Privacy-Aware Guessing Efficiency

Shahab Asoodeh, Mario Diaz, Fady Alajaji, and Tamás Linder

***Abstract*—We investigate the problem of guessing a discrete random variable Y under a privacy constraint dictated by another correlated discrete random variable X, where both guessing efficiency and privacy are assessed in terms of the probability of correct guessing. We define $\hbar(P_{XY}, \varepsilon)$ as the maximum probability of correctly guessing Y given an auxiliary random variable Z, where the maximization is taken over all $P_{Z|Y}$ ensuring that the probability of correctly guessing X given Z does not exceed ε. We show that the map $\varepsilon \mapsto \hbar(P_{XY}, \varepsilon)$ is strictly increasing, concave, and piecewise linear, which allows us to derive a closed form expression for $\hbar(P_{XY}, \varepsilon)$ when X and Y are connected via a binary-input binary-output channel. For $\{(X_i, Y_i)\}_{i=1}^n$ being pairs of independent and identically distributed binary random vectors, we similarly define $\underline{h}_n(P_{X^nY^n}, \varepsilon)$ under the assumption that Z^n is also a binary vector. Then we obtain a closed form expression for $\underline{h}_n(P_{X^nY^n}, \varepsilon)$ for sufficiently large, but nontrivial values of ε.**

I. Introduction and Preliminaries

Given private information, represented by a random variable X, non-private observable information, say Y, is generated via a fixed channel $P_{Y|X}$. Consider two communicating agents Alice and Bob, where Alice observes Y and wishes to disclose it to Bob as accurately as possible in order to receive a payoff, but in such a way that X is kept almost private from him. Given the joint distribution P_{XY}, Alice chooses a random mapping $P_{Z|Y}$, a so-called privacy filter, to generate a new random variable Z, called the *displayed data*, such that Bob can *guess* Y from Z with as small error probability as possible while Z cannot be used to efficiently guess X.

The tradeoff between utility and privacy was addressed from an information-theoretic viewpoint in [1]–[5], where both utility and privacy were measured in terms of information-theoretic quantities. In particular, in [2] both utility and privacy were measured in terms of the mutual information I. Specifically, the so-called *rate-privacy function* $g(P_{XY}, \varepsilon)$ was defined as the maximum of $I(Y; Z)$ over all $P_{Z|Y}$ such that $I(X; Z) \leq \varepsilon$. In the most stringent privacy setting $\varepsilon = 0$, called *perfect privacy*, it was shown that $g(P_{XY}, 0) > 0$ if and only if X is weakly independent of Y, that is, if the set of vectors $\{P_{X|Y}(\cdot|y) : y \in \mathcal{Y}\}$ is linearly dependent. In [4], an equivalent result was obtained in terms of the singular values of the operator $f \mapsto \mathbb{E}[f(X)|Y]$. Although a connection between this information-theoretic privacy measure and a coding theorem is established in [2] and [6], the use of mutual information as a privacy measure is not satisfactorily motivated in an *operational* sense. To find a measure of privacy with a clear operational meaning, in this paper we take an estimation-theoretic approach and define both privacy and utility measures in terms of the probability of guessing correctly.

This work was supported in part by NSERC of Canada.

The authors are with the Department of Mathematics and Statistics, Queen's University, Canada. Emails: {asoodehshahab, fady, linder}@mast.queensu.ca, 13madt@queensu.ca.

Given discrete random variables $U \in \mathcal{U}$ and $V \in \mathcal{V}$, the probability of correctly guessing U given V is defined as

$$\mathsf{P_c}(U|V) := \max_g \Pr(U = g(V)) = \sum_{v \in \mathcal{V}} \max_{u \in \mathcal{U}} P_{UV}(u, v),$$

where the first maximum is taken over all functions $g : \mathcal{V} \to \mathcal{U}$. It is easy to show that $\mathsf{P_c}$ satisfies the data processing inequality, i.e., $\mathsf{P_c}(U|W) \leq \mathsf{P_c}(U|V)$ for U, V and W which form the Markov chain $U \multimap V \multimap W$. Thus, we measure privacy in terms of $\mathsf{P_c}(X|Z)$ which quantifies the advantage of an adversary observing Z in guessing X in a single shot attempt.

A similar operational measure of privacy was recently proposed in [7], where $P_{Z|X}$ is said to be ε-private if $\log \frac{\mathsf{P_c}(U|Z)}{\mathsf{P_c}(U)} \leq \varepsilon$ for *all* auxiliary random variables U satisfying $U \multimap X \multimap Z$. This requirement guarantees that no *randomized* function of X can be efficiently estimated from Z, which leads to a strong privacy guarantee. In [8], maximal correlation [9] was proposed as another measure of privacy. Operational interpretations corresponding to this privacy measure are given in [10] for the discrete case and in [11] for a continuous setup.

To quantify the conflict between utility and privacy, we define the *privacy-aware guessing function* $\hbar$ as

$$\hbar(P_{XY}, \varepsilon) := \sup_{\substack{P_{Z|Y}: X \multimap Y \multimap Z, \\ \mathsf{P_c}(X|Z) \leq \varepsilon}} \mathsf{P_c}(Y|Z). \qquad (1)$$

Due to the data processing inequality, we can restrict the privacy threshold ε to the interval $[\mathsf{P_c}(X), \mathsf{P_c}(X|Y)]$, where $\mathsf{P_c}(X)$ is the probability of correctly guessing X in the absence of any side information. For ε close to $\mathsf{P_c}(X)$, the privacy guarantee $\mathsf{P_c}(X|Z) \leq \varepsilon$ intuitively means that it is nearly as hard to guess X observing Z as it is without observing Z.

We derive functional properties of the map $\varepsilon \mapsto \hbar(P_{XY}, \varepsilon)$. In particular, we show that it is strictly increasing, concave, and piecewise linear. Piecewise linearity (Theorem 1), which is the most important and technically difficult result in the paper, allows us to derive a tight upper bound on $\hbar(P_{XY}, \varepsilon)$ for general P_{XY}. As a consequence of concavity, we derive a closed form expression for $\hbar(P_{XY}, \varepsilon)$ for any $\varepsilon \in [\mathsf{P_c}(X), \mathsf{P_c}(X|Y)]$ when X and Y are both binary. It is shown (Theorem 2) that either the Z-channel or the *reverse* Z-channel achieves $\hbar(P_{XY}, \varepsilon)$ in this case depending on the backward channel.

We also consider the vector case for a pair of binary random vectors (X^n, Y^n) under an additional constraint that Z^n is a binary random vector. Here, Z^n is revealed publicly and the goal is to guess Y^n under the privacy constraint $\mathsf{P_c}(X^n|Z^n) \leq \varepsilon^n$. This model can be viewed as a privacy-constrained version of the *correlation distil-*

lation problem studied in [12]. Suppose Alice and Bob respectively observe Y^n and Z^n, where $\{(Y_i, Z_i)\}_{i=1}^n$ is independent and identically distributed (i.i.d.) according to the joint distribution P_{YZ}, and assume that they are to design non-constant Boolean functions f and g such that $\Pr(f(Y^n) = g(Z^n))$ is maximized. A dimension-free upper bound for this probability was given in [12]. Now suppose P_{YZ} is not given and Alice is to design $P_{Z|Y}$ (for a fixed $\mathcal{Y}$-marginal) that maximizes $\mathsf{P_c}(f(Y^n)|Z^n)$ for a given function f while $\mathsf{P_c}(X^n|Z^n) \le \varepsilon^n$. We show (Theorem 3) that if $\{(X_i, Y_i)\}_{i=1}^n$ is i.i.d. according to P_{XY} with $|\mathcal{X}| = |\mathcal{Y}| = 2$ and $P_{Y|X}$ is a binary symmetric channel, then the maximum of $\mathsf{P_c}(Y^n|Z^n)$ under the privacy constraint $\mathsf{P_c}(X^n|Z^n) \le \varepsilon^n$ admits a closed form expression for sufficiently large but nontrivial ε. This then provides a lower bound for the privacy-constrained correlation distillation problem due to the trivial fact that $\mathsf{P_c}(f(Y^n)|Z^n) \ge \mathsf{P_c}(Y^n|Z^n)$ for any function f.

We omit the proof of most of the results due to space limitations. The proofs are available in [13].

II. Scalar Case

Suppose X and Y are discrete random variables with finite alphabets $\mathcal{X} = \{1, \dots, M\}$ and $\mathcal{Y} = \{1, \dots, N\}$, respectively, and with joint distribution $\mathsf{P} = \{P_{XY}(x, y), x \in \mathcal{X}, y \in \mathcal{Y}\}$, whose marginals over $\mathcal{X}$ and $\mathcal{Y}$ are $(p_1, \dots, p_M)$ and $(q_1, \dots, q_N)$, respectively. Let X represent the private data and Y represent a non-private measurement of X, which, upon passing it via a privacy filter $P_{Z|Y}$, is publicly displayed as Z. In order to quantify the conflict between privacy with respect to X and utility with respect to Y, the so-called rate-privacy function $g(\mathsf{P}, \varepsilon)$ was introduced in [2]. In what follows, we use Arimoto's mutual information to generalize this definition.

A. The Utility-Privacy Function of Order (ν, μ)

Let $H_\nu(X)$ and $H_\nu^{\mathsf{A}}(X|Z)$ denote respectively the Rényi entropy of order ν and Arimoto's conditional entropy of order ν [14], defined for $\nu > 1$ as

$$H_\nu(X) := \frac{1}{1-\nu} \log \left(\sum_{x \in \mathcal{X}} P_X^\nu(x) \right),$$

and

$$H_\nu^{\mathsf{A}}(X|Z) := \frac{\nu}{1-\nu} \log \left(\sum_{z \in \mathcal{Z}} \left[\sum_{x \in \mathcal{X}} P_{XZ}^\nu(x, z) \right]^{1/\nu} \right).$$

We define (by continuity) $H_1(X) = H(X)$, $H_1^{\mathsf{A}}(X|Z) = H(X|Z)$, $H_\infty(X) = -\log \mathsf{P_c}(X)$, and $H_\infty^{\mathsf{A}}(X|Z) = -\log \mathsf{P_c}(X|Z)$. *Arimoto's mutual information of order* $\nu \ge 1$ is defined as (see, e.g., [14])

$$I_\nu^{\mathsf{A}}(X;Z) := H_\nu(X) - H_\nu^{\mathsf{A}}(X|Z).$$

Thus $I_1^{\mathsf{A}}(X;Z) = I(X;Z)$.

Definition 1. *For a given joint distribution* P *and a pair* (ν, μ)*,* $\nu, \mu \in [1, \infty]$*, the utility-privacy function of order* (ν, μ) *is*

$$g^{(\nu,\mu)}(\mathsf{P}, \varepsilon) := \max_{P_{Z|Y} \in \mathscr{D}^\nu(\mathsf{P}, \varepsilon)} I_\mu^{\mathsf{A}}(Y;Z),$$

where

$$\mathscr{D}^\nu(\mathsf{P}, \varepsilon) := \{P_{Z|Y} : X -\!\!\circ\!\!- Y -\!\!\circ\!\!- Z, I_\nu^{\mathsf{A}}(X;Z) \le \varepsilon\}.$$

Note that $\mathscr{D}^\nu(\mathsf{P}, \varepsilon)$ cannot be empty since all channels $P_{Z|Y}$ with Z independent of X satisfy $I_\nu^{\mathsf{A}}(X;Z) = 0$, and so they belong to $\mathcal{D}^\nu(\mathsf{P}, \varepsilon)$ for any $\varepsilon \ge 0$. Using a similar technique as in [15], one can show that $\varepsilon \mapsto g^{(\nu,\mu)}(\mathsf{P}, \varepsilon)$ is strictly increasing for any $\nu, \mu \ge 1$. It is also worth mentioning that an application of Minkowski's inequality implies that the map $P_{Z|Y} \mapsto \exp\left\{\frac{(\nu-1)}{\nu} I_\nu^{\mathsf{A}}(Y;Z)\right\}$ is convex for $\nu \ge 1$, and thus the maximum in the definition of $g^{(\nu,\mu)}(\mathsf{P}, \varepsilon)$ is achieved at the boundary of the feasible set where $I_\nu^{\mathsf{A}}(X;Z) = \varepsilon$. We denote $g^{(\infty,\infty)}(\mathsf{P}, \varepsilon)$ and $g^{(1,1)}(\mathsf{P}, \varepsilon)$ respectively by $g^\infty(\mathsf{P}, \varepsilon)$ and $g(\mathsf{P}, \varepsilon)$. Since $I_\infty(Y;Z) = \log \frac{\mathsf{P_c}(Y|Z)}{\mathsf{P_c}(Y)}$, $g^\infty(\mathsf{P}, \varepsilon)$ can be equivalently described as the smallest $\Gamma \ge 0$ such that $\mathsf{P_c}(Y|Z) \le \mathsf{P_c}(Y)2^\Gamma$, for every $P_{Z|Y}$ satisfying $\mathsf{P_c}(X|Z) \le \mathsf{P_c}(X)2^\varepsilon$. We note that for small ε the condition $I_\infty^{\mathsf{A}}(X;Z) \le \varepsilon$ intuitively means that it is nearly as hard for an *adversary* observing Z to predict X as it is without Z. Therefore, $g^\infty(\mathsf{P}, 0)$ quantifies the efficiency of guessing Y from Z such that $\mathsf{P_c}(X|Z) = \mathsf{P_c}(X)$. It is thus interesting to obtain a necessary and sufficient condition for P under which $g^\infty(\mathsf{P}, 0) > 0$. We obtain such a condition for the special case of binary X and Y in the next section.

In general, the map $\nu \mapsto I_\nu^{\mathsf{A}}(X;Z)$ is not monotonic[1] and hence $P_{Z|Y}$ might belong to $\mathscr{D}^\nu(\mathsf{P}, \varepsilon)$ but not to $\mathscr{D}^\mu(\mathsf{P}, \varepsilon)$ for $\mu < \nu$. Nevertheless, the following lemma allows us to obtain upper and lower bounds for $g^{(\nu,\mu)}(\mathsf{P}, \cdot)$ in terms of $g^\infty(\mathsf{P}, \cdot)$.

Lemma 1. *Let* (X, Y) *be a pair of random variables having joint distribution* P *and* $\nu, \mu \in (1, \infty)$*. Then*

$$g^{(\nu,\mu)}(\mathsf{P}, \varepsilon) \le g^\infty(\mathsf{P}, \psi(\nu, \varepsilon)) + H_\mu(Y) - H_\infty(Y),$$

where $\psi(\nu, \varepsilon) := \frac{\nu-1}{\nu}\varepsilon + \frac{1}{\nu} H_\infty(X)$*. Furthermore, we have for* $\varepsilon \ge H_\nu(X) - H_\infty(X)$ *that*

$$g^{(\nu,\mu)}(\mathsf{P}, \varepsilon) \ge \frac{\mu}{\mu-1} g^\infty(\mathsf{P}, \varphi(\nu, \varepsilon)) - \frac{1}{\mu-1} H_\infty(Y),$$

where $\varphi(\nu, \varepsilon) := \varepsilon - H_\nu(X) + H_\infty(X)$*.*

This lemma shows that the family of functions $g^{(\nu,\mu)}(\mathsf{P}, \varepsilon)$ for $\nu, \mu > 1$ can be bounded from above and below by $g^\infty(\mathsf{P}, \delta)$, where δ depends on ε and ν. The case $\nu = \mu = 1$ is studied in [2]. As a result, in the following section we only focus on $g^\infty(\mathsf{P}, \varepsilon)$. It turns out that it is easier to study $\hbar(\mathsf{P}, \varepsilon)$, defined in (1), instead. It is straightforward to verify that

$$g^\infty(\mathsf{P}, \varepsilon) = \log \frac{\hbar(\mathsf{P}, 2^\varepsilon \mathsf{P_c}(X))}{\mathsf{P_c}(Y)},$$

and hence all the results for $\hbar(\mathsf{P}, \varepsilon)$ can be translated to results for $g^\infty(\mathsf{P}, \varepsilon)$. In particular, perfect privacy $g^\infty(\mathsf{P}, 0)$ corresponds to $\hbar(\mathsf{P}, \mathsf{P_c}(X))$. Notice that $\hbar(\mathsf{P}, \mathsf{P_c}(X)) > \mathsf{P_c}(Y)$ is equivalent to $g^\infty(\mathsf{P}, 0) > 0$. As opposed to $I_\nu(X;Z)$ with $1 \le \nu < \infty$, $I_\infty(X;Z) = 0$ does not

[1] It is relatively easy to show that if X is uniformly distributed, then $I_\nu^{\mathsf{A}}(X;Z)$ coincides with Sibson's mutual information of order ν [14] which is known to be increasing in ν [16, Theorem 4]. Consequently, $\nu \mapsto I_\nu^{\mathsf{A}}(X;Z)$ is increasing over $(1, \infty]$ if X is uniformly distributed.

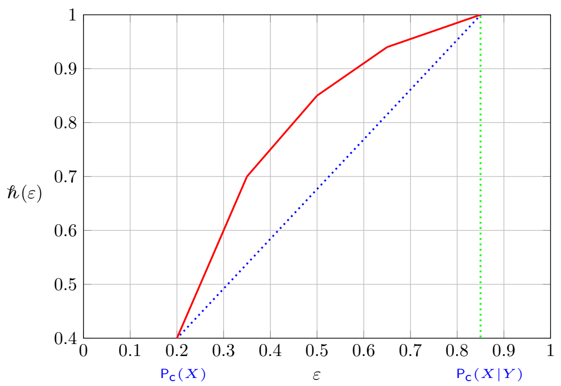


Fig. 1. Typical graph of $\hbar(\varepsilon)$. The dotted line represents the chord connecting $(p, \hbar(p))$ and $(\mathsf{P_c}(X|Y), 1)$ which can be viewed as a trivial lower bound for $\hbar(\cdot)$.

necessarily imply the independence of X and Z (unless X is uniformly distributed). In particular, the weak independence[2] argument from [2, Lemma 10] (see also [4]) cannot be applied for g^{∞}. For the sake of brevity, we simply write $\hbar(\varepsilon)$ for $\hbar(\mathsf{P}, \varepsilon)$ when there is no risk of confusion.

B. Privacy-Aware Guessing Function

It is clear from (1) that $\mathsf{P_c}(Y) \leq \hbar(\varepsilon) \leq 1$, and $\hbar(\varepsilon) = 1$ if and only if $\varepsilon \geq \mathsf{P_c}(X|Y)$. A direct application of the Support Lemma [17, Lemma 15.4] shows that it is enough to consider random variables Z supported on $\mathcal{Z} = \{1, \ldots, N+1\}$. Thus, the privacy filter $P_{Z|Y}$ can be realized by an $N \times (N+1)$ stochastic matrix F. Let $\mathcal{F}$ be the set of all such matrices. Then both utility $\mathcal{U}(\mathsf{P}, F) = \mathsf{P_c}(Y|Z)$ and privacy $\mathcal{P}(\mathsf{P}, F) = \mathsf{P_c}(X|Z)$ are functions of $F \in \mathcal{F}$ and we can express $\hbar(\varepsilon)$ as

$$\hbar(\varepsilon) = \max_{\substack{F \in \mathcal{F}, \\ \mathcal{P}(\mathsf{P}, \varepsilon) \leq \varepsilon}} \mathcal{U}(\mathsf{P}, F).$$

It can be verified that $F \mapsto \mathcal{P}(\mathsf{P}, F)$ and $F \mapsto \mathcal{U}(\mathsf{P}, F)$ are continuous convex functions over $\mathcal{F}$. It can also be shown that the set

$$\mathcal{R} := \{(\mathcal{P}(\mathsf{P}, F), \mathcal{U}(\mathsf{P}, F)) : \ F \in \mathcal{F}\}$$

is convex. Furthermore, since the graph of $\hbar(\varepsilon)$ is the upper boundary of $\mathcal{R}$, we conclude that $\varepsilon \mapsto \hbar(\varepsilon)$ is concave, and so it is strictly increasing and continuous on $[\mathsf{P_c}(X), \mathsf{P_c}(X|Y)]$. As a consequence, for every $\varepsilon \in [\mathsf{P_c}(X), \mathsf{P_c}(X|Y)]$ there exists G such that $\mathcal{P}(\mathsf{P}, G) = \varepsilon$ and $\mathcal{U}(\mathsf{P}, G) = \hbar(\varepsilon)$. We call such a privacy filter G optimal at ε.

The following theorem reveals that $\hbar(\cdot)$ is a piecewise linear function, as depicted in Fig. 1.

Theorem 1. *The function $\hbar : [\mathsf{P_c}(X), \mathsf{P_c}(X|Y)] \to \mathbb{R}$ is piecewise linear, i.e., there exist $K \geq 1$ and thresholds $\mathsf{P_c}(X) = \varepsilon_0 < \varepsilon_1 < \ldots < \varepsilon_K = \mathsf{P_c}(X|Y)$ such that $\hbar$ is linear on $[\varepsilon_{i-1}, \varepsilon_i]$ for all $1 \leq i \leq K$.*

Consider the map $\mathcal{H} : \mathcal{F} \to [0,1] \times [0,1]$ given by $\mathcal{H}(\mathsf{P}, F) = (\mathcal{P}(\mathsf{P}, F), \mathcal{U}(\mathsf{P}, F))$. Let $\mathcal{D} := \{D \in \mathcal{M}_{N \times N+1} : \|D\| = 1\}$, where $|| \cdot ||$ denotes the Euclidean norm on $\mathcal{M}_{N \times (N+1)}$, the set of real matrices of size $N \times (N+1)$. For $G \in \mathcal{F}$ define

$$\mathcal{D}(G) := \{D \in \mathcal{D} : G + tD \in \mathcal{F} \text{ for some } t > 0\}.$$

The proof of the previous theorem is heavily based on the following technical, yet intuitive, result: for every $G \in \mathcal{F}$, there exists $\delta > 0$ such that $\mathcal{H}$ is linear on $[G, G + \delta D]$ for *every* $D \in \mathcal{D}(G)$.

The proof technique allows us to derive the slope of $\hbar$ on $[\varepsilon_{i-1}, \varepsilon_i]$, given the family of optimal filters at a single point $\varepsilon \in [\varepsilon_{i-1}, \varepsilon_i]$. For example, since the family of optimal filters at $\varepsilon = \mathsf{P_c}(X|Y)$ is easily obtainable, it is then possible to compute $\hbar$ on the last interval. In the binary case, this observation and the concavity of $\hbar$ allow us to show that $\hbar$ is linear on its entire domain $[\mathsf{P_c}(X), \mathsf{P_c}(X|Y)]$.

C. Binary Case

Assume now that X and Y are both binary. Let $\mathsf{BIBO}(\alpha, \beta)$ denote a binary input binary output channel from X to Y with $P_{Y|X}(\cdot|0) = (\bar{\alpha}, \alpha)$ and $P_{Y|X}(\cdot|1) = (\beta, \bar{\beta})$, where $\bar{x} := 1 - x$ for $x \in [0,1]$. Notice that if $X \sim \mathsf{Bernoulli}(p)$ with $p \in [\frac{1}{2}, 1)$, then $\mathsf{P_c}(X) = p$ and hence $\hbar(p)$ corresponds to the maximum of $\mathsf{P_c}(Y|Z)$ under perfect privacy $\mathsf{P_c}(X|Z) = p$. Furthermore, if $P_{Y|X} = \mathsf{BIBO}(\alpha, \beta)$ with $\alpha, \beta \in [0, \frac{1}{2})$, then we have $\mathsf{P_c}(X|Y) = \max\{\bar{\alpha}\bar{p}, \beta p\} + \bar{\beta} p$. Notice that if $\bar{\alpha}\bar{p} \leq \beta p$, then $\mathsf{P_c}(X|Y) = \mathsf{P_c}(X) = p$.

The binary symmetric channel with crossover probability α, denoted by $\mathsf{BSC}(\alpha)$, and also the Z-channel with crossover probability β, denoted by $\mathsf{Z}(\beta)$, are both examples of $\mathsf{BIBO}(\alpha, \beta)$, corresponding to $\alpha = \beta$ and $\alpha = 0$, respectively. Let $q := \Pr(Y = 1)$. We say that perfect privacy yields a non-trivial utility if $\mathsf{P_c}(Y|Z) > \mathsf{P_c}(Y)$ for some Z such that $\mathsf{P_c}(X|Z) = \mathsf{P_c}(X)$, or equivalently, if $\hbar(p) > \max\{\bar{q}, q\}$. The following lemma determines $\hbar(p)$ in the non-trivial case $\bar{\alpha}\bar{p} > \beta p$.

Lemma 2. *Let $X \sim \mathsf{Bernoulli}(p)$ with $p \in [\frac{1}{2}, 1)$ and $P_{Y|X} = \mathsf{BIBO}(\alpha, \beta)$ with $\alpha, \beta \in [0, \frac{1}{2})$ such that $\bar{\alpha}\bar{p} > \beta p$. Then*

$$\hbar(p) = \begin{cases} 1 - \zeta q & \text{if } \alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2, \\ q & \text{otherwise,} \end{cases}$$

where $q = \alpha\bar{p} + \bar{\beta} p$ and

$$\zeta := \frac{\bar{\alpha}\bar{p} - \beta p}{\bar{\beta} p - \alpha\bar{p}}. \tag{2}$$

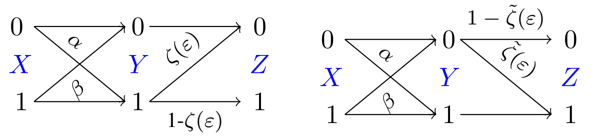
Notice that $1 - \zeta q > \bar{q}$ if and only if $\zeta < 1$, which occurs if and only if $p \in (\frac{1}{2}, 1)$. Also, it is straightforward to show that $1 - \zeta q > q$ if and only if $\alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2$. In particular, we have the following necessary and sufficient condition for non-trivial utility under perfect privacy.

Corollary 1. *Let $X \sim \mathsf{Bernoulli}(p)$ with $p \in [\frac{1}{2}, 1)$ and $P_{Y|X} = \mathsf{BIBO}(\alpha, \beta)$ with $\alpha, \beta \in [0, \frac{1}{2})$ such that $\bar{\alpha}\bar{p} > \beta p$. Then $g^{\infty}(\mathsf{P}, 0) > 0$ if and only if $\alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2$ and $p \in (\frac{1}{2}, 1)$.*

Remark that the condition $\alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2$ can be equivalently written as

$$P_{X|Y}(0|1) P_{X|Y}(0|0) < P_{X|Y}(1|0) P_{X|Y}(1|1).$$

[2]Using a similar proof as in [2], it can be shown that $g^{(\nu,\mu)}(\mathsf{P}, 0) > 0$ for $\nu, \mu \in [1, \infty)$ if and only if X is weakly independent of Y.



(a) $\alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2$ (b) $\alpha\bar{\alpha}\bar{p}^2 \geq \beta\bar{\beta}p^2$

Fig. 2. The optimal privacy filters for $P_{Y|X} = \mathsf{BIBO}(\alpha, \beta)$.

The following theorem establishes the linear behavior of $\hbar$ when $P_{Y|X} = \mathsf{BIBO}(\alpha, \beta)$.

Theorem 2. *Let $X \sim \mathsf{Bernoulli}(p)$ for $p \in [\frac{1}{2}, 1)$ and $P_{Y|X} = \mathsf{BIBO}(\alpha, \beta)$ with $\alpha, \beta \in [0, \frac{1}{2})$. If $\bar{\alpha}\bar{p} > \beta p$, then for any $\varepsilon \in [p, \bar{\alpha}\bar{p} + \bar{\beta}p]$, we have the following:*

- *If $\alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2$, then*
$$\hbar(\varepsilon) = 1 - \zeta(\varepsilon)q,$$
where $q = \alpha\bar{p} + \bar{\beta}p$ and
$$\zeta(\varepsilon) := \frac{\bar{\alpha}\bar{p} + \bar{\beta}p - \varepsilon}{\bar{\beta}p - \alpha\bar{p}}. \tag{3}$$
Furthermore, $\hbar(\varepsilon)$ is achieved by the Z-channel $\mathsf{Z}(\zeta(\varepsilon))$ (as shown in Fig. 2).
- *If $\alpha\bar{\alpha}\bar{p}^2 \geq \beta\bar{\beta}p^2$, then*
$$\hbar(\varepsilon) = 1 - \tilde{\zeta}(\varepsilon)\bar{q},$$
where
$$\tilde{\zeta}(\varepsilon) := \frac{\bar{\alpha}\bar{p} + \bar{\beta}p - \varepsilon}{\bar{\alpha}\bar{p} - \beta p}.$$
Moreover, $\hbar(\varepsilon)$ is achieved by a reverse Z-channel with crossover probability $\tilde{\zeta}(\varepsilon)$ (as shown in Fig. 2).

Proof Sketch. Recall that $\varepsilon \mapsto \hbar(\varepsilon)$ is concave, and thus its graph lies above the segment connecting $(p, \hbar(p))$ to $(\mathsf{P_c}(X|Y), 1)$. In particular,
$$\hbar(\varepsilon) \geq \hbar(p) + (\varepsilon - p)\left[\frac{1 - \hbar(p)}{\mathsf{P_c}(X|Y) - p}\right].$$
By Lemma 2, the above inequality becomes
$$\begin{aligned}\hbar(\varepsilon) \quad \geq \quad & \hbar(p) + \frac{q(\varepsilon - p)}{\bar{\beta}p - \alpha\bar{p}} 1_{\{\alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2\}} \\ & + \frac{\bar{q}(\varepsilon - p)}{\bar{\alpha}\bar{p} - \beta p} 1_{\{\alpha\bar{\alpha}\bar{p}^2 \geq \beta\bar{\beta}p^2\}}.\end{aligned} \tag{4}$$
Since $\varepsilon \mapsto \hbar(\varepsilon)$ is piecewise linear, its right derivative exists at $\varepsilon = \mathsf{P_c}(X|Y)$. Using the geometric properties of $\mathcal{H}$ used to prove Theorem 1, we can show that
$$\begin{aligned}\hbar'(\mathsf{P_c}(X|Y)) \quad = \quad & \frac{q}{\bar{\beta}p - \alpha\bar{p}} 1_{\{\alpha\bar{\alpha}\bar{p}^2 < \beta\bar{\beta}p^2\}} \\ & + \frac{\bar{q}}{\bar{\alpha}\bar{p} - \beta p} 1_{\{\alpha\bar{\alpha}\bar{p}^2 \geq \beta\bar{\beta}p^2\}},\end{aligned}$$
which is equal to the slope of the chord connecting $(p, \hbar(p))$ to $(\mathsf{P_c}(X|Y), 1)$ described in (4). The concavity of $\hbar(\cdot)$ thus implies that the inequality (4) is indeed equality. ∎

Under the hypotheses of the previous theorem, for every $\varepsilon \in [\mathsf{P_c}(X), \mathsf{P_c}(X|Y)]$ there exists a Z-channel that achieves $\hbar(\varepsilon)$. It can be shown that Z-channel is the *only* binary filter with this property. It is also worth mentioning that even if $P_{Y|X}$ is symmetric (i.e., $\alpha = \beta$), the optimal filter cannot be symmetric, unless X is uniform, in which case $\mathsf{BSC}(0.5\zeta(\varepsilon))$ is also optimal.

III. I.I.D. Binary Symmetric Vector Case

We next study privacy aware guessing for a pair of binary random *vectors* (X^n, Y^n) with $X^n, Y^n \in \{0, 1\}^n$. Recall that in this case it is sufficient to consider auxiliary random variables having supports of cardinality $2^n + 1$. However, this condition may be practically inconvenient. Moreover, in the scalar binary case examined in the last section we observed that a binary Z was sufficient to achieve $\hbar(\varepsilon)$. Hence, it is natural to require the privacy filters to produce also binary random vectors, i.e., $Z^n \in \{0, 1\}^n$, which leads to the following definition. Recall that the data processing inequality implies that $\mathsf{P_c}(X^n) \leq \mathsf{P_c}(X^n|Z^n) \leq \mathsf{P_c}(X^n|Y^n)$ and hence we can assume $\mathsf{P_c}(X^n) \leq \varepsilon^n \leq \mathsf{P_c}(X^n|Y^n)$.

Definition 2. *For a given pair of binary random vectors (X^n, Y^n), we define $\underline{h}_n(\varepsilon)$ for $\varepsilon \in [\mathsf{P_c}^{1/n}(X^n), \mathsf{P_c}^{1/n}(X^n|Y^n)]$, as*
$$\underline{h}_n(\varepsilon) := \max\ \mathsf{P_c}^{1/n}(Y^n|Z^n), \tag{5}$$
where the maximum is taken over all (not necessarily memoryless) channels $P_{Z^n|Y^n}$ such that $Z^n \in \{0, 1\}^n$, $X^n \multimap Y^n \multimap Z^n$, and $\mathsf{P_c}(X^n|Z^n) \leq \varepsilon^n$.

Note that this definition does not make any assumption about the privacy filters $P_{Z^n|Y^n}$ except that $Z^n \in \{0, 1\}^n$. From an implementation point of view, the simplest privacy filter is a memoryless one such that Z_k is a noisy version of Y_k for $k = 1, \ldots, n$. More precisely, we are interested in a *single* BIBO channel $P_{Z|Y}$ which, given Y_k, generates Z_k according to
$$P_{Z^n|Y^n}(z^n|y^n) = \prod_{k=1}^{n} P_{Z|Y}(z_k|y_k).$$
Now, let $h_n^{\mathsf{i}}(\varepsilon)$ be defined as $\max \mathsf{P_c}^{1/n}(Y^n|Z^n)$, where the maximum is taken over such memoryless privacy filters satisfying $\mathsf{P_c}(X^n|Z^n) \leq \varepsilon^n$. Let $\oplus$ denote mod 2 addition. In what follows, we study $\underline{h}_n$ and h_n^{i} for the following setup:

a) $X_1, \ldots, X_n$ are i.i.d. $\mathsf{Bernoulli}(p)$ random variables with $p \geq \frac{1}{2}$,

b) $Y_k = X_k \oplus V_k$ for $k = 1, \ldots, n$, where $V_1, \ldots, V_n$ are i.i.d. $\mathsf{Bernoulli}(\alpha)$ random variables independent of X^n, such that $\alpha < \frac{1}{2}$.

We first determine $h_n^{\mathsf{i}}(\varepsilon)$ for this model and show that (as expected) $h_n^{\mathsf{i}}(\varepsilon)$ is independent of n. According to this model, $\mathsf{P_c}(X^n) = p^n$ and $\mathsf{P_c}(X^n|Y^n) = \bar{\alpha}^n$, and thus $p \leq \varepsilon \leq \bar{\alpha}$.

Proposition 1. *If (X^n, Y^n) satisfies a) and b) with $p \in [\frac{1}{2}, 1)$ and $\alpha \in [0, \frac{1}{2})$ such that $\bar{\alpha} > p$, then*
$$h_n^{\mathsf{i}}(\varepsilon) = \hbar(\varepsilon) = 1 - \zeta(\varepsilon)q,$$
for all $\varepsilon \in [p, \bar{\alpha}]$, where $\zeta(\varepsilon)$ is given in (3) and $q = \alpha\bar{p} + \bar{\alpha}p$.

Note that the proposition reduces to Theorem 2 for $n = 1$. However, for $n \geq 2$, we have $h_n^{\mathsf{i}}(\varepsilon) < \underline{h}_n(\varepsilon) \leq$

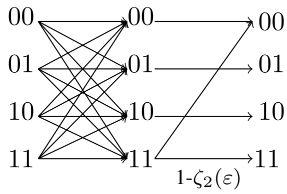


Fig. 3. The optimal privacy filter for $\underline{h}_2(\varepsilon)$ for $\varepsilon \in [\varepsilon_{\mathsf{L}}, \bar{\alpha})$, where $\zeta_2(\varepsilon)$ is defined in (6).

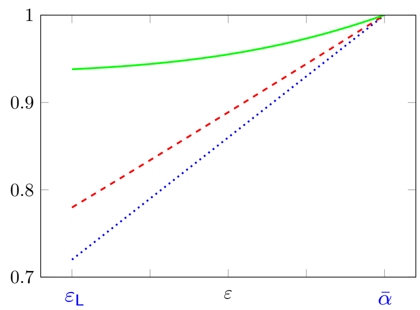


Fig. 4. The graphs of $\underline{h}_{10}$ (solid curve), $\underline{h}_2$ (dashed curve), and h^{i} (dotted line) given in Theorem 3 and Proposition 1 for i.i.d. (X^n, Y^n) with $X \sim \mathsf{Bernoulli}(0.6)$ and $P_{Y|X} = \mathsf{BSC}(0.2)$.

$\hbar(P_{X^nY^n}, \varepsilon)$, as implied by the following theorem. A channel W is said to be a 2^n-ary Z-channel, denoted by $\mathsf{Z}_n(\gamma)$, if the input and output alphabets are $\{0,1\}^n$ and $\mathsf{W}(a|a) = 1$ for $a \neq \mathbf{1}$, $\mathsf{W}(\mathbf{0}|\mathbf{1}) = \gamma$, and $\mathsf{W}(\mathbf{1}|\mathbf{1}) = \bar{\gamma}$, where $\mathbf{0} = (0, 0, \dots, 0)$ and $\mathbf{1} = (1, 1, \dots, 1)$.

Theorem 3. *Assume that (X^n, Y^n) satisfies a) and b) with $p \in [\frac{1}{2}, 1)$ and $\alpha \in [0, \frac{1}{2})$ such that $\bar{\alpha} > p$. Then, there exists $p \leq \varepsilon_{\mathsf{L}} < \bar{\alpha}$ such that*

$$\underline{h}_n^n(\varepsilon) = 1 - \zeta_n(\varepsilon) q^n,$$

for $\varepsilon \in [\varepsilon_{\mathsf{L}}, \bar{\alpha}]$, where $q = \alpha\bar{p} + \bar{\alpha}p$ and

$$\zeta_n(\varepsilon) := \frac{\bar{\alpha}^n - \varepsilon^n}{(\bar{\alpha}p)^n - (\alpha\bar{p})^n}. \tag{6}$$

Moreover, the channel $\mathsf{Z}_n(\zeta_n(\varepsilon))$ achieves $\underline{h}_n(\varepsilon)$ in this interval (see Fig. 3 for the case $n = 2$).

The memoryless privacy filter assumed in $h_n^{\mathsf{i}}(\varepsilon)$ is simple to implement. However, it is clear from Theorem 3 that this simple filter is not optimal even when (X^n, Y^n) is i.i.d. since $\underline{h}_n(\varepsilon)$ is a function of n, while $h_n^{\mathsf{i}}(\varepsilon)$ is not. The following corollary bounds the loss resulting from using a simple memoryless filter instead of an optimal one for $\varepsilon \in [\varepsilon_{\mathsf{L}}, \bar{\alpha}]$. Clearly, for $n = 1$, there is no gap as $\underline{h}_1(\varepsilon) = h_1^{\mathsf{i}}(\varepsilon)$.

Corollary 2. *Let (X^n, Y^n) satisfy a) and b) with $p \in [\frac{1}{2}, 1)$ and $\alpha \in [0, \frac{1}{2})$ such that $\bar{\alpha} > p$. If $p > \frac{1}{2}$ and $\alpha > 0$, then for $\varepsilon \in [\varepsilon_{\mathsf{L}}, \bar{\alpha}]$ and sufficiently large n*

$$\underline{h}_n(\varepsilon) - h_n^{\mathsf{i}}(\varepsilon) \geq (\bar{\alpha} - \varepsilon)[\Phi(1) - \Phi(n)], \tag{7}$$

where

$$\Phi(n) := \frac{q^n \bar{\alpha}^{n-1}}{(\bar{\alpha}p)^n - (\alpha\bar{p})^n}.$$

If $p = \frac{1}{2}$, then

$$h_n^{\mathsf{i}}(\varepsilon) \leq \underline{h}_n(\varepsilon) \leq h_n^{\mathsf{i}}(\varepsilon) + \frac{\alpha}{2\bar{\alpha}}, \tag{8}$$

for every $n \geq 1$ and $\varepsilon \in [\varepsilon_{\mathsf{L}}, \bar{\alpha}]$.

Since $\Phi(n) \downarrow 0$ as $n \to \infty$, (7) implies that, as expected, the gap between the performance of the optimal privacy filter and the optimal memoryless privacy filter increases as n increases. This observation is numerically illustrated in Fig. 4, where $\underline{h}_n(\varepsilon)$ is plotted as a function of ε for $n = 2$ and $n = 10$. Moreover, (8) implies that when $p = \frac{1}{2}$ and α is small, then $\underline{h}_n(\varepsilon)$ can be approximated by $h_n^{\mathsf{i}}(\varepsilon)$. Thus, we can approximate the optimal filter $\mathsf{Z}_n(\zeta_n(\varepsilon))$ with a simple memoryless filter given by $Z_k = Y_k \oplus W_k$, where $W_1, \dots, W_n$ are i.i.d. $\mathsf{Bernoulli}(0.5\zeta(\varepsilon))$ random variables that are independent of (X^n, Y^n).

REFERENCES

[1] H. Yamamoto, "A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers," *IEEE Trans. Inf. Theory*, vol. 29, no. 6, pp. 918–923, Nov. 1983.

[2] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Information extraction under privacy constraints," *Information*, vol. 7, 2016. [Online]. Available: http://www.mdpi.com/2078-2489/7/1/15

[3] S. Asoodeh, F. Alajaji, and T. Linder, "Notes on information-theoretic privacy," in *Proc. 52nd Annual Allerton Conf. Com, Control, and Computing*, Sept. 2014, pp. 1272–1278.

[4] F. P. Calmon, A. Makhdoumi, and M. Médard, "Fundamental limits of perfect privacy," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2015, pp. 1796–1800.

[5] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Médard, "From the information bottleneck to the privacy funnel," in *Proc. IEEE Inf. Theory Workshop (ITW)*, 2014, pp. 501–505.

[6] C. T. Li and A. E. Gamal, "Extended Gray-Wyner system with complementary causal side information," 2017. [Online]. Available: arXiv:1701.03207v1

[7] I. Issa, S. Kamath, and A. B. Wagner, "An operational measure of information leakage," in *Proc. Annual Conference on Information Science and Systems (CISS)*, March 2016, pp. 234–239.

[8] A. Makhdoumi and N. Fawaz, "Privacy-utility tradeoff under statistical uncertainty," in *Proc. 51st Annual Allerton Conf. Comm, Control, and Computing*, Oct 2013, pp. 1627–1634.

[9] H. Gebelein, "Das statistische problem der korrelation als variations- und eigenwert-problem und sein zusammenhang mit der ausgleichungsrechnung," *Zeitschrift fur angew. Math. und Mech.*, no. 21, pp. 364–379, 1941.

[10] F. P. Calmon, M. Varia, M. Médard, M. M. Christiansen, K. R. Duffy, and S. Tessaro, "Bounds on inference," in *Proc. 51st Annual Allerton Conf. Comm, Control, and Computing*, Oct 2013, pp. 567–574.

[11] S. Asoodeh, F. Alajaji, and T. Linder, "Privacy-aware MMSE estimation," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, July 2016, pp. 1989–1993.

[12] H. S. Witsenhausen, "On sequence of pairs of dependent random variables," *SIAM Journal on Applied Mathematics*, vol. 28, no. 2, pp. 100–113, 1975.

[13] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Estimation efficiency under privacy constraints," *To be submitted.*

[14] S. Verdú, "α-mutual information," in *Proc. Inf. Theory and Applications Workshop (ITA), 2015*, Feb. 2015, pp. 1–6.

[15] H. Witsenhausen and A. Wyner, "A conditional entropy bound for a pair of discrete random variables," *IEEE Trans. Inf. Theory*, vol. 21, no. 5, pp. 493–501, Sep. 1975.

[16] S. W. Ho and S. Verdú, "Convexity/concavity of Rényi entropy and α-mutual information," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2015, pp. 745–749.

[17] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Cambridge University Press, 2011.