Multicoding Schemes for Interference Channels

Ritesh Kolte, Student Member, IEEE, Ayfer Özgür, Member, IEEE, and Haim Permuter, Senior Member, IEEE

Abstract—The best known inner bound for the two-user discrete memoryless interference channel is the Han–Kobayashi rate region. The coding schemes that achieve this region are based on rate-splitting and superposition coding. In this paper, we develop a multicoding scheme to achieve the same rate region. A key advantage of the multicoding nature of the proposed coding scheme is that it can be naturally extended to more general settings, such as when encoders have state information or can overhear each other. In particular, we extend our coding scheme to characterize the capacity region of the state-dependent deterministic Z-interference channel when noncausal state information is available at the interfering transmitter. We specialize our results to the case of the linear deterministic model with ON/OFF interference, which models a wireless system where a cognitive transmitter is noncausally aware of the times it interferes with a primary transmission. For this special case, we provide an explicit expression for the capacity region and discuss some interesting properties of the optimal strategy. We also extend our multicoding scheme to find the capacity region of the deterministic Z-interference channel when the signal of the interfering transmitter can be overheard at the other transmitter (also known as unidirectional partial cribbing).

Index Terms—Interference channel, multicoding, Z-interference channel, partial cribbing, state information.

I. INTRODUCTION

THE discrete memoryless interference channel (DM-IC) is the canonical model for studying the effect of interference in wireless systems. The capacity of this channel is only known in some special cases e.g. class of deterministic ICs [3], [4], strong interference conditions [5]–[7], degraded conditions [8], [9] and a class of semideterministic ICs [10]. Characterizing the capacity region in the general case has been one of the long standing open problems in information theory. The best known achievable rate region is the so-called Han-Kobayashi scheme, which can be achieved by using schemes that are based on the concepts of rate-splitting and superposition coding [11], [12]. Rate-splitting refers to the technique of splitting the message at a transmitter into a common and a private part, where the common part is decoded at all the receivers and the private part is decoded only at the intended receiver. The two parts of the message are then combined into a single signal using superposition coding, first introduced in [13] in the context of the broadcast channel. In all the special cases where the capacity is known, the Han-Kobayashi region equals the capacity region. However, it has been very recently shown that this inner bound is not tight in general [14].

The first result we present in this paper is to show that the Han-Kobayashi region can be achieved by a multicoding scheme. This scheme does not involve any explicit rate-splitting. Instead, the codebook at each encoder is generated as a multicodebook, i.e. there are multiple codewords corresponding to each message, and the encoder chooses one by checking an appropriate joint typicality condition. The auxiliary random variable in this scheme is not explicitly associated with a designated part of the message to be decoded at both receivers. It’s role is reminiscent of the encoding for state-dependent channels in [16], and the alternative proof of Marton’s achievable rate region for the broadcast channel given in [17]. A key advantage of the multicoding nature is that the new scheme requires almost no change to be applicable to setups in which the canonical interference channel model is augmented to incorporate additional node capabilities such as cognition and state-dependence, while extending the original Han-Kobayashi scheme to such setups requires one to have a multicoding structure at each level of the superposition coding, along with having to change the decoding procedure significantly. We demonstrate this by providing schemes for the aforementioned augmented settings that are nearly identical to the new achievability scheme for the canonical interference channel.

The first setting we consider is when the interference channel is state-dependent and the state-information is available non-causally to one of the transmitters (cognitive transmitter). We know that for a point-to-point state-dependent channel with non-causal state information at the encoder, the optimal achievability scheme due to Gelfand and Pinsker uses multicoding at the encoders. Hence, for state-dependent

1A similar idea, combined with the block-Markov operation of backward encoding, has been recently used in [15] to develop an achievability scheme called distributed-decode-forward for broadcast traffic on relay networks.
interference channels with noncausal state information at the encoders too, we would like to use the idea of multicoding. Since the new achievability scheme that we present for the canonical interference channel already involves multicoding, it requires almost no change to be applicable to the state-dependent setting; the only modification that needs to happen is that the encoder now includes the state sequence in the joint typicality check that it performs. Note that extending the original Han-Kobayashi scheme to this setting on the other hand would require to generate multiple superposition codebooks, each corresponding to a possible realization of the state sequence. In this paper, we focus for simplicity on the case when the state-information is available only at one transmitter, and the cross-link from the other transmitter is weak enough to be ignored, giving rise to the so called Z-interference channel topology. We are able to prove the optimality of the new scheme for this setup when the interference is deterministic. Note however that as mentioned earlier the multicoding-based scheme for the canonical interference channel can be applied near-verbatim even for the general state-dependent interference channel with noncausal state information at the transmitters. The reason we focus on the deterministic Z-interference channel is that we are able to also prove a converse in this case and therefore obtain an exact capacity characterization.

We then specialize our capacity characterization for the state-dependent deterministic Z-interference channel to the case where the channels are governed by the linear deterministic model of [18]. In the recent literature, this model has proven extremely useful for approximating the capacity of wireless networks and developing insights for the design of optimal communication strategies. We consider a linear deterministic Z-interference channel, in which the state of the channel denotes whether the interference link is present or not. When the transmitters are base-stations and the receivers are end-users, this can model the scenario where one of the transmitters is cognitive, for example it can be a central controller that knows when the other Tx-Rx pair will be scheduled to communicate on the same frequency band. When the two Tx-Rx pairs are scheduled to communicate on the same frequency band, this gives an interference channel; when they communicate on different frequency bands each pair gets a clean channel free of interference. Moreover, the cognitive transmitter can know the schedule ahead of time, i.e. the times at which its transmission will be interfering with the second Tx-Rx pair. For this special case, we identify auxiliary random variables and provide an explicit expression for the capacity region. This explicit capacity characterization allows us to identify interesting properties of the optimal strategy. In particular, with single bit level for the linear deterministic channels (which would imply low to moderate SNR for the corresponding Gaussian channels), the sum rate is maximized when the interfering transmitter remains silent (transmits 0’s) at times when it interferes with the second transmission. It then treats these symbols as stuck to 0 and performs Gelfand-Pinsker coding. The second transmitter observes a clean channel at all times and communicates at the maximal rate of 1 bit per channel use. This capacity characterization also reveals that when all nodes are provided with the state information the sum-capacity cannot be further improved. Thus, for this channel, the sum-capacity when all nodes have state information is the same as that when only the interfering encoder has state information.

State-dependent interference channels where the state information is known only to some of the transmitters have received a lot of attention recently, which we mention in this paragraph. Given the inherent difficulty of the problem, works such as [19]–[22] have considered special cases and proposed particular coding schemes. However, exact capacity characterizations have proven difficult. Another line of related work has been the study of cognitive state-dependent ICs [23]–[26]. Here, the term “cognitive” is usually used to mean that the cognitive transmitters know not only the state of the channel but also messages of other transmitters. Note that this assumption is significantly stronger than assuming state information at the transmitter as we do here.

The second setting we consider is when one of the transmitters has the capability to overhear the signal transmitted by the other transmitter, which can be used to induce cooperation between the two transmitters. This is different from having orthogonal communication links (or conferencing) between the encoders, as studied in [27]. Instead, overhearing exploits the natural broadcasting nature of the wireless medium to establish cooperation without requiring any dedicated resources. A variety of different models have been used to capture overhearing [28]–[30], and are known by different names such as cribbing, source cooperation, generalized feedback, cognition etc. We use “partial cribbing” to model the overhearing, in which some deterministic function of the signal transmitted by the non-cognitive transmitter is available at the cognitive transmitter in a strictly causal fashion. Again, for simplicity, we focus on the case of the Z-interference channel, where the cross-link between the non-cognitive transmitter and its undesired receiver is weak enough to be ignored. For this setting, we develop a simple achievability scheme by combining our multicoding-based scheme with block-Markov coding and show that it is optimal for deterministic configurations.

Finally, to further illustrate the point that simple schemes can be obtained for augmented scenarios, we describe two extensions which introduce even more complexity in the model. In the first extension, a third message is introduced in the state-dependent Z-interference channel, which is to be communicated from the interfering transmitter to the interfered receiver. The second extension combines the state-dependent Z-IC and the Z-IC with unidirectional partial cribbing. In both extensions, we are able to obtain simple optimal schemes by naturally extending the multicoding-based achievability schemes.

Organize: We describe the models considered in this paper formally in Section II. The alternate achievability scheme that achieves the Han-Kobayashi region is presented in Sections III. Section IV describes the results concerning the state-dependent setup and section V describes the results concerning the cribbing setup. The two extensions are described in Section VI and we end the paper with a short discussion in Section VII.
II. MODEL

Capital letters, small letters and capital calligraphic letters denote random variables, realizations and alphabets respectively. The tuple \((x(1), x(2), \ldots, x(n))\) and the set \([a, a+1, \ldots, b]\) are denoted by \(x^n\) and \([a : b]\) respectively, and \(T^n_{\epsilon}\) stands for the \(\epsilon\)-strongly typical set of length-\(n\) sequences.

We now describe the channel models considered in this paper.

A. Canonical Interference Channel

The two-user discrete memoryless interference channel \(p(y_1, y_2|x_1, x_2|x_1, x_2)\) is depicted in Fig. 1. Each sender \(j \in \{1, 2\}\) wishes to communicate a message \(M_j\) to the corresponding receiver.

A \((n, 2^{nR_1}, 2^{nR_2}, \epsilon)\) code for the above channel consists of the encoding and decoding functions:

\[
\begin{align*}
  f_{j, i} & : [1 : 2^{nR_1}] \rightarrow X_j, \quad j \in \{1, 2\}, 1 \leq i \leq n, \\
  g_j & : Y^n \rightarrow [1 : 2^{nR_j}], \quad j \in \{1, 2\},
\end{align*}
\]

such that

\[
\Pr \left\{ g(Y^n) \neq M_j \right\} \leq \epsilon, \quad j \in \{1, 2\},
\]

where \(M_1\) and \(M_2\) are assumed to be distributed uniformly in \([1 : 2^{nR_1}]\) and \([1 : 2^{nR_2}]\) respectively. A rate pair \((R_1, R_2)\) is said to be achievable if for every \(\epsilon > 0\), there exists a \((n, 2^{nR_1}, 2^{nR_2}, \epsilon)\) code for sufficiently large \(n\). The capacity region is defined to be the closure of the achievable rate region.

B. State-Dependent Z-Interference Channel

The state-dependent Z-interference channel (S-D Z-IC), described by \(p(y_1|x_1, s)p(y_2|x_1, x_2, s)\) and memoryless state \(p(s)\), is depicted in Fig. 2. The states are assumed to be known noncausally at encoder 1. Each sender \(j \in \{1, 2\}\) wishes to communicate a message \(M_j\) at rate \(R_j\) to the corresponding receiver. For this setting, a \((n, 2^{nR_1}, 2^{nR_2}, \epsilon)\) code consists of the encoding and decoding functions:

\[
\begin{align*}
  f_{j, i} & : [1 : 2^{nR_1}] \times \mathcal{S}^n \rightarrow X_j, \quad 1 \leq i \leq n, \\
  f_{2,i} & : [1 : 2^{nR_1}] \rightarrow X_2, \quad 1 \leq i \leq n, \\
  g_j & : Y^n \rightarrow [1 : 2^{nR_j}], \quad j \in \{1, 2\},
\end{align*}
\]

such that

\[
\Pr \left\{ g(Y^n_j) \neq M_j \right\} \leq \epsilon, \quad j \in \{1, 2\}.
\]

The probability of error, achievable rate pairs \((R_1, R_2)\) and the capacity region are defined in a similar manner as before.

A special case of the S-D Z-IC, which is the injective deterministic S-D Z-IC, is depicted in Fig. 3. The channel output \(Y_1\) is a deterministic function \(y_1(x_1, s)\) of the channel input \(X_1\) and the state \(S\). At receiver 2, the channel output \(Y_2\) is a deterministic function \(y_2(x_2, T_1)\) of the channel input \(X_2\) and the interference \(T_1\), which is assumed to be a deterministic function \(t_1(x_1, S)\). We also assume that if \(y_2\) is given, \(y_2(x_2, t_1)\) is an injective function of \(t_1\), i.e., there exists some function \(g\) such that \(t_1 = g(y_2, x_2)\).

We consider a special case of the injective deterministic S-D Z-IC in detail, which is the modulo-additive S-D Z-IC, depicted in Fig. 4. All channel inputs and outputs take values in the same finite alphabet \(\mathcal{X}\). The state \(S\) is \(\text{Ber}(\lambda)\).

Assume that the state \(S\) is i.i.d. \(\text{Ber}(\lambda)\). A generalization of this model that incorporates multiple levels is also considered subsequently.
C. Z-Interference Channel With Partial Cribbing

The discrete memoryless deterministic Z-interference channel is depicted in Fig. 5. The channel output $Y_1$ is a deterministic function $y_1(X_1)$ of the channel input $X_1$. At receiver 2, the channel output $Y_2$ is a deterministic function $y_2(X_2, T_1)$ of the channel input $X_2$ and the interference $T_1$, which is assumed to be a deterministic function $t_1(X_1)$. We also assume that if $x_2$ is given, $y_2(x_2, t_1)$ is an injective function of $t_1$, i.e. there exists some function $g$ such that $t_1 = g(y_2, x_2)$. Each sender $j \in \{1, 2\}$ wishes to communicate a message $M_j$ at rate $R_j$ to the corresponding receiver.

We assume that encoder 1 can overhear the signal from transmitter 2 strictly causally, which is modeled as partial cribbing with a delay [31]. The partial cribbing signal, which is a function of $X_2$ is denoted by $Z_2$. So $X_{1i}$ is a function of $(M_1, Z_2^{i-1})$ and $X_2i$ is a function of $M_2$.

A $(n, 2^{nR_1}, 2^{nR_2}, \epsilon)$ code for this setting consists of

$$
\begin{align*}
&f_{1,i} : [1 : 2^{nR_1}] \times Z_2^{i-1} \to X_1, \quad 1 \leq i \leq n, \\
f_{2,i} : [1 : 2^{nR_2}] \to X_2, \quad 1 \leq i \leq n, \\
g_j : X_j^n \to [1 : 2^{nR_j}], \quad j \in \{1, 2\},
\end{align*}
$$

such that

$$
\Pr \left\{ g(Y_j^n) \neq M_j \right\} \leq \epsilon, \quad j \in \{1, 2\}.
$$

The probability of error, achievable rate pairs $(R_1, R_2)$ and the capacity region are defined in a similar manner as before.

III. CANONICAL INTERFERENCE CHANNEL

A. Preliminaries

The currently best known achievable rate region for the 2-user DM-IC was provided by Han and Kobayashi in [11], using a scheme based on rate-splitting and superposition coding. An alternative achievable rate region that included the Han-Kobayashi rate region was proposed in [12], using another scheme that used rate-splitting and superposition coding. Using the terminology introduced in [32], the encoding in [11] can be described as employing homogeneous superposition coding, while that in [12] can be described as employing heterogeneous superposition coding. It was then proved in [33] that the two regions are, in fact, equivalent and given by the following compact representation (see also [34], [35]).

**Theorem 1 (Han-Kobayashi Region):** A rate pair $(R_1, R_2)$ is achievable for the DM-IC $p(y_1, y_2|x_1, x_2)$ if

$$
\begin{align*}
R_1 &< I(X_1; Y_1|U_2, Q), \\
R_2 &< I(X_2; Y_2|U_1, Q), \\
R_1 + R_2 &< I(X_1; Y_1|U_1, U_2) + I(X_2; U_1; Y_2|Q), \\
R_1 + R_2 &< I(X_1, U_2; Y_1|U_1) + I(X_2, U_1; Y_2|U_2, Q), \\
R_1 + R_2 &< I(X_1, U_2; Y_1|U_1, U_2, Q) + I(X_2; U_1; Y_2|U_2, Q),
\end{align*}
$$

for some pmf $p(q)p(u_1, x_1|q)p(u_2, x_2|q)$, where $|U_1| \leq |X_1| + 4$, $|U_2| \leq |X_2| + 4$ and $|Q| \leq 4$.

B. Outline of the New Achievability Scheme

We first describe the alternative achievable scheme informally and discuss the similarities and differences with the existing achievability schemes. The later subsections describe and analyze the scheme formally.

Encoder $j$, where $j \in \{1, 2\}$ prepares two codebooks:

- A transmission multicodebook, which is a set of codewords $\{x_j^n(\cdot, \cdot)\}$ formed using the transmission random variable $X_j$. This set is partitioned into a number of bins (or subcodebooks), where the bin-index corresponds to the message,

- A coordination codebook which is a set of codewords $\{u_j^n(\cdot)\}$ formed using the auxiliary random variable $U_j$.

Given a message, one codeword $x_j^n$ from the corresponding bin in the transmission multicodebook is chosen so that it is jointly typical with some codeword $u_j^n$ in the coordination codebook. The codeword $x_j^n$ so chosen forms the transmission codeword.

At a decoder, the desired message is decoded by using joint typicality decoding, which uses the coordination codebook and the transmission multicodebook of the corresponding encoder and the coordination codebook of the other encoder. Thus, a receiver makes use of the interference via its knowledge of the coordination codebook at the interfering transmitter.

From the above description, it can be seen that the coordination codebook does not carry any explicit part of the message. Its purpose is to ensure that the transmission codeword from a given bin is well-chosen, i.e. it is beneficial to the intended receiver and also the unintended receiver. To the best of our knowledge, this is the first time an auxiliary random variable (which is not the time-sharing random variable) appears in one of the best known achievability schemes without being explicitly associated with any message.

C. Achievability Scheme

Choose a pmf $p(u_1, x_1)p(u_2, x_2)$ and $0 < \epsilon' < \epsilon$.  

2The term “multicodebook” refers to the fact that there are multiple codewords corresponding to each message.
Codebook Generation:
- Encoder 1 generates a coordination codebook consisting of $2^{nR_1}$ codewords $u_1^n(l_{1c})$, $l_{1c} \in [1 : 2^{nR_1}]$ i.i.d. according to $\prod_{i=1}^{n} p(u_{1i})$. It also generates a transmission multicodebook consisting of $2^{n(R_1+R_{1p})}$ codewords $x_1^n(m_1, l_{1p})$, $m_1 \in [1 : 2^{nR_1}]$, $l_{1p} \in [1 : 2^{nR_{1p}}]$ i.i.d. according to $\prod_{i=1}^{n} p(x_{1i})$.
- Similarly, encoder 2 generates a coordination codebook consisting of $2^{nR_2}$ codewords $u_2^n(l_{2c})$, $l_{2c} \in [1 : 2^{nR_2}]$ i.i.d. according to $\prod_{i=1}^{n} p(u_{2i})$. It also generates a transmission multicodebook consisting of $2^{n(R_2+R_{2p})}$ codewords $x_2^n(m_2, l_{2p})$, $m_2 \in [1 : 2^{nR_2}]$, $l_{2p} \in [1 : 2^{nR_{2p}}]$ i.i.d. according to $\prod_{i=1}^{n} p(x_{2i})$.

Encoding:
- To transmit message $m_1$, encoder 1 finds a pair $(l_{1c}, l_{1p})$ such that
  \[ (u_1^n(l_{1c}), x_1^n(m_1, l_{1p})) \in T_{e_1}^{(n)} \]
  and transmits $x_1^n(m_1, l_{1p})$. If it cannot find such a pair, it transmits $x_1^n(m_1, 1)$. 
- Similarly, to transmit message $m_2$, encoder 2 finds a pair $(l_{2c}, l_{2p})$ such that
  \[ (u_2^n(l_{2c}), x_2^n(m_2, l_{2p})) \in T_{e_2}^{(n)} \]
  and transmits $x_2^n(m_2, l_{2p})$. If it cannot find such a pair, it transmits $x_2^n(m_2, 1)$.

The codebook generation and encoding process are illustrated in Fig. 6.

Decoding:
- Decoder 1 finds the unique $\hat{m}_1$ such that
  \[ (u_1^n(l_{1c}), x_1^n(\hat{m}_1, l_{1p}), u_2^n(l_{2c}), y_1^n) \in T_{e_1}^{(n)} \]
  for some $(l_{1c}, l_{1p}, l_{2c})$. If none or more than one such $\hat{m}_1$ are found, then decoder 1 declares error.
- Decoder 2 finds the unique $\hat{m}_2$ such that
  \[ (u_2^n(l_{2c}), x_2^n(\hat{m}_2, l_{2p}), u_1^n(l_{1c}), y_2^n) \in T_{e_2}^{(n)} \]
  for some $(l_{2c}, l_{2p}, l_{1c})$. If none or more than one such $\hat{m}_2$ are found, then decoder 2 declares error.

Discussion: Before providing the formal analysis of the probability of error to show that the coding scheme described above achieves the Han-Kobayashi region, we discuss the connection between the new scheme and the scheme from [12] which motivates the equivalence of their rate regions.

Consider the set of codewords used at encoder 1. While this set resembles a multicodebook, it can be reduced to a standard codebook (one codeword per message) by stripping away the codewords in each bin that are not jointly typical with any of the $u_1^n$ codewords, and therefore are never used by the transmitters. In other words, after we generate the multicodebook in Fig. 6, we can form a smaller codebook by only keeping one codeword per message which is jointly typical with one of the $u_1^n$ codewords (i.e., those codewords highlighted in Fig. 6). Note that this reduced codebook indeed has a superposition structure. Each of the $2^{nR_1}$ remaining codewords $x_1^n$ is jointly typical with one of the $2^{nR_{1p}} u_1^n$ codewords, and when $n$ is large there will be exactly $2^{n(R_1-R_{1p})}$ $x_1^n$ codewords that are typical with each $u_1^n$ codeword, i.e., these $2^{n(R_1-R_{1p})}$ $x_1^n$ codewords will look as if they were generated i.i.d. from $p(x_1|u_1)$. Therefore, the $u_1^n$ codewords can be indeed thought as the cloud centers in this superposition codebook and $x_1^n$’s as the satellite codewords. Therefore, our multicodebook construction can be viewed as an equivalent way to generate a superposition codebook as in [33]. This reveals that both the codebook structure and the decoding in our scheme are similar to that in the Han-Kobayashi scheme and therefore the two achievable rate regions are, not surprisingly, equal.

However, note that for broadcast channels, combining Marton coding (which employs multicoding) [17] with Gelfand-Pinsker coding (which also employs multicoding) is more straightforward than combining superposition coding with Gelfand-Pinsker coding. The former has been shown to be optimal in some cases [36]. Since our codebook construction for the interference channel also has the flavor of multicoding, extending this construction to setups where multicoding is required is also quite straightforward. As mentioned in the introduction, we exploit this to develop simple achievability schemes for more general setups described in later sections.
Probability of Error: Due to the symmetry of the code, the average probability of error $P(E)$ is equal to $P(E|M_1, M_2)$, so we can assume $(M_1, M_2) = (1, 1)$ and analyze $P(E|1, 1)$. Let $(L_{1c}, L_{1p}, L_{2c}, L_{2p})$ denote the indices chosen during encoding by encoder 1 and encoder 2.

We now define events that cover the set of events in decoding message $m_1$:

$$E_1 \triangleq \{(U_1^n(l_{1c}), X_1^n(1, l_{1p})) \notin T^{(n)}_c \text{ for all } l_{1c}, l_{1p}\},$$

$$E_2 \triangleq \{(U_1^n(L_{1c}), X_1^n(1, L_{1p}), U_2^n(L_{2c}), Y_1^n) \notin T^{(n)}_c \},$$

$$E_3 \triangleq \{(U_1^n(L_{1c}), X_1^n(m_1, l_{1p}), U_2^n(L_{2c}), Y_1^n) \in T^{(n)}_c \text{ for some } m_1 \neq 1, \text{ for some } l_{1p}, l_{1c}\},$$

$$E_4 \triangleq \{(U_1^n(L_{1c}), X_1^n(m_1, l_{1p}), U_2^n(L_{2c}), Y_1^n) \in T^{(n)}_c \text{ for some } m_1 \neq 1, \text{ for some } l_{1p}, l_{1c}\},$$

$$E_5 \triangleq \{(U_1^n(l_{1c}), X_1^n(m_1, l_{1p}), U_2^n(l_{2c}), Y_1^n) \in T^{(n)}_c \text{ for some } m_1 \neq 1, \text{ for some } l_{1p}, l_{1c}\},$$

$$E_6 \triangleq \{(U_1^n(l_{1c}), X_1^n(m_1, l_{1p}), U_2^n(l_{2c}), Y_1^n) \in T^{(n)}_c \text{ for some } m_1 \neq 1, \text{ for some } l_{1c}, l_{1p}, l_{2c}\}.$$

Consider also the event $E'_1$, analogous to $E_1$, which is defined as follows.

$$E'_1 \triangleq \{(U_2^n(l_{2c}), X_2^n(l_{2p})) \notin T^{(n)}_c \text{ for all } l_{2c}, l_{2p}\}.$$

Since an error for $m_1$ occurs only if at least one of the above events occur, we use the union bound to get the following upper bound on the average probability of error in decoding $m_1$:

$$P(E_1) + P(E'_1) + P(E_2 \cap E_1') \cap E_1'^c) + P(E_3) + P(E_4) + P(E_5) + P(E_6).$$

By the mutual covering lemma [37, Ch. 8], $P(E_1) \to 0$ as $n \to \infty$ if

$$R_{1p} + R_{1c} < I(U_1; X_1) + \delta(e'),$$

where $\delta(e') \to 0$ as $e' \to 0$.

Similarly, we get $P(E'_1) \to 0$ as $n \to \infty$ if

$$R_{2p} + R_{2c} < I(U_2; X_2) + \delta(e').$$

By the conditional typicality lemma, $P(E_2 \cap E_1' \cap E_1'^c)$ tends to zero as $n \to \infty$.

For $P(E_3) \to 0$, we can use the packing lemma from [37, Ch. 3] to get the condition

$$R_1 + R_{1p} < I(X_1; U_1, U_2, Y_1) - \delta(e),$$

where $\delta(e) \to 0$ as $e \to 0$.

For $P(E_4) \to 0$, we can again use the packing lemma to get the condition

$$R_1 + R_{1p} + R_{2c} < I(X_1; U_2; U_1, Y_1) - \delta(e).$$

For $P(E_5) \to 0$, we apply the multivariate packing lemma from the Appendix as shown in (54) to get the condition

$$R_1 + R_{1p} + R_{1c} < I(U_1; X_1) + I(U_1, X_1; U_2, Y_1) - \delta(e).$$

Finally, for $P(E_6) \to 0$ as $n \to \infty$, another application of the multivariate packing lemma as shown in (55) gives the condition

$$R_1 + R_{1p} + R_{1c} + R_{2c} < I(U_1; X_1) + I(U_2; Y_1) + I(U_1, X_1; U_2, Y_1) - \delta(e).$$

A similar analysis leads to the following additional conditions for the probability of error in decoding $m_2$ to vanish as $n \to \infty$.

$$R_2 + R_{2p} < I(X_2; U_2, U_1, Y_2) - \delta(e),$$

$$R_2 + R_{2p} + R_{1c} < I(X_2, U_1; U_2, Y_2) - \delta(e),$$

$$R_2 + R_{2p} + R_{2c} < I(U_2; X_2) + I(U_2, X_2; U_1, Y_2) - \delta(e),$$

$$R_2 + R_{2p} + R_{2c} + R_{1c} < I(U_2; X_2) + I(U_1; Y_2) + I(U_2, X_2; U_1, Y_2) - \delta(e).$$

Hence the probability of error vanishes as $n \to \infty$ if the conditions (2)-(11) are satisfied. For the sake of brevity, let us first denote the RHS of the conditions (2)-(11) by $a, b, c, d, e, f, g, h, i, j$ respectively (ignoring the $\delta(e')$ and $\delta(e)$ terms).

We then note the following relations among these terms which can be proved using the chain rule of mutual information, the Markov chains $U_1 - X_1 - (U_2, X_2, Y_1, Y_2)$ and $U_2 - X_2 - (U_1, X_1, Y_1, Y_2)$ and the independence of $(U_1, X_1)$ and $(U_2, X_2)$.

$$e - a \leq \min(c, d),$$

$$f - a \leq d \leq f,$$

$$c \leq e \leq f,$$

$$i \leq b \leq \min(g, h),$$

$$j - b \leq h \leq j,$$

$$g \leq i \leq j.$$

We now employ Fourier-Motzkin elimination on the conditions (2)-(11) and $R_{1c}, R_{1p}, R_{2c}, R_{2p} \geq 0$ to eliminate $R_{1c}, R_{1p}, R_{2c}, R_{2p}$. The set of relations (12) can be used to simplify this task by recognizing redundant constraints. At the end, we get the following achievable region:

$$R_1 < e - a,$$

$$R_2 < i - b,$$

$$R_1 + R_2 < c + j - a - b,$$

$$R_1 + R_2 < d + h - a - b,$$

$$R_1 + R_2 < f + g - a - b,$$

$$2R_1 + R_2 < c + h + f - 2a - b,$$

$$R_1 + 2R_2 < d + g + j - a - 2b.$$

Using the same facts as those used to prove (12), we can show that the above region is the same as the Han-Kobayashi region. For the sake of completeness, we show this explicitly in Appendix I.
Then, by including a time-sharing random variable $Q$, we get that the following region is achievable:

$$R_1 < I(X_1; Y_1|U_2, Q),$$
$$R_2 < I(X_2; Y_2|U_1, Q),$$
$$R_1 + R_2 < I(X_1; Y_1|U_1, U_2, Q) + I(X_2, U_1; Y_2|U_2, Q),$$
$$R_1 + 2R_2 < I(X_1; Y_1|U_1, U_2, Q) + I(X_2, U_1; Y_2|U_2, Q)$$
+ $I(X_2, U_1; Y_1|U_1, Q)$,
$$R_1 + 2R_2 < I(X_2; Y_2|U_1, U_2, Q) + I(X_1, U_2; Y_1|U_1, Q)$$
+ $I(X_2, U_1; Y_2|U_2, Q),$ (14)

for pmf $p(q)p(u_1, x_1|q)p(u_2, x_2|q)$. This region is identical to the region in (1).

IV. STATE-DEPENDENT INTERFERENCE CHANNELS

In this section, we focus on the particular setup of the state-dependent Z-interference channel (S-D Z-IC) with noncausal state information at the interfering transmitter, as depicted in Fig. 2. We provide a simple achievability scheme for this setup, that is obtained by modifying only slightly the achievability scheme that we have developed for the general interference channel. The modification, apart from the choice of pmf, is simply that the informed encoder now includes the observed state sequence when it performs the joint typicality encoding, the rest of the scheme description remains the same. This scheme is shown to be optimal for the deterministic case. The auxiliary random variable used for encoding at the interfering transmitter now implicitly captures some part of the message as well as some part of the state sequence realization.

We mention here that the advantages of multicoading in terms of getting simple achievability schemes are not restricted to the deterministic Z-setups. They hold for general setups, and the reason we focus on the deterministic Z-setups is that we can prove converses, which suggests the merits of the new scheme. The achievability scheme can also be viewed as a generalization of the schemes presented in [22] and [38].

After characterizing the capacity region of the deterministic S-D Z-IC, we investigate a special case in detail: the modulo-additive S-D Z-IC. The modulo-additive channel is motivated by the linear deterministic model which has gained popularity over the recent years for studying wireless networks [18]. For this case (which can be thought of as a linear deterministic model with only one bit level), we obtain an explicit description of the capacity region and furthermore, show that the capacity region is also achieved by the standard Gelfand-Pinsker coding over the first link and treating interference as noise over the second link. Following this, the modulo-additive S-D Z-IC with multiple levels is considered and some discussion is provided about the capacity region and the performance of simple achievability schemes.

To summarize, this section contains the following contributions:

- An achievable rate region for the S-D Z-IC,
- Capacity region of the injective deterministic S-D Z-IC,
- Modulo-additive S-D Z-IC: optimality of treating interference-as-noise and other properties.

A. Results for the State-Dependent Channel

The following theorem provides an inner bound to the capacity region of the S-D Z-IC in Fig. 2.

**Theorem 2:** A rate pair $(R_1, R_2)$ is achievable for the S-D Z-IC in Fig. 2 if

$$R_1 < I(U; Y_1|Q) - I(U; S|Q),$$
$$R_2 < I(Y_2; V|Q),$$
$$R_1 + 2R_2 < I(V, X_2; Y_2|Q) - I(V; S|Q),$$
$$R_1 + 2R_2 < I(U; Y_1|Q) + I(V, X_2; Y_2|Q)$$
$$- I(U; S|Q) - I(U; S; V|Q).$$ (15)

for some pmf $p(u, v|x, s)p(x_1|u, v, s, q)p(x_2|q)$.

For the injective deterministic S-D Z-IC, we can identify natural choices for the auxiliary random variables in Theorem 2 that, in fact, yield the capacity region. This result is stated in the following theorem.

**Theorem 3:** The capacity region of the injective deterministic S-D Z-IC in Fig. 3 is the set of rate pairs $(R_1, R_2)$ that satisfy

$$R_1 \leq H(Y_1|S, Q),$$
$$R_2 \leq H(Y_2|T_1, Q),$$
$$R_2 \leq H(Y_2|Q) - I(T_1; S|Q),$$
$$R_1 + 2R_2 \leq H(Y_1|T_1, S, Q) + H(Y_2|Q) - I(T_1; S|Q),$$ (16)

for some pmf $p(q)p(x_1|s, q)p(x_2|q)$, where $|Q| \leq 4$.

**Remark 1:** Note that the capacity region remains unchanged even if the first receiver is provided with the state information. The proof of this theorem is presented in subsection IV-C.

B. Proof of Theorem 2

Fix $p(u, v|x)p(x_1|u, v, s)p(x_2)$ and choose $0 < \epsilon' < \epsilon$.

**Codebook Generation:**

- Encoder 2 generates $2^{nR_2}$ codewords $x_2^n(m_2), m_2 \in [1 : 2^{nR_2}]$ i.i.d. according to $p(x_2)$.
- Encoder 1 generates $2^{n(R_1+R_2)}$ codewords $u^n(m_1, 1)$ i.i.d. according to $p(u)$, where $m_1 \in [1 : 2^{nR_1}]$ and $l_1 \in [1 : 2^{nR_2}]$. Encoder 1 also generates $2^{nR_2}$ codewords $v^n(l_2), l_2 \in [1 : 2^{nR_2}]$ i.i.d. according to $p(v)$.

**Encoding:**

- To transmit message $m_2$, encoder 2 transmits $x_2^n(m_2)$.
- Assume that the message to be transmitted by encoder 1 is $m_1$. After observing $s^n$, it finds a pair $(l_1, l_2)$ such that $(u^n(m_1, l_1), v^n(l_2), s^n) \in T_{e^n}$.

**Decoding:**

- Decoder 1 finds a unique $\hat{m}_1$ such that $(u^n(\hat{m}_1, l_1), y_1^n) \in T^n_{e^n}$ for some $l_1$.
- Decoder 2 finds a unique $\hat{m}_2$ such that $(x_2^n(\hat{m}_2), v^n(l_2), y_2^n) \in T^n_{e^n}$ for some $l_2$. 
Probability of Error: Due to the symmetry of the code, the average probability of error \( P(\mathcal{E}) \) is equal to \( P(\mathcal{E}|M_1, M_2) \), so we can assume \((M_1, M_2) = (1, 1)\) and analyze \( P(\mathcal{E}|1, 1) \). Let \((L_1, L_2)\) denote the pair of indices chosen by encoder 1 such that \((U^n(1, L_1), V^n(L_2), S^n) \in T_1^n\).

We now define events that cover the error event:

\[
\begin{align*}
\mathcal{E}_1 & \triangleq \{(U^n(1, l_1), V^n(l_2), S^n) \notin T_1^n \text{ for all } l_1, l_2\}, \\
\mathcal{E}_2 & \triangleq \{(U^n(1, L_1), Y^n_1) \notin T_1^n\}, \\
\mathcal{E}_3 & \triangleq \{(U^n(m_1, l_1), Y^n_1) \in T_1^n \text{ for some } m_1 \neq 1, l_1\}, \\
\mathcal{E}_4 & \triangleq \{(X^n_2(1), V^n(L_2), Y^n_2) \notin T_1^n\}, \\
\mathcal{E}_5 & \triangleq \{(X^n_2(m_2), V^n(l_2), Y^n_2) \in T_1^n \text{ for some } m_2 \neq 1, l_2\}.
\end{align*}
\]

Since an error occurs only if at least one of the above events occur, we have the following upper bound on the average probability of error:

\[
P(\mathcal{E}) \leq P(\mathcal{E}_1) + P(\mathcal{E}_2 \cap \mathcal{E}_3^c) + P(\mathcal{E}_4 \cap \mathcal{E}_5^c) + P(\mathcal{E}_5).
\]

Similar to the proof of the mutual covering lemma [37, Ch. 8], we can show that \( P(\mathcal{E}_1) \rightarrow 0 \) as \( n \rightarrow \infty \) if

\[
\begin{align*}
R'_1 & > I(U; S) + \delta(\epsilon'),
\end{align*}
\]

\[
\begin{align*}
R'_2 & > I(V; S) + \delta(\epsilon'),
\end{align*}
\]

\[
\begin{align*}
R'_1 + R'_2 & > I(U; S) + I(U; S; V) + \delta(\epsilon'),
\end{align*}
\]

where \( \delta(\epsilon') \rightarrow 0 \) as \( \epsilon' \rightarrow 0 \).

By the conditional typicality lemma [37, Ch. 2], \( P(\mathcal{E}_2 \cap \mathcal{E}_3^c) \) and \( P(\mathcal{E}_4 \cap \mathcal{E}_5^c) \) both tend to zero as \( n \rightarrow \infty \).

By the packing lemma [37, Ch. 3], for \( P(\mathcal{E}_3) \rightarrow 0 \), we require

\[
R_1 + R'_1 < I(U; Y_1) - \delta(\epsilon),
\]

and for \( P(\mathcal{E}_5) \rightarrow 0 \), we require

\[
R_2 < I(X_2; Y_2|V) - \delta(\epsilon),
\]

\[
R_2 + R'_2 < I(V; X_2; Y_2) - \delta(\epsilon),
\]

where \( \delta(\epsilon) \rightarrow 0 \) as \( \epsilon \rightarrow 0 \). Hence, \( P(\mathcal{E}) \rightarrow 0 \) if (17), (18), (19), (20), (22) are satisfied. Allowing coding-time sharing with a time-sharing random variable \( Q \) and eliminating \( R'_1, R'_2 \) via Fourier-Motzkin elimination, we obtain the region (15). \( \blacksquare \)

C. Proof of Theorem 3

Achievability follows from Theorem 2 by choosing \( U = Y_1 \) and \( V = T_1 \). These choices are valid since encoder 1 knows \((M_1, S^n)\), which determines \( T^n_1 \) and \( Y^n_1 \). We now prove the converse.

Given a sequence of codes that achieves reliable communication (i.e., \( P_e^{(n)} \rightarrow 0 \) as \( n \rightarrow \infty \)) at rates \((R_1, R_2)\), we have, by Fano’s inequality:

\[
\begin{align*}
H(M_1|Y^n_1) & \leq n\epsilon_n, \\
H(M_2|Y^n_2) & \leq n\epsilon_n,
\end{align*}
\]

where \( \epsilon_n \rightarrow 0 \) as \( n \rightarrow \infty \).

Using these, we can establish an upper bound on \( R_1 \) as follows,

\[
\begin{align*}
nR_1 & = H(M_1) \\
& = H(M_1|S^n) \\
& \leq I(M_1; Y^n_1|S^n) + n\epsilon_n \\
& \leq H(Y^n_1|S^n) + n\epsilon_n \\
& \leq \sum_{i=1}^n H(Y_{1i}|S_i) + n\epsilon_n.
\end{align*}
\]

A simple upper bound on \( R_2 \) is established in the following:

\[
\begin{align*}
nR_2 & = H(M_2) \\
& = H(M_2|T^n_1) \\
& \leq I(M_2; Y^n_2|T^n_1) + n\epsilon_n \\
& \leq I(Y^n_2|T^n_1) + n\epsilon_n \\
& \leq \sum_{i=1}^n H(Y_{2i}|T_{1i}) + n\epsilon_n.
\end{align*}
\]

For the second upper bound on \( R_2 \), consider the following:

\[
\begin{align*}
nR_2 & = H(M_2) \\
& = H(M_2|T^n_1) \\
& \leq I(M_2; Y^n_2|T^n_1) + n\epsilon_n \\
& \leq I(Y^n_2|T^n_1) + n\epsilon_n \\
& \leq \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - H(T^n_1|M_2) \\
& \overset{(a)}{=} \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - H(T^n_1|M_2) \\
& \overset{(b)}{=} \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - H(T^n_1) \\
& \leq \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - I(T^n_1; S^n) \\
& = \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - H(S^n) + H(T^n_1|S^n) \\
& \leq \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - H(S^n) + \sum_{i=1}^n H(T_{1i}|S_i) \\
& \overset{(c)}{=} \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - \sum_{i=1}^n H(S_i) + \sum_{i=1}^n H(T_{1i}|S_i) \\
& = \sum_{i=1}^n H(Y_{2i}) + n\epsilon_n - \sum_{i=1}^n I(T_{1i}; S_i)
\end{align*}
\]

where step \((a)\) follows by the injectivity property, step \((b)\) follows because \( T^n_1 \) is independent of \( M_2 \), and step \((c)\) follows because \( S^n \) is an i.i.d. sequence.
The capacity region when $\mathcal{X} = \{0, 1\}$ and $S$ is i.i.d. $\text{Ber}(\tfrac{1}{2})$ is shown in Figure 7.

**Proof of Theorem 4:** Consider the capacity region stated in Theorem 3. Let $p_{1,0,}$ $p_{1,1}$ and $p_{2}$, all in $\mathcal{P}_X$, be used to denote the pmf’s $p(x_1|s = 0, q)$, $p(x_1|s = 1, q)$ and $p(x_2|q)$ respectively. Evaluating each of the constraints in $(16)$ gives us the following expression for the capacity region:

$$
R_1 < (1 - \lambda) H(p_{1,0}) + \lambda H(p_{1,1}),
$$
$$
R_2 < H(p_2),
$$
$$
R_2 < H((1 - \lambda)p_2 + \lambda \tilde{p}) + \lambda H(p_{1,1})
$$
$$
- H(\lambda p_{1,1} + (1 - \lambda) \delta_0),
$$
$$
R_1 + R_2 < (1 - \lambda) H(p_{1,0}) + H((1 - \lambda)p_2 + \lambda \tilde{p})
$$
$$
+ \lambda H(p_{1,1}) - H(\lambda p_{1,1} + (1 - \lambda) \delta_0),
$$

where $\tilde{p} \in \mathcal{P}_X$ is a pmf that is defined as

$$
\tilde{p}(k) = \sum_{i=0}^{\lfloor |\mathcal{X}| - 1 \rfloor} p_{1,1}(i)p_2(k - i), \quad 0 \leq k \leq |\mathcal{X}| - 1,
$$

and $k - i$ should be understood to be $(k - i) \mod |\mathcal{X}|$.

Firstly, we note that $p_{1,0}$ should be chosen as the pmf of the uniform distribution to maximize $H(p_{1,0})$, thus maximizing the RHS of the constraints in $(25)$. Similarly, $p_2$ should also be chosen to be the pmf of the uniform distribution. Then, we can also remove the first constraint on $R_2$, since it is rendered redundant by the other constraint on $R_2$. Thus, the capacity region is given by the convex closure of $(R_1, R_2)$ satisfying

$$
R_1 < (1 - \lambda) \log(|\mathcal{X}|) + \lambda H(p_{1,1}),
$$
$$
R_2 < \log(|\mathcal{X}|) + \lambda H(p_{1,1}) - H(\lambda p_{1,1} + (1 - \lambda) \delta_0),
$$
$$
R_1 + R_2 < (2 - \lambda) \log(|\mathcal{X}|) + \lambda H(p_{1,1})
$$
$$
- H(\lambda p_{1,1} + (1 - \lambda) \delta_0),
$$

for $p_{1,1} \in \mathcal{P}_X$.

For any $p$, the region in $(24)$ is contained in the region in $(26)$ for $p_{1,1} = p$. Hence, the convex closure of $(24)$ is contained in the convex closure of $(26)$.

However, also note that the region in $(26)$ for any $p_{1,1}$ is contained in the convex hull of two regions, one obtained by setting $p = p_{1,1}$ in $(24)$ and the other obtained by setting $p = \delta_0$ in $(24)$. Hence, the convex closure of $(26)$ is also contained in the convex closure of $(24)$. This concludes the proof of Theorem 4.

**Remark 2:** The optimal sum-rate $(2 - \lambda) \log(|\mathcal{X}|)$ is achieved by choosing $p = \delta_0$. This corresponds to setting the transmitted symbols of the first transmitter to 0 when $S = 1$ so that it does not interfere with the second transmission. The first transmitter then treats these symbols as stuck to 0 and performs Gelfand-Pinsker coding. The second transmitter transmits at rate $\log(|\mathcal{X}|)$ bits/channel use. It can be easily verified that this is also the optimal sum-rate when all nodes are provided with the state information. Thus, for this channel, the sum-capacity when all nodes have state information is the same as that when only encoder 1 has state information.

**Remark 3:** Finally, we note that there is also another way to achieve the capacity region of the modulo additive S-D Z-IC.
For this, first recall that to get the capacity region expression in Theorem 3, we set the auxiliary random variables $U$ and $V$ in the expression in Theorem 2 to $Y_1$ and $T_1$ respectively. Another choice, which corresponds to standard Gelfand-Pinsker coding for the first transmitter-receiver pair and treating interference as noise at the second receiver is to choose $V = \phi$ in Theorem 2. This gives us the following achievable region:

$$R_1 < I(U; Y_1 | Q) - I(U; S | Q),$$
$$R_2 < I(X_2; Y_2 | Q),$$

for some pmf $p(q)p(u | s, q)p(x_1 | u, s, q)p(x_2 | q)$. We can now see that for the modulo-additive S-D Z-IC, the capacity region is also achieved by making the following choices in the above region: $p(u | s = 0)$ to be the uniform pmf over $X$, $p(u | s = 1)$ to be $p$, $p(x_1 | u, s)$ to be $\delta_{p_0}$ (i.e. $X_1 = U$) and $p(x_2)$ to be the uniform pmf over $X$. Thus, the capacity region of the modulo-additive S-D Z-IC can also be achieved by treating interference as noise at the second receiver.

**E. Multiple-Level Modulo-Additive S-D Z-IC**

The linear deterministic model introduced in [18] consists of multiple bit levels that roughly correspond to bits communicated at different power levels. The modulo-additive S-D Z-IC that we looked at in the previous subsection is a special case in which the number of levels is one. Extending the model to have multiple bit levels raises some interesting questions which we consider in this subsection.

More specifically, consider the model depicted in Fig. 8, which can be thought of as three copies of the model in Fig. 4, which are however related by the common state affecting them. For simplicity, we restrict attention to the case when the alphabet on each level, denoted by $X_i$ is the binary alphabet, i.e. $\{0, 1\}$, and the state is $\text{Ber}(0.5)$. Let $L$ denote the number of bit levels.

This model also falls under the injective-deterministic setup for which we have completely characterized the capacity region. So the capacity region can be easily computed, as we indeed do in the following. This evaluation also allows us to immediately compare the capacity region with the rates achieved by some straightforward achievability schemes that we can employ. In particular, consider the following two simple achievability schemes:

- **“Separation”**: The simplest strategy one can employ is to separately consider each level and communicate over it independently of the other levels. This gives us that the rate pairs $(R_1, R_2)$ satisfying

$$R_1 < \frac{L}{2} + \sum_{i=1}^{L} \frac{1}{2} \mathcal{H}(p_i),$$
$$R_2 < L - \sum_{i=1}^{L} \mathcal{H}(p_i + \delta_0),$$

for some $p_1, p_2, \ldots, p_L \in \mathcal{P}_X$ are achievable.

- **“Communicate state”**: Alternatively, by noticing that strictly better rates could have been achieved if decoder 2 also had access to the state information, we can reserve one level to communicate the state from encoder 1 to decoder 2. This is done by ensuring that encoder 1 transmits a 1 on this reserved level whenever the state is 1, and encoder 2 constantly transmits a 0 on this level. The nodes communicate on the remaining levels keeping in mind that now decoder 2 also has state information.

Note that while no communication can happen between encoder 2 and decoder 2 on the reserved level, encoder 1 can still communicate with decoder 1 at rate 0.5 on this level by treating it as a channel with stuck bits (bit equals 1 whenever state equals 1). This strategy provides us the following achievable region:

$$R_1 < \frac{L}{2} + \frac{1}{2} \mathcal{H}(p),$$
$$R_2 < L - 1 - \frac{1}{2} \mathcal{H}(p),$$

for some $p \in \mathcal{P}_X^{L-1}$.

We can expect that the suboptimality of reserving one level for communicating the state should become relatively small as the number of levels increases i.e. at high SNR. This is corroborated by the numerical analysis, shown in Figs. 9 and 10, in which we can see that there is a marked improvement in the rates achieved by this scheme relative to the capacity region as we increase the number of levels.
encoder 2 only splits its message into a cooperative and non-cooperative private part, encoder 1 does not perform any rate-splitting at all. Encoder 2 needs to perform rate-splitting so that the partial cribbing can be exploited, but other than that, the achievability scheme is the same as the one we have developed for the canonical interference channel, except that the codewords are now generated conditioned on the cribbed sequence from the previous block. By focusing on the specific configuration of the Z-interference channel, we are able to prove the optimality of an achievability scheme that is simpler than the highly involved achievability schemes for the general case that are currently known.

A. Result for Partial Cribbing

Theorem 5: The capacity region of the injective deterministic Z-interference channel with unidirectional partial cribbing, depicted in Fig. 5, is given by the convex closure of \( (R_1, R_2) \) satisfying

\[
R_1 \leq H(Y_1|W), \quad R_2 \leq \min \left( H(Y_2), H(Y_2, Z_2|T_1, W) \right), \quad R_1 + R_2 \leq H(Y_1|T_1, W) + \min \left( H(Y_2), H(Y_2, Z_2|W) \right),
\]

(30)

for \( p(\omega)p(\gamma|x_1|w)p(\gamma|x_2|w) \), where \( W \) is an auxiliary random variable whose cardinality can be bounded as \( |W| \leq |Y_2| + 3 \).

The proof of this theorem is presented below.

B. Proof of Theorem 5

Achievability: Choose a pmf \( p(\omega)p(\gamma_u, \gamma_c, x_1|w) \) \( p(\gamma|x_2, \gamma|w) \) and \( 0 < \epsilon < \epsilon', \) where for the sake of generality, we use the auxiliary random variables \( U_c \) and \( U_e \). In the injective deterministic case at hand, they can be set to \( Y_1 \) and \( T_1 \) respectively.

Codebook Generation: The communication time is divided into \( B \) blocks, each containing \( n \) channel uses, and an independent random code is generated for each block \( b \in [1 : B] \). Whenever it is clear from the context, we suppress the dependence of codewords on \( b \) to keep the notation simple. The messages in block \( B \) are fixed apriori, so a total of \( B - 1 \) messages are communicated over the \( B \) blocks. The resulting rate loss can be made as negligible as desired by choosing a sufficiently large \( B \).

We split \( R_2 \) as \( R_2^c + R_2^d \), which corresponds to the split of message 2 into two parts, one that will be sent cooperatively by both transmitters to receiver 2 and the other non-cooperatively only by transmitter 2 to receiver 2. For each block \( b \), let \( m_2^c, b \in [1 : 2^nR_2^c] \) and \( m_2^d, b \in [1 : 2^nR_2^d] \). For each block \( b \in [1 : B] \), we generate \( 2^nR_2^c \) codewords \( \omega \) i.i.d. according to \( p(\omega) \).

- For each \( \omega \) in block \( b \), we generate \( 2^nR_2^c \) codewords \( \{ z_2^c(\omega, m_2^c, b) \} \) i.i.d. according to \( p(z_2|\omega) \).

Then for each \( (\omega, z_2^c) \), we generate \( 2^nR_2^d \) codewords \( \{ x_2^d(\omega, z_2^c, m_2^d, b) \} \) i.i.d. according to \( p(x_2|z_2, \omega) \).

- For each \( \omega \) in block \( b \), we generate \( 2^nR_c \) codewords \( \{ u_2^c(\omega, l_c) \} \) i.i.d. according to \( p(u_c|\omega) \), where

from 2 to 3. Indeed, since all the levels are affected by the same state, the entropy of the state becomes small compared to the communication rates as the SNR increases, so it is not a big overhead to explicitly communicate the state to decoder 2 at high SNR. However, at low SNR, the figures show that the overhead incurred is quite high due to which this approach is significantly suboptimal, while the simple scheme of treating the levels separately results in achieving very close to the entire capacity region.

V. INTERFERENCE CHANNELS WITH PARTIAL CRIBBING

In this section, we focus on deterministic Z-interference channels when the interfering transmitter can overhear the signal transmitted by the other transmitter after it passes through some channel. This channel is also modeled as a deterministic channel, dubbed as partial cribbing in [31]. Deterministic models, in particular linear deterministic models [18], have gained popularity due to the observation that they are simpler to analyze and are provably close in performance to Gaussian models.

There have been quite a few very sophisticated achievability schemes designed for interference channels with causal cribbing encoders, however optimality of the achievable rate regions has not been addressed. In the most general interference channel model with causal cribbing [29], each encoder needs to split its message into four parts: a common part to be sent cooperatively, a common part to be sent non-cooperatively, a private part to be sent cooperatively and a private part to be sent non-cooperatively. Further, because of the causal nature of cribbing, achievable schemes usually involve block-Markov coding, so that each encoder also needs to consider the cooperative messages of both encoders from the previous block. Motivated by the alternative achievable scheme we have presented earlier for the general interference channel, we present a simple optimal achievability scheme that minimizes the rate-splitting that is required. Specifically, while

![Fig. 10. Comparison of the different rate regions for 3-level binary moduloadditive S-D Z-IC.](image-url)
\[ I_c \in [1 : 2^{nR_c}]. \] We also generate \( 2^{n(R_1 + R_3)} \) codewords \( \{u_d(m_{1,b}, l_d)\} \) i.i.d. according to \( p(u_d) \), where \( m_{1,b} \in [1 : 2^{nR_1}] \) and \( l_d \in [1 : 2^{nR_3}] \).  

**Encoding:** Let us assume for now that as a result of the cribbing, encoder 1 knows \( m_{1,b-1}^i \) at the end of block \( b - 1 \). Then in block \( b \), both encoders can encode \( m_{2,b-1}^i \) using \( w_n(m_{2,b-1}^i) \) where \( w_n \) is from the code for block \( b \).

- To transmit message \( m_{1,b} \), encoder 1 finds a pair \( (l_d, l_c) \) such that
  \[ (w_n(m_{2,b-1}^i), u_c^n(w_n, l_c), u_d^n(m_{1,b}, l_d)) \in T_{e}^{(n)}. \]
  It transmits \( x_n^a \) that is generated i.i.d. according to \( p(x_1|w, u_d, u_c) \).
- To transmit message \( m_{2,b} = (m_{2,b}, m_{2,b}^i) \), encoder 2 encodes \( m_{2,b}^i \) as \( z_n^a(w_n, m_{2,b}^i) \) and then transmits \( x_n^a(w_n, z_n^a, m_{2,b}^i) \).

We fix apriori the messages in block \( B \) to be \( m_{1,B} = 1 \), \( m_{2,B} = 1 \) and \( m_{2,B}^i = 1 \). Also, to avoid mentioning edge cases explicitly, whenever \( m_{1,0}, m_{2,0} \) or \( m_{2,0}^i \) appear, we assume that all are fixed to 1.

**Decoding:**
- **Encoder 1:** At the end of block \( b \), assuming it has already decoded \( m_{2,b-1}^i \) at the end of block \( b - 1 \), encoder 1 decodes \( m_{2,b}^i \) by finding the unique \( \hat{m}_{2,b}^i \) such that the sequence \( z_n^a \) it has observed via cribbing is equal to \( z_n^a(w_n, \hat{m}_{2,b}^i) \).
- **Decoder 1:** In each block \( b \), decoder 1 finds the unique \( \hat{m}_{1,b} \) such that \( (u_d^n(m_{1,b}, l_d), y_n^i) \in T_{e}^{(n)} \) for some \( l_d \).
- **Decoder 2:** Decoder 2 performs backward decoding as follows:
  - In block \( B \), decoder 2 finds a unique \( m_{2,B-1}^i \) such that the condition (31) is satisfied for some \( l_c \).
    \begin{align*}
    (w_n(m_{2,B-1}^i), z_n^a(w_n, 1), x_n^a(w_n, z_n^a, 1), u_c^n(w_n, l_c), y_n^i) \in T_{e}^{(n)}
    \end{align*}
  (31)
  - In block \( b \), assuming \( m_{2,b}^i \) has been decoded correctly, it finds the unique \( \hat{m}_{2,b-1}^i \) such that the condition (32) is satisfied for some \( l_c \).
    \begin{align*}
    (w_n(\hat{m}_{2,b-1}^i), z_n^a(w_n, m_{2,b}^i), x_n^a(w_n, z_n^a, m_{2,b}^i), 
    m_{2,b}^i), u_c^n(w_n, l_c), y_n^i) \in T_{e}^{(n)}
    \end{align*}
    (32)

**Probability of Error:** To get a vanishing probability of error, we can impose the conditions described in the following list.

- Similar to the proof of the mutual covering lemma [37, Ch. 8], we can show that the following conditions are sufficient for the success of encoding at the first transmitter:
  \[ R_d > I(U_d; W) + 2\delta(\epsilon'). \]
  (33)
  \[ R_d + R_c > I(U_d; U_c, W) + 2\delta(\epsilon'). \]
  (34)

- For the decoding at encoder 1 to succeed:
  \[ R'_1 < H(Z_2|W) - 2\delta(\epsilon). \]
  (35)

\[^2\]Note that the \( u_d^n \) codewords are generated independently of the \( w_n \) codewords.

\[ R_1 + R_d < I(U_d; Y_1) - 2\delta(\epsilon). \]
(36)

- For the backward decoding at decoder 2 to succeed, it is sufficient that the following conditions are satisfied:
  \[ R'_2 < I(X_2; Z_2) - 2\delta(\epsilon), \]
  (37)
  \[ R''_2 + R_c < I(U_c, X_2; Z_2) - 2\delta(\epsilon), \]
  (38)
  \[ R'_2 + R''_2 + R_c < I(W, U_c, X_2; Y_2) - 2\delta(\epsilon). \]
  (39)

Noting that \( R'_2 + R''_2 = R_2 \), eliminating \( (R_d, R_c, R'_2, R''_2) \) from (33)-(39) via Fourier-Motzkin elimination, and substituting \( U_d = Y_1 \) and \( U_c = T_1 \), we get the achievable region in (30) with the following additional bound on \( R_1 \):
\[ R_1 < H(Y_1|W, T_1) + H(Y_2|W, Z_2). \]

To conclude the proof of achievability, we show that this bound is rendered redundant by \( R_1 < H(Y_1|W) \) which can be proved by the following chain of inequalities:
\[ H(Y_1|W, T_1) + H(Y_2|W, Z_2) \leq H(Y_1|W) \]
(40)

First, an upper bound on \( R_1 \) is established in (40).
\[ nR_1 = H(M_1) \leq H(M_1|Z_1) \leq I(M_1; Y_1^n) + n\epsilon_n \leq H(Y_1^n) + n\epsilon_n \leq \sum_{i=1}^{n} H(Y_1|W_i) + n\epsilon_n, \]
(40)
and the second one in (42) below:

\[ nR_2 = H(M_2 | M_1) \]
\[ = H(M_2, Z_2^n | M_1) \]
\[ = H(Z_2^n | M_1) + H(M_2 | M_1, Z_2^n) \]
\[ \leq H(Z_2^n | M_1) + I(M_2; Y_2^n | M_1, Z_2^n, T_1^n) + n\epsilon_n \]
\[ \leq H(Z_2^n) + H(Y_2^n | Z_2^n, T_1^n) + n\epsilon_n \]
\[ \leq \sum_{i=1}^{n} H(Z_{2i} | W_i) + \sum_{i=1}^{n} H(Y_{2i} | W_i, T_{1i}, Z_{2i}) + n\epsilon_n, \]
\[ \leq n \]
\[ \sum_{i=1}^{n} H(Y_{2i}) + n\epsilon_n, \]
\[ \sum_{i=1}^{n} H(Y_{2i}) + n\epsilon_n, \] (42)

where step (a) follows because \( T_1^n \) is a function of \( X_1^n \), which is a function of \( (M_1, Z_2^n) \), and step (b) follows because \( Z_{2i} - W_i - T_{1i} \).

Finally, we establish two bounds on the sum-rate, the first one in (43) below:

\[ n(R_1 + R_2) \]
\[ = H(M_1 | Z_1^n) + H(M_2, Z_2^n) \]
\[ = H(M_1, T_1^n | Z_1^n) + H(Z_2^n) + H(M_2 | Z_2^n) \]
\[ = H(T_1^n | Z_1^n) + H(M_1 | T_1^n, Z_1^n, Z_2^n) + H(Z_2^n) + H(M_2 | Z_2^n) \]
\[ \leq H(Y_1^n | X_1^n, Z_1^n) + I(M_1; Y_1^n | T_1^n, Z_1^n) + H(Z_2^n) \]
\[ + I(M_2; Y_2^n | T_1^n, Z_2^n) \]
\[ \leq H(Y_2^n | X_2^n, Z_2^n) + H(Y_2^n | T_1^n, Z_2^n) + H(Y_1^n | X_1^n, Z_1^n) \]
\[ \leq H(Y_2^n | X_2^n, Z_2^n) + H(Y_1^n | T_1^n, Z_2^n) + H(Z_2^n) \]
\[ \leq \sum_{i=1}^{n} H(Y_{1i} | T_{1i}, W_i) + \sum_{i=1}^{n} H(Y_{2i} | Z_{2i} | W_i) + n\epsilon_n, \]
\[ \leq \sum_{i=1}^{n} H(Y_{1i} | T_{1i}, W_i) + \sum_{i=1}^{n} H(Y_{2i} | Z_{2i} | W_i) + n\epsilon_n, \] (43)

where step (a) uses the fact that \( H(T_1^n | Z_2^n) \leq H(Y_2^n | X_2^n, Z_2^n) \), which is proved below:

\[ H(T_1^n | Z_2^n) \leq H(Y_2^n | X_2^n, Z_2^n) \]
\[ \leq H(Y_2^n | X_2^n) \]
\[ \leq H(Y_2^n | Z_2^n), \]

and step (b) follows because \( M_2 - (X_2^n, Z_2^n) - Y_2^n \).

The second bound on the sum-rate is established in (44) as follows:

\[ n(R_1 + R_2) \]
\[ = H(M_1, T_1^n | Z_1^n) + H(M_2) \]
\[ \leq H(T_1^n | Z_1^n) + H(M_1 | T_1^n, Z_1^n, Z_2^n) + I(M_2; Y_2^n) + n\epsilon_n \]
\[ \leq H(Y_2^n | X_2^n, Z_2^n) + I(M_1; Y_1^n | T_1^n, Z_1^n) + H(Y_2^n) \]
\[ \leq H(Y_2^n | T_1^n | Z_2^n) + H(Y_2^n) + n\epsilon_n \]
\[ \leq \sum_{i=1}^{n} H(Y_{1i} | T_{1i}, W_i) + \sum_{i=1}^{n} H(Y_{2i}) + n\epsilon_n. \] (44)

In (40)-(44), we can introduce a time-sharing random variable \( Q \) uniformly distributed on \([1 : n]\). Defining \( W \) to be \((W_Q, Q)\) and \((X_1, X_2, Y_1, Y_2)\) to be \((X_Q, X_{2Q}, Y_Q, Y_{2Q})\), we get the required bounds on the rates.

We also require the Markov relationship \( X_1 - W - X_2 \) to be satisfied. Since \( W_i \) is chosen to be \( Z_{2i} - 1 \), we immediately have \( X_{1i} - W_i - X_{2i} \) and hence \( X_1 - W - X_2 \). Finally, the bound on the cardinality of \( W \) can be established using the standard convex cover method.

VI. EXTENSIONS

A. STATE-DEPENDENT Z-CHANNEL

The result in Theorem 3 can be extended easily to Z-channels in which transmitter 1 also wishes to communicate a message \( M_{21} \) at rate \( R_{21} \) to receiver 2, as shown in Fig. 11.

**Theorem 6:** The capacity region of the injective deterministic state-dependent Z-channel is the set of rate pairs \((R_1, R_{21}, R_2)\) that satisfy

\[ R_1 \leq H(Y_1 | S, Q), \]
\[ R_2 \leq H(Y_2 | T_1, Q), \]
\[ R_{21} \leq H(T_1 | S, Q), \]
\[ R_1 + R_{21} \leq H(Y_1 | T_1, S, Q), \]
\[ R_2 + R_{21} \leq H(Y_2 | Q) - I(T_1; S | Q), \]
\[ R_1 + R_{21} + R_2 \leq H(Y_1 | T_1, S, Q) + H(Y_2 | Q) - I(T_1; S | Q), \]

for some pmf \( p(q) p(x_1 | q) p(x_2 | q) \), where \( |Q| \leq 6 \).

The achievability scheme for this case is similar to the achievability scheme for the Z-IC described in Section IV-C except that now the \( \nu \) codewords at encoder 1 are also binned, and this bin-index corresponds to the message \( M_{21} \), that is to be communicated from transmitter 1 to receiver 2. The converse can be established by following similar steps as the converse for the Z-IC.
Unidirectional Partial Cribbing.

Fig. 12. State-Dependent Injective Deterministic Z-Interference Channel with Unidirectional Partial Cribbing.

Remark 4: Since broadcast channel and multiple-access channel are special cases of the Z-channel, Theorem 6 also provides the capacity region for the deterministic broadcast channel and the injective deterministic multiple-access channel.

B. State-Dependent Injective Deterministic Z-IC With Unidirectional Partial Cribbing

To illustrate further the advantages of the multicoding scheme, we consider a model that combines the state-dependent Z-IC and the Z-IC with unidirectional partial cribbing, as depicted in Fig. 12. We can combine the achievability schemes for the two component setups from Sections IV-C and V-B in a straightforward manner to get an achievable scheme for this setup. It turns out that this is capacity-achieving, resulting in the following theorem.

Theorem 7: The capacity region of the channel in Fig. 12 is given by the convex closure of rate pairs \((R_1, R_2)\) satisfying

\[

e - a = I(U_1; X_1; U_2, Y_1) \leq (a) I(X_1; U_2, Y_1) \leq (b) I(X_1; Y_1 U_2),
\]

where step \((a)\) follows since \(U_1 - X_1 - (U_2, Y_1)\) is a Markov chain, and step \((b)\) follows since \(X_1\) is independent of \(U_2\).

- Similarly,

\[
i - b = I(X_2; Y_2 U_1).
\]

Appendix I

Equivalence to Han-Kobayashi Region

We show that the region in (13) is equivalent to the Han-Kobayashi region.

- Consider the upper bound on \(R_1\):

\[
e - a = I(U_1; X_1; U_2, Y_1) \leq (a) I(X_1; U_2, Y_1) \leq (b) I(X_1; Y_1 U_2),
\]

where step \((a)\) follows since \(U_1 - X_1 - (U_2, Y_1)\) is a Markov chain, and step \((b)\) follows since \(X_1\) is independent of \(U_2\).

- Similarly,

\[
i - b = I(X_2; Y_2 U_1).
\]

- Consider the first upper bound on the sum-rate \(c + j - a - b:\)

\[
c + j - a - b = I(X_1; U_1, U_2, Y_1) + I(U_2; X_2) + I(U_1; Y_2) + I(U_2, X_2; U_1, Y_2) - I(U_2; X_2) - I(U_1; X_1)
\]

\[
\leq (a) I(X_1; U_2, Y_1 U_1) + I(U_1; Y_2) + I(U_2, X_2; U_1, Y_2)
\]

\[
\leq (b) I(X_1; U_2, Y_1 U_1) + I(U_1; Y_2) + I(X_2; U_1, Y_2)
\]

\[
\leq (c) I(X_1; U_2, Y_1 U_1) + I(U_1; Y_2) + I(X_2; Y_2 U_1)
\]

\[
\leq (d) I(X_1; Y_1 U_1 U_2) + I(X_2; U_1, Y_2),
\]

where step \((a)\) follows by the chain rule of mutual information, step \((b)\) follows by the Markov chain \(U_2 - X_2 - (U_1, Y_2)\), step \((c)\) follows since \(U_1\) and \(X_2\) are
independent and step (d) follows by the independence of $U_2$ and $(U_1, X_1)$.

- By similar steps, $f + g - a - b = I(X_1; U_2; Y_1) + I(X_2; U_2; Y_2)
- I(U_1; X_1) - I(U_2; X_2)$
- The remaining upper-bound on the sum-rate $d + h - a - b$
can be simplified as follows:

$$d + h - a - b = I(X_1; U_2; U_1, Y_1) + I(X_2; U_2; Y_2) + I(U_1, X_1) + I(U_2; X_1)
- I(U_1; X_1) - I(U_2; X_2)$$

which follows by the chain rule of mutual information and the independence of $(U_1, X_1)$ and $(U_2, X_2)$.

- The upper bound on $2R_1 + R_2$ can be simplified as follows:

$$e + h + f - 2a - b = I(X_1; U_2, Y_1) + I(X_2, U_2; Y_2) + I(U_1, X_1)
+ I(U_2; Y_1) + I(U_1, X_2; U_2, Y_1) - 2I(U_1; X_1)
- I(U_2; X_2)$$

Lemma 1: Consider the following four assumptions.

(A) Let $\{U, X_1, X_2, ..., X_K, Y\}$ be random variables that have some joint distribution $p(U, X_1, X_2, ..., X_K, Y)$.

(B) Let $(\tilde{U}^n, \tilde{Y}^n) \sim p(\tilde{u}^n, \tilde{y}^n)$ be a pair of arbitrarily distributed sequences.

(C) For each $j \in [1 : K]$, let $\{X_j^n(m_j), m_j \in A_j\}$, where $|A_j| \leq 2^nK$, be random sequences each distributed according to $\prod_{i=1}^{n} p_{X_j}(x_{ji} | \tilde{u}_i)$.

(D) For each $j \in [1 : K]$ and each $m_j$, assume that $X_j^n(m_j)$ is pairwise conditionally independent of $(\ldots, X_{j-1}^{n-1}(m_{j-1}), X_{j+1}^{n+1}(m_{j+1}), \ldots)$ given $\tilde{Y}_n$ for all $(\ldots, m_{j-1}, m_{j+1}, \ldots)$, but arbitrarily dependent on other $X_j^n(\cdot)$ sequences.

Then there exists $\delta(\epsilon)$ that tends to zero as $\epsilon \to 0$ such that

$$\sum_{j=1}^{K} R_j < \sum_{j=1}^{K} H(X_j | U) - H(X_1, X_2, ..., X_K | U, Y) - \delta(\epsilon).$$

(52)
Proof: The proof follows on similar lines as that of the packing lemma in [37, Ch. 3].

Consider a fixed tuple \((\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_K)\). The chain of inequalities resulting in (53), as shown at the bottom of the previous page, bounds the probability of \((\tilde{U}^n, \{X^n_\phi(\tilde{m}_j)\}_{j=1}^K, \tilde{Y}^n)\) being jointly typical, where \((a), (b)\) and \((c)\) follow from assumptions \((C)\) and \((D)\). Then we can apply the union bound over all possible tuples \((\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_K)\) to get the condition (52).

Example 1: For the case \(U = \phi\) and random variables \(X_1, X_2, Y\), the condition (52) can be expressed as follows:

\[
R_1 + R_2 < H(X_1) + H(X_2) - H(X_1, X_2) - \delta(\epsilon) - \delta(\epsilon)
\]

where

\[
R_1 + R_2 < H(X_1) + H(X_2) - H(X_1, X_2, Y) - \delta(\epsilon)
\]

Example 2: For the case \(U = \phi\) and random variables \(X_1, X_2, X_3, Y\), the condition (52) can be expressed as follows:

\[
R_1 + R_2 + R_3 < H(X_1) + H(X_2) + H(X_3)
\]

\[
- H(X_1, X_2, X_3) - \delta(\epsilon)
\]

\[
= I(X_1; X_2) + I(X_3; Y)
\]

\[
+ I(X_1, X_2; X_3, Y) - \delta(\epsilon).
\]

REFERENCES


Ritesh Kolte (S’13) is a doctoral candidate in Electrical Engineering at Stanford University. His research interests are in the general area of information theory, optimization, statistical learning and inference, signal processing. He received the Bachelor of Technology and Master of Technology degrees from the Indian Institute of Technology Bombay in 2010 and the Master of Science degree from Stanford University in 2012, all in Electrical Engineering; and is a recipient of the Stanford Graduate Fellowship.

Ayfer Özgür (M’06) received her B.Sc. degrees in electrical engineering and physics from Middle East Technical University, Turkey, in 2001 and the M.Sc. degree in communications from the same university in 2004. From 2001 to 2004, she worked as hardware engineer for the Defense Industries Development Institute in Turkey. She received her Ph.D. degree in 2009 from the Information Processing Group at EPFL, Switzerland. In 2010 and 2011, she was a post-doctoral scholar with the Algorithmic Research in Network Information Group at EPFL. She is currently an Assistant Professor in the Electrical Engineering Department at Stanford University. Her research interests include network communications, wireless systems, and information and coding theory. Dr. Özgür received the EPFL Best Ph.D. Thesis Award in 2010 and a NSF CAREER award in 2013.

Haim Permuter (M’08–SM’13) received his B.Sc. (summa cum laude) and M.Sc.(summa cum laude) degrees in Electrical and Computer Engineering from the Ben-Gurion University, Israel, in 1997 and 2003 respectively, and the Ph.D. degree in Electrical Engineering from Stanford University, California in 2008.

Between 1997 and 2004, he was an officer at a research and development unit of the Israeli Defense Forces. Since 2009 he is with the department of Electrical and Computer Engineering at Ben-Gurion University where he is currently an associate professor.

Prof. Permuter is a recipient of several awards, among them the Fulbright Fellowship, the Stanford Graduate Fellowship (SGF), Allon Fellowship, and the U.S.-Israel Binational Science Foundation Bergmann Memorial Award. He is currently serving on the editorial boards of the IEEE TRANSACTIONS ON INFORMATION THEORY.