Interference Channels With Correlated Receiver Side Information

Nan Liu, *Member, IEEE*, Deniz Gündüz, *Member, IEEE*, Andrea J. Goldsmith, *Fellow, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—The problem of joint source-channel coding in transmitting independent sources over interference channels with correlated receiver side information is studied. When each receiver has side information correlated with its own desired source, it is shown that source-channel separation is optimal. When each receiver has side information correlated with the interfering source, sufficient conditions for reliable transmission are provided based on a joint source-channel coding scheme using the superposition encoding and partial decoding idea of Han and Kobayashi. When the receiver side information is a deterministic function of the interfering source, source-channel separation is again shown to be optimal. In addition to these source-channel coding problems, a new channel model that generalizes the classical interference channel is introduced: the interference channel with message side information. Achievable rate regions are given and a single letter characterization of the capacity region for a special class of Z-interference channels is provided. Using this capacity result and the optimality of source-channel separation, we demonstrate that our sufficient conditions for reliable transmission when each receiver has side information correlated with the interfering source are also necessary for some special cases. As a by-product, the capacity region of a class of Z-channels with degraded message sets is also provided.**

***Index Terms*—Interference channel, joint source-channel coding, receiver side information, source-channel separation theorem.**

I. Introduction

THE wireless medium is shared by multiple communication systems operating simultaneously, which leads to interference among users transmitting over the same frequency band. In the simple scenario of two transmitter-receiver pairs, the interference channel [1] models two simultaneous transmissions interfering with each other. In the classical interference channel, the messages intended for each receiver are independent of each other, and the receivers decode based only on their own received signals. On the other hand, in applications such as sensor networks, it is reasonable to assume that the receivers have access to their own correlated observations about the underlying source sequences as well. These correlated observations at the receivers can be exploited to improve the system performance. Our goal is to understand the most effective ways to exploit the correlated side information at the receivers in order to manage interference in wireless networks.

Toward this end, we study an interference channel with independent sources, while each receiver has access to side information correlated with the sources. We constrain the analysis to independent sources, rather than considering arbitrary correlation between the source signals as in [2] and [3], as this allows us to isolate the effect of correlated side information on interference management and not to consider gains possible due to correlated channel inputs [4]. Our aim is to characterize the necessary and sufficient conditions such that reliable transmission of the sources can be achieved. Since the capacity region of the underlying interference channel is unknown in general, exact characterization of the necessary and sufficient conditions is not possible in the presence of receiver side information. Thus, in this work, we provide a set of sufficient conditions for the most general scenario, and matching necessary conditions for some special cases.

In certain scenarios our results are limited to proving the optimality of source-channel separation without identifying the necessary and sufficient conditions in a computable single-letter form. This stems from the difficulty in identifying the channel capacity for the general interference channel. However, a separation result provides us insights for the design of the communication system. Moreover, we are able to provide single-letter matching conditions for some special scenarios in which the capacity region of the underlying interference channel is known.

The study of communication with receiver side information has received considerable attention. In a point-to-point scenario, the availability of correlated side information at the receiver is considered in [5]. It is shown that the source-channel separation theorem applies in this simple setting and that Slepian-Wolf source coding followed by channel coding is optimal. However, it is known that the source-channel separation theorem does not generalize to multi-user scenarios [1], [4], [6], and necessary and sufficient conditions for reliable transmission in the case of correlated sources and correlated receiver side information are not known in the most general setting. Broad-

Manuscript received October 13, 2008; revised May 28, 2010. Date of current version November 19, 2010. This work was supported in part by the National Science Foundation under Grants CCF-07-28208 and CNS-09-05398, by the Defense Advanced Research Projects Agency ITMANET program under Grant 1105741-1-TFIND, and by the U.S. Army Research Office under MURI Award W911NF-05-1-0246.

N. Liu was with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA. She is now with the School of Information Science and Engineering, Southeast University, 210096 Nanjing, China (e-mail: nanliu@seu.edu.cn).

D. Gündüz was with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA and the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA. He is now with the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), 08860 Barcelona, Spain (e-mail: deniz.gunduz@cttc.es).

A. J. Goldsmith is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: andrea@wsl.stanford.edu).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Communicated by I. Kontoyiannis, Associate Editor for Shannon Theory.

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

Digital Object Identifier 10.1109/TIT.2010.2080930

cast channel with correlated receiver side information has been studied in [7]–[13], and interference channel with correlated receiver side information has been studied in [6], which considers the case where each receiver has access to side information correlated with the interfering transmitter's source. Necessary and sufficient conditions for this setup are characterized under the strong source-channel interference conditions, which generalize the usual strong interference conditions by considering correlated side information as well. The result of [6] shows that interference cancelation is optimal even when the underlying channel interference is not strong, as long as the overall source-channel interference is.

In this paper, we extend the scenario studied in [6] to more general interference channels. We first consider the case in which each receiver has side information correlated with the source sequence it wants to decode. We prove the optimality of source-channel separation in this situation; that is, the optimal performance can be achieved by first compressing each of the sources conditioned on the correlated receiver side information, and then transmitting the compressed bits over the channel using an optimal interference channel code.

Next, we consider the scenario in which each receiver has side information correlated with the interfering transmitter's source. As an example of such a model and to illustrate the benefits of side information correlated with the interfering source, consider the extreme case in which each receiver has access to the message of the interfering transmitter. Note that this setup is equivalent to the restricted two-way channel model of Shannon, whose capacity is characterized in [1]. In this case, each receiver can excise the interference from the undesired transmitter since the message causing the interference is exactly known. Here, we consider the more general case of arbitrary correlation between the receiver side information and the interfering source. First, we provide sufficient conditions for reliable transmission by proposing a joint source-channel coding scheme based on the message splitting technique introduced by Han and Kobayashi [14], in which part of the message is decoded by the unintended receiver as well.

Then we digress from the joint source-channel coding problem and turn our attention to a closely related channel coding problem that generalizes the classical interference channel model: the interference channel with message side information (IC-SI), in which a portion of each user's message is available at the nonintended receiver. We provide a general achievability scheme for this channel model, and characterize the capacity region of the IC-SI for a special class of Z-interference channels.

Finally we go back to the source-channel coding problem and prove the optimality of source-channel separation when the side information is a deterministic function of the interfering source. Here, the separation theorem is presented using the n-letter expression for the capacity region of the underlying IC-SI. Then, for the special class of Z-interference channels mentioned above and when the receiver side information is a deterministic function of the interfering source, we provide single-letter necessary and sufficient conditions for reliable transmission based on source-channel separation and the capacity region of the underlying IC-SI. This setting constitutes an example for which the general sufficient conditions we provide are also necessary, proving their tightness for certain special cases.

The rest of the paper is organized as follows. In Section II, we present the system model. In Section III, we prove the optimality of source-channel separation when the side information is correlated with the desired source. The case in which the side information is correlated with the interfering source is considered in Section IV. In Section IV-A, we provide general sufficient conditions for reliable transmission. In Section IV-B, we consider a related channel coding problem by introducing the IC-SI and characterize the capacity region of this channel model for a special class of Z-interference channels. Then, in Section IV-C, we go back to the source-channel coding problem and focus on the special case in which the receiver side information is a deterministic function of the interfering source. We prove the optimality of source-channel separation in this scenario. We further show that for the special class of Z-interference channels studied in Section IV-B, the sufficient conditions for reliable transmission proposed in Section IV-A are also necessary. Discussion is presented in Section V, where in Section V-A, we comment on the most general case of correlated side information at the receiver, and in Section V-B, using the capacity results of Section IV-B, we characterize the capacity region of a related channel coding problem, i.e., a class of Z-channels with degraded message sets. This is followed by conclusions in Section VI.

For notational convenience, we drop the subscripts on probability distributions unless the arguments of the distributions are not lower case versions of the corresponding random variables. We use calligraphic, upper case and lower case variables to represent sets, random variables and deterministic numbers, respectively.

II. System Model

An interference channel is composed of two transmitter-receiver pairs. The underlying discrete memoryless channel is characterized by the transition probability $p(y_1, y_2|x_1, x_2)$ from the finite input alphabet $\mathcal{X}_1 \times \mathcal{X}_2$ to the finite output alphabet $\mathcal{Y}_1 \times \mathcal{Y}_2$. Transmitter k has access to the source sequence $\{U_{k,i}\}_{i=1}^{\infty}$, $k = 1, 2$. Consider side information sequences $\{V_{k,i}\}_{i=1}^{\infty}$, such that the source and the side information sequences are independent and identically distributed (i.i.d.) and are drawn according to joint distribution $p(u_1, v_1)p(u_2, v_2)$ over a finite alphabet $\mathcal{U}_1 \times \mathcal{V}_1 \times \mathcal{U}_2 \times \mathcal{V}_2$; that is, the two source-side information pairs are independent of each other while the components of each pair can have arbitrary correlation.

For $k = 1, 2$, Transmitter k observes U_k^n and wishes to transmit it noiselessly to Receiver k over n uses of the channel[1]. The encoding function at Transmitter k is

$$f_k^n : \mathcal{U}_k^n \rightarrow \mathcal{X}_k^n.$$

We assume that the side information vector $V_{\pi(k)}^n$ is available at receiver k, where $\pi(\cdot)$ is a permutation of $\{1, 2\}$. Depending on the scenario, $\pi(\cdot)$ specifies whether the side information is correlated with the desired source or with the interfering source.

[1]Here we use the notation $U_k^n = (U_{k,1}, \ldots, U_{k,n})$, and similar notation applies for other length-n sequences.

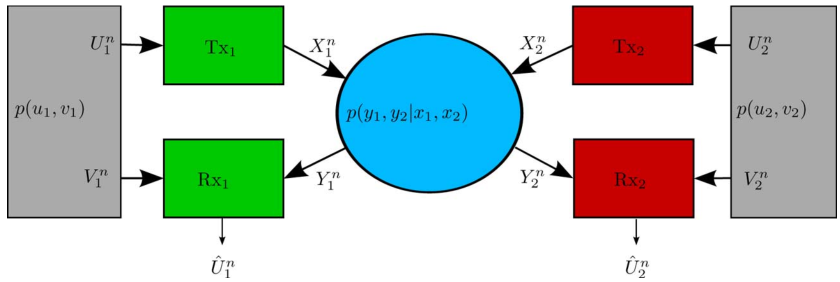

Fig. 1. Interference channel model in which the receivers have access to side information correlated with the sources they want to receive.

The decoding function at receiver k reconstructs its estimate $\hat{U}_k$ from its channel output and the side information vector using the decoding function

$$g_k^n : \mathcal{Y}_k^n \times \mathcal{V}_{\pi(k)}^n \rightarrow \mathcal{U}_k^n.$$

The probability of error for this system is defined as

$$P_e^n = \Pr\{(U_1^n, U_2^n) \neq (\hat{U}_1^n, \hat{U}_2^n)\} = \sum_{(u_1^n, u_2^n) \in \mathcal{U}_1^n \times \mathcal{U}_2^n} p(u_1^n, u_2^n)$$
$$\Pr\left\{(\hat{U}_1^n, \hat{U}_2^n) \neq (u_1^n, u_2^n) \middle| (U_1^n, U_2^n) = (u_1^n, u_2^n)\right\}.$$

Definition 1: We say that a source pair (U_1, U_2) can be *reliably* transmitted over a given interference channel in the presence of side information $(V_{\pi(1)}, V_{\pi(2)})$, if there exists a sequence of encoders and decoders $(f_1^n, f_2^n, g_1^n, g_2^n)$ such that $P_e^n \rightarrow 0$ as $n \rightarrow \infty$.

In the following sections, we focus on two scenarios in particular. In the first one, each receiver has side information correlated with its desired source, i.e., $\pi(k) = k$, $k = 1, 2$. In the second scenario, each receiver has side information correlated with the interfering source, i.e., $\pi(1) = 2$ and $\pi(2) = 1$. In both cases, we want to exploit the availability of correlated side information at the receivers. In the first case, each transmitter needs to transmit less information to its intended receiver due to the availability of correlated side information. In the latter case, while the side information is not helpful in reducing the amount of information to be transmitted, it can be used to mitigate the effects of interference.

III. Side Information Correlated With the Desired Source

In this section, we consider an interference channel in which each receiver has side information correlated with the source it wants to decode, i.e., receiver k has access to side information V_k (see Fig. 1). For this special case, we prove that the source-channel separation theorem applies; that is, it is optimal for the transmitters first to compress their sources conditioned on the side information at the corresponding receiver, and then to transmit the compressed bits over the channel using an optimal interference channel code. Note that, in the general case, we do not have a single-letter characterization of the capacity region of the interference channel, yet we can still prove the optimality of source-channel separation. In the proof, we use the n-letter expression for the capacity region, which was also used in [15] to prove the optimality of source-channel separation for a multiple access channel with receiver side information and feedback. The main result of this section is the following theorem.

Theorem 1: Sources U_1 and U_2 can be transmitted reliably to their respective receivers over the discrete memoryless interference channel $p(y_1, y_2|x_1, x_2)$ with side information V_k at receiver k, $k = 1, 2$, if

$$(H(U_1|V_1), H(U_2|V_2)) \in int(\mathcal{C}_{IC}) \quad (1)$$

where $int(\cdot)$ denotes the *interior*, and $\mathcal{C}_{IC}$ denotes the capacity region of the underlying interference channel.

Conversely, if $(H(U_1|V_1), H(U_2|V_2)) \notin \mathcal{C}_{IC}$, then sources U_1 and U_2 cannot be transmitted reliably.

Proof: A proof of Theorem 1 is given in Appendix A. ∎

IV. Side Information Correlated With the Interfering Source

Now we consider the case in which Receiver 1 has access to V_2 while Receiver 2 has access to V_1, i.e., each receiver has side information correlated with the interfering transmitter's source (see Fig. 2). We investigate how the side information about the interference helps in decoding the desired information.

A. Sufficient Conditions for Reliable Transmission

We first provide sufficient conditions for reliable transmission of the sources. In the spirit of the well-known Han-Kobayashi scheme for the classical interference channel [14], we provide sufficient conditions for reliable transmission of the sources by proposing a joint source-channel coding scheme that requires the receivers to decode part of the interference with the help of their side information. In the Han-Kobayashi scheme, each transmitter splits its message into two parts to allow the nonintended receiver to decode part of the interference. In our scheme, each transmitter enables a quantized version of its source to be decoded by both receivers, where the unintended receiver uses its correlated side information as well as the channel output to decode the interference corresponding to this quantized part.

Theorem 2: Sources U_1 and U_2 can be transmitted reliably over the interference channel $p(y_1, y_2|x_1, x_2)$ with side information V_1 at Receiver 2 and V_2 at Receiver 1 if there exist random variables W_1 and W_2 such that

$$H(U_1) < I(X_1; V_2, Y_1|W_2, Q)$$
$$H(U_2) < I(X_2; V_1, Y_2|W_1, Q)$$
$$H(U_1) < I(W_2, X_1; V_2, Y_1|Q) - I(U_2; W_2|Q)$$
$$H(U_2) < I(W_1, X_2; V_1, Y_2|Q) - I(U_1; W_1|Q)$$

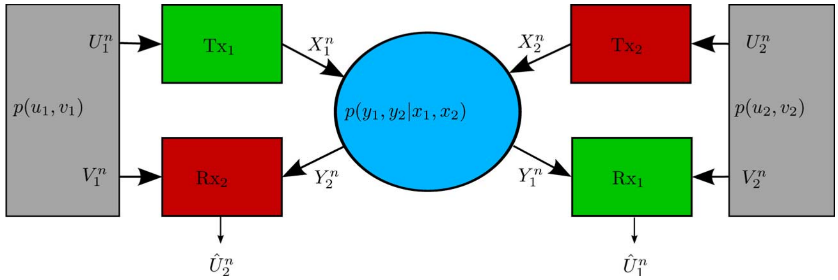

Fig. 2. Interference channel model in which the receivers have access to side information correlated with the sources of the interfering transmitters.

$$
\begin{aligned}
H(U_1)+H(U_2) &< I(X_1;V_2,Y_1|W_1,W_2,Q)\\
&\quad + I(W_1,X_2;V_1,Y_2|Q)\\
H(U_1)+H(U_2) &< I(X_2;V_1,Y_2|W_1,W_2,Q)\\
&\quad + I(W_2,X_1;V_2,Y_1|Q)\\
H(U_1)+H(U_2) &< I(W_1,X_2;V_1,Y_2|W_2,Q)\\
&\quad + I(W_2,X_1;V_2,Y_1|W_1,Q)\\
H(U_1)+H(U_2) &< I(W_2,X_1;V_2,Y_1|Q)-I(U_1;W_1|Q)\\
&\quad + I(W_1,X_2;V_1,Y_2|W_2,Q)\\
H(U_1)+H(U_2) &< I(W_1,X_2;V_1,Y_2|Q)-I(U_2;W_2|Q)\\
&\quad + I(W_2,X_1;V_2,Y_1|W_1,Q)\\
2H(U_1)+H(U_2) &< I(W_2,X_1;V_2,Y_1|Q)\\
&\quad + I(X_1;V_2,Y_1|W_1,W_2,Q)\\
&\quad + I(W_1,X_2;V_1,Y_2|W_2,Q)\\
H(U_1)+2H(U_2) &< I(W_1,X_2;V_1,Y_2|Q)\\
&\quad + I(X_2;V_1,Y_2|W_1,W_2,Q)\\
&\quad + I(W_2,X_1;V_2,Y_1|W_1,Q)
\end{aligned}
$$

for some $p(q)$, $p(w_1,x_1|u_1,q)$, and $p(w_2,x_2|u_2,q)$, where the entropies and mutual information terms are evaluated using the joint distribution

$$
\begin{aligned}
&p(q,u_1,v_1,u_2,v_2,w_1,w_2,x_1,x_2,y_1,y_2)\\
&\quad = p(q)p(u_1,v_1)p(u_2,v_2)p(w_1,x_1|u_1,q)p(w_2,x_2|u_2,q)\\
&\quad\quad \cdot p(y_1,y_2|x_1,x_2). \qquad (2)
\end{aligned}
$$

Proof: A proof of Theorem 2 is given in Appendix B. ■

Remark 1: In the special case of no receiver side information, i.e., $V_1=V_2=\emptyset$, by choosing the distribution in (2) such that the Markov Chain conditions $U_1 \rightarrow Q \rightarrow (W_1,X_1)$ and $U_2 \rightarrow Q \rightarrow (W_2,X_2)$ are satisfied, and defining $R_1 = H(U_1)$ and $R_2 = H(U_2)$, the sufficient conditions in Theorem 2 reduce to the Han-Kobayashi rate region in the form expressed in [16, Th. 2].

We do not know whether the sufficient conditions provided in Theorem 2 are too strong, leading to pessimistic results in general. However, we show in Section IV-C that the conditions in Theorem 2 are also necessary for some special cases, showing their tightness at least in certain scenarios.

B. Interference Channel With Message Side Information

In this subsection, we digress from the source-channel coding problem of Fig. 2 and consider a related channel coding problem. Apart from being a problem of interest in its own right, we will later use the capacity pertaining to this channel coding problem in identifying a source-channel separation theorem under the assumption of deterministic side information in Section IV-C.

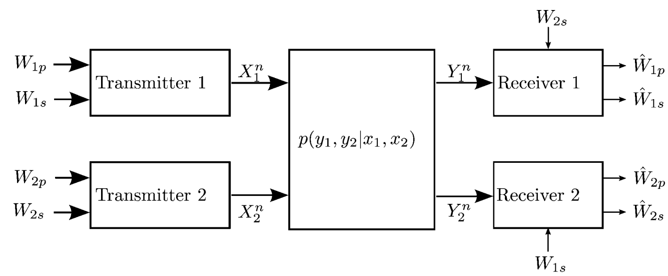

Fig. 3. Interference channel with message side information at the receivers.

We consider the communication model depicted in Fig. 3. In this model, which again is called the *interference channel with message side information* (IC-SI), Transmitter k, $k=1,2$, has two messages W_{ks} and W_{kp} at rates R_{ks} and R_{kp}, respectively. All messages are independent. Receiver k is interested in obtaining both messages W_{ks} and W_{kp}. Unlike the classical interference channel, Receiver 2 has access to W_{1s} and Receiver 1 has access to W_{2s} as side information. One can see that this channel model is a generalization of the classical interference channel in which W_{1s} and W_{2s} are constants.

A $\left(2^{nR_{1s}},2^{nR_{1p}},2^{nR_{2s}},2^{nR_{2p}},n\right)$ code for an IC-SI consists of two encoding functions

$$
f_k^n:\{1,2,\ldots,2^{nR_{ks}}\}\times\{1,2,\ldots,2^{nR_{kp}}\}\rightarrow\mathcal{X}_k^n, \quad \text{for } k=1,2
$$

and two decoding functions

$$
\begin{aligned}
g_k^n:\mathcal{Y}_k^n\times\{1,2,\ldots,2^{nR_{ls}}\}&\rightarrow\{1,2,\ldots,2^{nR_{ks}}\}\\
&\quad\times\{1,2,\ldots,2^{nR_{kp}}\}, \quad \text{for } (k,l)=(1,2),(2,1).
\end{aligned}
$$

The average probability of error for the $\left(2^{nR_{1s}},2^{nR_{1p}},2^{nR_{2s}},2^{nR_{2p}},n\right)$ code is defined as

$$
\begin{aligned}
P_e^n = \frac{1}{2^{n(R_{1s}+R_{1p}+R_{2s}+R_{2p})}} &\sum_{w_{1s}=1}^{2^{nR_{1s}}}\sum_{w_{1p}=1}^{2^{nR_{1p}}}\sum_{w_{2s}=1}^{2^{nR_{2s}}}\sum_{w_{2p}=1}^{2^{nR_{2p}}}\\
&\Pr\{g_1^n(Y_1^n,w_{2s})\neq(w_{1s},w_{1p}) \text{ or } g_2^n(Y_2^n,w_{1s})\\
&\quad\neq(w_{2s},w_{2p})|(w_{1s},w_{1p},w_{2s},w_{2p}) \text{ is sent}\}.
\end{aligned}
$$

Definition 2: A rate quadruplet $(R_{1s}, R_{1p}, R_{2s}, R_{2p})$ is said to be *achievable* for the IC-SI if there exists a sequence of $\left(2^{nR_{1s}}, 2^{nR_{1p}}, 2^{nR_{2s}}, 2^{nR_{2p}}, n\right)$ codes for which $P_e^n \to 0$ as $n \to \infty$.

Definition 3: The *capacity region* of the IC-SI is defined as the closure of the set of all achievable rate quadruplets $(R_{1s}, R_{1p}, R_{2s}, R_{2p})$, and is denoted by $\mathcal{C}_{IC-SI}$.

For the IC-SI, we propose the following achievable scheme that is based on the Han-Kobayashi scheme for the interference channel: the codebook at Transmitter k is generated such that the inner codebook carries the side information W_{ks} at Receiver l, as well as a portion of its private message W_{kp}, while the outer codebook carries the remaining part of W_{kp}, $(k,l) = (1,2),(2,1)$. The intuition behind this coding scheme is the following: since W_{ks} is already available at the nonintended receiver, the codeword corresponding to this message can be directly decoded and hence, it does not cause any additional interference. In particular, Transmitter 1 splits its private message W_{1p} into (W_{1p1}, W_{1p2}), where $W_{1p1} \in \{1, \ldots, \lfloor 2^{nR_{1p1}} \rfloor\}$, and $W_{1p2} \in \{1, \ldots, \lfloor 2^{nR_{1p2}} \rfloor\}$. Here W_{1p1} is to be decoded by Receiver 2 as well, even though it is not intended for this receiver. Overall, (W_{1s}, W_{1p1}) constitute the message that is decoded by both receivers, while W_{1p2} is decoded by only Receiver 1. A similar message splitting is also applied at Transmitter 2.

By standard error analysis of the above scheme, and noting that W_{ks} is already available at Receiver l, $(k,l) = (1,2),(2,1)$, we obtain the achievable rate region corresponding to the above achievable scheme for the IC-SI in the following lemma.

Lemma 1: Let $\mathcal{P}$ be the set of probability distributions $p(\cdot)$ in the form

$$p(q, w_1, w_2, x_1, x_2, y_1, y_2) = p(q)p(w_1, x_1|q)p(w_2, x_2|q) \, p(y_1, y_2|x_1, x_2).$$

For a fixed $p \in \mathcal{P}$, define $\mathcal{R}(p)$ as the set of rate tuples $(R_{1s}, R_{1p}, R_{2s}, R_{2p})$ that satisfy

$$\begin{aligned}
R_{1p} + R_{1s} &\leq I(X_1; Y_1|W_2, Q) \\
R_{2p} + R_{2s} &\leq I(X_2; Y_2|W_1, Q) \\
R_{1p} + R_{2p} + R_{1s} &\leq I(X_2; Y_2|W_1, W_2, Q) \\
&\quad + I(W_2, X_1; Y_1|Q) \\
R_{1p} + R_{2p} + R_{2s} &\leq I(X_1; Y_1|W_1, W_2, Q) \\
&\quad + I(W_1, X_2; Y_2|Q) \\
R_{1p} + R_{2p} &\leq I(W_2, X_1; Y_1|W_1, Q) \\
&\quad + I(W_1, X_2; Y_2|W_2, Q) \\
2R_{1p} + R_{2p} + R_{1s} &\leq I(X_1; Y_1|W_1, W_2, Q) \\
&\quad + I(W_2, X_1; Y_1|Q) \\
&\quad + I(W_1, X_2; Y_2|W_2, Q) \\
R_{1p} + 2R_{2p} + R_{2s} &\leq I(X_2; Y_2|W_1, W_2, Q) \\
&\quad + I(W_1, X_2; Y_2|Q) \\
&\quad + I(W_2, X_1; Y_1|W_1, Q) \\
R_{1s}, R_{1p}, R_{2s}, R_{2p} &\geq 0.
\end{aligned}$$

Then we have

$$\bigcup_{p \in \mathcal{P}} \mathcal{R}(p) \subseteq \mathcal{C}_{IC-SI}.$$

Proof: We skip the details of the proof as it closely resembles the proof given in [16] for the classical interference channel. ■

Next, we show that the coding scheme of Lemma 1 is capacity-achieving for a special class of interference channels. In particular, we consider a special class of Z-interference channels. The Han-Kobayashi scheme is shown to be capacity-achieving for this class of interference channels in [17] when there is no message side information. We show here that this optimality extends to the message side information scenario with the coding scheme of Lemma 1.

For Z-interference channels, $p(y_1, y_2|x_1, x_2)$ can be written as $p(y_2|x_1, x_2) \cdot p(y_1|x_1)$, i.e., the channel between X_1 and Y_1 is a single user channel characterized by $p(y_1|x_1)$. This corresponds to an interference channel in which only the second transmitter-receiver pair faces interference. In particular, the members of the class of Z-interference channels we consider satisfy the following conditions.

1) For any positive integer n, $H(Y_2^n|X_2^n = x_2^n)$, when evaluated with the distribution $\sum_{x_1^n} p(x_1^n)p(y_2^n|x_1^n, x_2^n)$, is independent of x_2^n for any $p(x_1^n)$.
2) Define τ as

$$\tau = \max_{p(x_1)p(x_2)} H(Y_2). \tag{3}$$

Then there exists a $p^*(x_2)$ such that $H(Y_2)$, when evaluated with the distribution $\sum_{x_1,x_2} p(x_1) \; p^*(x_2)p(y_2|x_1, x_2)$, is equal to τ for any $p(x_1)$.

Condition 1 specifies that the channel $p(y_2|x_1, x_2)$ is invariant, in terms of the conditional output entropy, with respect to the input sequence of Transmitter 2, i.e., x_2^n. Intuitively, Condition 2 specifies that no matter how tightly packed the codewords in the codebook of Transmitter 1 are, by spacing out the codewords in the codebook of Transmitter 2, we can always fill up the entire, or maximal, output space at Receiver 2. Please refer to [17] for the detailed intuition behind these conditions and examples of Z-interference channels that satisfy these two conditions.

In the next lemma, we provide a single-letter characterization for the capacity region of this class of Z-interference channels with message side information. Since Receiver 1 does not face interference, it does not benefit from the message side information W_{2s}. Hence, without loss of generality, we assume $R_{2s} = 0$.

Lemma 2: The capacity region of Z-interference channels satisfying Conditions 1 and 2, with message side information W_{1s} at Receiver 2, is characterized by

$$R_{1p} + R_{1s} \leq I(X_1; Y_1), \tag{4}$$

$$R_{2p} \leq I(W, X_2; Y_2) \tag{5}$$

$$R_{1p} + R_{2p} \leq I(X_1; Y_1|W) + I(W, X_2; Y_2) \tag{6}$$

for some $p(w)p(x_1|w)$, where the mutual informations and entropies are evaluated with the joint distribution of the form

$$p(w,x_1,x_2,y_1,y_2) = p(w)p(x_1|w)p^*(x_2)p(y_1|x_1)p(y_2|x_1,x_2).$$

Proof: A proof of Lemma 2 is given in Appendix C. For the proof of achievability, the general coding scheme of Lemma 1 is used. A specialized converse is developed and further shown to coincide with the achievable region using the fact that the channels satisfy Conditions 1 and 2. ■

Comparing these results in the case of side information at the receiver with the traditional Z-interference channel [17], the rate of W_{1p} takes the place of W_1, which means that the message that causes interference is reduced from W_1 to W_{1p}. Due to the fact that W_{1s} is available at Receiver 2, W_{1s} does not cause any interference and therefore its rate can be made as large as possible within the constraint of the capacity of the channel $p(y_1|x_1)$ depicted by (4).

C. Source Transmission Over Interference Channel With Deterministic Side Information

Now, we go back to the source-channel coding problem of Fig. 2 and focus on a special case in which the side information sequences V_1 and V_2 are deterministic functions of the sources U_1 and U_2, respectively, i.e.,

$$V_{k,i} = h_k(U_{k,i}), \qquad k=1,2, \quad i=1,2,\ldots$$

for some deterministic functions h_1 and h_2, or equivalently we have $H(V_k|U_k)=0$ for $k=1,2$.

The first result of this subsection states that when the side information at each receiver is a deterministic function of the interfering source, the source-channel separation theorem applies; that is, it is optimal to first perform source coding and encode V_k^n into message W_{ks}, and the remaining part of U_k^n, denoted by $U_k^n|V_k^n$, into message W_{kp}, $k=1,2$, and then to transmit these messages optimally over the underlying IC-SI. We will use the capacity region defined in Section IV-B for the IC-SI to express this separation theorem.

Theorem 3: Sources U_1 and U_2 can be transmitted reliably to their respective receivers over the discrete memoryless interference channel characterized by $p(y_1,y_2|x_1,x_2)$ with side information $V_1=h_1(U_1)$ at Receiver 2, and side information $V_2=h_2(U_2)$ at Receiver 1, if

$$(H(V_1),H(U_1|V_1),H(V_2),H(U_2|V_2)) \in int(\mathcal{C}_{IC-SI}). \quad (7)$$

Conversely, if $(H(V_1),H(U_1|V_1),H(V_2),H(U_2|V_2)) \notin \mathcal{C}_{IC-SI}$, then sources U_1 and U_2 cannot be transmitted reliably.

Proof: A proof of Theorem 3 is given in Appendix D. Similarly to Theorem 1, Theorem 3 is proved by using the n-letter characterization of $\mathcal{C}_{IC-SI}$. ■

The benefits of considering the side information samples as deterministic functions of the source samples are two-fold. Firstly, transmitters know the side information and they can use this knowledge to minimize the amount of interference they cause. Due to this fact, we are able to achieve any point in the capacity region of the IC-SI. Secondly, encoding V_k, $k=1,2$ into the codebook at Transmitter k not only helps reduce the interference at the other receiver, but also does not place any extra burden on Receiver k to decode V_k, as V_k is a deterministic function of U_k. This fact enables the converse proof of the given source-channel separation theorem.

Note that both of the separation results in Theorem 1 and Theorem 3 are based on n-letter capacity expressions, hence they cannot be computed in general. However, if a single-letter characterization of $\mathcal{C}_{IC}$ or $\mathcal{C}_{IC-SI}$ is known for the underlying interference channel, we would obtain the necessary and sufficient conditions in Theorem 1 and Theorem 3, respectively, in a single-letter form. Having established $\mathcal{C}_{IC-SI}$ for a special class of interference channels in Section IV-B, our second result of this subsection is to provide single letter necessary and sufficient conditions for reliable source transmission when the underlying interference channel falls into this special class.

Corollary 1: For Z-interference channels satisfying Conditions 1 and 2 and side information $V_1=h_1(U_1)$ at Receiver 2, U_1 and U_2 can be transmitted reliably to their respective receivers if

$$H(U_1) < I(X_1;Y_1) \quad (8)$$
$$H(U_2) < I(W,X_2;Y_2) \quad (9)$$
$$H(U_1|V_1)+H(U_2) < I(W,X_2;Y_2)+I(X_1;Y_1|W) \quad (10)$$

for some $p(w)p(x_1|w)$, where the mutual informations and entropies are evaluated with a joint distribution of the form

$$p(u_1,v_1,u_2,w,x_1,x_2,y_1,y_2) = p(u_1,v_1) \times p(u_2)p(w)p(x_1|w)p^*(x_2)p(y_1|x_1)p(y_2|x_1,x_2). \quad (11)$$

Conversely, if the sources can be transmitted reliably, then the inequalities (8)–(10) hold with $<$ replaced by $\le$ for some joint distribution of the form given in (11).

Proof: Corollary 1 follows directly from combining Theorem 3 and Lemma 2. ■

Corollary 1 shows how the side information $V_1=h_1(U_1)$ about the interference U_1 helps in reliable transmission, and determines the most efficient way of using this side information: Transmitter 1 performs a separation-based encoding scheme. It first splits its source U_1^n into V_1^n and a remaining part using entropy-achieving data compression techniques, and thus obtains two messages $W_{V_1^n}$ and $W_{U_1^n|V_1^n}$. Then, it further splits message $W_{U_1^n|V_1^n}$ into two parts W_{inner} and W_{outer}, at rates γ and $H(U_1|V_1)-\gamma$, respectively. Next, it performs superposition encoding, transmitting $W_{V_1^n}$ and W_{inner} through the inner code at rate $H(V_1)+\gamma$, and W_{outer} through the outer code at rate $H(U_1|V_1)-\gamma$. Transmitter 2 performs separation-based source-channel coding, first mapping U_2^n into a message W_2 and then mapping W_2 into a codeword of an i.i.d. codebook generated with distribution $p^*(x_2)$. Receiver 1 decodes both the inner and the outer codes. Receiver 2 knows the side information V_1^n and hence sees an inner codebook at an effective rate of γ only. It decodes the inner codeword and the codeword of Transmitter

2 jointly using the received signal and the available side information about the interference while treating the outer codeword of Transmitter 1 as noise.

Next, we show that the sufficient conditions in Corollary 1 can also be obtained from Theorem 2. In Theorem 2, let $\mathcal{V}_2 = \mathcal{W}_2 = \mathcal{Q} = \emptyset$ and $p(x_2) = p^*(x_2)$. Also choose $W_1 = (V_1, \tilde{W}_1)$ such that the distribution in (2) is of the form

$$p(u_1, v_1, u_2, \tilde{w}_1, x_1, x_2, y_1, y_2) = p(u_1, v_1)p(u_2)p(\tilde{w}_1, x_1) p^*(x_2)p(y_1, y_2|x_1, x_2).$$

Using Condition 2 and the fact that $H(U_1) - H(V_1) = H(U_1|V_1)$, we obtain the sufficient conditions in Corollary 1. Hence, the third result of this subsection is that we have found a special case of Fig. 2 where the sufficient conditions for reliable transmission described in Theorem 2 are also necessary, i.e., the sufficient conditions in Theorem 2 are tight in certain scenarios.

The intuition obtained from the special case studied in this subsection is that one should put as much information as possible about the side information within the inner codebook in order to minimize the impact of interference when the side information about the interference is available at the receiver.

V. Discussion

A. The General Correlated Side Information Case

In Sections III and IV, we have studied the two cases in which the side information is correlated with the desired source and the side information is correlated with the interfering source, respectively. In the most general case, the side information would be correlated with both the desired and interfering source. We do not consider this most general setup to keep the presentation simple; however, we expect the intuition developed in Sections III and IV to remain true, that is the transmitters should transmit the minimum amount of information that is required by the intended receiver to decode the source signal, and should include as much as possible the part of the source that the unintended receiver knows in the codeword that carries the common information decoded by both receivers.

B. Z-Channel With Degraded Message Sets

From the proof of [9, Th. 3], there is some interesting intuition that when the receiver has some side information about the undesired message, we can set up a new scenario in which the receiver does not have access to the side information, and is required to decode it. Then, when we remove the rate constraint associated with decoding of the side information at the receiver in the capacity region of the new scenario, we get the capacity results of the original scenario. Based on this intuition, the solution given in (4)–(6) resembles the solution of the following problem.

The channel is described by two transition probabilities $p(y_1|x_1)$ and $p(y_2|x_1, x_2)$, and satisfies both Conditions 1 and 2. There are three independent messages W_{1c}, W_{1p} and W_2. Transmitter 1 has messages W_{1c} and W_{1p} and Transmitter 2 has message W_2. W_{1c} needs to be decoded at both receivers, while W_{1p} and W_2 need to be decoded only at Receiver 1 and Receiver 2, respectively. Compared to the definition of the Z-channel in [18], W_{1c} is not only intended for Receiver 2, but also for Receiver 1. From another viewpoint, we have one receiver that decodes both W_{1c} and W_{1p}, and the other that decodes W_{1c}, in addition to W_2. In other words, this channel model can be seen as a broadcast channel (from Transmitter 1 to Receivers 1 and 2) with degraded message sets plus an extra link from Transmitter 2 to Receiver 2. Therefore, we call this channel model the *Z-channel with degraded message sets*. The Z-channel with degraded message sets includes the Z-interference channel as a special case, when the rate of W_{1c} is zero.

The capacity region for the Z-channel with degraded message sets when the underlying Z-interference channel satisfies Conditions 1 and 2 can be characterized as follows:

$$\begin{aligned} R_{1p} &\leq I(X_1; Y_1|W) + \gamma \\ R_{1c} + R_{1p} &\leq I(X_1; Y_1) \\ R_{1c} &\leq I(W; Y_2|X_2) - \gamma \\ R_2 + R_{1c} &\leq \tau - H(Y_2|W, X_2) - \gamma \end{aligned}$$

for some $p(w)p(x_1|w)$ and $\gamma \geq 0$ where the mutual informations and entropies are evaluated using $p(w, x_1, x_2, y_1, y_2) = p(w)p(x_1|w)p^*(x_2)p(y_1|x_1)p(y_2|x_1, x_2)$. The proof of this result follows from arguments very similar to those used in the scenario of IC-SI considered in Lemma 2.

VI. Conclusions

We have studied the problem of joint source-channel coding in interference channels with correlated receiver side information. In the case when the receiver side information is correlated with its desired source, we have shown that separate design of source and channel codes is optimal. Note that the optimal channel coding scheme is not known for interference channels in general, and hence, we have used the n-letter expression for the capacity region to prove the optimality of separation. In separate source-channel coding, each user transmits only the part of the sources that is not already known by their corresponding receivers. Since the interfered receiver does not have any side information about the source, there is no advantage of using a joint source-channel coding scheme to reduce the interference.

For the case in which the receiver side information is correlated with the interfering source, we have provided sufficient conditions for reliable transmission by proposing a joint source-channel coding scheme based on the idea of superposition encoding and partial decoding of Han and Kobayashi. As a special case, we have focused on the scenario in which the side information at the receiver is a deterministic function of the interfering source. By introducing the channel coding problem of the IC-SI, we have shown that source-channel separation is optimal for this situation as well. Finally, for a class of Z-interference channels for which superposition encoding and partial decoding is optimal in the absence of receiver side information, when the receiver facing interference has access to a deterministic function of the interfering source, using the capacity region of the underlying IC-SI and the optimality of source-channel separation, we have shown that the provided sufficient conditions are also necessary. Hence, our proposed sufficient conditions are tight at least in some special cases.

APPENDIX A
PROOF OF THEOREM 1

The achievability part of the proof is straightforward. If (1) holds, then there exists a rate pair $(R_1,\ R_2)$ in the interior of $\mathcal{C}_{IC}$ such that $H(U_k|V_k) \le R_k$ for $k=1,2$. Each transmitter compresses its source with respect to the side information at its own receiver. This can be done at rate R_k due to the Slepian-Wolf theorem. Then the compressed bits can be transmitted reliably over the channel since $(R_1,\ R_2)$ is in the capacity region of the underlying interference channel. However, we note here that we need to be careful in the error probability analysis as the capacity characterization for the interference channel is for average error probability criteria, not for maximal error probability. See [6] for a similar analysis in a multiple access channel scenario.

To prove the converse, we first make use of the infinite letter expression for the capacity region of the interference channel given in [19]. We define

$$E_n \triangleq \left\{ \left(\frac{1}{n}I(X_1^n;Y_1^n), \frac{1}{n}I(X_2^n;Y_2^n)\right): \right.$$
$$\left. p(x_1^n,x_2^n)=p(x_1^n)p(x_2^n) \right\}.$$

Then

$$\mathcal{C}_{IC} = \lim_{n\to\infty} E_n \tag{12}$$

where the limit is defined as in [1, Th. 5]. $\mathcal{C}_{IC}$ is a closed convex set in the Euclidean plane.

From Fano's inequality [20], we have, for $k=1,2$

$$H(U_k^n|\hat{U}_k^n) \le n\delta(P_e^n) \tag{13}$$

where $\delta(x)$ is a nonnegative function approaching zero as $x\to 0$.

Next, we write the following chain of inequalities:

$$\begin{aligned}
\frac{1}{n}I(X_1^n;Y_1^n) &\ge \frac{1}{n}I(U_1^n;Y_1^n) && (14)\\
&= \frac{1}{n}I(U_1^n,V_1^n;Y_1^n) && (15)\\
&\ge \frac{1}{n}I(U_1^n;Y_1^n|V_1^n) &&\\
&= \frac{1}{n}\left[H(U_1^n|V_1^n) - H(U_1^n|V_1^n,Y_1^n)\right] &&\\
&\ge H(U_1|V_1) - H(U_1^n|\hat{U}_1^n) && (16)\\
&\ge H(U_1|V_1) - \delta(P_e^n) && (17)
\end{aligned}$$

where (14) follows since $U_1^n \to X_1^n \to Y_1^n$ form a Markov chain; similarly (15) follows since $V_1^n\to U_1^n\to Y_1^n$ form a Markov chain, (16) follows as conditioning reduces entropy, and finally (17) follows from (13). Similarly, we can also show

$$\frac{1}{n}I(X_2^n;Y_2^n) \ge H(U_2|V_2) - \delta(P_e^n)$$

where the joint probability distribution factors as $p(x_1^n)p(x_2^n)$.

From the capacity region given in (12), we see that $(H(U_1|V_1) - \delta(P_e^n), H(U_2|V_2) - \delta(P_e^n)) \in \mathcal{C}_{IC}$ for all n. Then, since $\delta(P_e^n)\to 0$ as $n\to\infty$, and from the compactness of the capacity region, we can conclude that $P_e^n \to 0$ implies that $(H(U_1|V_1), H(U_2|V_2)) \in \mathcal{C}_{IC}$. This completes the proof.

APPENDIX B
PROOF OF THEOREM 2

Let the set of strongly typical n-tuples according to $p(x)$ be denoted by $T^n(X)$ where we have followed [21, Convention 2.11]. The definitions and notations of strong typicality can be extended to joint and conditional distributions in a similar manner [21].

Now, we start the achievability proof. Fix a joint distribution as in (2), where $p(u_1,v_1)$, $p(u_2,v_2)$, $p(y_1,y_2|x_1,x_2)$ are given while we are free to choose $p(q)$, $p(w_1,x_1|u_1,q)$, and $p(w_2,x_2|u_2,q)$.

Codebook Generation: First, generate one random n-sequence q^n in an i.i.d. fashion according to $p(q)$.

Next, for Transmitter 1, generate a codebook of size L_1 with $\frac{1}{n}\log L_1 > I(U_1;W_1|Q)+4\epsilon$, in which the codewords are generated i.i.d. with distribution $p(w_1|q)$. This codebook is denoted by $\mathcal{C}_w^1$.

For each possible source output u_1^n, count the number of codewords in $\mathcal{C}_w^1$ that are jointly typical with u_1^n. If there are at least $L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}$ codewords in $\mathcal{C}_w^1$ jointly typical with u_1^n, choose one uniformly at random, and call it $w_1^n(u_1^n)$. Intuitively, $w_1^n(U_1^n)$ is a compressed version of U_1^n that the receiver 2 decodes. This is similar to decoding part of the interference in the Han-Kobayashi scheme for the interference channel [14]. If there are fewer than $L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}$ codewords of $\mathcal{C}_w^1$ jointly typical with u_1^n, randomly choose one codeword from $\mathcal{C}_w^1$ to be $w_1^n(u_1^n)$. The reason why we require the number of codewords jointly typical with u_1^n to be large is to benefit the probability of error calculation later on in the proof. In a similar fashion, we generate $\mathcal{C}_w^2$.

Define $F(u_1^n,\ u_2^n)$ as the event that the number of $w_1^n\in\mathcal{C}_w^1$ jointly typical with u_1^n is larger than $L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}$ and the number of $w_2^n\in\mathcal{C}_w^2$ jointly typical with u_2^n is larger than $L_2 2^{-nI(U_2;W_2|Q)-2n\epsilon}$. We have

Lemma 3:

$$\Pr\{F^c(U_1^n,U_2^n)\} \le 3\epsilon \tag{18}$$

where superscript "c" denotes the complement.

Proof: For each $(q^n,u_1^n,u_2^n)\in T^n(QU_1U_2)$, define the random variable $\nu(i,u_1^n)$ as follows: $\nu(i,u_1^n)$ is 1 if the i-th codeword of $\mathcal{C}_w^1$ is jointly typical with u_1^n and 0 otherwise. Then

$$\begin{aligned}
2^{-nI(U_1;W_1|Q)-n\epsilon} &\le \mathbf{E}[\nu(i,u_1^n)|q^n]\\
&= \Pr\{\nu(i,u_1^n)=1|q^n\}\\
&\le 2^{-nI(U_1;W_1|Q)+n\epsilon} && (19)
\end{aligned}$$

$$\mathbf{V}[\nu(i,u_1^n)|q^n] \le \mathbf{E}^2[\nu(i,u_1^n)|q^n] \le \mathbf{E}[\nu(i,u_1^n)] \tag{20}$$

where $\mathbf{E}$ and $\mathbf{V}$ denote the expectation and variance, respectively. Further, define random variable $N(u_1^n)$ as the number of codewords in $\mathcal{C}_w^1$ that are jointly typical with u_1^n, i.e.,

$$N(u_1^n) = \sum_{i=1}^{L_1} \nu(i, u_1^n).$$

Then, from (19) and (20), we have

$$\begin{aligned} L_1 2^{-nI(U_1;W_1|Q)-n\epsilon} &\leq \mathbf{E}[N(u_1^n)|q^n] \\ &= \sum_{i=1}^{L_1} \mathbf{E}[\nu(i,u_1^n)|q^n] \\ &\leq L_1 2^{-nI(U_1;W_1|Q)+n\epsilon} \qquad (21) \\ \mathbf{V}[N(u_1^n)|q^n] &= \sum_{i=1}^{L_1} \mathbf{V}[\nu(i,u_1^n)|q^n] \leq \mathbf{E}[N(u_1^n)|q^n]. \qquad (22) \end{aligned}$$

Hence, we have

$$\begin{aligned} &\Pr\{N(u_1^n) \leq L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}|q^n\} \\ &\quad = \Pr\{\mathbf{E}[N(u_1^n)|q^n] - N(u_1^n) \\ &\qquad\quad \geq \mathbf{E}[N(u_1^n)|q^n] - L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}|q^n\} \\ &\quad \leq \Pr\Big\{\mathbf{E}[N(u_1^n)|q^n] - N(u_1^n) \geq L_1 2^{-nI(U_1;W_1|Q)-n\epsilon} \\ &\qquad\quad - L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}|q^n\Big\} \qquad (23) \\ &\quad \leq \Pr\Big\{\big|\mathbf{E}[N(u_1^n)|q^n] - N(u_1^n)\big| \geq L_1 2^{-nI(U_1;W_1|Q)-n\epsilon} \\ &\qquad\quad - L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}|q^n\Big\} \\ &\quad \leq \frac{\mathbf{V}[N(u_1^n)|q^n]}{\left(L_1 2^{-nI(U_1;W_1|Q)-n\epsilon} - L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}\right)^2} \qquad (24) \\ &\quad \leq \frac{\mathbf{E}[N(u_1^n)|q^n]}{\left(L_1 2^{-nI(U_1;W_1|Q)-n\epsilon} - L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}\right)^2} \qquad (25) \\ &\quad \leq \frac{L_1 2^{-nI(U_1;W_1|Q)+n\epsilon}}{\left(L_1 2^{-nI(U_1;W_1|Q)-n\epsilon} - L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}\right)^2} \qquad (26) \\ &\quad \leq \epsilon \qquad (27) \end{aligned}$$

where (23) and (26) follows from (21), (24) follows from Chebyshev's inequality, (25) follows from (22), and (27) is true when n is large enough. The same analysis applies for u_2^n.

Hence, we have proved that

$$\begin{aligned} &\Pr\{F^c(u_1^n,u_2^n)|q^n\} \\ &\quad = \Pr\Big\{N(u_1^n) \leq L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon} \\ &\qquad\qquad \text{or } N(u_2^n) \leq L_2 2^{-nI(U_2;W_2|Q)-2n\epsilon}|q^n\Big\} \\ &\quad \leq 2\epsilon \qquad (28) \end{aligned}$$

for all $(q^n,u_1^n,u_2^n) \in T^n(QU_1U_2)$ and all sufficiently large n. This means that

$$\begin{aligned} &\Pr\{F^c(U_1^n,U_2^n)\} \\ &\quad = \sum_{q^n,u_1^n,u_2^n} \Pr\Big\{F^c(U_1^n,U_2^n)|(U_1^n,U_2^n,Q^n) = (u_1^n,u_2^n,q^n)\Big\} \\ &\qquad \cdot \Pr\{(U_1^n,U_2^n,Q^n) = (u_1^n,u_2^n,q^n)\} \\ &\quad = \sum_{(q^n,u_1^n,u_2^n)\in T^n(QU_1U_2)} \Pr\{F^c(U_1^n,U_2^n)|(U_1^n,U_2^n,Q^n) \\ &\qquad\qquad = (u_1^n,u_2^n,q^n)\} \\ &\qquad \cdot \Pr\{(U_1^n,U_2^n,Q^n) = (u_1^n,u_2^n,q^n)\} \\ &\qquad + \sum_{(q^n,u_1^n,u_2^n)\notin T^n(QU_1U_2)} \Pr\{F^c(U_1^n,U_2^n)|(U_1^n,U_2^n,Q^n) \\ &\qquad\qquad = (u_1^n,u_2^n,q^n)\} \\ &\qquad \cdot \Pr\{(U_1^n,U_2^n,Q^n) = (u_1^n,u_2^n,q^n)\} \\ &\quad \leq 2\epsilon + \Pr\{(Q^n,U_1^n,U_2^n) \notin T^n(QU_1U_2)\} \qquad (29) \\ &\quad \leq 3\epsilon \qquad (30) \end{aligned}$$

where (29) follows from (28), and (30) follows when n is large enough from the asymptotic equipartition property (AEP) [20]. ∎

Lemma 3 demonstrates that with probability close to one, the number of sequences jointly typical with U_1^n and U_2^n in codebooks $\mathcal{C}_w^1$ and $\mathcal{C}_w^2$ are larger than $L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}$ and $L_2 2^{-nI(U_2;W_2|Q)-2n\epsilon}$, respectively. This fact will be used in the probability of error calculation.

For each possible u_1^n sequence, generate one x_1^n sequence in an i.i.d. fashion, conditioned on $w_1^n(u_1^n)$, u_1^n and q^n, according to $p(x_1|u_1,w_1,q)$. This x_1^n sequence is denoted by $x_1^n(u_1^n,\ w_1^n(u_1^n))$. The collection of all x_1^n sequences will be denoted as the codebook $\mathcal{C}_x^1$. Similarly, we generate the codebook $\mathcal{C}_x^2$.

Encoding: When Transmitter 1 observes the sequence u_1^n, it transmits $x_1^n(u_1^n, w_1^n(u_1^n))$. Similarly for Transmitter 2.

Decoding: Receiver 1 finds the unique pair (u_1^n, w_2^n), $u_1^n \in \mathcal{U}_1^n$, $w_2^n \in \mathcal{C}_w^2$, such that $(u_1^n, w_1^n(u_1^n), x_1^n(u_1^n, w_1^n(u_1^n)), w_2^n, y_1^n, v_2^n)$ are jointly typical and declares the first component of the pair as the transmitted source. If there are more than one pair, and the first component of the pairs are the same, then the decoder declares the transmitted source to be the first component. If there are more than one pair, and the first component of the pairs are not the same, an error is declared. Also, if no such pair exists, an error is declared. Similarly for Receiver 2.

Probability of Error Calculation: Denote by $E(u_1^n, w_2^n)$ the event $(u_1^n, w_1^n(u_1^n), X_1^n(u_1^n, w_1^n(u_1^n)), w_2^n, Y_1^n, V_2^n) \in T^n(U_1W_1X_1W_2Y_1V_2|q^n)$ for $(u_1^n, w_2^n) \in \mathcal{U}_1^n \times \mathcal{C}_w^2$. Further, denote by $G(u_1^n, u_2^n)$ the event $(u_1^n, u_2^n, w_1^n(u_1^n), w_2^n(u_2^n)) \in T^n(U_1U_2W_1W_2|q^n)$.

Then, the probability of error at Receiver 1 conditioned on $Q^n = q^n$, denoted by P_e^1, is given by

$$\begin{aligned}
&\Pr\left\{E^c(U_1^n, w_2^n(U_2^n)) \text{ or } \bigcup_{(u_1^n, w_2^n): u_1^n \neq U_1^n} E(u_1^n, w_2^n)\right\} \\
&\leq \Pr\Bigg\{E^c(U_1^n, w_2^n(U_2^n)) \text{ or } F^c(U_1^n, U_2^n) \text{ or } G^c(U_1^n, U_2^n) \\
&\qquad \text{or} \bigcup_{(u_1^n, w_2^n): u_1^n \neq U_1^n} E(u_1^n, w_2^n)\Bigg\} \\
&\leq \Pr\left\{E^c(U_1^n, w_2^n(U_2^n)) \text{ or } F^c(U_1^n, U_2^n) \text{ or } G^c(U_1^n, U_2^n)\right\} \\
&\quad + \Pr\left\{\bigcup_{(u_1^n, w_2^n): u_1^n \neq U_1^n} E(u_1^n, w_2^n)\middle| E \cap F \cap G\right\} \\
&\leq \Pr\left\{F^c(U_1^n, U_2^n)\right\} + \Pr\left\{G^c(U_1^n, U_2^n)|F\right\} \\
&\quad + \Pr\left\{E^c(U_1^n, w_2^n(U_2^n))|F \cap G\right\} \\
&\quad + \mathbf{E}\left\{\sum_{(u_1^n, w_2^n): u_1^n \neq U_1^n} \Pr\left\{E(u_1^n, w_2^n)|E \cap F \cap G\right\}\right\} \quad (31)
\end{aligned}$$

where we have used the short-hand E, F and G to denote events $E(U_1^n, w_2^n(U_2^n))$, $F(U_1^n, U_2^n)$ and $G(U_1^n, U_2^n)$, respectively.

The first term in (31) is bounded by 3ϵ as shown by (18). From the achievability results of multi-terminal rate-distortion theory [22], the second term in (31) is bounded by ϵ for sufficiently large n. The third term in (31) is bounded by ϵ for sufficiently large n based on the AEP [20]. Hence, from now on, we will concentrate on the fourth term in (31).

The fourth term in (31) may be upper bounded by the sum of the following four terms, which will be denoted by a_1, a_2, a_3, and a_4, respectively:

$$a_1 \triangleq \mathbf{E}\left\{\sum_{\substack{u_1^n \neq U_1^n \\ w_1^n(u_1^n) \neq w_1^n(U_1^n)}} \Pr\left\{E(u_1^n, w_2^n(U_2^n))|E \cap F \cap G\right\}\right\}$$

$$a_2 \triangleq \mathbf{E}\left\{\sum_{\substack{u_1^n \neq U_1^n \\ w_1^n(u_1^n) \neq w_1^n(U_1^n) \\ w_2^n \neq w_2^n(U_2^n)}} \Pr\left\{E(u_1^n, w_2^n)|E \cap F \cap G\right\}\right\}$$

$$a_3 \triangleq \mathbf{E}\left\{\sum_{\substack{u_1^n \neq U_1^n \\ w_1^n(u_1^n) = w_1^n(U_1^n)}} \Pr\left\{E(u_1^n, w_2^n(U_2^n))|E \cap F \cap G\right\}\right\}$$

$$a_4 \triangleq \mathbf{E}\left\{\sum_{\substack{u_1^n \neq U_1^n \\ w_1^n(u_1^n) = w_1^n(U_1^n) \\ w_2^n \neq w_2^n(U_2^n)}} \Pr\left\{E(u_1^n, w_2^n)|E \cap F \cap G\right\}\right\}.$$

First, we upper bound a_1. Define the set

$$\begin{aligned}
\mathcal{B}_1 = \{u_1^n \in \mathcal{U}_1^n : u_1^n \neq U_1^n, w_1^n(u_1^n) \neq w_1^n(U_1^n), \\
(u_1^n, w_1^n(u_1^n)) \in T^n(U_1 W_1 | Y_1^n V_2^n w_2^n(U_2^n) q^n)\}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\mathbf{E}\left\{|\mathcal{B}_1|\big|E \cap F \cap G\right\} \\
&\leq 2^{nH(U_1|Y_1,V_2,W_2,Q)+n\epsilon} 2^{nH(W_1|U_1,Y_1,V_2,W_2,Q)+n\epsilon} \\
&\quad \cdot 2^{-nH(W_1|U_1,Q)+n\epsilon} \qquad (32)
\end{aligned}$$

which follows because there are fewer than $2^{nH(U_1|Y_1,V_2,W_2,Q)+n\epsilon}$ u_1^n's in $T^n(U_1|Y_1^n V_2^n w_2^n(U_2^n) q^n)$, and for each u_1^n, the probability that $w_1^n(u_1^n)$ is in $T^n(W_1|U_1^n Y_1^n V_2^n w_2^n(U_2^n) q^n)$ is less than $2^{nH(W_1|U_1,Y_1,V_2,W_2,Q)+n\epsilon} 2^{-nH(W_1|U_1,Q)+n\epsilon}$ due to the symmetry of the random codebook generation. Hence, we may write

$$\begin{aligned}
a_1 &= \mathbf{E}\left\{\sum_{u_1^n \in \mathcal{B}_1} \Pr\left\{E(u_1^n, w_2^n(U_2^n))|E \cap F \cap G\right\}\right\} && (33) \\
&\leq \mathbf{E}\left\{|\mathcal{B}_1| \max_{u_1^n \in \mathcal{B}_1} \Pr\left\{E(u_1^n, w_2^n(U_2^n))|E \cap F \cap G\right\}\right\} \\
&= \mathbf{E}\Big\{|\mathcal{B}_1| \max_{u_1^n \in \mathcal{B}_1} \Pr\{X_1^n(u_1^n, w_1^n(u_1^n)) \in T^n(X_1|u_1^n \\
&\qquad w_1^n(u_1^n) w_2^n(U_2^n) Y_1^n V_2^n q^n)|E \cap F \cap G\}\Big\} && (34) \\
&\leq \mathbf{E}\Big\{|\mathcal{B}_1| \max_{u_1^n \in \mathcal{B}_1} 2^{nH(X_1|U_1,W_1,W_2,Y_1,V_2,Q)+n\epsilon} \\
&\qquad 2^{-nH(X_1|U_1,W_1,Q)+n\epsilon}\big|E \cap F \cap G\Big\} && (35) \\
&\leq 2^{nH(U_1)} 2^{-nI(U_1,W_1,X_1;Y_1,V_2|W_2,Q)+5n\epsilon} && (36) \\
&\leq 2^{nH(U_1)} 2^{-nI(X_1;Y_1,V_2|W_2,Q)+5n\epsilon} && (37)
\end{aligned}$$

where (34) follows from the definition of event $E(u_1^n, w_2^n)$ and $\mathcal{B}_1$, (35) follows from the fact that there are fewer than $2^{nH(X_1|U_1,W_1,W_2,Y_1,V_2,Q)+n\epsilon}$ x_1^n's in $T^n(X_1|u_1^n w_1^n(u_1^n) w_2^n(U_2^n) Y_1^n V_2^n q^n)$ and of these x_1^n's, each has a probability of less than $2^{-nH(X_1|U_1,W_1,Q)+n\epsilon}$ of being chosen as $x_1^n(u_1^n, w_1^n(u_1^n))$; (36) follows from (32); and (37) follows because the distribution in (2) satisfies the Markov chain relationship $(U_1, W_1) \to (X_1, W_2, Q) \to (V_2, Y_1)$. Next, we upper bound a_2. Define the set

$$\begin{aligned}
\mathcal{B}_2 = \{u_1^n \in \mathcal{U}_1^n, w_2^n \in \mathcal{C}_w^2 : u_1^n \neq U_1^n, w_1^n(u_1^n) \neq w_1^n(U_1^n), w_2^n \\
\neq w_2^n(U_2^n), (u_1^n, w_1^n(u_1^n), w_2^n) \in T^n(U_1 W_1 W_2 | Y_1^n V_2^n q^n)\}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\mathbf{E}\{|\mathcal{B}_2|\big|E \cap F \cap G\} \\
&\leq 2^{nH(W_2|Y_1,V_2,Q)+n\epsilon} 2^{-nH(W_2|Q)+n\epsilon} 2^{nH(U_1|W_2,Y_1,V_2,Q)+n\epsilon} \\
&\quad (L_2 - 1) 2^{nH(W_1|U_1,W_2,Y_1,V_2,Q)+n\epsilon} 2^{-nH(W_1|U_1,Q)+n\epsilon}.
\end{aligned}$$

Similarly to (33)–(36), we may write

$$\begin{aligned} a_2 &= \mathbf{E}\left\{\sum_{(u_1^n,w_2^n)\in\mathcal{B}_2} \Pr\left\{E(u_1^n,w_2^n)|E\cap F\cap G\right\}\right\} \\ &\le 2^{nH(U_1)}L_2 2^{-nI(U_1,W_1,X_1,W_2;V_2,Y_1|Q)+7n\epsilon} \\ &= 2^{nH(U_1)}L_2 2^{-nI(X_1,W_2;V_2,Y_1|Q)+7n\epsilon} \qquad (38) \end{aligned}$$

where (38) follows using the same reasoning as for (37). Next, we upper bound a_3. Define the set

$$\begin{aligned} \mathcal{B}_3 = \big\{u_1^n \in \mathcal{U}_1^n : u_1^n \ne U_1^n, w_1^n(u_1^n) = w_1^n(U_1^n),\\ u_1^n \in T^n(U_1|w_1^n(U_1^n)Y_1^n V_2^n w_2^n(U_2^n)q^n)\big\}. \end{aligned}$$

Then, we have

$$\begin{aligned} &\mathbf{E}\left\{|\mathcal{B}_3|\big|E\cap F\cap G\right\} \\ &\quad\le 2^{nH(U_1|W_1,Y_1,V_2,W_2,Q)+n\epsilon}\frac{1}{2^{-nI(U_1;W_1|Q)-2n\epsilon}L_1} \end{aligned}$$

which follows from the fact that we always choose randomly from at least $L_1 2^{-nI(U_1;W_1|Q)-2n\epsilon}$ choices to get $w_1^n(u_1^n)$. Similarly to (33)–(36), we may write

$$\begin{aligned} a_3 &= \mathbf{E}\left\{\sum_{u_1^n\in\mathcal{B}_3} \Pr\left\{E(u_1^n,w_2^n(U_2^n))|E\cap F\cap G\right\}\right\} \\ &\le \frac{2^{nH(U_1)}}{L_1}2^{-nI(U_1,X_1;Y_1,V_2|W_1,W_2,Q)+5n\epsilon} \\ &\le \frac{2^{nH(U_1)}}{L_1}2^{-nI(X_1;Y_1,V_2|W_1,W_2,Q)+5n\epsilon} \qquad (39) \end{aligned}$$

where (39) follows using the same reasoning as for (37). Finally, we upper bound a_4. Define the set

$$\begin{aligned} \mathcal{B}_4 = \{u_1^n \in \mathcal{U}_1^n, w_2^n \in \mathcal{C}_w^2 : u_1^n \ne U_1^n, w_1^n(u_1^n) = w_1^n(U_1^n),\\ w_2^n \ne w_2^n(U_2^n), (u_1^n,w_2^n) \in T^n(U_1W_2|w_1^n(U_1^n)Y_1^nV_2^nq^n)\}. \end{aligned}$$

Then, we have

$$\begin{aligned} &\mathbf{E}\left\{|\mathcal{B}_4|\big|E\cap F\cap G\right\} \\ &\quad\le 2^{nH(W_2|Y_1,V_2,W_1,Q)+n\epsilon}2^{-nH(W_2|Q)+n\epsilon}(L_2-1) \\ &\quad\quad 2^{nH(U_1|W_1,W_2,Y_1,V_2,Q)+n\epsilon}\frac{1}{2^{-nI(U_1;W_1|Q)-2n\epsilon}L_1}. \end{aligned}$$

Similarly to (33)–(36), we may write

$$\begin{aligned} a_4 &= \mathbf{E}\left\{\sum_{(u_1^n,w_2^n)\in\mathcal{B}_4} [E(u_1^n,w_2^n)|E\cap F\cap G]\right\} \\ &\le \frac{L_2}{L_1}2^{nH(U_1)}2^{-nI(U_1,X_1,W_2;Y_1,V_2|W_1,Q)+7n\epsilon} \\ &\le \frac{L_2}{L_1}2^{nH(U_1)}2^{-nI(X_1,W_2;Y_1,V_2|W_1,Q)+7n\epsilon} \qquad (40) \end{aligned}$$

where (40) follows using the same reasoning as for (37).

We have similar probability of error calculations at Receiver 2. Since

$$P_e^n \le \mathbf{E}_{Q^n}[P_e^1 + P_e^2],$$

for this achievability scheme, as long as the following equations are satisfied:

$$\begin{aligned} H(U_1) &\le I(X_1;V_2,Y_1|W_2,Q) \\ H(U_1) - \frac{1}{n}\log L_1 &\le I(X_1;V_2,Y_1|W_1,W_2,Q) \\ H(U_1) + \frac{1}{n}\log L_2 &\le I(W_2,X_1;V_2,Y_1|Q) \\ H(U_1) + \frac{1}{n}\log L_2 - \frac{1}{n}\log L_1 &\le I(W_2,X_1;V_2,Y_1|W_1,Q) \\ H(U_2) &\le I(X_2;V_1,Y_2|W_1,Q) \\ H(U_2) - \frac{1}{n}\log L_2 &\le I(X_2;V_1,Y_2|W_1,W_2,Q) \\ H(U_2) + \frac{1}{n}\log L_1 &\le I(W_1,X_2;V_1,Y_2|Q) \\ H(U_2) + \frac{1}{n}\log L_1 - \frac{1}{n}\log L_2 &\le I(W_1,X_2;V_1,Y_2|W_2,Q) \\ \frac{1}{n}\log L_1 &\ge I(U_1;W_1|Q) \text{ and} \\ \frac{1}{n}\log L_2 &\ge I(U_2;W_2|Q) \end{aligned}$$

for some $p(q)$, $p(w_1,x_1|u_1,q)$, and $p(w_2,x_2|u_2,q)$, the probability of error is arbitrarily small for sufficiently large n.

By Fourier-Motzkin elimination, we obtain the sufficient conditions given in Theorem 2.

Appendix C
Proof of Lemma 2

Due to the fact that the proof of this lemma is very similar to the proof of the capacity region in [17], we omit certain details. For notational convenience, denote the channel of $p(y_1|x_1)$ as $\bar{V}_1$ and the channel $p(y_2|x_1,x_2)$ as $\bar{V}_2$, where

$$\begin{aligned} \bar{V}_1(a|b) &= \Pr\{Y_1 = a|X_1 = b\} \\ &\text{and} \\ \bar{V}_2(c|b,d) &= \Pr\{Y_2 = c|X_1 = b, X_2 = d\}. \end{aligned}$$

A) Converse Result: The converse result derived in this subsection is valid for any Z-interference channel satisfying Condition 1. The tool that we are using comes from the following lemma.

Lemma 4 [21, pp. 314, eq. (3.34)]: For any n, and any random variables Y^n and Z^n and W, we have

$$\begin{aligned} &H(Z^n|W) - H(Y^n|W) \\ &\quad= \sum_{i=1}^n \Big(H(Z_i|Y^{i-1},Z_{i+1},Z_{i+2},\ldots,Z_n,W) \\ &\qquad\qquad - H(Y_i|Y^{i-1},Z_{i+1},Z_{i+2},\ldots,Z_n,W)\Big). \end{aligned}$$

Since the rate triplet $(R_{1s},\ R_{1p},\ R_{2p})$ is achievable, there exist two sequences of codebooks 1 and 2, denoted by $\mathcal{C}_1^n$ and $\mathcal{C}_2^n$, of rate $R_{1s}+R_{1p}$ and R_{2p}, and probability of error P_e^n, where $P_e^n \to 0$ as $n\to\infty$. Let us define X_1^n and X_2^n to be uniformly distributed on codebooks 1 and 2, respectively. Let Y_1^n be connected via $\bar{V}_1^n$ to X_1^n and Y_2^n be connected via $\bar{V}_2^n$ to X_1^n and X_2^n.

We start the converse with Fano's inequality [20]:

$$\begin{aligned}nR_{1p} &= H(W_{1p})\\ &\leq I(W_{1p};Y_1^n)+n\delta(P_e^n) &(41)\\ &\leq I(W_{1p};Y_1^n|W_{1s})+n\delta(P_e^n) &(42)\\ &= H(Y_1^n|W_{1s})-H(Y_1^n|W_{1s},W_{1p},X_1^n)+n\delta(P_e^n) &(43)\\ &= H(Y_1^n|W_{1s})-H(Y_1^n|X_1^n)+n\delta(P_e^n) &(44)\\ &= H(Y_1^n|W_{1s})-\sum_{i=1}^n H(Y_{1i}|X_{1i})+n\delta(P_e^n) &(45)\end{aligned}$$

where in (41), $\delta(x)$ is a nonnegative function approaching zero as $x\longrightarrow 0$, (42) follows from the fact that W_{1s} and W_{1p} are independent, (43) follows from the fact that without loss of generality, we may consider deterministic encoders, (44) follows from the Markov chain relationship $(W_{1s},W_{1p})\longrightarrow X_1^n\longrightarrow Y_1^n$, and (45) follows from the memoryless nature of $\bar{V}_1^n$. We also have

$$\begin{aligned}nR_{1s}+nR_{1p} &= H(W_{1p},W_{1s})\\ &\leq I(W_{1p},W_{1s};Y_1^n)+n\delta(P_e^n)\\ &\leq I(X_1^n;Y_1^n)+n\delta(P_e^n) &(46)\\ &\leq \sum_{i=1}^n I(X_{1i};Y_{1i})+n\delta(P_e^n) &(47)\end{aligned}$$

where (46) follows from the data processing inequality [20]. Furthermore, we have

$$\begin{aligned}nR_{2p} &= H(W_{2p})=H(W_{2p}|W_{1s}) &(48)\\ &\leq I(W_{2p};Y_2^n|W_{1s})+n\delta(P_e^n)\\ &\leq I(X_2^n;Y_2^n|W_{1s})+n\delta(P_e^n) &(49)\\ &= H(Y_2^n|W_{1s})-H(Y_2^n|X_2^n,W_{1s})+n\delta(P_e^n)\\ &\leq \sum_{i=1}^n H(Y_{2i})-H(Y_2^n|X_2^n,W_{1s})+n\delta(P_e^n) &(50)\\ &\leq n\tau-H(Y_2^n|X_2^n,W_{1s})+n\delta(P_e^n) &(51)\end{aligned}$$

where (48) follows from the independence of W_{2p} and W_{1s}, (49) follows from the Markov chain relationship $W_{2p}\longrightarrow(X_2^n,W_{1s})\longrightarrow Y_2^n$, (50) follows from the fact that conditioning reduces entropy, and (51) follows from the definition of τ in (3).

Let us define another channel, $\hat{V}_2:\mathcal{X}_1\longrightarrow\mathcal{Y}_2$, as

$$\hat{V}_2(t|x_1)=V_2(t|x_1,\bar{x}_2)$$

where $\bar{x}_2$ is an arbitrary element in $\mathcal{X}_2$. Further, let us define another sequence of random variables, T^n, which is connected via $\hat{V}_2^n$, the memoryless channel $\hat{V}_2$ used n times, to X_1^n, i.e., $T_i\longrightarrow X_{1i}\longrightarrow T_{\{i\}^c},X_{1\{i\}^c},X_2^n,Y_1^n,Y_2^n$. Also define $\bar{x}_2^n$ as the n-sequence with $\bar{x}_2$ repeated n times. Since the channel under consideration satisfies condition 1 , we have

$$H(Y_2^n|X_2^n,W_{1s})=H(T^n|W_{1s}). \qquad (52)$$

By applying Lemma 4, we have

$$\begin{aligned}&H(T^n|W_{1s})-H(Y_1^n|W_{1s})\\ &\quad=\sum_{i=1}^n H(T_i|Y_1^{i-1},T_{i+1},T_{i+2},\ldots,T_n,W_{1s})\\ &\qquad-H(Y_{1i}|Y_1^{i-1},T_{i+1},T_{i+2},\ldots,T_n,W_{1s}). \qquad (53)\end{aligned}$$

Furthermore, since conditioning reduces entropy, we can write

$$\begin{aligned}H(Y_1^n|W_{1s}) &= \sum_{i=1}^n H(Y_{1i}|Y_1^{i-1},W_{1s})\\ &\geq \sum_{i=1}^n H(Y_{1i}|Y_1^{i-1},T_{i+1},T_{i+2},\ldots,T_n,W_{1s}).\end{aligned} \qquad (54)$$

Define the following auxiliary random variables:

$$W_i=Y_1^{i-1},T_{i+1},T_{i+2},\ldots,T_n,W_{1s},\qquad i=1,2,\ldots,n.$$

Further define Q as a random variable that is uniform on the set $\{1,2,\ldots,n\}$ and independent of everything else. Also, define the following auxiliary random variables:

$$W=(W_Q,Q),\quad X_1=X_{1Q},\quad Y_1=Y_{1Q}\quad\text{and}\quad T=T_Q.$$

Then, from (53) and (54), we have

$$n^{-1}\left(H(T^n|W_{1s})-H(Y_1^n|W_{1s})\right)=H(T|W)-H(Y_1|W) \qquad (55)$$

$$n^{-1}H(Y_1^n|W_s)\geq H(Y_1|W). \qquad (56)$$

Due to the memoryless nature of $\bar{V}_1^n$ and $\hat{V}_2^n$, the fact that Q is independent of everything else, and the Markov chain relationship $T_i\longrightarrow X_{1i}\longrightarrow Y_{1i}$, for $i=1,2,\ldots,n$, the joint distribution of W, X_1, Y_1, T satisfies

$$p(w,x_1,y_1,t)=p(w)p(x_1|w)V_1(y_1|x_1)V_2(t|x_1,\bar{x}_2). \qquad (57)$$

From (55) and (56), we may conclude that there exists a number $\gamma\geq 0$ such that

$$\begin{aligned}\frac{1}{n}H(T^n|W_{1s}) &= H(T|W)+\gamma, &(58)\\ \frac{1}{n}H(Y_1^n|W_{1s}) &= H(Y_1|W)+\gamma. &(59)\end{aligned}$$

By combining (45), (47), (51), (52), (57)–(59), and allowing $n\longrightarrow\infty$, we obtain the following converse result: for any Z-interference channel that satisfies Condition 1 and the case where Receiver 2 has side information W_{1s}, the achievable rate triplets $(R_{1s},\ R_{1p},\ R_{2p})$ must satisfy

$$\begin{aligned}R_{1p} &\leq H(Y_1|W)+\gamma-H(Y_1|X_1) &(60)\\ R_{1s}+R_{1p} &\leq I(X_1;Y_1) &(61)\\ R_{2p} &\leq \tau-H(T|W)-\gamma &(62)\end{aligned}$$

for some $\gamma\geq 0$ and distribution $p(w)p(x_1|w)$, where the mutual informations and entropies are evaluated using $p(w,x_1,y_1,t)=p(w)p(x_1|w)V_1(y_1|x_1)V_2(t|x_1,\bar{x}_2)$.

B) Achievability Result: We use the coding scheme of Lemma 1 which is valid for any IC-SI. We design a codebook at Transmitter 1 such that the inner codebook carries the side information at the Receiver 2, i.e., W_{1s}, and part of W_{1p}, and the outer codebook carries the remaining part of W_{1p}. More specifically, the inner codebook is of rate $R_{1s}+\gamma$, and the outer codebook is of rate $R_{1p}-\gamma$. Then, we have the achievable rate region as the union over all $p(w)p(x_1|w)p(x_2)$ of

$$\begin{aligned} R_{1p} &\le H(Y_1|W)+\gamma-H(Y_1|X_1) && (63)\\ R_{1s}+R_{1p} &\le I(X_1;Y_1) && (64)\\ R_{2p} &\le I(X_2;Y_2|W) && (65)\\ R_{2p} &\le I(W,X_2;Y_2)-\gamma && (66) \end{aligned}$$

where the mutual informations are evaluated using

$$p(w,x_1,x_2,y_1,y_2)= p(w)p(x_1|w)p(x_2)V_1(y_1|x_1)V_2(y_2|x_1,x_2).$$

C) Capacity Region: Making use of Conditions 1 and 2 in the exact same way as in [17, Sec. VII-C], we can show that the converse result in (60)–(62) and the achievability result in (63)–(66) are the same for Z-interference channels satisfying Conditions 1 and 2, and hence the capacity region in this case is given in Lemma 2.

Appendix D
Proof of Theorem 3

We use the n-letter characterization of $\mathcal{C}_{IC-SI}$ provided in the next lemma. Define $\mathcal{G}^n$ as

$$\mathcal{G}^n=\Bigg\{(R_{1s},R_{1p},R_{2s},R_{2p}): \begin{aligned}[t] &R_{1p}\le \frac{1}{n}I(X_1^n;Y_1^n|S_{1s}^n,S_{2s}^n)\\ &R_{1s}+R_{1p}\le \frac{1}{n}I(X_1^n;Y_1^n|S_{2s}^n)\\ &R_{2p}\le \frac{1}{n}I(X_2^n;Y_2^n|S_{1s}^n,S_{2s}^n)\\ &R_{2s}+R_{2p}\le \frac{1}{n}I(X_2^n;Y_2^n|S_{1s}^n)\\ &R_{1s},R_{1p},R_{2s},R_{2p}\ge 0\\ &\text{for any } p^n(s_{1s}^n)p^n(s_{2s}^n)p^n(x_1^n|s_{1s}^n)p^n(x_2^n|s_{2s}^n)\Bigg\}. \end{aligned}$$

Lemma 5: The capacity region of the IC-SI as defined in Section IV-B is

$$\mathcal{C}_{IC-SI}=\lim_{n\longrightarrow\infty} \quad \mathcal{G}^n \tag{67}$$

where the limit of the region is as defined in [1, Th. 5].

Proof: We first start with the proof of achievability. Fix distributions $p(s_{1s})$, $p(x_1|s_{1s})$, $p(s_{2s})$, and $p(x_2|s_{2s})$. For codebook at Transmitter k, $k=1,2$, we generate an inner codebook of $2^{NR_{ks}}$ i.i.d. codewords of length N with probability $\prod_{i=1}^{N}p(s_{ks,i})$. Then, for each codeword of the inner codebook, we generate an outer codebook of $2^{NR_{kp}}$ i.i.d. codewords of length N with probability $\prod_{i=1}^{N}p(x_{k,i}|s_{ks,i})$. For $W_{ks}=w_{ks}$ and $W_{kp}=w_{kp}$, Transmitter k sends the w_{kp}-th codeword of the w_{ks}-th outer codebook. For decoding, Receiver 1 finds the codeword in all possible outer codebooks that is jointly typical with the received sequence and the w_{2s}-th codeword of the inner codebook of Transmitter 2. Receiver 2 operates similarly. The probability of error analysis follows from standard arguments [20], and we can show that the probability of error can be driven to zero as $N\longrightarrow\infty$, as long as the rates satisfy the following conditions:

$$\begin{aligned} R_{1p} &\le I(X_1;Y_1|S_{1s},S_{2s})\\ R_{1s}+R_{1p} &\le I(X_1;Y_1|S_{2s})\\ R_{2p} &\le I(X_2;Y_2|S_{1s},S_{2s})\\ R_{2s}+R_{2p} &\le I(X_2;Y_2|S_{1s}). \end{aligned}$$

For each n, similarly to [1, Th. 5], by treating the interference channel $p^n(y_1^n,y_2^n|x_1^n,x_2^n)$, which is a product channel of $p(y_1,y_2|x_1,x_2)$, as a memoryless channel, we conclude that the rates satisfying the following conditions are achievable for any n:

$$\begin{aligned} R_{1p} &\le \frac{1}{n}I(X_1^n;Y_1^n|S_{1s}^n,S_{2s}^n)\\ R_{1s}+R_{1p} &\le \frac{1}{n}I(X_1^n;Y_1^n|S_{2s}^n)\\ R_{2p} &\le \frac{1}{n}I(X_2^n;Y_2^n|S_{1s}^n,S_{2s}^n)\\ R_{2s}+R_{2p} &\le \frac{1}{n}I(X_2^n;Y_2^n|S_{1s}^n) \end{aligned}$$

i.e., any rate quadruplet $(R_{1s},R_{1p},R_{2s},R_{2p})\in\mathcal{G}^n$ is achievable. By the definition of the capacity region, the limiting points of $\mathcal{G}^n$ are also achievable, and thus, we have proved the achievability of all the points in $\mathcal{C}_{IC-SI}$.

We next prove the converse. For any $\left(2^{nR_{1s}},2^{nR_{1p}},2^{nR_{2s}},2^{nR_{2p}},n\right)$ code, denote its input to the channel as random variables X_1^n and X_2^n and the output of the channel as random variables Y_1^n and Y_2^n.

Arbitrarily choose $M_{1s}\triangleq 2^{nR_{1s}}$ n-letter sequences $u_1^{1s},u_2^{1s},\ldots,u_{M_{1s}}^{1s}$ all in $\mathcal{X}_1^n$, and $M_{2s}\triangleq 2^{nR_{2s}}$ n-letter sequences $u_1^{2s},u_2^{2s},\ldots,u_{M_{2s}}^{2s}$ all in $\mathcal{X}_2^n$. We then form a one-to-one correspondence between W_{1s}, W_{2s} and S_{1s}^n, S_{2s}^n, respectively, by

$$p^n(S_{1s}^n=u^n|W_{1s}=w_{1s}) =\begin{cases}1, & \text{if } u^n=u_{w_{1s}}^{1s}, \quad w_{1s}=1,2,\ldots,M_{1s}\\ 0, & \text{otherwise}\end{cases} \tag{68}$$

$$p^n(S_{2s}^n=u^n|W_{2s}=w_{2s}) =\begin{cases}1, & \text{if } u^n=u_{w_{2s}}^{2s}, \quad w_{2s}=1,2,\ldots,M_{2s}\\ 0, & \text{otherwise.}\end{cases} \tag{69}$$

By Fano's inequality [20], we have

$$\begin{aligned} nR_{1p} &= H(W_{1p})=H(W_{1p}|W_{1s},W_{2s})\\ &= I(W_{1p};Y_1^n|W_{1s},W_{2s})\\ &\quad + H(W_{1p}|Y_1^n,W_{1s},W_{2s}) \end{aligned}$$

$$\begin{aligned}&\le I(W_{1p};Y_1^n|W_{1s},W_{2s})+H(W_{1p}|Y_1^n,W_{2s})\\&\le I(W_{1p};Y_1^n|W_{1s},W_{2s})+n\delta(P_e^n) &(70)\\&\le I(X_1^n;Y_1^n|W_{1s},W_{2s})+n\delta(P_e^n) &(71)\\&= I(X_1^n;Y_1^n|S_{1s}^n,S_{2s}^n)+n\delta(P_e^n) &(72)\end{aligned}$$

where $\delta(x)$ in (70) is a nonnegative function approaching zero as $x\to 0$, (71) follows from the data processing inequality [20] as the distributions factor as $p(w_{1p})\ p(w_{1s})\ p(x_1^n|w_{1p},w_{1s})\ p(w_{2p})\ p(w_{2s})\ p(x_2^n|w_{2p},w_{2s})\ p(y_1^n|x_1^n,x_2^n)$ and the Markov chain relationship $(W_{1p},W_{1s})\to(X_1^n,W_{2s})\to Y_1^n$ is satisfied, and finally (72) follows from the definitions of the sequences S_{1s}^n and S_{2s}^n in (68) and (69), respectively. We also have

$$\begin{aligned}nR_{1s}+nR_{1p}&=H(W_{1s},W_{1p})\\&=H(W_{1s},W_{1p}|W_{2s})\\&=I(W_{1s},W_{1p};Y_1^n|W_{2s})\\&\quad+H(W_{1s},W_{1p}|Y_1^n,W_{2s})\\&\le I(W_{1s},W_{1p};Y_1^n|W_{2s})+n\delta(P_e^n)\\&\le I(X_1^n;Y_1^n|W_{2s})+n\delta(P_e^n) &(73)\\&=I(X_1^n;Y_1^n|S_{2s}^n)+n\delta(P_e^n) &(74)\end{aligned}$$

where (73) follows using the same reasoning as for (71), and (74) follows using the same reasoning as for (72).

Similarly, we have

$$\begin{aligned}nR_{2p}&\le I(X_2^n;Y_2^n|S_{1s}^n,S_{2s}^n)+n\delta(P_e^n)\\nR_{2s}+nR_{2p}&\le I(X_2^n;Y_2^n|S_{1s}^n)+n\delta(P_e^n).\end{aligned}$$

Hence, we have proved that for all n, $(R_{1s}-\delta(P_e^n),R_{1p}-\delta(P_e^n),R_{2s}-\delta(P_e^n),R_{2p}-\delta(P_e^n))\in\mathcal{G}^n$.

Since the region $\mathcal{C}_{IC-SI}$ as defined in (67) contains $\mathcal{G}^n$ for every n [1, Th. 5], we have

$$(R_{1s}-\delta(P_e^n),R_{1p}-\delta(P_e^n),R_{2s}-\delta(P_e^n),R_{2p}-\delta(P_e^n))\in\mathcal{C}_{IC-SI} \quad (75)$$

for all n. For codes where $P_e^n\to 0$ as $n\to\infty$, we have

$$(R_{1s},R_{1p},R_{2s},R_{2p})\in\mathcal{C}_{IC-SI} \quad (76)$$

since $\mathcal{C}_{IC-SI}$ is closed [1, Th. 5]. This concludes the converse part of the proof. ■

Note that in the proof of achievability, the auxiliary random variables S_{1s}^n and S_{2s}^n take on the meaning of inner codewords at Transmitters 1 and 2, respectively.

Now that we have the n-letter characterization of the capacity region of the IC-SI in Lemma 5, we are ready to prove Theorem 3.

The achievability part of the proof is straightforward. If (7) holds, then there exists a rate quadruplet $(R_{1s},R_{1p},R_{2s},R_{2p})$ in the interior of $\mathcal{C}$ such that $H(V_k)\le R_{ks}$ and $H(U_k|V_k)\le R_{kp}$ for $k=1,2$. Transmitter k first compresses V_k into index W_{ks} with rate $H(V_k)$, and then $U_k|V_k=v_k$ into index $W_{kp}(v_k)$ into rate $H(U_k|V_k)$, for all v_k in the typical set. Then the indices can be transmitted reliably over the channel since $(R_{1s},R_{1p},R_{2s},R_{2p})$ is in the capacity region of the underlying interference channel with message side information W_{1s} at Receiver 2 and W_{2s} at Receiver 1.

To prove the converse, we write

$$\begin{aligned}nH(U_1|V_1)&=H(U_1^n|V_1^n)=H(U_1^n|V_1^n,V_2^n)\\&=I(U_1^n;Y_1^n|V_1^n,V_2^n)+H(U_1^n|Y_1^n,V_1^n,V_2^n)\\&\le I(U_1^n;Y_1^n|V_1^n,V_2^n)+H(U_1^n|Y_1^n,V_2^n)\\&\le I(U_1^n;Y_1^n|V_1^n,V_2^n)+n\delta(P_e^n) &(77)\\&\le I(X_1^n;Y_1^n|V_1^n,V_2^n)+n\delta(P_e^n) &(78)\end{aligned}$$

where (77) follows from Fano's inequality, and (78) follows from the data processing inequality, in other words, from the Markov chain relationship $(U_1^n,V_1^n)\to(X_1^n,V_2^n)\to Y_1^n$. We can also write

$$\begin{aligned}nH(V_1)+nH(U_1|V_1)&=nH(U_1,V_1)\\&=nH(U_1) &(79)\\&=H(U_1^n)\\&=H(U_1^n|V_2^n)\\&=I(U_1^n;Y_1^n|V_2^n)+H(U_1^n|Y_1^n,V_2^n)\\&\le I(U_1^n;Y_1^n|V_2^n)+n\delta(P_e^n) &(80)\\&\le I(X_1^n;Y_1^n|V_2^n)+n\delta(P_e^n) &(81)\end{aligned}$$

where (79) follows because V_1 is a deterministic function of U_1, (80) follows from Fano's inequality, and (81) follows from the same reasoning as applied to (78). Similarly, we have

$$\begin{aligned}nH(U_2|V_2)&\le I(X_2^n;Y_2^n|V_1^n,V_2^n)+n\delta(P_e^n) &(82)\\nH(V_2)+nH(U_2|V_2)&\le I(X_2^n;Y_2^n|V_1^n)+n\delta(P_e^n). &(83)\end{aligned}$$

Hence, from (78), (81)–(83), we have

$$\big(H(V_1)-\delta(P_e^n),H(U_1|V_1)-\delta(P_e^n),H(V_2)-\delta(P_e^n),\\ H(U_2|V_2)-\delta(P_e^n)\big)\in\mathcal{G}^n$$

which by the same reasoning as applied to (75) and (76), for codes where $P_e^n\to 0$ as $n\to\infty$, we have

$$(H(V_1),H(U_1|V_1),H(V_2),H(U_2|V_2))\in\mathcal{C}_{IC-SI}.$$

This concludes the proof.

REFERENCES

[1] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Math. Stat. Prob.*, Berkeley, CA, 1961, vol. 1, pp. 611–644.
[2] M. Salehi and E. Kurtas, "Interference channels with correlated sources," in *Proc. IEEE Int. Symp. Information Theory*, San Antonio, TX, Jan. 1993.
[3] W. Liu and B. Chen, "Interference channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, to be published.
[4] T. M. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.
[5] S. Shamai and S. Verdú, "Capacity of channels with uncoded side information," *Eur. Trans. Telecommun. and Related Technol.*, vol. 6, no. 5, pp. 587–600, Sep.-Oct. 1995.
[6] D. Gündüz, E. Erkip, A. Goldsmith, and H. V. Poor, "Source and channel coding for correlated sources over multiuser channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 3927–3944, Sep. 2009.
[7] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.

[8] D. Gündüz and E. Erkip, "Reliable cooperative source transmission with side information," in *Proc. IEEE Information Theory Workshop*, Bergen, Norway, Jul. 2007.
[9] G. Kramer and S. Shamai, "Capacity for classes of broadcast channels with receiver side information," in *Proc. IEEE Information Theory Workshop*, Lake Tahoe, CA, Sep. 2007.
[10] Y. Wu, "Broadcasting when receivers know some messages a priori," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Nice, France, Jun. 2007.
[11] L. L. Xie, "Network coding and random binning for multi-user channels," in *Proc. Canadian Workshop on Information Theory*, Edmonton, AB, Canada, Jun. 2007.
[12] F. Xue and S. Sandhu, "PHY-layer network coding for broadcast channel with side information," in *Proc. IEEE Information Theory Workshop*, Lake Tahoe, CA, Sep. 2007.
[13] W. Kang and G. Kramer, "Broadcast channel with degraded source random variables and receiver side information," in *Proc. IEEE Int. Symp. Information Theory*, Toronto, ON, Canada, Jul. 2008.
[14] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.
[15] D. Gündüz and E. Erkip, "Lossless transmission of correlated sources over a multiple access channel with side information," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 2007.
[16] H. Chong, M. Motani, H. Garg, and H. E. Gamal, "On the Han-Kobayashi region for the interference channel," *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 3188–3195, Jul. 2008.
[17] N. Liu and A. Goldsmith, "Capacity regions and bounds for a class of Z-interference channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 4986–4994, Nov. 2009.
[18] S. Vishwanath, N. Jindal, and A. Goldsmith, "The Z channel," in *Proc. IEEE Globecom*, San Francisco, CA, Dec. 2003.
[19] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Information Theory*, Tsahkadsor, Armenian S.S.R, 1971, pp. 23–52.
[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.
[21] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless System*. New York: Academic, 1981.
[22] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

Nan Liu (S'04–M'07) received the B.Eng. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2001 and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park, in 2007.

From 2007 to 2008, she was a Postdoctoral Scholar in the Wireless Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA. In 2009, she became a Faculty Member in the School of Information Science and Engineering, Southeast University, Nanjing, China. Her research interests are in network information theory for wireless networks and wireless communication theory.

Deniz Gündüz (S'03–M'08) received the B.S. degree in electrical and electronics engineering from the Middle East Technical University in 2002 and the M.S. and Ph.D. degrees in electrical engineering from Polytechnic Institute of New York University, Brooklyn, in 2004 and 2007, respectively.

He is currently a Research Associate at CTTC, Barcelona, Spain. He also holds a Visiting Research Collaborator position at Princeton University, Princeton, NJ. Previously he was a Consulting Assistant Professor at the Department of Electrical Engineering, Stanford University, Stanford, CA, and a Postdoctoral Research Associate at the Department of Electrical Engineering, Princeton University. In 2004, he was a summer Researcher in the Laboratory of Information Theory (LTHI), EPFL, Lausanne, Switzerland. His research interests lie in the areas of communication theory and information theory with special emphasis on joint source-channel coding, cooperative communications, and cross-layer design.

Dr. Gündüz is the recipient of a Marie Curie Reintegration Grant, funded by the European Union's Seventh Framework Programme (FP7), the 2008 Alexander Hessel Award of the Polytechnic Institute of New York University, given to the best Ph.D. dissertation, and a recipient of the Best Student Paper Award at the 2007 IEEE International Symposium on Information Theory.

Andrea J. Goldsmith (S'90–M'95–SM'99–F'05) received the Ph.D. degree in electrical engineering from the University of California, Berkeley, CA, in 1994.

She is currently a Professor of electrical engineering at Stanford University, Stanford, CA. Previously, she was an Assistant Professor of electrical engineering at Caltech. While on leave from Stanford, she founded Quantenna Communications, Inc. and served as its CTO, and she has held other industry positions at Maxim Technologies, Memorylink Corporation, and AT&T Bell Laboratories. She has authored the book *Wireless Communications* (Cambridge University Press, 2005) and coauthored the book *MIMO Wireless Communications* (Cambridge University Press, 2007) . Her research includes work on wireless information and communication theory, MIMO systems, and multihop networks, cognitive radios, sensor networks, "green" wireless system design, and applications of communications and signal processing to neuroscience.

Dr. Goldsmith is a Fellow of Stanford and has received several awards for her work, including the IEEE Communications Society and Information Theory Society joint paper award, the National Academy of Engineering Gilbreth Lecture Award, the IEEE Wireless Communications Technical Committee Recognition Award, the Alfred P. Sloan Fellowship, and the Silicon Valley/San Jose Business Journal's Women of Influence Award. She currently serves on the Steering Committee for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and as Editor for the *Journal on Foundations and Trends in Communications and Information Theory* and *Networks*. She was previously an Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, the IEEE TRANSACTIONS ON COMMUNICATIONS, and for the IEEE WIRELESS COMMUNICATIONS MAGAZINE. She has served on the Board of Governors and as a Distinguished Lecturer for both the IEEE Information Theory and Communications Societies. She also served as President of the Information Theory Society and founded its Student Committee. At Stanford, she received the inaugural University Postdoc Mentoring Award and served as Chair of its Faculty Senate.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. His research interests are in the areas of stochastic analysis, statistical signal processing, and information theory, and their applications in wireless networks and related fields. Among his publications in these areas are *Quickest Detection* (Cambridge University Press, 2009), coauthored with O. Hadjiliadis, and *Information Theoretic Security* (Now Publishers, 2009), coauthored with Y. Liang and S. Shamai.

Dr. Poor is a member of the National Academy of Engineering, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U.K.). He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society in 2004–2007 and as the Editor-in-Chief, and recently as General Co-Chair of the 2009 IEEE International Symposium on Information Theory, Seoul, Korea. He is the recipient of the 2005 IEEE Education Medal, the 2008 Aaron D. Wyner Distinguished Service Award of the IEEE Information Theory Society, the 2009 Edwin Howard Armstrong Achievement Award of the IEEE Communications Society, and the 2011 IEEE Eric E. Sumner Award.