What can Information Guess? Guessing Advantage vs. Rényi Entropy for Small Leakages

Julien Béguinot and Olivier Rioul

LTCI, Télécom Paris, Institut Polytechnique de Paris, France *firstname.lastname*@telecom-paris.fr

Abstract—We leverage the Gibbs inequality and its natural generalization to Rényi entropies to derive closed-form parametric expressions of the optimal lower bounds of ρth-order guessing entropy (guessing moment) of a secret taking values on a finite set, in terms of the Rényi-Arimoto α-entropy. This is carried out in an non-asymptotic regime when side information may be available. The resulting bounds yield a theoretical solution to a fundamental problem in side-channel analysis: Ensure that an adversary will not gain much guessing advantage when the leakage information is sufficiently weakened by proper countermeasures in a given cryptographic implementation. Practical evaluation for classical leakage models show that the proposed bounds greatly improve previous ones for analyzing the capability of an adversary to perform side-channel attacks.

I. INTRODUCTION

Guessing entropy [1], also known as *guesswork* [2], is perhaps the most popular security metric in the context of side-channel attacks of embedded cryptographic devices, such as the cryptographic microcontrollers used in banking smartcards [3]–[5]. Such attacks exploit *leakage* information to recover the secret key, byte by byte in a divide-and-conquer strategy, in which each subkey byte $K \in \{1, \ldots, M\}$ (typically $M = 128$ or 256) is targeted independently of the others. The secret key K is generally assumed uniformly distributed, but in a more general framework, any type of secret $X \in \{1, \ldots, M\}$ (e.g., passwords, sensitive personal information, etc.) can be targeted from some disclosed side information Y.

Guessing entropy [1] $G(X|Y)$, or more generally, ρth order guessing moments [6] $G_\rho(X|Y)$, relate to the number of tries that the attacker has to make to find the actual secret X for a given leakage side information Y, thereby estimating the brute force effort to find X by exhaustive search. The popularity of G_ρ (particularly $G = G_1$) as a security criterion comes from the fact that it is particularly informative, as it is computed from whole key ranking distribution [7]. The adversary's *guessing advantage* is defined as

$$\Delta G_\rho(X; Y) \triangleq G_\rho(M) - G_\rho(X|Y) \tag{1}$$

where $G_\rho(M) = G_\rho(K)$ is for a blind estimation (no leakage) of a uniformly distributed M-ary secret K.

Side channel leakage was evaluated by information theoretic measures such as mutual information $I(X; Y)$ [3], [8]–[10], maximal leakage [11], [12], and more generally, Arimoto's α-information [13] $I_\alpha(X; Y) = H_\alpha(X) - H_\alpha(X|Y)$ or Sibson's α-information [14]–[17]. The latter two α-informations coincide for uniform secrets: $I_\alpha(K; Y) = \log M - H_\alpha(K|Y)$ [18].

In many practical cases, such as for protected implementations with masking or low SNR noise [10], [12], [19]–[21], the information leakage is *small*, which means that the conditional Rényi-Arimoto entropy [22] $H_\alpha(X|Y)$ approaches its maximum value $\log M$ in the case of uniform secrets. Thus, to evaluate the impact of leakage in cryptographic implementations, a significant quantity is the "*information advantage*"

$$\Delta H_\alpha(X; Y) \triangleq \log M - H_\alpha(X|Y). \tag{2}$$

A fundamental problem in side-channel analysis is to ensure that an adversary will not gain much guessing advantage when the leakage is sufficiently weakened by proper countermeasures in some cryptographic implementation [4], [5]. In other words, it is important to upper bound $\Delta G_\rho(X; Y)$ in terms of $\Delta H_\alpha(X; Y)$, or equivalently, to lower bound $G_\rho(X|Y)$ in terms of $H_\alpha(X|Y)$. As explained in [4], [5], this is also useful to practically evaluate guessing advantage for a full key (e.g., 256-bit key) whose direct evaluation is not tractable.

Many such bounds have been derived in the literature. Massey's original inequality [1] is for $\alpha = \rho = 1$ and was improved by Rioul [23] as an asymptotically optimal inequality [5] as $M \to +\infty$. Arikan's inequalities [6] are for $\alpha = \frac{1}{1+\rho}$ and exhibit asymptotic equivalence. More general bounds for various ranges of α and ρ were established in [23]. These works, however, can only be optimal as $M \to +\infty$ and do not consider the nonasymptotic scenario for small leakage, i.e., around the corner $(H_\alpha(X|Y) \approx \log M, G_\rho(X|Y) \approx G_\rho(M))$ for relatively small M.
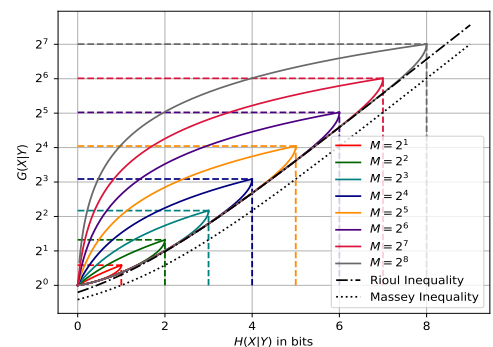

Fig. 1: Optimal joint range region between $H(X|Y)$ and $G(X|Y)$ for different values of M. The (optimal) *upper* bound is that of Mc Eliece and Yu [24]. The optimal lower bound is derived in this paper. The black dotted and dash-dotted curves correspond to Massey's [1] and Rioul's [23] inequalities, that do not depend on M.

As an illustration for $\alpha = \rho = 1$, Figure 1 shows that for any fixed value of M, there is a multiplicative gap of approximately $e/2$ in guessing entropy from this corner point $(\log M, \frac{M+1}{2})$ compared to Rioul's inequality. Sason and Verdú [25] improved Arikan's inequalities in a non-asymptotic regime for any ranking function but did not obtain the exact locus of attainable values of G_ρ vs. H_α, nor closed-form expressions for lower bounds of G_ρ vs. H_α.

In this paper, we leverage the Gibbs inequality [26] and its natural generalization to Rényi entropies [27] to derive closed-form parametric expressions of the optimal lower bounds of $G_\rho(X|Y)$ vs. $H_\alpha(X|Y)$ (or upper bounds of $\Delta G_\rho(X;Y)$ vs. $\Delta H_\alpha(X;Y)$). We obtain an explicit, easily computable, first-order bound of the form $\Delta G_\rho(X;Y) \leqslant c\sqrt{\Delta H_\alpha(X;Y)}$ for small leakages. We then evaluate and refine our bounds in the Hamming weight and random probing leakage models.

II. NOTATIONS

The secret is modeled as a random variable X with pmf p_X taking values in a finite set $\{1, \ldots, M\}$. The side channel $X \to Y$ leaks some side information Y, which is modeled as an arbitrary random variable (discrete or continuous). For most cryptographic applications, X is a uniformly distributed key $K \sim \mathcal{U}(M)$. In this case the side channel is noted $K \to Y$.

We let $\alpha > 0$ and $\rho > 0$ be the entropy and guessing entropy orders, respectively. The Rényi entropy of order α is $H_\alpha(X) \triangleq -\alpha' \log \|p_X\|_\alpha$ where $\|p\|_\alpha \triangleq (\sum p^\alpha)^{1/\alpha}$ denotes the "α-norm", the logarithm is taken to an arbitrary base, and α' is the Hölder conjugate of α: $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$, i.e. $\alpha' = \frac{\alpha}{\alpha-1}$ or $(\alpha'-1)(\alpha-1) = 1$. We also write $H_\alpha(p) = -\alpha' \log \|p\|_\alpha$. The limiting case $\alpha \to 1$ is the Shannon entropy $H(X) = H_1(X)$. The limiting case $\alpha \to \infty$ is the min-entropy $H_\infty(X) \triangleq -\log \max p_X$.

The ρ-guessing entropy (a.k.a. guessing ρth-order moment) is $G_\rho(X) \triangleq \min_\sigma \sum_{i=1}^M p_X(\sigma(i)) i^\rho = \min_\sigma \mathbb{E}(\sigma(X)^\rho)$, where the minimum is over all possible permutations of the secret values. We also write $G_\rho(p)$ when X has pmf p. By the rearrangement inequality, since the sequence i^ρ is increasing, the optimal guessing strategy is such that the $p_X(\sigma(i))$ are arranged in decreasing order. When $X = K$ is uniformly distributed, its ρ-guessing entropy is also noted $G_\rho(M) \triangleq \frac{1}{M} \sum_{i=1}^M i^\rho$.

In the presence of leakage Y, the most natural definition of conditional α-entropy is that of Arimoto [13], [22]: $H_\alpha(X|Y) \triangleq -\alpha' \log \mathbb{E}_Y \|p_{X|Y}\|_\alpha$. In particular $H_\infty(X|Y) \triangleq -\log \mathbb{E}_Y \max p_{X|Y}$ where $\mathbb{E}_Y \max p_{X|Y}$ is the optimal success probability as given by the MAP rule. Following Hirche [28] we also define

$$K_\alpha(X) \triangleq \|p_X\|_\alpha = \exp\left(\frac{1-\alpha}{\alpha} H_\alpha(X)\right) \qquad (3)$$

$$K_\alpha(X|Y) \triangleq \mathbb{E}_Y \|p_{X|Y}\|_\alpha = \exp\left(\frac{1-\alpha}{\alpha} H_\alpha(X|Y)\right) \qquad (4)$$

so that $K_\alpha(X|Y) = \mathbb{E}_y K_\alpha(X|Y = y)$. Likewise, the conditional ρ-guessing entropy is $G_\rho(X|Y) \triangleq \mathbb{E}_y G_\rho(X|Y = y)$. In particular for $\rho = 1$, $G(X|Y) \triangleq \mathbb{E}_y G(X|Y = y)$.

III. OPTIMAL LOWER BOUND ON $G_\rho(X|Y)$ vs. $H_\alpha(X|Y)$

A. Case $\rho = 1$: $G(X|Y)$ vs. $H(X|Y)$

Theorem 1: The optimal lower bound on $G(X|Y)$ vs. $H(X|Y)$ is given by the parametric curve for $\gamma \in (0,1)$:

$$\begin{cases} G(X|Y) = \frac{1}{1-\gamma} - \frac{M\gamma^M}{1-\gamma^M} \\ H(X|Y) = \log(\gamma \frac{1-\gamma^M}{1-\gamma}) - (\log \gamma)(\frac{1}{1-\gamma} - \frac{M\gamma^M}{1-\gamma^M}) \end{cases} \qquad (5)$$

where the limiting case $\gamma \to 1$ gives $G = \frac{M+1}{2}$ and $H = \log M$ attained for the uniform distribution.

The optimal upper bound on $\Delta G(X;Y) = \frac{M+1}{2} - G(X|Y)$ vs. $\Delta H(X;Y) = \log M - H(X|Y)$ is given by the parametric curve for $\mu \in (0, +\infty)$:

$$\begin{cases} \Delta G(X;Y) = \frac{1}{2}(M \coth(M\mu) - \coth(\mu)) \\ \Delta H(X;Y) = \log \frac{M \sinh \mu}{\sinh(M\mu)} + 2\mu(\log e)\Delta G(X;Y). \end{cases} \qquad (6)$$

Proof: Here $\alpha = \rho = 1$ is the classical situation studied by Massey [1], where the essential ingredient is the *Gibbs inequality* [26]: In the unconditional case, for any pmf q,

$$H(X) \leqslant -\mathbb{E}_X \log q(X) \qquad (7)$$

with equality iff $q = p_X$. One may always assume that $p_X(x)$ is nonincreasing in x, in which case $G(X) = \mathbb{E}(X)$. Therefore, we choose q such that $\log q(x) = a + bx$ for real constants a, b, so that $\mathbb{E}_X \log q(X) = a + bG(X)$. To allow equality in the Gibbs inequality ($q = p_X$), it is necessary that $q(x)$ be nonincreasing in x, i.e., $b \leqslant 0$. Now $q(x)$ rewrites as a truncated geometric pmf $q(x) = c_\gamma \gamma^x$ where $\gamma = \exp(b) \in (0,1]$ and $c_\gamma = (\sum_{i=1}^M \gamma^i)^{-1} > 0$ is a normalization factor. Thus one obtains

$$H(X) \leqslant -\log c_\gamma - (\log \gamma)G(X). \qquad (8)$$

for all $\gamma \in (0,1)$, with equality iff $p_X = q$ where $H(X) = H(q)$ and $G(X) = G(q)$.

In the conditional case, we similarly have $H(X|Y = y) \leqslant -\log c_\gamma - (\log \gamma)G(X|Y = y)$ for every y. Taking the expectation over Y yields the same inequality for conditioned entropies:

$$H(X|Y) \leqslant -\log c_\gamma - (\log \gamma)G(X|Y), \qquad (9)$$

that is, $G(X|Y) \geqslant -(\log \gamma)^{-1}(H(X|Y) + \log c_\gamma)$. The equality case still corresponds to $H(X|Y) = H(q)$ and $G(X|Y) = G(q)$. Since q approaches the uniform distribution as $\gamma \to 1$ and the Dirac distribution as $\gamma \to 0^+$, all possible values of entropies are attainable.

We thus obtain the following parameterization of the optimal lower bound on $G(X|Y)$ vs. $H(X|Y)$:

$$\begin{cases} G(X|Y) = G(q) = c_\gamma \sum_{i=1}^M i\gamma^i \\ H(X|Y) = H(q) = -\log c_\gamma - G(q) \log \gamma \end{cases} \qquad (10)$$

where $\gamma \in (0,1]$. The case $\gamma = 1$ gives $H = \log M$ and $G = \frac{M+1}{2}$ attained for the uniform distribution. For $0 < \gamma < 1$, a straightforward calculation gives (5). Setting $\mu \triangleq -\frac{1}{2} \ln \gamma \in (0, \infty)$, i.e., $\gamma = e^{-2\mu}$ gives (6) for $\Delta G(X;Y) = \frac{M+1}{2} - G(X|Y)$ and $\Delta H(X;Y) = \log M - H(X|Y)$. $\blacksquare$

B. $G_\rho(X|Y)$ vs. $H(X|Y)$

Theorem 2: The optimal lower bound of $G_\rho(X|Y)$ vs. $H(X|Y)$ is given by the parametric curve for $\gamma \in (0, 1]$:

$$\begin{cases} G_\rho(X|Y) = (\sum_{i=1}^{M} i^\rho \gamma^{i^\rho})(\sum_{i=1}^{M} \gamma^{i^\rho})^{-1} \\ H(X|Y) = \log(\sum_{i=1}^{M} \gamma^{i^\rho}) - (\log \gamma)\frac{\sum_{i=1}^{M} i^\rho \gamma^{i^\rho}}{\sum_{i=1}^{M} \gamma^{i^\rho}} \end{cases} \quad (11)$$

Proof: The proof is analog to the case of the preceding subsection, where $\log q(x) = a + bx^\rho$, $q(x) = c_\gamma \gamma^{x^\rho}$, $\gamma = \exp b \in (0, 1]$ and $c_\gamma = (\sum_{i=1}^{M} \gamma^{i^\rho})^{-1} > 0$. Again q approaches the uniform distribution as $\gamma \to 1$ and the Dirac distribution as $\gamma \to 0^+$, hence all possible values of entropies are attainable. One readily obtains

$$\begin{cases} G_\rho(X|Y) = G_\rho(q) = c_\gamma \sum_{i=1}^{M} i^\rho \gamma^{i^\rho} \\ H(X|Y) = H(q) = -\log c_\gamma - G_\rho(q) \log \gamma \end{cases} \quad (12)$$

which gives (11). ∎

C. $G_\rho(X|Y)$ vs. $H_\alpha(X|Y)$

The most general optimal lower bound is given by the following Theorem. It generalizes Theorems 1 and 2, which can be recovered in the limiting case $\alpha \to 1$.

Theorem 3: When $0 < \alpha < 1$, the optimal lower bound of $G_\rho(X|Y)$ vs. $H_\alpha(X|Y)$ is given by the parametric curve for $\gamma \in (0, \infty)$:

$$\begin{cases} G_\rho(X|Y) = 1 + \gamma^{-1}\big(\frac{\sum_{i=1}^{M}(1-\gamma+\gamma i^\rho)^{\alpha'}}{\sum_{i=1}^{M}(1-\gamma+\gamma i^\rho)^{\alpha'-1}} - 1\big) \\ H_\alpha(X|Y) = \alpha' \log \sum_{i=1}^{M}(1-\gamma+\gamma i^\rho)^{\alpha'-1} \\ \qquad\qquad + (1-\alpha')\log \sum_{i=1}^{M}(1-\gamma+\gamma i^\rho)^{\alpha'}. \end{cases} \quad (13)$$

When $\alpha > 1$, the optimal lower bound of $G_\rho(X|Y)$ in terms of $H_\alpha(X|Y)$ is given by the parametric curve for $\gamma \in (0, 1)$:

$$\begin{cases} G_\rho(X|Y) = \gamma^{-1}\big(1 - \frac{\sum_{i=1}^{M}(1-\gamma i^\rho)_+^{\alpha'}}{\sum_{i=1}^{M}(1-\gamma i^\rho)_+^{\alpha'-1}}\big) \\ H_\alpha(X|Y) = \alpha' \log \sum_{i=1}^{M}(1-\gamma i^\rho)_+^{\alpha'-1} \\ \qquad\qquad + (1-\alpha')\log \sum_{i=1}^{M}(1-\gamma i^\rho)_+^{\alpha'} \end{cases} \quad (14)$$

where $x_+ = \max(x, 0)$ denotes the positive part of x.

The proof relies on the following Gibbs inequality for Rényi entropies [27, Prop. 8]:

Lemma 1 (Generalized Gibbs Inequality): For any pmf q,

$$H_\alpha(X) \lesseqgtr -\alpha' \log \mathbb{E}_X q_\alpha^{1/\alpha'}(X) \quad (15)$$

with equality iff $p_X = q$. Here q_α is the *escort distribution* [27] of q, defined by $q_\alpha(x) = q^\alpha(x)/\|q\|_\alpha^\alpha$.
The proof given [27] was for random variables having pdfs with respect to the Lebesgue measure, but applies verbatim to discrete random variables having pmfs with respect to the counting measure on $\{1, 2, \ldots, M\}$. The Gibbs inequality (15) can be easily rewritten directly in terms of $q(x)$ as

$$H_\alpha(X) \leqslant (1-\alpha)H_\alpha(q) - \alpha' \log \mathbb{E}_X q^{\alpha-1}(X). \quad (16)$$

In terms of (3) it also rewrites

$$K_\alpha(X) \lesseqgtr \mathbb{E}_X q_\alpha^{1/\alpha'}(X) = \frac{\mathbb{E}_X q^{\alpha-1}(X)}{\|q\|_\alpha^{\alpha-1}} \quad (17)$$

where $\lesseqgtr$ denotes $\geqslant$ for $0 < \alpha < 1$ and $\leqslant$ for $\alpha > 1$.

Proof of Theorem 3: First consider the unconditional case. One may always assume that $p_X(x)$ is nonincreasing in x, in which case $G_\rho(X) = \mathbb{E}(X^\rho)$. Therefore, we wish to choose q such that $q^{\alpha-1}(x) = a + bx^\rho$ for real constants a, b, i.e., $q(x) = (a + bx^\rho)^{\alpha'-1}$, so that $\mathbb{E}_X q^{\alpha-1}(X) = a + b\,G_\rho(X)$.

When $0 < \alpha < 1$, $\alpha' < 0$, to allow equality in the Gibbs inequality ($p_X = q$), it is necessary that $a + bx^\rho \geqslant 0$ with nonempty support and that $(a + bx^\rho)^{\alpha'-1}$ is nonincreasing for $x = 1, 2, \ldots, M$. This gives the conditions $b \geqslant 0$ and $a > -b$. Rewriting $a + bx^\rho = (a + b) + b(x^\rho - 1)$ we obtain $q(x) = c_\gamma(1 + \gamma(x^\rho - 1))^{\alpha'-1}$ where $\gamma \triangleq \frac{b}{a+b} \in [0, +\infty)$ and $c_\gamma = (\sum_{i=1}^{M}(1 + \gamma(i^\rho - 1))^{\alpha'-1})^{-1} > 0$ is a normalization factor. Note that q is the uniform distribution when $\gamma = 0$ and approaches the Dirac distribution as $\gamma \to +\infty$, hence all possible values of entropies are attainable.

When $\alpha > 1$, $\alpha' > 0$, we choose the positive part $q(x) = (a + bx^\rho)_+^{\alpha'-1}$. Again to allow equality in the Gibbs inequality ($p_X = q$), it is necessary $q(x)$ has nonempty support and is nonincreasing for $x = 1, 2, \ldots, M$. This gives the conditions $b \leqslant 0$ and $a > -b$. Factoring out $a > 0$ we obtain $q(x) = c_\gamma(1 - \gamma x^\rho)_+^{\alpha'-1}$ where $\gamma \triangleq \frac{-b}{a} \in (0, 1)$ and $c_\gamma = (\sum_{i=1}^{M}(1 - \gamma x^\rho)_+^{\alpha'-1})^{-1} > 0$ is a normalization factor. Note that q approaches the uniform distribution when $\gamma \to 0$ and the Dirac distribution when $\gamma \to 1$—in fact, it is the Dirac distribution for all $\gamma \in [2^{-\rho}, 1)$. Hence all possible values of entropies are again attainable.

In both cases, the Gibbs inequality (17) takes the form[1] $K_\alpha(X) \lesseqgtr \varphi\big(G_\rho(X)\big)$ for some linear function φ, with equality iff $p_X = q$, in which case $H_\alpha(X) = H_\alpha(q)$ and $G_\rho(X) = G_\rho(q)$. In the conditional case, we similarly have $K_\alpha(X|Y = y) \lesseqgtr \varphi\big(G_\rho(X|Y = y)\big)$ for every y, and taking the expectation over Y yields the same inequality for conditioned entropies: $K_\alpha(X|Y) \lesseqgtr \varphi\big(G_\rho(X|Y)\big)$. The equality case still corresponds to $H_\alpha(X|Y) = H_\alpha(q)$ and $G_\rho(X|Y) = G_\rho(q)$.

When $0 < \alpha < 1$, $\mathbb{E}_X q^{\alpha-1}(X) = c_\gamma^{\alpha-1}(1 + \gamma(G_\rho(X) - 1))$ and the equality case of the Gibbs inequality (16) rewrites $H_\alpha(q) = -\frac{\alpha'}{\alpha}\log \mathbb{E}_X q^{\alpha-1}(X) = -\frac{\alpha'}{\alpha}\log\big(1 - \gamma + \gamma G_\rho(q)\big) - \log c_\gamma$. Since by definition $H_\alpha(q) = -\alpha'\log\|q\|_\alpha$ we obtain the parametric curve

$$\begin{cases} G_\rho(X|Y) = \gamma^{-1}(\|q\|_\alpha^\alpha c_\gamma^{1-\alpha} - 1) + 1 \\ H_\alpha(X|Y) = -\alpha' \log \|q\|_\alpha. \end{cases} \quad (18)$$

Substituting $\|q\|_\alpha^\alpha = c_\gamma^\alpha \sum_{i=1}^{M}(1 + \gamma(i^\rho - 1))^{\alpha'}$ gives (13).

When $\alpha > 1$, $\mathbb{E}_X q^{\alpha-1}(X) = c_\gamma^{\alpha-1}(1 - \gamma G_\rho(X))$, and the equality case of the Gibbs inequality (16) rewrites $H_\alpha(q) = -\frac{\alpha'}{\alpha}\log \mathbb{E}_X q^{\alpha-1}(X) = -\frac{\alpha'}{\alpha}\log(1 - \gamma G_\rho(q)) - \log c_\gamma$. We similarly obtain the parametric curve

$$\begin{cases} G_\rho(X|Y) = \gamma^{-1}(1 - \|q\|_\alpha^\alpha c_\gamma^{1-\alpha}) \\ H_\alpha(X|Y) = -\alpha' \log \|q\|_\alpha \end{cases} \quad (19)$$

---

[1]For $\alpha > 1$, the Gibbs inequality takes this form because when $p_X = q$, these pmfs have the same support, so that equality also holds in the inequality $\mathbb{E}_X(a + bX^\rho)_+ \geqslant \mathbb{E}_X(a + bX^\rho) \iff \mathbb{E}_X q^{\alpha-1}(X) \geqslant a + b G_\rho(X)$.

Substituting $\|q\|_\alpha^\alpha = c_\gamma^\alpha \sum_{i=1}^M (1-\gamma i^\rho)_+^{\alpha'}$ gives (14). ∎

Remark 1: Sason and Verdú [25] stated an implicit lower bound on guessing moment vs. Rényi entropy in the unconditional case, by specifying only the minimizing pmf (equation (59) in [25]) for a guessing strategy which is not necessarily optimal. Their minimizing pmf is not the same as in the above proof when $0 < \alpha < 1$ because it is does not satisfy the constraint that $p_X(x)$ should be decreasing in x. Additionally, it can be checked that it cannot approach the Dirac distribution, hence does not provide the full range of the entropy values.

As an important consequence, an explicit first-order upper bound on $\Delta G_\rho(X;Y)$ can be obtained, which is easy to compute for any adversary observing small leakages.

Corollary 1: As $\Delta H_\alpha(X;Y) \to 0$, up to first order,

$$\Delta G_\rho(X;Y) \lesssim \sqrt{\frac{2(G_{2\rho}(M) - G_\rho^2(M))}{\alpha}} \sqrt{\frac{\Delta H_\alpha(X;Y)}{\log e}}. \tag{20}$$

In particular, $\Delta G(X;Y) \lesssim \sqrt{\frac{M^2-1}{6\alpha}} \sqrt{\frac{\Delta H_\alpha(X;Y)}{\log e}}$.

Proof: One has $\Delta G_\rho(q) \to 0$ and $\Delta H_\alpha(q) \to 0$ when q approaches the uniform distribution, i.e., when $\gamma \to 0$ in (13) or (14). Taylor expansion about $\gamma = 0$ in both cases gives

$$\begin{cases} \Delta G_\rho(X;Y) = \gamma|1 - \alpha'|(G_{2\rho}(M) - G_\rho^2(M)) + O(\gamma^2) \\ \frac{\Delta H_\alpha(X;Y)}{\log e} = \frac{|\alpha'(1-\alpha')|}{2}(G_{2\rho}(M) - G_\rho^2(M))\gamma^2 + O(\gamma^3) \end{cases} \tag{21}$$

which yields (20). This is also valid for $\alpha = 1$ by taking the limit $\alpha \to 1$. ∎

Fig. 2 illustrates the results of this section. Next we give two examples for which explicit upper and lower bounds can be derived directly.

Example 1 (Binary Random Variables): Let $\mathcal{B}(p)$ be the Bernoulli distribution with parameter $p \in [0,1]$ and define $h_\alpha(p) \triangleq H_\alpha(\mathcal{B}(p)) = \frac{1}{1-\alpha}\log((1-p)^\alpha + p^\alpha)$ and $k_\alpha(p) \triangleq K_\alpha(\mathcal{B}(p)) = ((1-p)^\alpha + p^\alpha)^{1/\alpha}$. We let h_α^{-1} and k_α^{-1} be their inverse when restricted to $[0, \frac{1}{2}]$. Note that k_α is concave increasing if $\alpha \in (0,1)$; and convex decreasing if $\alpha > 1$.

For $M = 2$ in the unconditional case (without side information) we simply have $G_\rho(X) = (1-p)+p2^\rho = 1+(2^\rho-1)p = 1 + (2^\rho - 1)k_\alpha^{-1}(K_\alpha(X))$. Since k_α^{-1} is convex, by Jensen's inequality, we obtain the desired lower bound explicitly. The upper bound is similarly obtained by taking the chord of k_α^{-1}:
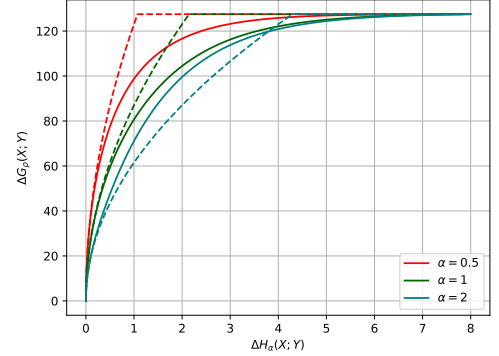
$$1 + (2^\rho - 1)h_\alpha^{-1}(H_\alpha(X|Y)) \leqslant G_\rho(X|Y)$$
$$\leqslant 1 + \frac{2^\rho-1}{2}\frac{\exp\left(\frac{1-\alpha}{\alpha}H_\alpha(X|Y)\right) - 1}{2^{\frac{1-\alpha}{\alpha}} - 1}. \tag{22}$$

Since $h_\alpha^{-1}(\log 2 - \delta) \approx \frac{1}{2} - \sqrt{\frac{1}{2\alpha}\frac{\delta}{\log e}}$ as $\delta \to 0$ we recover Corollary 1 as $\delta = \Delta H_\alpha(X;Y) \to 0$, which for $M = 2$ takes the form

$$\Delta G_\rho(X;Y) \lesssim (2^\rho - 1)\sqrt{\frac{\Delta H_\alpha(X;Y)}{2\alpha \log e}}. \tag{23}$$
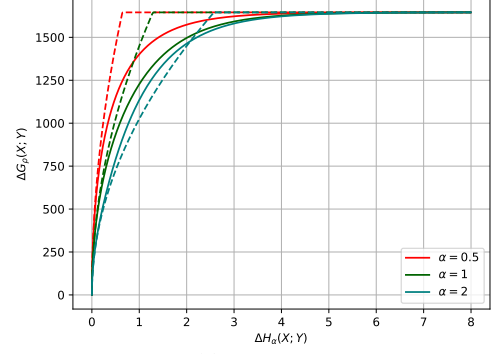
(a) $\rho = 1.5$.

(b) $\rho = 1$.

Fig. 2: *Solid*: Illustration of Theorems 1,2,3 (upper bounds of ΔG_ρ vs. ΔH_α) for various values of α and ρ when $M = 2^8$. *Dashed*: First-order lower bound $\Delta G_\rho \leqslant c\sqrt{\Delta H_\alpha}$ from Corollary 1.

In the other direction we also obtain at first order that

$$\Delta G_\rho(X;Y) \gtrsim \frac{2^\rho-1}{2}\frac{1-\alpha}{\alpha}\frac{2^{\frac{1-\alpha}{\alpha}}}{2^{\frac{1-\alpha}{\alpha}}-1}\frac{\Delta H_\alpha(X;Y)}{\log e} \tag{24}$$

which reads $\Delta G_\rho(X;Y) \gtrsim \frac{2^\rho-1}{2}\frac{\Delta H(X;Y)}{\log e}$ in the limiting case $\alpha \to 1$.

Example 2 (Guessing vs. Min-Entropy): When $\alpha = \infty$ explicit bounds can also be derived. Sason and Verdù [25, Theorem 9] provided the joint range between probability of error $\epsilon_{X|Y}$ and ρth-order guessing moments. Since $H_\infty(X|Y) = -\log(1 - \epsilon_{X|Y})$, their results can be easily rewritten in terms of min-entropy. This gives

$$K\sum_{i=1}^{\lfloor K^{-1} \rfloor} i^\rho + (1 - K\lfloor K^{-1}\rfloor)(1 + \lfloor K^{-1}\rfloor)^\rho \tag{25}$$

$$\leqslant G_\rho(X|Y) \leqslant 1 + \frac{M}{M-1}(G_\rho(M) - 1)(1 - K)$$

where $K = K_\infty(X|Y) = \exp(-H_\infty(X|Y))$. For $\rho = 1$, as shown in [29], this simplifies to $(\lfloor K^{-1}\rfloor + 1)(1 - \frac{\lfloor K^{-1}\rfloor}{2}K) \leqslant G(X|Y) \leqslant 1 + \frac{M}{2}(1 - K)$.

For small leakages of uniform secrets with maximal leakage $I_\infty(K;Y) = H_\infty(K) - H_\infty(K|Y) \leqslant \log\frac{M}{M-1}$, this rewrites $\frac{1}{2}(e^{I_\infty(K;Y)} - 1) \leqslant \Delta G(K;Y) \leqslant \frac{M-1}{2}(e^{I_\infty(K;Y)} - 1)$. As $I_\infty(K;Y) \to 0$ this yields the first order approximation,

$$\frac{1}{2}\frac{I_\infty(K;Y)}{\log e} \lesssim \Delta G(K;Y) \lesssim \frac{M-1}{2}\frac{I_\infty(K;Y)}{\log e}. \tag{26}$$

The upper bound is linear in $I_\infty(K;Y)$ which agrees with the fact that in (20) the term in $\sqrt{\Delta H_\alpha(K;Y)} = \sqrt{I_\alpha(K;Y)}$ vanishes as $\alpha \to +\infty$.

D. Evaluation in the Hamming Weight Lekage Model

A standard leakage model in side-channel analysis [3], [4], [8]–[10], [17], [20], [21] is the Hamming weight leakage model $Y = w_H(K) + N$, in which the Hamming weight $w_H(K)$ of the binary representation of the secret byte $K \sim \mathcal{U}(M)$, $M = 2^n$ leaks under additive Gaussian noise $N \sim \mathcal{N}(0, \sigma)$. Fig. 3 compares our bounds to previous ones and to the numerical evaluation of the actual guessing advantage. While the case $\alpha = 1$ (Thm. 1) is slightly better than $\alpha = \frac{1}{2}$ (Thm. 3), both bounds are very close to the exact value as noise increases, with huge gaps compared to previous Arikan's [6] and Rioul's [23] bounds which were used in [4], [5] for practical evaluation of full-key guessing entropy.
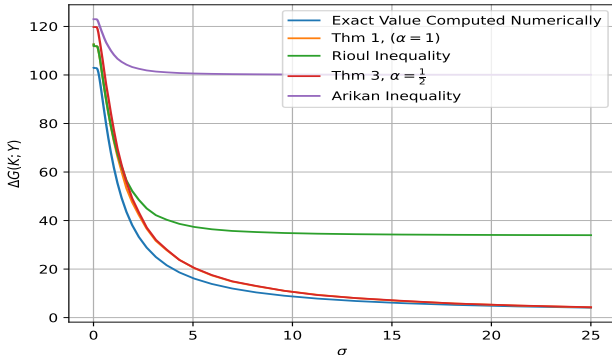
Fig. 3: Bound on the guessing advantage in the Hamming weight leakage model for increasing noise variance σ^2.

E. Evaluation in the Random Probing Model

The random probing model [30]–[32] is a well-known setting for security evaluation of cryptographic circuits. The side-channel is noiseless but has a random state: $Y = (Z, f_Z(K))$ where f_z is a deterministic leakage function depending on state z. Typically Y is either erased or leaks $f(K, T)$ with publicly available plain or cyphertext T.

Theorem 4: In the random probing model,

$$1 + \frac{M-1}{2} \frac{\exp(\frac{1-\alpha}{\alpha} H_\alpha(K|Y))-1}{M^{(1-\alpha)/\alpha}-1} \lessgtr G(K|Y) \lessgtr \frac{1+\exp(H_\alpha(K|Y))}{2}. \tag{27}$$

where $\lessgtr$ denotes $\geqslant$ for $\alpha \geqslant 1/2$ and $\leqslant$ for $\alpha \leqslant 1/2$. In the limiting case $\alpha = 1$,

$$\frac{1+\exp H(K|Y)}{2} \leqslant G(K|Y) \leqslant 1 + \frac{M-1}{2} \frac{H(K|Y)}{\log M}. \tag{28}$$

Proof: First, consider a fixed function $f = f_z$ with pre-image cardinality $M_y = |f^{-1}(\{y\})|$. Then $H_\alpha(K|Y = y) = \log M_y$ and $G(K|Y = y) = \frac{M_y+1}{2} = \frac{\exp(H_\alpha(K|Y=y))+1}{2} = \frac{K_\alpha(K|Y=y)^{\frac{\alpha}{1-\alpha}}+1}{2}$. If $\alpha = 1$, since $x \mapsto \frac{1+\exp(x)}{2}$ is convex we obtain by Jensen's inequality that $G(K|Y) \geqslant \frac{\exp(H_\alpha(K|Y))+1}{2}$. If $\alpha \neq 1$, $x \mapsto \frac{1+x^{\frac{\alpha}{1-\alpha}}}{2}$ is convex if $\alpha \geqslant \frac{1}{2}$ and concave

otherwise, hence $G(K|Y) \lessgtr \frac{K_\alpha(K|Y)^{\frac{\alpha}{1-\alpha}}+1}{2}$. The other bound is obtained by taking the chord of $x \mapsto \frac{1+x^{\frac{\alpha}{1-\alpha}}}{2}$. The region when f_z is obtained at random is obtained by taking the convex hull of the region (K_α, G), which is already convex. ∎

Fig. 4 shows that Theorem 4 greatly improves Theorem 3 in the random probing model. Interestingly, the case $\alpha = \frac{1}{2}$ gives equality $\Delta G(K;Y) = \frac{M}{2}(1 - \exp(-I_{\frac{1}{2}}(K;Y))) \approx \frac{M}{2} \frac{I_{\frac{1}{2}}(K;Y)}{\log e}$ where the approximation holds for small leakages.
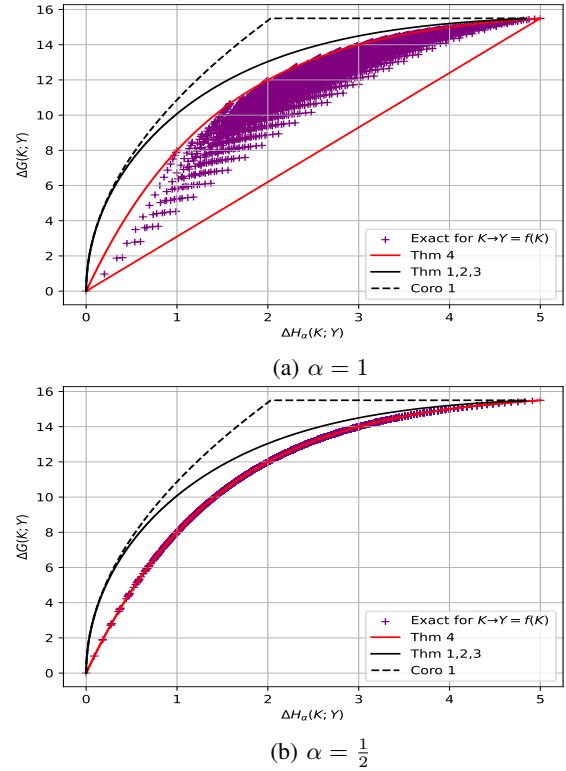
(a) $\alpha = 1$

(b) $\alpha = \frac{1}{2}$

Fig. 4: Improved bound ($M = 32$) from Theorem 4 (solid, red) compared to Theorem 3 (solid, black), Corollary 1 (dashed), and scatter plot of exact values when $Y = f(K)$ for arbitrarily given functions f.

IV. CONCLUSION

We derived closed-form optimal regions with explicit bounds on the guessing advantage ΔG_ρ of a secret random variable in terms of the α-information advantage ΔH_α. An important outcome is that it decreases as the square root $\Delta G_\rho \leqslant c\sqrt{\Delta H_\alpha}$ for small leakages. Simulations in the classical Hamming weight leakage model show that our bounds are much tighter than previous ones, especially for large noise (small leakage). We further sharpened the bounds in the random probing model where the guessing advantage is actually *equal* to a function of the Rényi-Arimoto entropy of order $1/2$.

As possible extensions of this work, it would be valuable to obtain sharpened bounds for any additive noise leakage model $Y = f(X) + Z$, for secrets with known (nonuniform) prior pmf p_X and generalize to negative values of α [33].

References

[1] J. L. Massey, "Guessing and entropy," in *Proceedings of 1994 IEEE International Symposium on Information Theory*, 1994, p. 204.

[2] J. O. Pliam, "Guesswork and variation distance as measures of cipher security," in *Selected Areas in Cryptography*, H. Heys and C. Adams, Eds., 2000, pp. 62–77.

[3] F.-X. Standaert, T. G. Malkin, and M. Yung, "A unified framework for the analysis of side-channel key recovery attacks," in *Advances in Cryptology - EUROCRYPT 2009*, A. Joux, Ed., 2009, pp. 443–461.

[4] M. O. Choudary and P. G. Popescu, "Back to Massey: Impressively fast, scalable and tight security evaluation tools," in *Proc. 19th Workshop on Cryptographic Hardware and Embedded Systems (CHES 2017)*, vol. LNCS 10529, 2017, pp. 367–386.

[5] A. Tănăsescu, M. O. Choudary, O. Rioul, and P. G. Popescu, "Tight and scalable side-channel attack evaluations through asymptotically optimal Massey-like inequalities on guessing entropy," *Entropy*, vol. 23, no. 11, pp. 1–10, 2021.

[6] E. Arikan, "An inequality on guessing and its application to sequential decoding," *IEEE Transactions on Information Theory*, vol. 42, no. 1, pp. 99–105, 1996.

[7] R. Poussier, F. Standaert, and V. Grosso, "Simple key enumeration (and rank estimation) using histograms: An integrated approach," in *Cryptographic Hardware and Embedded Systems - CHES 2016*, ser. LNCS, vol. 9813. Springer, 2016, pp. 61–81.

[8] É. de Chérisey, S. Guilley, O. Rioul, and P. Piantanida, "An information-theoretic model for side-channel attacks in embedded hardware," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 310–315.

[9] ——, "Best information is most successful: Mutual information and success rate in side-channel analysis," *IACR Transactions on Cryptographic Hardware and Embedded Systems (CHES 2019)*, pp. 49–79, Feb. 2019.

[10] W. Cheng, Y. Liu, S. Guilley, and O. Rioul, "Attacking masked cryptographic implementations: Information-theoretic bounds," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 654–659.

[11] I. Issa, A. B. Wagner, and S. Kamath, "An operational approach to information leakage," *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1625–1657, 2019.

[12] J. Béguinot, Y. Liu, O. Rioul, W. Cheng, and S. Guilley, "Maximal leakage of masked implementations using Mrs. Gerber's lemma for min-entropy," in *2023 IEEE International Symposium on Information Theory (ISIT)*, 2023, pp. 654–659.

[13] S. Arimoto, "Information measures and capacity of order α for discrete memoryless channels," *Topics in information theory (Colloquia Mathematica Societatis Janos Bolyai)*, vol. 16, pp. 41–52, 1977.

[14] R. Sibson, "Information radius," *Zeitschrift Wahrscheinlichkeitstheorie Verwandte Gebiete*, vol. 14, pp. 149–160, Jun. 1969.

[15] S. Verdú, "α-mutual information," in *Proc. Information Theory and Applications Workshop*, La Jolla, CA, Feb. 2015, pp. 1–6.

[16] J. Liao, L. Sankar, O. Kosut, and F. P. Calmon, "Robustness of maximal α-leakage to side information," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 642–646.

[17] Y. Liu, W. Cheng, S. Guilley, and O. Rioul, "On conditional alpha-information and its application to side-channel analysis," in *2021 IEEE Information Theory Workshop (ITW)*, 2021, pp. 1–6.

[18] M. Hayashi and V. Y. F. Tan, "Equivocations, exponents, and second-order coding rates under various Rényi information measures," *IEEE Transactions on Information Theory*, vol. 63, no. 2, pp. 975–1005, 2017.

[19] Y. Ishai, A. Sahai, and D. Wagner, "Private Circuits: Securing Hardware against Probing Attacks," in *CRYPTO*, ser. Lecture Notes in Computer Science, vol. 2729. Springer, August 17–21 2003, pp. 463–481, Santa Barbara, California, USA.

[20] L. Masure, O. Rioul, and F.-X. Standaert, "A nearly tight proof of Duc et al.'s conjectured security bound for masked implementations," in *Smart Card Research and Advanced Applications (CARDIS 2022)*, ser. LNCS, vol. 13820. Springer Nature, 2023, pp. 69–81.

[21] J. Béguinot, W. Cheng, S. Guilley, Y. Liu, L. Masure, O. Rioul, and F.-X. Standaert, "Removing the field size loss from Duc et al.'s conjectured bound for masked encodings," in *Constructive Side-Channel Analysis and Secure Design (COSADE 2023)*, ser. LNCS, vol. 13979. Springer Nature, 2023, pp. 86–104.

[22] S. Fehr and S. Berens, "On the conditional Rényi entropy," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 6801–6810, 2014.

[23] O. Rioul, "Variations on a theme by Massey," *IEEE Transactions on Information Theory*, vol. 68, no. 5, pp. 2813–2828, 2022.

[24] R. J. McEliece and Z. Yu, "An inequality on entropy," in *Proceedings of 1995 IEEE International Symposium on Information Theory*, 1995, p. 329.

[25] I. Sason and S. Verdú, "Improved bounds on lossless source coding and guessing moments via Rényi measures," *IEEE Transactions on Information Theory*, vol. 64, no. 6, pp. 4323–4346, 2018.

[26] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1st Ed. 1990, 2nd Ed. 2006.

[27] O. Rioul, "Rényi entropy power and normal transport," in *International Symposium on Information Theory and Its Applications, ISITA 2020, Kapolei, HI, USA, October 24-27, 2020*, 2020, pp. 1–5.

[28] C. Hirche, "Rényi bounds on information combining," in *IEEE International Symposium on Information Theory, ISIT 2020, Los Angeles, CA, USA, June 21-26, 2020*, 2020, pp. 2297–2302.

[29] O. Rioul, "The interplay between error, total variation, alpha-entropy and guessing: Fano and Pinsker direct and reverse inequalities," *Entropy*, vol. 25, no. 7, p. 978, 2023.

[30] A. Duc, S. Dziembowski, and S. Faust, "Unifying Leakage Models: From Probing Attacks to Noisy Leakage," in *Advances in Cryptology - EUROCRYPT 2014, Copenhagen, Denmark, May 11-15, 2014*, ser. LNCS, vol. 8441. Springer, 2014, pp. 423–440.

[31] S. Belaïd, J. Coron, E. Prouff, M. Rivain, and A. R. Taleb, "Random probing security: Verification, composition, expansion and new constructions," in *Advances in Cryptology - CRYPTO 2020, Santa Barbara, CA, USA, August 17-21, 2020*, ser. LNCS, vol. 12170. Springer, 2020, pp. 339–368.

[32] G. Cassiers, S. Faust, M. Orlt, and F. Standaert, "Towards tight random probing security," in *Advances in Cryptology - CRYPTO 2021, Virtual Event, August 16-20, 2021, Proceedings, Part III*, ser. LNCS, vol. 12827. Springer, 2021, pp. 185–214.

[33] A. R. Esposito, A. Vandenbroucque, and M. Gastpar, "On Sibson's α-mutual information," in *IEEE International Symposium on Information Theory, ISIT 2022, Espoo, Finland, June 26 - July 1, 2022*, 2022, pp. 2904–2909.