Parallel Gaussian Networks with a Common State-Cognitive Helper

Ruchen Duan, *Member, IEEE*, Yingbin Liang, *Member, IEEE*, Ashish Khisti, *Member, IEEE* and Shlomo Shamai (Shitz), *Fellow, IEEE*

***Abstract*—State-dependent parallel networks with a common state-cognitive helper is studied, in which K transmitters wish to send K messages to their corresponding receivers over K state-corrupted parallel channels, and a helper who knows the state information noncausally wishes to assist these receivers to cancel state interference. Furthermore, the helper also has its own message to be sent simultaneously to its corresponding receiver. Since the state information is known only to the helper, but not to other transmitters, transmitter-side state cognition and receiver-side state interference are mismatched. Our focus is on the high state power regime, i.e., the state power goes to infinity. Three (sub)models are studied. Model I serves as a basic model, which consists of only one transmitter-receiver (with state corruption) pair in addition to a helper that assists the receiver to cancel state in addition to transmitting its own message. Model II consists of two transmitter-receiver pairs in addition to a helper, and only one receiver is interfered by a state sequence. Model III generalizes model I to include multiple transmitter-receiver pairs with each receiver corrupted by independent state. For all models, inner and outer bounds on the capacity region are derived, and comparison of the two bounds yields characterization of either full or partial boundary of the capacity region under various channel parameters.**

***Index Terms*—Capacity region, channel state, parallel channel, helper, dirty paper coding, Gel'fand-Pinsker scheme, noncausal state information.**

I. INTRODUCTION

State-dependent network models, in which receivers are interfered by random state sequences, have recently been intensively studied. In many such models studied before, some or all of the transmitters know the states that interfere their targeted receivers noncausally, and can hence exploit the state information in encoding of messages as in [1], [2]. In this way, the state interference can be efficiently or even fully cancelled at receivers. For example, the state-dependent broadcast channel has been studied in, e.g., [3]–[7], in which the transmitter knows the state noncausally and can exploit such information to select the codeword to be sent in the channel. In [8], the state-dependent relay channel is studied, in which the source node knows the state and can use such information for encoding. In [3], [9], the multiple access channel (MAC) with the receiver being corrupted by one state variable is studied. In such a model, both transmitters are assumed to know the state sequence noncausally, and can use the state information to independently encode their own messages. Similarly, in [10], the state-dependent cognitive MAC is studied, in which one transmitter knows both messages as well as the state, and can hence use state information to encode both messages. As a similar situation, the state-dependent interference channel is studied in [11]–[14], in which the state information is known at both transmitters and can hence be exploited for encoding their messages, respectively.

A common nature that the above models share is that for each message to be transmitted, at least one transmitter in the system knows both the message and the state, and can incorporate the state information in encoding of the message so that state interference at the corresponding receiver can be cancelled. However, in practice, it is often the case that transmitters that have messages intended for receivers do not know the state, whereas some third-party nodes know the state, but do not know the message. In such a mismatched case, state information cannot be exploited in encoding of messages, but can still serve to cancel the state interference at receivers. A number of previously studied models capture such mismatched property. For example, in [15], a transmitter sends a message to a state-dependent receiver, and a helper knows the state noncausally and can help the transmission. Lattice coding is designed in [15] for the helper to assist state cancelation at the receiver, and is shown to be optimal under certain channel conditions. In [16], [17], the state-dependent relay channel is studied, and the case with the state noncausally known only at the relay is the mismatched scenario. Furthermore, in [18], the state-dependent MAC channel is studied with the state known at only one transmitter. In such a case, the other transmitter's message cannot be encoded with the information of the state. In [19]–[21], the MAC is corrupted by two states that are respectively known at the two transmitters. In such a case, neither message can be encoded with the full information of the state.

In this paper, we focus on the mismatched scenarios, where

The material in this paper was presented in part at the IEEE Information Theory Workshop, Seville, Spain, September 2013, and in part at the IEEE International Symposium on Information Theory, Honolulu, HI, USA, June-July 2014.

The work of R. Duan and Y. Liang was supported by the National Science Foundation under Grant CCF-12-18451 and by the National Science Foundation CAREER Award under Grant CCF-10-26565. The work of A. Khisti was supported by the Canada Research Chair's Program. The work of S. Shamai(Shitz) was supported by the Israel Science Foundation (ISF), and the European Commission in the framework of the Network of Excellence in FP7 Wireless COMmunications NEWCOM++.

R. Duan is with Samsung Semiconductor Inc., San Diego, CA 92121 USA (email: r.duan@samsung.com). Y. Liang is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (email: yliang06@syr.edu). Ashish Khisti is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S3G4, Canada (email: akhisti@comm.utoronto.ca). Shlomo Shamai (Shitz) is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion city, Haifa 32000, Israel (email: sshlomo@ee.technion.ac.il).

the state is known at only a helper which does not know messages. Furthermore, we are interested in the following issues that are not captured in the previously studied models: (1) when there are multiple state-dependent transmitter-receiver links, how should the helper trade off among helping multiple state-interfered receivers; (2) when the helper has its own message intended for a separate receiver (not state-dependent), how should the helper trade off between sending its own message and assisting state-dependent receivers; and (3) under what channel conditions, the above two tradeoffs are optimal (i.e., achieve the boundary of the capacity region).

More specifically, we study a class of state-dependent parallel networks with a common state-cognitive helper (see Fig. 2). In our model, K transmitters wish to send K messages respectively to K receivers over K parallel channels, and the receivers are corrupted by states. The channel state is known to neither the transmitters nor the receivers, but to a helper noncausally. The helper hence assists these transmitter-receiver pairs to cancel state interference. Furthermore, the helper also has its own message to be sent simultaneously to its corresponding receiver. Since the state information is known only to the helper, but not to the corresponding transmitters, transmitter-side state cognition and receiver-side state interference are mismatched. Our goal is to investigate such a mismatched scenario in high state power regime, i.e., as the power of the state sequences go to infinity. This is the most challenging scenario with state interference, because the state cannot be cancelled by simple reversion to achieve non-zero rate. More sophisticated schemes must be designed to cancel state interference. Such a model also suggests to exploit the state cognition for improving communication rates other than the traditional message cognition studied in the context of cognitive channels and networks.

A. Practical Implication of the Model

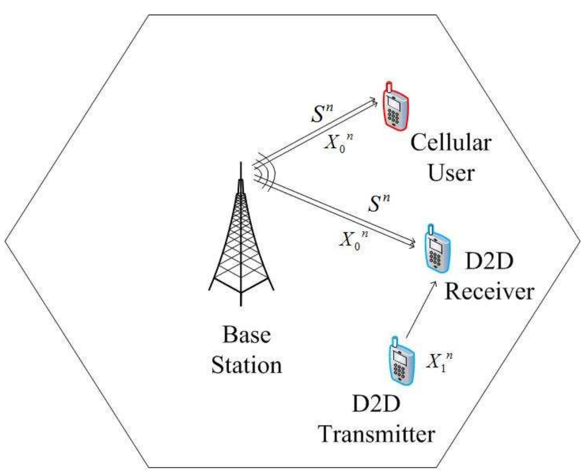


Fig. 1: A practical example for the parallel Gaussian networks with a common state-cognitive helper.

The model we study is well justified in practical wireless networks, and implies a new perspective of interference cancelation. We illustrate the idea via a simple example (see Fig. 1). Consider a cellular network that incorporates device-to-device (D2D) communications. It is typical that the cellular base station causes interference to D2D transmissions. In fact, the base station itself knows such interference noncausally, because the interference is the signal that the base station sends to cellular receivers. Thus, the interference can be viewed as the noncausal state sequence (denoted as S^n in Fig. 1). The base station is then able to exploit such interference (i.e., state) information and send a help signal (denoted by X_0^n in Fig. 1) to assist D2D users to cancel the interference. Although the help signal X_0^n may also cause interference to the cellular receiver, our results in this paper show that small power of X_0^n can cancel large interference S^n (even with infinite power). Thus, the network can still have substantial gain in throughputs. We further note that although the base station may employ Han-Kobayashi scheme [22] for interference cancelation, this requires that the base station share codebooks of cellular users with D2D receivers, and is not feasible in practice.

In this paper, we are interested in high state-power regime, which is well justified by the above example, because the base station typically transmit at a large power. In such a regime, the cellular user can easily decode the state S^n (that carries the base station's information to the cellular user) and can then subtract it from its received signal. Thus, S^n does not appear in the cellular user's output in our models described in Section II. Furthermore, the helper signal X_0^n by nature can not only cancel the state interference but also convey additional information to the cellular user. Hence, in our models, we allow X_0^n to carry information for the cellular user to capture the tradeoff between the above two roles that such a signal can play.

Such a simple scenario can be further extended. For example, there can be a number of D2D transceiver pairs that are interfered by cellular transmissions. Typically, D2D transceivers are not located together, and cellular interference to different D2D transceivers are independent. This is because the base station can transmit independent signals to cellular users located in different areas via directional antennas or via different relay nodes. In such cases, as we describe above, the base station (which knows the interference noncausally) can serve as a helper to assist all D2D users to cancel the interference. Our model described in Section II captures such a more general scenario.

B. Main Contributions

In this paper, we study three (sub)models of the state-dependent parallel networks with a common helper which characterize the main features of the motivated practical model. Model I serves as a basic model, which consists of only one state-corrupted receiver ($K = 1$) and a helper that assists this receiver to cancel state interference in addition to transmitting its own message. We dispense the state signal at the helper's corresponding receiver, because the state actually contains messages for this receiver and is hence decodable with large power. Instead, we require the helper to transmit its own message using the helper signal, which indicates the trade-off between transmitting message and assisting other interfered receivers. Same settings are also applied for model II and model III. Our study of this model provides necessary techniques to deal with state in the mismatched context for studying more complicated models II and III. In fact, this model can be viewed as the state-dependent Z-interference channel, in which the interference is only at receiver 1 caused by the helper. In contrast to the state-dependent Z-interference channel studied previously in [23], which assumes that state interference at both receivers are known to both (corresponding) transmitters, our model assumes that state interference is known noncausally only to the helper, not to the corresponding transmitter 1.

The challenge to design capacity-achieving schemes for model I lies in: (1) due to the mismatched structure, it is difficult for the helper to fully cancel the state at the receiver, and (2) the helper needs to resolve the tension between transmitting its own message and helping receiver 1 to cancel its interference. In this paper, we design an adapted dirty paper coding scheme for state cancelation in the mismatched context, in which correlation between the state variable and the state-cancelation variable is a design parameter, and can be chosen to optimize the rate region. This is in contrast to classical dirty paper coding [2], in which such a correlation parameter is fixed for fully canceling the state. We further design a layered coding scheme, in which the adapted dirty paper coding is superposed onto the helper's transmission of its own message. In particular, the design parameters in superposition enable the helper to judiciously trade off between transmitting its own message and helping to cancel the state.

Based on such a layered coding scheme, we derive achievable regions for both the discrete memoryless and Gaussian channels. We further derive an outer bound for the Gaussian channel in high state power regime. By comparing the inner and outer bounds, we characterize the boundary of the capacity region either fully or partially for all Gaussian channel parameters in high state power regime. Our result also implies that the capacity region is strictly inside the capacity region of the corresponding channel without state [24] due to infinite state power. This is in contrast to the results for Costa type of dirty paper channels, for which dirty paper coding achieves the capacity of the corresponding channels without state.

We then further study model II, which consists of two transmitter-receiver pairs in addition to the helper, and only one receiver is interfered by a state sequence. Here, the challenge lies in the fact that the helper inevitably causes interference to receiver 2 while assisting receiver 1 to cancel the state. For this model, we start with the scenario with the helper fully assisting the receivers without transmitting its own message. We first derive an outer bound on the capacity region. We then develop a two-layer dirty paper coding scheme with one layer helping receiver 1 to cancel state via dirty paper coding, and with the other layer of dirty paper coding canceling the interference caused by the helper in assisting receiver 1. By comparing inner and outer bounds, we characterize two segments of the capacity region boundary. One segment corresponds to the case, in which our scheme achieves the point-to-point channel capacity for receiver 2 and certain positive rate for receiver 1. This implies that the helper is able to assist receiver 1 without causing interference to receiver 2 effectively. The other segment corresponds to the case, in which our scheme achieves the best single-user rate for receiver 1 with assistance of the helper, while receiver 2 treats the helper's signal as noise. Such a scheme is guaranteed by our outer bound to be the best to achieve the sum capacity under certain channel parameters. We further extend these results to the scenario with the helper sending its own message in addition to assisting the two receivers.

We finally study model III, in which a common helper assists multiple transmitter-receiver pairs with each receiver corrupted by an independently distributed state sequence. We note that this model is more general than model I, but does not include model II as a special case. This is because model III has each receiver (excluding the helper) being corrupted by an infinitely powered state sequence, and hence never reduces to the model II, in which receiver 2 is not corrupted by a state sequence. This also leads to different technical challenges to characterize the capacity for model III due to the compound state interference. The same technical challenge is also reflected in the studies [25]–[27] of the state-dependent compound channel, for which the capacity is not known in general. As for model II, we also start with the scenario, in which the helper fully assists other users without sending its own message. We first derive a useful outer bound, which captures the sum rate limit due to the common helper. We then derive an inner bound based on a time-sharing scheme, in which the helper alternatively assists receivers. Somewhat interestingly, such a time-sharing scheme achieves the sum capacity under many channel parameters, although each individual transmitter may not be able to achieve its individual best rate. This is because these transmitters effectively have larger power during their transmissions in the time-sharing scheme so that the transmission rate matches the outer bound on the sum rate. We also characterize the full capacity region under certain channel parameters. We then extend our results to the general scenario with the helper also transmitting its own message.

The rest of the paper is organized as follows. In Section II, we describe the channel model. In Sections III, IV, and V, we present our results for models I, II, and III, respectively. Finally, in Section VII, we conclude the paper with a few remarks.

II. Channel Model

In this paper, we investigate the state-dependent parallel network with a common state-cognitive helper (see Figure 2), in which K transmitters wish to send K messages to their corresponding receivers over state-corrupted parallel channels, and a helper who knows the state information noncausally wishes to assist these receivers to cancel state interference. Furthermore, the helper also has its own message to be sent simultaneously to its corresponding receiver.

More specifically, each transmitter (say transmitter k) has an encoder $f_k : \mathcal{W}_k \to \mathcal{X}_k^n$, which maps a message $w_k \in \mathcal{W}_k$ to a codeword $x_k^n \in \mathcal{X}_k^n$ for $k = 1, \ldots, K$. The K inputs $x_1^n, \ldots, x_K^n$ are transmitted over K parallel channels,

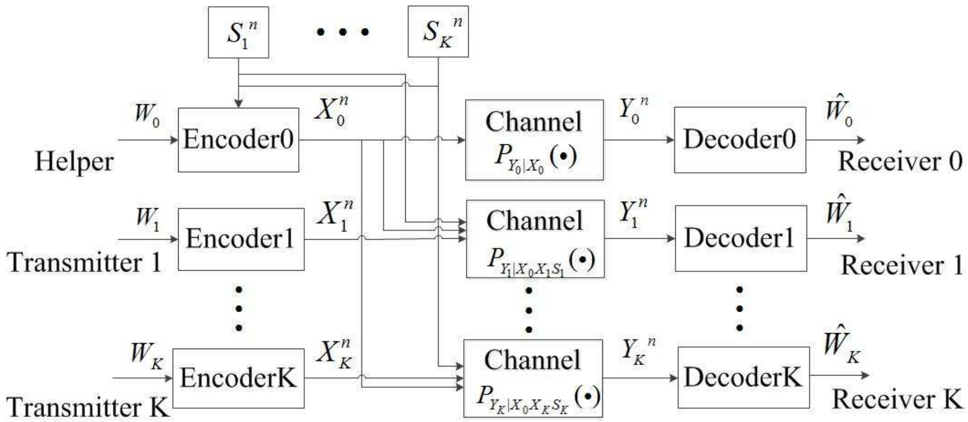


Fig. 2: Parallel Gaussian channel model with a common state-cognitive helper.

respectively. Each receiver (say receiver k) is interfered by an i.i.d. state sequence S_k^n for $k = 1, \dots, K$, which is known at none of transmitters $1, \dots, K$ and receivers $1, \dots, K$. A common helper is assumed to know all state sequences S_k^n for $k = 1, \dots, K$ noncausally. Thus, the encoder at the helper, $f_0 : \mathcal{W}_0 \times \{\mathcal{S}_1^n, \dots, \mathcal{S}_K^n\} \to \mathcal{X}_0^n$, maps a message $w_0 \in \mathcal{W}_0$ and the state sequences $(s_1^n, \dots, s_K^n) \in \mathcal{S}_1^n \times \dots \times \mathcal{S}_K^n$ to a codeword $x_0^n \in \mathcal{X}_0^n$. The entire channel transition probability is given by $P_{Y_0|X_0} \prod_{k=1}^K P_{Y_k|X_0,X_k,S_k}$. There are $K+1$ decoders with each at one receiver, $g_k : \mathcal{Y}_k^n \to \mathcal{W}_k$, maps a received sequence y_k^n into a message $\hat{w}_k \in \mathcal{W}_k$ for $k = 0, 1, \dots, K$.

Remark 1. *Without state interference, our model becomes the $K+1$-user Z-interference channel, in which the signal of the helper interferences all remaining K receivers.*

The average probability of error for a length-n code is defined as

$$P_e^{(n)} = \frac{1}{|\mathcal{W}_0||\mathcal{W}_1|\dots|\mathcal{W}_K|} \sum_{w_0=1}^{|\mathcal{W}_0|} \sum_{w_1=1}^{|\mathcal{W}_1|} \cdots \sum_{w_k=1}^{|\mathcal{W}_k|} Pr\{(\hat{w}_0, \hat{w}_1, \dots, \hat{w}_K) \neq (w_0, w_1, \dots, w_K)\}. \quad (1)$$

A rate tuple $(R_0, R_1, \dots, R_K)$ is *achievable* if there exists a sequence of message sets $\mathcal{W}_k^{(n)}$ with $|\mathcal{W}_k^{(n)}| = 2^{nR_k}$ for $k = 0, 1, \dots, K$, and encoder-decoder tuples $(f_0^{(n)}, f_1^{(n)}, \dots, f_K^{(n)}, g_0^{(n)}, g_1^{(n)}, \dots, g_K^{(n)})$ such that the average error probability $P_e^{(n)} \to 0$ as $n \to \infty$. The *capacity region* is defined to be the closure of the set consisting of all achievable rate tuples $(R_0, R_1, \dots, R_K)$.

In this paper, we study the following three Gaussian channel models.

In model I, $K = 1$, i.e., the helper assists one transmitter-receiver pair. The channel outputs at receiver 0 and 1 for one symbol time are given by

$$Y_0 = X_0 + N_0, \quad (2a)$$
$$Y_1 = X_0 + X_1 + S_1 + N_1. \quad (2b)$$

In model II, $K = 2$, in which one helper assists two transmitter-receiver pairs, and only one receiver is interfered by a state sequence. The channel outputs at receivers 0, 1 and 2 for one symbol time are given by

$$Y_0 = X_0 + N_0, \quad (3a)$$
$$Y_1 = X_0 + X_1 + S_1 + N_1, \quad (3b)$$
$$Y_2 = X_0 + X_2 + N_2. \quad (3c)$$

In model III, K is general, in which a common helper assists multiple transmitter-receiver pairs with each receiver corrupted by an independently distributed state sequence. This model is more general than model I, but does not include model II as a special case (due to infinite state power). The channel outputs at receivers 0 and receivers $1, \dots, K$ for one symbol time are given by

$$Y_0 = X_0 + N_0, \quad (4a)$$
$$Y_k = X_0 + X_k + S_k + N_k, \quad for \quad k = 1, \dots, K \quad (4b)$$

In the above three models, the noise variables $N_0, N_1 \dots, N_K$ and the state variables $S_1, \dots, S_K$ are Gaussian distributed with distributions $N_0, \dots, N_K \sim \mathcal{N}(0, 1)$ and $S_k \sim \mathcal{N}(0, Q_k)$ for $k = 1, \dots, K$, and all of the variables are independent and are i.i.d. over channel uses. The channel inputs $X_0, X_1, \dots, X_K$ are subject to the average power constraints $\frac{1}{n}\sum_{i=1}^n X_{ki}^2 \leqslant P_k$ for $k = 0, 1, \dots, K$.

We are interested in the regime of high state power, i.e., as $Q_k \to \infty$ for $k = 1, \dots, K$. Our goal is to design helper strategies in order to cancel the high power state interference and to further characterize the capacity region in this regime.

III. Model I: $K = 1$

Model I with $K = 1$ is a basic model, in which the helper assists one transmitter-receiver pair. Understanding this model will help the study of the general parallel network. In this section, we first develop outer and inner bounds on the capacity region, and then characterize the boundary of the capacity region based on these bounds.

A. Outer Bound

In this subsection, we provide an outer bound on the capacity region in high state power regime.

Proposition 1. *For the Gaussian channel of model I, an outer bound on the capacity region for the regime when $Q_1 \to \infty$ consists of rate pairs (R_0, R_1) satisfying:*

$$R_1 \leqslant \frac{1}{2}\log(1+P_1) \tag{5a}$$

$$R_0 + R_1 \leqslant \frac{1}{2}\log(1+P_0). \tag{5b}$$

The bound (5a) on R_1 follows simply from the capacity of the point-to-point channel between transmitter 1 and receiver 1 without signal and state interference. The bound (5b) on the sum rate is limited only by the power P_0 of the helper, and does not depend on the power P_1 of transmitter 1. Intuitively, this is because P_0 is split for transmission of W_0 and for helping transmission of W_1 by removing state interference, and hence P_0 determines a trade-off between R_0 and R_1. On the other hand, improving the power P_1, although may improve R_1, can also cause more interference for receiver 1 to decode the auxiliary variable for canceling state and interference. Thus, the balance of the two effects determines that P_1 does not affect the sum rate.

Proof. The proof is detailed in Appendix A. □

We further note that although the sum-rate upper bound (5b) can be achieved easily by keeping transmitter 1 silent (i.e., R_0 achieves the sum rate bound with $R_1 = 0$), we are interested in characterizing the capacity region (i.e., the trade-off between R_0 and R_1) rather than a single point that achieves the sum-rate capacity. In the next section, we characterize such optimal trade-off based on the sum-rate bound.

Remark 2. *The outer bound in Proposition 1 is strictly inside an achievable rate region of the corresponding channel without state interference (i.e., the Z-interference channel) [24]. This implies that the capacity region of our model is strictly inside that of the corresponding channel without state. This suggests that state interference does cause performance degradation for systems with mismatched state cognition and interference in high state power regime. This is in contrast to the results on Costa-type dirty paper channels [2], for which dirty paper coding achieves the capacity of the corresponding channels without state.*

B. Inner Bound

The major challenge in designing an achievable scheme arises from the mismatched property due to transmitter-side state cognition and receiver-side state interference, i.e., state interference to receiver 1 is known noncausally only to the helper, not to the corresponding transmitter 1. Since we study the regime with large state power, transmitter 1 can send information to receiver 1 only if the helper assists to cancel the state. Thus, the helper needs to resolve the tension between transmitting its own message to receiver 0 and helping receiver 1 to cancel its interference. A simple scheme of time-sharing between the two transmitters in general is not optimal.

We design a layered coding scheme as follows. The helper splits its signal into two parts in a layered fashion: one (represented by X_0' in Proposition 2) for transmitting its own message and the other (represented by U in Proposition 2) for helping receiver 1 to remove both state and signal interference. In particular, the second part of the scheme applies a single-bin dirty paper coding scheme, in which transmission of W_1 and treatment of state interference for decoding W_1 are performed separately by transmitter 1 and the helper. In traditional (multi-bin) dirty paper coding as in [2], the bin number carries the information of the message, and the index within each bin carries the information about the state. In our model, the helper knows only the state, not the message (of transmitter 1), and hence can only encode the state into the index within a single bin. For the state-dependent Gaussian channel, the single-bin dirty paper coding can be understood as quantization of the state interference. Such a scheme was also used by other studies, e.g., [28, Theorem 1]. Based on such a scheme, we obtain the following achievable rate region for the discrete memoryless channel, which is useful for deriving an inner bound for the Gaussian channel.

Proposition 2. *For the discrete memoryless channel of model I, an inner bound on the capacity region consists of rate pairs (R_0, R_1) satisfying:*

$$R_0 \leqslant I(X_0'; Y_0) \tag{6a}$$

$$R_1 \leqslant I(X_1; Y_1|U) \tag{6b}$$

$$R_1 \leqslant I(X_1 U; Y_1) - I(U; S_1 X_0') \tag{6c}$$

for some distribution $P_{S_1} P_{X_0'} P_{U|S_1 X_0'} P_{X_0|U S_1 X_0'} P_{X_1} P_{Y_0|X_0} P_{Y_1|S_1 X_0 X_1}$.

Proof. The proof is detailed in Appendix B. □

Based on Proposition 2, we have the following simpler inner bound.

Corollary 1. *For the discrete memoryless channel of model I, an inner bound on the capacity region consists of rate pairs (R_0, R_1) satisfying:*

$$R_0 \leqslant I(X_0'; Y_0) \tag{7a}$$

$$R_1 \leqslant I(X_1; Y_1|U) \tag{7b}$$

for some distribution $P_{S_1} P_{X_0'} P_{U|S_1 X_0'} P_{X_0|U S_1 X_0'} P_{X_1} P_{Y_0|X_0} P_{Y_1|S_1 X_0 X_1}$ *that satisfies*

$$I(U; Y_1) \geqslant I(U; S_1 X_0'). \tag{8}$$

Proof. The region follows from Proposition 2 because (6c) is redundant due to the condition (8). □

The inner bound in Corollary 1 corresponds to an intuitive achievable scheme based on successive cancelation. Namely, the condition guarantees that receiver 1 decodes the auxiliary random variable U first, and then removes it from its output and decodes the message, which results in the bound (7b). In particular, cancelation of U leads to cancelation of signal and state interference at receiver 1.

We next derive an inner bound for the Gaussian channel of model I based on Corollary 1.

Proposition 3. *For the Gaussian channel of model I, in the regime when $Q_1 \to \infty$, an inner bound on the capacity region consists of rate pairs (R_0, R_1) satisfying:*

$$R_0 \leqslant \frac{1}{2}\log\left(1 + \frac{\bar{\beta}P_0}{\beta P_0 + 1}\right) \tag{9a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1 + \frac{P_1}{1 + (1 - \frac{1}{\alpha})^2 \beta P_0}\right) \tag{9b}$$

for some real constants $\alpha > 0$ and $0 \leqslant \beta \leqslant 1$ that satisfy $\alpha \leqslant \frac{2\beta P_0}{\beta P_0 + P_1 + 1}$.

Proof. Proposition 3 follows from Corollary 1 by choosing the joint Gaussian distribution for random variables as follows:

$$U = X_0'' + \alpha(S_1 + X_0'), \; X_0 = X_0' + X_0''$$
$$X_0' \sim \mathcal{N}(0, \bar{\beta}P_0), \; X_0'' \sim \mathcal{N}(0, \beta P_0)$$
$$X_1 \sim \mathcal{N}(0, P_1)$$

where X_0', X_0'', X_1 and S_1 are independent, $\alpha > 0$, $0 \leqslant \beta \leqslant 1$, and $\bar{\beta} = 1 - \beta$. □

We note that in Proposition 3, the parameter α captures correlation between the state variable S_1 and the auxiliary variable U for dealing with the state, and can be chosen to optimize the rate region. This is in contrast to the classical dirty paper coding [2], in which such correlation parameter is fixed for state cancelation. Therefore, although Corollary 1 may provide a smaller inner bound than that given in Proposition 2, it can be shown that two inner bounds are equivalent for our chosen auxiliary random variables and input distribution after optimizing over α.

C. Capacity Region

In this section, we characterize the boundary points of the capacity region for the Gaussian channel of model I based on the inner and outer bounds given in Propositions 3 and 1, respectively. We divide the Gaussian channel into three cases based on the conditions on the power constraints: (1) $P_1 \geqslant P_0 + 1$; (2) $P_0 - 1 \leqslant P_1 < P_0 + 1$ and (3) $0 \leqslant P_1 < P_0 - 1$. For each case, we optimize the dirty paper coding parameter α in the inner bound in Proposition 3 to find achievable rate points that lie on the sum-rate upper bound (5b) in order to characterize the boundary points of the capacity region.

Case 1: $P_1 \geqslant P_0 + 1$. The capacity region is fully characterized in the following theorem.

Theorem 1. *For the Gaussian channel of model I, in the regime when $Q_1 \to \infty$, if $P_1 \geqslant P_0 + 1$, the capacity region consists of the rate pairs (R_0, R_1) satisfying*

$$R_0 + R_1 \leqslant \frac{1}{2}\log(1 + P_0). \tag{10}$$

Proof. Let $\tilde{P}_1$ be the actual power for transmitting W_1. Then the inner bound (9b) on R_1 is optimized when $\alpha = \frac{2\beta P_0}{\beta P_0 + \tilde{P}_1 + 1}$. By setting $\tilde{P}_1 = \beta P_0 + 1$, the inner bound given in Proposition 3 matches the outer bound given in Proposition 1, and hence is the capacity region. □

The capacity region of case 1 is illustrated in Fig. 3.

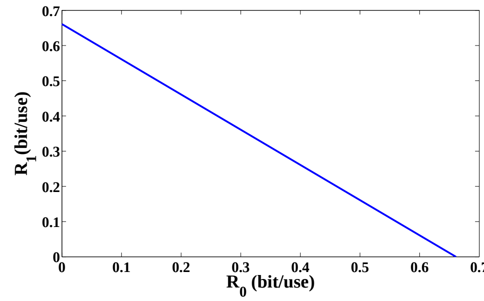


Fig. 3: The capacity region for case 1 with $P_0 = 1.5$ and $P_1 = 3$.

Theorem 1 implies that when P_1 is large enough, the power of the helper limits the system performance. Furthermore, since P_1 for transmission of W_1 causes interference to receiver 1 to decode the auxiliary variable for canceling state and interference, beyond a certain value, increasing P_1 does not improve the rate region any more. Theorem 1 also suggests that in order to achieve different points on the boundary of the capacity region (captured by the parameter β), different amounts of power $\tilde{P}_1$ should be applied.

Case 2: $P_0 - 1 \leqslant P_1 < P_0 + 1$. We summarize the capacity result in the following theorem.

Theorem 2. *Consider the Gaussian channel of model I in the regime when $Q_1 \to \infty$, and $P_0 - 1 \leqslant P_1 < P_0 + 1$. If $P_1 \geqslant 1$, the rate points (R_0, R_1) on the line A-B (see Fig. 4 (a) and Fig. 5 (a)) are on the capacity region boundary. More specifically, the points A and B are characterized as:*

$$\textit{Point } A: \left(\frac{1}{2}\log(1 + P_0), 0\right)$$
$$\textit{Point } B: \left(\frac{1}{2}\log(1 + \frac{P_0 - P_1 + 1}{P_1}), \frac{1}{2}\log P_1\right)$$

If $P_1 < 1$ the rate point A (see Fig. 4 (b) and Fig. 5 (b)) is on the capacity region boundary, and is characterized as:

$$\textit{Point } A: \left(\frac{1}{2}\log(1 + P_0), 0\right)$$

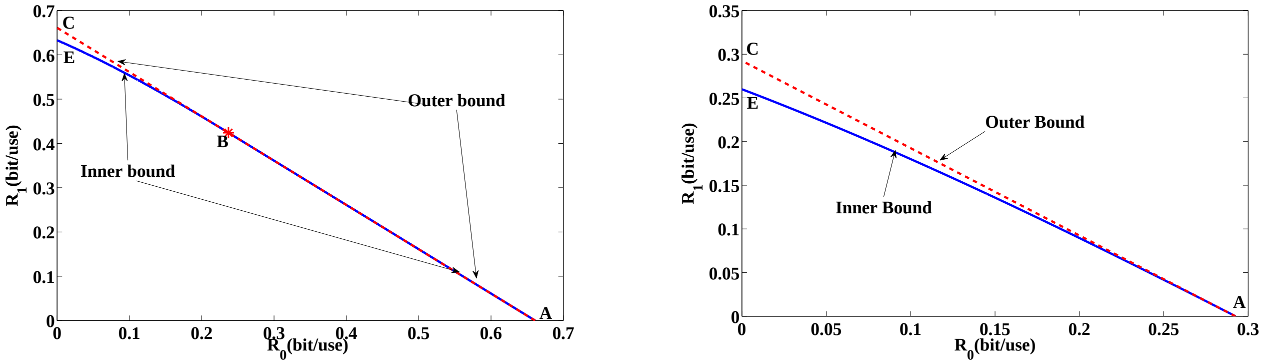
Proof. We first set $\alpha = \frac{2\beta P_0}{\beta P_0 + P_1 + 1}$, and then substitute α into (9b) and obtain the following inner bound:

$$R_0 \leqslant \frac{1}{2}\log\left(1 + \frac{\bar{\beta}P_0}{\beta P_0 + 1}\right) \tag{13a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1 + \frac{4\beta P_0 P_1}{4\beta P_0 + (P_1 + 1 - \beta P_0)^2}\right). \tag{13b}$$

When $P_1 \geqslant 1$, by setting $\beta = \frac{P_1 - 1}{P_0}$, we obtain an achievable rate point B given by $\left(\frac{1}{2}\log(1 + \frac{P_0 - P_1 + 1}{P_1}), \frac{1}{2}\log P_1\right)$, which is also on the outer bound. It is also clear that the point A given by $\left(\frac{1}{2}\log(1 + P_0), 0\right)$ is achievable by setting $\beta = 0$, which is also on the outer bound. Thus, the line $A - B$ is on the boundary of the capacity region due to time sharing. □

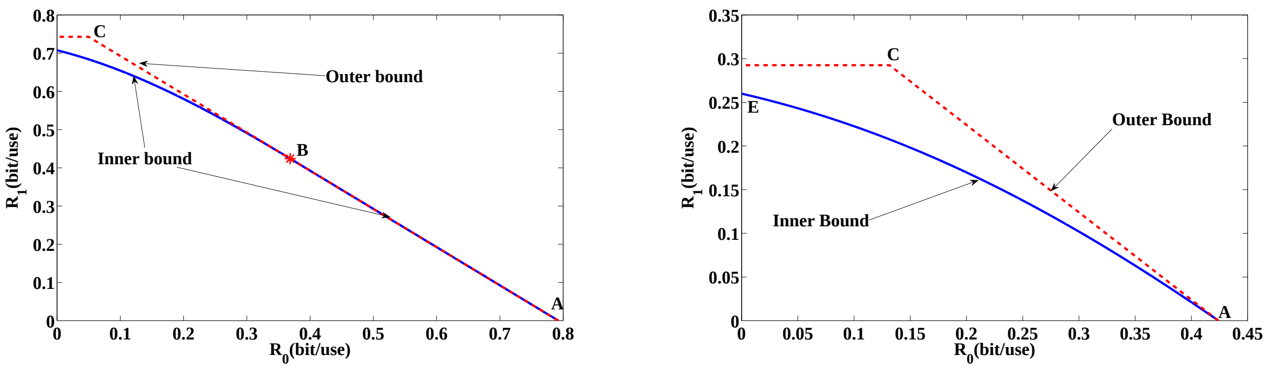
For this case, if $P_1 \geqslant 1$, i.e., P_1 is larger than the noise power, inner and outer bounds match over the line A-B as



(a) $P_1 \geqslant 1$ with $P_0 = 1.5$ and $P_1 = 1.8$

(b) $P_1 < 1$ with $P_0 = 0.5$ and $P_1 = 0.8$

Fig. 4: Inner and outer bounds for case 2 with $P_0 < P_1$, which match partially on the boundaries



(a) $P_1 \geqslant 1$ with $P_0 = 2$ and $P_1 = 1.8$

(b) $P_1 < 1$ with $P_0 = 0.8$ and $P_1 = 0.5$

Fig. 5: Inner and outer bounds for case 2 with $P_0 \geqslant P_1$, which match partially on the boundaries

illustrated in Fig. 4 (a) and Fig. 5 (a), and thus optimal trade-off between R_0 and R_1 is achieved over the points on the line A-B. If $P_1 < 1$, the inner and outer bounds match only at the rate point A as illustrated in Fig. 4 (b) and Fig. 5 (b), which achieves the sum-rate capacity. We further note that Fig. 4 is different from Fig. 5 in the outer bound. Fig. 5 corresponds to the case with $P_0 \geq P_1$, and hence the capacity region is also upper bounded by the point-to-point capacity of R_1. Such a bound is redundant in Fig. 4 which corresponds to the case with $P_0 < P_1$, because P_0 is not large enough to perfectly cancel state and signal interference at receiver 1. However, in case 3, we show that this point-to-point capacity of R_1 is achievable simultaneously with a certain positive R_0.

Case 3: $0 \leqslant P_1 < P_0 - 1$. We first summarize the capacity results in the following theorem.

Theorem 3. *Consider the Gaussian channel of model I in the regime when $Q_1 \to \infty$, and $P_1 < P_0 - 1$. If $P_1 \geqslant 1$, the rate points (R_0, R_1) on the line A-B (see Fig. 6 (a)) are on the boundary of the capacity region. More specifically, the points A and B are characterized as:*

$$Point\ A: \left(\frac{1}{2}\log(1+P_0), 0\right)$$
$$Point\ B: \left(\frac{1}{2}\log(1+\frac{P_0-P_1+1}{P_1}), \frac{1}{2}\log P_1\right)$$

And the rate points (R_0, R_1) on the line D-E (see Fig. 6 (a)) are on the boundary of the capacity region. The points D and E are characterized as:

$$Point\ D: \left(\frac{1}{2}\log(\frac{P_0+1}{P_1+2}), \frac{1}{2}\log(1+P_1)\right)$$
$$Point\ E: \left(0, \frac{1}{2}\log(1+P_1)\right)$$

If $P_1 < 1$, then point A (see Fig. 6 (b)) is on the capacity region boundary. The point A is characterized as:

$$Point\ A: \left(\frac{1}{2}\log(1+P_0), 0\right)$$

And the rate points (R_0, R_1) on the line D-E (see Fig. 6 (b)) are on the boundary of the capacity region. The points D and

E are characterized as:

$$\begin{aligned}&Point\ D: \left(\frac{1}{2}\log(\frac{P_0+1}{P_1+2}), \frac{1}{2}\log(1+P_1)\right)\\&Point\ E: \left(0, \frac{1}{2}\log(1+P_1)\right)\end{aligned}$$

Proof. For case 3, the inner bound boundary given in Proposition 3 is characterized by segment I consisting of rate points satisfying:

$$R_0 \leqslant \frac{1}{2}\log\left(1+\frac{\bar{\beta}P_0}{1+\beta P_0}\right) \tag{18a}$$

$$R_1 \leqslant \frac{1}{2}\log(1+\beta P_0) \tag{18b}$$

for $0 \leqslant \beta \leqslant \frac{P_1+1}{P_0}$; and segment II consisting of rate points satisfying

$$R_0 \leqslant \frac{1}{2}\log\left(1+\frac{\bar{\beta}P_0}{1+\beta P_0}\right) \tag{19a}$$

$$R_1 \leqslant \frac{1}{2}\log(1+P_1) \tag{19b}$$

for $\frac{P_1+1}{P_0} \leqslant \beta \leqslant 1$. Segment I is obtained by setting $\alpha = \frac{2\beta P_0}{\beta P_0+P_1+1}$, and segment II is obtained by setting $\alpha = 1$.

For segment I, if $P_1 \geqslant 1$, the line A-B is on the boundary of the capacity region as shown in Fig. 6 (a). If $P_1 < 1$, only point A is on the capacity boundary as shown in Fig. 6 (b). For segment II, it is clear that the point-to-point channel capacity for R_1 is achievable. Furthermore, by setting $\beta = \frac{P_1+1}{P_0}$, the point D is achievable. Thus, the line $D-E$ as shown in Fig. 6 (a) and (b) is on the boundary of the capacity region. □

Similarly to case 2, the inner and outer bounds match partially over the sum rate bound, i.e., the two bounds match over the line A-B (see Fig. 6 (a)) if $P_1 \geqslant 1$ and match at only the point A (see Fig. 6 (b)) if $P_1 < 1$. However, differently from case 2, the inner and outer bounds also match when $R_1 = \frac{1}{2}\log(1+P_1)$ over the line D-E (see Fig. 6 (a) and (b)). This is because the power P_0 of the helper in this case is large enough to fully cancel state and signal interference so that transmitter 1 is able to reach its maximum point-to-point rate to receiver 1 without interference. Furthermore, the helper is also able to simultaneously transmit its own message at a certain positive rate.

IV. Model II: $K = 2$

In this section, we consider the Gaussian channel of model II with $K = 2$, in which only receiver 1 is interfered by an infinite power state. We first study the scenario, in which the helper devotes to help two users without transmitting its own message, i.e., $W_0 = \phi$. We then extend the result to the more general scenario, in which the helper also has its own message destined for the corresponding receiver in addition to helping the two users, i.e., $W_0 \neq \phi$.

A. Scenario with Dedicated Helper ($W_0 = \phi$)

In this subsection, we study the scenario in which only receiver 1 is corrupted by a state sequence, and the helper (without transmission of its own messages) fully assists to cancel such state interference. Here, the challenge lies in the fact that the helper needs to assist receiver 1 to remove the state interference, but such signal inevitably causes interference to receiver 2. In this subsection, we first derive an outer bound, and then derive an inner bound based on the helper using a layered coding scheme with one layer assisting the state-interfered receiver, and the other layer canceling the interference in order to address the challenge mentioned above. We then characterize segments of the capacity boundary and the sum capacity under certain channel parameters.

We first derive a useful outer bound for Model II.

Proposition 4. *For the Gaussian channel of model II with $W_0 = \phi$, in the regime when $Q_1 \to \infty$, an outer bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leqslant \min\left\{\frac{1}{2}\log(1+P_0), \frac{1}{2}\log(1+P_1)\right\} \tag{20a}$$

$$R_2 \leqslant \frac{1}{2}\log(1+P_2) \tag{20b}$$

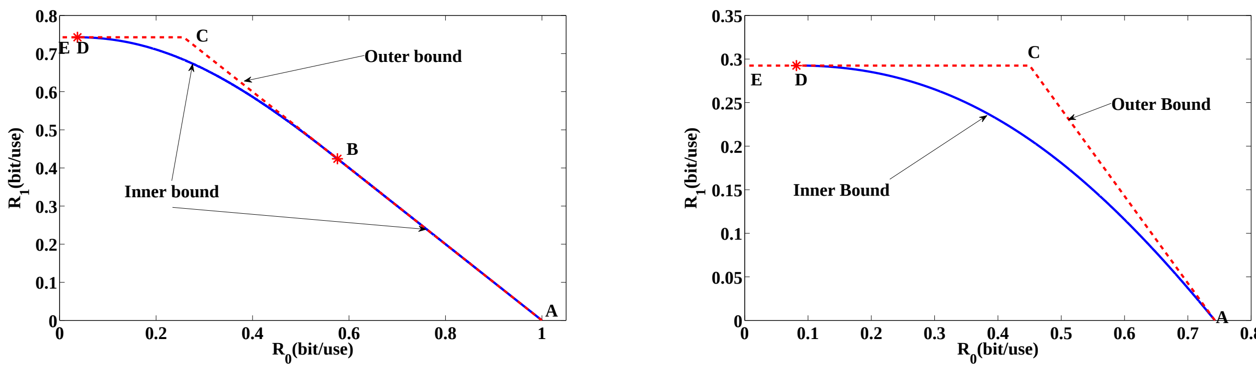
$$R_1 + R_2 \leqslant \frac{1}{2}\log(1+P_0+P_2). \tag{20c}$$

Proof. The proof is detailed in Appendix C. □

We note that (20a) represents the best single-user rate of receiver 1 with the helper dedicated to help it as shown in Proposition 1, (20b) is the point-to-point capacity for receiver 2, and (20c) implies that although the two transmitters communicate over parallel channels to their corresponding receivers, due to the shared common helper, the sum rate is still subject to a certain rate limit.

We next describe our idea to design an achievable scheme. We first note that although receiver 2 is not interfered by the state, the signal that the helper sends to assist receiver 1 to deal with the state still causes unavoidable interference to receiver 2. A natural idea to optimize the transmission rate to receiver 2 is simply to keep the helper silent. In this case, without the helper's assistance, receiver 1 gets zero rate due to infinite state power. Here, we design a novel scheme, which enables the point-to-point channel capacity for receiver 2 and a certain positive rate for receiver 1 simultaneously. Consequently, the helper is able to assist receiver 1 without causing interference to receiver 2. In our achievable scheme, the signal of the helper is split into two parts, represented by U and V as in Proposition 5. Here, U is designed to help receiver 1 to cancel the state while treating V as noise, and V is designed to help receiver 2 to cancel the interference caused by U. Since there is no state interference at receiver 2, U is decoded only at receiver 1. Based on such an achievable scheme, we obtain the following achievable region.

Proposition 5. *For the discrete memoryless channel of model II with $W_0 = \phi$, an achievable region consists of the rate pair*



(a) $P_1 \geqslant 1$ with $P_0 = 3$ and $P_1 = 1.8$

(b) $P_1 < 1$ with $P_0 = 1.5$ and $P_1 = 0.5$

Fig. 6: Inner and outer bounds for case 3, which match partially on the boundaries.

(R_1, R_2) satisfying

$$R_1 \leqslant I(X_1; Y_1 U) \tag{21a}$$
$$R_1 \leqslant I(X_1 U; Y_1) - I(U; S_1) \tag{21b}$$
$$R_2 \leqslant I(X_2; Y_2 V) \tag{21c}$$
$$R_2 \leqslant I(X_2 V; Y_2) - I(V; U S_1) \tag{21d}$$

for some distribution $P_{S_1 U V X_0 X_1 X_2} = P_{S_1} P_{UVX_0|S_1} P_{X_1} P_{X_2}$, where U and V are auxiliary random variables.

Proof. The proof is detailed in Appendix D. □

A straight-forward but more convenient subregion of the above inner bound is as follows.

Corollary 2. *For the discrete memoryless channel of model II with $W_0 = \phi$, an achievable region consists of the rate pair (R_1, R_2) satisfying*

$$R_1 \leqslant I(X_1; Y_1|U), \tag{22a}$$
$$R_2 \leqslant I(X_2; Y_2|V), \tag{22b}$$

for some distribution $P_{S_1 U V X_0 X_1 X_2} = P_{S_1} P_{UVX_0|S_1} P_{X_1} P_{X_2}$, where U and V are auxiliary random variables satisfyinf that

$$I(U; Y_1) \geqslant I(U; S_1), \tag{23a}$$
$$I(V; Y_2) \geqslant I(V; U S_1). \tag{23b}$$

Proof. The region follows from Proposition 5 because (21b) and (21d) are redundant due to conditions (23a) and (23b) □

Following from the above achievable region, we obtain an achievable region for the Gaussian channel by setting an appropriate joint input distribution.

Proposition 6. *For the Gaussian channel of model II with $W_0 = \phi$, in the regime when $Q_1 \to \infty$, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leqslant \frac{1}{2} \log \left(1 + \frac{P_1}{(1 - \frac{1}{\alpha})^2 P_{01} + P_{02} + 1} \right) \tag{24a}$$
$$R_2 \leqslant \frac{1}{2} \log \left(1 + \frac{P_2}{1 + \frac{(\beta-1)^2 P_{02} P_{01}}{P_{02} + \beta^2 P_{01}}} \right) \tag{24b}$$

where $P_{01}, P_{02} \geqslant 0$, $P_{01} + P_{02} \leqslant P_0$, $0 < \alpha \leqslant \frac{2P_{01}}{1 + P_0 + P_1}$, and $P_{02}^2 + 2\beta P_{01} P_{02} \geqslant \beta^2 P_{01}(P_{02} + P_2 + 1)$.

Proof. The region follows from Corollary 2 by choosing jointly Gaussian distribution for random variables as follows:

$$U = X_{01} + \alpha S_1, \; V = X_{02} + \beta X_{01}$$
$$X_0 = X_{01} + X_{02}$$
$$X_{01} \sim \mathcal{N}(0, P_{01}), \; X_{02} \sim \mathcal{N}(0, P_{02})$$
$$X_1 \sim \mathcal{N}(0, P_1), \; X_2 \sim \mathcal{N}(0, P_2)$$

where X_{01}, X_{02}, X_1, X_2 and S_1 are independent. □

Comparing the inner and outer bounds given in Propositions 6 and 4, respectively, we characterize two segments of the boundary of the capacity region, over which the two bounds meet.

Theorem 4. *Consider the Gaussian channel of model II with $W_0 = \phi$, in the regime when $Q_1 \to \infty$, the rate points on the line A-B (see Fig. 7) are on the capacity region boundary. More specifically, if $\frac{1}{2}(1 + P_0 + P_1) \geqslant \frac{P_0^2}{P_0 + P_2 + 1}$, points A and B are characterized as*

$$Point\ A: \left(0, \frac{1}{2} \log(1 + P_2) \right)$$
$$Point\ B: \Big(\frac{1}{2} \log \big(1 + \frac{4 P_1 P_0^2}{(1 + P_0 + P_1)^2 (1 + P_0 + P_2) - 4 P_1 P_0^2} \big), \frac{1}{2} \log(1 + P_2) \Big).$$

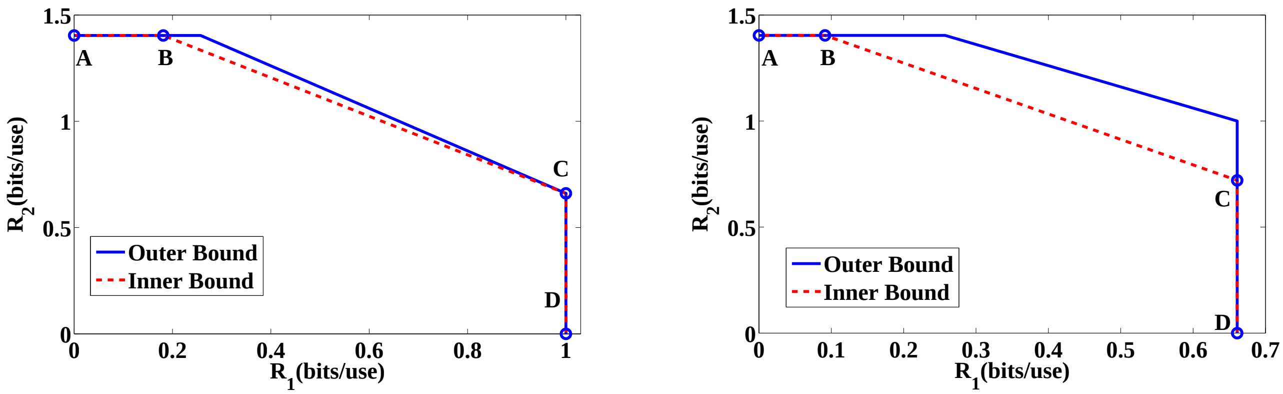
If $\frac{1}{2}(1 + P_0 + P_1) < \frac{P_0^2}{P_0 + P_2 + 1}$, points A and B are characterized as

$$Point\ A: \left(0, \frac{1}{2} \log(1 + P_2) \right)$$
$$Point\ B: \left(\frac{1}{2} \log \left(1 + \frac{P_1 (P_0 + P_2 + 1)}{P_0 + (P_0 + 1)(P_2 + 1)} \right), \frac{1}{2} \log(1 + P_2) \right).$$

Furthermore, the rate points on the line C-D (see Fig. 7) are also on the capacity region boundary. If $P_1 \geqslant P_0 + 1$, the points C and D are characterized as

$$Point\ C: \left(\frac{1}{2} \log(1 + P_0), \frac{1}{2} \log \left(1 + \frac{P_2}{P_0 + 1} \right) \right)$$
$$Point\ D: \left(\frac{1}{2} \log(1 + P_0), 0 \right),$$

as illustrated in Fig. 7 (a).



(a) $P_1 > P_0 + 1$ $\qquad$ $0 \leqslant P_1 \leqslant P_0 - 1$

Fig. 7: Segments of the capacity boundary for the Gaussian channel of model II

If $P_1 \leqslant P_0 - 1$, the points C and D is characterized as

$$\text{Point } C: \left(\frac{1}{2}\log(1+P_1), \frac{1}{2}\log\left(1+\frac{P_2}{P_1+2}\right)\right)$$
$$\text{Point } D: \left(\frac{1}{2}\log(1+P_1), 0\right),$$

as illustrated in Fig. 7 (b).

Proof. We first show that the line A-B is achievable. The point A is achievable by keeping the helper silent. To show that the point B is achievable, we set $\beta = 1$ in Proposition 6, and hence the achievable rate R_2 in (24b) reaches the point-to-point channel capacity, and the condition $P_{02}^2 + 2\beta P_{01}P_{02} \geqslant \beta^2 P_{01}(P_{02}+P_2+1)$ becomes $P_{01} \leqslant \frac{P_0^2}{P_0+P_2+1}$. We set $P_{01} = \frac{P_0^2}{P_0+P_2+1}$.

If $\frac{1}{2}(1+P_0+P_1) \geqslant \frac{P_0^2}{P_0+P_2+1}$, we have $\frac{2P_{01}}{1+P_0+P_1} \leqslant 1$. Thus, setting $\alpha = \frac{2P_{01}}{1+P_0+P_1}$, (24a) and (24b) imply that the point B is achievable.

If $\frac{1}{2}(1+P_0+P_1) \leqslant \frac{P_0^2}{P_0+P_2+1}$, we have $\frac{2P_{01}}{1+P_0+P_1} \geqslant 1$. By setting $\alpha = 1$, (24a) and (24b) imply that the point B is achievable.

We next show that the line C-D is on the capacity boundary.

As implied by Theorems 1 and 3, the only possible cases that the outer bound (20a) (i.e. the maximum rate R_1 with the helper fully assisting receiver 1) can be achieved are when $P_1 \leqslant P_0 - 1$ and $P_1 \geqslant P_0 + 1$.

If $P_1 \geqslant P_0 + 1$, setting the actual transmission power of transmitter 1 as $\tilde{P}_1 = P_0 + 1$, $P_{01} = P_0$, $\alpha = \frac{P_0}{1+P_0}$ and $\beta = 0$, then (24a) and (24b) imply that the rate point C is achievable. This point also achieves the sum capacity. It is obvious that point D is achievable, and hence the points on the line C-D are on the capacity boundary.

If $P_1 \leqslant P_0 - 1$, by setting $\beta = 0$, $\alpha = 1$ and $P_{01} = \tilde{P}_0 = P_1 + 1$ (where $\tilde{P}_0$ is the actual transmission power of the helper), then (24a) and (24b) imply that the rate point C is achievable. In particular, the actual power the helper uses is P_1+1 rather than P_0, because larger P_0 does not help receiver 1 to decode more, but increases interference to receiver 2. It is clear that the point D is achievable. Hence, the points on the line C-D are on the capacity boundary. ☐

The capacity result for the line A-B in Theorem 4 indicates that our coding scheme effectively enables the helper to assist receiver 1 without causing interference to receiver 2. Hence, R_2 achieves the corresponding point-to-point channel capacity, while transmitter 1 and receiver 1 communicate at a certain positive rate R_1 with the assistance of the helper.

The capacity result for the line C-D in Theorem 4 can be achieved based on a scheme, in which the helper assists receiver 1 to deal with the state and receiver 2 treats the helper's signal as noise. Such a scheme is guaranteed to be the best by the outer bound if receiver 1's rate is maximized.

Remark 3. *Theorem 4 implies that if $P_1 \geqslant P_0 + 1$, the sum capacity is achieved by the point C as illustrated in Fig. 7 (a).*

Corollary 3. *For the Gaussian channel of model II with $W_0 = \phi$, in the regime when $Q_1 \to \infty$, if $P_1 \geqslant P_0 + 1$, the sum capacity is given by $\frac{1}{2}\log(1+P_0+P_2)$.*

B. Scenario with Non-Dedicated Helper ($W_0 \neq \phi$)

In this subsection, we study the scenario, in which the helper also has its own message to transmit in addition to assisting the state-corrupted receivers, i.e., $W_0 \neq \phi$. The results we present below extend those in the preceding subsection for the scenario with $W_0 = \phi$. The proof techniques are similar and hence are omitted.

We first provide an outer bound for the Gaussian channel, which generalizes Propositions 1 and 4.

Proposition 7. *For the Gaussian channel of model II with $W_0 \neq \phi$, an outer bound on the capacity region for the regime* *when $Q_1 \to \infty$ consists of rate pairs (R_1, R_2) satisfying:*

$$R_0 \leqslant \frac{1}{2}\log(1+P_0) \tag{29a}$$

$$R_1 \leqslant \min\left\{\frac{1}{2}\log(1+P_0), \frac{1}{2}\log(1+P_1)\right\} \tag{29b}$$

$$R_2 \leqslant \frac{1}{2}\log(1+P_2) \tag{29c}$$

$$R_0+R_1+R_2 \leqslant \frac{1}{2}\log(1+P_0+P_2). \tag{29d}$$

We next present an achievable region for the discrete memoryless channel which generalizes Proposition 5 and Corollary 2.

Proposition 8. *For the discrete memoryless channel of model II with $W_0 \neq \phi$, an achievable region consists of the rate tuples (R_0, R_1, R_2) satisfying*

$$R_0 \leqslant I(X_{00}; Y_0) \tag{30a}$$

$$R_1 \leqslant I(X_1; Y_1|U) \tag{30b}$$

$$R_2 \leqslant I(X_2; Y_2|V) \tag{30c}$$

for some distribution $P_{S_1 UV X_{00} X_0 X_1 X_2} = P_{X_{00}} P_{S_1} P_{UVX_0|S_1 X_{00}} P_{X_1} P_{X_2}$, where U and V are auxiliary random variables that satisfy

$$I(U;Y_1) \geqslant I(U; S_1 X_{00}) \tag{31a}$$

$$I(V;Y_2) \geqslant I(V; US_1X_{00}). \tag{31b}$$

In the above proposition, the variable X_{00} is an auxiliary random variable representing the helper's own information for its intended receiver.

We then obtain an achievable region for the Gaussian channel by setting an appropriate joint distribution in Proposition 8.

Proposition 9. *For the Gaussian channel of model II with $W_0 \neq \phi$, an inner bound on the capacity region for the regime when $Q_1 \to \infty$ consists of rate tuples (R_0, R_1, R_2) satisfying:*

$$R_0 \leqslant \frac{1}{2}\log\left(1+\frac{P_{00}}{P_{01}+P_{02}+1}\right) \tag{32a}$$

$$R_1 \leqslant \frac{1}{2}\log\left(1+\frac{P_1}{(1-\frac{1}{\alpha})^2 P_{01}+P_{02}+1}\right) \tag{32b}$$

$$R_2 \leqslant \frac{1}{2}\log\left(1+\frac{P_2}{1+\frac{(\beta-1)^2 P_{02}(P_{00}+P_{01})}{P_{02}+\beta^2(P_{00}+P_{01})}}\right). \tag{32c}$$

where $P_{00}, P_{01}, P_{02} \geqslant 0$, $P_{00}+P_{01}+P_{02} \leqslant P_0$, $0<\alpha \leqslant \frac{2P_{01}}{1+P_{01}+P_{02}+P_1}$, and $P_{02}^2 + 2\beta P_{02}(P_{01}+P_{00}) \geqslant \beta^2(P_{00}+P_{01})(P_{02}+P_2+1)$.

Proof. The region follows from Proposition 8 by choosing jointly Gaussian distribution for random variables as follows:

$$U = X_{01} + \alpha(S_1 + X_{00}),\ V = X_{02} + \beta(X_{00}+X_{01})$$
$$X_0 = X_{00} + X_{01} + X_{02}$$
$$X_{00} \sim \mathcal{N}(0, P_{00}),\ X_{01} \sim \mathcal{N}(0, P_{01}),\ X_{02} \sim \mathcal{N}(0, P_{02})$$
$$X_1 \sim \mathcal{N}(0, P_1),\ X_2 \sim \mathcal{N}(0, P_2)$$

where X_{00}, X_{01}, X_{02}, X_1, X_2 and S_1 are independent. ☐

By comparing the inner and outer bounds, we obtain segments on the capacity region boundary as generalization of Theorem 4.

Theorem 5. *Consider the Gaussian channel of model II with $W_0 \neq \phi$, in the regime when $Q_1 \to \infty$, for any given achievable rate R_0, correspondingly a certain power P_{00}, the rate points on the line A-B are on the capacity region boundary. More specifically, if $\frac{1}{2}(1+P_1+P_0-P_{00}) > \frac{P_0^2}{P_0+P_2+1} - P_{00}$, the points A and B are characterized as:*

$$Point\ A: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),\ 0,\ \frac{1}{2}\log(1+P_2)\Big)$$
$$Point\ B: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),$$
$$\frac{1}{2}\log\big(1+\frac{P_1}{\frac{(1+P_0-P_{00}+P_1)^2(1+P_0+P_2)}{4(P_0^2-P_{00}(P_0+P_2+1))}-P_1}\big),\ \frac{1}{2}\log(1+P_2)\Big)$$

If $\frac{1}{2}(1+P_1+P_0-P_{00}) \leqslant \frac{P_0^2}{P_0+P_2+1} - P_{00}$, the points A and B are characterized as:

$$Point\ A: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),\ 0,\ \frac{1}{2}\log(1+P_2)\Big)$$
$$Point\ B: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),$$
$$\frac{1}{2}\log\big(1+\frac{P_1(P_0+P_2+1)}{P_0+(P_0+1)(P_2+1)}\big),\ \frac{1}{2}\log(1+P_2)\Big)$$

Furthermore, the rate points on the line C-D characterized below is on the capacity region boundary. If $P_1 \geqslant P_0-P_{00}+1$, the points C and D are characterized as:

$$Point\ C: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),$$
$$\frac{1}{2}\log(1+P_0-P_{00}),\ \frac{1}{2}\log\big(1+\frac{P_2}{P_0+1}\big)\Big)$$
$$Point\ D: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),\ \frac{1}{2}\log(1+P_0-P_{00}),\ 0\Big)$$

If $P_1 \leqslant P_0-P_{00}-1$, the points C and D are characterized as:

$$Point\ C: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),$$
$$\frac{1}{2}\log(1+P_1),\ \frac{1}{2}\log\big(1+\frac{P_2}{P_1+2}\big)\Big)$$
$$Point\ D: \Big(\frac{1}{2}\log\big(1+\frac{P_{00}}{P_0-P_{00}+1}\big),\ \frac{1}{2}\log(1+P_1),\ 0\Big)$$

Theorem 5 implies the following characterization of the sum capacity.

Corollary 4. *For the Gaussian channel of model II with $W_0 \neq \phi$, in the regime when $Q_1 \to \infty$, for a given $0 \leqslant P_{00} \leqslant P_0$, if $P_1 \geqslant P_0 - P_{00}+1$, the sum capacity is given by $\frac{1}{2}\log(1+P_0+P_2)$.*

V. Model III: General K

In this section, we consider the Gaussian channel of model III with $K \geqslant 2$, in which there are multiple receivers with each interfered by an independent state. We first study the scenario, in which the helper dedicates to help two users without transmitting its own message, i.e., $K=2$ and $W_0=\phi$. We then extend the result to the more general scenario, in which the helper dedicates to help more than two users without transmitting its own message, i.e., $K>2$ and $W_0=\phi$.

Finally, we study the case, in which the helper also has its own message destined for the corresponding receiver in addition to helping K $(K \geqslant 2)$ users, i.e., $K \geqslant 2$ and $W_0 \neq \phi$.

A. Scenario with Two State-Corrupted Receivers and Dedicated Helper ($K = 2$, $W_0 = \phi$)

In this subsection, we study the scenario with $K = 2$, which is more instructional. The case with $K \geqslant 2$ is relegated to Section V-B. We note that model III is more general than model I, but does not include model II as a special case, because model II has one receiver that is not corrupted by state, but each receiver (excluding the helper) in model III is corrupted by an infinitely powered state sequence. Hence for model III, the challenge lies in the fact that the helper needs to assist multiple receivers to cancel interference caused by independent states. In this subsection, we first derive an outer bound on the capacity region, and then derive an inner bound based on a time-sharing scheme for the helper. Somewhat interestingly, comparing the inner and outer bounds concludes that the time-sharing scheme achieves the sum capacity under certain channel parameters, and we hence characterize segments of the capacity region boundary corresponding to the sum capacity under these channel parameters.

We first derive an outer bound on the capacity region.

Proposition 10. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime when $Q_1, Q_2 \to \infty$, an outer bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leqslant \frac{1}{2}\log(1 + P_1) \tag{37}$$

$$R_2 \leqslant \frac{1}{2}\log(1 + P_2) \tag{38}$$

$$R_1 + R_2 \leqslant \frac{1}{2}\log(1 + P_0). \tag{39}$$

Proof. The proof is detailed in Appendix E. □

We note that although the two transmitters transmit over parallel channels, the above outer bound suggests that their sum rate is still subject to a certain constraint determined by the helper's power. This implies that it is not possible for one common helper to cancel the two independent high-power states simultaneously (i.e., using the common resource). This fact also suggests that a time-sharing scheme, in which the helper alternatively assists each receiver, can be desirable to achieve the sum rate upper bound (i.e., to achieve the sum capacity).

We hence design the following time-sharing achievable scheme. The helper splits its transmission duration into two time slots with the fraction γ of the total time duration for assisting receiver 1 and the fraction $1-\gamma$ for assisting receiver 2. Each transmitter transmits only during the time slot that it is assisted by the helper, and keeps silent while the helper assisting the other transmitter. We note that the power constraints for transmitters 1 and 2 in their corresponding transmission time slots are $\frac{P_1}{\gamma}$ and $\frac{P_2}{1-\gamma}$, respectively.

Now at each transmission slot, the channel consists of one transmitter-receiver pair with the receiver corrupted by a infinite-power state, and one helper that assists the receiver to cancel the state interference. Such a model is a special case of model I studied in Section III (with the helper not having its own message). Hence, following from Proposition 3, we can write the achievable rate as a function of P and P_0 respectively at the transmitter and the helper as follows:

$$R(P, P_0) := \begin{cases} \frac{1}{2}\log(1 + P_0), & P \geqslant P_0 + 1 \\ \frac{1}{2}\log\left(1 + \frac{4P_0P}{4P_0 + (P_0 - P - 1)^2}\right), & P_0 - 1 \leqslant P \leqslant P_0 + 1 \\ \frac{1}{2}\log(1 + P), & P \leqslant P_0 - 1. \end{cases} \tag{40}$$

By employing the time-sharing scheme between the helper assisting one receiver and the other alternatively, we obtain the following achievable region.

Proposition 11. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime with $Q_1, Q_2 \to \infty$, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \leqslant \gamma R\left(\frac{P_1}{\gamma}, P_0\right) \tag{41a}$$

$$R_2 \leqslant (1-\gamma) R\left(\frac{P_2}{1-\gamma}, P_0\right) \tag{41b}$$

where $0 \leqslant \gamma \leqslant 1$ is the time-sharing coefficient, and the function $R(\cdot,\cdot)$ is defined in (40).

We note that following from (40), the best possible single-user rate is $\frac{1}{2}\log(1+P_0)$, which can be achieved if $P \geqslant P_0+1$. This best rate may not be possible if P is not large enough. Interestingly, in a time-sharing scheme, both transmitters can simultaneously achieve the best single user rate $\frac{1}{2}\log(1+P_0)$ over their transmission fraction of time, because both of their powers get boosted over a certain fraction of time, although neither power is larger than $P_0 + 1$. In this way, the sum rate upper bound (39) can be achieved. Such a fact also suggests that in the helper-power limited regime, letting the helper simultaneously assist all users typically does not achieve further multiplexing gain. The following theorem characterizes the sum capacity of the channel for the scenario described above.

Theorem 6. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime with $Q_1, Q_2 \to \infty$, if $P_1 + P_2 \geqslant P_0 + 1$, then the sum capacity equals $\frac{1}{2}\log(1 + P_0)$. The rate points that achieve the sum capacity (i.e. on the capacity region boundary) are characterized as $(R_1, R_2) = \Big(\gamma R(\frac{P_1}{\gamma}, P_0), (1-\gamma)R(\frac{P_2}{1-\gamma}, P_0)\Big)$ for $\gamma \in \Big(\max(1 - \frac{P_2}{P_0+1}, 0), \min(\frac{P_1}{P_0+1}, 1)\Big)$.*

Proof. The proof is detailed in Appendix F. □

The above theorem implies the following characterization of the full capacity region under certain parameters.

Corollary 5. *For the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, in the regime with $Q_1, Q_2 \to \infty$, if*

$P_1,\ P_2 \geqslant P_0+1$, then the capacity region consists of the rate pair (R_1,R_2) satisfying $R_1+R_2 \leqslant \frac{1}{2}\log(1+P_0)$.

We next provide channel examples to understand the outer and inner bounds respectively in Proposition 10 and 11, and in the sum capacity in Theorem 6. It can be seen that the power constraints fall into four cases, among which we consider the following three cases: case 1. $P_1 \geqslant P_0, P_2 \geqslant P_0$; case 2. $P_1 \geqslant P_0, P_2 < P_0$; and case 3. $P_1 < P_0, P_2 < P_0$ by noting that case 4 is opposite to case 2 and is omitted due to symmetry of the two transmitters.

- Case 1: $P_1 \geqslant P_0, P_2 \geqslant P_0$. We consider an example channel with $P_0 = 1$, $P_1 = 1.8$ and $P_2 = 1.5$. Fig. 8 (a) plots the inner and outer bounds on the capacity region. In particular, the two bounds meet over the line segment B-C, which corresponds to the rate points $(R_1,R_2) = \Big(\gamma R(\frac{P_1}{\gamma},P_0),(1-\gamma)R(\frac{P_2}{1-\gamma},P_0)\Big)$ for $\gamma \in \Big(\max(1-\frac{P_2}{P_0+1},0),\min(\frac{P_1}{P_0+1},1)\Big)$ as characterized in Theorem 6. All these rate points achieve the sum capacity. It can also be seen that although neither transmitter achieves the best possible single-user rate, the sum capacity can be achieved due to the time-sharing scheme. We also note that, in this case, if the conditions in Corollary 5 are satisfied, the full capacity region is characterized.
- Case 2: $P_1 > P_0, P_2 \leqslant P_0$. We consider an example channel with $P_0 = 2$, $P_1 = 2.5$ and $P_2 = 0.8$. Fig. 8 (b) plots the inner and outer bounds on the capacity region. Similarly to case 1, the two bounds meet over the line segment B-C as characterized in Theorem 6, and the points on such a line segment achieve the sum capacity. Differently from case 1, transmitter 2 achieves its point-to-point channel capacity indicated by the point A in Figure 8 (b). This is consistent with the single user rate provided in (40) for the case with $P_2 \leqslant P_0-1$.
- Case 3: $P_1 < P_0, P_2 < P_0$. We consider an example channel with $P_0 = 4$, $P_1 = 3$ and $P_2 = 3$. Figure 8 (c) plots the inner and outer bounds on the capacity region. The points on the line segment B-C achieve the sum capacity as characterized in Theorem 6, and the points A and D respectively achieve the point-to-point capacity for two transceiver pairs. This is consistent with the single-user rate provided in (40) for the case with $P_1,P_2 \leqslant P_0-1$.

B. Scenario with Multiple State-Corrupted Receivers and Dedicated Helper ($K \geqslant 2$, $W_0 = \phi$)

In this subsection, we study the scenario, with $K \geqslant 2$ state-corrupted receivers, in which the helper is dedicated to assist these receivers, i.e., $W_0 = \phi$. The results we present below extend those in the preceding subsection for the scenario with $K = 2$. The proof techniques are similar and hence are omitted.

We first provide an outer bound on the capacity region as follows, which is an extension of Proposition 10.

Proposition 12. *For the Gaussian channel of model III with $K \geqslant 2$ and $W_0 = \phi$, in the regime when $Q_1,\ldots,Q_K \to \infty$, an outer bound on the capacity region consists of rate tuples $(R_0,\ldots,R_K)$ satisfying:*

$$\begin{aligned} &R_k \leqslant \frac{1}{2}\log(1+P_k) \quad for\ k=1,\ldots,K \\ &\sum_{k=1}^{K} R_k \leqslant \frac{1}{2}\log(1+P_0). \end{aligned}$$

In order to derive an inner bound, we employ the following time-sharing scheme. We divide the entire transmission time into K slots, and during each slot the helper assists one receiver. We provide the corresponding achievable region as follows as an extension of Proposition 11.

Proposition 13. *For the Gaussian channel of model III with $K \geqslant 2$ and $W_0 = \phi$, in the regime when $Q_1,\ldots,Q_K \to \infty$, an inner bound consists of rate tuples $(R_1,\ldots,R_K)$ satisfying:*

$$R_k \leqslant \gamma_k R\left(\frac{P_k}{\gamma_k},P_0\right), \quad for\ k=1,\ldots,K$$

for some γ_k that satisfy

$$\sum_{k=1}^{K}\gamma_k = 1,\ and\ \gamma_k \geqslant 0 \quad for\ k=1,\ldots,K,$$

where $R(\cdot,\cdot)$ is the function defined in (40)*.*

By comparing the inner and outer bounds, we obtain the following result on the capacity as an extension of Theorem 6.

Theorem 7. *For the Gaussian channel of model III with $K \geqslant 2$ and $W_0 = \phi$, in the regime with $Q_1,\ldots,Q_k \to \infty$, the sum capacity equals $\frac{1}{2}\log(1+P_0)$. The rate points that achieve the sum capacity on the capacity region boundary are characterized as*

$$R_k \leqslant \frac{\gamma_k}{2}\log(1+P_0) \qquad k=1,\ldots,K$$

where

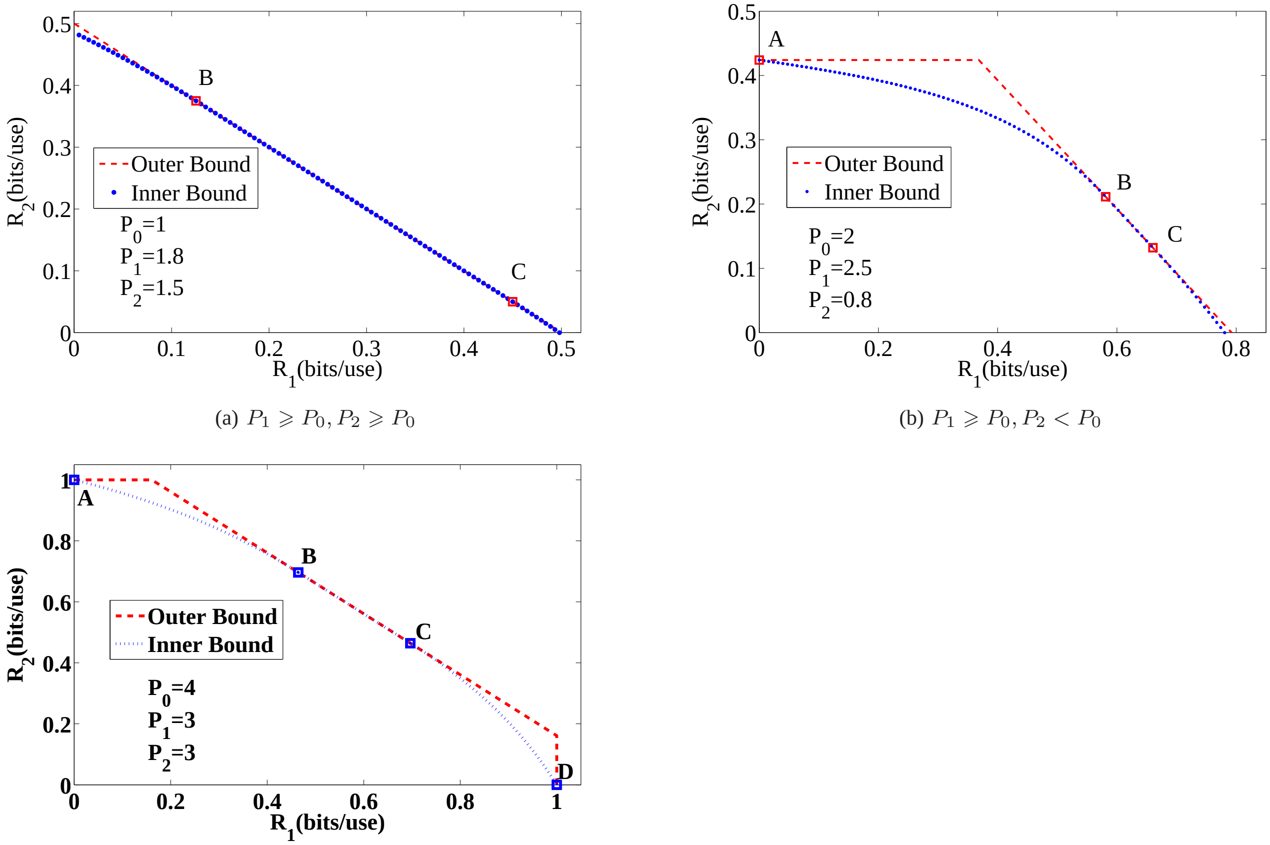
$$\begin{aligned} &\sum_{k=1}^{K}\gamma_k = 1, \ \ \gamma_k \geqslant 0 \\ &\frac{P_k}{\gamma_k} \geqslant P_0+1 \quad for\ k=1,\ldots,K. \end{aligned}$$

C. Scenario with Non-Dedicated Helper ($W_0 \neq \phi$)

In this subsection, we study the scenario, in which the helper also has its own message for the corresponding receiver in addition to assisting the state-corrupted receivers, i.e., $W_0 \neq \phi$. The results we present below extend those in the preceding subsection for the scenario with $W_0 = \phi$ as well as the results in Section III for model I. The proof techniques combine those in Sections V-A and III, and some are omitted.

We first provide an outer bound on the capacity region as follows, which is an extension of Proposition 10.

Proposition 14. *For the Gaussian channel of model III with $W_0 \neq \phi$, in the regime when $Q_1,\ldots,Q_K \to \infty$, an*



(a) $P_1 \geqslant P_0, P_2 \geqslant P_0$

(b) $P_1 \geqslant P_0, P_2 < P_0$

(c) $P_1 < P_0, P_2 < P_0$

Fig. 8: Inner and outer bounds for the Gaussian channel of model III with $K = 2$ and $W_0 = \phi$, which match partially.

outer bound on the capacity region consists of rate tuples $(R_0, \ldots, R_K)$ *satisfying:*

$$R_k \leqslant \frac{1}{2}\log(1+P_k) \quad for\ k = 0, \ldots, K$$
$$\sum_{k=0}^{K} R_k \leqslant \frac{1}{2}\log(1+P_0).$$

Proof. The proof is detailed in Appendix G. $\square$

In order to derive an inner bound, we employ the following time-sharing scheme. We divide the entire transmission time into K slots, and during each slot the helper assists one receiver and simultaneously transmits its own message. We provide the corresponding achievable region as follows as an extension of Proposition 11.

Proposition 15. *For the Gaussian channel of model III with* $W_0 \neq \phi$*, in the regime when* $Q_1, \ldots, Q_K \to \infty$*, an inner bound consists of rate tuples* $(R_0, \ldots, R_K)$ *satisfying:*

$$R_0 \leqslant \sum_{k=1}^{K} \frac{\gamma_k}{2} \log\left(1 + \frac{\bar{\beta}_k P_0}{\beta_k P_0 + 1}\right)$$
$$R_k \leqslant \gamma_k R\left(\frac{P_k}{\gamma_k}, \beta_k P_0\right) \quad for\ k = 1, \ldots, K$$

for some parameters that satisfy

$$\sum_{k=1}^{K} \gamma_k = 1, \quad \gamma_k \geqslant 0,$$
$$\bar{\beta}_k + \beta_k = 1, \quad \beta_k \geqslant 0 \quad for\ k = 1, \ldots, K,$$

where $R(\cdot,\cdot)$ *is the function defined in* (40)*.*

By comparing the inner and outer bounds, we obtain the following result on the capacity as an extension of Theorem 6.

Theorem 8. *For the Gaussian channel of model III with* $W_0 \neq \phi$*, in the regime with* $Q_1, \ldots, Q_k \to \infty$*, the sum capacity equals* $\frac{1}{2}\log(1+P_0)$*. The rate points that achieve the sum capacity on the capacity region boundary are characterized as*

$$R_0 \leqslant \sum_{k=1}^{K} \frac{\gamma_k}{2} \log\left(1 + \frac{\bar{\beta}_k P_0}{\beta_k P_0 + 1}\right)$$
$$R_k \leqslant \frac{\gamma_k}{2}\log(1+\beta_k P_0) \quad for\ k = 1, \ldots, K$$

where

$$\sum_{k=1}^{K}\gamma_k = 1,\ \ \gamma_k \geqslant 0$$
$$\frac{P_k}{\gamma_k} \geqslant \beta_k P_0 + 1$$
$$\bar{\beta}_k + \beta_k = 1,\ \ \beta_k \geqslant 0,\ \ for\ k = 1, \ldots, K.$$

VI. Further Discussion

In this paper, we focus on state-dependent Gaussian models in the large state-power regime, i.e., $Q \to \infty$. In this case, the state cannot be canceled by direct reversion as the helper has only finite amount of power. Thus, we design achievable schemes that precode the state into helper signals via single-bin dirty paper coding so that the state-interfered receiver can cancel the state interference. Two outer bounds are useful for us to understand the optimality of our achievable schemes. One is the capacity of the corresponding channel models without state, which captures how the transmitters' power limits the transmission rate if the state can be fully cancelled. We further develop a new outer bound in this paper, which captures how the helper's power limits the transmission rate. Based on these outer bounds, we characterize under what channel parameters the developed achievable schemes are optimal.

Such a study also provides some useful understanding for further exploring these models in the finite state-power regime. Since the state power is finite, direct reversion can help to partially or fully cancel the state. Hence, a natural achievable scheme is to combine direct reversion with single-bin dirty paper coding that we apply in this paper to cancel the state interference. For outer bounds, other than outer bounds used in this paper that capture impact of the transmitters' power and the helper's power on the transmission rate, it is desirable to derive more refined outer bounds that can capture how the state power limits the transmission rates. Such outer bounds (together with the inner bounds) will suggest under what channel parameters, only partially canceling state interference is already optimal.

In this paper, we study two types of models, i.e., model II with two users respectively being/not being corrupted by state, and model III (including model I as a special case) with all users being corrupted by state. Since the state power is infinite, models II and III do not contain each other as a special case. Although these two models capture different communication scenarios, they can be unified by a general model described as follows. Suppose there are two groups of transmitter-receiver pairs. All receivers in group 1 are corrupted by independent states, whereas all receivers in group 2 are not corrupted by state. Similarly to models II and III, a helper transmitter knows the states noncausally and wishes to help all receivers in addition to transmitting its own information. It is clear that such a model includes both model II and model III as special cases. Our analysis for models II and III can be naturally extended to study the above general model. First, the achievable scheme can be designed by integrating our achievable schemes for models II and III together. More specifically, for the receivers in group 1, time-sharing can be employed for the helper to assist one transmitter-receiver pair at a time. Then, between the two groups, superposition coding as for model II can be employed to design the achievable scheme. In order to design outer bounds, for each pair of receivers with one from group 1 and one from group 2, the outer bound that we derive for model II is applicable and serves as one upper bound. Then, combining all such upper bounds together provides an outer bound for the more general model. Based on such inner and outer bounds, one can analyze under what channel parameters the capacity region of the general model can be characterized.

VII. Conclusion

In this paper, we proposed and studied three models of parallel communication networks with a state-cognitive helper. For each model, there is unique challenge to design capacity-achieving schemes for the helper to trade off among multiple functions. For model I, we designed an adapted dirty paper coding together with superposition coding for the helper to trade off between assisting to cancel the state and transmitting its own message. We showed that such a scheme achieves the full capacity region or segments on the capacity region boundary for all channel parameters. For model II, we designed a multi-layer scheme, such that the helper assists receiver 1 to cancel the infinite-power state while simultaneously eliminating its interference to receiver 2. Such a scheme achieves two segments on the capacity region. Over one segment, the helper is capable to fully cancel the interference that it causes to receiver 2, and simultaneously assists receiver 1 to achieve a certain positive rate. In the second segment, the sum capacity is obtained with the helper dedicated to help receiver 1. For model III, we employed a time-sharing scheme such that the helper alternatively assists each receiver, and we showed that such a scheme achieves the sum capacity for certain channel parameters. These results suggest that with only state information of parallel users, a helper can still assist these users to cancel infinite-power state in an optimal way. We anticipate that the techniques that we develop in this paper will be useful for studying various other multi-user state-dependent models with state-cognitive helpers.

As we mentioned in Section I-A, large state interference and state-cognitive helpers are well justified in practical wireless networks, and such a model implies a promising perspective of interference cancelation in wireless networks. Hence, the schemes developed here can be potentially used to greatly improve the throughput of wireless networks.

Appendix A
Proof of Proposition 1

The first bound follows easily from the single-user rate bound of receiver 1 as follows.

$$\begin{aligned}
nR_1 &\leqslant I(W_1;Y_1^n) + n\epsilon_n \\
&\leqslant I(W_1;Y_1^n S_1^n X_0^n) + n\epsilon_n \\
&= I(W_1;Y_1^n|S_1^n X_0^n) + n\epsilon_n \\
&\leqslant h(Y_1^n|S_1^n X_0^n) - h(Y_1^n|W_1 S_1^n X_1^n X_0^n) + n\epsilon_n \\
&= h(X_1^n + N_1^n) - h(N_1^n) + n\epsilon_n \\
&\leqslant \frac{n}{2}\log(1+P_1)
\end{aligned} \tag{42}$$

We then bound the sum rate as follows. For the message W_0, based on Fano's inequality, we have

$$\begin{aligned} nR_0 &\leqslant I(W_0;Y_0^n)+n\epsilon_n \\ &= h(Y_0^n)-h(Y_0^n|W_0)+n\epsilon_n, \end{aligned} \tag{43}$$

where $\epsilon_n \to 0$ as $n \to \infty$.

For the message W_1, based on Fano's inequality, we have

$$\begin{aligned} nR_1 &\leqslant I(W_1;Y_1^n)+n\epsilon_n \\ &= h(Y_1^n)-h(Y_1^n|W_1)+n\epsilon_n \\ &\leqslant h(Y_1^n)-h(Y_1^n|W_1X_1^n)+n\epsilon_n \\ &= h(Y_1^n)-h(X_0^n+S_1^n+N_1^n)+n\epsilon_n \\ &\leqslant h(Y_1^n)-h(X_0^n+S_1^n+N_1^n|W_0Y_0^n)+n\epsilon_n \end{aligned} \tag{44}$$

Summation of (43) and (44) yields

$$\begin{aligned} n(R_0+R_1) &\leqslant h(Y_0^n)+h(Y_1^n) \\ &\quad - h(Y_0^n, X_0^n+S_1^n+N_1^n|W_0)+n\epsilon_n \\ &= h(Y_0^n)+h(Y_1^n) \\ &\quad - h(X_0^n+N_0^n, X_0^n+S_1^n+N_1^n|W_0)+n\epsilon_n \end{aligned} \tag{45}$$

Since the two receivers perform decoding independently, the capacity region of the channel depends on only the marginal distributions of (X_0,Y_0) and (X_0,X_1,S,Y_1). It is clear that setting $N_1=N_0$ does not change the two marginal distributions respectively involving Y_0 and Y_1, and hence does not affect the capacity region. Thus,

$$\begin{aligned} & n(R_0+R_1) \\ \leqslant & h(Y_0^n)+h(Y_1^n) \\ & - h(X_0^n+N_1^n, X_0^n+S_1^n+N_1^n|W_0)+n\epsilon_n \\ \leqslant & h(Y_0^n)+h(Y_1^n)-h(S_1^n, X_0^n+N_1^n|W_0)+n\epsilon_n \\ \leqslant & h(Y_0^n)+h(Y_1^n)-h(S_1^n)-h(N_1^n)+n\epsilon_n \\ \leqslant & \frac{n}{2}\log(1+P_0) \\ & + \frac{n}{2}\log\left(1+\frac{P_0+2\sqrt{P_0Q_1}+P_1+1}{Q_1}\right)+n\epsilon_n \end{aligned} \tag{46}$$

As $Q_1 \to \infty$, the second term of the above bound goes to 0, and we have

$$R_0+R_1 \leqslant \frac{1}{2}\log(1+P_0). \tag{47}$$

APPENDIX B
PROOF OF PROPOSITION 2

We use random codes and fix the following joint distribution:

$$\begin{aligned} P_{S_1X_0'UX_0X_1Y_0Y_1} =& P_{S_1}P_{X_0'}P_{U|S_1X_0'}P_{X_0|US_1X_0'} \\ & \cdot P_{X_1}P_{Y_0|X_0}P_{Y_1|X_0X_1S_1}. \end{aligned}$$

Let $T_\epsilon^n(P_{S_1X_0'UX_0X_1Y_0Y_1})$ denote the strongly joint ϵ-typical set (see, e.g., [29, Sec. 10.6] and [30, Sec. 1.3] for definition) based on the above distribution. For a given sequence x^n, let $T_\epsilon^n(P_{U|X}|x^n)$ denote the set of sequences u^n such that (u^n,x^n) is jointly typical based on the distribution P_{XU}.

1) Codebook Generation
   - Generate $2^{n\tilde{R}}$ i.i.d. codewords $u^n(t)$ according to $P(u^n)=\prod_{i=1}^n P_U(u_i)$ for the fixed marginal probability P_U as defined, in which $t\in[1,2^{n\tilde{R}}]$.
   - Generate 2^{nR_0} i.i.d codewords $x_0^{'n}(w_0)$ according to $P(x_0^{'n})=\prod_{i=1}^n P_{X_0'}(x_{0i}')$ for the fixed marginal probability $P_{X_0'}$ as defined, in which $w_0\in[1,2^{nR_0}]$.
   - Generate 2^{nR_1} i.i.d. codewords $x_1^n(w_1)$ according to $P(x_1^n)=\prod_{i=1}^n P_{X_1}(x_{1i})$ for the fixed marginal probability P_{X_1} as defined, in which $w_1\in[1,2^{nR_1}]$.

2) Encoding
   - Encoder at the helper: Given w_0, map w_0 into $x_0^{'n}(w_0)$. For each $x_0^{'n}(w_0)$, select $\tilde{t}$ such that $(u^n(\tilde{t}),s_1^n,x_0^{'n}(w_0)) \in T_\epsilon^n(P_{S_1}P_{X_0'}P_{U|S_1X_0'})$. If $u^n(\tilde{t})$ cannot be found, set $\tilde{t}=1$. Then map $(s_1^n,u^n(\tilde{t}),x_0^{'n}(w_0))$ into $x_0^n=f_0^{(n)}(x_0^{'n}(w_0),s_1^n,u^n(\tilde{t}))$. Based on the rate distortion type of argument [29, Sec. 10.5] or the Covering Lemma [31, Sec. 3.7], it can be shown that such $u^n(\tilde{t})$ exists with high probability for large n if

$$\tilde{R} > I(U;S_1X_0'). \tag{48}$$

   - Encoder 1: Given w_1, map w_1 into $x_1^n(w_1)$.

3) Decoding
   - Decoder 0: Given y_0^n, find $\hat{w}_0$ such that $(x_0^{'n}(\hat{w}_0),y_0^n) \in T_\epsilon^n(P_{X_0'Y_0})$. If no or more than one $\hat{w}_0$ can be found, declare an error. It can be shown that the decoding error is small for sufficient large n if

$$R_0 \leqslant I(X_0';Y_0). \tag{49}$$

     The proof for the above bound (and the similar bounds in the sequel) follows the standard techniques as given in [29, Sec. 7.7], and hence is omitted.
   - Decoder 1: Given y_1^n, find a pair $(\hat{t},\hat{w}_1)$ such that $(u^n(\hat{t}),x_1^n(\hat{w}_1),y_1^n)\in T_\epsilon^n(P_{UX_1Y_1})$. If no or more than one such pair can be found, then declare an error. It can be shown that decoding is successful with small probability of error for sufficiently large n if the following conditions are satisfied

$$R_1 \leqslant I(X_1;Y_1|U), \tag{50}$$
$$\tilde{R} \leqslant I(U;Y_1|X_1), \tag{51}$$
$$R_1+\tilde{R} \leqslant I(UX_1;Y_1). \tag{52}$$

We note that (51) corresponds to the decoding error for the index t, which is not the message of interest. Hence, the bound (51) can be removed. Hence, combining (48), (49), (50), and (52), and eliminating $\tilde{R}$, we obtain the desired achievable region.

Appendix C
Proof of Proposition 4

The single rate bounds follow from Proposition 1 and the point-to-point channel capacity. For the sum rate bound, based on Fano's inequality, we have

$$\begin{aligned}
&n(R_1+R_2)\\
\leqslant& I(W_1;Y_1^n)+I(W_2;Y_2^n)+n\epsilon_n\\
=&h(Y_1^n)-h(Y_1^n|W_1)+h(Y_2^n)-h(Y_2^n|W_2)+n\epsilon_n\\
\overset{(a)}{=}&h(Y_1^n)-h(Y_1^n|W_1X_1^n)+h(Y_2^n)-h(Y_2^n|W_2X_2^n)+n\epsilon_n\\
=&h(Y_1^n)-h(X_0^n+S_1^n+N_1^n)\\
&+h(Y_2^n)-h(X_0^n+N_2^n)+n\epsilon_n\\
\leqslant& h(Y_1^n)-h(X_0^n+S_1^n+N_1^n|X_0^n+N_1^n)\\
&+h(Y_2^n)-h(X_0^n+N_2^n)+n\epsilon_n
\end{aligned}$$

where (a) follows from the fact that X_1^n is a function of W_1, and X_2^n is a function of W_2, and they are independent from X_0^n, state and noise. As argued in Appendix A, setting $N_1^n = N_2^n$ does not change the capacity region. Thus,

$$\begin{aligned}
&n(R_1+R_2)\\
\leqslant& h(Y_1^n)-h(X_0^n+S_1^n+N_1^n,X_0^n+N_1^n)+h(Y_2^n)+n\epsilon_n\\
=&h(Y_1^n)-h(S_1^n,X_0^n+N_1^n)+h(Y_2^n)+n\epsilon_n\\
=&h(Y_1^n)-h(S_1^n)-h(X_0^n+N_1^n|S_1^n)+h(Y_2^n)+n\epsilon_n\\
\leqslant& h(Y_1^n)-h(S_1^n)-h(X_0^n+N_1^n|S_1^n,X_0^n)+h(Y_2^n)+n\epsilon_n\\
\overset{(b)}{=}&h(X_0^n+X_1^n+S_1^n+N_1^n)-h(S_1^n)-h(N_1^n)\\
&+h(X_0^n+X_2^n+N_1^n)+n\epsilon_n\\
\leqslant&\frac{n}{2}\log 2\pi e(P_1+P_0+2\sqrt{P_0Q_1}+Q_1+1)-\frac{n}{2}\log(2\pi eQ_1)\\
&+\frac{n}{2}\log 2\pi e(P_0+P_2+1)-\frac{n}{2}\log(2\pi e)+n\epsilon_n\\
=&\frac{n}{2}\log\left(\frac{P_1+P_0+2\sqrt{P_0Q_1}+Q_1+1}{Q_1}\right)\\
&+\frac{n}{2}\log(P_0+P_2+1)+n\epsilon_n\\
\to&\frac{n}{2}\log(P_0+P_2+1)\ \ as\ Q_1\to\infty
\end{aligned}$$

where (b) follows from the fact that X_0^n and S_1^n are independent from N_1^n.

Appendix D
Proof of Proposition 5

We use random codes and fix the following joint distribution:

$$\begin{aligned}
P_{S_1UVX_0X_1X_2Y_1Y_2}=&P_{VUS_1}P_{X_0|VUS_1}\\
&\cdot P_{X_1}P_{X_2}P_{Y_1|X_0X_1S_1}P_{Y_2|X_0X_2}. \qquad (53)
\end{aligned}$$

Let $T_\epsilon^n(P_{S_1UVX_0X_1X_2Y_1Y_2})$ denote the strongly joint ϵ-typical set based on the above distribution.

1) Codebook Generation
   - Generate $2^{n\tilde{R}_1}$ i.i.d. codewords $u^n(t)$ according to $P(u^n)=\prod_{i=1}^n P_U(u_i)$ for the fixed marginal probability P_U as defined, in which $t\in[1,2^{n\tilde{R}_1}]$.
   - Generate $2^{n\tilde{R}_2}$ i.i.d. codewords $v^n(k)$ according to $P(v^n)=\prod_{i=1}^n P_V(v_i)$ for the fixed marginal probability P_V as defined, in which $k\in[1,2^{n\tilde{R}_2}]$.
   - Generate 2^{nR_1} i.i.d. codewords $x_1^n(w_1)$ according to $P(x_1^n)=\prod_{i=1}^n P_{X_1}(x_{1i})$ for the fixed marginal probability P_{X_1} as defined, in which $w_1\in[1,2^{nR_1}]$.
   - Generate 2^{nR_2} i.i.d. codewords $x_2^n(w_2)$ according to $P(x_2^n)=\prod_{i=1}^n P_{X_2}(x_{2i})$ for the fixed marginal probability P_{X_2} as defined, in which $w_2\in[1,2^{nR_2}]$.
2) Encoding
   - Encoder at the helper: Given s_1^n, find $\tilde{t}$, such that $(u^n(\tilde{t}),s_1^n)\in T_\epsilon^n(P_{S_1U})$. Such $u^n(\tilde{t})$ exists with high probability for large n if
   $$\tilde{R}_1\geqslant I(U;S_1). \qquad (54)$$
   - For each $\tilde{t}$ selected, select $\tilde{k}$, such that $(v^n(\tilde{k}),u^n(\tilde{t}),s_1^n)\in T_\epsilon^n(P_{VUS_1})$. Such $v^n(\tilde{k})$ exists with high probability for large n if
   $$\tilde{R}_2\geqslant I(V;S_1U). \qquad (55)$$
   - Map (s_1^n,u^n,v^n) into $x_0^n=f_0^{(n)}(u^n(\tilde{t}),v^n(\tilde{k}),s_1^n)$.
   - Encoder 1: Given w_1, map w_1 into $x_1^n(w_1)$.
   - Encoder 2: Given w_2, map w_2 into $x_2^n(w_2)$.
3) Decoding
   - Decoder 1: Given y_1^n, find $(\hat{w}_1,\hat{t})$ such that $(x_1^n(\hat{w}_1),u^n(\hat{t}),y_1^n)\in T_\epsilon^n(P_{X_1UY_1})$. If no or more than one $\hat{w}_1$ can be found, declare an error. One can show that the decoding error is small for sufficient large n if
   $$\begin{aligned}R_1&\leqslant I(X_1;Y_1U) &(56)\\ R_1+\tilde{R}_1&\leqslant I(X_1U;Y_1). &(57)\end{aligned}$$
   - Decoder 2: Given y_2^n, find $(\hat{w}_2,\hat{k})$ such that $(x_2^n(\hat{w}_2),v^n(\hat{k}),y_2^n)\in T_\epsilon^n(P_{X_2VY_2})$. If no or more than one $\hat{w}_2$ can be found, declare an error. One can show that the decoding error is small for sufficient large n if
   $$\begin{aligned}R_2&\leqslant I(X_2;Y_2V), &(58)\\ R_2+\tilde{R}_2&\leqslant I(X_2V;Y_2). &(59)\end{aligned}$$

Combining (54)-(59), and eliminating $\tilde{R}_1$ and $\tilde{R}_2$, we obtain the desired achievable region.

Appendix E
Proof of Proposition 10

The bounds on R_1 and R_2 follow from the point-to-point channel capacity. For the sum rate bound, based on the Fano's inequality, we have

$$\begin{aligned}
&n(R_1+R_2)\\
\leqslant& I(W_1;Y_1^n)+I(W_2;Y_2^n)+n\epsilon_n\\
=&h(Y_1^n)-h(Y_1^n|W_1)+h(Y_2^n)-h(Y_2^n|W_2)+n\epsilon_n\\
\overset{(a)}{=}&h(Y_1^n)-h(Y_1^n|W_1X_1^n)+h(Y_2^n)-h(Y_2^n|W_2X_2^n)+n\epsilon_n\\
=&h(Y_1^n)-h(X_0^n+S_1^n+N_1^n)+h(Y_2^n)\\
&-h(X_0^n+S_2^n+N_2^n)+n\epsilon_n\\
\leqslant& h(Y_1^n)-h(X_0^n+S_1^n+N_1^n|X_0^n+N_1^n)+h(Y_2^n)\\
&-h(X_0^n+S_2^n+N_2^n|X_0^n+N_2^n,X_0^n+S_1^n+N_1^n)\\
&+h(X_0^n+N_1^n)-h(X_0^n+N_1^n)+n\epsilon_n
\end{aligned}$$

where (a) follows from the fact that X_1^n is a function of W_1, X_2^n is a function of W_2, and they are independent from X_0^n, S_1^n, S_2^n, N_1^n and N_2^n. Since receivers 1 and 2 decode based on the marginal distributions only, setting $N_1^n=N_2^n$ does not affect the channel capacity. Therefore,

$$\begin{aligned}
&n(R_1+R_2)\\
\leqslant& h(Y_1^n)-h(X_0^n+S_1^n+N_1^n,X_0^n+S_2^n+N_1^n,X_0^n+N_1^n)\\
&+h(Y_2^n)+h(X_0^n+N_1^n)+n\epsilon_n\\
=&h(Y_1^n)-h(S_1^n,S_2^n,X_0^n+N_1^n)\\
&+h(Y_2^n)+h(X_0^n+N_1^n)+n\epsilon_n\\
=&h(Y_1^n)-h(S_1^n)-h(S_2^n)-h(X_0^n+N_1^n|S_1^n,S_2^n)\\
&+h(Y_2^n)+h(X_0^n+N_1^n)+n\epsilon_n\\
\leqslant& h(Y_1^n)-h(S_1^n)-h(S_2^n)-h(X_0^n+N_1^n|S_1^n,S_2^n,X_0^n)\\
&+h(Y_2^n)+h(X_0^n+N_1^n)+n\epsilon_n\\
\overset{(b)}{=}&h(Y_1^n)-h(S_1^n)-h(S_2^n)-h(N_1^n)\\
&+h(Y_2^n)+h(X_0^n+N_1^n)+n\epsilon_n\\
\overset{(c)}{=}&\sum_{i=1}^n[h(Y_{1i}|Y_1^{i-1})-h(S_{1i})-h(S_{2i})-h(N_{1i})\\
&+h(Y_{2i}|Y_2^{i-1})+h(X_{0i}+N_{1i}|X_0^{i-1}+N_1^{i-1})]+n\epsilon_n\\
\leqslant&\sum_{i=1}^n[h(Y_{1i})-h(S_{1i})-h(S_{2i})-h(N_{1i})\\
&+h(Y_{2i})+h(X_{0i}+N_{1i})]+n\epsilon_n\\
=&\sum_{i=1}^n[h(X_{0i}+X_{1i}+S_{1i}+N_{1i})-h(S_{1i})\\
&-h(S_{2i})-h(N_{1i})+h(X_{0i}+X_{2i}+S_{2i}+N_{1i})\\
&+h(X_{0i}+N_{1i})]+n\epsilon_n
\end{aligned}\tag{60}$$

We then derive the items respectively. The first term in (60) can be derived as

$$\begin{aligned}
&\sum_{i=1}^n h(X_{0i}+X_{1i}+S_{1i}+N_{1i})\\
\overset{(d)}{\leqslant}&\frac{1}{2}\sum_{i=1}^n\log 2\pi e(E(X_{0i}+X_{1i}+S_i+N_i)^2)\\
\leqslant&\frac{1}{2}\sum_{i=1}^n\log 2\pi e\Big(E[X_{0i}^2]+E(X_{0i}S_i)\\
&+E[S_i^2]+E[X_{1i}^2]+E[N_i^2]\Big)\\
\overset{(e)}{\leqslant}&\frac{n}{2}\log 2\pi e\left(\frac{1}{n}\sum_{i=1}^n E[X_{0i}^2]+\frac{2}{n}\sum_{i=1}^n E(X_{0i}S_i)\right.\\
&\left.+\frac{1}{n}\sum_{i=1}^n E[S_i^2]+\frac{1}{n}\sum_{i=1}^n E[X_{1i}^2]+\frac{1}{n}\sum_{i=1}^n E[N_i^2]\right)\\
\overset{(f)}{\leqslant}&\frac{n}{2}\log 2\pi e\left(P_0+Q+P_1+1+\frac{2}{n}\sum_{i=1}^n E(X_{1i}S_i)\right)\\
\leqslant&\frac{n}{2}\log 2\pi e\Big(P_0+P_1+Q+1+2\sqrt{P_0Q}\Big)
\end{aligned}\tag{61}$$

where (d) follows from the fact that the Gaussian distribution maximizes the entropy given the variance of the random variable, (e) follows from the concavity of the logarithm function and Jensen's inequality, and (f) follows from the power constraints. Similarly, we have

$$\begin{aligned}
&\sum_{i=1}^n h(X_{0i}+X_{2i}+S_{1i}+N_{1i})\\
&\qquad\leqslant\frac{n}{2}\log 2\pi e(P_2+P_0+2\sqrt{P_0Q_2}+Q_2+1)\\
&\sum_{i=1}^n h(X_{0i}+N_{1i})\leqslant\frac{n}{2}\log(P_0+1)
\end{aligned}$$

And hence, we have

$$\begin{aligned}
&n(R_1+R_2)\\
\leqslant&\frac{n}{2}\log 2\pi e(P_1+P_0+2\sqrt{P_0Q_1}+Q_1+1)\\
&-\frac{n}{2}\log(2\pi eQ_1)-\frac{n}{2}\log(2\pi eQ_2)-\frac{n}{2}\log(2\pi e)\\
&+\frac{n}{2}\log 2\pi e(P_2+P_0+2\sqrt{P_0Q_2}+Q_2+1)\\
&+\frac{n}{2}\log 2\pi e(P_0+1)+n\epsilon_n\\
\leqslant&\frac{n}{2}\log\left(\frac{P_1+P_0+2\sqrt{P_0Q_1}+Q_1+1}{Q_1}\right)\\
&+\frac{n}{2}\log\left(\frac{P_2+P_0+2\sqrt{P_0Q_2}+Q_2+1}{Q_2}\right)\\
&+\frac{n}{2}\log(P_0+1)+n\epsilon_n\\
\to&\frac{n}{2}\log(P_0+1)\quad as\ Q_1\to\infty,Q_2\to\infty
\end{aligned}$$

where (b) follows from the fact that X_0^n, S_1^n and S_2^n are independent from N_1^n.

Appendix F
Proof of Theorem 6

The proof contains two parts: 1. we first show that if $P_1 + P_2 \geqslant P_0 + 1$, then the sum capacity can be obtained; 2. we further characterize the time allocation parameters γ that achieves the sum capacity.

1. For a given P_0, we consider the following two cases.

a). If the power constraint satisfies $P_1 + P_2 = P_0 + 1$, by applying Proposition 11, and by setting $\gamma = \frac{P_1}{P_1+P_2}$, the point $(R_1, R_2) = (\frac{P_1}{2(P_1+P_2)} \log(1+P_0), \frac{P_2}{2(P_1+P_2)} \log(1+P_0))$ is achievable, which achieves the sum rate outer bound in Proposition 10.

b). If $P_1+P_2 \geqslant P_0+1$, we set the actual transmission power $\tilde{P}_1$ and $\tilde{P}_2$ of transmitters 1 and 2 to satisfy $\tilde{P}_1+\tilde{P}_2 = P_0+1$, $\tilde{P}_1 \leqslant P_1$ and $\tilde{P}_2 \leqslant P_2$. Then following a), the sum capacity is obtained.

2. In order for each transmitter to achieve the sum capacity during its own transmission slot, (40) together with (41a) and (41b) imply that

$$\frac{P_1}{\gamma} \geqslant P_0 + 1 \tag{62}$$

$$\frac{P_2}{1-\gamma} \geqslant P_0 + 1. \tag{63}$$

It is clear that (62) implies

$$\gamma \leqslant \frac{P_1}{P_0+1},$$

and (63) implies

$$\gamma \geqslant 1 - \frac{P_2}{P_0+1}.$$

Considering $0 \leq \gamma \leq 1$, we obtain the desired bounds on γ.

Appendix G
Proof of Proposition 14

The individual rate can be bounded by the point-to-point channel capacity. We then bound the sum rate. By following the Fano's inequality, we have

$$\begin{aligned}
&\sum_{k=1}^{K} nR_k \\
&\leqslant \sum_{k=1}^{K} I(W_k; Y_k^n) \\
&= \sum_{k=1}^{K} [h(Y_k^n) - h(Y_k^n|W_k)] \\
&\overset{(a)}{=} \sum_{k=1}^{K} [h(Y_k^n) - h(Y_k^n|W_k X_k^n)] \\
&= \sum_{k=1}^{K} [h(Y_k^n) - h(X_0^n + S_k^n + N_k^n)] \\
&\leqslant \sum_{k=1}^{K} [h(Y_k^n) - h(X_0^n + S_k^n + N_k^n | X_0^n + N_1^n, \\
&\quad X_0^n + S_{k-1}^n + N_{k-1}^n, \ldots, X_0^n + S_1^n + N_1^n)] \\
&\quad + h(X_0^n + N_1^n) - h(X_0^n + N_1^n) \\
&= \sum_{k=1}^{K} h(Y_k^n) \\
&\quad - h(X_0^n + S_1^n + N_1^n, \ldots, X_0^n + S_K^n + N_K^n, X_0^n + N_1^n) \\
&\quad + h(X_0^n + N_1^n)
\end{aligned}$$

where (a) follows from that X_k^n is function of W_k, and they are independent from X_0^n, state and noise. Because the decoders decode based on the marginal distribution only, we can set $N_k^n = N_1^n$ for $k = 1, \ldots, K$, therefore,

$$\begin{aligned}
&\sum_{k=1}^{K} nR_i \\
&\leqslant \sum_{k=1}^{K} h(Y_k^n) - h(S_1^n, \ldots, S_K^n, X_0^n + N_1^n) + h(X_0^n + N_1^n) \\
&= \sum_{k=1}^{K} [h(Y_k^n) - h(S_k^n)] \\
&\quad - h(X_0^n + N_1^n | S_1^n, \ldots, S_K^n) + h(X_0^n + N_1^n) \\
&\leqslant \sum_{k=1}^{K} [h(Y_k^n) - h(S_k^n)] \\
&\quad - h(X_0^n + N_1^n | S_1^n, \ldots, S_K^n, X_0^n) + h(X_0^n + N_1^n) \\
&\overset{(b)}{=} \sum_{k=1}^{K} [h(X_0^n + X_k^n + S_k^n + N_k^n) - h(S_k^n)] \\
&\quad - h(N_1^n) + h(X_0^n + N_1^n) \\
&\leqslant \sum_{k=1}^{K} \left[\frac{n}{2} \log 2\pi e \left(P_k + P_0 + 2\sum_{i=1}^{n} E(X_{0i}S_{ki}) + Q_k + 1 \right) \right. \\
&\quad \left. - \frac{n}{2} \log(2\pi e Q_k) \right] - \frac{n}{2} \log(2\pi e) + \frac{n}{2} \log 2\pi e(P_0 + 1) \\
&\leqslant \sum_{k=1}^{K} \frac{n}{2} \log \left(\frac{P_k + P_0 + 2\sqrt{P_0 Q_k} + Q_k + 1}{Q_k} \right) \\
&\quad + \frac{n}{2} \log(P_0 + 1) \\
&\to \frac{n}{2} \log(P_0 + 1) \text{ as } Q_k \to \infty
\end{aligned}$$

References

[1] S. Gel'fand and M. Pinsker, "Coding for channels with ramdom parameters," *Probl. Contr. Inf. Theory*, vol. 9, no. 1, pp. 19–31, January 1980.

[2] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[3] S. Gel'fand and M. Pinsker, "On Gaussian channels with random parameters," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Russian, Jun. 1984.

[4] R. Khosravi-Farsani and F. Marvasti, "Capacity bounds for multiuser channels with non-causal channel state information at the transmitters," in *Proc. IEEE Information Theory Workshop (ITW)*, Paraty , Brazil, Oct. 2011.

[5] A. Lapidoth and L. Wang, "The state-dependent semideterministic broadcast channel," *IEEE Trans. Inform. Theory*, vol. 59, no. 4, pp. 2242–2251, Apr. 2013.

[6] Y. Steinberg, "Correspondence coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 8, August 2005.

[7] Y. Steinberg and S. Shamai (Shitz), "Achievable rates for the broadcast channel with states known at the transmitter," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, July 2005.

[8] A. Zaidi, S. Shamai (Shitz), P. Piantanida, and L. Vandendorpe, "Bounds on the capacity of the relay channel with noncausal state at source," *IEEE Trans. Inform. Theory*, vol. 59, no. 5, pp. 2639–2672, May 2013.

[9] Y.-H. Kim, A. Sutivong, and S. Sigurjónsson, "Multiple user writing on dirty paper," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Chicago, Illinois, Jul. 2004.

[10] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verdú, "Cooperative multiple-access encoding with states available at one transmitter," *IEEE Trans. Inform. Theory*, vol. 54, no. 10, pp. 4448–4469, October 2008.

[11] L. Zhang, J. Jiang, and S. Cui, "Gaussian interference channel with state information," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4058–4071, Aug. 2013.

[12] R. Duan and Y. Liang, "Bounds and capacity theorems for cognitive interference channels with state," *IEEE Trans. Inform. Theory*, vol. 61, no. 1, pp. 280–304, Jan. 2015.

[13] R. Duan, Y. Liang, and S. Shamai (Shitz), "On the capacity region of Gaussian interference channels with state," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Istanbul, Turkey, Jul. 2013.

[14] S. Ghasemi-Goojani and H. Behroozi, "On the achievable rate-regions for state-dependent Gaussian interference channel," Available at http://arxiv.org/abs/1301.5535, submitted in January 2013.

[15] S. Mallik and R. Koetter, "Helpers for cleaning dirty papers," in *Proc. IEEE International ITG Conference on Source and Channel Coding (SCC)*, Ulm, Germany, Jan 2008.

[16] B. Akhbari, M. Mirmohseni, and M. R. Aref, "Compress-and-forward strategy for the relay channel with non-causal state information," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, Jul. 2009.

[17] A. Zaidi, S. P. Kotagiri, J. N. Laneman, and L. Vandendorpe, "Cooperative relaying with state available noncausally at the relay," *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2272–2298, May 2011.

[18] S. P. Kotagiri and J. N. Laneman, "Multiaccess channels with state known to some encoders and independent messages," *EURASIP Journal on Wireless Communications and Networking*, 2008.

[19] I.-H. Wang, "Distributed interference cancellation in multiple access channels," *IEEE Trans. Inform. Theory*, vol. 58, no. 5, pp. 2781–2787, May 2012.

[20] T. Philosof and R. Zamir, "The rate loss of single letter characterization for the "dirty" multiple access channel," in *Proc. IEEE Information Theory Workshop (ITW)*, May 2008, pp. 31–35.

[21] T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, "Lattice strategies for the dirty multiple access channel," *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 5006–5035, August 2011.

[22] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. 27, no. 1, pp. 49–60, January 1981.

[23] S. Hajizadeh, M. Monemizadeh, E. Bahmani, G. A. Hodtani, and M. Joneidi, "State-dependent Z channel," Available at http://arxiv.org/abs/1301.6272.

[24] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inform. Theory*, vol. 50, no. 6, pp. 1345–1356, June 2004.

[25] P. Mitran, N. Devroye, and V. Tarokh, "On compound channels with side-information at the transmitter," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1745–1755, April 2006.

[26] A. J. Khisti, U. Erez, A. Lapidoth, and G. Wornell, "Carbon copying onto dirty paper," *IEEE Trans. Inform. Theory*, vol. 53, no. 5, pp. 1814–1827, May 2007.

[27] P. Piantanida and S. Shamai (Shitz), "On the capacity of compound state-dependent channels with states known at the transmitter," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, June 2010, pp. 624 –628.

[28] A. Khisti, S. Diggavi, and G. Wornell, "Secret-key agreement with channel state information at the transmitter," *IEEE Trans. on Forensics and Security, Special Issue on Physical Layer Security*, vol. 6, no. 3, pp. 672 – 681, June 2011.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory, Second Edition*. New York: Wiley, 2006.

[30] G. Kramer, *Topics in Multi-User Information Theory*. Hanover, MA: Now Publishers Inc, 2008.

[31] A. El Gamal and Y.-H. Kim, *Network Information Theory*. New York: Cambridge University Press, 2011.

Ruchen Duan (S'11, M'15) received the B.E. degree from Beijing University of Posts and Telecommunications, Beijing, China in 2010, and the Ph.D. degree in Electrical Engineering from Syracuse Universit in 2015. She received the Syracuse University Fellowship Award for years 2010-2012. She is now a senior engineer at Samsung Semiconductor Inc.. Ruchen's research interests focus on information theory and wireless communications.

Yingbin Liang (S'01, M'05) received the Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2005. In 2005-2007, she was working as a postdoctoral research associate at Princeton University. In 2008-2009, she was an assistant professor at the Department of Electrical Engineering at the University of Hawaii. Since December 2009, she has been on the faculty at Syracuse University, where she is an associate professor. Dr. Liang's research interests include information theory, wireless communications and networks, and machine learning.

Dr. Liang was a Vodafone Fellow at the University of Illinois at Urbana-Champaign during 2003-2005, and received the Vodafone-U.S. Foundation Fellows Initiative Research Merit Award in 2005. She also received the M. E. Van Valkenburg Graduate Research Award from the ECE department, University of Illinois at Urbana-Champaign, in 2005. In 2009, she received the National Science Foundation CAREER Award, and the State of Hawaii Governor Innovation Award. More recently, her paper received the 2014 EURASIP Best Paper Award for the EURASIP Journal on Wireless Communications and Networking. She is currently serving as an Associate Editor for the Shannon Theory of the IEEE Transactions on Information Theory.

Ashish Khisti (M'08) is an Associate Professor at the University of Toronto and holds a Canada Research Chair in Wireless Networks. He obtained his BASc degree from the Engineering Sciences program (Electrical Engineering Option) at the same university in 2002, and his SM and PhD degrees from the Massachusetts Institute of Technology (MIT) in Electrical Engineering and Computer Science in 2004 and 2009 respectively. His research interests include Information Theory, Physical Layer Security and Error Control Coding for Multimedia Applications. He also actively consults telecommunication companies. He is a recipient of the HP-IRP award from Hewlett-Packard and an Ontario Early Researcher Award. He presently serves as a Associate Editor in Shannon Theory for the IEEE Transactions on Information Theory.

Shlomo Shamai (Shitz) (S'80, M'82, SM'88, F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975-1985 he was with the Communications Research Labs, in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now a Technion Distinguished Professor, and holds the William Fondiller Chair of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is an IEEE Fellow, a member of the Israeli Academy of Sciences and Humanities and a foreign member of the US National Academy of Engineering. He is the recipient of the 2011 Claude E. Shannon Award and the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering.

He has been awarded the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and is a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 and 2015 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++, NEWCOM#) Best Paper Awards, the 2010 Thomson Reuters Award for International Excellence in Scientific Research, the 2014 EURASIP Best Paper Award (for the EURASIP Journal on Wireless Communications and Networking), and the 2015 IEEE Communications Society Best Tutorial Paper Award. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE Transactions on Information Theory, and has also served twice on the Board of Governors of the Information Theory Society. He has served on the Executive Editorial Board of the IEEE Transactions on Information Theory