

# On the Optimality of Treating Interference as Noise for Parallel Deterministic Interference Networks

Hua Sun and Syed A. Jafar  
Center for Pervasive Communications and Computing (CPCC)  
University of California Irvine, Irvine, CA 92697

**Abstract**—It has been shown recently by Geng et al. that in a  $K$  user Gaussian interference network, 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 signal strengths measured in dB scale), then power control and treating interference as noise (TIN) is sufficient to achieve the entire generalized degrees of freedom (GDoF) region. Motivated by the intuition that the deterministic model of Avestimehr et al. (ADT deterministic model) is particularly suited for exploring the optimality of TIN, the results of Geng et al. are first re-visited under the ADT deterministic model, and corresponding TIN optimality results are obtained. Next, we focus on the extension of these results to ADT deterministic parallel interference networks, from a sum-capacity perspective. To this end, we interpret the explicit characterization of the sum-capacity of a TIN optimal network (without parallel channels) as a minimum weighted matching problem in combinatorial optimization, and obtain a simple characterization in terms of a partition of the interference network into vertex-disjoint cycles. Aided by insights from the cyclic partition, the sum-capacity optimality of TIN for  $K$  user parallel interference networks is characterized for the ADT deterministic model. Subject to a *mild* invertibility condition the optimality of TIN is shown to extend to parallel networks in a separable fashion.

## I. INTRODUCTION

Treating interference as noise (TIN) is a strategy that is universally applied in wireless networks to deal with interference from users that are far away. Interestingly, it is also known to be capacity optimal when the interference is sufficiently weak. Most relevant to this work is the recent result by Geng et al. in [1], where a broadly applicable condition is identified and shown to be sufficient (also conjectured to be necessary in almost all cases) for TIN to achieve the generalized degrees of freedom (GDoF) region.

The ADT deterministic model [2] captures much of the essence of the GDoF framework — the diversity of signal strengths — but is less useful when the finer details such as the channel phase or the distinction between rational and irrational realizations become critical. The main idea motivating this work is that the ADT deterministic model is well suited for studying those robust regimes where the finer aspects of channel realizations are not relevant. Given this insight, and since the regime where TIN is optimal is arguably the most robust regime, it follows that the ADT deterministic model should suffice to identify this regime in the GDoF sense and

to study its properties. As initial verification of this insight, we begin by exploring the TIN optimality result of Geng et al. in the ADT framework, where both the optimality conditions and the GDoF region are mapped easily and become more transparent. Encouraged by this insight, we proceed to explore the optimality of TIN for parallel interference networks.

Optimality of TIN for parallel Gaussian interference networks is an intriguing question for the following reasons. On the one hand, with the exception of the MAC-Z-BC network (which contains the multiple access channel, Z-channel and broadcast channel as special cases), it is known that all parallel Gaussian networks are in general inseparable [3], [4], [5]. On the other hand, for the 2 user interference network, extensions to parallel channels have been made from an exact *sum-capacity* perspective in [6] and from a GDoF perspective in [7]. In both cases, the results support separability of TIN optimal sub-channels. However, the insights from the 2 user setting do not directly extend to the  $K$  user setting. For example, the GDoF region for the TIN optimal 2 user interference network is polymatroidal, whereas the GDoF region of TIN optimal  $K$  user interference networks, with  $K \geq 3$ , is no longer polymatroidal. The distinction is particularly significant for parallel channels. Given the significant challenges in going beyond 2 users, it is most intriguing if the separability of parallel Gaussian interference networks will hold in the regime where TIN is sum-GDoF optimal. In this work, we explore the answer of this question under the ADT deterministic model, from the sum-capacity perspective.

The focus on sum-capacity motivates us to first seek a more explicit characterization. To this end, we show that the sum-capacity characterization for a  $K$  user interference network is essentially a minimum weighted matching problem in combinatorial optimization. Consequently, the sum-capacity is characterized in terms of a partition of the interference network into disjoint cycles. Aided by the insights from the cyclic partition approach, a separable extension of the optimality of TIN to ADT deterministic parallel interference networks is obtained subject to a mild invertibility condition. While we focus here only on the ADT deterministic model, extensions to the Gaussian setting can be found in the full paper [8].

## II. SYSTEM MODEL, DEFINITIONS, AND NOTATION

Consider the  $K$  user ADT deterministic interference network, with  $M$  parallel sub-channels. Over the  $t$ -th channel

use, the signal sent from Transmitter  $i$ ,  $X_i^{[m]}(t)$ , as observed at Receiver  $k$ ,  $Y_k^{[m]}(t)$ , over the  $m$ -th sub-channel, is scaled up by a nonnegative integer value  $n_{ki}^{[m]}$ , called the signal strength level. The channel may be written as

$$Y_k^{[m]}(t) = [2^{n_{k1}^{[m]}} X_1^{[m]}(t)] \oplus \dots \oplus [2^{n_{kK}^{[m]}} X_K^{[m]}(t)],$$

$$\forall k \in [K] \triangleq \{1, 2, \dots, K\}, m \in [M] \triangleq \{1, 2, \dots, M\}$$

where addition is performed on each bit (modulo two). We assume the real-valued channel input is positive and has peak power constraint 1, then it can be written in base 2 as  $X_i^{[m]} = 0.X_{i,(1)}^{[m]}X_{i,(2)}^{[m]}X_{i,(3)}^{[m]} \dots$  where the channel index is omitted for compactness.  $\mathbf{X}_i(t) = [X_i^{[1]}(t), \dots, X_i^{[M]}(t)]$ ,  $\mathbf{Y}_k(t) = [Y_k^{[1]}(t), \dots, Y_k^{[M]}(t)]$  are the vectors containing the transmitted symbols from Transmitter  $i$  and received signals observed at Receiver  $k$  over all sub-channels, respectively. All channels are fixed and known at all transmitters and receivers. At Transmitter  $i$ , an independent message  $W_i$  uniformly distributed over the message index set  $\{1, 2, \dots, [2^{nR_i}]\}$  is mapped to the transmitted codeword  $[\mathbf{X}_i(1), \mathbf{X}_i(2), \dots, \mathbf{X}_i(n)]$  (abbreviated as  $\mathbf{X}_i^n$ ) over  $n$  channel uses. At Receiver  $k$ , the received signal  $[\mathbf{Y}_k(1), \mathbf{Y}_k(2), \dots, \mathbf{Y}_k(n)]$  (abbreviated as  $\mathbf{Y}_k^n$ ) is used to produce the estimate  $\hat{W}_k$  of the message  $W_k$ . The probability of error for Receiver  $k$  is given by the probability that  $\hat{W}_k$  is not equal to  $W_k$ . A rate tuple  $(R_1, R_2, \dots, R_K)$  is said to be achievable if we have an encoding and decoding mapping such that the probability of error for each receiver approaches zero as  $n$  approaches infinity. The capacity region  $\mathcal{C}$  is the closure of the set of all achievable rate tuples. The sum-capacity is defined as  $\mathcal{C}_\Sigma = \max_{\mathcal{C}} \sum_{k=1}^K R_k$ .

We introduce a directed graph representation that will be useful to efficiently present the results in this work. The directed graph representation of the  $K$  user interference network consists of  $K$  vertices,  $V_1, V_2, \dots, V_K$ , one for each user. Since the vertices correspond directly to users, we will also refer to them as users. For all  $(i, j) \in [K] \times [K]$ , there is a directed edge  $e_{ij}$  from user  $j$  to user  $i$ , with weight  $w(e_{ij})$  defined as  $w(e_{ij}) = n_{ij}$ , if  $i \neq j$  and  $w(e_{ij}) = 0$ , if  $i = j$ .

We are particularly interested in the notion of cycles on this directed graph. We define a cycle,  $\pi$ , as a cyclically ordered subset of users, without repetitions. The set of all cycles is denoted as  $[\Pi]$ . The cardinality of a cycle, denoted as  $|\pi|$  is the number of users that it involves,  $|\pi| = \sum_{V_k \in \pi} 1, \forall \pi \in [\Pi]$ . A cycle with only one user is a trivial cycle. Two cycles  $\pi_p, \pi_q$ , are said to be disjoint if they contain no common user, denoted as  $\pi_p \cap \pi_q = \phi$ .

Introducing a slight abuse of notation in the interest of conciseness, the same cycle,  $\pi$ , can also be equivalently represented as a set of edges representing a closed path where no user is visited more than once. The weight of a cycle, denoted as  $w(\pi)$ , is the sum of the weights of all the edges traversed in completing the cycle,  $w(\pi) = \sum_{e_{ij} \in \pi} w(e_{ij}), \forall \pi \in [\Pi]$ . Note that the weight of a trivial cycle is zero. Intuitively, the weight of a cycle is the accumulation of the strengths of interference terms encountered in the cycle.

**Cyclic Partition:** A subset of the set of all cycles,  $\Pi \subset [\Pi]$ , is said to be a cyclic partition if  $\pi_p \cap \pi_q = \phi, \forall \pi_p, \pi_q \in \Pi$  and  $\sum_{\pi \in \Pi} |\pi| = K$ . In other words, a cyclic partition is a disjoint cyclic cover of the  $K$  users.

**Cyclic Partition Bound:** For any cyclic partition  $\Pi$ , define the corresponding cyclic partition bound,  $\mathcal{C}_\Sigma^\Pi$ , as  $\sum_{k=1}^K R_k \leq \sum_{k=1}^K n_{kk} - w(\Pi)$  where  $w(\Pi) = \sum_{\pi \in \Pi} w(\pi)$  is the net weight of the cyclic partition, representing the total interference encountered in this partition. Since there are many cyclic partitions, each of which gives rise to a cyclic partition bound, let us denote the tightest of these bounds as the *best cyclic partition bound*,  $\mathcal{C}_\Sigma^{\Pi^*}$ . A cyclic partition that produces the best cyclic partition bound is labeled an *optimal* cyclic partition, and denoted by  $\Pi^*$ . For example, when  $K = 6$ , one possible cyclic partition is  $\Pi = \{\{1, 3, 5\}, \{4, 2\}, \{6\}\}$  which decomposes the users into three cycles, such that each user is represented in exactly one cycle. The corresponding cyclic partition bound is  $\sum_{k=1}^6 R_k \leq \sum_{k=1}^6 n_{kk} - (n_{13} + n_{35} + n_{51}) - (n_{42} + n_{24}) - (0)$ .

**Participating Edge:** Edge  $e_{ij}$  is a participating edge for the cyclic partition  $\Pi$  if  $i \neq j$  and  $e_{ij} \in \pi$  for some  $\pi \in \Pi$ .

**Cyclic Predecessor:** Under cyclic partition  $\Pi$ , the cyclic predecessor for user  $k$  is user  $\Pi(k)$ , if  $e_{\Pi(k)k}$  is a participating edge for  $\Pi$ . Note that if user  $k$  belongs to a trivial cycle in  $\Pi$  then  $\Pi(k) = \phi$ .

### III. OPTIMALITY OF TIN THROUGH THE ADT MODEL

Following Geng et al.'s result [1] on the optimality of TIN for the  $K$  user Gaussian interference network with one sub-channel, it is not difficult to obtain a corresponding TIN optimality result for the ADT deterministic model. The sub-channel index is omitted in this section for compactness.

**Theorem 1:** In a  $K$  user ADT deterministic interference network, where the channel strength level from Transmitter  $i$  to Receiver  $j$  is equal to  $n_{ji}$ ,  $\forall i, j \in [K]$ , if the following condition is satisfied

$$n_{ii} \geq \max_{j: j \neq i} \{n_{ji}\} + \max_{k: k \neq i} \{n_{ik}\}, \quad \forall i, j, k \in [K], \quad (1)$$

then power control and treating interference as noise can achieve the whole capacity region. Moreover, the capacity region is given by

$$\mathcal{C}_{\text{TIN}} = \left\{ (R_1, R_2, \dots, R_K) \in \mathbb{R}_+^K : \sum_{V_k \in \pi} R_k \leq \sum_{V_k \in \pi} n_{kk} - w(\pi), \quad \forall \pi \in [\Pi] \right\}$$

The proof details for Theorem 1 are omitted because they parallel those for Theorem 1 in [1]. Here we only provide some interesting observations. Each of the bounds defining the capacity region represents the sum-capacity of a cyclic interference sub-network contained in the  $K$  user fully connected interference network. A cyclic sub-network is comprised of a cyclically ordered subset of users where each user causes interference only to the preceding user and suffers interference only from the following user in the cycle. The sum-capacity of a cyclic interference sub-network is simply the sum of all

desired link strengths minus the sum of all cross link strengths. Note that there exists many such sum-capacity bounds, as there are many distinct cycles for all subsets of users. Out of these cycle bounds, all but the tightest are redundant. Nevertheless, when considered together, the cycle bounds describe the precise capacity region of the fully connected network whenever condition (1) is satisfied. This remarkable aspect greatly simplifies the proof of the outer bound, because only cyclic interference networks need to be considered.

*Remark: Henceforth, we refer to (1) as the TIN optimality condition. A network (sub-channel) is called TIN optimal if the TIN optimality condition (1) is satisfied over the network (sub-channel).*

We now switch our attention from capacity region to sum-capacity, where an explicit characterization is obtained. The problem is highly non-trivial because the capacity region is not polymatroidal. Toward this end, we use a minimum weighted matching problem to obtain a simple characterization of the sum-capacity in terms of a partition of the interference network into vertex-disjoint cycles. The main result is that the sum-capacity is always given by a cyclic partition bound. This is stated in the following theorem.

**Theorem 2:** For TIN optimal ADT deterministic interference networks,  $\mathcal{C}_\Sigma = \mathcal{C}_\Sigma^{\Pi^*}$ .

Due to space limitation, for a detailed proof, we refer the readers to our full paper available online [8]. The importance of Theorem 2 is that it gives an explicit analytical formula of the sum-capacity achieved by TIN, which we shall use for parallel networks.

#### IV. OPTIMALITY OF TIN FOR PARALLEL NETWORKS

As we move from the single sub-channel case to multiple parallel sub-channels, the outer bound proof becomes significantly more challenging. Whereas formerly it was sufficient to only consider each cyclic sub-network obtained by eliminating all other users, messages and links, this is no longer possible for parallel interference networks. For example, a different cycle may be active in each sub-channel, however one cannot eliminate a different set of links for each sub-channel. As an outer bounding argument, eliminating a link is justified by including a genie that takes all the messages originating at the transmitter of that link, and provides them to the receiver of that link, so that the receiver can reconstruct and subtract the transmitted symbols from its received signal. However, in a parallel channels setting, the message information provided by the genie allows a receiver to reconstruct and subtract the transmitted symbols from a transmitter on *all* sub-channels. Thus, if a link from Transmitter  $i$  to Receiver  $j$  is removed for one sub-channel, it must be removed for all sub-channels. This makes it impossible to reduce a fully connected parallel interference network directly into *different* cyclic sub-networks over each sub-channel. As such, for parallel interference networks, the reduction to cyclic networks is in general no longer an option, and the entire network must be directly considered for the outer bound. After exposing this added source of difficulty, we proceed to the parallel networks.

#### A. ADT Deterministic Model

While we deal with multiple parallel sub-channels in this section, it is worth highlighting that we assume throughout each sub-channel is TIN optimal by itself. What we wish to explore is whether *collectively* such parallel channels remain separable and therefore TIN optimal. Let us start with a few relevant definitions.

For the definitions that have been introduced, we will add a superscript to indicate the sub-channel index, for example cyclic partition  $\Pi^{[m]}$ , cyclic predecessor  $\Pi^{[m]}(k)$ , and cyclic partition bound  $\mathcal{C}_\Sigma^{\Pi^{[m]}}$ . Note that many cyclic partitions are possible for each sub-channel, and a different cyclic partition may be used for each sub-channel.

**Participating Input and Output Levels ( $X_{i,u}^{[m]}, Y_{k,u}^{[m]}$ ):** For the  $m$ -th sub-channel, we define participating input levels  $X_{i,u}^{[m]} \triangleq 0.X_{i,(1)}^{[m]}, \dots, X_{i, \left( n_{\Pi^{[m]}(i)}^{[m]} \right)}^{[m]}$  to be the bits that are sent from Transmitter  $i$  and observed at its predecessor Receiver  $\Pi^{[m]}(i)$ . The received signal levels resulting from all interfering  $X_{i,u}^{[m]}$  are defined as the participating output levels  $Y_{k,u}^{[m]} \triangleq \sum_{i=1, i \neq k}^K 2^{n_{ki}^{[m]}} X_{i,u}^{[m]}$  where the summation is bit-wise modulo two. We can also write  $X_{i,u}$  in a vector form as  $X_{i,u}^{[m]} = [X_{i,(1)}^{[m]}, \dots, X_{i, \left( n_{\Pi^{[m]}(i)}^{[m]} \right)}^{[m]}]$ . Similar vector notation

is used for  $Y_{k,u}^{[m]}$  when the vector form is clearer.

**Invertibility:** The  $m$ -th sub-channel is said to be invertible if the mapping from  $\mathbf{X}_u^{[m]} \triangleq (X_{1,u}^{[m]}, \dots, X_{K,u}^{[m]})$  to  $\mathbf{Y}_u^{[m]} \triangleq (Y_{1,u}^{[m]}, \dots, Y_{K,u}^{[m]})$  is invertible for an optimal cyclic partition  $\Pi^{[m]*}$ . Mathematically, we require  $H(\mathbf{X}_u^{[m]} | \mathbf{Y}_u^{[m]}) = 0$ .

The significance of these definitions will become clear with the statement of the result, illustrative examples, and finally from the details of the proof. Perhaps the most intriguing is invertibility. At this point it suffices to say that it is a “mild” property and is easily testable for a given problem instance. The mildness of this property will be explicitly addressed in Section IV-B. With these definitions, we are now ready to state the main result of this section in the following theorem.

**Theorem 3:** In a  $K$  user ADT deterministic interference network with  $M$  sub-channels, if each sub-channel is individually TIN optimal and invertible, then even collectively for all the sub-channels of the parallel interference network, the sum-capacity is achieved by a separate TIN solution over each sub-channel.

The proof of Theorem 3 is deferred to Section IV-C. At this point it is important to understand the statement of the theorem and its limitations through illustrative examples.

**Example 1:** Consider the  $K = 3$  user ADT deterministic interference network with  $M = 3$  parallel sub-channels, shown in Figure 1(a). It is readily verified that each sub-channel by itself is TIN optimal. Therefore, according to Theorem 1, TIN is optimal for each sub-channel by itself. For the 3 sub-channels, consider the optimal cyclic partitions  $\Pi^{[1]*} = \{\{1, 2, 3\}\}$ ,  $\Pi^{[2]*} = \{\{3, 2, 1\}\}$ ,  $\Pi^{[3]*} = \{\{1\}, \{2, 3\}\}$ . The weights of the participating edges are

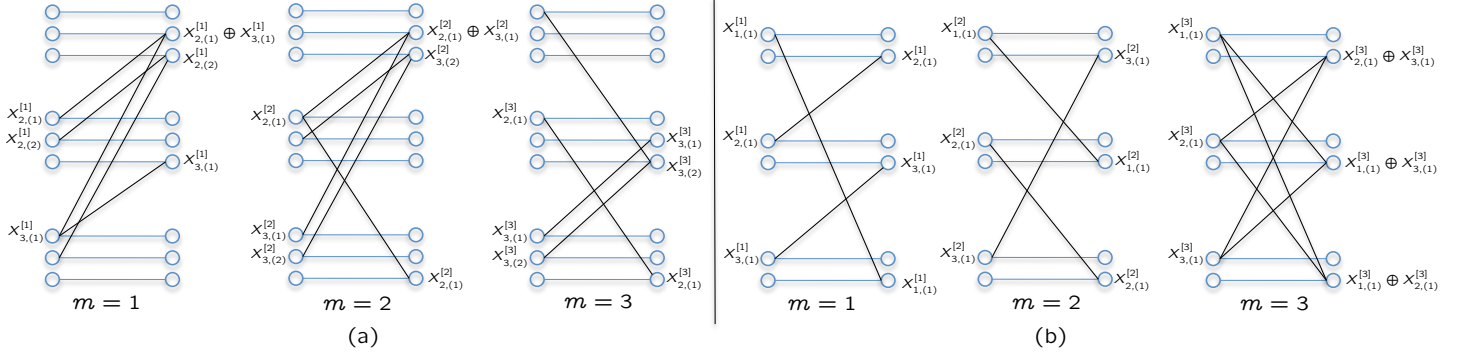


Fig. 1. Two 3 user ADT deterministic interference networks with 3 sub-channels each. Each sub-channel is TIN optimal individually. (a) For the optimal cyclic partitions  $\Pi^{[1]*} = \{\{1, 2, 3\}\}$ ,  $\Pi^{[2]*} = \{\{3, 2, 1\}\}$ ,  $\Pi^{[3]*} = \{\{1\}, \{2, 3\}\}$ , the participating input and output levels,  $X_{i,u}^{[m]}, Y_{i,u}^{[m]}, i, m \in \{1, 2, 3\}$  are labeled and the mapping from  $(X_{1,u}^{[m]}, X_{2,u}^{[m]}, X_{3,u}^{[m]})$  to  $(Y_{1,u}^{[m]}, Y_{2,u}^{[m]}, Y_{3,u}^{[m]})$  is easily verified to be invertible for each sub-channel. (b) For the optimal cyclic partitions  $\Pi^{[1]*} = \{\{1, 2, 3\}\}$ ,  $\Pi^{[2]*} = \{\{3, 2, 1\}\}$  and  $\Pi^{[3]*} = \{\{1, 2, 3\}\}$ , participating inputs and outputs  $X_{i,u}^{[m]}, Y_{i,u}^{[m]}, i, m \in \{1, 2, 3\}$  are labeled. In this case, the mapping from  $(X_{1,u}^{[3]}, X_{2,u}^{[3]}, X_{3,u}^{[3]})$  to  $(Y_{1,u}^{[3]}, Y_{2,u}^{[3]}, Y_{3,u}^{[3]})$  is not invertible.

$w(\Pi^{[1]*}) = w(\{e_{12}^{[1]}, e_{23}^{[1]}, e_{31}^{[1]}\}) = n_{12}^{[1]} + n_{23}^{[1]} + n_{31}^{[1]} = 3$ ,  $w(\Pi^{[2]*}) = w(\{e_{32}^{[2]}, e_{21}^{[2]}, e_{13}^{[2]}\}) = n_{32}^{[2]} + n_{21}^{[2]} + n_{13}^{[2]} = 3$ ,  $w(\Pi^{[3]*}) = w(\{e_{11}^{[3]}, e_{23}^{[3]}, e_{32}^{[3]}\}) = 0 + n_{23}^{[3]} + n_{32}^{[3]} = 3$ . Then according to Theorem 2, the sum-capacity values for each sub-channel by itself are given by  $\mathcal{C}_{\Sigma}^{[m]} = \sum_{i=1}^3 n_{ii}^{[m]} - w(\Pi^{[m]*}) = 9 - 3 = 6$ ,  $m = 1, 2, 3$ . What we wish to know is if TIN continues to be the sum-capacity optimal scheme for all 3 sub-channels collectively. Let us check for invertibility for each sub-channel. According to the definitions, the participating inputs for sub-channel 1 are  $X_{1,u}^{[1]} = [X_{1,(1)}^{[1]}, \dots, X_{1,(n_{31}^{[1]})}^{[1]}] = \phi$ ,  $X_{2,u}^{[1]} = [X_{2,(1)}^{[1]}, \dots, X_{2,(n_{12}^{[1]})}^{[1]}] = [X_{2,(1)}^{[1]}, X_{2,(2)}^{[1]}]$ ,  $X_{3,u}^{[1]} = [X_{3,(1)}^{[1]}, \dots, X_{3,(n_{23}^{[1]})}^{[1]}] = [X_{3,(1)}^{[1]}]$  and the participating outputs for sub-channel 1 are  $Y_{1,u}^{[1]} = [X_{2,(1)}^{[1]} \oplus X_{3,(1)}^{[1]}, X_{2,(2)}^{[1]}]$ ,  $Y_{2,u}^{[1]} = [X_{3,(1)}^{[1]}]$  and  $Y_{3,u}^{[1]} = \phi$ . It is now trivial to verify that from  $(Y_{1,u}^{[1]}, Y_{2,u}^{[1]}, Y_{3,u}^{[1]})$ , we can recover  $(X_{1,u}^{[1]}, X_{2,u}^{[1]}, X_{3,u}^{[1]})$ . Therefore, sub-channel 1 is invertible. Similarly, the participating inputs and outputs for sub-channels 2 and 3 are shown in Figure 1(a) and it is easily verified that sub-channels 2 and 3 are invertible as well. Therefore, we conclude that separate TIN is optimal for this parallel network, and, the sum-capacity of the 3 sub-channels collectively, is the sum of their individual sum-capacities. In other words, the sum-capacity is  $6 + 6 + 6 = 18$  and is achieved by separate TIN on each sub-channel.

To also expose the limitation of Theorem 3, the next example illustrates a relatively rare situation where invertibility is not satisfied, and so Theorem 3 cannot be applied. We do not know whether separate TIN is optimal here.

**Example 2:** Consider the 3 user ADT deterministic interference network with 3 sub-channels, as shown in Figure 1(b), with the optimal cyclic partitions  $\Pi^{[1]*} = \{\{1, 2, 3\}\}$ ,  $\Pi^{[2]*} = \{\{3, 2, 1\}\}$  and  $\Pi^{[3]*} = \{\{1, 2, 3\}\}$  for the first, second and third sub-channel, respectively. It is easy to verify that all 3 sub-channels are TIN optimal individually. However, with the participating inputs and outputs  $X_{i,u}^{[m]}, Y_{i,u}^{[m]}$  shown in the figure, it is also easy to see while the first two sub-channels are invertible, the third sub-channel is not.

### B. Mildness of Invertibility Condition

The intuition behind the mildness of the invertibility condition is analogous to the commonly encountered issue of invertibility of channel matrices in wireless networks, i.e., the property is satisfied everywhere except over an algebraic variety of lower dimension than the parameter space, and therefore is increasingly likely to be true when the parameter space is a large field. In particular, we expect invertibility to hold for an increasing fraction of channel realizations in the deterministic setting as the signal levels  $n_{ij}$  are drawn from larger sets (parallel to power increasing in the Gaussian setting for the GDoF metric). To strengthen this intuition, we take a closer look at the invertibility condition. We present 3 classes of networks that satisfy the invertibility property, thus separate TIN is sum-capacity optimal when each sub-channel is individually TIN optimal. The proofs and related discussions are relegated to the full paper [8].

**Network 1: (3 User)** 3 user ADT deterministic parallel interference networks where each sub-channel satisfies  $n_{12}^{[m]} + n_{23}^{[m]} + n_{31}^{[m]} \neq n_{21}^{[m]} + n_{32}^{[m]} + n_{13}^{[m]}, \forall m \in [M]$ .

**Network 2: (Acyclic Graph)**  $K$  user parallel ADT deterministic interference networks, where for each sub-channel, the bipartite graph is acyclic. The undirected bipartite graph is defined by viewing the cross links between the participating input and output levels as the edges.

**Network 3: (Dominant Partitions)**  $K$  user ADT deterministic parallel interference networks where in each sub-channel  $n_{\Pi^{[m]*}(k)k}^{[m]} > n_{jk}^{[m]}, \forall j, k \in [K], j \notin \{k, \Pi^{[m]*}(k)\}, m \in [M]$ .

### C. Proof of Theorem 3

Theorem 2 provides the achievable rate  $\sum_{m=1}^M \mathcal{C}_{\Sigma}^{\Pi^{[m]*}} = \sum_{m=1}^M \left[ \sum_{i=1}^K n_{ii}^{[m]} - w(\Pi^{[m]*}) \right]$  by separate TIN over each sub-channel. We only need to prove that it is an outer bound, assuming that each sub-channel is invertible. Consider the optimal cyclic partition for each sub-channel. Then by definition,  $w(\Pi^{[m]*}) = \sum_{i=1}^K n_{\Pi^{[m]*}(i)i}^{[m]}$ . We define  $i_{\max}^{[m]} \triangleq \arg\max_{j \neq i} n_{ji}^{[m]}$  to be the user that receives the most interference from Transmitter  $i$  in sub-channel  $m$ . Writing the binary

$$X_i^{[m]} = \sum_{b=1}^{n_{ii}^{[m]}} X_{i,(b)}^{[m]} 2^{-b} = \sum_{b=1}^{n_{\Pi^{[m]*}(i)i}^{[m]}} X_{i,(b)}^{[m]} 2^{-b} + \sum_{b=n_{\Pi^{[m]*}(i)i}^{[m]}+1}^{n_{i_{\max}^{[m]}}^{[m]}} X_{i,(b)}^{[m]} 2^{-b} + \sum_{b=n_{i_{\max}^{[m]}}^{[m]}+1}^{n_{ii}^{[m]}} X_{i,(b)}^{[m]} 2^{-b} \triangleq X_{i,u}^{[m]} + X_{i,v}^{[m]} + X_{i,q}^{[m]} \quad (2)$$

$$n(R_i - \epsilon) \leq I(W_i; \mathbf{Y}_i^n, \mathbf{X}_{i,u}^n, \mathbf{X}_{i,v}^n) = H(\mathbf{Y}_i^n, \mathbf{X}_{i,u}^n, \mathbf{X}_{i,v}^n) - H(\mathbf{Y}_i^n, \mathbf{X}_{i,u}^n, \mathbf{X}_{i,v}^n | W_i) \quad (3)$$

$$= H(\mathbf{X}_{i,u}^n | \mathbf{X}_{i,v}^n) + H(\mathbf{X}_{i,v}^n) + H(\mathbf{Y}_i^n | \mathbf{X}_{i,u}^n, \mathbf{X}_{i,v}^n) - H(\mathbf{Y}_i^n | W_i) \quad (4)$$

$$\leq H(\mathbf{X}_{i,u}^n | \mathbf{X}_{i,v}^n) + n \sum_{m=1}^M (n_{i_{\max}^{[m]}}^{[m]} - n_{\Pi^{[m]*}(i)i}^{[m]}) + n \sum_{m=1}^M (n_{ii}^{[m]} - n_{i_{\max}^{[m]}}^{[m]}) - H(\mathbf{Y}_i^n | W_i) \quad (5)$$

$$= H(\mathbf{X}_{i,u}^n | \mathbf{X}_{i,v}^n) - H(\mathbf{Y}_i^n | W_i) + n \sum_{m=1}^M (n_{ii}^{[m]} - n_{\Pi^{[m]*}(i)i}^{[m]}) \quad (6)$$

$$\sum_{i=1}^K n(R_i - \epsilon) \leq \sum_{i=1}^K H(\mathbf{X}_{i,u}^n | \mathbf{X}_{i,v}^n) - \sum_{i=1}^K H(\mathbf{Y}_i^n | W_i) + n \sum_{i=1}^K \sum_{m=1}^M (n_{ii}^{[m]} - n_{\Pi^{[m]*}(i)i}^{[m]}) \quad (7)$$

$$\leq H(\mathbf{X}_u^n | \mathbf{X}_v^n) - \sum_{i=1}^K H(\mathbf{Y}_i^n | W_i, \mathbf{X}_v^n) + n \sum_{m=1}^M \left( \sum_{i=1}^K n_{ii}^{[m]} - \sum_{i=1}^K n_{\Pi^{[m]*}(i)i}^{[m]} \right) \quad (8)$$

$$= H(\mathbf{X}_u^n | \mathbf{X}_v^n) - H(\mathbf{Y}_u^n | \mathbf{X}_v^n) + n \sum_{m=1}^M \left[ \sum_{i=1}^K n_{ii}^{[m]} - w(\Pi^{[m]*}) \right] = n \sum_{m=1}^M \mathcal{C}_{\Sigma}^{\Pi^{[m]*}} \quad (9)$$

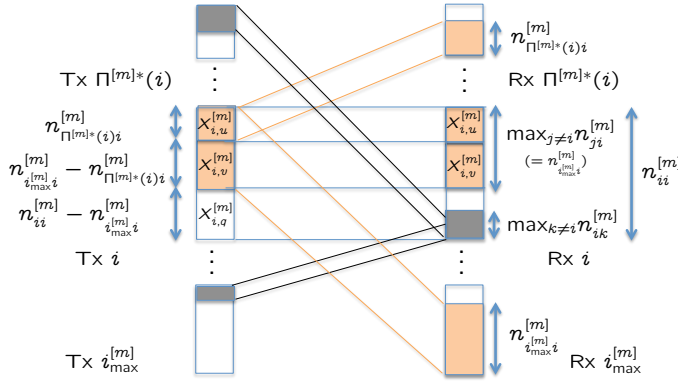


Fig. 2. The signal levels of user  $i$ . The signal levels that cause interference ( $X_{i,u}^{[m]}, X_{i,v}^{[m]}$ ) suffer no interference at the desired receiver.

expansion of the channel input, we have (2) shown at the top of this page, where  $X_{i,u}^{[m]}, X_{i,v}^{[m]}, X_{i,q}^{[m]}$  are the bits that interfere at Receiver  $\Pi^{[m]*}(i)$ , the other bits that interfere at Receiver  $i_{\max}^{[m]}$  and the remaining input bits, respectively (see Figure 2).  $\mathbf{X}_{i,u} = [X_{i,u}^{[1]}, \dots, X_{i,u}^{[M]}]$  is the stack of  $X_{i,u}^{[m]}$  for all sub-channels. Similar notation is used for  $\mathbf{X}_{i,v}$  with  $v$  replacing  $u$ . Give  $\mathbf{X}_{i,u}, \mathbf{X}_{i,v}$  through a genie to Receiver  $i$ . Then from Fano's inequality, we have (3) shown at the top of this page, where the second term in (5) follows from the fact that the entropy of  $X_{i,v}^{[m]}$  is smaller than the number of bits therein and the third term in (5) is due to the property that the signal levels in  $\mathbf{Y}_i$  that receive  $\mathbf{X}_{i,u}, \mathbf{X}_{i,v}$  do not suffer interference (see Figure 2), because each sub-channel is TIN optimal.

Adding (6) for  $i \in \{1, \dots, K\}$ , we have (7) (see above), where in (8),  $\mathbf{X}_u^n = (\mathbf{X}_{1,u}^n, \dots, \mathbf{X}_{K,u}^n)$  is the collection of  $\mathbf{X}_{i,u}^n$  for all users. Similar notations are used for  $\mathbf{X}_v^n$  and  $\mathbf{Y}_u^n$ . Then (8) follows from the independence of  $\mathbf{X}_i$  and the fact that conditioning does not increase entropy. The second term of (9)

follows from the definition of  $Y_{k,u}^{[m]} = \sum_{i=1, i \neq k}^K 2^{n_{ki}^{[m]}} X_{i,u}^{[m]}$  and the fact that given the desired message  $W_k$  and  $X_{i,v}^{[m]}$ , the only thing left in  $Y_k^{[m]}$  is  $Y_{k,u}^{[m]}$ . The last step is due to the invertibility assumption. Normalizing it by  $n$  and applying the limit  $n \rightarrow \infty$ , we arrive at the desired outer bound.

## V. DISCUSSIONS

In the context of  $K$  user parallel ADT deterministic interference networks when each sub-channel satisfies the TIN optimal condition, we show that separate TIN over each sub-channel is optimal from the perspective of sum-capacity, subject to a mild invertibility condition. The results translate directly into GDoF results for  $K$  user parallel Gaussian interference networks, as shown in the full paper [8].

## REFERENCES

- [1] C. Geng, N. Naderializadeh, S. Avestimehr, and S. Jafar, "On the Optimality of Treating Interference as Noise," *ArXiv:1305.4610*, May 2013.
- [2] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. on Inf. Theory*, vol. 57, pp. 1872–1905, 2011.
- [3] V. Cadambe and S. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Trans. on Inf. Theory*, vol. 55, no. 9, pp. 3983–3990, 2009.
- [4] —, "Sum-capacity and the unique separability of the parallel Gaussian MAC-Z-BC network," in *ISIT*, 2010, pp. 2318–2322.
- [5] L. Sankar, X. Shang, E. Erkip, and H. Poor, "Ergodic Fading Interference Channels: Sum-Capacity and Separability," *IEEE Trans. on Inf. Theory*, vol. 57, pp. 2605–2626, May 2011.
- [6] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Noisy-Interference Sum-Rate Capacity of Parallel Gaussian Interference Channels," *IEEE Trans. on Inf. Theory*, vol. 57, no. 1, pp. 210–226, 2011.
- [7] S. Karmakar and M. K. Varanasi, "The generalized degrees of freedom region of the MIMO interference channel and its achievability," *IEEE Trans. on Inf. Theory*, vol. 58, no. 12, pp. 7188–7203, 2012.
- [8] H. Sun and S. A. Jafar, "On the Optimality of Treating Interference as Noise for  $K$  user Parallel Gaussian Interference Networks," *ArXiv http://arxiv.org/abs/1401.2692*, Jan, 2014.