Discrete Signaling and Treating Interference as Noise for the Gaussian Interference Channel

Min Qiu, Yu-Chih Huang, and Jinhong Yuan

Abstract

The two-user Gaussian interference channel (G-IC) is revisited, with a particular focus on practically amenable discrete input signalling and treating interference as noise (TIN) receivers. The corresponding deterministic interference channel (D-IC) is first investigated and coding schemes that can achieve the entire capacity region of D-IC under TIN are proposed. These schemes are then systematically translated into multi-layer superposition coding schemes based on purely discrete inputs for the real-valued G-IC. Our analysis shows that the proposed scheme is able to achieve the entire capacity region to within a constant gap for all channel parameters. To the best of our knowledge, this is the first constant-gap result under purely discrete signalling and TIN for the entire capacity region and all the interference regimes. Furthermore, the approach is extended to obtain coding scheme based on discrete inputs for the complex-valued G-IC. For such a scenario, the minimum distance and the achievable rate of the proposed scheme under TIN are analyzed, which takes into account the effects of random phase rotations introduced by the channels. Simulation results show that our scheme is capable of approaching the capacity region of the complex-valued G-IC and significantly outperforms Gaussian signalling with TIN in various interference regimes.

Index Terms

Interference channel, discrete inputs, treating interference as noise.

This work was presented in part at the 2019 IEEE Global Communications Conference (GLOBECOM) [1] and will be presented in part at the 2020 IEEE Internal Symposium on Information Theory (ISIT) [2].

M. Qiu, and J. Yuan are with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, NSW, 2052 Australia (e-mail: min.qiu@unsw.edu.au; j.yuan@unsw.edu.au).

Y.-C. Huang is with the Institute of Communications Engineering, National Chiao Tung University, Hsinchu City 30010, Taiwan (e-mail: jerryhuang@nctu.edu.tw).
I. Introduction

Interference is one of the key challenges in wireless networks where multiple uncoordinate transmissions share and compete for the same medium resource [3]. To study this problem, it is essential to start with one of the most fundamental channel models: the two-user Gaussian interference channel (G-IC). For this channel, after a long pursuit [4], [5], the capacity region can now be characterized to within 1 bit/s/Hz [6]. For some special cases where the interference are either strong (including moderately strong and very strong) [7]–[9] or very weak (and symmetric) [10], the exact characterizations of the capacity regions are also available. The key ingredients for deriving these results are a tight converse bound [6] and the use of Han-Kobayashi (HK) scheme [11], [12] along with Gaussian signaling. The main idea of the HK scheme is to split the message at each transmitter into a common message and a private message, while the common message needs to be successfully decoded and subtracted out first at both intended and unintended receivers. However, such a successive interference cancellation procedure would introduce extra decoding latency and complexity and may compromise the security of the transmissions.

Treating interference as noise (TIN) is more appealing in practice due to its low complexity and latency. It is well-known that when the interference is sufficiently weak (i.e., the very weak interference regime), Gaussian signaling with TIN is constant-gap optimal for the two-user G-IC [10], [13], [14] and it is optimal in the $K$-user G-IC from a generalized degrees of freedom perspective [15]. However, for other interference regimes, adopting Gaussian signaling with TIN usually achieves significantly suboptimal results due to the excessive interference. Apart from the regular interference channel, the application of TIN has also been a subject of study for other channels such as the X channel [16], the parallel interference channel [17], the interfering multiple access channel [18]. Essentially, all these studies demonstrate a negative trend of TIN as the interference will significantly degrade the achievable rate except that when the interference is sufficiently small such that each user’s desired signal strength is no less than the sum of the strongest interference strengths from and to this user [15]. On the other hand, encouraging results can be found in [19], [20] where the capacity region of the interference channel is shown to be achievable with each receiver performing single-user decoding, i.e., TIN. However, due to the multi-letter nature of the results in [19], [20], the capacity region is hard to compute and the capacity-achieving input distributions are difficult to find. Nonetheless, these results reveal that the suboptimality of TIN is not fundamental to the problem itself; but merely a limitation
to the existing schemes. This motivates our study of finding input distributions that are optimal under TIN.

Since most of the classical TIN results adopt Gaussian input distributions, one may start suspecting that the Gaussian input distributions are the main source of the suboptimality of TIN. Indeed, although being the best input distribution for an additive noise channel, Gaussian is also the worst additive noise [21]. Recently, there have been some researches in this direction [22]–[24] that confirm the above suspect. In [22], it was shown that it is possible to achieve higher rate when one user adopts discrete inputs and the other user adopts Gaussian inputs. Furthermore, Dytso et al. [23] showed that employing mixed pulse amplitude modulation (PAM) and Gaussian inputs at each user can achieve the capacity region of the real-valued G-IC within a gap of at most $O(\log_2(\log_2(\min(\text{SNR}, \text{INR}))/\eta))$ up to a Lebesgue measure $\eta \in (0, 1]$, where SNR is the signal to noise ratio and INR is the interference to noise ratio. The rationale behind this success is that under TIN, the structure of discrete interference can be harnessed by carefully designing the power allocation for the discrete and continuous parts of the mixed inputs. When only purely discrete PAM is employed, it was further proved in [24] that the sum capacity of the symmetric real-valued G-IC for any interference level regime can be achieved to within a gap of $O(\log_2(\log_2(\text{SNR})))$ under TIN. This gap has been further shrunk in [25] to a constant number of bits under the assumption that the channel gains are powers of 2. The approach in [24] was to construct a scheme with TIN to achieve the sum capacity of the symmetric real-valued deterministic interference channel (D-IC) [26], i.e., the linear deterministic approximation of the G-IC [27], and translate the scheme for real-valued G-IC. However, it remains unclear whether it is possible to construct practical schemes based on purely discrete input and TIN to achieve the entire capacity region of the asymmetric G-IC to within a constant gap for all channel parameters. Moreover, for the complex G-IC, since all the links have their own channel phases, it does not seem possible to simultaneously compensate all the channel phase distortions, even with full channel knowledge at the transmitters and the receivers. As a result, the performance of using discrete modulations could be severely affected by the phase distortions.

In this work, we continue the quest of designing (asymptotically) optimal input distributions that can achieve the capacity region to within a constant gap for all channel parameters. In

---

1 The asymmetric very strong interference regime and some subregimes of the symmetric weak interference regime can be achieved by purely discrete inputs to within a constant gap [23].
particular, for practical relevance, we focus solely on purely discrete input distributions at the encoders and TIN at the decoders. Moreover, we restrict our scheme to a single time-slot (no symbol extensions). Our goal here is not to obtain sharpened bounds on the achievable rate of discrete inputs, but rather to further push the effort of discrete signaling with TIN in other interference regimes and show its (constant-gap) optimality. Specifically, we focus on the not very weak, not very strong and mixed interference regimes as for other regimes, such results have been shown. The main contributions of the papers are as follows:

- To obtain insights into the asymmetric G-IC, we first look into the corresponding D-IC model [26] and provide a systematic way to construct schemes that are proven to achieve the entire capacity region for all channel parameters under the considered interference regime. Different from the results in [26] which achieve the capacity region of D-IC with HK schemes, our achievable scheme is based on TIN. We would also like to emphasize that our scheme is a non-trivial generalization of [24] since the channel parameters are different for two users and we focus on achieving the entire capacity region rather than a single rate pair in the symmetric case.

- We then systematically translate the proposed scheme from the D-IC into a multi-layer superposition coding scheme based on PAM signaling for the real-valued G-IC. To analyze the performance of our scheme, we use an Ozarow-type bound [28] together with a detailed analysis of the minimum distance of the superimposed constellation at each receiver. This allows us to prove that the proposed scheme is able to achieve the entire capacity region to within a constant gap, regardless of channel parameters. To the best of our knowledge, this is the first time that discrete signaling with TIN is proved to be constant-gap optimal for all the interference regimes. It is also worth noting that a two-layer scheme based on PAM inputs was mentioned in [23, Sec. VIII-C] that may be good for the moderately weak interference regime. Our results can be deemed as a significant extension of the two layer scheme to multi-layer and to cover all the interference regimes. Moreover, we provide a complete understanding and new insights based on thorough analysis for the multi-layer inputs schemes.

- We also extend our scheme to a multi-layer superposition coding based on quadratic amplitude modulation (QAM) for the complex G-IC by translating our capacity-achieving scheme from the D-IC. The achievable rate is similarly bounded by our derived variant of the
Ozarow-type bound. Although obtaining a closed form expression for the minimum distance of the superimposed constellation is difficult due to random phase distortions experienced by different links, we still manage to get crude analysis of the minimum distance, from which some insights into the gap between the achievable rate and the capacity of the complex G-IC can be drawn. Simulation results show that our scheme is capable of approaching the (upper bound) capacity region of the complex G-IC [6] with discrete inputs and TIN and it significantly outperforms Gaussian inputs with TIN.

A. Notations

This paper uses the following notations. \( \mathbb{Z} \), \( \mathbb{N} \), \( \mathbb{R} \) and \( \mathbb{C} \) represent the sets of integers, natural numbers, real numbers and complex numbers, respectively. Random variables are written in uppercase Sans Serif font, e.g., \( X \). For \( x \in \mathbb{R} \), \( \lfloor x \rfloor = k \in \mathbb{Z} \) gives the nearest integer \( k \leq x \). For a set \( S \), \( |S| \) outputs the cardinality of \( S \). For some \( a, b \in \mathbb{Z} \) and \( b > a \), the integer set \( \{a, a+1, \ldots, b\} \) is represented by \( [a : b] \). For \( x \in \mathbb{C} \), \( \Re(x) \) and \( \Im(x) \) represent the real and imaginary part of \( x \), respectively. The binary field and the collection of binary matrices of size \( m \times n \) are denoted by \( \mathbb{F}_2 \) and \( \mathbb{F}_{2^m}^{m,n} \), respectively. Unless specified otherwise, we use \( \text{PAM}(|\Lambda|, d_{\min}(\Lambda)) \) to represent a conventional PAM \( \Lambda \) with mean \( \mathbb{E}[\Lambda] = 0 \), cardinality \( |\Lambda| \), minimum distance \( d_{\min}(\Lambda) \) and with average energy \( \mathbb{E}[\|\Lambda\|^2] = d_{\min}^2(\Lambda) \frac{|\Lambda|^2-1}{12} \). Similarly, we use \( \text{QAM}(|\Lambda|, d_{\min}(\Lambda)) \) to represent a conventional QAM \( \Lambda \) with mean \( \mathbb{E}[\Lambda] = 0 \), cardinality \( |\Lambda| \), minimum distance \( d_{\min}(\Lambda) \) and with average energy \( \mathbb{E}[\|\Lambda\|^2] = d_{\min}^2(\Lambda) \frac{|\Lambda|^2-1}{6} \).

II. System Model

The two-user complex-valued G-IC is described by the following input-output relationship

\[
Y_1 = h_{11}X_1 + h_{12}X_2 + Z_1, \tag{1}
\]

\[
Y_2 = h_{21}X_1 + h_{22}X_2 + Z_2, \tag{2}
\]

where for \( k, i \in \{1, 2\} \), \( h_{ki} \) is the channel between user \( k \)'s transmitter and receiver \( i \), \( X_k \) is user \( k \)'s signal intended for receiver \( k \) and is subject to a unit power constraint \( \mathbb{E}[\|X_k\|^2] \leq 1 \), and \( Z_k \sim \mathcal{CN}(0, 1) \) is the additive white Gaussian noise. The channel \( h_{ki} = |h_{ki}|e^{j\theta_{ki}} \) is characterized by their amplitudes and phases. We assume that the channels are fixed and the channel magnitudes are known to all transmitters and receivers while the channel phases are only known to receivers, in order to focus on the impact of random phase rotations to the discrete input distributions.
We define user \(k\)’s SNR as \(\text{SNR}_k \triangleq |h_{kk}^2|\), user 1’s INR as \(\text{INR}_1 \triangleq |h_{12}^2|\) and user 2’s INR as \(\text{INR}_2 \triangleq |h_{21}^2|\)².

For the real-valued G-IC, the channel model is identical to the complex setting except that all the signals are real numbers, the noise becomes \(Z_k \sim \mathcal{N}(0,1)\), and there is no channel phase rotations.

For both real and complex G-IC, a set of outer bounds have been established in the literature. To be specific, the outer bounds for the weak and mixed interference regimes are characterized in [6] and that for the strong interference regime is in [8], [9]. We will compare the achieve rate of our scheme with these bounds.

III. THE LINEAR DETERMINISTIC INTERFERENCE CHANNEL

In this section, we first look into the linear D-IC as an approximation to the G-IC model and propose a family of capacity achieving schemes. The schemes obtained here will be systematically translated into coding schemes for real and complex G-IC in Sec. IV and Sec. V, respectively.

A. Channel Model

The channel model for the two-user D-IC is defined as [26]

\[ \begin{align*}
Y_1 &= S^{q-n_{11}}X_1 \oplus S^{q-n_{12}}X_2, \\
Y_2 &= S^{q-n_{21}}X_1 \oplus S^{q-n_{22}}X_2,
\end{align*} \]

where the multiplication and summation are over \(\mathbb{F}_2\), \(n_{kk} \triangleq \lceil \log_2 \text{SNR}_k \rceil\) for \(k \in \{1, 2\}\), \(n_{12} \triangleq \lceil \log_2 \text{INR}_1 \rceil\), and \(n_{21} \triangleq \lceil \log_2 \text{INR}_2 \rceil\) are for the complex channel setting; while \(n_{kk} \triangleq \)

²We consider \(\text{SNR}_k \geq 1\) and \(\text{INR}_k \geq 1\) as the capacities of user 1 and user 2’s direct links for \(\text{SNR}_k < 1\) and \(\text{INR}_k < 1\) are at most 1 bit for the complex G-IC and \(\frac{1}{2}\) bits for the real G-IC. Hence, the gap between the rate pair of any achievable scheme and the capacity region under \(\text{SNR}_k < 1\) and \(\text{INR}_k < 1\) is already bounded.
\[ \left\lfloor \frac{1}{2} \log_2 \text{SNR}_k \right\rfloor \] for \( k \in \{1, 2\} \), \( n_{12} \triangleq \left\lfloor \frac{1}{2} \log_2 \text{INR}_1 \right\rfloor \), and \( n_{21} \triangleq \lfloor \log_2 \frac{1}{2} \text{INR}_2 \rfloor \) are adopted for the real channel setting, \( q = \max\{n_{11}, n_{12}, n_{21}, n_{22}\} \), \( S \) is a \( q \times q \) shift matrix,

\[
S = \begin{bmatrix}
0 & 0 & 0 & \ldots & 0 \\
1 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & \ldots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & 1 & 0
\end{bmatrix},
\]

and \( X_k, Y_k \in \mathbb{F}_2^q \) are two binary column vectors representing the discrete channel inputs and outputs, respectively, for user \( k \). Each entry of the input column vector represents a power level. The most significant bit of \( X_k \) is called the highest power level while the least significant bit is called the lowest level. We note that \( S^n X_k \) for some natural number \( n \leq q \) will down-shift \( X_k \) by \( n \) positions, which mimics the situation that the lowest \( n \) bits of \( X_k \) are already below the noise level from the receiver perspective.

According to [26], the capacity region of the D-IC is the set of non-negative rate pair \( (r_1, r_2) \) satisfying the following:

\[
\begin{align*}
    r_k & \leq n_{kk}, \quad k = 1, 2 \\
    r_1 + r_2 & \leq \max\{n_{11} - n_{12}, 0\} + \max\{n_{22}, n_{12}\} \\
    r_1 + r_2 & \leq \max\{n_{22} - n_{21}, 0\} + \max\{n_{11}, n_{21}\} \\
    r_1 + r_2 & \leq \max\{n_{21}, \max\{n_{11} - n_{12}, 0\}\} \\
    & \quad + \max\{n_{12}, \max\{n_{22} - n_{21}, 0\}\} \\
    2r_1 + r_2 & \leq \max\{n_{11}, n_{21}\} + \max\{n_{11} - n_{12}, 0\} \\
    & \quad + \max\{n_{12}, \max\{n_{22} - n_{21}, 0\}\} \\
    r_1 + 2r_2 & \leq \max\{n_{22}, n_{12}\} + \max\{n_{22} - n_{21}, 0\} \\
    & \quad + \max\{n_{21}, \max\{n_{11} - n_{12}, 0\}\}.
\end{align*}
\]

Most importantly, it is shown that the capacity gap between the D-IC and the G-IC can be upper bounded by a constant which is independent of SNR and INR [26]. We emphasize here that the above capacity region of the D-IC can be achieved by Han-Kobayashi scheme [26]. In what follows, we will construct schemes to achieve the capacity region with TIN.
B. Proposed Scheme and the Main Result

For $k \in \{1, 2\}$, let $U_k$ be user $k$’s message vector of length $r_k$ with i.i.d. entries drawn independently and uniformly distributed over $\mathbb{F}_2$. And let $X_k = E_k U_k$, for some $E_k \in \mathbb{F}_2^{q \times r_k}$. We also let $A_1 \triangleq S^{q-n_{11}}, B_1 \triangleq S^{q-n_{12}}, A_2 \triangleq S^{q-n_{22}},$ and $B_2 \triangleq S^{q-n_{21}}$ represent the channels of the D-IC. Here, we aim to design the generator matrices $E_1$ and $E_2$ which produce the discrete input distributions to achieve the capacity region of the D-IC.

The achievable rate of user 1 with single-user decoding (i.e., TIN) can be derived as

$$I(X_1; Y_1) = H(Y_1) - H(Y_1|X_1)$$
$$= H(S^{q-n_{11}} E_1 U_1 \oplus S^{q-n_{12}} E_2 U_2) - H(S^{q-n_{12}} E_2 U_2)$$
$$= \text{rank}([S^{q-n_{11}} E_1 S^{q-n_{12}} E_2]) - \text{rank}(S^{q-n_{12}} E_2)$$
$$= \text{rank}([A_1 E_1 B_1 E_2]) - \text{rank}(B_1 E_2),$$

(12)

where the multiplication and addition are over $\mathbb{F}_2$.

Similarly, the achievable rate of user 2 with single-user decoding is given by

$$I(X_2; Y_2) = H(Y_2) - H(Y_2|X_2)$$
$$= H(S^{q-n_{22}} E_2 U_2 \oplus S^{q-n_{21}} E_1 U_1) - H(S^{q-n_{21}} E_1 U_1)$$
$$= \text{rank}([S^{q-n_{22}} E_2 S^{q-n_{21}} E_1]) - \text{rank}(S^{q-n_{21}} E_1)$$
$$= \text{rank}([A_2 E_2 B_2 E_1]) - \text{rank}(B_2 E_1).$$

(13)

From this point onwards, the problem becomes jointly designing $E_1$ and $E_2$ such that

$$(I(X_1; Y_1), I(X_2; Y_2)) = (r_1, r_2),$$

with $r_1$ and $r_2$ lying inside the capacity region defined in (6)–(11).

We state the main result of this section in the following.

**Theorem 1.** In any interference regime of the D-IC, for any target rate pair $(r_1, r_2)$ lying inside the capacity region, there exist a pair of input distributions $(X_1, X_2)$ such that $(r_1, r_2)$ can be achieved with TIN, i.e., $(I(X_1; Y_1), I(X_2; Y_2)) = (r_1, r_2)$.

Since TIN has been proved to be constant-gap optimal in the very weak interference regime [15] and in the very strong interference regime [23], here we provide the proof for the rest of the interference regimes.
C. Proof of Theorem 1

We assume \( n_{11} \geq n_{22} \) without loss of generality. To clearly express the idea of our approach in the interest of space, we give the full proof for two typical cases and defer the proof for the rest of the cases to Appendix A.

We first look at the weak interference regime, which is defined as \( n_{11} > n_{21}, n_{22} > n_{12} \) according to [6]. To distinguish the considered case with the very weak interference regime, \( n_{11}, n_{22}, n_{12}, \) and \( n_{21} \) shall not satisfy \( n_{11} > n_{12} + n_{21} \) and \( n_{22} > n_{12} + n_{21} \) simultaneously [15].

Weak 1: \( n_{11} > n_{22} > n_{12} > n_{21} \)

1) : We consider the subcase \( n_{11} > n_{12} + n_{21} > n_{22} \). The capacity region is reduced to the set of non-negative rate pair \( (r_1, r_2) \) satisfying

\[
\begin{align*}
    r_1 &\leq n_{11} \\
    r_2 &\leq n_{22} \\
    r_1 + r_2 &\leq n_{11} \\
    2r_1 + r_2 &\leq 2n_{11} \\
    r_1 + 2r_2 &\leq n_{11} - n_{12} + 2n_{22} - n_{21}
\end{align*}
\]

By inspecting the above capacity region, we obtain four corner points on the capacity region \((n_{11}, 0), (n_{11} + n_{12} + n_{21} - 2n_{22}, 2n_{22} - n_{12} - n_{21}), (n_{11} - n_{12} - n_{21}, n_{22}), (0, n_{22})\).

1a) To achieve the rate pairs between point \((n_{11}, 0)\) and \((n_{11} + n_{12} + n_{21} - 2n_{22}, 2n_{22} - n_{12} - n_{21})\), we propose

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{t_1, r_1} \\
F_{1,2} \\
0^{t_2, r_1} \\
F_{1,3}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{22} - n_{21} - t_2, r_2} \\
0^{n_{12} + n_{21} - n_{22}, r_2} \\
0^{n_{22} - n_{12} - t_1, r_2} \\
F_{2,2} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix}
\]

where \( F_{1,1} \in \mathbb{F}_{2}^{n_{21} - t_1, r_1}, F_{1,2} \in \mathbb{F}_{2}^{n_{11} - n_{21} - n_{12}, r_1}, F_{1,3} \in \mathbb{F}_{2}^{n_{11} - t_2, r_1}, F_{2,1} \in \mathbb{F}_{2}^{t_2, r_2}, F_{2,2} \in \mathbb{F}_{2}^{t_1, r_2} \) and all the rows from any one of these submatrices are linearly independent. It should be noted that \( F_{i,j} \) and \( F_{i',j'} \) are always linearly independent as long as \((i, j) \neq (i', j')\). The two independent variables \( t_1 \in [0 : n_{22} - n_{12}], t_2 \in [0 : n_{22} - n_{21}] \) are tunable parameters allowing our scheme to
achieve all the integer rate pairs between corner points \((n_{11}, 0)\) and \((n_{11} + n_{12} + n_{21} - 2n_{22}, 2n_{22} - n_{12} - n_{21})\).

Substituting \(E_1\) and \(E_2\) into the first term of the last equation in (12) leads to

\[
\text{rank}([A_1 E_1 B_1 E_2]) = \text{rank} \left( \begin{bmatrix} F_{1,1} & 0_{t_1, r_1} & 0_{n_{11} - n_{12}, r_2} \\ F_{1,2} & 0_{t_2, r_1} & F_{2,1} \\ F_{1,3} & 0_{n_{22} - n_{21} - t_2, r_2} & 0_{n_{12} + n_{21} - n_{22}, r_2} \end{bmatrix} \right) = \min \{n_{11} - t_1, r_1 + r_2\}. \tag{20}\]

Note that for our proposed \(E_1\), the rank of \(A_1 E_1\) is equal to \(r_1\), i.e., we have

\[
\text{rank}(A_1 E_1) = n_{11} - t_1 - t_2 = r_1. \tag{21}\]

Substituting (21) into (20) gives

\[
\text{rank}([A_1 E_1 B_1 E_2]) = \min \{n_{11} - t_1, n_{11} - t_1 - t_2 + r_2\}. \tag{22}\]

The last term of the last equation in (12) becomes

\[
\text{rank}(B_1 E_2) = \text{rank}(F_{2,1}) = \min \{t_2, r_2\}. \tag{23}\]

Substituting (20), (22) and (23) into (12), we obtain user 1’s rate as

\[
I(X_1; Y_1) = n_{11} - t_1 - t_2. \tag{24}\]

For user 2, substituting \(E_1\) and \(E_2\) into (13) gives

\[
\text{rank}([A_2 E_2 B_2 E_1]) = \text{rank} \left( \begin{bmatrix} 0_{n_{11} - n_{22}, r_2} \& 0_{n_{11} - n_{21}, r_1} \\ F_{2,1} \& 0_{n_{22} - n_{21} - t_2, r_2} \\ 0_{n_{12} + n_{21} - n_{22}, r_2} \& F_{1,1} \\ 0_{n_{22} - n_{12} - t_1, r_2} \& 0_{t_1, r_1} \end{bmatrix} \right) = \min \{n_{21} + t_2, r_2 + r_1\}. \tag{25}\]
Similar to the case for user 1, for our designed \( E_2 \), it holds that the rank of \( A_2 E_2 \) is equal to
\[
\text{rank}(A_2 E_2) = t_1 + t_2 = r_2. \tag{26}
\]

Note that our joint design of \( E_1 \) and \( E_2 \) ensures that conditions (21) and (26) are satisfied simultaneously. Due to (26), the rank of \( [A_2 E_2 B_2 E_1] \) can be written as
\[
\text{rank}([A_2 E_2 B_2 E_1]) = \min\{n_{21} + t_2, t_1 + t_2 + r_1\}. \tag{27}
\]

The last term of the last equation in (13) becomes
\[
\text{rank}(B_2 E_1) = \text{rank}(F_{1,1}) = \min\{n_{21} - t_1, r_1\}. \tag{28}
\]

Hence, user 2’s rate is obtained as
\[
I(X_2; Y_2) = t_1 + t_2. \tag{29}
\]

It can be easily verified that the rate pair \((n_{11} - t_1 - t_2, t_1 + t_2)\) satisfying (16) and hence achieves the corner points \((n_{11}, 0)\) and \((n_{11} + n_{12} + n_{21} - 2n_{22}, 2n_{22} - n_{12} - n_{21})\) as well as all the integer rate pairs between them.

**Remark 1.** *In the above scheme, each signal corresponding to the submatrix in \( E_1 \) or \( E_2 \) is uniquely associated with one power level (i.e., the position of the corresponding submatrix in \( E_1 \) or \( E_2 \)) and different signals have different power levels. We refer to such a scheme as a type I scheme. One can see from the above derivations that our designed scheme in the above example enforces two properties:*

**P1.** *For each received signal, the desired signal is placed at a set of signal levels (i.e., the position of the corresponding submatrix in \( [A_1 E_1 B_2 E_2] \) or \( [A_2 E_2 B_2 E_1] \) that is disjoint with the set of signal levels occupied by the interference;*

**P2.** *The rank of \( A_k E_k \) equals to \( r_k \), i.e., \( \text{rank}(A_k E_k) = r_k \) for \( k \in \{1, 2\} \).*

In fact, all the type I schemes proposed in this paper satisfy the above two properties. Consequently, user 1’s mutual information in (12) can then be simplified to
\[
I(X_1; Y_1) = \text{rank}([A_1 E_1 B_1 E_2]) - \text{rank}(B_1 E_2)
= \text{rank}(A_1 E_1) + \text{rank}(B_1 E_2) - \text{rank}(B_1 E_2)
= \text{rank}(A_1 E_1). \tag{30}
\]
Similarly, for user 2, the mutual information in (13) can be simplified to

\[ I(X_2; Y_2) = \text{rank}(A_2 E_2). \]  

(31)

It should be noted that for arbitrary \( E_1 \) and \( E_2 \), the above two properties may not hold in general.

1b) To achieve the capacity region between points \((n_{11} + n_{12} + n_{21} - 2n_{22}, 2n_{22} - n_{12} - n_{21})\) and \((n_{11} - n_{12} - n_{21}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix}
0^{t_3, r_1} \\
F_{1,1} \\
0^{n_{22} - n_{12}, r_1} \\
F_{1,2} \\
0^{n_{22} - n_{21} + t_3, r_1} \\
F_{1,3}
\end{bmatrix}
\quad \text{and} \quad
E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{12} + n_{21} - n_{22} - t_3, r_2} \\
F_{2,2} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
\]  

(32)

where \( F_{1,1} \in F_2^{n_{12} + n_{21} - n_{22} - t_3, r_1} \), \( F_{1,2} \in F_2^{n_{11} - n_{12} - n_{21}, r_1} \), \( F_{1,3} \in F_2^{n_{12} + n_{21} - n_{22} - t_3, r_1} \), \( F_{2,1} \in F_2^{n_{22} - n_{21} + t_3, r_2} \), \( F_{2,2} \in F_2^{n_{22} - n_{12}, r_2} \), \( t_3 \in [0: n_{12} + n_{21} - n_{22}] \).

We design \( E_1 \) and \( E_2 \) such that \( \text{P1.} \) and \( \text{P2.} \) hold, which can be seen by noting that user 1 and user 2’s signals are disjoint in \([A_1 E_1 \ B_1 E_2]\), i.e.,

\[
[A_1 E_1 \ B_1 E_2] = \begin{bmatrix}
0^{t_3, r_1} \\
F_{1,1} \\
0^{n_{22} - n_{12}, r_1} \\
F_{1,2} \\
0^{n_{22} - n_{21} + t_3, r_1} \\
F_{1,3}
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
F_{2,1} \\
0^{n_{12} + n_{21} - n_{22} - t_3, r_2}
\end{bmatrix},
\]  

(33)

and

\[ \text{rank}(A_1 E_1) = n_{11} + n_{12} + n_{21} - 2n_{22} - 2t_3 = r_1. \]  

(34)

As a result, user 1’s rate can be directly obtained by using (30) as

\[ I(X_1; Y_1) = \text{rank}(A_1 E_1) = n_{11} + n_{12} + n_{21} - 2n_{22} - 2t_3. \]  

(35)
For user 2, notice that

\[
\begin{bmatrix}
A_2 & E_2 \\
B_2 & E_1
\end{bmatrix} = \begin{bmatrix}
0_{n_{11}-n_{22},r_1} & 0_{n_{11}-n_{21}+t_3,r_2} \\
F_{2,1} & 0_{n_{12}+n_{21}-n_{22}-t_3,r_1} \\
0_{n_{12}+n_{21}-n_{22}-t_3,r_1} & F_{1,1} \\
F_{2,2} & 0_{n_{22}-n_{12},r_2}
\end{bmatrix},
\]  

\hspace{1cm} (36)

and

\[\text{rank}(A_2E_2) = 2n_{22} - n_{12} - n_{21} + t_3 = r_2.\]  

\hspace{1cm} (37)

Hence, P1. and P2. still hold. User 2’s rate can then be directly obtained by using (31) as

\[I(X_2; Y_2) = \text{rank}(A_2E_2) = 2n_{22} - n_{12} - n_{21} + t_3.\]  

\hspace{1cm} (38)

The achievable rate pair \((n_{11} + n_{12} + n_{21} - 2n_{22} - 2t_3, 2n_{22} - n_{12} - n_{21} + t_3)\) satisfies (18) and also achieves the corner points \((n_{11} + n_{12} + n_{21} - 2n_{22}, 2n_{22} - n_{12} - n_{21})\) and \((n_{11} - n_{12} - n_{21}, n_{22})\).

Note that one can also obtain the same achievable rate pair by computing the ranks in (12) and (13) one by one, albeit with more computing steps.

\(1c\) To achieve the capacity region between \((n_{11} - n_{12} - n_{21}, n_{22})\) and \((0, n_{22})\), we propose

\[E_1 = \begin{bmatrix}
0_{n_{12},r_1} \\
0_{t_4,r_1} \\
F_{1,1}
\end{bmatrix}, E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0_{n_{11}-n_{22},r_2}
\end{bmatrix},\]  

\hspace{1cm} (39)

where \(F_{1,1} \in \mathbb{F}_{2}^{n_{11}-n_{12}-n_{21}-t_4,r_1}, F_{2,1} \in \mathbb{F}_{2}^{n_{12},r_2}, F_{2,2} \in \mathbb{F}_{2}^{n_{22}-n_{12},r_2}, t_4 \in [0 : n_{11} - n_{12} - n_{21}]\).

For user 1, we note that P1. and P2. hold since

\[\begin{bmatrix}
A_1 & E_1 \\
B_1 & E_2
\end{bmatrix} = \begin{bmatrix}
0_{n_{21},r_1} & 0_{n_{11}-n_{12},r_2} \\
0_{t_4,r_1} & F_{1,1}
\end{bmatrix}.\]  

\hspace{1cm} (40)

Thus, user 1’s achievable rate can be obtained from (30) as

\[I(X_1; Y_1) = n_{11} - n_{12} - n_{21} - t_4.\]  

\hspace{1cm} (41)
For user 2, $\mathbf{P}_1$ and $\mathbf{P}_2$ still hold since

$$[\mathbf{A}_2 \mathbf{E}_2 \mathbf{B}_2 \mathbf{E}_1] = \begin{bmatrix} 0^{n_{11}-n_{22},r_1} \\ \mathbf{F}_{2,1} \\ \mathbf{F}_{2,2} \end{bmatrix} 0^{n_{11},r_2}.$$  \hfill (42)

Thus, user 2’s achievable rate can be obtained from (31) as

$$I(X_2; Y_2) = n_{22}.$$  \hfill (43)

The achievable rate pair satisfies (15) and also achieves the corner points $(n_{11} - n_{12} - n_{21}, n_{22})$ and $(0, n_{22})$.

2) : We now consider $n_{11} < n_{12} + n_{21}, n_{11} + n_{22} - n_{12} - 2n_{21} < 0$. The corner points on the corresponding capacity region are $(n_{11}, 0), (n_{11} + n_{12} - n_{22}, 2(n_{22} - n_{12})), (2(n_{11} - n_{12}), n_{22} + n_{12} - n_{11}), (0, n_{22})$. Here, the idea of avoiding interference by carefully design signal levels may no longer be sufficient to achieve the desired rate pair. Therefore, we propose another type of schemes, which we refer to as the type II schemes.

Under the conditions outlined at the beginning of Sec. [III-C2] we further consider the subcase $2(n_{12} + n_{21} - n_{22}) - n_{11} < 0$ which implies that $2n_{11} + n_{22} - 2n_{12} - 2n_{21} > 0$ because $n_{11} > n_{22}$.

2a) To achieve the region between $(n_{11}, 0), (n_{11} + n_{12} - n_{22}, 2(n_{22} - n_{12}))$, we propose

$$\mathbf{E}_1 = \begin{bmatrix} \mathbf{F}_{1,1} \\ \mathbf{F}_{1,2} \\ \mathbf{F}_{1,3} \\ \mathbf{F}_{1,4} \\ 0^{2n_{22} + n_{11} - 2n_{12} - 2n_{21} - t_2, r_1} \end{bmatrix}, \mathbf{E}_2 = \begin{bmatrix} 0^{n_{12} - n_{21}, r_2} \\ 0^{t_2, r_2} \\ \mathbf{F}_{2,1} \\ \mathbf{F}_{2,2} \\ \mathbf{F}_{2,3} \\ \mathbf{F}_{2,4} \\ 0^{n_{11} - n_{22}, r_2} \end{bmatrix},$$  \hfill (44)

where $\mathbf{F}_{1,1} \in \mathbb{F}_2^{n_{11}-n_{21},r_1}, \mathbf{F}_{1,2} \in \mathbb{F}_2^{t_1,r_2}, \mathbf{F}_{1,3} \in \mathbb{F}_2^{2n_{21}+n_{12}-n_{11}-n_{22}-t_1,r_1}, \mathbf{F}_{1,4} \in \mathbb{F}_2^{t_2,r_1}, \mathbf{F}_{2,1} \in \mathbb{F}_2^{2n_{21}+n_{12}-n_{11}-n_{22}-t_2,r_2}, \mathbf{F}_{2,2} \in \mathbb{F}_2^{2n_{22}+n_{11}-2n_{12}-2n_{21}-t_2,r_2}, \mathbf{F}_{2,3} \in \mathbb{F}_2^{2n_{22}+n_{11}-2n_{12}-2n_{21}-t_2,r_2}, \mathbf{F}_{2,4} \in \mathbb{F}_2^{n_{11}-n_{22},r_2},$ and $t_1 \in [0 : 2n_{21} + n_{12} - n_{11} - n_{22}], t_2 \in [0 : 2n_{22} + n_{11} - 2n_{12} - 2n_{21}].$
Substituting $E_1$ and $E_2$ into the first term of the last equation in (12) leads to

$$\text{rank}([A_1E_1 B_1E_2]) = \text{rank}\left(\begin{bmatrix} F_{1,1} & 0^{n_{11}-n_{12},r_2} \\ F_{1,2} & 0^{t_1,r_2} \\ F_{1,3} & F_{2,1} \\ F_{1,4} & 0^{t_2,r_2} \\ F_{1,5} & 0^{n_{12}+n_{21}-n_{22},r_2} \\ F_{1,6} \end{bmatrix}\right)$$

$$= \min\{n_{11}, r_1 + r_2\}.$$  \hspace{1cm} (45)

Different from the type I scheme, here we introduce a correlation in $E_1$, which can be seen by noting that $F_{1,3}$ appears in two different levels. Moreover, we still design the generator matrices such that $P_2$ holds and thus

$$\text{rank}(A_1E_1) = \text{rank}\left(\begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \\ F_{1,5} \\ F_{1,6} \end{bmatrix}\right)$$

$$= (a) n_{11} + n_{12} - n_{22} + t_1 + t_2 = r_1,$$  \hspace{1cm} (46)

where $(a)$ follows that these two $F_{1,3}$ are exactly the same matrix (linearly dependent). Then, the rank of $[A_1E_1 B_1E_2]$ in (45) can be written as

$$\text{rank}([A_1E_1 B_1E_2]) = \min\{n_{11}, n_{11} + n_{12} - n_{22} + t_1 + t_2 + r_2\}.$$ \hspace{1cm} (47)
The last term of the last equation in (12) becomes
\[ \text{rank}(B_1E_2) = \text{rank} \left( \begin{bmatrix} F_{2,1} \\ F_{2,2} \end{bmatrix} \right) \]
\[ = \min \{ n_{22} - n_{12} - t_1 - t_2, r_2 \}. \] (48)

Substituting (45)-(48) into (12) gives
\[ I(X_1; Y_1) = n_{11} + n_{12} - n_{22} + t_1 + t_2. \] (49)

For user 2, substituting \( E_1 \) and \( E_2 \) into (13) gives
\[ \text{rank}([A_2E_1B_2E_1]) = \text{rank} \left( \begin{bmatrix} 0^{n_{11}-n_{22},r_2} \\ 0^{n_{12}-n_{21},r_2} \\ 0^{t_1,r_2} \\ F_{2,1} \\ 0^{t_2,r_2} \\ F_{2,2} \end{bmatrix} \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \\ 0^{2n_{22}+n_{11}-2n_{12}-2n_{21}-t_2,r_1} \\ F_{1,5} \\ F_{1,4} \\ F_{1,3} \end{bmatrix} \right) \]
\[ = \min \{ n_{22} + n_{21} - n_{12} - t_1 - t_2, r_2 + r_1 \}. \] (50)

Similar to user 1, we design the generator matrices such that \( P2. \) holds. As a result, the rank of \([A_2E_2B_2E_1]\) in (50) can be written as
\[ \text{rank}([A_2E_2B_2E_1]) = \min \{ n_{22} + n_{21} - n_{12} - t_1 - t_2, 2(n_{22} - n_{12} - t_1 - t_2) + r_1 \}. \] (51)
The last term of the last equation in (13) becomes

\[
\text{rank}(B_2E_1) = \text{rank} \begin{pmatrix} \mathbf{F}_{1,1} \\ \mathbf{F}_{1,2} \\ \mathbf{F}_{1,3} \\ \mathbf{F}_{1,4} \\ \mathbf{F}_{1,5} \\ \mathbf{F}_{1,3} \end{pmatrix} = \min\{n_{21} + n_{12} - n_{22} + t_1 + t_2, r_1\}. \tag{52}
\]

Hence, user 2’s rate is obtained as

\[I(X_2; Y_2) = 2(n_{22} - n_{12} - t_1 - t_2). \tag{53}\]

The achievable rate pair is \((n_{11} + n_{12} - n_{22} + t_1 + t_2, 2(n_{22} - n_{12} - t_1 - t_2))\), which lies in between \((n_{11}, 0)\) and \((n_{11} + n_{12} - n_{22}, 2(n_{22} - n_{12}))\).

**Remark 2.** The only difference between a type I scheme and a type II scheme is that in a type II scheme, there are some submatrices associated with two different signal levels. For this kind of submatrix, it aligns with one of the non-zero submatrix from the other user on one signal level and aligns with an all-zero submatrix on another signal level.\(^3\) For example, \([\mathbf{F}_{1,3} \mathbf{F}_{2,1}]\) is in a higher signal level than \([\mathbf{F}_{1,3} 0]\) in (45). This introduces a correlation between two different signal levels associated with that submatrix. The main purpose is to strike a balance between maximizing the rank of the desired signals and minimizing the rank of the interference. It is also worth pointing out that since the two correlated submatrices are in fact the same matrices, the one aligned with a non-zero submatrix does not contribute to the rank of \(E_1\) or \(E_2\). In this regard, the calculation of the rank for \([\mathbf{A}_1 \mathbf{E}_1 \mathbf{B}_2 \mathbf{E}_2]\) or \([\mathbf{A}_2 \mathbf{E}_2 \mathbf{B}_2 \mathbf{E}_1]\) under a type II scheme can be made equivalently to that of a type I scheme as if there are no aligned submatrices. This, with property **P2.** induced by our design, guarantees that the conditions (30) and (31) still hold for all type II schemes.

\(^3\)For the ease of presentation, we use the term “matrix alignment” to represent that two submatrices are in the same signal level in the D-IC. The reader should not confuse this matrix alignment with any form of interference alignment \([29], [30]\).
2b) To achieve the region between \((n_{11} + n_{12} - n_{22}, 2(n_{22} - n_{12}))\) and \((2(n_{11} - n_{12}), n_{22} + n_{12} - n_{11})\), we propose

\[
E_4 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
F_{1,3} \\
0^{t_3, r_1} \\
F_{1,4} \\
0^{t_4, r_1} \\
F_{1,5} \\
F_{1,6} \\
0^{t_5, r_1} \\
F_{1,7} \\
F_{1,8} \\
F_{1,9}
\end{bmatrix} ,
E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{12} - n_{21} - t_3, r_2} \\
F_{2,2} \\
F_{2,3} \\
F_{2,4} \\
F_{2,5} \\
0^{2n_{21} + n_{12} - n_{11} - n_{22} - t_4, r_2} \\
0^{n_{11} + n_{22} - 2n_{12} - 2n_{21}, r_2} \\
F_{2,6} \\
0^{n_{12} - n_{21} - t_5, r_2} \\
F_{2,5} \\
0^{2n_{21} + n_{12} - n_{11} - n_{22} - t_4, r_2} \\
F_{2,7} \\
F_{2,8} \\
F_{2,9} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix} ,
\]

where \(F_{1,1} \in F_2^{t_4, r_1}, F_{1,2}, F_{1,5}, F_{1,9} \in F_2^{2n_{21} + n_{12} - n_{11} - n_{22} - t_4, r_1}, F_{1,3}, F_{1,6} \in F_2^{2n_{11} + n_{22} - 2n_{12} - 2n_{21}, r_1}, F_{1,4} \in F_2^{n_{12} - n_{21} - t_3, r_1}, F_{1,7} \in F_2^{n_{11} - n_{21} - t_5, r_1}, F_{1,8} \in F_2^{t_4, r_1}, F_{2,1} \in F_2^{t_4, r_2}, F_{2,2}, F_{2,5}, F_{2,8} \in F_2^{t_5, r_2}, F_{2,3}, F_{2,9} \in F_2^{2n_{21} + n_{12} - n_{11} - n_{22} - t_4, r_2}, F_{2,4}, F_{2,7} \in F_2^{2n_{22} + n_{11} - 2n_{12} - 2n_{21}, r_2}, F_{2,6} \in F_2^{t_5, r_2}, \) and \(t_3, t_5 \in [0 : n_{12} - n_{21}], t_4 \in [0 : 2n_{21} + n_{12} - n_{11} - n_{22}].\) Here, the value of \(t_5\) depends on the value of \(t_3,\) i.e., \(t_5 = 0\) when \(t_3 < n_{12} - n_{21}\) and \(t_5\) can take any value from \([0 : n_{12} - n_{21}]\) when \(t_3 = n_{12} - n_{21}.\)
For user 1, note that
\[
\text{rank}
\left[
\begin{array}{cccc}
F_{1,1} & 0^{n_{11}-n_{12},r_2} \\
F_{1,2} & F_{2,1} \\
F_{1,3} & 0^{n_{11}-n_{21},r_2} \\
o^{4,r_1} & F_{2,2} \\
F_{1,4} & F_{2,3} \\
o^{4,r_1} & F_{2,4} \\
F_{1,5} & 0^{n_{11}-n_{21}-t_3,r_2} \\
F_{1,6} & F_{2,5} \\
o^{4,r_1} & 0^{n_{11}-n_{12}-t_4,r_2} \\
F_{1,7} & F_{2,6} \\
F_{1,8} & 0^{n_{11}-n_{21}-t_5,r_2} \\
F_{1,9} & F_{2,7} \\
o^{2,r_1} & 0^{n_{11}-n_{21}-t_6,r_2}
\end{array}
\right]
\right) = \text{rank}
\left[
\begin{array}{cccc}
F_{2,1} & 0^{n_{11}-n_{12},r_2} \\
F_{2,2} & F_{2,1} \\
F_{2,3} & 0^{n_{11}-n_{21},r_2} \\
F_{2,4} & F_{2,2} \\
F_{2,5} & F_{2,3} \\
F_{2,6} & F_{2,4} \\
F_{2,7} & F_{2,5} \\
F_{2,8} & 0^{n_{11}-n_{21}-t_4,r_2} \\
F_{2,9} & F_{2,7} \\
o^{2,r_1} & 0^{n_{11}-n_{21}-t_6,r_2}
\end{array}
\right].
\tag{55}
\]  

Note that it is easy to verify that the above rank is equal to the rank of the above matrix with the upper \(F_{1,5}\) and the lower \(F_{2,5}\) replaced by 0. Moreover, we have that
\[
\text{rank}(B_1E_2) = \text{rank}
\left[
\begin{array}{cccc}
F_{2,1} & 0^{n_{11}-n_{12},r_2} \\
F_{2,2} & F_{2,1} \\
F_{2,3} & 0^{n_{11}-n_{21},r_2} \\
F_{2,4} & F_{2,2} \\
F_{2,5} & F_{2,3} \\
F_{2,6} & F_{2,4} \\
F_{2,7} & F_{2,5} \\
F_{2,8} & 0^{n_{11}-n_{21}-t_4,r_2} \\
F_{2,9} & F_{2,7} \\
o^{2,r_1} & 0^{n_{11}-n_{21}-t_6,r_2}
\end{array}
\right] = \text{rank}
\left[
\begin{array}{cccc}
F_{2,1} & 0^{n_{11}-n_{12},r_2} \\
F_{2,2} & F_{2,1} \\
F_{2,3} & 0^{n_{11}-n_{21},r_2} \\
F_{2,4} & F_{2,2} \\
F_{2,5} & F_{2,3} \\
F_{2,6} & F_{2,4} \\
F_{2,7} & F_{2,5} \\
F_{2,8} & 0^{n_{11}-n_{21}-t_4,r_2} \\
F_{2,9} & F_{2,7} \\
o^{2,r_1} & 0^{n_{11}-n_{21}-t_6,r_2}
\end{array}
\right].
\tag{56}
\]

These indicate that evaluating the rate of this type II scheme is equivalent to evaluate a corresponding type I scheme with the upper \(F_{1,5}\) and the lower \(F_{2,5}\) replaced by 0. Hence, we can again use the property \(P2\) to obtain user 1’s achievable rate by using (30) as
\[
I(X_1;Y_1) = \text{rank}(A_1E_1) = n_{11} + n_{12} - n_{22} - t_3 - t_4 - t_5.
\tag{57}
\]

This can be easily seen by the fact that Gaussian elimination does not alter the rank. However, we opt not to use the term “Gaussian elimination” deliberately to avoid causing the confusion that we are doing SIC, which we do not.
For user 2, we notice that

\[
\begin{pmatrix}
0^{n_{11} - n_{22}, r_2} \\
F_{2,1} \\
0^{n_{12} - n_{21} - t_3, r_2} \\
F_{2,2} \\
F_{2,3} \\
F_{2,4} \\
F_{2,5} \\
0^{2n_{21} + n_{12} - n_{11} - n_{22} - t_4, r_2} \\
F_{2,6} \\
0^{n_{12} - n_{21} - t_5, r_2} \\
F_{2,7} \\
0^{n_{21} + n_{12} - n_{11} - n_{22} - t_4, r_2} \\
F_{2,8} \\
0^{2n_{22} + n_{11} - 2n_{12} - 2n_{21}, r_1} \\
F_{2,9} \\
F_{1,1} \\
F_{1,2} \\
0^{t_4, r_1} \\
F_{1,4} \\
0^{t_5, r_1} \\
F_{1,5} \\
0^{t_4, r_1} \\
F_{1,5}
\end{pmatrix}
\]

\[
\text{rank}([A_2E_2 \ B_2E_1]) = \text{rank}
\]

Moreover, the dependence of \(t_5\) on \(t_3\) ensures that \(F_{2,6}\) and \(F_{1,4}\) are disjoint while more integer rate pairs between the two neighboring corner points can be achieved. Similar to user 1, this with property \(P2\) allows us to obtain user 2’s achievable rate by using (31) as

\[
I(X_2; Y_2) = \begin{cases} 
2(n_{22} - n_{12}) + t_3 + t_4, & t_5 = 0, \ t_3 < n_{12} - n_{21} \\
2n_{22} - n_{21} - n_{12} + t_5 + t_4, & t_5 \geq 0, \ t_3 = n_{12} - n_{21}
\end{cases}
\] (59)
2c) To achieve the region between \((2(n_{11} - n_{12}), n_{22} + n_{12} - n_{11})\) and \((0, n_{22})\), we propose

\[
E_1 = \begin{bmatrix}
0^{6,r_1} \\
F_{1,1} \\
0^{7,r_1} \\
F_{1,2} \\
0^{n_{22} - n_{21}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
F_{2,4} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
\] (60)

where \(F_{1,1}, F_{1,4} \in \mathbb{F}_2^{2n_{21}+n_{12} - n_{11} - n_{22}, r_1}\), \(F_{1,2}, F_{1,3} \in \mathbb{F}_2^{2n_{11} + n_{22} - 2n_{12} - 2n_{21}, r_1}\), \(F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{21}, r_5}\), \(F_{2,2}, F_{2,5} \in \mathbb{F}_2^{6, r_2}\), \(F_{2,3}, F_{2,4} \in \mathbb{F}_2^{2n_{11} + n_{22} - n_{11} - n_{22}, r_2}\), \(F_{2,4} \in \mathbb{F}_2^{7, r_2}\), \(F_{2,6} \in \mathbb{F}_2^{n_{12} - n_{21}, r_2}\), \(F_{2,7} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2}\), and \(t_6 \in [0 : 2n_{21} + n_{12} - n_{11} - n_{22}], t_7 \in [0 : 2n_{11} + n_{22} - 2n_{12} - 2n_{21}]\).

For user 1, we design the generator matrices such that \(P2\) holds. Moreover, by noting that

\[
\text{rank}([A_1 E_1 B_1 E_2]) = \text{rank} \left( \begin{bmatrix}
0^{6,r_1} \\
F_{1,1} \\
0^{7,r_1} \\
F_{1,2} \\
0^{n_{22} - n_{21}, r_1}
\end{bmatrix} \right) = \text{rank} \left( \begin{bmatrix}
0^{n_{11} - n_{12}, r_2} \\
0^{2n_{21}+n_{12} - n_{11} - n_{22}, r_1} \\
0^{7,r_1} \\
F_{1,3} \\
0^{6,r_1} \\
F_{1,4} \\
0^{n_{12} - n_{21}, r_1}
\end{bmatrix} \right) \] (61)

Note that the above rank is equal to the rank of the above matrix with the lower \(F_{2,3}\) replaced by 0. User 1’s achievable rate is then obtained as

\[
I(X_1; Y_1) = \text{rank}(A_1 E_1) = 2(n_{11} - n_{12} - t_6 - t_7).
\] (62)
For user 2, we design the generator matrices to ensure that \( P_2 \) also holds. Moreover, we note that
\[
\text{rank}(\begin{bmatrix} A_2 E_2 & B_2 E_1 \end{bmatrix}) = \text{rank}\left( \begin{bmatrix} 0^{n_{11}-n_{22},r_2} & 0^{n_{11}-n_{21},r_1} \\ F_{2,1} & 0^{t_6,r_1} \\ F_{2,2} & F_{1,1} \\ F_{2,3} & 0^{t_7,r_1} \\ F_{2,4} & 0^{n_{22}-n_{21},r_1} \\ 0^{2n_{12}-2n_{21}-2n_{22}-t_7,r_2} & F_{1,2} \\ F_{2,5} & 0^{n_{21}+n_{12}-n_{11}-n_{22},r_1} \\ F_{2,6} & F_{2,7} \end{bmatrix} \right),
\]
(63)

Similar to user 1, user 2’s achievable rate is obtained as
\[
I(X_2; Y_2) = \text{rank}(A_2 E_2) = n_{22} + n_{12} - n_{11} + t_6 + t_7.
\]
(64)

In the interest of space, for the other cases, we focus on achieving the non-trivial corner points on the capacity region (all the corner points exclude \((n_{11}, 0), (0, n_{22})\) since they correspond to the single user channel capacity without interference) and provide our choices of \( E_1 \) and \( E_2 \). We note that once all the corner points are achieved, the capacity region between any two neighbouring corner points can be achieved by time sharing.

The rest of the cases are discussed in Appendix A. The details of subcases and a pointer to the proof of each subcase are provided in Table I.

### IV. The Real-Valued Gaussian Interference Channel

In this section, we propose purely discrete input distributions for the real-valued G-IC by systematically translating the schemes for the D-IC into the G-IC. The constant-gap optimality of the proposed discrete input distributions is then shown.

#### A. Proposed Scheme and the Main Result

Recall that we have \( n_{11} = \lfloor \frac{1}{2} \log_2 \text{SNR}_1 \rfloor, n_{12} = \lfloor \frac{1}{2} \log_2 \text{INR}_1 \rfloor, n_{22} = \lfloor \frac{1}{2} \log_2 \text{SNR}_2 \rfloor, n_{21} = \lfloor \frac{1}{2} \log_2 \text{INR}_2 \rfloor \) and \( q = \max\{n_{11}, n_{12}, n_{22}, n_{21}\} \) according to the D-IC setting. Then, we define
### TABLE I
INTERFERENCE REGIMES AND SCHEMES

| Interference regimes | Subregimes | Scheme |
|----------------------|------------|--------|
| Weak: \( n_{12} < n_{22}, n_{21} < n_{11} \) | Weak 1: \( n_{11} > n_{22} > n_{12} > n_{21} \) | Appendix A-A |
|                     | Weak 2: \( n_{11} > n_{22} > n_{21} > n_{12} \) | Appendix A-B |
|                     | Weak 3: \( n_{11} > n_{21} > n_{22} > n_{12} \) | Appendix A-C |
| Strong: \( n_{11} < n_{21}, n_{22} < n_{12} \) | Strong 1: \( n_{12} > n_{21} > n_{11} > n_{22} \) | Appendix A-D |
|                     | Strong 2: \( n_{21} > n_{12} > n_{11} > n_{22} \) | Appendix A-E |
|                     | Strong 3: \( n_{21} > n_{11} > n_{12} > n_{22} \) | Appendix A-F |
| Mixed: \( n_{12} < n_{22}, n_{21} > n_{11} \) or \( n_{12} > n_{22}, n_{21} < n_{11} \) | Mixed 1: \( n_{11} > n_{12} > n_{22} > n_{21} \) | Appendix A-G |
|                     | Mixed 2: \( n_{11} > n_{21} > n_{12} > n_{22} \) | Appendix A-H |
|                     | Mixed 3: \( n_{11} > n_{12} > n_{21} > n_{22} \) | Appendix A-I |
|                     | Mixed 4: \( n_{12} > n_{11} > n_{21} > n_{22} \) | Appendix A-J |
|                     | Mixed 5: \( n_{12} > n_{11} > n_{22} > n_{21} \) | Appendix A-K |
|                     | Mixed 6: \( n_{21} > n_{11} > n_{22} > n_{12} \) | Appendix A-L |

The difference between the actual channel value and the corresponding quantized value as

\[
\begin{align*}
\beta_{11} & \triangleq \frac{1}{2} \log_2 \text{SNR}_1 - n_{11}, \\
\beta_{12} & \triangleq \frac{1}{2} \log_2 \text{INR}_1 - n_{12}, \\
\beta_{22} & \triangleq \frac{1}{2} \log_2 \text{SNR}_2 - n_{22}, \\
\beta_{21} & \triangleq \frac{1}{2} \log_2 \text{INR}_2 - n_{21},
\end{align*}
\]

where \( \beta_{11}, \beta_{12}, \beta_{22}, \beta_{21} \in [0, 1] \). The proposed distributions are systematically translated from the proposed distributions in D-IC into a multi-layer superposition PAM signaling where each PAM’s power level and cardinality can be directly derived from \( E_1 \) and \( E_2 \) in our proposed scheme for the D-IC. Specifically, for \( k \in \{1, 2\} \), user \( k \)'s signal is

\[
X_k = \gamma_k \sum_{i_k=1}^{L_k} 2^{\text{row}(E_k[\ldots,F_{k,i_k}^T]^T)} \rho_{k,i_k} F_{k,i_k},
\]

Here, \( \gamma_k \) are normalization factors to ensure the power constraint \( \mathbb{E}[\|X_k\|^2] \leq 1 \). In this section, we particularly choose \( \gamma_k = 2^{-q} \), with which the power constraints are satisfied as shown in Lemma 1 in Appendix B. In Sec. IV-B we will show that this choice of \( \gamma_k \) is sufficient to prove the constant-gap optimality of our scheme. In practice and in our simulation, larger normalization factors can be chosen such that we “top up” the power constraint \( \mathbb{E}[\|X_k\|^2] = 1 \). \( L_k \) is the number
of non-zero submatrices in $E_k$. $\text{row}(.)$ outputs the number of rows. $F_{k,i_k}$ is a random variable associated with $F_{k,i_k}$ with $i_k$ being the index of the non-zero submatrix in $E_k$. $E_k \setminus [\ldots, F_{k,i_k}^T]^T$ represents the submatrix of $E_k$ containing all the rows below $F_{k,i_k}$ (in other words, removing all the rows above and including $F_{k,i_k}$). We refer to $2^{\text{row}(E_k \setminus [\ldots, F_{k,i_k}^T]^T)}$ as the power level associated with $F_{k,i_k}$ and $\rho_{k,i_k} \in [1, 2)$ is the power scaling factor for $F_{k,i_k}$. The parameter $\rho_{k,i_k}$ is introduced to align the power of the desired and that of the interference signal at a receiver when their corresponding matrix blocks in the D-IC are at the same signal level. Note that this situation appears in the type II scheme only.

In a type I scheme, $F_{k,i_k}$ is uniformly distributed over $\text{PAM}(2^{\text{rank}(F_{k,i_k}) - 1}, 1)$ and $\rho_{k,i_k} = 1$ (i.e., no power scaling). Each PAM is associated with one power level. In a type II scheme, $F_{k,i_k}$ is uniformly distributed over $\text{PAM}(2^{\text{rank}(F_{k,i_k}) - 2}, 1)$. Notice that there are some PAM signals that are associated with two power levels in type II schemes. These PAM signals correspond to the same submatrices and thus are identical signal but with two different power levels. As discussed in Remark 2, there are two submatrices align in one signal level while the submatrix with two power levels does not align with any other non-zero submatrix in a lower or higher signal level. However, since channel gains are not necessarily powers of 2, we introduce the adjustment constant to enforce perfect alignment for those signal levels. Specifically, when the aligned submatrix is at the higher level from receiver 1’s perspective in the D-IC (e.g., $[F_{1,3} \ F_{2,1}]$ is at a higher signal level than $[F_{1,3} \ 0]$ in (45)), the power level for $F_{k,i_k}$ is slightly adjusted by applying the following power scaling factor

$$\rho_{k,i_k} = 2^{\max\{\beta_{11}, \beta_{12}\} - \beta_{1k}}.$$  \hspace{1cm} (70)

On the other hand, when the aligned submatrix is at the higher level from the receiver 2’s perspective (e.g., $[F_{2,5} \ F_{1,1}]$ is at a higher level than $[F_{2,5} \ 0]$ in (58)), the power level of $F_{k,i_k}$ is slightly adjusted by the power scaling factor

$$\rho_{k,i_k} = 2^{\max\{\beta_{21}, \beta_{22}\} - \beta_{2k}}.$$  \hspace{1cm} (71)

Otherwise, $\rho_{k,i_k} = 1$.

Here, the reduction on the order of the cardinality of a PAM signal and the introduction of the power scaling factors are to reduce the extra interference between each PAM signals caused by the mismatch between the expected channel gains (powers of 2) and the actual channel gains. The reason of this will soon become clear in the minimum distance analysis later.
Remark 3. In our scheme, we translate each non-zero submatrix in $E_k$ into a PAM constellation. Note that each submatrix can be further partitioned into a number of submatrices and can be translated into the superposition of multiple independent PAM constellations. In the extreme case, each non-zero row could be translated into a binary phase shift keying signal. However, for each independent PAM modulation, we have to added a one-bit (two-bit for type II schemes) guard interval as mentioned above. Hence, in our translation, we tend to keep the number of independent PAM modulations, i.e., $L_k$, small. The largest number of $L_k$ among all of our schemes for the D-IC in Sec. III-C and Appendix A is 10.

We state the main result of this section in the following.

Theorem 2. There exist a pair of purely discrete input distributions $(X_1, X_2)$ such that the whole capacity region of the two-user real-valued G-IC can be achieved to within a constant gap by using TIN, where the gap is independent of all channel parameters.

In the next section, we provide the proof for Theorem 2.

B. Proof of Theorem 2

It has been shown in [26] that the capacity region of the D-IC $C_{D-IC}$ and that of the G-IC $C_{G-IC}$ satisfy $C_{G-IC} \subseteq C_{D-IC} + c$ for some constant $c > 0$ for both the real- and complex-valued G-IC. In what follows, we will show that the rate region achieved by our discrete input distribution given in (69) with TIN satisfies $C_{D-IC} \subseteq R_{TIN}^{G-IC} + c'$ for some constant $c' > 0$. Hence, the entire capacity region of real-valued G-IC can be achieved by our scheme to within a constant gap, i.e., $C_{G-IC} \subseteq R_{TIN}^{G-IC} + c''$ for some constant $c'' > 0$.

We first analyze the achievable rate of user 1 under our scheme with discrete signaling and TIN. User 2’s achievable rate can be similarly analyzed by following the same line of proof for user 1 and the results are deferred to Sec. IV-B3.

First, note that user 1’s mutual information is

$$I(X_1; Y_1) = h(Y_1) - h(Y_1 | X_1)$$

$$= h(h_{11}X_1 + h_{12}X_2 + Z_1) - h(h_{12}X_2 + Z_1)$$

$$= h(h_{11}X_1 + h_{12}X_2 + Z_1) - h(Z_1) - (h(h_{12}X_2 + Z_1) - h(Z_1))$$

$$= I(h_{11}X_1 + h_{12}X_2; h_{11}X_1 + h_{12}X_2 + Z_1) - I(h_{12}X_2; h_{12}X_2 + Z_1).$$  (72)
To bound $I(h_{11}X_1 + h_{12}X_2; h_{11}X_1 + h_{12}X_2 + Z_1)$, we note that

$$h_{11}X_1 + h_{12}X_2 = \sqrt{\text{SNR}_1}X_1 + \sqrt{\text{INR}_1}X_2$$

$$= 2^{n_{11} + \beta_{11} - q} \sum_{i=1}^{L_1} 2^{\text{row}(E_i \setminus \ldots \set{T})} \rho_{1,i} F_{1,i} + 2^{n_{12} + \beta_{12} - q} \sum_{j=1}^{L_2} 2^{\text{row}(E_j \setminus \ldots \set{T})} \rho_{2,j} F_{2,j}$$

$$= X_A + X_B,$$  \hspace{1cm} (73)

where (a) follows from the definitions of

$$X_A = \sum_{k=1}^{2} 2^{n_{1k} + \beta_{1k} - q} \sum_{i_k \in A_k} 2^{\text{row}(E_k \setminus \ldots \set{T})} \rho_{k,i_k} F_{k,i_k},$$  \hspace{1cm} (74)

$$A_k \triangleq \{i_k | n_{1k} - q + \text{row}(E_k \setminus \ldots \set{T}) \geq 0\},$$  \hspace{1cm} (75)

$$X_B = \sum_{k=1}^{2} 2^{n_{1k} + \beta_{1k} - q} \sum_{i_k \in B_k} 2^{\text{row}(E_k \setminus \ldots \set{T})} \rho_{k,i_k} F_{k,i_k},$$  \hspace{1cm} (76)

$$B_k \triangleq \{i_k | n_{1k} - q + \text{row}(E_k \setminus \ldots \set{T}) < 0\},$$  \hspace{1cm} (77)

where $X_A$ and $X_B$ represent the superpositions of all signals above and below the noise level, respectively, from receiver 1’s perspective. Note that $|A_k| \leq L_1$ and $|A_2| \leq L_2$. In what follows, we give an example to show the signals above and below the noise level for a type II scheme.

**Example 1.** Consider the case in Sec. III-C2(a). In this case, $\gamma_1 = \gamma_2 = 2^{-n_{11}}$. We assume $\beta_{11} > \beta_{12}$.

*From receiver 1’s perspective, the signal above the noise level is*

$$X_A = 2^{\beta_{11}} (F_{1,6} + 2^{n_{12} + n_{21} - n_{22} - t_1} F_{1,5} + 2^{n_{22} + n_{11} - n_{12} - n_{21} - t_2} F_{1,4} + (2^{n_{11} - n_{21}} + 2^{n_{22} + n_{11} - n_{12} - n_{21}}) F_{1,3} + 2^{n_{21} - t_1} F_{1,2} + 2^{n_{21}} F_{1,1}) + 2^{\beta_{12}} (2^{n_{12} + n_{21} - n_{22}} F_{2,2} + 2^{n_{22} + n_{11} - n_{12} - n_{21} + \beta_{11} - \beta_{12}} F_{2,1}),$$  \hspace{1cm} (78)

where $\rho_{2,1} = 2^{\beta_{11} - \beta_{12}}$ is the power scaling factor for $F_{2,1}$ and the rest of the power scaling factors are 1.

*The signal below the noise level is*

$$X_B = 2^{\beta_{12}} (2^{n_{12} - n_{22}} F_{2,4} + 2^{(n_{12} + n_{21} - n_{22}) - n_{11}} F_{2,3}).$$  \hspace{1cm} (79)
With (73), the mutual information \( I(h_{11}X_1 + h_{12}X_2; h_{11}X_1 + h_{12}X_2 + Z_1) \) can be bounded by

\[
I(h_{11}X_1 + h_{12}X_2; h_{11}X_1 + h_{12}X_2 + Z_1) = h(X_A + X_B + Z_1) - h(Z_1)
\]

\[
\geq h(X_A + Z_1) - h(Z_1)
\]

\[
\geq I(X_A; X_A + Z_1)
\]

\[
\geq H(X_A) - \frac{1}{2} \log_2 2\pi e \left( \frac{1}{d_{\text{min}}^2(X_A)} + \frac{1}{12} \right),
\]

(80)

where we note that the lower bound in (b) does not result in too much loss as \( X_B \) is already under the noise level as we will show in the following; (c) follows from an Ozarow-type bound \[28\] for the achievable rate of a uniform input distribution over a one-dimensional constellation in \[23\] Prop. 1.

**Remark 4.** Although there exist better bounds for the mutual information for discrete inputs (e.g., in \[23\], \[31\]), we opt to use a type of Ozarow-Wyner bound \[28\] due to its simplicity for enabling closed-form analytical computation (see also \[23\] Sec. II-A).

To bound \( I(h_{12}X_2; h_{12}X_2 + Z_1) \) from (73), we note that \( h_{12}X_2 = X_{A,2} + X_{B,2} \), where \( X_{A,2} = 2^{n_{12} + \beta_{12} - q} \sum_{i_2 \in A_2} 2^{\text{row}(E_2[,F_{2,i_2}^T])} \rho_{2,i_2} \mathcal{F}_{2,i_2} \) represents user 2’s signal above the noise level and \( X_{B,2} = 2^{n_{12} + \beta_{12} - q} \sum_{i_2 \in B_2} 2^{\text{row}(E_2[,F_{2,i_2}^T])} \rho_{2,i_2} \mathcal{F}_{2,i_2} \) is the part below the noise level according to (74)-(77). Hence,

\[
I(h_{12}X_2; h_{12}X_2 + Z_1) = h(h_{12}X_2 + Z_1) - h(X_{B,2} + Z_1) + h(X_{B,2} + Z_1) - h(Z_1)
\]

\[
= I(X_{A,2}; h_{12}X_2 + Z_1) + I(X_{B,2}; X_{B,2} + Z_1)
\]

\[
\leq H(X_{A,2}) - \frac{1}{2} \log_2 (1 + 2\mathbb{E} ||X_{B,2}||^2)]
\]

\[
= H(X_{A,2}) - \frac{1}{2} \log_2 (1 + 2\mathbb{E} 2^{n_{12} + \beta_{12} - q} \sum_{j \in B_2} 2^{\text{row}(E_2[,F_{j}^T])} \rho_{2,j} \mathcal{F}_{2,j} ||^2)]
\]

\[
\leq H(X_{A,2}) - \frac{1}{2} \log_2 (1 + 2^{\beta_{12}} \mathbb{E} ||X_2||^2)
\]

\[
\leq H(X_{A,2}) - \frac{1}{2} \log_2 3,
\]

(81)

where \( \frac{1}{2} \log_2 3 \) is the maximum number of loss bits due to our way of characterizing the signals below the noise level.

Substituting (80) and (81) into (72) gives

\[
I(X_1; Y_1) \geq H(X_A) - H(X_{A,2}) - \frac{1}{2} \log_2 2\pi e \left( \frac{1}{d_{\text{min}}^2(X_A)} + \frac{1}{12} \right) - \frac{1}{2} \log_2 3.
\]

(82)
In what follows, we analyze the cardinality and the minimum distance of $X_A$. We present the detailed analysis for our proposed two types of schemes separately. This together with the converse bound in [6] will allow us to complete the proof of Theorem 2.

Before proceeding, we define the inter-constellation distance, which will allow us to introduce the concept of constellation interception.

**Definition 1.** Consider a superimposed constellation $P_1A_1 + P_2A_2$ with $P_1, P_2 \in \mathbb{R}$ and $P_1 > P_2 > P_1 > 0, P_1, P_2 \in \mathbb{R}$. Let $\lambda_1, \lambda_2 \in \Lambda_2$ and $\lambda_1 > \lambda_2$. The inter-constellation distance is defined as the minimum of the set of distances from the leftmost constellation point of sub-constellation $P_1A_1 + P_2\lambda_1$ to the rightmost constellation point of sub-constellation $P_1A_1 + P_2\lambda_2$. It is computed as

$$d_{IC}(P_1A_1 + P_2A_2) \triangleq \min_{\lambda_1, \lambda_2 \in \Lambda_2} \{\min\{P_1A_1 + P_2\lambda_1\} - \max\{P_1A_1 + P_2\lambda_2\}\}.$$ 

An illustration for the distance $\min\{P_1A_1 + P_2\lambda_1\} - \max\{P_1A_1 + P_2\lambda_2\}$ is shown in Fig. 1.

With $d_{IC}$, we can define whether the two constellations intercept or not. Specifically, when $d_{IC}(P_1A_1 + P_2\lambda_2) < 0$, we say the sub-constellation $P_1A_1 + P_2\lambda_1$ intercepts with the sub-constellation $P_1A_1 + P_2\lambda_2$. In the example shown in Fig. 1, it is clearly that $d_{IC}(P_1A_1 + P_2\lambda_2) > 0$ and thus the sub-constellation $P_1A_1 + P_2\lambda_1$ does not intercept with $P_1A_1 + P_2\lambda_2$ for all $\lambda_1, \lambda_2 \in \Lambda_2$.

1) : When type 1 scheme is used, $X_A$ in (74) can be written into the following format in ascending order based on the power level

$$X_A = \sum_{l=1}^{L} P_lV_l \in \sum_{l=1}^{L} 2^{\sum_{i=1}^{l} a_i + m_{l-1} + \beta_i} \Lambda_l \triangleq \Lambda_{\Sigma}, \quad (83)$$

where $V_l$ is uniformly distributed over a uniform PAM$(2^{m_l-1}, 1)$, representing either $F_{1,i_1}$ or $F_{2,i_2}$ with $i_1 \in A_1, i_2 \in A_2$ according to (75). $m_l$ is equal to either rank($F_{1,i_1}$) or rank($F_{2,i_2}$).
and $m_0 = 0$, $\rho_l = 1$ for type I scheme, $P_l \triangleq 2^\sum_{i=1}^{l} \alpha_i + m_{l-1} + \beta_l$ is the overall power coefficient including the power level and the channel gain, $\alpha_l$ is a non-negative integer, representing the number of rows of the all-zero submatrix that is in the signal level between the submatrices associated with $V_l$ and $V_{l-1}$ in the D-IC, and $\alpha_1 = 0$ (because $V_1$ has the lowest power level), $\beta_l \in \{\beta_{11}, \beta_{12}\}$ is the channel difference associated with $V_l$, $L \in \mathbb{N}$ is the total number of PAM signals above the noise level, $\Lambda_\Sigma$ is the overall constellation.

We then have the following proposition for $X_A$.

**Proposition 1.** $X_A$ is uniformly distributed over the superimposed constellation $\Lambda_\Sigma$ defined in (83) satisfying

\[
\begin{align*}
  i) \quad |\Lambda_\Sigma| &= 2^{\sum_{l=1}^{L} m_l - L}, \\
  ii) \quad 1 \leq d_{\min}(\Lambda_\Sigma) < 2, \\
  iii) \quad 1 - 2^{\sum_{l=1}^{L} (m_l + \alpha_l) - 1} < \lambda < 2^{\sum_{l=1}^{L} (m_l + \alpha_l) - 1} - 1, \forall \lambda \in \Lambda_\Sigma.
\end{align*}
\]

**Proof:** See Appendix B-A. □

The intuition behind how the ‘−1’ on the cardinality allows us to guarantee non-vanishing minimum distance is illustrated in Fig. 2, where we visualize the desired signals (user 1’s signals) and interference (user 2’s signals) as well as their signal levels from the D-IC model. Specifically, each non-zero submatrix is represented by a grid where its color and pattern are used for distinguishing different users’ signals. The all-zero submatrices are left as blank. The position of each submatrix is determined by its signal level in the D-IC. As it can be seen in Fig. 2(a), when the channel gain of each signal is power of 2 with $\beta_l = 0$, the minimum distance is not vanished since the desired signal and the interference are disjoint. When the channel gains are not powers of 2 as shown in Fig. 2(b), the interference could be stronger than expected, which shrinks the minimum distance. In Fig. 2(c), the ‘−1’ bit on the cardinality serves as a guard interval to maintain a large minimum distance. Since the power spacing between the least significant bit in the desired signal and most significant bit in the interference signals is determined by the difference $\beta_{11} - \beta_{12} \in (-1, 1)$, adding a 1-bit guard interval suffices to guarantee a constant minimum distance even in the worst case.

With Proposition 1, (74), (75) and (82), user 1’s achievable rate under type I schemes (i.e.,
all the type I schemes in Appendix (A) and in Sec. (III-C1) is lower bounded by
\[
I(X_1; Y_1) \geq \sum_{k=1}^{2} \left( \sum_{i_k \in A_k} \text{rank}(F_{i_k}) - |A_k| \right) - \left( \sum_{i_2 \in A_2} \text{rank}(F_{i_2}) - |A_2| \right) - \frac{1}{2} \log_2 2\pi e \left( \frac{1}{d_{\text{min}}^2(X_A)} + \frac{1}{12} \right) - \frac{1}{2} \log_2 3
\]
\[
= \sum_{i_1 \in A_1} \text{rank}(F_{i_1}) - |A_1| - \frac{1}{2} \log_2 2\pi e \left( \frac{1}{d_{\text{min}}^2(X_A)} + \frac{1}{12} \right) - \frac{1}{2} \log_2 3
\]
\[
(d) \quad r_1 - |A_1| - \frac{1}{2} \log_2 2\pi e \left( \frac{13}{12} \right) - \frac{1}{2} \log_2 3,
\]
where \((d)\) follows that \(\sum_{i_1 \in A_1} \text{rank}(F_{i_1})\) corresponds to \(\text{rank}(A_1E_1)\) in the D-IC, which is equal to user 1’s target rate \(r_1\) due to property P2. in Remark 1, and \(|A_1|\) is a constant due to the \(-1\) on the cardinality of user 1’s signal above the noise level.

**Remark 5.** For some cases, it is possible to reduce the gap in (84) by using less guard bits. Consider \(\beta_{11} < \beta_{12}\) as an example. Let \(\Lambda_l = \text{PAM}(2^{m_l}, 1)\) and \(\Lambda_{l+1} = \text{PAM}(2^{m_{l+1}}, 1)\) be user 1 and user 2’s PAM signals, respectively, for some \(l \in [1 : L]\). According to (21b) from [23] Proposition 2], the minimum distance of the superposition of \(\Lambda_l\) and \(\Lambda_{l+1}\) satisfies
\[
d_{\text{min}}(2\sum_{i=1}^{l} \alpha_i + m_{l-1} + \beta_{11} \Lambda_l + 2\sum_{i=1}^{l+1} \alpha_i + m_{l-1} + \beta_{12} \Lambda_{l+1}) = 2\sum_{i=1}^{l} \alpha_i + m_{l-1} + \beta_{11} d_{\text{min}}(\Lambda_l)\]
without reducing \(m_l\). Moreover, even when \(\beta_{11} > \beta_{12}\) but with \(\alpha_{l+1} \geq 1\), the above condition is still satisfied. The introduction of the ‘\(-1\)’ to the cardinalities of all PAM signals is to universally lower bound
\(d_{\min}(\Lambda_\Sigma)\) by a constant regardless of the values of \(\beta_l\) and \(\alpha_l\), which significantly simplifies the proof.

2) : Following Sec. [IV-B1] when type II scheme is used, \(X_A\) in (74) can be written into the following format

\[
X_A = \sum_{i=1}^{L} P_i V_i + \sum_{l' \in \Phi} P_{l'} V_{l'}
\]

\[
\in \sum_{i=1}^{L} 2^{\sum_{i=1}^{L} \alpha_i + m_{i-1} + \beta_l} P_i \Lambda_i + \sum_{l' \in \Phi} 2^{\sum_{i=1}^{\phi} \alpha_i + m_{i-1} + \beta_{l'}} P_{l'} \Lambda_{l'} \triangleq \Lambda_\Sigma,
\]

(85)

where \(V_i\) is uniformly distributed over \(\Lambda_i\), a uniform PAM\((2^{m_{i-2}}, 1)\), representing user 1 or user 2’s signal above the noise level, i.e., either \(F_1, i_1\) or \(F_2, i_2\) with \(i_1 \in \mathcal{A}_1, i_2 \in \mathcal{A}_2\) according to (75) while \(V_1, \ldots, V_L\) are \(L\) distinct random variables and \(P_1, \ldots, P_L\) are \(L\) distinct constants and \(P_l\) is the power coefficient associated with \(V_l\). Here, there are some \(V_{l'}\), for \(l' \in \Phi \subseteq [1 : L]\) associated with two power coefficients \(P_{l'} \neq P_{l''}\), where \(\Phi\) is the collection of \(l'\) which varies from scheme to scheme, and \(\overline{\rho} \in [1 : L], \overline{\rho} \not\in \Phi\) is the index of the random variable \(V_{l'}\) whose associated submatrix is aligned with the submatrix associated with \(V_{l''}\) in the D-IC. \(\beta_{l'} \neq \beta_{l''} \in \{\beta_{11}, \beta_{12}\}\) are the channel differences associated with \(V_{l'}\) and \(V_{l''}\), respectively. The power scaling factors for \(V_{l'}\) and \(V_{l''}\) are \((\rho_{l'}, \rho_{l''}) = (2^{\max(\beta_{11}, \beta_{12})-\beta_{l'}}, 2^{\max(\beta_{11}, \beta_{12})-\beta_{l''}})\) when \(2^{\sum_{i=1}^{\phi} \alpha_i + m_{i-1}} > 2^{\sum_{i=1}^{\phi} \alpha_i + m_{i-1}}\) according to (70) and \((\rho_{l'}, \rho_{l''}) = (1, 2^{\max(\beta_{22}, \beta_{21})-\beta_{l'}})\) when \(2^{\sum_{i=1}^{\phi} \alpha_i + m_{i-1}} < 2^{\sum_{i=1}^{\phi} \alpha_i + m_{i-1}}\) according to (71). Otherwise, \(\rho_l = 1\) for \(V_l\). The power scaling factors are to ensure that both \(V_{l'}\) and \(V_{l''}\) still share the same power coefficients \(P_{l'}\) even when the channel gains are not powers of 2.

In Example 1 \(V_{l'} = F_{1,3}, P_{l'} = 2^{\beta_{11} + n_{11} - n_{21}}, \rho_{l'} = 1\) while \(V_{l''} = F_{2,1}, P_{l''} = 2^{\beta_{11} + n_{12} + n_{22} - n_{12} - n_{21}}, \rho_{l''} = 2^{\beta_{11} - \beta_{12}}\). Note that here \(P_{l'}\) is the power coefficient of \(F_{2,1}\) and also one of the power coefficient of \(F_{1,3}\).

We emphasize here that the submatrix alignments have two scenarios depending on the values of \(P_{l'}\) and \(P_{l''}\). To be specific, \(P_{l'} > P_{l''}\) corresponds to the scenario where the submatrix associated with two power levels is aligned with the other non-zero submatrix in its higher signal level and does not align to any non-zero submatrix in its lower signal level in the D-IC from receiver 1’s perspective. \(P_{l'} < P_{l''}\) corresponds to the scenario where the submatrix associated with two power levels is aligned with the other non-zero submatrix in its lower signal level and does not align to any non-zero submatrix in its higher signal level in the D-IC from receiver 1’s perspective.
For the ease of presentation, from now on we refer to the first and the second scenarios as high alignment and low alignment, respectively. Furthermore, we emphasize that within a scheme, the high alignment is only from one receiver’s perspective while it becomes low alignment from another receiver’s perspective (see (45) and (50) for the scheme with a high alignment at receiver 1 and a low alignment at receiver 2). Otherwise, if either high alignment or low alignment is from both receivers’ perspectives, it can be easily checked that this scheme can be transformed into a type I scheme.

In what follows, the minimum distance and the cardinality of a superimposed constellation with one low alignment and high alignment are analyzed in 2a) and 2b), respectively. Based on the results in 2a) and 2b), the minimum distance and the cardinality of $X_A$ for type II scheme with one matrix alignment and two matrix alignments are then analyzed in 2c) and 2d), respectively. The scenario in 2c) covers all type II schemes for achieving all the corner points on the capacity region of the D-IC; hence, it suffices the purpose of approaching the entire capacity region of the real-valued G-IC. In 2d), we showcase that the type II schemes for achieving integer rate points between two neighboring corner points on the capacity region of the D-IC, can approach the capacity region of the real-valued G-IC without time-sharing.

2a) For a type II scheme with a low alignment at receiver 1, a submatrix $F_{\Sigma,1}$ of $[A_1E_1 B_1E_2]$ that contains the aligned matrices has the following form

$$F_{\Sigma,1} = \begin{bmatrix} \Pi_{1,F}[F_1 \ 0^{m_1,r_1'}] \\ \Pi_{2,F}[F_2 \ 0^{m_2,r_2'}] \\ \Pi_{3,F}[F_3 \ 0^{m_3,r_3'}] \\ F_{\Sigma,2} \\ \Pi_{2,F}[F_2 F_4] \end{bmatrix}, \quad (86)$$

where for $l \in [1 : 3]$, $F_l \in F_2^{m_l,r_{kl'}}$, $F_4 \in F_2^{m_2,r_{k2'}}$ and $F_2$ is the submatrix associated with two power levels, $m_l \in \mathbb{Z}^+$, $r_{kl'} \neq r_{k2'} \in \{r_1, r_2\}$ and their values depend on the scheme. Specifically, $\Pi_{l,F}[F_a \ F_b]$ determines the positions of $F_a$ and $F_b$. If $\Pi_{l,F}[F_a \ F_b] = [F_a \ F_b]$ then the number of columns for $F_a$ and $F_b$ are $r_1$ and $r_2$, respectively, otherwise if $\Pi_{l,F}[F_a \ F_b] = [F_b \ F_a]$ then the number of columns for $F_a$ and $F_b$ are $r_2$ and $r_1$, respectively. $F_{\Sigma,2}$ is a matrix of size $m_0 \times (r_1 + r_2)$ (may contain some all-zero submatrices) representing all the signals above $F_4$ and below $F_3$, respectively. An example of a type II scheme with a low alignment at receiver 1 can be found in (61).
The superimposed signal corresponds to (86) ordered from the lowest power level to the highest power level is

\[ F_{\Sigma,1} = 2^{\beta_4} F_4 + 2^{m_2} F_{\Sigma,2} + 2^{m_2+m_0+\beta_3} F_3 + (2^{\beta_2} + 2^{m_2+m_0+m_3+\beta_2}) \rho_2 F_2 \]

\[ + 2^{m_0+m_3+2m_2+\beta_1} F_1 \]

\[ \in 2^{\beta_4} \Lambda_4 + 2^{m_2} \Lambda_{\Sigma,2} + 2^{m_2+m_0+\beta_3} \Lambda_3 + (2^{\beta_2} + 2^{m_2+m_0+m_3+\beta_2}) \rho_2 \Lambda_2 \]

\[ + 2^{m_0+m_3+2m_2+\beta_1} \Lambda_1 = \Lambda_{\Sigma,1}, \tag{87} \]

where for \( l \in [1 : 3], F_l \) is a random variable uniformly distributed over a uniform PAM \( \Lambda_l = PAM(2^{m_l-2}, 1) \), \( F_4 \) is a random variable uniformly distributed over a uniform PAM \( \Lambda_4 = PAM(2^{m_4-2}, 1) \), \( \beta_l, \beta_4 \in \{\beta_{11}, \beta_{12}\} \) are the channel differences associated with \( F_l \) and \( F_4 \), respectively, and their values depend on whether they belong to user 1 or user 2 and \( \beta_2 \neq \beta_4 \), the power scaling factors are \( \rho_1 = \rho_3 = \rho_4 = 1 \) and thus are omitted from the equation above while \( \rho_2 = 2^{\max\{\beta_{21}, \beta_{22}\} - \beta_2} \) according to (71) because a low alignment for \( F_2 \) at receiver 1 means a high alignment for \( F_2 \) at receiver 2. \( F_{\Sigma,2} \) is a random variable uniformly distributed over a superimposed constellation \( \Lambda_{\Sigma,2} \), where each user’s PAM signal in \( \Lambda_{\Sigma,2} \) has its own value of \( \beta \). We consider that \( 1 \leq d_{\min}(\Lambda_{\Sigma,2}) < 2 \) and for all \( \lambda \in \Lambda_{\Sigma,2} \), the condition \( 1 - 2^{m_0-1} \leq \lambda \leq 2^{m_0-1} - 1 \) holds for all values of \( \beta \) in \( \Lambda_{\Sigma,2} \). Then, we have the following proposition.

**Proposition 2.** For the random variable \( F_{\Sigma,1} \) and the superimposed constellation \( \Lambda_{\Sigma,1} \) defined in (87) with \( \Lambda_l \) and \( \Lambda_{\Sigma,2} \) being non-empty sets for \( l \in [1 : 4] \), \( F_{\Sigma,1} \) is uniformly distributed over \( \Lambda_{\Sigma,1} \) satisfying

i) \( |\Lambda_{\Sigma,1}| = |\Lambda_4| \cdot |\Lambda_{\Sigma,2}| \cdot |\Lambda_3| \cdot |\Lambda_2| \cdot |\Lambda_1| \)

ii) \( 1 \leq d_{\min}(\Lambda_{\Sigma,1}) < 2 \)

iii) \( 1 - 2^{m_0+m_3+2m_2+m_1-1} < \lambda < 2^{m_0+m_3+2m_2+m_1-1} - 1, \forall \lambda \in \Lambda_{\Sigma,1}. \)

**Proof:** We analyze the minimum distance for each layer of superposition. First, note that

\[ d_{\min}(2^{\beta_4} \Lambda_4 + 2^{m_2} \Lambda_{\Sigma,2}) = 2^{\beta_4}, \tag{88} \]

\[ d_{\min}(2^{m_2} \Lambda_{\Sigma,2} + 2^{m_2+m_0+\beta_3} \Lambda_3) = 2^{m_2} d_{\min}(\Lambda_{\Sigma,2}), \tag{89} \]

\[ d_{\min}(2^{m_2+m_0+\beta_3} \Lambda_3 + (2^{\beta_2} + 2^{m_2+m_0+m_3+\beta_2}) \rho_2 \Lambda_2) = 2^{m_2+m_0+\beta_3}, \tag{90} \]

\[ d_{\min}((2^{\beta_2} + 2^{m_2+m_0+m_3+\beta_2}) \rho_2 \Lambda_2 + 2^{m_0+m_3+2m_2+\beta_1} \Lambda_1) = (2^{\beta_2} + 2^{m_2+m_0+m_3+\beta_2}) \rho_2, \tag{91} \]
where (88) and (89) can be shown by applying Lemma 2 in Appendix B since the following conditions hold for $\Lambda_\Sigma,2$

$$\max\{2^{m_2}\Lambda_\Sigma,2\} < 2^{m_2}(2^{m_0-1} - 1),$$  \hfill (92)

$$\min\{2^{m_2}\Lambda_\Sigma,2\} > 2^{m_2}(1 - 2^{m_0-1}),$$  \hfill (93)

and (90) and (91) follow directly from [23] Prop. 2 as $\Lambda_1, \Lambda_2,$ and $\Lambda_3$ are uniform PAMs.

Then with (88)-(91) and by Lemma 3, we obtain that

$$d_{\min}(\Lambda_\Sigma,1) = 2^{\beta_4} \in [1, 2).$$  \hfill (94)

Furthermore, since there is no overlapping in $\Lambda_\Sigma,1$, the cardinality thus satisfies

$$|\Lambda_\Sigma,1| = |A_4| \cdot |\Lambda_\Sigma,2| \cdot |A_3| \cdot |A_2| \cdot |A_1|. $$  \hfill (95)

By considering the extreme case of $\beta_l = 1, l \in [1 : 4]$, we can obtain an upper on $\max\{\Lambda_\Sigma,1\}$ and a lower bound on $\min\{\Lambda_\Sigma,1\}$ as

$$\max\{\Lambda_\Sigma,1\} \leq 2\max\{A_4\} + \max\{2^{m_2}\Lambda_\Sigma,2\} + 2^{m_2+m_0+1}\max\{A_3\} + (2 + 2^{m_2+m_0+m_3+1})\max\{A_2\} + 2^{m_0+m_3+2m_2+1}\max\{A_1\}$$

$$< (2^{m_2-2} - 1) + 2^{m_2}(2^{m_0-1} - 1) + 2^{m_2+m_0}(2^{m_3-2} - 1) + (1 + 2^{m_2+m_0+m_3})(2^{m_2-2} - 1)$$

$$+ 2^{m_0+m_3+2m_2}(2^{m_1-2} - 1)$$

$$< 2^{m_0+m_3+2m_2+m_1-1} - 1,$$  \hfill (96)

$$\min\{\Lambda_\Sigma,1\} = -\max\{\Lambda_\Sigma,1\}$$

$$> 1 - 2^{m_0+m_3+2m_2+m_1-1}.$$  \hfill (97)

This completes the proof. \hfill □

Note that when $F_\Sigma,1$ and $\Lambda_\Sigma,1$ defined in (87) are with either one or more of the following conditions: 1) $\Lambda_1 = \{0\}$; 2) $\Lambda_3 = \{0\}$; 3) $\Lambda_\Sigma,2 = \{0\}$ it can be easily shown that results of Proposition 2 still hold.

In Fig. 3, we give a visual illustration on the effect of ‘−2’ bits. As it can be seen from Fig. 3(a), part of the desired signal and the interference signals intercept even though when the power level is power of 2. This is due to combining the two power levels of one signal into one

\footnotetext{For the ease of presentation, we use $\{0\}$ to represent $\emptyset$ as this does not change our scheme while allowing us to directly substitute $|\{0\}| = 1$ into cardinality calculations (e.g., condition $i$) in Proposition 2 for convenience.
(i.e., the power level for the pink part is combined with the power level of the second red part). Hence, the ‘−2’ bits serves a larger guard interval to avoid the carry over from the signal with two power levels to the signal above. Consequently, this case can be viewed as a type I scheme as shown in Fig. 3(c).

![Diagram of desired and interference signals](image)

**Fig. 3.** The desired and interference signals observed by receiver 1 when (a) the power level is power of 2 and no reduction on the cardinality, (b) the power level is not power of 2 and no reduction on the cardinality, (c) the power level is not power of 2 and with ‘−2’ reduction on the cardinality of $\Lambda_2$.

**Remark 6.** We introduce the ‘−2 bits’ reduction to the cardinalities of all the PAM signals for a type II scheme to simplify the presentation of our scheme under G-IC without losing the constant-gap optimality. In practice, the reduction on the cardinality can be less.

2b) For a type II scheme with high alignment at receiver 1, a submatrix $F_{\Sigma_3}$ of $[A_1 E_1 B_1 E_2]$ that contains the aligned matrices has the following form

$$F_{\Sigma_3} = \begin{bmatrix}
\Pi_{1, F}[F_1 O^{m_1, r'_{1}}] \\
\Pi_{2, F}[F_2 F_4] \\
\Pi_{3, F}[F_3 O^{m_3, r'_{3}}] \\
F_{\Sigma_2} \\
\Pi_{2, F}[F_2 O^{m_2, r'_{2}}]
\end{bmatrix},$$

where the settings are almost the same as in (86) except that the position of $\Pi_{2, F}[F_4 F_2]$ is in a higher signal level than in (86). The corresponding superimposed signal for (98) ordered from the lowest power level to the highest power level is

$$F_{\Sigma_3} = 2^{m_2}F_{\Sigma, 2} + 2^{m_0 + m_2 + \beta_3}F_3 + 2^{m_0 + m_2 + m_3 + \beta_4}F_4 + (2^{\beta_2} + 2^{m_0 + m_2 + m_3 + \beta_2})\rho_2 F_2.$$
\[
\sum 2^{m_0+m_3+2m_2+\beta_1} F_1 + 2^{m_0+m_2+\beta_2} \Lambda_3 + 2^{m_0+m_2+m_3+\beta_4} \rho_4 \Lambda_4 + (2^{\beta_2} + 2^{m_0+m_2+m_3+\beta_2}) \rho_2 \Lambda_2 \\
+ 2^{m_0+m_3+2m_2+\beta_1} \Lambda_1 \triangleq \Lambda_{\Sigma, 3},
\]
where we have used the same notations with those in (87) except now \((\rho_2, \rho_4) = (2^{\max\{\beta_1, \beta_2\}} - \beta_2, 2^{\max\{\beta_1, \beta_2\}} - \beta_4)\) are for slightly adjusting the power levels. Then we have the following proposition.

**Proposition 3.** For the random variable \(F_{\Sigma, 3}\) and the superimposed constellation defined in (99) with \(\Lambda_I\) and \(\Lambda_{\Sigma, 2}\) being non-empty sets for \(l \in [1 : 4]\), \(F_{\Sigma, 3}\) is uniformly distributed over \(\Lambda_{\Sigma, 3}\) satisfying

- \(i)\) \(|\Lambda_{\Sigma, 3}| = |\Lambda_4| \cdot |\Lambda_{\Sigma, 2}| \cdot |\Lambda_3| \cdot |\Lambda_2| \cdot |\Lambda_1|\),
- \(ii)\) \(1 \leq d_{\min}(\Lambda_{\Sigma, 3}) < 2\),
- \(iii)\) \(1 - 2^{m_0+m_3+2m_2+m_1-1} < \lambda < 2^{m_0+m_3+2m_2+m_1-1} - 1, \forall \lambda \in \Lambda_{\Sigma, 3}\).

**Proof:** We first analyze the minimum distance for each layer of superposition. We define

\[
\Lambda_A \triangleq 2^{m_0+m_2+m_3+\beta_4} \rho_4 \Lambda_4 + (2^{\beta_2} + 2^{m_0+m_2+m_3+\beta_2}) \rho_2 \Lambda_2 = 2^{\max\{\beta_1, \beta_1\}} \left(2^{m_0+m_2+m_3} \Lambda_4 + (1 + 2^{m_0+m_2+m_3}) \Lambda_2\right),
\]

\[
\Lambda_B \triangleq 2^{m_2} \Lambda_{\Sigma, 2} + 2^{m_2+m_0+\beta_3} \Lambda_3
\]

Thanks to the introduction of \(\rho_2\) and \(\rho_4\), according to Lemma \(4\) in Appendix \(B\), we can guarantee that the minimum distance of \(\Lambda_A\) satisfies

\[
d_{\min}(\Lambda_A) = 2^{\max\{\beta_1, \beta_1\}}.
\]

Moreover, using Lemma \(2\) in Appendix \(B\) together with (102), we obtain that

\[
d_{\min}(\Lambda A + 2^{m_0+2m_2+m_3+\beta_1} \Lambda_1) = d_{\min}(\Lambda A).
\]

As for \(\Lambda_B\), we know that \(d_{\min}(\Lambda_B) = 2^{m_2} d_{\min}(\Lambda_{\Sigma, 2}) \in [2^{m_2}, 2^{m_2+1})\) from (89). It remains to be shown that the minimum distance of \(\Lambda_A + \Lambda_B\) is bounded by a constant.

Again, by Lemma \(4\) in Appendix \(B\), \(\Lambda_A\) can be decomposed into

\[
\Lambda_A = \{\Lambda'_t\}, t \in [1 - 2^{m_2-2} : 2^{m_2-2} - 1],
\]

where \(\Lambda'_t = \text{PAM}(2^{m_2-2} - |t|, 2^{\max\{\beta_1, \beta_1\}})\) with mean \(E[\Lambda'_t] = 2^{\max\{\beta_1, \beta_1\}} t (2^{m_0+m_2+m_3} + \frac{1}{2})^6\)

\(^6\)For convenience, when \(2^{m_2-2} - |t| = 1\), we use \(\text{PAM}(1, d)\) with \(E[\Lambda_t] = a\) to denote \(\Lambda_t = \{a\}\) for some \(a \in \mathbb{R}\).
and

\[
d_{\text{Edge}}(\Lambda'_t, \Lambda'_{t+1}) \triangleq \min\{\Lambda'_t\} - \max\{\Lambda'_t\}
\]

\[
= \begin{cases} 
2^{\max\{\beta_{11}, \beta_{12}\}}(2 + 2^{m_0 + m_2 + m_3} - 2^{m_2 - 2} + t), & t \geq 0 \\
2^{\max\{\beta_{11}, \beta_{12}\}}(1 + 2^{m_0 + m_2 + m_3} - 2^{m_2 - 2} - t), & t < 0
\end{cases}
\]  (105)

is the distance between the edge points of two neighboring PAMs $\Lambda'_t$ and $\Lambda'_{t+1}$. An illustration of superimposed constellation $\Lambda_A$, sub-constellation $\Lambda'_t$ and $d_{\text{Edge}}(\Lambda'_t, \Lambda'_{t+1})$ are shown in Fig. 4.

![Fig. 4. An illustration of $\Lambda_A$, $\Lambda'_t$ and $d_{\text{Edge}}(\Lambda'_t, \Lambda'_{t+1})$.](image)

Using Lemma 2 in Appendix B we obtain that:

\[
d_{\min}(\Lambda'_t + \Lambda_B) = 2^{\max\{\beta_{11}, \beta_{12}\}}.
\]  (106)

Hence, the minimum distance of the superposition of any sub-constellation of $\Lambda_A$ and $\Lambda_B$ is still a constant.

Next, we compute the distance from the leftmost constellation point of $\Lambda'_t + \Lambda_B$ to the rightmost constellation point of $\Lambda'_t + \Lambda_B$ for $t \in [1 - 2^{m_2 - 2}; 2^{m_2 - 2} - 2]$ as

\[
d_{\text{Edge}}(\Lambda'_t + \Lambda_B, \Lambda'_{t+1} + \Lambda_B) = \min\{\Lambda'_t + \Lambda_B\} - \max\{\Lambda'_t + \Lambda_B\}
\]

\[
= \min\{\Lambda'_t + \Lambda_B\} - \max\{\Lambda'_t\} + \min\{\Lambda_B\} - \max\{\Lambda_B\}
\]

\[
\geq \min\{\Lambda'_t\} - \max\{\Lambda'_t\} + \min\{\Lambda_B\} - \max\{\Lambda_B\}
\]

\[
= 2^{\max\{\beta_{11}, \beta_{12}\}}(2 + 2^{m_0 + m_2 + m_3} - 2^{m_2 - 2})
\]

\[
+ 2^{m_2 + m_0 + \beta_3}(1 - 2^{m_3 - 2}) + 2^{m_2 + 1}(1 - 2^{m_0 - 1})
\]

\[
\geq 2^{1 + \max\{\beta_{11}, \beta_{12}\}}.
\]  (107)
Analogues to the inter-constellation in Definition 1, the positivity of this metric indicates that the \( \Lambda_A + \Lambda_B \) and \( \Lambda_{A+1} + \Lambda_B \) do not intercept with each other. Since \( \Lambda_A \) can be decomposed into the form in (104), the minimum distance of \( \Lambda_A + \Lambda_B \) satisfies

\[
d_{\min}(\Lambda_A + \Lambda_B) = \min \left\{ \min_t \{d_{\text{Edge}}(\Lambda_A + \Lambda_B, \Lambda_{A+1} + \Lambda_B)\}, \min_t \{d_{\min}(\Lambda_A + \Lambda_B)\}, \min_t \{d_{\min}(\Lambda'_A)\} \right\}
\]

\[
\leq d_{\min}(\Lambda_A) = 2^{\max\{\beta_{11}, \beta_{12}\}},
\]

(108)

where \((e)\) is due to (102), (106) and (107).

Now, with (89), (103), (108) and by Lemma 2 and Lemma 3 in Appendix B, we obtain that

\[
d_{\min}(\Lambda_{A,3}) = 2^{\max\{\beta_{11}, \beta_{12}\}} \in [1, 2).
\]

(109)

Since there is no overlapping in \( \Lambda_{A,3} \), the cardinality thus satisfies

\[
|\Lambda_{A,3}| = |\Lambda_4| \cdot |\Lambda_{A,2}| \cdot |\Lambda_3| \cdot |\Lambda_2| \cdot |\Lambda_1|. 
\]

(110)

By considering the extreme case such that \( \beta_l = 1 \) for \( l \in [1 : 4] \), we can obtain an upper bound on \( \max\{\Lambda_{A,3}\} \) and a lower bound on \( \min\{\Lambda_{A,3}\} \) as

\[
\max\{\Lambda_{A,3}\} \leq \max\{2^{m_2} \Lambda_{A,2}\} + 2^{m_2 + m_0} \max\{\Lambda_3\} + 2^{m_2 + m_0 + m_3} \max\{\Lambda_4\} + (1 + 2^{m_2 + m_0 + m_3}) \max\{\Lambda_2\} + 2^{m_0 + m_3 + 2m_2} \max\{\Lambda_1\} < 2^{m_2}(2^{m_0 - 1} - 1) + 2^{m_2 + m_0}(2^{m_3 - 2} - 1) + 2^{m_2 + m_0 + m_3}(2^{m_2 - 2} - 1) + (1 + 2^{m_2 + m_0 + m_3})(2^{m_2 - 2} - 1) + 2^{m_0 + m_3 + 2m_2}(2^{m_1 - 2} - 1)
\]

\[
< 2^{m_0 + m_3 + 2m_2 + m_1 - 1} - 1,
\]

(111)

\[
\min\{\Lambda_{A,3}\} = - \max\{\Lambda_{A,3}\} > 1 - 2^{m_0 + m_3 + 2m_2 + m_1 - 1}.
\]

(112)

This completes the proof.

Note that when \( F_{\Sigma,3} \) and \( \Lambda_{A,3} \) defined in (99) are with either one or more of the following conditions: 1) \( \Lambda_1 = \{0\} \); 2) \( \Lambda_3 = \{0\} \); 3) \( \Lambda_{A,2} = \{0\} \), it can be easily shown that the results of Proposition 3 still hold.

We now apply Positions 1, 3 to analyze the minimum distance and the cardinality of \( X_A \) under one alignment and two matrix alignments. In general, the signal above the noise level can be written as

\[
X_A \in \Lambda_{A,a} + 2^{m_a} \Lambda_{A,b} + 2^{m_a + m_0 + m_3 + 2m_2 + m_1} \Lambda_{A,b} \triangleq \Lambda_{A}, k \in \{1, 3\},
\]

(113)
where \( k = 1 \) indicates that there is a low alignment inside \( \Lambda_{\Sigma,1} \) and thus \( \Lambda_{\Sigma,1} \) has the form of \( 87 \) while \( k = 3 \) indicates that there is a high alignment inside \( \Lambda_{\Sigma,3} \) and thus \( \Lambda_{\Sigma,3} \) has the form of \( 99 \). \( \Lambda_{\Sigma,a} \) and \( \Lambda_{\Sigma,b} \) are two superimposed constellations including the channel differences \( \beta, m_a \) is the number of rows of the matrix associated with \( \Lambda_{\Sigma,a} \) in the D-IC. The submatrices associated with \( \Lambda_{\Sigma,a} \) and \( \Lambda_{\Sigma,b} \) have higher and lower signal levels than that of the submatrices associated with \( \Lambda_{\Sigma,1} \), respectively, in the D-IC.

2c) First, we consider the scenario where there is only one matrix alignment from receiver 1’s perspective in a type II scheme. Note that all the type II schemes for achieving all the corner points of the capacity region of the D-IC (i.e., all the type II schemes for user 1 and user 2 in Appendix A) only have one matrix alignment. Since the matrix alignment happens in the submatrix associated with either \( \Lambda_{\Sigma,1} \) or \( \Lambda_{\Sigma,3} \) in (113), then \( \Lambda_{\Sigma,a} \) and \( \Lambda_{\Sigma,b} \) have the following properties according to Proposition 1

\[
1 \leq d_{\min}(\Lambda_{\Sigma,a}) < 2, \max\{\Lambda_{\Sigma,a}\} < 2^{m_a-1} - 1, \min\{\Lambda_{\Sigma,a}\} > 1 - 2^{m_a-1}.
\]

(114)

\[
1 \leq d_{\min}(\Lambda_{\Sigma,b}) < 2.
\]

(115)

By Lemma 2 in Appendix B, Proposition 2 and Proposition 3, we obtain the followings

\[
d_{\min}(\Lambda_{\Sigma,a} + 2^{m_a} \Lambda_{\Sigma,k}) = d_{\min}(\Lambda_{\Sigma,a})
\]

(116)

\[
d_{\min}(2^{m_a} \Lambda_{\Sigma,k} + 2^{m_a+m_3+2m_2+m_1} \Lambda_{\Sigma,b}) = 2^{m_a} d_{\min}(\Lambda_{\Sigma,k}).
\]

(117)

Finally, with (116) and (117) and by applying Lemma 3, we have that

\[
d_{\min}(\Lambda_{\Sigma}) \geq 1.
\]

(118)

When \( X_A \) and \( \Lambda_{\Sigma} \) defined in (113) with either one or more than one of the following conditions:

1) \( \Lambda_{\Sigma,a} = \{0\} \); 2) \( \Lambda_{\Sigma,b} = \{0\} \), it can be easily shown that (118) still holds.

Therefore, \( X_A \) is always uniformly distributed over \( \Lambda_{\Sigma} \) in (85) with \( d_{\min}(\Lambda_{\Sigma}) \geq 1 \) and cardinality \( |\Lambda_{\Sigma}| = \prod_{l=1}^{L} |\Lambda_l| \). The achievable rate of user 1 is

\[
I(X_1; Y_1) \geq r_1 - 2|A_1| - \frac{1}{2} \log_2 2\pi e \left( \frac{13}{12} \right) - \frac{1}{2} \log_2 3,
\]

(119)

where \( 2|A_1| \) is a constant due to the \(-2\) on the cardinality of user 1’s signals above the noise level.

2d) Now we consider the case where there are two matrix alignments from receiver 1’s perspective in a type II scheme. In particular, we consider that one of the matrix alignment is
low alignment as this covers all the type II schemes we presented in this work with two matrix
alignments. The only instance of such case is the type II scheme in Sec. III-C2b (for both users).
First, we note that the scenario of low alignment in (86) can be regarded as a type I scheme after
the 2-bit guard bits being introduced to each PAM modulation of $A_{\Sigma,1}$ in (87) (also illustrated
in Fig. 3). Therefore, if one of the matrix alignment is low alignment regardless whether the
other matrix alignment is high low alignment, then this case is equivalent to the case of one
matrix alignment discussed in Sec. IV-B2c. Thus, under this case, user 1’s achievable is also
lower bounded by (119).

Remark 7. To achieve the rate points between any two neighboring corner points, it is possible
for a type II scheme to have two high alignments or more than two matrix alignments. Here,
we only cover the cases of one matrix alignment and partially cover the cases of two matrix
alignments to demonstrate that it is possible to achieve those rate points by our schemes with
TIN and without time-sharing. Although we do not dig further in the interest of space, we do
believe that our analysis is generalizable to other type II schemes with two high alignments or
more than two matrix alignments.

3) User 2: The analysis for user 2’ achievable rate is similar to that (i.e., (72) - (82)) for
user 1 and the scenarios considered from Sec. IV-B1 to Sec. IV-B2d also apply to user 2. To
avoid repetition, we present the final form of user 2’s rate as

$$I(X_2; Y_2) \geq r_2 - 2|A_2| - \frac{1}{2} \log_2 2\pi e \left( \frac{1}{d_{\min}(X_C)} + \frac{1}{12} \right) - \frac{1}{2} \log_2 3,$$

(120)

where $X_C$ is the signals above the noise level from receiver 2’s point of view; $r_2$ corresponds
to user 2’s target rate in Theorem 1. Following the same steps for user 1, user 2’s rate is lower
bounded by

$$I(X_2; Y_2) \geq r_2 - 2|A_2| - \frac{1}{2} \log_2 2\pi e \left( \frac{13}{12} \right) - \frac{1}{2} \log_2 3.$$

(121)

As it can be seen from (119) and (121) that $r_1 - I(X_1; Y_1)$ and $r_2 - I(X_2; Y_2)$ are upper
bounded by two constants, respectively. Hence, our scheme is able to achieve every corner point
of $C_{G-IC}$ to within a constant gap. This, together with time-sharing, shows that our proposed
scheme achieves a rate region $R_{G-IC}^{TIN}$ satisfying $C_{G-IC} \subseteq R_{G-IC}^{TIN} + c''$ for some constant $c'' > 0.$
This completes the proof of Theorem 2.
V. THE COMPLEX GAUSSIAN INTERFERENCE CHANNEL

In this section, we translate the proposed scheme from the D-IC to the discrete input distribution for the complex G-IC and analyze the achievable rate pair.

A. Proposed Scheme

For the complex G-IC, we have that $n_{11} = \lfloor \log_2 \text{SNR}_1 \rfloor, n_{12} = \lfloor \log_2 \text{INR}_1 \rfloor, n_{22} = \lfloor \log_2 \text{SNR}_2 \rfloor, n_{21} = \lfloor \log_2 \text{INR}_2 \rfloor$ and $q = \max\{n_{11}, n_{12}, n_{22}, n_{21}\}$. We again define the differences between the actual channel gains and the quantized channel gains as

\begin{align}
\beta_{11} &\triangleq \log_2 \text{SNR}_1 - n_{11}, \\
\beta_{12} &\triangleq \log_2 \text{INR}_1 - n_{12}, \\
\beta_{22} &\triangleq \log_2 \text{SNR}_2 - n_{22}, \\
\beta_{21} &\triangleq \log_2 \text{INR}_2 - n_{21},
\end{align}

where $\beta_{11}, \beta_{12}, \beta_{22}, \beta_{21} \in [0, 1)$. We then translate our input distributions for D-IC in Sec. III-C to obtain the following proposed input signaling, which is a multi-layer superposition of QAM constellation for each user

\begin{equation}
X_k = \gamma_k \sum_{i_k=1}^{L_k} 2^{-\frac{\text{row}(\text{rank}(F_{k,i_k}))+1}{2}} F_{k,i_k}, k \in \{1, 2\},
\end{equation}

where the notations and their definitions here follow from those in (69) except that $F_{k,i_k}$ is a random variable uniformly distributed over $\text{QAM}(2^{\text{rank}(F_{k,i_k})}, 1)$, and $\gamma_k = 2^{-q}$ to ensure $\mathbb{E}[\|X_k\|^2] \leq 1$. When using discrete signaling in the complex G-IC, the most difficult problem to deal with is the channel phase distortions in two direct links and cross links. Thus, we ignore the fine tunes, including $-1$ or $-2$ to the cardinality and the power scaling factors $\rho_{k,i_k}$, as they may not be effective.

We analyze the achievable rate of user 1 under our scheme first. User 2’s achievable rate can be similarly analyzed by following the same line of analysis for user 1 and is thus omitted.
At receiver 1, the received signal is given in (1) where
\[ h_{11}X_1 + h_{12}X_2 = \sqrt{\text{SNR}_1} e^{j\theta_{11}}X_1 + \sqrt{\text{INR}_1} e^{j\theta_{12}}X_2 \]
\[ = 2^{n_{11} + \beta_{11} - q} \sum_{i=1}^{L_1} 2^{\text{row}(\mathbf{E}_1[i], \mathbf{F}_1^T[j])} e^{j\theta_{11}} F_{1,i} + 2^{n_{12} + \beta_{12} - q} \sum_{j=1}^{L_2} 2^{\text{row}(\mathbf{E}_2[j], \mathbf{F}_2^T[j])} e^{j\theta_{12}} F_{2,j} \]
\[ = 2^{n_{1k} + \beta_{1k} - q} \sum_{k=1}^{2} 2^{\text{row}(\mathbf{E}_1[k], \mathbf{F}_1^T[k])} e^{j\theta_{1k}} F_{k,i_k} + 2^{n_{2k} + \beta_{2k} - q} \sum_{k=1}^{2} 2^{\text{row}(\mathbf{E}_2[k], \mathbf{F}_2^T[k])} e^{j\theta_{2k}} F_{2,j_k}, \]
where \( X_A \) and \( X_B \) represent the superpositions of all signals above and below the noise level, respectively, from receiver 1’s perspective and \( A_k, B_k \) follow (75) and (77), respectively.

Similar to the rate analysis for real setting in (80) and (81), user 1’s achievable rate in the complex G-IC is lower bounded by
\[ I(X_1; Y_1) \geq I(X_A; X_A + Z_1) - H(X_{A,2}) - \log_3 3 \]
\[ \geq a \frac{4}{\pi d_{\text{min}}^2(X_A)} + \frac{1}{4} - H(X_{A,2}) - \log_3 3 \]
\[ \geq a \frac{4}{\pi d_{\text{min}}^2(X_A)} + \frac{1}{4} + \log_3 3, \]
where \( \log_3 3 \) is the maximum number of loss bits due to signals below the noise level, \( (a) \) follows by applying an Ozarow-type [28] for a uniform input distribution over a two-dimensional constellation in Lemma 5 in Appendix B and \( (b) \) holds only if the probability of overlapping in \( X_A \) is zero, which will be proved in Proposition 4. Then the gap between the capacity and the achievable rate solely depends on the minimum distance of the signal above the noise level from receiver 1’s point of view. Similar to the real setting, we focus on the analysis of the minimum distance of \( X_A \) only and omit the details for user 2 to avoid unnecessary repetition.

B. Minimum Euclidean Distance Analysis

The signals above the noise level for the complex G-IC in (1) can be written in the following form based on (127) and (75)
\[ X_A = 2^{\beta_{11}} e^{j\theta_{11}} \sum_{i_1 \in A_1} P_{1,i_1} F_{1,i_1} + 2^{\beta_{12}} e^{j\theta_{12}} \sum_{i_2 \in A_2} P_{2,i_2} F_{2,i_2}, \]
(129)
where \( P_{k,i,k} \triangleq 2^{n_{1,k}} \cdot 2^{g + \text{row}(E_k) \cdot [\ldots F_{k,i,k}]^T} \) and \( F_{k,i,k} \) is uniformly distributed over \( \Lambda_{k,i,k} = \text{QAM}(2^{\text{rank}(F_{k,i,k})}, 1) \) for \( k \in \{1, 2\} \). Moreover, since any phase rotation at the receiver does not lose information, we equivalently consider \( e^{-j\theta_{11}} X_A \)

\[
e^{-j\theta_{11}} X_A \in 2^{\beta_{11}} \sum_{i_1 \in A_1} P_{1,i_1} \lambda_{1,i_1} + 2^{\beta_{12}} e^{j\theta} \sum_{i_2 \in A_2} P_{2,i_2} \lambda_{2,i_2} \in \Lambda_{\Sigma},
\]

where \( \theta \triangleq \theta_{12} - \theta_{11} \sim \text{Unif}[0, 2\pi] \) is the phase difference between \( h_{11} \) and \( h_{12} \).

For any \( \lambda_{k,i,k} \in \Lambda_{k,i,k} \), we let \( \mathfrak{R}(\lambda_{k,i,k}), \mathfrak{I}(\lambda_{k,i,k}) \in \{ \pm \frac{1}{2}, \ldots, \pm 2^{\frac{\log_2|\Lambda_{k,i,k}|}{2}} - \frac{1}{2} \} \) represent the real and imaginary part of a constellation point \( \lambda_{k,i,k} \), respectively. For any \( \lambda \neq \lambda' \in \Lambda_{\Sigma} \), the square Euclidean distance between \( \lambda \) and \( \lambda' \) is

\[
d^2(\lambda, \lambda') = \left( \mathfrak{R} \left( 2^{\beta_{11}} \sum_{i_1 \in A_1} P_{1,i_1} (\lambda_{1,i_1} - \lambda'_{1,i_1}) + 2^{\beta_{12}} e^{j\theta} \sum_{i_2 \in A_2} P_{2,i_2} (\lambda_{2,i_2} - \lambda'_{2,i_2}) \right) \right)^2 + \left( \mathfrak{I} \left( 2^{\beta_{11}} \sum_{i_1 \in A_1} P_{1,i_1} (\lambda_{1,i_1} - \lambda'_{1,i_1}) + 2^{\beta_{12}} e^{j\theta} \sum_{i_2 \in A_2} P_{2,i_2} (\lambda_{2,i_2} - \lambda'_{2,i_2}) \right) \right)^2
\]

\[
= 2^{\beta_{11}} (\Delta_{R_1}^2 + \Delta_{I_1}^2) + 2^{\beta_{12}} (\Delta_{R_2}^2 + \Delta_{I_2}^2) + 2^{\beta_{11} + \beta_{12} + 1} \cos \theta (\Delta_{R_1} \Delta_{R_2} + \Delta_{I_1} \Delta_{I_2})
\]

\[
+ 2^{\beta_{11} + \beta_{12} + 1} \sin \theta (\Delta_{R_2} \Delta_{I_1} - \Delta_{R_1} \Delta_{I_2})
\]

\[
= (2^{\beta_{12}} \Delta_{R_2} + 2^{\beta_{11}} \Delta_{R_1} \cos \theta + 2^{\beta_{11}} \Delta_{I_1} \sin \theta)^2 + (2^{\beta_{12}} \Delta_{I_2} - 2^{\beta_{11}} \Delta_{R_1} \sin \theta + 2^{\beta_{11}} \Delta_{I_1} \cos \theta)^2,
\]

where we define the followings for notation simplicity for \( k \in \{1, 2\} \)

\[
\Delta_{R_k} \triangleq \sum_{i,k \in A_k} P_{k,i,k} \mathfrak{R}(\lambda_{k,i,k} - \lambda'_{k,i,k}) \in \mathbb{Z},
\]

\[
\Delta_{I_k} \triangleq \sum_{i,k \in A_k} P_{k,i,k} \mathfrak{I}(\lambda_{k,i,k} - \lambda'_{k,i,k}) \in \mathbb{Z},
\]

and \( \mathfrak{R}(\lambda_{k,i,k} - \lambda'_{k,i,k}), \mathfrak{I}(\lambda_{k,i,k} - \lambda'_{k,i,k}) \in \{ 0, \pm 1, \ldots, \pm 2^{\frac{\log_2|\Lambda_{k,i,k}|}{2}} - 1 \} \).

The minimum Euclidean distance of \( X_A \) is then given by

\[
d_{\min}(\Lambda_{\Sigma}) = \min\{d(\lambda, \lambda')\}.
\]

We define the outage probability \( \eta \triangleq \Pr\{d_{\min}(\Lambda_{\Sigma}) < d_0\} \) for a target minimum distance \( d_0 > 0 \).

According to (131), it is obvious that \( d_{\min}(\Lambda_{\Sigma}) = 0 \) if and only if

\[
2^{\beta_{12}} \Delta_{R_2} = -2^{\beta_{11}} \Delta_{R_1} \cos \theta - 2^{\beta_{11}} \Delta_{I_1} \sin \theta,
\]

\[
2^{\beta_{12}} \Delta_{I_2} = 2^{\beta_{11}} \Delta_{R_1} \sin \theta - 2^{\beta_{11}} \Delta_{I_1} \cos \theta.
\]
In what follows, we show that this event has measure zero.

**Proposition 4.** For the channel model in Sec. II and by using the proposed scheme, the channels \((h_{11}, h_{12}) \in \mathbb{C}^2\) with \(\theta_{11}, \theta_{12} \in [0, 2\pi]\) such that \(d_{\min}(\Lambda_{\Sigma}) = 0\) for \(\Lambda_{\Sigma}\) in (130) have Lebesgue measure zero.

**Proof:** We prove the proposition from receiver 1’s perspective. The proof for receiver 2 is similar to receiver 1 and thus is omitted.

According to the conditions for \(d_{\min}(\Lambda_{\Sigma}) = 0\) in (135) and (136), we have

\[
\sin \theta = 2^{\beta_{12} - \beta_{11}} \frac{\Delta_{R_1} \Delta_{I_2} - \Delta_{I_1} \Delta_{R_2}}{\Delta_{R_1}^2 + \Delta_{I_1}^2} \in [-1, 1],
\]

(137)

\[
\cos \theta = -2^{\beta_{12} - \beta_{11}} \frac{\Delta_{R_1} \Delta_{R_2} + \Delta_{I_1} \Delta_{I_2}}{\Delta_{R_1}^2 + \Delta_{I_1}^2} \in [-1, 1].
\]

(138)

Since (137) and (138) need to satisfy \(\sin^2 \theta + \cos^2 \theta = 1\), we thus obtain

\[
2^{2(\beta_{12} - \beta_{11})} \frac{\Delta_{R_2}^2 + \Delta_{I_2}^2}{\Delta_{R_1}^2 + \Delta_{I_1}^2} = 1.
\]

(139)

Substituting (139) into (137) and (138) gives

\[
\sin \theta = \frac{\Delta_{R_1} \Delta_{I_2} - \Delta_{I_1} \Delta_{R_2}}{\sqrt{\Delta_{R_1}^2 + \Delta_{I_1}^2}} \in [-1, 1],
\]

(140)

\[
\cos \theta = -\frac{\Delta_{R_1} \Delta_{R_2} + \Delta_{I_1} \Delta_{I_2}}{\sqrt{\Delta_{R_1}^2 + \Delta_{I_1}^2}} \in [-1, 1].
\]

(141)

Again since \(\Delta_{R_k}\) and \(\Delta_{I_k}\) only take value from a subset of the integer set as shown in (132) and (133), the solution set of \(\theta\) to (131) is a discrete set and thus countable. Hence, \(d_{\min}(\Lambda_{\Sigma}) = 0\) has measure zero.

To obtain a closed-form expression for the minimum in (134) is difficult. In what follows, we use some examples to show the minimum distance of \(\Lambda_{\Sigma}\) for a number of channel settings.

**Example 2.** Consider a superimposed constellation \(\Lambda_{\Sigma} = \sum_{l=1}^{3} e^{j \theta_l} P_l \Lambda_l\), where \(\theta_1 = \theta_3 = \theta_{11}, \theta_2 = \theta_{12}\), for \(l \in \{1, 3\}\), \(\Lambda_l = QAM(2^{m_l}, 1)\), and \((P_1, P_2, P_3) = (1, 2^{m_1}, 2^{m_2 + m_3})\). The outage probability \(\eta\) versus target minimum distance \(d_{\delta}\) for various channel settings are shown in Fig. 3 where the legend shows the values of \((m_1, m_2, m_3)\).

From the figure, it can be seen that for a given target outage probability \(\eta\), \(d_{\min}(\Lambda_{\Sigma})\) is reduced by at most about a factor of 2 when the superimposed constellation size \(|\Lambda_{\Sigma}| = 2^{m_1 + m_2 + m_3}\) is increased from \(2^{10}\) to \(2^{20}\), which is equivalent to about doubling \(\max\{\text{SNR}_1, \text{INR}_1\}\) in dB. This
Fig. 5. Minimum distance and the outage.

is because $|\Lambda_\Sigma|$ is at most $\max\{n_{11}, n_{12}\} = \max\{\log_2 \text{SNR}_1 - \beta_{1}, \log_2 \text{INR}_1 - \beta_{12}\}$. Moreover, the minimum distance does not reduced much when the overall constellation size is increased from $2^{30}$ to $2^{40}$.

C. Simulation Results

In this section, we evaluate a number of achievable rate pairs $(I(X_1; Y_1), I(X_2; Y_2))$ under the complex G-IC by Monte Carlo simulation.

We consider two cases: $(\text{SNR}_1, \text{INR}_1, \text{SNR}_2, \text{INR}_2) = (49, 37, 43, 31)$ and $(25, 30, 13, 17)$ dB, corresponding to $(n_{11}, n_{12}, n_{22}, n_{21}) = (16, 12, 14, 10)$ and $(8, 10, 4, 2)$, respectively. The first case belongs to the case of Weak 1-2 in Sec. III-C2 and the second case is Mixed 5-1 in Sec. A-K1. The achievable rate pairs $(I(X_1; Y_1), I(X_2; Y_2))$ are averaged over 50000 samples of random channel phases. To put the results of the proposed scheme in context, we also include the capacity outer bound of the complex G-IC from [6], the capacity region of complex D-IC from [26], HanKobayashi achievable region with Gaussian signaling from [6] and the achievable rate of Gaussian signaling with TIN.

As shown in Fig. 6 and Fig. 7 our scheme with purely discrete inputs and single-user TIN decoding can operate close to the outer bound of the capacity region of the complex G-IC and that of the complex D-IC for both cases. Notably, our scheme significantly outperforms the conventional approach using Gaussian signaling with single-user TIN decoding in the weak
Fig. 6. Achievable rate pairs for the weak interference regime.

Fig. 7. Achievable rate pairs for the mixed interference regime.

interference regime (Fig. 6) and for user 1 in the mixed interference regime (Fig. 7). The reason
that user 2’s achievable rate is similar to that of the Gaussian TIN is because the interference experience by user 2 is already very weak. In summary, our results indicate that although suffering from random phase distortion introduced by complex channel coefficients, by properly designing the input distributions, the proposed scheme with low-complexity TIN decoding can still be very promising.

VI. CONCLUDING REMARKS

In this work, we studied the problem of using discrete signaling with TIN for the real and complex G-IC. Most importantly, we constructed coding schemes with TIN to achieve the entire capacity region of the D-IC for all cases under weak, strong and mixed interference regimes. We then translated the schemes from the D-IC into the G-IC and provided a systematic way to design discrete input signaling. Achievable rates of our schemes were analyzed for the real and complex G-IC under TIN. For the real G-IC, we proved that our scheme is able to achieve the entire capacity region to within a constant gap, regardless of all channel parameters. For the complex G-IC, we investigated the impact of phase distortions on the minimum distance of the underlying modulations and showed that our scheme significantly outperforms the existing scheme with TIN. We remark that when translating schemes in the D-IC to that for the G-IC, there are many parameters one can tune to get improved results. However, since our motivation is to showcase the usefulness of discrete input signalling under TIN, we focus on a simple and systematic translation that is analytically tractable and we leave meticulous optimization for future work.

APPENDIX A

PROOF OF THEOREM 1 CONT.

In this section, for the sake of completeness, we provide our choice of $E_1$ and $E_2$ for achieving all the non-trivial corner points (exclude $(n_{11}, 0)$ and $(0, n_{22})$) on the corresponding capacity region of the asymmetric D-IC. We treat each interference regime and its subregimes separately. We note that the schemes proposed in Sec. III-C and here in Appendix A together cover all the scenarios of the not very weak, not very strong and mixed interference regimes.

Recall that for the weak interference regime, i.e., $n_{11} > n_{21}, n_{22} > n_{12}$ [6], we only need to consider the regime under $n_{22} < n_{12} + n_{21}$. This is because with our assumption $n_{11} > n_{22}$ along with $n_{22} > n_{12} + n_{21}$ imply that $n_{11} > n_{22} > n_{12} + n_{21}$, which corresponds to very
weak interference regime \[15\]. For the strong interference regime, i.e., \(n_{11} < n_{21}, n_{22} < n_{12}\) \[6\], we exclude the very strong interference regime \[8\], i.e., the channel condition shall not satisfy \(n_{12} > n_{11} + n_{22}\) and \(n_{21} > n_{11} + n_{22}\) simultaneously. The mixed interference regime is defined as \(n_{12} < n_{22}, n_{21} > n_{11}\) or \(n_{12} > n_{22}, n_{21} < n_{11}\) according to \[6\].

A. Weak 1: \(n_{11} > n_{22} > n_{12} > n_{21}\)

In this subsection, we provide the generator matrices for achieving the non-trivial corner points for the rest of the regimes in Weak 1 which have not appeared in Sec. \[\text{III-C}\]

1) : Now we consider the regime of \(n_{11} < n_{12} + n_{21}\) and \(n_{11} + n_{22} - 2n_{12} - n_{21} > 0\), where the second inequality implies that \(n_{11} + n_{22} - n_{12} - 2n_{21} > 0\). The non-trivial corner points on the corresponding capacity region are \((2n_{11} - n_{12} - n_{21}, 2(n_{21} + n_{12} - n_{11}))\), \((2(n_{12} + n_{21} - n_{22}), 2n_{22} - n_{12} - n_{21})\).

To achieve the target rate pair \((2n_{11} - n_{12} - n_{21}, 2(n_{21} + n_{12} - n_{11}))\), we propose

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{n_{12} + n_{21} - n_{11}, r_1} \\
F_{1,2}
\end{bmatrix}, 
E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{11} + n_{22} - n_{12} - 2n_{21}, r_2} \\
0^{n_{12} + n_{21} - n_{22}, r_2} \\
0^{n_{11} + n_{22} - 2n_{12} - n_{21}, r_2}
\end{bmatrix},
\tag{142}
\]

where \(F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}, F_{2,1}, F_{2,2} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{11}, r_2}\).

To achieve \((2(n_{12} + n_{21} - n_{22}), 2n_{22} - n_{12} - n_{21})\), we propose

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{n_{11} + n_{22} - 2n_{12} - n_{21}, r_1} \\
0^{n_{12} + n_{21} - n_{11}, r_1} \\
0^{n_{11} + n_{22} - n_{12} - 2n_{21}, r_1}
\end{bmatrix}, 
E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0^{n_{12} + n_{21} - n_{22}, r_2} \\
F_{2,3} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
\tag{143}
\]

where \(F_{1,1}, F_{1,2} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{12}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{11}, r_1}, F_{2,2} \in \mathbb{F}_2^{n_{11} + n_{22} - n_{12} - 2n_{21}, r_2}, F_{2,3} \in \mathbb{F}_2^{n_{12} - n_{12}, r_2}\).

2) : Now we consider the regime of \(n_{11} < n_{12} + n_{21}\) and \(n_{11} + n_{22} - n_{12} - 2n_{21} > 0 > n_{11} + n_{22} - 2n_{12} - n_{21}\). The non-trivial corner points on the corresponding capacity region are \((n_{11} + n_{12} - n_{22}, 2(n_{22} - n_{12})), (2(n_{11} - n_{12}), n_{22} + n_{12} - n_{11})\).
To achieve the target rate pair \((n_{11} + n_{12} - n_{22}, 2(n_{22} - n_{12}))\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{22} - n_{12}, r_1} \\ F_{1,2} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{2n_{12} + n_{21} - n_{11} - n_{22}, r_2} \\ F_{2,2} \\ 0^{n_{11} - n_{22}}, r_2 \end{bmatrix}
\] (144)

where \(F_{1,1} \in \mathbb{P}_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{P}_2^{n_{22} - n_{12}, r_1}, F_{2,1}, F_{2,2} \in \mathbb{P}_2^{n_{11} - n_{22}, r_2}\).

To achieve the rate pair \((2(n_{11} - n_{12}), n_{22} + n_{12} - n_{11})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{2n_{12} + n_{21} - n_{11} - n_{22}, r_1} \\ 0^{n_{22} - n_{12}, r_1} \\ 0^{n_{11} + n_{22} - 2n_{21} - n_{12}, r_1} \\ F_{1,2} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{11} - n_{12}, r_2} \\ F_{2,3} \\ F_{2,4} \end{bmatrix}
\] (145)

where \(F_{1,1}, F_{1,2} \in \mathbb{P}_2^{n_{11} - n_{12}, r_1}, F_{2,1} \in \mathbb{P}_2^{n_{12} + n_{21} - n_{11}, r_2}, F_{2,2} \in \mathbb{P}_2^{n_{11} + n_{22} - 2n_{21} - n_{12}, r_2}, F_{2,3} \in \mathbb{P}_2^{n_{21} + n_{22} - n_{11} - n_{22}, r_2}, F_{2,4} \in \mathbb{P}_2^{n_{22} - n_{12}, r_2}\).

3) When \(n_{11} < n_{12} + n_{21}\) and \(n_{11} + n_{22} - n_{12} - 2n_{21} < 0\), the non-trivial corner points are the same as in Weak 1-2 in Appendix A-A2.

3a) Under the conditions in Appendix A-A3, we further consider \(2(n_{12} + n_{21} - n_{22}) - n_{11} > 0\).

To achieve \((n_{11} + n_{12} - n_{22}, 2(n_{22} - n_{12}))\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{12} - n_{21}, r_2} \\ F_{2,2} \\ 0^{n_{11} - n_{22}, r_2} \end{bmatrix}
\] (146)

where \(F_{1,1} \in \mathbb{P}_2^{n_{11} - n_{21}, r_1}, F_{1,2} \in \mathbb{P}_2^{n_{22} - n_{12}, r_1}, F_{1,3} \in \mathbb{P}_2^{2n_{12} + 2n_{21} - n_{11} - 2n_{22}, r_1}, F_{1,4} \in \mathbb{P}_2^{n_{11} - n_{21}, r_1}, F_{2,1}, F_{2,2} \in \mathbb{P}_2^{n_{22} - n_{12}, r_2}\).
$3b$: To achieve $(2(n_{11} - n_{12}), n_{22} + n_{12} - n_{11})$, we need to further consider $3n_{12} + n_{21} - 2n_{11} - n_{22} > 0$ under the conditions of $3a$ above. The generator matrices are given by

$$
E_1 = \begin{bmatrix}
F_{1,1} & 0^{n_{22} - n_{21}, r_1} \\
0^{2n_{21} + 2n_{12} - n_{11} - 2n_{22}, r_1} & F_{1,2} \\
0^{n_{22} - n_{12}, r_1} & 0^{n_{12} - n_{21}, r_1} \\
F_{1,3} & F_{1,4}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
F_{2,4}
\end{bmatrix},
$$

(147)

where $F_{1,1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{21}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{11} - n_{12}, r_2}, F_{2,3} \in \mathbb{F}_2^{3n_{12} + 2n_{21} - 2n_{11} - n_{22}, r_2}, F_{2,4} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2}$.

Then for $3n_{12} + n_{21} - 2n_{11} - n_{22} < 0$, the generator matrices for achieving $(2(n_{11} - n_{12}), n_{22} + n_{12} - n_{11})$ are

$$
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
0^{n_{22} - n_{21}, r_1} \\
0^{2n_{21} + 2n_{12} - n_{11} - 2n_{22}, r_1} \\
0^{n_{22} - n_{12}, r_1} \\
0^{n_{12} - n_{21}, r_1} \\
F_{1,3} \\
F_{1,4}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0^{2n_{11} + n_{22} - 2n_{21} - 2n_{12}, r_2} \\
F_{2,3} \\
F_{2,4} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
$$

(148)

where $F_{1,1}, F_{1,4} \in \mathbb{F}_2^{2n_{21} + n_{12} - n_{11} - n_{22}, r_1}, F_{1,2}, F_{1,3} \in \mathbb{F}_2^{2n_{11} + n_{22} - 2n_{21} - 2n_{12}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{21}, r_2}, F_{2,2} \in \mathbb{F}_2^{2n_{21} + n_{12} - n_{11} - n_{22}, r_2}, F_{2,3} \in \mathbb{F}_2^{n_{21} - n_{22}, r_2}, F_{2,4} \in \mathbb{F}_2^{n_{11} - n_{22}, r_2}$.

$B. \text{ Weak 2: } n_{11} > n_{22} > n_{21} > n_{12}$

Note that when $n_{11} > n_{12} + n_{21}$, this regime is the same as in Sec III-C1. And for the regime with $n_{11} < n_{12} + n_{21}$ and $n_{11} + n_{22} - 2n_{21} - n_{12} > 0$, the generator matrices are the same as that in the regime in Appendix A-A1 because the inequality $n_{11} + n_{22} - 2n_{21} - n_{12} > 0$ implies $n_{11} + n_{22} - 2n_{21} - n_{12} > 0$.

$I)$: Consider $n_{11} < n_{12} + n_{21}$ and $n_{11} + n_{22} - 2n_{21} - n_{12} < 0$. The non-trivial corner points on the capacity region are $(n_{11} + n_{21} - n_{22}, 2(n_{22} - n_{21})), (2(n_{11} - n_{21}), n_{22} + n_{21} - n_{11})$. 
Ia): Under the condition in Appendix [A-B1], we further consider $3n_{21} + n_{12} - 2n_{22} - n_{11} < 0$. The generator matrices for achieving $(n_{11} + n_{21} - n_{22}, 2(n_{22} - n_{21}))$ are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
F_{1,3} \\
F_{1,4}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0^{n_{12} + n_{21} - n_{22}, r_2} \\
0^{n_{21} - n_{22}, r_2}
\end{bmatrix},
\]

(149)

where $F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{2n_{21} + n_{12} - n_{22}, r_1}, F_{1,3} \in \mathbb{F}_2^{3n_{21} - n_{12}, r_1}, F_{2,1}, F_{2,4} \in \mathbb{F}_2^{2n_{21} + n_{12} - n_{22}, r_2}, F_{2,2}, F_{2,3} \in \mathbb{F}_2^{3n_{21} - 2n_{22} - n_{11}, r_2}.$

When $3n_{21} + n_{12} - 2n_{22} - n_{11} > 0$, the generator matrices for achieving $(n_{11} + n_{21} - n_{22}, 2(n_{22} - n_{21}))$ are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
F_{1,3} \\
F_{1,4}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{12} + n_{21} - n_{22}, r_2} \\
0^{n_{21} - n_{22}, r_2} \\
F_{2,2}
\end{bmatrix},
\]

(150)

where $F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{2n_{21} - n_{22}, r_1}, F_{1,3} \in \mathbb{F}_2^{3n_{21} - n_{12} - 2n_{22}, r_1}, F_{1,4} \in \mathbb{F}_2^{3n_{11} - n_{21}, r_1}, F_{2,1}, F_{2,2} \in \mathbb{F}_2^{n_{12} - n_{22}, r_2}.$

Ib): Under the condition in Appendix [A-B1], when $2n_{12} + n_{21} - n_{11} - n_{22} < 0$, the generator matrices for achieving $(2(n_{11} - n_{21}), n_{22} + n_{21} - n_{11})$ are

\[
E_1 = \begin{bmatrix}
0^{2n_{21} + n_{12} - n_{11} - n_{22}, r_1} \\
F_{1,1} \\
0^{n_{11} + n_{22} - 2n_{12} - n_{21}, r_1} \\
0^{n_{12} + n_{11} - n_{21}, r_1} \\
0^{n_{11} + 2n_{22} - n_{12} - 3n_{21}, r_1} \\
0^{n_{22} - n_{12}, n_{11}} \\
F_{1,2}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0^{n_{11} - n_{21}, r_2} \\
F_{2,3} \\
F_{2,4} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
\]

(151)

where $F_{1,1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}, F_{2,1}, F_{2,4} \in \mathbb{F}_2^{n_{22} - n_{21}, r_2}, F_{2,2} \in \mathbb{F}_2^{2n_{21} + n_{12} - n_{11} - n_{22}, r_2}, F_{2,3} \in \mathbb{F}_2^{n_{11} - n_{12}, r_2}.$
When $2n_{12} + n_{21} - n_{11} - n_{22} > 0$ and $2(n_{21} + n_{12} - n_{11}) - n_{22} < 0$, the generator matrices for achieving $(2(n_{11} - n_{21}), n_{22} + n_{21} - n_{11})$ are

$$E_1 = \begin{bmatrix} 0^{n_{21} - n_{12}, r_1} \\ F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ 0^{n_{22} - n_{12}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ F_{2,3} \\ F_{2,4} \\ F_{2,5} \\ 0^{n_{11} - n_{21}, r_2} \\ F_{2,6} \\ 0^{n_{11} - n_{22}, r_2} \end{bmatrix},$$

where $F_{1,1} \in F_2^{2(n_{12} + n_{21} - n_{11} - n_{22}, r_1}, F_{1,2} \in F_2^{2(n_{11} + n_{22} - 2n_{12} - 2n_{21}, r_1}, F_{1,3} \in F_2^{n_{11} - n_{21}, r_1}, F_{2,1}, F_{2,5} \in F_2^{\min\{2n_{12} + n_{21} - n_{11} - n_{22}, n_{22} - n_{21}\}, r_2}, F_{2,2}, F_{2,4} \in F_2^{\min\{2(n_{21} + n_{12} - n_{22}) - n_{11}, n_{21} - n_{12}\}, r_2}, F_{2,3} \in F_2^{\min\{n_{11} + n_{22} - 3n_{12} - n_{21}\}, r_2}, F_{2,6} \in F_2^{n_{22} - n_{12}, r_2}.$

When $2(n_{21} + n_{12} - n_{11}) - n_{22} > 0$ and $3n_{21} + 2n_{12} - 2n_{11} - 2n_{22} > 0$, the generator matrices for achieving $(2(n_{11} - n_{21}), n_{22} + n_{21} - n_{11})$ are

$$E_1 = \begin{bmatrix} 0^{n_{21} - n_{12}, r_1} \\ F_{1,1} \\ 0^{2(n_{21} + n_{12} - n_{11}) - n_{22}, r_1} \\ F_{1,1} \\ 0^{n_{21} - n_{12}, r_1} \\ 0^{n_{22} - n_{21}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ F_{2,3} \\ F_{2,4} \\ F_{2,5} \\ F_{2,6} \\ F_{2,7} \\ 0^{n_{11} - n_{21}, r_2} \\ 0^{n_{11} - n_{22}, r_2} \end{bmatrix},$$

where $F_{1,1} \in F_2^{n_{11} - n_{21}, r_1}, F_{1,2} \in F_2^{n_{22} - n_{21}, r_2}, F_{2,1}, F_{2,7} \in F_2^{\min\{3n_{21} + 2n_{12} - 2n_{11} - 2n_{22}, n_{21} - n_{12}\}, r_2}, F_{2,2}, F_{2,6} \in F_2^{\min\{3n_{21} + 2n_{12} - 2n_{11} - 2n_{22}, n_{21} - n_{12}\}, r_2}, F_{2,3}, F_{2,5} \in F_2^{\min\{n_{11} - n_{21}, 2n_{11} + 2n_{22} - 3n_{12} - 3n_{21}\}, r_2}, F_{2,4} \in F_2^{\min\{3n_{12} + n_{21} - n_{11} - 2n_{22}, 3n_{11} + 2n_{22} - 3n_{12} - 3n_{21}\}, r_2}, F_{2,8} \in F_2^{n_{22} - n_{12}, r_2}.$
When \(2(n_{21} + n_{12} - n_{11}) - n_{22} > 0\) and \(3n_{21} + 2n_{12} - 2n_{11} - 2n_{22} < 0\), the generator matrices for achieving \((2(n_{11} - n_{21}), n_{22} + n_{21} - n_{11})\) are

\[
E_1 = \begin{bmatrix} 0^{n_{21}-n_{12}, r_1} \\ F_{1,1} \\ 0^{2(n_{21}+n_{12}-n_{11})-n_{22}, r_1} \\ F_{1,1} \\ 0^{n_{21}-n_{12}, r_1} \\ 0^{n_{22}-n_{12}, r_1} \\ F_{1,2} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ F_{2,3} \\ F_{2,4} \\ F_{2,5} \\ F_{2,6} \\ F_{2,7} \\ 0^{n_{11}-n_{12}, r_2} \\ F_{2,8} \\ 0^{n_{11}-n_{22}, r_2} \end{bmatrix},
\]

where \(F_{1,1}, F_{1,2} \in \mathbb{F}_2^{n_{11}-n_{21}, r_1}, F_{2,1}, F_{2,7} \in \mathbb{F}_2^{2(n_{21}+n_{12}-n_{11})-n_{22}, r_2}, F_{2,2}, F_{2,6} \in \mathbb{F}_2^{\min\{2n_{11}+2n_{22}-3n_{21}-2n_{12}, n_{11}-n_{21}\}, r_2}, F_{2,3}, F_{2,5} \in \mathbb{F}_2^{\min\{n_{11}-n_{21}, 2n_{12}+2n_{21}-n_{11}-2n_{22}\}, r_2}, F_{2,4} \in \mathbb{F}_2^{\min\{3n_{12}+n_{21}-n_{11}-2n_{22}, n_{12}+3n_{21}-n_{11}-2n_{22}\}, r_2}, F_{2,9} \in \mathbb{F}_2^{n_{22}-n_{12}, r_2}.

C. Weak 3: \(n_{11} > n_{21} > n_{22} > n_{12}\)

1) : Consider \(n_{11} > n_{12} + n_{21} > n_{22}\), the non-trivial corner points on the capacity region are \((n_{11} + n_{12} - n_{22}, n_{22} - n_{12}), (n_{11} - n_{22} - n_{12}, n_{22})\).

To achieve \((n_{11} + n_{12} - n_{22}, n_{22} - n_{12})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0^{n_{22}-n_{12}, r_1} \\ F_{1,3} \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0^{n_{12}, r_2} \\ F_{2,1} \\ 0^{n_{11}-n_{22}, r_2} \end{bmatrix},
\]

where \(F_{1,1} \in \mathbb{F}_2^{n_{21}-n_{22}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{21}, r_1}, F_{1,3} \in \mathbb{F}_2^{n_{31}-n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22}-n_{12}, r_2}.

To achieve \((n_{11} - n_{22} - n_{12}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{22}, r_1} \\ F_{1,2} \\ 0^{n_{12}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{11}-n_{22}, r_2} \end{bmatrix},
\]

where \(F_{1,1} \in \mathbb{F}_2^{n_{21}-n_{22}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{11}-n_{12}-n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{22}-n_{12}, r_2}.

2) : For $n_{11} < n_{12} + n_{21}$, the non-trivial corner points on the capacity region are $(2n_{11} - n_{22} - n_{21}, n_{22} + n_{21} - n_{11}), (n_{21} - n_{22}, n_{22})$.

2a): Under the above condition in Appendix A-C2, we further consider $n_{11} + n_{22} - 2n_{12} - n_{21} > 0$. To achieve corner point $(2n_{11} - n_{22} - n_{21}, n_{22} + n_{21} - n_{11})$, we propose

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{n_{12} + n_{21} - n_{11}, r_1} \\
F_{1,2} \\
0^{n_{12} + n_{21} - 2n_{12} - n_{21}, r_1} \\
0^{n_{12} + n_{21} - n_{11}, r_1}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{11} - n_{21}, r_2} \\
F_{2,2} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
(157)
\]

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}, F_{1,2}, F_{1,3} \in \mathbb{F}_2^{n_{12} + n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{11}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2}$.

To achieve $(n_{21} - n_{22}, n_{22})$, we propose

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{n_{11} - n_{12} + n_{22} - n_{21}, r_1} \\
F_{1,2} \\
0^{n_{12} + n_{21} - n_{11}, r_1}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
(158)
\]

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2}$.

2b): Under the condition in Appendix A-C2, we further consider $n_{11} + n_{22} - 2n_{12} - n_{21} < 0$ and $2n_{12} + 2n_{21} - 2n_{11} - n_{22} > 0$. To achieve corner point $(2n_{11} - n_{22} - n_{21}, n_{22} + n_{21} - n_{11})$, we propose

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
0^{n_{22} - n_{12}, r_1}
\end{bmatrix},
E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
F_{2,4} \\
F_{2,5} \\
0^{n_{11} - n_{21}, r_2} \\
F_{2,6} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},
(159)
\]

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}, F_{1,2}, F_{1,3} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{\min\{2n_{12} + 2n_{21} - 2n_{11} - n_{22}, n_{11} - n_{21}\}, r_2}, F_{2,2} \in \mathbb{F}_2^{\min\{2n_{12} + 2n_{21} - 2n_{11} - n_{22}, n_{11} - n_{21}\}, r_2}, F_{2,3} \in \mathbb{F}_2^{\min\{2n_{12} + 2n_{21} - 2n_{11} - n_{22}, n_{11} - n_{21}\}, r_2}, F_{2,4} \in \mathbb{F}_2^{\min\{3n_{12} + 2n_{21} - 2n_{11} - 2n_{22}, n_{11} - n_{21}\}, r_2}, F_{2,5} \in \mathbb{F}_2^{\min\{2n_{12} + 2n_{21} - 2n_{11} - n_{22}, n_{22} - n_{12}\}, r_2}, F_{2,6} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2}$. 
2c): Under the condition in Appendix A-C2 we further consider $n_{11} + n_{22} - 2n_{12} - n_{21} < 0$ and $2n_{12} + 2n_{21} - 2n_{11} - n_{22} < 0$. To achieve corner point $(2n_{11} - n_{22} - n_{21}, n_{22} + n_{21} - n_{11})$, we propose

$$\begin{bmatrix}
F_{1,1} \\
0^{n_{22} - n_{12}, r_1} \\
F_{1,2} \\
F_{1,3} \\
F_{1,2} \\
0^{n_{22} - n_{12}, r_1} \\
F_{1,4}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
0^{n_{11} - n_{21}, r_2} \\
F_{2,4} \\
0^{n_{11} - n_{22}, r_2}
\end{bmatrix},$$

(160)

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}$, $F_{1,2} \in \mathbb{F}_2^{2n_{12} + n_{21} - n_{11} - n_{22}, r_1}$, $F_{1,3} \in \mathbb{F}_2^{2n_{11} + n_{22} - 2n_{12} - 2n_{21}, r_1}$, $F_{1,4} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}$, $F_{2,1}, F_{2,3} \in \mathbb{F}_2^{n_{12} - n_{11} - n_{22}, r_2}$, $F_{2,2} \in \mathbb{F}_2^{3n_{12} + n_{21} - n_{11} - 2n_{22}, r_2}$, $F_{2,4} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2}$.

The generator matrices for achieving $(n_{21} - n_{22}, n_{22})$ under the regimes given in 2b) and 2c) are the same as in (158).

D. Strong 1: $n_{12} > n_{21} > n_{11} > n_{22}$

Consider $n_{12} > n_{11} + n_{22} > n_{21}$, the non-trivial corner points on the capacity are $(n_{11}, n_{21} - n_{11}), (n_{21} - n_{22}, n_{22})$.

1a): Under the conditions given in Appendix A-D we further consider $n_{21} < n_{22} + n_{11} < n_{12}$.

To achieve $(n_{11}, n_{21} - n_{11})$, we propose

$$\begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
0^{n_{21} - n_{12}, r_1} \\
0^{n_{12} - n_{21}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{12} - n_{22}, r_2} \\
0^{n_{11}, r_2}
\end{bmatrix},$$

(161)

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}$, $F_{1,2} \in \mathbb{F}_2^{n_{11} + n_{22} - n_{21}, r_1}$, $F_{2,1} \in \mathbb{F}_2^{n_{21} - n_{12}, r_2}$.

To achieve corner point $(n_{21} - n_{22}, n_{22})$, we propose

$$\begin{bmatrix}
F_{1,1} \\
0^{n_{12} + n_{22} - n_{21}, r_1} \\
F_{2,1} \\
F_{2,2} \\
0^{n_{21} - n_{22}, r_1} \\
0^{n_{11} + n_{22} - n_{21}, r_2}
\end{bmatrix},$$

(162)

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, n_{12}}$, $F_{2,1} \in \mathbb{F}_2^{n_{22}, n_{12}}$, $F_{2,2} \in \mathbb{F}_2^{n_{12} - n_{11} - n_{22}, n_{12}}$. 
\textbf{Ib):} Now consider $n_{21} < n_{11} + n_{22} < n_{12}$ and $2n_{11} + n_{22} - n_{12} - n_{21} > 0$. The generator matrices for achieving corner point $(n_{11}, n_{21} - n_{11})$ are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
F_{1,3} \\
F_{1,4} \\
0^{n_{12} - n_{11}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{12} - n_{21}, r_2} \\
0^{n_{11} + n_{22} - n_{12}, r_2} \\
0^{n_{12} - n_{22}, r_2}
\end{bmatrix},
\] (163)

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{11}, r_1}$, $F_{1,2} \in \mathbb{F}_2^{n_{21} - n_{11}, r_1}$, $F_{1,3} \in \mathbb{F}_2^{n_{12} - n_{21}, r_1}$, $F_{1,4} \in \mathbb{F}_2^{n_{11} + n_{22} - n_{12} - n_{21}, r_1}$, $F_{2,1} \in \mathbb{F}_2^{n_{12} - n_{21}, r_2}$.

Now to achieve corner point $(n_{21} - n_{22}, n_{22})$, we need to further consider two subregimes.

When $2n_{11} + n_{22} - n_{12} - n_{21} > 0 > n_{11} + 2n_{22} - n_{12} - n_{21}$, the generator matrices are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{n_{11} + n_{22} - n_{21}, r_1} \\
0^{n_{21} - n_{11}, r_1} \\
0^{n_{12} - n_{21}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
0^{n_{12} + n_{21} - n_{11} - 2n_{22}, r_2} \\
0^{n_{12} - n_{21}, r_2} \\
F_{2,2}
\end{bmatrix},
\] (164)

where $F_{1,1} \in \mathbb{F}_2^{n_{11} + n_{22} - n_{12}, r_1}$, $F_{1,2} \in \mathbb{F}_2^{n_{11} + n_{22} - n_{11} - 2n_{22}, r_1}$, $F_{2,1} \in \mathbb{F}_2^{n_{12} - n_{11}, r_2}$, $F_{2,2} \in \mathbb{F}_2^{n_{11} + n_{22} - n_{12}, r_2}$.

When $2n_{11} + n_{22} - n_{12} - n_{21} > n_{11} + 2n_{22} - n_{12} - n_{21} > 0$ and $n_{12} + n_{21} - n_{11} - 2n_{22} < 0$, the generator matrices for achieving corner point $(n_{21} - n_{22}, n_{22})$ are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{n_{11} + n_{22} - n_{21}, r_1} \\
0^{n_{21} - n_{11}, r_1} \\
0^{n_{12} - n_{21}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
F_{2,2}
\end{bmatrix},
\] (165)

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}$, $F_{2,1} \in \mathbb{F}_2^{n_{12} - n_{11}, r_2}$, $F_{2,2} \in \mathbb{F}_2^{n_{21} - n_{22}, r_2}$, $F_{2,3} \in \mathbb{F}_2^{n_{22} + n_{11} - n_{12} - n_{21}, r_2}$. 
1c): When \( n_{12} < n_{12} < n_{22} + n_{11} \) and \( 2n_{11} + n_{22} - n_{12} - n_{21} < 0 \) which implies \( 2n_{22} + n_{11} - n_{12} - n_{21} < 0 \), to achieve corner point \((n_{11}, n_{21} - n_{11})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ 0^{n_{12}+n_{21}-2n_{11}-n_{22},r_1} \\ 0^{n_{12}-n_{21},r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0^{n_{12}-n_{21},r_2} \\ F_{2,1} \\ F_{2,2} \\ 0^{n_{11}+n_{22}-n_{12},r_2} \\ 0^{n_{12}-n_{22},r_2} \end{bmatrix}, \quad (166)
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{21}-n_{22},r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12}-n_{21},r_1}, F_{1,3} \in \mathbb{F}_2^{n_{11}+n_{22}-n_{12},r_1}, F_{2,1} \in \mathbb{F}_2^{n_{11}+n_{22}-n_{12},r_2}, F_{2,2} \in \mathbb{F}_2^{n_{12}+n_{21}-2n_{11}-n_{22},r_2} \).

To achieve corner point \((n_{21} - n_{22}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0^{n_{12}+n_{22}-n_{12},r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{11}+n_{22}-n_{12}-n_{21},r_2} \\ 0^{n_{12}-n_{21},r_2} \end{bmatrix}, \quad (167)
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{11}+n_{22}-n_{12},r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12}+n_{21}-2n_{22}-n_{11},r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12}-n_{11},r_2}, F_{2,2} \in \mathbb{F}_2^{n_{11}+n_{22}-n_{12},r_2} \).

E. Strong 2: \( n_{21} > n_{12} > n_{11} > n_{22} \)

The non-trivial corner points on the capacity region are \((n_{11}, n_{12} - n_{11}), (n_{12} - n_{22}, n_{22})\).

1a): For \( n_{12} < n_{11} + n_{22} < n_{21} \), the generator matrices for achieving \((n_{11}, n_{12} - n_{11})\) are

\[
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{21}-n_{11}-n_{22},r_1} \\ 0^{n_{22},n_{21}} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{11}+n_{22}-n_{12},r_2} \\ 0^{n_{12}-n_{21},n_{21}} \end{bmatrix}, \quad (168)
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{11}-r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12}-n_{11},r_2} \).

Then, the generator matrices for achieving \((n_{12} - n_{22}, n_{22})\) are

\[
E_1 = \begin{bmatrix} 0^{n_{11}+n_{22}-n_{12},r_1} \\ F_{1,1} \\ 0^{n_{21}-n_{11}-n_{22},r_1} \\ 0^{n_{22},n_{21}} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{12}-n_{22},r_2} \\ 0^{n_{21}-n_{12},r_2} \end{bmatrix}, \quad (169)
\]
where $F_{1,1} \in \mathbb{F}_{2}^{n_{12} - n_{21}, r_1}$, $F_{2,1} \in \mathbb{F}_{2}^{n_{12} - n_{11}, r_2}$, $F_{2,2} \in \mathbb{F}_{2}^{n_{11} + n_{22} - n_{12}, r_2}$.

1b): Now consider $n_{12} < n_{21} < n_{11} + n_{22}$ and $n_{12} + n_{21} - 2n_{11} - n_{22} > 0 \Rightarrow n_{12} + n_{21} - 2n_{22} - n_{11} > 0$. The generator matrices for achieving $(n_{11}, n_{12} - n_{11})$ are

$$
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \\ 0^{n_{21} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ F_{2,3} \\ 0^{n_{21} - n_{12}, r_1} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{21} - n_{12}, r_2} \end{bmatrix},
$$

(170)

where $F_{1,1}, F_{1,4} \in \mathbb{F}_{2}^{n_{11} + n_{22} - n_{21}, r_1}$, $F_{1,2} \in \mathbb{F}_{2}^{n_{11} + n_{21} - n_{12}, r_1}$, $F_{1,3} \in \mathbb{F}_{2}^{n_{12} + n_{21} - 2n_{22} - n_{11}, r_1}$, $F_{2,1} \in \mathbb{F}_{2}^{n_{11} + n_{22} - n_{21}, r_1}$, $F_{2,2} \in \mathbb{F}_{2}^{n_{12} + n_{21} - 2n_{11} - n_{22}, r_2}$.

Then, the generator matrices for achieving $(n_{12} - n_{22}, n_{22})$ are

$$
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{21} - n_{12}, r_1} \\ F_{1,2} \\ F_{1,1} \\ 0^{n_{21} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ F_{2,3} \\ F_{2,4} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{21} - n_{12}, r_2} \end{bmatrix},
$$

(171)

where $F_{1,1} \in \mathbb{F}_{2}^{n_{11} + n_{22} - n_{21}, r_1}$, $F_{1,2} \in \mathbb{F}_{2}^{n_{12} + n_{21} - n_{11} - 2n_{22}, r_1}$, $F_{2,1}, F_{2,3} \in \mathbb{F}_{2}^{n_{11} + n_{22} - n_{21}, r_2}$, $F_{2,2} \in \mathbb{F}_{2}^{n_{11} + n_{22} - n_{21}, r_2}$, $F_{2,4} \in \mathbb{F}_{2}^{n_{11} + n_{21} - 2n_{11} - n_{22}, r_2}$.

1c): Now consider $2n_{11} + n_{22} - n_{12} - n_{21} > 0$. The generator matrices for achieving $(n_{11}, n_{12} - n_{11})$ are

$$
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \\ F_{1,5} \\ 0^{n_{21} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{2n_{11} + n_{22} - n_{12} - n_{21}, r_2} \\ F_{2,1} \\ 0^{n_{21} - n_{12}, r_2} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{21} - n_{12}, r_2} \end{bmatrix},
$$

(172)

where $F_{1,1}, F_{1,5} \in \mathbb{F}_{2}^{\min\{n_{21} - n_{22}, 2n_{11} + n_{22} - n_{12}, n_{21}\}, r_1}$, $F_{1,2}, F_{1,4} \in \mathbb{F}_{2}^{\min\{2n_{21} + n_{12} - 2n_{11} - 2n_{22}, n_{12} - n_{11}\}, r_1}$, $F_{1,3} \in \mathbb{F}_{2}^{\min\{2(n_{12} + n_{21} - n_{22}) - 3n_{11}, |n_{11} + 2(n_{22} - n_{21})|\}, r_1}$, $F_{2,1} \in \mathbb{F}_{2}^{n_{11} - n_{12}, r_2}$. 
Then, the generator matrices for achieving $(n_{12} - n_{22}, n_{22})$ are

$$E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{21} - n_{12}, r_1} \\ F_{1,2} \\ 0^{n_{21} - n_{11} - n_{22}, r_1} \\ 0^{n_{22}, r_1} \end{bmatrix}, E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{21} - n_{12}, r_2} \end{bmatrix}, \quad (173)$$

where $F_{1,1} \in F_2^{n_{12} - n_{22}, r_1}, F_{2,1} \in F_2^{n_{11} - n_{12}, r_2}, F_{2,2} \in F_2^{n_{11} + n_{22} - n_{12}, r_2}$.

**F. Strong 3**: $n_{21} > n_{11} > n_{12} > n_{22}$

For this regime, the non-trivial corner point on the capacity region is $(n_{11} - n_{22}, n_{22})$. The case here is further divided into four subregimes in the following.

1a): When $n_{21} > n_{11} + n_{22}$, the generator matrices are

$$E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{22}, r_1} \\ F_{1,2} \\ 0^{n_{21} - n_{11} - n_{22}, r_1} \\ 0^{n_{22}, r_1} \end{bmatrix}, E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{21} - n_{12}, r_2} \end{bmatrix}, \quad (174)$$

where $F_{1,1} \in F_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in F_2^{n_{11} - n_{22}, r_2}, F_{2,1} \in F_2^{n_{22}, r_2}$.

1b): When $n_{21} < n_{11} + n_{22}$ and $n_{12} + n_{21} - n_{11} - 2n_{22} > 0$, the generator matrices are

$$E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0^{n_{21} - n_{11}, r_1} \\ F_{1,3} \\ F_{1,2} \\ 0^{n_{21} - n_{11}, r_1} \end{bmatrix}, E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{21} - n_{12}, r_2} \end{bmatrix}, \quad (175)$$

where $F_{1,1} \in F_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in F_2^{n_{11} + n_{22} - n_{21}, r_1}, F_{1,3} \in F_2^{n_{11} + n_{21} - n_{22} - n_{12}, r_1}, F_{2,1} \in F_2^{n_{11} + n_{22} - n_{21}, r_2}, F_{2,2} \in F_2^{n_{21} - n_{11}, r_2}$. 
Ic): When \( n_{21} < n_{11} + n_{22}, n_{12} + n_{21} - n_{11} - 2n_{22} < 0 \) and \( 3n_{22} + 2(n_{11} - n_{12} - n_{21}) > 0 \), the generator matrices are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
0^{n_{21} - n_{11}, r_1} \\
0^{n_{11} + 2n_{22} - n_{12} - n_{21}, r_1} \\
F_{1,2} \\
0^{n_{21} - n_{11}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
F_{2,4} \\
F_{2,5} \\
0^{n_{12} - n_{22}, r_2}
\end{bmatrix},
\]

(176)

where \( F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, \ F_{1,2} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}, \ F_{2,1}, F_{2,4} \in \mathbb{F}_2^{n_{12} - n_{22}, r_2}, \ F_{2,2}, F_{2,5} \in \mathbb{F}_2^{n_{21} - n_{11}, r_2}, \ F_{2,3} \in \mathbb{F}_2^{3n_{22} + 2(n_{11} - n_{12} - n_{21}), r_2}.

Id): When \( n_{21} < n_{11} + n_{22}, n_{12} + n_{21} - n_{11} - 2n_{22} < 0 \) and \( 3n_{22} + 2(n_{11} - n_{12} - n_{21}) < 0 \), the generator matrices are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
0^{n_{21} - n_{11}, r_1} \\
0^{n_{11} + 2n_{22} - n_{12} - n_{21}, r_1} \\
F_{1,2} \\
0^{n_{21} - n_{11}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
F_{2,4} \\
F_{2,5} \\
0^{n_{12} - n_{22}, r_2}
\end{bmatrix},
\]

(177)

where \( F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, \ F_{1,2} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}, \ F_{2,1}, F_{2,4} \in \mathbb{F}_2^{\min(n_{12} - n_{22}, n_{11} + 2n_{22} - n_{12} - n_{21}), r_2}, \ F_{2,2} \in \mathbb{F}_2^{2n_{12} + n_{21} - n_{11} - 3n_{22}, r_2}, \ F_{2,3} \in \mathbb{F}_2^{\min(n_{12} - n_{22}, n_{21} - n_{11}, n_{11} + 2n_{22} - n_{12} - n_{21}, 2(n_{12} + n_{21} - n_{11}) - 3n_{22}), r_2}, \ F_{2,4} \in \mathbb{F}_2^{2n_{21} + n_{12} - 2n_{11} - 2n_{22}, r_2}, \ F_{2,5} \in \mathbb{F}_2^{\min(n_{21} - n_{11}, n_{11} + 2n_{22} - n_{12} - n_{21}), r_2}.

G. Mixed 1: \( n_{11} > n_{12} > n_{22} > n_{21} \)

I) : Consider \( n_{11} > n_{12} + n_{21} > n_{22} \), the non-trivial corner points are \( (n_{11} + n_{21} - n_{22}, n_{22} - n_{21}), (n_{11} - n_{21} - n_{22}, n_{22}) \).
To achieve \((n_{11} + n_{21} - n_{22}, n_{22} - n_{21})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0^{n_{22} - n_{21}, r_1} \\ F_{1,3} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{21}, r_2} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{11} - n_{12}, r_2} \end{bmatrix}
\] (178)

where \(F_{1,1} \in \mathbb{F}_2^{n_{21}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{21} - n_{12}, r_1}, F_{1,3} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{32}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{21}, r_2}\).

To achieve \((n_{11} - n_{21} - n_{22}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} 0^{n_{21}, r_1} \\ F_{1,1} \\ 0^{n_{22}, r_1} \\ F_{1,2} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{12} - n_{22}, r_2} \\ F_{2,2} \end{bmatrix}
\] (179)

where \(F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{21} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22}, r_2} \) and \(F_{2,2} \in \mathbb{F}_2^{n_{11} - n_{12}, r_2}\).

2) : Now for \(n_{22} < n_{11} < n_{12} + n_{21}\), the non-trivial corner points are \((2n_{11} - n_{12} - n_{22}, n_{12} + n_{22} - n_{11}), (n_{12} - n_{22}, n_{22})\).

2a): Under the conditions in Appendix A-G2 we further consider \(n_{11} + n_{22} - 2n_{21} - n_{12} > 0\).

To achieve \((2n_{11} - n_{12} - n_{22}, n_{12} + n_{22} - n_{11})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{12} + n_{21} - n_{11}, r_1} \\ 0^{n_{11} + n_{22} - 2n_{21} - n_{12}, r_1} \\ F_{1,2} \\ 0^{n_{12} + n_{22} - n_{12}, r_1} \\ F_{1,3} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{11} - n_{12}, r_2} \\ 0^{n_{12} - n_{22}, r_2} \\ 0^{n_{11} - n_{12}, r_2} \end{bmatrix}
\] (180)

where \(F_{1,1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{12} - n_{12}, r_1}, F_{1,3} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}, F_{2,1}, F_{2,3} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{11}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{11} + n_{22} - 2n_{21} - n_{12}, r_2}\).

To achieve \((n_{12} - n_{22}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} 0^{n_{11} - n_{12}, r_1} \\ 0^{n_{12} + n_{21} - n_{11}, r_1} \\ 0^{n_{11} + n_{22} - 2n_{21} - n_{12}, r_1} \\ 0^{n_{21}, r_1} \\ F_{1,1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{12} - n_{22}, r_2} \\ F_{2,2} \\ F_{2,3} \\ 0^{n_{11} - n_{12}, r_2} \end{bmatrix}
\] (181)

where \(F_{1,1} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{11}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{11} + n_{22} - 2n_{21} - n_{12}, r_2}, F_{2,3} \in \mathbb{F}_2^{n_{12}, r_2}\).
2b): Under the conditions in Appendix A-G2 we further consider the subregime of $n_{11} + n_{22} - 2n_{21} - n_{12} < 0$ and $2n_{21} + 2n_{12} - 2n_{11} - n_{22} < 0$. To achieve the first corner point $(2n_{11} - n_{12} - n_{22}, n_{12} + n_{22} - n_{11})$, we propose the following generator matrices

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
F_{1,2} \\
0^{n_{22} - n_{21}, r_1} \\
F_{1,3} \\
0^{n_{22} - n_{21}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0^{2n_{11} + n_{22} - 2n_{12} - 2n_{21}, r_2} \\
F_{2,3} \\
0^{n_{12} - n_{22}, r_2} \\
0^{n_{11} - n_{12}, r_2}
\end{bmatrix},
\]

where $F_{1,1}, F_{1,4} \in \mathbb{F}_2^{2n_{21} + n_{12} - n_{11} - n_{22}, r_1}$, $F_{1,2}, F_{1,3} \in \mathbb{F}_2^{2n_{11} + n_{22} - 2n_{12} - 2n_{21}, r_1}$, $F_{1,5} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}$, $F_{2,1}, F_{2,3} \in \mathbb{F}_2^{2n_{22} - n_{12}, r_2}$, $F_{2,2} \in \mathbb{F}_2^{2n_{21} + n_{12} - n_{11} - n_{22}, r_2}$.

2c): Under the conditions in Appendix A-G2 and with $n_{11} + n_{22} - 2n_{21} - n_{12} < 0$ and $2n_{21} + 2n_{12} - 2n_{11} - n_{22} > 0$, the generator matrices for achieving the first corner point $(2n_{11} - n_{12} - n_{22}, n_{12} + n_{22} - n_{11})$ are

\[
E_1 = \begin{bmatrix}
F_{1,1} \\
0^{n_{22} - n_{21}, r_1} \\
F_{1,1} \\
0^{n_{22} - n_{12}, r_1} \\
F_{1,2} \\
0^{2n_{21} + 2n_{12} - 2n_{11} - n_{22}, r_1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
F_{2,3} \\
0^{n_{12} - n_{22}, r_2} \\
0^{n_{11} - n_{12}, r_2}
\end{bmatrix},
\]

where $F_{1,1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}$, $F_{1,3} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}$, $F_{2,1}, F_{2,3} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2}$, $F_{2,2} \in \mathbb{F}_2^{n_{11} - n_{12}, r_2}$, $F_{2,4} \in \mathbb{F}_2^{2n_{21} + 2n_{12} - 2n_{11} - n_{22}, r_2}$.

2d): For the regimes given in 2a)-2c) in Appendix A-G, the generator matrices for achieving the second corner point $(n_{12} - n_{22}, n_{22})$ are the same

\[
E_1 = \begin{bmatrix}
0^{n_{11} - n_{12}, r_1} \\
0^{n_{12} + n_{21} - n_{11}, r_1} \\
0^{n_{11} + n_{22} - n_{12} - n_{21}, r_1} \\
F_{1,1}
\end{bmatrix}, \quad E_2 = \begin{bmatrix}
F_{2,1} \\
F_{2,2} \\
0^{n_{12} - n_{22}, r_2} \\
0^{n_{11} - n_{12}, r_2}
\end{bmatrix},
\]

(184)
where $F_{1,1} \in F_{2}^{n_{12}-n_{22}, r_1}$, $F_{2,1} \in F_{2}^{n_{22}-n_{21}, r_2}$, $F_{2,2} \in F_{2}^{n_{21}, r_2}$.

**H. Mixed 2: $n_{11} > n_{21} > n_{12} > n_{22}$**

1) We first consider $0 < n_{12} + n_{21} - n_{11} < n_{22}$. The non-trivial corner points are $(2n_{11} - n_{12} - n_{21}, n_{12} + n_{21} - n_{11}), (n_{12} + n_{21} - 2n_{22}, n_{22})$.

When $2(n_{12} + n_{21} - n_{11}) - n_{22} > 0$, the generator matrices for achieving $(2n_{11} - n_{12} - n_{21}, n_{12} + n_{21} - n_{11})$ are

$$E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{11}+n_{22}-n_{12}-n_{21}, r_1} \\ 0^{n_{12}-n_{22}, r_2} \\ 0^{n_{11}+n_{12}-n_{21}, r_2} \end{bmatrix},$$

(185)

where $F_{1,1} \in F_{2}^{n_{21}-n_{22}, r_1}$, $F_{1,2}, F_{1,3} \in F_{2}^{n_{11}+n_{22}-n_{12}-n_{21}, r_1}$, $F_{1,4} \in F_{2}^{n_{12}-n_{21}, r_1}$, $F_{2,1}, F_{2,3} \in F_{2}^{3n_{12}+3n_{21}-3n_{11}-2n_{22}, r_2}$.

When $2(n_{12} + n_{21} - n_{11}) - n_{22} < 0$, the generator matrices for achieving $(2n_{11} - n_{12} - n_{21}, n_{12} + n_{21} - n_{11})$ are

$$E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \\ F_{1,5} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{11}+n_{22}-n_{12}-n_{21}, r_2} \\ 0^{n_{11}-n_{21}, r_2} \end{bmatrix},$$

(186)

where $F_{1,1} \in F_{2}^{n_{21}-n_{22}, r_1}$, $F_{1,2} \in F_{2}^{n_{12}+n_{21}-n_{11}, r_1}$, $F_{1,3} \in F_{2}^{n_{12}+2(n_{11} - n_{12} - n_{21}), r_1}$, $F_{1,4} \in F_{2}^{n_{11}+n_{21}-n_{12}-n_{21}, r_1}$, $F_{1,5} \in F_{2}^{n_{12}-n_{22}, r_1}$, $F_{2,1} \in F_{2}^{n_{12}+n_{21}-n_{11}, r_2}$.

To achieve $(n_{12} + n_{21} - 2n_{22}, n_{22})$, we propose

$$E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{11}+n_{22}-n_{12}-n_{21}, r_1} \\ 0^{n_{11}+n_{22}-n_{12}-n_{21}, r_1} \\ 0^{n_{11}+n_{22}-n_{12}-n_{21}, r_1} \\ F_{1,2} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{11}-n_{21}, r_2} \\ 0^{n_{11}-n_{21}, r_2} \end{bmatrix},$$

(187)

where $F_{1,1} \in F_{2}^{n_{21}-n_{22}, r_1}$, $F_{1,2} \in F_{2}^{n_{12}-n_{22}, r_1}$, $F_{2,1} \in F_{2}^{n_{22}, r_2}$.
2) : For $n_{12} + n_{21} > n_{11} + n_{22}$, the non-trivial corner point is $(n_{11} - n_{22}, n_{22})$. We further consider two subregimes under this condition in the following.

2): When $2n_{22} + n_{11} - n_{12} - n_{21} > 0$, to achieve this corner point, we propose

$$E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ F_{2,3} \end{bmatrix}, \quad (188)$$

where $F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}$, $F_{1,2} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{11} - n_{22}, r_1}$, $F_{1,3} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}$, $F_{2,1}, F_{2,3} \in \mathbb{F}_2^{\min(2n_{22} + n_{11} - n_{12} - n_{21}, n_{12} + n_{21} - n_{11} - n_{22}), r_2}$, $F_{2,2} \in \mathbb{F}_2^{3n_{22} + 2(n_{11} - n_{12} - n_{21}), r_2}$.

2b): When $2n_{22} + n_{11} - n_{12} - n_{21} < 0$, to achieve $(n_{11} - n_{22}, n_{22})$, we propose

$$E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \\ F_{1,4} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{12} + n_{21} - n_{11} - 2n_{22}, r_2} \\ 0^{n_{11} - n_{21}, n_{11}} \\ 0^{n_{11} - n_{12}, n_{11}} \end{bmatrix}, \quad (189)$$

where $F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}$, $F_{1,2} \in \mathbb{F}_2^{n_{22}, r_1}$, $F_{1,3} \in \mathbb{F}_2^{n_{12} + n_{21} - n_{11} - 2n_{22}, r_1}$, $F_{1,4} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}$, $F_{2,1} \in \mathbb{F}_2^{n_{22}, r_2}$.

3) : When $n_{11} > n_{12} + n_{21}$, the corner point is $(n_{11} - 2n_{22}, n_{22})$. To achieve this corner point, we propose

$$E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ F_{1,3} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0^{n_{12} - 2n_{22}, r_2} \\ 0^{n_{11} - n_{12}, r_2} \end{bmatrix}, \quad (190)$$

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}$, $F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{12} - n_{21}, r_1}$, $F_{1,3} \in \mathbb{F}_2^{n_{12} - n_{22}, r_1}$, $F_{2,1} \in \mathbb{F}_2^{n_{22}, r_2}$.

1. **Mixed 3**: $n_{11} > n_{12} > n_{21} > n_{22}$

The capacity regions and the achievable schemes for this regime are the same as those in Mixed 2 in Appendix A-H.
J. Mixed 4: $n_{12} > n_{11} > n_{21} > n_{22}$

For this regime, the non-trivial corner point on the capacity region is $(n_{11} - n_{22}, n_{22})$.

1a): When $n_{12} > n_{11} + n_{22}$, the generator matrices for achieving the corner point $(n_{11} - n_{22}, n_{22})$ are

$$E_1 = \begin{bmatrix} F_{1,1} \\ 0_{n_{22}, r_1} \\ F_{1,2} \\ 0_{n_{12} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0_{n_{12} - n_{11}, n_{22}, r_2} \end{bmatrix}$$

where $F_{1,1} \in \mathbb{F}_2^{n_{22} + n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12} - n_{22} - n_{11}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12} - n_{11}, r_2}$.

1b): We then consider the regime of $n_{12} < n_{11} + n_{22}$ and $2n_{22} + n_{11} - n_{12} - n_{21} < 0$. The generator matrices for achieving corner point $(n_{11} - n_{22}, n_{22})$ are

$$E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0_{n_{22}, r_1} \\ F_{1,3} \\ 0_{n_{12} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0_{n_{12} + n_{21} - 2n_{22} - n_{11}, r_2} \\ F_{2,2} \\ 0_{n_{12} - n_{11}, r_2} \\ 0_{n_{11} - n_{21}, r_2} \end{bmatrix}$$

where $F_{1,1} \in \mathbb{F}_2^{n_{22} + n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12} + n_{21} - 2n_{22} - n_{11}, r_1}, F_{1,3} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12} - n_{11}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{12} + n_{11} - n_{12}, r_2}$.

1c): When $n_{12} < n_{11} + n_{22}$ and $2n_{22} + n_{11} - n_{12} - n_{21} > 0$, the generator matrices for achieving corner point $(n_{11} - n_{22}, n_{22})$ are

$$E_1 = \begin{bmatrix} F_{1,1} \\ 0_{n_{22}, r_1} \\ F_{1,2} \\ 0_{n_{12} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ F_{2,3} \\ 0_{n_{12} - n_{11}, r_2} \end{bmatrix}$$

where $F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12} - n_{11}, r_2}, F_{2,2} \in \mathbb{F}_2^{n_{21} - n_{22}, r_2}, F_{2,3} \in \mathbb{F}_2^{n_{21} - n_{22}, r_2}$.
K. Mixed 5: \( n_{12} > n_{11} > n_{22} > n_{21} \)

1) : Consider \( n_{12} > n_{11} + n_{22} - n_{21} \), then the corner points are \( (n_{11}, n_{12} - n_{11}), (n_{12} - n_{22}, n_{22}) \).

To achieve corner point \((n_{11}, n_{12} - n_{11})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0_{n_{12} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0_{n_{11} + n_{22} - n_{12} - n_{21}, r_2} \\ 0_{n_{21}, n_{12}} \\ 0_{n_{12} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}^n_{2, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{21} - n_{12}, r_1}, F_{2,1} \in \mathbb{F}_{2}^{n_{11} - n_{22}, r_2} \).

To achieve corner point \((n_{12} - n_{22}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} 0_{n_{21}, r_1} \\ 0_{n_{11} + n_{22} - n_{12} - n_{21}, r_1} \\ F_{1,1} \\ 0_{n_{12} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0_{n_{12} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{21}, r_2}, F_{2,2} \in \mathbb{F}_{2}^{n_{11} + n_{22} - n_{12} - n_{21}, r_2} \).

2) : When \( n_{12} < n_{11} + n_{22} - n_{21} \), the corner points are \((n_{11}, n_{22} - n_{21}), (n_{11} - n_{21}, n_{22})\).

To achieve \((n_{11}, n_{22} - n_{21})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0_{n_{12} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ 0_{n_{12} + n_{21} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{21}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{21}, r_2} \).

To achieve \((n_{11} - n_{21}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} 0_{n_{21}, r_1} \\ F_{1,1} \\ 0_{n_{12} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0_{n_{12} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{21}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{21}, r_2}, F_{2,2} \in \mathbb{F}_{2}^{n_{21}, r_2} \).
L. Mixed 6: \( n_{21} > n_{11} > n_{22} > n_{12} \)

1) : When \( n_{11} + n_{22} - n_{12} - n_{21} > 0 \), the non-trivial corner points are \((n_{11}, n_{21} - n_{11}), (n_{21} - n_{22}, n_{22})\). To achieve \((n_{11}, n_{21} - n_{11})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0^{n_{21} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} O^{n_{12}, r_2} \\ 0^{n_{11} + n_{22} - n_{12} - n_{21}, r_2} \\ F_{2,1} \\ 0^{n_{21} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{21} - n_{11}, r_1} \).

To achieve \((n_{21} - n_{22}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{11} + n_{22} - n_{12} - n_{21}, r_1} \\ 0^{n_{12}, r_1} \\ 0^{n_{21} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{21} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{21} - n_{22}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12}, r_1}, F_{2,2} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2} \).

2) : When \( n_{11} + n_{22} - n_{12} - n_{21} < 0 \), the corner points are \((n_{11}, n_{22} - n_{12}), (n_{11} - n_{12}, n_{22})\).

To achieve \((n_{11}, n_{22} - n_{12})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ F_{1,2} \\ 0^{n_{21} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0^{n_{12}, n_{21}} \\ F_{2,1} \\ 0^{n_{21} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, F_{1,2} \in \mathbb{F}_2^{n_{12}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2} \).

To achieve \((n_{11} - n_{12}, n_{22})\), we propose

\[
E_1 = \begin{bmatrix} F_{1,1} \\ 0^{n_{12} + n_{21} - n_{11}, r_1} \end{bmatrix}, \quad E_2 = \begin{bmatrix} F_{2,1} \\ F_{2,2} \\ 0^{n_{21} - n_{22}, r_2} \end{bmatrix},
\]

where \( F_{1,1} \in \mathbb{F}_2^{n_{11} - n_{12}, r_1}, F_{2,1} \in \mathbb{F}_2^{n_{12}, r_1}, F_{2,2} \in \mathbb{F}_2^{n_{22} - n_{12}, r_2} \).

**APPENDIX B**

**USEFUL LEMMAS**

**Lemma 1.** The normalization factor in (69) for user \( k \in \{1, 2\} \), i.e., \( \gamma_k = \frac{1}{\sqrt{E[|x_k|^2]}} \) satisfy the following condition

\[
\gamma_k > 2^{-q}.
\]
Proof: According to (69), we have
\[
\gamma_k = \frac{1}{\sqrt{\mathbb{E}[\| \sum_{i_k=1}^{L_k} 2^{\text{row}(E_k[\ldots,F_{k,i_k}^T])} \rho_{k,i_k} F_{k,i_k} \|^2]}}
\]
\[
\underset{(a)}{\geq} \frac{1}{\max_{i_k \in [1:L_k]} \{\rho_{k,i_k}\}} \sqrt{\mathbb{E}[\| 2^{\text{row}(E'_k)} E'_k \|^2]}
\]
\[
> \frac{1}{2} \sqrt{\frac{12}{2^{2q} - 1}}
\]
\[
> 2^{-q}, \quad (203)
\]
where (a) follows that \(\mathbb{E}[\| \sum_{i_k=1}^{L_k} 2^{\text{row}(E_k[\ldots,F_{k,i_k}^T])} F_{k,i_k} \|^2] \leq \mathbb{E}[\| 2^{\text{row}(E'_k)} E'_k \|^2]\) since \(E'_k \in \mathbb{P}^{q,r}\) is a full rank matrix and \(E'_k\) is uniformly distributed over PAM(\(2^q, 1\)). \qed

In what follows, we provide some properties of superimposed constellations.

**Lemma 2.** Let \(P_2 > P_1 > 0\) be two real constants. Let \((\Lambda_1, \Lambda_2)\) be two one-dimensional constellations (not necessarily regular) with \(d_{\min}(\Lambda_1) > 0\) and \(d_{\min}(\Lambda_2) > 0\), respectively. When the following condition holds
\[
P_1(\min\{\Lambda_1\} - \max\{\Lambda_1\}) + P_2d_{\min}(\Lambda_2) > 0,
\]
the minimum distance of the superimposed constellation \(P_1\Lambda_1 + P_2\Lambda_2\) satisfies
\[
d_{\min}(P_1\Lambda_1 + P_2\Lambda_2) = \min\{P_1(\min\{\Lambda_1\} - \max\{\Lambda_1\}) + P_2d_{\min}(\Lambda_2), P_1d_{\min}(\Lambda_1)\}.
\]

**Proof:** First, we let \(\lambda_1, \lambda_2 \in \Lambda_2\) and assume \(\lambda_1 > \lambda_2\) without loss of generality. Then, we compute the inter-constellation distance metric of \(P_1\Lambda_1 + P_2\Lambda_2\) according to Definition [I] as
\[
d_{\text{IC}}(P_1\Lambda_1 + P_2\Lambda_2) = \min_{\lambda_1, \lambda_2 \in \Lambda_2} \{P_1 \min\{\Lambda_1\} + P_2\lambda_1 - P_1 \max\{\Lambda_1\} - P_2\lambda_2\}
\]
\[
= P_1(\min\{\Lambda_1\} - \max\{\Lambda_1\}) + P_2d_{\min}(\Lambda_2).
\]
Hence, when \(d_{\text{IC}}(P_1\Lambda_1 + P_2\Lambda_2) > 0\), the minimum distance \(d_{\min}(P_1\Lambda_1 + P_2\Lambda_2)\) is either \(d_{\text{IC}}(P_1\Lambda_1 + P_2\Lambda_2)\) or \(d_{\min}(P_1\Lambda_1)\). \qed

Note that when \(d_{\text{IC}}(P_1\Lambda_1 + P_2\Lambda_2) < 0\), Lemma 2 does not hold anymore.

**Lemma 3.** Consider a superimposed constellation
\[
\Lambda_\Sigma = \sum_{l=1}^L P_l\Lambda_l,
\]
where \( L \in \mathbb{Z}^+, L > 1, P_l \in \mathbb{R}^+, P_l > P_{l-1}, \) and \( \Lambda_l \) is a one-dimensional constellation (not necessarily regular) with \( d_{\min}(\Lambda_l) > 0. \) If the following condition holds for all \( l \in [1 : L - 1] \)

\[
d_{\min}(P_l\Lambda_l + P_{l+1}\Lambda_{l+1}) = P_ld_{\min}(\Lambda_l),
\]

(204)

the minimum distance of \( \Lambda_{\Sigma} \) satisfies

\[
d_{\min}(\Lambda_{\Sigma}) = P_1d_{\min}(\Lambda_1).
\]

**Proof:** We proof this lemma by induction.

First, substituting \( l = 1 \) into (204) gives

\[
d_{\min}(P_1\Lambda_1 + P_2\Lambda_2) = P_1d_{\min}(\Lambda_1).
\]

Next, we assume that for \( l > 2, \)

\[
d_{\min}\left(\sum_{i=1}^{l} P_i\Lambda_i\right) = P_1d_{\min}(\Lambda_1).
\]

(205)

According to Lemma 2, we note from (204) that

\[
P_l(\min\{\Lambda_l\} - \max\{\Lambda_l\}) + P_{l+1}d_{\min}(\Lambda_{l+1}) \geq P_l d_{\min}(\Lambda_l),
\]

(206)

By summing the inequality in (206) from 1 to \( l, \) we get

\[
\sum_{i=1}^{l} (P_i(\min\{\Lambda_i\} - \max\{\Lambda_i\}) + P_{i+1}d_{\min}(\Lambda_{i+1})) \geq \sum_{j=1}^{l} P_jd_{\min}(\Lambda_j),
\]

\[
\Rightarrow \min\left\{\sum_{i=1}^{l} P_i\Lambda_i\right\} - \max\left\{\sum_{i=1}^{l} P_i\Lambda_i\right\} + P_{l+1}d_{\min}(\Lambda_{l+1}) \geq P_1d_{\min}(\Lambda_1).\]

(207)

With (207) and Lemma 2, we obtain that

\[
d_{\min}\left(\sum_{i=1}^{l} P_i\Lambda_i + P_{l+1}\Lambda_{l+1}\right)
\]

\[
= \min\left\{\min\left\{\sum_{i=1}^{l} P_i\Lambda_i\right\} - \max\left\{\sum_{i=1}^{l} P_i\Lambda_i\right\} + P_{l+1}d_{\min}(\Lambda_{l+1}), d_{\min}\left(\sum_{i=1}^{l} P_i\Lambda_i\right)\right\}
\]

\[
\overset{\text{(205)}}{=} \min\left\{\min\left\{\sum_{i=1}^{l} P_i\Lambda_i\right\} - \max\left\{\sum_{i=1}^{l} P_i\Lambda_i\right\} + P_{l+1}d_{\min}(\Lambda_{l+1}), P_1d_{\min}(\Lambda_1)\right\}
\]

\[
\overset{\text{(207)}}{=} P_1d_{\min}(\Lambda_1).
\]

(208)

This completes the proof.
Lemma 4. A superimposed constellation $2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2$, where $\Lambda_1, \Lambda_2$ are PAM($2^{m_2}, 1$) with $E[\Lambda_1] = 0, E[\Lambda_2] = 0$ and $m_1 \geq m_2, m_1, m_2 \in \mathbb{Z}^+$, has the following properties:

i) the minimum distance of $2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2$ satisfies

$$d_{\min}(2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2) = 1,$$  \hfill (209)

ii) $2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2$ can be decomposed into the following $2^{m_2+1} - 1$ subsets

$$2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2 = \{\Lambda'_t\}, t \in [1 - 2^{m_2} : 2^{m_2} - 1],$$  \hfill (210)

where $\Lambda'_t = \text{PAM}(2^{m_2} - |t|, 1)$ with mean $E[\Lambda'_t] = t(2^{m_1} + \frac{1}{2})$ and the distance between the edge points of two neighboring PAMs $\Lambda_t$ and $\Lambda_{t+1}$ is

$$d_{\text{Edge}}(\Lambda'_t, \Lambda'_{t+1}) \triangleq \min\{\Lambda'_{t+1}\} - \max\{\Lambda'_t\} = \begin{cases} 2 + 2^{m_1} - 2^{m_2} + t, & t \geq 0, \\ 1 + 2^{m_1} - 2^{m_2} - t, & t < 0. \end{cases}$$  \hfill (211)

Before we proceed to the proof, as an example, we show a sketch of the superimposed constellation $2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2$ and its subset $\Lambda'_t$ in Fig. 8.

![Fig. 8. An illustration of set $2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2$ and its subset $\Lambda'_t$.](image)

**Proof:** Recall that $\text{PAM}(2^{m_2}, 1) = \{\pm \frac{1}{2}, \ldots, \pm \frac{2^{m_2}-1}{2}\}$. For any $\lambda_{1,1}, \lambda_{1,2} \in \Lambda_1, \lambda_{2,1}, \lambda_{2,2} \in \Lambda_2$ and $(\lambda_{1,1}, \lambda_{2,2}) \neq (\lambda_{1,2}, \lambda_{2,2})$, the minimum distance is

$$d_{\min}(2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2) = \min\{|2^{m_1}(\lambda_{1,1} - \lambda_{1,2}) + (2^{m_1} + 1)(\lambda_{2,1} - \lambda_{2,2})|\}$$

$$= \min\{|2^{m_1}\Delta_1 + (2^{m_1} + 1)\Delta_2|\}$$

$$= 1.$$  \hfill (212)

where $\Delta_1 \triangleq \lambda_{1,1} - \lambda_{1,2}, \Delta_2 \triangleq \lambda_{2,1} - \lambda_{2,2}, \Delta_1, \Delta_2 \in \{\pm 1, \ldots, \pm (2^{m_2} - 1)\}$ and the last inequality follows that since $2^{m_1} \Delta_1 + (2^{m_1} + 1)\Delta_2 \in \mathbb{Z} \setminus \{0\}$ because $\Delta_1 < 2^{m_1}$ and $\Delta_2 < 2^{m_1}$ and the minimum is taken when $2^{m_1} \Delta_1 + (2^{m_1} + 1)\Delta_2 = \pm 1$, e.g., $\Delta_1 = 1, \Delta_2 = -1$ or $\Delta_1 = -1, \Delta_2 = 1$.

To prove the second property, we first write down the element of set $\lambda \in 2^{m_1} \Lambda_1 + (2^{m_1} + 1) \Lambda_2$

$$\lambda = 2^{m_1}(\lambda_1 + \lambda_2) + \lambda_2,$$  \hfill (213)
where \( \lambda_1 \in \Lambda_1, \lambda_2 \in \Lambda_2, \lambda_1 + \lambda_2 \in [-2^{m_2} + 1 : 2^{m_2} - 1] \). By letting \( t \triangleq \lambda_1 + \lambda_2 \), we can then divide \( \{ \lambda \} \) into \( 2^{m_2+1} - 1 \) subsets based on the value of \( t \). The \( t \)-th subset is

\[
\Lambda'_t \triangleq \{ 2^m t + \lambda_2 \},
\]

as we note that

\[
\min\{\Lambda'_{t+1}\} \geq 2^m(t + 1) - \frac{2^{m_2} - 1}{2} > 2^m t + \frac{2^{m_2} - 1}{2} \geq \max\{\Lambda'_t\},
\]

for a given \( t \in [-2^{m_2} + 1 : 2^{m_2}] \) as \( m_1 \geq m_2 \). Next, we need to determine the values of each elements in \( \Lambda'_t \). Since \( t \) and \( \lambda_2 \) are not independent of each other, hence \( \lambda_1 \) and \( \lambda_2 \) take values from a subset of \( \{ \pm \frac{1}{2}, \ldots, \pm \frac{2^{m_2} - 1}{2} \} \) depending on the value of the \( t \). When \( t > 0 \), we have

\[
\lambda_2 = t - \lambda_1 \geq t - \max\{\lambda_1\} = -\frac{2^{m_2} - 1}{2} + t.
\]

Therefore, \( \lambda_1 \) and \( \lambda_2 \) only take values from set \( \{-\frac{2^{m_2} - 1}{2} + t, -\frac{2^{m_2} - 3}{2} + t, \ldots, \frac{2^{m_2} - 1}{2}\} \). Similarly, When \( t < 0 \), we have

\[
\lambda_2 = t - \lambda_1 \leq t - \min\{\lambda_1\} = \frac{2^{m_2} - 1}{2} + t.
\]

Therefore, \( \lambda_1 \) and \( \lambda_2 \) only take values from set \( \{-\frac{2^{m_2} - 1}{2}, -\frac{2^{m_2} - 3}{2}, \ldots, \frac{2^{m_2} - 1}{2} + t\} \). Finally, when \( t = 0 \), \( \lambda_1 \) and \( \lambda_2 \) take values from set \( \{-\frac{2^{m_2} - 1}{2}, -\frac{2^{m_2} - 3}{2}, \ldots, \frac{2^{m_2} - 1}{2}\} \). From here, one can see that (214) can be written as

\[
\Lambda'_t = \begin{cases} 
2^m t + \{ \lambda_2 | \lambda_2 \in \{-\frac{2^{m_2} - 1}{2} + t, -\frac{2^{m_2} - 3}{2} + t, \ldots, \frac{2^{m_2} - 1}{2}\}, t \geq 0 \\
2^m t + \{ \lambda_2 | \lambda_2 \in \{-\frac{2^{m_2} - 1}{2}, -\frac{2^{m_2} - 3}{2}, \ldots, \frac{2^{m_2} - 1}{2} + t\}, t < 0 
\end{cases}.
\]

From (218), one can see that \( \Lambda'_t \) is \( \text{PAM}(2^{m_2} - |t|, 1) \) with mean \( \mathbb{E}[\Lambda'_t] = t(2^{m_1} + \frac{1}{2}) \).

Furthermore, the distance between the edge points of two neighboring PAMs \( \Lambda_t \) and \( \Lambda_{t+1} \) is

\[
d_{\text{Edge}}(\Lambda'_t, \Lambda'_{t+1}) = \min\{\Lambda'_{t+1}\} - \max\{\Lambda'_t\} = \begin{cases} 2 + 2^{m_1} - 2^{m_2} + t, & t \geq 0 \\
1 + 2^{m_1} - 2^{m_2} - t, & t < 0 \end{cases}.
\]

This completes the proof.

\[ \square \]

A. Proof of Proposition 7

Recall that \( P_t \triangleq 2^{\sum_{i=1}^l (\alpha_i + m_{i-1}) + \beta_t} \) and \( |\Lambda_t| = 2^{m_{l-1}} \) for \( l \in [1 : L] \). We then note that for \( l \in [1 : L - 1] \)

\[
P_{l+1} |\Lambda_l| = 2^{\sum_{i=1}^l (\alpha_i + m_{i-1}) + \beta_l} \cdot 2^{m_l - 1} \\
= 2^{-1 + \sum_{i=1}^l (\alpha_i + m_l) + \beta_l} \leq 2^{\sum_{i=1}^{l+1} (\alpha_i + m_{i-1}) + \beta_{l+1}} \triangleq P_{l+1}.
\]
Since $\Lambda_l = \text{PAM}(2^{m_i-1}, 1)$, we can directly use [23, Prop. 2] to obtain that
\[ d_{\text{min}}(P_l \Lambda_l + P_{l+1} \Lambda_{l+1}) = P_l \Lambda_l. \]

With the above condition and by applying Lemma 3, we arrive at
\[ d_{\text{min}} \left( \sum_{i=1}^{L} P_i \Lambda_i \right) = d_{\text{min}}(\Lambda_{\Sigma}) = P_1 = 2^{\alpha_1 + \beta_1} \in [1, 2). \]

Now, since none of the constellation points are overlapped, hence
\[ |\Lambda_{\Sigma}| = \prod_{l=1}^{L} |\Lambda_l|. \]

By considering the extreme case of $\beta_l = 1$, we can obtain an upper bound on $\max\{\Lambda_{\Sigma}\}$ and a lower bound on $\min\{\Lambda_{\Sigma}\}$ as
\begin{align*}
\max\{\Lambda_{\Sigma}\} &\leq \sum_{l=1}^{L} 2^{\sum_{i=1}^{l} (\alpha_i + m_{i-1})+1} \max\{\Lambda_l\} \\
&< \sum_{l=1}^{L} 2^{\sum_{i=1}^{l} (\alpha_i + m_{i-1})+1} \cdot \frac{2^{m_l-1} - 1}{2} \\
&< 2^{\sum_{l=1}^{L} (\alpha_l + m_l)-1} - 1 \\
\min\{\Lambda_{\Sigma}\} &= -\max\{\Lambda_{\Sigma}\} \\
&> 1 - 2^{\sum_{l=1}^{L} (\alpha_l + m_l)-1}.
\end{align*}

This completes the proof.

The following lemma is a modification of [23, Prop. 2] to encompass two-dimensional discrete constellations with irregular shapes (i.e., not necessarily carved from lattices).

**Lemma 5.** Let $X$ be a discrete random variable uniformly distributed over a two-dimensional constellation $A$ with minimum distance $d_{\text{min}}(A) > 0$. Let $Z \sim \mathcal{CN}(0, 1)$ and independent of $X$. Then
\[ I(X; X + Z) \geq H(X) - \log_2 2\pi e \left( \frac{4}{\pi d_{\text{min}}^2(A)} + \frac{1}{4} \right). \]  

(220)

**Proof:** Let $X' = X + U$, where $U$ is independent of $X$ and is uniformly distributed over a sphere $B(0, \frac{d_{\text{min}}(A)}{2})$ with radius of $\frac{d_{\text{min}}(A)}{2}$ and centred at 0. Clearly, $X', X, Y$ form a Markov chain in the following order
\[ X' \rightarrow X \rightarrow Y. \]  

(221)
Therefore, from the data processing inequality [21], we have

\[ I(X; Y) \geq I(X'; Y) \]

\[ = h(X') - h(X'|Y) \]

\[ = H(X) + h(U) - h(X'|Y) \]

\[ = H(X) + \log_2(\text{Vol}(B(0, \frac{d_{\text{min}}(A)}{2}))) - h(X'|Y) \]

\[ = H(X) + \log_2(\pi \frac{d_{\text{min}}^2(A)}{4}) - h(X'|Y). \] (222)

Note that

\[ h(X'|Y = y) = - \int p(x'|y) \log_2 p(x'|y) dx' \leq - \int p(x'|y) \log_2 q_y(x') dx', \] (223)

for any valid distribution \( q_y(x') \). We pick

\[ q_y(x') = \left( \frac{1}{\sqrt{2\pi s}} \right) e^{-\frac{(x'_l - ky_l)^2}{2s^2}}, \] (224)

where \( x'_l \) and \( y_l \) are the \( l \)th elements of \( x' \) and \( y \), respectively. Plugging this choice into (223) gives

\[ h(X'|Y = y) \leq \left( \ln 2\pi s^2 + \frac{1}{s^2} \mathbb{E}[\|X' - ky\|^2|Y = y] \right) \log_2 e. \] (225)

Thus,

\[ h(X'|Y) \leq \left( \ln 2\pi s^2 + \frac{1}{s^2} \mathbb{E}[\|X' - kY\|^2] \right) \log_2 e. \] (226)

Now, choosing \( k = \frac{\mathbb{E}[\|X\|^2]}{1 + \mathbb{E}[\|X\|^2]} \), we have

\[ \mathbb{E}[\|X' - kY\|^2] = \mathbb{E}[\|X + U - k(X + Z)\|^2] \]

\[ = (1 - k)^2 \mathbb{E}[\|X\|^2] + \sigma^2(B(0, \frac{d_{\text{min}}(A)}{2})) + k^2 \]

\[ = \frac{\mathbb{E}[\|X\|^2]}{1 + \mathbb{E}[\|X\|^2]} + \frac{d_{\text{min}}^2(A)}{16} \] (227)

Hence, (226) becomes

\[ h(X'|Y) \leq \left( \ln 2\pi s^2 + \frac{1}{s^2} \left( \frac{\mathbb{E}[\|X\|^2]}{1 + \mathbb{E}[\|X\|^2]} + \frac{d_{\text{min}}^2(A)}{16} \right) \right) \log_2 e. \] (228)
We choose $s^2 = \frac{\mathbb{E}[\|X\|^2]}{1 + \mathbb{E}[\|X\|^2]} + \frac{d^2_{\min}(A)}{16}$ to obtain

$$h(X'\mid Y) \leq \log_2 2\pi e \left( \frac{\mathbb{E}[\|X\|^2]}{1 + \mathbb{E}[\|X\|^2]} + \frac{d^2_{\min}(A)}{16} \right)$$

$$\leq \log_2 2\pi e \left( 1 + \frac{d^2_{\min}(A)}{16} \right). \quad (229)$$

Plugging (229) into (222) results in

$$I(X; Y) \geq H(X) + \log_2 \left( \frac{d^2_{\min}(A)}{4} \right) - \log_2 2\pi e \left( 1 + \frac{d^2_{\min}(A)}{16} \right)$$

$$= H(X) - \log_2 2\pi e \left( \frac{4}{\pi d^2_{\min}(A)} + \frac{1}{4} \right). \quad (230)$$

This completes the proof.

**REFERENCES**

[1] M. Qiu, Y.-C. Huang, and J. Yuan, “Multiuser MISO broadcast channels with imperfect CSI: Discrete signaling without SIC,” in *Proc. IEEE Globecom*, 2019, pp. 1–6.

[2] M. Qiu, Y.-C. Huang, and J. Yuan, “On discrete signaling and treating interference as noise for complex Gaussian interference channels,” in *Proc. IEEE ISIT, to appear*, 2020.

[3] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. USA: Cambridge University Press, 2012.

[4] H. Sato, “Two-user communication channels,” *IEEE Trans. Inf. Theory*, vol. 23, no. 3, p. 295304, May 1977.

[5] A. Carleial, “Interference channels,” *IEEE Trans. Inf. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.

[6] 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.

[7] A. Carleial, “A case where interference does not reduce capacity (corresp.),” *IEEE Trans. Inf. Theory*, vol. 21, no. 5, p. 569570, Sep. 1975.

[8] H. Sato, “The capacity of the Gaussian interference channel under strong interference,” *IEEE Trans. Inf. Theory*, vol. 27, no. 6, pp. 786–788, Nov. 1981.

[9] M. Costa and A. E. Gamal, “The capacity region of the discrete memoryless interference channel with strong interference,” *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 710–711, Sep. 1987.

[10] V. S. Annapureddy and V. V. Veeravalli, “Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region,” *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.

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

[12] H. Chong, M. Motani, H. K. Garg, and H. El Gamal, “On the HanKobayashi region for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3188–3195, Jul. 2008.

[13] X. Shang, G. Kramer, and B. Chen, “A new outer bound and the noisy-interference sumrate capacity for Gaussian interference channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.

[14] A. S. Motahari and A. K. Khandani, “Capacity bounds for the Gaussian interference channel,” *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.

[15] C. Geng, N. Naderializadeh, A. S. Avestimehr, and S. A. Jafar, “On the optimality of treating interference as noise,” *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1753–1767, Apr. 2015.
[16] C. Geng, H. Sun, and S. A. Jafar, “On the optimality of treating interference as noise: General message sets,” *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3722–3736, Jul. 2015.

[17] H. Sun and S. A. Jafar, “On the optimality of treating interference as noise for K-user parallel Gaussian interference networks,” *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1911–1930, Apr. 2016.

[18] H. Joudeh and B. Clerckx, “On the optimality of treating inter-cell interference as noise in uplink cellular networks,” *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7208–7232, Nov. 2019.

[19] R. Dobrushin, “A general formulation of the fundamental theorem of Shannon in the theory of information,” *Amer. Math. Soc. Trans.*, vol. 33, no. 6, pp. 323–438, 1963.

[20] R. Ahlswede, “Multi-way communication channels,” in *Proc. IEEE ISIT*, Mar. 1973, pp. 23–52.

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

[22] K. Moshksar, A. Ghasemi, and A. K. Khandani, “An alternative to decoding interference or treating interference as gaussian noise,” *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 305–322, Jan. 2015.

[23] A. Dytso, D. Tuninetti, and N. Devroye, “Interference as noise: Friend or foe?” *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3561–3596, Jun. 2016.

[24] S. Li, Y.-C. Huang, T. Liu, and H. D. Pfister, “On the limits of treating interference as noise for two-user symmetric Gaussian interference channels,” in *Proc. IEEE ISIT*, Jun. 2015, pp. 1711–1715.

[25] S. Li, *Matrix Analysis of Communication and Brain Networks*. Texas A & M University, USA: Doctoral dissertation, Dec. 2016.

[26] G. Bresler and D. Tse, “The two-user Gaussian interference channel: a deterministic view,” *Europ. Trans. Telecommun.*, vol. 19, no. 4, pp. 333–354, 2008.

[27] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, “Wireless network information flow: A deterministic approach,” *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.

[28] L. H. Ozarow and A. D. Wyner, “On the capacity of the Gaussian channel with a finite number of input levels,” *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, 1990.

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

[30] A. S. Motahari, S. Oveis-Gharan, M. Maddah-Ali, and A. K. Khandani, “Real interference alignment: Exploiting the potential of single antenna systems,” *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4799–4810, Aug. 2014.

[31] M. U. Baig, A. Hst-Madsen, and A. Nosratinia, “Discrete modulation for interference mitigation,” *IEEE Trans. Inf. Theory*, vol. 66, no. 5, pp. 3026–3039, 2019.