On Simultaneously Generating Multiple Keys in Joint Source-Channel Model

Wenwen Tu, Mario Goldenbaum, Lifeng Lai and H. Vincent Poor

***Abstract*—In this paper, the problem of simultaneously generating multiple keys over a cascade of a noiseless channel and a wiretap channel is considered. The problem consists of three legitimate parties (i.e., Alice, Bob, and Carol), where Alice and Bob wish to agree with Carol on independent secret keys. Alice and Bob are connected via a noiseless channel, and Bob is connected with Carol via a wiretap channel, while there is no direct connection between Alice and Carol. To Alice and Carol, Bob acts as a relay. Under this model, we first provide a full characterization of the secret-key capacity region for the case when Eve has no side information. The result shows that there exists a trade-off between the individual secret-key rates. Then we generalize the obtained result into the case when Eve has side information, and fully characterize the corresponding secret-key capacity region.**

***Index Terms*—Cascaded channels, correlated sources, key capacity region, simultaneous key generation, wiretap channel.**

I. Introduction

Enabling communication parties to share a common secret key is a fundamental problem in cryptography. Recently, the paradigm of secret key generation via public discussion, under both source and channel models, has received significant attention [1]–[7]. Under the source model, the legitimate terminals have access to correlated random sequences, from which they can generate a secret key by exchanging messages over a public noiseless channel fully accessible to an eavesdropper [8]–[11]. On the other hand, under the channel model the legitimate terminals usually have no access to correlated random sequences, but they can utilize the differences between the channel connected to the legitimate receiver and the channel connected to the eavesdropper to generate a secret key [12]–[15]. As the problem is typically approached either from a source or a channel perspective, Khisti et al. recently introduced a new *joint source-channel model* for the problem of key generation [16]. Under this model, they provided both a lower and an upper bounds for the key capacity. Furthermore, [16] also contains a full characterization of the key capacity when certain markov chain conditions are satisfied.

One important assumption in the existing works is that the public discussion is directly available to *all* legitimate users. While it is important to assume that the public discussion is available to Eve (so that the generated key is secure in the worst case scenario), there are some practical scenarios in which the public discussion is directly received only by a subset of the legitimate users. For example, in key generation over wireless networks [17], public discussion messages are transmitted over wireless channels. Hence, it is reasonable to assume that public discussion messages are directly received only by neighboring legitimate users. In this case, the assumption that the public discussion is directly available to all legitimate users is too optimistic.

To gain some understanding of scenarios with limited direct access to the public discussion by certain legitimate users, we consider an extension of the joint source-channel model of [16]. In our model, there are three legitimate users: Alice, Bob, and Carol. Alice and Bob are connected by a noiseless public channel (Eve can observe this noiseless channel), and Bob is connected with Carol via a noisy channel (Eve can also eavesdrop on this channel). However, Alice has no direct connection with Carol and therefore Carol does not have direct access to the public discussion messages sent by Alice. This network setting captures many relevant scenarios such as the scenario where Alice is a server who connects with the base station Bob over an optical fiber, which can be viewed as noiseless, and Carol is a wireless user. Furthermore, we assume that Alice and Carol have access to correlated random sources.

Under this network topology, we consider the problem of simultaneously generating two secret keys: One between Alice and Carol, and one between Bob and Carol. The problem of simultaneously generating multiple keys is well motivated in applications in which multiple keys are needed for different communication sessions [18]–[22]. In our setup, we require that the key generated by Alice and Carol is secure from Bob and Eve, while the key generated by Bob and Carol is secure from Alice and Eve. We first consider a case where Eve has no side information, and fully characterize the secret-key capacity region. Then, we generalize the considered model to the case when Eve has side information, and obtain a full characterization of the corresponding capacity region as well. It turns out that if we only care about the key between Alice and Carol, the considered model can be simplified to the source model with one-way limited-rate public discussion as studied in [9], and we show that our result recovers the result in [9]. On the other hand, if we only care about the

W. Tu and L. Lai are with Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA. The work of W. Tu and L. Lai was supported by the National Science Foundation CAREER Award under Grant CCF-1318980 and by the National Science Foundation under Grant ECCS-1408114. Email: {wtu,llai}@wpi.edu.

M. Goldenbaum and H. V. Poor are with Department of Electrical Engineering, Princeton University, Princeton, NJ. The work of M. Goldenbaum was supported in part by the German Research Foundation (DFG) under grant GO 2669/1-1. The work of H. V. Poor was supported by the National Science Foundation under Grants CMMI-1435778 and CNS-1456793. Email:{goldenbaum,poor}@princeton.edu.

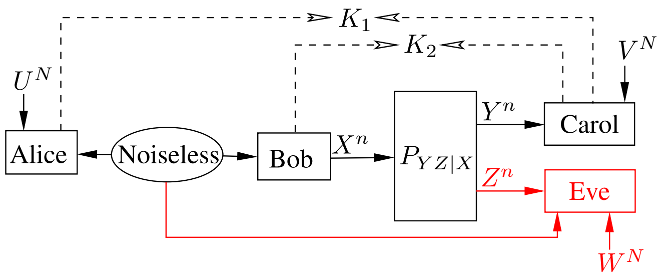

Fig. 1. System model: Alice and Bob wish to share individual secret keys K_1 and K_2 with Carol respectively. The link between Alice and Bob is assumed to be noiseless, while Bob and Carol are connected via a wiretap channel. Besides, Alice, Carol and Eve have access to correlated sequences U^N, V^N and W^N respectively.

key between Bob and Carol, the model can be viewed as a wiretap channel [14] and our result recovers that of the wiretap channel. Furthermore, there is a trade-off between the two cases so that Alice and Bob cannot attain their maximal secret-key rates simultaneously.

In addition to the work mentioned above, our work is related to recent papers on simultaneously generating multiple keys in networks consisting of trusted and untrusted parties [18]–[22]. The main differences between our model and models in these papers are: 1) we consider a joint source-channel model; and 2) we assume that the public discussion is not directly available to all users.

The remainder of the paper is organized as follows. The system model and the problem setup are introduced in Section II. In Section III, we consider the case when Eve has no side information. Then, we generalize our obtained results to the case when Eve has side information in Section IV. Finally, we offer our concluding remarks in Section V.

II. System Model and Problem Statement

As illustrated in Fig. 1, we consider a scenario in which Alice and Carol wish to agree on a secret key K_1 taking values from $\mathcal{K}_1$, while Bob wishes to agree with Carol on a secret key K_2 taking values from $\mathcal{K}_2$. Under this model, K_1 is required to be kept confidential from Bob and Eve, while K_2 is required to be kept confidential from Alice and Eve.

Unlike Bob who can communicate with Carol over a noisy channel eavesdropped by Eve, Alice has no direct connection with Carol and therefore she needs assistance from Bob. The link between Bob and Carol is modeled as a wiretap channel $(\mathcal{X}, P_{YZ|X}, \mathcal{Y}, \mathcal{Z})$, where $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ denote finite channel input and output alphabets. Alice and Bob can communicate through a noiseless link. However, any message exchanged over this noiseless link will also be perfectly overheard by Eve .

Alice, Carol and Eve are assumed to have access to three correlated random sequences U^N, V^N and W^N, $N \in \mathbb{N}$, which are generated according to a given joint probability mass function (PMF)

$$P_{U^N V^N W^N}(u^N, v^N, w^N) = \prod_{i=1}^{N} P_{UVW}(u_i, v_i, w_i), \qquad (1)$$

where U, V and W take values from the finite alphabets $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$, respectively.

Definition 1. *An (N, n) key-agreement protocol for the joint source-channel model is as follows.*

- ***Step** 0). Alice generates a random variable F_0, and Bob generates another random variable F_0'. F_0 and F_0' are mutually independent, and are independent with all other random variables in the model.*
- ***Step** 1). Alice and Bob exchange messages f_1 and f_1', where $f_1 \triangleq f_1(F_0, U^N)$ and $f_1' \triangleq f_1'(F_0')$, over the noiseless channel.*
- ***Step** i). Alice and Bob exchange messages $f_i(F_0, U^N, \mathbf{f}'^{i-1})$ and $f_i'(F_0', \mathbf{f}^{i-1})$, in which $\mathbf{f}^{i-1} \triangleq (f_1, \cdots, f_{i-1})$ and $\mathbf{f}'^{i-1}$ is defined in a similar manner.*
- ***Step** k). (After Alice and Bob finish their discussion) Denote $\mathbf{F} := (\mathbf{f}^{k-1}, \mathbf{f}'^{k-1})$. Bob generates another independent random variable F_b and transmits $X^n(\mathbf{F}, F_b)$ into the wiretap channel.*
- ***Final step**). Alice computes a key via a function $K_1 \triangleq K_1(U^N, \mathbf{F}, F_0)$; Bob computes a key via a function $K_2 \triangleq K_2(\mathbf{F}, F_0', F_b)$; Carol computes two keys via functions $K_1' \triangleq K_1'(Y^n, V^N), K_2' \triangleq K_2'(Y^n, V^N)$.*

Here, the use of random variables F_0 and F_0' enables the messages exchanged over the public noiseless channel to be random functions of U^N, while F_b ensures that Bob can use stochastic coding to generate his own key with Carol. Throughout the paper, for notational convenience, we let $\beta = \frac{n}{N}$.

Definition 2. *A secret-key rate pair (R_1, R_2) is said to be achievable if $\forall \epsilon > 0$ there exists an $n(\epsilon) \in \mathbb{N}$ and a sequence of (N, n) codes such that $\forall n \geq n(\epsilon)$, we have*

$$\Pr\{K_i \neq K_i'\} \leq \epsilon, \quad i = 1, 2, \qquad (2)$$

$$\frac{1}{n} I(K_1; \mathbf{F}, F_0', F_b) \leq \epsilon, \qquad (3)$$

$$\frac{1}{n} I(K_2; \mathbf{F}, F_0, U^N) \leq \epsilon, \qquad (4)$$

$$\frac{1}{n} I(K_1, K_2; \mathbf{F}, Z^n, W^N) \leq \epsilon, \qquad (5)$$

$$\frac{1}{n} H(K_i) \geq \frac{1}{n} \log |\mathcal{K}_i| - \epsilon, \quad i = 1, 2, \qquad (6)$$

$$\frac{1}{n} H(K_1) \geq R_1 - \epsilon, \ \frac{1}{n} H(K_2) \geq R_2 - \epsilon. \qquad (7)$$

Here, (2) indicates that the keys generated at the key generating parties should be the same with high probability, (3) means that K_1 is required to be secure from Bob, (4) means that K_2 should be secure from Alice, (5) implies that (K_1, K_2) should be jointly secure from Eve, and (6) indicates that the generated keys should be nearly uniformly distributed.

Definition 3. *The secret-key capacity region $\mathcal{C}$ is defined as:*

$$\mathcal{C} \triangleq \left\{(R_1, R_2) \in \mathbb{R}_+^2 \,|\, (R_1, R_2) \, \text{is achievable}\right\}.$$

Furthermore, we use C_1 to denote the maximal value of R_1 (Key capacity of K_1), C_2 to denote the maximal value of R_2

(Key capacity of K_2) and C_{sum} to denote the maximal value of $R_1 + R_2$ (Sum capacity of (K_1, K_2)).

III. Capacity Region with No Side Information at Eve

In this section, to facilitate the presentation and understanding of our scheme, we first consider the special case when Eve has no side information, i.e., the case where $\mathcal{W} = \emptyset$, and denote the corresponding secret-key capacity region by $\mathcal{C}_0$. For this case, we fully characterize $\mathcal{C}_0$. The results will be extended to the general model with side information at Eve in Section IV.

For auxiliary random variables M, S and T satisfying $M - U - V$ and $T - S - X - (Y, Z)$, define

$$\begin{aligned} &\mathcal{R}(P_{M|U}, P_{TS}P_{X|S}) \triangleq \\ &\quad \{(R_1, R_2) : R_1 \leq \frac{1}{\beta} I(M;V), \\ &\qquad\qquad R_2 \leq \big[I(S;Y|T) - I(S;Z|T)\big]^+, \qquad (8) \\ &\qquad \text{s.t.} \quad I(M;U) - I(M;V) \leq \beta I(T;Y).\} \qquad (9) \end{aligned}$$

Here, $[x]^+ = \max\{0, x\}$. Furthermore, the notation $M - U - V$ means that random variables (M, U, V) form a Markovian chain in that order. $T - S - X - (Y, Z)$ (and other similar relationships throughout the paper) is defined in a similar manner.

We have the following result.

Theorem 1. *The secret-key capacity region for the case with no side information at Eve is*

$$\mathcal{C}_0 = \bigcup_{P_{M|U}, P_{TS}P_{X|S}} \mathcal{R}(P_{M|U}, P_{TS}P_{X|S}). \qquad (10)$$

Proof. The proof contains two parts: converse and achievability. In the converse proof presented in Appendix A-A, we show that (10) is an outer bound. In the achievability part, we show that for any given $(P_{M|U}P_{UV}, P_{TS}P_{X|S})$, rate pair (R_1, R_2) with

$$\begin{aligned} &R_1 = \frac{1}{\beta} I(M;V) - \epsilon, \;\; R_2 = \big[I(S;Y|T) - I(S;Z|T)\big]^+ - \epsilon \\ &\text{s.t.} \quad I(M;U) - I(M;V) \leq \beta I(T;Y), \qquad (11) \end{aligned}$$

is achievable, and hence the region specified in (10) is achievable. Detailed proof of the achievability part is provided in Appendix A-B. Here, we provide a high level idea of how the achievability scheme works. As illustrated in Fig. 2, the codebook construction is a combination of source coding techniques and channel coding techniques. From Alice and Carol's perspective, the noisy channel $P_{Y|X}$ acts as a noiseless channel with rate $I(T;Y)$. This guarantees that, if messages sent by Alice have a rate less than $I(T;Y)$, they can be correctly decoded by Carol using Y^n with high probability. As a result, the key generation model between Alice, Carol and Eve can be viewed as a source model (with no side information at Eve) using one way public discussion with rate constraint, and the rate $\frac{1}{\beta} I(M;V) - \epsilon$ is achievable using techniques for this model. In particular, to generate K_1, we generate $2^{N(I(M;U)+\epsilon)}$ sequences M^Ns, and randomly assign them into $2^{N(I(M;U)-I(M;V)+2\epsilon)}$ bins (we choose the number of bins to guarantee that its rate is less than $I(T;Y)$). Alice then sends the bin index to Carol through Bob. With this bin index along with its source observation V^N, Carol will be able to decode M^N. We will obtain a key at the rate of $\frac{1}{\beta} I(M;V) - \epsilon$ by setting the sub-bin index of the decoded M^N as the key value of K_1. At Bob's side, Bob chooses T^n to convey the received message to Carol, while using S^n generated by the selected T^n to generate his own key with Carol: We generate $2^{n(I(T;Y)-\epsilon)}$ sequences T^ns. For each T^n we generate $2^{n(I(S;Y|T)-\epsilon)}$ sequences S^ns and randomly assign them into $2^{n(I(S;Y|T)-I(S;Z|T)-2\epsilon)}$ bins. We set the bin index of S^n as the key value of K_2. □

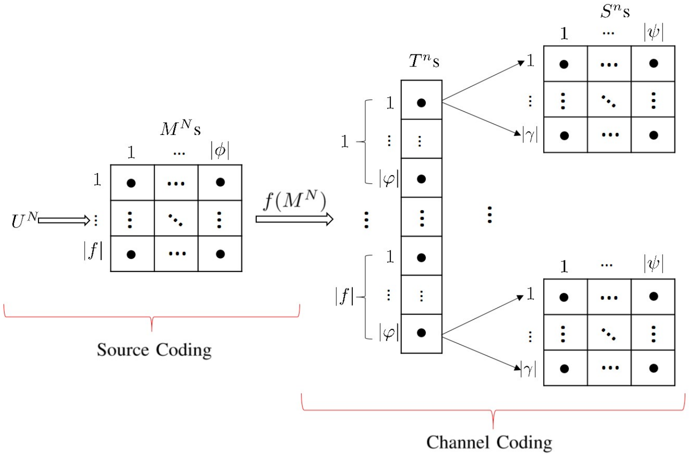

Fig. 2. Codebook construction

Corollary 1. *The secret-key capacity of K_1 for the case with no side information at Eve is*

$$\begin{aligned} &C_1 = \max_{M-U-V} \frac{1}{\beta} I(M;V) \\ &\textit{s.t.} \quad I(M;U) - I(M;V) \leq \max_{P_X} \beta I(X;Y). \qquad (12) \end{aligned}$$

Proof. According to Theorem 1, the following rate is achievable for K_1

$$\begin{aligned} &R_1 \leq \frac{1}{\beta} I(M;V), \\ &\text{s.t.} \quad I(M;U) - I(M;V) \leq \beta I(T;Y). \qquad (13) \end{aligned}$$

Due to $T - X - Y$, we conclude that (13) is contained in the region

$$\begin{aligned} &\{R_1 \in \mathbb{R}_+ : R_1 \leq \frac{1}{\beta} I(M;V), \\ &\quad \text{s.t.} \quad I(M;U) - I(M;V) \leq \beta I(X;Y).\}. \qquad (14) \end{aligned}$$

And (14) is achievable by setting $T = X$. Hence, (12) can be obtained via maximizing (14). □

Corollary 1 shows that if one only cares about the key K_1, the channel between Alice and Carol can be viewed as a noiseless channel with rate constraint $R = \max_{P_X} \beta I(X;Y)$ and our problem is equivalent to the problem of generating a single key with one-way public discussion subject to rate constraint as studied in [9, Sec. II. Case 6]. Our result is consistent with [9, Thm. 2.4]. Note that even though we allow multiple rounds discussion over the public noiseless channel in our model, the public discussion is between Alice and Bob, not between Alice and Carol. And Carol is connected to this noiseless channel via a wiretap channel, which is a one-way link. Thus, the link between Alice and Carol can be viewed as a one-way channel with rate constraint.

Corollary 2. *The secret-key capacity of K_2 for the case with no side information at Eve is*

$$C_2 = \max_{P_{SX}} \{I(S;Y) - I(S;Z)\}. \tag{15}$$

Proof. According to Theorem 1, we have

$$\begin{aligned} C_2 &= \max_{P_{TS}P_{X|S}} \{I(S;Y|T) - I(S;Z|T)\} \\ &= \max_{P_{TS}P_{X|S}} \sum_t P_T(t)\big[I(S;Y|T=t) - I(S;Z|T=t)\big] \\ &\overset{(a)}{\leq} \max_{P_{TS}P_{X|S}} \max_t \big[I(S;Y|T=t) - I(S;Z|T=t)\big] \\ &= \max_{P_{SX}} \{I(S;Y) - I(S;Z)\}. \end{aligned} \tag{16}$$

The equality of (a) can be obtained by setting T be some constant. □

Corollary 2 shows that if one only cares about K_2, the key capacity is the same as the capacity of a discrete memoryless wiretap channel. This implies that the correlated sources (U^N, V^N) do not help in increasing R_2, as we require K_2 to be secure from Alice.

Corollary 3. *The sum capacity of (K_1, K_2) for the case with no side information at Eve is*

$$\begin{aligned} C_{\text{sum}} = &\max_{\substack{M-U-V \\ T-S-X-Y,Z}} \{I(S;Y|T) - I(S;Z|T) + \frac{1}{\beta} I(M;V)\}, \\ &\textit{s.t.} \quad I(M;U) - I(M;V) \leq \beta I(T;Y). \end{aligned} \tag{17}$$

Proof. It's a direct result from Theorem 1. □

The plot of $\mathcal{C}_0$ is as shown in Fig. 3, where $\mathcal{C}_0 = \mathcal{R}_1 \bigcup \mathcal{R}_2 \bigcup \mathcal{R}_3$. $\mathcal{R}_1$ is the region where there exists a P_{T^*} such that $\beta I(T^*;Y) \geq H(U|V) = \max\{I(M;U) - I(M;V)\}$ ($\mathcal{R}_1$ vanishes if $H(U|V) \geq \max_{P_X} \beta I(X;Y)$). One doesn't need to sacrifice R_1 in order to obtain a larger R_2 at least when $R_2 \leq \max_{P_{S|T^*}P_{S|X}} \{I(S;Y|T^*) - I(S;Z|T^*)\}$. $\mathcal{R}_3$ is the region obtained when $P_T = \arg\max_{P_T} \max_{P_{S|T}P_{X|S}} \{I(S;Y|T) - I(S;Z|T)\}$ and $I(M;U) - I(M;V) \leq \beta I(T;Y)$ ($\mathcal{R}_3$ vanishes if T is a constant). And in $\mathcal{R}_3$, one doesn't need to sacrifice R_2 in order to obtain a larger R_1. Obviously, in $\mathcal{R}_2$, there exists a tradeoff between R_1 and R_2, and C_{sum} is obtained in this region.

Note that our model is related to the setup in [16], especially when one only cares about C_{sum}. The major difference is that we consider the achievable region while [16] equivalently focuses only on the sum capacity. In addition, Alice and Bob are connected by a noiseless channel in our model while the setup in [16] can be viewed as that Alice and Bob are combined into one terminal. What's more, we require that K_1 is concealed from Bob and K_2 is concealed from Alice while these requirements don't exist in [16].

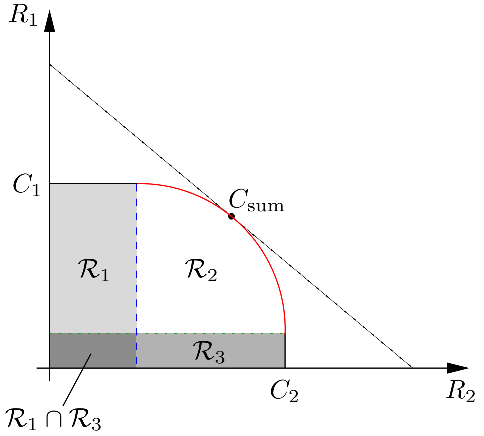

Fig. 3. Secret-key capacity region $\mathcal{C} \triangleq \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3$.

IV. Capacity Region with Side Information at Eve

In this section, we generalize the results obtained in Section III to a more general case when Eve has access to side information, i.e. $\mathcal{W} \neq \emptyset$. Under this model, we fully characterize the corresponding secret-key capacity region.

For auxiliary random variables L, M, S and T satisfying $L - M - U - (V,W)$ and $T - S - X - (Y,Z)$, we define

$$\begin{aligned} &\mathcal{R}(P_{M|U}P_{L|M}, P_{TS}P_{X|S}) \triangleq \\ &\quad \{(R_1,R_2) : R_1 \leq \frac{1}{\beta}\big[I(M;V|L) - I(M;W|L)\big]^+, \\ &\qquad R_2 \leq \big[I(S;Y|T) - I(S;Z|T)\big]^+, \qquad (18) \\ &\qquad \text{s.t.} \quad I(M;U) - I(M;V) \leq \beta I(T;Y).\} \qquad (19) \end{aligned}$$

Then, we have the following result.

Theorem 2. *In the joint source-channel model with side information at Eve, the secret-key capacity region is*

$$\mathcal{C} = \bigcup_{P_{M|U}P_{L|M}, P_{TS}P_{X|S}} \mathcal{R}(P_{M|U}P_{L|M}, P_{TS}P_{X|S}). \tag{20}$$

Proof. This proof contains two parts: converse and achievabily. In the converse part, we will show that any achievable pair (R_1, R_2) is included in $\mathcal{C}$. The converse proof is provided in Appendix C-A. In the achievability part, we will show that for any given $(P_{M|U}P_{L|M}, P_{TS}P_{X|S})$, $\mathcal{R}(P_{M|U}P_{L|M}, P_{TS}P_{X|S})$ is an achievable region. Details of the achievability proof are provided in Appendix C-B. Here, we provide a high level idea of how it works. The codebook of the achievability scheme is illustrated in Fig. 4. To Alice and Carol, the noisy channel between Bob and Carol acts as a noiseless channel with rate constraint

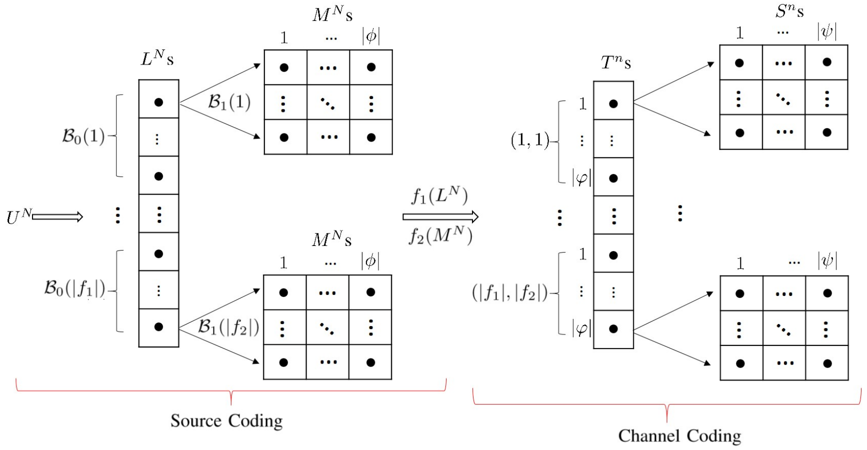

Fig. 4. Codebook construction

$I(T;Y)$, such that the key generation model between Alice, Carol and Eve can be viewed as a source model (with side information at Eve) using one way public discussion with rate constraint. Thus the rate $\frac{1}{\beta}[I(M;V|L) - I(M;W|L) - 2\epsilon]$ is achievable using techniques for this model. In particular, we generate $2^{n(I(L;U)+\epsilon)}$ sequences L^ns, and randomly assign them into $2^{n(I(L;U)-I(L;V)+2\epsilon)}$ bins. For each L^n, we generate $2^{n(I(M;U|L)+\epsilon)}$ sequences M^ns, and randomly assign them into $2^{n(I(M;U|L)-I(M;V|L)+2\epsilon)}$ bins. Within each bin, assign each M^n into $2^{n(I(M;V|L)-I(M;W|L)-2\epsilon)}$ subbins, and set the subbin index as the key value of K_1. At Bob's side, Bob uses $2^{n(I(T;Y)-\epsilon)}$ sequences T^ns to convey the message sent by Alice to Carol. For each T^n, generate $2^{n(I(S;Y|T)-\epsilon)}$ sequences S^ns and randomly assign them into $2^{n(I(S;Y|T)-I(S;Z|T)-2\epsilon})$ bins. We set the bin index of S^n as the key value of K_2. After selecting T^n, Bob randomly selects a S^n to send to Carol, in which way Bob can establish a key with Carol of a rate $I(S;Y|T) - I(S;Z|T) - 2\epsilon$. □

Similar to Corollaries 1 and 3, we have the following corollaries with regard to K_1 and K_2.

Corollary 4. *The secret-key capacity of K_1 for the case with side information at Eve is*

$$\begin{aligned} C_1 = & \max_{L-M-U-(V,W)} \frac{1}{\beta}\big[I(M;V|L) - I(M;W|L)\big] \\ s.t. \quad & I(M;U) - I(M;V) \leq \max_{P_X} \beta I(X;Y). \qquad (21) \end{aligned}$$

We note that, in general, the secret-key capacity with side information at Eve under multiple rounds of public discussion is still unknown [11]. The reason why we are able to characterize the key capacity of K_1 in our model is that, even though we allow multiple rounds of discussion over the public noiseless channel, the public discussion is between Alice and Bob, not between Alice and Carol. In our model, Carol is connected to this noiseless channel via a wiretap channel, which is a one-way link. Since Bob observes no randomness correlated with (U,V,W) in advance, compared with the case of one-way discussion between Alice and Bob, the multiple rounds of discussion between Alice and Bob does not increase the key rate between Alice and Carol. Thus, the link between Alice and Carol can be viewed as a one-way channel with rate constraint.

Corollary 5. *The sum capacity of (K_1, K_2) for the case with side information at Eve is*

$$\begin{aligned} & C_{\text{sum}} = \\ & \max_{\substack{P_{L|M}P_{M|U} \\ P_{TS}P_{X|S}}} I(S;Y|T) - I(S;Z|T) + \frac{1}{\beta}\big[I(M;V|L) - I(M;W|L)\big], \\ & s.t. \quad I(M;U) - I(M;V) \leq \beta I(T;Y). \qquad (22) \end{aligned}$$

V. Concluding Remarks

We have introduced the problem of simultaneously generating multiple secret keys under a cascade model of a noiseless channel and a wiretap channel, using joint correlated sources and channels, to gain some understanding of key generation models with limited access to the public discussion channel. We have fully characterized the secret-key capacity region of the corresponding generated keys under the case when Eve has no side information, and generalized the result to the more general case when Eve has side information.

Appendix A
Proof of Theorem 1

A. Converse

Here, we provide the converse proof of Theorem 1. Before going further, we first introduce a lemma from [11], which will be used frequently in the following.

Lemma 1 (Lemma 4.1 of [11])**.** *For arbitrary random variables U, V and sequences of random variables Y^n, Z^n we have*

$$\begin{aligned} & I(U;Y^n|V) - I(U;Z^n|V) \\ & = \sum_{i=1}^{n} \Big[I(U;Y_i|Y^{i-1}Z_{i+1}^n, V) - I(U;Z_i|Y^{i-1}Z_{i+1}^n, V)\Big]. (23) \end{aligned}$$

Converse of Theorem 1: In this part, we will show that any achievable pair (R_1, R_2) must be in the union defined by the right hand side of (10).

According to the setup, the following Markov relationships are true:

$$\begin{aligned} & V^N - U^N - \mathbf{F} - (Y^n, Z^n), && (24)\\ & V^N - U^N - (\mathbf{F}, K_2) - (Y^n, Z^n). && (25) \end{aligned}$$

Let $\epsilon > 0$ be arbitrary, we have

$$\begin{aligned} H(K_1) &= H(K_1|Y^n, V^N) + I(K_1; Y^n, V^N)\\ &\le I(K_1; Y^n, V^N) + n\epsilon\\ &= I(K_1; Y^n) + I(K_1; V^N|Y^n) + n\epsilon\\ &\le I(K_1; \mathbf{F}) + I(K_1; V^N|Y^n) + n\epsilon\\ &\le \sum_{i=1}^{N} I(K_1; V_i|Y^n, V^{i-1}) + 2n\epsilon\\ &\le \sum_{i=1}^{N} I(K_1, U_{i+1}^n, V^{i-1}, Y^n; V_i) + 2n\epsilon\\ &= \sum_{i=1}^{N} I(M_i; V_i) + 2n\epsilon\\ &= \sum_{i=1}^{N} I(M_Q; V_Q|Q = i) + 2n\epsilon\\ &= N\sum_{i=1}^{N} \frac{1}{N} I(M_Q; V_Q|Q = i) + 2n\epsilon\\ &= NI(M_Q; V_Q|Q) + 2n\epsilon\\ &= NI(M_Q, Q; V_Q) - NI(Q; V_Q) + 2n\epsilon\\ &= NI(M; V) + 2n\epsilon, && (26) \end{aligned}$$

in which $M_i := (K_1, U_{i+1}^n, V^{i-1}, Y^n)$, $M := (M_Q, Q)$, and Q is an independent random variable uniformly distributed over $[1 : N]$.

Thus, we have

$$R_1 \le \frac{1}{\beta} I(M; V) + 2\epsilon. \qquad (27)$$

Furthermore, $M - U - V$ is true as

$$\begin{aligned} & && (U^{i-1}, U_{i+1}^N, V^{i-1}) - U_i - V_i\\ &\Rightarrow && (U^N, V^{i-1}) - U_i - V_i\\ &\Rightarrow && (K_1, \mathbf{F}, U_{i+1}^N, V^{i-1}) - U_i - V_i\\ &\overset{(a)}{\Rightarrow} && (K_1, Y^n, U_{i+1}^N, V^{i-1}) - U_i - V_i\\ &\Leftrightarrow && M_i - U_i - V_i, && (28) \end{aligned}$$

in which (a) is true as Y^n can be seen as a function of $(\mathbf{F}, \theta)$ (θ is some random variable which is independent with all variables in (28)).

Now, we prove (8). We have

$$\begin{aligned} H(K_2) &\le H(K_2) - I(K_2; Z^n, \mathbf{F}) + 2n\epsilon\\ &\overset{(a)}{=} H(K_2) - I(K_2; Z^n, \mathbf{F}, V^N) + 2n\epsilon\\ &= H(K_2|Y^n, \mathbf{F}, V^N) + I(K_2; Y^n, \mathbf{F}, V^N)\\ &\quad - I(K_2; Z^n, \mathbf{F}, V^N) + 2n\epsilon\\ &\le I(K_2; Y^n, \mathbf{F}, V^N) - I(K_2; Z^n, \mathbf{F}, V^N) + 3n\epsilon\\ &= I(K_2; Y^n|\mathbf{F}, V^N) - I(K_2; Z^n|\mathbf{F}, V^N) + 3n\epsilon\\ &= \sum_{i=1}^{n} \Big[I(K_2; Y_i|Y^{i-1}, Z_{i+1}^n, \mathbf{F}, V^N)\\ &\qquad - I(K_2; Z_i|Y^{i-1}, Z_{i+1}^n, \mathbf{F}, V^N) \Big] + 3n\epsilon\\ &= \sum_{i=1}^{n} [I(S_i; Y_i|T_i) - I(S_i; Z_i|T_i)] + 3n\epsilon\\ &= \sum_{i=1}^{n} \Big[I(S_J; Y_J|T_J, J = i)\\ &\qquad - I(S_J; Z_J|T_J, J = i) \Big] + 3n\epsilon\\ &= n\,[I(S; Y|T) - I(S; Y|T)] + 3n\epsilon. && (29) \end{aligned}$$

Here, $S_i := (K_2, V^N, Y^{i-1}, Z_{i+1}^n, \mathbf{F})$, $T_i := (V^N, Y^{i-1}, Z_{i+1}^n, \mathbf{F})$, $S := (S_J, J)$, $T := (T_J, J)$, and J is an independent random variable uniformly distributed over $[1 : n]$. (a) is true due to

$$\begin{aligned} & && \begin{cases} V^N - \mathbf{F} - K_2\\ V^N - (\mathbf{F}, K_2) - Z^n \end{cases}\\ &\Rightarrow && V^N - \mathbf{F} - (Z^n, K_2)\\ &\Rightarrow && V^N - (\mathbf{F}, Z^n) - K_2\\ &\Leftrightarrow && (V^N, Z^n, \mathbf{F}) - (\mathbf{F}, Z^n) - K_2. && (30) \end{aligned}$$

Hence, we have

$$R_2 \le I(S; Y|T) - I(S; Y|T) + 3\epsilon. \qquad (31)$$

Furthermore, we can easily show that $T - S - X - (Y, Z)$.

Now, to prove (9), we first have

$$\begin{aligned} & I(U^N; Y^n) - I(V^N; Y^n)\\ &\le I(\mathbf{F}; Y^n) - I(V^N; Y^n)\\ &= I(\mathbf{F}, V^N; Y^n) - I(V^N; Y^n|\mathbf{F}) - I(V^N; Y^n)\\ &= I(\mathbf{F}; Y^n|V^N) - I(V^N; Y^n|\mathbf{F})\\ &= I(\mathbf{F}; Y^n|V^N)\\ &= \sum_{i=1}^{n} I(\mathbf{F}; Y_i|Y^{i-1}, V^N)\\ &\le \sum_{i=1}^{n} I(\mathbf{F}, Y^{i-1}, Z_{i+1}^n, V^N; Y_i)\\ &= \sum_{i=1}^{n} I(T_i; Y_i)\\ &= nI(T; Y). && (32) \end{aligned}$$

On the other hand, we have

$$
\begin{aligned}
& I(U^N;Y^n)-I(V^N;Y^n)\\
&= I(U^N;Y^n,K_1)-I(V^N;Y^n,K_1)\\
&\quad -I(U^N;K_1|Y^n)+I(V^N;K_1|Y^n)\\
&= I(U^N;Y^n,K_1)-I(V^N;Y^n,K_1)\\
&\quad +H(K_1|Y^n,U^N)-H(K_1|Y^n,V^N)\\
&\geq I(U^N;Y^n,K_1)-I(V^N;Y^n,K_1)-n\epsilon\\
&= \sum_{i=1}^{N}\Big(I(Y^n,K_1;U_i|U_{i+1}^N,V^{i-1})\\
&\quad -I(Y^n,K_1;V_i|U_{i+1}^N,V^{i-1})\Big)-n\epsilon\\
&= \sum_{i=1}^{N}\Big(I(Y^n,K_1,U_{i+1}^N,V^{i-1};U_i)\\
&\quad -I(Y^n,K_1,U_{i+1}^N,V^{i-1};V_i)\Big)-n\epsilon\\
&= \sum_{i=1}^{N}I(M_i;U_i)-I(M_i;V_i)-n\epsilon\\
&= N\big(I(M;U)-I(M;V)\big)-n\epsilon. \qquad (33)
\end{aligned}
$$

Combining (32) and (33), we have

$$I(M;U)-I(M;V)\leq \beta I(T;Y)+\beta\epsilon. \qquad (34)$$

■

B. Achievability

In this part, we will show that $\mathcal{R}(P_{M|U},P_{TS}P_{X|S})$ is an achievable region. It suffices to show that there exists at least one scheme such that the pair (R_1,R_2) with

$$
\begin{aligned}
&R_1=\frac{1}{\beta}[I(M;V)-\epsilon], R_2=\big[I(S;Y|T)-I(S;Z|T)\big]^+-\epsilon\\
&\text{s.t.}\quad I(M;U)-I(M;V)<\beta I(T;Y), \qquad (35)
\end{aligned}
$$

is achievable. Without loss of generality, we assume $I(S;Y|T)-I(S;Z|T)>0$.

Codebook Construction:

$\mathcal{C}_A$ at Alice. Given $P_{M|U}P_{UV}$ (suppose $I(M;U)-I(M;V)<\beta I(T;Y)$), randomly and independently generate 2^{NR_0} sequences M^Ns according to $\prod_{i=1}^{N}P_M(M_i)$. These sequences are indexed by (f,ϕ) with $f\in[1:2^{NR_{01}}]$, $\phi\in[1:2^{NR_{02}}]$,

$\mathcal{C}_B$ at Bob. Given $P_{TS}P_{X|S}P_{YZ|X}$ randomly and independently generate $2^{nR_{11}}$ sequences T^ns according to $\prod_{i=1}^{n}P_T(T_i)$. These sequences are indexed by (f,φ) with $\varphi\in[1:2^{nR_{12}}]$. For each $T^n(f,\varphi)$, randomly and independently generate $2^{nR_{13}}$ sequences S^ns which are indexed by (γ,ψ) with $\gamma\in[1:2^{nR_{14}}]$ and $\psi\in[1:2^{nR_{15}}]$, according to $\prod_{i=1}^{n}P_{S|T}(S_i|T_i)$.

Here, we set

$$
\begin{aligned}
&R_0=I(M;U)+\epsilon, &(36)\\
&R_{01}=I(M;U)-I(M;V)+2\epsilon, &(37)\\
&R_{02}=I(M;V)-\epsilon, &(38)\\
&R_{11}=I(T;Y)-\epsilon, &(39)\\
&R_{12}=I(T;Y)-\epsilon-\frac{1}{\beta}\left(I(W;U)-I(W;V)+2\epsilon\right), &(40)\\
&R_{13}=I(S;Y|T)-\epsilon, &(41)\\
&R_{14}=I(S;Z|T)+\epsilon, &(42)\\
&R_{15}=I(S;Y|T)-I(S;Z|T)-2\epsilon. &(43)
\end{aligned}
$$

Encoding: After observing sequence U^N, Alice selects one M^N that is jointly P_{MU} typical with U^N in $\mathcal{C}_A$. If there are more than one of such M^Ns, randomly select one from these sequences. If there is no such sequence, randomly select one from the whole codebook. We denote the selected sequence by $M^N(f,\phi)$. Alice sends the index f to Bob. Upon receiving f, Bob refers to $\mathcal{C}_B$, randomly generates a value for φ, and then looks into the sequences S^n generated by $T^n(f,\varphi)$, randomly selects one $S^n(\gamma,\psi)$, and finally transmits it to Carol via the channel $P_{X|S}P_{YZ|X}$.

Decoding: Upon receiving sequence Y^n, Carol first tries to find a unique $T^n(\hat{f},\hat{\varphi})$ that is jointly typical with Y^n in $\mathcal{C}_B$: If there are more than one of such T^ns, randomly selects one; If there exists no such T^n, declares an error. Then Carol looks into those S^ns generated by $T^n(\hat{f},\hat{\varphi})$, trying to find a unique $S^n(\hat{\gamma},\hat{\psi})$ that is jointly typical with $(T^n(\hat{f},\hat{\varphi}),Y^n)$: If there are more than one of such S^ns, randomly selects one; If there exists no such T^n, declares an error. Meantime, after decoding $\hat{f}$, Carol tries to find a unique $M^N(\hat{f},\hat{\phi})$ that is jointly typical with V^N.

Key Generation: Alice sets $K_1=\phi$; Bob sets $K_2=\psi$; Carol sets $K_1'=\hat{\phi}$ and $K_2'=\hat{\psi}$.

Key Rates Analysis: According to the codebook constructed above, we know ϕ and ψ are uniformly distributed in $[1:2^{NR_{02}}]$ and $[1:2^{nR_{14}}]$, respectively. Thus,

$$
\begin{aligned}
R_1&=\frac{N}{n}R_{02}=\frac{1}{\beta}[I(M;V)-\epsilon], &(44)\\
R_2&=I(S;Y|T)-I(S;Z|T)-2\epsilon. &(45)
\end{aligned}
$$

Error Analysis: Denote

$$
\begin{aligned}
\xi&\triangleq\{K_1\neq K_1' \text{ or } K_2\neq K_2'\}, &(46)\\
\xi_1&\triangleq\{T^n(f,\varphi)\neq T^n(\hat{f},\hat{\varphi})\}, &(47)\\
\xi_2&\triangleq\{S^n(\gamma,\psi)\neq S^n(\hat{\gamma},\hat{\psi})\}, &(48)\\
\xi_3&\triangleq\{M^N(f,\phi)\neq M^N(\hat{f},\hat{\phi})\}. &(49)
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
\Pr\{\xi\} &\leq \Pr\{\xi_1\cup\xi_2\cup\xi_3\}\\
&= \Pr\{\xi_1\}+\Pr\{\xi_2|\xi_1^c\}+\Pr\{\xi_3|(\xi_2\cup\xi_1)^c\}\\
&\overset{(a)}{=} \Pr\{\xi_1\}+\Pr\{\xi_2|\xi_1^c\}+\Pr\{\xi_3|\xi_1^c\}, \qquad (50)
\end{aligned}
$$

in which (a) is true since ξ_2 and ξ_3 are independent given ξ_1^c according to the above encoding approach. In the following, we bound each term in (50) one by one.

In our scheme, each T^n is randomly and independently generated according to $\prod_{i=1}^{n} P_T(T_i)$ and the total number of T^ns is $2^{nR_{11}}$. Furthermore, Y^n is equivalently generated by $T^n(f, \varphi)$ according to $\prod_{i=1}^{n} P_{Y|T}(Y_i|T_i)$, with $P_{Y|T} = P_{S|T}P_{X|S}P_{Y|X}$. Hence, it's easy to show that with high probability, $(T^n(f, \varphi), Y^n)$ is jointly typical and there will be no other T^ns to be jointly typical with Y^n (one may refer to Chapter 7 in [23]). Thus,

$$\Pr\{\xi_1\} \leq \epsilon/3,$$

when n is sufficiently large.

Given ξ_1^c, which is equivalent to that $T^n(f, \varphi)$ is given, there are $2^{nR_{13}}$ S^ns and that S^ns are randomly and independently generated by $T^n(f, \varphi)$ according to $\prod_{i=1}^{n} P_{S|T}(S_i|T_i)$. In addition, Y^n is equivalently generated by $S^n(\gamma, \psi)$ according to $\prod_{i=1}^{n} P_{Y|S}(Y_i|S_i)$, with $P_{Y|S} = P_{X|S}P_{Y|X}$, we can show that with high probability, $T^n(f, \varphi), (S^n(\gamma, \psi)$ and $Y^n)$ are jointly typical and there will be no other S^ns that are jointly typical with Y^n according to the Packing Lemma [24]. Thus, we can conclude

$$\Pr\{\xi_2|\xi_1^c\} \leq \epsilon/3,$$

when n is sufficiently large.

Since there are 2^{NR_0} M^Ns, and that M^Ns are randomly and independently generated according to $\prod_{i=1}^{N} P_M(M_i)$, we can show that with high probability there exists at least one M^N that is jointly typical with U^N (also jointly typical with V^N since $M \to U \to V$). Besides, given $T^n(f, \varphi)$, which indicates f is given, there are total $2^{NR_{02}}$ $M^N(f, \cdot)$s, thus, with high probability there will be no other M^Ns that are jointly typical with V^N. Then, we have

$$\Pr\{\xi_2|\xi_1^c\} \leq \epsilon/3,$$

when N is sufficiently large.

Hence,

$$\Pr\{\xi\} \leq \epsilon. \tag{51}$$

Information Leakage Analysis: Since ϕ and f are independent, and that $\phi \to f \to Z^n$,

$$I(K_1; f, Z^n|\mathcal{C}_A, \mathcal{C}_B) = I(\phi; f, Z^n|\mathcal{C}_A, \mathcal{C}_B) = I(\phi; f|\mathcal{C}_A) = 0.$$

To bound $I(K_2; U^N, f, Z^n|\mathcal{C}_A, \mathcal{C}_B)$, we have

$$\begin{aligned} I(K_2; U^N, f, Z^n|\mathcal{C}_A, \mathcal{C}_B) &= I(\psi; U^N, f, Z^n|\mathcal{C}_A, \mathcal{C}_B) \\ &\overset{(a)}{=} I(\psi; f, Z^n|\mathcal{C}_A, \mathcal{C}_B) \\ &\leq I(\psi; T^n, Z^n|\mathcal{C}_B) \\ &= I(\psi; T^n|\mathcal{C}_B) + I(\psi; Z^n|T^n, \mathcal{C}_B) \\ &= I(\psi; Z^n|T^n, \mathcal{C}_B), \end{aligned} \tag{52}$$

in which (a) is true due to

$$\begin{aligned} &\begin{cases} U^N - f, \psi - Z^n, \\ U^N - f - \psi \end{cases} \\ \Rightarrow \quad & U^N - f - \psi, Z^n \\ \Rightarrow \quad & U^N - f, Z^n - \psi \\ \Leftrightarrow \quad & U^N, Z^n - f, Z^n - \psi. \end{aligned} \tag{53}$$

Now, we have

$$\begin{aligned} & I(\psi; Z^n|T^n, \mathcal{C}_B) \\ &= H(Z^n|T^n, \mathcal{C}_B) - H(Z^n|T^n, \psi, \mathcal{C}_B) \\ &= H(Z^n|T^n, \mathcal{C}_B) - H(S^n, Z^n|T^n, \psi, \mathcal{C}_B) \\ &\quad + H(S^n|Z^n, T^n, \psi, \mathcal{C}_B) \\ &= H(Z^n|T^n, \mathcal{C}_B) - H(S^n|T^n, \psi, \mathcal{C}_B) \\ &\quad - H(Z^n|S^n, T^n, \psi, \mathcal{C}_B) + H(S^n|Z^n, T^n, \psi, \mathcal{C}_B) \\ &= H(Z^n|T^n, \mathcal{C}_B) - H(S^n|T^n, \psi, \mathcal{C}_B) \\ &\quad - H(Z^n|S^n, \mathcal{C}_B) + H(S^n|Z^n, T^n, \psi, \mathcal{C}_B). \end{aligned} \tag{54}$$

We can easily obtain that

$$\begin{aligned} H(Z^n|T^n, \mathcal{C}_B) &\leq nH(Z|T) + n\epsilon, \\ H(Z^n|S^n, \mathcal{C}_B) &\geq nH(Z|S) - n\epsilon, \end{aligned} \tag{55}$$

and according to Lemma 2 below, we have

$$\begin{aligned} I(\psi; Z^n|T^n, \mathcal{C}_B) &\leq n(H(Z|T) - H(Z|S) + I(S; Z|T) + 3\epsilon) \\ &= 3n\epsilon. \end{aligned} \tag{56}$$

Thus, we have

$$I(K_2; U^N, f, Z^n|\mathcal{C}_A, \mathcal{C}_B) \leq 3n\epsilon.$$

Lemma 2. *If $R_{15} + I(S; Z|T) < \frac{1}{n}H(S^n|T^n, \mathcal{C}_B)$, then*

$$\frac{1}{n}H(S^n|Z^n, T^n, \psi, \mathcal{C}_B) \leq \frac{1}{n}H(S^n|T^n, \psi, \mathcal{C}_B) - I(S; Z|T) + \epsilon.$$

Proof. See appendix B. □

Finally, using standard information theoretic arguments, we can conclude that there exists a particular code such that (35) is achievable and hence $\mathcal{R}(P_{M|U}, P_{TS}P_{X|S})$ is an achievable region.

Appendix B
Proof of Lemma 2

The proof here follows similar steps as those in the proof of [24, Lemma 22.3].

Given T^n, denote $\mathcal{T}_\epsilon^n(SZ|T^n)$ as the set of pairs (S^n, Z^n) which are jointly typical with T^n. Define

$$E_1 = \begin{cases} 1, & (S^n, Z^n) \in \mathcal{T}_\epsilon^n(SZ|T^n); \\ 0, & (S^n, Z^n) \notin \mathcal{T}_\epsilon^n(SZ|T^n). \end{cases}$$

Then, according to the Law of Large Numbers, we have

$$\Pr\{E_1 = 0\} \overset{n \to \infty}{\longrightarrow} 0, \tag{57}$$

since $T - S - Z$.

Thus, we have

$$\begin{aligned}
&H(S^n|Z^n,T^n,\psi,\mathcal{C}_B)\\
&\leq H(S^n,E_1|Z^n,T^n,\psi,\mathcal{C}_B)\\
&= H(E_1|Z^n,T^n,\psi,\mathcal{C}_B) + H(S^n|Z^n,E_1,T^n,\psi,\mathcal{C}_B)\\
&\leq 1+\Pr\{E_1=0\}H(S^n|Z^n,E_1=0,T^n,\psi,\mathcal{C}_B)\\
&\quad +\Pr\{E_1=1\}H(S^n|Z^n,E_1=1,T^n,\psi,\mathcal{C}_B)\\
&\leq 1+\Pr\{E_1=0\}H(S^n|Z^n,E_1=0,T^n,\psi,\mathcal{C}_B)\\
&\quad +\sum_{z^n,t^n,\psi}\Pr\{z^n,t^n,\psi|E_1=1\}H(S^n|z^n,E_1=1,t^n,\psi,\mathcal{C}_B)\\
&\leq n\epsilon +\sum_{z^n,t^n,\psi}\Pr\{z^n,t^n,\psi|E_1=1\}H(S^n|z^n,E_1=1,t^n,\psi,\mathcal{C}_B).
\end{aligned}$$

Now, given t^n,ψ,z^n and $E_1=1$, define Num(z^n,t^n) as the number of $S^n\in S^n(\cdot,\psi|t^n)\cap\mathcal{T}_\epsilon^n(S|z^n)$ ($S^n(\cdot,\psi|t^n)$ denotes the sequences S^ns generated by t^n, with second index ψ), we can easily show that

$$\begin{aligned}
&\mathbb{E}(\text{Num}(z^n,t^n)) = 2^{-nI(S;Z|T)}|S^n(\cdot,\psi|t^n)|,\\
&\text{Var}(\text{Num}(z^n,t^n)) \leq 2^{-nI(S;Z|T)}|S^n(\cdot,\psi|t^n)|,
\end{aligned}\tag{58}$$

where

$$\log|S^n(\cdot,\psi|t^n)| = H(S^n|T^n,\mathcal{C}_B)-nR_{15}.$$

Thus, we have

$$\begin{aligned}
&\Pr\{\text{Num}(z^n,t^n)\geq 2\mathbb{E}(\text{Num}(z^n,t^n))\}\\
&\leq 2^{-(H(S^n|T^n,\mathcal{C}_B)-nR_{15}-nI(S;Z|T))}.
\end{aligned}\tag{59}$$

Then, we have

$$\begin{aligned}
&H(S^n|z^n,E_1=1,t^n,\psi,\mathcal{C}_B)\\
&\leq n\epsilon + H(S^n|T^n,\mathcal{C}_B)-nR_{15}-nI(S;Z|T)\\
&= n\epsilon + H(S^n|T^n,\psi,\mathcal{C}_B)-nI(S;Z|T).
\end{aligned}\tag{60}$$

Hence, we have

$$H(S^n|Z^n,T^n,\psi,\mathcal{C}_B)\leq 2n\epsilon+H(S^n|T^n,\psi,\mathcal{C}_B)-nI(S;Z|T).$$

Appendix C
Proof of Theorem 2

A. Converse

Similar to the converse proof of Theorem 1, we will show that for any achievable pair (R_1,R_2), there exists $(P_{M|U}P_{L|M},P_{TS}P_{X|S})$ s.t. $(R_1,R_2)\in\mathcal{R}(P_{M|U}P_{L|M},P_{TS}P_{X|S})$.

First, we have

$$\begin{aligned}
H(K_1) &= H(K_1|Y^n,V^N)+I(K_1;Y^n,V^N)\\
&\leq I(K_1;Y^n,V^N)+n\epsilon\\
&\leq I(K_1;Y^n,V^N)-I(K_1;Z^n,W^N,\mathbf{F})+2n\epsilon\\
&\leq I(K_1;Y^n,V^N)-I(K_1;W^N,\mathbf{F})+2n\epsilon\\
&\overset{(a)}{=} I(K_1;Y^n,V^N)-I(K_1;Y^n,W^N,\mathbf{F})+2n\epsilon\\
&\leq I(K_1;Y^n,V^N)-I(K_1;Y^n,W^N)+2n\epsilon\\
&\leq I(K_1;V^N|Y^n)-I(K_1;W^N|Y^n)+2n\epsilon\\
&= \sum_{i=1}^N\big[I(K_1;V_i|V^{i-1},W_{i+1}^N,Y^n)\\
&\qquad -I(K_1;W_i|V^{i-1},W_{i+1}^N,Y^n)\big]+2n\epsilon\\
&= \sum_{i=1}^N\big[I(M_i;V_i|L_i)-I(M_i;W_i|L_i)\big]+2n\epsilon\\
&= N\big[I(M;V|L)-I(M;W|L)\big]+2n\epsilon,
\end{aligned}\tag{61}$$

in which $M_i := (K_1,V^{i-1},W_{i+1}^N,Y^n)$, $L_i := (V^{i-1},W_{i+1}^N,Y^n)$ and $M:=(M_Q,Q)$, $L:=(L_Q,Q)$. (a) is true because of

$$\begin{aligned}
& W^N-U^N-\mathbf{F}-Y^n\\
\Rightarrow\quad & (U^N,W^N)-\mathbf{F}-Y^n\\
\Rightarrow\quad & (K_1,W^N)-\mathbf{F}-Y^n\\
\Rightarrow\quad & K_1-(W^N,\mathbf{F})-Y^n.
\end{aligned}\tag{62}$$

Thus, we have

$$R_1\leq\frac{1}{\beta}\big[I(M;V|L)-I(M;W|L)\big]+2\epsilon.\tag{63}$$

Furthermore, similar to (28), we can show that $L-M-U-(V,W)$.

The derivation of R_2 is exactly the same as in (31), thus, we have

$$R_2\leq I(S;Y|T)-I(S;Y|T)+3\epsilon,\tag{64}$$

where $S := (K_2,V^N,Y^{J-1},Z_{J+1}^n,\mathbf{F},J)$, and $T := (V^N,Y^{J-1},Z_{J+1}^n,\mathbf{F},J)$.

Next, we show (19). From (33), we conclude

$$\begin{aligned}
&I(U^N;Y^n)-I(V^N;Y^n)\\
&\geq\sum_{i=1}^N\Big[I(Y^n,K_1,U_{i+1}^N,V^{i-1};U_i)\\
&\qquad -I(Y^n,K_1,U_{i+1}^N,V^{i-1};V_i)\Big]-n\epsilon.
\end{aligned}\tag{65}$$

Now, since

$$\begin{aligned}
& W_{i+1}^N-U_{i+1}^N-(U^N,V^i)\\
\Rightarrow\quad & W_{i+1}^N-U_{i+1}^N-(K_1,\mathbf{F},U_i,V^i)\\
\Rightarrow\quad & W_{i+1}^N-U_{i+1}^N-(K_1,Y^n,U_i,V^i)\\
\Rightarrow\quad & W_{i+1}^N-(Y^n,K_1,U_{i+1}^N,V^{i-1})-(U_i,V_i)\\
\Rightarrow\quad & \begin{cases} W_{i+1}^N-(Y^n,K_1,U_{i+1}^N,V^{i-1})-U_i\\ W_{i+1}^N-(Y^n,K_1,U_{i+1}^N,V^{i-1})-V_i\end{cases},
\end{aligned}\tag{66}$$

and

$$\begin{aligned} & (U^N, V^{i-1}, W_{i+1}^N) - U_i - V_i \\ \Rightarrow\quad & (K_1, \mathbf{F}, U_{i+1}^N, V^{i-1}, W_{i+1}^N) - U_i - V_i \\ \Rightarrow\quad & (K_1, Y^n, U_{i+1}^N, V^{i-1}, W_{i+1}^N) - U_i - V_i \\ \Rightarrow\quad & U_{i+1}^N - (Y^n, K_1, V^{i-1}, W_{i+1}^N, U_i) - V_i, \end{aligned} \tag{67}$$

we have

$$\begin{aligned} & \sum_{i=1}^{N} \Big[I(Y^n, K_1, U_{i+1}^N, V^{i-1}; U_i) \\ & \qquad - I(Y^n, K_1, U_{i+1}^N, V^{i-1}; V_i) \Big] \\ = & \sum_{i=1}^{N} \Big[I(Y^n, K_1, U_{i+1}^N, V^{i-1}, W_{i+1}^N; U_i) \\ & \qquad - I(Y^n, K_1, U_{i+1}^N, V^{i-1}, W_{i+1}^N; V_i) \Big] \\ = & \sum_{i=1}^{N} \Big[I(Y^n, K_1, V^{i-1}, W_{i+1}^N; U_i) \\ & \qquad - I(Y^n, K_1, V^{i-1}, W_{i+1}^N; V_i) \Big] \\ & + \sum_{i=1}^{N} \Big[I(U_{i+1}^N; U_i | Y^n, K_1, V^{i-1}, W_{i+1}^N) \\ & \qquad - I(U_{i+1}^N; V_i | Y^n, K_1, V^{i-1}, W_{i+1}^N) \Big] \\ = & \sum_{i=1}^{N} \Big[I(Y^n, K_1, V^{i-1}, W_{i+1}^N; U_i) \\ & \qquad - I(Y^n, K_1, V^{i-1}, W_{i+1}^N; V_i) \Big] \\ & + \sum_{i=1}^{N} \Big[I(U_{i+1}^N; Y^n, K_1, V^{i-1}, W_{i+1}^N, U_i, V_i) \\ & \qquad - I(U_{i+1}^N; Y^n, K_1, V^{i-1}, W_{i+1}^N, V_i) \Big] \\ \geq & \sum_{i=1}^{N} \Big[I(Y^n, K_1, V^{i-1}, W_{i+1}^N; U_i) \\ & \qquad - I(Y^n, K_1, V^{i-1}, W_{i+1}^N; V_i) \Big] \\ = & \sum_{i=1}^{N} \Big[I(M_i; U_i) - I(M_i; V_i) \Big] \\ = & N \Big[I(M; U) - I(M; V) \Big]. \end{aligned} \tag{68}$$

Thus, it follows

$$I(U^N; Y^n) - I(V^N; Y^n) \geq N \Big[I(M; U) - I(M; V) \Big] - n\epsilon.$$

On the other hand, same as (32), we conclude

$$I(U^N; Y^n) - I(V^N; Y^n) \leq nI(T; Y).$$

Hence,

$$\begin{aligned} & N \Big[I(M; U) - I(M; V) \Big] - n\epsilon \leq nI(T; Y) \\ \Rightarrow\quad & I(M; U) - I(M; V) \leq \beta I(T; Y) + \beta\epsilon. \end{aligned} \tag{69}$$

Combining the fact that ϵ in each term is an arbitrary small number, we can conclude that there exists such $(P_{M|U} P_{L|M}, P_{TS} P_{X|S})$ that $(R_1, R_2) \in \mathcal{R}(P_{M|U} P_{L|M}, P_{TS} P_{X|S})$.

B. Achievability

It suffices to show that the pair (R_1, R_2) with

$$R_1 = \frac{1}{\beta} I(M; V|L) - I(M; W|L) - \epsilon, \tag{70}$$

$$R_2 = I(S; Y|T) - I(S; Z|T) - \epsilon, \tag{71}$$

$$\text{s.t.} \quad I(M; U) - I(M; V) < \beta I(T; Y), \tag{72}$$

is achievable.

Given $(P_{M|U} P_{L|M}, P_{TS} P_{X|S})$, without loss of generality, we assume $I(M; V|L) - I(M; W|L) > 0$ and $I(S; Y|T) - I(S; Z|T) > 0$.

Codebook Construction:

$\mathcal{C}_A$ at Alice. Given P_L, randomly and independently generate $2^{NR_{10}}$ sequences L^Ns according to $\prod_{i=1}^{N} P_L(L_i)$, and assign each L^N into $2^{NR_{11}}$ bins indexed by $f_1(L^N)$ with $f_1 \in [1 : 2^{NR_{11}}]$, using a uniform distribution and denote the corresponding bin by $\mathcal{B}_0(f_1)$.

For each L^N, randomly and independently generate $2^{NR_{12}}$ sequences M^Ns according to $\prod_{i=1}^{N} P_{M|L}(M_i|L_i)$. Assign each M^N into $2^{NR_{13}}$ bins indexed $f_2(M^N)$ with $f_2 \in [1 : 2^{NR_{13}}]$, using a uniform distribution and denote the corresponding bin by $\mathcal{B}_1(f_2)$. Within each bin $\mathcal{B}_1(f_2)$, randomly assign each M^N into $2^{NR_{14}}$ sub-bins indexed $\phi(M^N)$ with $\phi \in [1 : 2^{NR_{14}}]$, using a uniform distribution and denote the corresponding sub-bin by $\mathcal{B}_1(f_2, \phi)$.

$\mathcal{C}_B$ at Bob. Given P_T, randomly and independently generate $2^{nR_{20}}$ sequences T^ns according to $\prod_{i=1}^{n} P_T(T_i)$, indexed by (f_1, f_2, φ). For each $T^n(f_1, f_2, \varphi)$, randomly and independently generate $2^{nR_{21}}$ sequences S^ns according to $\prod_{i=1}^{n} P_{S|T}(S_i|T_i)$, and assign each S^n in to $2^{nR_{22}}$ bins indexed by $\psi(S^n)$ with $\psi \in [1 : 2^{nR_{22}}]$, using a uniform distribution and denote the corresponding bin by $\mathcal{B}_2(\psi)$. Here, we set

$$\begin{aligned} R_{10} &= I(L; U) + \epsilon, \\ R_{11} &= I(L; U) - I(L; V) + 2\epsilon, \\ R_{12} &= I(M; U|L) + \epsilon, \\ R_{13} &= I(M; U|L) - I(M; V|L) + 2\epsilon, \\ R_{14} &= I(M; V|L) - I(M; W|L) - 2\epsilon, \\ R_{20} &= I(T; Y) - \epsilon, \\ R_{21} &= I(S; Y|T) - \epsilon, \\ R_{22} &= I(S; Y|T) - I(S; Z|T) - 2\epsilon. \end{aligned} \tag{73}$$

Encoding: With the observed sequence U^N, Alice looks into $\mathcal{C}_A$, looking for a L^N that is jointly typical with U^N according to P_{LU}. If there are more than one such sequence, randomly select one (suppose $L^N = l^N$ is selected); If Alice can't find it, declares an error. Then, Alice looks into those M^Ns generated by L^N, looking for a M^N that is jointly typical with (L^N, U^N) according to P_{LMU}. If there are more than one such sequence, randomly select one (suppose $M^N = m^N$ is selected); If there exists no such sequence, declare an error. Finally, Alice sends $(f_1(L^N), f_2(M^N))$ to Bob.

Upon receiving (f_1, f_2), Bob first randomly generates a value for φ and selects a sequence $T^n(f_1, f_2), \varphi)$ in $\mathcal{C}_B$. Then Bob randomly selects one $S^n = s^n$ from those S^ns generated by $T^n(f_1, f_2), \varphi)$, and transmits it to Carol via the channel $P_{X|S}P_{YZ|X}$.

Decoding: Upon receiving Y^n, Carol first tries to decode $(\hat{T}^n, \hat{S}^n)$ using the same method as described in the proof of Theorem 1.

After decoding $\hat{T}^n$ Carol will obtain corresponding values for (f_1, f_2). Then Carol refers to $\mathcal{C}_A$, looking for a unique $\hat{L}^N$ in $\mathcal{B}_0(f_1)$ that is jointly typical with V^N. If Carol can't find it, randomly selects one $\hat{L}^N$. Then, Carol turns to those M^Ns generated by $\hat{L}^N$, looking for a unique $\hat{M}^N$ which is jointly typical with $(\hat{L}^N, V^N)$ according to P_{LMV}. If Carol can't find it, randomly selects one.

Key Generation: Alice sets $K_1 = \phi(M^N)$; Bob sets $K_2 = \psi(S^n)$; Carol sets $\hat{K}_1 = \phi(\hat{M}^N)$ and $\hat{K}_2 = \psi(\hat{S}^n)$.

Key Rates Analysis: According to the above constructed codebook, ϕ and ψ are uniformly distributed over $[1 : 2^{NR_{14}}]$ and $[1 : 2^{NR_{22}}]$ respectively, thus

$$\begin{aligned} R_1 &= \frac{1}{\beta}\big[I(M;V|L) - I(M;W|L) - 2\epsilon\big], \\ R_2 &= I(S;Y|T) - I(S;Z|T) - 2\epsilon. \end{aligned} \tag{74}$$

Error Analysis: Note that

$$\begin{aligned} H(f_1, f_2) &= \frac{1}{\beta}\big[R_{11} + R_{13}\big] \\ &= \frac{1}{\beta}\big[I(M;U) - I(M;V) + 4\epsilon\big]. \end{aligned} \tag{75}$$

Thus, with the same reason as discussed in the proof of Theorem 1, Carol will decode (T^n, S^n, L^N) correctly with high probability. Now, we show Carol can also decode M^N correctly.

Since there are in total $2^{NR_{12}}$ M^Ns generated by given L^N, there must exist at least one M^N that is jointly typical with U^N with high probability, according to the covering lemma [24] . Furthermore, in bin $\mathcal{B}_1(f_2)$, there are approximately

$$2^{NR_{12}}/2^{NR_{12}} = 2^{N(I(M;V|L)-\epsilon)} \tag{76}$$

M^N sequences. This guarantees that, with high probability, there is no other $\widetilde{M}^N$ to be jointly typical with (L^N, V^N), according to the packing lemma [24]. Thus, we can conclude that

$$\Pr\{K_1 \neq \hat{K}_1 \text{ or } K_2 \neq \hat{K}_2 | \mathcal{C}_A, \mathcal{C}_B\} \leq \epsilon, \tag{77}$$

when n is sufficiently large.

Information Leakage Analysis: Similar to the leakage analysis in the proof of Theorem 1, we can also obtain that

$$I(K_2; U^N, W^N, f_1, f_2, Z^n | \mathcal{C}_A, \mathcal{C}_B) \leq n\epsilon, \tag{78}$$

since

$$I(K_2; U^N, W^N, f_1, f_2, Z^n | \mathcal{C}_A, \mathcal{C}_B) \leq I(K_2; T^n, Z^n | \mathcal{C}_A, \mathcal{C}_B).$$

Now, we bound $I(K_1; f_1, f_2, \psi, W^N, Z^n | \mathcal{C}_A, \mathcal{C}_B)$ from above. Since,

$$\begin{aligned} & I(K_1; f_1, f_2, \psi, W^N, Z^n | \mathcal{C}_A, \mathcal{C}_B) \\ = \quad & I(K_1; f_1, f_2, W^N | \mathcal{C}_A, \mathcal{C}_B) \\ = \quad & I(K_1; f_1, f_2, W^N | \mathcal{C}_A), \end{aligned}$$

we can obtain that

$$I(K_1; f_1, f_2, W^N | \mathcal{C}_A) \leq n\epsilon,$$

using the same argument that is used in the achievability proof of Theorem 22.4 in [24].

Finally, following standard information theoretic arguments, we can conclude that there exists at least one scheme such that (R_1, R_2) specified in (70) and (71) are achievable, and hence $\mathcal{R}(P_{M|U}P_{L|M}, P_{TS}P_{X|S})$ is achievable.

References

[1] S. Watanabe and Y. Oohama, "Secret key agreement from vector Gaussian sources by rate limited public communication," *IEEE Trans. Inform. Forensics and Security*, vol. 6, pp. 541–550, Sept. 2011.

[2] A. Agrawal, Z. Rezki, A. Khisti, and M.-S. Alouini, "Noncoherent capacity of secret-key agreement with public discussion," *IEEE Trans. Inform. Forensics and Security*, vol. 6, pp. 565–574, Sept. 2011.

[3] N. Wang, N. Zhang, and T. A. Gulliver, "Cooperative key agreement for wireless networking: Key rates and practical protocol design," *IEEE Trans. Inform. Forensics and Security*, vol. 9, pp. 272–284, Jan. 2014.

[4] C. W. Wong, T. F. Wong, and J. M. Shea, "Secret-sharing LDPC codes for the BPSK-constrained Gaussian wiretap channel," *IEEE Trans. Inform. Forensics and Security*, vol. 6, pp. 551–564, Sept. 2011.

[5] K. Chen, B. B. Natarajan, and S. Shattil, "Secret key generation rate with power allocation in relay-based LTE-A networks," *IEEE Trans. Inform. Forensics and Security*, vol. 10, pp. 2424–2434, Nov. 2015.

[6] M. F. Haroun and T. A. Gulliver, "Secret key generation using chaotic signals over frequency selective fading channels," *IEEE Trans. Inform. Forensics and Security*, vol. 10, pp. 1764–1775, Aug. 2015.

[7] S. Tomasin and A. Dall'Arche, "Resource allocation for secret key agreement over parallel channels with full and partial eavesdropper CSI," *IEEE Trans. Inform. Forensics and Security*, vol. 10, pp. 2314–2324, Nov. 2015.

[8] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3047–3061, Dec. 2004.

[9] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inform. Theory*, vol. 46, pp. 344–366, Mar. 2000.

[10] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, pp. 733–742, May 1993.

[11] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography, Part I: Secret sharing," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1121–1132, July 1993.

[12] A. Khisti, S. N. Diggavi, and G. W. Wornell, "Secret-key agreement with channel state information at the transmitter," *IEEE Trans. Inform. Forensics and Security*, vol. 6, pp. 672–681, Sept. 2011.

[13] I. Csiszar and P. Narayan, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inform. Theory*, vol. 54, pp. 2437–2452, Jun. 2008.

[14] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339–348, May 1978.

[15] A. Zibaeenejad, “Key generation over wiretap models with non-causal side information,” *IEEE Trans. Inform. Forensics and Security*, vol. 10, pp. 1456–1471, Jul. 2015.

[16] A. Khisti, S. N. Diggavi, and G. W. Wornell, “Secret-key generation using correlated sources and channels,” *IEEE Trans. Inform. Theory*, vol. 58, pp. 652–670, Feb. 2012.

[17] C. Ye, S. Mathur, A. Reznik, W. Trappe, and N. Mandayam, “Information-theoretic key generation from wireless channels,” *IEEE Trans. Inform. Forensics and Security*, vol. 5, pp. 240–254, Jun. 2010.

[18] H. Zhang, L. Lai, Y. Liang, and H. Wang, “The capacity region of the source-type model for secret key and private key generation,” *IEEE Trans. Inform. Theory*, vol. 60, pp. 6389–6398, Jul. 2014.

[19] L. Lai and L. Huie, “Simultaneously generating multiple keys in many to one networks,” in *Proc. IEEE Intl. Symposium on Inform. Theory*, (Istanbul, Turkey), pp. 2394–2398, July 2013.

[20] C. Ye and P. Narayan, “The secret key-private key capacity region for three terminals,” in *Proc. IEEE Intl. Symposium on Inform. Theory*, (Adelaide, Australia), pp. 2142–2146, Sept. 4-9, 2005.

[21] C. Ye and P. Narayan, “Secret key and private key constructions for simple multiterminal source models,” *IEEE Trans. Inform. Theory*, vol. 58, pp. 639–651, Feb. 2012.

[22] P. Xu, Z. Ding, X. Dai, and G. Karagiannidis, “Simultaneously generating secret and private keys in a cooperative pairwise independent network,” *IEEE Trans. Inform. Forensics and Security*. To appear.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 2012.

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