Private Broadcasting with Probing Constraint

Y. Ozan Basciftci,
Dep. of Electrical & Computer Eng.
The Ohio State University
Columbus, Ohio, USA
Email: basciftci.1@osu.edu

C. Emre Koksal,
Dep. of Electrical & Computer Eng.
The Ohio State University
Columbus, Ohio, USA
Email: koksal@ece.osu.edu

***Abstract*—We consider a three-receiver wiretap channel in which the transmitter aims to send a common message to legitimate receivers 1 and 2 while keeping it secret from the receiver 3 (eavesdropper). The transmitter takes cost constrained actions to probe the transmitter-to-receiver 1 and transmitter-to-receiver 2 channels. Under the probing cost constraint, we provide a lower bound to the secrecy capacity of this setting. The achievability strategy employs a block Markov coding strategy in which the channel state sequence of the previous block is mapped to a key which secures the part of the confidential message to be transmitted in the current block. The challenge in using the state sequence as a source of the key is the fact that the legitimate receivers are not aware of each other's channel state. To that end, we employ a novel strategy in which the transmitter takes XOR of individual keys generated from the state sequences of transmitter-to receiver 1 and transmitter-to-receiver 2 channels and send the XOR-ed key along with the confidential message. We furthermore analyze the trade-off between the achievable secrecy rate and the probing cost constraint. We conclude that there is a linear dependency between the key rate and probing cost constraint.**

I. Introduction

We consider the three-receiver wiretap channel with state $S \triangleq (S_1, S_2, S_3)$, as depicted in Figure I, in which the transmitter aims to communicate a common message to receivers 1 and 2 while hiding it from receiver 3 (eavesdropper). The transmitter takes cost constrained actions to observe transmitter-to-receiver 1 channel state S_1 and transmitter-to-receiver 2 channel state S_2 *causally*, whereas each receiver only observes the state of its own channel causally without a cost. In this work, we establish a lower bound to the secrecy capacity of this setting. Note that the secrecy capacity is a function of the channel observing cost. We investigate the trade-off between the lower bound and the channel observing cost constraint at the transmitter, namely *probing cost constraint*.

The derivation of the lower bound includes block Markov coding [9] in which communication time is separated to b transmission blocks of length n and a sequence of independent messages are transmitted over these b blocks. The state sequence available at the transmitter at the previous block is mapped to key K and this key is XOR-ed with the part of the confidential message in the current block. The challenge in using the state sequence available at the transmitter to generate a key is that the legitimate receivers are not aware of each other's channel states.

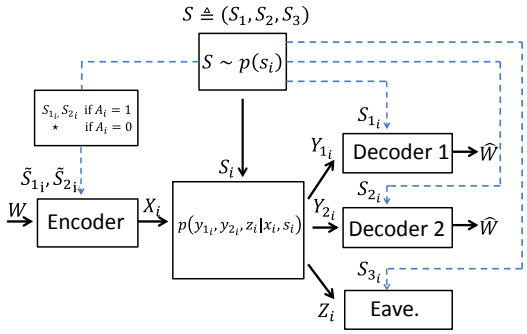

Fig. 1. Wiretap channel with state $S \triangleq (S_1, S_2, S_3)$. Action A_i denotes the probing decision of the transmitter at time i.

To that end, we propose a novel strategy to establish a secret key between the legitimate terminals. The transmitter maps observed transmitter-to-receiver 1 channel state sequence $\tilde{S}_1^n$ to $K_1 \in [1 : 2^{nR_K}]$ and observed transmitter-to-receiver 2 channel state sequence $\tilde{S}_2^n$ to $K_2 \in [1 : 2^{nR_K}]$, where R_K is the key rate. The transmitter picks key K as K_1, i.e., $K = K_1$ and sends $K_1 \oplus K_2$ in addition to the confidential message. Receiver 2 can decode key K_1 by $K_1 = K_1 \oplus K_2 \oplus K_2$. Note that we do not make a noiseless public channel assumption in this the paper, which is a common assumption in the literature [8] for establishing a key between multiple terminals. Hence, the transmitter needs to send the information related to the key, which is $K_1 \oplus K_2$, along with the confidential message throughout the wiretap channel.

In our scheme, the transmitter needs to acquire the state information of the transmitter-to-receiver 1 and transmitter-to-receiver 2 channels to generate key K. The probing of the channel states, however, is a costly process and results in the consumption of system resources, e.g., power. To investigate the trade-off between key rate R_K and the probing cost constraint, we adopt the probing cost model in [6] in which the transmitter chooses to probe or not to probe the channel in each channel use. In [6], the authors aim to maximize the achievable rate when the average probing cost is constrained by a probing cost constraint. In this paper, we show that key rate R_K is a linear and increasing function of the probing cost constraint.

Related Work: Motivated by the wiretap channel, introduced in [1], later generalized to the non-degraded [2] and the Gaussian settings [3], there has been a flurry of recent studies on wireless secrecy. Here, we will discuss only those that are directly related to our work.

In [10], the authors study the wiretap channel with a single legitimate receiver, where the legitimate pair obtains the channel state causally with no probing cost. The authors also employ block Markov coding approach to exploit channel state information to generate a secret key. In Proposition 1 of [10], it is shown that if a key rate is smaller than the entropy of the channel state conditioned on the received signal at the eavesdropper, the key is kept secret from the eavesdropper. In this paper, we extend this proposition to the multiple receiver case. The main challenge in deriving the conditions that makes the key secure is that there is no coordination between the legitimate receivers while generating K_1 and K_2 since they do not observe each other's channel state. To address the challenge, the receivers 1 and 2 independently generate K_1 and K_2, respectively. We show that in addition to the individual entropies of the legitimates receivers' channel states, the joint entropy of the channel states conditioned on the eavesdropper signal bound the key rate so that the key is hidden from the eavesdropper.

In [11], the authors study the private broadcasting of a common message to multiple receivers in the presence of an eavesdropper as in our setting. The authors, however, assume that the eavesdropper has a complete knowledge of the channel state information. For this reason, the state information available at the legitimate pairs can not be exploited to generate a secret key.

II. Problem Setup

We first provide the notation to be used throughput the paper. The uppercase letters (eg. $X, Y, \dots$) denote random variables and realizations of random variables are denoted by lowercase letters (eg. $x, y, \dots$). Let discrete random variable X have a probability mass function $p(x)$, i.e., $X \sim p(x)$ and $\epsilon \in (0,1)$. We define the set of ϵ-typical length-n sequences as

$$T_\epsilon^n = \left\{ x^n : \left| \frac{|i : \{x_i = x\}|}{n} - p(x) \right| \leq \epsilon p(x), \forall x \in \mathcal{X} \right\} \quad (1)$$

where $\mathcal{X}$ is the sample space of X.

Consider a three-receiver discrete memoryless broadcast channel (DM-BC) with discrete memoryless (DM) state shown in Figure I. We assume that the channel model $(\mathcal{X} \times \mathcal{S}, p(y_1, y_2, z|x, s)p(s), \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{Z}_1)$ consists of a finite input alphabet $\mathcal{X}$, a finite output alphabet $\mathcal{Y}_2 \times \mathcal{Y}_2 \times \mathcal{Z}$, a finite state alphabet $\mathcal{S}$, and a collection of conditional pmfs $p(y_1, y_2, z|x_i, s)$ on $\mathcal{Y}_2 \times \mathcal{Y}_2 \times \mathcal{Z}$. The channel is memoryless and described by $p(y_1^n, y_2^n, z^n|x^n, s^n) = \prod_{i=1}^n p_{Y_1,Y_2,Z|X,S}(y_{1_i}, y_{2_i}, z_i|x_i, s_i)$. The channel state evolves in a memoryless fashion as specified by $p(s^n) = \prod_{i=1}^n p_S(s_i)$ where $S \triangleq (S_1, S_2, S_3)$ and S_1, S_2, and S_3 denote the channel states of transmitter-to-receiver 1, transmitter-to-receiver 2, and transmitter-to-eavesdropper channels, respectively.

By employing a $(2^{nR_s}, n)$ code, the transmitter wishes to send message $w \in \{1, \dots, 2^{nR_s}\} = [1 : 2^{nR_s}]$ to receiver 1 and 2 and to keep w secret from the eavesdropper. A $(2^{nR_s}, n)$ code consists of

- Probing Logic: The transmitter picks an action sequence A^n where $p(a^n) = \prod_{i=1}^n p(a_i)$ and $A^n \in \mathcal{A}^n = \{0,1\}^n$. The action sequence satisfies the constraint
$$E\left[\sum_{i=1}^n \Lambda(A_i)\right] \leq \Gamma, \quad (2)$$
where Γ is the probing constraint and $\Lambda(\cdot)$ is the cost function. We assume that $\Lambda(0) = 0$ and $\Lambda(1) = B$, where $B > 0$. Event $\{A_i = 1\}$ denotes that the transmitter observes the transmitter-to-receiver 1 and transmitter-to-receiver 2 channel states perfectly, whereas $\{A_i = 0\}$ corresponds to the lack of observation i.e.,
$$(\tilde{S}_{1_i}, \tilde{S}_{2_i}) = h(S_{1_i}, S_{2_i}, A_i) = \begin{cases} (S_{1_i}, S_{2_i}), & \text{if } A_i = 1. \\ \star, & \text{if } A_i = 0. \end{cases}$$
where h is the probing function, $\star$ represents the lack of knowledge of the channel state information[1], and $\tilde{S}_{1_i}$ and $\tilde{S}_{1_i}$ denote the channel state information available at the transmitter 1 and transmitter 2, respectively for i-th time instant.
- Encoding: The stochastic encoder maps message $w \in [1 : 2^{nR_s}]$ to a sequence of channel inputs $X^n(w) \in \mathcal{X}^n$ where $p(x^n) = \prod_{i=1}^n p(x_i|\tilde{s}^i, x^{i-1}, w)$. We assume that the transmitter obtains the channel state information causally i.e., it acquires $\tilde{S}^i$ at i-th time instant. Define $\tilde{S}_i \triangleq (\tilde{S}_{1_i}, \tilde{S}_{2_i})$.
- Decoding: The decoding function 1 (respectively, 2) is described by function $f_1 : \mathcal{Y}_1^n \times \mathcal{S}_1^n \times \mathcal{A}^n \to [1 : 2^{nR_s}]$ (respectively, $f_2 : \mathcal{Y}_2^n \times \mathcal{S}_2^n \times \mathcal{A}^n \to [1 : 2^{nR_s}]$) that estimates w and outputs $\tilde{w}_1$ (respectively, $\tilde{w}_2$). We assume that all terminals have a causal access to it's own channel state and the action sequence. For instance; the eavesdropper acquires S_3^i and A^i at time instant i.

We assume that message W is uniformly distributed on $[1 : 2^{nR_s}]$. The probability of error is defined by $P_e^n \triangleq \mathbb{P}\left[\cup_{i\in\{1,2\}}\{W \neq \tilde{W}_i\}\right]$.

The uncertainty of message W at the eavesdropper is measured by equivocation rate $R_e \triangleq \frac{1}{n}H(W|Z^n, A^n, S_3^n)$. Rate R_s is achievable if, for any $\epsilon > 0$, there exists sequence of $(2^{nR_s}, n)$ codes such that $P_e^n \leq \epsilon$ and $R_e \geq R_s - \epsilon$ for sufficiently large n. The supremum of the achievable rates is referred to the secrecy capacity.

[1]Without loss of generality, we assume that $\{\star\} \notin \mathcal{S}$

III. Main Result

Theorem 1. *The secrecy capacity of DM-BC $p(y_1, y_2, z|x, s)$ is bounded below by*

$$C_s \geq \max\Big\{\big\{R - I(U;Z,S,A) + \min\{\frac{1}{2}I(U;Z,S,A), R_{Key1}\}\big\}, \min\big\{R_{Key2}, \frac{1}{2}R\big\}\Big\}, \quad (3)$$

where the maximization is over all joint distribution of the form

$$\begin{aligned} &p_{U,S,A,\tilde{S},V,X,Y_1,Y_2,Z}(u,s,a,\tilde{s},v,x,y_1,y_2,z) \\ &= p_U(u)p_S(s)p_A(a)1_{\tilde{s}=h(a,s_1,s_2)}1_{v=g(u,\tilde{s})}p_{X|V,\tilde{S}}(x|v,\tilde{s}) \\ &\quad \times p_{Y_1,Y_2,Z|X,S}(y_1,y_2,z|x,s) \end{aligned} \quad (4)$$

for some $p_U(u), p_A(a), g, p_{X|V,\tilde{S}}(x|v,\tilde{s})$ such that $E[\Lambda(A)] \leq \Gamma$, and

$$R = \min\{I(U;Y_1,S_1,A), I(U;Y_2,S_2,A)\} \quad (5)$$

$$R_{Key1} = \min\big\{H(\tilde{S}_1|Z,S_3,A), H(\tilde{S}_2|Z,S_3,A), \frac{1}{2}H(\tilde{S}_1,\tilde{S}_2|Z,S_3,A)\big\} \quad (6)$$

$$R_{Key2} = \min\big\{H(\tilde{S}_1|Z,U,S_3,A), H(\tilde{S}_2|Z,U,S_3,A), \frac{1}{2}H(\tilde{S}_1,\tilde{S}_2|Z,U,S_3,A)\big\} \quad (7)$$

□

Remark 1. ***(Comparison of the key rate for two cases: probing is free and probing is costly)*** *We compare the key rate expressions* (6)-(7) *with the ones for the case in which the encoder always probes the channel. Specifically, we analyze the key rate for two cases: 1) The encoder cannot probe the channel all the time due to the probing constraint 2, 2) The encoder always probes the channel, i.e., the probing is free. Without loss of generality, we pick the joint entropy term in* (6)*,*

$$H(\tilde{S}_1,\tilde{S}_2|Z,S_3,A) = H(\tilde{S}_1,\tilde{S}_2|Z,S_3,A=1)\mathbb{P}(A=1) \quad (8)$$

$$= H(S_1,S_2|Z,S_3,A=1)\mathbb{P}(A=1) \quad (9)$$

where (S,Z) in (9) *is distributed with*

$$p_{S,Z,A}(s,z|A=1) = \sum_u p_U(u)p_S(s)1_{v=g(u,s_1,s_2)} \times p_{X|V,\tilde{S}}(x|v,s_1,s_2)p_{Z|X,S}(z|x,s) \quad (10)$$

In the scenario where the encoder always probes the channel, we analyze the associated term $H(S_1,S_2|Z,S_3)$ where (S,Z) is distributed with

$$f_{S,Z}(s,z) = \sum_u f_U(u)p_S(s)1_{v=g_1(u,s_1,s_2)} \times f_{X|V,S}(x|v,s_1,s_2)p_{Z|X,S}(z|x,s) \quad (11)$$

As seen in (10) *and* (11)*, if the encoding strategy in two cases is kept same, i.e., $f_U(u) = p_U(u)$, $f_{X|V,S}(x|v,s_1,s_2) = p_{X|V,\tilde{S}}(x|v,s_1,s_2)$, $g(u,s_1,s_2) = g_1(u,s_1,s_2)$, then $H(S_1,S_2|Z,S_3,A=1) = H(S_1,S_2|Z,S_3)$.*[2] *Hence, the ratio of the key rate in the case 1 over the one in case 2 is $P(A=1)$ for the same encoding strategy.*

Note that the probing constraint in Theorem 1 can be written as $P(A=1) \leq \frac{\Gamma}{B}$. If $P(A=1) = \frac{\Gamma}{B}$ then, we observe that as Γ increases, the key rate increases in (6) *(see* (9)*) and approaches to the key rate that is achieved when the encoder always probes the channel.* □

Here we give the proof sketch of Theorem 1. The complete proof will be available in [12]. The encoding scheme contains b transmission blocks each of which has n channel uses. The encoder skips the first transmission block and a sequence of $(b-1)$ messages $w(j), j \in [1 : b-1]$, each is selected independently from a random variable uniformly distributed over $[1 : 2^{nR_s}]$, is transmitted over $(b-1)$ blocks.

Key Generation: At the end of block j, the decoder 1 (resp. decoder 2) randomly and independently assigns each $\tilde{s}_1^n(j) \in \tilde{\mathcal{S}}_1^n$ (resp. $\tilde{s}_2^n(j) \in \tilde{\mathcal{S}}_2^n$) to bin $\mathcal{B}_1(k_1(j)), k_1(j) \in [1 : 2^{nR_K}]$ (resp. $\mathcal{B}_1(k_2(j)), k_2(j) \in [1 : 2^{nR_K}]$), where $\tilde{\mathcal{S}}_1 \triangleq \mathcal{S} \cup \{\star\}$ Also, the process of assigning the sequences into bins in decoder 1 is independent of that in decoder 2 i.e., events $\{s_1^n \in \mathcal{B}_1(k_1)\}$ and $\{s_2^n \in \mathcal{B}_2(k_2)\}$ are independent for all $s_1^n \in \tilde{\mathcal{S}}_1^n$, $s_2^n \in \tilde{\mathcal{S}}_2^n$, and $k_1, k_2 \in [1 : 2^{nR_K}]$. Partitioning is done uniformly such that

$$\mathbb{P}[s_1^n \in \mathcal{B}_1(k_1)] = \mathbb{P}[s_2^n \in \mathcal{B}_2(k_2)] = 2^{-nR_K} \quad (12)$$

for all $s_1^n, s_2^n, k_1,$ and k_2. Note that decoder 1 (resp. decoder 2) is not aware of $k_2(j)$ (resp. $k_1(j)$). Bin index for decoder 1, $k_1(j)$ is used as a key in block $(j+1)$.

Encoding: Due to the space constraints, we only outline the achievability proof of the first rate in the maximization term (3) for the case when $\min\{I(U;Y_1,S_1,A), I(U;Y_2,S_2,A)\} \geq I(U;Z,S,A)$. Message $w(j) \in [1 : 2^{nR_s}]$ is partitioned into two independent sub-messages $w_0(j) \in [1 : 2^{nR_{s_1}}]$ and $w_0(j) \in [1 : 2^{nR_{s_1}}]$, where $R_s = R_{s_1} + R_{s_2}$. The second part of message is secured with one-time pad with the key generated at block $j-1$ i.e., $w_1'(j) = w_1(j) \oplus k_1(j-1)$.

The encoder generates codebook $\mathcal{C} = \{u^n(l) : l \in [1 : 2^{n\tilde{R}}]\}$ where $\tilde{R} < \min\{I(U;Y_1,S_1,A), I(U;Y_2,S_2,A)\}$ and each codeword $u^n(l), l \in [1 : 2^{nR_s}]$ is picked independently from distribution $\prod_{i=1}^n p_U(u_i(l))$. The codebook is partitioned into bins $\{\mathcal{C}(m_0,m_1,k)\}$, $m_0 \in [1 : 2^{nR_0}]$, $m_1 \in [1 : 2^{nR_K}]$ and $k \in [1 : 2^{nR_K}]$.

To send message message $w(j)$, codeword $u^n(L)$ is selected uniformly randomly from bin $\mathcal{C}(w_0(j), w_1'(j), k_1(j-1) \oplus k_2(j-1))$. Then, the encoder employs Shannon strategy [7] such that $u_i(L)$ along with $\tilde{s}_{i_1}, \tilde{s}_{i_2}$ is passed through function g to generate v_i. Stochastic mapping $p_{X|V,\tilde{S}}(x|v,\tilde{s}_{1_i},\tilde{s}_{2_i})$ is used to map $\tilde{s}_{1_i}, \tilde{s}_{2_i}, v_i$ to channel input x_i. This mapping process is repeated for each time instant i independently.

Decoding and Secrecy Analysis: Using a typical set decoding, both decoders decode L with an arbitrarily low proba-

[2] The entropy term on the left is computer over (10) in whereas the one on the right is computed over (11)

bility error since $\tilde{R} < \min\{I(U;Y_1,S_1,A), I(U;Y_2,S_2,A)\}$. Decoder 2 decrypts the key and message as $k_1(j-1) = k_1(j-1)\oplus k_2(j-1)\oplus k_2(j-1)$ and $w_1(j) = w_1'(j)\oplus k_1(j-1)$.

To show that the first part of the message $w_0(j)$ is kept secure, we can benefit from the ideas in Wyner wiretap coding [1]. Here, $w_1(j)\oplus k_1(j-1)$, and $k_1(j-1)\oplus k_2(j-1)$ can be considered as the amount of randomization added to $w_0(j)$ to keep $w_0(j)$ secure. To show the second part of message $w_1(j)$ is kept secure, we need to show key $k_1(j-1)$, which is one-time padded with $w_1(j)$, is hidden from the eavesdropper. Next proposition presents conditions under which K_1 and K_2 are kept secure from the eavesdropper.

Proposition 1. *Suppose that $R_K < H(\tilde{S}_1|Z,S_3,A)$, $R_K < H(\tilde{S}_2|Z,S_3,A)$, $2R_K < H(\tilde{S}_1,\tilde{S}_2|Z,S_3,A)$, and $\tilde{R} > I(U;Z,S,A)$. For any $\delta>0$ and $j\in[1:b]$,*

1) $H(K_1(j),K_2(j)|\mathcal{B}_1,\mathcal{B}_2) \geq n(2R_K-\delta)$
2) $I(K_1(j),K_2(j);Z^n(j),S_3(j)|\mathcal{C}) \leq n\delta$
3) $I(K_1(j),K_2(j);Z^{jn},S_3(j)|\mathcal{C}) \leq n\delta$

for sufficiently large n.

IV. Conclusion

In this paper, we established a lower bound to the secrecy capacity of the three-receiver wiretap channel with state, subject to a probing cost incurred for the transmitter to acquire the channel states of the legitimate receivers. In our achievability strategy, we used a block Markov coding strategy in which the channel state sequence of the previous block is mapped to a key. We show that the achievable key rate increases linearly with constraint on the probing cost constraint.

V. Appendix: Proof of Proposition 1

We leave the proofs of part 2 and part 3 to the technical report. Here, we prove the generalized version of the first part of Proposition 1 . When $R_{K_1} = R_{K_2} = R_K$, Proposition 2 reduces to the first part of Proposition 1.

Proposition 2. *Suppose that $R_1 < H(S_1)$, $R_2 < H(S_2)$, and $R_1+R_2 < H(S_1,S_2)$. For any $\delta>0$ and $j\in[1:b]$, $H(K_1(j),K_2(j)|\mathcal{B}_1,\mathcal{B}_2) \geq n(R_1+R_2-\delta)$ for sufficiently large n.*

Proof. For the sake of simplicity, we omit block index j in the proof. Hence, $K_1 \triangleq K_1(j)$, $K_2 \triangleq K_2(j)$, $S_1^n \triangleq S_1^n(j)$, and $S_2^n \triangleq S_2^n(j)$. Consider

$$H(K_1,K_2|\mathcal{B}_1,\mathcal{B}_2) \tag{13}$$
$$\geq \mathbb{P}\left[(S_1^n,S_2^n)\in T_\epsilon^n\right] \times H\left(K_1,K_2|\mathcal{B}_1,\mathcal{B}_2,(S_1^n,S_2^n)\in T_\epsilon^n\right)$$
$$\geq (1-\epsilon_n)\times H\left(K_1,K_2|\mathcal{B}_1,\mathcal{B}_2,(S_1^n,S_2^n)\in T_\epsilon^n\right) \tag{14}$$

where $\epsilon_n\to 0$ as $n\to\infty$ and $T_\epsilon^n = T_\epsilon^n(S_1,S_2)$.

$$H\left(K_1,K_2|\mathcal{B}_1,\mathcal{B}_2,(S_1^n,S_2^n)\in T_\epsilon^n\right) \tag{15}$$
$$= \mathbb{E}\left[-2^{n(R_1+R_2)}\log\left(\hat{p}_{K_1,K_2}(1,1)\right)\hat{p}_{K_1,K_2}(1,1)\right] \tag{16}$$

where $\hat{p}_{K_1,K_2}(1,1) \triangleq p_{K_1,K_2}(1,1|\mathcal{B}_1,\mathcal{B}_2,(S_1^n,S_2^n)\in T_\epsilon^n)$. In (16), the expectation is over random variables $\mathcal{B}_1$ and $\mathcal{B}_2$, and $\hat{p}_{K_1,K_2}(1,1)$ is the function of $\mathcal{B}_1$ and $\mathcal{B}_2$.

Now we will analyze the first and second moments of $\hat{p}_{K_1,K_2}(1,1)$. We can write $\hat{p}_{K_1,K_2}(1,1)$ as

$$\hat{p}_{K_1,K_2}(1,1) \tag{17}$$
$$= \mathbb{P}(S_1^n\in\mathcal{B}_1(1),S_2^n\in\mathcal{B}_2(1)|\mathcal{B}_1,\mathcal{B}_2,(S_1^n,S_2^n)\in T_\epsilon^n)$$
$$= \mathbb{P}(S_1^n\in\mathcal{B}_1(1),S_2^n\in\mathcal{B}_2(1)|\,(S_1^n,S_2^n)\in T_\epsilon^n)$$
$$= \sum_{s_1^n\in B_1(1)}\sum_{s_2^n\in B_2(1)} p_{S_1^n,S_2^n}(s_1^n,s_2^n|\,(S_1^n,S_2^n)\in T_\epsilon^n) \tag{18}$$
$$= \sum_{(s_1^n,s_2^n)\in T_\epsilon^n} \frac{p_{S_1^n,S_2^n}(s_1^n,s_2^n)}{\mathbb{P}\left((S_1^n,S_2^n)\in T_\epsilon^n\right)} I_{s_1^n\in\mathcal{B}_1(1)} I_{s_2^n\in\mathcal{B}_2(1)}. \tag{19}$$

The expectation of $\hat{p}_{K_1,K_2}(1,1)$ is

$$\mathbb{E}\left[\hat{p}_{K_1,K_2}(1,1)\right] \tag{20}$$
$$= \sum_{(s_1^n,s_2^n)\in T_\epsilon^n} \frac{p_{S_1^n,S_2^n}(s_1^n,s_2^n)}{\mathbb{P}\left((S_1^n,S_2^n)\in T_\epsilon^n\right)} \mathbb{E}\left[I_{s_1^n\in\mathcal{B}_1(1)} I_{s_2^n\in\mathcal{B}_2(1)}\right]$$
$$= 2^{-n(R_1+R_2)}. \tag{21}$$

where (21) follows from the fact that $\mathcal{B}_1(1)$ and $\mathcal{B}_2(1)$ are independent random variables. Next we upper bound $\text{VAR}\left[\hat{p}_{K_1,K_2}(1,1)\right]$ as

$$\text{VAR}\left[\hat{p}_{K_1,K_2}(1,1)\right] \tag{22}$$
$$= \sum_{(s_1^n,s_2^n)\in T_\epsilon^n}\sum_{(s_3^n,s_4^n)\in T_\epsilon^n} \Bigg(\frac{p_{S_1^n,S_2^n}(s_1^n,s_2^n)p_{S_1^n,S_2^n}(s_3^n,s_4^n)}{\mathbb{P}\left((S_1^n,S_2^n)\in T_\epsilon^n\right)^2}$$
$$\times \text{COV}\left(I_{s_1^n\in B_1(1)}I_{s_2^n\in B_2(1)}, I_{s_3^n\in B_1(1)}I_{s_4^n\in B_2(1)}\right)\Bigg) \tag{23}$$
$$\leq \sum_{(s_1^n,s_2^n)\in T_\epsilon^n}\sum_{(s_3^n,s_4^n)\in T_\epsilon^n} \Bigg(\frac{p_{S_1^n,S_2^n}(s_1^n,s_2^n)p_{S_1^n,S_2^n}(s_3^n,s_4^n)}{(1-\epsilon_n)^2}$$
$$\times \text{COV}\left(I_{s_1^n\in B_1(1)}I_{s_2^n\in B_2(1)}, I_{s_3^n\in B_1(1)}I_{s_4^n\in B_2(1)}\right)\Bigg) \tag{24}$$

The covariance term in (24) can be written as

$$\text{COV}\left(I_{s_1^n\in B_1(1)}I_{s_2^n\in B_2(1)}, I_{s_3^n\in B_1(1)}I_{s_4^n\in B_2(1)}\right) \tag{25}$$
$$= \begin{cases} 2^{-n(R_1+R_2)}(1-2^{-n(R_1+R_2)}) \text{ if } s_1^n=s_3^n, s_2^n=s_4^n \\ 2^{-n(R_1+2R_2)}(1-2^{-nR_1}) \text{ if } s_1^n=s_3^n, s_2^n\neq s_4^n \\ 2^{-n(2R_1+R_2)}(1-2^{-nR_2}) \text{ if } s_1^n\neq s_3^n, s_2^n=s_4^n \\ 0 \text{ if } s_1^n\neq s_3^n, s_2^n\neq s_4^n \end{cases} \tag{26}$$

We continue bounding $\text{VAR}\left[\hat{p}_{K_1,K_2}(1,1)\right]$ with the following:

$$\text{VAR}\left[\hat{p}_{K_1,K_2}(1,1)\right]$$
$$\leq \frac{2^{-n(R_1+R_2)}(1-2^{-n(R_1+R_2)})}{(1-\epsilon_n)^2} \sum_{s_1^n,s_2^n\in T_\epsilon^n} p_{S_1^n,S_2^n}(s_1^n,s_2^n)^2$$
$$+ \frac{2^{-n(R_1+2R_2)}(1-2^{-nR_1})}{(1-\epsilon_n)^2} \times \sum_{s_1^n,s_2^n,s_3^n,s_4^n\in A_1} p_{S_1^n,S_2^n}(s_1^n,s_2^n)p_{S_1^n,S_2^n}(s_3^n,s_4^n)$$
$$+ \frac{2^{-n(2R_1+R_1)}(1-2^{-nR_2})}{(1-\epsilon_n)^2}$$

$$\times \sum_{s_1^n,s_2^n,s_3^n,s_4^n \in A_2} p_{S_1^n,S_2^n}(s_1^n,s_2^n) p_{S_1^n,S_2^n}(s_3^n,s_4^n) \quad (27)$$

$$\leq \frac{2^{-n(R_1+R_2)}}{(1-\epsilon_n)^2} 2^{-2n(H(S_1,S_2)-\delta_1(\epsilon))} \sum_{s_1^n,s_2^n \in T_\epsilon^n} 1$$
$$+ \frac{2^{-n(R_1+2R_2)}}{(1-\epsilon_n)^2} 2^{-2n(H(S_1,S_2)-\delta_1(\epsilon))} \sum_{s_1^n,s_2^n,s_3^n,s_4^n \in A_1} 1$$
$$+ \frac{2^{-n(2R_1+R_1)}}{(1-\epsilon_n)^2} 2^{-2n(H(S_1,S_2)-\delta_1(\epsilon))} \sum_{s_1^n,s_2^n,s_3^n,s_4^n \in A_2} 1 \quad (28)$$

where $\delta_1(\epsilon) = \epsilon H(S_1,S_2)$ and sets A_1 and A_2 are defined as

$$\{s_1^n,s_2^n,s_3^n,s_4^n : s_1^n,s_2^n \in T_\epsilon^n, s_3^n,s_4^n \in T_\epsilon^n, s_1^n = s_3^n, s_2^n \neq s_4^n\}$$
$$\{s_1^n,s_2^n,s_3^n,s_4^n : s_1^n,s_2^n \in T_\epsilon^n, s_3^n,s_4^n \in T_\epsilon^n, s_1^n \neq s_3^n, s_2^n = s_4^n\}$$

respectively. To upper bound (28), we bound the sizes of A_1 and A_2. With the following analysis, we upper bound the size of A_1:

$$|A_1| = |s_1^n,s_2^n,s_4^n : s_2^n \in T_\epsilon^n(S_2|s_1^n), s_4^n \in T_\epsilon^n(S_2|s_1^n), s_2^n \neq s_4^n|$$
$$\leq |s_1^n,s_2^n,s_4^n : s_2^n \in T_\epsilon^n(S_2|s_1^n), s_4^n \in T_\epsilon^n(S_2|s_1^n)|$$
$$\leq \sum_{s_1^n \in T_\epsilon^n} |s_2^n,s_4^n : s_2^n \in T_\epsilon^n(S_2|s_1^n), s_4^n \in T_\epsilon^n(S_2|s_1^n)| \quad (29)$$
$$\leq 2^{n(H(S_1)+2H(S_2|S_1)+\delta_2(\epsilon))} = 2^{n(H(S_1,S_2)+H(S_2|S_1)+\delta_2(\epsilon))} \quad (30)$$

where $\delta_2(\epsilon) = \epsilon\left(H(S_1,S_2) + H(S_2|S_1)\right)$. With a similar analysis, we show that

$$|A_2| \leq 2^{n(H(S_1,S_2)+H(S_1|S_2)+\delta_3(\epsilon))} \quad (31)$$

where $\delta_3(\epsilon) = \epsilon\left(H(S_1,S_2) + H(S_1|S_2)\right)$.

Next, we combine the bounds on $|A_1|$ and $|A_2|$ with (28) as

$$\text{VAR}\left[\hat{p}_{K_1,K_2}(1,1)\right] \leq \frac{2^{-n(R_1+R_2)}}{(1-\epsilon_n)^2} 2^{-n(H(S_1,S_2)-3\delta_1(\epsilon))}$$
$$+ \frac{2^{-n(R_1+2R_2)}}{(1-\epsilon_n)^2} 2^{-n(H(S_1)-\delta_4(\epsilon))}$$
$$+ \frac{2^{-n(2R_1+R_2)}}{(1-\epsilon_n)^2} 2^{-n(HS_2)-\delta_5(\epsilon))} \quad (32)$$

where $\delta_4(\epsilon) = \epsilon(3H(S_1,S_2) + H(S_2|S_1))$ and $\delta_5(\epsilon) = \epsilon(3H(S_1,S_2) + H(S_1|S_2))$.

By the Chebyshev inequality and by the bound in (32),

$$\mathbb{P}\left(|\hat{p}_{K_1,K_2}(1,1) - \mathbb{E}\left[\hat{p}_{K_1,K_2}(1,1)\right]| \geq \epsilon \mathbb{E}\left[\hat{p}_{K_1,K_2}(1,1)\right]\right)$$
$$\leq \frac{\text{VAR}\left[\hat{p}_{K_1,K_2}(1,1)\right]}{\epsilon^2 \mathbb{E}\left[\hat{p}_{K_1,K_2}(1,1)\right]^2} = \frac{\text{VAR}\left[\hat{p}_{K_1,K_2}(1,1)\right]}{\epsilon^2 2^{-2n(R_1+R_2)}} \quad (33)$$
$$\leq \frac{2^{-n(H(S_1,S_2)-R_1-R_2-3\delta_1(\epsilon))}}{\epsilon^2(1-\epsilon_n)^2} + \frac{2^{-n(H(S_1)-R_1-\delta_4(\epsilon))}}{\epsilon^2(1-\epsilon_n)^2}$$
$$+ \frac{2^{-n(H(S_2)-R_2-\delta_5(\epsilon))}}{\epsilon^2(1-\epsilon_n)^2}. \quad (34)$$

If we choose ϵ small enough such that the followings are satisfied

$$3\delta_1(\epsilon) < H(S_1,S_2) - (R_1+R_2) \quad (35)$$
$$\delta_4(\epsilon) < H(S_1) - R_1 \quad (36)$$
$$\delta_5(\epsilon) < H(S_2) - R_2 \quad (37)$$

then, (34) $\to 0$ as $n \to \infty$

We are now ready to prove the lemma.

$$H(K_1,K_2|B_1,B_2) \geq (1-\epsilon_n)\mathbb{P}(\mathcal{A})$$
$$\times \mathbb{E}\left[-2^{n(R_1+R_2)} \log\left(\hat{p}_{K_1,K_2}(1,1)\right) \hat{p}_{K_1,K_2}(1,1) | \mathcal{A}\right]$$
$$\geq n\left((R_1+R_2) - \epsilon(R_1+R_2) - \frac{(1-\epsilon)\log(1+\epsilon)}{n}\right)$$
$$\times (1-\epsilon_n)(1-\mathbb{P}(\mathcal{A}^c)) \quad (38)$$
$$\geq n\Bigg((R_1+R_2) - (R_1+R_2)(\mathbb{P}(\mathcal{A}^c)+\epsilon_n) + \epsilon(R_1+R_2)$$
$$- \frac{(1-\epsilon)\log(1+\epsilon)}{n}\Bigg) \quad (39)$$
$$\geq n(R_1+R_2-\delta) \quad (40)$$

where

$$\mathcal{A} = \left\{|\hat{p}_{K_1,K_2}(1,1) - \mathbb{E}\left[\hat{p}_{K_1,K_2}(1,1)\right]| < \epsilon \mathbb{E}\left[\hat{p}_{K_1,K_2}(1,1)\right]\right\}.$$

Note that in (34) we show that $\mathbb{P}(\mathcal{A}^c) \to 0$ as $n \to \infty$ if conditions (35)-(37) are satisfied. In the derivation above, (38) and (39) follow from a simple algebra and (40) follows from the fact that term $(R_1+R_2)(\mathbb{P}(\mathcal{A}^c)+\epsilon_n) + \epsilon(R_1+R_2) - \frac{(1-\epsilon)\log(1+\epsilon)}{n}$ can be made arbitrarily small for sufficiently large n and for sufficiently small ϵ. □

REFERENCES

[1] A. D.Wyner, "The wire-tap channel". *Bell Syst. Tech. J.*, 54(8):1355–1387, October 1975.

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

[3] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.

[4] Y. Liang, H. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Trans. on Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, June 2008.

[5] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5059–5067, Nov. 2008.

[6] H. Asnani, H. H. Permuter, and T. Weissman, "Probing capacity, *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 73177332, Nov. 2011

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

[8] I. Csiszar and P. Narayan, "Common randomness and secret key generation with a helper, *IEEE Trans. Inf. Theory*, vol. 46, Mar. 2000.

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

[10] Y.-K. Chia and A. El Gamal, "Wiretap channel with causal state information, *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2838-2849, 2012.

[11] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels, *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453-2469, June 2008.

[12] Y. O. Basciftci and C. E. Koksal, "Private Broadcasting with Probing Constraint", http://www2.ece.ohio-state.edu/~koksal/papers/control_secrecy_asilomar_10.pdf