An Adaptive Composition Theorem for Maximal Leakage for Binary Inputs

Ibrahim Issa
Electrical and Computer Engineering Department
American University of Beirut, Beirut, Lebanon
ibrahim.issa@aub.edu.lb

Aaron B. Wagner
School of Electrical and Computer Engineering
Cornell University, Ithaca, New York 14850, USA
wagner@cornell.edu

2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834392

***Abstract*—Given a binary random variable X representing sensitive information and n noisy observations $Y_1, Y_2, \ldots, Y_n$ available to an adversary, we analyze the maximal leakage $\mathscr{L}(X \to Y^n)$ in the following setting modeling adaptive attacks. At each stage i, the adversary may choose an action to interact with the system containing X to obtain Y_i. The action may depend on previous realizations of the observations, but the leakage at each stage is limited. We derive an adaptive composition theorem wherein $\mathscr{L}(X \to Y^n)$ is bounded in terms of the leakage of each stage. Furthermore, we show that the bound is achieved for $\mathscr{L}(X \to Z^n)$ where $(Z_1, Z_2, \ldots, Z_n)$ are conditionally independent given X and each Z_i corresponds to the output of a binary erasure channel with the appropriate parameter; moreover, $X - Z^n - Y^n$ can be coupled as a Markov chain for any feasible Y^n. As a corollary of this result and the asymptotic analysis of composition by Wu *et al.*, we show that the binary erasure channel maximizes the Chernoff information between the "rows" of binary-input channels given a maximal leakage constraint. On the other hand, we show that the binary symmetric channel minimizes the Chernoff information for a given maximal leakage constraint.**

***Index Terms*—composition, adaptive, maximal leakage, Chernoff information**

I. Introduction

Consider a system containing sensitive data X and an adversary interacting with the system to receive a noisy observation, Y. In many scenarios, an adversary may generate several "attacks" (e.g., averaging attacks in side channels, multiple queries to a database, etc.), receiving a sequence of observations $Y_1, Y_2, \ldots, Y_n$. Thus even if the information leakage from X to Y, denoted by $L(X \to Y)$, is limited, it is essential to analyze the degradation of privacy/security guarantees under multiple observations, $L(X \to Y^n)$. In the privacy and security literature, results of this form (bounding $L(X \to Y^n)$ in terms of $L(X \to Y)$) are termed *composition theorems*. A smart adversary may adapt their attacks to previous observations, in which case we speak of *adaptive composition theorems*, which are the focus of this paper.

Tight adaptive composition theorems have been derived for (approximate) differential privacy in [1], where an equivalent characterization of differential privacy in terms of a binary hypothesis testing problem was used to demonstrate the existence of a "dominating" mechanism. Composition theorems for variants of differential privacy have also been studied [2]–[4]. Wu *et al.* [5] derived (non-adaptive) asymptotic composition theorems, assuming $Y_1, Y_2, \ldots, Y_n$ are conditionally i.i.d given X, for several information-theoretic measures, namely Sibson mutual information [6], [7], Arimoto mutual information [8], and α-maximal leakage [9]. In particular, they showed that $L(X \to Y^n)$ (where L is any of the mentioned measures) converges exponentially fast to its corresponding limit. Moreover, the rate of convergence for all these measures is the same, namely it is the Chernoff information among all pairs of distinct distributions $P_{Y|X}(.|x)$ and $P_{Y|X}(.|x')$.

In this paper, we consider maximal leakage (also studied in [5] as it is equal to Sibson mutual information of order ∞), which is a security metric that emerged in the computer security [10]–[12] and information theory literature [13], [14]. It is given by

$$\mathscr{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} P_{Y|X}(y|x), \tag{1}$$

where we assume the alphabets are finite and X has full support (throughout the paper). We consider an adaptive setting where at each stage i, an attacker chooses an action to interact with the system to obtain Y_i. The action may depend on previous realizations of the observations, but the leakage at each stage is limited. We derive a tight finite-n bound on $\mathscr{L}(X \to Y^n)$ in terms of the leakages at each stage, under the assumption that X is binary. We show that the bound is in fact achievable for a "non-adaptive" mechanism, in particular, it is achievable for $P_{Z^n|X}$ where $(Z_1, \ldots, Z_n)$ are conditionally independent given X and each Z_i corresponds to the output of a binary erasure channel (BEC). To that end, for any $P_{Y|X}$ we find the largest $\alpha \in [0, 1]$ for which it is possible to couple $X - Z - Y$ as a Markov chain where Z is the output of a BEC with parameter α. This is then extended into coupling $X - Z^n - Y^n$ as a Markov chain with the appropriate parameters for any feasible $P_{Y^n|X}$.

As a consequence of this result and the asymptotic result of Wu *et al.* [5], it turns out that the BEC maximizes the Chernoff information between the "rows" of binary-input channels given a maximal leakage constraint. On the other hand, we show that the binary symmetric channel minimizes the Chernoff information for a given maximal leakage constraint.

II. Main Result

Notation: Given a joint distribution P_{XY} on alphabets $\mathcal{X} \times \mathcal{Y}$, we use $\mathscr{L}(P_{Y|X})$ to denote $\mathscr{L}(X \to Y)$ when X has full support. The M-ary input erasure channel ($M \in \mathbb{N}$) with parameter $\alpha \in [0,1]$ is denoted by M-EC(α), i.e., $P_{Z|X} = M\text{-EC}(\alpha)$ indicates $|\mathcal{X}| = M$, $Z \in \mathcal{Z} = \mathcal{X} \cup \{e\}$, and $P_{Z|X}(e|x) = \alpha$ and $P_{Z|X}(x|x) = 1-\alpha$ for all $x \in \mathcal{X}$. For the special case where $M = 2$, we denote it by BEC(α).

Throughout this paper, we assume all alphabets are finite.

Definition 1 ($\overrightarrow{\ell}$-Leakage Adaptive Mechanism)**.** *Given $n \in \mathbb{N}$, $\overrightarrow{\ell} = (\ell_1, \ell_2, \ldots, \ell_n) \in \mathbb{R}_+^n$, alphabets $\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_n$, and a conditional distribution $P_{Y^n|X}$ where $X \in \mathcal{X}$ and $Y^n \in \mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n$, we say $P_{Y^n|X}$ is an $\overrightarrow{\ell}$-*leakage adaptive mechanism *($\overrightarrow{\ell}$-LAM) if for all $i \in \{1, 2, \ldots, n\}$, and all $(y_1, y_2, \ldots, y_{i-1}) \in \mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_{i-1}$, we have*

$$\mathscr{L}\left(P_{Y_i|X, Y_1=y_1, Y_2=y_2, \ldots, Y_{i-1}=y_{i-1}}\right) \le \ell_i. \tag{2}$$

Wu *et al.* [5] studied the asymptotic growth of $\mathscr{L}\left(P_{Y^n|X}\right)$ when $(Y_1, Y_2, \ldots, Y_n)$ are conditionally i.i.d given X, i.e., $P_{Y^n|X} = \prod_{i=1}^n P_{Y_i|X}$ and for all i, $P_{Y_i|X} = P_{Y|X}$ for some fixed channel $P_{Y|X}$. Such (non-adaptive) mechanisms can be seen as a subset of ℓ-LAMs with $\ell = \mathscr{L}\left(P_{Y|X}\right)$.

Our main result tightly upper-bounds $\mathscr{L}\left(P_{Y^n|X}\right)$ for any n and $\overrightarrow{\ell}$, assuming X is binary.

Theorem 1. *Suppose $\mathcal{X} = \{0,1\}$ and let $X \sim P_X$ have full support. Consider $n \in \mathbb{N}$, $\overrightarrow{\ell} \in [0,1]^n$, alphabets $\mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_n$, and $\overrightarrow{\ell}$-LAM $P_{Y^n|X}$. Let $Z_1, Z_2, \ldots, Z_n$ be conditionally independent given X, with $P_{Z_i|X} = \text{BEC}(2 - e^{\ell_i})$. Then, there exists a coupling of (X, Y^n, Z^n) such that $X - Z^n - Y^n$ is a Markov chain (i.e., the channel $P_{Y^n|X}$ is degraded with respect to $P_{Z^n|X}$). Consequently,*

$$\mathscr{L}(X \to Y_1, Y_2, \ldots, Y_n) \le \mathscr{L}(X \to Z_1, Z_2, \ldots, Z_n) \tag{3}$$

$$= \log\left(2 - \prod_{i=1}^n (2 - e^{\ell_i})\right). \tag{4}$$

Notably, the theorem states that a *non-adaptive* mechanism (namely, conditionally independent outputs of binary erasure channels) achieves the maximum total leakage $\mathscr{L}\left(P_{Y^n|X}\right)$ among $\overrightarrow{\ell}$-LAMs. A similar phenomenon occurs in the context of composition theorems for (approximate) differential privacy [15], as well as composition theorems for f-differential privacy [4], wherein a non-adaptive mechanism maximizes the privacy degradation. The latter work defines privacy in terms of the trade-off function resulting from a binary hypothesis testing problem. The tools developed in these papers [4], [15] could be used to prove the inequality in Theorem 1 (herein, we provide a more explicit proof through the coupling of (X, Y^n, Z^n)). This common phenomenon is due to the existence of a "dominating" mechanism for both differential privacy and f-differential privacy (for every valid choice of f), analogous to the BEC in our context.

A. Proof of Theorem 1

For the case $n = 1$, the theorem follows from the following two lemmas.

Lemma 1. *If $\mathcal{X} = \{0,1\}$, then for any conditional distribution $P_{Y|X}$,*

$$\mathscr{L}\left(P_{Y|X}\right) = \log\left(1 + d_{TV}\left(P_{Y|X=0}, P_{Y|X=1}\right)\right)$$

$$= \log\left(2 - \sum_{y \in \mathcal{Y}} \min_{x \in \{0,1\}} P_{Y|X}(y|x)\right), \tag{5}$$

where d_{TV} is the total variation distance.

Proof of Lemma 1: The first equality is due to Sibson [6]. For the second equality, note that

$$\begin{aligned}
&\sum_{y \in \mathcal{Y}} \min_{x \in \{0,1\}} P_{Y|X}(y|x) \\
&= \frac{1}{2} \sum_{y \in \mathcal{Y}} \min_{x \in \{0,1\}} P_{Y|X}(y|x) + \frac{1}{2} \sum_{y \in \mathcal{Y}} \max_{x \in \{0,1\}} P_{Y|X}(y|x) \\
&\quad + \frac{1}{2} \sum_{y \in \mathcal{Y}} \left(\min_{x \in \{0,1\}} P_{Y|X}(y|x) - \max_{x \in \{0,1\}} P_{Y|X}(y|x) \right) \\
&= \frac{1}{2} \sum_{y \in \mathcal{Y}} \left(P_{Y|X}(y|0) + P_{Y|X}(y|1) \right) \\
&\quad - \frac{1}{2} \sum_{y \in \mathcal{Y}} \left| P_{Y|X}(y|0) - P_{Y|X}(y|1) \right| \\
&= 1 - d_{TV}\left(P_{Y|X=0}, P_{Y|X=1}\right).
\end{aligned}$$

■

Lemma 2. *Fix an arbitrary finite alphabet $\mathcal{X}$ and let $P_{Z|X} = M\text{-EC}(\alpha)$ for some[1] $\alpha \in [0,1)$ where $M = |\mathcal{X}|$. Consider an arbitrary finite alphabet $\mathcal{Y}$ and a conditional distribution $P_{Y|X}$. Then there exists a coupling of (X, Y, Z) such that $X - Z - Y$ is a Markov chain if and only if $\sum_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} P_{Y|X}(y|x) \ge \alpha$. Moreover, the following $P_{Y|Z}$ yields the desired coupling:*

$$P_{Y|Z}(y|e) = \frac{1}{\alpha'} \min_{z \in \mathcal{X}} P_{Y|X}(y|z), \tag{6}$$

$$P_{Y|Z}(y|z) = \frac{1}{1-\alpha} \left(P_{Y|X}(y|z) - \frac{\alpha}{\alpha'} \min_{z' \in \mathcal{X}} P_{Y|X}(y|z') \right), \tag{7}$$

for $z \in \mathcal{X}$, where $\alpha' := \sum_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} P_{Y|X}(y|x) \ge \alpha$.

Remark 1. *For the purposes of the proof Theorem 1, only the ($\Leftarrow$) implication of the lemma for $M = 2$ is needed. However, the equivalence is simple to prove in general and may be of independent interest.*

[1]For $\alpha = 1$, the statement is trivially true as the output of Z of an M-EC(1) is a constant and $\sum_y \min_x P_{Y|X}(y|x) = 1$ implies that X is independent from Y

Proof of Lemma 2: ($\Rightarrow$) : If $X - Z - Y$ is a Markov chain, then

$$\begin{aligned} P_{Y|X}(y|x) &= \sum_{z\in\mathcal{X}\cup\{e\}} P_{Y|Z}(y|z)P_{Z|X}(z|x) \\ &= P_{Y|Z}(y|x)(1-\alpha) + P_{Y|Z}(y|e)\cdot\alpha \\ &\geq P_{Y|Z}(y|e)\cdot\alpha. \end{aligned}$$

Consequently,

$$\sum_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} P_{Y|X}(y|x) \geq \alpha \sum_{y\in\mathcal{Y}} P_{Y|Z}(y|e) = \alpha.$$

($\Leftarrow$) : Let $Z' = \mathrm{BEC}(\alpha')$. Define $X - Z' - \tilde{Y}$, with $\tilde{Y} \in \mathcal{Y}$ as follows:

$$\begin{aligned} P_{\tilde{Y}|Z'}(y|e) &= \frac{1}{\alpha'} \min_{z'\in\mathcal{X}} P_{Y|X}(y|z'), \\ P_{\tilde{Y}|Z'}(y|z') &= \frac{1}{1-\alpha'}\left(P_{Y|X}(y|z') - \min_{\tilde{z}\in\mathcal{X}} P_{Y|X}(y|\tilde{z})\right), \tilde{z}\in\mathcal{X}. \end{aligned}$$

Note that all the above terms are non-negative. Moreover,

$$\sum_{y\in\mathcal{Y}} P_{\tilde{Y}|Z'}(y|e) = \frac{\sum_{y\in\mathcal{Y}} \min_{z'\in\mathcal{X}} P_{Y|X}(y|z)}{\alpha'} = 1.$$

Similarly,

$$\begin{aligned} \sum_{y\in\mathcal{Y}} P_{\tilde{Y}|Z'}(y|z') &= \frac{\sum_{y\in\mathcal{Y}} \left(P_{Y|X}(y|z') - \min_{\tilde{z}\in\mathcal{X}} P_{Y|X}(y|\tilde{z})\right)}{1-\alpha'} \\ &= 1. \end{aligned}$$

Hence, the conditional distribution above is valid. Now note that for all $y \in \mathcal{Y}$ and $x \in \mathcal{X}$,

$$\begin{aligned} &P_{\tilde{Y}|X}(y|x) \\ &= \sum_{z'\in\mathcal{X}\cup\{e\}} P_{Z'|X}(z'|x)P_{\tilde{Y}|Z'}(y|z') \\ &= (1-\alpha')\, P_{\tilde{Y}|Z'}(y|x) + \alpha' P_{\tilde{Y}|Z'}(y|e) \\ &= P_{Y|X}(y|x) - \min_{x'\in\mathcal{X}} P_{Y|X}(y|x') + \min_{x'\in\mathcal{X}} P_{Y|X}(y|x') \\ &= P_{Y|X}(y|x). \end{aligned}$$

Hence, $X - Z' - Y$ can be coupled as a Markov chain. Finally, one can obtain $X - Z - Z' - Y$ as a Markov chain by defining $P_{Z'|Z}$ as

$$P_{Z'|Z}(z'|z) = \begin{cases} 1, & \text{if } z = e \text{ and } z' = e, \\ \frac{\alpha'-\alpha}{1-\alpha}, & \text{if } z \neq e \text{ and } z' = e \\ \frac{1-\alpha'}{1-\alpha}, & \text{if } z = z' \text{ and } z' \neq e \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

and the induced conditional $P_{Y|Z}$ is as given in equations (6) and (7). Indeed, for $z = e$,

$$\begin{aligned} P_{Y|Z}(y|e) &= \sum_{z'\in\mathcal{X}\cup\{e\}} P_{Y|Z'}(y|z')P_{Z'|Z}(z'|z) \\ &= P_{Y|Z'}(y|e)\cdot 1 \\ &= \frac{1}{\alpha'} \min_{z'\in\mathcal{X}} P_{Y|X}(y|z'), \end{aligned}$$

as given in equation (6). Similarly, for $z \neq e$,

$$\begin{aligned} P_{Y|Z}(y|z) &= \sum_{z'\in\mathcal{X}\cup\{e\}} P_{Y|Z'}(y|z')P_{Z'|Z}(z'|z) \\ &= P_{Y|Z'}(y|z)P_{Z'|Z}(z|z) + P_{Y|Z'}(y|e)P_{Z'|Z}(e|z) \\ &= \frac{1}{1-\alpha'}\left(P_{Y|X}(y|z) - \min_{\tilde{z}\in\mathcal{X}} P_{Y|X}(y|\tilde{z})\right)\frac{1-\alpha'}{1-\alpha} \\ &\quad + \frac{1}{\alpha'}\left(\min_{\tilde{z}\in\mathcal{X}} P_{Y|X}(y|\tilde{z})\right)\frac{\alpha'-\alpha}{1-\alpha} \\ &= \frac{1}{1-\alpha}\left(P_{Y|X}(y|z) - \frac{\alpha}{\alpha'}\min_{\tilde{z}\in\mathcal{X}} P_{Y|X}(y|\tilde{z})\right), \end{aligned}$$

as given in equation (7). ∎

Now consider the $n = 1$ case in Theorem 1, that is, we have $\mathscr{L}\left(P_{Y|X}\right) \leq \ell$. It follows from Lemma 1 that $\sum_{y\in\mathcal{Y}} \min_{x\in\{0,1\}} P_{Y|X}(y|x) \geq 2 - e^{\ell}$. Hence, from Lemma 2, $X - Z - Y$ can be coupled as a Markov chain where $P_{Z|X} = \mathrm{BEC}(2-e^{\ell})$ and $P_{Y|Z}$ is as given in (6) and (7) with $\alpha = 2 - e^{\ell}$ and $\alpha' = 2 - e^{\mathscr{L}(P_{Y|X})}$.

Now consider the general case. Let $\tilde{Y}^n \in \mathcal{Y}_1 \times \mathcal{Y}_2 \times \ldots \times \mathcal{Y}_n$ and define the joint

$$\begin{aligned} &P_{XZ^n\tilde{Y}^n} := \\ &P_X P_{Z_1|X} P_{\tilde{Y}_1|Z_1} P_{Z_2|X} P_{\tilde{Y}_2|Z_2,\tilde{Y}_1} \ldots P_{Z_n|X} P_{\tilde{Y}_n|Z_n,\tilde{Y}^{n-1}}, \end{aligned}$$

where $P_{Z_i|X} = \mathrm{BEC}(2-e^{\ell_i})$, and $P_{\tilde{Y}_i|Z_i,\tilde{Y}^{i-1}=y^{i-1}}$ is defined as in equations (6) and (7) of Lemma 2, where $\alpha = 2 - e^{\ell_i}$ and $\alpha' = 2 - \exp\{\mathscr{L}\left(P_{Y_i|X,Y_1=y_1,Y_2=y_2,\ldots,Y_{i-1}=y_{i-1}}\right)\}$. By construction, $X - Z^n - \tilde{Y}^n$ is a Markov chain, and

$$\begin{aligned} P_{\tilde{Y}^n|X}(y^n|x) &= \sum_{z^n} P_{\tilde{Y}^n,Z^n|X}(y^n, z^n|x) \\ &= \sum_{z^n}\prod_{i=1}^{n} P_{Z_i|X}(z_i|x)P_{\tilde{Y}_i|Z_i,\tilde{Y}^{i-1}}(y_i|z_i, y^{i-1}) \\ &= \prod_{i=1}^{n}\sum_{z_i} P_{Z_i|X}(z_i|x)P_{\tilde{Y}_i|Z_i,\tilde{Y}^{i-1}}(y_i|z_i, y^{i-1}) \\ &\overset{(a)}{=} \prod_{i=1}^{n} P_{Y_i|X,Y^{i-1}}(y_i|x, y^{i-1}) \\ &= P_{Y^n|X}(y^n|x), \end{aligned}$$

where (a) follows from Lemma 2. Hence this coupling yields that $X - Z^n - Y^n$ is a Markov chain.

To show the equality $\mathscr{L}\left(X \to Z_1, Z_2, \ldots, Z_n\right) = \log\left(2 - \prod_{i=1}^{n}(2-e^{\ell_i})\right)$, note that for the n-fold BEC, for any output z^n, if there exists i such that $z_i \neq e$, then $P_{Z^n|X}(z^n|0) = 0$ or $P_{Z^n|X}(z^n|1) = 0$. Hence,

$$\begin{aligned} \sum_{\substack{z^n \in \\ \{0,1,e\}^n}} \min_{x\in\{0,1\}} P_{Z^n|X}(z^n|x) &= \min_{x\in\{0,1\}} P_{Z^n|X}((e,e,\ldots,e)|x) \\ &= \prod_{i=1}^{n}(2-e^{\ell_i}). \end{aligned}$$

Hence, by Lemma 1,

$$\mathscr{L}\left(P_{Z^n|X}\right) = \log\left(2 - \prod_{i=1}^{n}(2 - e^{\ell_i})\right).$$

III. Chernoff Information and Total Variation Distance

Definition 2 ([16]). *The* Chernoff information *between two probability mass functions, P and Q, over the same alphabet $\mathcal{X}$ is given as follows. First, for all $x \in \mathcal{X}$ and $\lambda \in [0,1]$, let:*

$$P_\lambda(x) = \frac{P(x)^\lambda Q(x)^{1-\lambda}}{\sum_{x' \in \mathcal{X}} P(x')^\lambda Q(x')^{1-\lambda}}. \tag{9}$$

Then the Chernoff information is given by

$$\mathscr{C}(P||Q) = D(P_{\lambda^*}||P) = D(P_{\lambda^*}||Q), \tag{10}$$

where λ^* is any value of λ such that the above two relative entropies are equal. Equivalently, the Chernoff information is also given by:

$$\mathscr{C}(P||Q) = -\min_{0 \le \lambda < 1} \log\left(\sum_x P(x)^\lambda Q(x)^{1-\lambda}\right). \tag{11}$$

Wu *et al.* [5] analyzed the case in which $Y_1, Y_2, \ldots, Y_n$ are conditionally i.i.d given X. In particular, it was shown that (assuming the rows of $P_{Y|X}$ are distinct)

$$\lim_{n\to\infty} \mathscr{L}\left(X \to Y^n\right) = \log|\mathcal{X}|, \tag{12}$$

and the convergence is exponential with a rate given by the minimum Chernoff information between any two "rows" of $P_{Y|X}$, that is,

$$\begin{aligned}&\lim_{n\to\infty} -\frac{1}{n}\log\left(\log|\mathcal{X}| - \mathscr{L}\left(X \to Y^n\right)\right)\\ &= \min_{x_1 \neq x_2} \mathscr{C}\left(P_{Y|X}(.|x_1)||P_{Y|X}(.|x_2)\right).\end{aligned} \tag{13}$$

Notably, it was shown [5] that $I(X;Y^n)$ converges to its limit, $H(X)$, at the same rate. Similarly, Sibson mutual information $I_\alpha^S(X;Y^n)$ and Arimoto mutual information $I_\alpha^S(X;Y^n)$ converge to their respective limits at the same rate for all $\alpha \in [0,\infty]$.

Theorem 1 and the result of Wu *et al.* [5] thus yield that the binary-input channel that has the highest Chernoff information between its rows (hence the fastest convergence of $\mathscr{L}\left(X \to Y^n\right)$ to its limit) for a given maximal leakage constraint is the binary erasure channel. This is formalized in the next subsection where we cast the optimization in terms of total variation distance (which, by virtue of Lemma 1, is equivalent to maximal leakage for binary-input channels). On the other hand, we show that the binary-input channel minimizing the Chernoff information for a given maximal leakage (or a given total variation distance between the two rows) is the binary symmetric channel.

A. Maximum Chernoff Information

Corollary 1. *Consider any $\alpha \in [0,1]$. Then,*

$$\max_{\mathcal{Y}} \max_{\substack{P,Q:\\ d_{TV}(P,Q)\le\alpha}} \mathscr{C}(P,Q) = -\log(1-\alpha), \tag{14}$$

where $\mathcal{Y}$ is an arbitrary finite alphabet, and P and Q are distributions over $\mathcal{Y}$. In particular, the maximum is achieved for $\mathcal{Y} = \{0,1,e\}$, and

$$\begin{bmatrix} P^\star \\ Q^\star \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 1-\alpha \\ 0 & \alpha & 1-\alpha \end{bmatrix},$$

i.e., $P^\star$ and $Q^\star$ are the rows of the $\mathrm{BEC}(1-\alpha)$.

Proof: Let $\mathcal{X} = \{0,1\}$. Fix any alphabet $\mathcal{Y}$ and any two distributions P and Q over $\mathcal{Y}$ with $d_{TV}(P,Q) = \alpha' \le \alpha$, and define the conditional distribution $W_{Y|X}$ such that

$$\begin{aligned} W_{Y|X=0} &= P \\ \text{and } W_{Y|X=1} &= Q. \end{aligned}$$

Define $W_{Y^n|X}$ as $\prod_{i=1}^n W_{Y_i|X}$, i.e., Y^n are i.i.d conditioned on X and for all i, $W_{Y_i|X} = W_{Y|X}$. Now, for $n \in \mathbb{N}$, let $\mathcal{P}_n(\alpha')$ be the set of $\vec{\ell}$-LAMs $\{P_{Y^n|X}\}$ with $\mathcal{Y}_i = \mathcal{Y}$ and $\ell_i = \ell := \log(1+\alpha')$ for all $i \in \{1,2,\ldots,n\}$. Note that $W_{Y^n|X} \in \mathcal{P}_n(\alpha')$ since, by Lemma 1,

$$\mathscr{L}\left(W_{Y|X}\right) = \log(1 + d_{TV}(P,Q)) = \log(1+\alpha') = \ell.$$

Then,

$$\begin{aligned}\mathscr{C}(P,Q) &\overset{(a)}{=} \lim_{n\to\infty} -\frac{1}{n}\log\left(\log 2 - \mathscr{L}\left(W_{Y^n|X}\right)\right)\\ &\le \lim_{n\to\infty} -\frac{1}{n}\log\left(\log 2 - \sup_{P_{Y^n|X}\in\mathcal{P}_n(\alpha')} \mathscr{L}\left(P_{Y^n|X}\right)\right)\\ &\overset{(b)}{=} \lim_{n\to\infty} -\frac{1}{n}\log\left(\log 2 - \log\left(2 - (1-\alpha')^n\right)\right)\\ &\overset{(c)}{=} -\log(1-\alpha')\\ &\le -\log(1-\alpha).\end{aligned}$$

where (a) follows from [5, Theorem 1] (cf. equation (13)), (b) follows from Theorem 1 and the fact that $e^{\ell_i} = 1 + \alpha'$ for all $i \in \{1,2,\ldots,n\}$, and (c) follows from l'Hopital's rule (applied twice).

On the other hand, it is easy to check that $d_{TV}(P^\star, Q^\star) = \alpha$ and $\mathscr{C}(P^\star, Q^\star) = -\log(1-\alpha)$. ∎

B. Minimum Chernoff Information

We now turn to minimizing the Chernoff information for a given total variation constraint. To that end, we will make use of the properties of Rényi divergence.

Definition 3. *The* Rényi divergence *of order α between probability distributions P and Q is defined for $\alpha \in [0,\infty)$, $\alpha \neq 1$ as:*

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1}\log\sum_{x\in\mathcal{X}} P(x)^\alpha Q(x)^{1-\alpha}, \tag{15}$$

where the continuous extension at $\alpha = 1$ is given by the standard Kullback-Leibler divergence

$$D(P||Q) = \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)}. \tag{16}$$

Proposition 3. *Consider any $\alpha \in [0,1]$. Then,*

$$\min_{\mathcal{Y}} \min_{\substack{P,Q:\\ d_{TV}(P,Q)\geq\alpha}} \mathscr{C}(P,Q) = -\frac{1}{2}\log\left((1-\alpha)(1+\alpha)\right), \tag{17}$$

where $\mathcal{Y}$ is an arbitrary alphabet, and P and Q are distributions over $\mathcal{Y}$. In particular, the minimum is achieved for $\mathcal{Y} = \{0,1\}$ and

$$\begin{bmatrix} P^\star \\ Q^\star \end{bmatrix} = \begin{bmatrix} (1+\alpha)/2 & (1-\alpha)/2 \\ (1-\alpha)/2 & (1+\alpha)/2 \end{bmatrix},$$

i.e., $P^\star$ and $Q^\star$ are the rows of the binary symmetric channel with parameter $(1-\alpha)/2$.

Proof: Consider any $\mathcal{Y}$, $|\mathcal{Y}| \geq 2$, and any two distributions P and Q over $\mathcal{Y}$. Let Q_0 be any measure such that $P \ll Q_0$ and $Q \ll Q_0$, e.g., $Q_0 = P + Q$. Let

$$A = \left\{ y \in \mathcal{Y} : \frac{dP}{dQ_0}(y) > \frac{dQ}{dQ_0}(y) \right\}.$$

(Note that here $\mathcal{Y}$ is not assumed to be finite, hence the use of the measure-theoretic notation.) Now let $\mathcal{Y}' = \{0,1\}$, and define P' and Q' over $\mathcal{Y}'$ such that

$$P'(0) = P(A) \text{ and } Q'(0) = Q(A).$$

Note that $d_{TV}(P',Q') = d_{TV}(P,Q)$. Moreover,

$$\begin{aligned}\mathscr{C}(P',Q') &= -\min_{0\leq\lambda\leq 1}(\lambda-1)D_\lambda(P',Q') \\ &\leq -\min_{0\leq\lambda\leq 1}(\lambda-1)D_\lambda(P,Q) = \mathscr{C}(P,Q),\end{aligned}$$

where the inequality follows from the data processing inequality for Rényi divergences [17, Theorem 9]. It is thus sufficient to consider distributions P and Q over $\mathcal{Y} = \{0,1\}$. Note that $D_\lambda(P,Q)$ is convex in the pair (P,Q) [17, Theorem 11], hence

$$\mathscr{C}(P,Q) = \sup_{0\leq\lambda\leq 1} -(\lambda-1)D_\lambda(P,Q)$$

is a supremum of convex functions, so it is also convex in (P,Q). Without loss of generality, suppose $P(0) \geq Q(0)$. In particular,

$$P(0) = Q(0) + P(0) - Q(0) = Q(0) + d_{TV}(P,Q) = Q(0) + \alpha',$$

where $\alpha' \geq \alpha$. Define the distributions

$$\begin{bmatrix} \hat{P} \\ \hat{Q} \end{bmatrix} = \begin{bmatrix} (1+\alpha')/2 & (1-\alpha')/2 \\ (1-\alpha')/2 & (1+\alpha')/2 \end{bmatrix},$$

$\bar{P} = 1 - P$ and $\bar{Q} = 1 - Q$. Then, by symmetry (cf. equations (11) and (10)),

$$\mathscr{C}(P,Q) = \mathscr{C}(\bar{P},\bar{Q}) = \mathscr{C}(\bar{Q},\bar{P}).$$

Moreover, note that for any distributions P and Q over $\{0,1\}$ with $P(0) = Q(0) + \alpha'$,

$$\hat{P} = (P + \bar{Q})/2 \text{ and } \hat{Q} = (Q + \bar{P})/2,$$

so that

$$\mathscr{C}(P,Q) = \frac{1}{2}\mathscr{C}(P,Q) + \frac{1}{2}\mathscr{C}(\bar{Q},\bar{P}) \geq \mathscr{C}(\hat{P},\hat{Q}),$$

where the inequality follows from convexity. Finally,

$$\begin{aligned}\mathscr{C}(\hat{P},\hat{Q}) &\overset{(a)}{=} -\frac{1}{2}\log\left((1-\alpha')(1+\alpha')\right) \\ &\overset{(b)}{\geq} -\frac{1}{2}\log\left((1-\alpha)(1+\alpha)\right)\end{aligned}$$

where (a) follows from equation (11) (the minimum is achieved for $\lambda^\star = 1/2$), and (b) follows from the fact that the function $f(t) = -\log((1-t)(1+t))$ is non-decreasing over $[0,1]$.

On the other hand, it is easy to check that $d_{TV}(P^\star,Q^\star) = \alpha$ and $\mathscr{C}(P^\star,Q^\star) = -\frac{1}{2}\log\left((1-\alpha)(1+\alpha)\right)$. ∎

References

[1] P. Kairouz, S. Oh, and P. Viswanath, "The composition theorem for differential privacy," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 4037–4049, Jun. 2017.

[2] C. Dwork and G. N. Rothblum, "Concentrated differential privacy." [Online]. Available: arXiv:1603.01887

[3] I. Mironov, "Rényi differential privacy," in *Proc. IEEE Comp. Sec. Found. Symp.*, 2017, pp. 263–275.

[4] J. Dong, A. Roth, and W. J. Su, "Gaussian differential privacy," *arXiv preprint arXiv:1905.02383*, 2019.

[5] B. Wu, A. B. Wagner, G. E. Suh, and I. Issa, "Strong asymptotic composition theorems for mutual information measures," *arXiv preprint arXiv:2005.06033*, 2020.

[6] R. Sibson, "Information radius," *Zeitschrift for Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 14, no. 2, pp. 149–160, 1969. [Online]. Available: http://link.springer.com/10.1007/BF00537520

[7] S. Verdú, "α-mutual information," in *Proc. Inf. Theory and Appl. (ITA) Workshop*, 2015.

[8] S. Arimoto, "Information measures and capacity of order α for discrete memoryless channels," *Topics in Information Theory Proc. Coll. Math Soc. Janos Bolyai*, pp. 41–52, 1975.

[9] J. Liao, O. Kosut, L. Sankar, and F. P. Calmon, "A tunable measure for information leakage," *arXiv:1806.03332 [cs, math]*, Jun. 2018, arXiv: 1806.03332. [Online]. Available: http://arxiv.org/abs/1806.03332

[10] C. Braun, K. Chatzikokolakis, and C. Palamidessi, "Quantitative notions of leakage for one-try attacks," *Electronic Notes in Theoretical Computer Science*, vol. 249, pp. 75–91, 2009.

[11] B. Espinoza and G. Smith, "Min-entropy as a resource," *Information and Computation*, vol. 226, pp. 57 – 75, 2013, special Issue: Information Security as a Resource.

[12] M. Alvim, K. Chatzikokolakis, C. Palamidessi, and G. Smith, "Measuring information leakage using generalized gain functions," in *Computer Security Foundations Symposium (CSF), 2012 IEEE 25th*, June 2012, pp. 265–279.

[13] I. Issa, A. B. Wagner, and S. Kamath, "An operational measure of information leakage," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1625–1657, Mar. 2020.

[14] I. Issa, S. Kamath, and A. B. Wagner, "An operational measure of information leakage," in *Proc. Conf. Inf. Sci. and Sys. (CISS)*, 2016, pp. 234–239.

[15] P. Kairouz, S. Oh, and P. Viswanath, "The composition theorem for differential privacy," in *International conference on machine learning*. PMLR, 2015, pp. 1376–1385.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., 2006.

[17] T. van Erven and P. Harremoës, "Rényi divergence and Kullback-Leibler divergence," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 3797–3820, Jul. 2014, arXiv: 1206.2459. [Online]. Available: http://arxiv.org/abs/1206.2459