

# On the Optimality of Treating Interference as Noise: General Message Sets

Chunhua Geng, *Student Member, IEEE*, Hua Sun, *Student Member, IEEE*, and Syed Ali Jafar, *Fellow, IEEE*

**Abstract**—In a  $K$ -user Gaussian interference channel, it has been shown that if for each user the desired signal strength is no less than the sum of the strengths of the strongest interference from this user and the strongest interference to this user (all values in decibel scale), then treating interference as noise (TIN) is optimal from the perspective of generalized degrees of freedom (GDoF) and achieves the entire channel capacity region to within a constant gap. In this paper, we show that for such TIN-optimal interference channels, even if the message set is expanded to include an independent message from each transmitter to each receiver, operating the new channel as the original interference channel and treating interference as noise is still optimal for the sum capacity up to a constant gap. Furthermore, we extend the result to the sum-GDoF optimality of TIN in the general setting of  $X$  channels with arbitrary numbers of transmitters and receivers.

**Index Terms**—Gaussian networks, generalized degrees of freedom (GDoF), sum capacity, treating interference as noise (TIN),  $X$  channels.

## I. INTRODUCTION

TREATING interference as noise (TIN) when it is sufficiently weak is an attractive interference management principle for wireless networks in practice due to its simplicity and robustness. Remarkably, TIN is also information-theoretically optimal when the interference is sufficiently weak. This is established in [1]–[6] from an exact capacity perspective, and in [7]–[12] from an approximate capacity perspective. Each approach has its merits – the former identifies relatively narrow regimes where TIN achieves exact capacity, whereas the latter identifies significantly broader regimes where TIN is approximately optimal. Most relevant to this work are the results by Geng et al. in [12] where it is shown that in a general  $K$ -user interference channel, if for each user the desired signal strength is no less than the sum of the strengths of the strongest interference from this user and the strongest interference to this user (all values in dB scale), then TIN achieves the entire generalized degrees

of freedom (GDoF) region and is optimal for the whole channel capacity region up to a constant gap of no more than  $\log_2(3K)$  bits per user.

In this paper we explore the sum-rate optimality of TIN when the message set is expanded to include an independent message from each transmitter to each receiver, i.e., the  $X$  channel setting [13]–[15]. Related prior works on the  $X$  setting in [16] and [17] have primarily focused on the case with 2 transmitters and 2 receivers. In [16], Huang, Cadambe and Jafar characterize the sum-GDoF for the symmetric  $X$  channel and identify sufficient conditions for TIN to achieve exact capacity in the asymmetric case. In [17], Niesen and Maddah-Ali characterize the capacity for the general asymmetric case within a constant gap subject to an outage set.

The main contribution of this work is to show that, for the  $K$ -user TIN-optimal interference channels identified by Geng et al. in [12], even if the message set is expanded to also include an independent message from each transmitter to each receiver, operating as the original interference channel and treating interference as noise at each receiver is still optimal for the *sum* capacity up to a constant gap (see Theorem 2 in Section III). We also extend the optimality of TIN to the general  $X$  channel with arbitrary numbers of transmitters and receivers (see Theorem 3 in Section III). In each case, the achievability argument follows directly from [12] because it is based only on operating the target network as an interference channel and treating interference as noise. The main difficulty lies in deriving tight information theoretical outer bounds. Recall that in [12], for interference channels, the converse is based on reducing the channel to a cyclic network. Each such reduction produces an outer bound and collectively these outer bounds suffice for the GDoF characterization in the interference channel setting. However, this is no longer true once the message set is expanded. While one can similarly obtain outer bounds on the sum rates of subsets of messages by considering all cyclic subnetworks, it is not hard to verify that these bounds do not suffice for our purpose. As illustrated in this work, to obtain desired outer bounds, the genie signal provided to each user should be chosen more judiciously. Notably, to complete the generalization to  $X$  channels, especially for the case where the number of receivers is larger than that of transmitters, due to the added difficulty, we also resort to deterministic channel models [16]–[22], in which certain combinatorial structure can be exploited to simplify the proof. More specifically, we use the fact that the sum capacity of Gaussian  $X$  channels is upper bounded by that of their carefully chosen deterministic counterparts up to a constant gap, which shares the same

Manuscript received May 4, 2014; revised December 6, 2014; accepted April 20, 2015. Date of publication May 13, 2015; date of current version June 12, 2015. This work was supported in part by ONR and in part by NSF under Grant CCF-1161418, Grant CCF-1317351, and Grant CCF-1319104. This paper was presented in part at the 2014 IEEE International Symposium on Information Theory.

The authors are with the Center of Pervasive Communications and Computing, Department of Electrical Engineering and Computer Science, University of California at Irvine, Irvine, CA 92697 USA (e-mail: chunhug@uci.edu; huas2@uci.edu; syed@uci.edu).

Communicated by S. Avestimehr, Associate Editor for Communications.

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

Digital Object Identifier 10.1109/TIT.2015.2432751

principle with [20]–[22]. Then by upper bounding the sum capacity of the deterministic channel appropriately, we obtain the desired converse for the original Gaussian channel.

## II. PRELIMINARIES

### A. Channel Model

Consider the wireless channel with  $M$  transmitters and  $N$  receivers, which can be described by the following input-output equations,<sup>1</sup>

$$Y_k(t) = \sum_{i=1}^M \tilde{h}_{ki} \tilde{X}_i(t) + Z_k(t), \quad \forall k \in [N], \quad (1)$$

where  $\tilde{h}_{ki}$  is the complex channel gain value from transmitter  $i$  to receiver  $k$ .  $\tilde{X}_i(t)$ ,  $Y_k(t)$  and  $Z_k(t)$  are the transmitted symbol of transmitter  $i$ , the received signal of receiver  $k$ , and the additive circularly symmetric complex Gaussian noise with zero mean and unit variance seen by receiver  $k$ , respectively, at each time index  $t$ . All the symbols are complex. Each transmitter  $i$  is subject to the power constraint  $E[|\tilde{X}_i(t)|^2] \leq P_i$ .

Following similar approaches in [7] and [12], we translate the standard channel model (1) into an equivalent normalized form that is more conducive for GDoF studies. We define<sup>2</sup>

$$\alpha_{ki} \triangleq \frac{\log(\max\{1, |\tilde{h}_{ki}|^2 P_i\})}{\log P}, \quad \forall i \in [M], \quad \forall k \in [N], \quad (2)$$

where  $P > 1$  is a nominal power value.

Now according to (2), we represent the original channel model (1) in the following form,

$$\begin{aligned} Y_k(t) &= \sum_{i=1}^M h_{ki} X_i(t) + Z_k(t) \\ &= \sum_{i=1}^M \sqrt{P^{\alpha_{ki}}} e^{j\theta_{ki}} X_i(t) + Z_k(t), \quad \forall k \in [N], \end{aligned} \quad (3)$$

where  $X_i(t) = \tilde{X}_i(t)/\sqrt{P_i}$  is the normalized transmit symbol of transmitter  $i$ , subject to the unit power constraint, i.e.,  $E[|X_i(t)|^2] \leq 1$ .  $\sqrt{P^{\alpha_{ki}}}$  and  $\theta_{ki}$  are the magnitude and the phase, respectively, of the channel between transmitter  $i$  and receiver  $k$ . The exponent  $\alpha_{ki}$  is called the channel strength level of the link between transmitter  $i$  and receiver  $k$ . As in [7] and [12], for the GDoF metric, we preserve the ratios  $\alpha_{ki}$  as all SNRs approach infinity. In the rest of the paper, we only consider the equivalent channel model in (3).

In the  $K$ -user interference channel where  $M = N = K$ , each transmitter intends to send one independent message to its corresponding receiver. Because we wish to prove the negative result that additional messages do not add to the sum-GDoF in a TIN-optimal network, the strongest result corresponds to the case where we include messages from every transmitter to every receiver. Therefore, we will consider the  $X$  channel setting. In the  $M \times N$   $X$  channel, transmitter  $i$  has message

$W_{ki}$  intended for receiver  $k$ , and all the messages  $W_{ki}$  are independent,  $\forall i \in [M], \forall k \in [N]$ . The size of the message set  $\{W_{ki}\}$  is denoted by  $|W_{ki}|$ . For codewords spanning  $n$  channel uses, the rates  $R_{ki} = \frac{\log |W_{ki}|}{n}$  are achievable if the probability of error of all messages can be made arbitrarily small simultaneously by choosing an appropriately large  $n$ . The channel capacity region  $\mathcal{C}$  is the closure of the set of all achievable rate tuples. Collecting the channel strength levels and phases in the sets

$$\alpha \triangleq \{\alpha_{ki}\}, \quad \theta \triangleq \{\theta_{ki}\}, \quad \forall i \in [M], \quad \forall k \in [N], \quad (4)$$

the capacity region is denoted as  $\mathcal{C}(P, \alpha, \theta)$ , which is a function of  $\alpha$ ,  $\theta$ , and  $P$ . The sum channel capacity is defined as

$$C_{\Sigma, X} \triangleq \max_{\mathcal{C}(P, \alpha, \theta)} \sum_{i=1}^M \sum_{k=1}^N R_{ki} \quad (5)$$

Then the GDoF region of the  $X$  channel as represented in (3) is given by

$$\begin{aligned} \mathcal{D}(\alpha, \theta) \triangleq & \left\{ (d_{11}, d_{12}, \dots, d_{NM}) : d_{ki} = \lim_{P \rightarrow \infty} \frac{R_{ki}}{\log P}, \right. \\ & \forall i \in [M], \forall k \in [N], \\ & \left. (R_{11}, R_{12}, \dots, R_{NM}) \in \mathcal{C}(P, \alpha, \theta) \right\}, \end{aligned} \quad (6)$$

and its sum-GDoF value is

$$d_{\Sigma, X} \triangleq \max_{\mathcal{D}(\alpha, \theta)} \sum_{i=1}^M \sum_{k=1}^N d_{ki} \quad (7)$$

### B. On the Optimality of TIN for Interference Channels

Let us first review the optimality of TIN for the  $K$ -user interference channel from the perspective of GDoF.

*Theorem 1 [12, Th. 1]:* In a  $K$ -user interference channel, where the channel strength level from transmitter  $i$  to receiver  $j$  is equal to  $\alpha_{ji}$ ,  $\forall i, j \in [K]$ , if the following condition is satisfied

$$\alpha_{ii} \geq \max_{j: j \neq i} \{\alpha_{ji}\} + \max_{k: k \neq i} \{\alpha_{ik}\}, \quad \forall i, j, k \in [K], \quad (8)$$

then power control and treating interference as noise achieves the whole GDoF region. Moreover, the GDoF region is the set of all  $K$ -tuples  $(d_1, d_2, \dots, d_K)$  satisfying

$$\text{individual bounds : } 0 \leq d_i \leq \alpha_{ii}, \quad \forall i \in [K] \quad (9)$$

$$\begin{aligned} \text{cycle bounds : } & \sum_{j=1}^m d_{i_j} \leq \sum_{j=1}^m (\alpha_{i_j i_{j+1}} - \alpha_{i_{j-1} i_j}), \\ & \forall (i_1, i_2, \dots, i_m) \in \Pi_K, \\ & \forall m \in \{2, 3, \dots, K\}, \end{aligned} \quad (10)$$

where  $\Pi_K$  is the set of all possible cyclic sequences of all subsets of  $[K]$  with cardinality no less than 2, and the modulo- $m$  arithmetic is implicitly used on user indices, e.g.,  $i_m = i_0$ .

To help interpret the results in Theorem 1, consider a 3-user interference channel as an example, where

$$\Pi_3 = \{(1, 2), (1, 3), (2, 3), (1, 2, 3), (1, 3, 2)\}$$

<sup>1</sup>Throughout this paper, for any positive integer  $Z$ ,  $[Z]$  denotes the set  $\{1, 2, \dots, Z\}$ .

<sup>2</sup>As noted in [12], avoiding negative  $\alpha_{ki}$  does not impact the GDoF or the constant gap result.

Note that each cyclic sequence in  $\Pi_K$  is essentially a cyclically ordered subset of user indices, without repetitions. Then according to Theorem 1, under the condition (8), the GDoF region  $(d_1, d_2, d_3)$  is fully characterized by the following inequalities together

$$\begin{aligned} 0 &\leq d_1 \leq \alpha_{11} \\ 0 &\leq d_2 \leq \alpha_{22} \\ 0 &\leq d_3 \leq \alpha_{33} \\ d_1 + d_2 &\leq (\alpha_{11} + \alpha_{22}) - (\alpha_{12} + \alpha_{21}) \\ d_1 + d_3 &\leq (\alpha_{11} + \alpha_{33}) - (\alpha_{13} + \alpha_{31}) \\ d_2 + d_3 &\leq (\alpha_{22} + \alpha_{33}) - (\alpha_{23} + \alpha_{32}) \\ d_1 + d_2 + d_3 &\leq (\alpha_{11} + \alpha_{22} + \alpha_{33}) - (\alpha_{12} + \alpha_{23} + \alpha_{31}) \\ d_1 + d_2 + d_3 &\leq (\alpha_{11} + \alpha_{22} + \alpha_{33}) - (\alpha_{13} + \alpha_{32} + \alpha_{21}) \end{aligned}$$

which is achievable via power control and TIN.

*Remark:* Theorem 1 claims that in the  $K$ -user interference channel, if for each user the desired signal strength is no less than the sum of the strengths of the strongest interference from this user and the strongest interference to this user (all values in dB scale), then TIN is GDoF-optimal. Furthermore, it is shown in [12] that under the same condition, TIN achieves the entire channel capacity region to within a gap of no larger than  $\log_2(3K)$  bits per user. Note that the gap is bounded by a constant for a fixed number of users, i.e., it does not depend on the channel strength parameters  $\alpha_{ij}$  and  $P$ .

### III. RESULTS

The main result of this paper is the following theorem.

*Theorem 2:* In a  $K$ -user interference channel, where the channel strength level from transmitter  $i$  to receiver  $j$  is equal to  $\alpha_{ji}$ ,  $\forall i, j \in [K]$ , when the following condition is satisfied,

$$\alpha_{ii} \geq \max_{j:j \neq i} \{\alpha_{ji}\} + \max_{k:k \neq i} \{\alpha_{ik}\}, \quad \forall i, j, k \in [K] \quad (11)$$

then even if the message set is increased to the  $X$  channel setting, operating the new channel as the original interference channel and treating interference as noise at each receiver still achieves the sum-GDoF. Furthermore, the same scheme is also optimal for the sum channel capacity up to a constant gap of no more than  $K \log_2[K(K+1)]$  bits.

The proof of Theorem 2 is presented in Section IV-A.

While the  $K$ -user interference channel is naturally associated with a  $K \times K$   $X$  channel setting, the  $X$  channel setting also allows for unequal numbers of transmitters and receivers. A natural question is whether such a generalization of the TIN-optimality result is possible for the  $X$  channel with  $M \neq N$ . The following theorem provides such a generalization.

*Theorem 3:* In an  $M \times N$   $X$  channel, where the channel strength level from transmitter  $i$  to receiver  $j$  is equal to  $\alpha_{ji}$  and  $\kappa \triangleq \min\{M, N\}$ , if there exist two permutations  $\Pi^T$  and  $\Pi^R$  for the transmitter and receiver indices, respectively, such that

$$\alpha_{\Pi_i^R \Pi_i^T} \geq \max_{j:j \neq i} \{\alpha_{\Pi_j^R \Pi_i^T}\} + \max_{k:k \neq i} \{\alpha_{\Pi_i^R \Pi_k^T}\}, \quad \forall i \in [\kappa], \quad \forall j \in [N], \quad \forall k \in [M], \quad (12)$$

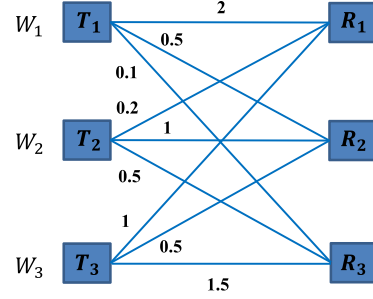


Fig. 1. A 3-user interference channel, where the value on each link denotes its channel strength level.

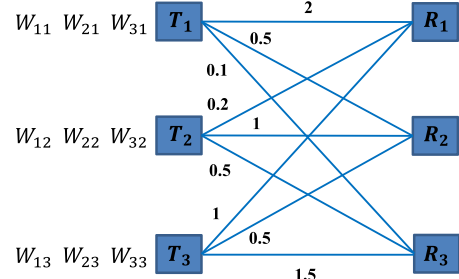


Fig. 2. A  $3 \times 3$   $X$  channel, which has the same channel strength levels as the 3-user interference channel in Fig. 1.

where  $\Pi_i^T$  ( $\Pi_i^R$ ) denotes the  $i$ -th element in the permutation of transmitters (receivers)  $\Pi^T$  ( $\Pi^R$ ), then operating the channel as a  $\kappa$ -user interference channel, in which each transmitter  $\Pi_i^T$  intends to deliver an independent message to receiver  $\Pi_i^R$ ,  $i \in [\kappa]$ , and treating interference as noise at each receiver is sum-GDoF optimal.<sup>3</sup>

The proof of Theorem 3 is given in Section IV-B.

*Example 1:* First, consider the 3-user interference channel illustrated in Fig. 1, where the value on each link denotes its channel strength level. In this interference channel, transmitter  $i$  intends to send an independent message  $W_i$  to its desired receiver  $i$ ,  $\forall i \in \{1, 2, 3\}$ . Hence there are 3 messages in this setting. It's easy to check that the TIN-optimality condition (11) is satisfied for each user. Then according to Theorem 1, it is not hard to verify that the sum-GDoF value of this interference channel is

$$d_{\Sigma, IC} = d_1 + d_2 + d_3 = 2.5$$

which is achieved by power control and TIN.

Next, let us expand the set of messages to the  $X$  channel setting, where each transmitter intends to send an independent message to each receiver as shown in Fig. 2. Therefore, there are totally 9 messages in this  $X$  channel. Theorem 2 claims that for this  $3 \times 3$   $X$  channel, the sum-GDoF value is still

$$d_{\Sigma, X} = \sum_{i=1}^3 \sum_{k=1}^3 d_{ki} = 2.5$$

which can be achieved by setting  $W_{ki} = \phi$  for  $i \neq k$  and  $\forall i, k \in \{1, 2, 3\}$ , sending only  $\{W_{11}, W_{22}, W_{33}\}$  through the channel and treating interference as noise at each receiver.

<sup>3</sup>Based on the proof in Section IV-B, it is not hard to verify that the same TIN scheme is also optimal to achieve the sum channel capacity to within a constant gap.

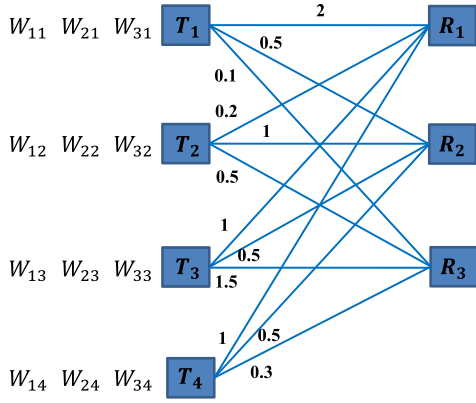


Fig. 3. A  $4 \times 3$   $X$  channel, which is obtained by adding another transmitter to the  $X$  channel in Fig. 2.

Finally, after adding another transmitter as depicted in Fig. 3, the number of the messages increases to 12 in this  $4 \times 3$   $X$  channel. It's easy to verify that (12) holds. Then according to Theorem 3, for this  $X$  channel and its reciprocal channel, the same TIN scheme is still optimal in terms of the sum-GDoF, whose value remains 2.5.

#### IV. PROOFS

##### A. Proof for Theorem 2

In the following, we first consider the sum-GDoF of the  $K \times K$   $X$  channel. Then we use the insight gained in the GDoF study to derive the constant gap result for the sum channel capacity.

1) *Proof for the Sum-GDoF*: The proof consists of two steps. First, we show that when the condition (11) is satisfied, for all individual and cycle bounds of a TIN-optimal  $K$ -user interference channel (see Theorem 1), if each  $d_i$  ( $\forall i \in [K]$ ) is replaced by  $\hat{d}_i = \sum_{j=1}^K d_{ij}$ , these bounds still hold for its counterpart  $X$  channel. Next, based on the first step, we prove that under the condition (11), the  $K$ -user interference channel and its counterpart  $X$  channel have the same sum-GDoF. Therefore, according to Theorem 1, we establish that power control and TIN achieves the sum-GDoF of the  $K \times K$   $X$  channel when the condition (11) holds.

Let's start with the first step. In the following, we first give an example of the  $3 \times 3$   $X$  channel to show how to obtain the outer bounds corresponding to the individual and cycle bounds in Theorem 1, where each  $d_i$  is replaced by  $\hat{d}_i = \sum_{j=1}^3 d_{ij}$ ,  $\forall i \in \{1, 2, 3\}$ . Then we generalize the proof to the  $K \times K$   $X$  channel.

*Example 2*: Consider a 3-user TIN-optimal interference channel. According to Theorem 1, we can obtain the entire GDoF region, which is characterized by certain individual and cycle bounds. To extend the result to the  $X$  channel setting, each of these bounds will be extended. To illustrate the key ideas in this example, we consider the following two bounds,

$$d_3 \leq \alpha_{33} \quad (13)$$

$$d_1 + d_2 \leq (\alpha_{11} + \alpha_{22}) - (\alpha_{12} + \alpha_{21}) \quad (14)$$

and intend to prove that in the counterpart  $3 \times 3$   $X$  channel, if we replace each  $d_i$  by  $\hat{d}_i = \sum_{j=1}^3 d_{ij}$ ,  $\forall i \in \{1, 2, 3\}$ , the above

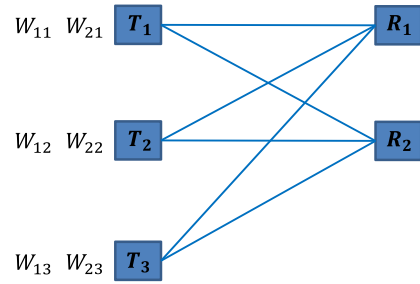


Fig. 4. The subnetwork with 3 transmitters, 2 receivers and 6 messages.

two bounds still hold, i.e.,

$$\begin{aligned} \hat{d}_3 &= d_{31} + d_{32} + d_{33} \leq \alpha_{33} \quad (15) \\ \hat{d}_1 + \hat{d}_2 &= d_{11} + d_{12} + d_{13} + d_{21} + d_{22} + d_{23} \\ &\leq (\alpha_{11} + \alpha_{22}) - (\alpha_{12} + \alpha_{21}) \quad (16) \end{aligned}$$

All the remaining bounds can be extended to the  $X$  channel similarly.

To prove (15), we just need to consider the multiple-access channel (MAC) consisting of all the transmitters and the receiver 3, then we have

$$R_{31} + R_{32} + R_{33} \leq \log_2(1 + P^{\alpha_{31}} + P^{\alpha_{32}} + P^{\alpha_{33}}) \quad (17)$$

Because (11) is satisfied, i.e.,  $\alpha_{33} \geq \alpha_{32}$  and  $\alpha_{33} \geq \alpha_{31}$ , in the GDoF sense we have

$$\hat{d}_3 = d_{31} + d_{32} + d_{33} \leq \alpha_{33} \quad (18)$$

To prove (16), consider the subnetwork consisting of all the transmitters and the receivers 1 and 2 as illustrated in Fig. 4, where we have eliminated the third receiver and its desired messages  $W_{31}, W_{32}, W_{33}$ . This cannot hurt the rates of the remaining messages, so the outer bound arguments remain valid. Define

$$S_1(t) = h_{21}X_1(t) + Z_2(t) \quad (19)$$

$$S_2(t) = h_{12}X_2(t) + Z_1(t) \quad (20)$$

For the receiver 1, we provide  $S_1^n$ ,  $W_{21}$  and  $W_{23}$  through a genie. From Fano's inequality, we have

$$\begin{aligned} n(R_{11} + R_{12} + R_{13} - \epsilon) &\leq I(W_{11}, W_{12}, W_{13}; Y_1^n, S_1^n, W_{21}, W_{23}) \quad (21) \\ &= I(W_{11}, W_{12}, W_{13}; Y_1^n, S_1^n | W_{21}, W_{23}) \quad (22) \\ &= I(W_{11}, W_{12}, W_{13}; S_1^n | W_{21}, W_{23}) \\ &\quad + I(W_{11}, W_{12}, W_{13}; Y_1^n | S_1^n, W_{21}, W_{23}) \quad (23) \end{aligned}$$

$$\begin{aligned} &= h(S_1^n | W_{21}, W_{23}) - h(S_1^n | W_{21}, W_{23}, W_{11}, W_{12}, W_{13}) \\ &\quad + h(Y_1^n | S_1^n, W_{21}, W_{23}) \\ &\quad - h(Y_1^n | S_1^n, W_{21}, W_{23}, W_{11}, W_{12}, W_{13}) \quad (24) \end{aligned}$$

$$\begin{aligned} &\leq h(S_1^n | W_{21}, W_{23}) - h(S_1^n | W_{21}, W_{23}, W_{11}, W_{12}, W_{13}, X_1^n) \\ &\quad + h(Y_1^n | S_1^n, W_{21}, W_{23}) \quad (25) \end{aligned}$$

$$\begin{aligned} &\leq h(S_1^n | W_{21}) - h(Z_2^n) + h(Y_1^n | S_1^n) - h(S_2^n | W_{12}) \quad (26) \end{aligned}$$

where (22) follows because all the messages are independent, (25) holds since adding conditioning does not increase

entropy and (26) holds because dropping conditioning (in the first and third terms) does not reduce entropy.

Due to symmetry, for the receiver 2, we similarly obtain

$$\begin{aligned} n(R_{21} + R_{22} + R_{23} - \epsilon) \\ \leq h(S_2^n | W_{12}) - h(Z_1^n) + h(Y_2^n | S_2^n) - h(S_1^n | W_{21}) \end{aligned} \quad (27)$$

Thus the sum rate is bounded as follows.

$$\begin{aligned} n\left(\sum_{i=1}^2 \sum_{j=1}^3 R_{ij} - 2\epsilon\right) \\ \leq h(Y_1^n | S_1^n) + h(Y_2^n | S_2^n) - h(Z_1^n) - h(Z_2^n) \quad (28) \\ \leq \sum_{t=1}^n \left[ h(Y_1(t) | S_1(t)) + h(Y_2(t) | S_2(t)) - h(Z_1(t)) \right. \\ \left. - h(Z_2(t)) \right] \quad (29) \end{aligned}$$

where the second inequality follows from the chain rule and the fact that dropping conditioning does not reduce entropy. Finally, because the circularly symmetric complex Gaussian distribution maximizes conditional differential entropy for a given covariance constraint, we obtain

$$\begin{aligned} \sum_{i=1}^2 \sum_{j=1}^3 R_{ij} - 2\epsilon \\ \leq \log_2 \left( 1 + P^{\alpha_{13}} + P^{\alpha_{12}} + \frac{P^{\alpha_{11}}}{1 + P^{\alpha_{21}}} \right) \\ + \log_2 \left( 1 + P^{\alpha_{23}} + P^{\alpha_{21}} + \frac{P^{\alpha_{22}}}{1 + P^{\alpha_{12}}} \right) \end{aligned} \quad (30)$$

Due to the condition (11), in the GDoF sense we obtain

$$\hat{d}_1 + \hat{d}_2 \leq (\alpha_{11} + \alpha_{22}) - (\alpha_{12} + \alpha_{21}) \quad (31)$$

which is the desired extension, (16), to the  $X$  channel setting of the original bound, (14), for the interference channel.  $\square$

Now let us consider the proof for the general  $K \times K$   $X$  channel. For the individual bounds in the  $K$ -user interference channel

$$d_i \leq \alpha_{ii}, \quad \forall i \in [K], \quad (32)$$

in its counterpart  $X$  channel, the corresponding bound comes from the MAC consisting of all the transmitters and the receiver  $i$ ,

$$\sum_{j=1}^K R_{ij} \leq \log_2 \left( 1 + \sum_{j=1}^K P^{\alpha_{ij}} \right) \quad (33)$$

According to (11), in the GDoF sense we have

$$\hat{d}_i = \sum_{j=1}^K d_{ij} \leq \alpha_{ii} \quad (34)$$

For any cycle bound in the interference channel

$$\begin{aligned} \sum_{j=1}^m d_{ij} \leq \sum_{j=1}^m (\alpha_{ij} - \alpha_{i_{j-1}j}), \\ \forall (i_1, i_2, \dots, i_m) \in \Pi_K, \quad \forall m \in \{2, 3, \dots, K\}, \end{aligned} \quad (35)$$

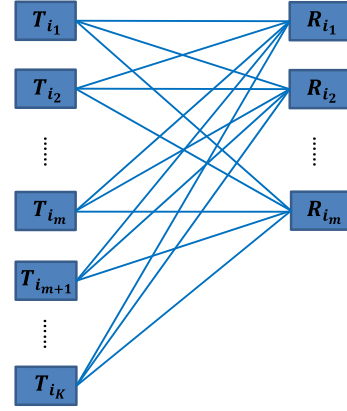


Fig. 5. A  $K \times m$   $X$  channel ( $K \geq m$ ).

consider the subnetwork consisting of all the transmitters and the receivers  $\{i_1, i_2, \dots, i_m\}$  as shown in Fig. 5. Eliminate all other receivers and their desired messages, which cannot hurt the rates of the remaining messages. For such a  $K \times m$   $X$  channel, define  $\mathcal{W} \triangleq \{W_{ijik}\}$ ,  $\mathcal{W}_{i_j}^* \triangleq \{W_{ij i_1}, W_{ij i_2}, \dots, W_{ij i_K}\}$ ,  $\mathcal{W}_{i_k}^\dagger \triangleq \{W_{i_1 i_k}, W_{i_2 i_k}, \dots, W_{i_m i_k}\}$ , and  $\mathcal{W}_{\mathcal{S}}^c \triangleq \mathcal{W} \setminus \mathcal{W}_{\mathcal{S}}$ , where  $\forall j \in [m], \forall k \in [K]$ , and  $\mathcal{S}$  is any subset of message indices. In words, the sets  $\mathcal{W}$ ,  $\mathcal{W}_{i_j}^*$ , and  $\mathcal{W}_{i_k}^\dagger$  represent all the messages delivered in this  $K \times m$   $X$  channel, all the messages intended to receiver  $i_j$ , and all the messages coming from transmitter  $i_k$ , respectively, and  $\mathcal{W}_{\mathcal{S}}^c$  is the complement of  $\mathcal{W}_{\mathcal{S}}$  in  $\mathcal{W}$ . For instance, when  $j, k \in \{1, 2\}$  and  $\mathcal{S} = \{i_1 i_1, i_1 i_2\}$ , then  $\mathcal{W}_{\mathcal{S}} = \{W_{i_1 i_1}, W_{i_1 i_2}\}$  and  $\mathcal{W}_{\mathcal{S}}^c = \{W_{i_2 i_1}, W_{i_2 i_2}\}$ . Modulo- $m$  arithmetic is used on the receiver indices, e.g.,  $i_0 = i_m$ . Lastly, to complete the setup, define

$$S_{ij}(t) = h_{ij-1i_j} X_{ij}(t) + Z_{ij-1}(t), \quad \forall j \in [m] \quad (36)$$

For the receiver  $i_1$ , we provide  $S_{i_1}^n$ ,  $\mathcal{W}_{i_2 i_2}^c \setminus \mathcal{W}_{i_1}^*$  through a genie. From Fano's inequality, we have

$$\begin{aligned} n\left(\sum_{k=1}^K R_{i_1 i_k} - \epsilon\right) \\ \leq I(\mathcal{W}_{i_1}^*; Y_{i_1}^n, S_{i_1}^n, \mathcal{W}_{i_2 i_2}^c \setminus \mathcal{W}_{i_1}^*) \end{aligned} \quad (37)$$

$$= I(\mathcal{W}_{i_1}^*; Y_{i_1}^n, S_{i_1}^n | \mathcal{W}_{i_2 i_2}^c \setminus \mathcal{W}_{i_1}^*) \quad (38)$$

$$= I(\mathcal{W}_{i_1}^*; S_{i_1}^n | \mathcal{W}_{i_2 i_2}^c \setminus \mathcal{W}_{i_1}^*) + I(\mathcal{W}_{i_1}^*; Y_{i_1}^n | S_{i_1}^n, \mathcal{W}_{i_2 i_2}^c \setminus \mathcal{W}_{i_1}^*) \quad (39)$$

$$= h(S_{i_1}^n | \mathcal{W}_{i_2 i_2}^c \setminus \mathcal{W}_{i_1}^*) - h(S_{i_1}^n | \mathcal{W}_{i_2 i_2}^c) \\ + h(Y_{i_1}^n | S_{i_1}^n, \mathcal{W}_{i_2 i_2}^c \setminus \mathcal{W}_{i_1}^*) - h(Y_{i_1}^n | S_{i_1}^n, \mathcal{W}_{i_2 i_2}^c) \quad (40)$$

$$\leq h(S_{i_1}^n | \mathcal{W}_{i_1}^\dagger \setminus W_{i_1 i_1}) - h(Z_{i_0}^n) + h(Y_{i_1}^n | S_{i_1}^n) \\ - h(S_{i_2}^n | \mathcal{W}_{i_2}^\dagger \setminus W_{i_2 i_2}) \quad (41)$$

where (38) follows because all the messages are independent, and in (41) we use the fact that dropping conditioning does not reduce entropy.

Similarly, for other receivers  $i_j$ ,  $\forall j \in \{2, 3, \dots, m-1\}$ , by providing  $S_{i_j}^n$ ,  $\mathcal{W}_{i_{j+1} i_{j+1}}^c \setminus \mathcal{W}_{i_j}^*$  through a genie we have

$$\begin{aligned} n\left(\sum_{k=1}^K R_{i_j i_k} - \epsilon\right) \leq h(S_{i_j}^n | \mathcal{W}_{i_j}^\dagger \setminus W_{i_j i_j}) - h(Z_{i_{j-1}}^n) \\ + h(Y_{i_j}^n | S_{i_j}^n) - h(S_{i_{j+1}}^n | \mathcal{W}_{i_{j+1}}^\dagger \setminus W_{i_{j+1} i_{j+1}}) \end{aligned} \quad (42)$$

Finally, for the receiver  $i_m$ , we provide  $S_{i_m}^n$ ,  $\mathcal{W}_{i_1 i_1}^c \setminus \mathcal{W}_{i_m}^*$  through a genie and obtain

$$n \left( \sum_{k=1}^K R_{i_m i_k} - \epsilon \right) \leq h(S_{i_m}^n | \mathcal{W}_{i_m}^\dagger \setminus \mathcal{W}_{i_m i_m}) - h(Z_{i_m-1}^n) + h(Y_{i_m}^n | S_{i_m}^n) - h(S_{i_1}^n | \mathcal{W}_{i_1}^\dagger \setminus \mathcal{W}_{i_1 i_1}) \quad (43)$$

Then taking the sum of  $n(\sum_{k=1}^K R_{i_j i_k} - \epsilon)$  for all  $j \in [m]$ , we have

$$n \left( \sum_{j=1}^m \sum_{k=1}^K R_{i_j i_k} - m\epsilon \right) \leq \sum_{j=1}^m \left[ h(Y_{i_j}^n | S_{i_j}^n) - h(Z_{i_j}^n) \right] \quad (44)$$

$$\leq \sum_{t=1}^n \sum_{j=1}^m \left[ h(Y_{i_j}(t) | S_{i_j}(t)) - h(Z_{i_j}(t)) \right] \quad (45)$$

where (45) follows the chain rule and the fact that dropping conditioning does not reduce entropy. Once again, using the fact that the circularly symmetric complex Gaussian distribution maximizes conditional differential entropy for a given covariance constraint and the condition (11), we can obtain the following desired outer bound in the GDoF sense, through the same set of manipulations as in Example 2,

$$\sum_{j=1}^m \hat{d}_{i_j} = \sum_{j=1}^m \sum_{k=1}^K d_{i_j i_k} \leq \sum_{j=1}^m (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}) \quad (46)$$

Now we can proceed to the second step to prove that under the condition (11), the  $K$ -user interference channel and its counterpart  $K \times K$   $X$  channel have the same sum-GDoF. According to Theorem 1, for the  $K$ -user interference channel, under the condition (11), to obtain its sum-GDoF  $d_{\Sigma, IC}$ , we need to solve the following linear programming (LP) problem

$$\max \sum_{i=1}^K d_i \quad (47)$$

$$\text{s.t. } 0 \leq d_i \leq \alpha_{ii}, \quad \forall i \in [K] \quad (48)$$

$$\sum_{j=1}^m d_{i_j} \leq \sum_{j=1}^m (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}),$$

$$\forall (i_1, i_2, \dots, i_m) \in \Pi_K, \quad \forall m \in \{2, 3, \dots, K\} \quad (49)$$

To get the sum-GDoF of its counterpart  $X$  channel  $d_{\Sigma, X}$ , we consider a similar LP problem. Note for this LP problem with the objective function  $\sum_{i=1}^K \hat{d}_i$ , from the first step, we know that at least it needs to follow two similar constraints to (48) and (49), in which each  $d_i$  is just replaced by  $\hat{d}_i$ . Thus we have  $d_{\Sigma, IC} \geq d_{\Sigma, X}$ . Obviously, in any case, the sum-GDoF of the  $K$ -user interference channel must be less than or equal to that of its counterpart  $X$  channel, i.e.,  $d_{\Sigma, IC} \leq d_{\Sigma, X}$ . Therefore, under the condition (11), we have established that the  $K$ -user interference channel and its counterpart  $X$  channel have the same sum-GDoF.

2) *Proof for the Constant Gap of Sum Capacity*: Based on the insight gained in the above GDoF study, for the TIN-optimal  $K \times K$   $X$  channel, we intend to characterize the sum channel capacity to within a constant gap of no more than  $K \log_2[K(K+1)]$  bits. To this end, first recall the achievability proof in [12]. By operating the  $K \times K$   $X$  channel as an interference channel, in which each transmitter  $i$  sends

one independent message  $W_i$  to its corresponding receiver  $i$  ( $\forall i \in [K]$ ), power control and TIN can achieve the following rate tuples  $(R_{1, \text{TIN}}, R_{2, \text{TIN}}, \dots, R_{K, \text{TIN}})$  satisfying

$$0 \leq R_{i, \text{TIN}} \leq \max \left\{ 0, \alpha_{ii} \log_2 P + \log_2 \left( \frac{1}{K} \right) \right\}, \quad \forall i \in [K], \quad (50)$$

$$\sum_{j=1}^m R_{i_j, \text{TIN}} \leq \max \left\{ 0, \sum_{j=1}^m \left[ (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}) \log_2 P + \log_2 \left( \frac{1}{K} \right) \right] \right\}, \quad (51)$$

for all cyclic sequences  $(i_1, i_2, \dots, i_m) \in \Pi_K$ ,  $\forall m \in \{2, 3, \dots, K\}$ .

Next, consider the converse. Define  $\hat{R}_i \triangleq \sum_{j=1}^K R_{i_j}$ . Start with the individual bounds,

$$\hat{R}_i \leq \log_2 \left( 1 + \sum_{j=1}^K P^{\alpha_{ij}} \right) \quad (52)$$

$$\leq \log_2 \left[ (K+1) P^{\alpha_{ii}} \right] \quad (53)$$

$$= \alpha_{ii} \log_2 P + \log_2(K+1) \quad (54)$$

Then for the cycle bounds, from (45), it is easy to obtain

$$\sum_{j=1}^m \hat{R}_{i_j} \leq \sum_{j=1}^m \log_2 \left[ \frac{(K+1) P^{\alpha_{i_j i_j}}}{P^{\alpha_{i_{j-1} i_j}}} \right] \quad (55)$$

$$= \sum_{j=1}^m \left[ (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}) \log_2 P + \log_2(K+1) \right] \quad (56)$$

for all cyclic sequences  $(i_1, i_2, \dots, i_m) \in \Pi_K$ ,  $\forall m \in \{2, 3, \dots, K\}$ .

Compare (50) with (54). Denote the difference between the achievable rate in (50) and the outer bound in (54) as  $\delta_{\hat{R}_i}$ . Consider the following two cases,

- $\alpha_{ii} \log_2 P + \log_2 \left( \frac{1}{K} \right) > 0$ : In this case, we have

$$\begin{aligned} \delta_{\hat{R}_i} &= \left[ \alpha_{ii} \log_2 P + \log_2(K+1) \right] \\ &\quad - \max \left\{ 0, \alpha_{ii} \log_2 P + \log_2 \left( \frac{1}{K} \right) \right\} \\ &= \left[ \alpha_{ii} \log_2 P + \log_2(K+1) \right] \\ &\quad - \left[ \alpha_{ii} \log_2 P + \log_2 \left( \frac{1}{K} \right) \right] \\ &= \log_2[K(K+1)] \end{aligned}$$

- $\alpha_{ii} \log_2 P + \log_2 \left( \frac{1}{K} \right) \leq 0$ : In this case, we have

$$\begin{aligned} \delta_{\hat{R}_i} &= \left[ \alpha_{ii} \log_2 P + \log_2(K+1) \right] \\ &\quad - \max \left\{ 0, \alpha_{ii} \log_2 P + \log_2 \left( \frac{1}{K} \right) \right\} \\ &= \alpha_{ii} \log_2 P + \log_2(K+1) \\ &\leq \log_2 K + \log_2(K+1) \\ &= \log_2[K(K+1)] \end{aligned}$$

In both cases, we obtain

$$\delta_{\hat{R}_i} \leq \log_2[K(K+1)], \quad \forall i \in [K] \quad (57)$$

Similarly, denote the difference between the achievable rate in (51) and the outer bound in (56) as  $\delta_{\sum_{j=1}^m \hat{R}_{i_j}}$ . Comparing (51) with (56), we always have

$$\delta_{\sum_{j=1}^m \hat{R}_{i_j}} \leq m \log_2[K(K+1)] \quad (58)$$

for all cyclic sequences  $(i_1, i_2, \dots, i_m) \in \Pi_K$ ,  $\forall m \in \{2, 3, \dots, K\}$ .

According to (57) and (58), we can characterize the sum channel capacity to within a constant gap of no more than  $K \log_2[K(K+1)]$  bits, which is only dependent on the number of users  $K$ .

### B. Proof for Theorem 3

It is easy to verify that when  $M \geq N$ , by defining  $\hat{d}_i = \sum_{j=1}^M d_{ij}$  ( $\forall i \in [N]$ ) and following the same argument as in the proof of Theorem 2, we can complete the proof. Therefore, hereafter we only consider the case where  $\kappa = M < N$ .

Without loss of generality, we assume that in the two permutations  $\Pi^T$  and  $\Pi^R$  which satisfy the condition (12),  $\Pi_i^T = i$ ,  $\forall i \in [M]$  and  $\Pi_j^R = j$ ,  $\forall j \in [N]$ , i.e., the following condition is satisfied,

$$\alpha_{ii} \geq \max_{j:j \neq i} \{\alpha_{ji}\} + \max_{k:k \neq i} \{\alpha_{ik}\}, \quad \forall i, k \in [M], \forall j \in [N] \quad (59)$$

In this case, similar to the proof of Theorem 2, the key step is to show that when (59) holds, for each individual bound and cycle bound in the  $M$ -user interference channel consisting of transmitters  $i \in [M]$  and receivers  $j \in [M]$ , if each  $d_i$  ( $\forall i \in [M]$ ) is replaced by<sup>4</sup>  $\bar{d}_i = \sum_{j=1}^N d_{ji}$ , the resulting bounds hold in the  $M \times N$   $X$  channel. Then based on the same argument of Theorem 2, we can prove the optimality of TIN for the sum-GDoF of the  $M \times N$   $X$  channel where  $M < N$ .

For the individual bounds, consider the degraded broadcast channel (BC) comprised of the transmitter  $i$  ( $\forall i \in [M]$ ) and all the receivers, eliminating all other transmitters and their messages. Since (59) is satisfied, the receiver  $i$  is the strongest receiver, and can decode all the messages from the transmitter  $i$ . Thus in the GDoF sense we have

$$\bar{d}_i = \sum_{j=1}^N d_{ji} \leq \alpha_{ii}, \quad \forall i \in [M] \quad (60)$$

Now the only task left is to prove that in the  $X$  channel setting, by replacing  $d_i$  with  $\bar{d}_i$ , all the cycle bounds still hold. Before exploring the proof details, let's see an intuitive sketch of proof first for a  $2 \times 4$   $X$  channel as illustrated in Fig. 6. For this  $X$  channel, when (59) is satisfied, we intend to prove

$$\bar{d}_1 + \bar{d}_2 \leq (\alpha_{11} + \alpha_{22}) - (\alpha_{12} + \alpha_{21})$$

<sup>4</sup>It is noteworthy that in the case of  $M < N$ ,  $\bar{d}_i = \sum_{j=1}^N d_{ji}$  ( $\forall i \in [M]$ ) is different from  $\hat{d}_i = \sum_{j=1}^M d_{ij}$  ( $\forall i \in [N]$ ) we defined in the case of  $M \geq N$ . It is easy to find that in the  $M \times N$   $X$  channel where  $M < N$ , to recover the corresponding individual and cycle bounds of the  $M$ -user interference channel,  $\bar{d}_i = \sum_{j=1}^N d_{ji}$  ( $\forall i \in [M]$ ) is the natural choice.

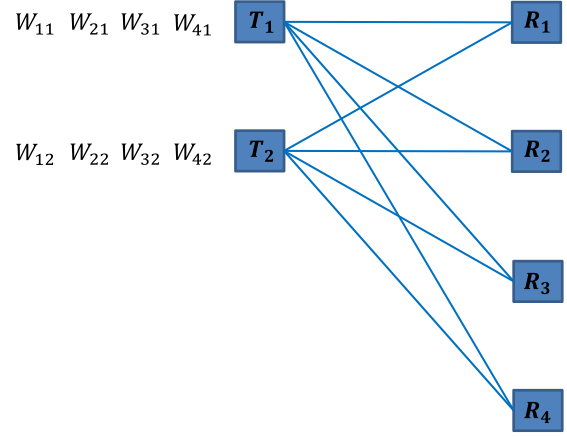


Fig. 6. A  $2 \times 4$   $X$  channel.

*An Intuitive Sketch of Proof for the Above Cycle Bound:* In this  $2 \times 4$   $X$  channel, we assume  $\alpha_{31} \geq \alpha_{41}$  and  $\alpha_{42} \geq \alpha_{32}$ . The proof for all the other cases follows similarly. Define the message set  $\tilde{\mathcal{W}} \triangleq \{W_{ki}\}$ ,  $\forall i \in \{1, 2\}$ ,  $\forall k \in \{1, 2, 3, 4\}$ . Also define

$$S_1(t) = h_{21}X_1(t) + Z_2(t)$$

$$S_2(t) = h_{12}X_2(t) + Z_1(t)$$

For the receiver 1, we provide  $S_1^n$  and  $W_{21}$  through a genie. From Fano's inequality, we obtain

$$n(R_{11} + R_{12} - \epsilon) \leq I(W_{11}, W_{12}; Y_1^n, S_1^n, W_{21}) \quad (61)$$

$$= I(W_{11}, W_{12}; Y_1^n, S_1^n | W_{21}) \quad (62)$$

$$= h(Y_1^n, S_1^n | W_{21}) - h(Y_1^n, S_1^n | W_{21}, W_{11}, W_{12}) \quad (63)$$

$$= h(S_1^n | W_{21}) + h(Y_1^n | S_1^n, W_{21}) - h(Y_1^n | W_{21}, W_{11}, W_{12}) - h(S_1^n | Y_1^n, W_{21}, W_{11}, W_{12}) \quad (64)$$

$$\leq h(S_1^n | W_{21}) + h(Y_1^n | S_1^n) - h(Y_1^n | W_{21}, W_{11}, W_{12}) - h(S_1^n | Y_1^n, W_{21}, W_{11}, W_{12}, W_{31}, W_{41}) \quad (65)$$

$$= h(S_1^n | W_{21}) + h(Y_1^n | S_1^n) - h(Y_1^n | W_{21}, W_{11}, W_{12}) - h(S_1^n | Y_1^n, W_{21}, W_{11}, W_{12}, W_{31}, W_{41}, X_1^n) \quad (66)$$

$$= h(S_1^n | W_{21}) + h(Y_1^n | S_1^n) - h(Y_1^n | W_{21}, W_{11}, W_{12}) - h(Z_2^n) \quad (67)$$

$$\leq h(S_1^n | W_{21}) + h(Y_1^n | S_1^n) - h(Y_3^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) - h(S_2^n | W_{12}) - h(Z_2^n) - n o(\log(P)) \quad (68)$$

where (62) holds since  $W_{21}$  is independent of  $W_{11}$  and  $W_{12}$ , (65) follows that dropping conditioning (in the second term) does not reduce entropy and adding conditioning (in the last term) does not increase entropy, (66) holds since we can reconstruct  $X_1^n$  from  $W_{11}$ ,  $W_{21}$ ,  $W_{31}$  and  $W_{41}$ , and in (68)  $\tilde{\mathcal{W}}_{\{31,41\}}^c$  denotes the complement of  $\{W_{31}, W_{41}\}$  in  $\tilde{\mathcal{W}}$ . *The last inequality in (68) is the key step of the proof.* Intuitively, it is due to the fact that out of the  $\alpha_{11} \log(P)$  bit levels of  $Y_1$  that are above the noise floor,  $S_2$  is contained in the lowest  $\alpha_{12} \log(P)$  bit levels of  $Y_1$ , whereas only the top  $\alpha_{31} \log(P)$  bit levels are seen by the receiver 3. Since  $\alpha_{11} \geq \alpha_{12} + \alpha_{31}$ , these bit levels do not overlap, i.e., they can be recovered from  $Y_1$  within a bounded entropy gap. As shown later, this argument becomes evident in a deterministic approach.

$$\begin{aligned}
 Y_i(t) &= \sum_{k=1}^M h_{ik} X_k(t) + Z_i(t) \\
 &= \sum_{k=1}^M \left[ (h_{ik}^R X_k^R(t) - h_{ik}^I X_k^I(t)) + j(h_{ik}^I X_k^R(t) + h_{ik}^R X_k^I(t)) \right] + Z_i(t), \quad \forall i \in [N]
 \end{aligned} \tag{73}$$

$$\begin{aligned}
 \hat{Y}_i(t) &= \sum_{k=1}^M \left[ \lfloor \text{sign}(\bar{X}_k^R(t)) h_{ik}^R \sum_{b=1}^{m_{ik}^R} \bar{X}_{k,b}^R(t) 2^{-b} \rfloor - \lfloor \text{sign}(\bar{X}_k^I(t)) h_{ik}^I \sum_{b=1}^{m_{ik}^I} \bar{X}_{k,b}^I(t) 2^{-b} \rfloor \right. \\
 &\quad \left. + j \left( \lfloor \text{sign}(\bar{X}_k^R(t)) h_{ik}^I \sum_{b=1}^{m_{ik}^I} \bar{X}_{k,b}^R(t) 2^{-b} \rfloor + \lfloor \text{sign}(\bar{X}_k^I(t)) h_{ik}^R \sum_{b=1}^{m_{ik}^R} \bar{X}_{k,b}^I(t) 2^{-b} \rfloor \right) \right], \quad \forall i \in [N]
 \end{aligned} \tag{74}$$

Next, consider the degraded BC comprised of the transmitter 1 and the receivers 3 and 4. Since  $\alpha_{31} \geq \alpha_{41}$ , we have

$$\begin{aligned}
 n(R_{31} + R_{41} - \epsilon) &\leq I(W_{31}, W_{41}; Y_3^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) \tag{69} \\
 &= h(Y_3^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) - h(Z_3^n) \tag{70}
 \end{aligned}$$

Adding (68) and (70), we obtain

$$\begin{aligned}
 n(R_{11} + R_{12} + R_{31} + R_{41} - 2\epsilon) \\
 \leq h(S_1^n | W_{21}) + h(Y_1^n | S_1^n) - h(S_2^n | W_{12}) - n o(\log(P))
 \end{aligned} \tag{71}$$

Similarly, we have

$$\begin{aligned}
 n(R_{21} + R_{22} + R_{32} + R_{42} - 2\epsilon) \\
 \leq h(S_2^n | W_{12}) + h(Y_2^n | S_2^n) - h(S_1^n | W_{21}) - n o(\log(P))
 \end{aligned} \tag{72}$$

Finally, through adding (71) and (72) together and some other manipulations, we can obtain the desired outer bound,

$$\begin{aligned}
 n(R_\Sigma - 4\epsilon) &\leq h(Y_1^n | S_1^n) + h(Y_2^n | S_2^n) - n o(\log(P)) \\
 \Rightarrow \bar{d}_1 + \bar{d}_2 &\leq (\alpha_{11} + \alpha_{22}) - (\alpha_{12} + \alpha_{21})
 \end{aligned}$$

In the following, in order to make the intuitive justification of the key step (68) rigorous, we take a *deterministic* approach [16]–[22]. We first show that the sum capacity of the original complex Gaussian  $X$  channel is upper bounded by that of one suitably-chosen deterministic channel up to a constant gap. Then by upper bounding that deterministic channel, we obtain the desired converse of the original Gaussian channel as well. Such a deterministic approach has been shown instrumental to provide approximate capacity characterization for various Gaussian networks, e.g., the 2-user interference channel [20],  $K$ -user interference channel [21] and  $K \times K \times K$  networks [22]. Also, several deterministic models have been introduced in previous works [17], [19], [21]. The one adopted in the following is mainly inspired by the Avestimehr-Diggavi-Tse linear deterministic model and the truncated deterministic model in [19].

Recall the original complex Gaussian  $M \times N$   $X$  channel. Denote

$$\begin{aligned}
 X_k(t) &= X_k^R(t) + jX_k^I(t) \\
 h_{ik} &= \sqrt{P^{\alpha_{ik}}} e^{j\theta_{ik}} = h_{ik}^R + jh_{ik}^I
 \end{aligned}$$

The input-output relationship can be written as the equation (73) given at the top of this page, where  $E[|X_i(t)|^2] \leq 1$  and  $Z_i(t) \sim \mathcal{CN}(0, 1)$ . To facilitate the following deterministic approach, by scaling the output, we may set

$$E[|X_i(t)|^2] \leq 2, \quad Z_i(t) \sim \mathcal{CN}(0, 2),$$

which does not affect the channel capacity of the Gaussian model (73).

In this paper, we consider the deterministic model (74) in top of this page, where  $\lfloor x \rfloor$  is the truncation function which maps  $x$  to its integer part,  $m_{ik}^R \triangleq \lfloor \log_2 |h_{ik}^R| \rfloor$ ,  $m_{ik}^I \triangleq \lfloor \log_2 |h_{ik}^I| \rfloor$ , the real and imaginary parts of the input signal  $\bar{X}_i(t) = \bar{X}_i^R(t) + j\bar{X}_i^I(t)$  both satisfy the unit *peak* power constraint, and  $\bar{X}_{i,b}^R(t)$  ( $\bar{X}_{i,b}^I(t)$ ) is the  $b$ -th bit in the fractional part of  $|\bar{X}_i^R(t)|$  ( $|\bar{X}_i^I(t)|$ ) in the binary expansion.<sup>5</sup> In the following, to distinguish the deterministic model in (74) from others, we call it the *truncated binary-expansion deterministic model*. The following lemma shows that the sum capacity of the Gaussian  $X$  channel in (73) is upper bounded by that of the truncated binary-expansion deterministic model in (74) up to a constant gap.

*Lemma 1:* The sum capacity of the complex Gaussian  $M \times N$   $X$  channel is upper bounded by the sum capacity of its corresponding truncated binary-expansion deterministic model up to a constant gap.

The proof for the above lemma follows [20] and is relegated to Appendix A for completeness.

Now define  $m_{ij} \triangleq \lfloor \frac{1}{2} \log_2 P^{\alpha_{ij}} \rfloor$ . Since  $P > 1$  and  $\alpha_{ii} \geq \alpha_{ij} + \alpha_{ki}$ ,  $\forall i \notin \{j, k\}$ , we have

$$\lfloor \frac{\alpha_{ii}}{2} \log_2 P \rfloor \geq \lfloor \frac{(\alpha_{ij} + \alpha_{ki})}{2} \log_2 P \rfloor \tag{75}$$

$$\Rightarrow \lfloor \frac{\alpha_{ii}}{2} \log_2 P \rfloor \geq \lfloor \frac{\alpha_{ij}}{2} \log_2 P \rfloor + \lfloor \frac{\alpha_{ki}}{2} \log_2 P \rfloor \tag{76}$$

$$\Rightarrow m_{ii} \geq m_{ij} + m_{ki} \quad \forall i, j, k, \quad i \notin \{j, k\} \tag{77}$$

<sup>5</sup>We can write the real-valued signal  $|\bar{X}_i^R|$  ( $|\bar{X}_i^R| \leq 1$ ) in terms of its binary expansion as

$$|\bar{X}_i^R| = \sum_{b=1}^{\infty} \bar{X}_{i,b}^R 2^{-b} = 0.\bar{X}_{i,1}^R \bar{X}_{i,2}^R \bar{X}_{i,3}^R \dots$$



In order to convey the key ingredients of the proof more clearly, next we give an example for the *real* Gaussian  $2 \times 4$   $X$  channel, and then generalize the proof to the *complex* Gaussian  $M \times N$  ( $M < N$ )  $X$  channel.

*Example 3:* Consider the real Gaussian  $X$  channel with 2 transmitters and 4 receivers, where (59) is satisfied. In this example we still assume  $\alpha_{31} \geq \alpha_{41}$  and  $\alpha_{42} \geq \alpha_{32}$ . As previously mentioned, the proof for all the other cases follows similarly. Also define the message set  $\tilde{\mathcal{W}} \triangleq \{W_{ki}\}$ ,  $\forall i \in \{1, 2\}$ ,  $\forall k \in \{1, 2, 3, 4\}$ .

Recall that for this  $2 \times 4$   $X$  channel, we intend to prove the following cycle bound

$$\bar{d}_1 + \bar{d}_2 \leq \frac{1}{2} \left[ (\alpha_{11} + \alpha_{22}) - (\alpha_{12} + \alpha_{21}) \right]$$

where the factor  $\frac{1}{2}$  is due to the fact that the Gaussian  $X$  channel is real-valued.

We start with the corresponding truncated binary-expansion deterministic model. Define

$$\hat{S}_1(t) = \lfloor \text{sign}(\bar{X}_1(t)) h_{21} \sum_{b=1}^{m_{21}} \bar{X}_{1,b}(t) 2^{-b} \rfloor \quad (78)$$

$$\hat{S}_2(t) = \lfloor \text{sign}(\bar{X}_2(t)) h_{12} \sum_{b=1}^{m_{12}} \bar{X}_{2,b}(t) 2^{-b} \rfloor. \quad (79)$$

Also define

$$\bar{X}_{31,s}(t) = \text{sign}(\bar{X}_1(t)) \sum_{b=1}^{m_{31}} \bar{X}_{1,b}(t) 2^{-b} \quad (80)$$

Thus the output of the receiver 1 can be written as

$$\begin{aligned} \hat{Y}_1(t) &= \lfloor \text{sign}(\bar{X}_1(t)) h_{11} \sum_{b=1}^{m_{11}} \bar{X}_{1,b}(t) 2^{-b} \rfloor \\ &\quad + \lfloor \text{sign}(\bar{X}_2(t)) h_{12} \sum_{b=1}^{m_{12}} \bar{X}_{2,b}(t) 2^{-b} \rfloor \end{aligned} \quad (81)$$

$$\begin{aligned} &= \lfloor \text{sign}(\bar{X}_1(t)) h_{11} \sum_{b=1}^{m_{31}} \bar{X}_{1,b}(t) 2^{-b} \rfloor + \hat{C}_1(t) \\ &\quad + \lfloor \text{sign}(\bar{X}_1(t)) h_{11} \sum_{b=m_{31}+1}^{m_{11}} \bar{X}_{1,b}(t) 2^{-b} \rfloor + \hat{S}_2(t) \end{aligned} \quad (82)$$

$$\begin{aligned} &= \underbrace{\lfloor h_{11} \bar{X}_{31,s}(t) \rfloor}_{\hat{Y}_{1,u}(t)} + \hat{C}_1(t) \\ &\quad + \underbrace{\lfloor \text{sign}(\bar{X}_1(t)) h_{11} \sum_{b=m_{31}+1}^{m_{11}} \bar{X}_{1,b}(t) 2^{-b} \rfloor}_{\hat{Y}_{1,l}(t)} + \hat{S}_2(t) \end{aligned} \quad (83)$$

where  $\hat{C}_1(t)$  may take a value from  $\{-1, 0, 1\}$ .

For the receiver 1, we have

$$\begin{aligned} n(R_{11} + R_{12} - \epsilon) &\leq I(W_{11}, W_{12}; \hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n, \hat{S}_1^n | W_{21}) \end{aligned} \quad (84)$$

$$\begin{aligned} &= H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n, \hat{S}_1^n | W_{21}) \\ &\quad - H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n, \hat{S}_1^n | W_{21}, W_{11}, W_{12}) \end{aligned} \quad (85)$$

$$\begin{aligned} &= H(\hat{S}_1^n | W_{21}) + H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | \hat{S}_1^n, W_{21}) \\ &\quad - H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | W_{21}, W_{11}, W_{12}) \\ &\quad - H(\hat{S}_1^n | \hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n, W_{21}, W_{11}, W_{12}) \end{aligned} \quad (86)$$

$$\begin{aligned} &\leq H(\hat{S}_1^n | W_{21}) + H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | \hat{S}_1^n) \\ &\quad - H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | W_{21}, W_{11}, W_{12}) \end{aligned} \quad (87)$$

where (87) follows that dropping conditioning does not reduce entropy. Now consider the last term in (87),

$$\begin{aligned} &H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | W_{21}, W_{11}, W_{12}) \\ &= H(\hat{Y}_{1,u}^n | W_{21}, W_{11}, W_{12}) \\ &\quad + H(\hat{Y}_{1,l}^n, \hat{C}_1^n | \hat{Y}_{1,u}^n, W_{21}, W_{11}, W_{12}) \end{aligned} \quad (88)$$

$$\begin{aligned} &= H(\bar{X}_{31,s}^n | W_{21}, W_{11}, W_{12}) \\ &\quad + H(\hat{Y}_{1,l}^n, \hat{C}_1^n | \hat{Y}_{1,u}^n, W_{21}, W_{11}, W_{12}) \end{aligned} \quad (89)$$

$$\geq H(\bar{X}_{31,s}^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) + H(\hat{S}_2^n | \hat{Y}_{1,u}^n, W_{21}, W_{11}, W_{12}) \quad (90)$$

$$= H(\hat{Y}_3^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) + H(\hat{S}_2^n | W_{12}) \quad (91)$$

where (89) holds since the function  $f : \bar{X}_{31,s} \rightarrow \hat{Y}_{1,u}$  is bijective when  $m_{11} \geq m_{31}$ , and (91) follows that conditioning on the messages  $\tilde{\mathcal{W}}_{\{31,41\}}^c$ , the function  $f : \bar{X}_{31,s} \rightarrow \hat{Y}_3$  is bijective. Note that the equations (88)-(91) correspond to the key step (68) in the intuitive proof given before.

Plugging (91) into (87), we have

$$\begin{aligned} n(R_{11} + R_{12} - \epsilon) &\leq H(\hat{S}_1^n | W_{21}) + H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | \hat{S}_1^n) \\ &\quad - H(\hat{S}_2^n | W_{12}) - H(\hat{Y}_3^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) \end{aligned} \quad (92)$$

Then, consider the degraded BC comprised of the transmitter 1 and the receivers 3 and 4. Since  $m_{31} \geq m_{41}$ , we have

$$n(R_{31} + R_{41} - \epsilon) \leq I(W_{31}, W_{41}; \hat{Y}_3^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) \quad (93)$$

$$= H(\hat{Y}_3^n | \tilde{\mathcal{W}}_{\{31,41\}}^c) \quad (94)$$

Combining (92) and (94), we obtain

$$\begin{aligned} n(R_{11} + R_{12} + R_{31} + R_{41} - 2\epsilon) &\leq H(\hat{S}_1^n | W_{21}) + H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | \hat{S}_1^n) - H(\hat{S}_2^n | W_{12}) \end{aligned} \quad (95)$$

Similarly, by considering the receiver 2 and the degraded BC comprised of the transmitter 2 and the receivers 3 and 4, we obtain

$$\begin{aligned} n(R_{21} + R_{22} + R_{32} + R_{42} - 2\epsilon) &\leq H(\hat{S}_2^n | W_{12}) + H(\hat{Y}_{2,u}^n, \hat{Y}_{2,l}^n, \hat{C}_2^n | \hat{S}_2^n) - H(\hat{S}_1^n | W_{21}) \end{aligned} \quad (96)$$

Adding (95) and (96), the sum capacity of this deterministic  $2 \times 4$   $X$  channel is upper bounded by

$$n(R_{\Sigma,D} - 4\epsilon) \leq H(\hat{Y}_{1,u}^n, \hat{Y}_{1,l}^n, \hat{C}_1^n | \hat{S}_1^n) + H(\hat{Y}_{2,u}^n, \hat{Y}_{2,l}^n, \hat{C}_2^n | \hat{S}_2^n) \quad (97)$$

$$\leq \sum_{t=1}^n \left[ H(\hat{Y}_{1,u}(t) | \hat{S}_1(t)) + H(\hat{Y}_{1,l}(t) | \hat{S}_1(t)) + H(\hat{C}_1(t)) + H(\hat{Y}_{2,u}(t) | \hat{S}_2(t)) + H(\hat{Y}_{2,l}(t) | \hat{S}_2(t)) + H(\hat{C}_2(t)) \right] \quad (98)$$

where the last inequality follows from the chain rule and the fact that dropping conditioning does not reduce entropy.

Next, for the term  $H(\hat{Y}_{1,u}(t) | \hat{S}_1(t)) + H(\hat{Y}_{1,l}(t) | \hat{S}_1(t))$ , we consider two cases,

- $m_{21} \geq m_{31}$ :

$$H(\hat{Y}_{1,u}(t) | \hat{S}_1(t)) + H(\hat{Y}_{1,l}(t) | \hat{S}_1(t)) \leq 0 + (m_{11} - m_{21}) + \text{constant} \quad (99)$$

$$= (m_{11} - m_{21}) + \text{constant} \quad (100)$$

where (99) follows that conditioning on  $\hat{S}_1$ , out of the received signal  $\hat{Y}_{1,l}$ , both the signals from transmitter 1 and 2 have at most  $m_{11} - m_{21}$  bit levels, and the signs of the signals and carry-overs due to the sum of two such signals can only induce a loss of constant bits.

- $m_{21} < m_{31}$ : Similarly, we have

$$H(\hat{Y}_{1,u}(t) | \hat{S}_1(t)) + H(\hat{Y}_{1,l}(t) | \hat{S}_1(t)) \leq (m_{31} - m_{21}) + (m_{11} - m_{31}) + \text{constant} \quad (101)$$

$$= (m_{11} - m_{21}) + \text{constant} \quad (102)$$

Due to symmetry, we always have

$$H(\hat{Y}_{2,u}(t) | \hat{S}_2(t)) + H(\hat{Y}_{2,l}(t) | \hat{S}_2(t)) \leq (m_{22} - m_{12}) + \text{constant} \quad (103)$$

Therefore,

$$n(R_{\Sigma,D} - 4\epsilon) \leq \sum_{t=1}^n \left[ (m_{11} - m_{21}) + (m_{22} - m_{12}) + \text{constant} \right] \quad (104)$$

According to Lemma 1, for the sum capacity of the original Gaussian  $X$  channel  $R_{\Sigma,G}$ , we have

$$R_{\Sigma,G} \leq R_{\Sigma,D} + \text{constant} \quad (105)$$

$$\leq (m_{11} - m_{21}) + (m_{22} - m_{12}) + \text{constant} \quad (106)$$

$$\leq \frac{1}{2} \left[ (\alpha_{11} - \alpha_{21}) + (\alpha_{22} - \alpha_{12}) \right] \log_2 P + \text{constant} \quad (107)$$

Finally, we obtain the desired GDoF cycle bound,

$$\bar{d}_1 + \bar{d}_2 \leq \frac{1}{2} \left[ (\alpha_{11} - \alpha_{21}) + (\alpha_{22} - \alpha_{12}) \right]. \quad (108)$$

□

Equipped with the bounding techniques in the above example, we can extend the proof to the complex Gaussian  $M \times N$  ( $M < N$ )  $X$  channels. To obtain an arbitrary desired cycle bound

$$\sum_{j=1}^m \bar{d}_{i_j} \leq \sum_{j=1}^m (\alpha_{i_j i_j} - \alpha_{i_{j-1} i_j}), \quad \forall (i_1, i_2, \dots, i_m) \in \Pi_M, \quad \forall m \in \{2, 3, \dots, M\}, \quad (109)$$

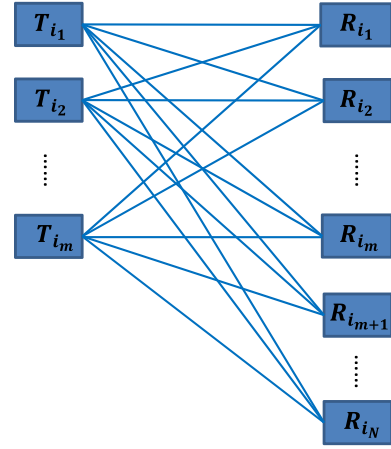


Fig. 7. An  $m \times N$   $X$  channel ( $m < N$ ).

consider the subnetwork consisting of all the receivers and the transmitters  $\{i_1, i_2, \dots, i_m\}$  in Fig. 7, eliminating all other transmitters and their messages. Note that in Fig. 7, for the receivers other than  $i_1, i_2, \dots, i_m$ , they are labelled as  $i_{m+1}, i_{m+2}, \dots, i_N$ , respectively, and the index ordering of these receivers does not affect the desired outer bound in (109). Define the message set  $\tilde{\mathcal{W}} \triangleq \{W_{ikij}, \forall j \in [m], \forall k \in [N]\}$ . Also define  $\mathcal{W} \triangleq \{W_{ikij}\}$ ,  $\mathcal{W}_{i_k}^* \triangleq \{W_{i_k i_1}, W_{i_k i_2}, \dots, W_{i_k i_m}\}$ , and  $\mathcal{W}_{i_j'} \triangleq \{W_{i_{m+1} i_j}, W_{i_{m+2} i_j}, \dots, W_{i_N i_j}\}$ ,  $\forall j, k \in [m]$ . Similarly,  $\mathcal{W}_{\mathcal{S}}^c$  denotes  $\mathcal{W} \setminus \mathcal{W}_{\mathcal{S}}$ , where  $\mathcal{S}$  is a subset of message indices.

To simplify the proof, we construct the following channel as shown in Fig. 8, which upper bounds the sum channel capacity of the original complex Gaussian  $X$  channel in Fig. 7:

- Step 1: We start with an  $m \times m$   $X$  channel with channel coefficients  $h_{ikij}, \forall k, j \in [m]$ .
- Step 2: For each transmitter  $i_j, \forall j \in [m]$ , we create another  $N - m$  virtual receivers. The virtual receiver  $R'_{ikij}, \forall k \in \{m+1, m+2, \dots, N\}$ , only connects to the transmitter  $i_j$  with the channel coefficient  $h_{ikij}$  and desires the message  $W_{ikij}$  from the transmitter  $i_j$ . Note now there are  $m \times N$  messages totally in the network.
- Step 3: For each receiver  $i_k, k \in [m]$ , we can rotate its channel output by multiplying it with  $e^{-j\theta_{i_k i_k}}$  to make the link from transmitter  $i_k$  to receiver  $i_k$  real-valued, which does not affect the capacity of this channel. Similarly, for each virtual receiver  $R'_{ikij}, j \in [m], k \in \{m+1, m+2, \dots, N\}$ , we can rotate its channel output by multiplying it with  $e^{-j\theta_{i_k i_j}}$  to make its only connected link real-valued, which again does not affect the channel capacity. Therefore, without loss of generality, we can assume that in Fig. 8, all the links in red are real-valued.
- Step 4: The input signal  $X_{i_j}(t)$  satisfies the power constraint  $E[|X_{i_j}(t)|^2] \leq 2, \forall j \in [m]$ , and the AWGN seen at all the receivers are independent and with zero mean and variance 2.

Then combining with Lemma 1, to obtain the desired cycle bound in (109), we only need to consider the truncated binary-expansion deterministic model of the constructed channel in Fig. 8.

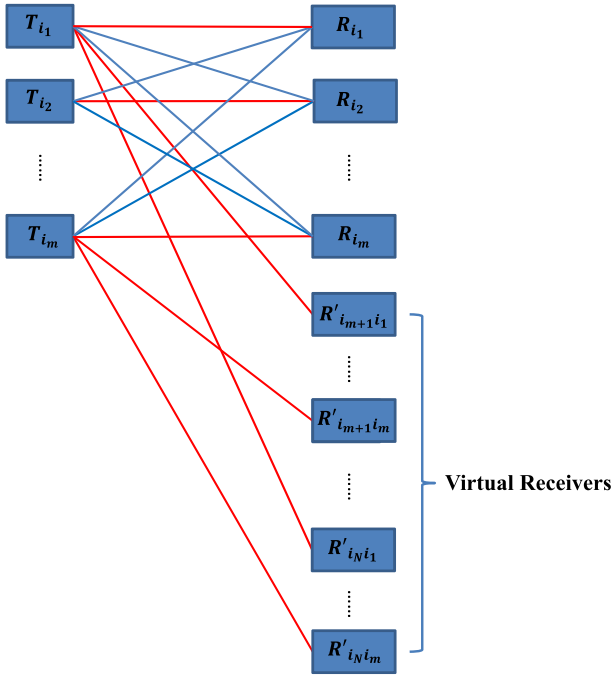


Fig. 8. The constructed channel which upper-bounds the sum channel capacity of the  $m \times N$  X channel in Fig. 7. The red links are real-valued by rotating the phase of the received signal at the corresponding receivers.

For the receivers  $i_j, \forall j \in [m]$ , the channel output can be written in the matrix form (110), as shown at the bottom of this page.

While for the virtual receivers, the channel output is given by the following equations

$$\hat{Y}_{ikj}(t) = \begin{pmatrix} \lfloor \text{sign}(\bar{X}_{ij}^R(t)) h_{ikij}^R \sum_{b=1}^{m_{ikij}^R} \bar{X}_{ij,b}^R(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{ij}^I(t)) h_{ikij}^I \sum_{b=1}^{m_{ikij}^I} \bar{X}_{ij,b}^I(t) 2^{-b} \rfloor \end{pmatrix}, \quad \forall j \in [m], \quad \forall k \in \{m+1, m+2, \dots, N\}. \quad (111)$$

Next, we present a lemma that will be used later.

**Lemma 2:** For any  $j \in [m]$ , define  $\hat{S}_j(t)$  and  $\hat{S}'_j(t)$  as (112) and (113), as shown at the bottom of this page, respectively, where the modulo- $m$  arithmetic is implicitly used

$$\hat{Y}_{ij}(t) = \begin{pmatrix} \hat{Y}_{ij}^R(t) \\ \hat{Y}_{ij}^I(t) \end{pmatrix} = \begin{pmatrix} \lfloor \text{sign}(\bar{X}_{ij}^R(t)) h_{ijij}^R \sum_{b=1}^{m_{ijij}^R} \bar{X}_{ij,b}^R(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{ij}^I(t)) h_{ijij}^I \sum_{b=1}^{m_{ijij}^I} \bar{X}_{ij,b}^I(t) 2^{-b} \rfloor \end{pmatrix} + \sum_{k=1, k \neq j}^m \begin{pmatrix} \lfloor \text{sign}(\bar{X}_{ik}^R(t)) h_{ijik}^R \sum_{b=1}^{m_{ijik}^R} \bar{X}_{ik,b}^R(t) 2^{-b} \rfloor - \lfloor \text{sign}(\bar{X}_{ik}^I(t)) h_{ijik}^I \sum_{b=1}^{m_{ijik}^I} \bar{X}_{ik,b}^I(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{ik}^R(t)) h_{ijik}^I \sum_{b=1}^{m_{ijik}^I} \bar{X}_{ik,b}^I(t) 2^{-b} \rfloor + \lfloor \text{sign}(\bar{X}_{ik}^I(t)) h_{ijik}^R \sum_{b=1}^{m_{ijik}^R} \bar{X}_{ik,b}^R(t) 2^{-b} \rfloor \end{pmatrix} \quad (110)$$

$$\hat{S}_j(t) = \begin{pmatrix} \text{sign}(\bar{X}_{ij}^R(t)) \sum_{b=1}^{\max\{m_{ij-1ij}^R, m_{ij-1ij}^I\}} \bar{X}_{ij,b}^R(t) 2^{-b} \\ \text{sign}(\bar{X}_{ij}^I(t)) \sum_{b=1}^{\max\{m_{ij-1ij}^R, m_{ij-1ij}^I\}} \bar{X}_{ij,b}^I(t) 2^{-b} \end{pmatrix}, \quad (112)$$

$$\hat{S}'_j(t) = \begin{pmatrix} \lfloor \text{sign}(\bar{X}_{ij}^R(t)) h_{ij-1ij}^R \sum_{b=1}^{m_{ij-1ij}^R} \bar{X}_{ij,b}^R(t) 2^{-b} \rfloor - \lfloor \text{sign}(\bar{X}_{ij}^I(t)) h_{ij-1ij}^I \sum_{b=1}^{m_{ij-1ij}^I} \bar{X}_{ij,b}^I(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{ij}^R(t)) h_{ij-1ij}^I \sum_{b=1}^{m_{ij-1ij}^I} \bar{X}_{ij,b}^I(t) 2^{-b} \rfloor + \lfloor \text{sign}(\bar{X}_{ij}^I(t)) h_{ij-1ij}^R \sum_{b=1}^{m_{ij-1ij}^R} \bar{X}_{ij,b}^R(t) 2^{-b} \rfloor \end{pmatrix} \quad (113)$$

on the user indices, e.g.,  $i_0 = i_m$ . Then  $f : \hat{S}_{ij}(t) \rightarrow \hat{S}'_j(t)$  is bijective.

The proof of Lemma 2 is presented in Appendix B.

Assume that in Fig. 8,  $R_{i_1}^{*}$  is the strongest virtual receiver connected to transmitter  $i_1$ , i.e.,  $|h_{i_1 i_1}^*| = \max_{j \in \{m+1, \dots, N\}} \{|h_{i_1 j}\}|$ . In the truncated binary-expansion deterministic model, the output of receiver  $i_1$  can be rewritten as (114)-(116), as shown in top of the next page, where  $\hat{C}_{i_1}^R(t)$  and  $\hat{C}_{i_1}^I(t)$  can both take values from  $\{-1, 0, 1\}$ . Define

$$\bar{X}_{i_1^*i_1, S}^R(t) = \text{sign}(\bar{X}_{i_1}^R(t)) \sum_{b=1}^{m_{i_1^*i_1}^R} \bar{X}_{i_1,b}^R(t) 2^{-b} \quad (117)$$

$$\bar{X}_{i_1^*i_1, S}^I(t) = \text{sign}(\bar{X}_{i_1}^I(t)) \sum_{b=1}^{m_{i_1^*i_1}^I} \bar{X}_{i_1,b}^I(t) 2^{-b} \quad (118)$$

Also define the sum of the terms in (115) and (116) as  $\hat{Y}_{i_1, l}(t)$  (see equation (119), as shown in top of the next page).

Then we have

$$\hat{Y}_{i_1}(t) = \hat{Y}_{i_1, u}(t) + \hat{Y}_{i_1, l}(t) + \hat{C}_{i_1}(t) \quad (120)$$

For the receiver  $i_1$ , starting from Fano's inequality, we obtain

$$n \left( \sum_{j=1}^m R_{i_1 j} - \epsilon \right) \leq I(\mathcal{W}_{i_1}^*; \hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n, \hat{S}_{i_1}^n | \mathcal{W}_{i_2}^c \setminus \mathcal{W}_{i_1}^*) \quad (121)$$

$$= H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n, \hat{S}_{i_1}^n | \mathcal{W}_{i_2}^c \setminus \mathcal{W}_{i_1}^*) - H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n, \hat{S}_{i_1}^n | \mathcal{W}_{i_2}^c) \quad (122)$$

$$= H(\hat{S}_{i_1}^n | \mathcal{W}_{i_2}^c \setminus \mathcal{W}_{i_1}^*) + H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n, \hat{S}_{i_1}^n | \mathcal{W}_{i_2}^c \setminus \mathcal{W}_{i_1}^*) - H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n | \mathcal{W}_{i_2}^c) - H(\hat{S}_{i_1}^n | \hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n, \mathcal{W}_{i_2}^c) \quad (123)$$

$$\leq H(\hat{S}_{i_1}^n | \mathcal{W}_{i_2}^c, \mathcal{W}_{i_3}^c, \dots, \mathcal{W}_{i_m}^c) + H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n | \hat{S}_{i_1}^n) - H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n | \mathcal{W}_{i_2}^c) \quad (124)$$

$$\hat{Y}_{i_1}(t) = \underbrace{\left( \begin{array}{c} \lfloor \text{sign}(\bar{X}_{i_1}^R(t)) h_{i_1 i_1}^R \sum_{b=1}^{m_{i_1^* i_1}^R} \bar{X}_{i_1, b}^R(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{i_1}^I(t)) h_{i_1 i_1}^R \sum_{b=1}^{m_{i_1^* i_1}^R} \bar{X}_{i_1, b}^I(t) 2^{-b} \rfloor \end{array} \right)}_{\hat{Y}_{i_1, u}(t)} + \underbrace{\left( \begin{array}{c} \hat{C}_{i_1}^R(t) \\ \hat{C}_{i_1}^I(t) \end{array} \right)}_{\hat{C}_{i_1}(t)} \quad (114)$$

$$+ \left( \begin{array}{c} \lfloor \text{sign}(\bar{X}_{i_1}^R(t)) h_{i_1 i_1}^R \sum_{b=m_{i_1^* i_1}^R+1}^{m_{i_1^* i_1}^R} \bar{X}_{i_1, b}^R(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{i_1}^I(t)) h_{i_1 i_1}^R \sum_{b=m_{i_1^* i_1}^R+1}^{m_{i_1^* i_1}^R} \bar{X}_{i_1, b}^I(t) 2^{-b} \rfloor \end{array} \right) \quad (115)$$

$$+ \sum_{k=2}^m \left( \begin{array}{c} \lfloor \text{sign}(\bar{X}_{i_k}^R(t)) h_{i_k i_k}^R \sum_{b=1}^{m_{i_k^* i_k}^R} \bar{X}_{i_k, b}^R(t) 2^{-b} \rfloor - \lfloor \text{sign}(\bar{X}_{i_k}^I(t)) h_{i_k i_k}^I \sum_{b=1}^{m_{i_k^* i_k}^I} \bar{X}_{i_k, b}^I(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{i_k}^R(t)) h_{i_k i_k}^I \sum_{b=1}^{m_{i_k^* i_k}^I} \bar{X}_{i_k, b}^R(t) 2^{-b} \rfloor + \lfloor \text{sign}(\bar{X}_{i_k}^I(t)) h_{i_k i_k}^R \sum_{b=1}^{m_{i_k^* i_k}^R} \bar{X}_{i_k, b}^I(t) 2^{-b} \rfloor \end{array} \right) \quad (116)$$

$$\hat{Y}_{i_1, l}(t) = \left( \begin{array}{c} \lfloor \text{sign}(\bar{X}_{i_1}^R(t)) h_{i_1 i_1}^R \sum_{b=m_{i_1^* i_1}^R+1}^{m_{i_1^* i_1}^R} \bar{X}_{i_1, b}^R(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{i_1}^I(t)) h_{i_1 i_1}^R \sum_{b=m_{i_1^* i_1}^R+1}^{m_{i_1^* i_1}^R} \bar{X}_{i_1, b}^I(t) 2^{-b} \rfloor \end{array} \right) \quad (119)$$

$$+ \sum_{k=2}^m \left( \begin{array}{c} \lfloor \text{sign}(\bar{X}_{i_k}^R(t)) h_{i_k i_k}^R \sum_{b=1}^{m_{i_k^* i_k}^R} \bar{X}_{i_k, b}^R(t) 2^{-b} \rfloor - \lfloor \text{sign}(\bar{X}_{i_k}^I(t)) h_{i_k i_k}^I \sum_{b=1}^{m_{i_k^* i_k}^I} \bar{X}_{i_k, b}^I(t) 2^{-b} \rfloor \\ \lfloor \text{sign}(\bar{X}_{i_k}^R(t)) h_{i_k i_k}^I \sum_{b=1}^{m_{i_k^* i_k}^I} \bar{X}_{i_k, b}^R(t) 2^{-b} \rfloor + \lfloor \text{sign}(\bar{X}_{i_k}^I(t)) h_{i_k i_k}^R \sum_{b=1}^{m_{i_k^* i_k}^R} \bar{X}_{i_k, b}^I(t) 2^{-b} \rfloor \end{array} \right)$$

where the last inequality follows from the fact that dropping conditioning does not reduce entropy.

Now consider the last term in (124),

$$\begin{aligned} H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n | \mathcal{W}_{i_2 i_2}^c) &= H(\hat{Y}_{i_1, u}^n | \mathcal{W}_{i_2 i_2}^c) + H(\hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n | \hat{Y}_{i_1, u}^n, \mathcal{W}_{i_2 i_2}^c) \quad (125) \\ &\geq H(\hat{Y}_{i_1, u}^n | \mathcal{W}_{i_2 i_2}^c) + H(\hat{S}_{i_2}^n | \hat{Y}_{i_1, u}^n, \mathcal{W}_{i_2 i_2}^c) \quad (126) \\ &= H(\bar{X}_{i_1^* i_1, S}^R | \bar{X}_{i_1^* i_1, S}^I | \mathcal{W}_{i_2 i_2}^c) + H(\hat{S}_{i_2}^n | \hat{Y}_{i_1, u}^n, \mathcal{W}_{i_2 i_2}^c) \quad (127) \\ &\geq H(\hat{Y}_{i_1^* i_1}^n | \tilde{\mathcal{W}} \setminus \mathcal{W}_{i_1}^c) + H(\hat{S}_{i_2}^n | W_{i_1 i_2}, W_{i_3 i_2}, \dots, W_{i_m i_2}) \quad (128) \end{aligned}$$

where (127) follows Lemma 2, i.e., both functions  $f: \hat{Y}_{i_1, u}^n \rightarrow \bar{X}_{i_1^* i_1, S}^R \times \bar{X}_{i_1^* i_1, S}^I$  and  $f: \hat{S}_{i_2}^n \rightarrow \hat{S}_{i_2}$  are bijective.

Plugging (128) into (124), we have

$$\begin{aligned} n \left( \sum_{j=1}^m R_{i_1 i_j} - \epsilon \right) &\leq H(\hat{S}_{i_1}^n | W_{i_2 i_1}, W_{i_3 i_1}, \dots, W_{i_m i_1}) \\ &\quad + H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n | \hat{S}_{i_1}^n) - H(\hat{S}_{i_2}^n | W_{i_1 i_2}, W_{i_3 i_2}, \dots, W_{i_m i_2}) \\ &\quad - H(\hat{Y}_{i_1^* i_1}^n | \tilde{\mathcal{W}} \setminus \mathcal{W}_{i_1}^c) \quad (129) \end{aligned}$$

Then consider the degraded BC comprised of the transmitter  $i_1$  and the virtual receivers  $\{R'_{i_{m+1} i_1}, R'_{i_{m+2} i_1}, \dots, R'_{i_N i_1}\}$ . Since  $R'_{i_{m+1} i_1}$  is the strongest receiver which can decode all the messages from the transmitter  $i_1$  to its connected virtual receivers, we have

$$\begin{aligned} n \left( \sum_{j=m+1}^N R_{i_1 i_j} - \epsilon \right) &\leq I(\mathcal{W}_{i_1}^c; \hat{Y}_{i_1^* i_1}^n | \tilde{\mathcal{W}} \setminus \mathcal{W}_{i_1}^c) \quad (130) \\ &= H(\hat{Y}_{i_1^* i_1}^n | \tilde{\mathcal{W}} \setminus \mathcal{W}_{i_1}^c) \quad (131) \end{aligned}$$

Adding (129) and (131), we have

$$\begin{aligned} n \left( \sum_{j=1}^m R_{i_1 i_j} + \sum_{j=m+1}^N R_{i_1 i_j} - 2\epsilon \right) &\leq H(\hat{S}_{i_1}^n | W_{i_2 i_1}, W_{i_3 i_1}, \dots, W_{i_m i_1}) + H(\hat{Y}_{i_1, u}^n, \hat{Y}_{i_1, l}^n, \hat{C}_{i_1}^n | \hat{S}_{i_1}^n) \\ &\quad - H(\hat{S}_{i_2}^n | W_{i_1 i_2}, W_{i_3 i_2}, \dots, W_{i_m i_2}) \quad (132) \end{aligned}$$

Similarly, for  $k \in \{2, 3, \dots, m-1\}$  we obtain

$$\begin{aligned} n \left( \sum_{j=1}^m R_{i_k i_j} + \sum_{j=m+1}^N R_{i_k i_j} - 2\epsilon \right) &\leq H(\hat{S}_{i_k}^n | W_{i_1 i_k}, \dots, W_{i_{k-1} i_k}, W_{i_{k+1} i_k}, \dots, W_{i_m i_k}) \\ &\quad + H(\hat{Y}_{i_k, u}^n, \hat{Y}_{i_k, l}^n, \hat{C}_{i_k}^n | \hat{S}_{i_k}^n) \\ &\quad - H(\hat{S}_{i_{k+1}}^n | W_{i_1 i_{k+1}}, \dots, W_{i_k i_{k+1}}, W_{i_{k+2} i_{k+1}}, \dots, W_{i_m i_{k+1}}) \quad (133) \end{aligned}$$

And from the receiver  $i_m$  and the degraded BC comprised of the transmitter  $i_m$  and all its connected virtual receivers, we have

$$\begin{aligned} n \left( \sum_{j=1}^m R_{i_m i_j} + \sum_{j=m+1}^N R_{i_m i_j} - 2\epsilon \right) &\leq H(\hat{S}_{i_m}^n | W_{i_1 i_m}, W_{i_2 i_m}, \dots, W_{i_{m-1} i_m}) \\ &\quad + H(\hat{Y}_{i_m, u}^n, \hat{Y}_{i_m, l}^n, \hat{C}_{i_m}^n | \hat{S}_{i_m}^n) \\ &\quad - H(\hat{S}_{i_1}^n | W_{i_2 i_1}, W_{i_3 i_1}, \dots, W_{i_m i_1}) \quad (134) \end{aligned}$$

Adding all the terms in (132), (133) and (134) together and applying the same argument used in Example 3, we have

$$n(R_{\Sigma, D} - 2m\epsilon) \leq \sum_{k=1}^m H(\hat{Y}_{i_k, u}^n, \hat{Y}_{i_k, l}^n, \hat{C}_{i_k}^n | \hat{S}_{i_k}^n) \quad (135)$$

$$\leq \sum_{k=1}^m \sum_{t=1}^n \left[ H(\hat{Y}_{i_k, u}(t) | \hat{S}_{i_k}(t)) + H(\hat{Y}_{i_k, l}(t) | \hat{S}_{i_k}(t)) + H(\hat{C}_{i_k}(t)) \right] \quad (136)$$

$$\leq n \sum_{k=1}^m \left[ 2(m_{i_k i_k} - m_{i_{k-1} i_k}) + \text{constant} \right] \quad (137)$$

$$\leq n \sum_{k=1}^m \left[ (a_{i_k i_k} - a_{i_{k-1} i_k}) \log_2 P + \text{constant} \right] \quad (138)$$

where (137) holds since  $m_{i_k i_j} - 1 \leq \max\{m_{i_k i_j}^R, m_{i_k i_j}^I\} \leq m_{i_k i_j}$ , and also note the modulo- $m$  arithmetic is implicitly used on user indices, e.g.,  $i_0 = i_m$ .

Finally according to Lemma 1, we can obtain the desired GDoF cycle bound for the original complex Gaussian  $X$  channel through some simple manipulations,

$$\sum_{j=1}^m \bar{d}_{i_j} \leq \sum_{j=1}^m (a_{i_j i_j} - a_{i_{j-1} i_j}) \quad (139)$$

and hence complete the whole proof.

## V. CONCLUSION

In this paper, we extend the optimality of TIN to more general classes of message sets. The main result is that for the TIN-optimal  $K$ -user interference channel, even if the message set expands to include the  $X$  setting where each transmitter has one independent message to each receiver, operating the new channel as the original interference channel and treating interference as noise at each receiver is still optimal to achieve the sum channel capacity to within a constant gap. Furthermore, the optimality of TIN for the general  $M \times N$   $X$  channel is also demonstrated.

We conclude with a comment on the necessity of the optimality conditions. In [12] it is conjectured that (8) is also necessary for TIN to be optimal for the *entire GDoF region* except for a set of channel gain values with measure zero. However, note that no claim is made for the necessity of condition (8) for the optimality of TIN for *sum-GDoF*. In fact it is easy to see that (8) is not necessary for the sum-GDoF optimality of TIN. For example, consider the 2-user interference channel with  $\alpha_{11} > \alpha_{12} + \alpha_{21} > \alpha_{22}$ , which violates (8), and whose optimal sum-GDoF value,  $\alpha_{11}$  (as shown in [7]), is trivially achieved by activating only user 1. It also should be noted that even in the  $2 \times 2$   $X$  channel, when  $\alpha_{11} > \alpha_{12} + \alpha_{21} > \alpha_{22}$ , using the bounding techniques presented in this paper it is not hard to prove that the optimal sum-GDoF value is still  $\alpha_{11}$ . Similarly, since our focus is only on sum-GDoF, the optimality conditions are only sufficient, but not necessary.

## APPENDIX A PROOF OF LEMMA 1

We begin with the general complex Gaussian  $M \times N$   $X$  channel, and convert it to the corresponding truncated binary-expansion deterministic channel step-by-step. In each step, we show that only a loss of constant bits is introduced. Here we follow the similar steps used in [20]. For the sake of simplicity, we define  $\mathcal{W}_i^* \triangleq \{W_{i1}, W_{i2}, \dots, W_{iM}\}$ , and we suppress the time index  $t$  if no confusion would be caused.

- **Step 1: Average power constraint to peak power constraint.** Recall that in the original complex Gaussian channels, by scaling the output, we set

$$E[|X_i(t)|^2] \leq 2, \quad Z_i(t) \sim \mathcal{CN}(0, 2)$$

Then for each input  $X_i = X_i^R + jX_i^I$ , we truncate both the real and imaginary parts to satisfy the peak power constraint of 1. Define the part of input  $X_i^R$  that exceeds the peak power constraint as

$$\tilde{X}_i^R = \lfloor X_i^R \rfloor = \text{sign}(X_i^R) \sum_{b=-\infty}^0 X_{i,b}^R 2^{-b}$$

and the remaining signal as

$$\bar{X}_i^R = X_i^R - \tilde{X}_i^R = \text{sign}(X_i^R) \sum_{b=1}^{\infty} X_{i,b}^R 2^{-b}$$

For the imaginary part of the input, we have the similar definitions for  $X_i^I$  with  $I$  replacing  $R$ . Then  $\bar{X}_i^R$  and  $\bar{X}_i^I$  satisfy the unit peak power constraint. Also define

$$\bar{X}_i = \bar{X}_i^R + j\bar{X}_i^I \quad (140)$$

$$\tilde{X}_i = \tilde{X}_i^R + j\tilde{X}_i^I \quad (141)$$

Letting  $\bar{Y}_i$  be the output of receiver  $i$  due to the truncated input  $\bar{X}_i$ , and  $\tilde{Y}_i$  be the difference between  $Y_i$  and  $\bar{Y}_i$ , for each receiver  $i \in [N]$  we have

$$I(\mathcal{W}_i^*; Y_i^n) \leq I(\mathcal{W}_i^*; \bar{Y}_i^n, \tilde{Y}_i^n) \quad (142)$$

$$= I(\mathcal{W}_i^*; \bar{Y}_i^n) + I(\mathcal{W}_i^*; \tilde{Y}_i^n | \bar{Y}_i^n) \quad (143)$$

$$\leq I(\mathcal{W}_i^*; \bar{Y}_i^n) + H(\tilde{Y}_i^n) \quad (144)$$

$$\leq I(\mathcal{W}_i^*; \bar{Y}_i^n) + \sum_{k=1}^M H(\tilde{X}_k^n) \quad (145)$$

$$\leq I(\mathcal{W}_i^*; \bar{Y}_i^n) + n \times \text{constant} \quad (146)$$

The last inequality follows from [20, Lemma 6], which says that for sufficiently large  $n$ , the entropy of  $\tilde{X}_k^n$  (the part of input  $X_k^n$  exceeding the unit peak power constraint) is no larger than a constant number multiplied by  $n$ .

- **Step 2: Truncate signals at noise level and remove noise.** Recall  $\lfloor \log_2 |h_{ik}^R| \rfloor = m_{ik}^R$  and  $\lfloor \log_2 |h_{ik}^I| \rfloor = m_{ik}^I$ . We have (147) at the top of next page. Next, define  $\varepsilon_i$  as the difference between  $\bar{Y}_i$  and  $\tilde{Y}_i$  as shown in (148) in next page, where  $\text{frac}(x)$  denotes the fractional part of  $x$ . Also note

$$|h_{ik}^R \sum_{b=m_{ik}^R+1}^{\infty} X_{k,b}^R 2^{-b}| \leq 2^{m_{ik}^R+1} 2^{-(m_{ik}^R)} = 2$$

$$\hat{Y}_i = \sum_{k=1}^M \left[ (\lfloor \text{sign}(X_k^R) h_{ik}^R \sum_{b=1}^{m_{ik}^R} X_{k,b}^R 2^{-b} \rfloor - \lfloor \text{sign}(X_k^I) h_{ik}^I \sum_{b=1}^{m_{ik}^I} X_{k,b}^I 2^{-b} \rfloor) + j(\lfloor \text{sign}(X_k^R) h_{ik}^I \sum_{b=1}^{m_{ik}^I} X_{k,b}^R 2^{-b} \rfloor + \lfloor \text{sign}(X_k^I) h_{ik}^R \sum_{b=1}^{m_{ik}^R} X_{k,b}^I 2^{-b} \rfloor) \right] \quad (147)$$

$$\begin{aligned} \varepsilon_i &= \bar{Y}_i - \hat{Y}_i \\ &= \sum_{k=1}^M \left\{ \left[ \text{sign}(X_k^R) h_{ik}^R \sum_{b=m_{ik}^R+1}^{\infty} X_{k,b}^R 2^{-b} - \text{sign}(X_k^I) h_{ik}^I \sum_{b=m_{ik}^I+1}^{\infty} X_{k,b}^I 2^{-b} \right. \right. \\ &\quad \left. \left. + \text{frac}(\text{sign}(X_k^R) h_{ik}^R \sum_{b=1}^{m_{ik}^R} X_{k,b}^R 2^{-b}) - \text{frac}(\text{sign}(X_k^I) h_{ik}^I \sum_{b=1}^{m_{ik}^I} X_{k,b}^I 2^{-b}) \right] \right. \\ &\quad \left. + j \left[ \text{sign}(X_k^R) h_{ik}^I \sum_{b=m_{ik}^I+1}^{\infty} X_{k,b}^R 2^{-b} + \text{sign}(X_k^I) h_{ik}^R \sum_{b=m_{ik}^R+1}^{\infty} X_{k,b}^I 2^{-b} \right. \right. \\ &\quad \left. \left. + \text{frac}(\text{sign}(X_k^R) h_{ik}^I \sum_{b=1}^{m_{ik}^I} X_{k,b}^R 2^{-b}) + \text{frac}(\text{sign}(X_k^I) h_{ik}^R \sum_{b=1}^{m_{ik}^R} X_{k,b}^I 2^{-b}) \right] \right\} + Z_i \\ &= \sum_{k=1}^M \hat{X}_k + Z_i \end{aligned} \quad (148)$$

Similarly, we have

$$\begin{aligned} |h_{ik}^I \sum_{b=m_{ik}^I+1}^{\infty} X_{k,b}^I 2^{-b}| &\leq 2 \\ |h_{ik}^I \sum_{b=m_{ik}^I+1}^{\infty} X_{k,b}^R 2^{-b}| &\leq 2 \\ |h_{ik}^R \sum_{b=m_{ik}^R+1}^{\infty} X_{k,b}^I 2^{-b}| &\leq 2 \end{aligned}$$

Finally, we obtain

$$\begin{aligned} I(\mathcal{W}_i^*; \bar{Y}_i^n) &\leq I(\mathcal{W}_i^*; \hat{Y}_i^n, \varepsilon_i^n) \quad (149) \\ &= I(\mathcal{W}_i^*; \hat{Y}_i^n) + I(\mathcal{W}_i^*; \varepsilon_i^n | \hat{Y}_i^n) \quad (150) \\ &= I(\mathcal{W}_i^*; \hat{Y}_i^n) + h(\varepsilon_i^n | \hat{Y}_i^n) - h(\varepsilon_i^n | \hat{Y}_i^n, \mathcal{W}_i^*) \quad (151) \\ &\leq I(\mathcal{W}_i^*; \hat{Y}_i^n) + h(\varepsilon_i^n) - h(Z_i^n) \quad (152) \\ &= I(\mathcal{W}_i^*; \hat{Y}_i^n) + I(\hat{X}_1^n, \hat{X}_2^n, \dots, \hat{X}_M^n; \varepsilon_i^n) \quad (153) \\ &\leq I(\mathcal{W}_i^*; \hat{Y}_i^n) + n \times \text{constant} \quad (154) \end{aligned}$$

where the last inequality is due to the fact that  $\hat{X}_1, \hat{X}_2, \dots, \hat{X}_M \mapsto \varepsilon_i$  forms a complex Gaussian MAC with a finite SNR independent of  $P$  for each transmitter [20].  $\square$

## APPENDIX B PROOF OF LEMMA 2

For notation brevity, we define

$$\check{X}_R \triangleq \text{sign}(\bar{X}_{ij}^R(t)) \sum_{b=1}^{\max\{m_{ij-1j}^R, m_{ij-1j}^I\}} \bar{X}_{ij,b}^R(t) 2^{-b} \quad (155)$$

$$\check{X}_I \triangleq \text{sign}(\bar{X}_{ij}^I(t)) \sum_{b=1}^{\max\{m_{ij-1j}^R, m_{ij-1j}^I\}} \bar{X}_{ij,b}^I(t) 2^{-b} \quad (156)$$

$$\begin{aligned} \check{S}_R &\triangleq \underbrace{\lfloor \text{sign}(\bar{X}_{ij}^R(t)) h_{ij-1j}^R \sum_{b=1}^{m_{ij-1j}^R} \bar{X}_{ij,b}^R(t) 2^{-b} \rfloor}_{\check{S}_{R,1}} \\ &\quad - \underbrace{\lfloor \text{sign}(\bar{X}_{ij}^I(t)) h_{ij-1j}^I \sum_{b=1}^{m_{ij-1j}^I} \bar{X}_{ij,b}^I(t) 2^{-b} \rfloor}_{\check{S}_{R,2}} \end{aligned} \quad (157)$$

$$\begin{aligned} \check{S}_I &\triangleq \underbrace{\lfloor \text{sign}(\bar{X}_{ij}^R(t)) h_{ij-1j}^I \sum_{b=1}^{m_{ij-1j}^R} \bar{X}_{ij,b}^R(t) 2^{-b} \rfloor}_{\check{S}_{I,1}} \\ &\quad + \underbrace{\lfloor \text{sign}(\bar{X}_{ij}^I(t)) h_{ij-1j}^R \sum_{b=1}^{m_{ij-1j}^I} \bar{X}_{ij,b}^I(t) 2^{-b} \rfloor}_{\check{S}_{I,2}} \end{aligned} \quad (158)$$

Note  $(\check{X}_R, \check{X}_I)$  and  $(\check{S}_R, \check{S}_I)$  can be seen as the input and output of the deterministic channel, respectively. Obviously, given one input  $(\check{X}_R, \check{X}_I)$ , we can only produce one output  $(\check{S}_R, \check{S}_I)$ . Next, we prove the other direction by contradiction. We assume there exist two different inputs  $(\check{X}_R^*, \check{X}_I^*)$  and  $(\check{X}_R^{**}, \check{X}_I^{**})$  that can generate the same output, i.e.,  $(\check{S}_R^*, \check{S}_I^*) = (\check{S}_R^{**}, \check{S}_I^{**})$ . In the following, without loss of generality, we assume

$$|h_{ij-1j}^R| \geq |h_{ij-1j}^I| \Rightarrow m_{ij-1j}^R \geq m_{ij-1j}^I. \quad (159)$$

We first consider the case where  $\text{sign}(h_{ij-1ij}^R) = \text{sign}(h_{ij-1ij}^I)$ . For the term  $\check{S}_R$ , we have the following subcases:

- $\check{S}_{R,1} = \check{S}_{R,1}^{**}$  and  $\check{S}_{R,2} = \check{S}_{R,2}^{**}$ . In this case, for the term  $\check{S}_I$ , if  $|h_{ij-1ij}^R| > |h_{ij-1ij}^I|$ , since  $(\check{X}_R^*, \check{X}_I^*)$  and  $(\check{X}_R^{**}, \check{X}_I^{**})$  are different, we have  $\check{S}_{I,1}^* = \check{S}_{I,1}^{**}$  and  $\check{S}_{I,2}^* \neq \check{S}_{I,2}^{**}$ , which contradicts the assumption that  $(\check{X}_R^*, \check{X}_I^*)$  and  $(\check{X}_R^{**}, \check{X}_I^{**})$  generate the same output; if  $|h_{ij-1ij}^R| = |h_{ij-1ij}^I|$ , since  $(\check{X}_R^*, \check{X}_I^*)$  and  $(\check{X}_R^{**}, \check{X}_I^{**})$  generate the same  $(\check{S}_R, \check{S}_I)$ , we have  $(\check{X}_R^*, \check{X}_I^*) = (\check{X}_R^{**}, \check{X}_I^{**})$ , which contradicts the assumption that  $(\check{X}_R^*, \check{X}_I^*)$  and  $(\check{X}_R^{**}, \check{X}_I^{**})$  are different.
- $\check{S}_{R,1} > \check{S}_{R,1}^{**}$  and  $\check{S}_{R,2} > \check{S}_{R,2}^{**}$ . In this case, for the term  $\check{S}_I$ , we have  $\check{S}_{I,1}^* \geq \check{S}_{I,1}^{**}$  and  $\check{S}_{I,2}^* > \check{S}_{I,2}^{**}$ , which contradicts the assumption that  $(\check{X}_R^*, \check{X}_I^*)$  and  $(\check{X}_R^{**}, \check{X}_I^{**})$  generate the same output.
- $\check{S}_{R,1} < \check{S}_{R,1}^{**}$  and  $\check{S}_{R,2} < \check{S}_{R,2}^{**}$ . In this case, for the term  $\check{S}_I$ , we have  $\check{S}_{I,1}^* \leq \check{S}_{I,1}^{**}$  and  $\check{S}_{I,2}^* < \check{S}_{I,2}^{**}$ , which contradicts the assumption that  $(\check{X}_R^*, \check{X}_I^*)$  and  $(\check{X}_R^{**}, \check{X}_I^{**})$  generate the same output.

For the other case where  $\text{sign}(h_{ij-1ij}^R) = -\text{sign}(h_{ij-1ij}^I)$ , we can follow the above argument to get the same conclusion.  $\square$

#### ACKNOWLEDGMENT

The authors would like to thank the Associate Editor, and the reviewers for their valuable and timely comments.

#### REFERENCES

- [1] 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.
- [2] 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.
- [3] X. Shang, G. Kramer, and B. Chen, "A new outer bound and the noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.
- [4] V. S. Annapureddy and V. V. Veeravalli, "Sum capacity of MIMO interference channels in the low interference regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2565–2581, May 2011.
- [5] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Noisy-interference sum-rate capacity of parallel Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 210–226, Jan. 2011.
- [6] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Capacity regions and sum-rate capacities of vector Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 5030–5044, Oct. 2010.
- [7] 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.
- [8] S. A. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric Gaussian  $K$  user interference channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3297–3303, Jul. 2010.
- [9] T. Gou and S. A. Jafar, "Capacity of a class of symmetric SIMO Gaussian interference channels within  $\mathcal{O}(1)$ ," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1932–1958, Apr. 2011.
- [10] S. Karmakar and M. K. Varanasi, "The capacity region of the MIMO interference channel and its reciprocity to within a constant gap," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 4781–4797, Aug. 2013.
- [11] S. Karmakar and M. K. Varanasi, "The generalized degrees of freedom region of the MIMO interference channel and its achievability," *IEEE Trans. Inf. Theory*, vol. 58, no. 12, pp. 7188–7203, Dec. 2012.
- [12] 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.
- [13] S. A. Jafar and S. Shamai (Shitz), "Degrees of freedom region for the MIMO  $X$  channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.
- [14] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO  $X$  channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.
- [15] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom of wireless  $X$  networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3893–3908, Sep. 2009.
- [16] C. Huang, S. A. Jafar, and V. R. Cadambe, "Interference alignment and the generalized degrees of freedom of the  $X$  channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5130–5150, Aug. 2012.
- [17] U. Niesen and M. A. Maddah-Ali, "Interference alignment: From degrees of freedom to constant-gap capacity approximations," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 4855–4888, Aug. 2013.
- [18] A. A. El Gamal and M. H. M. Costa, "The capacity region of a class of deterministic interference channels," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 343–346, Mar. 1982.
- [19] 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.
- [20] G. Bresler and D. N. C. Tse, "The two-user Gaussian interference channel: A deterministic view," *Eur. Trans. Telecommun.*, vol. 19, no. 4, pp. 333–354, Jun. 2008.
- [21] M. Anand and P. R. Kumar, "A digital interface for Gaussian relay and interference networks: Lifting codes from the discrete superposition model," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2548–2564, May 2011.
- [22] I. Shomorony and A. S. Avestimehr, "On the role of deterministic models in  $K \times K \times K$  wireless networks," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sep. 2012, pp. 60–64.

**Chunhua Geng** (S'12) received the B.E. degree in Communications Engineering from Beijing Jiaotong University, Beijing, China, in 2007, and the M.S. degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2010. He is currently pursuing the Ph.D. degree in the Department of Electrical Engineering and Computer Science, University of California Irvine, Irvine, CA, USA. His research interests include network information theory, communication theory, information theoretic security, and wireless communications.

**Hua Sun** (S'12) received the B.E. degree in Communications Engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2011, the M.S. degree in Electrical and Computer Engineering from University of California Irvine, USA, in 2013. He is currently pursuing the Ph.D. degree at the University of California Irvine. His research interests include multiuser information theory and wireless communications.

**Syed Ali Jafar** (S'99–M'04–SM'09–F'14) received his B. Tech. from IIT Delhi, India, in 1997, M.S. from Caltech, USA, in 1999, and Ph.D. from Stanford, USA, in 2003, all in Electrical Engineering. His industry experience includes positions at Lucent Bell Labs, Qualcomm Inc. and Hughes Software Systems. He is a Professor in the Department of Electrical Engineering and Computer Science at the University of California Irvine, Irvine, CA USA. His research interests include multiuser information theory, wireless communications and network coding.

Dr. Jafar received the NSF CAREER award in 2006, the ONR Young Investigator Award in 2008, and the School of Engineering Maseeh Outstanding Research Award in 2010, his co-authored papers received Best Journal Paper Awards from the IEEE Information Theory Society in 2009, from the IEEE Communication Society in 2013, and from the IEEE Signal Processing Society in 2015, and Best Conference Paper Awards at IEEE GLOBECOM 2012 and at IEEE GLOBECOM 2014. Dr. Jafar received the UC Irvine EECS Professor of the Year award five times, in 2006, 2009, 2011, 2012, and 2014, from the Engineering Students Council and the Teaching Excellence Award in 2012 from the School of Engineering. He was a University of Canterbury Erskine Fellow in 2010 and an IEEE Communications Society Distinguished Lecturer for 2013–2014. Dr. Jafar was recognized as a Thomson Reuters Highly Cited Researcher and included by Sciencewatch among The World's Most Influential Scientific Minds in 2014. He served as Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS 2004–2009, for IEEE COMMUNICATIONS LETTERS 2008–2009 and for IEEE TRANSACTIONS ON INFORMATION THEORY 2009–2012. Dr. Jafar was elevated to IEEE Fellow, Class of 2014, for contributions to analyzing the capacity of wireless communication networks.