Sum Capacity of Interference Channels With a Local View: Impact of Distributed Decisions

Vaneet Aggarwal, Member, IEEE, Youjian Liu, Member, IEEE, and Ashutosh Sabharwal, Senior Member, IEEE

Abstract—Due to the large size of wireless networks, it is often impractical for nodes to track changes in the complete network state. As a result, nodes have to make distributed decisions about their transmission and reception parameters based on their local view of the network. In this paper, we characterize the impact of distributed decisions on the global network performance in terms of achievable sum rates. We first formalize the concept of local view by proposing a protocol abstraction using the concept of local message passing. In the proposed protocol, nodes forward information about the network state to other neighboring nodes, thereby allowing network-state information to trickle to all the nodes. The protocol proceeds in rounds, where all transmitters send a message followed by a message by all receivers. The number of rounds then provides a natural metric to quantify the extent of local information at each node. We next study two network connectivities, Z-channel, and a three-user double Z-channel. In each case, we characterize achievable sum rate with partial message passing leading to two main results. First, in many cases, nodes can make distributed decisions with only local information about the network and can still achieve the same sum capacity as can be attained with global information irrespective of the actual channel gains. We label such schemes as universally optimal. Second, for the case of three-user double Z-channel, we show that universal optimality is not achievable if the per node information is below a threshold. In fact, distributed decisions can lead to unbounded losses compared to full information case for some channel gains.

Index Terms—Distributed decisions, double Z-channel, interference channel, local view, message passing, universally optimal strategy, Z-channel.

I. INTRODUCTION

Due to mobility, the network connectivity and channel gains in a wireless network are constantly time varying. For nodes to make optimal decisions about their transmission and reception parameters, like rate, power, codebooks, and decoders, they require full knowledge of the state of the network (defined as the network connectivity and the channel gains on each link) to compute the capacity region and thus their own operational point. However, in large wireless networks, centralizing complete information about the network state implies prohibitive overhead and, thus, seldom performed in any practical network. As a result, nodes have to make distributed decisions about their transmission and reception parameters based on their limited local view of the network state. In this situation, the driving question is: Can and when do distributed decisions lead to globally optimal network operation?

In this paper, we consider single-hop interference channels [3]–[5], where the receiver for each transmitter has a direct connection to its receiver but otherwise the network connectivity is arbitrary. The network state is not known to any node in the network. As a result, none of the nodes know the set of jointly achievable rates and the associated capacity-achieving transmission schemes. In contrast, prior work in quantifying the network capacity implicitly assumes that each node in the network knows the full network state perfectly, e.g., [1]–[3].

To understand network performance with partial information about the network, we need a natural metric to quantify extent of network information each node has about the network. Toward that end, we propose a protocol abstraction in the form of a local message-passing protocol, where the nodes propagate messages related to network-state information. The protocol abstraction is inspired by the fact that in a network, the only feasible mechanism available for nodes to learn any information is to pass messages to their neighbors. In fact, local message passing is the building block in all network protocols, such as medium access, routing and gossiping. Inspired by the belief propagation algorithm commonly used in LDPC decoding [6], the proposed message-passing algorithm proceeds in rounds, where one round consists of a forward and a reverse phase. In the forward phase, each transmitter sends a message and in the reverse phase, each receiver responds with a message. Each message constitutes of only the new information and, thus, is similar to extrinsic information in belief propagation. While there are many parallels between the proposed algorithm and belief propagation, we will not explore them further in this paper.

The message-passing protocol exposes the fundamental capacity problem of interest. With each round the nodes learn more about the network but they do not have full network information till the message-passing protocol has terminated. As a result, in the intermediate rounds before the termination of the protocol, not only the nodes have incomplete network information, but they may also have mismatched view of the network. Thus, we propose to use the number of protocol rounds as a
proxy to quantify extent of network information each node has about the network state.

The characterization of capacity under partial network-state information is nontrivial since the exact network capacity with full information is still unsolved. Thus, we consider two special cases: Z-channel and three-user double Z-channel (two Z channels stacked on top of each other). In each case, we focus on the sum-rate point on the capacity boundary. Each of the network connectivities has the largest network diameter for a given number of users, and thus, the message passing requires the maximum number of rounds to ensure that all nodes have full network information. For the cases when the protocol has not terminated, the resulting network problem is often labeled as the hidden-node problem in the network protocol literature [21]. While there is a rich body of literature to design protocols to counter the issues related to hidden nodes, the authors are not aware of any information theoretic capacity analyses with hidden nodes.

We seek universally optimal strategies, where each node decides its action based only on its own view of the network but the resulting network sum rate is equal to the sum capacity with global information at all the nodes for all network states. Our results are derived for both deterministic and Gaussian channels and are summarized as follows.

1) Z-channel: For the Z-channel, the message-passing protocol requires three full rounds to terminate. In the case of deterministic Z-channel, we show that a unique universally optimal scheme exists with only 1.5 rounds of the message-passing protocol, even though one of the transmitters does not know all the channel gains. The key feature of the scheme is that transmitters are politely as greedy as possible but do not hurt transmission of any other flow about which they have knowledge. The deterministic case is extended to the case of the Gaussian Z-channel, where we show that the sum capacity within 2 bits can be obtained with 1.5 rounds of message passing. Note that at least two full rounds are required for all nodes to learn the full network in a Z-channel.

2) Double Z-channel: For the case of deterministic double Z-channel, four rounds of message passing are needed. We build on the Z-channel result to propose a distributed rate allocation for double Z-channel and characterize the resulting sum rate after 1.5 rounds and 2.5 rounds, which are in general different. Our result shows the growth in achievable sum rate with more information about the network. In this case, 2.5 rounds suffice to obtain a universally optimal scheme. However, the scheme is no longer unique. For the case of 1.5 rounds, we also show a converse result that there exists no distributed scheme which can be universally optimal. In fact, our proposed scheme is optimal for some network states but can have arbitrarily large losses in other cases. Thus, this is the first indication that partial information can significantly reduce network capacity and that loss in network capacity is unavoidable, i.e., every strategy will be suboptimal in a certain regime of channel gains.

In order to prove the aforementioned results, the capacity region for the deterministic three-user double Z-channel is found in this paper for a general class of deterministic channels which is later specialized to the deterministic models. The aforementioned results have further been extended to a Gaussian model. For the Gaussian double Z-channel, the sum capacity within 4 bits can be achieved with 2.5 rounds of message passing. However, for 1.5 rounds, sum-rate optimality can only be guaranteed for a subset of channel gains. To derive the Gaussian result, we derived novel outer bounds for the sum capacity of the Gaussian double Z-channel. Interestingly, we find that treating interference as noise which is optimal for weak interference for the Z-channel is only optimal for the double Z-channel in the cases of very weak interference.

We make two salient observations about the message-passing protocol and its relation to other metrics. First, after d full rounds of message passing, each transmitter knows all channels which are 2d hops away and each receiver knows links up to 2d − 1 hops away. After d +0.5 rounds, receiver information increases to 2d +1 hops. Thus, the number of protocol rounds is directly related to a common method to specify local information in network protocol analyses. Second, the messages can be easily related to practical network operation. For example, 1.5 rounds is a common choice in cellular systems, translating to beacons from the base stations, feedback by the mobile units followed by last half round of a message from the base station indicate rate and power decisions.

The rest of the paper is organized as follows. In Section II, we describe the channel models. In Section III, we give a general message-passing protocol for a K-user interference channel. In Section IV, we characterize the sum rate with partial information at the nodes for a deterministic and Gaussian Z-channel. In Section V, we find the capacity region for a deterministic double Z interference channel and outer bounds for the Gaussian deterministic channel. The capacity region is found for a general class of deterministic channels on the lines of [14] which is then specialized to a specific deterministic channel model considered in this paper. We also derive results for 1.5 and 2.5 rounds of message passing for both deterministic and Gaussian channels. In Section VI, we conclude the paper.

II. PROBLEM FORMULATION

A. Channel Models

We will consider two models for interference channels with K transmitters \( \{T_i\}_{i=1}^K \) and K receivers \( \{D_i\}_{i=1}^K \); a deterministic model [3] and the additive noise Gaussian model described as follows. We assume that both transmitters and receivers can transmit messages, i.e., the network is bidirectional in nature.

1) Deterministic Model: In a deterministic interference channel, the inputs of \( k \)th transmitter at time \( i \) are denoted by \( X_{ki} = [X_{ki1}, X_{ki2}, \ldots, X_{kiq}]^T \in \{0,1\}^q \), \( k = 1, 2, \ldots, K \), such that \( X_{ki1} \) and \( X_{kiq} \) are the most and the least significant bits, respectively. The received signal of user \( j \) at time \( i \) is denoted by the vector \( Y_{ji} = [Y_{iji1}, Y_{iji2}, \ldots, Y_{ijiq}]^T \in \{0,1\}^q \). Specifically, the received signal \( Y_{ji} \) of an interference channel is given by

\[
Y_{ji} = \bigoplus_{k=1}^K S_k^{q-n_{ki}} X_{ki}
\]
where $\oplus$ denotes the XOR operation, and $S^{\sigma_{m,n}}$ is a $q \times q$ shift matrix with entries $S_{mn}$ that are nonzero only for $(m,n) = (q-nk+j, n), n = 1, 2, \ldots, nk$. We will also use $X_k^n$, $Y_k^n$, etc., to denote $(X_{k1}, \ldots, X_{kn})$, $(Y_{k1}, \ldots, Y_{kn})$, etc. Associated with each transmitter $k$ and receiver $j$, is a nonnegative integer $n_{kj}$ that defines the number of bit levels of $X_k$ observed at receiver $j$. The maximum number of bits supported by any link is $q = \max_{k,j} n_{kj}$. The network can be represented by a square matrix $H$ whose $(i,j)$th entry is $H_{ij} = n_{ij}$. We note that $H$ need not be symmetric.

2) Gaussian Model: In a Gaussian interference channel, the inputs of $k$th transmitter at time $i$ are denoted by $X_{ki} \in \mathbb{C}$, $k = 1, 2, \ldots, K$, and the outputs at $j$th receiver in time $i$ can be written as $Y_{ji} \in \mathbb{C}, j = 1, 2, \ldots, K$. The received signal $Y_{ji}$, $j = 1, 2, \ldots, K$ is given by

$$Y_{ji} = \sum_{k=1}^{K} h_{kj}X_{ki} + Z_{ji} \quad (2)$$

where $h_{kj} \in \mathbb{R}^+$ is the channel gains associated with each transmitter $k$ and receiver $j$, and $Z_{ji}$ are additive white complex Gaussian random variables of unit variance. We will also use $X_k^n$, $Y_k^n$, etc., to denote $(X_{k1}, \ldots, X_{kn})$, $(Y_{k1}, \ldots, Y_{kn})$, etc. Furthermore, the input $X_{ki}$ has an average power constraint of unity, i.e., $\mathbb{E}(|X_{ki}|^2) \leq 1$ (where $\mathbb{E}$ denotes the expectation of the random variable). Only in Section III, we will make an exception where no power constraint will be imposed on messages sent by transmitters or receivers.

Like the deterministic case, we represent the network state by a square matrix $H$ whose $(i,j)$th entry is $H_{ij} = |h_{ij}|^2$. Thus, we will use the matrix $H$ for both the deterministic and the Gaussian model, where the usage will be clear from the context.

B. Per Node Local View

Our objective is to understand the impact of nodes’ decisions on network sum rate, when the decisions are based on their partial information about the matrix $H$. For transmitters, we will denote this partial information about the network as $N_k$ and as $N'_k$ for the receivers. If the nodes know nothing about the network matrix $H$ (i.e., no information about its size or entries), then $N_k = N'_k = \emptyset$ (empty set), which is equal to assuming that there is no other node in the network. On the other hand, if the nodes know everything about the network, then $N_k = N'_k = H$ and is also the most commonly assumed scenario in most information-theoretic analyses [4], [17], [20].

We now define network state and network connectivity. We assume that there is a direct link between every transmitter $T_i$ and its intended receiver $D_i$. On the other hand, if a cross-link between transmitter $i$ and receiver $j$ does not exist, then $H_{ij} \equiv 0$. Given a network, its connectivity is a set of edges $E = \{(T_i, D_j)\}$ such that a link $T_i \rightarrow D_j$ is not identically zero. Then, the set of network states $G$ is the set of all weighted graphs defined on $E$. For the deterministic model, the set of network states can be written as $G(E) = \{H : H_{ij} \equiv 0 \text{ if } (T_i, D_j) \notin E \text{ else } H_{ij} \in \{0, 1, \ldots, q\}\}$ and in the Gaussian model as

$$G(E) = \{H : H_{ij} \equiv 0 \text{ if } (T_i, D_j) \notin E \text{ else } H_{ij} \in \{0, 1, \ldots, q\}\}.$$

Note that the channel gain can be zero but not guaranteed to be if the node pair $(T_i, D_j) \notin E$.

Our main focus is the case where $N_k$, $N'_k$ for each $k$ is only a subset of the whole matrix. Thus, each node knows the network matrix partially. In fact, it is quite possible that nodes know only a few entries of the matrix and do not know the size of the whole matrix $H$, i.e., network size. As we will see later, this partial knowledge of the network matrix at each node leads to the case where each node’s knowledge about the network is mismatched from other nodes in the network. That is, each node possibly knows a different part of the whole matrix $H$. We will study the achievable sum rate as the network information at each node grows from no information to full information. In Section III, we will define a special trajectory of sequence of growing network information which is directly connected to protocols in practical systems and is also related to commonly used metric of “number of hops” to denote amount of side information at each node. To aid analysis, we will assume that all nodes are provided some side information $SI$ about the network state before the onset of the protocol. Thus, nodes may have nonzero information about the network before even a single message is sent. We next define the sum capacity.

C. Sum Capacity

First consider the $K$-user deterministic interference channel. For each user $k$, let the message index $m_k$ be uniformly distributed over $\{1, 2, \ldots, 2^nR_k\}$, where $R_k$ is chosen by a node based on its local view $N_k$ and side information about the network $SI$. The message is encoded as $X_k^n$ using the encoding functions $e_k(m_k|N_k, SI) : \{1, 2, \ldots, 2^nR_k\} \mapsto \{0, 1\}^{nq}$, which depend on the local view $N_k$ and side information about the network $SI$. The message is decoded at the receiver using the decoding function $d_k(Y_k^n|N'_k, SI) : \{0, 1\}^{nq} \mapsto \{1, 2, \ldots, 2^nR_k\}$, where $N'_k$ is the receiver local view and $SI$ is the side information. The corresponding probability of decoding error $\lambda_k(n)$ is defined as $Pr[m_k \neq d_k(Y_k^n|N'_k, SI)]$. A rate tuple $(R_1, R_2, \ldots, R_K)$ is said to be achievable if there exists a sequence of codes such that the error probabilities $\lambda_1(n), \ldots, \lambda_K(n)$ go to zero as $n$ goes to infinity for all network states consistent with the side information. For instance, the encoding scheme chosen at a certain encoder must still give small error probabilities for every possible network state possible with that local and side information.

Now, consider the $K$-user Gaussian interference channel. For each user $k$, we again assume that the message index $m_k$ is uniformly distributed over $\{1, 2, \ldots, 2^nR_k\}$, where $R_k$ is chosen by a node based on its local view $N_k$ and side information about the network $SI$. Furthermore, we use the same notation for encoding and decoding functions. Thus, encoding functions are

1The model is inspired by fading channels, where the existence of a link is based on its average channel gain. On the average, the link gain may be above noise floor but its instantaneous value can be below noise floor.
The sum capacity in both cases is defined as

$$C_{\text{sum}} = \sup \left\{ \sum_{i=1}^{K} R_i : (R_1, \ldots, R_K) \in \mathcal{C} \right\}.$$  

We note that all encoding and decoding functions depend only on the local and side information at the transmitters and receivers about the network. In this case, the nodes have to operate with the local knowledge $N_k, N'_k$ and the side information $SI$ so that the probability of error at the receivers go to zero as $n$ goes to infinity for all $H \in \mathcal{G}(E)$, leading to a compound channel capacity formulation. In this paper, the optimal sum capacity refers to the sum capacity with the full state information at all the nodes, and the capacity region $\mathcal{C}$ is the closure of the set of achievable rate tuples with full state information at all nodes.

When the local information about the network is mismatched, the nodes can take actions which can work against each other and in the process reduce the sum rate of the network. This issue of making distributed decisions about rate, power, codebooks and decoder is fundamental in most networks and is the main topic of study in this paper.

### III. LEARNING NETWORK-STATE INFORMATION

In this section, we describe a protocol that uses local message passing to propagate network-state information to the nodes in the network. The protocol is described in terms of entries of matrix $H = [H_{ij}]$ and, thus, applies to both deterministic and Gaussian models. Our motivation is to find the most relevant cases of local view $\{N_k, N'_k\}$ that we should consider.

#### A. Why Learn the Network?

Before we consider the details of learning the network-state information, it is important to understand why we might need to even estimate and propagate network-state information. One could adopt a “noncoherent” approach, where no resources are wasted in estimating any channel or network connectivity and nodes code such that reliable communication is possible without any network-state information, i.e., $N_k = N'_k = \emptyset$. However, in compound capacity formulation, the capacity region with no information (local or side) about the network is a singleton, where the only possible rate tuple is all zero-rate tuple. This follows directly from point-to-point Gaussian channel, where the compound capacity is zero if the link state is unknown to the transmitter since in the worst case the link gain can be zero. Thus, to achieve a nonzero rate, the network information at nodes should be nontrivial.

The obvious next question is what cases of network information $\{N_k, N'_k\}$ should one consider. For a $K$-user network, the matrix has $K^2$ entries, which implies that the per node information can be any of $2^{K^2}$ cases. Thus, there are $2^{2K^3}$ possible combinations of side information cases. This large number of cases quickly becomes intractable. However, we contend that most of these side information cases are not of practical interest.

A common metric to capture extent of network view at each node is number of hops (e.g., see [22] and references therein). That is equivalent to each node $k$ knowing a submatrix of $H$. The metric is motivated by message-passing algorithms which broadcast and forward information about a local state. A clear advantage of this metric that it greatly reduces the number of local information subcases one needs to consider, and there is a direct relation with actual protocols which gather this side information. We propose to adopt a related metric that is equally concise and tightly related to protocols in many operational networks.

#### B. Message-Passing Protocol

For nodes to learn and propagate the network state, they have to communicate with each other. This internode communication is possible only with nodes to which there is a direct link, i.e., messages have to be exchanged locally and those messages are then processed and propagated to other nodes. This obvious construct of local message passing is central to all multihop network protocols. We assume that the message-passing protocol runs over an out-of-band, bidirectional, noiseless message-passing channel, which is separate from the payload channel. The message-passing channel and the payload channel share the same topology (i.e., connectivity graph), which is motivated by current wireless systems. We assume that the connectivity of message-passing channel is given by $E$, while some of the channel gains can be 0 for the payload channel. In our development, the only practical reality we will be concerned with is that direct communication is possible only between neighbors and its impact on amount of network-state information at each node. Hence, we will simplify some of the implementation complexities as follows.

The proposed message-passing protocol proceeds in rounds, where each full round has two phases: a forward phase where all transmitters broadcast a message and a reverse phase where all receivers broadcast a message each. We assume that all messages are scheduled so that there are no collisions at any of the nodes in the receiving mode due to simultaneous transmissions. Finally, the broadcast messages can only be heard by nodes to which the sending nodes has direct links (the links that are in the network connectivity $E$); thus, no extra feedback or Genie channels are available.

The message broadcasted by the transmitter $k$ in round $t$ (transmitters are data sources) is labeled $m_{k,t}$, which is received by all the receivers $j$ who have direct links to transmitter $k$. Analogously, the message broadcasted by the receiver $k$ at round $t$ is labeled $M_{k,t}$, which is received by all the transmitters $j$ who have a direct link to receiver $k$. We assume that the message-passing channel accepts tuples as inputs. It is akin to perfect out-of-band feedback channels, or conferencing channels.
Recall that each node’s information about the network is represented by either $N_k$ (for transmitters) or $N'_k$ (for receivers). Instead of assuming that the nodes have no information to start with (i.e. $N_k = N'_k = \emptyset$), we will consider special subcases, where all nodes have some minimal side information $\mathbf{S}_i$ about the network. The assumption of side information is largely for analytical simplification. This will only change the contents of messages and not the message computation and passing rules. For the following description, we will assume that at time $t = 0$, each node knows the size of the network or $\mathbf{S}_i = \{K\}$. Thus, the nodes know the size of the matrix $H$ but do not know any of its entries. An alternate case will also be considered where the size of network $K$ will be assumed unknown at the onset of the protocol but nodes will have a priori knowledge of the form of the matrix $H$.

The message-passing protocol with knowledge of side information $\mathbf{S}_i = \{K\}$ is described as follows.\[1) \textit{Round 1 (Forward):} \textit{Since none of the entries in the matrix }\]

$H$ in known, the first message from each transmitter is a known training signal along with the transmitter identity. Thus, $m_{k,1} = \{\psi_k, T_k\}$, where $\psi_k$ is the training signal from transmitter $k$. At the end of the transmitter messages, the receiver $j$ knows the nonzero elements of column $j$ of matrix $H$ learnt via channel estimation (however may not know the value of $j$).

$\textit{Round 1 (Reverse):}$ The receiver $k$ broadcasts $M_{k,1} = \bigcup_{i \in E_k} \{(H_k, T_i, D_k)\}$, where $E_k$ is the set of vertices connected to receiver $k$. The transmitter $T_j$ can receive $M_{k,1}$ if it has a direct link to receiver $k$. This completes the first round.

$2) \textit{Roundt} > 1$: In round $t > 1$, nodes only forward new information, which is computed as follows. In the forward phase for transmitters, the broadcast message is

$$m_{k,t} = \bigcup_{j \in J_k} M_{j,t-1} \setminus \bigcap_{t'=2}^{t-1} \bigcup_{j \in J_k} M_{j,t'} \bigcup_{t'=2}^{t-1} \bigcup_{j \in J_k} M_{j,t'} \bigcup_{t'=2}^{t-1} \bigcup_{j \in J_k} M_{j,t'}$$

(4)

where $J_k$ is the set of vertices connected to transmitter $k$. The message $m_{k,t}$ is a concatenated version of its received messages from previous round minus the messages it has broadcast in previous transmissions and those that are already known to all of its neighbors.

In response, the receivers broadcast following in the reverse phase:

$$M_{k,t} = \bigcup_{j \in E_k} m_{j,t} \setminus \bigcap_{t'=1}^{t-1} \bigcup_{j \in E_k} m_{j,t'} \bigcup_{t'=1}^{t-1} \bigcup_{j \in E_k} m_{j,t'}$$

(5)

The message $M_{k,t}$ is the concatenation of its received message minus its previously broadcasts messages and after removing what is known to all its neighboring transmitters. The messages $m_{k,t}$ and $M_{k,t}$ are similar to the extrinsic information in belief propagation with the main difference being that the messages are broadcasts.

3) $\textbf{Stopping rule:}$ If a transmitter or receiver has no new updates, it sends a NULL message $\phi$ in its assigned time slot. Thus, nodes only forward information when new information is received and send “nothing” otherwise. When all the neighbors of a node send a NULL message, each node stops sending any new messages (even NULL messages).

The aforementioned protocol is similar to message passing for belief propagation on factor graphs, often used in LDPC decoding [6]. Belief propagation is closer to gossiping in networks [7], where a node can talk only one node at any time. Our proposed approach exploits the broadcast nature of wireless and, hence, is closer to broadcasting in networks [8]. While connections between broadcast-based generalization of belief propagation [9] and proposed message passing are of interest, they are beyond the scope of this paper. In Sections IV and V, we assume that the protocol is used for $d+1/2$ rounds for $d \geq 0$. The extra half round at the end is the forward message, which will aid in the receivers knowing more than any of the transmitters it is connected to. Thus, the receiver will know the codebooks used by the transmitters allowing it to decode the messages.

Before we proceed, we quickly note our use of word “message.” In this paper, we use messages for transmissions, which may or may not depend on information bits. In contrast, the usual information-theoretic parlance, message often only refers to the “raw” information bits (the information-bearing message) sent from sender to the receiver [23]. In our case, the messages do carry information but about the network state. Thus, they are similar to control messages, like training, feedback, ARQ, etc., in networks.

C. Notes on the Message-Passing Protocol

It is instructive to consider an example to understand the properties of the message-passing protocol; Fig. 1 shows the steps for the Z-channel. Some key facts are as follows.

1) $\textit{Form of messages:}$ Each nonnull message is an unordered set of channel coefficients along with the location in the matrix indicated by the transmitter–receiver identity. This ensures that the set intersection and union in (4) and (5) are well defined.

2) $\textit{Guaranteed termination in K+1 rounds:}$ Since the channel estimation is assumed to be perfect and messages face no losses, the message-passing protocol is guaranteed to converge in $K + 1$ rounds. However, the actual number of rounds depends on the graph diameter. If the interference network is fully connected, then message passing terminates after two rounds. However, for a stacked Z-network, full $K + 1$ rounds are required.

3) $\textit{Impact of side information:}$ The messages can be reduced in length if any side information about the network is known. In the preceding discussion, we assumed that the network size is known, and hence, the size of the matrix is assumed known. As a result, we do not need to estimate the network size.

4) $\textit{Full and partial network information:}$ The case of full information is equivalent to the case of message passing operating till termination. With fewer than maximum number of rounds, at least one of the nodes may not have full network information. For example, in Fig. 1, less than two
rounds imply that there is some node (e.g., $T_1$), which does not have full information.

5) **Mismatched local views**: The protocol naturally exposes one of the key issues in networks that different nodes will have different information about the network state if the protocol is not carried to its completion. For large networks, taking the protocol to completion would imply a large number of rounds and is, thus, impractical. For example, after 1.5 rounds (first full round + only forward phase in second round) in the Z-channel, $T_2$ does not know the $T_1 \rightarrow D_1$ channel, while other nodes know the whole network. We will thus use number of rounds as a proxy for amount of local information, with each extra round providing increased information about the network.

6) **Relation to hop length**: By observing the time line in message passing in Fig. 1, it is straightforward to conclude that different nodes learn about the whole network at different times. Each full round (except the first) allows every node to learn two hops of extra information. In $d$ rounds, a transmitter will know all routes, which are $2d$-hop long rooted at that transmitter, and a receiver knows routes of length $2d-1$. That is why, in Fig. 1, $T_2$ needs two rounds to learn about $H_{11}$, since it is three hops away.

7) **Relation to practical protocols**: The messages can be translated into practical network operation. First round is training, like physical layer preamble in most networks. And rest of the rounds can be understood as channel feedback, like often studied in [10]. In practice, networks operate with very few rounds of message passing. For cellular networks, 1.5 rounds is a common choice, roughly translating to beacons from base stations, feedback from mobiles, and a rate-allocation decision from the base station. Thus, the case of 1.5 rounds will be of special interest to us.

### D. Optimal Strategies

We now formally define the concept of universally optimal strategy with partial information. Suppose that the transmitter $i$ knows local information $N_i$, the receiver $i$ knows local information $N'_i$, and all the nodes know side information $\text{SI}$.

**Definition 1 (Approximate Universal Optimality)**: An approximate universally optimal strategy with partial information at nodes $(N_i$ at transmitter $i$, $N'_i$ at receiver $i$, and $\text{SI}$ at all the nodes) is defined as the strategy that each of the transmitter $i$ uses based on its local information $N_i$ and side information $\text{SI}$, such that following holds. There exist a sequence of codes having rates $R_i$ at the transmitter $i$ such that the error probabilities at the receivers, $\lambda_1(n), \ldots, \lambda_K(n)$, tend to zero as $n$ tends to infinity, satisfying

$$\sum_i R_i \geq C_{\text{sum}} - \tau$$

for all the sets of network states consistent with the side information. Here, $C_{\text{sum}}$ is the sum capacity of the whole network with the full information and $\tau$ is a fixed constant independent of the channel gains. A strategy is called universally optimal if $\tau = 0$.

Thus, an (approximate) universally optimal strategy is one where for all network states consistent with the side information, decisions based on local information and the side information lead to (approximately) globally optimal solutions. We will use the notion of universal optimality for the deterministic model. Since we do not know the exact capacity region for the Gaussian interference channel, the notion of universal optimality will be replaced by approximate universal optimality.

In this paper, we will assume that the local information at the nodes is obtained by a message-passing protocol when run for $d$ or $d+5$ rounds for $d \geq 0$. In the case of $d+5$ rounds, the last 0.5 round of message passing represents the message from the transmitters to the receivers but no message in the reverse direction. This is to ensure that the receivers know more than the transmitters so that reliable decoding can take place.

### IV. Two-User Z-Channel

The smallest possible network of interest is a two-user interference channel. There are three network connectivities in this case, a fully connected bipartite graph (interference channel), a Z-channel (one cross-link is missing) and two decoupled flows (two point-to-point links). We assume that all the nodes know that there are a total of two nodes in the network, or $\text{SI} =
\{K = 2\}. From the message-passing protocol in Fig. 1, it is clear that 1.5 rounds are sufficient for every node in the network to learn the whole of the state information for both fully connected bipartite graph and two decoupled flow connectivities. So the strategies decided with local view result in globally optimal strategy decisions. Hence, we will focus our interest on the Z-channel, where 1.5 rounds do not result in full information at all the nodes; \( T_2 \) does not know \( H_{11} \) after 1.5 rounds.

### A. Deterministic Z-Channel

In a deterministic Z-channel, \( K = 2 \) and \( n_{21} = 0 \). Specifically, the received signal \( Y_{ji}, j = 1,2 \), of a Z-channel is given by

\[
Y_{1i} = S^{n_{11}} X_{1i} \tag{6a}
\]

\[
Y_{2i} = S^{n_{12}} X_{1i} \oplus S^{n_{22}} X_{2i}. \tag{6b}
\]

The network-state message passing is described in Fig. 1. After two full rounds, every node in the network knows the complete network state, i.e., the matrix \( H \) is known completely to all four nodes. In this case, the achievable capacity region is also known exactly [5], [11], [14].

**Theorem 1** ([5], [11], [14]): The deterministic channel capacity region for a two-user Z-channel is the set of nonnegative rates \( (R_1, R_2) \) satisfying

\[
R_1 \leq n_{11} \tag{7a}
\]

\[
R_2 \leq n_{22} \tag{7b}
\]

\[
R_1 + R_2 \leq \max(n_{22}, n_{12}, n_{11} + n_{22} - n_{12}). \tag{7c}
\]

Since our main interest is in the case of partial information, we ask if fewer than two full rounds suffice to achieve the sum capacity. The following theorem proves that 1.5 rounds are sufficient to achieve the sum capacity for all two-user \( H \), i.e., a universally optimal rate allocation exists for all \( H \). Note that we assume that the network size of \( K = 2 \) is known to all the nodes. Furthermore, in 1.5 rounds of message passing, all nodes know that \( n_{21} = 0 \). Thus, all nodes know that the connectivity is that of a Z-channel.

**Theorem 2**: The sum capacity for a Z-channel can be achieved without completing the full message-passing algorithm. To be precise, only the first 1.5 rounds are required to achieve the full-knowledge sum capacity with the side information that the network size is \( K = 2 \).

**Proof**: With the side information and 1.5 rounds of message passing, all the nodes know if they are the top user or the bottom user of the Z-channel. In Appendix A, we show that each transmitter can use only their local view to decide their rate and codebooks such that full-knowledge sum capacity is achievable.

**Corollary 1**: There exist a universally optimal strategy with the local information \( N_i \) and \( N_i' \) provided by 1.5 rounds of message passing and the side information that \( K = 2 \).

**Proof**: With 1.5 rounds of message passing and side information \( K = 2 \), all the nodes know the network connectivity. For all the connectivity choices, a strategy can be found by the nodes based on the channel gains they know that would achieve the full-information sum capacity. For example, if the network is fully connected, the nodes can use the node indices to order themselves since the labels \( T_i \) are unique and thus the nodes can compute an optimal strategy. If the network is a Z-channel, the strategy of Z-channel described in the proof of Theorem 2 can be used.

Thus, there is no loss in the performance even if the second transmitter \( T_2 \) does not know about the whole network state, and schemes with partial information exist, which are universally optimal for all Z-channel \( H \in G(E_Z) \), where \( E_Z \) is the \( E_Z \) is the Z-channel connectivity. It is very instructive to closely study the structure of the rate-allocation scheme.

Since transmitter \( T_2 \) does not know the direct channel \( H_{11} \), it in fact chooses to ignore the presence of the other transmitter. As a result, \( T_2 \) acts in a greedy fashion and sends at full rate of \( H_{22} \) bits. On the other hand, \( T_1 \) knows the whole network and it interferes with \( D_2 \). It can then choose a transmission scheme that sends information below the noise floor of \( T_2 \)'s transmission and, if possible, above \( T_2 \)'s signal.

The schemes which achieve the optimal sum capacity follow two rules. First, if a transmitter does not have enough information about other links, it acts greedily and sends at its maximum possible link rate. This is the case for the \( T_2 \), which does not know about the \( T_1 \rightarrow D_1 \) link. Second, if the link does know who it is causing interference to and by how much, then it ensures that it only sends at rates and powers, which do not impede on the success of other flows. This is the case for transmitter \( T_1 \), which knows that it is causing interference at receiver \( D_2 \). In short, the transmitters act greedily and politely by maximizing their individual rate but constraining themselves not to hurt other flows about which they have sufficient information.

We next show that the rate allocation in the aforementioned strategy is the only unique possible way to obtain a universally optimal strategy with 1.5 rounds of message passing.

**Theorem 3**: For the network connectivity of Z-channel, there exists a unique distributed rate-allocation strategy that is universally optimal with 1.5 rounds of the message-passing protocol and is given by the strategy of Theorem 2.

**Proof**: To show the uniqueness, we first consider the strategy of the second user. Since the second user does not know \( n_{11} \), \( T_2 \) has to transmit at \( n_{22} \) to avoid being suboptimal for the case of \( n_{11} = 0 \). This implies that irrespective of the channel gains \( n_{22} \) and \( n_{12} \), no strategy that involves the rate of the second user other than \( n_{22} \) can be universally optimal.

Now given that the aforementioned \( T_2 \) strategy of sending at full rate, the first user cannot assign any other rate than in the proof of Theorem 2, since only this rate would result in the optimality of the sum rate given that the second user is sending at \( n_{22} \).

Perhaps the most interesting aspect of Theorem 2 is that locally optimal rates achieve globally optimal sum capacity. The reason is that the proposed distributed scheme achieves a corner point on the capacity region of Z-channel, which explains the
uniqueness as shown in Theorem 3. Thus, an interesting observation is that corner points need less information about the network at some of the nodes. We will observe this fact again for a bigger network in Section V.

B. Gaussian Z-Channel

In a Gaussian Z-channel, \( K = 2 \) and \( h_{21} = 0 \). The received signal \( Y_{ji}, j = 1,2 \), of a Z-channel is given by

\[
Y_{1i} = h_{11}X_{1i} + Z_{1i} \quad (8a)
\]

\[
Y_{2i} = h_{12}X_{1i} + h_{22}X_{2i} + Z_{2i} \quad (8b)
\]

Also, let \( \text{SNR}_i = \frac{|h_i|^2}{i \in \{1,2\} \text{ and } \text{INR}_2 = \frac{|h_{12}|^2}{2} \). We have assumed that all \( h_{ij} \)'s are real and positive and, hence, use \( h_{ii} = \sqrt{\text{SNR}_i} \) and \( h_{12} = \sqrt{\text{INR}_2} \).

After two full rounds, every node in the network knows the complete network state, i.e., the matrix \( H \) is known completely to all four nodes. In this case, an upper bound on the capacity region is given in the following theorem.

**Theorem 4** ([4], [18]–[20]): The channel capacity region for a two-user Gaussian Z-channel is upper bounded by the region formed by the set of nonnegative rates \( (R_1, R_2) \) satisfying

\[
R_1 \leq \log (1 + \text{SNR}_1) \quad (9a)
\]

\[
R_2 \leq \log (1 + \text{SNR}_2) \quad (9b)
\]

If \( \text{INR}_2 \leq \text{SNR}_1 \),

\[
R_1 + R_2 \leq \log (1 + \text{SNR}_1) + \log \left(1 + \frac{\text{SNR}_2}{1 + \text{INR}_2}\right). \quad (10)
\]

If \( \text{INR}_2 \geq \text{SNR}_1 \),

\[
R_1 + R_2 \leq \log (1 + \text{SNR}_2 + \text{INR}_2). \quad (11)
\]

We now focus on 1.5 rounds of message passing in the Gaussian channel. In Section IV-A, we showed that for the deterministic Z-channel, the sum capacity can be achieved with 1.5 rounds of message passing. We now show that sum capacity within 2 bits can be achieved for a two-user Gaussian Z-channel with 1.5 rounds of message passing. Thus, there exist an approximate universality optimal strategy with \( \tau = 2 \).

**Theorem 5**: The sum capacity for a Gaussian two-user Z-channel can be achieved within 2 bits with the local information \( N_i \) and \( N'_i \) at the nodes obtained with 1.5 rounds of message passing and the side information about the network size, i.e., \( SI = \{K = 2\} \).

**Proof:** With 1.5 rounds of message passing, the two nodes know the network connectivity, i.e., if they are the first user (that causes interference to the other) or the second.

The second transmitter uses a codebook of rate

\[
R_2(\text{SNR}_2, \text{INR}_2) = \log \left(1 + \frac{\text{SNR}_2 + \text{INR}_2}{1 + 2 \text{INR}_2}\right), \quad (12)
\]

with a power level of \( P_2 = 1 \) to transmit. The first user, however, uses a common and a private message with rates

\[
R_{1,c}(\text{SNR}_1, \text{SNR}_2, \text{INR}_2) = \begin{cases} 
0 & \text{if } \text{INR}_2 < \text{SNR}_2 \\
\log (1 + \min \left(\frac{\text{SNR}_1 \cdot \text{INR}_2}{1 + \text{SNR}_1}, \frac{1}{1 + \text{SNR}_1} \right)) & \text{if } \text{INR}_2 \geq \text{SNR}_2, \text{INR}_2 \geq \text{SNR}_1 \\
\log (1 + 1 + \text{INR}_2 + \text{SNR}_1) & \text{if } \text{INR}_2 \geq \text{SNR}_2, \text{INR}_2 < \text{SNR}_1 \\
\log (1 + \text{SNR}_1 / (1 + \text{INR}_2)) & \text{otherwise} \end{cases} \quad (13)
\]

Furthermore, the power levels of \( P_{1,c}(\text{SNR}_1, \text{SNR}_2, \text{INR}_2) = \begin{cases} 
0 & \text{if } \text{INR}_2 < \text{SNR}_2 \\
1 & \text{if } \text{INR}_2 \geq \text{SNR}_2, \text{INR}_2 \geq \text{SNR}_1 \\
\frac{1}{1 + \text{INR}_2} & \text{if } \text{INR}_2 \geq \text{SNR}_2, \text{INR}_2 < \text{SNR}_1 \\
\frac{1}{1 + \text{INR}_2} & \text{otherwise} \end{cases} \quad (15)
\]

are used for the common and private parts, respectively. The common and private parts are added together and transmitted.

In Appendix B, we prove that the aforementioned rates can be decoded and yield the sum capacity within 2 bits.

**Corollary 2:** There exist an approximately universally optimal strategy with \( \tau = 2 \) with the local information \( N_i \) and \( N'_i \) provided by 1.5 rounds of message passing and the side information that \( K = 2 \).

**Proof:** With 1.5 rounds of message passing and side information \( K = 2 \), all nodes know the global connectivity. For all the connectivity choices, a strategy can be found by the nodes based on the channel gains they know that would achieve the global information sum capacity within 2 bits. For example, if the network is fully connected, then the nodes can use the node identities to order themselves and can use an appropriate (approximately) optimal strategy. If the network is a Z-channel, the strategy of Z-channel described in Theorem 5 can be used.

For a general Gaussian Z-channel, the capacity region is not known exactly. There exist achievable schemes that can achieve the region as shown in Fig. 2. The achievable point corresponding to the maximum \( R_1 \) when \( R_1 = \log (1 + h_{11}^2) \) is known exactly [20] but the maximal rate point corresponding to \( R_2 = \log (1 + h_{21}^2) \) is only known approximately. Here, we achieve a point approximate to the maximal rate point corresponding to \( R_2 = \log (1 + h_{21}^2) \).

While there is no loss in rate due to lack of knowledge in the deterministic Z-channel, it is no longer true for the case of the Gaussian Z-channel. For example, the exact capacity region is known for some regimes, like the strong interference case. However, the optimal sum rate point is not achieved by the distributed scheme with 1.5 rounds of knowledge. In this case, lack of full knowledge at \( T_2 \) implies that it is not aware that it is part.
of a strong interference channel. As a result, the $T_2$ backs off
a little on the rate thinking that it may not be able to cancel all
interference.

V. THREE-USER DOUBLE Z-CHANNEL

In this section, we will describe our results for three-user
double Z-channel. We will first find the capacity region for a
general class of deterministic channels in which the interference
at the receivers is a deterministic function of the inputs,
and the received signal is a deterministic function of the direct
signal and the interference and, therefore, specialize it to the
deterministic model of Section II. We then derive new genie-aided
outer bounds for the Gaussian double Z-channel. We will fur-
ther provide an achievable strategy with 1.5 rounds of the mes-
sage-passing algorithm that achieves the sum capacity in some
cases and has an unbounded gap in certain other cases. We also
show that the loss is unavoidable in the sense that there does
not exist any strategy with local information that can be uni-
versally optimal even with additional side information of network
connectivity $\mathcal{SI} = \mathcal{E}_{3Z}$ ($\mathcal{E}_{3Z}$ is the network connectivity
representing the three-user double Z-channel), which is more than
the side information $\mathcal{SI} = \{K = 3\}$. The achievability is extended
to the Gaussian model. Furthermore, we show that 1.5 round
strategy can be tweaked for the first user to derive a 2.5 strategy,
which is optimal for the deterministic case and is approximately
optimal in terms of the sum capacity for the Gaussian case with
side information $\mathcal{SI} = \{K = 3\}$. This also proves that there
exist a universally optimal/approximately universally optimal
strategy with the local information obtained by 2.5 rounds of
message passing and the side information of $\{K = 3\}$. That
is because with 2.5 rounds of messages, in all network connec-
tivities except double Z-channel, all the nodes would know the
whole state in 2.5 rounds and can hence take the same decision
as the centralized optimal solution.

A. Channel Models and Messaging Passing

In a deterministic double Z-channel, $K = 3$ and $n_{13} = n_{21} = n_{31} = n_{32} = 0$. Specifically, the received signal $Y_{ji}$,
$j = 1, 2, 3$, of a double Z-channel is given by

$$Y_{1i} = h_{11}X_{1i} + Z_{1i} \quad (17a)$$

$$Y_{2i} = h_{12}X_{1i} + h_{22}X_{2i} + Z_{2i} \quad (17b)$$

$$Y_{3i} = h_{13}X_{1i} + h_{23}X_{2i} + h_{33}X_{3i} + Z_{3i} \quad (17c)$$

In a Gaussian double Z-channel, and $h_{13} = h_{21} = h_{31} = h_{32} = 0$. The received signal $Y_{ji}$, $j = 1, 2, 3$, of a double
Z-channel is given by

$$Y_{1i} = h_{11}X_{1i} + Z_{1i} \quad (18a)$$

$$Y_{2i} = h_{12}X_{1i} + h_{22}X_{2i} + Z_{2i} \quad (18b)$$

$$Y_{3i} = h_{23}X_{2i} + h_{33}X_{3i} + Z_{3i} \quad (18c)$$

Also, let $\text{SNR}_i = |h_{ii}|^2$, $i \in \{1, 2, 3\}$ and $\text{INR}_{i+1} = |h_{i(i+1)}|^2$, $i \in \{1, 2\}$. Furthermore, we assume that all $h_{ij}$’s are real and
positive.

If the side information is $\mathcal{SI} = \{K = 3\}$, message passing
converges in four rounds. The details of the protocol are as fol-
 lows. (To simplify the notation, the node identities appended to
each channel gain are not shown and implied by the channel
subscripts.)

1) Round 1: $m_{k,1} = \psi_{k}$. $M_{1,1} = \{H_{11}\}$, $M_{2,1} = \{H_{12}, H_{22}\}$, and $M_{3,1} = \{H_{23}, H_{33}\}$.

2) Round 2: $m_{1,2} = \{H_{11}, H_{12}, H_{22}\}$, $m_{2,2} = \{H_{12}, H_{22}, H_{23}, H_{33}\}$, and $m_{3,2} = \emptyset$. $M_{1,2} = \emptyset$, $M_{2,2} = \{H_{11}, H_{23}, H_{33}\}$, and $M_{3,2} = \{H_{12}, H_{22}\}$.

3) Round 3: $m_{1,3} = \{H_{23}, H_{33}\}$, $m_{2,3} = \{H_{11}\}$, and $m_{3,3} = \emptyset$. $M_{1,3} = M_{2,3} = \emptyset$, and $M_{3,3} = \{H_{11}\}$.

4) Round 4: No new information is to be sent by any trans-
mitter, and hence, the algorithm halts by transmitters
sending a silent message $\emptyset$.

Note that $H_{ij}$ is replaced with $n_{ij}$ in the deterministic or $h_{ij}^2$
for the Gaussian model.

The local view of each node after 1.5 and 2.5 rounds of
message passing is shown in Fig. 3. It is clear that with fewer rounds
of message passing, the local view of each node is less than full.
The challenge for the node is that they have to make decisions
on their transmission parameters based only on their local view.
Thus, we will ask how close we can get to sum capacity with
1.5 and 2.5 rounds of message passing. Before we derive sum
capacity with partial information, we will need the full capacity
region for the case of full information. In the next section, we
derive the new capacity results.

B. Capacity Regions With Full Information

The capacity region for a general three-user interference
channel is open. It has been solved in certain special cases of
deterministic model in [11]–[13]. In this section, we provide
the capacity region for a three-user double Z-channel, which
has not been considered before.

Theorem 6 (Double Z Deterministic Channel Capacity Re-
 gion): The deterministic channel capacity region for a three-
user double Z interference channel is the set of nonnegative rates
$(R_1, R_2, R_3)$ satisfying

$$R_i \leq n_{ii}, \ i = 1, 2, 3,$$

$$R_1 + R_2 \leq \max(n_{11}, n_{12}, n_{22}, n_{11} + n_{22} - n_{12}),$$

$$R_2 + R_3 \leq \max(n_{22}, n_{23}, n_{33}, n_{22} + n_{33} - n_{23}),$$

$$R_1 + R_2 + R_3 \leq \max(n_{33}, n_{23} + (n_{11} - n_{12})^+ + \max(n_{12}, n_{22} - n_{23}).$$

Proof: The proof is provided in Appendix C for a class of
deterministic channels, which is much broader than the deter-
mministic model, along the lines of the class of deterministic

Fig. 2. Outer and inner bounds in the Gaussian Z-channel.
channels for two user interference channels proposed in [14]. This region can be specialized for the deterministic model in (17a)–(17c) to obtain the aforementioned result.

Note that the bounds on $R_i$ are the single-user bounds, and the $R_i + R_j$ bounds are due to the two two-user $Z$-channels, one consisting of $\{T_1, T_2, D_1, D_2\}$ and the other consisting of $\{T_2, T_3, D_2, D_3\}$. Finally, the sum bound $R_1 + R_2 + R_3$ is due to the common transmitter–receiver pair $\{T_2, D_2\}$ in the two $Z$-channels.

For the Gaussian channel, we provide an outer bound to the capacity region for the three-user $Z$-channel. We divide the region of the channels to four cases depicting the strong/weak interference from the first two transmitters.

1) $\text{INR}_2 \geq \text{SNR}_1$ and $\text{INR}_3 \geq \text{SNR}_2$: In this case, an outer bound on the rate region is given as follows:

$$R_1 \leq \log(1 + \text{SNR}_1)$$
$$R_2 \leq \log(1 + \text{SNR}_2)$$
$$R_3 \leq \log(1 + \text{SNR}_3)$$
$$R_1 + R_2 \leq \log(1 + \text{SNR}_2 + \text{INR}_2)$$
$$R_2 + R_3 \leq \log(1 + \text{SNR}_3 + \text{INR}_3).$$

2) $\text{INR}_2 \geq \text{SNR}_1$ and $\text{INR}_3 \leq \text{SNR}_2$: In this case, an outer bound on the rate region is given as follows:

$$R_1 \leq \log(1 + \text{SNR}_1)$$
$$R_2 \leq \log(1 + \text{SNR}_2)$$
$$R_3 \leq \log(1 + \text{SNR}_3)$$
$$R_1 + R_2 \leq \log(1 + \text{SNR}_2 + \text{INR}_2)$$
$$R_2 + R_3 \leq \log(1 + \text{SNR}_2 + \text{INR}_3).$$

Further, if $(\text{INR}_2 + 1)\text{INR}_3 \leq \text{SNR}_2$

$$R_1 + R_2 + R_3 \leq \log(1 + \text{SNR}_1) + \log(1 + \text{SNR}_2 + \text{INR}_3)$$

else if $(\text{INR}_2 + 1)\text{INR}_3 \geq \text{SNR}_2$

$$R_1 + R_2 + R_3 \leq \log(1 + \text{SNR}_1 + \text{INR}_3) + \log(1 + \text{INR}_2).$$

3) $\text{INR}_2 \leq \text{SNR}_1$ and $\text{INR}_3 \geq \text{SNR}_2$: In this case, an outer bound on the rate region is given as follows:

$$R_1 \leq \log(1 + \text{SNR}_1)$$
$$R_2 \leq \log(1 + \text{SNR}_2)$$
$$R_3 \leq \log(1 + \text{SNR}_3)$$
$$R_1 + R_2 \leq \log(1 + \text{SNR}_1) + \log(1 + \text{SNR}_2 + \text{INR}_2)$$
$$R_2 + R_3 \leq \log(1 + \text{SNR}_3 + \text{INR}_3).$$

4) $\text{INR}_2 \leq \text{SNR}_1$ and $\text{INR}_3 \leq \text{SNR}_2$: In this case, an outer bound on the rate region is given as follows:

$$R_1 \leq \log(1 + \text{SNR}_1)$$
$$R_2 \leq \log(1 + \text{SNR}_2)$$
$$R_3 \leq \log(1 + \text{SNR}_3)$$
$$R_1 + R_2 \leq \log(1 + \text{SNR}_1) + \log(1 + \text{SNR}_2 + \text{INR}_2)$$
$$R_2 + R_3 \leq \log(1 + \text{SNR}_3 + \text{INR}_3).$$

Furthermore, if $(\text{INR}_2 + 1)\text{INR}_3 \leq \text{SNR}_2$

$$R_1 + R_2 + R_3 \leq \log(1 + \text{SNR}_1) + \log(1 + \text{SNR}_2 + \text{INR}_3) + \log(1 + \frac{\text{SNR}_2}{1 + \text{INR}_2})$$

else if $(\text{INR}_2 + 1)\text{INR}_3 \geq \text{SNR}_2$

$$R_1 + R_2 + R_3 \leq \log(1 + \text{SNR}_1) + \log(1 + \text{INR}_3 + \text{SNR}_3).$$

Theorem 7 (Double Z Gaussian Channel Outer Bound):
The capacity region of the Gaussian double $Z$-channel is outer bounded by the region formed by $(R_1, R_2, R_3)$ satisfying (20)–(27).

Proof: All the single-user bounds and the bounds on $R_1 + R_2$ and $R_2 + R_3$ follow from the two-user $Z$-channel. The new
bounds for $R_1 + R_2 + R_3$ in the second and the fourth cases are new and given in Appendix D.

We note that in the case of weak interference for the two $Z$-channels $(T_1, T_2, R_1, R_2)$ and $(T_2, T_3, R_2, R_3)$, it is not always optimal to treat interference as noise unlike the case in the two-user $Z$-channel. However, there is a region in which the interference is very weak, which is when $\text{INR}_3(\text{INR}_2 + 1) \leq \text{SNR}_2$, where treating interference as noise is optimal.

**Lemma 1:** Let $K \geq 3$ and consider a symmetric $K$-user $Z$-channel, where $(K-1)$ $Z$-channels are stacked one over the other, with $\text{SNR}_i = \text{SNR}$ and $\text{INR}_i = \text{INR}$. When $\text{INR}(\text{INR} + 1) \leq \text{SNR}$, the sum rate is outer bounded by $\log(1 + \text{SNR}) + (K-1) \log \left( 1 + \frac{\text{SNR}}{1 + \text{INR}} \right)$. Thus, in this regime, the sum capacity can be achieved by treating interference as noise. However, as can be seen in the case of three-user system, the previous fact is not true in general for $\text{INR} < \text{SNR} < \text{INR}(\text{INR} + 1))$.

**Proof:** This proof follows using the same techniques as in the special case of very weak interference in Appendix D and is thus omitted.

C. Deterministic Model: 1.5 Rounds

In this section, we will study the achievable sum rate after 1.5 rounds of message passing. Recall that the local view of each node is given in Fig. 3(a) and the nodes have to base their choice of rates and transmission strategy only on their local views. We first describe an achievable rate, building on the $Z$-channel allocation strategy discussed in Theorem 2. Note that to derive the achievable rate, we will only assume side information about the network size, i.e., $K = 3$. However, we will show the converse for 1.5 rounds with extra side information about connectivity $\mathcal{S}_1 = \{E_{12Z}\}$, which makes our converse stronger than one with only the network size as side information.

1) From the point of view of the first transmitter, it knows the upper part of the $Z$-channel $(n_{11}, n_{12}, n_{22})$ and only extra cross-link possible is a link from the second transmitter to the third receiver. Thus, the first transmitter knows that it is a $Z$-channel with the third user alone or a double $Z$-channel. Since $T_3$ does not know which of the two cases are applicable, it assumes the worst case that second transmitter $T_2$ will send at full rate and acts as if it is a $Z$-channel consisting of the first two users only. Note that this strategy is optimal if network connectivity turned out a $Z$-channel with a decoupled transmitter $T_3$, instead of a double $Z$-channel. Furthermore, this rate can be decoded at the receiver even if it is a double $Z$-channel.

2) From the point of the second transmitter, it knows $n_{12}, n_{22}, n_{23},$ and $n_{33}$. Hence, the only cross-link possible that it does not know is $n_{33}$. Note that while $n_{33}$ is actually zero, the second transmitter does not know about it. Now suppose that the second transmitter sends as if there are only two users (the second and the third) and hence functions as the upper user of the $Z$-channel. It assumes that if it is a double $Z$, the first transmitter $T_1$ will send as if the second transmitter $T_2$ is sending at full rate and thus will take care of interference itself. Hence, the data can be decoded at the second receiver $D_2$. If the network connectivity is cyclic with the presence of $n_{33}$, then every user will use this strategy and back off, and thus, the data can still be decoded.

3) The third transmitter $T_3$ knows $n_{33}$ and $n_{23}$. Thus, the cross-channel gains that the third transmitter does not know about are $n_{12}$ and $n_{21}$. Even though $n_{21} = 0$, the third transmitter does not know about it. The third transmitter transmits at $n_{33}$. In case both the links $n_{12}$ and $n_{21}$ are present, the second transmitter knows the whole state and can shut down or select the rate since it would know the strategy of other users. In case only $n_{21}$ is present, then the second transmitter knows it is the one-to-many configuration and there exist an optimal strategy in this case where the first and the third user send at full rate, while the second transmitter backs off. The only other case is the double $Z$-channel, where we use this strategy and show that it is still achievable.

To summarize, this achievability scheme reduces for a double $Z$-channel to using the same strategy at all the users as in the two-user $Z$-channel with the relevant knowledge. More specifically, the first transmitter assumes that it is the top transmitter in a two-user $Z$-channel (consisting of the first and the second user), the second transmitter assumes that it is the top user of a two-user $Z$-channel (consisting of the second and the third user) and the third transmitter assumes that it is the bottom user of a two-user $Z$-channel (consisting of the second and the third user).

**Theorem 8 (Achievable Rate With 1.5 Rounds):** The aforementioned scheme can achieve the following sum rate with 1.5 rounds of message passing for a double $Z$-channel:

$$
\min\left(\max(n_{22}, n_{23}, n_{33}, n_{22} + n_{33} - n_{23}), n_{22} + n_{33}\right)
+ \min\left(n_{11}, \max(n_{11}, n_{12}) - \min(n_{12}, n_{22})\right).
$$

(28)

**Proof:** We show that each transmitter uses network-state information obtained from the first round to decide the transmission strategy. The third transmitter sends at full rate, $n_{33}$. The second transmitter does not know $n_{11}$ and uses the strategy as if it was a $Z$-channel consisting of only the second and the third user. If $n_{23} \leq n_{33}$, the second transmitter will send at a rate of $(n_{22} - n_{23})$. If $n_{23} > n_{33}$, the second transmitter sends at a rate of \(\min\left(\max(n_{22}, n_{23}) - n_{33}, n_{22}\right)\). Thus, the second transmitter sends at a rate of \(\min\left(n_{22}, \max(n_{22}, n_{23}) - \min(n_{23}, n_{33})\right) = \min\left(\max(n_{22}, n_{23}, n_{33}, n_{22} + n_{33} - n_{23}), n_{22} + n_{33}\right) - n_{33}\). The first transmitter transmits as if it was a $Z$-channel consisting of first two users and considering that the second user sends at a rate of $n_{22}$ and, hence, sends at a rate of $\min(n_{11}, \max(n_{11}, n_{12}) - \min(n_{12}, n_{22}))$. Hence, the aforementioned sum rate can be achieved.

So, the obvious next question is how well does the aforementioned scheme perform compared to the sum capacity in Theorem 6. We show that the gap from the sum capacity can be anywhere from zero to arbitrarily large. First, we classify all those network states in which the aforementioned distributed scheme achieves the sum capacity.
Theorem 9 (Achieving Sum Capacity): The sum capacity can be achieved with 1.5 rounds of message passing for a double Z-channel if any of the following are true:

1) \( n_{23} \geq n_{22} + n_{33} \)
2) \( n_{12} \geq n_{11} + n_{22} \)
3) \( n_{23} \leq n_{33} \) and \( n_{22} \geq n_{23} + n_{12} \).

Proof:
1) \( n_{23} \geq n_{22} + n_{33} \): In this case, (28) reduces to the sum rate of \( n_{33} + \min(n_{11} + n_{22}, n_{23}, n_{12}, n_{11} + n_{22} - n_{12}) \), which is optimal since the upper bounds on \( R_3 \) and \( R_1 + R_2 \) in Theorem 6 are identical to the aforementioned expression.
2) \( n_{12} \geq n_{11} + n_{22} \): In this case, (28) reduces to the sum rate of \( n_{11} + \max(n_{22}, n_{23}, n_{33}, n_{22} + n_{33} - n_{23}) \), which is optimal since it is identical to upper bounds on \( R_1 \) and \( R_2 + R_3 \) in Theorem 6.
3) \( n_{23} \leq n_{33} \) and \( n_{22} \geq n_{23} + n_{12} \): In this case, the achievable sum rate and the outer bound both reduce to \( n_{22} + n_{33} - n_{23} + (n_{11} - n_{12}) \). Hence, the sum rate is optimal since this matches the \( R_1 + R_2 + R_3 \) outer bound in Theorem 6.

We next construct an example where the loss can be arbitrarily large. Let \( n_{11} = n_{12} = n_{22} = n_{23} = n_{33} = x \). The distributed scheme achieves a sum rate of \( x \) related to rate tuple of \((0,0,x)\) as shown in Fig. 4(a). However, the rate tuple \( (x,0,x) \) is in the capacity region and relates to the rate allocation shown in Fig. 4(b). By taking \( x \) large enough, the achievable sum rate can be made arbitrarily far from the outer bound. Thus, the aforementioned distributed strategy is not universally optimal, i.e., it does not achieve the sum capacity in all network states. The next result shows that no distributed scheme can be universally optimal with 1.5 rounds of knowledge, even with extra side information.

Theorem 10 (Loss is Inevitable): There exist no universally optimal strategy with the local information obtained by 1.5 rounds of message passing for the double Z-channel even with the side information \( S_I = E_{3Z} \), the network connectivity of the double Z-channel. Thus, there is no universally optimal strategy with the local information obtained by 1.5 rounds of message passing and the side information of \( K = 3 \).

Proof: We first assume that the each node is given the information that the network connectivity is a double Z-channel and the information about its relative placement in the network. So, the nodes can choose strategy tailored for the double Z-channel. Thus, the set of network states consistent with the local information only have \( n_{11}, n_{22}, n_{33}, n_{12}, \) and \( n_{23} \) as parameters.

We will now prove the theorem by contradiction. Suppose that there is a universally optimal strategy. Consider the third transmitter’s rate and codebook selection with local information. The third transmitter does not know \( n_{12} \) and \( n_{31} \) and has to come up with a scheme that achieves the sum capacity with global knowledge irrespective of the values of unknown channel gains. If \( n_{22} = n_{11} = 0 \), the third transmitter has to chose \( R_3 = n_{33} \) to achieve full information sum capacity. Hence, if there exists a universally optimal strategy, \( R_3 = n_{33} \). (Even though the transmitter knows \( n_{23} \), it cannot use this extra knowledge to make a decision on the rate allocation.)

The second transmitter does not know \( n_{11} \). Hence, if there exist an universally optimal strategy, it should work even if \( n_{11} = 0 \). For \( n_{11} = 0 \) and \( R_3 = n_{33} \), the only way the second user can use optimal rate allocation is to transmit at a rate as in the proof of Theorem 8.

The first transmitter does not know \( n_{33} \). Hence, its optimal strategy should work even if \( n_{33} = 0 \). If \( n_{33} = 0 \), \( R_2 = n_{22} \), and hence, the first user will have to transmit at a rate as in Theorem 8.

Thus, we see that if there exist an universally optimal strategy, the rates of the users have to be the same as in Theorem 8. We note that the sum rate in Theorem 8 is not optimal in general and, hence, leads to a contradiction. Thus, there is no universally optimal strategy.

Theorem 10 is the key result in this paper. It shows that no distributed scheme, which only relies on its local view after 1.5 rounds, can be guaranteed to be globally optimal for all channel conditions even with the global knowledge of connectivity. Distributed schemes can be optimal for some values of network matrices \( H \in G_{E_{3Z}} \) but not all of them simultaneously. Thus, there is no universally optimal scheme with the local information provided by 1.5 rounds of knowledge and the side information of network connectivity.

The fundamental reason for this unavoidable loss for some channel gains is severely incomplete view of the network at different transmitters. As shown in Fig. 4, the loss in spatial reuse is due to mismatched knowledge in the deterministic double Z-channel. Here, \( T_1 \) is backing off for \( T_2 \), which in turn is backing off for \( T_3 \). In this example, \( T_1 \) could have sent at full rate but ended up being too conservative due to its lack of knowledge about the state of \( T_3 \), as shown in Fig. 4(b). As a result, \( T_1 \) tailors its action to the worst-case scenario, which is \( T_2 \) sending at full rate. The reader is reminded that each node is only allowed to adapt its transmission based on its own local knowledge and cannot base its decision on what is not known. Thus, once a node defines a rate-allocation policy based on its local information (e.g., 1.5 rounds), it has to use the same allocation for all the possible values of other channel gains that are not known with the local and side information.

We observe that the previous proof can also be extended to show that there does not exist any strategy that will perform within a bounded gap (independent of channel gains) from the optimal sum capacity. Thus, an approximately universally optimal strategy also does not exist. To prove that claim, assume \( n_{ij} = c_{ij}L \), where \( L \) can be taken as large as possible, \( c_{ij} \geq 0 \). With the local channel knowledge at the nodes, \( R_3 \geq c_{ij}L - \Theta(1) \) (where \( \Theta(1) \) represents a function that is independent of \( L \)) since that is the only way an approximately optimal strategy
can exist when $c_{22} = c_{11} = 0$. Given this strategy of the third user, $R_2 \geq 1 - \min(c_{22}, c_{33}, c_{22}, c_{33}, c_{22} + c_{33}) - 1$ since otherwise the strategy will not be approximately universally optimal for $c_{11} = 0$. A similar strategy goes for the first user. We can also show that this strategy will be the unbounded rate away from optimal as $L \rightarrow \infty$, thus proving the nonexistence of approximately universally optimal strategies.

D. Gaussian Channel: 1.5 Rounds

The achievability strategy for Gaussian channels follows the same technique as in the deterministic model. As mentioned for the deterministic model, the achievable scheme used by each of the users will work as an achievable strategy for all possible global network connectivities if they see the same local network connectivity. The details are as follows.

The third transmitter makes a codebook of rate

$$R_3(SNR_3, INR_3) = \log(1 + SNR_3 / (1 + 2 INR_3))$$

and uses a power level of $P_3 = 1$ to transmit. The second user, however, uses a common and a private message with rates

$$R_{2, c}(SNR_2, SNR_3, INR_3) = \begin{cases} 0 & \text{if } INR_3 < SNR_3 \\ \log \left( 1 + \frac{1 + INR_3}{1 + INR_3 + SNR_3} \min \left( SNR_2, \frac{\text{INR}_2}{1 + \text{INR}_2} \right) \right) & \text{if } INR_3 \geq SNR_3, INR_3 \geq SNR_2 \\ \log \left( 1 + \frac{1 + INR_3}{1 + INR_3 + SNR_3} \min \left( \frac{\text{INR}_2}{1 + \text{INR}_2}, \frac{\text{INR}_3}{1 + \text{INR}_3} \right) \right) & \text{if } INR_3 \geq SNR_3, INR_3 < SNR_2 \end{cases}$$

Furthermore, the power levels of

$$P_{2, c}(SNR_2, SNR_3, INR_3) = \begin{cases} 0 & \text{if } INR_3 < SNR_3 \\ \frac{INR_3}{1 + \text{INR}_3} & \text{if } INR_3 \geq SNR_3, INR_3 \geq SNR_2 \\ \frac{INR_3}{1 + \text{INR}_3} & \text{if } INR_3 \geq SNR_3, INR_3 < SNR_2 \end{cases}$$

are used for the common and private parts. The common and the private parts are added together and transmitted.

The first user uses a common and a private message with rates

$$R_{1, c}(SNR_1, SNR_2, INR_2) = \begin{cases} 0 & \text{if } INR_2 < SNR_2 \\ \log \left( 1 + \min(SNR_1, \frac{\text{INR}_1}{1 + \text{INR}_1}) \right) & \text{if } INR_2 \geq SNR_2, INR_2 \geq SNR_1 \\ \log \left( 1 + \frac{1 + \text{INR}_2}{1 + \text{INR}_2 + SNR_2} \right) & \text{if } INR_2 \geq SNR_2, INR_2 < SNR_1 \end{cases}$$

Furthermore, the power levels of

$$P_{1, c}(SNR_1, SNR_2, INR_2) = \begin{cases} 0 & \text{if } INR_2 < SNR_2 \\ \frac{INR_2}{1 + \text{INR}_2} & \text{if } INR_2 \geq SNR_2, INR_2 \geq SNR_1 \\ \frac{INR_2}{1 + \text{INR}_2} & \text{if } INR_2 \geq SNR_2, INR_2 < SNR_1 \end{cases}$$

are used for the common and private parts. The common and the private parts are added together and transmitted.

We will show in Appendix E that the aforementioned rates can be decoded by the receivers. We now see the various cases when the achievable strategy will be bounded distance from the sum capacity.

**Theorem 11:** The sum capacity within 4 bits can be achieved with the first full round and half of the second round of message passing when any of the following conditions holds:

1) $\text{INR}_3 \geq \text{SNR}_3$, $\frac{\text{INR}_2}{1 + \text{INR}_2} \geq \text{SNR}_2$

2) $\text{INR}_2 \geq \text{SNR}_2$, $\frac{\text{INR}_3}{1 + \text{INR}_3} \geq \text{SNR}_1$

3) $\text{INR}_2 < \text{SNR}_2$ and $\text{INR}_3 < \text{SNR}_3$, $\text{SNR}_2 \geq \text{INR}_3(\text{INR}_2 + 1)$

**Proof:**

1) $\text{INR}_3 \geq \text{SNR}_3$, $\frac{\text{INR}_2}{1 + \text{INR}_2} \geq \text{SNR}_2$. In this case, $R_2 = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + \text{INR}_2 + \text{SNR}_2} \right)$ which yields $R_1 + R_2$ within 2 bits of optimal as in the two-user Z-channel and since $R_3$ is within 1 bit of $\log(1 + \text{SNR}_3)$, and $R_1 + R_2 + R_3$ is within 3 bits of the outer bound.

2) $\text{INR}_2 \geq \text{SNR}_2$, $\frac{\text{INR}_3}{1 + \text{INR}_3} \geq \text{SNR}_1$. In this case, $R_2 = \log(1 + \text{INR}_3)$. Furthermore, $R_2 + R_3$ is within 2 bits of that in the case of two-user Z-channel, which is further within 2 bits of optimal; thus leading to the sum rate within 4 bits of optimal.

3) $\text{INR}_2 < \text{SNR}_2$ and $\text{INR}_3 < \text{SNR}_3$, $\text{SNR}_2 \geq \text{INR}_3(\text{INR}_2 + 1)$. The achievable sum rate in this case reduces to

$$\log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_2} \right) + \log \left( 1 + \frac{\text{INR}_2}{1 + 2 \text{INR}_2} \right)$$

If $\text{INR}_2 \geq \text{SNR}_3$, the outer bound is $\log(1 + \text{SNR}_2 + \text{INR}_2) + \log \left( 1 + \frac{\text{INR}_2}{1 + \text{INR}_2} \right)$. However, if $\text{INR}_2 < \text{SNR}_3$, the outer bound is $\log(1 + \text{SNR}_2 + \text{INR}_2) + \log \left( 1 + \frac{\text{INR}_2}{1 + \text{INR}_2} \right) + \log \left( 1 + \frac{\text{INR}_2}{1 + \text{INR}_2} \right)$. We note that in both these cases, the achievability is within 4 bits in both the cases.

As in the deterministic model, there are cases when the sum rate achieved with 1.5 rounds can be arbitrarily far from the optimal. As an example, consider $\text{INR}_2 = \text{SNR}_1 = \text{SNR}_2 = \text{INR}_3 = \text{SNR}_3 = x$. The achievable sum rate is $\log(1 + 6x + 11x^2 + 5x^3) / (1 + 3x + x^2)$. However, with full information, rate pair of $(\log(1 + x), 0, \log(1 + x))$ can be achieved. For $x \geq 2$,
log[(1 + 6x + 11x^2 + 5x^3)/(1 + 3x + x^2)] ≤ 2log(1 + x) and the difference grows unbounded with x thus proving that the difference between achievability and outer bound can be unbounded in some cases.

We also note that the strong converse mentioned for the deterministic case also holds in the Gaussian case that there is no approximately universally optimal strategy with 1.5 rounds of message passing. Since the proof uses the same ideas, it is omitted.

E. Both Channels With 2.5 Rounds

We first note that with 2.5 rounds of message passing, all the nodes know the state information except the third transmitter which does not know $H_{13}$. Thus, the side information of $\mathbf{SI} = \{K = 3\}$ with 2.5 rounds of message passing and the side information of network connectivity, $\mathbf{SI} = \{E\}$, with 2.5 rounds of message passing are equivalent.

We first consider the capacity region of the deterministic double Z-channel as shown in Fig. 5. The region consists of single-user constraints, the constraints on $R_1 + R_2, R_2 + R_3$, and a constraint on $R_1 + R_2 + R_3$ depicting the various planes in the figure. Note that there is a segment on the optimal face of $R_1 + R_2 + R_3$ that has $R_3 = \gamma_{33}$ as marked in the figure. Any point on this segment can be achieved with 2.5 rounds of message passing. This is because the third transmitter sends at a rate of $\gamma_{33}$, while the first two transmitters know all the channel gains to select the policy to operate at any point on this line.

Thus, there exist universally optimal strategy with the local information at the nodes obtained by 2.5 rounds of message passing and the side information about network size $\mathbf{SI} = \{K = 3\}$. So with 2.5 rounds, we can prove existence of a universally optimal strategy with less side information compared to the converse for 1.5 rounds in Theorem 10.

**Theorem 12:** There exists a universally optimal strategy with the local information at the nodes obtained by 2.5 rounds of message passing when each node is provided the side information that there are only three nodes in the network.

$$R_{1,c} = \begin{cases} \min(\log(1 + \text{SNR}_3), \log(1 + \text{INR}_2 + \text{SNR}_2) - R_2) & \text{if } \gg \gg \gg \\ \min(\log(1 + \text{SNR}_3), \log(1 + \text{INR}_2) - R_2) & \text{if } \gg \gg \gg < \\ \min(\log(1 + \frac{\text{INR}_2}{1 + \text{SNR}_2 + \text{INR}_3}), \log(1 + \frac{\text{INR}_2}{1 + \text{SNR}_2 + \text{INR}_3}}) - R_2) & \text{if } \gg \gg \gg < \\ \min(\log(1 + \frac{\text{INR}_2}{1 + \text{SNR}_2 + \text{INR}_3}) - R_2) & \text{if } \gg \gg \gg < \\ \log(1 + \text{INR}_2 + \frac{\text{SNR}_2}{1 + \text{INR}_2} - R_{2,p}) & \text{if } \gg \gg \gg < \end{cases}$$

**Proof:** With 2.5 rounds of message passing, all the nodes would know the network connectivity in their connected component. In all the cases except the double Z-channel, all the nodes would know all the channel gains of the connected component in which they are and, hence, can use an optimal strategy by ordering the nodes based on the node identities and the network states. For the double Z-channel connectivity, the strategy described before can be used to obtain a universally optimal strategy.

We now consider the Gaussian channel model. As can be seen in the deterministic model that any strategy following the rate allocation on a line is optimal, we consider one of the corner points on the line corresponding to higher $R_2$ and provide an achievability strategy that is within 4 bits of sum capacity for 2.5 rounds of message passing. In this strategy, the second and the third user use the same policy as with 1.5 rounds and, therefore, do not change the strategy. However, the first transmitter knows all the channels and, hence, changes its strategy.

The first user uses a common and a private message with rates as given in (38), shown at the bottom of the page, where the four inequalities in the $R$ condition represents the order in $\text{INR}_2$ versus $\text{SNR}_2$, $\text{INR}_3$ versus $\text{SNR}_3$, $\text{INR}_2$ versus $\text{SNR}_1$, and $\text{INR}_3$ versus $\text{SNR}_2$, respectively. $\gg \gg$ represents that it can be either
side of the inequality. For example, \( \leq \leq \geq \) represents \( \text{INR}_2 < \text{SNR}_2, \text{INR}_3 \geq \text{SNR}_3, \text{INR}_2 < \text{SNR}_3, \) and \( \text{INR}_3 \geq \text{SNR}_2. \)

\[
R_{1,p} = \begin{cases} 
0 & \text{if } \text{INR}_2 \geq \text{SNR}_1 \\
\log(1 + \text{SNR}_1/(1 + \text{INR}_2)) & \text{if } \text{INR}_2 < \text{SNR}_1 
\end{cases} 
(39)
\]

Furthermore, the power levels of

\[
P_{1,c} = \begin{cases} 
1 & \text{if } \text{INR}_2 \geq \text{SNR}_1 \\
\text{INR}_2/(1 + \text{INR}_2) & \text{if } \text{INR}_2 < \text{SNR}_1 
\end{cases}
(40)
\]

\[
P_{1,p} = \begin{cases} 
0 & \text{if } \text{INR}_2 \geq \text{SNR}_1 \\
1/(1 + \text{INR}_2) & \text{if } \text{INR}_2 < \text{SNR}_1 
\end{cases}
(41)
\]

are used for the common and private parts. The common and private parts are added together and transmitted. It is straightforward to see in all the cases that the aforementioned rates can be decoded by all the users. Furthermore, as shown in the next theorem, the aforementioned strategy achieves the sum capacity within 4 bits.

**Theorem 13:** The aforementioned strategy achieves a sum rate that is within 4 bits of the full-information sum capacity for all \( H \in \mathcal{G}(E_{32}). \)

**Proof:** We will show in Appendix F that the sum capacity within 4 bits can be achieved by splitting the channel gain regimes into 14 cases.

**Corollary 3:** There exists an approximately universally optimal strategy with \( \tau = 4 \) with the local information at the nodes obtained by 2.5 rounds of message passing when each node is provided the side information that there are only three nodes in the network, i.e., \( \mathcal{SI} = \{K = 3\}. \)

**Proof:** With 2.5 rounds of message passing, all the nodes would know the network connectivity in their connected component. In all the cases except the double Z-channel, all the nodes would know all the channel gains of the connected component in which they are and, hence, can use an optimal strategy by ordering the nodes based on the node identities and the network states. For the double Z-channel connectivity, the strategy described in this section can be used to obtain the approximately universally optimal strategy.

**VI. CONCLUSIONS**

Almost all networks operate with partial network information at different nodes, requiring nodes to make distributed decisions. While a rich literature exists on design of network protocols and their analysis, there is no prior work to understand the impact of distributed decisions on the Shannon-theoretic capacity region. In this paper, we laid foundation to characterize partial network information and studied the impact in several network connectivities. Seeking universal optimality, where local decisions with certain side information are always globally optimal, we discovered that there appears to be a critical minimum information required for the network to allow globally optimal decisions. Our current approach is compound capacity based and our next step is to understand impact of partial information on fading interference channels.

![Fig. 6](image1.png)

*Fig. 6.* In this case, \( n_{12} \leq n_{22} \). The bold lines denote the active bits.

![Fig. 7](image2.png)

*Fig. 7.* In this case, \( n_{12} > n_{22} \). The bold lines denote the active bits.

**APPENDIX A**

**PROOF OF THEOREM 2**

We note that at the end of first round, the first transmitter knows all the channel gains, while the second transmitter does not know one of the channel gain \( n_{11} \). Let the strategy that the transmitter uses be as follows.

1) Transmitter 2, which is not producing interference, sends it maximum possible rate of \( n_{22}. \)

2) Transmitter 1 assumes that transmitter 2 is sending at full rate. Thus, transmitter 1 sends at a rate of \( (n_{11} - n_{12})^+ \) if \( n_{12} \leq n_{22} \) thus not sending on any link that produces interference. However, it sends at a rate of \( \min(\max(n_{11}, n_{12} - n_{22}, n_{11}) \) if \( n_{12} > n_{22} \), transmitting at the noninterfering links to the signal of the second transmitter communicating at the rate \( n_{22}. \)

We now show that this strategy can achieve the sum rate as follows.

If \( n_{12} \leq n_{22} \), the sum rate in \( (7c) \) simplifies as

\[
R_1 + R_2 \leq (n_{11} - n_{12})^+ + n_{22}. \quad (42)
\]

Hence, the first transmitter will send at a rate of \( (n_{11} - n_{12})^+. \)

Furthermore, since the first transmitter knows \( n_{11} \) and \( n_{22} \), it can send data on the links at which it is not generating any interference and, thus, achieve the sum capacity (see Fig. 6).

Now consider the case when \( n_{12} > n_{22} \), in which case the sum rate in \( (7c) \) simplifies as

\[
R_1 + R_2 \leq \min(\max(n_{11}, n_{12} - n_{22}, n_{11} + n_{22}), \quad (43)
\]

Thus, the first transmitter sends at a rate of \( \min(\max(n_{11}, n_{12} - n_{22}, n_{11}) = (n_{11} - n_{12})^+ + \min(n_{12} - n_{22}, n_{11}). \)

This is achieved by sending along the dimensions that cannot even be heard at receiver 2, which are \((n_{11} - n_{12})^+\) in number. In addition, among the dimensions that can be heard by the second receiver, the data are sent along the dimensions which do not produce an interference to the direct signal. These are \(\min(n_{12} - n_{22}, n_{11})\) in number; see Fig. 7. The extra half round is required for the receiver to learn the strategy used by the transmitters. Thus, the maximum sum rate can be achieved with 1.5 rounds.
APPENDIX B
PROOF OF THEOREM 5

A) The rates mentioned in (12)–(16) can be decoded at the receivers: In this section, we will show that the rates mentioned in (12)–(16) can be decoded at the receivers. In order to see the decoding process at the two users, consider the following scenarios,

1) $\text{INR}_2 < \text{SNR}_2$. The first receiver will be able to decode the data as the rate is supported by the power sent. The second receiver treats the first user’s signal as interference. The interference power is $\text{INR}_2/(1 + \text{INR}_2)$. Thus, the second receiver is able to decode its data treating this power as interference.

2) $\text{INR}_2 \geq \text{SNR}_2, \text{INR}_2 \geq \text{SNR}_1$. Since the received power from the signal of the first transmitter is $\text{SNR}_3$ and rate $\leq \log(1 + \text{SNR}_3)$, the first receiver is able to decode. At the second receiver, the two signals are decoded jointly. We have to verify

$$R_1 \leq \log(1 + \text{INR}_2) \quad (44a)$$

$$R_2 \leq \log(1 + \text{SNR}_2) \quad (44b)$$

$$R_1 + R_2 \leq \log(1 + \text{INR}_2 + \text{SNR}_2). \quad (44c)$$

The first two hold. For the third,

$$R_1 + R_2 \leq \log \left( 1 + \frac{\text{INR}_2}{1 + \text{SNR}_2 + \text{INR}_2} \right)$$

$$+ \log(1 + \text{SNR}_2/(1 + \text{SNR}_2 + 2\text{SNR}_2))$$

$$= \log(1 + \text{INR}_2 + \text{SNR}_2).$$

Hence proved.

3) $\text{INR}_2 \geq \text{SNR}_2, \text{INR}_2 < \text{SNR}_1$. In this case, there is both a public and a private message from the first user. At the first receiver, we need to decode the public data treating the other as noise. Furthermore, the private data can be decoded. To check the first, we need to see

$$R_{1,e} \leq \log \left( 1 + \frac{\text{SNR}_1 \text{INR}_2}{1 + \text{SNR}_2 + \text{INR}_2} \right). \quad (45)$$

Thus,

$$\frac{\text{INR}_2^2}{\text{SNR}_1 \text{INR}_2} \leq \frac{\text{SNR}_2 + \text{INR}_2}{1 + \text{INR}_2 + \text{SNR}_2}. \quad (46)$$

It is enough to prove $\text{INR}_2^2/(1 + \text{INR}_2 + \text{SNR}_2) \leq \text{INR}_2$, since $\text{INR}_2 \leq \text{SNR}_1$ will prove the rest. The first part trivially holds too.

At the second receiver, we need to show that we can jointly decode the common message of user 1 and the data of user 2 treating the private message from the first user as noise. For this, we need the following:

$$R_{1,e} \leq \log \left( 1 + \frac{\text{INR}_2^2/(1 + \text{INR}_2)}{1 + \text{INR}_2 + \text{SNR}_2} \right)$$

$$R_2 \leq \log \left( 1 + \frac{\text{SNR}_2}{1 + \text{INR}_2 + \text{SNR}_2} \right)$$

$$R_1 + R_2 \leq \log \left( 1 + \frac{\text{SNR}_2 + \text{INR}_2^2/(1 + \text{INR}_2)}{1 + \text{INR}_2 + \text{SNR}_2} \right). \quad (47)$$

We note that all these three conditions are satisfied.

B) Difference between achievability and outer bound: To show that the achievable sum rate is within 2 bits of the outer bound, we consider the following regimes.

1) $\text{INR}_2 < \text{SNR}_2, \text{INR}_2 < \text{SNR}_1$. Achievable sum rate

$$R_{ac} = \log(1 + \frac{\text{SNR}_1}{1 + \text{INR}_2})$$

$$+ \log \left( 1 + \text{SNR}_2 \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2} \right). \quad (48)$$

Outer bound on the sum rate:

$$R_{co} = \log(1 + \text{SNR}_2) + \log \left( 1 + \frac{\text{SNR}_2}{1 + \text{INR}_2} \right)$$

$$R_{co} - R_{ac} = \log(1 + \text{SNR}_2 + \text{INR}_2)$$

$$- \log(1 + \text{SNR}_2 + \text{SNR}_1)$$

$$- \log \left( 1 + \text{SNR}_2 \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2} \right) \leq \log(1 + \text{SNR}_2 + \text{SNR}_2)$$

$$\leq \log(1 + 2\text{SNR}_2) - \log(1 + \text{SNR}_2/2) \leq 2. \quad (50d)$$

2) $\text{INR}_2 < \text{SNR}_2, \text{INR}_2 \geq \text{SNR}_1$. Achievable sum rate

$$R_{ac} = \log(1 + \frac{\text{SNR}_1}{1 + \text{INR}_2})$$

$$+ \log \left( 1 + \text{SNR}_2 \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2} \right). \quad (51)$$

Outer bound on the sum rate:

$$R_{co} = \log(1 + \text{SNR}_2 + \text{INR}_2)$$

$$R_{co} - R_{ac} = \log(1 + \text{SNR}_2 + \text{SNR}_2)$$

$$- \log(1 + \text{SNR}_2 + \text{SNR}_1) + \log(1 + \text{INR}_2)$$

$$- \log \left( 1 + \text{SNR}_2 \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2} \right) \leq \log(1 + \text{SNR}_2 + \text{SNR}_2)$$

$$- \log(1 + \text{SNR}_2 + \text{SNR}_2)$$

$$- \log \left( 1 + \frac{\text{SNR}_2}{2} \right) \leq \log(1 + \text{SNR}_2 + \text{SNR}_2) - \log(1 + (\text{SNR}_2 + \text{INR}_2)/4) \leq 2. \quad (53e)$$

3) $\text{INR}_2 \geq \text{SNR}_2, \text{INR}_2 \geq \text{SNR}_1$. 
Achievable sum rate

\[
R_{ac} = \log \left(1 + \min \left(\frac{SNR_1, INR_2}{1 + \frac{SNR_1}{1 + \frac{SNR_1}{1 + \frac{INR_2}{1 + 2INR_2}}}}\right) + \log(1 + SNR_2)/(1 + 2INR_2))\right).
\]  

Outer bound on the sum rate

\[
R_{co} = \log(1 + SNR_2 + INR_2).
\]  

If \(\frac{INR_2}{1 + \frac{INR_2}{1 + 2INR_2}} \leq SNR_1\), achievability matches the outer bound. However, if \(\frac{INR_2}{1 + \frac{INR_2}{1 + 2INR_2}} \geq SNR_1\), consider \(R_{co} = \log(1 + SNR_1) + \log(1 + SNR_2)\) to show that \(R_{co} - R_{ac} \leq 1\).

4) \(SNR_2 \geq SNR_1, INR_2 < SNR_1\).

Achievable sum rate

\[
R_{ac} = \log \left(1 + \frac{INR_2^2}{1 + 2INR_2 + SNR_2(1 + INR_2)}\right) + \log(1 + SNR_1)/(1 + INR_2) + \log \left(1 + \frac{SNR_1}{1 + 2INR_2}\right).
\]  

Outer bound on the sum rate:

\[
R_{co} = \log(1 + SNR_1) + \log \left(1 + \frac{SNR_2}{1 + INR_2}\right).
\]

Note that

\[
R_{ac} = \log \left(1 + \frac{INR_2^2}{1 + 2INR_2 + SNR_2(1 + INR_2)}\right) + \log \left(1 + \frac{SNR_1}{1 + INR_2}\right) + \log \left(1 + \frac{SNR_1}{1 + 2INR_2}\right) \geq \log \left(1 + \frac{SNR_1}{1 + INR_2}\right) + \log \left(1 + \frac{INR_2^2}{1 + 2INR_2}\right) = \log \left(1 + \frac{SNR_1}{1 + 2INR_2}\right) + \log \left(1 + \frac{INR_2}{1 + 2INR_2}\right) \geq \log \left(1 + \frac{SNR_1}{1 + INR_2}\right) - \log(1 + 2INR_2) + \log(1 + SNR_1) - \log(1 + 2INR_2) \geq \log(1 + SNR_1) - 1.
\]

Thus,

\[
R_{co} - R_{ac} \leq \log(1 + SNR_1) + \log \left(1 + \frac{SNR_2}{1 + INR_2}\right) - \log(1 + SNR_1) + \log \left(1 + \frac{SNR_2}{1 + INR_2}\right) + 1 \leq 2.
\]

APPENDIX C

CLASS OF THREE-USER DETERMINISTIC CHANNELS

In this appendix, we will introduce a class of deterministic double Z-channels and find a capacity region for this class of channels.

We consider a class of deterministic discrete memoryless IFC's in which the outputs \(Y_1, Y_2,\) and \(Y_3\), and the interferences \(V_1\) and \(V_2\) are deterministic functions of the inputs \(X_1, X_2,\) and \(X_3\) as follows:

\[
Y_1 = f_1(X_1) \\
Y_2 = f_2(X_2, V_1) \\
Y_3 = f_3(X_3, V_2) \\
V_1 = g_1(X_1) \\
V_2 = g_2(X_2)
\]

where \(f_1(\cdot), f_2(\cdot), f_3(\cdot), g_1(\cdot),\) and \(g_2(\cdot)\) are the deterministic functions. Furthermore, let \(f_1(\cdot), f_2(\cdot),\) and \(f_3(\cdot)\) satisfy

\[
H(Y_1|X_1) = 0 \\
H(Y_2|X_2) = H(V_1) \\
H(Y_3|X_3) = H(V_2).
\]

Theorem 14: The capacity region for the class of deterministic double Z-channels is given by

\[
R_1 \leq H(Y_1) \\
R_2 \leq H(Y_2) \\
R_3 \leq H(Y_3) \\
R_1 + R_2 \leq H(Y_1|V_1) + H(V_2) \\
R_2 + R_3 \leq H(Y_2|V_1) + H(V_2) + H(Y_3).
\]

The remaining part of the section is devoted to the proof of this theorem.

A) Han–Kobayashi region for a general DMC: The achievability of the aforementioned theorem will follow by specializing the Han–Kobayashi region to the class of deterministic channels. In this section, we will describe the Han–Kobayashi region for a general discrete memoryless channel. Let \(U_{11}, U_{12}, U_{22}, U_{23},\) and \(U_3\) be the auxiliary variables and \(Q\) be the time-sharing auxiliary variable satisfying the following conditions.

1) \(U_{11}, U_{12}, U_{22}, U_{23},\) and \(U_3\) are conditionally independent given \(Q\).

2) \(X_1 = f_1(U_{11}, U_{12}, Q), X_2 = f_2(U_{22}, U_{23}, Q),\) and \(X_3 = f_3(U_3, Q)\).
3) \( \Pr\{Y_1 = y_1, Y_2 = y_2, Y_3 = y_3\} = w(y_1, y_2, y_3|x_1, x_2, x_3) \).

**Theorem 15:** Under the aforementioned constraints, the Han–Kobayashi rate region is given by 
\( R_{11}, R_{12}, R_{22}, R_{23}, R_3 \), where \( R_{11}, R_{12}, R_{22}, R_{23}, R_3 \) satisfy the following set of inequalities:

\[
\begin{align*}
R_{11} &\leq a_1 \\
R_{12} &\leq b_1 \\
R_{11} + R_{12} &\leq c_1 \\
R_{12} &\leq d_1 \\
R_{22} &\leq e_1 \\
R_{23} &\leq f_1 \\
R_{12} + R_{22} &\leq g_1, \\
R_{12} + R_{23} &\leq h_1, \\
R_{22} + R_{23} &\leq i_1 \\
R_{12} + R_{22} + R_{23} &\leq j_1 \\
R_{23} + R_3 &\leq k_1 \\
R_{23} &\leq l_1 \\
R_3 &\leq m_1 \\
-R_{11} &\leq 0 \\
-R_{12} &\leq 0 \\
-R_{22} &\leq 0 \\
-R_{23} &\leq 0 \\
-R_3 &\leq 0
\end{align*}
\]

where

\[
\begin{align*}
a_1 &= I(Y_1; U_{11} | U_{12}, Q) \\
b_1 &= I(Y_1; U_{12} | U_{11}, Q) \\
c_1 &= I(Y_1; U_{12}, U_{11} | Q) \\
d_1 &= I(Y_2; U_{12} | U_{22}, U_{23}, Q) \\
e_1 &= I(Y_2; U_{22} | U_{12}, U_{23}, Q) \\
f_1 &= I(Y_2; U_{23} | U_{12}, U_{22}, Q, Q) \\
g_1 &= I(Y_2; U_{12}, U_{22} | U_{23}, Q) \\
h_1 &= I(Y_2; U_{12}, U_{23} | U_{22}, Q) \\
i_1 &= I(Y_2; U_{22} | U_{12}, U_{23}, Q) \\
j_1 &= I(Y_2; U_{12}, U_{22} | U_{23}, Q) \\
k_1 &= I(Y_3; U_{23} | U_{23}, Q) \\
l_1 &= I(Y_3; U_{23} | U_{23}, Q) \\
m_1 &= I(Y_3; U_{23} | U_{23}, Q).
\end{align*}
\]

**Proof:** The proof is an easy extension of [16] and is therefore omitted.

**Theorem 16:** The aforementioned rate region can be simplified to

\[
\begin{align*}
R_1 &\leq \min(c_1, a_1 + d_1) \\
R_2 &\leq \min(e_1, c_1 + i_1) \\
R_3 &\leq m_1 \\
R_1 + R_2 &\leq a_1 + \min(j_1, g_1 + l_1) \\
R_2 + R_3 &\leq k_1 + e_1
\end{align*}
\]

\[2R_1 + R_2 \leq 2a_1 + g_1 + h_1 \]  
\[R_1 + R_2 + R_3 \leq a_1 + g_1 + k_1 \]
\[-R_1 \leq 0 \]
\[-R_2 \leq 0 \]
\[-R_3 \leq 0. \]

**Proof:** The proof follows by repeated use of the Fourier–Motzkin elimination algorithm as described in the following. In (63a)–(63r), replace \( R_{11} \) by \( R_1 - R_{12} \) and eliminate \( R_{12} \) using the Fourier–Motzkin elimination algorithm. Furthermore, substitute \( R_{22} = R_2 - R_{23} \) and eliminate \( R_{23} \) using the Fourier–Motzkin elimination algorithm. After these steps, we obtain the equivalent region as follows:

\[
\begin{align*}
R_1 &\leq \min(c_1, a_1 + b_1, a_1 + d_1, a_1 + h_1), \\
R_2 &\leq \min(i_1, e_1 + f_1, e_1 + h_1), \\
R_3 &\leq \min(k_1, m_1), \\
R_1 + R_2 &\leq \min(a_1 + j_1 + a_1 + f_1 + g_1, \\
a_1 + g_1 + h_1, a_1 + e_1 + h_1), \\
a_1 + g_1 + l_1)
\end{align*}
\]

Since \( c_1 \leq a_1 + b_1, d_1 \leq h_1, \) and \( d_1 \leq g_1, \) the \( R_1 \) equation is the same as in the statement of the theorem. Since \( i_1 \leq e_1 + f_1 \leq e_1 + h_1, \) the \( R_2 \) equation is the same as in the statement of the theorem. Furthermore, \( m_1 \leq k_1 \) gives \( R_3 \) same as in the statement of the theorem. As \( j_1 \leq f_1 + g_1, j_1 \leq e_1 + h_1, \) and \( f_1 \leq h_1, \) the \( R_1 + R_2 \) bound is also the same as in the statement of the theorem. The rest of the statements directly follows.

**B) Specializing the Han–Kobayashi achievability for model in Theorem 14:** To show the achievable region, we specialize the Han–Kobayashi rate region by taking the time-sharing variable as trivial, and using \( U_{12} = V_1, U_{23} = V_2, \) \( X_2 = h_1(U_{22}, U_{23}), X_1 = h_2(U_{11}, U_{12}), U_{22} = g_3(X_2), U_{11} = g_4(X_1), \) and \( U_3 = X_3 \) for some deterministic functions \( h_1, h_2, g_3, g_4. \) With these substitutions, we now show that the Han–Kobayashi rate region reduces to that in the statement of Theorem 14.

We consider all the equations in the Han–Kobayashi region one by one.

From (65), \( R_1 \leq \min(c_1, a_1 + d_1) \). Here, \( c_1 = I(Y_1; X_1) = H(Y_1) \) and \( a_1 + d_1 = H(Y_1 | V_1) + H(Y_2 | X_2) = H(Y_1, V_1). \) Thus, the aforementioned expression reduces to \( R_1 \leq H(Y_1) \) that is the same as in the statement of Theorem 14.

From (65), \( R_2 \leq \min(i_1, e_1 + l_1). \) Here, \( i_1 = H(Y_2 | V_1) \) and \( e_1 + l_1 = H(Y_2 | V_1 | V_2) + H(Y_2 | V_3 | X_3) = H(Y_2 | V_1, V_2) + H(Y_2) \geq H(Y_2, V_2 | V_1) \geq H(Y_2, V_2 | V_1) \). Thus, this bound also reduces to the same as in the statement of Theorem 14.

From (65), \( R_3 \leq m_1 = H(Y_3 | V_2). \)
From (65), \( R_1 + R_2 \leq a_1 + \min(j_1, g_1 + l_1) \). Here, \( j_1 = H(Y_2) \leq H(Y_2, V_2) = H(Y_2|V_2) + H(X_3|X_3) = g_1 + l_1 \). This reduces to \( R_1 + R_2 \leq H(Y_1|V_1) + H(Y_2|V_1, V_2) \) as in the statement of Theorem 14.

From (65), \( R_3 + R_2 \leq a_1 + g_1 + h_1 = 2H(Y_1|V_1) + H(Y_2|V_2) + H(Y_2|V_2) \). We will now show that this constraint is looser than the sum of the constraints on \( R_1 \) and \( R_1 + R_2 \), which proves that this is not a limiting condition on the rate region:

\[
2H(Y_1|V_1) + H(Y_2|V_1) + H(Y_2|V_2) \leq H(Y_2|V_2) + H(Y_2|V_2) + H(Y_2|V_2)
\]

(67a)

\[
= 2H(Y_1|V_1) + H(Y_2|V_1) + H(Y_2|V_2) + H(Y_2|V_2)
\]

(67b)

\[
= 2H(Y_1|V_1) + H(Y_2|V_1) + H(Y_2|V_2) + H(Y_2|V_2) + H(Y_2|V_2)
\]

(67c)

\[
\geq H(Y_1) + H(Y_2) + H(Y_1) + H(Y_1)
\]

(67d)

where (a) follows since \( V_2 \) and \( V_2 \) are independent. Thus, the bound of \( 2R_1 + R_2 \) is redundant.

From (65), \( R_1 + R_2 + R_3 \leq a_1 + g_1 + h_1 = H(Y_1|V_1) + H(Y_2|V_2) + H(Y_3|V_2) \) that proves the last condition in Theorem 14. This completes the proof of the achievability of the rate region.

C) Converse for Theorem 14: The individual bounds on \( R_1, R_2, R_3, \) and \( R_1 + R_2 + R_3 \) follow the same steps as in [14] and are omitted.

For \( R_2 + R_3 \),

\[
n(R_2 + R_3) \leq I(X_2; Y_2) + I(X_3; Y_3) + n \epsilon
\]

(67a)

\[
\leq I(X_2^2; Y_2^2| V_1^1) + I(X_3^3; Y_3^3) + n \epsilon
\]

(67b)

\[
\leq I(X_2^2; Y_2^2| V_1^1) + I(X_3^3; Y_3^3) + n \epsilon
\]

(67c)

\[
\leq H(Y_2^2| X_2^2) + H(Y_3^3) - H(Y_2^2| V_1^1) + I(X_3^3; Y_3^3) + n \epsilon
\]

(67d)

\[
\leq H(Y_2^2) + H(Y_2^2| V_1^1) + H(Y_3^3)
\]

(67e)

\[
= I(X_2^2; Y_2^2| V_1^1) + I(X_3^3; Y_3^3) + n \epsilon
\]

(67f)

where (a) follows since \( Y_2 \) and \( Y_2 \) imply \( H(Y_2) \leq H(Y_2, V_2) = H(Y_2|V_2) + H(X_3|X_3) = g_1 + l_1 \). This reduces to \( H(Y_1) + H(Y_2) + H(Y_2) \) in the statement of Theorem 14.

\[
\leq H(Y_1) + H(Y_2) + H(Y_2) + H(Y_1) + H(Y_2)
\]

(67g)

\[
\leq H(Y_1) + H(Y_2) + H(Y_2) + H(Y_1) + H(Y_2)
\]

(67h)

\[
\leq H(Y_1) + H(Y_2) + H(Y_2) + H(Y_1) + H(Y_2)
\]

(67i)

\[
\leq H(Y_1) + H(Y_2) + H(Y_2) + H(Y_1) + H(Y_2)
\]

(67j)

Taking \( n \to \infty \) and by the convexity properties of the region, the bound in the statement of Theorem 14 is obtained.

APPENDIX D

OUTER BOUNDS FOR THE GAUSSIAN THREE-USER DOUBLE Z-CHANNEL

In this section, we prove the sum rate bounds in the second and the four cases, more precisely, (22)–(23), (26)–(27).

At \( SNR \geq SNR_1 \) and \( SNR_2 \leq SNR_3 \). Note that there is a strong interference between first two users. Hence, both \( X_1 \) and \( X_2 \) can be decoded from \( Y_2 \). Thus,

\[
n(R_1 + R_2 + R_3) \leq I(X_1^1, X_2^2; Y_2^2) + I(X_3^3) + n \epsilon
\]

(70a)

\[
= I(X_1^1, X_2^2; Y_2^2) + I(X_3^3) + n \epsilon
\]

(70b)

\[
= I(X_1^1, X_2^2; Y_2^2) + I(X_3^3) + n \epsilon
\]

(70c)

\[
= I(X_1^1, X_2^2; Y_2^2) + I(X_3^3) + n \epsilon
\]

(70d)

\[
\leq -h(Y_2^2) + h(Y_2^2| Y_1^1) + h(Y_3^3) + n \epsilon
\]

(70e)

\[
\leq -h(Y_2^2) + h(Y_2^2| Y_1^1) + h(Y_3^3) + n \epsilon
\]

(70f)

The last step follows since Gaussian maximizes conditional entropy as well as marginal entropy [17]. Furthermore, choosing \( \epsilon \) arbitrarily small,

\[
(R_1 + R_2 + R_3)
\]

\[
\leq \epsilon
\]

(71a)

\[
= \epsilon
\]

(71b)

\[
= \epsilon
\]

(71c)

\[
= \epsilon
\]

(71d)

\[
= \epsilon
\]

(71e)

\[
= \epsilon
\]

(71f)

The last step follows since Gaussian maximizes conditional entropy as well as marginal entropy [17]. Furthermore, choosing \( \epsilon \) arbitrarily small,

\[
(R_1 + R_2 + R_3)
\]

\[
\leq \epsilon
\]

(71a)

\[
= \epsilon
\]

(71b)

\[
= \epsilon
\]

(71c)

\[
= \epsilon
\]

(71d)

\[
= \epsilon
\]

(71e)

\[
= \epsilon
\]

(71f)

The last step follows since Gaussian maximizes conditional entropy as well as marginal entropy [17]. Furthermore, choosing \( \epsilon \) arbitrarily small,

\[
(R_1 + R_2 + R_3)
\]

\[
\leq \epsilon
\]

(71a)

\[
= \epsilon
\]

(71b)

\[
= \epsilon
\]

(71c)

\[
= \epsilon
\]

(71d)

\[
= \epsilon
\]

(71e)

\[
= \epsilon
\]

(71f)

The last step follows since Gaussian maximizes conditional entropy as well as marginal entropy [17]. Furthermore, choosing \( \epsilon \) arbitrarily small,

\[
(R_1 + R_2 + R_3)
\]

\[
\leq \epsilon
\]

(71a)

\[
= \epsilon
\]

(71b)

\[
= \epsilon
\]

(71c)

\[
= \epsilon
\]

(71d)

\[
= \epsilon
\]

(71e)

\[
= \epsilon
\]

(71f)

The last step follows since Gaussian maximizes conditional entropy as well as marginal entropy [17]. Furthermore, choosing \( \epsilon \) arbitrarily small,

\[
(R_1 + R_2 + R_3)
\]

\[
\leq \epsilon
\]

(71a)

\[
= \epsilon
\]

(71b)

\[
= \epsilon
\]

(71c)

\[
= \epsilon
\]

(71d)

\[
= \epsilon
\]

(71e)

\[
= \epsilon
\]

(71f)

The last step follows since Gaussian maximizes conditional entropy as well as marginal entropy [17]. Furthermore, choosing \( \epsilon \) arbitrarily small,

\[
(R_1 + R_2 + R_3)
\]

\[
\leq \epsilon
\]

(71a)

\[
= \epsilon
\]

(71b)

\[
= \epsilon
\]

(71c)

\[
= \epsilon
\]

(71d)

\[
= \epsilon
\]

(71e)

\[
= \epsilon
\]

(71f)
For \((\text{INR}_2 + 1)\text{INR}_3 \geq \text{SNR}_2\), we can choose \(V = c_1\sqrt{\text{SNR}_2} X_1 + \sqrt{\text{SNR}_3} Z_2\), where \(c_1 \leq 1\) is chosen to make the previous expression unit variance. With this choice, \(\begin{align}
I(X_1; Y_2; Z_2; S_G) & = I(X_1; Y_2; S_G | X_2; S_G) \\
& = I(X_1; Y_2; S_G | X_2)\quad(74a)
\end{align}\)

For the second term \(I(X_1; Y_2; S_G | X_2), S_G\) is the deterministic function of \(X_2; S_G\) and \(Y_2; S_G\) and this term is therefore 0. The third term \(I(X_2; Y_2; S_G) = I(X_2; Y_2; S_G)\) is \(0\) as \(X_2 = S_G - Y_2\) is a Markov chain. Thus,
\(\begin{align}
I(X_1; Y_2; S_G | X_2) & = I(X_1; Y_2; S_G | X_2; S_G) \\
& = I(X_1; Y_2; S_G) + I(X_2; S_G | Y_2, S_G)\quad(74b)
\end{align}\)

This gives the previous sum rate.

### B) \(\text{INR}_2 \leq \text{SNR}_1\) and \(\text{INR}_3 \leq \text{SNR}_2\):
By Fano’s inequality,
\(\begin{align}
n(I(X_1^2; Y_2^2) + I(X_2^2; Y_2^2) + I(X_3^2; Y_3^2) + n\epsilon)\quad(76)
\end{align}\)

Note that \(X_3 = \sqrt{\text{SNR}_3} X_2 + \sqrt{\text{SNR}_3} Z_3, Y_2 = \sqrt{\text{SNR}_2} X_2 + \sqrt{\text{SNR}_2} Z_2\), and \(Y_1 = \sqrt{\text{SNR}_1} X_1 + Z_1\).
Let \(S_1 = \sqrt{\text{SNR}_1} X_1 + V_1\) and \(S_2 = \sqrt{\text{SNR}_3} X_2 + V_2\) where \(V_1\) and \(V_2\) are mutual independent complex Gaussians of unit variance and independent of all \(X_1\). Not that \(I(X_2^2; Y_2^2) = I(X_2^2; S_1^2 + \sqrt{\text{SNR}_2} X_2^2)\) since the distributions remain the same.
\(\begin{align}
n(R_1 + R_2 + R_3) & = I(X_1^2; Y_1^2) + I(X_2^2; Y_2^2) + I(X_3^2; Y_3^2) + n\epsilon \\
& \leq I(X_1^2; Y_1^2 | S_1^2) + I(X_2^2; S_1^2 + \sqrt{\text{SNR}_2} X_2^2) + I(X_2^2; Y_2^2) + I(X_3^2; Y_3^2) + n\epsilon \\
& + I(X_1^2; Y_1^2 | S_1^2) + I(X_2^2; S_1^2 + \sqrt{\text{SNR}_2} X_2^2) + I(X_2^2; Y_2^2) + I(X_3^2; Y_3^2) + n\epsilon \\
& = I(X_1^2; S_1^2) + I(X_1^2; Y_1^2 | S_1^2) + I(X_2^2, S_2^2) + I(X_2^2; Y_2^2) + I(X_3^2; Y_3^2) + n\epsilon \quad(77c)
\end{align}\)

The last step follows by choosing \(V_1^2\) independent of \(Z_2^2\). Furthermore, since conditional and marginal entropies are maximized by Gaussians,
\(\begin{align}
n(R_1 + R_2 + R_3) & \leq h(Y_1^2 | S_1^2) - h(Z_1^2, V_1^2) - h(V_2^2) \\
& + h(S_1^2 + \sqrt{\text{SNR}_2} X_2^2 | S_2^2) + h(Y_3^2) + n\epsilon \quad(78a)
\end{align}\)

\(\begin{align}
& \leq h(Y_1^2 | S_1^2) - h(Z_1^2, V_1^2) - h(V_2^2) \\
& + h(S_1^2 + \sqrt{\text{SNR}_2} X_2^2 | S_2^2) + h(Y_3^2) + n\epsilon \\
& = \text{log}(1 + \text{SNR}_1) + h(Y_3^2) - h(S_1^2) - h(V_2^2) \\
& + h(S_1^2 + \sqrt{\text{SNR}_2} X_2^2 | S_2^2)\quad(80a)
\end{align}\)

### APPENDIX E
**Proof That Rates Can Be Decoded With 1.5 Rounds of Message Passing**

To show this, we divide the range of \(\text{INR}_2, \text{SNR}_2, \text{INR}_3, \text{SNR}_3\) into the following nine cases.

\(\text{SNR}_1, \text{SNR}_2, \text{SNR}_3, \text{SNR}_4\). Note that the last expression \(I(X_1; S_1; X_2, S_2; S_3; S_2, S_2)\) is similar to that in Appendix D-A, and thus, \(S_2\) can be chosen as in Appendix D-A. \(S_1 + \sqrt{\text{SNR}_2} X_2\) plays the role of \(Y_2; S_2, S_2\); the only difference is that \(Z_2\) is replaced by \(V_2\). Thus, the same steps give the required bounds on \(R_1 + R_2 + R_3\).
A) $\text{INR}_2 \geq \text{SNR}_2$, $\text{INR}_3 \geq \text{SNR}_3$, $\text{INR}_2 \geq \text{SNR}_3$, $\text{INR}_3 \geq \text{SNR}_2$. The first transmitter makes a codebook of rate

$$R_1 = \log \left( 1 + \min \left( \frac{\text{INR}_3}{\text{SNR}_1}, \frac{\text{INR}_3}{1 + \text{INR}_2 + \text{INR}_3} \right) \right)$$

(81)

and uses a power level of 1 to transmit. The second transmitter makes a codebook of rate

$$R_2 = \log \left( 1 + \frac{1 + \text{INR}_3}{1 + 2 \text{INR}_2} \times \right.$$  

$$\min \left( \frac{\text{SNR}_2}{1 + \text{INR}_2 + \text{INR}_3} \right) \right)$$

(82)

and uses a power level of 1 to transmit. The third transmitter makes a codebook of rate

$$R_3 = \log (1 + \text{SNR}_3/(1 + \text{INR}_3) \text{/(1 + 2INR}_3))$$

(83)

and uses a power level of 1 to transmit. At receiver 1, $R_1$ can be decoded since $R_1 \leq \log(1 + \text{SNR}_1)$. At receiver 2, $R_1$ and $R_2$ are decoded jointly. We see that

$$R_1 \leq \log(1 + \text{INR}_2)$$

(84a)

$$R_2 \leq \log(1 + \text{SNR}_2)$$

(84b)

$$R_1 + R_2 \leq \log(1 + \text{INR}_2 + \text{SNR}_2).$$

(84c)

The $R_1 + R_2$ equation holds as

$$R_1 + R_2 \leq \log \left( 1 + \frac{\text{INR}_3}{1 + \text{INR}_2 + \text{INR}_3} \right) +$$

$$\log \left( 1 + \frac{1 + \text{INR}_3}{1 + 2 \text{INR}_2} \times \right.$$  

$$\min \left( \frac{\text{SNR}_2}{1 + \text{INR}_2 + \text{INR}_3} \right) \right)$$

$$= \log(1 + \text{INR}_2 + \text{SNR}_2).$$

At receiver 3, $R_2$ and $R_3$ are decoded jointly. We see that

$$R_2 \leq \log(1 + \text{INR}_3)$$

(85a)

$$R_3 \leq \log(1 + \text{SNR}_3)$$

(85b)

$$R_2 + R_3 \leq \log \left( 1 + \frac{\text{INR}_3}{1 + \text{INR}_2 + \text{SNR}_3} \right) =$$

$$\log(1 + \text{INR}_3 + \text{SNR}_3).$$

B) $\text{INR}_2 \geq \text{SNR}_2$, $\text{INR}_3 \geq \text{SNR}_3$, $\text{INR}_2 \geq \text{SNR}_3$, $\text{INR}_3 < \text{SNR}_2$. The first transmitter makes a codebook of rate

$$R_1 = \log \left( 1 + \min \left( \frac{\text{SNR}_1}{1 + \text{INR}_2 + \text{INR}_3} \right) \right)$$

(86)

and uses a power level of 1 to transmit.

The second transmitter makes two codebooks, the first one of rate

$$R_{2,c} = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2} \times \right.$$  

$$\frac{\text{INR}_3}{1 + 2 \text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)} \right),$$

(87)

and the second of rate

$$R_{2,p} = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2} \times \right.$$  

$$\frac{\text{SNR}_2/2}{1 + 2 \text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right).$$

(88)

The third transmitter makes a codebook of rate

$$R_3 = \log(1 + \text{SNR}_3/(1 + \text{INR}_3) \text{/(1 + 2INR}_3))$$

(89)

and uses a power level of 1 to transmit. At receiver 1, $R_1$ can be decoded since $R_1 \leq \log(1 + \text{SNR}_1)$. At receiver 2, $R_1$, $R_{2,c}$, and $R_{2,p}$ are decoded jointly. We see that

$$R_1 \leq \log(1 + \text{INR}_2)$$

(90a)

$$R_{2,c} \leq \log \left( 1 + \frac{\text{INR}_3}{1 + \text{SNR}_2} \right)$$

(90b)

$$R_{2,p} \leq \log \left( 1 + \frac{\text{SNR}_2}{1 + \text{INR}_3} \right)$$

(90c)

$$R_1 + R_{2,c} \leq \log \left( 1 + \text{INR}_2 \right.$$  

$$\text{ INR}_3 \left. \right) \text{/(1 + 2INR}_3 \left. \right) \right)$$

(90d)

$$R_1 + R_{2,p} \leq \log \left( 1 + \text{INR}_2 \right.$$  

$$\text{ INR}_3 \left. \right) \text{/(1 + 2INR}_3 \left. \right) \right)$$

(90e)

$$R_{2,c} + R_{2,p} \leq \log(1 + \text{SNR}_2)$$

(90f)

$$R_1 + R_{2,c} + R_{2,p} \leq \log(1 + \text{INR}_2 + \text{SNR}_2).$$

(90g)

The first three subequations hold trivially. For the fourth equation, we see that

$$R_1 + R_{2,c} \leq \log \left( 1 + \frac{\text{INR}_3}{1 + \text{SNR}_3(1 + \text{INR}_3)} \right)$$

$$+ \log \left( 1 + \frac{\text{INR}_3}{1 + 2 \text{INR}_2 + \text{SNR}_3(1 + \text{INR}_3)} \right)$$

(91)

$$= \log \left( 1 + \frac{\text{INR}_3}{1 + 2 \text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)} \right)$$

$$+ \frac{\text{INR}_3}{1 + 2 \text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)} \times$$

$$\frac{\text{INR}_3}{1 + 2 \text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)} \right),$$

(92)

Thus, we need to show that

$$\frac{\text{INR}_3}{1 + 2 \text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)}$$

$$+ \frac{\text{INR}_3}{1 + 2 \text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)} \times$$

$$\frac{\text{INR}_3}{1 + 2 \text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)} \right) \leq 0.$$

(93)

Note that LHS decreases with $\text{SNR}_3$, and hence, it is enough to show that the previous expression is negative for $\text{SNR}_3 = 0$. With $\text{SNR}_3 = 0$, we find that the the previous expression decreases with $\text{SNR}_2$, and thus, we only need to show for a low
value of $\text{SNR}_2$ and as $\text{SNR}_2 \geq \text{INR}_3$, it is enough to show for $\text{INR}_3 = \text{SNR}_3$. The previous expression reduces to

$$A = - \frac{\text{INR}_3(1 + \text{INR}_3 + 2\text{INR}_2 + \text{INR}_3\text{INR}_2)}{(1 + 2\text{INR}_2)(1 + \text{INR}_3)}$$

$$+ \frac{\text{INR}_3\text{SNR}_2(1 + \text{INR}_3 + 2\text{INR}_2)}{1 + 2\text{INR}_2} \leq \text{INR}_2^2 \leq 0.$$  

We now show that $A$ decreases with $\text{INR}_2$. To see that, differentiating w.r.t. $\text{INR}_2$, we obtain

$$\frac{dA}{d\text{INR}_2} = \frac{1}{(1 + 2\text{INR}_2)^2} (4\text{INR}_3/(1 + \text{INR}_3) - 2\text{INR}_2).$$

As $\text{INR}_3 \leq \text{INR}_2$, the previous equation is negative; thus, it is enough to prove $A \leq 0$ for $\text{INR}_2 = \text{INR}_3$ which can be shown easily.

To show $R_1 + R_2 \leq \log(1 + \text{INR}_3 + \text{SNR}_2/(1 + \text{INR}_3))$, we will show that $R_1 + R_2 \leq \log(1 + \text{INR}_2 + \text{SNR}_2/(1 + \text{INR}_3)) - \text{INR}_2$ which will prove the claim.

To see this, it is enough to prove that

$$\text{INR}_2 \leq \frac{\text{SNR}_2 - \text{INR}_2}{(1 + \text{SNR}_2 + 2\text{INR}_2 + \text{SNR}_2\text{INR}_2)(1 + \text{INR}_3)},$$

$$= \text{INR}_2 \frac{\text{SNR}_2 - \text{INR}_2}{(1 + \text{SNR}_2 + 2\text{INR}_2 + \text{SNR}_2\text{INR}_2)(1 + \text{INR}_3)} \leq \text{INR}_2 (1 + \text{INR}_3).$$

This is equivalent to proving

$$\text{INR}_2 = \text{SNR}_2 - \frac{\text{INR}_2}{(1 + \text{SNR}_2 + 2\text{INR}_2 + \text{SNR}_2\text{INR}_2)(1 + \text{INR}_3)}$$

$$= \text{INR}_2 \text{SNR}_2 - \frac{\text{INR}_2 \text{SNR}_2 - \text{INR}_2^2 \text{INR}_3}{(1 + \text{INR}_3)} \leq \text{SNR}_2 - \text{INR}_2(1 + \text{INR}_3).$$

which is true. Hence proved.

To show $R_{2,e} + R_{2,f} \leq \log(1 + \text{SNR}_2)$, it is enough to prove that

$$\text{SNR}_2 \leq \frac{\text{INR}_2^2}{(1 + 2\text{INR}_2 + \text{SNR}_2)(1 + \text{INR}_3)} + \text{SNR}_2$$

$$\leq \text{INR}_2^2$$

or

$$\text{INR}_3 \leq \text{SNR}_2(1 + \text{SNR}_2)$$

which holds since $\text{SNR}_2 \geq \text{INR}_3$.

To show $R_1 + R_{2,e} + R_{2,f} \leq \log(1 + \text{INR}_2 + \text{SNR}_2)$, it is sufficient to show that

$$\text{INR}_2 + \text{SNR}_2/(1 + \text{INR}_3) - d$$

$$+ \frac{\text{INR}_2}{1 + 2\text{INR}_2} \frac{\text{SNR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)}$$

$$+ (\text{SNR}_2/1 + \text{INR}_3) - d \frac{\text{INR}_2}{1 + 2\text{INR}_2} \leq \text{INR}_2^2$$

(100)

where $d = \text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)/(1 + \text{INR}_3)$.

Note that the previous expression is equivalent to show

$$- \text{SNR}_2\text{INR}_3/(1 + \text{INR}_3) - d$$

$$+ \frac{\text{SNR}_2}{1 + 2\text{INR}_2} \frac{\text{INR}_2^2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)}$$

$$+ (\text{SNR}_2/1 + \text{INR}_3) - d \frac{\text{INR}_2}{1 + 2\text{INR}_2} \leq \text{INR}_2^2$$

(101)

Since the left-hand side decreases with $\text{SNR}_3$, it is enough to show the previous expression for $\text{SNR}_3 = 0$. Note that the previous expression is equivalent to show

$$- \text{SNR}_2\text{INR}_3/(1 + \text{INR}_3) - d$$

$$+ \text{SNR}_2/(1 + \text{INR}_3) - d \frac{\text{INR}_2}{1 + 2\text{INR}_2} \leq \text{INR}_2^2$$

(102)

Since this decreases with $\text{SNR}_2$, it is enough to show the previous expression for $\text{SNR}_2 = \text{INR}_3$ as $\text{SNR}_2 \geq \text{INR}_3$:

$$- \text{SNR}_2\text{INR}_3/(1 + \text{INR}_3) - d$$

$$+ \frac{\text{SNR}_2}{1 + 2\text{INR}_2} \frac{\text{INR}_2^2}{1 + 2\text{INR}_2 + \text{INR}_3}$$

$$= (1 - \frac{\text{INR}_2}{1 + \text{SNR}_2})$$

$$- \frac{\text{SNR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2\text{INR}_2} \leq 0.$$  

Thus, it suffices to show

$$A = - \frac{\text{INR}_3(1 + 3\text{INR}_2 + \text{INR}_3 + 2\text{INR}_2(2 + \text{INR}_3))}{\text{INR}_3(1 + 3\text{INR}_2 + \text{INR}_3 + 2\text{INR}_2(2 + \text{INR}_3))}$$

$$< 0.$$  

Thus, it is enough to prove $A \leq 0$ for $\text{INR}_2 = \text{INR}_3$, which is easy to see.

At receiver 3, $R_{2,e}$ and $R_3$ are decoded jointly treating $R_{2,f}$ as noise. We see that

$$R_{2,e} \leq \log(1 + \text{INR}_3/(1 + 2\text{INR}_3))$$

$$R_3 \leq \log(1 + \text{SNR}_3 \frac{1 + \text{INR}_3}{1 + 2\text{INR}_3})$$
\[ R_{2,c} + R_3 \leq \log \left( \frac{1 + \frac{\text{INR}_2}{1 + \text{INR}_3}}{1 + 2\text{INR}_3} + \frac{1 + \frac{\text{SNR}_3}{1 + \frac{1 + \text{INR}_3}{1 + \text{INR}_3}}}{1 + 2\text{INR}_3} \right). \]  

(105c)

To show \( R_{2,c} + R_3 \leq \log \left( \frac{1 + \frac{\text{INR}_2}{1 + \text{INR}_3} + \text{SNR}_3(1 + \text{INR}_3)/(1 + 2\text{INR}_3)}{1 + 2\text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)/(1 + 2\text{INR}_3)} \right) \), we see that \( R_{2,c} + R_3 \leq \log \left( \frac{1 + \frac{\text{INR}_2}{1 + \text{INR}_3} + \text{SNR}_3(1 + \text{INR}_3)/(1 + 2\text{INR}_3)}{1 + 2\text{INR}_3 + \text{SNR}_3(1 + \text{INR}_3)/(1 + 2\text{INR}_3)} \right) \) as was in the case of a two-user channel.

C) \( \text{INR}_2 \geq \text{SNR}_3, \text{INR}_3 \geq \text{SNR}_3, \text{INR}_3 < \text{SNR}_3, \text{INR}_3 \geq \text{SNR}_2 \): The first transmitter makes two codebooks, the first one of rate
\[ R_{1,c} = \log \left(1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_2)} \right), \]  

(106)

and the second of rate
\[ R_{1,p} = \log \left(1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right). \]  

(107)

The transmitter sends the first one at a power of \( \frac{\text{INR}_2}{1 + \text{INR}_2} \) and the second one at a power of \( 1/(1 + \text{INR}_2) \), adds them up, and transmits.

The second transmitter makes a codebook of rate
\[ R_2 = \log \left(1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_2)} \right), \]  

(108)

and uses a power level of 1 to transmit. The third transmitter makes a codebook of rate
\[ R_3 = \log \left(1 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right), \]  

(109)

and uses a power level of 1 to transmit.

The first receiver is able to decode following the same steps as in the two-user interference channel given earlier. The second receiver decodes \( R_{1,c} \) and \( R_2 \) treating \( R_{1,p} \) as noise. The decoding occurs when
\[ R_{1,c} \leq \log \left(1 + \frac{\text{INR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(110a)

\[ R_2 \leq \log \left(1 + \frac{\text{SNR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(110b)

\[ R_{1,c} + R_2 \leq \log \left(1 + \frac{\text{SNR}_2 + \text{SNR}_3}{1 + \frac{\text{SNR}_2 + \text{SNR}_3}{1 + \text{INR}_1}} \right). \]  

(110c)

These conditions can be easily shown to be satisfied.

The third receiver does a joint decoding and can decode \( R_3 \), and the rate constraints are satisfied as in the two-user case since powers remain the same as those while rates now are even lesser.

D) \( \text{INR}_2 \geq \text{SNR}_2, \text{INR}_3 \geq \text{SNR}_3, \text{INR}_2 < \text{SNR}_3, \text{INR}_3 < \text{SNR}_2 \): The first transmitter makes two codebooks, the first one of rate
\[ R_{1,c} = \log \left(1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_2)} \right), \]  

(111)

and the second of rate
\[ R_{1,p} = \log \left(1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right). \]  

(112)

The transmitter sends the first one at a power of \( \frac{\text{INR}_2}{1 + \text{INR}_2} \) and the second one at a power of \( 1/(1 + \text{INR}_2) \), adds them up, and transmits.

The second transmitter makes two codebooks, the first one of rate
\[ R_{2,c} = \log \left(1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_2)} \right), \]  

(113)

and the second of rate
\[ R_{2,p} = \log \left(1 + \frac{\text{SNR}_2}{1 + 2\text{INR}_2} \right). \]  

(114)

The third transmitter makes a codebook of rate
\[ R_3 = \log \left(1 + \frac{\text{SNR}_3}{1 + \frac{\text{SNR}_3}{1 + \text{INR}_2}} \right), \]  

(115)

and uses a power level of 1 to transmit.

As in the two-user case, the first receiver is able to decode the public message treating private as noise and then decode the private message. The third receiver can decode \( R_{2,c} \) and \( R_3 \) jointly treating \( R_{2,p} \) as noise which can be shown following the same steps as in the two-user case. However, the calculations at the second receiver are different. The second receiver decodes \( R_{1,c} \), \( R_{2,c} \), and \( R_{2,p} \) jointly treating \( R_{1,p} \) as noise. The decoding happens when the following are satisfied:
\[ R_{1,c} \leq \log \left(1 + \frac{\text{INR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(116a)

\[ R_{2,c} \leq \log \left(1 + \frac{\text{SNR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(116b)

\[ R_{2,p} \leq \log \left(1 + \frac{\text{SNR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(116c)

\[ R_{1,c} + R_{2,c} \leq \log \left(1 + \frac{\text{INR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(116d)

\[ R_{1,c} + R_{2,c} \leq \log \left(1 + \frac{\text{INR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(116e)

\[ R_{2,c} + R_{2,p} \leq \log \left(1 + \frac{\text{SNR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(116f)

\[ R_{1,c} + R_{2,c} + R_{2,p} \leq \log \left(1 + \frac{\text{SNR}_2}{1 + \frac{\text{SNR}_2}{1 + \text{INR}_1}} \right), \]  

(116g)

It is straightforward to see that the individual constraints are satisfied. To see the constraint on \( R_{1,c} + R_{2,c} \), we write the satisfying equation and note as in second case that it is enough to use \( \text{SNR}_3 = 0 \) and \( \text{SNR}_3 = \text{INR}_3 \). Thus, it is sufficient to prove that
\[ \frac{\text{INR}_2^2}{1 + 2\text{INR}_2 + \text{INR}_3(1 + \text{INR}_2)} + \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{INR}_3} \frac{\text{INR}_2^2}{1 + 2\text{INR}_2 + \text{INR}_3(1 + \text{INR}_2)} \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{INR}_3} + \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{INR}_3(1 + \text{INR}_2)} \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{INR}_3} = \text{INR}_3. \]
\[ \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2} \left( \frac{\text{INR}_2^2}{1 + \text{INR}_2^3} + \frac{\text{INR}_3^2}{1 + \text{INR}_3^2} \right) \leq 1 + \text{INR}_2 \]

This is equivalent to
\[ -1 + \frac{1 + \text{INR}_2^3}{1 + 2 \text{INR}_3^3} - (1 + 2 \text{INR}_2 + \text{INR}_3(1 + 2 \text{INR}_2)) \times \frac{\text{INR}_2^2}{(1 + \text{INR}_2^3)(1 + 2 \text{INR}_3^3)} \leq 0. \]

The previous expression increases with \( \text{INR}_3 \), and hence, it is sufficient to show that the previous expression holds as \( \text{INR}_2 \rightarrow \infty \), at which, also, this is \( \leq 0 \) and, hence, holds.

To show \( R_{1c} + R_{2p} \leq \log \left( 1 + \frac{\text{INR}_2^2}{1 + 2 \text{INR}_2^2 + \text{SNR}_2(1 + 2 \text{INR}_2)} \right) \), it is sufficient to show that
\[ \frac{\text{INR}_2^2}{1 + 2 \text{INR}_2^2 + \text{SNR}_2(1 + 2 \text{INR}_2)} \times \left( \frac{\text{INR}_2}{1 + \text{INR}_2 + \text{SNR}_3} \right) \leq \frac{1 + \text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_2^2} \]

This reduces to
\[ \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2} \leq \text{SNR}_2(1 + \text{INR}_2) \]

which trivially holds. Note that more precisely,
\[ R_{1c} + R_{2p} \leq \log \left( 1 + \frac{\text{INR}_2^2}{1 + \text{INR}_2 + \text{SNR}_3} \right) - d \]

where \( d = \frac{1 + \text{INR}_2^3}{1 + 2 \text{INR}_2^3 + \text{SNR}_2(1 + 2 \text{INR}_2)} \).

The result \( R_{2c} + R_{2p} \leq \log \left( 1 + \frac{\text{INR}_2^2}{1 + \text{INR}_2 + \text{SNR}_3} \right) \) holds by using the same steps as in the two-user Z-channel, since cancelling \( \frac{1 + \text{INR}_2^3}{1 + 2 \text{INR}_2^3 + \text{SNR}_2(1 + 2 \text{INR}_2)} \) on both sides, the LHS is now even smaller than was earlier.

Finally, we need to show \( R_{1c} + R_{2c} + R_{2p} \leq \log \left( 1 + \frac{\text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_2^3} \right) \). To show this, it is sufficient to prove that
\[ \log \left( 1 + \frac{\text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_2^3} \right) - d \]
\[ + \log \left( 1 + \frac{\text{INR}_2^2}{2 \text{INR}_2 + 1 + 2 \text{INR}_3 + \text{SNR}_3(1 + 2 \text{INR}_3)} \right) \]
\[ \leq \log \left( 1 + \frac{\text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_2^3} \right) . \]

As before, it is sufficient to prove this for \( \text{SNR}_3 = 0 \). Hence, it is sufficient to prove that
\[ -d \left( 1 + \frac{1 + \text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + 2 \text{INR}_3} \right) \]
\[ + \frac{1 + \text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + 2 \text{INR}_3} \]
\[ + \frac{\text{SNR}_2(1 + \text{INR}_2) \text{INR}_3^2}{(1 + \text{INR}_3)(1 + 2 \text{INR}_2) \text{INR}_3} \]
\[ \times \left( 1 - \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \right) \leq 0. \]

As \( d \) increases with \( \text{SNR}_2 \), the previous expression decreases with \( \text{SNR}_2 \), and thus, it is enough to show that the previous expression holds for \( \text{SNR}_2 = \text{INR}_3 \). Thus,
\[ \frac{1 + \text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \times \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \right) \]
\[ + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \times \left( 1 + \frac{\text{INR}_2}{(1 + 2 \text{INR}_2)^2} \right) \]
\[ - \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \right) \leq 0 \]

or
\[ -\frac{1 + \text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \times \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \right) \]
\[ - \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \right) \leq 0 \]

which is true since the previous expression reduces to
\[ -\frac{1 + \text{INR}_2^2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \times \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + 1 + \text{INR}_3} \right) \]

E) \( \text{INR}_2 \geq \text{SNR}_2 \), \( \text{INR}_3 < \text{SNR}_3 \), \( \text{INR}_2 \geq \text{SNR}_3 \): The first transmitter makes a codebook of rate
\[ R_1 = \log \left( 1 + \min \left( \frac{\text{SNR}_2, 1}{1 + \frac{\text{INR}_2}{1 + \text{INR}_2}} \right) \right) \]

and uses a power level of 1 to transmit.

The second transmitter makes a codebook of rate
\[ R_2 = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \]

and uses a power level of \( 1/(1 + \text{INR}_3) \) to transmit.

The third transmitter makes a codebook of rate
\[ R_3 = \log \left( 1 + \frac{\text{SNR}_3(1 + \text{INR}_3)}{1 + 2 \text{INR}_3} \right) \]

and uses a power level of 1 to transmit.

The first receiver is able to decode as the rate \( \leq \log(1 + \text{SNR}_1) \). The third user is able to decode treating \( R_2 \) as noise. Furthermore, the second user decodes \( R_1 \) and \( R_2 \) jointly. The decoding occurs since
\[ R_1 \leq \log(1 + \text{INR}_2) \]
\[ R_2 \leq \log(1 + \text{SNR}_2(1 + \text{INR}_3)) \]
\[ R_1 + R_2 \leq \log \left( 1 + \frac{\text{SNR}_2}{1 + \text{INR}_3} \right) . \]

To show the sum rate, it is enough to prove
\[ \frac{\text{INR}_2}{1 + \text{SNR}_2 + 2 \text{INR}_2} + \frac{1 + \text{INR}_2}{1 + 2 \text{INR}_2} + \frac{\text{INR}_2}{1 + \text{SNR}_2 + 2 \text{INR}_2} \]
\[ \leq \frac{\text{INR}_2(1 + \text{SNR}_2 + 2 \text{INR}_2)}{1 + \text{SNR}_2 + 2 \text{INR}_2} \]

This is equivalent to proving
\[ (1 + \text{INR}_2)(1 + \text{SNR}_2 + 2 \text{INR}_2) \leq \text{INR}_2(1 + 2 \text{INR}_2)(1 + \text{INR}_3) + 1 \]
\[ + \text{SNR}_2 + 2 \text{INR}_2 + \text{SNR}_2 \text{INR}_2 \]
which trivially holds.

F) \( \text{INR}_2 \geq \text{SNR}_2, \text{INR}_3 < \text{SNR}_3, \text{INR}_2 \leq \text{SNR}_3 \): The first transmitter makes two codebooks, the first one of rate
\[
R_{1,c} = \log \left( 1 + \frac{\text{INR}_3^2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \tag{131}
\]
and the second of rate
\[
R_{1,p} = \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right). \tag{132}
\]
The transmitter sends the first one at a power of \( \text{INR}_2/(1 + \text{INR}_2) \) and the second one at a power of \( 1/(1 + \text{INR}_2) \), adds them up, and transmits.

The second transmitter makes a codebook of rate
\[
R_2 = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \tag{133}
\]
and uses a power level of \( 1/(1 + \text{INR}_2) \) to transmit.

The third transmitter makes a codebook of rate
\[
R_3 = \log \left( 1 + \frac{\text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_3} \right) \tag{134}
\]
and uses a power level of 1 to transmit.

The first decoder can decode in the same way as in the two-user Z-channel, and the third user is able to decode treating \( R_2 \) as noise. The second receiver can decode \( R_{1,c} \) and \( R_2 \) jointly treating \( R_{1,p} \) as noise since
\[
R_{1,c} \leq \log \left( 1 + \frac{\text{INR}_2^2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \tag{135a}
\]
\[
R_2 \leq \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right) \tag{135b}
\]
\[
R_{1,c} + R_2 \leq \log \left( 1 + \frac{\text{INR}_3}{1 + 2\text{INR}_2} \right) \tag{135c}
\]
These conditions are straightforward to verify and are, therefore, omitted.

G) \( \text{INR}_2 < \text{SNR}_2, \text{INR}_3 \geq \text{SNR}_3, \text{INR}_2 \geq \text{SNR}_2 \): The first transmitter makes a codebook of rate
\[
R_1 = \log \left( 1 + \frac{\text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_3} \right) \tag{136}
\]
and uses a power level of \( 1/(1 + \text{INR}_3) \) to transmit.

The second transmitter makes a codebook of rate
\[
R_2 = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \tag{137}
\]
and uses a power level of 1 to transmit.

The third transmitter makes a codebook of rate
\[
R_3 = \log \left( 1 + \frac{\text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_3} \right) \tag{138}
\]
and uses a power level of 1 to transmit.

The first and third receivers are able to decode by the same scheme of the two-user case. The second receiver can decode treating \( R_1 \) as noise.

H) \( \text{INR}_2 < \text{SNR}_2, \text{INR}_3 \geq \text{SNR}_3, \text{INR}_3 < \text{SNR}_2 \): The first transmitter makes a codebook of rate
\[
R_1 = \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_2} \right) \tag{139}
\]
and uses a power level of \( 1/(1 + \text{INR}_2) \) to transmit.

The second transmitter makes two codebooks, the first one of rate
\[
R_{2,c} = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \tag{140}
\]
and the second of rate
\[
R_{2,p} = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right). \tag{141}
\]
The transmitter sends the first one at a power of \( \text{INR}_3/(1 + \text{INR}_3) \) and the second one at a power of \( 1/(1 + \text{INR}_3) \), adds them up, and transmits.

The third transmitter makes a codebook of rate
\[
R_3 = \log \left( 1 + \frac{\text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_3} \right) \tag{142}
\]
and uses a power level of 1 to transmit.

The receiver treats \( R_1 \) as noise and the equation for decoding at second receiver follows the two-user equations.

I) \( \text{INR}_2 < \text{SNR}_2 \) and \( \text{INR}_3 < \text{SNR}_3 \): The first transmitter makes a codebook of rate
\[
R_1 = \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_2} \right) \tag{143}
\]
and uses a power level of \( 1/(1 + \text{INR}_2) \) to transmit.

The second transmitter makes a codebook of rate
\[
R_2 = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \tag{144}
\]
and uses a power level of \( 1/(1 + \text{INR}_2) \) to transmit.

The third transmitter makes a codebook of rate
\[
R_3 = \log \left( 1 + \frac{\text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_3} \right) \tag{145}
\]
and uses a power level of 1 to transmit.

Here, the second user can treat first user’s data as noise. Thus, all are able to decode.

APPENDIX F

SUM RATE IS WITHIN 4 BITS WITH 2.5 ROUNDS OF MESSAGE PASSING

A) \( \text{INR}_2 \geq \text{SNR}_2, \text{INR}_3 \geq \text{SNR}_3, \text{INR}_2 \geq \text{SNR}_3, \text{INR}_3 \geq \text{SNR}_2 \): Note that we only need to consider the case when \( \frac{\text{SNR}_3}{1 + \text{INR}_3} \leq \frac{\text{SNR}_2}{1 + \text{INR}_2} \leq \frac{\text{SNR}_3}{1 + \text{INR}_2} \), since otherwise the 1.5 round scheme was within 4 bits of optimal.

The rates allocated to the users in this case are as follows:
\[
R_1 = \min \{ \log (1 + \text{SNR}_1), \log (1 + \text{INR}_2 + \text{SNR}_2) - R_2 \} \tag{146a}
\]
\[
R_2 = \log \left( \frac{1 + \text{INR}_2}{1 + 2\text{INR}_2 + \text{SNR}_2(1 + \text{INR}_3)} \right) \tag{146b}
\]
\[
R_3 = \log \left( \frac{1 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_3} \right). \tag{146c}
\]
For \( \log (1 + \text{INR}_2 + \text{SNR}_2) - R_2 \geq \log (1 + \text{SNR}_1) \), the sum rate is within 3 bits since \( R_1 \) is optimal and \( R_2 + R_3 \) is within 3 bits (2 due to \( Z \), 1 additional due to user 2 backing off). For \( \log (1 + \text{INR}_2 + \text{SNR}_2) - R_2 \leq \log (1 + \text{SNR}_1) \), \( R_1 + R_2 \) is optimal and \( R_3 \) is within 1 bit.
B) $\text{INR}_2 \geq \text{SNR}_2$, $\text{INR}_3 \geq \text{SNR}_3$, $\text{INR}_2 \geq \text{SNR}_3$, $\text{INR}_3 < \text{SNR}_2$: The rates allocated to the users in this case are as follows:

$$R_1 = \min \left( \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right), \log \left( 1 + \frac{\text{INR}_2}{1 + \text{INR}_2} \right) \right)$$

$$R_2 = \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \times \frac{\text{SNR}_3}{1 + \text{SNR}_2 + \text{INR}_3} \right)$$

$$R_3 = \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right). \quad (147a)$$

For $\log(1 + \text{INR}_2) - R_2 \geq \log(1 + \text{SNR}_1)$, the sum rate is within 3 bits since $R_1$ is optimal and $R_2 + R_3$ is within 4 bits (2 due to $Z$ and 2 additional due to user 2 backing off). For $\log(1 + \text{INR}_2) - R_2 \leq \log(1 + \text{SNR}_1)$, $R_1 + R_2$ is within 1 bit and $R_3$ is within 1 bit making the overall sum rate within 2 bits.

C) $\text{INR}_2 \geq \text{SNR}_2$, $\text{INR}_3 \geq \text{SNR}_3$, $\text{INR}_2 < \text{SNR}_1$. $\text{INR}_3 \geq \text{SNR}_2$: The rates allocated to the users in this case are as follows:

$$R_1 = \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right) + \min \left( \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \right), \log \left( 1 + \frac{\text{SNR}_2 + \text{INR}_3}{1 + \text{SNR}_2 + \text{INR}_3} \right) \right) \quad (147b)$$

$$R_2 = \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \times \frac{\text{SNR}_3}{1 + \text{SNR}_2 + \text{INR}_3} \right)$$

$$R_3 = \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right). \quad (147c)$$

For $\log(1 + \text{INR}_2) - R_2 \geq \log(1 + \text{SNR}_1)$, $R_1$ is optimal and $R_2 + R_3$ is within 4 bits (2 due to $Z$ and 2 additional due to user 2 backing off). For $\log(1 + \text{INR}_2) - R_2 \leq \log(1 + \text{SNR}_1)$, $R_1 + R_2$ is within 1 bit and $R_3$ is within 1 bit making the overall sum rate within 2 bits.

D) $\text{INR}_2 \geq \text{SNR}_2$, $\text{INR}_3 \geq \text{SNR}_3$, $\text{INR}_2 < \text{SNR}_1$, $\text{INR}_3 < \text{SNR}_2$: The rates allocated to the users in this case are as follows:

$$R_1 = \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right) + \min \left( \log \left( 1 + \frac{\text{SNR}_2}{1 + \text{INR}_2 + \text{INR}_3} \right), \log \left( 1 + \frac{\text{INR}_3}{1 + 2\text{INR}_2} \right) \right)$$

$$R_2 = \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \times \frac{\text{SNR}_3}{1 + \text{SNR}_2 + \text{INR}_3} \right)$$

$$R_3 = \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right). \quad (150a)$$

If $\log(1 + \text{INR}_2) - R_2 \geq \log(1 + \text{SNR}_1)$, $R_1 + R_2$ is within 1 bit of optimal and $R_3$ is within 1 bit making the overall sum rate within 2 bits.

Thus, the first transmitter transmits at the optimal rate and $R_2$ is within 4 bits.

If $\log(1 + \text{INR}_2) - R_2 \leq \log(1 + \text{SNR}_1)$, $R_1 + R_2$ is within 1 bit of optimal and $R_3$ is within 1 bit making the overall sum rate within 2 bits.

E) $\text{INR}_2 \geq \text{SNR}_2$, $\text{INR}_3 < \text{SNR}_3$, $\text{INR}_2 \geq \text{SNR}_1$: The rates allocated to the users in this case are as follows:

$$R_1 = \min \left( \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right), \log \left( 1 + \frac{\text{SNR}_2}{1 + \text{INR}_2} \right) \right) \quad (151a)$$

$$R_2 = \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \times \frac{\text{SNR}_3}{1 + \text{SNR}_2 + \text{INR}_3} \right)$$

$$R_3 = \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right). \quad (151b)$$

If $\log(1 + \text{INR}_1) \leq \log(1 + \text{INR}_2 + \text{SNR}_3) - R_2$, then user 1 transmits at the optimal rate and $R_2 + R_3$ is within 3 bits of the optimal sum rate and, thus, we achieve within 3 bits of the sum capacity.

If $\log(1 + \text{INR}_1) \geq \log(1 + \text{INR}_2 + \text{SNR}_3)$, $R_1 + R_2$ is within 1 bit of optimal and $R_3$ is within 1 bit, and thus, the sum rate is within 2 bits of optimal.

F) $\text{INR}_2 \geq \text{SNR}_2$, $\text{INR}_3 < \text{SNR}_3$, $\text{INR}_2 < \text{SNR}_1$: The rates allocated to the users in this case are as follows:

$$R_1 = \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right) + \min \left( \log \left( 1 + \frac{\text{SNR}_2}{1 + \text{INR}_2 + \text{INR}_3} \right), \log \left( 1 + \frac{\text{INR}_3}{1 + 2\text{INR}_2} \right) \right)$$

$$R_2 = \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \times \frac{\text{SNR}_3}{1 + \text{SNR}_2 + \text{INR}_3} \right)$$

$$R_3 = \log \left( 1 + \frac{\text{INR}_3}{1 + \text{INR}_3} \right). \quad (152b)$$

If $\log(1 + \text{INR}_2) - R_2 \geq \log(1 + \text{SNR}_1)$, $R_1 + R_2$ is within 1 bit of optimal and $R_3$ is within 1 bit of optimal. Thus, the sum rate is within 2 bits.
If $\log \left( 1 + \frac{1}{1 + \text{SNR}_2} \right) \leq \log \left( 1 + \frac{\text{INR}_2}{1 + 2 \text{SNR}_2} \right) - R_2$, then $R_1 = \log \left( 1 + \frac{1}{1 + \text{SNR}_2} \right)$ and $R_1 + R_2$ is within 3 bits of the optimal, and thus, the sum rate is within 3 bits of optimal.

If $\log \left( 1 + \frac{1}{1 + \text{SNR}_2} \right) \geq \log \left( 1 + \frac{\text{INR}_2}{1 + 2 \text{SNR}_2} \right) - R_2$, $R_1 = \log \left( 1 + \frac{1}{1 + \text{SNR}_2} \right) + \log \left( 1 + \frac{1}{1 + \text{SNR}_2} \right)$, which is within 2 bits of optimal as in case D. Thus, the overall sum rate is within 3 bits of optimal.

G. $\text{INR}_2 < \text{SNR}_2$. The rates allocated to the users in this case are as follows:

$$R_1 = \min \{ \log(1 + \text{SNR}_1) \},$$
$$R_2 = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{SNR}_2} \right),$$
$$R_3 = \log \left( 1 + \frac{1 + \text{INR}_3}{1 + 2 \text{SNR}_3} \right).$$

If $\log \left( 1 + \frac{1}{1 + \text{SNR}_1} \right) \leq \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{SNR}_2} \right) - R_2$, the first user sends at the optimal rate and $R_1 + R_2$ is within 3 bits of optimal, and thus, the sum capacity is within 3 bits of optimal.

If $\log \left( 1 + \frac{1}{1 + \text{SNR}_1} \right) \geq \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{SNR}_2} \right) - R_2$, then $R_1 + R_2$ is within 3 bits of optimal, and thus, the sum capacity is within 1 bit of optimal.

H. $\text{INR}_2 < \text{SNR}_2$. The rates allocated to the users in this case are as follows:

$$R_1 = \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right) + \min \{ \log(1 + \text{SNR}_2) \},$$
$$R_2 = \log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{SNR}_2} \right) \times \min \{ \text{SNR}_2, \frac{\text{INR}_3}{1 + \text{SNR}_3} \},$$
$$R_3 = \log \left( 1 + \frac{1 + \text{INR}_3}{1 + 2 \text{SNR}_3} \right).$$

We consider the four cases when the corresponding term in $R_1$ is the minimum.

1) Minimum is $\log(1 + \text{SNR}_1)$: $R_1$ is optimal and $R_2 + R_3$ is within 4 bits.

2) Minimum is $\log \left( 1 + \frac{1 + \text{INR}_2}{1 + 2 \text{SNR}_2} \right)$: $R_1 + R_2 \geq \log \left( 1 + \frac{1 + \text{INR}_2 + \text{INR}_3}{1 \nabla \text{INR}_3} \right)$.

3) Minimum is $\log \left( 1 + \frac{1 + \text{INR}_2 + \text{SNR}_2(1 + \text{INR}_2)}{1 + 2 \text{INR}_2} \right)$: $R_1$.

4) Minimum is $\log \left( 1 + \frac{1 + \text{INR}_3}{1 + 2 \text{SNR}_3} \right)$.

The preceding expression is within 2 bits of the optimal sum capacity.
log \left( 1 + \frac{INR_2^2 + SNR_3(INR_2 + 1 + INR_2)}{1 + 2INR_2} \right) - R_{2,c},

log \left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)}{1 + 2INR_2} \right) - R_{2,p}, \log
\left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)}{1 + 2INR_2} \right) - R_2 \right) \quad (160a)

R_2 = \log \left( 1 + \frac{1 + INR_2}{1 + 2INR_2} \times \frac{INR_3^2 + SNR_3(1 + INR_3)}{1 + 2INR_3 + SNR_3(1 + INR_3)} \right) + \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) \quad (160b)

R_3 = \log \left( 1 + \frac{1 + SNR_3(1 + INR_3)(1 + 2INR_3)}{1 + 2INR_3 + SNR_3(1 + INR_3)} \right). \quad (160c)

We consider the four cases when the corresponding term in R_{1,c} is the minimum.

1) Minimum is \( \log \left( 1 + \frac{INR_2^2 + SNR_3(INR_2 + 1 + INR_2)}{1 + 2INR_2} \right) \): \( R_1 = \log \left( 1 + \frac{INR_2^2}{1 + 2INR_2 + 1 + INR_3} \right) \) within 1 bit of \( \log(1 + SNR_1) \) (will be shown in Appendixes F–N). Moreover, \( R_2 + R_3 \) is within 2 bits of \( \log(1 + INR_3 + SNR_3) \) (shown in Appendixes F–I). \( R_{2,p} \) is within 1 bit of \( \log(1 + SNR_2) - \log(1 + 1 + INR_3) \). Thus, \( R_2 + R_3 \) is within 3 bits of optimal and \( R_1 \) within 1 bit of optimal; thus, the overall sum rate is within 4 bits of optimal.

2) Minimum is \( \log \left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)(1 + 2INR_2)}{1 + 2INR_2 + 1 + INR_3} \right) \): \( R_{2,c} \). In this case,

\( R_{1,c} + R_2 \geq \log \left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)}{1 + 2INR_2} \right) \).

This sum rate will be larger than in the fourth case, and is therefore, omitted here.

3) Minimum is \( \log \left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)(1 + 2INR_2)}{1 + 2INR_2 + 1 + INR_3} \right) \): \( R_{2,p} \).

In this case,

\( R_1 + R_2 + R_3 = \log \left( 1 + \frac{SNR_1}{1 + INR_2} \right) + \log \left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)(1 + 2INR_2)}{1 + 2INR_2 + 1 + INR_3} \right) + \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) \geq \log \left( 1 + \frac{SNR_1}{1 + INR_2} \right) + \log \left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)(1 + 2INR_2)}{1 + 2INR_2 + 1 + INR_3} \right) + \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) - 2 \).

We consider the following two cases.

1) If \( INR_3(INR_2 + 1) \geq \frac{SNR_2}{1 + INR_2} \):

\( \log \left( 1 + \frac{SNR_1}{1 + INR_2} \right) + \log \left( 1 + \frac{INR_2^2 + SNR_3(1 + INR_2)(1 + 2INR_2)}{1 + 2INR_2 + 1 + INR_3} \right) \geq \log \left( 1 + \frac{SNR_1}{1 + INR_2} \right) + \log \left( 1 + \frac{INR_2^2}{1 + 2INR_2 + 1 + INR_3} \right) + \log(1 + SNR_3) - 1. \)

Thus, the rate within 3 bits of sum capacity can be achieved.

2) If \( INR_3(INR_2 + 1) \leq \frac{SNR_2}{1 + INR_2} \):

\( \log \left( 1 + \frac{INR_2^2 + SNR_2(1 + INR_2)(1 + 2INR_2)}{1 + 2INR_2 + 1 + INR_3} \right) \geq \log \left( 1 + \frac{INR_2^2}{1 + 2INR_2 + 1 + INR_3} \right) \).

Thus, the rate within 3 bits of sum capacity can be achieved.

4) Minimum is \( \log \left( 1 + \frac{INR_2^2 + SNR_2(1 + INR_2)}{1 + 2INR_2 + 1 + INR_3} \right) \): \( R_2 \). In this case, \( R_1 + R_3 \) is within 1 bit of optimal \( \log(1 + SNR_1) + \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) \) and \( R_3 \) is also within 1 bit of optimal; thus, the sum capacity within 2 bits can be achieved.

K) \( INR_2 < SNR_2, INR_3 < SNR_3, INR_2 \geq SNR_1, INR_3 \geq SNR_2 \): The rates allocated to the users in this case are as follows:

\( R_1 = \min \{ \log(1 + SNR_1) \}

\log \left( 1 + \frac{SNR_2}{1 + INR_2} \right) - R_2 \right) \quad (161a)

\( R_2 = \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) \quad (161b)

\( R_3 = \log \left( 1 + \frac{1 + SNR_3}{1 + 2INR_3} \right) \).

\( \frac{SNR_1}{1 + INR_2} \leq \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) - R_2 \), \( R_2 + R_3 \) is within 3 bits of optimal, and thus, the sum rate within 3 bits can be achieved.

\( \frac{SNR_1}{1 + INR_2} \geq \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) - R_2 \).

\( R_1 + R_2 + R_3 \geq \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) \geq \log(1 + SNR_1) + R_3 \geq \log(1 + SNR_1) + \log(1 + SNR_3) - 1 \geq \log(1 + SNR_3) + \log(1 + SNR_3 + INR_3) - 2. \)

Thus, the sum capacity within 2 bits can be achieved.

L) \( INR_2 < SNR_2, INR_3 < SNR_3, INR_2 \geq SNR_1, INR_3 < SNR_2 \): Note that we will focus on \( SNR_2 \leq INR_3(INR_2 + 1) \), since for \( SNR_2 \geq INR_3(INR_2 + 1) \), 1.5 round was already within 4 bits. The rates allocated to the users in this case are as follows:

\( R_1 = \min \{ \log(1 + SNR_1) \}

\log \left( 1 + \frac{SNR_2}{1 + INR_2} \right) - R_2 \right) \quad (162a)

\( R_2 = \log \left( 1 + \frac{1 + INR_2 \cdot SNR_3}{1 + 2INR_2 + 1 + INR_3} \right) \quad (162b)

\( R_3 = \log \left( 1 + \frac{1 + INR_3}{1 + 2INR_3} \right) \).

(162c)
If \( \log(1 + \text{SNR}_3) \leq \log \left( 1 + \text{INR}_2 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right) - R_2, R_2 + R_3 \) is within 3 bits of optimal, and thus, the sum rate within 3 bits can be achieved.

If \( \log(1 + \text{SNR}_1) \geq \log \left( 1 + \text{INR}_2 + \frac{\text{SNR}_3}{1 + \text{INR}_3} \right) - R_2, \)

\[
R_1 + R_2 + R_3 \geq \log \left( 1 + \frac{\text{INR}_2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} \right) + R_3
\]

\[
\geq \log(1 + \text{INR}_2) + R_3 \geq \log(1 + \text{INR}_2) + \log(1 + \text{SNR}_3) - 1.
\]

Thus, the sum rate within 2 bits can be achieved.

\( \text{If } \text{INR}_2 < \text{SNR}_2, \text{INR}_3 < \text{SNR}_3, \text{INR}_2 < \text{SNR}_3, \text{INR}_3 \geq \text{SNR}_2 \): The rates allocated to the users in this case are as follows:

\[
R_1 = \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right) + \min \left( \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \right), \log(1 + \frac{\text{INR}_2^2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} - \text{INR}_2) \right)
\]

\[
R_2 = \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \right) + \log(1 + \frac{\text{SNR}_1}{1 + \text{INR}_3}),
\]

\[
R_3 = \log \left( 1 + \frac{\text{INR}_2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} \right).
\]

We subdivide this into the following two cases. If

\[
\log \left( 1 + \frac{\text{INR}_2^2}{1 + \text{INR}_3} \right) \leq \log \left( 1 + \frac{\text{INR}_2^2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} \right) - R_2,
\]

then

\[
R_1 \geq \log \left( 1 + \frac{\text{SNR}_3}{1 + \text{INR}_2} \right) \left( 1 + \frac{\text{INR}_2^2}{1 + 2\text{INR}_2} \right) - 1.
\]

Thus, the sum rate is within 4 bits of the optimal.

\[
\text{If } \log \left( 1 + \frac{\text{INR}_2^2}{1 + \text{INR}_3} \right) \geq \log \left( 1 + \frac{\text{INR}_2^2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} \right)
\]

\[
R_1 + R_2 = \log \left( 1 + \frac{\text{INR}_2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} \right) + \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_3} \right)
\]

\[
\geq \log(1 + \text{INR}_2) + \log \left( 1 + \frac{\text{INR}_2^2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} \right) - \log(1 + \text{INR}_2)
\]

\[
\geq \log(1 + \text{SNR}_1).
\]

Thus, \( R_1 + R_2 + R_3 \geq \log(1 + \text{SNR}_1) + \log(1 + \text{INR}_2) + \log(1 + \text{SNR}_3 + \text{INR}_3) - 1 \).

The rates allocated to the users in this case are as follows:

\[
R_1 = \log \left( 1 + \frac{\text{SNR}_1}{1 + \text{INR}_2} \right) + \min \left( \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \right), \log(1 + \frac{\text{INR}_2^2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} - \text{INR}_2) \right)
\]

\[
R_2 = \log \left( 1 + \frac{\text{INR}_2}{1 + 2\text{INR}_2} \right) + \log(1 + \frac{\text{SNR}_1}{1 + \text{INR}_3}),
\]

\[
R_3 = \log \left( 1 + \frac{\text{INR}_2 + \text{SNR}_3(1 + \text{INR}_3)}{1 + 2\text{INR}_2} \right).
\]

References

[1] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” IEEE Trans. Inf. Theory, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[2] L. Xie and P. R. Kumar, “A network information theory for wireless communication: Scaling laws and optimal operation,” IEEE Trans. Inf. Theory, vol. 50, no. 5, pp. 748–767, May 2004.

[3] 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.

[4] R. H. Elkik, 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.

[5] G. Bresler and D. Tse, “The two-user Gaussian interference channel: A deterministic view,” Euro. Trans. Telecommun., vol. 19, no. 4, pp. 333–345, Jun. 2008.

[6] T. Richardson and R. Urbanke, Modern Coding Theory. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[7] D. Shah, “Gossip algorithms,” Found. Trends Netw., vol. 3, no. 1, pp. 1–125, 2008.

[8] J. Wu, Handbook of Theoretical and Algorithmic Aspects of Sensor, Ad Hoc Wireless, and Peer to Peer Networks. Boca Raton, FL: Auerbach, 2006.

[9] D. Biskick, “Gaussian Belief Propagation: Theory and Applications,” Ph.D. thesis, Hebrew Univ., Jerusalem, Israel, 2009.

[10] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, “An overview of limited feedback in wireless communication systems,” IEEE J. Sel. Areas Commun., vol. 26, no. 8, pp. 1341–1365, Oct. 2008, Special issue on Exploiting Limited Feedback in Tomorrow’s Wireless Networks.

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

[12] B. Bandemer, G. Vazquez-Vilar, and A. El Gamal, “On the sum capacity of a class of cyclically semidefinite symmetric interference channels,” presented at the IEEE Int. Symp. Inf. Theory Seoul, Korea, Jun. 2009.

[13] 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.

[14] A. El Gamal and M. Costa, “The capacity region of a class of deterministic interference channels,” IEEE Trans. Inf. Theory, vol. 28, no. 2, pp. 343–346, Mar. 1982.
[15] T. S. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” IEEE Trans. Inf. Theory., vol. 27, no. 1, pp. 49–60, Jan. 1981.

[16] K. Kobayashi and T. S. Han, “A further consideration on the HK and the CMG regions for the interference channel,” presented at the Inf. Theory Appl. Workshop, San Diego, CA, 2007.

[17] 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.

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

[19] M. H. M. Costa, “On the Gaussian interference channel,” IEEE Trans. Inf. Theory, vol. 31, no. 5, pp. 607–615, Sep. 1985.

[20] G. Kramer, “Outer bounds on the capacity of Gaussian interference channels,” IEEE Trans. Inf. Theory, vol. 50, no. 3, pp. 581–586, Mar. 2004.

[21] F. A. Tobagi and L. Kleinrock, “Packet switching in radio channels: Part II. The hidden node problem in carrier sense multiple-access and the busy-tone solution,” IEEE Trans. Commun., vol. 23, no. 12, pp. 1417–1433, Dec. 1975.

[22] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, and E. Knightly, “Distributed multi-hop scheduling and medium access with delay and throughput constraints,” in Proc. ACM Int. Conf. Mobile Comput. Netw., Rome, Italy, 2001, pp. 200–209.

[23] T. M. Cover and J. A. Thomas, Elements of Information Theory. New York: Wiley, 1991.

Vaneet Aggarwal (S’08–M’11) received the B.Tech. degree in 2005 from the Indian Institute of Technology, Kanpur, India, and the M.A. and Ph.D. degrees in 2007 and 2010, respectively, from Princeton University, Princeton, NJ, all in electrical engineering.

He is currently a Senior Member of Technical Staff at AT&T Labs-Research, Florham Park, NJ. His research interests include applications of information and coding theory to wireless systems and quantum error correction.

Dr. Aggarwal is the recipient of Princeton University’s Porter Ogden Jacobus Honorific Fellowship in 2009.

Youjian Liu (S’98–M’01) received the B.E. degree in electrical engineering from Beijing University of Aeronautics and Astronautics, China, in 1993, the M.S. degree in electronics from Beijing University, China, in 1996, and the M.S. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, OH, in 1998 and 2001, respectively.

He has been an Associate Professor in the Department of Electrical and Computer Engineering, University of Colorado, Boulder, since August 2002. From January 2001 to August 2002, he worked on spacetime communications for 3G mobile communication systems as a Member of the Technical Staff in CDMA System Analysis and Algorithms Group, Wireless Advanced Technology Laboratory, Lucent Technologies, Bell Labs Innovations, Whippany, NJ. His current research interests include network communications, information theory, and coding theory.

Ashutosh Sabharwal (S’91–M’99–SM’04) received the B.Tech. degree from the Indian Institute of Technology, New Delhi, in 1993 and the M.S. and Ph.D. degrees from The Ohio State University, Columbus, in 1995 and 1999, respectively.

He is currently an Associate Professor in the Department of Electrical and Computer Engineering, Rice University, Houston, TX. His research interests include the areas of information theory and communication algorithms for wireless systems.

Dr. Sabharwal was the recipient of Presidential Dissertation Fellowship Award in 1998.