Wiretap Channel with Causal State Information

Yeow-Khiang Chia and Abbas El Gamal
Department of Electrical Engineering
Stanford University
Email: ykchia@stanford.edu, abbas@ee.stanford.edu

***Abstract*—A lower bound on the secrecy capacity of the wiretap channel with state information available causally at both the encoder and decoder is established. The lower bound is shown to be strictly larger than that for the noncausal case by Liu and Chen. Achievability is proved using block Markov coding, Shannon strategy, and key generation from common state information. The state sequence available at the end of each block is used to generate a key, which is used to enhance the transmission rate of the confidential message in the following block. An upper bound on the secrecy capacity when the state is available noncausally at the encoder and decoder is established and is shown to coincide with the lower bound for several classes of wiretap channels with state.**

I. Introduction

Consider the 2-receiver wiretap channel with state depicted in Figure 1. The sender X wishes to send a message to the legitimate receiver Y while keeping it asymptotically secret from the eavesdropper Z. The secrecy capacity for this channel can be defined under various scenarios of state information availability at the encoder and decoder. When the state information is not available at either party, the problem reduces to the classical wiretap channel for the channel averaged over the state and the secrecy capacity is known [1]. When the state is available only at the decoder, the problem reduces to the wiretap channel with augmented receiver (Y, S).

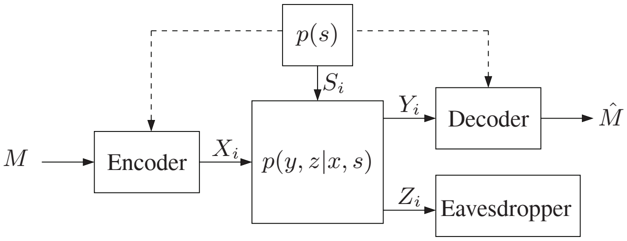

Fig. 1: Wiretap channel with State

The interesting scenarios to consider therefore are when the state information is available at the encoder and may or may not be available at the decoder. This raises the question of how the encoder and decoder can make use of the state information to increase the secrecy rate. In [2], Chen and Vinck established a lower bound on the secrecy capacity when the state information is available noncausally only at the encoder. The lower bound is established using a combination of Gelfand–Pinsker coding and Wyner wiretap coding. Subsequently, Liu and Chen [3] used the same techniques to establish a lower bound on the secrecy capacity when the state information is available noncausally at both the encoder and decoder. In a related direction, Khisti, Diggavi, and Wornell [4] considered the problem of secret key agreement, first studied in [5] and [6], for the wiretap channel with state and established the secret key capacity when the state is available causally or noncausally at the encoder and decoder. The key is generated in two parts; the first using a wiretap channel code while treating the state sequence as a time-sharing sequence, and the second part is generated from the state itself.

In this paper, we consider the wiretap channel with state information available *causally* at the encoder and decoder. We establish a lower bound that can be strictly larger than the lower bound for the noncausal case in [3]. Our achievability scheme is quite different from the scheme in [3]. We use block Markov coding, Shannon strategy for channels with state [7], and secret key agreement from state information, which builds on the work in [4]. However, unlike [4], we are not directly interested in the size of the secret key, but rather in using the secret key generated from the state sequence in one transmission block to increase the secrecy rate in the following block. This block Markov scheme causes additional information leakage through the correlation between the secret key generated in a block and the received sequences at the eavesdropper in subsequent blocks. We show that this leakage is asymptotically negligible. Although a similar block Markov coding scheme was used in [8] to establish the secrecy capacity of the degraded wiretap channel with rate limited secure feedback, in their setup no information about the key is leaked to the eavesdropper because the feedback link is assumed to be secure.

We also establish an upper bound on the secrecy

978-1-4244-7892-7/10/$26.00 ©2010 IEEE

capacity of the wiretap channel with state information available noncausally at the encoder and decoder. We show that the upper bound coincides with the aforementioned lower bound for several classes of channels. Thus, the secrecy capacity for these classes does not depend on whether the state information is known causally or noncausally at the encoder.

In the following section, we provide the needed definitions. In Section III, we present the lower bound. The upper bound and secrecy capacity results are presented in Section IV. The omitted proofs and other results are given in the extended version posted online at ArXiv: http://arxiv.org/abs/1001.2327.

II. Problem Definition

Consider a discrete memoryless wiretap channel (DM-WTC) with discrete memoryless state (DM) $(\mathcal{X}\times\mathcal{S}, p(y,z|x,s)p(s), \mathcal{Y}, \mathcal{Z})$ consisting of a finite input alphabet $\mathcal{X}$, finite output alphabets $\mathcal{Y}$, $\mathcal{Z}$, a finite *state* alphabet $\mathcal{S}$, a collection of conditional pmfs $p(y,z|x,s)$ on $\mathcal{Y}\times\mathcal{Z}$, and a pmf $p(s)$ on the state alphabet $\mathcal{S}$. The sender X wishes to send a confidential message $M \in [1:2^{nR}]$ to the receiver Y while keeping it secret from the eavesdropper Z with either causal or noncausal state information available at both the encoder and decoder.

A $(2^{nR}, n)$ code for the DM-WTC with causal state information at the encoder and decoder consists of: (i) a message set $[1:2^{nR}]$, (ii) an encoder that generates a symbol $X_i(m)$ according to a conditional pmf $p(x_i|m, s^i, x^{i-1})$ for $i \in [1:n]$; and a decoder that assigns an estimate $\hat{M}$ or an error message to each received sequence pair (y^n, s^n). We assume throughout that the message M is uniformly distributed over the message set. The probability of error is defined as $P_e^{(n)} = \mathrm{P}\{\hat{M} \neq M\}$. The information leakage rate at the eavesdropper Z, which measures the amount of information about M that leaks out to the eavesdropper, is defined as $R_L = \frac{1}{n}I(M;Z^n)$. A secrecy rate R is said to be achievable if there exists a sequence of codes with $P_e^{(n)} \to 0$ and $R_L \to 0$ as $n \to \infty$. The secrecy capacity $C_{\mathrm{S-CSI}}$ is the supremum of the set of achievable rates.

We also consider the case when the state information is available noncausally at the encoder. The only change in the above definitions is that the encoder now generates a codeword $X^n(m)$ according to the conditional pmf $p(x^n|m, s^n)$, i.e., the stochastic mapping is allowed to depend on the entire state sequence instead of just the past and present state sequence. The secrecy capacity for this scenario is denoted by $C_{\mathrm{S-NCSI}}$.

The notation used in this paper will follow that of El Gamal–Kim Lectures on Network Information Theory [9, Lecture 1].

III. Lower Bound

The main result in this paper is the following lower bound on the secrecy capacity of the DM-WTC with causal state information available causally at both the encoder and decoder.

Theorem 1: The secrecy capacity of the DM-WTC with state information available causally at the encoder and decoder is lower bounded as

$$C_{\mathrm{S-CSI}} \geq \max\{R_{\mathrm{S-CSI-1}}, R_{\mathrm{S-CSI-2}}\}, \text{ where}$$

$$\begin{aligned} R_{\mathrm{S-CSI-1}} &= \max_{p(v|s)p(x|v,s)} \min\{I(V;Y|S) - I(V;Z|S) \\ &\qquad + H(S|Z), I(V;Y|S)\}, \\ R_{\mathrm{S-CSI-2}} &= \max_{p(v)p(x|v,s)} \min\{H(S|Z,V), I(V;Y|S)\}. \end{aligned}$$

Note that if $S = \emptyset$, the expression above reduces to the secrecy capacity for the wiretap channel. In [3], the authors established the following lower bound for the noncausal case

$$\begin{aligned} & C_{\mathrm{S-NCSI}} \\ &\geq \max_{p(u|s)p(x|u,s)} (I(U;Y,S) - \max\{I(U;Z), I(U;S)\}) \\ &= \max_{p(u|s)p(x|u,s)} \min\{I(U;Y|S) - I(U;Z|S) \\ &\qquad + I(S;U|Z), I(U;Y|S)\}. \end{aligned} \tag{1}$$

Clearly, $R_{\mathrm{S-CSI-1}}$ is at least as large as this lower bound. Hence, our lower bound is at least as large as this lower bound (1). In the extended version of this paper, we show that (1) is as large as $R_{\mathrm{S-CSI-1}}$. To show that $R_{\mathrm{S-CSI-2}}$ can be strictly larger than $R_{\mathrm{S-CSI-1}}$, consider the channel with $\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{S} \in \{0,1\}$ and $p(y,z|x,s) = p(y,z|x)$ and $p(y,z|x) = p(z|x)p(y|z)$, where $p(z|x) = 1$ for $x = z$ and 0 otherwise, and $p(y|z) = 0.9$ for $y = z$ and 0.1 otherwise. The state S is an i.i.d process with $H(S) = 1 - H(0.1)$. In the extended version, we show that $R_{\mathrm{S-CSI-2}} = 1 - H(0.1) > R_{\mathrm{S-CSI-1}}$.

Proof of Theorem 1

Proof of Theorem 1 follows by proving the achievability of $R_{\mathrm{S-CSI-1}}$ and $R_{\mathrm{S-CSI-2}}$ separately. The proof of achievability for $R_{\mathrm{S-CSI-1}}$ is split into two cases (Cases 1 and 2) while $R_{\mathrm{S-CSI-2}}$ is proved in Case 3.

Using the functional representation lemma [10], we can show that it suffices to perform the maximization for $R_{\mathrm{S-CSI-1}}$ over $p(u), p(x|v,s)$, and functions $v(u,s)$. Thus, we prove achievability for the equivalent characterization $R_{\mathrm{S-CSI-1}}$

$$\begin{aligned} R_{\mathrm{S-CSI-1}} \geq & \max_{p(u),v(u,s),p(x|v,s)} \min\{I(U;Y,S) \\ & - I(U;Z,S) + H(S|Z), I(U;Y,S)\}. \end{aligned} \tag{2}$$

Case 1: $R_{\text{S-CSI-1}}$ with $I(U;Y,S) > I(U;Z,S)$

Codebook generation: Split message M_j into two independent messages $M_{j0} \in [1 : 2^{nR_0}]$ and $M_{j1} \in [1 : 2^{nR_1}]$, thus $R = R_0 + R_1$. Let $\tilde{R} \geq R$. The codebook generation consists of two steps.

Message codeword generation: We randomly and independently generate $2^{n\tilde{R}}$ sequences $u^n(l)$, $l \in [1 : 2^{n\tilde{R}}]$, each according to $\prod_{i=1}^n p(u_i)$ and partition them into 2^{nR_0} equal-size bins $\mathcal{C}(m_0)$, $m_0 \in [1 : 2^{nR_0}]$. Partition the sequences within each bin $\mathcal{C}(m_0)$ into 2^{nR_K} equal size sub-bins, $\mathcal{C}(m_0, m_1)$, $m_1 \in [1 : 2^{nR_1}]$.

Key codebook generation: We randomly and uniformly partition the set of s^n sequences into 2^{nR_K} bins $\mathcal{B}(k)$, $k \in [1 : 2^{nR_K}]$.

Encoding: We send $b-1$ messages over b n-transmission blocks. In the first block, we randomly select a $u^n(L)$ sequence. The encoder then computes $v_i = v(u_i(L), s_i)$, $i \in [1 : n]$, and transmits a randomly generated sequence X^n according to $\prod_{i=1}^n p(x_i|s_i, v_i)$. At the end of the first block, the encoder and decoder declare $k_1 \in [1 : 2^{nR_K}]$ such that $\mathbf{s}(1) \in \mathcal{B}(k_1)$ as the key to be used in block 2.

Encoding in block $j \in [2 : b]$ proceeds as follows. To send message $m_j = (m_{j0}, m_{j1})$ and given key k_{j-1}, the encoder computes $m'_{j1} = m_{j1} \oplus k_{j-1}$. To ensure secrecy, we must have $R_1 \leq R_K$ [11]. The encoder then selects a random sequence $u^n(L) \in \mathcal{C}(m_{j0}, m'_{j1})$. The encoder then uses Shannon strategy as depicted in Figure 2. At time $i \in [(j-1)n+1 : jn]$, it computes $v_i = v(u_i(L), s_i)$, and transmits $X_i \sim p(x_i|s_i, v_i)$.

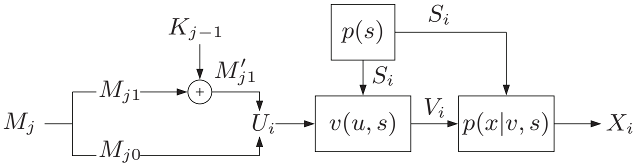

Fig. 2: Encoding in block j.

Decoding and analysis of the probability of error: At the end of block j, the decoder declares that $\hat{l}$ is sent if it is the unique index such that $(u^n(\hat{l}), \mathbf{Y}(j), \mathbf{S}(j)) \in \mathcal{T}_\epsilon^{(n)}$, otherwise it declares an error. It then finds the indices $(\hat{m}_{j0}, \hat{m}'_{j1})$ such that $u^n(l) \in \mathcal{C}(\hat{m}_{j0}, \hat{m}'_{j1})$. Finally, it recovers $\hat{m}_{j1}$ by computing $\hat{m}_{j1} = \hat{m}'_{j1} \oplus k_{j-1}$.

To analyze the error probability, let $\epsilon'' > \epsilon' > \epsilon > 0$, and define the following events for $j \in [2 : b]$:

$$\begin{aligned}
&\mathcal{E}(j) = \{\hat{M}_j \neq M_j\},\\
&\mathcal{E}_1(j) = \{(U^n(L), \mathbf{S}(j)) \notin \mathcal{T}_{\epsilon'}^n\},\\
&\mathcal{E}_2(j) = \{(U^n(L), \mathbf{S}(j), \mathbf{Y}(j)) \notin \mathcal{T}_{\epsilon''}^n\},\\
&\mathcal{E}_3(j) = \{(U^n(\hat{l}), \mathbf{S}(j), \mathbf{Y}(j)) \in \mathcal{T}_{\epsilon''}^n \text{ for some } \hat{l} \neq L\}.
\end{aligned}$$

The probability of error is upper bounded as

$$\mathrm{P}(\mathcal{E}) = \mathrm{P}\{\cup_{j=2}^b \mathcal{E}(j)\} \leq \sum_{j=2}^b \mathrm{P}(\mathcal{E}(j)).$$

Each probability of error term can be upper bounded as

$$\begin{aligned}
\mathrm{P}(\mathcal{E}(j)) \leq \mathrm{P}(\mathcal{E}_1(j)) &+ \mathrm{P}(\mathcal{E}_2(j) \cap \mathcal{E}_1^c(j))\\
&+ \mathrm{P}(\mathcal{E}_3(j) \cap \mathcal{E}_2^c(j)).
\end{aligned}$$

Now, $\mathrm{P}(\mathcal{E}_1(j)) \to 0$ as $n \to \infty$ by Law of Large Numbers (LLN) since $\mathrm{P}\{(U^n(L) \in \mathcal{T}_\epsilon^{(n)})\} \to 1$ as $n \to \infty$ and $\mathbf{S}(j) \sim \prod_{i=1}^n p(s_i) = \prod_{i=1}^n p(s_i|u_i)$ by independence. The term $\mathrm{P}(\mathcal{E}_2(j) \cap \mathcal{E}_1^c(j)) \to 0$ as $n \to \infty$ by LLN since $(U^n(L), \mathbf{S}(j) \in \mathcal{T}_{\epsilon'}^n$ and $Y^n \sim \prod_{i=1}^n p(y_i|u_i, s_i)$. For the last term, consider

$$\mathrm{P}(\mathcal{E}_3(j) \cap \mathcal{E}_2^c(j)) = \sum_l p(l)\, \mathrm{P}(\mathcal{E}_3(j) \cap \mathcal{E}_2^c(j) | L = l).$$

Note that L is independent of the transmission codebook sequences U^n and the current state sequence $\mathbf{S}(j)$. Therefore, by the packing lemma [9, Lecture 3], $\mathrm{P}(\mathcal{E}_3(j) \cap \mathcal{E}_2^c(j) | L = l) \to 0$ as $n \to \infty$ if $\tilde{R} < I(U;Y,S) - \delta(\epsilon'')$. Hence, $\mathrm{P}(\mathcal{E}_3(j) \cap \mathcal{E}_2^c(j)) \to 0$ as $n \to \infty$ if $\tilde{R} < I(U;Y,S) - \delta(\epsilon'')$.

Analysis of the information leakage rate: Note that M_{j0} is transmitted using Wyner wiretap coding. Hence, it can be kept secret from eavesdropper if $I(U;Y,S) - I(U;Z,S) > 0$. The new part of the proof is to show that M_{j1} can be kept secret from the eavesdropper. This involves showing that the eavesdropper has negligible information about K_{j-1}, which is correlated with its received sequences in blocks $j-1$ to b. We show that the eavesdropper has negligible information about K_{j-1} provided $R_K < H(S|Z)$. We will need the following.

Proposition 1: If $R_K < H(S|Z) - 4\delta(\epsilon)$ and $\tilde{R} \geq I(U;Z,S)$, then the following holds for every $j \in [1 : b]$.

1) $H(K_j|\mathcal{C}) \geq n(R_K - \delta(\epsilon))$.
2) $I(K_j; \mathbf{Z}(j)|\mathcal{C}) \leq 2n\delta(\epsilon)$.
3) $I(K_j; \mathbf{Z}^j|\mathcal{C}) \leq n\delta'(\epsilon)$, where $\delta(\epsilon) \to 0$ and $\delta'(\epsilon) \to 0$ as $\epsilon \to 0$.

The proof is given in the extended version.

We are now ready to upper bound the leakage rate averaged over codes. Consider

$$\begin{aligned}
I(M_2, M_3, \ldots, M_b; \mathbf{Z}^b|\mathcal{C}) &= \sum_{j=2}^b I(M_j; \mathbf{Z}^b|\mathcal{C}, M_{j+1}^b)\\
&\overset{(a)}{\leq} \sum_{j=2}^b I(M_j; \mathbf{Z}^b|\mathcal{C}, \mathbf{S}(j), M_{j+1}^b)\\
&\overset{(b)}{=} \sum_{j=2}^b I(M_j; \mathbf{Z}^j|\mathcal{C}, \mathbf{S}(j)),
\end{aligned}$$

where (a) follows by the independence of M_j and $(\mathbf{S}(j), M_{j+1}^b)$, and (b) follows by the Markov Chain relation $(\mathbf{Z}_{j+1}^b, M_{j+1}^b, \mathcal{C}) \to (\mathbf{Z}^j, \mathbf{S}(j), \mathcal{C}) \to (M_j, \mathcal{C})$. Hence, it suffices to upper bound each individual term $I(M_j; \mathbf{Z}^j|\mathcal{C}, \mathbf{S}(j))$. Consider

$$\begin{aligned} I(M_j; \mathbf{Z}^j|\mathcal{C}, \mathbf{S}(j)) &= I(M_{j0}, M_{j1}; \mathbf{Z}^j|\mathcal{C}, \mathbf{S}(j)) \\ &= I(M_{j0}, M_{j1}; \mathbf{Z}^{j-1}|\mathcal{C}, \mathbf{S}(j)) \\ &\quad + I(M_{j0}, M_{j1}; \mathbf{Z}(j)|\mathcal{C}, \mathbf{S}(j), \mathbf{Z}^{j-1}). \end{aligned}$$

Note that the first term is equal to zero by the independence of M_j and past transmissions, the codebook, and state sequence. For the second term, we have

$$\begin{aligned} &I(M_{j0}, M_{j1}; \mathbf{Z}(j)|\mathcal{C}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\quad = I(M_{j0}; \mathbf{Z}(j)|\mathcal{C}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\qquad + I(M_{j1}; \mathbf{Z}(j)|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}). \end{aligned}$$

Consider the first term

$$\begin{aligned} &I(M_{j0}; \mathbf{Z}(j)|\mathcal{C}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &= I(M_{j0}, L; \mathbf{Z}(j)|\mathcal{C}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\quad - I(L; \mathbf{Z}(j)|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\leq I(U^n; \mathbf{Z}(j)|\mathcal{C}, \mathbf{S}(j), \mathbf{Z}^{j-1}) - H(L|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\quad + H(L|\mathbf{Z}(j), M_{j0}, \mathbf{S}(j)) \\ &\leq \sum_{i=1}^{n} \left(H(\mathbf{Z}_i(j)|\mathcal{C}, \mathbf{S}_i(j)) - H(\mathbf{Z}_i(j)|\mathcal{C}, U_i, \mathbf{S}_i(j))\right) \\ &\quad - H(L|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) + H(L|\mathbf{Z}(j), M_{j0}, \mathbf{S}(j)\big) \\ &\overset{(a)}{\leq} nI(U; Z|S) - H(L|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\quad + H(L|M_{j0}, \mathbf{S}(j), \mathbf{Z}(j)) \\ &\overset{(b)}{\leq} nI(U; Z|S) - H(L|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\quad + n(\tilde{R} - R_0 - I(U; Z, S) + \delta(\epsilon)) \\ &\overset{(c)}{=} n(\tilde{R} - R_0) - H(L|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) + n\delta(\epsilon) \\ &= n(\tilde{R} - R_0) - H(M_{j1} \oplus K_{j-1}|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\quad - H(L|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}, M_{j1} \oplus K_{j-1}) + n\delta(\epsilon) \\ &\leq n(\tilde{R} - R_0) - H(M_{j1} \oplus K_{j-1}|\mathcal{C}, \\ &\quad M_{j0}, \mathbf{S}(j), K_{j-1}, \mathbf{Z}^{j-1}) - n(\tilde{R} - R_0 - R_K) + n\delta(\epsilon) \\ &\overset{(d)}{=} nR_K - H(M_{j1} \oplus K_{j-1}|\mathcal{C}, M_{j0}, \mathbf{S}(j), K_{j-1}) + n\delta(\epsilon) \\ &= nR_K - H(M_{j1}|\mathcal{C}, M_{j0}, \mathbf{S}(j), K_{j-1}) + n\delta(\epsilon) = n\delta(\epsilon), \end{aligned}$$

where (a) follows from the fact that $H(\mathbf{Z}_i(j)|\mathcal{C}, \mathbf{S}_i(j)) \leq H(\mathbf{Z}_i(j)|\mathbf{S}_i(j)) = H(Z|S)$ and $H(\mathbf{Z}_i(j)|\mathcal{C}, U_i, \mathbf{S}_i(j)) = H(Z|U, S)$. Step (b) follows by Lemma 1 in [12] which requires that (i) $\mathrm{P}\{(U^n(L), \mathbf{S}(j), \mathbf{Z}(j)) \in \mathcal{T}_\epsilon^{(n)}\} \to 1$ as $n \to \infty$, and (ii) $\tilde{R} - R_0 \geq I(U; Z, S)$; where (i) can be shown using the same steps as in the analysis of probability of error. Step (c) follows by the independence of U and S. Step (d) follows from the Markov Chain relation $(\mathbf{Z}^{j-1}, M_{j0}, \mathbf{S}(j)) \to (K_{j-1}, M_{j0}, \mathbf{S}(j)) \to (M_{j1} \oplus K_{j-1}, M_{j0}, \mathbf{S}(j))$. The last step follows by the fact that M_{j1} is independent of $(\mathcal{C}, M_{j0}, \mathbf{S}(j), K_{j-1})$ and uniformly distributed over $[1 : 2^{nR_K}]$. Next, consider the second term

$$\begin{aligned} &I(M_{j1}; \mathbf{Z}(j)|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\leq I(U^n; \mathbf{Z}(j)|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\qquad - H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\qquad + H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j}) \\ &\overset{(a)}{\leq} nI(U; Z|S) - H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\qquad + H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j}) \\ &\leq nI(U; Z|S) - H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\qquad + H(L|M_{j0}, \mathbf{S}(j), \mathbf{Z}(j)) \\ &\overset{(b)}{\leq} nI(U; Z|S) - H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\qquad + n(\tilde{R} - R_0) - nI(U; Z, S) + n\delta(\epsilon) \\ &= n(\tilde{R} - R_0) - H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) + n\delta(\epsilon), \end{aligned}$$

where (a) follows from the same steps used in bounding $I(M_{j0}; \mathbf{Z}(j)|\mathcal{C}, \mathbf{S}(j), \mathbf{Z}^{j-1})$; (b) follows from Lemma 1 in [12]. Next consider

$$\begin{aligned} &H(L|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &= H(M_{j1} \oplus K_{j-1}|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &\qquad + H(L|\mathcal{C}, M_{j0}, M_{j1}, M_{j1} \oplus K_{j-1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \\ &= H(K_{j-1}|\mathcal{C}, M_{j0}, M_{j1}, \mathbf{S}(j), \mathbf{Z}^{j-1}) + n(\tilde{R} - R_0 - R_K) \\ &= H(K_{j-1}|\mathcal{C}, \mathbf{Z}^{j-1}) + n(\tilde{R} - R_0 - R_K). \end{aligned}$$

From Proposition 1, $H(K_{j-1}|\mathcal{C}, \mathbf{Z}^{j-1}) \geq n(R_K - \delta(\epsilon) - \delta'(\epsilon))$, which implies that

$$I(M_{j1}; \mathbf{Z}(j)|\mathcal{C}, M_{j0}, \mathbf{S}(j), \mathbf{Z}^{j-1}) \leq n(\delta'(\epsilon) + 2\delta(\epsilon)).$$

This completes the analysis of information leakage rate.

Combining the rate constraints and performing Fourier-Motzkin elimination gives the bounds in (2).

Case 2: $R_{\text{S-CSI-1}}$ with $I(U; Y, S) \leq I(U; Z, S)$

In this case, only the key from the previous block is used to encrypt the message transmitted to the eavesdropper. Part of the key is also used to generate uncertainty at the eavesdropper about the codeword sent, to ensure that a sufficiently large secret key rate is achieved in the current block. The proof is given in the extended version of the paper.

Case 3: $R_{\text{S-CSI-2}}$

Achievability of $R_{\text{S-CSI-2}}$ uses the same techniques as for Case 2 of $R_{\text{S-CSI-1}}$. However, here the key generated in a block is used only to encrypt the message in the following block. The eavesdropper may be able to decode the codeword transmitted in a block, which would reduce the key rate generated at the end of that block. This is compensated for by the fact that the entire key is used for encryption.

IV. Upper Bound and Secrecy Capacities

We establish the following upper bound on the secrecy capacity of the wiretap channel with noncausal state information available at both the encoder and decoder (which holds also for the causal case).

Theorem 2: The following is an upper bound to the secrecy capacity of the DM-WTC with state noncausally available at the encoder and decoder

$$C_{\text{S-NCSI}} \leq \min\{I(V_1;Y|U,S) - I(V_1;Z|U,S) \\ +H(S|Z,U), I(V_2;Y|S)\}.$$

for some U, V_1 and V_2 such that $p(u,v_1,v_2,x|s) = p(u|s)p(v_1|u,s)p(v_2|v_1,s)p(x|v_2,s)$.

The achievable rate $R_{\text{S-CSI-1}}$ in Theorem 1 coincide with Theorem 2 when $I(U;Y|S) \geq I(U;Z|S)$ for U such that $(U,S) \to (X,S) \to (Y,Z)$ form a Markov chain, i.e., when Y is *less noisy* than Z for every state $s \in \mathcal{S}$ [13].

Theorem 3: The secrecy capacity for the DM-WTC with the state information available causally or noncausally at the encoder and decoder when Y is less noisy than Z is

$$C_{\text{S-CSI}} = C_{\text{S-NCSI}} = \max_{p(x|s)} \min\{I(X;Y|S) \\ - I(X;Z|S) + H(S|Z), I(X;Y|S)\}.$$

Setting $p(y,z|x,s) = p(y,z|x)$ and considering the case when Z is a degraded version of Y, Theorem 3 specializes to the secrecy capacity for the wiretap channel with a key [14]

$$C_{\text{S-CSI}} = C_{\text{S-NCSI}} \\ = \max_{p(x)} \min\{I(X;Y) - I(X;Z) + H(S), I(X;Y)\}.$$

In addition, $R_{\text{S-CSI-2}}$ in Theorem 1 and Theorem 2 coincide for the following.

Theorem 4: The secrecy capacity for the DM-WTC with the state information available causally or noncausally at the encoder and decoder when $p(y,z|x,s) = p(y,z|x)$ and Z is less noisy than Y is

$$C_{\text{S-CSI}} = C_{\text{S-NCSI}} = \max_{p(x)} \min\{H(S), I(X;Y)\}.$$

Proofs of Theorems 2, 3, 4 and other secrecy capacity results are given in the extended version.

V. Conclusion

We established bounds on the secrecy capacity of the wiretap channel with state information causally available at the encoder and decoder. We showed that our lower bound can be strictly larger than the best known lower bound for the noncausal state information case. The upper bound holds when the state information is available noncausally at the encoder and decoder. We showed that the bounds are tight for several classes of wiretap channels.

As we have seen, the secrecy capacity for several special classes of the wiretap channels with state available at both the encoder and the legitimate receiver does not depend on whether the state is available causally or noncausally. An interesting question posed by an anonymous reviewer is whether this observation holds in general for our setup.

References

[1] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[2] Y. Chen and A. J. Han Vinck, "Wiretap channel with side information," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 395–402, 2006.

[3] W. Liu and B. Chen, "Wiretap channel with two-sided state information," in *Proc. 41st Asilomar Conf. Signals, Systems and Comp.*, Pacific Grove, CA, Nov. 2007, pp. 893–897.

[4] A. Khisti, S. N. Diggavi, and G. W. Wornell, "Secret key agreement using asymmetry in channel state knowledge," in *Proc. IEEE International Symposium on Information Theory*, Seoul, South Korea, July 2009, pp. 2286–2290.

[5] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, 1993.

[6] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—I: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, 1993.

[7] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, vol. 2, pp. 289–293, 1958.

[8] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y. H. Kim, "Wiretap channel with rate-limited feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5353–5361, December 2009.

[9] A. El Gamal and Y. H. Kim, "Lectures on network information theory," 2010, online: http://arxiv.org/abs/1001.3404.

[10] F. M. J. Willems and E. C. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 313–327, 1985.

[11] C. E. Shannon, "Communication theory of secrecy systems," *Bell Systems Tech. J.*, vol. 28, pp. 656–715, 1949.

[12] Y. K. Chia and A. El Gamal, "3-receiver broadcast channels with common and confidential messages," 2009, submitted to IEEE Trans. Inf. Theory. online: http://arxiv.org/abs/0910.1407.

[13] J. Körner and K. Marton, "Comparison of two noisy channels," in *Topics in Information Theory (Second Colloq., Keszthely, 1975)*, 1977, pp. 411–423.

[14] H. Yamamoto, "Rate-distortion theory for the Shannon cipher system," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 827–835, 1997.