Communication Subject to State Obfuscation

Author(s):
Wang, Ligong; Wornell, Gregory W.

Publication date:
2020-02-26

Permanent link:
https://doi.org/10.3929/ethz-b-000402954

International Zurich Seminar on Information and Communication (IZS), February 26 – 28, 2020

Communication Subject to State Obfuscation

Ligong Wang* and Gregory W. Wornell†
*ETIS—Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS, Cergy-Pontoise, France
†Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA

***Abstract*—We consider communication over a state-dependent discrete memoryless channel subject to a constraint that requires that the output sequence be nearly independent of the state. We consider three scenarios for the transmitter: where it knows the state, where it does not know the state and can use a stochastic encoder, and where it does not know the state and must use a deterministic encoder. For the state, we assume it to be either independent and identically distributed across channel uses or randomly generated but constant over all channel uses. We present single-letter capacity formulas for all except one combination of the above scenarios, and also solve some illustrative examples.**

I. Introduction

State-dependent channels have been extensively studied in Information Theory [1]–[3]. The current work considers communication over a state-dependent channel, with an additional requirement that the channel state should remain unknown to the receiver. A potential application for such a model is a scenario where the transmitter wishes to conceal its physical location: its location may affect the statistics of the channel to the receiver, hence can be modeled as a channel state.

The problem we study is closely related to "state masking" and, to a lesser extent, "state amplification" [4]–[8]. Consider a state-dependent discrete memoryless channel (DMC) where, given input $X = x$ and state $S = s$, the probability for the output Y to equal y is given by $W(y|x, s)$. Assume that the state is independent and identically distributed (IID) across channel uses according to a known distribution. The state-masking constraint considered in [4] is

$$\lim_{n\to\infty} \frac{1}{n} I(S^n; Y^n) \leq E \tag{1}$$

for some parameter E, where n denotes the number of times the channel is used. When channel-state information (CSI) is available noncausally to the transmitter (meaning the transmitter knows the realization of S^n before sending any input to the channel), a communication rate R is achievable under the above constraint if, and only if [4, Theorem 2]

$$R \leq I(U; Y) - I(U; S) \tag{2}$$

for some auxiliary random variable U such that $U \multimap (X, S) \multimap Y$ form a Markov chain, and that

$$I(S; U, Y) \leq E. \tag{3}$$

Note that (2) is the Gel'fand-Pinsker rate expression [2], while the condition (3) concerns $I(S; U, Y)$ and not $I(S; Y)$.

In the current paper we are interested in problems where the states must be almost completely concealed from the receiver, namely, where the limit in (1) must equal zero. Our result when CSI is available to the transmitter then follows almost immediately from [4]. We also consider situations where CSI is not available and derive similar capacity formulas. Interestingly, capacity differs between the cases where the transmitter must use a deterministic encoder and where it may use a stochastic encoder (that is not known to the receiver). Furthermore, keeping in mind that the state may be used to model the transmitter's physical location, we study models where the state remains constant during the entire transmission, instead of being IID. When CSI is available to the transmitter, or when CSI is not available and the transmitter must use a deterministic encoder, the capacity turns out to be the same as in the IID-state case. When CSI is not available and transmitter may use a stochastic encoder, however, capacity is different.

We consider IID states in Section II and constant states in Section III, and then conclude with some remarks.

II. IID States

Consider a DMC with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$ that is affected by a random state S which takes value in the set $\mathcal{S}$. The sets $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{S}$ are all assumed to be finite. The channel law is, given the input $x \in \mathcal{X}$ and state $s \in \mathcal{S}$, the probability of the output being $y \in \mathcal{Y}$ is $W(y|x, s)$.

In this section, we assume that the states are drawn IID across channel uses according to a probability mass function P_S.

The message to be communicated is drawn from the set $\{1, \ldots, \lfloor 2^{nR} \rfloor\}$, where n denotes the total number of channel uses, and R the rate of communication in bits per channel use. The message is fed to an encoder, which in turn produces the channel input sequence x^n. We consider both cases where the state realizations are known and unknown to the transmitter, respectively. When the states are unknown to the transmitter, we further distinguish between deterministic and stochastic encoders; details are provided below. In all cases, the receiver tries to guess the message based on its observations y^n.

The *state-obfuscation* constraint we impose is

$$\lim_{n\to\infty} \frac{1}{n} I(S^n; Y^n) = 0, \tag{4}$$

where the mutual information is computed for the joint distribution induced by the encoder and a uniformly drawn message. As will become clear via our achievability proofs, all results in this section will continue to hold when we replace (4) by the stronger condition

$$I(S^n; Y^n) = 0 \quad \text{for every } n. \tag{5}$$

In each of the following cases, we define capacity as the supremum over all rates R for which a sequence of encoder-decoder pairs can be constructed such that the probability of a guessing error by the decoder tends to zero as n grows to infinity, and such that (4) is satisfied.

A. With CSI

Assume that the state realizations are available to the encoder. In the case of *noncausal* CSI, the encoder is a (possibly random) mapping from the message m and the state sequence s^n to the input sequence x^n. In the case of *causal* CSI, the encoder is a sequence of (possibly random) mappings from m and s^i to x_i, with $i \in \{1, \ldots, n\}$.

Theorem 1: When the transmitter has either noncausal or causal CSI, the capacity is

$$C_{\text{CSI}}^{\text{IID}} = \sup I(U;Y), \tag{6}$$

where the supremum is taken over joint probability distributions of the form

$$P_S(s)P_U(u)P_{X|US}(x|u,s)W(y|x,s) \tag{7}$$

subject to

$$I(S;U,Y) = 0. \tag{8}$$

Proof: The noncausal case follows from [4, Theorem 2] by noting that (8) requires that U be independent of S. It thus remains only to prove the achievability part for the causal case. To this end, fix any joint distribution of the form (7). For each message $m \in \{1, \ldots, \lfloor 2^{nR} \rfloor\}$, randomly generate a vector $u^n(m)$ by choosing each entry IID according to P_U. To send m, the encoder randomly picks its input at time i to be x_i with probability $P_{X|US}(x_i|u_i(m), s_i)$. Each vector $u^n(m)$ is reveaed to the receiver, but the transmitter's choice of x_i is *not* revealed to the receiver. A standard argument shows that the probability of decoding error can be made arbitrarily close to zero as n grows large provided that $R < I(U;Y)$.

We next examine the constraint (4). Note that (8) implies

$$P_{Y|US}(y|u,s) = P_{Y|U}(y|u) \quad \text{for all } s,u,y. \tag{9}$$

When the code is used to transmit a uniformly chosen message, the probability of $Y^n = y^n$ and $S^n = s^n$, for any y^n and s^n, can be written as

$$\begin{aligned} P_{S^nY^n}(s^n,y^n) &= \sum_{m=1}^{\lfloor 2^{nR} \rfloor} \frac{1}{\lfloor 2^{nR} \rfloor} \prod_{i=1}^{n} P_S(s_i) P_{Y|US}(y|u_i(m),s_i) \\ &= \prod_{i=1}^{n} P_S(s_i) \sum_{m=1}^{\lfloor 2^{nR} \rfloor} \frac{P_{Y|U}(y_i|u_i(m))}{\lfloor 2^{nR} \rfloor}. \end{aligned} \tag{10}$$

Clearly, we have $I(S^n;Y^n) = 0$ for every n. ∎

B. No CSI, Deterministic Encoder

We next consider the case where no CSI is available to the encoder, and where the encoder must be deterministic. Thus, the transmitted sequence x^n is a deterministic function of the message m.

Theorem 2: When the transmitter has no CSI and cannot use a stochastic encoder, the capacity is

$$C_{\text{det}}^{\text{IID}} = \sup I(X;Y), \tag{11}$$

where the supremum is taken over joint distributions of the form

$$P_S(s)P_X(x)W(y|x,s) \tag{12}$$

subject to

$$I(S;X,Y) = 0. \tag{13}$$

Proof: For achievability, we generate each codeword IID according to P_X. The analysis is essentially identical to that in the proof of Theorem 1 and hence omitted.

For converse, by the fact that X^n is a deterministic function of the message M, and by Fano's inequality, we have

$$H(X^n|Y^n) \le H(M|Y^n) \le n\epsilon_n, \tag{14}$$

for some $\epsilon_n \downarrow 0$ as $n \to \infty$. We thus have

$$\begin{aligned} I(S^n;X^n,Y^n) &= I(S^n;X^n|Y^n) + I(S^n;Y^n) \\ &\le H(X^n|Y^n) + I(S^n;Y^n) \\ &\le 2n\epsilon_n, \end{aligned} \tag{15}$$

where the last step follows by the constraint (4). We also have

$$\begin{aligned} I(S^n;X^n,Y^n) &= H(S^n) - H(S^n|X^n,Y^n) \\ &= \sum_{i=1}^{n} H(S_i) - H(S_i|X^n,Y^n,S^{i-1}) \\ &\ge \sum_{i=1}^{n} I(S_i;X_i,Y_i) \\ &\ge nI(S;\bar{X},\bar{Y}), \end{aligned} \tag{16}$$

where $\bar{X}$ denotes a random variable whose distribution is the average of the marginal distributions for every X_i, $i = 1, \ldots, n$, and $\bar{Y}$ is the output corresponding to $\bar{X}$. Here, the last step follows because the distributions for S_i are identical, and by the convexity of mutual information in the conditional distribution of (X_i, Y_i) given S_i. Combining (15) and (16) we obtain

$$I(S;\bar{X},\bar{Y}) \le 2\epsilon_n. \tag{17}$$

On the other hand, by the standard converse proof procedure (see, e.g., [9]),

$$R \le I(\bar{X},\bar{Y}) + \epsilon_n. \tag{18}$$

Combining (17) and (18) we obtain that

$$C_{\text{det}}^{\text{IID}} \le \liminf_{n\to\infty} \sup_{P_n} I(X;Y) \tag{19}$$

where the mutual information is computed according to a distribution of the form

$$P_S(s)P_X(x)W(y|x,s) \tag{20}$$

subject to

$$\lim_{n\to\infty} I(S;X,Y) = 0. \tag{21}$$

The converse to the theorem follows by invoking continuity properties of mutual information. ∎

Remark 1: Theorem 2 is equivalent to saying that the transmitter can only use those input symbols that are not affected by S, namely, it can only use x if $W(\cdot|x,s_1) = W(\cdot|x,s_2)$ for all $s_1, s_2 \in \mathcal{S}$.

C. No CSI, Stochastic Encoder

Next we consider the case where the transmitter has no CSI, but is allowed to use a stochastic encoder. The receiver knows the distribution according to which the codebook is chosen, but not the actual choice by the transmitter. Thus, the encoder is a *random* mapping from message m to input sequence x^n, while the decoder is, as before, a mapping from y^n to its guess of m.

Theorem 3: When the transmitter has no CSI but can use a stochastic encoder, the capacity is

$$C_{\text{sto}}^{\text{IID}} = \sup I(U;Y), \tag{22}$$

where the supremum is taken over joint distributions of the form

$$P_S(s)P_U(u)P_{X|U}(x|u)W(y|x,s) \tag{23}$$

subject to

$$I(S;U,Y) = 0. \tag{24}$$

Proof: The achievability part is similar to the previous cases and is omitted. To prove the converse part, we first use Fano's inequality to obtain

$$\begin{aligned} n(R-\epsilon_n) &\le I(M;Y^n) \\ &\le \sum_{i=1}^{n} I(M,Y^{i-1};Y_i). \end{aligned} \tag{25}$$

We also have

$$\begin{aligned} I(S^n;M,Y^n) &= I(S^n;M|Y^n) + I(S^n;Y^n) \\ &\le H(M|Y^n) + I(S^n;Y^n) \\ &\le 2n\epsilon_n, \end{aligned} \tag{26}$$

where the last step follows by Fano's inequality and the constraint (4). On the other hand,

$$\begin{aligned} I(S^n;M,Y^n) &= \sum_{i=1}^{n} I(S_i;M,Y^n,S^{i-1}) \\ &\ge \sum_{i=1}^{n} I(S_i;M,Y^{i-1},Y_i). \end{aligned} \tag{27}$$

Let $U_i \triangleq (M,Y^{i-1})$, $i = 1,\dots,n$. We have shown

$$\sum_{i=1}^{n} I(U_i;Y_i) \ge n(R-\epsilon_n) \tag{28}$$

$$\sum_{i=1}^{n} I(S;U_i,Y_i) \le 2n\epsilon_n, \tag{29}$$

where $\epsilon_n \downarrow 0$ as $n \to \infty$. Note that U_i is independent of S_i because S^n is IID. The rest of the proof is similar to that for Theorem 2. ∎

The next example shows that $C_{\text{sto}}^{\text{IID}}$ can be larger than $C_{\text{det}}^{\text{IID}}$.

Example 1: Consider a channel where $\mathcal{X} = \mathcal{Y} = \{0,1,2\}$ and $\mathcal{S} = \{0,1\}$. The channel law is, when $S = 0$, $Y = X$ with probability one; when $S = 1$, $Y = 0$ if $X = 0$, but the other two symbols are reversed: $Y = 2$ if $X = 1$ and $Y = 1$ if $X = 2$ (all with probability one). A deterministic encoder can only use the input symbol 0, hence it cannot send any information. A stochastic encoder can choose $U \in \{0,1\}$ uniformly, $X = 0$ if $U = 0$, and $X = 1$ or 2 equally likely if $U = 1$. This achieves one bit per channel use. One can verify that this is in fact optimal.

D. A Consequence

A simple consequence to the above results is that the capacity in every case is upper-bounded by the worst-state capacity over all $s \in \mathcal{S}$.

Corollary 4: In all settings above, capacity is upper-bounded by

$$\min_s \sup_{P_X} I(X;Y|S=s). \tag{30}$$

Proof: It suffices to consider the CSI case, since clearly

$$C_{\text{CSI}}^{\text{IID}} \ge C_{\text{sto}}^{\text{IID}} \ge C_{\text{det}}^{\text{IID}}. \tag{31}$$

Recall that, in the formula (6), S must be independent of the pair (U,Y). It follows that

$$I(U;Y) = I(U;Y|S=s) \tag{32}$$

for every $s \in \mathcal{S}$. Hence

$$\begin{aligned} C_{\text{CSI}}^{\text{IID}} &\le \sup_{P_U,P_{X|US}} \min_s I(U;Y|S=s) \\ &\le \min_s \sup_{P_U,P_X|U} I(U;Y|S=s) \\ &\le \min_s \sup_{P_X} I(X;Y|S=s), \end{aligned} \tag{33}$$

where the last step follows because $U \multimap\!\!- (X,S) \multimap\!\!- Y$ form a Markov chain. ∎

Example 2: Consider a channel where $\mathcal{X} = \mathcal{Y} = \mathcal{S} = \{0,1\}$. Assume that P_S is uniform. When $S = 0$, the channel is a perfect bit pipe: $Y = X$ with probability one; when $S = 1$, it is a Z-channel with $1 \to 0$ cross-over probability $p \in (0,1)$ (see [9]). Corollary 4 implies that $C_{\text{CSI}}^{\text{IID}}$ cannot exceed the capacity of the Z-channel. We show that they are equal. Let U be a binary random variable with the capacity-achieving input distribution of the Z-channel. Let $P_{X|US}$ be such that

$$P_{X|US}(1|0,s) = 0, \qquad s = 1,2 \tag{34a}$$

$$P_{X|US}(1|1,0) = 1-p \tag{34b}$$

$$P_{X|US}(1|1,1) = 1, \tag{34c}$$

namely, when $S = 1$, we choose $X = U$ with probability one; when $S = 0$, X is produced by passing U through the above Z-channel. By this choice, we have the same Z-channel from U to Y irrespectively of the value of S, hence $I(S;U,Y) = 0$, whereas $I(U;Y)$ equals the capacity of the Z-channel.

One can show that $C_{\text{sto}}^{\text{IID}} = C_{\text{det}}^{\text{IID}} = 0$. We delay the proof to the end of the next section, when we return to this example.

III. Constant States

Consider the same DMC as described in the first paragraph of Section II. We now assume the state to be constant instead of IID. This means the state is generated randomly according to P_S before communication starts, and remains the same throughout the n channel uses when transmission takes place. The decoder is, as in Section II, a mapping from y^n to a guess of the message. For state obfuscation, we now require

$$\lim_{n\to\infty} I(S;Y^n) = 0. \tag{35}$$

All our claims in this section will continue to hold under the stronger condition that removes the limit in (35). In all cases below, capacity is defined as the supremum over all rates for which one can find a sequence of encoder-decoder pairs such that (35) is satisfied while the decoding error probability will approach zero when n grows large.[1]

A. With CSI

When CSI is available to the transmitter, the encoder is a possibly random mapping from (s,m) to x^n, where m denotes the message and x^n the input sequence. The capacity in this case is the same for constant and IID states.

Theorem 5: For any DMC described by transition law $W(\cdot|\cdot,\cdot)$ and state distribution P_S, the capacity when S is constant and when CSI is available to the transmitter is

$$C_{\text{CSI}}^{\text{const}} = C_{\text{CSI}}^{\text{IID}}, \tag{36}$$

where $C_{\text{CSI}}^{\text{IID}}$ is given by Theorem 1.

Proof: The achievability proof is essentially the same as that for Theorem 1. We note that, since by the choice of joint distribution, the pair (U,Y) is independent of S, we can use typicality to treat (u^n,y^n), even though the state is constant and not ergodic.

To prove the converse, we define auxiliary random variables

$$U_i \triangleq M, Y^{i-1}, \quad i = 1,\dots,n. \tag{37}$$

Using Fano's inequality and the chain rule, we have

$$\begin{aligned} n(R-\epsilon_n) &\le I(M;Y^n) \\ &\le \sum_{i=1}^{n} I(M,Y^{i-1};Y_i) \\ &= \sum_{i=1}^{n} I(U_i;Y_i). \end{aligned} \tag{38}$$

We next show that $I(S;U_i,Y_i)$ must be close to zero for every i. Clearly, it is enough to show that $I(S;M,Y^n)$ must approach zero as n grows large. To this end, define a binary random variable F that equals 0 when decoding is correct and equals 1 when decoding is incorrect. Then we have

$$\begin{aligned} I(S;M,Y^n) &= I(S;Y^n) + I(S;M|Y^n) \\ &\le I(S;Y^n) + I(S;M,F|Y^n) \\ &= I(S;Y^n) + I(S;F|Y^n) + I(S;M|Y^n,F) \\ &\le I(S;Y^n) + H(F) + I(S;M|Y^n,F). \end{aligned} \tag{39}$$

[1]Since the state remains constant during communication, our definition requires that the error probability be small for every possible realization of S.

The first two terms on the right-hand side of (39) both tend to zero as n grows large, the first by (35), and the second because the probability of a decoding error must tend to zero. For the last term, let ϵ denote the probability of a decoding error, then we have

$$\begin{aligned} &I(S;M|Y^n,F) \\ &= (1-\epsilon)\sum_{y^n} \Pr(Y^n=y^n|F=0) I(S;M|Y^n=y^n,F=0) \\ &\quad + \epsilon \sum_{y^n} \Pr(Y^n=y^n|F=1) I(S;M|Y^n=y^n,F=1) \\ &\le (1-\epsilon)\cdot 0 + \epsilon \cdot \log|\mathcal{S}| \\ &= \epsilon \cdot \log|\mathcal{S}|, \end{aligned} \tag{40}$$

which also must tend to zero as n grows large. Hence we have shown that, as n grows large, the right-hand side of (39) must tend to zero, and consequently $I(S;U_i,Y_i)$ must tend to zero for every i. This, together with (38) and a continuity argument, completes the converse proof. ∎

B. No CSI, Deterministic Encoder

Assume that the encoder must be a deterministic mapping that maps the message m to an input sequence x^n. The capacity is again the same as in the IID-state case.

Theorem 6: For any $W(\cdot|\cdot,\cdot)$ and P_S, the capacity in the current setting is

$$C_{\text{det}}^{\text{const}} = C_{\text{det}}^{\text{IID}}. \tag{41}$$

Proof: The achievability is essentially the same as before. For converse, we have, for every $i \in \{1,\dots,n\}$,

$$\begin{aligned} I(S;X_i,Y_i) &\le I(S;X_i,Y^n) \\ &= I(S;Y^n) + I(S;X_i|Y^n) \\ &\le I(S;Y^n) + H(X_i|Y^n). \end{aligned} \tag{42}$$

Since the encoder is deterministic, the decoder should be able to correctly guess every X_i from Y^n (by first guessing M). By Fano's inequality, $H(X_i|Y^n)$ must vanish together with the error probability. Hence, for every i,

$$\lim_{n\to\infty} I(S;X_i,Y_i) = 0. \tag{43}$$

Next consider the communication rate R. For some vanishing ϵ_n,

$$\begin{aligned} n(R-\epsilon_n) &\le I(X^n;Y^n) \\ &\le I(X^n,S;Y^n) \\ &\le \sum_{i=1}^{n} I(X_i,S;Y_i) \\ &\le \sum_{i=1}^{n} I(X_i;Y_i) + I(S;X_i,Y_i). \end{aligned} \tag{44}$$

Combining (43) and (44) completes the converse. ∎

C. No CSI, Stochastic Encoder

When the transmitter has no CSI, a stochastic encoder is a random mapping that maps m to x^n. The decoder knows the distribution used by the stochastic encoder, but not which codebook is chosen. Denote the capacity in this case subject to (35) by $C_{\text{sto}}^{\text{const}}$. We have not been able to find a single-letter expression for $C_{\text{sto}}^{\text{const}}$. One can verify that the achievability part of Theorem 3 is still valid. We can thus order the capacities in various cases as

$$C_{\text{det}}^{\text{IID}} = C_{\text{det}}^{\text{const}} \leq C_{\text{sto}}^{\text{IID}} \leq C_{\text{sto}}^{\text{const}} \leq C_{\text{CSI}}^{\text{IID}} = C_{\text{CSI}}^{\text{const}}. \tag{45}$$

That the first inequality above can be strict was demonstrated by Example 1. The other two inequalities can also be strict, as we show via the next two examples.

Example 3: Let $\mathcal{X} = \mathcal{Y} = \mathcal{S} = \{0, 1\}$. When $S = 0$ the channel is a noiseless bit pipe; when $S = 1$ the bit is flipped at the output with probability one.

We have $C_{\text{sto}}^{\text{IID}} = 0$ because, without CSI and when the states are IID, it is impossible for the transmitter to send any information, even without the constraint (4). We show that

$$C_{\text{sto}}^{\text{const}} = 1 \text{ bit}. \tag{46}$$

Consider the following simple scheme. The transmitter generates a random variable B uniformly over $\{0, 1\}$. To send $(n-1)$ information bits over n channel uses, it sends B followed by the XOR of each information bit and B. The output string is then IID and uniform irrespectively of the value of S. To decode, the receiver obtains $B \oplus S$ from the first bit, and computes its XOR with the next $(n-1)$ received bits to recover the information bits.

Example 4: Consider the same channel as in Example 2, except now the state remains the same for all n channel uses. Recall that $C_{\text{CSI}}^{\text{IID}}$ equals the capacity of the Z-channel; by Theorem 5, so does $C_{\text{CSI}}^{\text{const}}$. We shall show that

$$C_{\text{sto}}^{\text{const}} = 0. \tag{47}$$

Together with (45), this will imply $C_{\text{det}}^{\text{IID}} = C_{\text{det}}^{\text{const}} = C_{\text{sto}}^{\text{IID}} = 0$. To show (47), consider any sequence of encoder-decoder pairs, and define

$$A_n \triangleq \sum_{i=1}^{n} X_i \tag{48}$$

$$B_n \triangleq \sum_{i=1}^{n} Y_i. \tag{49}$$

Further define

$$\alpha \triangleq P\text{-}\limsup_{n\to\infty} \frac{A_n}{n}, \tag{50}$$

where $P\text{-}\limsup$ denotes the limit-supremum in probability: α is the smallest real number for which the probability that $\frac{A_n}{n} > \alpha$ tends to zero as $n \to \infty$. Assume that $\alpha > 0$. Note that, when $S = 0$, $B_n = A_n$ with probability one. Thus we have

$$\limsup_{n\to\infty} \Pr\left(\frac{B_n}{n} \geq \left(1 - \frac{p}{2}\right)\alpha \middle| S = 0 \right) > 0. \tag{51}$$

When $S = 1$, B_n is conditionally a binomial distribution with parameters A_n and p, so its limit-supremum in probability given $S = 1$ must equal $(1-p)\alpha$, therefore

$$\lim_{n\to\infty} \Pr\left(\frac{B_n}{n} \geq \left(1 - \frac{p}{2}\right)\alpha \middle| S = 1 \right) = 0. \tag{52}$$

It follows from (51) and (52) that the total variation distance between the conditional distributions of B_n conditional on $S = 0$ and $S = 1$, respectively, cannot approach zero as n grows large. By Pinsker's Inequality [9], this further implies that $I(S; B_n)$ cannot approach zero, and therefore $I(S; Y^n)$ cannot approach zero either. Thus the assumption that $\alpha > 0$ is incompatible with the requirement (35). But having $\alpha = 0$ clearly does not permit communication at a positive rate. We have thus proven (47).

IV. Concluding Remarks

We have presented information-theoretic capacity expressions for several instances of communication subject to state obfuscation. The case where the state remains constant during transmission time and is unknown to the transmitter, and where the transmitter can use a stochastic encoder, is yet unsolved. We have demonstrated via examples that the capacity in this case differs from both the IID-state no-CSI stochastic-encoder case and the constant-state with-CSI case.

To analyze real-life scenarios where the transmitter wishes to guarantee a low probability of geolocation by the receiver, one may replace the abstract models considered in the current paper by specific channel models. For example, in line-of-sight multiple-antenna wireless communication, the state S may correspond to the phase difference between observation at receive antennas. For free-space optical communication, S may correspond to attenuation of the transmitted signal. Examples 2 and 4 may be considered a first step along the latter direction.

References

[1] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Research and Development*, vol. 2, pp. 289–293, 1958.

[2] S. I. Gel'fand and M. S. Pinsker, "Coding for channels with random parameters," *Prob. Contr. and Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

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

[4] N. Merhav and S. Shamai, "Information rates subject to state masking," *IEEE Trans. Inform. Theory*, vol. 53, pp. 2254–2261, June 2007.

[5] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inform. Theory*, vol. 54, pp. 1850–1859, May 2008.

[6] O. O. Koyluoglu, R. Soundararajan, and S. Vishwanath, "State amplification under masking constraints," in *Proc. 49th Allerton Conf. Comm., Contr. and Comp.*, (Monticello, IL), Sept. 28–30, 2011.

[7] T. Courtade, "Information masking and amplification: The source coding setting," in *Proc. IEEE Int. Symp. Inform. Theory*, (Cambridge, MA, USA), July 1–6 2012.

[8] M. Dikshtein and S. Shamai, "Broadcasting information subject to state masking." 2018, arXiv:1810.11781.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, second ed., 2006.