS, $\hat{\xi} = \mathbf{g}^H \boldsymbol{\xi}$ is the equivalent additive noise. Now, the multiple-antenna multi-user MAC model is converted into a single-antenna multi-user MAC system as described in (15), which can be readily solved when $N = 2$ and can also be iteratively solved when $N \geq 3$ as discussed in Sec. IV-A .[3] Note that the design of the precoding vector $\mathbf{p}_k$ and receiver vector $\mathbf{g}$ has already been considered in [42] and [43] for various design criteria and the design is beyond the scope of this paper.

V. Simulation Results and Discussions

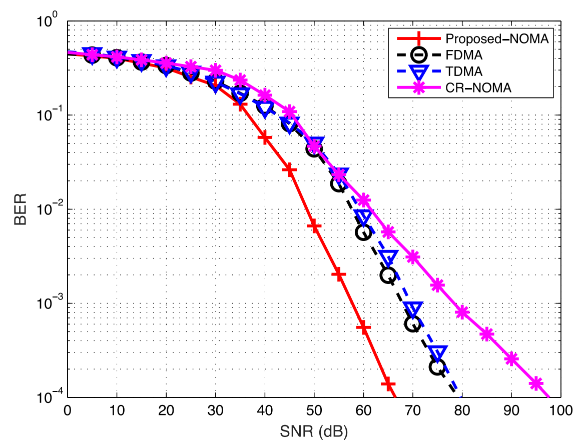
In this section, we carry out computer simulations to verify the effectiveness of our NOMA design relative to the CR-NOMA design proposed in [27] and the OMA methods including TDMA and frequency-division multiple access (FDMA) schemes in various channel conditions and system configurations. Without loss of generality, we assume that $P_1 = P_2 = 1$ and the system signal-to-noise ratio (SNR) is defined by $\rho \triangleq 1/2\sigma^2$. All channels are subject to Rayleigh fading such that $h_k \sim \mathcal{CN}(0, 2\delta_k^2)$, $k = 1, 2$.

For the sake of comparison, we assume that both users transmit alternatively by using half of the total time slots or half of the available frequency band in TDMA and FDMA, respectively. In both methods, to maintain the same data rate for each user, we should increase the constellation sizes by using M_1^4- and M_2^4-ary QAM constellations instead. There is no interference occurring at the receiver side since the channels are orthogonal to each other. It is worth mentioning that, for both TDMA and FDMA, we assume that the instantaneous transmit power of each user remains the same as in the NOMA. For FDMA method, as the total bandwidth of each user is halved, the equivalent noise at the receiver side also reduces by half compared with the TDMA method. Therefore, we would expect the FDMA scheme has an around 3dB SNR gain compared with TDMA method. For the CR-NOMA, we let each user transmit at the maximum allowable power by using constellations $\left\{\exp(\frac{j2\pi k}{N})\right\}_{k=0}^{N-1}$ and $\left\{\exp(\frac{j2\pi k + j\pi}{N})\right\}_{k=0}^{N-1}$ as proposed in [27] for users S_1 and S_2, respectively. Note that comparing the proposed design using QAM and CR-NOMA with PSK seems to be unfair. However, to the best of our knowledge, there is no other benchmark available in open literature and the CR-NOMA with PSK is the best benchmark NOMA scheme that we can find, where the uniquely decoding property can be guaranteed. Furthermore, the design principle of CR-NOMA with PSK constellations provided in [27] cannot be directly extended to the QAM constellations.

We first compare the average BER of all the schemes where the variances of the channel coefficients are the same,



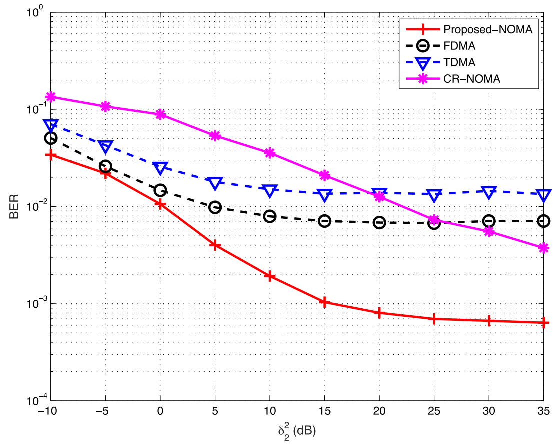
(a)



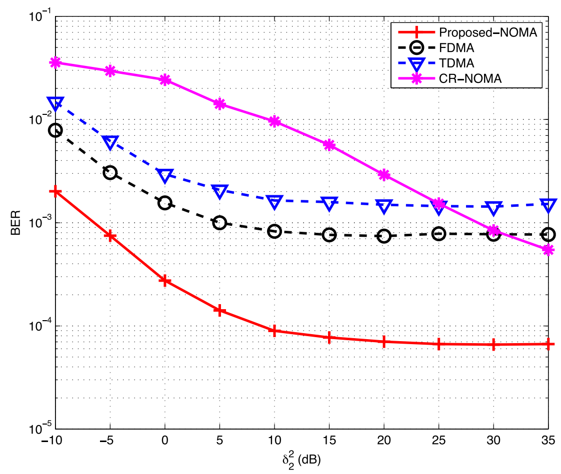
(b)

Fig. 1. Comparison between the Proposed-NOMA, CR-NOMA, TDMA and FDMA methods where 64-QAM is used for our case and 64-PSK is used for CR-based method: (a) $(\delta_1^2, \delta_2^2) = (1, 1)$, (b) $(\delta_1^2, \delta_2^2) = (1, 1/64)$.

[3] It is worth pointing out that such an ideal precoding cannot be obtained in some cases. In those cases, the precoders of the users and combining vector at the receiver should be jointly designed, which is a challenging problem since the geometric structure of the received sum-constellation is hard to characterize in a high-dimensional space. We believe that addressing this challenging problem will require the development of a totally new design framework, which can constitute another full paper and is thus left as our future work.

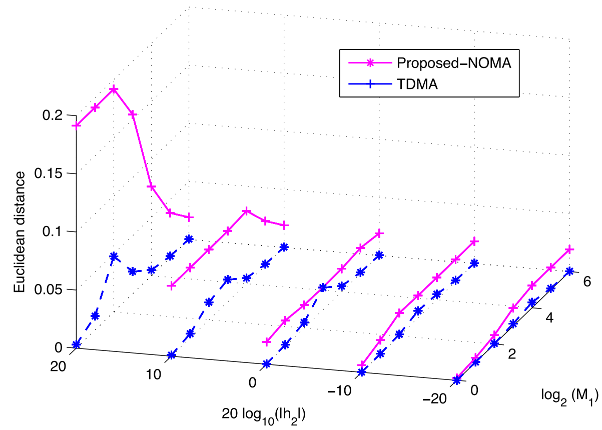


(a)

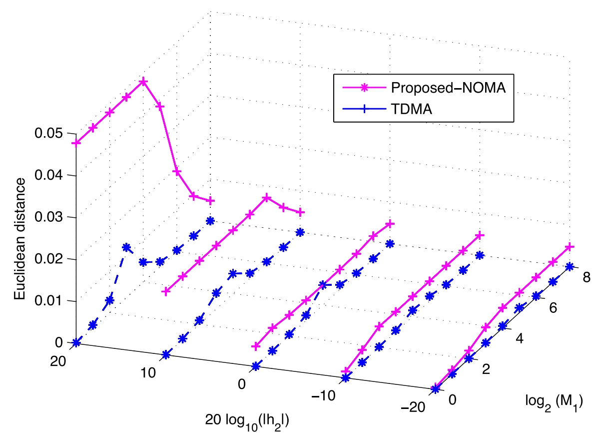


(b)

Fig. 2. Comparison between the Proposed-NOMA with CR-NOMA, TDMA, and FDMA methods, 64-QAM are used for our case and 64-PSK are used for CR based method with (a) $\rho = 40$dB. (b) $\rho = 50$dB.



(a)



(b)

Fig. 3. Comparison of the minimum Euclidean distance of the proposed-NOMA (i.e., d_{noma} given in (13)) and that of TDMA method (d_{oma} given in (14)) with $|h_1| = 1$ for different $|h_2|$ and M_1 (a) $M = 64$, (b) $M = 256$.

i.e., $(\delta_1^2, \delta_2^2) = (1, 1)$ in Fig. 1(a). In the simulation, without loss of generality, we assume that each user adopts 64-QAM for the proposed NOMA design and 64-PSK is used by each user in CR-NOMA. Meanwhile, for TDMA and FDMA methods, each user uses 4096-QAM. As can be observed from Fig. 1(a) that, the proposed NOMA design outperforms all the designs in moderate and high SNR regimes. In addition, the FDMA method has a better error performance than the TDMA scheme as expected. The CR-NOMA has the highest BER due to the fact that the PSK constellation has a smaller Euclidean distance under the same power constraint compared with QAM constellation.

In the following simulation, we take the near-far effect into consideration by letting $(\delta_1^2, \delta_2^2) = (1, 1/64)$ as shown in Fig. 1(b). Likewise, the proposed NOMA design has the lowest BER compared with all the benchmark schemes. Also, we can observe that the gap between the proposed NOMA and the FDMA as well as TDMA is larger than that in the case of equal channel gain. For example, at the BER 10^{-3}, the proposed NOMA has around 5dB SNR gain in Fig. 1(a), while the SNR gain is approximately 10dB in Fig. 1(b). Interestingly, we also observe that the error performance of CR-NOMA improves substantially compared to TDMA and FDMA in this case with near-far effect.

From both Figs. 1(a) and 1(b), we can observe that the performance gain of NOMA is highly related to the relative strength of the channel coefficients. To show this phenomenon clearly, we now study the BER against the relative strength of the channel coefficients under different SNRs. More specifically, in Fig.2(a), we set the variance of user S_1 as $\delta_1^2 = 1$, and we plot the BER against the variance of user S_2, i.e., δ_2^2, in dB. It can be observed from Fig. 2(a) that, for $\rho = 40$dB (i.e., the SNR is relatively low relative to the target transmission rate), our proposed NOMA scheme outperforms all the benchmark schemes. When δ_2^2 is less than 1 (i.e., less than 0dB), the error performance is mainly limited by user S_1 and even if δ_2^2 equals to 1, the BER gain of the proposed NOMA method is still marginal. However, with the increase of δ_2^2, the BER gain of the proposed NOMA method increases and finally gets saturated. Actually, when δ_2^2 is extremely large, the BER of the proposed NOMA is close to the system with one user transmitting with 64-QAM in both orthogonal blocks, while for the OMA method, it saturates as one user transmits using 4096-QAM in one block. This validates our observation that the proposed NOMA has a higher SNR gain when there is near-far effect. With the increase of δ_2^2, the performance of CR-NOMA improves dramatically and it eventually outperforms the OMA methods. However, the BER performance is poor when the channel gains of the two users are close. This is due to the fact that with the same spectral efficiency, a PSK constellation has a smaller minimum Euclidean distance than a QAM constellation. Moreover, the sum-constellation of two PSK constellations at the receiver does not have a good geometric structure. In Fig. 2(b), we can see that with the near-far effect, the BER gain of the proposed NOMA also become more significant. The BER gain of the proposed NOMA is evident even if $\delta_2^2 = 1$, which coincides well with the phenomenon observed in Fig. 1.

At last, we compare the minimum Euclidean distance of the proposed NOMA design with that of TDMA method in Fig. 3. It can be observed that our proposed NOMA design achieves larger minimum Euclidean distance than TDMA method in all simulated cases, which validates the result presented in Corollary 2. We can also see that the stronger the near-far effect is, the larger the performance gap between the proposed NOMA and TDMA. More importantly, for the proposed NOMA, when we enlarge M_1, there is a large interval in which the minimum Euclidean distance of NOMA (i.e., d_{noma}) will remain almost unchanged, while that of TDMA only has one peak among the considered range of M_1. This indicates that the proposed NOMA has a larger degree of freedom in adaptive rate allocation than that of TDMA under the condition of causing nearly no degradation of system error performance.

VI. Conclusions

In this paper, we have presented a practical design framework for the non-orthogonal multiple access (NOMA) scheme in a classical two-user multiple access channel (MAC) with quadrature amplitude modulation (QAM) constellations at both users, the sizes of which are not necessarily the same. More specifically, we aimed to maximize the minimum Euclidean distance of the sum-constellation at the receiver by adjusting the instantaneous transmit power and phase of each user under an individual average power constraint. The design objective was formulated into a *mixed continuous-discrete* optimization problem. By introducing a new mathematical concept termed *punched Farey sequence* and investigating its fundamental properties, we managed to attain a compact closed-form solution. Computer simulations were conducted to verify our derivation under various channel configurations, and the simulation results demonstrated that our proposed NOMA scheme outperforms OMA and existing NOMA significantly and the performance gap can be further enlarged when there is a near-far effect between the users.

Appendix

A. Proof of Property 1

1) We first prove that $\frac{n_1}{m_1} < \frac{n_1+n_2}{m_1+m_2}$, which can be showed by calculating $\frac{n_1+n_2}{m_1+m_2} - \frac{n_1}{m_1} = \frac{m_1\ n_2 - m_2\ n_1}{m_1(m_1+m_2)} > 0$, since $\frac{n_1}{m_1} < \frac{n_2}{m_2}$. The rest cases can be proved in a similar fashion and hence are omitted.

2) We now prove that $m_1\ n_2 - m_2\ n_1 = 1$ and it also gives the construction of the term which succeeds $\frac{n_1}{m_1}$ in $\mathfrak{P}_K^L$. First of all, since $\langle m_1, n_1\rangle = 1$, the following equation

$$m_1 n - m n_1 = 1 \tag{18}$$

has integer solutions in m, n such that $m = m_0 + r m_1, n = n_0 + r n_1$ for any integer r, where m_0, n_0 is a particular set of solutions to (18) and $\langle m, n\rangle = 1$ [35, Th. 25]. As $\frac{n_1}{m_1} \in \mathfrak{P}_K^L$, we have $0 \le n_1 \le L$ and $0 \le m_1 \le K$. Then, we can choose m, n satisfying either condition:

$$\textit{Case 1: } K - m_1 < m \le K, \quad \text{and } 0 < n \le L; \tag{19a}$$

$$\textit{Case 2: } 0 < m \le K - m_1, \quad \text{and } L - n_1 < n \le L. \tag{19b}$$

Now, since $\frac{n}{m}$ is in its lowest terms (i.e., $\langle m, n\rangle = 1$), and for either case we have $0 < m \le K$, $0 < n \le L$, we conclude that $\frac{n}{m}$ is a fraction of $\mathfrak{P}_K^L$. In what follows, we will show that either Case 1 or 2 will generate the next term which comes after $\frac{n_1}{m_1}$ in $\mathfrak{P}_K^L$.

Case 1: From (18), $\frac{n}{m} = \frac{n_1}{m_1} + \frac{1}{m_1\ m} > \frac{n_1}{m_1}$, hence $\frac{n}{m}$ comes after $\frac{n_1}{m_1}$ in $\mathfrak{P}_K^L$. Then, if $\frac{n}{m}$ is not $\frac{n_2}{m_2}$, it will come after $\frac{n_2}{m_2}$, and then

$$\begin{aligned}\frac{n_2}{m_2} - \frac{n_1}{m_1} &= \frac{m_1 n_2 - m_2 n_1}{m_1 m_2} \ge \frac{1}{m_1 m_2};\\ \frac{n}{m} - \frac{n_2}{m_2} &= \frac{m_2 n - m n_2}{m_2 m} \ge \frac{1}{m_2 m}.\end{aligned} \tag{20}$$

As a result, by jointly considering (18) and (20), we have $\frac{1}{m_1\ m} \overset{(a)}{=} \frac{m_1\ n - m n_1}{m_1\ m} = \frac{n}{m} - \frac{n_1}{m_1} = \frac{n}{m} - \frac{n_2}{m_2} + \frac{n_2}{m_2} - \frac{n_1}{m_1} \overset{(b)}{\ge} \frac{1}{m_2\ m} + \frac{1}{m_1\ m_2} = \frac{m_1 + m}{m_1\ m_2\ m} \overset{(c)}{>} \frac{K}{m_1\ m_2\ m} \overset{(d)}{\ge} \frac{1}{m_1\ m}$, where (a) follows from (18); inequality (b) holds since (20); inequality (c) follows from (19a) and (d) is true since $\frac{n_2}{m_2} \in \mathfrak{P}_K^L$. This is a contradiction, and therefore $\frac{n}{m}$ must be $\frac{n_2}{m_2}$, and hence $m_1\ n_2 - m_2\ n_1 = 1$.

Case 2: As in Case 1, according to (18), $\frac{n}{m} = \frac{n_1}{m_1} + \frac{1}{m_1\ m} > \frac{n_1}{m_1}$, and hence $\frac{n}{m}$ comes after $\frac{n_1}{m_1}$ in $\mathfrak{P}_K^L$. As a result, if $\frac{n}{m}$ is not $\frac{n_2}{m_2}$, it comes after $\frac{n_2}{m_2}$, and we have

$$\begin{aligned}\frac{m_1}{n_1} - \frac{m_2}{n_2} &= \frac{m_1 n_2 - m_2 n_1}{n_1 n_2} \ge \frac{1}{n_1 n_2};\\ \frac{m_2}{n_2} - \frac{m}{n} &= \frac{m_2 n - m n_2}{n_2 n} \ge \frac{1}{n_2 n}.\end{aligned} \tag{21}$$

Likewise, we conclude that $\frac{1}{n_1\ n} \overset{(a)}{=} \frac{m_1\ n - m n_1}{n_1\ n} = \frac{m_1}{n_1} - \frac{m}{n} = \frac{m_1}{n_1} - \frac{m_2}{n_2} + \frac{m_2}{n_2} - \frac{m}{n} \overset{(b)}{\ge} \frac{1}{n_1\ n_2} + \frac{1}{n_2\ n} = \frac{n + n_1}{n_1\ n_2\ n} \overset{(c)}{>} \frac{L}{n_1\ n_2\ n} \overset{(d)}{\ge} \frac{1}{n_1\ n}$, where (a) follows from (18), inequality (b) holds as a result of (21), and inequality (c) follows from (19b), and (d) is true since $\frac{n_2}{m_2} \in \mathfrak{P}_K^L$. There is also a contradiction, therefore $\frac{n}{m}$ must be $\frac{n_2}{m_2}$, and thus $m_1\ n_2 - m_2\ n_1 = 1$. 3) From the above proof, i.e., (19a) and (19b), we can observe that: (a) If $n_1 + n_2 \le L$, then $m_1 + m_2 > K$; (b) If $m_1 + m_2 \le K$, then $n_1 + n_2 > L$.

4) From (19a) and (19b), we can observe that, $m_1 + m_2 > 0$ (i.e., $m_1 + m_2 \ge 1$) and $n_1 + n_2 > 0$ (i.e., $n_1 + n_2 \ge 1$).

First, we consider the case $\frac{n_1}{m_1} = \frac{0}{1}$. As $n_1 = 0$, then by solving (18) of Case 1 in the above discussion, we have $n_2 = 1$. By (19a), we attain $K-1 < m_2 \le K$, i.e., $m_2 = K$. Now, we attain two adjacent terms $\frac{n_1}{m_1} = \frac{0}{1}$ and $\frac{n_2}{m_2} = \frac{1}{K}$ such that $n_1 + n_2 = 1$ and $m_1 + m_2 = K + 1$. Similarly, we can find adjacent terms $\frac{n_1}{m_1} = \frac{L}{1}$ an $\frac{n_2}{m_2} = \frac{1}{0}$ such that $m_1 + m_2 = 1$ and $n_1 + n_2 = L + 1$.

This completes the proof of Property 1. □

B. Proof of Property 2

By Property 1, we have $m_1\ n_2 - m_2\ n_1 = 1$, $m_2\ n_3 - m_3\ n_2 = 1$. Then, solving the following equations: $m_3\ m_1\ n_2 - m_3\ m_2\ n_1 = m_3$, $m_1\ m_2\ n_3 - m_1\ m_3\ n_2 = m_1$, $n_3\ m_1\ n_2 - n_3\ m_2\ n_1 = n_3$, $n_1\ m_2\ n_3 - n_1\ m_3\ n_2 = n_1$, for m_2, n_2, we attain $m_2(m_1\ n_3 - m_3\ n_1) = m_1 + m_3$, $n_2(m_1\ n_3 - m_3\ n_1) = n_1 + n_3$. As $m_1\ n_3 - m_3\ n_1 \neq 0$, we have $\frac{n_2}{m_2} = \frac{n_1+n_3}{m_1+m_3}$. The property is proved. □

C. Proof of Property 3

From the assumption, we have $\frac{n_2}{m_2} - \frac{n_1+n_3}{m_1+m_3} = \frac{m_1\ n_2+m_3\ n_2-m_2\ n_1-m_2\ n_3}{m_2(m_1+m_3)} = \frac{m_1\ n_2-m_2\ n_1-1}{m_2(m_1+m_2)} \ge 0$; $\frac{n_2+n_4}{m_2+m_4} - \frac{n_3}{m_3} = \frac{m_3\ n_2+m_3\ n_4-m_2\ n_3-m_4\ n_3}{m_3(m_2+m_4)} = \frac{m_3\ n_4-m_4\ n_3-1}{m_3(m_2+m_4)} \ge 0$. The completes the proof. □

D. Proof of Proposition 2

Recall that $d(m,n) = \big||\tilde{h}_1|\tilde{w}_1\ n - |\tilde{h}_2|\tilde{w}_2\ m\big|$. Therefore, for $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_1}{m_1}, \frac{n_2}{m_2}\right)$, we have $d(m_1,n_1) - d(m_2,n_2) = \big||\tilde{h}_1|\tilde{w}_1\ n_1 - |\tilde{h}_2|\tilde{w}_2\ m_1\big| - \big||\tilde{h}_1|\tilde{w}_1\ n_2 - |\tilde{h}_2|\tilde{w}_2\ m_2\big| = -|\tilde{h}_1|\tilde{w}_1\ n_1 + |\tilde{h}_2|\tilde{w}_2\ m_1 - |\tilde{h}_1|\tilde{w}_1\ n_2 + |\tilde{h}_2|\tilde{w}_2\ m_2 = (m_1 + m_2)|\tilde{h}_1|\tilde{w}_1\left(\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} - \frac{n_1+n_2}{m_1+m_2}\right)$. The results in Proposition 2 can be readily obtained, and we complete the proof. □

E. Proof of of Proposition 3

Proof: As $\frac{n_1}{m_1}$ and $\frac{n_4}{m_4}$ are arbitrarily chosen, Proposition 3 is equivalent to

1) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}\right)$, then $d(m_2,n_2) < d(m_3,n_3)$, $d(m_2,n_2) < d(m_4,n_4)$, and $d(m_2,n_2) < d(m_1,n_1)$;
2) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2+n_3}{m_2+m_3}, \frac{n_3}{m_3}\right)$, then $d(m_3,n_3) < d(m_2,n_2)$, $d(m_3,n_3) < d(m_1,n_1)$, and $d(m_3,n_3) < d(m_4,n_4)$.

First, by Proposition 2, we have

1) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}\right)$, then $d(m_2,n_2) < d(m_3,n_3)$ and $d(m_2,n_2) < d(m_4,n_4)$;
2) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2+n_3}{m_2+m_3}, \frac{n_3}{m_3}\right)$, then $d(m_3,n_3) < d(m_1,n_1)$ and $d(m_3,n_3) < d(m_2,n_2)$.

Then, we want to show that:

1) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}\right)$, then $d(m_2,n_2) < d(m_1,n_1)$;
2) If $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2+n_3}{m_2+m_3}, \frac{n_3}{m_3}\right)$, then $d(m_3,n_3) < d(m_4,n_4)$.

The first case can be proved by considering $d(m_1,n_1) - d(m_2,n_2) = \big||\tilde{h}_1|\tilde{w}_1\ n_1 - |\tilde{h}_2|\tilde{w}_2\ m_1\big| - \big||\tilde{h}_1|\tilde{w}_1\ n_2 - |\tilde{h}_2|\tilde{w}_2\ m_2\big| = |\tilde{h}_1|\tilde{w}_1\left(\frac{|\tilde{h}_2|w_2}{|\tilde{h}_1|w_1}(m_1 - m_2) - (n_1 - n_2)\right)$. As $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2}{m_2}, \frac{n_2+n_3}{m_2+m_3}\right)$, then $d(m_1,n_1) - d(m_2,n_2) \ge 0$ is true if $\frac{n_2}{m_2}(m_1 - m_2) - (n_1 - n_2) \ge 0$ and $\frac{n_2+n_3}{m_2+m_3}(m_1 - m_2) - (n_1 - n_2) \ge 0$. We know that $\frac{n_2}{m_2}(m_1 - m_2) - (n_1 - n_2) = \left(m_1\ n_2 - m_2 n_1\right)/m_2 > 0$, and $\frac{n_2+n_3}{m_2+m_3}(m_1 - m_2) - (n_1 - n_2) = \big((n_2 + n_3)(m_1 - m_2) - (m_2 + m_3)(n_1 - n_2)\big)/(m_2 + m_3) = \big((m_1 + m_3)n_2 - m_2(n_1 + n_3) + m_1\ n_3 - m_3\ n_1\big)/(m_2 + m_3) > 0$, where the inequality is always true by Property 3. Likewise, the second case can be proved by considering $d(m_4,n_4) - d(m_3,n_3) = \big||\tilde{h}_1|\tilde{w}_1\ n_4 - |\tilde{h}_2|\tilde{w}_2\ m_4\big| - \big||\tilde{h}_1|\tilde{w}_1\ n_3 - |\tilde{h}_2|\tilde{w}_2\ m_3\big| = |\tilde{h}_1|\tilde{w}_1\left(\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1}(m_3 - m_4) - (n_3 - n_4)\right)$. As $\frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \in \left(\frac{n_2+n_3}{m_2+m_3}, \frac{n_3}{m_3}\right)$, then $d(m_4,n_4) - d(m_3,n_3) \ge 0$ is true if $\frac{n_3}{m_3}(m_3 - m_4) - (n_3 - n_4) \ge 0$ and $\frac{n_2+n_3}{m_2+m_3}(m_3 - m_4) - (n_3 - n_4) \ge 0$. We know that $\frac{n_3}{m_3}(m_3 - m_4) - (n_3 - n_4) = \big(n_3(m_3 - m_4) - m_3(n_3 - n_4)\big)/m_3 = (m_3\ n_4 - m_4\ n_3)/m_3 > 0$, and $\frac{n_2+n_3}{m_2+m_3}(m_3 - m_4) - (n_3 - n_4) = \big((n_2 + n_3)(m_3 - m_4) - (m_2 + m_3)(n_3 - n_4)\big)/(m_2 + m_3) = \big(m_3(n_2 + n_4) - n_3(m_2 + m_4) + m_2\ n_4 - m_4\ n_2\big)/(m_2 + m_3) > 0$, where the inequality is always true by Property 3. We complete the proof. □

F. Proof of Lemma 1

According to proposition 3 and notice that $\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \left(\frac{b_k}{a_k}, \frac{b_k+b_{k+1}}{a_k+a_{k+1}}\right) \cup \left(\frac{b_k+b_{k+1}}{a_k+a_{k+1}}, \frac{b_{k+1}}{a_{k+1}}\right)$, 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 3 (Sub-Problem 1): The optimization problem is stated as follows:

$$\begin{aligned} g_1\Big(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\Big) = \max_{(\tilde{w}_1,\tilde{w}_2)}\ & |\tilde{h}_2|\tilde{w}_2 a_k - |\tilde{h}_1|\tilde{w}_1 b_k \\ \text{s.t.}\ & \frac{b_k}{a_k} \le \frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} < \frac{b_k + b_{k+1}}{a_k + a_{k+1}}, \\ & 0 < \tilde{w}_1 \le 1,\ 0 < \tilde{w}_2 \le 1. \end{aligned} \tag{22}$$

■

Problem 4 (Sub-Problem 2): We aim to solve the following optimization problem:

$$\begin{aligned} g_2\Big(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\Big) = \max_{(\tilde{w}_1,\tilde{w}_2)}\ & |\tilde{h}_1|\tilde{w}_1 b_{k+1} - |\tilde{h}_2|\tilde{w}_2 a_{k+1} \\ \text{s.t.}\ & \frac{b_k + b_{k+1}}{a_k + a_{k+1}} \le \frac{|\tilde{h}_2|\tilde{w}_2}{|\tilde{h}_1|\tilde{w}_1} \le \frac{b_{k+1}}{a_{k+1}}, \\ & 0 < \tilde{w}_1 \le 1, 0 < \tilde{w}_2 \le 1. \end{aligned} \tag{23}$$

■

We know that the constraint of (22) is equivalent to $\frac{|\tilde{h}_2|(a_k+a_{k+1})}{|\tilde{h}_1|(b_k+b_{k+1})}\tilde{w}_2 < \tilde{w}_1 \le \frac{a_k|\tilde{h}_2|}{b_k|\tilde{h}_1|}\tilde{w}_2$, $0 < \tilde{w}_1 \le 1$, $0 < \tilde{w}_2 \le 1$. We can find that the objective function is a linear decreasing function of $\tilde{w}_1$. Then, we let $\tilde{w}_1 = \frac{|\tilde{h}_2|(a_k+a_{k+1})}{|\tilde{h}_1|(b_k+b_{k+1})}\tilde{w}_2$, and the objective function can be reformulated by $|\tilde{h}_2|\tilde{w}_2\ a_k - |\tilde{h}_1|\tilde{w}_1\ b_k = \big(a_k(b_k + b_{k+1}) - b_k(a_k + a_{k+1})\big)\frac{|\tilde{h}_2|\tilde{w}_2}{b_k+b_{k+1}} \overset{(a)}{=} \frac{|\tilde{h}_2|\tilde{w}_2}{b_k+b_{k+1}}$, where (a) follows from Property 1. Now, the constraints on $\tilde{w}_2$ are $0 < \tilde{w}_2 \le 1$, $0 < \tilde{w}_2 \le \frac{|\tilde{h}_1|(b_k+b_{k+1})}{|\tilde{h}_2|(a_k+a_{k+1})}$.

Therefore, the solution to (22) can be given as follows:

$$g_1\Big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\Big) = \begin{cases} \dfrac{|\tilde{h}_2|}{b_k+b_{k+1}}, \quad \text{with } (\tilde{w}_1,\tilde{w}_2)=(\dfrac{|\tilde{h}_2|(a_k+a_{k+1})}{|\tilde{h}_1|(b_k+b_{k+1})},1), \\ \quad \text{if } \dfrac{|\tilde{h}_2|}{|\tilde{h}_1|}\le \dfrac{b_k+b_{k+1}}{a_k+a_{k+1}}; \\ \dfrac{|\tilde{h}_1|}{a_k+a_{k+1}}, \quad \text{with } (\tilde{w}_1,\tilde{w}_2)=(1,\dfrac{|\tilde{h}_1|(b_k+b_{k+1})}{|\tilde{h}_2|(a_k+a_{k+1})}), \\ \quad \text{if } \dfrac{|\tilde{h}_2|}{|\tilde{h}_1|}> \dfrac{b_k+b_{k+1}}{a_k+a_{k+1}}. \end{cases} \tag{24}$$

Likewise, we note that the constraint of (23) is equivalent to $\frac{(b_k+b_{k+1})|\tilde{h}_1|}{(a_k+a_{k+1})|\tilde{h}_2|}\tilde{w}_1 \le \tilde{w}_2 \le \frac{b_{k+1}|\tilde{h}_1|}{a_{k+1}|\tilde{h}_2|}\tilde{w}_1,\ 0<\tilde{w}_1\le 1, 0<\tilde{w}_2\le 1$. By letting $\tilde{w}_2 = \frac{(b_k+b_{k+1})|\tilde{h}_1|}{(a_k+a_{k+1})|\tilde{h}_2|}\tilde{w}_1$, the objective function can be reformulated by $|\tilde{h}_1|\tilde{w}_1\ b_{k+1} - |\tilde{h}_2|\tilde{w}_2\ a_{k+1} = \big(b_{k+1}(a_k+a_{k+1}) - a_{k+1}(b_k+b_{k+1})\big)\frac{|\tilde{h}_1|\tilde{w}_1}{a_k+a_{k+1}} = \frac{|\tilde{h}_1|\tilde{w}_1}{a_k+a_{k+1}}$. The constraints on $\tilde{w}_1$ are $0<\tilde{w}_1\le 1$ and $0<\tilde{w}_1<\frac{(a_k+a_{k+1})|h_2|}{(b_k+b_{k+1})|\tilde{h}_1|}$. Thus, we have

$$g_2\Big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\Big) = \begin{cases} \dfrac{|\tilde{h}_2|}{b_k+b_{k+1}}, \quad \text{with } (\tilde{w}_1,\tilde{w}_2)=(\dfrac{|\tilde{h}_2|(a_k+a_{k+1})}{|\tilde{h}_1|(b_k+b_{k+1})},1), \\ \quad \text{if } \dfrac{|\tilde{h}_2|}{|\tilde{h}_1|}\le \dfrac{b_k+b_{k+1}}{a_k+a_{k+1}}; \\ \dfrac{|\tilde{h}_1|}{a_k+a_{k+1}} \quad \text{with } (\tilde{w}_1,\tilde{w}_2)=(1,\dfrac{|\tilde{h}_1|(b_k+b_{k+1})}{|\tilde{h}_2|(a_k+a_{k+1})}), \\ \quad \text{if } \dfrac{|\tilde{h}_2|}{|\tilde{h}_1|}> \dfrac{b_k+b_{k+1}}{a_k+a_{k+1}}. \end{cases} \tag{25}$$

Combining the two cases, we have the result in Lemma 1, and we complete the proof. □

G. Proof of Theorem 1

First of all, the feasible region of Problem 1 is $\mathcal{U} = \big\{(\tilde{w}_1,\tilde{w}_2) : 0<\tilde{w}_1\le 1, 0<\tilde{w}_2\le 1\big\}$ while the feasible region of problem in (11) is $\mathcal{A}_k$ such that $\mathcal{U} = \bigcup_{k=1}^{C-1}\mathcal{A}_k$. We note that $d^* = \max_{(\tilde{w}_1,\tilde{w}_2)\in\mathcal{U}} \min_{(m,n)\in\mathbb{Z}^2_{(M_1-1,M_2-1)}\setminus\{(0,0)\}} d(m,n)$ and $g\big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\big) = \max_{(\tilde{w}_1,\tilde{w}_2)\in\mathcal{A}_k} \min_{(m,n)\in\mathbb{Z}^2_{(M_1-1,M_2-1)}\setminus\{(0,0)\}} d(m,n)$ for $k=1,2,\ldots,C-1$, then we have $d^* = \max_{k=1,2,\ldots,C-1} g\big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\big)$. Now, we consider each case separately as follows:

1) If $\frac{|\tilde{h}_2|}{|\tilde{h}_1|}\le\frac{1}{M_2}$, we have $\frac{|\tilde{h}_2|}{|\tilde{h}_1|}\le\frac{1}{M_2}=\frac{b_1+b_2}{a_1+a_2}\le\frac{b_k+b_{k+1}}{a_k+a_{k+1}}, k=1,\ldots,C-1$. By Lemma 1, for each Farey interval, we can attain that $g\big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\big) = \frac{|\tilde{h}_2|}{b_k+b_{k+1}}$, for $k=1,\ldots,C-1$. As a consequence, the minimum Euclidean distance d^* can be attained by taking the maximum value of the objective function over all the possible intervals, given by: $d^* = \max_{k=1,2,\ldots,C-1} g\big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\big) = \max\Big\{\frac{|\tilde{h}_2|}{b_1+b_2},\ldots,\frac{|\tilde{h}_2|}{b_{C-1}+b_C}\Big\} = \frac{|\tilde{h}_2|}{b_1+b_2} = |\tilde{h}_2|$, where the optimality is attained when $(\tilde{w}_1^*,\tilde{w}_2^*) = (M_2\frac{|\tilde{h}_2|}{|\tilde{h}_1|},1)$ with the help of Property 1, and hence $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*}=M_2$.

2) If $\frac{1}{M_2}<\frac{|\tilde{h}_2|}{|\tilde{h}_1|}\le\frac{M_1}{M_2}$, we can suppose that $\frac{b_{\ell_1}+b_{\ell_1+1}}{a_{\ell_1}+a_{\ell_1+1}} < \frac{|\tilde{h}_2|}{|\tilde{h}_1|} \le \frac{b_{\ell_1+1}+b_{\ell_1+2}}{a_{\ell_1+1}+a_{\ell_1+2}}$, where ℓ_1 can be determined upon the knowledge of $\frac{|\tilde{h}_2|}{|\tilde{h}_1|}$. Then, with the help of Lemma 1, we have

$$g\Big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\Big) = \begin{cases} \dfrac{|\tilde{h}_1|}{a_k+a_{k+1}}, & k=1,\ldots,\ell_1; \\ \dfrac{|\tilde{h}_2|}{b_k+b_{k+1}}, & k=\ell_1+1,\ldots,C-1. \end{cases}$$

First, for $a_k+a_{k+1}, k=1,\ldots,\ell_1$, we have the following two cases: (a) If $a_k+a_{k+1}\ge M_2$, then we have $\frac{1}{a_k+a_{k+1}}\le\frac{1}{M_2}$; (b) If $a_k+a_{k+1}<M_2$ (i.e., $a_k+a_{k+1}\le M_2-1$), then by Property 1, we have $b_k+b_{k+1}\ge M_1$ (i.e., $b_k+b_{k+1}>M_1-1$). From the assumption, we have $\frac{b_k+b_{k+1}}{a_k+a_{k+1}} \le \frac{b_{\ell_1}+b_{\ell_1+1}}{a_{\ell_1}+a_{\ell_1+1}} < \frac{|\tilde{h}_2|}{|\tilde{h}_1|} \le \frac{M_1}{M_2}$. Therefore, we have $\frac{1}{a_k+a_{k+1}} \le \frac{M_1}{M_2(b_k+b_{k+1})} \le \frac{1}{M_2}$. Combining the above two cases, we have

$$\frac{1}{a_k+a_{k+1}} \le \frac{1}{M_2}, \quad k=1,\ldots,\ell_1. \tag{26}$$

Next, consider $b_k+b_{k+1}, k=\ell_1+1,\ldots,C-1$ and we can show that: (a) If $b_k+b_{k+1}<M_1$ (i.e., $b_k+b_{k+1}\le M_1-1$), then by Property 1, we have $a_k+a_{k+1}\ge M_2$ (i.e., $a_k+a_{k+1}>M_2-1$). As a consequence, we have $\frac{|\tilde{h}_2|}{|\tilde{h}_1|} \le \frac{b_{\ell_1+1}+b_{\ell_1+2}}{a_{\ell_1+1}+a_{\ell_1+2}} \le \frac{b_k+b_{k+1}}{a_k+a_{k+1}} \le \frac{b_k+b_{k+1}}{M_2}$; (b) If $b_k+b_{k+1}\ge M_1$, then we have $\frac{|\tilde{h}_2|}{|\tilde{h}_1|}\le\frac{M_1}{M_2}\le\frac{b_k+b_{k+1}}{M_2}$. Combining both cases, we have

$$\frac{|\tilde{h}_1|}{M_2} \ge \frac{|\tilde{h}_2|}{b_k+b_{k+1}}, \quad k=\ell_1+1,\ldots,C-1. \tag{27}$$

Now, with the help of (26) and (27), the overall minimum Euclidean distance is given by $d^* = \max_{k=1,2,\ldots,C-1} g\big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\big) = \max\ \big\{\frac{|\tilde{h}_1|}{a_1+a_2},\ldots,\frac{|\tilde{h}_1|}{a_{\ell_1}+a_{\ell_1+1}},\frac{|\tilde{h}_2|}{b_{\ell_1+1}+b_{\ell_1+2}},\ldots,\frac{|\tilde{h}_2|}{b_{C-1}+b_C}\big\} = \max\big\{\frac{|\tilde{h}_1|}{M_2},\frac{|\tilde{h}_2|}{b_{\ell_1+1}+b_{\ell_1+2}},\ldots,\frac{|\tilde{h}_2|}{b_{C-1}+b_C}\big\} = \frac{|\tilde{h}_1|}{M_2}$, where the optimality is attained when $(\tilde{w}_1^*,\tilde{w}_2^*) = (1,\frac{|\tilde{h}_1|}{M_2|\tilde{h}_2|})$ and as a result we have $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*}=M_2$.

3) If $\frac{M_1}{M_2}<\frac{|\tilde{h}_2|}{|\tilde{h}_1|}\le M_1$, we can suppose that $\frac{b_{\ell_2}+b_{\ell_2+1}}{a_{\ell_2}+a_{\ell_2+1}} \le \frac{|\tilde{h}_2|}{|\tilde{h}_1|} < \frac{b_{\ell_2+1}+b_{\ell_2+2}}{a_{\ell_2+1}+a_{\ell_2+2}}$. With the help of Lemma 1, we have

$$g\Big(\frac{b_k}{a_k},\frac{b_{k+1}}{a_{k+1}}\Big) = \begin{cases} \dfrac{|\tilde{h}_1|}{a_k+a_{k+1}}, & k=1,\ldots,\ell_2; \\ \dfrac{|\tilde{h}_2|}{b_k+b_{k+1}}, & k=\ell_2+1,\ldots,C-1. \end{cases}$$

We first show that, for $b_k+b_{k+1}, k=\ell_2+1,\ldots,C-1$, (a) If $b_k+b_{k+1}\ge M_1$, then we have $\frac{1}{b_k+b_{k+1}}\le\frac{1}{M_1}$; (b) If $b_k+b_{k+1}<M_1$, then by Property 1, we have $a_k+a_{k+1}\ge M_2$. From the assumption, we have $\frac{b_k+b_{k+1}}{a_k+a_{k+1}} \ge \frac{b_{\ell_2+1}+b_{\ell_2+2}}{a_{\ell_2+1}+a_{\ell_2+2}} >$

$\frac{|\tilde{h}_2|}{|\tilde{h}_1|} > \frac{M_1}{M_2}$. Therefore, we have $\frac{1}{b_k+b_{k+1}} < \frac{M_2}{M_1(a_k+a_{k+1})} \leq \frac{1}{M_1}$. By jointly considering both cases, we have

$$\frac{1}{b_k+b_{k+1}} \leq \frac{1}{M_1}, \quad \text{for } k = \ell_2+1, \ldots, C-1. \tag{28}$$

Next, we consider $a_k+a_{k+1}, k = 1, \ldots, \ell_2$, (a) If $a_k+a_{k+1} < M_2$, then by Property 1, we have $b_k + b_{k+1} \geq M_1$; As a result, $\frac{M_1}{a_k+a_{k+1}} \leq \frac{b_k+b_{k+1}}{a_k+a_{k+1}} \leq \frac{b_{\ell_2}+b_{\ell_2+1}}{a_{\ell_2}+a_{\ell_2+1}} \leq \frac{|\tilde{h}_2|}{|\tilde{h}_1|}$. (b) If $a_k + a_{k+1} \geq M_2$, then $\frac{M_1}{a_k+a_{k+1}} \leq \frac{M_1}{M_2} < \frac{|\tilde{h}_2|}{|\tilde{h}_1|}$. Combining both cases, we conclude that

$$\frac{|\tilde{h}_1|}{a_k+a_{k+1}} \leq \frac{|\tilde{h}_2|}{M_1}, \quad k = 1, \ldots, \ell_2. \tag{29}$$

Therefore, with the help of (28) and (29), the overall minimum Euclidean distance is $d^* = \max_{k=1,2,\ldots,C-1} g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \max\left\{\frac{|\tilde{h}_1|}{a_1+a_2}, \ldots, \frac{|\tilde{h}_1|}{a_{\ell_2}+a_{\ell_2+1}}, \frac{|\tilde{h}_2|}{b_{\ell_2+1}+b_{\ell_2+2}}, \ldots, \frac{|\tilde{h}_2|}{b_{C-1}+b_C}\right\} = \max\left\{\frac{|\tilde{h}_1|}{a_1+a_2}, \ldots, \frac{|\tilde{h}_1|}{a_{\ell_2}+a_{\ell_2+1}}, \frac{|\tilde{h}_2|}{M_1}\right\} = \frac{|\tilde{h}_2|}{M_1}$, where the optimality is attained when $(\tilde{w}_1^*, \tilde{w}_2^*) = \left(\frac{|\tilde{h}_2|}{M_1|\tilde{h}_1|}, 1\right)$ and as a result, $d^* = \frac{|\tilde{h}_2|}{M_1}$ and $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*} = \frac{1}{M_1}$.

4) If $M_1 < \frac{|\tilde{h}_2|}{|\tilde{h}_1|}$, then $\frac{b_k+b_{k+1}}{a_k+a_{k+1}} \leq M_1 < \frac{|\tilde{h}_2|}{|\tilde{h}_1|}$, for $k = 1, \ldots, C-1$. By using Lemma 1, $g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \frac{|\tilde{h}_1|}{a_k+a_{k+1}}$ for $k = 1, \ldots, C-1$, and $d^* = \max_{k=1,2,\ldots,C-1} g\left(\frac{b_k}{a_k}, \frac{b_{k+1}}{a_{k+1}}\right) = \max\left\{\frac{|\tilde{h}_1|}{a_1+a_2}, \ldots, \frac{|\tilde{h}_1|}{a_{C-1}+a_C}\right\} = \frac{|\tilde{h}_1|}{a_{C-1}+a_C} = |\tilde{h}_1|$, where the optimality is attained when $(\tilde{w}_1, \tilde{w}_2) = (1, M_1\frac{|\tilde{h}_1|}{|\tilde{h}_2|})$ with the help of Property 1, and as a result, $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*} = \frac{1}{M_1}$.

The solution to Problem 1 can be summarized as

- If $\frac{|\tilde{h}_2|}{|\tilde{h}_1|} \leq \frac{1}{M_2}$, then $(\tilde{w}_1^*, \tilde{w}_2^*) = \left(M_2\frac{|\tilde{h}_2|}{|\tilde{h}_1|}, 1\right)$, $d^* = |\tilde{h}_2|$, and $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*} = M_2$;
- If $\frac{1}{M_2} < \frac{|\tilde{h}_2|}{|\tilde{h}_1|} \leq \frac{M_1}{M_2}$, then $(\tilde{w}_1^*, \tilde{w}_2^*) = \left(1, \frac{|\tilde{h}_1|}{M_2|\tilde{h}_2|}\right)$, $d^* = \frac{|\tilde{h}_1|}{M_2}$, and $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*} = M_2$;
- If $\frac{M_1}{M_2} < \frac{|\tilde{h}_2|}{|\tilde{h}_1|} \leq M_1$, then $(\tilde{w}_1^*, \tilde{w}_2^*) = \left(\frac{|\tilde{h}_2|}{M_1|\tilde{h}_1|}, 1\right)$, $d^* = \frac{|\tilde{h}_2|}{M_1}$, and $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*} = \frac{1}{M_1}$;
- If $M_1 < \frac{|\tilde{h}_2|}{|\tilde{h}_1|}$, then $(\tilde{w}_1^*, \tilde{w}_2^*) = \left(1, M_1\frac{|\tilde{h}_1|}{|\tilde{h}_2|}\right)$, $d^* = |\tilde{h}_1|$, and $\frac{|\tilde{h}_1|\tilde{w}_1^*}{|\tilde{h}_2|\tilde{w}_2^*} = \frac{1}{M_1}$.

From the previous assumption, we know that $\tilde{w}_1 = \sqrt{\frac{2(M_1^2-1)}{3\ P_1}}w_1$, $\tilde{w}_2 = \sqrt{\frac{2(M_2^2-1)}{3\ P_2}}w_2$, $|\tilde{h}_1| = \sqrt{\frac{3\ P_1}{2(M_1^2-1)}}|h_1|$, and $|\tilde{h}_2| = \sqrt{\frac{3\ P_2}{2(M_2^2-1)}}|h_2|$. After some algebraic manipulations, the conclusion in Theorem 1 can be readily obtained and we complete the proof of the theorem. □

H. Proof of Corollary 1

Without loss of generality, we consider $\frac{|h_2|}{|h_1|} \leq \sqrt{\frac{P_1(M_2^2-1)}{P_2\ M_2^2(M_1^2-1)}}$, and therefore $|h_1|w_1^*s_1 + |h_2|w_2^*s_2 = \sqrt{\frac{3\ P_2\ M_2^2}{2(M_2^2-1)}}\frac{|h_2|}{|h_1|}|h_1|s_1 + \sqrt{\frac{3\ P_2}{2(M_2^2-1)}}|h_2|s_2 = \sqrt{\frac{3\ P_2}{2(M_2^2-1)}}|h_2|(M_2\ s_1 + s_2)$. Recall that $s_1 \in \mathcal{A}_{M_1} = \{\pm(2k-1)\}_{k=1}^{M_1/2}$ and $s_2 \in \mathcal{A}_{M_2} = \{\pm(2k-1)\}_{k=1}^{M_2/2}$, and therefore $M_2\ s_1+s_2 \in \mathcal{A}_{M_1\ M_2} = \{\pm(2k-1)\}_{k=1}^{M_1\ M_2/2}$. The quadrature component of the sum-constellation is identical to that of the in-phase component. Hence, the sum-constellation is an $M_1^2\ M_2^2$-QAM constellation with a minimum Euclidean distance d_{noma}. The case $\frac{|h_2|}{|h_1|} > \sqrt{\frac{P_1(M_2^2-1)}{P_2\ M_2^2(M_1^2-1)}}$ can be proved in a similar manner and hence is omitted for brevity. □

I. Proof of Corollary 2

Recall that d_{noma} and d_{oma} given in (13) and (14), respectively. We consider the following cases one by one as follows:

1) If $\frac{M_2^2(M_1^2-1)}{M_2^2-1} \leq \frac{P_1|h_1|^2}{P_2|h_2|^2}$, we have $d_{\text{noma}} = \sqrt{\frac{3\ P_2}{2(M_2^2-1)}}|h_2|$, and then $\frac{d_{\text{noma}}}{d_{\text{oma},2}} = \sqrt{M_2^2+1} > 1$.

2) If $\frac{M_2^2(M_1^2-1)}{M_1^2(M_2^2-1)} \leq \frac{P_1|h_1|^2}{P_2|h_2|^2} < \frac{M_2^2(M_1^2-1)}{M_2^2-1}$, we attain $d_{\text{noma}} = \sqrt{\frac{3\ P_1}{2M_2^2(M_1^2-1)}}|h_1|$ and then we consider the following two scenarios: (a) For $M_2 \leq M_1$, we conclude $\frac{d_{\text{noma}}}{d_{\text{oma},1}} = \sqrt{\frac{M_1^2+1}{M_2^2}} > 1$; (b) For $M_2 > M_1$, we attain $\frac{d_{\text{noma}}}{d_{\text{oma},2}} = \sqrt{\frac{P_1|h_1|^2(M_2^4-1)}{P_2|h_2|^2\ M_2^2(M_1^2-1)}}$. As $\frac{P_1|h_1|^2}{P_2|h_2|^2} \geq \frac{M_2^2(M_1^2-1)}{M_1^2(M_2^2-1)}$, we attain $\frac{d_{\text{noma}}}{d_{\text{oma},2}} \geq \sqrt{\frac{M_2^2+1}{M_1^2}} > 1$.

3) If $\frac{M_1^2-1}{M_1^2(M_2^2-1)} \leq \frac{P_1|h_1|^2}{P_2|h_2|^2} < \frac{M_2^2(M_1^2-1)}{M_1^2(M_2^2-1)}$, we have $d_{\text{noma}} = \sqrt{\frac{3\ P_2|h_2|^2}{2\ M_1^2(M_2^2-1)}}$. Likewise, we consider the following two scenarios: (a) For $M_1 \leq M_2$, then $\frac{d_{\text{noma}}}{d_{\text{oma},2}} = \sqrt{\frac{M_2^2+1}{M_1^2}} > 1$. (b) For $M_1 > M_2$, then $\frac{d_{\text{noma}}}{d_{\text{oma},1}} = \sqrt{\frac{P_2|h_2|^2(M_1^4-1)}{P_1|h_1|^2\ M_1^2(M_2^2-1)}}$. As $\frac{P_1|h_1|^2}{P_2|h_2|^2} < \frac{M_2^2(M_1^2-1)}{M_1^2(M_2^2-1)}$, we have $\frac{d_{\text{noma}}}{d_{\text{oma},1}} > \sqrt{\frac{M_1^2+1}{M_2^2}} > 1$.

4) If $\frac{P_1|h_1|^2}{P_2|h_2|^2} < \frac{M_1^2-1}{M_1^2(M_2^2-1)}$, we attain $\sqrt{\frac{3\ P_1}{2(M_1^2-1)}}|h_1|$, and hence $\frac{d_{\text{noma}}}{d_{\text{oma},1}} = \sqrt{M_1^2+1} > 1$.

From the above discussion, we can conclude that $d_{\text{noma}} > d_{\text{oma}}$ and this completes the proof. □

References

[1] 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.

[2] 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.

[3] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5G networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, Oct. 2017.

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

[5] P. A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[6] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2524–2530, Oct. 1998.

[7] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

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

[9] 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.

[10] 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.

[11] 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.

[12] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 611–615.

[13] Z. Ding, L. Dai, and H. V. Poor, "MIMO-NOMA design for small packet transmission in the Internet of Things," *IEEE Access*, vol. 4, pp. 1393–1405, 2016.

[14] Z. Dong, Y.-Y. Zhang, J.-K. Zhang, and X.-C. Gao, "Quadrature amplitude modulation division for multiuser MISO broadcast channels," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 8, pp. 1551–1566, Dec. 2016.

[15] W. Bao, H. Chen, Y. Li, and B. Vucetic, "Joint rate control and power allocation for non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2798–2811, Dec. 2017.

[16] Y. Yu, H. Chen, Y. Li, Z. Ding, L. Song, and B. Vucetic, "Antenna selection for MIMO nonorthogonal multiple access systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3158–3171, Apr. 2017.

[17] Y. Yu, H. Chen, Y. Li, Z. Ding, and B. Vucetic, "On the performance of non-orthogonal multiple access in short-packet communications," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 590–593, Mar. 2017.

[18] 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," *IEEE Commun. Mag.*, vol. 55, no. 10, pp. 176–183, Oct. 2017.

[19] Z. Dong, H. Chen, J.-K. Zhang, and L. Huang, "On non-orthogonal multiple access with finite-alphabet inputs in Z-channels," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 12, pp. 2829–2845, Dec. 2017.

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

[21] 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.

[22] 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.

[23] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1037–1076, 1973.

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

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

[26] A. Lozano, A. M. Tulino, and S. Verdú, "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] J. Harshan and B. S. 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.

[28] 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.

[29] 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.

[30] 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.

[31] A. Dytso, D. Tuninetti, and N. Devroye, "On the two-user interference channel with lack of knowledge of the interference codebook at one receiver," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1257–1276, Mar. 2015.

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

[33] 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.

[34] X. Xiao, Q. Huang, and E. Viterbo, "Joint optimization scheme and sum constellation distribution for multi-user Gaussian multiple access channels with finite input constellations," in *Proc. Austral. Commun. Theory Workshop (AusCTW)*, Jan. 2016, pp. 130–135.

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

[36] 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.

[37] 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.

[38] 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.

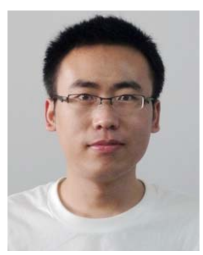
[39] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[40] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2628–2639, Jul. 2012.

[41] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, "Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2413–2424, Oct. 2017.

[42] J. Choi, "On the power allocation for MIMO-NOMA systems with layered transmissions," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3226–3237, May 2016.

[43] M. F. Hanif, Z. Ding, T. Ratnarajah, and G. K. Karagiannidis, "A minorization-maximization method for optimizing sum rate in the downlink of non-orthogonal multiple access systems," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 76–88, Jan. 2016.



Zheng Dong (M'18) 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, Canada, in 2016. His research interests include signal processing and information theory. He received the Outstanding Bachelor Thesis Award from Shandong University and the Outstanding Master Thesis Award from Shandong Province.

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 Research Fellow with 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, non-orthogonal multiple access, millimeter-wave wireless communications, and wireless energy harvesting and transfer. He was a recipient of the Chinese Government Award for Outstanding Self-Financed Students Abroad. He received the Outstanding Bachelor Thesis Award from Shandong University and the Outstanding Master Thesis Award from Shandong Province.

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

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. He holds research positions with McMaster University and Harvard University, Cambridge, MA, USA. 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 has co-authored the paper that received the IEEE Signal Processing Society Best Young Author Award in 2008. He has served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS and the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He currently serves as an Associate Editor for the *Journal of Electrical and Computer Engineering*.

Lei Huang (M'07–SM'14) was born in 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, Shenzhen Graduate School, Harbin Institute of Technology. Since 2014, he has been with the Department of Information Engineering, Shenzhen University, where he is currently a Distinguished Professor. His research interests include spectral estimation, array signal processing, statistical signal processing, and their applications in radar, and navigation and wireless communications.

He has been an Elected Member of the Sensor Array and Multichannel Technical Committee of the IEEE Signal Processing Society since 2016. He has been on the editorial boards of the IEEE TRANSACTIONS ON SIGNAL PROCESSING since 2015, the *Elsevier-Digital Signal Processing* since 2012, and the *IET Signal Processing* since 2017. He was an Elected IET Fellow in 2018.

Branka Vucetic (SM'00–F'03) has held research and academic positions in Yugoslavia, Australia, U.K., and China. She is currently an ARC Laureate Fellow, a Professor of telecommunications, and the Director of the Centre of Excellence in Telecommunications, The University of Sydney. Her research interests include coding, communication theory, and signal processing and their applications in wireless networks, and industrial Internet of Things.

She has co-authored four books and over 400 papers in telecommunications journals and conference proceedings. She is a fellow of the Australian Academy of Technological Sciences and Engineering.