Partial Strong Converse for the Non-Degraded Wiretap Channel

Yi-Peng Wei Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
ypwei@umd.edu *ulukus@umd.edu*

***Abstract*—We prove the partial strong converse property for the discrete memoryless non-degraded wiretap channel, for which we require the leakage to the eavesdropper to vanish but allow an asymptotic error probability $\epsilon \in [0,1)$. We use a recently developed technique based on information spectrum and a recursive bounding method to evaluate the probability of correct decoding. We show that when the transmission rate is above the secrecy capacity, the probability of correct decoding decays to zero exponentially. Therefore, the maximum transmission rate is the same for $\epsilon \in [0,1)$, and the partial strong converse property holds.**

I. Introduction

We consider the discrete memoryless non-degraded wiretap channel, in which a transmitter wishes to send messages to a legitimate receiver while keeping the messages secret from an eavesdropper. The wiretap channel was first studied in [1] with the assumption that the wiretap channel is degraded, and later on the secrecy capacity of *non-degraded* wiretap channel was shown in [2]. The general formula for the wiretap channel can be found in [3]. Although [1] and [2] provide the secrecy capacity for the wiretap channel, the proof relies on Fano's inequality and therefore only a weak converse can be shown.

Strong converse property was first proposed in [4] for the point-to-point channel, and has received significant attention recently due to the study of finite block-length channel coding rate [5]–[7]. For the point-to-point channel, strong converse property states that when the transmission rate is above the capacity, the asymptotic error probability goes to 1. It also implies that if we allow a potentially non-zero asymptotic error probability $\epsilon \in [0,1)$, the maximal transmission rate is the same as the capacity, which only allows $\epsilon = 0$. In [8], the authors build equivalent conditions for the strong converse property with the information spectrum method for the point-to-point channel, and later on in [9, Sec. 3.7], the author extends it to channels with cost constraints.

The maximal transmission rate for the wiretap channel is constrained by two conditions: reliability and security. We use the asymptotic error probability ϵ for the reliability condition and variational distance δ for the secrecy constraint[1]. In [10], the authors extend the method in [9, Sec. 3.7] to show the strong converse property for $(\epsilon, \delta) \in [0,1) \times \{0\}$ for the degraded wiretap channel, and name it *partial strong converse* to account for the strict secrecy constraint. In [11], the authors utilize the relationship between the wiretap channel with feedback and secret key agreement [12] to show that the strong converse property holds when $\epsilon + \delta < 1$ for the degraded wiretap channel. To the best of our knowledge, there is no strong converse property reported for the general *non-degraded* wiretap channel.

Recently, a new strong converse technique has been proposed in [13]–[16]. This technique is based on a novel usage of information spectrum method [9] and a new recursive bounding method. It lower bounds the exponent function of the probability of correct decoding, and therefore shows that the probability of correct decoding goes to zero exponentially when the rate is above the capacity. This technique is general and has been applied to broadcast channels in [13], [14], state dependent channels in [15] and for Wyner-Ziv coding in [16]. In this work, we show that the partial strong converse property holds for the discrete memoryless non-degraded wiretap channel based on this new technique.

II. Problem setting and Main Results

A. System Model and Definitions

A wiretap channel consists of a transmitter (Alice) who wishes to send a message uniformly distributed in $\mathcal{M}_n$ to a legitimate receiver (Bob) secretly in the presence of an eavesdropper (Eve) through a channel $W^n : \mathcal{X}^n \to \mathcal{Y}^n \times \mathcal{Z}^n$. $\mathcal{X}$ denotes the input alphabet for Alice, while $\mathcal{Y}$ and $\mathcal{Z}$ denote the output alphabets for Bob and Eve, respectively. $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ are finite. Here, we consider the discrete memoryless channel, and therefore, we have

$$W^n(y^n, z^n|x^n) = \prod_{t=1}^{n} W(y_t, z_t|x_t), \tag{1}$$

where $W(y,z|x)$ is the conditional probability mass function (pmf). Moreover, we define

$$W_1(y|x) \triangleq \sum_{z\in\mathcal{Z}} W(y,z|x), \;\; W_2(z|x) \triangleq \sum_{y\in\mathcal{Y}} W(y,z|x). \tag{2}$$

In the following, X^n denotes a random variable taking values in $\mathcal{X}^n$, and the elements of $\mathcal{X}^n$ are denoted by x^n. The pmf of random variable X^n is denoted by p_{X^n}. Similar notation

This work was supported by NSF Grants CNS 13-14733, CCF 14-22111, CCF 14-22129, and CNS 15-26608.

[1]There are various kinds of secrecy constraints [3, Proposition 1]. For instance, [1] and [2] use normalized mutual information, and [10] and [11] use variational distance.

also applies to other random variables. To satisfy the secrecy constraint, we require

$$\lim_{n\to\infty}\frac{1}{n}I(M_n;Z^n)=0. \tag{3}$$

The encoder $\phi^{(n)}$ maps the message $m\in\mathcal{M}_n$ to a codeword $x^n\in\mathcal{X}^n$. We allow the encoder $\phi^{(n)}$ to be a stochastic encoder and denote it as $\phi^{(n)}=\{\phi^{(n)}(x^n|m)\}_{(m,x^n)\in\mathcal{M}_n\times\mathcal{X}^n}$, where $\phi^{(n)}(x^n|m)$ is a conditional pmf. The decoder is denoted by $\psi^{(n)}$ such that $\psi^{(n)}:\mathcal{Y}^n\to\mathcal{M}_n$. The joint pmf on $\mathcal{M}_n\times\mathcal{X}^n\times\mathcal{Y}^n\times\mathcal{Z}^n$ is given by

$$p_{M_nX^nY^nZ^n}(m,x^n,y^n,z^n)=\frac{1}{|\mathcal{M}_n|}\phi^{(n)}(x^n|m)\prod_{t=1}^{n}W(y_t,z_t|x_t). \tag{4}$$

The average probability of correct decoding is given by

$$\mathrm{P}_c^{(n)}=\mathrm{P}_c^{(n)}(\phi^{(n)},\psi^{(n)})\triangleq\Pr\{\psi^{(n)}(Y^n)=M_n\}. \tag{5}$$

The average probability of error is $\mathrm{P}_e^{(n)}=1-\mathrm{P}_c^{(n)}$. For fixed $\epsilon\in[0,1)$, a rate R is *ϵ-achievable* if there exists a sequence of codes $\{\phi^{(n)},\psi^{(n)}\}_{n=1}^{\infty}$ such that

$$\limsup_{n\to\infty}\mathrm{P}_e^{(n)}\le\epsilon,\quad\liminf_{n\to\infty}\frac{1}{n}\log|\mathcal{M}_n|\ge R. \tag{6}$$

In this work, we consider the partial strong converse property. Therefore, we require the code to satisfy (3). Let $\mathcal{C}_s(\epsilon|W)$ denote the maximal *ϵ-achievable* rate satisfying (3) through the wiretap channel $W(y,z|x)$. From [2], we have

$$\mathcal{C}_s(0|W)=\max_{p\in\mathcal{P}(W)}I(U;Y)-I(U;Z), \tag{7}$$

where $\mathcal{P}(W)$ is defined as

$$\mathcal{P}(W)\triangleq\big\{p_{UXYZ}(u,x,y,z):|\mathcal{U}|\le|\mathcal{X}|,\ p_{YZ|X}(y,z|x)=W(y,z|x),U\to X\to(Y,Z)\big\}. \tag{8}$$

In the following, we will evaluate the asymptotic behavior of $\mathrm{P}_c^{(n)}$. Under the condition that (3) is satisfied, we define the following quantity for the exponent of the probability of correct decoding:

$$G^{(n)}(R|W)\triangleq\min_{\substack{(\phi^{(n)},\psi^{(n)}):\\ \frac{1}{n}\log|\mathcal{M}_n|\ge R.}}\left(-\frac{1}{n}\right)\log\mathrm{P}_c^{(n)}(\phi^{(n)},\psi^{(n)}) \tag{9}$$

$$G(R|W)=\inf_{n\ge1}G^{(n)}(R|W) \tag{10}$$

B. Main Result

Theorem 1. *For a discrete memoryless non-degraded wiretap channel $W(y,z|x)$, the partial strong converse property holds, i.e., for $\epsilon\in[0,1)$, for any code $(\phi^{(n)},\psi^{(n)})$ satisfying* (3)*, we have*

$$\mathcal{C}_s(\epsilon|W)=\mathcal{C}_s(0|W). \tag{11}$$

The proof is provided in Section III. We apply the new strong converse proof technique developed in [13]–[16]. We lower bound the probability of correct decoding exponent $G(R|W)$ as defined in (10). We show that $G(R|W)>0$ when the rate is above $\mathcal{C}_s(0|W)$. Therefore, $\mathrm{P}_c^{(n)}$ decays to zero exponentially, and the partial strong converse property holds.

III. Proof of the Main Result

Lemma 1. *For any $\eta>0$ and for any $(\phi^{(n)},\psi^{(n)})$ satisfying $\frac{1}{n}\log|\mathcal{M}_n|\ge R$ and* (3)*, we have*

$$\begin{aligned}\mathrm{P}_c^{(n)}(\phi^{(n)},\psi^{(n)})\le p_{M_nX^nY^nZ^n}\Bigg\{&\\ 0\le&\frac{1}{n}\log\frac{W_1^n(Y^n|X^n)}{q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}(Y^n|X^n,Z^n,M_n)}+\eta,\\ 0\le&\frac{1}{n}\log\frac{W_2^n(Z^n|X^n)}{q^{(\mathrm{ii})}_{Z^n|X^nY^nM_n}(Z^n|X^n,Y^n,M_n)}+\eta,\\ R\le&\frac{1}{n}\log\frac{p_{Y^n|M_n}(Y^n|M_n)}{q^{(\mathrm{iii})}_{Y^n}(Y^n)}+\eta\Bigg\}+3e^{-n\eta}.\end{aligned} \tag{12}$$

In (12)*, we can choose any conditional pmf $q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}$ on $\mathcal{Y}^n$ given (X^n,Z^n,M_n), $q^{(\mathrm{ii})}_{Z^n|X^nY^nM_n}$ on $\mathcal{Z}^n$ given (X^n,Y^n,M_n) and any pmf $q^{(\mathrm{iii})}_{Y^n}$ on $\mathcal{Y}^n$.*

The proof of Lemma 1 can be found in Section IV-A.

Lemma 2. *For any $\eta>0$ and for any $(\phi^{(n)},\psi^{(n)})$ satisfying $\frac{1}{n}\log|\mathcal{M}_n|\ge R$ and* (3)*, we have*

$$\begin{aligned}\mathrm{P}_c^{(n)}(\phi^{(n)},\psi^{(n)})\le p_{M_nX^nY^nZ^n}\Bigg\{&\\ 0\le&\frac{1}{n}\log\frac{W_1^n(Y^n|X^n)}{q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}(Y^n|X^n,Z^n,M_n)}+\eta,\\ 0\le&\frac{1}{n}\log\frac{W_2^n(Z^n|X^n)}{q^{(\mathrm{ii})}_{Z^n|X^nY^nM_n}(Z^n|X^n,Y^n,M_n)}+\eta,\\ R\le&\frac{1}{n}\log\frac{q^{(\mathrm{iv})}_{Y^n|M_n}(Y^n|M_n)q^{(\mathrm{iii})}_{Z^n}(Z^n)}{q^{(\mathrm{iv})}_{Z^n|M_n}(Z^n|M_n)q^{(\mathrm{iii})}_{Y^n}(Y^n)}\\ &+\frac{1}{n}\log\frac{p_{Y^n|M_n}(Y^n|M_n)}{q^{(\mathrm{iv})}_{Y^n|M_n}(Y^n|M_n)}\\ &+\frac{1}{n}\log\frac{q^{(\mathrm{iv})}_{Z^n|M_n}(Z^n|M_n)}{q^{(\mathrm{iii})}_{Z^n}(Z^n)}+\eta\Bigg\}+3e^{-n\eta}.\end{aligned} \tag{13}$$

In (13)*, $q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}$, $q^{(\mathrm{ii})}_{Z^n|X^nY^nM_n}$ and $q^{(\mathrm{iii})}_{Y^n}$ are chosen to be the same as* (12)*. We can choose arbitrary conditional pmf $(q^{(\mathrm{iv})}_{Y^n|M_n},q^{(\mathrm{iv})}_{Z^n|M_n})$ and arbitrary pmf $q^{(\mathrm{iii})}_{Z^n}$ in* (13)*.*

The proof of Lemma 2 follows from Lemma 1.

Later on, we will identify the following auxiliary random variables. For $t=1,2,\cdots,n$, set

$$\mathcal{U}_t\triangleq\mathcal{Y}^{t-1}\mathcal{Z}_{t+1}^n\mathcal{M}_n,\quad U_t\triangleq(Y^{t-1},Z_{t+1}^n,M_n)\in\mathcal{U}_t, \tag{14}$$

$$\mathcal{V}_t\triangleq\mathcal{Y}^{t-1}\mathcal{Z}_{t+1}^n,\qquad V_t\triangleq(Y^{t-1},Z_{t+1}^n)\in\mathcal{V}_t, \tag{15}$$

$$\hat{\mathcal{U}}_t\triangleq\mathcal{Y}^{t-1}\mathcal{M}_n,\qquad \hat{U}_t\triangleq(Y^{t-1},M_n)\in\hat{\mathcal{U}}_t. \tag{16}$$

Lemma 3. *For any $\eta > 0$ and for any $(\phi^{(n)}, \psi^{(n)})$ satisfying $\frac{1}{n}\log|\mathcal{M}_n| \geq R$ and* (3)*, we have*

$$\begin{aligned}
\mathrm{P}_c^{(n)}(\phi^{(n)},\psi^{(n)}) \leq p_{M_n X^n Y^n Z^n}\Bigg\{ & \\
0 \leq & \frac{1}{n}\sum_{t=1}^{n}\log\frac{W_1(Y_t|X_t)}{q^{(\mathrm{i})}_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t)} + \eta, \\
0 \leq & \frac{1}{n}\sum_{t=1}^{n}\log\frac{W_2(Z_t|X_t)}{q^{(\mathrm{ii})}_{Z_t|X_tY_tU_t}(Z_t|X_t,Y_t,U_t)} + \eta, \\
R \leq & \frac{1}{n}\sum_{t=1}^{n}\log\frac{q^{(\mathrm{iv})}_{Y_t|U_t}(Y_t|U_t)q^{(\mathrm{iii})}_{Z_t|V_t}(Z_t|V_t)}{q^{(\mathrm{iv})}_{Z_t|U_t}(Z_t|U_t)q^{(\mathrm{iii})}_{Y_t|V_t}(Y_t|V_t)} \\
& + \frac{1}{n}\sum_{t=1}^{n}\log\frac{p_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q^{(\mathrm{iv})}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)} + \eta\Bigg\} + 3e^{-n\eta}. \quad (17)
\end{aligned}$$

In (17)*, we can choose arbitrary conditional pmfs: $q^{(\mathrm{i})}_{Y_t|X_tZ_tU_t}, q^{(\mathrm{ii})}_{Z_t|X_tY_tU_t}, q^{(\mathrm{iii})}_{Y_t|V_t}, q^{(\mathrm{iii})}_{Z_t|V_t}, q^{(\mathrm{iii})}_{Z_t|\tilde{U}_t}, q^{(\mathrm{iv})}_{Y_t|U_t}, q^{(\mathrm{iv})}_{Z_t|U_t}$, $q^{(\mathrm{iv})}_{Y_t|\hat{U}_t}, q^{(\mathrm{iv})}_{Z_t|\tilde{U}_t}$, for $t = 1, 2, \cdots, n$.*

Proof: We first consider $q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}$ in (13). We choose

$$\begin{aligned}
& q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}(Y^n|X^n,Z^n,M_n) \\
&= \prod_{t=1}^{n} q^{(\mathrm{i})}_{Y_t|X^nY^{t-1}Z^nM_n}(Y_t|X^n,Y^{t-1},Z^n,M_n) \quad (18)\\
&= \prod_{t=1}^{n} q^{(\mathrm{i})}_{Y_t|X_tY^{t-1}Z_t^nM_n}(Y_t|X_t,Y^{t-1},Z_t^n,M_n) \quad (19)\\
&= \prod_{t=1}^{n} q^{(\mathrm{i})}_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t), \quad (20)
\end{aligned}$$

where (18) is due to the chain rule of probability. We choose the arbitrary conditional pmf $q^{(\mathrm{i})}_{Y_t|X^nY^{t-1}Z^nM_n} = q^{(\mathrm{i})}_{Y_t|X_tY^{t-1}Z_t^nM_n}$ in (19). (20) holds according to (14). $q^{(\mathrm{ii})}_{Z^n|X^nY^nM_n}$ in (13) can be processed similarly such that $q^{(\mathrm{ii})}_{Z^n|X^nY^nM_n} = \prod_{t=1}^{n} q^{(\mathrm{ii})}_{Z_t|X_tY_tU_t}$.

We then consider $\frac{q^{(\mathrm{iv})}_{Y^n|M_n}}{q^{(\mathrm{iv})}_{Z^n|M_n}}$ in (13). We have

$$\begin{aligned}
\frac{q^{(\mathrm{iv})}_{Y^n|M_n}(Y^n|M_n)}{q^{(\mathrm{iv})}_{Z^n|M_n}(Z^n|M_n)} &= \frac{q^{(\mathrm{iv})}_{Y^n|Z^nM_n}(Y^n|Z^n,M_n)}{q^{(\mathrm{iv})}_{Z^n|Y^nM_n}(Z^n|Y^n,M_n)} \quad (21)\\
&= \prod_{t=1}^{n}\frac{q^{(\mathrm{iv})}_{Y_t|Y^{t-1}Z^nM_n}(Y_t|Y^{t-1},Z^n,M_n)}{q^{(\mathrm{iv})}_{Z_t|Y^nZ_{t+1}^nM_n}(Z_t|Y^n,Z_{t+1}^n,M_n)} \quad (22)\\
&= \prod_{t=1}^{n}\frac{q^{(\mathrm{iv})}_{Y_t|Y^{t-1}Z_{t+1}^nM_n}(Y_t|Y^{t-1},Z_{t+1}^n,M_n)}{q^{(\mathrm{iv})}_{Z_t|Y^{t-1}Z_{t+1}^nM_n}(Z_t|Y^{t-1},Z_{t+1}^n,M_n)} \quad (23)\\
&= \prod_{t=1}^{n}\frac{q^{(\mathrm{iv})}_{Y_t|U_t}(Y_t|U_t)}{q^{(\mathrm{iv})}_{Z_t|U_t}(Z_t|U_t)}. \quad (24)
\end{aligned}$$

In (21), the equality holds for every $q^{(\mathrm{iv})}_{Y^nZ^nM_n}$. (22) comes from the chain rule of probability. Since we can choose arbitrary conditional pmf, we pick $q^{(\mathrm{iv})}_{Y_t|Y^{t-1}Z^nM_n} = q^{(\mathrm{iv})}_{Y_t|Y^{t-1}Z_{t+1}^nM_n}$ and $q^{(\mathrm{iv})}_{Z_t|Y^nZ_{t+1}^nM_n} = q^{(\mathrm{iv})}_{Z_t|Y^{t-1}Z_{t+1}^nM_n}$. Therefore, (23) holds. Finally, according to (14), we have (24).

For $\frac{q^{(\mathrm{iii})}_{Z^n}}{q^{(\mathrm{iii})}_{Y^n}}$ in (13), we have

$$\begin{aligned}
\frac{q^{(\mathrm{iii})}_{Z^n}(Z^n)}{q^{(\mathrm{iii})}_{Y^n}(Y^n)} &= \frac{q^{(\mathrm{iii})}_{Z^n|Y^n}(Z^n|Y^n)}{q^{(\mathrm{iii})}_{Y^n|Z^n}(Y^n|Z^n)} \quad (25)\\
&= \prod_{t=1}^{n}\frac{q^{(\mathrm{iii})}_{Z_t|Y^nZ_{t+1}^n}(Z_t|Y^n,Z_{t+1}^n)}{q^{(\mathrm{iii})}_{Y_t|Y^{t-1}Z^n}(Y_t|Y^{t-1},Z^n)} \quad (26)\\
&= \prod_{t=1}^{n}\frac{q^{(\mathrm{iii})}_{Z_t|Y^{t-1}Z_{t+1}^n}(Z_t|Y^{t-1},Z_{t+1}^n)}{q^{(\mathrm{iii})}_{Y_t|Y^{t-1}Z_{t+1}^n}(Y_t|Y^{t-1},Z_{t+1}^n)} \quad (27)\\
&= \prod_{t=1}^{n}\frac{q^{(\mathrm{iii})}_{Z_t|V_t}(Z_t|V_t)}{q^{(\mathrm{iii})}_{Y_t|V_t}(Y_t|V_t)}. \quad (28)
\end{aligned}$$

Note that (28) holds according to (15).

For $\frac{p_{Y^n|M_n}}{q^{(\mathrm{iv})}_{Y^n|M_n}}$ in (13), we have

$$\begin{aligned}
\frac{p_{Y^n|M_n}(Y^n|M_n)}{q^{(\mathrm{iv})}_{Y^n|M_n}(Y^n|M_n)} &= \prod_{t=1}^{n}\frac{p_{Y_t|Y^{t-1}M_n}(Y_t|Y^{t-1},M_n)}{q^{(\mathrm{iv})}_{Y_t|Y^{t-1}M_n}(Y_t|Y^{t-1},M_n)} \quad (29)\\
&= \prod_{t=1}^{n}\frac{p_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q^{(\mathrm{iv})}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}, \quad (30)
\end{aligned}$$

where (30) holds by (16).

Finally, for $\frac{q^{(\mathrm{iv})}_{Z^n|M_n}}{q^{(\mathrm{iii})}_{Z^n}}$ in (13), since $q^{(\mathrm{iv})}_{Z^n|M_n}$ is arbitrary, we choose $q^{(\mathrm{iv})}_{Z^n|M_n} = q^{(\mathrm{iii})}_{Z^n}$. We have

$$\frac{q^{(\mathrm{iv})}_{Z^n|M_n}(Z^n|M_n)}{q^{(\mathrm{iii})}_{Z^n}(Z^n)} = 1. \quad (31)$$

By (24), (28), (30), (31) and (13), we have (17). ■

For $t = 1, 2, \cdots, n$, let $\mathcal{Q}(\mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$ be a set of all pmfs on $\mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} = \mathcal{M}_n \times \mathcal{X} \times \mathcal{Y}^t \times \mathcal{Z}^{n-t+1}$. The following notation simplifies the expressions:

$$\begin{aligned}
\mathcal{Q}_t &= \mathcal{Q}(\mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}), \quad (32)\\
q_t &= q_{U_tX_tY_tZ_t} \in \mathcal{Q}_t, \quad (33)\\
\mathcal{Q}^n &\triangleq \prod_{t=1}^{n}\mathcal{Q}_t, \quad (34)\\
q^n &\triangleq \{q_t\}_{t=1}^{n} \in \mathcal{Q}^n. \quad (35)
\end{aligned}$$

By the properties of pmf in Lemma 3, we have the following lemma.

Lemma 4. *For any $\eta > 0$ and for any $(\phi^{(n)}, \psi^{(n)})$ satisfying $\frac{1}{n}\log|\mathcal{M}_n| \geq R$ and* (3)*, we have*

$$\begin{aligned}
\mathrm{P}_c^{(n)}(\phi^{(n)},\psi^{(n)}) \leq p_{M_n X^n Y^n Z^n}\Bigg\{ & \\
0 \leq & \frac{1}{n}\sum_{t=1}^{n}\log\frac{W_1(Y_t|X_t)}{q_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t)} + \eta, \\
0 \leq & \frac{1}{n}\sum_{t=1}^{n}\log\frac{W_2(Z_t|X_t)}{q_{Z_t|X_tY_tU_t}(Z_t|X_t,Y_t,U_t)} + \eta,
\end{aligned}$$

$$R \le \frac{1}{n}\sum_{t=1}^{n}\log\frac{q_{Y_t|U_t}(Y_t|U_t)q_{Z_t|V_t}(Z_t|V_t)}{q_{Z_t|U_t}(Z_t|U_t)q_{Y_t|V_t}(Y_t|V_t)} + \frac{1}{n}\sum_{t=1}^{n}\log\frac{p_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)} + \eta\Bigg\} + 3e^{-n\eta}. \quad (36)$$

In (36)*, for* $t = 1, 2, \cdots, n$*, the conditional pmfs:* $q_{Y_t|X_tZ_tU_t}, q_{Z_t|X_tY_tU_t}, q_{Y_t|V_t}, q_{Z_t|V_t}, q_{Y_t|U_t}, q_{Z_t|U_t}, q_{Y_t|\hat{U}_t}$*, are induced by the joint pmf* $q_t = q_{U_tX_tY_tZ_t} \in \mathcal{Q}_t$ *in* (33)*.*

We will use Chernoff-Cramer bound to upper bound (36).

Lemma 5 (Chernoff-Cramer)**.** *For any real valued random variable* A *and any* $\theta > 0$*, we have*

$$\Pr\{A \ge a\} \le \exp\{-[\theta a - \log \mathrm{E}[\exp(\theta A)]]\}. \quad (37)$$

By Lemma 4 and 5, we have the following lemma.

Lemma 6. *For any* $\alpha, \beta, \theta > 0$*, any* $q^n \in \mathcal{Q}^n$*, and for any* $(\phi^{(n)}, \psi^{(n)})$ *satisfying* $\frac{1}{n}\log|\mathcal{M}_n| \ge R$ *and* (3)*, we have*

$$\mathrm{P}_c^{(n)}(\phi^{(n)}, \psi^{(n)}) \le 4\exp\left\{-n\frac{\theta R - \frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n)}{1+(1+\alpha+\beta)\theta}\right\}, \quad (38)$$

where

$$\begin{aligned}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n) \triangleq \log \mathrm{E}_{p^{(n)}}\Bigg[&\left\{\prod_{t=1}^{n}\frac{W_1^{\alpha\theta}(Y_t|X_t)}{q^{\alpha\theta}_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t)}\right\}\\ &\times\left\{\prod_{t=1}^{n}\frac{W_2^{\beta\theta}(Z_t|X_t)}{q^{\beta\theta}_{Z_t|X_tY_tU_t}(Z_t|X_tY_tU_t)}\right\}\\ &\times\left\{\prod_{t=1}^{n}\frac{q^{\theta}_{Y_t|U_t}(Y_t|U_t)q^{\theta}_{Z_t|V_t}(Z_t|V_t)}{q^{\theta}_{Z_t|U_t}(Z_t|U_t)q^{\theta}_{Y_t|V_t}(Y_t|V_t)}\right\}\\ &\times\left\{\prod_{t=1}^{n}\frac{p^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}\right\}\Bigg]. \end{aligned} \quad (39)$$

Note that for $t = 1, 2, \cdots, n$*, the conditional pmfs:* $q_{Y_t|X_tZ_tU_t}, q_{Z_t|X_tY_tU_t}, q_{Y_t|V_t}, q_{Z_t|V_t}, q_{Y_t|U_t}, q_{Z_t|U_t}, q_{Y_t|\hat{U}_t}$*, are induced by the joint pmf* $q_t = q_{U_tX_tY_tZ_t} \in \mathcal{Q}_t$ *in* (33)*.* $p^{(n)} = p_{M_nX^nY^nZ^n}$*.*

Proof: We define the random variables B_1, B_2 and B_3 as

$$B_1 \triangleq \frac{1}{n}\sum_{t=1}^{n}\log\frac{W_1(Y_t|X_t)}{q_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t)}, \quad (40)$$

$$B_1 \triangleq \frac{1}{n}\sum_{t=1}^{n}\log\frac{W_2(Z_t|X_t)}{q_{Z_t|X_tY_tU_t}(Z_t|X_t,Y_t,U_t)}, \quad (41)$$

$$B_3 \triangleq \frac{1}{n}\sum_{t=1}^{n}\log\frac{q_{Y_t|U_t}(Y_t|U_t)q_{Z_t|V_t}(Z_t|V_t)}{q_{Z_t|U_t}(Z_t|U_t)q_{Y_t|V_t}(Y_t|V_t)} + \frac{1}{n}\sum_{t=1}^{n}\log\frac{p_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)} - R. \quad (42)$$

Then, by (36) in Lemma 4, we have

$$\mathrm{P}_c^{(n)}(\phi^{(n)}, \psi^{(n)}) \le p_{M_nX^nY^nZ^n}\{B_1 \ge -\eta, B_2 \ge -\eta, B_3 \ge -\eta\} + 3e^{-n\eta} \quad (43)$$

$$\le p_{M_nX^nY^nZ^n}\{\alpha B_1 + \beta B_2 + B_3 \ge -(\alpha+\beta+1)\eta\} + 3e^{-n\eta} \quad (44)$$

$$= p_{M_nX^nY^nZ^n}\{n(\alpha B_1 + \beta B_2 + B_3) \ge -n(\alpha+\beta+1)\eta\} + 3e^{-n\eta}. \quad (45)$$

By identifying $A = n(\alpha B_1 + \beta B_2 + B_3)$, $a = -n(\alpha+\beta+1)\eta$ and applying (37) in Lemma 5, we have

$$\mathrm{P}_c^{(n)}(\phi^{(n)}, \psi^{(n)}) \le \exp\{-[\theta(-n(\alpha+\beta+1)\eta) - \log \mathrm{E}_{p^{(n)}}[\exp(n\theta(\alpha B_1 + \beta B_2 + B_3))]]\} + 3e^{-n\eta} \quad (46)$$

$$= \exp\Big\{n\big[\theta(\alpha+\beta+1)\eta - \theta R + \frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n)\big]\Big\} + 3e^{-n\eta} \quad (47)$$

$$= 4\exp\{-n\eta\}, \quad (48)$$

where (48) holds by choosing

$$-\eta = \theta(\alpha+\beta+1)\eta - \theta R + \frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n). \quad (49)$$

Therefore,

$$\eta = \frac{\theta R - \frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n)}{1+(1+\alpha+\beta)\theta}. \quad (50)$$

By putting η in (50) into (48), we get (38), which completes the proof. ■

Let $\mathcal{P}^{(n)}(W)$ be a set of all pmfs $p_{M_nX^nY^nZ^n}$ on $\mathcal{M}_n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ with the form:

$$p_{M_nX^nY^nZ^n}(m, x^n, y^n, z^n) = p_{M_n}(m)p_{X^n|M_n}(x^n|m)\prod_{t=1}^{n}W(y_t, z_t|x_t). \quad (51)$$

Denoting $p_{M_nX^nY^nZ^n} \in \mathcal{P}^{(n)}(W)$ by $p^{(n)}$, we assume that $p_{U_tX_tY_tZ_t} = p_{M_nX_tY^tZ_t^n}$ is a marginal pmf induced by $p^{(n)}$. Moreover, we denote $p_t = p_{U_tX_tY_tZ_t}$. The main difference between $p^{(n)} \in \mathcal{P}^{(n)}(W)$ and $q^n \in \mathcal{Q}^n$ is that p_t, $t = 1, 2, \cdots, n$, are consistent with $p^{(n)}$ while q_t may not be consistent.

In order to bound the exponent function in (38), we define the communication potential $\overline{\Omega}^{(\alpha,\beta,\theta)}(W)$ as follows:

$$\overline{\Omega}^{(\alpha,\beta,\theta)}(W) \triangleq \sup_{n\ge1}\max_{p^{(n)}\in\mathcal{P}^{(n)}(W)}\min_{q^n\in\mathcal{Q}^n}\frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n). \quad (52)$$

By the definition of communication potential in (52) and (38) in Lemma 6, we have the following corollary.

Corollary 1. *For $R > 0$ and $\theta > 0$, we have*

$$G(R|W) \geq \frac{\theta R - \overline{\Omega}^{(\alpha,\beta,\theta)}(W)}{1 + (1+\alpha+\beta)\theta}. \tag{53}$$

Now, we derive an upper bound for $\overline{\Omega}^{(\alpha,\beta,\theta)}(W)$.

To simplify the notation, define a function of (u_t, x_t, y_t, z_t) on $\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ by

$$\begin{aligned} &f^{(\alpha,\beta,\theta)}_{p_t||q_t}(x_t, y_t, z_t|u_t) \\ &\triangleq \frac{W_1^{\alpha\theta}(y_t|x_t)}{q^{\alpha\theta}_{Y_t|X_tZ_tU_t}(y_t|x_t,z_t,u_t)} \frac{W_2^{\beta\theta}(z_t|x_t)}{q^{\beta\theta}_{Z_t|X_tY_tU_t}(z_t|x_t,y_t,u_t)} \\ &\quad \times \frac{q^{\theta}_{Y_t|U_t}(y_t|u_t) q^{\theta}_{Z_t|V_t}(z_t|v_t)}{q^{\theta}_{Z_t|U_t}(z_t|u_t) q^{\theta}_{Y_t|V_t}(y_t|v_t)} \frac{p^{\theta}_{Y_t|\hat{U}_t}(y_t|\hat{u}_t)}{q^{\theta}_{Y_t|\hat{U}_t}(y_t|\hat{u}_t)}. \end{aligned} \tag{54}$$

Then, by definition in (39), we have

$$\exp\{\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n)\} = \mathrm{E}_{p^{(n)}}\left[\prod_{t=1}^n f^{(\alpha,\beta,\theta)}_{p_t||q_t}\right] \tag{55}$$

$$\begin{aligned} &= \sum_{m,x^n,y^n,z^n} p_{M_nZ^n}(m,z^n) p_{X^nY^n|M_nZ^n}(x^n,y^n|m,z^n) \\ &\quad \times \prod_{t=1}^n f^{(\alpha,\beta,\theta)}_{p_t||q_t}(x_t,y_t,z_t|u_t). \end{aligned} \tag{56}$$

In the following, we are going to choose q_t in Lemma 4. We first define two auxiliary distributions $p^{(\alpha,\beta,\theta;q^t)}_{X^tY^t|M_nZ^n}$ and $p^{(\alpha,\beta,\theta;q^t)}_{M_nZ^n}$. For each $t = 1, 2, \cdots, n$, we define a conditional pmf of (X^t, Y^t) given (M_n, Z^n) by

$$\begin{aligned} &p^{(\alpha,\beta,\theta;q^t)}_{X^tY^t|M_nZ^n} \\ &\triangleq \left\{p^{(\alpha,\beta,\theta;q^t)}_{X^tY^t|M_nZ^n}(x^t,y^t|m,z^n)\right\}_{(x^t,y^t,m,z^n)\in\mathcal{X}^t\times\mathcal{Y}^t\times\mathcal{M}_n\times\mathcal{Z}^n}, \\ &p^{(\alpha,\beta,\theta;q^t)}_{X^tY^t|M_nZ^n}(x^t,y^t|m,z^n) \\ &\triangleq C_t^{-1}(m,z^n) p_{X^tY^t|M_nZ^n}(x^t,y^t|m,z^n) \\ &\quad \times \prod_{i=1}^t f^{(\alpha,\beta,\theta)}_{p_i||q_i}(x_i,y_i,z_i|u_i), \end{aligned} \tag{57}$$

where

$$\begin{aligned} C_t(m,z^n) &\triangleq \sum_{x^t,y^t} p_{X^tY^t|M_nZ^n}(x^t,y^t|m,z^n) \\ &\quad \times \prod_{i=1}^t f^{(\alpha,\beta,\theta)}_{p_i||q_i}(x_i,y_i,z_i|u_i) \end{aligned} \tag{58}$$

are constants for normalization. From (56) and (58), we have

$$\begin{aligned} &\exp\{\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n)\} \\ &\quad = \sum_{m,z^n} p_{M_nZ^n}(m,z^n) C_n(m,z^n) && (59) \\ &\quad = \sum_{m,z^n} p_{M_nZ^n}(m,z^n) \prod_{t=1}^n \frac{C_t(m,z^n)}{C_{t-1}(m,z^n)} && (60) \\ &\quad = \sum_{m,z^n} p_{M_nZ^n}(m,z^n) \prod_{t=1}^n \Phi^{(\alpha,\beta,\theta)}_{t,q^t}(m,z^n) && (61) \end{aligned}$$

where $C_0(m, z^n) = 1$ in (60). By defining

$$\Phi^{(\alpha,\beta,\theta)}_{t,q^t}(m,z^n) \triangleq \frac{C_t(m,z^n)}{C_{t-1}(m,z^n)}, \tag{62}$$

we have (61).

Next, we define a pmf of (M_n, Z^n) on $\mathcal{M} \times \mathcal{Z}^n$ by

$$p^{(\alpha,\beta,\theta;q^t)}_{M_nZ^n}(m,z^n) = \tilde{C}_t^{-1} p_{M_nZ^n}(m,z^n) \prod_{i=1}^t \Phi^{(\alpha,\beta,\theta)}_{i,q^i}(m,z^n). \tag{63}$$

$$\tilde{C}_t = \sum_{m,z^n} p_{M_nZ^n}(m,z^n) \prod_{i=1}^t \Phi^{(\alpha,\beta,\theta)}_{i,q^i}(m,z^n) \tag{64}$$

are constants for normalization. By (61) and (64), we have

$$\begin{aligned} &\exp\{\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n)\} \\ &\quad = \tilde{C}_n = \prod_{t=1}^n \frac{\tilde{C}_t}{\tilde{C}_{t-1}} = \prod_{t=1}^n \Lambda^{(\alpha,\beta,\theta)}_{t,q^t}, \end{aligned} \tag{65}$$

where in (65), $\tilde{C}_0 = 1$ and we define $\Lambda^{(\alpha,\beta,\theta)}_{t,q^t} \triangleq \frac{\tilde{C}_t}{\tilde{C}_{t-1}}$. The following two lemmas help us gain more insights on $\Phi^{(\alpha,\beta,\theta)}_{t,q^t}$ and $\Lambda^{(\alpha,\beta,\theta)}_{t,q^t}$.

Lemma 7. *For each $t = 1, 2, \cdots, n$, and for any $(m, z^n, x^t, y^t) \in \mathcal{M}_n \times \mathcal{Z}^n \times \mathcal{X}^t \times \mathcal{Y}^t$, we have*

$$\begin{aligned} &p^{(\alpha,\beta,\theta;q^t)}_{X^tY^t|M_nZ^n}(x^t,y^t|m,z^n) = (\Phi^{(\alpha,\beta,\theta)}_{t,q^t}(m,z^n))^{-1} \\ &\quad \times p^{(\alpha,\beta,\theta;q^{t-1})}_{X^{t-1}Y^{t-1}|M_nZ^n}(x^{t-1},y^{t-1}|m,z^n) \\ &\quad \times p_{X_tY_t|X^{t-1}Y^{t-1}M_nZ^n}(x_t,y_t|x^{t-1},y^{t-1},m,z^n) \\ &\quad \times f^{(\alpha,\beta,\theta)}_{p_t||q_t}(x_t,y_t,z_t|u_t). \end{aligned} \tag{66}$$

$$\begin{aligned} &\Phi^{(\alpha,\beta,\theta)}_{t,q^t}(m,z^n) \\ &= \sum_{x^t,y^t} p^{(\alpha,\beta,\theta;q^{t-1})}_{X^{t-1}Y^{t-1}|M_nZ^n}(x^{t-1},y^{t-1}|m,z^n) \\ &\quad \times p_{X_tY_t|X^{t-1}Y^{t-1}M_nZ^n}(x_t,y_t|x^{t-1},y^{t-1},m,z^n) \\ &\quad \times f^{(\alpha,\beta,\theta)}_{p_t||q_t}(x_t,y_t,z_t|u_t). \end{aligned} \tag{67}$$

The proof of Lemma 7 is similar to [14, Lemma 4].

Lemma 8.

$$\begin{aligned} \Lambda^{(\alpha,\beta,\theta)}_{t,q^t} &= \sum_{m,z^n} p^{(\alpha,\beta,\theta;q^{t-1})}_{M_n,Z^n}(m,z^n) \Phi^{(\alpha,\beta,\theta)}_{t,q^t}(m,z^n) && (68) \\ &= \sum_{m,z^n} \sum_{x^t,y^t} p^{(\alpha,\beta,\theta;q^{t-1})}_{M_nZ^n}(m,z^n) \\ &\quad \times p^{(\alpha,\beta,\theta;q^{t-1})}_{X^{t-1}Y^{t-1}|M_nZ^n}(x^{t-1},y^{t-1}|m,z^n) \\ &\quad \times p_{X_tY_t|X^{t-1}Y^{t-1}M_nZ^n}(x_t,y_t|x^{t-1},y^{t-1},m,z^n) \\ &\quad \times f^{(\alpha,\beta,\theta)}_{p_t||q_t}(x_t,y_t,z_t|u_t). && (69) \end{aligned}$$

The proof of Lemma 8 is similar to [14, Lemma 5].

We introduce the definitions for the following lemma.

$\Omega^{(\alpha,\beta,\lambda)}_q(XYZ|U)$

$$\triangleq \log \mathrm{E}_q\left[\left\{\frac{W_1^{\alpha}(Y|X)}{q^{\alpha}_{Y|XZU}(Y|X,Z,U)}\frac{W_2^{\beta}(Z|X)}{q^{\beta}_{Z|XYU}(Z|X,Y,U)} \times \frac{q_{Y|U}(Y|U)q_{Z|V}(Z|V)}{q_{Z|U}(Z|U)q_{Y|V}(Y|V)}\right\}^{\lambda}\right] \tag{70}$$

$$\Omega^{(\alpha,\beta,\lambda)}(W) \triangleq \max_{q\in\mathcal{Q}} \Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U) \tag{71}$$

$$\mathcal{Q} \triangleq \{q_{UXYZ} : |\mathcal{U}| \le |\mathcal{Y}| + |\mathcal{Z}| - 1\} \tag{72}$$

The following lemma upper bounds the communication potential, $\overline{\Omega}^{(\alpha,\beta,\theta)}(W)$.

Lemma 9. *For $\theta \in (0,1)$, set $\lambda = \frac{\theta}{1-\theta}$. We have*

$$\overline{\Omega}^{(\alpha,\beta,\theta)}(W) \le \frac{1}{1+\lambda}\Omega^{(\alpha,\beta,\lambda)}(W), \tag{73}$$

where $\Omega^{(\alpha,\beta,\lambda)}(W)$ is defined in (71)*.*

Proof:

$$\min_{q^n\in\mathcal{Q}^n} \frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n)$$
$$\le \frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n) = \frac{1}{n}\sum_{t=1}^{n}\log\Lambda^{(\alpha,\beta,\theta)}_{t,q^t} \tag{74}$$

The equality in (74) can be obtained from (65). In the following, we upper bound $\Lambda^{(\alpha,\beta,\theta)}_{t,q^t}$. We first define the following pmf by observing (69) in Lemma 8.

$$p^{(\alpha,\beta,\theta;q^{t-1})}_{M_nX_tY^tZ_t^n}(m,x_t,y^t,z_t^n) = p^{(\alpha,\beta,\theta;q^{t-1})}_{U_tX_tY_tZ_t}(u_t,x_t,y_t,z_t) \tag{75}$$

$$\triangleq \sum_{x^{t-1},z^{t-1}} p^{(\alpha,\beta,\theta;q^{t-1})}_{M_nZ^n}(m,z^n) \times p^{(\alpha,\beta,\theta;q^{t-1})}_{X^{t-1}Y^{t-1}|M_nZ^n}(x^{t-1},y^{t-1}|m,z^n) \times p_{X_tY_t|X^{t-1}Y^{t-1}M_nZ^n}(x_t,y_t|x^{t-1},y^{t-1},m,z^n). \tag{76}$$

Then, by (69) in Lemma 8, we have

$$\Lambda^{(\alpha,\beta,\theta)}_{t,q^t} = \sum_{u_t,x_t,y_t,z_t} p^{(\alpha,\beta,\theta;q^{t-1})}_{U_tX_tY_tZ_t}(u_t,x_t,y_t,z_t) \times f^{(\alpha,\beta,\theta)}_{p_t||q_t}(x_t,y_t,z_t|u_t). \tag{77}$$

For $t = 1,2,\cdots,n$, we choose q_t as

$$q_{U_tX_tY_tZ_t}(u_t,x_t,y_t,z_t) = p^{(\alpha,\beta,\theta;q^{t-1})}_{U_tX_tY_tZ_t}(u_t,x_t,y_t,z_t), \tag{78}$$

and the following conditional pmfs: $q_{Y_t|X_tZ_tU_t}, q_{Z_t|X_tY_tU_t}$, $q_{Y_t|V_t}, q_{Z_t|V_t}, q_{Y_t|U_t}, q_{Z_t|U_t}, q_{Y_t|\hat{U}_t}$, are induced by $q_{U_tX_tY_tZ_t}$. Then, we can bound each $\Lambda^{(\alpha,\beta,\theta)}_{t,q^t}$ as follows.

$$\Lambda^{(\alpha,\beta,\theta)}_{t,q^t} = \mathrm{E}_{q_t}\left[\left\{\frac{W_1^{\alpha\theta}(Y_t|X_t)}{q^{\alpha\theta}_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t)}\frac{W_2^{\beta\theta}(Z_t|X_t)}{q^{\beta\theta}_{Z_t|X_tY_tU_t}(Z_t|X_t,Y_t,U_t)}\right\} \times \left\{\frac{q^{\theta}_{Y_t|U_t}(Y_t|U_t)q^{\theta}_{Z_t|V_t}(Z_t|V_t)}{q^{\theta}_{Z_t|U_t}(Z_t|U_t)q^{\theta}_{Y_t|V_t}(Y_t|V_t)}\right\}\left\{\frac{p^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}\right\}\right] \tag{79}$$

$$\le \left\{\mathrm{E}_{q_t}\left[\left(\frac{W_1^{\alpha\theta}(Y_t|X_t)}{q^{\alpha\theta}_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t)}\frac{W_2^{\beta\theta}(Z_t|X_t)}{q^{\beta\theta}_{Z_t|X_tY_tU_t}(Z_t|X_t,Y_t,U_t)} \times \frac{q^{\theta}_{Y_t|U_t}(Y_t|U_t)q^{\theta}_{Z_t|V_t}(Z_t|V_t)}{q^{\theta}_{Z_t|U_t}(Z_t|U_t)q^{\theta}_{Y_t|V_t}(Y_t|V_t)}\right)^{\frac{1}{1-\theta}}\right]\right\}^{1-\theta} \times \left\{\mathrm{E}_{q_t}\left[\left(\frac{p^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}\right)^{\frac{1}{\theta}}\right]\right\}^{\theta} \tag{80}$$

$$\le \left\{\mathrm{E}_{q_t}\left[\left(\frac{W_1^{\alpha}(Y_t|X_t)}{q^{\alpha}_{Y_t|X_tZ_tU_t}(Y_t|X_t,Z_t,U_t)}\frac{W_2^{\beta}(Z_t|X_t)}{q^{\beta}_{Z_t|X_tY_tU_t}(Z_t|X_t,Y_t,U_t)} \times \frac{q_{Y_t|U_t}(Y_t|U_t)q_{Z_t|V_t}(Z_t|V_t)}{q_{Z_t|U_t}(Z_t|U_t)q_{Y_t|V_t}(Y_t|V_t)}\right)^{\frac{\theta}{1-\theta}}\right]\right\}^{1-\theta} \tag{81}$$

$$= \exp\left\{(1-\theta)\Omega_{q_t}^{(\alpha,\beta,\frac{\theta}{1-\theta})}(X_tY_tZ_t|U_t)\right\} \tag{82}$$

$$= \exp\left\{\frac{1}{1+\lambda}\Omega_{q_t}^{(\alpha,\beta,\lambda)}(X_tY_tZ_t|U_t)\right\} \tag{83}$$

$$\le \exp\left\{\frac{1}{1+\lambda}\hat{\Omega}_n^{(\alpha,\beta,\lambda)}(W)\right\} \tag{84}$$

$$= \exp\left\{\frac{1}{1+\lambda}\Omega^{(\alpha,\beta,\lambda)}(W)\right\} \tag{85}$$

where (80) is due to Holder's inequality. (81) is due to the following equalities:

$$\mathrm{E}_{q_t}\left[\left(\frac{p^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q^{\theta}_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}\right)^{\frac{1}{\theta}}\right] = \mathrm{E}_{q_t}\left[\left(\frac{p_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}{q_{Y_t|\hat{U}_t}(Y_t|\hat{U}_t)}\right)\right] \tag{86}$$

$$= \sum_{y,\hat{u}} q_{Y_t\hat{U}_t}(y,\hat{u}) \times \frac{p_{Y_t|\hat{U}_t}(y|\hat{u})}{q_{Y_t|\hat{U}_t}(y|\hat{u})} \tag{87}$$

$$= \sum_{y,\hat{u}} q_{\hat{U}_t}(\hat{u}) \times p_{Y_t|\hat{U}_t}(y|\hat{u}) \tag{88}$$

$$= \sum_{\hat{u}} q_{\hat{U}_t}(\hat{u}) \times \sum_{y} p_{Y_t|\hat{U}_t}(y|\hat{u}) = 1 \tag{89}$$

and (82) is obtained by applying the definition in (70), and (83) is from λ in Lemma 9. For (84), we set

$$\hat{\Omega}_n^{(\alpha,\beta,\lambda)}(W) \triangleq \max_{q\in\hat{\mathcal{Q}}_n(W)} \Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U), \tag{90}$$

$$\hat{\mathcal{Q}}_n(W) \triangleq \{q = q_{UXYZ} : |\mathcal{U}| \le |\mathcal{M}_n||\mathcal{Y}^{n-1}||\mathcal{Z}^{n-1}|\}. \tag{91}$$

Hence, (84) holds by definition. Finally, by the analysis similar to that in [15, Lemma 10], it can be shown that (85) holds.

Therefore, by (85), we know (74) can be continued as

$$\min_{q^n\in\mathcal{Q}^n} \frac{1}{n}\Omega^{(\alpha,\beta,\theta)}_{p^{(n)}||q^n}(X^nY^nZ^n|M_n) \le \frac{1}{1+\lambda}\Omega^{(\alpha,\beta,\lambda)}(W). \tag{92}$$

Since (92) holds for any $n \ge 1$ and any $p^{(n)}$, we get (73). ■

Corollary 2. *For $\theta \in (0,1)$, set $\lambda = \frac{\theta}{1-\theta}$. Then, we have*

$$G(R|W) \ge \frac{\lambda R - \Omega^{(\alpha,\beta,\lambda)}(W)}{1+(2+\alpha+\beta)\lambda}. \tag{93}$$

Proof:

$$G(R|W) \geq \frac{\theta R - \overline{\Omega}^{(\alpha,\beta,\theta)}(W)}{1+(1+\alpha+\beta)\theta} \tag{94}$$

$$\geq \frac{\frac{\lambda}{1+\lambda}R - \frac{1}{1+\lambda}\Omega^{(\alpha,\beta,\lambda)}(W)}{1+(1+\alpha+\beta)\frac{\lambda}{1+\lambda}} \tag{95}$$

$$= \frac{\lambda R - \Omega^{(\alpha,\beta,\lambda)}(W)}{1+(2+\alpha+\beta)\lambda} \tag{96}$$

where (94) follows from (53) in Corollary 1, and (95) is from (73) in Lemma 9. ■

From (96), we have the following definitions:

$$F^{(\alpha,\beta,\lambda)}(R|W) \triangleq \frac{\lambda R - \Omega^{(\alpha,\beta,\lambda)}(W)}{1+(2+\alpha+\beta)\lambda} \tag{97}$$

$$F(R|W) \triangleq \sup_{\alpha,\beta,\lambda>0} F^{(\alpha,\beta,\lambda)}(R|W) \tag{98}$$

We have the following lemma to lower bound $G(R|W)$.

Lemma 10.

1) $\Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U)$ *is a convex function of* $\lambda > 0$.
2) *For every* $q \in \mathcal{Q}$, *we have*

$$\begin{aligned}\lim_{\lambda\to 0^+} &\frac{\Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U)}{\lambda} \\ &= -\alpha D(q_{Y|XZU}||W_1|q_{XZU}) \\ &\quad -\beta D(q_{Z|XYU}||W_2|q_{XYU}) \\ &\quad + I_q(U;Y|V) - I_q(U;Z|V).\end{aligned} \tag{99}$$

3) *If* $R > \mathcal{C}_s(0|W)$, *then we have* $F(R|W) > 0$.

In Corollary 2, we already show $G(R|W) > F(R|W)$. Moreover, when $R > \mathcal{C}_s(0|W)$, we have $F(R|W) > 0$. We know the probability of correct decoding will decay exponentially when $R > \mathcal{C}_s(0|W)$. Therefore, the partial strong converse property holds.

IV. APPENDIX

A. Proof of Lemma 1

Define $\mathcal{A}_1(m)$, $\mathcal{A}_2(m)$ and $\mathcal{A}_3(m)$ as below:

$$\mathcal{A}_1(m) \triangleq \left\{(x^n,y^n,z^n): \frac{W_1^n(y^n|x^n)}{q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}(y^n|x^n,z^n,m)} \geq e^{-n\eta}\right\}, \tag{100}$$

$$\mathcal{A}_2(m) \triangleq \left\{(x^n,y^n,z^n): \frac{W_2^n(z^n|x^n)}{q^{(\mathrm{ii})}_{Z^n|X^nY^nM_n}(z^n|x^n,y^n,m)} \geq e^{-n\eta}\right\}, \tag{101}$$

$$\mathcal{A}_3(m) \triangleq \left\{(x^n,y^n,z^n): p_{Y^n|M_n}(y^n|m) \geq |\mathcal{M}_n|e^{-n\eta}q^{(\mathrm{iii})}_{Y^n}(y^n)\right\}. \tag{102}$$

Let $\mathcal{A}(m) \triangleq \mathcal{A}_1(m)\cap\mathcal{A}_2(m)\cap\mathcal{A}_3(m)$. Moreover, set $\mathcal{D}(m)$ to be the decoding region of message m for the code $(\phi^{(n)},\psi^{(n)})$, and denote $\underline{a} = (m,x^n,y^n,z^n)$.

For the probability of correct decoding, we have

$$\begin{aligned}&\mathrm{P}_c^{(n)}(\phi^{(n)},\psi^{(n)}) \\ &= \sum_{m\in\mathcal{M}_n}\sum_{\substack{(x^n,y^n,z^n)\in\mathcal{A}(m),\\ y^n\in\mathcal{D}(m)}} p_{M_nX^nY^nZ^n}(\underline{a}) \\ &\quad + \sum_{m\in\mathcal{M}_n}\sum_{\substack{(x^n,y^n,z^n)\in\mathcal{A}^c(m),\\ y^n\in\mathcal{D}(m)}} p_{M_nX^nY^nZ^n}(\underline{a})\end{aligned} \tag{103}$$

$$\leq \Delta_0 + \Delta_1 + \Delta_2 + \Delta_3, \tag{104}$$

where

$$\Delta_0 \triangleq \sum_{m\in\mathcal{M}_n}\sum_{(x^n,y^n,z^n)\in\mathcal{A}(m)} p_{M_nX^nY^nZ^n}(\underline{a}), \tag{105}$$

$$\Delta_1 \triangleq \sum_{m\in\mathcal{M}_n}\sum_{(x^n,y^n,z^n)\in\mathcal{A}_1^c(m)} p_{M_nX^nY^nZ^n}(\underline{a}), \tag{106}$$

$$\Delta_2 \triangleq \sum_{m\in\mathcal{M}_n}\sum_{(x^n,y^n,z^n)\in\mathcal{A}_2^c(m)} p_{M_nX^nY^nZ^n}(\underline{a}) \tag{107}$$

$$\Delta_3 \triangleq \sum_{m\in\mathcal{M}_n}\sum_{\substack{(x^n,y^n,z^n)\in\mathcal{A}_3^c(m)\\ y^n\in\mathcal{D}(m)}} p_{M_nX^nY^nZ^n}(\underline{a}). \tag{108}$$

By definition, Δ_0 corresponds to the first term of (12). Now, we show $\Delta_i \leq e^{-n\eta}$, $i=1,2,3$. For Δ_1, we have

$$\Delta_1 = \sum_{m\in\mathcal{M}_n}\sum_{(x^n,y^n,z^n)\in\mathcal{A}_1^c(m)} p_{M_nX^nY^nZ^n}(\underline{a}) \tag{109}$$

$$\leq e^{-n\eta}\sum_{m\in\mathcal{M}_n}\sum_{(x^n,y^n,z^n)\in\mathcal{A}_1^c(m)} p_{M_nX^nZ^n}(m,x^n,z^n)q^{(\mathrm{i})}_{Y^n|X^nZ^nM_n}(y^n|x^n,z^n,m) \tag{110}$$

$$\leq e^{-n\eta}. \tag{111}$$

$\Delta_2 \leq e^{-n\eta}$, can be shown similarly.

For Δ_3, we have

$$\Delta_3 = \sum_{m\in\mathcal{M}_n}\sum_{\substack{(x^n,y^n,z^n)\in\mathcal{A}_3^c(m)\\ y^n\in\mathcal{D}(m)}} p_{M_nX^nY^nZ^n}(\underline{a}) \tag{112}$$

$$= \frac{1}{|\mathcal{M}_n|}\sum_{m\in\mathcal{M}_n}\sum_{\substack{y^n\in\mathcal{D}(m)\\ p_{Y^n|M_n}(y^n|m)<|\mathcal{M}_n|e^{-n\eta}q^{(\mathrm{iii})}_{Y^n}}} p_{Y^n|M_n}(y^n|m) \tag{113}$$

$$\leq e^{-n\eta}\sum_{m\in\mathcal{M}_n}\sum_{y^n\in\mathcal{D}(m)} q^{(\mathrm{iii})}_{Y^n}(y^n) \leq e^{-n\eta}. \tag{114}$$

B. Proof of Lemma 10

To simplify the notation, define

$$\underline{a} \triangleq (u,x,y,z), \underline{A} \triangleq (U,X,Y,Z), \tag{115}$$

$$\underline{\mathcal{A}} \triangleq \mathcal{U}\times\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}, \tag{116}$$

$$\rho(\underline{a}) \triangleq \alpha\log\frac{W_1(y|x)}{q_{Y|XZU}(y|x,z,u)} + \beta\log\frac{W_2(z|x)}{q_{Z|XYU}(z|x,y,u)}$$

$$+\log\frac{q_{Y|U}(y|u)q_{Z|V}(z|v)}{q_{Z|U}(z|u)q_{Y|V}(y|v)}, \tag{117}$$

$$\xi(\lambda)\triangleq\Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U). \tag{118}$$

From (70), we have

$$\Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U)=\xi(\lambda)=\log\left[\sum_{\underline{a}\in\underline{\mathcal{A}}}q_{\underline{A}}(\underline{a})e^{\lambda\rho(\underline{a})}\right]. \tag{119}$$

To show the convexity in λ, we calculate the second derivative.

$$\xi'(\lambda)=e^{-\xi(\lambda)}\left[\sum_{\underline{a}\in\underline{\mathcal{A}}}q_{\underline{A}}(\underline{a})\rho(\underline{a})e^{\lambda\rho(\underline{a})}\right], \tag{120}$$

$$\xi''(\lambda)=e^{-2\xi(\lambda)}\times\left[\sum_{\underline{a},\underline{b}\in\underline{\mathcal{A}}}q_{\underline{A}}(\underline{a})q_{\underline{A}}(\underline{b})\frac{\{\rho(\underline{a})-\rho(\underline{b})\}^2}{2}e^{\lambda\{\rho(\underline{a})+\rho(\underline{b})\}}\right]. \tag{121}$$

From (121), since $\xi''(\lambda)\geq 0$, $\Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U)$ is a convex function of λ. Hence, part 1) holds.

For part 2), consider $\lambda=0$ in (120), we have

$$\begin{aligned}\xi'(0)=&-\alpha D(q_{Y|XZU}||W_1|q_{XZU})\\&-\beta D(q_{Z|XYU}||W_2|q_{XYU})\\&+I_q(U;Y|V)-I_q(U;Z|V).\end{aligned} \tag{122}$$

For part 3), let us consider $R>\mathcal{C}_s(0|W)$. It can be shown similar to that in [15, Appendix B] that

$$\mathcal{C}_s(0|W)=\inf_{\alpha,\beta>0}\tilde{\mathcal{C}}^{(\alpha,\beta)}(W), \tag{123}$$

$$\begin{aligned}\tilde{\mathcal{C}}^{(\alpha,\beta)}(W)\triangleq\max_{q\in\mathcal{Q}}\big\{&-\alpha D(q_{Y|XZU}||W_1|q_{XZU})\\&-\beta D(q_{Z|XYU}||W_2|q_{XYU})\\&+I_q(U;Y|V)-I_q(U;Z|V)\big\}.\end{aligned} \tag{124}$$

Then, there exists $\alpha,\beta,\epsilon>0$, such that $R\geq\tilde{\mathcal{C}}^{(\alpha,\beta)}(W)+\epsilon$. Define the following function:

$$\begin{aligned}\zeta(\lambda)\triangleq\xi(\lambda)-\lambda\Big[&-\alpha D(q_{Y|XZU}||W_1|q_{XZU})\\&-\beta D(q_{Z|XYU}||W_2|q_{XYU})\\&+I_q(U;Y|V)-I_q(U;Z|V)+\frac{\epsilon}{2}\Big].\end{aligned} \tag{125}$$

$\zeta(\lambda)$ has the following properties:

$$\zeta(0)=0,\ \zeta'(0)=-\frac{\epsilon}{2}<0,\ \zeta''(\lambda)=\xi''(\lambda)\geq 0. \tag{126}$$

From (126), there exists $f(\epsilon)>0$ such that $\zeta(\lambda)\leq 0$ for $\lambda\in(0,f(\epsilon)]$. Now, we focus on $\lambda\in(0,f(\epsilon)]$ and for every $q\in\mathcal{Q}$, we have

$$\begin{aligned}\Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U)\leq\lambda\Big[&-\alpha D(q_{Y|XZU}||W_1|q_{XZU})\\&-\beta D(q_{Z|XYU}||W_2|q_{XYU})\\&+I_q(U;Y|V)-I_q(U;Z|V)+\frac{\epsilon}{2}\Big].\end{aligned} \tag{127}$$

Therefore, for any $\lambda\in(0,f(\epsilon)]$, we have

$$\begin{aligned}\Omega^{(\alpha,\beta,\lambda)}(W)&=\max_{q\in\mathcal{Q}}\Omega_q^{(\alpha,\beta,\lambda)}(XYZ|U)\\&\leq\lambda\max_{q\in\mathcal{Q}}\Big[-\alpha D(q_{Y|XZU}||W_1|q_{XZU})\\&\qquad-\beta D(q_{Z|XYU}||W_2|q_{XYU})\\&\qquad+I_q(U;Y|V)-I_q(U;Z|V)+\frac{\epsilon}{2}\Big]. \qquad (128)\\&=\lambda\left(\tilde{\mathcal{C}}^{(\alpha,\beta)}(W)+\frac{\epsilon}{2}\right). \qquad (129)\end{aligned}$$

Now, we can start bounding $F(R|W)$ as follows:

$$\begin{aligned}&F(R|W)\\&\geq\sup_{\lambda\in(0,f(\epsilon)]}F^{(\alpha,\beta,\lambda)}(R|W)=\sup_{\lambda\in(0,f(\epsilon)]}\frac{\lambda R-\Omega^{(\alpha,\beta,\lambda)}(W)}{1+(2+\alpha+\beta)\lambda} \qquad (130)\\&\geq\sup_{\lambda\in(0,f(\epsilon)]}\lambda\frac{R-(\tilde{\mathcal{C}}^{(\alpha,\beta)}(W)+\frac{\epsilon}{2})}{1+(2+\alpha+\beta)\lambda} \qquad (131)\\&\geq\sup_{\lambda\in(0,f(\epsilon)]}\frac{\lambda\epsilon}{2+(4+2\alpha+2\beta)\lambda}>0. \qquad (132)\end{aligned}$$

The inequality in (130) is from the definition of $F(R|W)$ in (98), while the inequality in (130) is from the definition of $F^{(\alpha,\beta,\lambda)}(R|W)$ in (97). Since $\lambda\in(0,f(\epsilon)]$, from (129) we have (131). Since we consider the case $R\geq\tilde{\mathcal{C}}^{(\alpha,\beta)}(W)+\epsilon$, (132) holds. From (132), part 3) holds.

REFERENCES

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

[2] I. Csiszar and J. Korner. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory*, 24(3):339–348, May 1978.

[3] M. R. Bloch and J. N. Laneman. Strong secrecy from channel resolvability. *IEEE Trans. Inf. Theory*, 59(12):8077–8098, Dec. 2013.

[4] J. Wolfowitz. The coding of messages subject to chance errors. *Illinois Journal of Mathematics*, 1(4):591–606, 1957.

[5] M. Hayashi. Information spectrum approach to second-order coding rate in channel coding. *IEEE Trans. Inf. Theory*, 55(11):4947–4966, Nov. 2009.

[6] Y. Polyanskiy, H. V. Poor, and S. Verdu. Channel coding rate in the finite blocklength regime. *IEEE Trans. Inf. Theory*, 56(5):2307–2359, May 2010.

[7] V. Y. F. Tan. Asymptotic estimates in information theory with non-vanishing error probabilities. *Foundations and Trends in Communications and Information Theory*, 11(1-2):1–184, Sep. 2014.

[8] S. Verdu and T. S. Han. A general formula for channel capacity. *IEEE Trans. Inf. Theory*, 40(4):1147–1157, Jul. 1994.

[9] T. S. Han. *Information-Spectrum Methods in Information Theory*. Springer-Verlag, Feb. 2003.

[10] V. Y. F. Tan and M. R. Bloch. Information spectrum approach to strong converse theorems for degraded wiretap channels. *IEEE Trans. Inf. Forensics and Security*, 10(9):1891–1904, Sep. 2015.

[11] M. Hayashi, H. Tyagi, and S. Watanabe. Strong converse for a degraded wiretap channel via active hypothesis testing. In *Allerton Conf.*, Sep. 2014.

[12] H. Tyagi and S. Watanabe. Converses for secret key agreement and secure computing. *IEEE Trans. Inf. Theory*, 61(9):4809–4827, Sep. 2015.

[13] Y. Oohama. Strong converse exponent for degraded broadcast channels at rates outside the capacity region. In *IEEE ISIT*, Jun. 2015.

[14] Y. Oohama. New strong converse for asymmetric broadcast channels. *http://arxiv.org/abs/1604.02901v3*, Apr. 2016.

[15] Y. Oohama. Strong converse exponent for state dependent channels with full state information at the sender and partial state information at the receiver. *https://arxiv.org/abs/1603.06344*, Mar. 2016.

[16] Y. Oohama. Exponent function for source coding with side information at the decoder at rates below the rate distortion function. *http://arxiv.org/abs/1601.05650*, Jan. 2016.