Capacity Characterization for State-Dependent Gaussian Channel with a Helper [1]

Yunhao Sun, [2] Ruchen Duan, [3] Yingbin Liang, [4] Ashish Khisti,[5] Shlomo Shamai (Shitz)[6]

Abstract

The state-dependent point-to-point Gaussian channel with a helper is first studied, in which a transmitter communicates with a receiver via a state-corrupted channel. The state is not known to the transmitter nor to the receiver, but known to a helper noncausally, which then wishes to assist the receiver to cancel the state. Differently from previous work that characterized the capacity only in the infinite state power regime, this paper explores the general case with arbitrary state power. A lower bound on the capacity is derived based on an achievable scheme that integrates direct state subtraction and single-bin dirty paper coding. By analyzing this lower bound and further comparing it with the existing upper bounds, the capacity of the channel is characterized for a wide range of channel parameters. Such an idea of characterizing the capacity is further extended to study the two-user state-dependent multiple access channel (MAC) with a helper. By comparing the derived inner and outer bounds, the channel parameters are partitioned into appropriate cases, and for each case, either segments on the capacity region boundary or the full capacity region are characterized.

1 Introduction

A type of helper-assisted state-dependent models have been an active research topic recently. The basic point-to-point model (see Figure 2) was studied in [1], in which a transmitter wishes to send the message W to a receiver over the state-corrupted channel, and a helper that knows the state information noncausally wishes to assist the receiver to cancel state

---

[1]The work of Y. Sun, 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++.

[2]Yunhao Sun is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (email: ysun33@syr.edu).

[3]Ruchen Duan is with Samsung Semiconductor Inc., San Diego, CA 92121 USA (email: r.duan@samsung.com).

[4]Yingbin Liang is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (email: yliang06@syr.edu).

[5]Ashish Khisti is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S3G4, Canada (email: akhisti@comm.utoronto.ca).

[6]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).

interference. The state information is not known at the transmitter nor at the receiver. Here, the transmitter that needs to send the message does not know the state, whereas the helper that knows the state does not know the message. Such a mismatched property differentiates this channel from the traditional state-dependent channel studied in [2, 3], where the transmitter knows both the message and the state. Such a model serves as a building block for a number of more general channel models studied later on, including the state-dependent multiple-access channel (MAC) [4, 5], its further generalization [6, 7], the state-dependent Z-interference channel [8], and the state-dependent relay channel [9, 10].

In [1], lower and upper bounds on the capacity for the point-to-point model were derived with the lower bound based on lattice coding. However, the capacity was established only in the asymptotic regime as the state power goes to infinity. Moreover, further studies of more general models did not provide further capacity results beyond the infinite state power regime, when the models are specialized to the point-to-point model here. Thus, the capacity in the regime with finite state power is left as an open problem for this type of state-dependent channels with a helper [7].

The focus in this paper is on the finite state power regime. Our main contribution lies in characterization of the capacity of the point-to-point model for a wide range of channel parameters, and demonstration of applicability of our idea to characterize the capacity region of the MAC model either partially or fully for all parameter regimes. There are two major challenges here: (1) the achievable schemes proposed previously may not be sufficiently good for finite state power regime although they are optimal for infinite state power regime; and (2) the derived lower bounds on the capacity tend to have complicated form to capture correlation of the helper's input and the state, and hence are difficult to analyze and compare with upper bounds which may also have complicated form as in [1, 7].

For the point-to-point model with a helper, our achievable scheme is based on integration of single-bin dirty paper coding and direct state subtraction (i.e., the helper directly cancels partial state in the received output) with optimal trade-off between the two schemes. Such a scheme is equivalent to the generalized dirty paper coding used in [10] for the state-dependent relay channel, which assumes that the relay input and state are correlated. A lower bound on the capacity is derived based on such a scheme, which takes a complicated form and involves various parameters (i.e., dirty paper parameter and parameter capturing trade-off between two schemes) to be optimized. Our major novelty lies in identifying two special cases to analytically optimize the lower bound so that the optimizing lower bound matches either the upper bound in [1] or the capacity of the channel without state for various channel parameters. We thus establish the capacity under these channel parameters.

Our capacity result can be summarized as follows. If the helper's power is relatively small (compared to the transmitter's power and state power), then the capacity is characterized as a function of the state power, the helper's power and the transmitter's power. In particular, the capacity is strictly less than the capacity of the channel without state, which implies that there exists no achievable scheme that fully cancels the state interference. Here, direct

[7]In fact, the interesting case in the finite power regime is when the helper's power is less than the state power. Otherwise, it is straightforward to cancel the state by setting the helper's signal to reverse the state.

state subtraction is necessary for the achievable scheme to be optimal. On the other hand, if the helper's power is larger than a threshold, then the channel achieves the capacity of the channel without state, which implies that the state can be fully canceled. Here, single-bin dirty paper coding is optimal and direct state subtraction is not necessary. Such characterization of the capacity reduces to the capacity result for infinite state power regime obtained in the previous studies [1, 5, 7, 8, 10].

We then generalize our analysis to the state-dependent Gaussian MAC model with a helper, in which two transmitters send two independent messages to a common state-corrupted receiver. The state sequence is not known at the transmitters nor at the receiver, but is noncausally known to a helper, which wishes to assist the receiver to cancel the state. We first derive an outer bound on the capacity region, which not only consists of the natural outer bound given by the capacity region of the channel without state, but also consists of bounds that capture the impact of the helper's power and the state power on the transmission rates, similarly to the upper bound in [1] for the point-to-point channel. We then derive an inner bound based on a scheme that integrates direct state cancelation and single-bin dirty paper coding. Since the inner and outer bounds are both characterized in complicated forms, direct comparison of the two regions is challenging. We thus first compare the bounds on the individual rates and sum rate separately, and characterize conditions on the channel parameters such that each individual rate and the sum rate separately achieves its corresponding outer bound. In this way, we characterize separate segments on the capacity region boundary. Intersections of conditions for the individual and sum rates then collectively characterize channel parameters under which multiple segments on the capacity boundary are obtained. Based on such an idea, we partition the channel parameters into appropriate cases, and characterize segments on the capacity region boundary for all these cases. Among these cases, we characterize the full capacity region for one case, which achieves the capacity region of the Gaussian MAC without state, i.e., the state is fully canceled. Such a case suggests that if the helper's power is large enough, then the state can always be canceled. However, if the helper's power is below a threshold, the helper can assist to fully cancel the state only when the state power is small enough.

1.1 Related Work

As mentioned earlier, the model we study was initially studied in [1]. A number of more general models were then further studied, which include the channel of interest here as a special case. More specifically, in [4, 5], the state-dependent multiple-access channel (MAC) was studied, which can be viewed as the model here with the helper also having its own message to the receiver. Two more general state-dependent MACs were studied in [6] and [7], which can be viewed as the MAC model in [4, 5] respectively with the helper further knowing the transmitter's message and with one more state corruption known at the transmitter. In [8], the state-dependent Z-interference channel was studied, which can be viewed as the model here with the helper also having a message to its own receiver. In [9, 10], the state-dependent relay channel was studied, which can be viewed as the model here with the helper also receiving information from the transmitter and serving as a relay. When these models reduce to the model here, the results in [5, 7, 8, 10] characterize the capacity of the Gaussian channel as the state power goes to infinity as in [1]. In particular, the achievable scheme in [7] is based on lattice coding similar to [1], and the scheme in [5, 8, 10] can be viewed as single-bin dirty paper coding (i.e., a special case of dirty paper coding [2, 3] with only one bin). In this paper, our focus is to characterize the capacity for the finite state power regime.

Various state-dependent MAC models were studied previously, which are related but different from the MAC model with a helper studied in this paper. State-dependent MAC models with state causally or strictly causally known at the transmitter were studied in [11–15], whereas our model assumes that the state is noncausally known at the helper. The two-user MAC with state noncausally known at the transmitters has been previously studied in various cases. [16, 17] studied the MAC model with state noncausally known at both transmitters, while [4, 5] assumed that the state is known only to one transmitter. [6] studied the cognitive MAC model in which one transmitter also knows the other transmitter's message in addition to the noncausal state information. Furthermore, [7, 18] studied the model with the receiver being corrupted by two independent states and each state is known noncausally to one transmitter. In all these two-user MAC models with noncausal state information, at least one transmitter knows the state information, and can hence encode messages by incorporating the state information. Our MAC model is different in that only an additional helper knows the state information and assists to cancel the state.

1.2 Practical Motivation

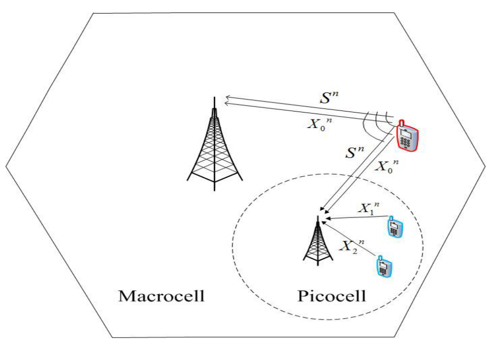

Figure 1: A practical example of the MAC model with a helper.

The type of state-dependent models with a helper can arise from a new perspective of interference cancelation in wireless networks. The idea can be illustrated via a simple MAC example (see Fig. 1). Consider multiple-access communications in a picocell located inside a macrocell of a cellular network. It is typical that a macrocell user causes interference to the picocell users. The macrocell user itself knows the interference that it causes to the picocell users noncausally because such interference is in fact the signal that this user sends to its own receiver (i.e., the base station in the macrocell). Thus, the interference is referred to as dirty interference (i.e., the noncausal state sequence in our model) and is denoted as S^n in Fig. 1. The macrocell user is then able to exploit such interference (i.e., state) information and send a help signal (denoted by X_0 in Fig. 1) to assist the picocell receiver to cancel the interference. In this way, a user can assist to cancel the interference that itself causes to other users by exploiting its knowledge about the interference. Although the help signal X_0 may also cause interference to the macrocell base station, as long as the power of X_0 is much less than the power of S, there is still significant gain in throughput. In fact, our results in this paper demonstrate that the interfering user can use a relatively small amount of power to completely cancel the interference that it causes to other users (e.g., the picocell users in our previous example) even if the interference is as large as infinite.

1.3 Organization

The rest of the paper is organized as follows. In Section 2, we present the point-to-point channel model and our characterization of the capacity for various channel parameters of this channel. In Section 3, we present the MAC model and our characterization of the capacity for various parameter regimes of this channel. In Section 4, we conclude the paper with several remarks.

2 Point-to-Point Channel with a Helper

2.1 Channel Model

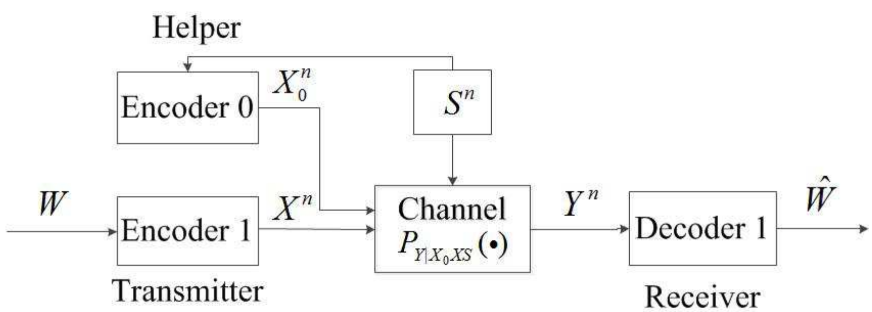

Figure 2: The state-dependent channel with a helper

We consider the state-dependent channel with a helper (see Figure 2), in which a transmitter sends a message to a receiver over the state-dependent channel, and a helper that knows the state sequence noncausally wishes to assist the transmission by canceling the state. More specifically, the transmitter has an encoder $f : \mathcal{W} \to \mathcal{X}^n$, which maps a message $w \in \mathcal{W}$ to a codeword $x^n \in \mathcal{X}^n$. The input x^n is transmitted over the channel, which is corrupted by an independent and identically distributed (i.i.d.) state sequence S^n. The state sequence is assumed to be known at neither the transmitter nor the receiver, but at a helper noncausally. Thus, the encoder at the helper, $f_0 : \mathcal{S}^n \to \mathcal{X}_0^n$, maps a state sequence $s^n \in \mathcal{S}^n$ to a codeword $x_0^n \in \mathcal{X}_0^n$ and sends it over the channel. The channel is characterized by the transition probability distribution $P_{Y|X_0,X,S}$. The decoder at the receiver, $g : \mathcal{Y}^n \to \mathcal{W}$, maps a received sequence y^n into a message $\hat{w} \in \mathcal{W}$.

We assume that the message is uniformly distributed over the set $\mathcal{W}$, and define the average probability of error for a length-n code as follows.

$$P_e = \frac{1}{|\mathcal{W}|} \sum_{w=1}^{|\mathcal{W}|} Pr\{\hat{w} \neq w\}. \tag{1}$$

A rate R is *achievable* if there exist a sequence of message sets $\mathcal{W}^{(n)}$ with $|\mathcal{W}^{(n)}| = 2^{nR}$ and encoder-decoder tuples $(f_0^{(n)}, f^{(n)}, g^{(n)})$ such that the average probability of error $P_e^{(n)} \to 0$ as $n \to \infty$. We define the *capacity* of the channel to be supremum of all achievable rates R.

In this paper, we focus on the state-dependent Gaussian channel, with input-output relationship for one channel use given by

$$Y = X_0 + X + S + N \tag{2}$$

where the noise variable N and the state variable S are Gaussian distributed with distributions $N \sim \mathcal{N}(0, 1)$ and $S \sim \mathcal{N}(0, Q)$, and both variables are i.i.d. over channel uses. The channel inputs X_0 and X are subject to the average power constraints

$$\frac{1}{n} \sum_{i=1}^{n} X_{0i}^2 \leqslant P_0 \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^{n} X_i^2 \leqslant P. \tag{3}$$

2.2 Achievable Scheme and Lower Bound

We adopt an achievable scheme that integrates (1) precoding state into a help signal using single-bin dirty paper coding and (2) directly subtracting state. The single-bin dirty paper coding is a special case of dirty paper coding with only one bin, because the bin number corresponds to the message index in dirty paper coding, and here the helper does not know the message to be sent. It has been shown in previous studies that only single-bin dirty paper coding is sufficient to achieve the capacity in the infinite state power regime. This is reasonable because direct state subtraction is not useful when the state power is infinite. However, for the finite state power regime, direct state subtraction can be more efficient and hence should be included in the achievable scheme. In order to achieve the best performance, the achievable scheme should include the two coding schemes with the best trade-off, which results the following achievable rate.

Proposition 1. *For the state-dependent Gaussian channel with a helper, the following rate is achievable*

$$R \leqslant \max_{\substack{(\alpha,\beta)\ s.t. \\ -\sqrt{\frac{P_0}{Q}} \leqslant \beta \leqslant \sqrt{\frac{P_0}{Q}}}} \min\{R_1(\alpha, \beta), R_2(\alpha, \beta)\}, \tag{4}$$

where

$$R_1(\alpha,\beta) = \frac{1}{2}\log\frac{P_0'(P_0'+(1+\beta)^2Q+P+1)}{P_0'Q(\alpha-1-\beta)^2+P_0'+\alpha^2Q}, \tag{5a}$$

$$R_2(\alpha,\beta) = \frac{1}{2}\log\left(1+\frac{P(P_0'+\alpha^2Q)}{P_0'Q(\alpha-1-\beta)^2+P_0'+\alpha^2Q}\right), \tag{5b}$$

and $P_0' = P_0 - \beta^2 Q$.

Proposition 1 is consistent with the achievable rate derived for the state-dependent relay channel in [10] with the noise power on the source to relay link set to infinity. However, the rate expression in [10] is much more complicated due to existence of the relay. It appears to mask the crucial elements required in obtaining the capacity results in the present model. For this reason, we provide a simple proof of achievability that emphasizes the interplay between state subtraction and dirty-paper coding. We would also like to note that the optimality results presented herein (in Section 2.3) are new, and were not known previously to the best of our knowledge.

Proof. We first derive an achievable rate based on single-bin Gel'fand-Pinsker binning scheme for the discrete memoryless state-dependent channel. For a given distribution P_{US}, a number of u^n is generated using the marginal distribution P_U, so that for any s^n, there exists a u^n that is jointly typical with s^n. The helper's input x_0^n is then created based on $P_{X_0|SU}$. The transmitter's input x^n is created based on P_X. The receiver jointly decodes both u^n and x^n. The following lemma characterizes the achievable rate based on the above scheme.

Lemma 1. *For the state-dependent point-to-point channel with a helper, the following rate is achievable*

$$R \leqslant \min\{I(UX;Y) - I(U;S), I(X;Y|U)\} \tag{6}$$

for some distribution $P_{X_0US}P_XP_{Y|X_0XS}$.

The above rate can also be derived from [8, Proposition 2] by setting $X_0' = \phi$. Detailed proof is relegated to Appendix A.

Proposition 1 then follows by evaluating the mutual information terms in (6) based on the following joint Gaussian distribution for the random variables

$$\begin{aligned} X_0 &= X_{00} + \beta S \\ U &= X_{00} + \alpha S \end{aligned} \tag{7}$$

where X_{00} is independent of S and $X_{00} \sim \mathcal{N}(0, P_0')$ with $P_0' = P_0 - \beta^2 Q$ and $-\sqrt{\frac{P_0}{Q}} \leqslant \beta \leqslant \sqrt{\frac{P_0}{Q}}$. □

We note that in (7), the helper's input X_0 contains two parts with X_{00} designed using single-bin dirty paper coding, and βS serving for state subtraction. The parameter β captures the trade-off between the two schemes. Furthermore, the achievable rate in (6) can be intuitively understood as follows. The first term

$$I(U, X; Y) - I(U; S) = I(U, X; Y) - I(U, X; S),$$

where (U, X) play the role of the auxiliary variable in Gel'fand-Pinsker scheme. The second term

$$I(X; Y|U) = I(U, X; Y) - I(U; Y)$$

can be interpreted as coding via (U, X) but paying the price needed to convey U to the receiver.

The achievable rate in Proposition 1 is optimized over α and β. The optimization is a max-min problem, i.e., maximization of minimum of $R_1(\alpha, \beta)$ and $R_2(\alpha, \beta)$. In general, such optimization cannot be solved analytically with close-form expressions. In order to obtain further insights of such a lower bound, we consider two special cases in which the optimization is solved analytically and the corresponding achievable rate turns out to achieve the capacity as we present in Section 2.3. The idea is to optimize $R_1(\alpha, \beta)$ and $R_2(\alpha, \beta)$ separately. For example, when $R_1(\alpha, \beta)$ is optimized, if $R_2(\alpha, \beta)$ at the optimizing values of α and β is greater than the optimal $R_1(\alpha, \beta)$, then the corresponding optimal $R_1(\alpha, \beta)$ is achievable. The same argument is applicable to optimizing $R_2(\alpha, \beta)$ instead. Such an idea yields the following two corollaries on the achievable rate.

Corollary 1. *For the state-dependent Gaussian channel with a helper, the following rate R is achievable*

$$R = \max_{-1 \leqslant \rho_{0S} \leqslant 1} \min\{R_1(\rho_{0S}), R_2(\rho_{0S})\} \tag{8}$$

where

$$R_1(\rho_{0S}) = \frac{1}{2} \log \left(1 + \frac{P}{Q + 2\rho_{0S}\sqrt{P_0 Q} + P_0 + 1}\right) + \frac{1}{2} \log(1 + P_0 - \rho_{0S}^2 P_0) \tag{9a}$$

$$R_2(\rho_{0S}) = \frac{1}{2} \log \left(1 + \frac{P((1 + P_0(1 - \rho_{0S}^2))^2 + (1 - \rho_{0S}^2)P_0(\sqrt{Q} + \rho_{0S}\sqrt{P_0})^2)}{(Q + 2\rho_{0S}\sqrt{P_0 Q} + P_0 + 1)(1 + P_0 - \rho_{0S}^2 P_0)}\right). \tag{9b}$$

Proof. It can be shown that $R_1(\alpha, \beta)$ is optimized by $\alpha = \frac{(1+\beta)P_0'}{P_0'+1}$. We further set $\beta = \rho_{0S}\sqrt{\frac{P_0}{Q}}$ to better illustrate the result, where $-1 \leqslant \rho_{0S} \leqslant 1$. Corollary 1 then follows by substituting α and β into (5a) and (5b). □

Corollary 2. *For the state-dependent Gaussian channel with a helper, the following rate R is achievable*

$$R = \min\left\{\frac{1}{2} \log \frac{P_0'(P_0' + \alpha^2 Q + P + 1)}{P_0' + \alpha^2 Q}, \frac{1}{2} \log(1 + P)\right\}. \tag{10}$$

for some $\alpha \in \Omega_\alpha = \{\alpha : 1 - \sqrt{\frac{P_0}{Q}} \leq \alpha \leq 1 + \sqrt{\frac{P_0}{Q}}\}$.

Proof. It can be shown that $R_2(\alpha, \beta)$ is optimized by setting $\beta = \alpha - 1$. Corollary 2 then follows by substituting β into (5a) and (5b). □

2.3 Capacity Characterization

In order to characterize the capacity, we first present two useful upper bounds on the capacity. In [1], the following upper bound on the capacity was derived.

Lemma 2. *The capacity of the state-dependent Gaussian channel with a helper is upper bounded as*

$$C \leqslant \max_{-1\leqslant\rho_{0S}\leqslant 1} \frac{1}{2}\log\left(1+\frac{P}{Q+2\rho_{0S}\sqrt{P_0Q}+P_0+1}\right)+\frac{1}{2}\log(1+P_0-\rho_{0S}^2P_0). \tag{11}$$

It is also clear that the capacity of the channel between the transmitter and receiver without state serves as an upper bound on the capacity of the state-dependent channel.

Lemma 3. *The capacity of the state-dependent Gaussian channel with a helper is upper bounded as*

$$C \leqslant \frac{1}{2}\log(1+P). \tag{12}$$

By comparing the achievable rate in Corollary 1 with the upper bound in Lemma 2, we characterize the capacity for various channel parameters in the following theorem.

Theorem 1. *For the state-dependent Gaussian channel with a helper, suppose ρ_{0S}^* maximizes $R_1(\rho_{0S})$ in* (9a)*. If the channel parameters satisfy the following condition:*

$$R_1(\rho_{0S}^*) \leqslant R_2(\rho_{0S}^*), \tag{13}$$

where $R_2(\rho_{0S})$ is given in (9b)*, then the channel capacity $\mathcal{C} = R_1(\rho_{0S}^*)$.*

Proof. Due to Corollary 1 and the condition (13), $R_1(\rho_{0S}^*)$ is achievable. Since such an achievable rate matches the upper bound in Lemma 2, it is thus the capacity of the channel. □

We note that for channels that satisfy the condition (13), the capacity $R_1(\rho_{0S}^*)$ is less than the capacity of the channel without the state. Thus, in such cases, the state interference cannot be fully canceled by any scheme.

Furthermore, by comparing the achievable rate in Corollary 2 with the upper bound in Lemma 3, we further characterize the capacity for an additional set of channel parameters.

Theorem 2. *For the state-dependent Gaussian channel with a helper, if the channel parameters satisfy the following condition:*

$$P_0'^2 \geq \alpha^2 Q(P+1-P_0') \tag{14}$$

where $P_0' = P_0 - (\alpha-1)^2Q$ holds for some $\alpha \in \Omega_\alpha = \{\alpha : 1-\sqrt{\frac{P_0}{Q}} \leq \alpha \leq 1+\sqrt{\frac{P_0}{Q}}\}$, then the channel capacity $\mathcal{C} = \frac{1}{2}\log(1+P)$.

Proof. Due to Corollary 2 and the condition (14), the rate $\frac{1}{2}\log(1+P)$ is achievable. Since such an achievable rate matches the upper bound in Lemma 3, it is thus the capacity of the channel. $\square$

It is clear that under the condition (14), the state-dependent channel achieves the capacity of the channel without state. Thus, the state can be fully cancelled even if the state-cognitive node (i.e., the helper) does not know the message.

We further note that as the state power Q goes to infinity, Theorems 1 and 2 collectively characterize the capacity established in the previous studies [1, 5, 7, 8, 10].

2.4 Numerical Result

In this section, we demonstrate our characterization of the capacity via numerical plots.

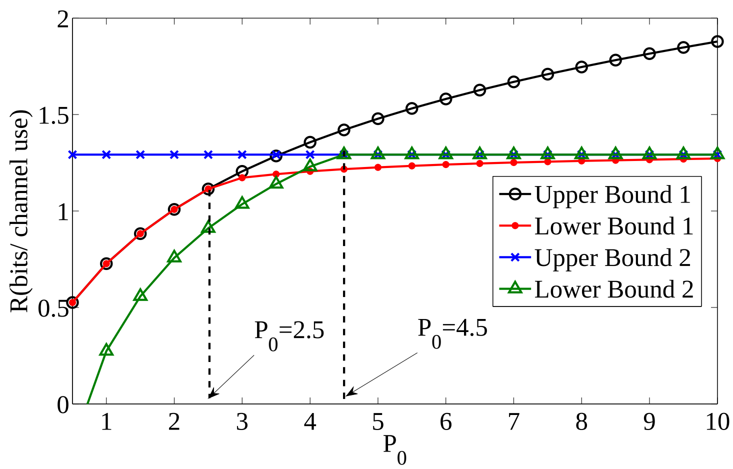

Figure 3: Lower and upper bounds on the capacity for the state-dependent channel with a helper

In Fig. 3, we fix $P = 5$, and $Q = 12$, and plot the lower bounds in Corollaries 1 and 2 and the upper bounds in Lemmas 2 and 3 as functions of the helper's power P_0. It can be seen that the lower bound 1 in Corollary 1 matches the upper bound 1 in Lemma 2 when $P_0 \le 2.5$, which corresponds to the capacity characterization in Theorem 1, and the lower bound 2 in Corollary 2 matches the upper bound 2 in Lemma 3 when $P_0 \ge 4.5$, which corresponds to the capacity characterization in Theorem 2. The numerical result also suggests that when P_0 is small, the channel capacity is limited by the helper's power and increases as the helper's power P_0 increases. However, as P_0 becomes large enough, the channel capacity is determined only by the transmitter's power P, in which case the state is perfectly canceled. We further note that the channel capacity without state can even be achieved when $P_0 < Q$ (e.g., $4.5 \le P_0 \le 10$). This implies that for these cases, the state is fully cancelled not only by state subtraction, but also by precoding the state via single-bin dirty paper coding. We finally note that a better achievable rate can be achieved by the convex envelop of the two lower bounds, which does not yield further capacity result and is not shown in Fig. 3.

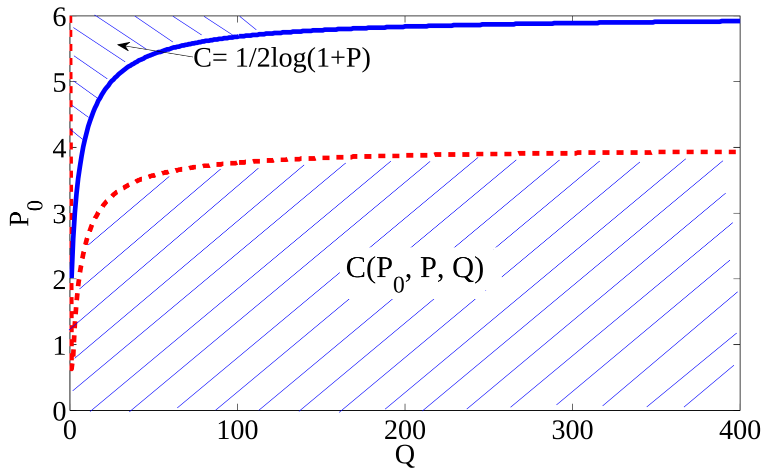

Figure 4: Ranges of parameters for which the capacity is characterized

In Fig. 4, we fix $P = 5$, and plot the range of the channel parameters (Q, P_0) for which we characterize the capacity. Each point in the figure corresponds to one parameter pair (Q, P_0). The upper shaded area corresponds to channel parameters that satisfy (14), i.e., P_0 is large enough compared to Q, and hence the capacity of the channel without state can be achieved. The lower shaded area corresponds to channel parameters that satisfy (13), and hence the capacity is characterized by a function of not only P, but also P_0 and Q.

3 MAC with a Helper

3.1 Channel Model

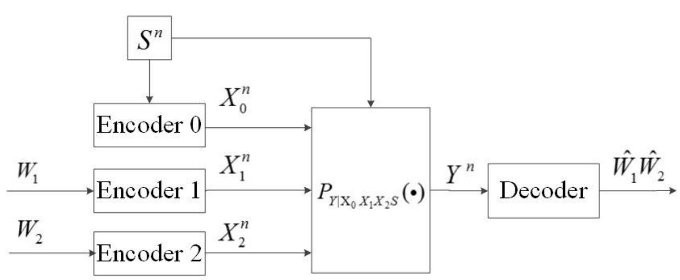

Figure 5: The state-dependent MAC with a helper

We consider the state-dependent MAC with a helper (as shown in Fig. 5), in which transmitter 1 sends a message W_1, and transmitter 2 sends a message W_2 to the receiver. The encoder $f_k : \mathcal{W} \to \mathcal{X}_k^n$ at transmitter k maps a message $w_k \in \mathcal{W}_k$ to a codeword $x_k^n \in \mathcal{X}_k^n$ for $k = 1, 2$. The two inputs x_1^n and x_2^n are transmitted over the MAC to a receiver, which is corrupted by an i.i.d. state sequence S^n. The state sequence is known to neither the transmitters nor the receiver, but is known to a helper noncausally. Hence, the helper assists the receiver to cancel the state interference. The encoder $f_0 : \mathcal{S}^n \to \mathcal{X}_0^n$ at the helper maps the state sequence $s^n \in \mathcal{S}^n$ into a codeword $x_0^n \in \mathcal{X}_0^n$. The channel transition probability is given by $P_{Y|X_0X_1X_2S}$. The decoder $g : \mathcal{Y}^n \to (\mathcal{W}_1, \mathcal{W}_2)$ at the receiver maps the received sequence y^n into two messages $\hat{w}_k \in \mathcal{W}_k$ for $k = 1, 2$.

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

$$P_e^{(n)} = \frac{1}{|\mathcal{W}_1||\mathcal{W}_2|} \sum_{w_1=1}^{|\mathcal{W}_1|} \sum_{w_2=1}^{|\mathcal{W}_2|} Pr\{(\hat{w}_1, \hat{w}_2) \neq (w_1, w_2)\}. \tag{15}$$

A rate pair (R_1, R_2) is *achievable* if there exist a sequence of message sets $\mathcal{W}_k^{(n)}$ with $|\mathcal{W}_k^{(n)}| = 2^{nR_k}$ for $k = 1, 2$, and encoder-decoder tuples $(f_0^{(n)}, f_1^{(n)}, f_2^{(n)}, g^{(n)})$ such that the average error probability $P_e^{(n)} \to 0$ as $n \to \infty$. We define the *capacity region* to be the closure of the set of all achievable rate pairs (R_1, R_2).

We focus on the state-dependent Gaussian channel with the output at the receiver for one channel use given by

$$Y = X_0 + X_1 + X_2 + S + N \tag{16}$$

where the noise variables $N \sim \mathcal{N}(0, 1)$, and $S \sim \mathcal{N}(0, Q)$. Both the noise variables and the state variable are i.i.d. over channel uses. The channel inputs X_0, X_1 and X_2 are subject to the average power constraints P_0, P_1 and P_2.

Our goal is to characterize the capacity region of the Gaussian channel under various channel parameters (P_0, P_1, P_2, Q).

3.2 Outer and Inner Bounds

We first provide an outer bound on the capacity region as follows, in which the first terms in the "min" improve the corresponding bounds give in [19].

Proposition 2. *An outer bound on the capacity region of the state-dependent Gaussian*

MAC with a helper consists of rate pairs (R_1, R_2) satisfying:

$$R_1 \leqslant \min\Big\{\frac{1}{2}\log(1+\frac{P_1}{Q+2\rho_{0S}\sqrt{P_0Q}+P_0+1})+\frac{1}{2}\log(1+P_0-\rho_{0S}^2P_0),$$
$$\frac{1}{2}\log(1+P_1)\Big\} \tag{17a}$$

$$R_2 \leqslant \min\Big\{\frac{1}{2}\log(1+\frac{P_2}{Q+2\rho_{0S}\sqrt{P_0Q}+P_0+1})+\frac{1}{2}\log(1+P_0-\rho_{0S}^2P_0),$$
$$\frac{1}{2}\log(1+P_2)\Big\} \tag{17b}$$

$$R_1+R_2 \leqslant \min\Big\{\frac{1}{2}\log(1+\frac{P_1+P_2}{Q+2\rho_{0S}\sqrt{P_0Q}+P_0+1})+\frac{1}{2}\log(1+P_0-\rho_{0S}^2P_0),$$
$$\frac{1}{2}\log(1+P_1+P_2)\Big\} \tag{17c}$$

for some ρ_{0S} that satisfies $-1 \leqslant \rho_{0S} \leqslant 1$.

Proof. See Appendix B. □

The second terms in the "min" in (17a)-(17c) capture the capacity region of the Gaussian MAC without state. If these bounds dominate the outer bound, then it is possible to design achievable schemes to fully cancel the state. Otherwise, if the first terms in the "min" in (17a)-(17c) dominate the outer bound, then the state cannot be fully canceled by any scheme, and the capacity region of the state-dependent MAC is smaller than that of the MAC without state.

We next derive an achievable region for the channel based on an achievable scheme that integrates direct state cancelation and single-bin dirty paper coding. In particular, since the helper does not know the messages, dirty paper coding naturally involves only one bin. More specifically, an auxiliary random variable (represented by U in Proposition 3) is generated to incorporate the state information so that the receiver decodes such variable first to cancel the state and then decode the transmitters' information. Based on such an achievable scheme, we derive the following inner bound on the capacity region.

Proposition 3. *For the discrete memoryless state-dependent MAC with a helper, an inner bound on the capacity region consists of rate pairs (R_1, R_2) satisfying:*

$$R_1 \le \min\{I(X_1;Y|X_2,U),\ I(U,X_1;Y|X_2)-I(U;S)\} \tag{18a}$$
$$R_2 \le \min\{I(X_2;Y|X_1,U),\ I(U,X_2;Y|X_1)-I(U;S)\} \tag{18b}$$
$$R_1+R_2 \le \min\{I(X_1,X_2;Y|U),\ I(U,X_1,X_2;Y)-I(U;S)\} \tag{18c}$$

for some distribution $P_SP_{U|S}P_{X_0|US}P_{X_1}P_{X_2}P_{Y|SX_0X_1X_2}$.

Proof. See Appendix C. □

Based on the above inner bound, we derive the following inner bound for the Gaussian channel.

Proposition 4. *For the state-dependent Gaussian MAC with a helper, an inner bound on the capacity region consists of rate pairs* (R_1, R_2) *satisfying:*

$$R_1 \leqslant \min\{f(\alpha,\beta,P_1), g(\alpha,\beta,P_1)\} \tag{19a}$$

$$R_2 \leqslant \min\{f(\alpha,\beta,P_2), g(\alpha,\beta,P_2)\} \tag{19b}$$

$$R_1 + R_2 \leqslant \min\{f(\alpha,\beta,P_1+P_2), g(\alpha,\beta,P_1+P_2)\} \tag{19c}$$

for some real constants α *and* β *satisfying* $-\sqrt{\frac{P_0}{Q}} \leqslant \beta \leqslant \sqrt{\frac{P_0}{Q}}$. *In the above bounds,*

$$f(\alpha,\beta,P) = \frac{1}{2}\log\frac{P_0'(P_0' + (1+\beta)^2Q + P + 1)}{P_0'Q(\alpha-1-\beta)^2 + P_0' + \alpha^2 Q}, \tag{20}$$

$$g(\alpha,\beta,P) = \frac{1}{2}\log\left(1 + \frac{P(P_0' + \alpha^2 Q)}{P_0'Q(\alpha-1-\beta)^2 + P_0' + \alpha^2 Q}\right), \tag{21}$$

where $P_0' = P_0 - \beta^2 Q$.

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

$$U = X_0' + \alpha S, \quad X_0 = X_0' + \beta S, \quad X_0' \sim \mathcal{N}(0, P_0'), \quad X_1 \sim \mathcal{N}(0, P_1), \quad X_2 \sim \mathcal{N}(0, P_2)$$

where X_0', X_1, X_2, S are independent. The constraint on β follows due to the power constraint on X_0. □

We note that the above construction of the input X_0 of the helper reflects two state cancelation schemes: the term βS represents direct cancelation of some state power in the output of the receiver; and the variable X_0' is used for dirty paper coding via generation of the state-correlated auxiliary variable U. Hence, the parameter β controls the balance of two schemes in the integrated scheme, and can be optimized to achieve the best performance. This scheme is also equivalent to the one with $U = X_0 + \alpha S$, where X_0 and S are correlated. While such approaches have been considered in the literature (see e.g., [4]), we believe that selecting U and X_0 successively provides a more operational meaning to the correlation structure.

3.3 Capacity Characterization

By comparing the inner and outer bounds provided in Section 3.2, we characterize the capacity region or segments on the capacity boundary in various channel cases. Our idea is to separately analyze the bounds (19a)-(19c) in the inner bound and characterize conditions on the channel parameters (P_0, P_1, P_2, Q) under which these bounds respectively meet the bounds (17a)-(17c) in the outer bound. In such cases, the corresponding segment on the capacity region is characterized.

We first consider the bound on R_1 in (19a). Let

$$\alpha_1 = \frac{(1+\beta_1)P_0'}{P_0'+1}, \quad \beta_1 = \rho_{0S}^*\sqrt{\frac{P_0}{Q}}. \tag{22}$$

Then $f(\alpha, \beta, P_1)$ takes the following form

$$f(\alpha_1, \beta_1, P_1) = \frac{1}{2}\log\left(1 + \frac{P_1}{Q + 2\rho_{0S}^*\sqrt{P_0 Q} + P_0 + 1}\right) + \frac{1}{2}\log(1 + P_0 - \rho_{0S}^{*2}P_0) \tag{23}$$

where $\rho_{0S}^* \in [-1, 1]$ maximizes

$$\frac{1}{2}\log\left(1 + \frac{P_1}{Q + 2\rho_{0S}\sqrt{P_0 Q} + P_0 + 1}\right) + \frac{1}{2}\log(1 + P_0 - \rho_{0S}^2 P_0).$$

In fact, α_1 is set to maximize $f(\alpha, \beta, P_1)$ for fixed β, and β_1 is set to maximize the function with α being replaced by α_1. If $f(\alpha_1, \beta_1, P_1) \leqslant g(\alpha_1, \beta_1, P_1)$, then $R_1 = f(\alpha_1, \beta_1, P_1)$ is achievable, and this meets the outer bound (the first term in "min" in (17a)). Thus, one segment of the capacity region is specified by $R_1 = f(\alpha_1, \beta_1, P_1)$.

Furthermore, we set $\beta = \alpha - 1$ and then obtain:

$$g(\alpha, \alpha-1, P_1) = \frac{1}{2}\log(1 + P_1). \tag{24}$$

If $g(\alpha, \alpha-1, P_1) \leqslant f(\alpha, \alpha-1, P_1)$, i.e., $P_0'^2 \geq \alpha^2 Q(P_1 + 1 - P_0')$ where $P_0' = P_0 - (\alpha-1)^2 Q$ holds for some $\alpha \in \Omega_\alpha = \{\alpha : 1 - \sqrt{\frac{P_0}{Q}} \leq \alpha \leq 1 + \sqrt{\frac{P_0}{Q}}\}$, then $R_1 = \frac{1}{2}\log(1 + P_1)$ is achievable, and this meets the outer bound (the second term in "min" in (17a)). This also equals the maximum rate for R_1 when the channel is not corrupted by state. Thus, one segment of the capacity region is specified by $R_1 = \frac{1}{2}\log(1 + P_1)$.

Similarly, following the above arguments, segments on the capacity region boundary corresponding to bounds on R_2 and $R_1 + R_2$ can be characterized.

Summarizing the above analysis, we obtain the following characterization of segments of the capacity region boundary.

Theorem 3. *The channel parameters (P_0, P_1, P_2, Q) can be partitioned into the sets $\mathcal{A}_1, \mathcal{B}_1, \mathcal{C}_1$, where*

$$\begin{aligned}
&\mathcal{A}_1 = \{(P_0, P_1, P_2, Q) : f(\alpha_1, \beta_1, P_1) \leqslant g(\alpha_1, \beta_1, P_1)\} \\
&\mathcal{C}_1 = \{(P_0, P_1, P_2, Q) : P_0'^2 \geq \alpha^2 Q(P_1 + 1 - P_0') \\
&\qquad \textit{where } P_0' = P_0 - (\alpha-1)^2 Q, \ \textit{for some } \alpha \in \Omega_\alpha\} \\
&\mathcal{B}_1 = (\mathcal{A}_1 \cup \mathcal{C}_1)^c.
\end{aligned}$$

If $(P_0, P_1, P_2, Q) \in \mathcal{A}_1$, then $R_1 = f(\alpha_1, \beta_1, P_1)$ captures one segment of the capacity region boundary, where the state cannot be fully canceled. If $(P_0, P_1, P_2, Q) \in \mathcal{C}_1$, then $R_1 = \frac{1}{2}\log(1+$

P_1) captures one segment of the capacity region boundary, where the state is fully canceled. If $(P_0, P_1, P_2, Q) \in \mathcal{B}_1$, R_1 segment of the capacity region boundary is not characterized.

The channel parameters (P_0, P_1, P_2, Q) can alternatively be partitioned into the sets $\mathcal{A}_2, \mathcal{B}_2, \mathcal{C}_2$, where

$$\begin{aligned}
&\mathcal{A}_2 = \{(P_0, P_1, P_2, Q) : f(\alpha_2, \beta_2, P_2) \leqslant g(\alpha_2, \beta_2, P_2)\}\\
&\mathcal{C}_2 = \{(P_0, P_1, P_2, Q) : P_0'^2 \geq \alpha^2 Q(P_2 + 1 - P_0')\\
&\qquad \text{where } P_0' = P_0 - (\alpha-1)^2 Q, \text{ for some } \alpha \in \Omega_\alpha\}\\
&\mathcal{B}_2 = (\mathcal{A}_2 \cup \mathcal{C}_2)^c,
\end{aligned}$$

where α_2, β_2 are defined similarly to (22) with P_1 being replaced by P_2. If $(P_0, P_1, P_2, Q) \in \mathcal{A}_2$, then $R_2 = f(\alpha_2, \beta_2, P_2)$ captures one segment of the capacity region boundary, where the state cannot be fully canceled. If $(P_0, P_1, P_2, Q) \in \mathcal{C}_2$, then $R_2 = \frac{1}{2}\log(1+P_2)$ captures one segment of the capacity region boundary, where the state is fully canceled.

Furthermore, the channel parameters (P_0, P_1, P_2, Q) can also be partitioned into the sets $\mathcal{A}_3, \mathcal{B}_3, \mathcal{C}_3$, where

$$\begin{aligned}
&\mathcal{A}_3 = \{(P_0, P_1, P_2, Q) :\\
&\qquad f(\alpha_3, \beta_3, P_1 + P_2) \leqslant g(\alpha_3, \beta_3, P_1 + P_2)\}\\
&\mathcal{C}_3 = \{(P_0, P_1, P_2, Q) : P_0'^2 \geq \alpha^2 Q(P_1 + P_2 + 1 - P_0')\\
&\qquad \text{where } P_0' = P_0 - (\alpha-1)^2 Q, \text{ for some } \alpha \in \Omega_\alpha\}\\
&\mathcal{B}_3 = (\mathcal{A}_3 \cup \mathcal{C}_3)^c,
\end{aligned}$$

where α_3, β_3 are defined similarly to (22) with P_1 being replaced by P_1+P_2. If $(P_0, P_1, P_2, Q) \in \mathcal{A}_3$, then $R_1 + R_2 = f(\alpha_3, \beta_3, P_1 + P_2)$ captures one segment of the sum capacity, where the state cannot be fully canceled. If $(P_0, P_1, P_2, Q) \in \mathcal{C}_3$, then $R_1 + R_2 = \frac{1}{2}\log(1 + P_1 + P_2)$ captures one segment of the sum capacity, where the state is fully canceled.

The above theorem describes three partitions of the channel parameters respectively characterizing segments on the capacity region corresponding to R_1, R_2 and $R_1 + R_2$. Then intersection of three sets (with each from one partition) collectively characterizes all segments on the capacity region boundary. For example, if a given channel parameter tuple satisfies $(P_0, P_1, P_2, Q) \in (\mathcal{C}_1 \bigcap \mathcal{C}_2 \bigcap \mathcal{A}_3)$, then following Theorem 3, line segments characterized by $R_1 = \frac{1}{2}\log(1 + P_1)$, $R_2 = \frac{1}{2}\log(1 + P_2)$, and $R_1 + R_2 = f(\alpha_3, \beta_3, P_1 + P_2)$ are on the capacity region boundary. Since parameters α and β that achieve these segments are not the same, the intersection of these segments are not on the capacity region boundary.

Fig. 6 lists all possible intersections of sets that the channel parameters can belong to. In principle, there should be $3^3 = 27$ cases. We further note that if $(P_0, P_1, P_2, Q) \in \mathcal{C}_3$, they must belong to $\mathcal{C}_1$ and $\mathcal{C}_2$. Hence, the total number of cases becomes $3^2 \times 2 + 1 = 19$. For each case in Fig. 6, we use the solid red lines to represent the segments on the capacity region that are characterized in Theorem 3, and we also mark the value of the capacity that each segment corresponds to as characterized in Theorem 3.

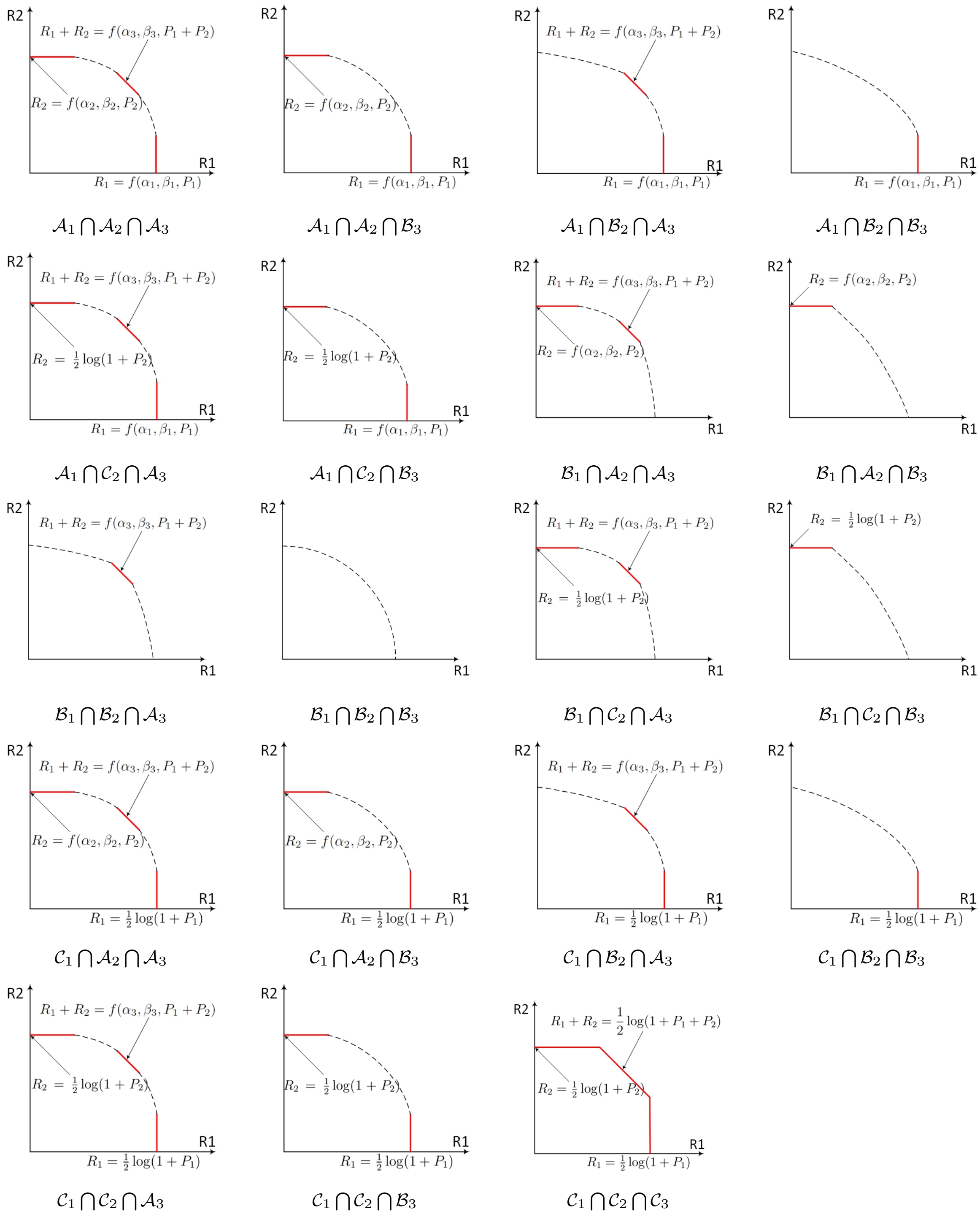

Figure 6: Segments of the capacity region for all cases of channel parameters

We note that for several cases, segments on the capacity region boundary are characterized to be strictly inside the capacity region of the MAC without the state, i.e., the state cannot be fully canceled. For example, for cases with $(P_0, P_1, P_2, Q) \in (\mathcal{A}_1 \bigcap \mathcal{A}_2 \bigcap \mathcal{A}_3)$ and $(\mathcal{A}_1 \bigcap \mathcal{C}_2 \bigcap \mathcal{A}_3)$, sum capacity segments are characterized to be smaller than the sum capacity of the MAC without state. These cases include mostly channel parameters with *finite* Q, and thus contain much larger sets of channel parameters than [19] that characterizes such sum capacity segment only for *infinite* Q.

We further note an interesting case (the last case in Fig. 6), for which the capacity region is fully characterized. We state this result in the following theorem.

Theorem 4. *If* $(P_0, P_1, P_2, Q) \in (\mathcal{C}_1 \bigcap \mathcal{C}_2 \bigcap \mathcal{C}_3)$*, i.e.,*

$$P_0'^2 \geq \alpha^2 Q(P_1 + P_2 + 1 - P_0'), \tag{25}$$

where $P_0' = P_0 - (\alpha - 1)^2 Q$ *for some*$\alpha \in \Omega_\alpha$*, then the capacity region of the state-dependent Gaussian MAC contains* (R_1, R_2) *satisfying*

$$\begin{aligned} R_1 &\leqslant \frac{1}{2}\log(1 + P_1) \\ R_2 &\leqslant \frac{1}{2}\log(1 + P_2) \\ R_1 + R_2 &\leqslant \frac{1}{2}\log(1 + P_1 + P_2) \end{aligned}$$

which achieves the capacity region of the Gaussian MAC without state.

Theorem 4 implies that the state is fully canceled if the channel parameters satisfy the condition (25). We further note two special sets of channel parameters in this case. First, if $P_0 \geq Q$, then $\alpha = 0 \in \Omega_\alpha$ and the condition clearly holds. This is not surprising because the helper has enough power to directly cancel the state. Secondly, if $P_1 + P_2 + 1 \leq P_0 < Q$, then the condition holds for $\alpha = 1 \in \Omega_\alpha$ for arbitrarily large Q. This implies that if the helper's power is above a certain threshold, then the state can always be canceled for arbitrary state power Q (even for infinite Q).

4 Conclusion

In this paper, we studied the state-dependent channel with a helper. Our achievable scheme is based on integration of state subtraction and single-bin dirty paper coding. By analyzing the corresponding lower bound on the capacity, and comparing to the upper bounds, we characterize the capacity for various channel parameters. We anticipate that our way of analyzing the lower bound and characterizing the capacity can be applied to characterizing the capacity for other state-dependent networks. We further point out closely related problems of state masking [20], state amplification [21], assisted interference suppression [22,23], which have a similar goal of minimizing the impact of the state on the output. It will be interesting to explore if the understanding here can shed any insight on these problems.

Appendix

A Proof of Lemma 1

We use random codes and fix the following joint distribution:

$$P_{SUX_0XY_0} = P_S P_{U|S} P_{X_0|US} \cdot P_X P_{Y|X_0XS}.$$

Let $T_\epsilon^n(P_{SUX_0XY})$ denote the strongly joint ϵ-typical set (see, e.g., [24, Sec. 10.6] and [25, 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} i.i.d. codewords $x^n(w)$ according to $P(x^n) = \prod_{i=1}^n P_X(x_i)$ for the fixed marginal probability P_X as defined, in which $w \in [1, 2^{nR}]$.
2. Encoding
   - Encoder at the helper: For given s^n, select $\tilde{t}$ such that $(u^n(\tilde{t}), s^n) \in T_\epsilon^n(P_S P_{U|S})$. If $u^n(\tilde{t})$ cannot be found, set $\tilde{t} = 1$. Then map $(s^n, u^n(\tilde{t}))$ into $x_0^n = f_0^{(n)}(s^n, u^n(\tilde{t}))$. Based on the rate distortion type of argument [24, Sec. 10.5] or the Covering Lemma [26, Sec. 3.7], it can be shown that such $u^n(\tilde{t})$ exists with high probability for large n if
   $$\tilde{R} > I(U;S). \tag{26}$$
   - Encoder at the transmitter: Given w, map w into $x^n(w)$.
3. Decoding
   - Decoder: Given y^n, find a pair $(\hat{t}, \hat{w})$ such that $(u^n(\hat{t}), x^n(\hat{w}), y^n) \in T_\epsilon^n(P_{UXY})$. 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
   $$R \leqslant I(X;Y|U), \tag{27}$$
   $$\tilde{R} \leqslant I(U;Y|X), \tag{28}$$
   $$R + \tilde{R} \leqslant I(UX;Y). \tag{29}$$

We note that (28) corresponds to the decoding error for the index t, which is not the message of interest. Hence, the bound (28) can be removed. Hence, combining (26), (27), and (29), and eliminating $\tilde{R}$, we obtain the desired achievable rate in Lemma 1.

B Proof of Proposition 2

The second bounds in "min" in (17a)-(17c) follow from the capacity of the Gaussian MAC without state. The remaining bounds arise due to capability of the helper for assisting state cancelation and are derived as follows.

Consider a $(2^{nR_1}, 2^{nR_2}, n)$ code with an average error probability $P_e^{(n)}$. The probability distribution on $W_1 \times W_2 \times S^n \times X_0^n \times X_1^n \times X_2^n \times Y^n$ is given by

$$P_{W_1 W_2 S^n X_1^n X_2^n X_0^n Y^n} = P_{W_1} P_{W_2} \left[\prod_{i=1}^{n} P_{S_i} \right] P_{X_1^n|W_1} P_{X_2^n|W_2} P_{X_0^n|S^n} \prod_{i=1}^{n} P_{Y_i|X_{1i}X_{2i}X_{0i}S_i}. \tag{30}$$

By Fano's inequality, we have

$$H(W_1 W_2 | Y^n) \leqslant n(R_1 + R_2) P_e^{(n)} + 1 = n\delta_n \tag{31}$$

where $\delta_n \to 0$ as $n \to +\infty$.

We first bound R_1 based on Fano's inequality as follows:

$$\begin{aligned}
nR_1 &\leqslant I(W_1; Y^n) + n\delta_n \\
&\leqslant I(X_1^n; Y^n) + n\delta_n \\
&= H(X_1^n) - H(X_1^n | Y^n) + n\delta_n \\
&\overset{(a)}{\leqslant} H(X_1^n | X_2^n) - H(X_1^n | X_2^n Y^n) + n\delta_n \\
&= I(X_1^n; Y^n | X_2^n) + n\delta_n \\
&= H(Y^n | X_2^n) - H(Y^n | X_1^n X_2^n)] + n\delta_n \\
&= H(Y^n | X_2^n) - H(S^n Y^n | X_1^n X_2^n) + H(S^n | X_1^n X_2^n Y^n) + n\delta_n \\
&= H(Y^n | X_2^n) - H(Y^n | S^n X_1^n X_2^n) - H(S^n | X_1^n X_2^n) + H(S^n | X_1^n X_2^n Y^n) + n\delta_n \\
&\leqslant H(Y^n | X_2^n) - H(Y^n | S^n X_0^n X_1^n X_2^n) - H(S^n) + H(S^n | X_1^n X_2^n Y^n) + n\delta_n \\
&\overset{(b)}{\leqslant} \sum_{i=1}^{n} [H(Y_i | X_{2i}) - H(Y_i | S_i X_{0i} X_{1i} X_{2i}) - H(S_i) + H(S_i | X_{1i} X_{2i} Y_i)] + n\delta_n
\end{aligned} \tag{32}$$

where (a) follows because X_1^n and X_2^n are independent, and (b) follows because S^n is an i.i.d. sequence.

We bound the first term in the above equation as

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} h(Y_i|X_{2i}) \\
&\quad\leqslant \frac{1}{2n}\sum_{i=1}^{n} \log 2\pi e(Var(X_{1i}+X_{0i}+S_i+N_i)) \\
&\quad= \frac{1}{2n}\sum_{i=1}^{n} \log 2\pi e(Var(X_{1i})+Var(X_{0i}+S_i)+Var(N_i)) \\
&\quad\leqslant \frac{1}{2n}\sum_{i=1}^{n} \log 2\pi e\Big(E[X_{1i}^2]+E[X_{0i}^2]+2E(X_{0i}S_i)+E[S_i^2]+E[N_i^2]\Big) \\
&\quad\leqslant \frac{1}{2}\log 2\pi e\left(\frac{1}{n}\sum_{i=1}^{n} E[X_{1i}^2]+\frac{1}{n}\sum_{i=1}^{n} E[X_{0i}^2]+\frac{2}{n}\sum_{i=1}^{n} E(X_{0i}S_i)+\frac{1}{n}\sum_{i=1}^{n} E[S_i^2]\right. \qquad (33)\\
&\qquad\qquad \left.+\frac{1}{n}\sum_{i=1}^{n} E[N_i^2])\right) \\
&\quad\leqslant \frac{1}{2}\log 2\pi e\left(P_1+P_0+Q+1+\frac{2}{n}\sum_{i=1}^{n} E(X_{0i}S_i)\right) \\
&\quad\leqslant \frac{1}{2}\log 2\pi e\Big(P_1+P_0+Q+1+2\rho_{0s}\sqrt{P_0Q}\Big) \qquad (34)
\end{aligned}
$$

where $\rho_{0s}=\frac{1}{n\sqrt{P_0Q}}\sum_{i=1}^{n} E(X_{0i}S_i)$.

It is easy to obtain bounds on the second and third terms in (32) as follows.

$$
\frac{1}{n}\sum_{i=1}^{n} H(Y_i|S_iX_{0i}X_{1i}X_{2i}) = \frac{1}{2}\log 2\pi e \qquad (35)
$$

$$
\frac{1}{n}\sum_{i=1}^{n} H(S_i) = \frac{1}{2}\log 2\pi e Q \qquad (36)
$$

We next bound the last term in (32) as follows.

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} h(S_i|X_{1i}X_{2i}Y_i) &= \frac{1}{n}\sum_{i=1}^{n} h(S_i|X_{0i}+S_i+N_i) \\
&\leq \frac{1}{n}\sum_{i=1}^{n} h(S_i-\alpha(X_{0i}+S_i+N_i)|X_{0i}+S_i+N_i) \\
&\leqslant \frac{1}{n}\sum_{i=1}^{n} h(S_i-\alpha(X_{0i}+S_i+N_i)) \\
&= \frac{1}{2}\log 2\pi e(\alpha^2P_0+(1-\alpha)^2Q-2\alpha(1-\alpha)\rho_{0s}\sqrt{P_0Q}+\alpha^2) \\
&= \frac{1}{2}\log 2\pi e\left(\frac{Q+(P_0-\rho_{0S}^2P_0)Q)}{Q+2\rho_{0S}\sqrt{P_0Q}+P_0+1}\right) \qquad (37)
\end{aligned}
$$

where the last equation follows by setting $\alpha = \frac{\rho_{0s}\sqrt{P_0 Q}+Q}{1+P_0+Q+2\rho_{0s}\sqrt{P_0 Q}}$ so that $S_i - \alpha(X_{0i}+S_i+N_i)$ and $X_{0i}+S_i+N_i$ are uncorrelated.

Combining the above four bounds, we obtain the following upper bound on R_1.

$$\begin{aligned}
R_1 \leqslant & \frac{1}{2}\log 2\pi e(1+P_0+P_1+Q+2\rho_{0s}\sqrt{P_0 Q}) - \frac{1}{2}\log 2\pi e - \frac{1}{2}\log 2\pi e Q \\
& + \frac{1}{2}\log 2\pi e\left(\frac{Q+(P_0-\rho_{0S}^2 P_0)Q)}{Q+2\rho_{0S}\sqrt{P_0 Q}+P_0+1}\right) \\
\leqslant & \frac{1}{2}\log(1+\frac{P_1}{Q+2\rho_{0S}\sqrt{P_0 Q}+P_0+1}) + \frac{1}{2}\log(1+P_0-\rho_{0S}^2 P_0) \qquad (38)
\end{aligned}$$

Similarly, we can derive an upper bound for R_2 as

$$R_2 \leqslant \frac{1}{2}\log(1+\frac{P_2}{Q+2\rho_{0S}\sqrt{P_0 Q}+P_0+1}) + \frac{1}{2}\log(1+P_0-\rho_{0S}^2 P_0). \qquad (39)$$

We further bound R_1+R_2 following similar arguments. We highlight some important steps below.

$$\begin{aligned}
n(R_1+R_2) & \leqslant I(W_1 W_2; Y^n) + n\delta_n \\
& \leqslant I(X_1^n X_2^n; Y^n) + n\delta_n \\
& = H(Y^n) - H(Y^n|X_1^n X_2^n) + n\delta_n \\
& = H(Y^n) - H(S^n Y^n|X_1^n X_2^n) + H(S^n|X_1^n X_2^n Y^n) + n\delta_n \\
& = H(Y^n) - H(Y^n|S^n X_1^n X_2^n) - H(S^n|X_1^n X_2^n) + H(S^n|X_1^n X_2^n Y^n) + n\delta_n \\
& \leqslant H(Y^n) - H(Y^n|S^n X_0^n X_1^n X_2^n) - H(S^n) + H(S^n|X_1^n X_2^n Y^n) + n\delta_n \\
& \leqslant \sum_{i=1}^{n}[H(Y_i) - H(Y_i|S_i X_{0i} X_{1i} X_{2i}) - H(S_i) + H(S_i|X_{1i} X_{2i} Y_i)] + n\delta_n \qquad (40)
\end{aligned}$$

The first term in (40) can be bounded as follows.

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} h(Y_i) & \leqslant \frac{1}{2n}\sum_{i=1}^{n}\log 2\pi e(Var(X_{1i}+X_{2i}+X_{0i}+S_i+N_i)) \\
& = \frac{1}{2n}\sum_{i=1}^{n}\log 2\pi e(Var(X_{1i})+Var(X_{2i})+Var(X_{0i}+S_i)+Var(N_i)) \\
& \leqslant \frac{1}{2n}\sum_{i=1}^{n}\log 2\pi e\Big(E[X_{1i}^2]+E[X_{2i}^2]+E[X_{0i}^2]+2E(X_{0i}S_i)+E[S_i^2]+E[N_i^2])\Big) \\
& \leqslant \frac{1}{2}\log 2\pi e\Bigg(\frac{1}{n}\sum_{i=1}^{n}E[X_{1i}^2]+\frac{1}{n}\sum_{i=1}^{n}E[X_{2i}^2]+\frac{1}{n}\sum_{i=1}^{n}E[X_{0i}^2]+\frac{2}{n}\sum_{i=1}^{n}E(X_{0i}S_i) \\
& \qquad +\frac{1}{n}\sum_{i=1}^{n}E[S_i^2]+\frac{1}{n}\sum_{i=1}^{n}E[N_i^2])\Bigg) \\
& \leqslant \frac{1}{2}\log 2\pi e\Big(P_1+P_2+Q+1+2\rho_{0s}\sqrt{P_0 Q}\Big) \qquad (41)
\end{aligned}$$

Other bounds in (40) can be bounded in the way as in (35), (36), and (37). Combining these bounds with (41), we obtain the following desired upper bound on $R_1 + R_2$.

$$R_1 + R_2 \leqslant \frac{1}{2} \log \left(1 + \frac{P_1 + P_2}{Q + 2\rho_{0S}\sqrt{P_0 Q} + P_0 + 1}\right) + \frac{1}{2} \log(1 + P_0 - \rho_{0S}^2 P_0) \tag{42}$$

C Proof of Proposition 3

We use random codes and fix the following joint distribution:

$$P_{SUX_0X_1X_2Y} = P_{SU} P_{X_0|SU} P_{X_1} P_{X_2} P_{Y|SX_0X_1X_2}.$$

Let $T_\epsilon^n(P_{SUX_0X_1X_2Y})$ denote the strongly joint ϵ-typical set 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 $2^{n\tilde{R}}$ codewords $u^n(v)$ with i.i.d. components based on P_U. Index these codewords by $v = 1, \ldots, 2^{n\tilde{R}}$.
   - Generate 2^{nR_1} codewords $x_1^n(w_1)$ with i.i.d. components based on P_{X_1}. Index these codewords by $w_1 = 1, \ldots, 2^{nR_1}$.
   - Generate 2^{nR_2} codewords $x_2^n(w_2)$ with i.i.d. components based on P_{X_2}. Index these codewords by $w_2 = 1, \ldots, 2^{nR_2}$.
2. Encoding:
   - Helper: Given s^n, find $\tilde{v}$, such that $(u^n(\tilde{v}), s^n) \in T_\epsilon^n(P_{SU})$. It can be shown that for large n, such $\tilde{v}$ exists with high probability if
   $$\tilde{R} \geqslant I(S; U). \tag{43}$$
   Then given $(u^n(\tilde{v}), s^n)$, generate x_0^n with i.i.d. components based on $P_{X_0|SU}$ for transmission.
   - Transmitter 1: Given w_1, map w_1 into $x_1^n(w_1)$ for transmission.
   - Transmitter 2: Given w_2, map w_2 into $x_2^n(w_2)$ for transmission.
3. Decoding:
   - Given y^n, find $(\hat{v}, \hat{w}_1, \hat{w}_2)$ such that $(u^n(\hat{v}), x_1^n(\hat{w}_1), x_2^n(\hat{w}_2), y^n) \in T_\epsilon^n(P_{UX_1X_2Y})$. If no or more than one $(\hat{w}_1, \hat{w}_2)$ can be found, declare an error.

It can be shown that for sufficiently large n, decoding is correct with high probability if

$$\begin{aligned}
R_1 &\leqslant I(X_1; Y|X_2, U) \\
\tilde{R} + R_1 &\leqslant I(U, X_1; Y|X_2) \\
R_2 &\leqslant I(X_2; Y|X_1, U) \\
\tilde{R} + R_2 &\leqslant I(U, X_2; Y|X_1) \\
R_1 + R_2 &\leqslant I(X_1, X_2; Y|U) \\
\tilde{R} + R_1 + R_2 &\leqslant I(U, X_1, X_2; Y)
\end{aligned}$$

We note that the event that multiple $\hat{v}$ with only single pair $(\hat{w}_1, \hat{w}_2)$ satisfy the above decoding requirement is not counted as an error event, because the index v is not the decoding requirement. Finally, combining the above bounds with (43) yields the desired achievable region.

References

[1] 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.

[2] 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.

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

[4] 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.

[5] A. Zaidi, S. P. Kotagiri, J. N. Laneman, and L. Vandendorpe, "Multiaccess channels with state known to one encoder: Another case of degraded message sets," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, Jul. 2009.

[6] 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.

[7] 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.

[8] R. Duan, Y. Liang, A.Khisti, and S. Shamai (Shitz), "Parallel Gaussian networks with a common state-cognitive helper," *IEEE Trans. Inform. Theory*, vol. 61, no. 12, pp. 6680–6699, December 2015.

[9] 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.

[10] 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 2010.

[11] S. Sigurjónsson and Y.-H. Kim, "On multiple user channels with state information at the transmitters," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, Sep. 2005.

[12] A. Lapidoth and Y. Steinberg, "The multiple-access channel with causal side information: Common state," *IEEE Trans. Inform. Theory*, vol. 59, no. 1, pp. 32–50, Jan. 2013.

[13] A. Zaidi, P. Piantanida, and S. Shamai (Shitz), "Capacity region of cooperative multiple access channel with states," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, pp. 6153–6174, October 2013.

[14] A. Lapidoth and Y. Steinberg, "The multiple-access channel with causal side information: Double state," *IEEE Trans. Inform. Theory*, vol. 59, no. 3, pp. 1379–1393, Mar. 2013.

[15] M. Li, O. Simeone, and A. Yener, "Multiple access channels with state causally known at transmitters," *IEEE Trans. Inform. Theory*, vol. 59, no. 3, pp. 1394–1404, Mar. 2013.

[16] 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.

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

[18] 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.

[19] R. Duan, Y. Liang, and S. Shamai (Shitz), "Dirty interference cancelation for multiple access channels," in *Proc. IEEE Int. Symp. Information Theory and Its Applications(ISITA)*, Melbourne, Australia, Oct. 2014.

[20] N. Merhav and S. Shamai (Shitz), "Information rates subject to state masking," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2254–2261, June 2007.

[21] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.

[22] P. Grover and A. Sahai, "Vector Witsenhausen counterexample as assisted interference suppression," *International Journal of Systems, Control and Communications*, 2010.

[23] C. Choudhuri and U. Mitra, “On Witsenhausen’s counterexample: The asymptotic vector case,” in *Proc. IEEE Information Theory Workshop (ITW)*, 2012.

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

[25] G. Kramer, “Topics in multi-user information theory,” *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 4-5, pp. 265–444, Now Publishers, Hanover, MA, USA, 2007.

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