Achieving Strong Secrecy in Discrete Memoryless Wiretap Channel via Extractors

Shahid Mehraj Shah

Dept. of ECE, IISc Bangalore, India Email: shahid@ece.iisc.ernet.in

Abstract

This monogram is simply the simplified proof of achievability part of secrecy capacity in stronger sense in [1] and [2].

Index Terms

Secret key, Physical Layer Security, Secrecy Capacity.

I. INTRODUCTION

This paper is just the simple way of re-writing the proof of achievability of secrecy capacity of DM-Wiretap channel in stronger sense, first proved by Imre Csiszar [2].

A note about the notation: capital letters, like W will denote a random variable and the corresponding small letter w its realization. An n-length vector (A_1, A_2, \ldots, A_n) will be denoted as A^n .

II. SOME PRELIMINARY RESULTS

To prove the main theorem, we need several lemmas. In this section we will simply state those lemmas. The proof of these lemmas can be found in the appendix.

Lemma 1 (Extractor Lemma): (i) For a PD P on a finite set \mathcal{U} and $\epsilon > 0$, if $P(u) \le 1/d$ for each $u \in \mathcal{U}$ then a randomly selected mapping $\pi : \mathcal{U} \to \{1, \ldots, k\}$ satisfies

$$\left| P(\pi^{-1}(i)) - \frac{1}{k} \right| \le \frac{\epsilon}{k}, \ i = 1, \dots, k,$$
 (1)

with probability at least

$$1 - 2ke^{-\epsilon^2 d/2k(1+\epsilon)}.$$
(2)

1

(ii) The weaker hypothesis $P(\{u : P(u) \le 1/d\}) \ge 1 - \eta$ still suffices for

$$\sum_{i=1}^{k} \left| P(\pi^{-1}(i)) - \frac{1}{k} \right| \le \frac{\epsilon}{k} \le \epsilon + 2\eta \tag{3}$$

to hold with probability at least

$$1 - 2ke^{-\epsilon^2(1-\eta)d/2k(1+\epsilon)}.$$
(4)

In particular, if each P in a family \mathcal{P} of PDs on \mathcal{U} satisfies that hypothesis then a randomly selected π as above will be an $(\epsilon + 2\eta)$ -extractor for the family \mathcal{P} with probability at least

$$1 - 2k|\mathcal{P}|e^{-\epsilon^2(1-\eta)d/2k(1+\epsilon)}.$$
(5)

Lemma 2 (Secrecy Lemma): Let U, V be random variables with values in finite sets U, V

1. The hypothesis

$$P_{UV}\left(\left\{(u,v): P_{U|V}(u|v) \le \frac{1}{d}\right\}\right) \ge 1 - \eta^2, \eta \le \frac{1}{3}$$
(6)

2

$$k\ln(2k|\mathcal{V}|) < \alpha^2 d, \alpha \le \frac{1}{6} \tag{7}$$

imply the existence of a mapping $\pi : \mathcal{U} \to \{1, \ldots, k\}$ satisfying the security bound

$$S(\pi(U)|V) \le (\alpha + 2\eta)\log(k) + h(\alpha + \eta).$$
(8)

with probability at least $1 - 2k|\mathcal{V}|e^{-\alpha^2 d/k}$.

2. If there exists a set $\mathcal{B} \subset \mathcal{U} \times \mathcal{V}$ such that

$$P_{UV}(u,v) < \frac{1}{\alpha |\mathcal{B}|} \text{ when } (u,v) \in \mathcal{B}$$
(9)

$$P_{UV}(\mathcal{B}) \ge 1 - (\eta^2 - \alpha^2) \tag{10}$$

with $\alpha \leq 1/6$, $\eta \leq 1/3$, then there exists a mapping $\pi : \mathcal{U} \to \{1, \ldots, k\}$ satisfying the security bound (8) provided that

$$k < \min\left[\alpha^{6} \min |\mathcal{B}_{v}|, \frac{1}{2|\mathcal{V}|}e^{1/\alpha}\right]$$
(11)

Corollary 3: For discrete memoryless multiple source (DMMS) with generic variables (X, Y) (i) to any δ there exists $\xi > 0$ such that for $k \le e^{n[H(X|Y)-\delta]}$ a randomly selected mapping $pi : \mathcal{X}^n \to \{1, \ldots, k\}$ gives

$$S(\pi(X^n)|Y^n) < e^{-\xi n}, \text{d.e.s.};$$
(12)

(ii) Let $Z^{(n)}$ be any random variable jointly distributed with (X^n, Y^n) that has at most e^{nr} possible values. Then if $k \leq e^{n[H(X|Y)-r-\delta]}$, assertion (i) remains valid for $S(\pi(X^n)|Y^n, Z^{(n)})$.

Lemma 4: Given finite sets \mathcal{X}, \mathcal{Y} , and H > 0, $\delta > 0$, for a randomly selected mapping $\pi : \mathcal{X}^n \to \{1, \ldots, k\}$ with $1/n \log k \ge H + \delta$ and a suitable decoder mapping $\phi : \mathcal{Y}^n \times \{1, \ldots, k\} \to \mathcal{X}^n$ (depending on π), it holds with probability approaching unity as $n \to \infty$ that, simultaneously for all DMMSs with geniric variables X, Y satisfying $H(X|Y) \le H$,

$$Pr\left\{\phi(Y^n, \pi(X^n)) \neq X^n\right\} \leqslant \epsilon_n,\tag{13}$$

where $\epsilon_n \to 0$ exponentially rapidly.

Lemma 5 (Typicality Lemma): It holds, assuming in (iii)-(iv) that n is sufficiently large:

(i)

$$\left| -\frac{1}{n} \log P_X^n(x^n) - H(X) \right| < \tau \text{ if } x^n \in \mathcal{T}_{[X]_{\xi}},\tag{14}$$

$$\left|-\frac{1}{n}\log P_X^n(x^n|z^n) - H(X|Z)\right| < \tau \text{ if } (x^n, z^n) \in \mathcal{T}_{[UX]_{\xi}},\tag{15}$$

(iii)

$$\left|\frac{1}{n}\log|T_{[X]_{\xi}}| - H(X)\right| < \tau \tag{16}$$

$$\left|\frac{1}{n}\log\left|T_{[UX]_{\xi}}(z^{n})\right| - H(X|Z)\right| < \tau \text{ if } \mathcal{T}_{[UX]_{\xi}}(z^{n}) \neq \emptyset$$
(17)

Lemma 6: The probability that in N independent trials an event of probability q occurs less/more than $\alpha q N$ times, according as $\alpha \ge 1$, is bounded above by $e^{-c(\alpha)Nq}$, where $c(\alpha) = \alpha \ln \alpha - \alpha + 1$.

Corollary 7A: Consider $N = 2^{nR}$ sequences $x_i^n \in \mathcal{X}^n$ independently drawn from the distribution P_X^n . (i) To any $\xi > 0$ there exists $\rho > 0$ such that all but a fraction $e^{-\rho n}$ of the sequences x_i^n are ξ - typical, d.e.s. (ii) If I(X;Y) < R, to any $\tau > 0$ there exists $\zeta > 0$ such that

$$\left|\frac{1}{n}\log\left|\left\{i:x_i^n\in\mathcal{T}_{[XY]_{\zeta}}(y^n)\right\}\right| - (R - I(X;Y))\right| < \tau, \text{ d.e.s.}$$

$$\tag{18}$$

Corollary 7B: Consider $N = 2^{nR}$ sequences $x_i^n \in \mathcal{X}^n$ drawn independently from the distribution $P_{X|U}^n(.|u^n)$, for given $u^n \in \mathcal{U}$. The following hold uniformly with respect to the choice of u^n .

(i) If u^n is ξ -typical and $\zeta - \xi$, there exists $\rho > 0$ such that all but a fraction $e^{(-n\rho)}$ of the x_i^n are jointly typical with u^n , d.e.s.

(ii) If I(X; Y|U) < R, to any $\tau > 0$ there exists $\sigma > 0$ such that

$$\left|\frac{1}{n}\log\left|\left\{i:x_i^n\in\mathcal{T}_{[UXY]_{\sigma}}(u^n,y^n)\right\}\right| - \left(R - I(X;Y|U)\right)\right| < \tau, \text{ d.e.s.}$$

$$\tag{19}$$

simultaneously for all $y^n \in \mathcal{Y}^n$ with $\mathcal{T}_{[UXY]_{\sigma}}(u^n, y^n) \neq \emptyset$, with $\zeta < \sigma$.

III. STRONG SECRECY

Theorem 1: For wiretap channel, the secrecy capacity achieved is

$$C_S = \max\left[I(X;Y) - I(X;Z)\right] \tag{20}$$

in strong sence.

Proof:Let $\delta > 0$ be a fixed small number and we select randomly $N = e^{n[I(X;Y)-\delta]}$ sequences $X_i^n \in \mathcal{X}^n$ from the distribution P_X^n . It is well known that with high probability, the resulting set $\mathcal{U} \triangleq \{X_1^n, X_2^n, \ldots, X_N^n\}$ is the codeword set of a code for channel $\{W_1\}$ whose average probability of error is exponentially small.

Let U be a random variable uniformaly distributed on \mathcal{U} , and let Y^n , Z^n be the ourputs of the channels $\{W_1\}$, $\{W_2\}$ corresponding to input U. Due to decodability of U from Y^n the achievability of $R \triangleq I(X;Y) - I(X;Z)$ (in the stronger sense) will follow if we prove the existence of a mapping $\pi : \mathcal{U} \to \{1, 2, \dots, k\}$ with $\frac{1}{n} \log k$ arbitrarily close to R s.t. $M = \pi(U)$ satisfies the secrecy condition, i.e.

$$S(M|Z^n) = \log(M) - H(M|Z^n) < \epsilon.$$
⁽²¹⁾

In this case $V = Z^n$, and we apply (*ii*) part to the set of all jointly ζ -typical pairs $(x_i^n, z^n) \in \mathcal{U} \times Z^n$ in the role of B. Now we know from secrecy lemma part (ii), the set in the lemma \mathcal{B} should satisfy property (9). Now from (15) we have

$$P_{X|Z}^{n}(x^{n}|z^{n}) < 2^{-n(H(X|Z)-\tau)}$$
(22)

Now from (17), we have

$$H(X|Z) > \frac{1}{n}\log|\mathcal{B}| - \tau \tag{23}$$

Now from (22) and (23) we have

$$P_{X|Z}^{n}(x^{n}|z^{n}) < 2^{-\left[\frac{1}{n}\log|\mathcal{B}|-\tau-\tau\right]}$$
$$= \frac{2^{-n\tau}}{|\mathcal{B}|}$$
(24)

Now we have shown that the condition (9) is satisfied with $\alpha = 2^{-n\tau}$, where τ is arbitrarily small if ζ is. Now we need to show that (10) is also satisfied. This condition can be equivalently written as $P_{XZ}(\mathcal{B}^c)$ is exponentially small, (A^c denotes complement of set A). To show this consider the following

$$P_U\left(\mathcal{U} - \mathcal{T}_{[X]_{xi}}\right) = \frac{1}{N} \left| \{i : x_i^n \notin \mathcal{T}_{[X]_{\zeta}}\} \right|$$
(25)

where A - B denotes all the elements that belong to A but not to B. Hence this equation is equivalent to saying that what is the probability that the codeword sequences are not typical. We know from corollary 7A that (25) is exponentially small for any $\xi > 0$, and if $x_i^n \in \mathcal{U} \cap \mathcal{T}_{[X]_{\xi}}, \xi < \zeta$, then

$$W_2\left(\overline{T}_{[XZ]_{\zeta}}(x_i^n)|x_i^n\right) \le 2|\mathcal{X}||\mathcal{Z}|e^{-2(\zeta-\xi)^2n}$$
(26)

by part (i) of typicality lemma.

Finally, we need a lower bound to $|\mathcal{B}_v|$, i.e., the number of those $x_i^n \in \mathcal{U}$ which are jointly typical with a given z^n in the projection of \mathcal{B} to \mathcal{Z}^n . As such z^n satisfy $\mathcal{T}_{XZ_{\zeta}}(z^n) \neq \emptyset$, corollary 7A applied with Z in the role of Y yields

$$\left|\frac{1}{n}\log\left|\left\{i:x_i^n\in\mathcal{T}_{[XZ]_{\zeta}}(z^n)\right\}\right|-I(X;Y)-\delta-I(X;Z)\right|<\tau, \text{ d.e.s.},$$
(27)

where τ is arbitrarily small if ζ is. This completes the consideration needed to apply part (ii) of Lemma 2 to this scenario, and it follows that k there can grow with an exponential rate arbitrarily close R = I(X;Y) - I(X;Z), if δ and ζ are small enough. \Box

IV. CONCLUSION

In this paper we have achieved secrecy rate equal to the main channel capacity by using previous secret messages as key for transmitting the current message. This can be done while still retaining *strong secrecy*.

APPENDIX A

PROOF OF LEMMAS

A. Proof of Lemma 1

: We fix $i \in \{1, ..., k\}$ for a random mapping π . Define an indicator random variable as

$$\chi(x) = \begin{cases} 1, & \text{if } \pi(u) = i \\ 0, & \text{otherwise} \end{cases}$$

It is easy to see that $\chi(u), u \in \mathcal{U}$, are independent, identically distributed (i.i.d.) random variables with $Pr\{\chi(u) = 1\} = 1/k$. Next we define following probability distribution as

$$P(\pi^{-1}(i)) = \sum_{u \in \mathcal{U}} P(u)\chi(u).$$
⁽²⁸⁾

Applying Chernoff bound on (28) we get, for any $\beta > 0$.

$$Pr\left\{P(\pi^{-1}(i)) > \frac{1+\epsilon}{k} > \right\} = Pr\left\{2^{\left(\beta d \sum_{u \in \mathcal{U}} P(u)\chi(u)\right)} > 2^{\left(\beta d \frac{1+\epsilon}{k}\right)}\right\}$$

$$\leq \frac{\mathbb{E}\left[2^{\beta d \sum_{u \in \mathcal{U}} P(u)\chi(u)}\right]}{2^{\beta d \frac{1+\epsilon}{k}}}$$

$$\stackrel{(a)}{=} 2^{-\beta d \frac{1+\epsilon}{k}} \mathbb{E}\left[\prod_{u \in \mathcal{U}} 2^{\beta d P(u)\chi(u)}\right]$$

$$\stackrel{(b)}{=} 2^{-\beta d \frac{1+\epsilon}{k}} \prod_{u \in \mathcal{U}} \left[\mathbb{E}\left(2^{\beta d P(u)\chi(u)}\right)\right]$$

$$\stackrel{(c)}{=} 2^{-\beta d \frac{1+\epsilon}{k}} \prod_{u \in \mathcal{U}} \left\{Pr\{\chi(u)=1\}2^{\beta d P(u)} + Pr\{\chi(u)=0\}2^{0}\right\}$$

$$= 2^{-\beta d \frac{1+\epsilon}{k}} \prod_{u \in \mathcal{U}} \left\{\frac{1}{k}2^{\beta d P(u)} + 1 - \frac{1}{k}\right\}$$

$$= 2^{-\beta d \frac{1+\epsilon}{k}} \prod_{u \in \mathcal{U}} \left\{1 + \frac{1}{k}\left[2^{\beta d P(u)} - 1\right]\right\}.$$
(29)

where (a) follows since \mathcal{U} is discrete set, (b) follows because $\chi(u)$ are i.i.d. and (c) follows since $Pr\{\chi(u) = 1\} = 1/k$.

Now it is by the hypothesis $P(u) \leq 1/d$ for all $u \in \mathcal{U}$, and noting that $2^x = e^{\ln(2^x)} = e^{x \ln 2}$ we have

$$2^{\beta d(u)} - 1 = \sum_{j=1}^{\infty} \frac{(\beta dP(u) \ln 2)^j}{j!}$$

= $\beta dP(u) \ln 2 \left[1 + \sum_{j=2}^{\infty} \frac{(\beta dP(u) \ln 2)^{j-1}}{j!} \right]$
 $\stackrel{(a)}{\leq} \beta dP(u) \ln 2 \left[1 + \sum_{j=2}^{\infty} \frac{(\beta \ln 2)^{j-1}}{j!} \right]$
 $\stackrel{(b)}{\leqslant} \beta dP(u) \ln 2 \left[1 + \sum_{j=2}^{\infty} \frac{1}{2} (\beta \ln 2)^{j-1} \right]$
 $\stackrel{(c)}{=} \beta dP(u) \ln 2 \left[1 + \frac{1}{2} \frac{\beta \ln 2}{1 - \beta \ln 2} \right]$
 $= \beta dP(u) \ln 2(1 + \beta^*),$ (30)

where

$$\beta^* = \frac{\beta \ln 2}{2(1 - \beta \ln 2)} \tag{31}$$

(a) follows from the hypothesis $P(u) \le 1/d$, (b) follows since j! > 2 for j > 2, (c) follows from sum of infinite geometric series, provided that $\beta^* \ln 2 < 1$.

Now we note that the for function $f(t) = 1 + t \ln 2 - 2^t$, $f'(t) = \ln 2 - 2^t \ln 2$ and $f''(t) = -2^{2t} (\ln 2)^2$. Now $f'(t) = 0 \Rightarrow t = 0$ and f''(0) < 0 hence the function f(t) is decreasing, hence for t > 0 f(t) < f(0) hence $1 + t \ln 2 < 2^t$. Now using this inequality and going back to equation (29) we have

$$2^{-\beta d \frac{1+\epsilon}{k}} \prod_{u \in \mathcal{U}} \left\{ 1 + \frac{1}{k} \left[2^{\beta d P(u)} - 1 \right] \right\}$$

$$\stackrel{(a)}{<} 2^{-\beta d \frac{1+\epsilon}{k}} \prod_{u \in \mathcal{U}} \left\{ 1 + \frac{1}{k} \left[\beta d P(u) \ln 2(1+\beta^*) \right] \right\}$$

$$\stackrel{(b)}{<} 2^{-\beta d \frac{1+\epsilon}{k}} \prod_{u \in \mathcal{U}} 2^{\frac{1}{k}\beta d P(u) \ln 2(1+\beta^*)}$$

$$= 2^{-\beta d \frac{1+\epsilon}{k}} 2^{\sum_{u \in \mathcal{U}} \frac{1}{k}\beta d P(u) \ln 2(1+\beta^*)}$$

$$\stackrel{(c)}{=} 2^{-\beta d \frac{1+\epsilon}{k}} 2^{\beta d(1+\beta^*)/k} = 2^{-\frac{\beta d}{k}(\epsilon-\beta^*)}$$
(32)

Now we let $\beta = \epsilon \log e/(1+\epsilon)$, whence $\beta^* = \frac{\epsilon \log e/(1+\epsilon)}{2(1-\epsilon \log e/(1+\epsilon))} = \frac{\epsilon \log e \ln 2}{2(1+\epsilon-\epsilon \log e \ln 2)} = \epsilon/2$, we finally get

$$Pr\left\{P(\pi^{-1}(i)) > \frac{1+\epsilon}{k} > \right\} \le 2^{-\frac{\epsilon^2 d \log e}{2k(1+\epsilon)}} = e^{-\frac{\epsilon^2 d}{2(1+\epsilon)k}}$$
(33)

Similarly we have

$$Pr\left\{P(\pi^{-1}(i)) < \frac{1-\epsilon}{k}\right\} = Pr\left\{2^{-\beta d\sum_{u\in\mathcal{U}}P(u)\chi(u)} > 2^{-\beta d\frac{1-\epsilon}{k}}\right\}$$
$$\leq 2^{\beta d\frac{1-\epsilon}{k}} \prod_{u\in\mathcal{U}}\left\{1 + \frac{1}{k}\left[2^{-\beta dP(u)} - 1\right]\right\}.$$
(34)

Now we bound the term $2^{-\beta dP(u)} - 1$ as

$$2^{-\beta dP(u)} - 1 = \sum_{j=1}^{\infty} \frac{(-\beta dP(u) \ln 2)^+}{j!}$$

$$\leq -\beta dP(u) \ln 2 \left(1 - \frac{1}{2}\beta dP(u) \ln 2\right)$$

$$\leq -\beta P(u) \ln 2 \left(1 - \frac{1}{2}\beta \ln 2\right).$$
(35)

Using (35) in (34) we get

$$Pr\left\{P(\pi^{-1}(i)) < \frac{1-\epsilon}{k}\right\} \le e^{-\epsilon^2 d/2k} \tag{36}$$

From (33) and (36), we get

$$\Pr\left\{ \left| P(\pi^{-1}(i)) - \frac{1}{k} \right| \le \frac{\epsilon}{k} \right\} \ge 1 - 2ke^{-\epsilon^2 d/2k(1+\epsilon)}.$$
(37)

B. Proof of lemma 2 (Secrecy Lemma)

Proof: We define the following set

$$\mathcal{G}_{v} = \{ u : P_{U|V}(u|v) \le 1/d \}.$$
(38)

We have

$$P_{UV}\left(\left\{(u,v): P_{U|V}(u|v) \le \frac{1}{d}\right\}\right) = \sum_{v \in \mathcal{V}} P_V(v) P_{U|V}(\mathcal{G}_v|v),\tag{39}$$

Now consider a set

$$\mathcal{E} \triangleq \left\{ v : P_{U|V}(\mathcal{G}_v|v) < 1 - \eta \right\}$$
(40)

Now we show that $P_V(\mathcal{E}) < \eta$.

Now as each probability distribution (PD) $P_{U|V}(.|v), v \notin \mathcal{E}$, satisfies hypothesis (ii) of lemma 1, i.e. the bound in (6), with arbitrary $\epsilon > 0$, holds simultaneously for each $P_{U|V}(.|v), v \notin \mathcal{E}$, in the role of P, at least with probability

$$1 - 2k|\mathcal{V}|e^{-\frac{\epsilon^2(1-\eta)}{2(1+\epsilon)}\frac{d}{k}} \tag{41}$$

For convenience we set $\epsilon = 2\alpha, \alpha \le 1/6$, then $\epsilon^2(1-\eta)/2(1+\epsilon) > \alpha^2$ (if $\eta \le 1/3$), and thus the final probability is at least

$$1 - 2k|\mathcal{V}|e^{-\alpha^2 d/k}.\tag{42}$$

Now we use lemma 8 to prove the security index bound. The variation distance bound (10) allows us to bound the difference of $(H(\pi(U)|V=v)$ from the entropy $\log k$ of the uniform distribution, e.g., by lemma 8. We use one more sharper bound of lemma 9, to obtain for $v \notin \mathcal{E}$

$$\log k - H(\pi(U)|V=v) \le \frac{\epsilon + 2\eta}{2} \log k + h(\frac{\epsilon + 2\eta}{2},$$
(43)

where $\epsilon = 2\alpha$. For $v \in \mathcal{E}$ we bound the entropy difference trivially by $\log k$. Thus by definition of security index, we get the required bound.

(ii) Since it is given that \mathcal{B} satisfies (9), we have some notations

$$\mathcal{B}' \triangleq \mathcal{B} \bigcap \{ \mathcal{U} \times \mathcal{C} \}, \text{ where } \mathcal{C} \triangleq \left\{ v : P_V(v) \ge \frac{\alpha^2 |\mathcal{B}_v|}{|\mathcal{B}|} \right\}.$$
(44)

Thus for $(u, v) \in \mathcal{B}'$ we have

$$P_{U|V}(u|v) = \frac{P_{UV}(u,v)}{P_V(v)}$$

$$\leq \frac{(\alpha|\mathcal{B}|)^{-1}}{\alpha^2|\mathcal{B}_v|.|\mathcal{B}|^{-1}}$$

$$\leq \frac{1}{\alpha^3 \min|\mathcal{B}_v|}$$
(45)

This implies that (6) holds with $d = \alpha^3 \min |\mathcal{B}_v|$, because $P_{UV}(\mathcal{B}' \ge P_{UV}(\mathcal{B} - P_V(\overline{\mathcal{C}}))$, where $P_V(\overline{\mathcal{C}}) < \alpha^2$ due to $\sum_{v \in \mathcal{V}} |\mathcal{B}_v| = |\mathcal{B}|$.

REFERENCES

[1] I. Csiszar and J. Körner, Information theory: coding theorems for discrete memoryless systems. Cambridge University Press, 2011.

[2] I. Csiszár, "Almost independence and secrecy capacity," Problemy Peredachi Informatsii, vol. 32, no. 1, pp. 48–57, 1996.