Two-Key Generation for a Cellular Model with a Helper

Huishuai Zhang and Yingbin Liang
Dept. of EECS
Syracuse University
Syracuse, NY 13244
Email: {hzhan23, yliang06}@syr.edu

Lifeng Lai
Dept of ECE
Worcester Polytechnic Institute
Worcester, MA 01609
Email: llai@wpi.edu

Shlomo Shamai (Shitz)
Dept of EE
Technion-Israel Institute of Technology
Technion city, Haifa 32000, Israel
Email: sshlomo@ee.technion.ac.il

***Abstract*—The problem of simultaneously generating two keys for a cellular model is investigated, in which each of four terminals, $\mathcal{X}_0$, $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$ observes one component of correlated sources. The terminal $\mathcal{X}_0$ wishes to generate secret keys K_1 and K_2 respectively, with terminals $\mathcal{X}_1$ and $\mathcal{X}_2$ under the help of terminal $\mathcal{X}_3$. They are allowed to communicate over a public channel. Both K_1 and K_2 are required to be concealed from an eavesdropper that has access to the public discussion. The key capacity region is established by designing a unified achievable strategy to achieve the cut-set outer bounds, which greatly simplifies the proof.**

***Keywords*—Cut-set bound, key capacity region, secret key generation, source model.**

I. Introduction

Recently, there is increasing interest on the topic of simultaneously generating multiple secret keys over source models. The model of generating a secret key and a private key among three terminals via public discussion was first studied by Ye and Narayan in [1], in which three terminals observe correlated source outputs. All three terminals wish to agree on a common secret key to be kept secure from an eavesdropper while two designated terminals wish to agree on a private key to be kept secure from both the third terminal and the eavesdropper. [1] provided an outer bound on the key capacity region and showed that it is achievable under a certain condition. More recently, [2], [3] further showed that the outer bound established in [1] is achievable in general by employing a random binning and joint decoding scheme. Using similar techniques, [4] established the key capacity region for a cellular model with three terminals.

It is of interest to explore whether the achievable techniques in [3] can be extended to other models with more than three terminals to establish multi-key capacity results. In this paper, we study a model of generating two keys over four terminals (see Fig. 1), in which each of the four terminals observes one component of a correlated vector source. Terminals $\mathcal{X}_0$ and $\mathcal{X}_1$ wish to generate a key K_1, and terminals $\mathcal{X}_0$ and $\mathcal{X}_2$ wish to generate another independent key K_2. The four terminals are allowed to communicate over a pubic channel. The two keys are required to be concealed from an eavesdropper which has access to the public discussion. Terminal $\mathcal{X}_3$, a dedicated helper, can help to generate the two keys. This model can be interpreted as a cellular network, in which a base station (i.e., terminal $\mathcal{X}_0$) wishes to establish independent secret keys with two mobile terminals $\mathcal{X}_1$ and $\mathcal{X}_2$.

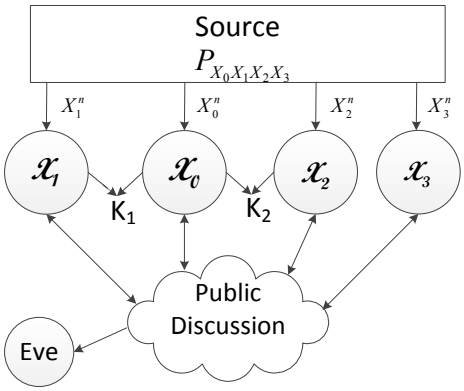

Fig. 1. System model

Our exploration of the above model focuses on two issues. (1) Although the key capacity for the three-terminal models was shown to be equal to the cut-set bound, it is not clear at the outset whether the cut-set bound can still be achieved for the four-terminal model studied here. In general, cut-set bound is less likely achievable as the system gets more complicated. (2) For the three-terminal model in [3], [4], there are three cuts for generating two keys, and schemes are designed to achieve corner points of the cut-set bound for each case of the source distribution. However, for the four-terminal model, there are six cuts for generating two keys, and these six cuts yield eight possible structures of the cut-set bound. It is not clear whether there exists a unified design of schemes to achieve all cut-set bound structures.

Our contribution in this paper lies in affirmative answers to both of the above issues. We establish the cut-set bound as the key capacity region for the four-terminal model by showing that the cut-set bound is indeed achievable. Furthermore, we establish a unified achievable strategy to achieve the corner points of cut-set bound structures corresponding to all source distributions. The schemes to achieve different structures vary only in the rate at which each terminal reveals information to public. Thus, the achievability proof is significantly simplified. More specifically, the achievable strategy is based on random binning and joint decoding. Given such a unified strategy, we derive the Slepian-Wolf conditions that guarantee correct

key agreement and further derive sufficient conditions that guarantee secrecy. Then for each individual case, it is sufficient to verify the public transmission rates of terminals satisfy the derived Slepian-Wolf conditions and secrecy conditions, which can be performed easily.

The paper is organized as follows. Section II introduces the system model. Section III provides our main results on the key capacity region. Section IV gives technical proofs. Section V concludes this paper with comments.

II. System Model

Consider a system (see Fig. 1) with four distinct terminals $\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{X}_3$, each of which observes one component of a discrete memoryless vector source with a joint distribution $P_{X_0X_1X_2X_3}$. Here, for simplicity, we also use $\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{X}_3$ to denote the finite alphabets, respectively. More specifically, terminal $\mathcal{X}_0$ observes n independently and identically distributed (i.i.d.) repetitions of X_0, denoted by X_0^n, and terminals $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{X}_3$ observe X_1^n, X_2^n and X_3^n, respectively. The four terminals can communicate interactively via a public channel with no rate constraints. All transmissions over the public channel are available to all parties including an eavesdropper. We assume that the eavesdropper does not observe any further information such as source sequences. In this model, terminal $\mathcal{X}_0$ wishes to establish two keys K_1 and K_2 respectively with terminals $\mathcal{X}_1$ and $\mathcal{X}_2$ under the help of terminal $\mathcal{X}_3$. Both K_1 and K_2 are required to be secure from the eavesdropper. Neither key needs to be secure from the helper $\mathcal{X}_3$, which is regarded as a trusted terminal.

We use $\mathbf{F}$ to denote all transmissions in public discussion. Each of the four terminals can reveal any causal information, i.e., a deterministic function of the corresponding terminal's source observations and the previous public transmissions, to public with multiple rounds.

We next introduce the mathematical conditions that a key pair (K_1, K_2) should satisfy. A random variable U is said to be *ϵ-recoverable* from another random variable V, if there exists a function f such that

$$\Pr\{U \neq f(V)\} < \epsilon. \tag{1}$$

In the system of interest, the key K_1 is required to be *ϵ-recoverable* at terminals $\mathcal{X}_0$ and $\mathcal{X}_1$ with the public transmission $\mathbf{F}$, i.e., it can be *ϵ-recoverable* from $(X_0^n, \mathbf{F})$ and $(X_1^n, \mathbf{F})$, respectively. The key K_2 is required to be *ϵ-recoverable* at terminals $\mathcal{X}_0$ and $\mathcal{X}_2$ with public transmission $\mathbf{F}$, i.e., it can be *ϵ-recoverable* from $(X_0^n, \mathbf{F})$ and $(X_2^n, \mathbf{F})$, respectively. Moreover, K_1 and K_2 are required to satisfy the secrecy condition

$$\frac{1}{n} I(K_1K_2; \mathbf{F}) < \epsilon \tag{2}$$

and the uniformity conditions

$$\frac{1}{n} H(K_1) \geq \frac{1}{n} \log |\mathcal{K}_1| - \epsilon \tag{3}$$

$$\frac{1}{n} H(K_2) \geq \frac{1}{n} \log |\mathcal{K}_2| - \epsilon \tag{4}$$

where $\mathcal{K}_1$ and $\mathcal{K}_2$ denote the alphabets of the random variables K_1 and K_2, respectively, and ϵ can be arbitrarily small as the sequence length n goes to infinity.

Definition 1. *A rate pair (R_1, R_2) is said to be achievable if for every $\epsilon > 0$, $\delta > 0$, and for sufficiently large n, a key pair (K_1, K_2) can be generated satisfying* (2)-(4) *and*

$$\frac{1}{n} H(K_1) > R_1 - \delta, \quad \textit{and} \quad \frac{1}{n} H(K_2) > R_2 - \delta. \tag{5}$$

Our goal is to characterize the *key capacity region* that contains all achievable rate pairs (R_1, R_2).

III. Main Results

In this section, we state our main result on the key capacity region for the model of interest with interpretation of the region as cut-set bounds.

For convenience, we introduce the following notations:

$$R_A := \min\{I(X_1X_3; X_0X_2), I(X_1; X_0X_2X_3)\}, \tag{6}$$

$$R_B := \min\{I(X_0X_1; X_2X_3), I(X_2; X_0X_1X_3)\}, \tag{7}$$

$$R_C := \min\{I(X_0; X_1X_2X_3), I(X_0X_3; X_1X_2)\}. \tag{8}$$

Theorem 1. *The key capacity region for the model described in Section II contains rate pairs (R_1, R_2) satisfying the following inequalities:*

$$R_1 \leq R_A, \tag{9}$$

$$R_2 \leq R_B, \tag{10}$$

$$R_1 + R_2 \leq R_C. \tag{11}$$

The structure of the key capacity region is illustrated as Fig. 2. The secrecy constraints on K_1 and K_2 are symmetric so that the bounds on R_1 and R_2 are also symmetric. These bounds can be intuitively understood as cut-set bounds. In particular, the upper bound on R_1 in (9) is due to the two cuts separating $\mathcal{X}_0$ and $\mathcal{X}_1$ for generating K_1. The upper bound on R_2 in (10) is due to the cut separating $\mathcal{X}_0$ and $\mathcal{X}_2$ for generating K_2. The sum rate bound (11) is due to the cut separating $\mathcal{X}_0$ and $(\mathcal{X}_1, \mathcal{X}_2)$ for generating the two keys K_1 and K_2 simultaneously.

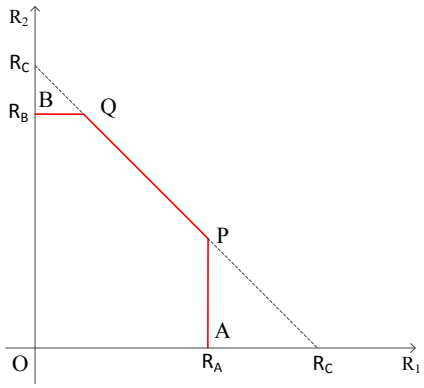

Fig. 2. The key capacity region of the model described in Sec. II

IV. Technical Proofs

Proof of Converse. First, if we need only to generate K_1, the model reduces to the secret key generation problem studied in [5]. The key capacity is shown to be $\min\{R_A, R_C\}$, which provides an upper bound (9) on R_1. Similarly, if we dedicate to generate K_2, the key capacity is shown in [5] to be $\min\{R_B, R_C\}$, which provides an upper bound (10) on R_2. For the sum rate bound, we consider an enhanced model by replacing terminals $\mathcal{X}_1$ and $\mathcal{X}_2$ with a super terminal $\mathcal{X}_s$ that observes both X_1^n and X_2^n. The secret key rate between $\mathcal{X}_0$ and $\mathcal{X}_s$ is upper bounded by R_C as given in [5], which yields the sum rate bound (11).

Proof of Achievability. The key capacity region is plotted in Fig. 2 as the pentagon O-A-P-Q-B-O, where the coordinates of points A and B are $(\min\{R_A, R_C\}, 0)$ and $(0, \min\{R_B, R_C\})$, respectively. The corner point A is achieved by letting $\mathcal{X}_2$ be a dedicated helper to generate K_1 following the omniscience scheme in [5]. Similarly, the corner point B is achieved by letting $\mathcal{X}_1$ be a dedicated helper to generate K_2 as shown in [5]. We note that the point P would collapse to the point A if $R_C \leq R_A$, and the point Q would collapse to the point B if $R_C \leq R_B$. It is thus sufficient to show that the points P and Q are achievable whenever they are different from the points A and B, respectively. Then the entire pentagon can be achieved by time sharing.

We note that since the secrecy constraints on K_1 and K_2 are symmetric, it is sufficient to show that the corner point P in Fig. 2 is achievable, and then the achievability of the point Q follows by symmetry. We assume that $R_A < R_C$, because otherwise the point P would collapse to the point A and has been justified to be achievable.

The key rate pair of the point P is given by $R_1 = R_A, R_2 = R_C - R_A$. Corresponding to different source distributions, each of R_A and R_C can take one of the two mutual information terms given in (6) and (8), respectively. Hence, the coordinates of the point P can take four forms. Correspondingly, we need to show the achievability for four cases, i.e., case 1 with $R_A = I(X_1; X_0X_2X_3)$ and $R_C = I(X_0; X_1X_2X_3)$, case 2 with $R_A = I(X_1; X_0X_2X_3)$ and $R_C = I(X_0X_3; X_1X_2)$, case 3 with $R_A = I(X_0X_2; X_1X_3)$ and $R_C = I(X_0; X_1X_2X_3)$, and case 4 with $R_A = I(X_0X_2; X_1X_3)$ and $R_C = I(X_0X_3; X_1X_2)$.

We construct a unified strategy to prove the achievability for all cases. More specifically, terminals $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$ reveal information to public in order to guarantee that terminal $\mathcal{X}_0$ recovers X_1^n and X_2^n. Then, K_1 is generated by terminals $\mathcal{X}_0$ and $\mathcal{X}_1$ based on X_1^n, and K_2 is generated by terminals $\mathcal{X}_0$ and $\mathcal{X}_2$ based on X_2^n. The schemes for the four cases are different only in the rate at which each terminal reveals information to public. In the following, we first describe the unified scheme, and then study each case one by one. In general, our scheme is based on random binning and joint decoding.

Codebook Generation: At terminal $\mathcal{X}_1$, randomly and independently assign a bin index f to each sequence $x_1^n \in \mathcal{X}_1^n$, where $f \in [1 : 2^{n\tilde{R}_1}]$. Then randomly and independently assign a sub-bin index ϕ to each sequence x_1^n, where $\phi \in [1 : 2^{nR_1}]$.

At terminal $\mathcal{X}_2$, randomly and independently assign a bin index g to each sequence $x_2^n \in \mathcal{X}_2^n$, where $g \in [1 : 2^{n\tilde{R}_2}]$. Then randomly and independently assign a sub-bin index ψ to each sequence x_2^n, where $\psi \in [1 : 2^{n\tilde{R}_2}]$.

At terminal $\mathcal{X}_3$, randomly and independently assign a bin index l to each sequence $x_3^n \in \mathcal{X}_3^n$, where $l \in [1 : 2^{n\tilde{R}_3}]$.

The codebook is revealed to all parties, i.e., terminals $\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3$, and the eavesdropper.

Encoding and Transmission: Given a sequence x_1^n, terminal $\mathcal{X}_1$ reveals the bin index $f = f(x_1^n)$ over the public channel to all parties.

Given a sequence x_2^n, terminal $\mathcal{X}_2$ reveals the bin index $g = g(x_2^n)$ over the public channel to all parties.

Given a sequence x_3^n, terminal $\mathcal{X}_3$ reveals the bin index $l = l(x_3^n)$ over the public channel to all parties.

Decoding: The decoding scheme is based on the joint typicality.

Terminal $\mathcal{X}_0$, given x_0^n and the bin indexes f, g and l, claims $\hat{x}_1^n$, $\hat{x}_2^n$ and $\hat{x}_3^n$ are observations of terminals $\mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{X}_3$, respectively, if there exists a unique tuple of sequences $(\hat{x}_1^n, \hat{x}_2^n, \hat{x}_3^n)$ such that $\hat{x}_1^n \in B_1(f)$, $\hat{x}_2^n \in B_2(g)$, $\hat{x}_3^n \in B_3(l)$ and $(x_0^n, \hat{x}_1^n, \hat{x}_2^n, \hat{x}_3^n) \in T_\epsilon^{(n)}(P_{X_0X_1X_2X_3})$.

Based on Slepian-Wolf coding theorem [5], [6], the decoding error can be arbitrarily small if the rates $\tilde{R}_1, \tilde{R}_2$ and $\tilde{R}_3$ satisfy the following Slepian-Wolf conditions:

$$\tilde{R}_1 > H(X_1|X_0X_2X_3), \tag{12}$$
$$\tilde{R}_2 > H(X_2|X_0X_1X_3), \tag{13}$$
$$\tilde{R}_3 > H(X_3|X_0X_1X_2), \tag{14}$$
$$\tilde{R}_1 + \tilde{R}_2 > H(X_1X_2|X_0X_3), \tag{15}$$
$$\tilde{R}_1 + \tilde{R}_3 > H(X_1X_3|X_0X_2), \tag{16}$$
$$\tilde{R}_2 + \tilde{R}_3 > H(X_2X_3|X_0X_1), \tag{17}$$
$$\tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3 > H(X_1X_2X_3|X_0). \tag{18}$$

Hence, (12)-(18) guarantee the correct key establishment.

Key Generation: Terminal $\mathcal{X}_1$ sets $K_1 = \phi(X_1^n)$. Terminal $\mathcal{X}_2$ sets $K_2 = \psi(X_2^n)$. Terminal $\mathcal{X}_0$ sets $\hat{K}_1 = \phi(\hat{X}_1^n)$ and $\hat{K}_2 = \psi(\hat{X}_2^n)$. If the decoding error vanishes asymptotically (i.e., (12)-(18) are satisfied), we have

$$\Pr\{K_1 = \hat{K}_1\} > 1 - \epsilon, \tag{19}$$
$$\Pr\{K_2 = \hat{K}_2\} > 1 - \epsilon. \tag{20}$$

Secrecy: We derive the sufficient conditions for achieving the secrecy requirement (2). Then these sufficient conditions need to be verified for each of the four cases later on.

We evaluate the key leakage rates averaged over the random codebook ensemble. Let $f := f(X_1^n)$, $g := g(X_2^n)$ and $l := l(X_3^n)$. Then, it is clear that $\mathbf{F} = \{f, g, l\}$. We further let $\phi := \phi(X_1^n)$ and $\psi := \psi(X_2^n)$. Hence, $K_1 = \phi$ and $K_2 = \psi$.

We first derive

$$\begin{aligned}
&I(K_1K_2;\mathbf{F}|\mathcal{C})\\
&= I(\phi,\psi;f,g,l|\mathcal{C})\\
&\le I(\psi;g|\mathcal{C}) + I(\psi;f,l|g,\mathcal{C}) + I(\phi;f,g,\psi,l|\mathcal{C})\\
&\le I(\psi;g|\mathcal{C}) + I(g,\psi;f,l|\mathcal{C}) + I(\phi;f|\mathcal{C}) + I(f,\phi;g,\psi,l|\mathcal{C}).
\end{aligned} \tag{21}$$

We next consider each of the four terms in (21). It can be shown that if

$$\tilde{R}_1 + R_1 < H(X_1) - 2\delta(\epsilon), \tag{22}$$

then

$$\frac{1}{n} I(\phi;f|\mathcal{C}) < \delta(\epsilon); \tag{23}$$

and if

$$\tilde{R}_2 + R_2 < H(X_2) - 2\delta(\epsilon), \tag{24}$$

then

$$\frac{1}{n} I(\psi;g|\mathcal{C}) < \delta(\epsilon). \tag{25}$$

In order to bound the second term in (21), we have the following derivation:

$$\begin{aligned}
&I(g,\psi;f,l|\mathcal{C})\\
&\quad\le I(X_2^n,g,\psi;f,l|\mathcal{C})\\
&\quad= I(X_2^n;f,l|\mathcal{C})\\
&\quad= I(X_2^n;X_1^n,X_3^n|\mathcal{C}) - I(X_2^n;X_1^n,X_3^n|f,l,\mathcal{C})\\
&\quad= H(X_1^n,X_3^n|\mathcal{C}) - H(X_1^n,X_3^n|X_2^n,\mathcal{C})\\
&\qquad - H(X_1^n,X_3^n|f,l,\mathcal{C}) + H(X_1^n,X_3^n|X_2^n,f,l,\mathcal{C})\\
&\quad\le H(X_1^n,X_3^n|\mathcal{C}) - H(X_1^n,X_3^n|X_2^n,\mathcal{C}) - [H(X_1^n,X_3^n|\mathcal{C})\\
&\qquad - n\tilde{R}_1 - n\tilde{R}_3] + H(X_1^n,X_3^n|X_2^n,f,l,\mathcal{C})\\
&\quad\le n[\tilde{R}_1 + \tilde{R}_3 - H(X_1X_3|X_2)] + H(X_1^n,X_3^n|X_2^n,f,l,\mathcal{C}).
\end{aligned}$$

It can be shown that if

$$\tilde{R}_1 + \tilde{R}_3 \le H(X_1X_3|X_2) - 2\delta(\epsilon), \tag{26}$$

then

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} H(X_1^n,X_3^n|X_2^n,f,l,\mathcal{C})\\
&\qquad < H(X_1X_3|X_2) - \tilde{R}_1 - \tilde{R}_3 + \delta(\epsilon).
\end{aligned} \tag{27}$$

Consequently,

$$\frac{1}{n} I(g,\psi;f,l|\mathcal{C}) < \delta(\epsilon). \tag{28}$$

Next we consider the last term in (21) as follows:

$$\begin{aligned}
&I(f,\phi;g,\psi,l|\mathcal{C})\\
&\quad\le I(X_1^n;g,\psi,l|\mathcal{C})\\
&\quad= I(X_1^n;X_2^n,X_3^n|\mathcal{C}) - I(X_1^n;X_2^n,X_3^n|g,\psi,l,\mathcal{C})\\
&\quad\le n[\tilde{R}_2 + R_2 + \tilde{R}_3 - H(X_2X_3|X_1)]\\
&\qquad + H(X_2^n,X_3^n|X_1^n,g,\psi,l,\mathcal{C}).
\end{aligned} \tag{29}$$

It can be shown that if

$$\tilde{R}_2 + R_2 + \tilde{R}_3 < H(X_2X_3|X_1) - 2\delta(\epsilon), \tag{30}$$

then

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{1}{n} H(X_2^n,X_3^n|X_1^n,g,\psi,l,\mathcal{C})\\
&\qquad < H(X_2X_3|X_1) - \tilde{R}_2 - R_2 - \tilde{R}_3 + \delta(\epsilon).
\end{aligned}$$

Consequently,

$$\frac{1}{n} I(f,\phi;g,\psi,l|\mathcal{C}) < \delta(\epsilon). \tag{31}$$

Therefore, (22), (24), (26) and (30) are sufficient conditions to guarantee the secrecy requirement (2).

Uniformity: Uniformity of keys is due to properties of random binning and typicality.

We next show the achievability of the point P for the four cases. For each case, it is sufficient to set the rates $\tilde{R}_1$, R_1, $\tilde{R}_2$, R_2 and $\tilde{R}_3$ to satisfy the Slepian-Wolf conditions (12)-(18) for guaranteeing correct key agreement and to satisfy the sufficient conditions (22), (24), (26) and (30) for guaranteeing secrecy.

Case 1: $R_A = I(X_1;X_0X_2X_3)$ and $R_C = I(X_0;X_1X_2X_3)$, which imply

$$H(X_3|X_0X_2) < H(X_3|X_1), \tag{32}$$
$$H(X_3|X_1X_2) < H(X_3|X_0). \tag{33}$$

Moreover, $R_A < R_C$ implies

$$H(X_2X_3|X_0) < H(X_2X_3|X_1). \tag{34}$$

The rate pair at the point P is given by $(I(X_1;X_0X_2X_3), H(X_2X_3|X_1) - H(X_2X_3|X_0))$. To achieve this rate pair, we set the binning rates in the achievable strategy as follows:

$$\tilde{R}_1 = H(X_1|X_0X_2X_3) + \epsilon, \tag{35}$$
$$R_1 = I(X_1;X_0X_2X_3) - 2\delta(\epsilon) - 2\epsilon, \tag{36}$$
$$\tilde{R}_2 = H(X_2X_3|X_0) - \tilde{R}_3 + \epsilon, \tag{37}$$
$$R_2 = H(X_2X_3|X_1) - H(X_2X_3|X_0) - 4\delta(\epsilon) - 3\epsilon, \tag{38}$$
$$\tilde{R}_3 = \min\{H(X_3|X_2), H(X_3|X_0)\} - 2\delta(\epsilon) - 2\epsilon. \tag{39}$$

It can be verified that the above rates satisfy the Slepian-Wolf conditions (12)-(18) if $\tilde{R}_3 > H(X_3|X_0X_1X_2)$. Otherwise, the rate pair can be easily achieved without the helper's assistance.

It can also be verified that the above rates (35)-(39) satisfy the sufficient conditions (22), (24), (26) and (30) for secrecy.

Case 2: $R_A = I(X_1;X_0X_2X_3)$ and $R_C = I(X_0X_3;X_1X_2)$, which imply the following two inequalities:

$$H(X_3|X_0X_2) < H(X_3|X_1), \tag{40}$$
$$H(X_3|X_0) < H(X_3|X_1X_2). \tag{41}$$

Here, $R_A < R_C$ is equivalent to

$$H(X_2|X_0X_3) < H(X_2|X_1). \tag{42}$$

The key rate pair at the point P in this case is given by $(I(X_1;X_0X_2X_3), H(X_2|X_1) - H(X_2|X_0X_3))$. To achieve

this rate pair, we set the binning rates in the achievable strategy as follows:

$$\tilde{R}_1 = H(X_1|X_0X_2X_3) + \epsilon, \tag{43}$$
$$R_1 = I(X_1; X_0X_2X_3) - 2\delta(\epsilon) - 2\epsilon, \tag{44}$$
$$\tilde{R}_2 = H(X_2|X_0X_3) + \epsilon, \tag{45}$$
$$R_2 = H(X_2|X_1) - H(X_2|X_0X_3) - 2\delta(\epsilon) - 2\epsilon, \tag{46}$$
$$\tilde{R}_3 = H(X_3|X_0) + \epsilon. \tag{47}$$

It can be verified that the above rates satisfy the Slepian-Wolf conditions (12)-(18) and the sufficient conditions (22), (24), (26) and (30) for secrecy.

Case 3: $R_A = I(X_0X_2; X_1X_3)$ and $R_C = I(X_0; X_1X_2X_3)$, which imply the following two inequalities:

$$H(X_3|X_0X_2) > H(X_3|X_1), \tag{48}$$
$$H(X_3|X_0) > H(X_3|X_1X_2). \tag{49}$$

Here, $R_A < R_C$ is equivalent to

$$H(X_2|X_0) < H(X_2|X_1X_3). \tag{50}$$

The rate pair of the point P in this case is given by $(I(X_0X_2; X_1X_3), H(X_2|X_1X_3) - H(X_2|X_0))$. To achieve this rate pair, we set the binning rates as follows:

$$\tilde{R}_1 = H(X_1X_3|X_0X_2) - H(X_3|X_1) + \epsilon, \tag{51}$$
$$R_1 = I(X_0X_2; X_1X_3) - 2\delta(\epsilon) - 2\epsilon, \tag{52}$$
$$\tilde{R}_2 = H(X_2|X_0) + \epsilon, \tag{53}$$
$$R_2 = H(X_2|X_1X_3) - H(X_2|X_0) - 2\delta(\epsilon) - 2\epsilon, \tag{54}$$
$$\tilde{R}_3 = H(X_3|X_1) + \epsilon. \tag{55}$$

It can be verified that the above rates satisfy the Slepian-Wolf conditions (12)-(18).

It can also be verified that the above rates satisfy the secrecy sufficient conditions (22), (24) and (30). The condition (26) is satisfied if

$$H(X_1X_3|X_0X_2) < H(X_1X_3|X_2). \tag{56}$$

Otherwise, we have the Markov chain $X_1X_3 - X_2 - X_0$. To show the secrecy, we derive a new condition to replace (26) to guarantee that $I(g, \psi; f, l|\mathcal{C})$ is asymptotically small.

$$\begin{aligned} &I(g, \psi; f, l|\mathcal{C}) \\ &\le I(g, \psi; X_1^n, X_3^n|\mathcal{C}) \\ &= I(X_2^n; X_1^n, X_3^n|\mathcal{C}) - I(X_2^n; X_1^n, X_3^n|g, \psi, \mathcal{C}) \\ &\le n[\tilde{R}_2 + R_2 - H(X_2|X_1X_3)] + H(X_2^n|X_1^n, X_3^n, g, \psi, \mathcal{C}). \end{aligned}$$

It can be shown that if

$$\tilde{R}_2 + R_2 \le H(X_2|X_1X_3) - 2\delta(\epsilon), \tag{57}$$

then

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n} H(X_2^n|X_1^n, X_3^n, f, l, \mathcal{C}) \\ &\quad < H(X_2|X_1X_3) - \tilde{R}_2 - R_2 + \delta(\epsilon). \end{aligned} \tag{58}$$

Thus, $I(g, \psi; f, l|C) < \delta(\epsilon)$. It is clear that rate settings (51)-(55) satisfy (57) and hence secrecy is guaranteed.

Case 4: $R_A = I(X_0X_2; X_1X_3)$ and $R_C = I(X_0X_3; X_1X_2)$, which imply the following two inequalities:

$$H(X_3|X_0X_2) > H(X_3|X_1), \tag{59}$$
$$H(X_3|X_1X_2) > H(X_3|X_0). \tag{60}$$

Then, we have $H(X_3|X_1) < H(X_3|X_0X_2) \le H(X_3|X_0) < H(X_3|X_1X_2)$, which yields contradiction. Thus, this case does not exist.

V. Conclusion

In this paper, we have studied the problem of generating a pair of keys for a cellular model with a helper. We have established the full key capacity region. This four-terminal model is more complicated to analyze than three-terminal model because the cut-set outer bound may take a number of structures due to different source distributions. Instead of designing a specific achievable scheme for each case one by one to achieve the cut-set bound, we have developed a unified strategy that achieves corner points for all cases, which significantly reduces the complexity of the achievability proof. It seems, with focus on cellular models, that generating secret keys between the cell-site and many users, extending what is done here with 2 keys, might be of future interest. It is also promising to use the unified strategy to solve the problem of generating asymmetric keys for cellular models, where keys have different secrecy requirements as proposed in [7].

Acknowledgment

The work of H. Zhang and Y. Liang was supported by a National Science Foundation CAREER Award under Grant CCF-10-26565 and by the National Science Foundation under Grant CNS-11-16932. The work of L. Lai was supported by a National Science Foundation CAREER Award under Grant CCF-13-18980 and by the National Science Foundation under Grant CNS-13-21223. The work of S. Shamai was supported by the Israel Science Foundation (ISF), and the European Commission in the framework of the Network of Excellence in Wireless COMmunications NEWCOM#.

References

[1] C. Ye and P. Narayan. The secret key-private key capacity region for three terminals. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, September 2005.

[2] H. Zhang, L. Lai, Y. Liang, and H. Wang. The secret key-private key generation over three terminals: Capacity region. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Honolulu, HI, USA, June 2014.

[3] 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*, 60(10):6389–6398, October 2014.

[4] H. Zhang, Y. Liang, and L. Lai. Key capacity region for a cellular source model. In *Proc. IEEE Information Theory Workshop (ITW)*, Hobart, Tasmania, Australia, November 2014.

[5] I. Csiszár and P. Narayan. Secrecy capacities for mulitple terminals. *IEEE Trans. Inform. Theory*, 50(12):3047–3061, December 2004.

[6] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, IT-19:471–480, 1973.

[7] H. Zhang, Y. Liang, and L. Lai. Helper-assisted asymmetric two key generation. In *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, USA, November 2014.