Optimal Accuracy-Privacy Trade-Off for Secure Computations

Patrick Ah-Fat and Michael Huth

***Abstract*— The purpose of secure multi-party computation is to enable protocol participants to compute a public function of their private inputs while keeping their inputs secret, without resorting to any trusted third party. However, opening the public output of such computations inevitably reveals some information about the private inputs. We propose a measure generalizing both Rényi entropy and g-entropy so as to quantify this information leakage. In order to control and restrain such information flows, we introduce the notion of function substitution, which replaces the computation of a function that reveals sensitive information with that of an approximate function. We exhibit theoretical bounds for the privacy gains that this approach provides and experimentally show that this enhances the confidentiality of the inputs while controlling the distortion of computed output values. Finally, we investigate the inherent compromise between accuracy of computation and privacy of inputs and we demonstrate how to realize such optimal trade-offs.**

***Index Terms*— Computational privacy, g-entropy, information flow, non-linear optimization, Rényi entropy.**

I. Introduction

WE STUDY the setting of functions f that map n integral inputs $x_1, \ldots, x_n$ into one integral output. Each input x_i is controlled by some agent i and its value is considered private to agent i. The computation of function f is *secure* if its evaluation protects the privacy of the inputs, so that agent j cannot learn more from this computation about the other values x_i than what agent j is able to infer from knowledge of her own input x_j and the publicly observable output $f(x_1, \ldots, x_n)$.

Secure Multi-party Computation (SMC) is a domain of cryptography that can implement such a black-box functionality: it enables protocol participants to compute a public function of their private inputs, such that no trusted third party is required, and that the confidentiality of the inputs is protected [1]–[6]. Recent advances in SMC have given birth to a variety of efficient protocols that achieve computational and information-theoretic security against passive and active adversaries [7]–[10].

Manuscript received February 1, 2018; revised July 25, 2018; accepted November 4, 2018. Date of publication December 12, 2018; date of current version April 19, 2019. This work was supported by the UK EPSRC under Grant EP/N020030/1 and Grant EP/N023242/1.

The authors are with the Department of Computing, Imperial College London, London SW7 2AZ, U.K. (e-mail: m.huth@imperial.ac.uk).

Communicated by R. Safavi-Naini, Associate Editor for Complexity and Cryptography.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2018.2886458

SMC therefore gives strong security guarantees, but it does allow inferences about other agents' input values based on the publicly observable output and one's own private input. This is referred to as the *acceptable information flow* in the SMC literature, in which this is therefore largely ignored. In fact, this so called acceptable information flow is oblivious to the manner in which a protocol realizes the aims of SMC.

Consequently, such information flows would also occur in the setting of outsourced computation. In this case, a trusted third party or a central authority (e.g. a national health agency) holds some records from different parties (e.g. some medical insurance companies), computes a function of those records and informs the parties of the result of the computation, such that no information leaks about the parties' inputs apart from the public output.

But we believe that such information flow is not always acceptable, e.g., in the medical domain with its strict privacy regulations. Moreover, we think that it is important to understand and quantify such information flow in order to

- better understand potential risks of using SMC in a specific application, say, a health-care consortium of insurers and hospitals
- devise methods that can mitigate or prevent such information flow.

The latter aim contains within it an inherent friction. The information flow whose existence only depends on knowledge of some private inputs and the public output can neither be mitigated against nor prevented by an SMC protocol that computed that function f. Rather, for such measures to be effective, we will need to modify the actual behavior of function f: we will instead use another function f' for which the acceptable information flow is absent, less pronounced or optimal according to some risk measure. The aforementioned friction consists of the need to shield against such undesired information flow for function f by replacing the latter with function f'. This substitution naturally introduces some inaccuracy in the value of the computed output, which will need to be controlled.

The notions of security developed for SMC are not directly helpful in understanding this friction and its inherent trade-off. SMC security neither reflects the amount of information that leaks from a computation once the public output is revealed, nor does it account for the ability of an attacker to influence such leakage before entering a protocol [11]–[14]. We therefore develop, in this paper, bespoke methods for understanding this better. Specifically:

1) We generalize a model of such information flow, developed in [15], to an entire family of conditional entropies. This subsumes the Shannon and min entropies as well as the notion of g-leakage in Computer Security.
2) We devise a method of distorting the output of a function f with so called *virtual inputs* so that the distorted function f' may be computed through standard means, such as SMC protocols.
3) We express a trade-off between privacy preservation of agents' inputs and output accuracy as a non-linear optimization problem, whose solution computes optimal virtual inputs.
4) We demonstrate that these optimizations can be solved for a large class of our conditional entropies, including the aforementioned ones.
5) We also offer theoretical insights that relate and characterize the relationship between the accuracy of the distorted function and the level of privacy that such distortions offer.

This work is motivated by, and applicable to, but not restricted to SMC. Our methods do not rely on the particular protocols used for SMC, but only on the abstract setting of a *black-box* function f to which parties i submit a private input x_i and all then learn the public output of f. We will therefore present the core of our technical development in this abstract setting, to stress that these results are orthogonal to the choice of an SMC protocol.

Naturally, the application context of an SMC may constrain or inform our approach. In a voting protocol, e.g., a mere deviation from the original function f would hardly be tolerable. But our approach may be used to enhance the privacy in less restrictive scenarios such as in the computation of statistical measures or financial audits.

Outline of Paper: We discuss related work in Section II. We present an intuitive idea of our approach and discuss its premisses in Section III. Needed technical background from information theory is covered in Section IV. Our development of a generalized conditional entropy is the subject of Section V. The development of our model for information flow for black-box functions and the attacker's entropy for that are described in Section VI. The method by which one can randomize black-box functions through virtual inputs is developed in Section VII and its theory is presented in Section VIII. Our approach to optimization of the trade-off between privacy and accuracy of black-box functions, and its contributions, are developed in Section IX. A discussion of our work is contained in Section X and Section XI concludes the paper.

II. Related Works

A. Information Flow in Programs

Information flow analysis in imperative programs has been explored with many different approaches. One of the fundamental concepts is that of security classes, introduced by Denning [16], which enables one to classify the variables of a program with respect to their level of confidentiality in order to form a lattice of information. Based on this classification, type systems [17] and semantic approaches [18] have been implemented in order to define the security of instructions involving such variables. The most basic model considers only two security classes L and H separating the variables with a low and high level of confidentiality respectively [16]. The security of a program is then expressed with the notion of non-interference between both classes [17], [19], [20]. However, as programs in practice may contain some interference, other quantitative approaches [21]–[26] have been proposed in order to measure the information flow that can arise between variables from different security classes. The computation of such quantitative information flows also includes the use of probabilistic instructions [18], [19], [27] that can randomize the algorithms and make programs non-deterministic and thus in some cases protect the confidentiality of information processed by variables in H.

B. Information Flow in SMC

The security of SMC protocols ensures by definition that the participants can compute the public output of a public function of their private inputs without learning anything about the other parties' input, apart from what is inferable from the so called acceptable information flow we already discussed.

In [15], we introduced a model of deceitful adversaries which enabled us to reason about the acceptable leakage, and to quantify, based on Shannon entropy, the information that such attackers can deduce from public outputs and their own private inputs. We also extended our model to a theoretic game that allows an attacker to evaluate the influence that he can have, depending on the input he provides to the SMC protocol. In this present work, we build on this model to develop an approach that can mitigate or prevent this information leakage. We are able to do this for a large class of conditional entropies, which subsumes both the notions of Rényi entropy and g-entropy. We then introduce the notion of an approximate function, a corresponding non-linear optimization problem, and we show how solving such optimization problems can address certain privacy concerns raised in [15].

C. Quantitative Information Flow and Differential Privacy

The aim of Quantitative Information Flow (QIF) [25], [28] is to provide frameworks and techniques based on information theory and probability theory for measuring the amount of information that leaks from a secret.

The principle of randomizing the output of a computation in order to protect the privacy of the data on which some calculations are performed is related to the concept of Differential Privacy (DP) [29], [30]. DP formalizes privacy concerns and introduces techniques that provide users of a database with the assurance that their personal details will not have a significant impact on the output of the queries performed on the database. More precisely, it proposes mechanisms which ensure that the outcome of the queries performed on two databases differing in at most one element will be statistically indistinguishable. Moreover, minimizing the distortion of the outcome of the queries while ensuring privacy is an important trade-off that governs DP.

To this extent, our approach resonates with the aims and concerns of DP. However, we identify two main differences between our work and DP which lie in the interpretation of privacy and utility. First, the privacy guarantee that DP offers is to make the outputs statistically indistinguishable whereas we are interested in minimizing the information flow of outputs. However, both ends are tightly related and recent works have shown that it is possible to determine an explicit upper bound on the information flow of an ε-differentially private mechanism which satisfies certain symmetry conditions [31]. On the other hand, our work and DP have different utility purposes. The accuracy of a query in DP is dependent on the mechanism that is used to achieve ε-DP, and the distortion on the output may be arbitrarily large — although with a small probability. In contrast, our work explicitly imposes a bound on the range of the additive noise. This divergence with DP may be crucial in a variety of scenarios which involve sensitive outputs with little or no tolerance of larger inaccuracies – such as auctions, e-voting or financial audits [32], [33].

Moreover, applying certain DP techniques directly, such as the Laplace mechanism [34], would not be sensible in our case: we need to work with integer values in order to align with the requirements of cryptographic protocols whose security relies on finite field theory. Other DP techniques, to mention the exponential mechanism [35], would technically be applicable to our case. However, their use would still neither allow us to restrict the noise to a given range, nor could we ensure that these techniques produce minimal information flow – which is the objective of our work.

In conclusion, our work differs from DP in regard to both privacy and utility: DP focuses on producing statistically indistinguishable outputs without imposing explicit bounds on the distortion whereas our mechanisms aim at minimizing the information flow given a certain range for the additive noise. It would certainly be interesting to investigate potential mechanisms on discrete domains that achieve ε-DP while restraining the distortion to a given range, to derive a bound on the information flow of such mechanisms courtesy of [31], and to then compare the results with the optimal information flow that the methods presented in our present work achieve. Such work is beyond the scope of this paper.

As an aside, we note that other recent works have explored new paradigms for unifying SMC and DP so as to propose more efficient protocols [36].

III. Methodology

In this section, we introduce an intuitive idea of our approach. We also articulate and discuss the premisses and the main contributions of our work.

We consider the scenario where several parties holding sensitive inputs wish to learn the result of a public function f of their private inputs. We assume that the public function f can be computed securely, i.e. that the output o of f can be computed without revealing any other information to the parties, apart from o itself. Such a secure functionality can be achieved e.g. by appealing to an SMC protocol, by using a trusted hardware, or by sending all the private inputs to a trusted third party who would compute and output the result of f. We will qualify such functions as black-boxes since our work will be orthogonal to the protocols used to compute those functions securely. In particular, the technicalities of our paper will not have any interaction with the domain of SMC, although the latter may benefit from our results.

In our setting, some attackers holding fixed inputs $\mathbf{x}_{\mathbb{A}}$ wish to learn as much information as possible on some targeted inputs $X_{\mathbb{T}}$ while the other inputs of the computation, called spectators' inputs, will be denoted as $X_{\mathbb{S}}$. We call $O = f(\mathbf{x}_{\mathbb{A}}, X_{\mathbb{T}}, X_{\mathbb{S}})$ the output of the function and we measure the information that the attackers have on $X_{\mathbb{T}}$ via $\mathrm{H}(X_{\mathbb{T}} \mid O, \mathbf{x}_{\mathbb{A}})$ for a certain entropy measure H. The first contribution of our work is to introduce a general notion of such an entropy H that generalizes the families of g-entropy and Rényi entropy, which we discuss next. Our second contribution is to propose a mechanism that enhances the inputs' privacy by introducing some random noise Φ, which we call virtual input, to the output of f. Intuitively, we build a function f' as $f'(\mathbf{x}_{\mathbb{A}}, X_{\mathbb{T}}, X_{\mathbb{S}}, \Phi) = h(f(\mathbf{x}_{\mathbb{A}}, X_{\mathbb{T}}, X_{\mathbb{S}}), \Phi)$ for a suitable function h, and we denote the output of f' as $O' = f'(\mathbf{x}_{\mathbb{A}}, X_{\mathbb{T}}, X_{\mathbb{S}}, \Phi)$. We prove that this mechanism can only enhance the targets' privacy, i.e. that $\mathrm{H}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}}) \geq \mathrm{H}(X_{\mathbb{T}} \mid O, \mathbf{x}_{\mathbb{A}})$ for the generalized entropy measure that we introduced, and we experimentally show the privacy gain that this method offers, which we further characterize when the functions $o \longmapsto h(o, \varphi)$ are injective. Our last main contribution is to propose some methods for computing an optimal distribution for the virtual input Φ so as to maximize $\mathrm{H}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}})$ averaged over all attackers' inputs $\mathbf{x}_{\mathbb{A}}$ given some bounds on the range of Φ. The aforementioned quantities $\mathrm{H}(X_{\mathbb{T}} \mid O, \mathbf{x}_{\mathbb{A}})$ and $\mathrm{H}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}})$ parameterized with bespoke attributes introduced in later sections will respectively be denoted as $\mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_{\mathbb{A}})$ and $\mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_{\mathbb{A}})$.

We compared our second contribution with some related works in Paragraph *c* of Section II and we now discuss our first contribution. Introducing a notion of entropy that generalizes both g-entropy and Rényi entropy presents two advantages. First, it enables us to formulate general and unifying theorems and to elaborate generic proofs in Section VIII. This also enables us, in Section IX, to approximate some optimal distributions for Φ which maximizes $\mathrm{H}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}})$ when H is non-differentiable, particularly when H is a g-entropy.

On the other hand, we are able to develop some results that are valid for a large range of entropies, and may thus benefit, in particular, the domains where Rényi entropy is used, which we quickly discuss in this paragraph. Rényi entropy has a major importance in privacy amplification [37], where Rényi entropy of order 2 is used to capture the fact that an attacker may have a statistical knowledge about a secret that any constraint based on Shannon entropy cannot convey [38]. This result was then generalized in [39], which derives a bound on the smooth entropy [40] using Rényi entropy of any order α greater than 1; this yields more precise bounds on the smooth entropy for large alphabets. The use of Rényi entropy in this result thus benefits the field of privacy amplification and some cryptographic techniques

where it can be applied such as key agreement from common information [41] and oblivious transfer [42]. Rényi entropy has also been used to quantify the randomness produced by universal hashing in the context of pseudo-random number generation [43].

Generalizing g-entropy and Rényi entropy may thus yield interesting results in some particular security scenarios, and may be beneficial in future research investigations that require Rényi entropy to characterize a given uncertainty property, such as in [38]. Although other more general notions of entropy have been proposed in different contexts such as relative perfect secrecy [44], a mathematical discussion on the natural properties that an entropy measure should satisfy [45] shows that our generalized entropy cannot both satisfy the data-processing inequality and be an expectation of its prior version. The definition of our general entropy thus violates the averaging property which does not have any operational significance for us, whereas the data-processing inequality, which we prove for our entropy in Theorem 1, is a fundamental result that supports our randomization mechanism.

IV. Background and Notations

We recall different notions of entropy used for quantifying information.

Notations: Let D be a discrete set. We write $\mathcal{P}(D)$ for the power set of D, and $|D|$ for the cardinality of set D. Let $\Omega(D)$ be the set of all probability distributions whose support is contained in D. Throughout, we present distributions as Python dictionaries with domain values as keys and associated probabilities as values. For example, $\{4\colon 1/2, 8\colon 1/2\}$ represents the uniform distribution over $\{4, 8\}$. For any integers a and b, we will write $[\![a, b]\!]$ for the set of consecutive integers ranging from a to b, namely $\{a, a+1, \cdots, b\}$. The set of positive integers will be denoted by $\mathbb{N}_{>0}$ while $\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$ will denote the set of non-negative and positive real numbers respectively. Let n be in $\mathbb{N}_{>0}$. A *linear distribution* over $[\![1, n]\!]$ will refer to the triangular distribution with mode n, i.e. to the distribution $\{k\colon \frac{2k}{n(n+1)} \mid 1 \leq k \leq n\}$ where $\frac{2}{n(n+1)}$ is a normalizing factor. Given random variable X and value x, the event "$X = x$" will be abbreviated by "x" when there is no ambiguity, and its probability will be denoted by $p(x)$. Similarly, we will abbreviate $\sum_{x \in D}$ by $\sum_x$ when the domain D is obvious from context.

We denote by $\langle x_i \rangle_{1 \leq i \leq n}$ the n-dimensional vector in $\mathbb{R}^n$ whose coordinates are $x_1, \cdots, x_n$; we abbreviate this by $\langle x_i \rangle_i$ when there is no ambiguity. For all vector v in $\mathbb{R}^n$ and all p in $\mathbb{R}_{\geq 0}$, the usual p-norm of v is denoted by $\|v\|_p$. Let log denote the logarithm in base 2 and let μ be the function defined for all non-negative real x as:

$$\mu(x) = \begin{cases} -x \cdot \log(x) & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases} \tag{1}$$

Shannon and min-entropy: Recall that for two random variables X and Y taking values in $\mathcal{X}$ and $\mathcal{Y}$, respectively, the *Shannon entropy* $\mathrm{H}_1(X)$ [46] of X and the Shannon entropy $\mathrm{H}_1(X \mid Y)$ of X given Y are given as:

$$\mathrm{H}_1(X) = -\sum_x p(x) \log p(x)$$
$$\mathrm{H}_1(X \mid Y) = \sum_y p(y)\, \mathrm{H}_1(X \mid y)$$

where $\mathrm{H}_1(X \mid y) = -\sum_x p(x \mid y) \log p(x \mid y)$. On the other hand, the *Bayes vulnerability* $\mathrm{V}_\infty(X)$ [26], [28], [40] expresses the probability of guessing a secret in one try. Similarly, the *conditional* Bayes vulnerability [47] $\mathrm{V}_\infty(X \mid Y)$ of X given Y reflects the average probability of guessing the secret X in one try. They are defined as:

$$\mathrm{V}_\infty(X) = \max_x p(x)$$
$$\mathrm{V}_\infty(X \mid Y) = \sum_y p(y)\, \mathrm{V}_\infty(X \mid y)$$

where $\mathrm{V}_\infty(X \mid y) = \max_x p(x \mid y)$. The *min-entropy* of X and *conditional* min-entropy of X given Y are defined as:

$$\mathrm{H}_\infty(X) = -\log \mathrm{V}_\infty(X)$$
$$\mathrm{H}_\infty(X \mid Y) = -\log \mathrm{V}_\infty(X \mid Y)$$

Rényi Entropy: A more general notion of entropy, called *Rényi entropy* [48], generalizes both notions of Shannon entropy and min-entropy. For sake of notational convenience, let us first define the *α-vulnerability* of X, for all positive real $\alpha \neq 1$, as: $\mathrm{V}_\alpha(X) = \left\| \langle p(x) \rangle_x \right\|_\alpha$ Using this notion, we may express the *Rényi entropy* $\mathrm{H}_\alpha(X)$ of X as:

$$\mathrm{H}_\alpha(X) = \frac{\alpha}{1-\alpha} \log \mathrm{V}_\alpha(X)$$

It is well-known and easily shown that the Rényi entropy $\mathrm{H}_\alpha(X)$ converges towards the min-entropy $\mathrm{H}_\infty(X)$ as α tends towards infinity. Moreover, an application of L'Hôpital's rule ensures that the Rényi entropy $\mathrm{H}_\alpha(X)$ converges towards the Shannon entropy $\mathrm{H}_1(X)$ as α tends towards 1.

However, although different notions of conditional Rényi entropy have been proposed, none of them has yet been commonly accepted as *the* conditional Rényi entropy [49]. Yet, one candidate seems to be particularly suitable for our needs: Arimoto's [50] notion of conditional Rényi entropy not only satisfies the natural properties of chain rule ($\mathrm{H}_\alpha(X \mid Y) \geq \mathrm{H}_\alpha(XY) - \log|\mathcal{Y}|$, where $\mathrm{H}_\alpha(XY)$ denotes the joint entropy of X and Y) and monotonicity ($\mathrm{H}_\alpha(X \mid Y) \leq \mathrm{H}_\alpha(X)$). But it is also compatible with both the Shannon entropy and the min-entropy in that we have the convergences $\mathrm{H}_\alpha(X \mid Y) \xrightarrow[\alpha \to 1]{} \mathrm{H}_1(X \mid Y)$ and $\mathrm{H}_\alpha(X \mid Y) \xrightarrow[\alpha \to \infty]{} \mathrm{H}_\infty(X \mid Y)$. Therefore, we will introduce and work with the notion of conditional Rényi entropy due to Arimoto [50]. For sake of notational consistency, let us define the α-vulnerability of X given Y as: $\mathrm{V}_\alpha(X \mid Y) = \sum_y p(y)\, \mathrm{V}_\alpha(X \mid y)$ where $\mathrm{V}_\alpha(X \mid y) = \left\| \langle p(x \mid y) \rangle_x \right\|_\alpha$. For $\alpha \neq 1$, we may now define the *conditional Rényi entropy* of X given Y as:

$$\mathrm{H}_\alpha(X \mid Y) = \frac{\alpha}{1-\alpha} \log \mathrm{V}_\alpha(X \mid Y)$$

g-Entropy: The g-entropy [51] measures the gain that someone might get from guessing a secret — in our case, the private

inputs of other parties. Since this is a relevant way of measuring risk of privacy violations, we wish that our approach and developed methods also support use of this notion of entropy.

Let $\mathcal{X}$ be the domain of X, where random variable X models a *secret*. Let $\mathcal{W}$ be the set of possible guesses for the value of X. A function g of type $\mathcal{W} \times \mathcal{X} \to [0, 1]$ is then called a *gain function*. This function assigns to each guess w in $\mathcal{W}$ and possible value x of the secret in $\mathcal{X}$ a reward $g(w, x)$ that an attacker would gain by guessing w when the secret value actually is x. Set $\mathcal{W}$ may be designed so that its elements refer to properties of secrets, values that are "close" to the secret or other means of expressing aspects of the secret.

For such a gain function g, the *g-vulnerability of* X, $\mathrm{V}_g(X)$, is the expected reward that an attacker would gain by selecting his best guess. It, and the *conditional* g-vulnerability $\mathrm{V}_g(X \mid Y)$ of X given Y are defined as:

$$\mathrm{V}_g(X) = \max_w \sum_x p(x)g(w, x)$$
$$\mathrm{V}_g(X \mid Y) = \sum_y p(y)\,\mathrm{V}_g(X \mid y)$$

where $\mathrm{V}_g(X \mid y) = \max_w \sum_x p(x \mid y)g(w, x)$. The *g-entropy* and *conditional g-entropy* are then defined as follows:

$$\mathrm{H}_g(X) = -\log \mathrm{V}_g(X)$$
$$\mathrm{H}_g(X \mid Y) = -\log \mathrm{V}_g(X \mid Y)$$

The g-entropy generalizes the min-entropy: for $\mathcal{W} = \mathcal{X}$ and gain function $id\colon \mathcal{X} \times \mathcal{X} \to [0, 1]$ — where $id(w, x) = 0$ if w is not equal to x, and $id(x, x) = 1$ for all x in $\mathcal{X}$, then $\mathrm{H}_{\mathsf{id}}(X)$ equals $\mathrm{H}_\infty(X)$ and $\mathrm{H}_{\mathsf{id}}(X \mid Y)$ equals $\mathrm{H}_\infty(X \mid Y)$.

V. Generalized Conditional Entropy

To get a very general definition of information leakage in Secure Multi-Party Computations (SMC), we define a more general notion of entropy that subsumes both Rényi entropy and g-entropy. For random variables X and Y with finite domain $\mathcal{X}$ and $\mathcal{Y}$, and a finite set $\mathcal{W}$ of possible guesses for X, we adapt the existing notions, such as the Bayesian vulnerability, to the presence of a gain function g and its set of guesses $\mathcal{W}$. We indicate that dependency by writing $V_{\alpha,g}$ and so forth, subsequently. We define properties of gain functions that are pertinent to our technical development.

Definition 1: Let $g\colon \mathcal{W} \times \mathcal{X} \to [0, 1]$ *be a gain function.*

1) *Then* g *is* positive *iff* $\forall x \in \mathcal{X}\colon \sum_w g(w, x) > 0$.
2) *Let* β *be in* $\mathbb{R}_{>0}$. *The gain function* g *is* β*-positive iff* $\forall x \in \mathcal{X}\colon \sum_w g(w, x) \geq \beta$.
3) *Function* g *is* unitary *iff* $\forall x \in \mathcal{X}\colon \sum_w g(w, x) = 1$.

We will only consider positive gain functions: a gain function that is not positive is the constant 0 function, and can only produce 0 vulnerabilities — as mentioned in [51]; β-positive gain functions will be useful in later sections. Note that, since $\mathcal{X}$ is finite, all positive gain function g have some $\beta > 0$ such that g is β-positive.

Let $g\colon \mathcal{W} \times \mathcal{X} \to [0, 1]$ be a gain function and $0 < \alpha \neq 1$. The (α, g)-vulnerability $\mathrm{V}_{\alpha,g}(X)$ of X and the conditional (α, g)-vulnerability $\mathrm{V}_{\alpha,g}(X \mid Y)$ of X given Y are defined as:

$$\mathrm{V}_{\alpha,g}(X) \coloneqq \left\| \left\langle \sum_x p(x)g(w, x) \right\rangle_w \right\|_\alpha \tag{2}$$
$$\mathrm{V}_{\alpha,g}(X \mid Y) \coloneqq \sum_y p(y)\,\mathrm{V}_{\alpha,g}(X \mid y) \tag{3}$$

where $\mathrm{V}_{\alpha,g}(X \mid y) \coloneqq \left\| \left\langle \sum_x p(x \mid y)g(w, x) \right\rangle_w \right\|_\alpha$. We now define the (α, g)-entropy of X and the conditional (α, g)-entropy of X as:

$$\mathrm{H}_{\alpha,g}(X) \coloneqq \frac{\alpha}{1-\alpha} \log \mathrm{V}_{\alpha,g}(X) \tag{4}$$
$$\mathrm{H}_{\alpha,g}(X \mid Y) \coloneqq \frac{\alpha}{1-\alpha} \log \mathrm{V}_{\alpha,g}(X \mid Y) \tag{5}$$

Again, we can easily verify that the (α, g)-entropies $\mathrm{H}_{\alpha,g}(X)$ and $\mathrm{H}_{\alpha,g}(X \mid Y)$ both converge towards their respective g-entropies as α tends towards infinity. We may thus define:

$$\mathrm{H}_{\infty,g}(X) \coloneqq \mathrm{H}_g(X)$$
$$\mathrm{H}_{\infty,g}(X \mid Y) \coloneqq \mathrm{H}_g(X \mid Y)$$

We now focus on the case when α tends towards 1 and we define, where μ is as in (1):

$$\mathrm{H}_{1,g}(X) \coloneqq \sum_w \mu\left(\sum_x p(x)g(w, x)\right)$$
$$\mathrm{H}_{1,g}(X \mid Y) \coloneqq \sum_w \mu\left(\sum_x p(x \mid y)g(w, x)\right)$$

For *unitary* gain functions g, it is easy to see that the (α, g)-entropies $\mathrm{H}_{\alpha,g}(X)$ and $\mathrm{H}_{\alpha,g}(X \mid Y)$ converge towards $\mathrm{H}_{1,g}(X)$ and $\mathrm{H}_{1,g}(X \mid Y)$, respectively, when α tends towards 1. The reason for this is that, when g is unitary, the (α, g)-vulnerabilities $\mathrm{V}_{\alpha,g}(X)$ and $\mathrm{V}_{\alpha,g}(X \mid Y)$ converge towards 1 as α tends towards 1. And then the claimed results follow from the application of L'Hôpital's rule, in a similar fashion as done for Rényi entropies. Let us formalize this:

Lemma 1: 1) *Let* $g\colon \mathcal{W} \times \mathcal{X} \to [0, 1]$ *be a gain function. Then* $\mathrm{H}_{\alpha,g}(X)$ *and* $\mathrm{H}_{\alpha,g}(X \mid Y)$ *converge for* $\alpha \to \infty$*:*

$$\lim_{\alpha\to\infty} \mathrm{H}_{\alpha,g}(X) = \mathrm{H}_{\infty,g}(X)$$
$$\lim_{\alpha\to\infty} \mathrm{H}_{\alpha,g}(X \mid Y) = \mathrm{H}_{\infty,g}(X \mid Y)$$

2) *Moreover, if* g *is unitary, then* $\mathrm{H}_{\alpha,g}(X)$ *and* $\mathrm{H}_{\alpha,g}(X \mid Y)$ *converge when* α *tends towards* 1*, and we then have:*

$$\lim_{\alpha\to 1} \mathrm{H}_{\alpha,g}(X) = \mathrm{H}_{1,g}(X)$$
$$\lim_{\alpha\to 1} \mathrm{H}_{\alpha,g}(X \mid Y) = \mathrm{H}_{1,g}(X \mid Y)$$

When the gain function g is id with $\mathcal{W} = \mathcal{X}$ as above, we obtain that for all positive reals α, the (α, id)-entropies agree with the Rényi entropies:

$$\mathrm{H}_{\alpha,\mathsf{id}}(X) = \mathrm{H}_\alpha(X)$$
$$\mathrm{H}_{\alpha,\mathsf{id}}(X \mid Y) = \mathrm{H}_\alpha(X \mid Y)$$

This result is immediate for all values of α different from 1. When α is equal to 1, this follows from the fact that

α \ g	$g = \mathsf{id}$	$g \in [0,1]^{\mathcal{W}\times\mathcal{X}}$
$\alpha = 1$	Shannon entropy	only for unitary g
$\alpha = \infty$	min-entropy	g-entropy
$\alpha \in]0,\infty]$	Rényi entropy	

Fig. 1. Summary of the different notions of entropy that our generalized measure of information flow $\mathrm{H}_{\alpha,g}$ subsumes.

id is a unitary gain function and that we can apply the previous result ensuring that when α tends towards 1, the (α, id)-entropies $\mathrm{H}_{\alpha,\mathsf{id}}(X)$ and $\mathrm{H}_{\alpha,\mathsf{id}}(X \mid Y)$, respectively, converge towards $\mathrm{H}_{1,\mathsf{id}}(X)$ (the Shannon entropy) and $\mathrm{H}_{1,\mathsf{id}}(X \mid Y)$ (the conditional Shannon entropy), respectively. We summarize those results and our discussion in Figure 1.

In conclusion, our new notion of entropy subsumes both the g-entropy and the whole family of Rényi entropies, including the Shannon entropy and the min-entropy. Therefore, all results that we develop in this paper will also be valid for all the different entropies mentioned earlier.

VI. Information Flow for Secure Multi-Party Computation

A. Model for Information Flow

Let us recall the technical setting and the assumptions introduced in [15] for studying and quantifying the information flow produced by public outputs in SMC, as this constitutes a basis for the remaining technical developments in this paper. Throughout this paper, we consider a set of $n > 1$ parties $\mathbb{P} = \{P_1, \cdots, P_n\}$ holding the respective inputs $x_1, \cdots, x_n$, each of them belonging to $\mathbb{Z}$. Let $f\colon \mathbb{Z}^n \to \mathbb{Z}$ be a function. Let o denote the output of the function applied with the parties' inputs, i.e. $o = f(x_1, \cdots, x_n)$. Both o and f are public and so known to all parties in $\mathbb{P}$. In order to study the aforementioned acceptable information leakage of this situation, we introduce the following model. Let $\mathbb{A}$ and $\mathbb{T}$ be two non-empty subsets of $\mathbb{P}$ and $\mathbb{S}$ be a possibly empty subset of $\mathbb{P}$ such that $(\mathbb{A}, \mathbb{T}, \mathbb{S})$ forms a partition of $\mathbb{P}$. Our attack models assumes that all parties in $\mathbb{A}$ are willing to collaborate between each other in order to maximize information leakage on inputs of the parties in $\mathbb{T}$. The sets $\mathbb{A}$, $\mathbb{T}$ and $\mathbb{S}$ will thus respectively be referred to as the sets of *attackers, targets and spectators*, respectively. We now define the attackers' input $\mathbf{x}_\mathbb{A} = \langle x_i \rangle_{i\in\mathbb{A}}$, the targets' input $\mathbf{x}_\mathbb{T} = \langle x_i \rangle_{i\in\mathbb{T}}$ and the spectators' input $\mathbf{x}_\mathbb{S} = \langle x_i \rangle_{i\in\mathbb{S}}$. By abuse of notation (or a reordering of arguments for f), we will also refer to the output specification of f as $o = f(\mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{S})$.

Let $a = |\mathbb{A}|$, $t = |\mathbb{T}|$ and $s = |\mathbb{S}|$ denote the cardinality of the respective sets. Let $D_\mathbb{A}$ be an element of $\mathcal{P}(\mathbb{Z}^a)$ and let us assume that the input vector of the parties in $\mathbb{A}$ is ranged in $D_\mathbb{A}$. Similarly, let $D_\mathbb{T}$ in $\mathcal{P}(\mathbb{Z}^t)$ and $D_\mathbb{S}$ in $\mathcal{P}(\mathbb{Z}^s)$ be the domain of the input vectors of the parties in $\mathbb{T}$ and $\mathbb{S}$ respectively. In other words, we assume that:

$$\mathbf{x}_\mathbb{A} \in D_\mathbb{A}, \quad \mathbf{x}_\mathbb{T} \in D_\mathbb{T}, \quad \mathbf{x}_\mathbb{S} \in D_\mathbb{S}$$

However, as those inputs are private, their exact value is not known to the other parties. In order to quantify the information leaks that output o produces, we model the parties' inputs as random variables $X_\mathbb{A}$, $X_\mathbb{T}$ and $X_\mathbb{S}$ respectively, following the respective probability distributions:

$$\pi_\mathbb{A} \in \Omega(D_\mathbb{A}), \quad \pi_\mathbb{T} \in \Omega(D_\mathbb{T}), \quad \pi_\mathbb{S} \in \Omega(D_\mathbb{S})$$

where $\Omega(X)$ is the set of discrete probability distributions over a finite set X. These probability distributions will model the beliefs that each set of parties has on the other parties' inputs. More precisely, the parties in $\mathbb{A}$ and $\mathbb{T}$ believe that random variable $X_\mathbb{S}$ is governed by $\pi_\mathbb{S}$, the parties in $\mathbb{A}$ and $\mathbb{S}$ believe that $X_\mathbb{T}$ follows $\pi_\mathbb{T}$, whereas the parties in $\mathbb{T}$ and $\mathbb{S}$ believe that $X_\mathbb{A}$ follows $\pi_\mathbb{A}$. We articulate the assumptions we make about these distributions:

Assumption 1: We assume that the parties' beliefs $\pi_\mathbb{A}$, $\pi_\mathbb{T}$ and $\pi_\mathbb{S}$ are public and are part of the common knowledge amongst all parties in $\mathbb{P}$. Moreover, our model assumes that the three groups of parties will not collaborate between each other and that their inputs are thus independent.

The independence of $X_\mathbb{A}$, $X_\mathbb{T}$ and $X_\mathbb{S}$ will play an important role in the proofs of the Theorems in Section VIII. The assumption that their probability distributions are public and part of the common knowledge ensures that all the parties will be able to access the same data produced by our measure of information flow in Section VI-B and Section VII, and will be able to reach a consensus regarding how to best protect the targeted inputs' privacy, as discussed in Section IX. These probability distributions can express a variety of beliefs from uniform to point mass distributions.

Lastly, let D_O in $\mathcal{P}(\mathbb{Z})$ be the output domain, defined as $D_O = \{f(\mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{S}) \mid \mathbf{x}_\mathbb{A} \in D_\mathbb{A}, \mathbf{x}_\mathbb{T} \in D_\mathbb{T}, \mathbf{x}_\mathbb{S} \in D_\mathbb{S}\}$. As a function of random variables, the output $o = f(\mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{S})$ will therefore be modeled by the random variable:

$$O_f = f(X_\mathbb{A}, X_\mathbb{T}, X_\mathbb{S}) \tag{6}$$

ranged in D_O. We sometimes write O when f is clear from context. In order to quantify the information that the attackers would learn about $X_\mathbb{T}$ when inputting a particular input $\mathbf{x}_\mathbb{A}$, we introduced in [15] the attackers' weighted average entropy $\mathsf{awae}^\mathbb{A}_\mathbb{T}$ defined for all $\mathbf{x}_\mathbb{A}$ in $D_\mathbb{A}$ as the conditional *Shannon* entropy of $X_\mathbb{T}$ given O and $\mathbf{x}_\mathbb{A}$, i.e.:

$$\mathsf{awae}^\mathbb{A}_\mathbb{T}(\mathbf{x}_\mathbb{A}) = \sum_o p(o \mid \mathbf{x}_\mathbb{A}) \sum_{\mathbf{x}_\mathbb{T}} \mu(p(\mathbf{x}_\mathbb{T} \mid o, \mathbf{x}_\mathbb{A})) \tag{7}$$

where μ was defined in (1).

A deceitful attacker, i.e. an attacker who is willing to lie about his honest and intended input in order to learn more information on the private inputs of his targets, will now be able to take advantage of this indicator in (7) in order to shape his input so as to maximize his information gain. Since the notion of $\mathsf{awae}^\mathbb{A}_\mathbb{T}$ in (7) is an instance of the conditional Shannon entropy, we need to widen the approach and analyses of [15] to make them compatible with more general notions of entropy. We develop this next.

B. General Attackers' Entropy

Function $\mathsf{awae}^\mathbb{A}_\mathbb{T}$ for measuring information leakage is dependent on some implicit parameters, namely the function f, the partition $(\mathbb{A}, \mathbb{T}, \mathbb{S})$ of $\mathbb{P}$ and the distributions $\pi_\mathbb{T}$

and $\pi_{\mathbb{S}}$ of the targets and spectators' inputs. Our technical development needs to make those parameters explicit, and we wish to integrate the generalized entropy notion presented in Section V. Therefore, we now define a higher-order function Awae which fulfills those requirements. Subsequently, we will work with a set of allowable guesses $\mathcal{W}$ for the targeted input $\mathbf{x}_{\mathbb{T}}$.

Definition 2: Let α be in $\mathbb{R}_{>0} \cup \{\infty\}$ and $g\colon \mathcal{W} \times D_{\mathbb{T}} \to [0,1]$ be a gain function. We introduce the higher-order function $\mathsf{Awae}_{\alpha,g}$ of type:

$$\mathsf{Awae}_{\alpha,g}\colon (\mathbb{Z}^n \to \mathbb{Z}) \times \mathcal{P}(\mathbb{P})^3 \times \Omega(\mathbb{Z})^2 \to (D_{\mathbb{A}} \to \mathbb{R}_{\geq 0})$$

that takes as arguments a function f of type $\mathbb{Z}^n \to \mathbb{Z}$, three disjoint sets of participants $(\mathbb{A}, \mathbb{T}, \mathbb{S})$ that form a partition of $\mathbb{P}$, the probability distribution $(\pi_{\mathbb{T}}, \pi_{\mathbb{S}})$ of the respective targets' and spectators' inputs, and returns a function $\mathsf{Awae}_{\alpha,g}(f, (\mathbb{A}, \mathbb{T}, \mathbb{S}), (\pi_{\mathbb{T}}, \pi_{\mathbb{S}}))$ of type $D_{\mathbb{A}} \to \mathbb{R}_{\geq 0}$, denoted as $\mathsf{awae}^{f}_{\alpha,g}$ and defined for all $\mathbf{x}_{\mathbb{A}}$ in $D_{\mathbb{A}}$ as the conditional (α, g)-entropy of $X_{\mathbb{T}}$ given O_f as in (6) and $\mathbf{x}_{\mathbb{A}}$:

$$\mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_{\mathbb{A}}) = \mathrm{H}_{\alpha,g}(X_{\mathbb{T}} \mid O_f, \mathbf{x}_{\mathbb{A}})$$

For subsequent theorems and proofs, we note that for $0 < \alpha \neq 1$ we have:

$$\mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_{\mathbb{A}}) = \frac{\alpha}{1-\alpha} \cdot \log \mathrm{V}_{\alpha,g}(X_{\mathbb{T}} \mid O_f, \mathbf{x}_{\mathbb{A}}) \tag{8}$$

where the (α, g)-vulnerability $\mathrm{V}_{\alpha,g}(X_{\mathbb{T}} \mid O_f, \mathbf{x}_{\mathbb{A}})$ can be written as:

$$\begin{aligned}
&\mathrm{V}_{\alpha,g}(X_{\mathbb{T}} \mid O, \mathbf{x}_{\mathbb{A}}) \\
&\quad = \sum_{o} p(o \mid \mathbf{x}_{\mathbb{A}}) \cdot \left\| \left\langle \sum_{\mathbf{x}_{\mathbb{T}}} p(\mathbf{x}_{\mathbb{T}} \mid o, \mathbf{x}_{\mathbb{A}}) \cdot g(w, \mathbf{x}_{\mathbb{T}}) \right\rangle_{w} \right\|_{\alpha} \\
&\quad = \sum_{o} \left\| \left\langle \sum_{\mathbf{x}_{\mathbb{T}}} p(\mathbf{x}_{\mathbb{T}}) \cdot p(o \mid \mathbf{x}_{\mathbb{T}}, \mathbf{x}_{\mathbb{A}}) \cdot g(w, \mathbf{x}_{\mathbb{T}}) \right\rangle_{w} \right\|_{\alpha}
\end{aligned} \tag{9}$$

This is so since the α-norm is homogeneous, even for $\alpha < 1$, and since $\mathbf{x}_{\mathbb{A}}$ and $\mathbf{x}_{\mathbb{T}}$ are independent random variables.

This new function $\mathsf{awae}^{f}_{\alpha,g}$ provides us with a generic way of measuring information flow. Indeed, it subsumes some notions of entropy that are widely used in cryptography. For example, when g equals id, this function corresponds to the conditional Rényi entropy. When α equals ∞, it corresponds to the conditional g-entropy. We also observe that when α equals 1 and g equals id, our new function $\mathsf{awae}^{f}_{1,\mathsf{id}}$ is identical to the function $\mathsf{awae}^{\mathbb{A}}_{\mathbb{T}}$ introduced in [15].

We now illustrate how our general measure of information flow in Secure Multi-Party Computations enables us to quantify the information that attackers can gain on their targets' inputs. In doing so, we also raise interesting concerns that will further motivate our present work. Let us consider an example.

Example 1: Let us consider 3 parties X, Y and Z holding the respective inputs x, y and z, and where $\mathbb{A} = \{X\}$ is attacking $\mathbb{T} = \{Y\}$ under spectator $\mathbb{S} = \{Z\}$. Let $D_{\mathbb{A}} = D_{\mathbb{T}} = D_{\mathbb{S}} = [\![1, 30]\!]$ and let us assume that $X_{\mathbb{T}}$ and $X_{\mathbb{S}}$ are uniformly distributed over this domain. Let $f\colon \mathbb{Z}^3 \to \mathbb{Z}$ be defined by $f(x, y, z) = x(2y + z) + 2z$.

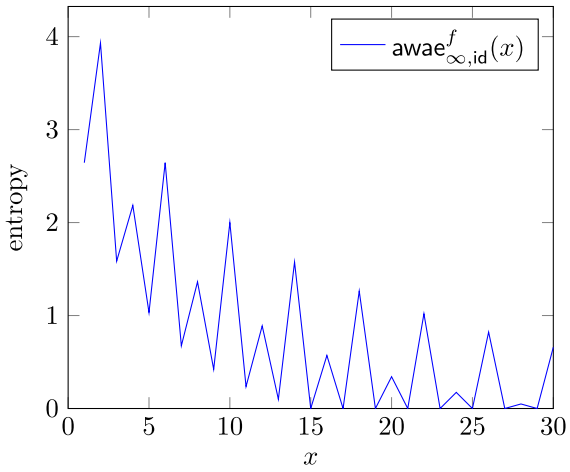

Fig. 2. Illustration of $\mathsf{awae}^{f}_{\infty,\mathsf{id}}$ in the computation of function $f(x, y, z) = x(2y+z)+2z$ with $\pi_{\mathbb{T}}$ and $\pi_{\mathbb{S}}$ uniform over $[\![1, 30]\!]$, when X attacks Y under spectator Z.

In this example, we will study the behavior of the conditional min-entropy of the targeted inputs. In other words, we will instantiate α with ∞ and g with id in order to study the function $\mathsf{awae}^{f}_{\infty,\mathsf{id}}$ which we plot in Figure 2. This plot clearly shows that some values of $\mathbf{x}_{\mathbb{A}}$ are more advantageous for attacker X in that they produce lower conditional entropies for his targeted input Y. For instance, inputting $x = 2$ would produce a high entropy and would not reveal much information about y. In contrast, input $x = 15$ would produce entropy 0, which means that X would learn the exact value of y from the output. Indeed, as X knows his own input, he knows that in this case, the output equals $o = f(15, y, z) = 30y + 17z$. We can check that for all z in $D_{\mathbb{S}}$ the function $f_z\colon y \mapsto f(15, y, z)$ is bijective from $D_{\mathbb{T}}$ to $f_z(D_{\mathbb{T}})$ as both sets have size 30. This thus ensures that attacker X can deduce the exact value of y when learning the output value.

We just saw that the choice of the attackers' input $\mathbf{x}_{\mathbb{A}}$ can have a dramatic influence on the entropy of the targeted input $\mathbf{x}_{\mathbb{T}}$. In particular, the attackers can harm the privacy of their targets by choosing some judicious inputs $\mathbf{x}_{\mathbb{A}}$. In order to mitigate against this privacy concern, we next introduce and study the notion of *approximate function*.

VII. Function Randomization via Virtual Inputs

We now consider the case where revealing the exact value of the output of f, namely $o = f(\mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}, \mathbf{x}_{\mathbb{S}})$, would be likely to jeopardize the privacy of the targeted input $\mathbf{x}_{\mathbb{T}}$. Thus, we would like to be able to replace the computation of f by the computation of an approximate function f', whose output should not only be a decent indicator of o, but should also enhance the privacy of $\mathbb{T}$'s input. This presents an inherent trade-off between the accuracy of the output and the privacy of the inputs. We will understand this trade-off in detail in Section IX.

In order to randomize the observed output, the function f' will take an additional argument φ, that may consist of a

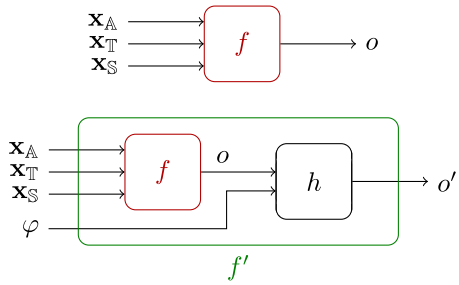

Fig. 3. Comparison of the black-box model for secure function f (top) with that of its approximation f' (bottom), as introduced in Definition 3. The virtual inputs φ and the output o of f are fed into function h within the black box to produce approximate output o'.

number of integer inputs, and that will act as a source of randomness that can distort the output to protect privacy of targeted inputs. Let us next formalize this notion of approximate function.

Definition 3: Let n be in $\mathbb{N}_{>0}$, v be in $\mathbb{N}$ and $f\colon \mathbb{Z}^n \to \mathbb{Z}$ be an n-ary function.

1) *Function $f'\colon \mathbb{Z}^n \times \mathbb{Z}^v \to \mathbb{Z}$ is an* approximation *of f or that f'* approximates *f iff there exists a function $h\colon \mathbb{Z} \times \mathbb{Z}^v \to \mathbb{Z}$ such that:*

$$\forall \mathbf{x} \in \mathbb{Z}^n, \forall \varphi \in \mathbb{Z}^v\colon f'(\mathbf{x}, \varphi) = h(f(\mathbf{x}), \varphi) \qquad (10)$$

2) *An approximation f' of f is a* close approximation *of f — or f'* closely approximates *f — iff for all φ in $\mathbb{Z}^v$, the function $h_\varphi\colon \mathbb{Z} \to \mathbb{Z}$ is injective, where h_φ is defined for all φ in $\mathbb{Z}^v$ as $\forall o \in \mathbb{Z}\colon h_\varphi(o) = h(o, \varphi)$.*

3) *We define $f^+\colon \mathbb{Z}^{n+1} \to \mathbb{Z}$, the* additive approximation *of f, for all $\mathbf{x}$ in $\mathbb{Z}^n$ and φ in $\mathbb{Z}$ as $f^+(\mathbf{x}, \varphi) = f(\mathbf{x}) + \varphi$.*

We illustrate the notion of approximate function f' for a function f in Figure 3. Function f' has all inputs of f and additional virtual inputs φ; and its black box contains "internal wirings" so that φ and the output o of f are fed into function h within that black box to produce approximate output o'. A *close* approximation f' of f requires all the functions h_φ to be injective, which makes sense for SMC as it enforces a correlation between the output of f and that of its approximation f'. Indeed, knowledge of o' and φ determine that of o, which prevents o' from being independent from o. We also note that the *additive* approximation f^+ of a function f closely approximates the latter.

The use of a substitute function f' aims to contain and limit the information that would flow from the computation of f by randomizing the output of f with an additional variable φ. Therefore, we need to understand and quantify the information flow that the computation of such an approximate function f' produces, and we need to study and represent the behavior of the additional variable φ that f' uses to randomize the output of f. To ensure the security of such approximations, variable φ is not held by any physical party; it is a *virtual input*, a concept we formalize next.

Definition 4: Let n and v be in $\mathbb{N}_{>0}$. A v-dimensional virtual input *φ is a vector in $\mathbb{Z}^v$,* independent of the other inputs*, and not held by any party in $\mathbb{P}$. As such, its value φ is kept secret and appears to all the parties in $\mathbb{P}$ as a random variable Φ on domain D_Φ following a probability distribution π_Φ, referred to as the* virtual distribution*. A set of* virtual parties *$\mathbb{V}$ is deemed to be the (virtual) owner of φ.*

In other words, the probability distribution π_Φ can be regarded as the prior belief that all the parties in $\mathbb{P}$ have on input φ. Note that all those parties in $\mathbb{P}$ will have the same public prior belief on φ, in the spirit of Assumption 1, and that $\mathbb{P}$ and $\mathbb{V}$ are mutually disjoint.

The set of parties $\mathbb{P}'$ for function f' is $\mathbb{P}' = \mathbb{P} \cup \mathbb{V}$. We now study the privacy that targeted parties gain when the computation of a function f is replaced by that of an approximation f', randomized by a virtual input φ.

Definition 5: Let $n > 1$ and v in $\mathbb{N}_{>0}$. Let $f\colon \mathbb{Z}^n \to \mathbb{Z}$ be a function and let $f'\colon \mathbb{Z}^n \times \mathbb{Z}^v \to \mathbb{Z}$ approximate f. Let a virtual input φ be in $\mathbb{Z}^v$ and let π_Φ be its probability distribution. Finally, let α be in $\mathbb{R} \cup \{\infty\}$ and g be a gain function of type $\mathcal{W} \times D_\mathbb{T} \to [0, 1]$. Using the joint probability distribution defined by $(\pi_\mathbb{S} \cdot \pi_\Phi)(\mathbf{x}_\mathbb{S}, \varphi) := \pi_\mathbb{S}(\mathbf{x}_\mathbb{S}) \cdot \pi_\Phi(\varphi)$ for all $\mathbf{x}_\mathbb{S}$ in $D_\mathbb{S}$ and φ in D_Φ, function $\mathsf{awae}_{\alpha,g}^{f',\pi_\Phi}\colon D_\mathbb{A} \to \mathbb{R}_{\geq 0}$ is given as:

$$\mathsf{awae}_{\alpha,g}^{f',\pi_\Phi} := \mathsf{Awae}_{\alpha,g}(f', (\mathbb{A}, \mathbb{T}, \mathbb{S} \cup \mathbb{V}), (\pi_\mathbb{T}, \pi_\mathbb{S} \cdot \pi_\Phi))$$

This function $\mathsf{awae}_{\alpha,g}^{f',\pi_\Phi}$ measures the privacy of the targets, given a certain approximate function and virtual input distribution. It will be particularly useful, for studying how privacy changes for different virtual input distributions. The assumption that for f' and f, the sets $\mathbb{A}$ and $\mathbb{T}$ are unchanged, does not compromise the security of our approach: an attacker for function f' could not really learn anything useful about the input of parties in $\mathbb{V}$, since these inputs are randomly drawn according to π_Φ. Let us illustrate the benefits offered by function substitution.

Example 2: Let us re-consider the scenario of Example 1, but now with the additive approximation f^+ of f. We will study the behavior of the conditional min-entropy of the targeted inputs when we approximate f with f^+. In other words, we will study the function $\mathsf{awae}_{\infty,\mathsf{id}}^{f^+,\pi_{\Phi i}}$ for the following distributions $\pi_{\Phi i}$:

$$\pi_{\Phi 1} = \{-2\colon 1/4, 0\colon 1/4, 2\colon 1/4, 4\colon 1/4\}$$

$$\pi_{\Phi 2} = \{-1\colon 1/4, 0\colon 1/4, 1\colon 1/4, 2\colon 1/4\}$$

$$\pi_{\Phi 3} = \{-3\colon 1/8, -2\colon 1/8, -1\colon 1/8, 0\colon 1/4, 1\colon 1/8, 2\colon 1/8, 3\colon 1/8\}$$

As seen in Figure 4, for all $1 \leq i \leq 3$, function $\mathsf{awae}_{\infty,\mathsf{id}}^{f^+,\pi_{\Phi i}}$ is above $\mathsf{awae}_{\infty,\mathsf{id}}^{f}$. This suggests that randomizing a computation effectively enhances the privacy of the targeted inputs.

The latter example indicates that function randomization indeed contributes to improving the privacy of the targets. In the next section, we want to formally investigate the privacy gains offered by function randomization. In particular, we would like to understand why substituting the computation of a function f by that of an approximation f' can only enhance the privacy of the targets, and we will further characterize this privacy gain for *close* approximations.

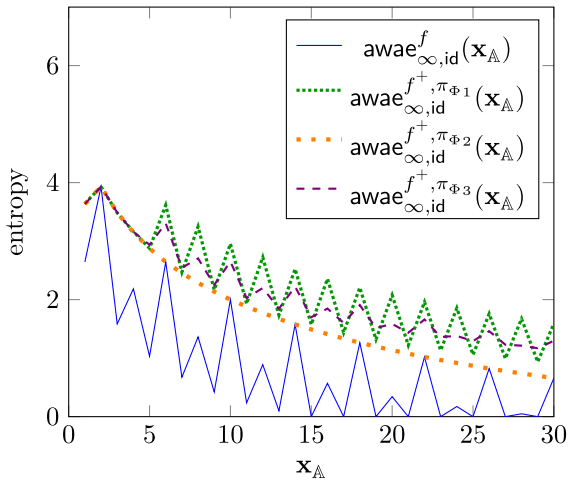

Fig. 4. Benefits of replacing the computation of f by that of its approximation f^+ in the computation of $f(x, y, z) = x(2y+z)+2z$ with $\pi_{\mathbb{T}}$ and $\pi_{\mathbb{S}}$ uniform over $[\![1, 30]\!]$. For all $1 \leq i \leq 3$, the function $\mathsf{awae}^{f^+,\pi_{\Phi i}}_{\infty,\mathsf{id}}$ is above $\mathsf{awae}^{f}_{\infty,\mathsf{id}}$.

VIII. Theory of Virtual Input Randomization

We first summarize the mathematical setting we study in the remainder of this paper:

Assumption 2: In the remainder of this paper, including lemmas and theorems, f' is an approximation of f, where φ is a virtual input with domain D_Φ. Moreover, $g\colon \mathcal{W} \times D_{\mathbb{T}} \to [0, 1]$ is a positive gain function, and β is a positive real.

The following theorem states that the computation of any approximate function f' will not produce a lower privacy for the targeted inputs than that produced by the computation of f.

Theorem 1: Let α be a positive real different from 1. *Then, we have:*

$$\forall \pi_\Phi \in \Omega(D_\Phi), \forall \mathbf{x}_{\mathbb{A}} \in D_{\mathbb{A}}\colon \mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_{\mathbb{A}}) \geq \mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_{\mathbb{A}}) \tag{11}$$

Proof: Let $D \coloneqq D_{\mathbb{A}} \times D_{\mathbb{T}} \times D_{\mathbb{S}}$ and $X \coloneqq (X_{\mathbb{A}}, X_{\mathbb{T}}, X_{\mathbb{S}})$. By definition, since f' approximates f, there is a function h such that $f'(\mathbf{x}, \varphi) = h(f(\mathbf{x}), \varphi)$ for all $\mathbf{x}$ in D and all φ in D_Φ. The random variable representing the output of f, namely $O = f(X)$ has domain D_O. Similarly, let $D_{O'}$ be the domain of the output of f', namely $O' = f'(X, \Phi) = h(f(X), \Phi) = h(O, \Phi)$. Let π_Φ be in $\Omega(D_\Phi)$ and $\mathbf{x}_{\mathbb{A}}$ be in $D_{\mathbb{A}}$. We recall that we have:

$$\mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_{\mathbb{A}}) = \frac{\alpha}{1-\alpha} \cdot \log \mathrm{V}_{\alpha,g}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}})$$

where:

$$\mathrm{V}_{\alpha,g}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}}) = \sum_{o'} \left\| \left\langle \sum_{\mathbf{x}_{\mathbb{T}}} p(\mathbf{x}_{\mathbb{T}}) \cdot p(o' \mid \mathbf{x}_{\mathbb{T}}, \mathbf{x}_{\mathbb{A}}) \cdot g(w, \mathbf{x}_{\mathbb{T}}) \right\rangle_w \right\|_\alpha \tag{12}$$

Applying Bayes Theorem twice, and as Φ is independent from $X_{\mathbb{A}}$, $X_{\mathbb{T}}$ and O, we obtain that:

$$\begin{aligned} p(o' \mid \mathbf{x}_{\mathbb{T}}, \mathbf{x}_{\mathbb{A}}) &= \sum_{\varphi} p(\varphi) \cdot p(o' \mid \mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}, \varphi) \\ &= \sum_{\varphi} p(\varphi) \cdot \sum_{o \in h_\varphi^{-1}(o')} p(o \mid \mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}) \\ &\quad \cdot p(o' \mid \mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}, \varphi, o) \end{aligned} \tag{13}$$

since $p(o' \mid \mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}, \varphi, o) \neq 0$ only when $h_\varphi(o) = o'$. Moreover, $p(o' \mid \mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}, \varphi, o) = 1$ when $h_\varphi(o) = o'$.

Case $\alpha > 1$: We can apply the triangular inequality twice from Equation (12) in order to obtain:

$$\mathrm{V}_{\alpha,g}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}}) \leq \sum_{o'} \sum_{\varphi} p(\varphi) \cdot \sum_{o \in h_\varphi^{-1}(o')} \left\| \left\langle \sum_{\mathbf{x}_{\mathbb{T}}} p(\mathbf{x}_{\mathbb{T}}) \cdot p(o \mid \mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}) \cdot g(w, \mathbf{x}_{\mathbb{T}}) \right\rangle_w \right\|_\alpha \tag{14}$$

For any φ in D_Φ, the collection of sets $(h_\varphi^{-1}(o'))_{o' \in D_{O'}}$ constitutes a partition of D_O. So there exists a unique o' in $D_{O'}$ such that $h_\varphi(o) = o'$. We can thus simplify the double summation over o' and o as a single sum over o:

$$\mathrm{V}_{\alpha,g}(X_{\mathbb{T}} \mid O', \mathbf{x}_{\mathbb{A}}) \leq \sum_{\varphi} p(\varphi) \cdot \sum_{o} \left\| \left\langle \sum_{\mathbf{x}_{\mathbb{T}}} p(\mathbf{x}_{\mathbb{T}}) \cdot p(o \mid \mathbf{x}_{\mathbb{A}}, \mathbf{x}_{\mathbb{T}}) \cdot g(w, \mathbf{x}_{\mathbb{T}}) \right\rangle_w \right\|_\alpha$$

Since α is greater than 1, the expression $\frac{\alpha}{1-\alpha}$ is negative and we get:

$$\mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_{\mathbb{A}}) \geq \mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_{\mathbb{A}}) \tag{15}$$

Case $\alpha < 1$: We can show that for all n in $\mathbb{N}_{>0}$, for all x and y in $(\mathbb{R}_{\geq 0})^n$, we have $\|x + y\|_\alpha \geq \|x\|_\alpha + \|y\|_\alpha$. This follows from Minkowski inequality in the case where α is lower than 1, since $x \mapsto x^\alpha$ is then concave on $\mathbb{R}_{\geq 0}$. This reversed triangular inequality reverses the inequality obtained in (14) and as $\frac{\alpha}{1-\alpha}$ is now positive, we find the same result as in (15). □

The proof of the previous theorem is based on the analysis of the formal expressions of $\mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_{\mathbb{A}})$ and $\mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_{\mathbb{A}})$ when $0 < \alpha \neq 1$. However, we can extend this result to $\alpha = 1$ and $\alpha = \infty$, by appealing to that result for positive $\alpha \neq 1$ and the continuity of inequalities under limits:

Corollary 1: 1) *We have that:* $\forall \pi_\Phi \in \Omega(D_\Phi), \forall \mathbf{x}_{\mathbb{A}} \in D_{\mathbb{A}}\colon \mathsf{awae}^{f',\pi_\Phi}_{\infty,g}(\mathbf{x}_{\mathbb{A}}) \geq \mathsf{awae}^{f}_{\infty,g}(\mathbf{x}_{\mathbb{A}})$.

2) *Moreover, if g is unitary, then we have:* $\forall \pi_\Phi \in \Omega(D_\Phi), \forall \mathbf{x}_{\mathbb{A}} \in D_{\mathbb{A}}\colon \mathsf{awae}^{f',\pi_\Phi}_{1,g}(\mathbf{x}_{\mathbb{A}}) \geq \mathsf{awae}^{f}_{1,g}(\mathbf{x}_{\mathbb{A}})$.

Proof: By virtue of Lemma 1, we know that letting α tend towards ∞ in Theorem 1 yields the result stated in item 1) above. Similarly, if g is unitary, Lemma 1 ensures that Theorem 1 implies the result stated in item 2) as $\alpha \to 1$. □

Concretely, the theorem states that learning a function of the output of f cannot leak more information on the inputs

of f than the output of f may leak already. This can be seen as a generalization of the data-processing inequality (DPI) and Item 1 of Corollary 1 enables us to recover the known DPI for g-entropy [51]. On the other hand, we are able to estimate an upper bound for the privacy of the inputs of the targeted parties, once a computation has been randomized. The next theorem states that, when replacing the computation of a function f by that of a *close* approximation f', the entropy gain provided by a virtual input cannot exceed the entropy of the distribution for the virtual inputs.

Theorem 2: Let f' be a close approximation of f and $0 < \alpha \neq 1$. Then, we have:

$$\forall \pi_\Phi \in \Omega(D_\Phi), \forall \mathbf{x}_\mathbb{A} \in D_\mathbb{A}: \quad \mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \leq \mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_\mathbb{A}) + \mathrm{H}_\alpha(\pi_\Phi) \tag{16}$$

where $\mathrm{H}_\alpha(\pi_\Phi)$ refers to the Rényi entropy of order α of the distribution π_Φ.

Proof: By definition, since f' closely approximates f, there exists some function h such that $f'(\mathbf{x}, \varphi) = h(f(\mathbf{x}), \varphi)$ for all $\mathbf{x}$ in D and φ in D_Φ. Let π_Φ be in $\Omega(D_\Phi)$ and $\mathbf{x}_\mathbb{A}$ be in $D_\mathbb{A}$. For sake of readability, we set $V' = \mathrm{V}_{\alpha,g}(X_\mathbb{T} \mid O', \mathbf{x}_\mathbb{A})$ and use V' in the arguments below. From Equation (13), we recall that:

$$\mathrm{V}' = \sum_{o'}\left(\sum_{w}\left[\sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot \sum_{\varphi} p(\varphi) \cdot p(o' \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \varphi) \cdot g(w, \mathbf{x}_\mathbb{T})\right]^\alpha\right)^{\frac{1}{\alpha}}$$

Case $\alpha > 1$: We know that $x \mapsto x^\alpha$ is convex on $\mathbb{R}_{\geq 0}$ and equals 0 at 0. We also know that $x \mapsto x^{\frac{1}{\alpha}}$ is increasing on $\mathbb{R}_{\geq 0}$ and thus:

$$\begin{aligned}\mathrm{V}' &\geq \sum_{o'}\left(\sum_{w}\sum_{\varphi}\left[\sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(\varphi) \cdot p(o' \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \varphi) \cdot g(w, \mathbf{x}_\mathbb{T})\right]^\alpha\right)^{\frac{1}{\alpha}} \\ &\geq \sum_{o'}\left(\sum_{w}\sum_{\varphi} p(\varphi)^\alpha \cdot \left[\sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o' \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \varphi) \cdot g(w, \mathbf{x}_\mathbb{T})\right]^\alpha\right)^{\frac{1}{\alpha}}\end{aligned} \tag{17}$$

Let us denote $\sum_\varphi p(\varphi)^\alpha$ by σ. For any φ, we have $p(\varphi)^\alpha = \sigma \cdot \frac{p(\varphi)^\alpha}{\sigma}$. But also $\sum_\varphi \frac{p(\varphi)^\alpha}{\sigma}$ equals 1. We also know that $x \mapsto x^{\frac{1}{\alpha}}$ is concave. Therefore, Jensen's inequality yields:

$$\mathrm{V}' \geq \sigma^{\frac{1}{\alpha}} \sum_{\varphi} \frac{p(\varphi)^\alpha}{\sigma} \sum_{o'}\left(\sum_{w}\left[\sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o' \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \varphi) \cdot g(w, \mathbf{x}_\mathbb{T})\right]^\alpha\right)^{\frac{1}{\alpha}} \tag{18}$$

Moreover, we have $p(o' \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \varphi) = p(O \in h_\varphi^{-1}(o') \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T})$ since O and Φ are independent. Furthermore, for all φ in D_Φ, we know that h_φ is injective. Thus, from (18) we get that:

$$\begin{aligned}\mathrm{V}' &\geq \sigma^{\frac{1}{\alpha}} \sum_{\varphi} \frac{p(\varphi)^\alpha}{\sigma} \sum_{o}\left(\sum_{w}\left[\sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}) \cdot g(w, \mathbf{x}_\mathbb{T})\right]^\alpha\right)^{\frac{1}{\alpha}} \\ &\geq \sigma^{\frac{1}{\alpha}} \sum_{o} \left\| \left\langle \sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o \mid \mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}) \cdot g(w, \mathbf{x}_\mathbb{T}) \right\rangle_w \right\|_\alpha\end{aligned}$$

and as $\frac{\alpha}{1-\alpha}$ is negative, the claim follows:

$$\mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \leq \mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_\mathbb{A}) + \mathrm{H}_\alpha(\pi_\Phi) \tag{19}$$

Case $\alpha < 1$: This is dual: $x \mapsto x^\alpha$ is concave, the inequality of (17) is reversed, $x \mapsto x^{\frac{1}{\alpha}}$ is convex, and the inequality in (18) is reversed, too. However, term $\frac{\alpha}{1-\alpha}$ is now positive. Thus, a dual argument shows that (19) holds. □

We can also extend the result of Theorem 2 to the limiting cases, i.e. to when α equals 1 or ∞.

Corollary 2: 1) *We have that:* $\forall \pi_\Phi \in \Omega(D_\Phi), \forall \mathbf{x}_\mathbb{A} \in D_\mathbb{A}$: $\mathsf{awae}^{f',\pi_\Phi}_{\infty,g}(\mathbf{x}_\mathbb{A}) \leq \mathsf{awae}^{f}_{\infty,g}(\mathbf{x}_\mathbb{A}) + \mathrm{H}_\infty(\pi_\Phi)$.

2) *Moreover, if g is unitary, then we have:* $\forall \pi_\Phi \in \Omega(D_\Phi), \forall \mathbf{x}_\mathbb{A} \in D_\mathbb{A}$: $\mathsf{awae}^{f',\pi_\Phi}_{1,g}(\mathbf{x}_\mathbb{A}) \leq \mathsf{awae}^{f}_{1,g}(\mathbf{x}_\mathbb{A}) + \mathrm{H}_1(\pi_\Phi)$.

Proof: Lemma 1 ensures the convergence of $\mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A})$ and of $\mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_\mathbb{A})$ towards their respective limiting values $\mathsf{awae}^{f',\pi_\Phi}_{\infty,g}(\mathbf{x}_\mathbb{A})$ and $\mathsf{awae}^{f}_{\infty,g}(\mathbf{x}_\mathbb{A})$ when α tends to ∞. Moreover, it is known that the Rényi entropy $\mathrm{H}_\alpha(\pi_\Phi)$ of order α converges to the min-entropy $\mathrm{H}_\infty(\pi_\Phi)$ as α tends towards ∞. Thus, the result stated in item 1) follows from Theorem 2 by letting α tend towards ∞. A similar argument concludes the proof for α tending towards 1, in the case that g is unitary. □

Although one could have proved Corollaries 1 and 2 with bespoke and somewhat different arguments, it is pleasing to see that our generalized conditional entropy makes such arguments uniform and reasonably simple. Given a close approximation f' of a function f, Theorems 1 and 2 and Corollaries 1 and 2 give us a lower bound and an upper bound for the entropy gain that a given virtual distribution provides. We can formalize this through a gain function $\Gamma_{\alpha,g}$, which indicates how much entropy we gain by adding a virtual input to the secure computation — as a function of the chosen probability distribution of this virtual input:

Definition 6: Let f' be a close approximation of a function f. Let α be in $\mathbb{R}_{>0} \cup \{\infty\}$. Let us further assume that either α is different from 1 or g is unitary. Then, we define the function $\Gamma_{\alpha,g}$ for all π_Φ in $\Omega(D_\Phi)$ and $\mathbf{x}_\mathbb{A}$ in $D_\mathbb{A}$ by:

$$\Gamma_{\alpha,g}(\pi_\Phi, \mathbf{x}_\mathbb{A}) \coloneqq \mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A}) - \mathsf{awae}^{f}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \tag{20}$$

Then, under the assumptions of Theorem 2, we can summarize our above results as follows:

Corollary 3: Let f' be a close approximation of f. Let α be in $\mathbb{R}_{>0} \cup \{\infty\}$. Let us further assume that either α is different

from 1 *or* g *is unitary. Then, we have:*

$$\forall \pi_\Phi \in \Omega(D_\Phi), \quad \forall \mathbf{x}_\mathbb{A} \in \Omega(D_\mathbb{A}): \\ 0 \leq \Gamma_{\alpha,g}(\pi_\Phi, \mathbf{x}_\mathbb{A}) \leq \mathrm{H}_\alpha(\pi_\Phi) \quad (21)$$

Proof: This result is a direct consequence of Theorems 1 and 2 and Corollaries 1 and 2. □

Let us now illustrate Theorems 1 and 2 by means of a worked example.

Example 3: Let us re-consider the scenario in Example 5 with the additive approximation f^+ of f; in particular, f' is a close approximation of f. We study the behavior of the conditional min-entropy of the targeted inputs when we approximate f with f^+. In other words, we study the function $\mathsf{awae}^{f^+,\pi_\Phi}_{\infty,\mathsf{id}}$ *for different distributions π_Φ. Since f^+ is a close approximation of f, Theorems 1 and 2 apply, and thus for all π_Φ in $\Omega(\mathbb{Z})$ and for all $\mathbf{x}_\mathbb{A}$ in $D_\mathbb{A}$, we have:*

$$\mathsf{awae}^{f}_{\infty,\mathsf{id}}(\mathbf{x}_\mathbb{A}) \leq \mathsf{awae}^{f^+,\pi_\Phi}_{\infty,\mathsf{id}}(\mathbf{x}_\mathbb{A}) \\ \leq \mathsf{awae}^{f}_{\infty,\mathsf{id}}(\mathbf{x}_\mathbb{A}) + \mathrm{H}_\infty(\pi_\Phi) \quad (22)$$

In order to illustrate this property, we choose different distributions for φ that all have equal min-entropy:

$$\pi_{\Phi 1} = \{-2\colon 1/4, 0\colon 1/4, 2\colon 1/4, 4\colon 1/4\}$$
$$\pi_{\Phi 2} = \{-1\colon 1/4, 0\colon 1/4, 1\colon 1/4, 2\colon 1/4\}$$
$$\pi_{\Phi 3} = \{-3\colon 1/8, -2\colon 1/8, -1\colon 1/8, 0\colon 1/4, \\ 1\colon 1/8, 2\colon 1/8, 3\colon 1/8\}$$

All those distributions have the same min-entropy which equals $-\log(1/4) = 2$. In Figure 5, we plot the functions $\mathsf{awae}^{f}_{\infty,\mathsf{id}}$, $\mathsf{awae}^{f}_{\infty,\mathsf{id}}+2$, *and* $\mathsf{awae}^{f^+,\pi_{\Phi i}}_{\infty,\mathsf{id}}$ *for all $1 \leq i \leq 3$. We can verify that Equation (22) indeed holds: for all $1 \leq i \leq 3$, the function* $\mathsf{awae}^{f^+,\pi_{\Phi i}}_{\infty,\mathsf{id}}$ *is contained between the functions* $\mathsf{awae}^{f}_{\infty,\mathsf{id}}$ *and* $\mathsf{awae}^{f}_{\infty,\mathsf{id}} + \mathrm{H}_\infty(\pi_{\Phi i})$.

Finally, note that although the three virtual distributions $\pi_{\Phi i}$ have equal min-entropy, they produce different values for $\mathsf{awae}^{f^+,\pi_{\Phi i}}_{\infty,\mathsf{id}}$. *From the plots we can clearly see, e.g., that $\pi_{\Phi 1}$ produces higher entropy values than $\pi_{\Phi 2}$ in general. This observation motivates us to seek optimal virtual distributions, which we focus on in the next section.*

IX. Optimal Trade-Off Between Accuracy and Privacy

So far, we developed a means of replacing a function f by an approximating function f' which resorts to additional, virtual inputs governed by some distribution. We showed that such approximations enable us to protect the privacy of the targeted inputs. These benefits are hampered by the fact that the approach introduces a distortion on the output for function f when computing with f' instead. The participants of the SMC computation from set $\mathbb{P}$ are either eager to learn the actual output of function f or they would tolerate only a certain difference between the outputs of f and f', and these demands would typically be informed by the use-context of the SMC computation.

Therefore, we need to have methods by which we can control the support and the distribution of the virtual input,

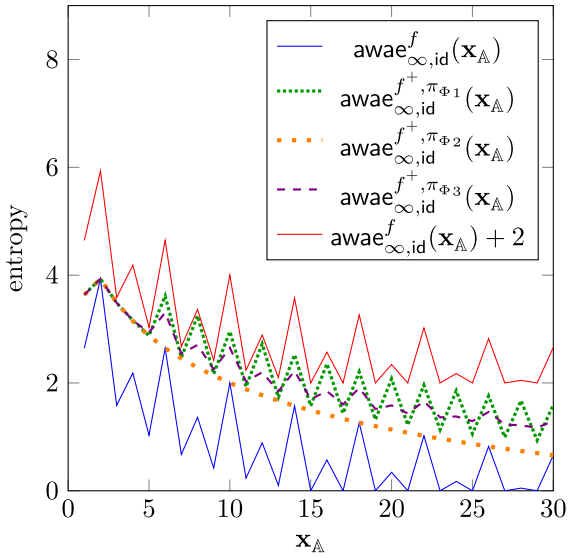

Fig. 5. Illustration of the bounds of $\mathsf{awae}^{f^+,\pi_{\Phi i}}_{\infty,\mathsf{id}}$ in the computation of $f(x, y, z) = x(2y + z) + 2z$ with $\pi_\mathbb{T}$ and $\pi_\mathbb{S}$ uniform over $[\![1, 30]\!]$. For all i in $[\![1, 3]\!]$, the function $\mathsf{awae}^{f^+,\pi_{\Phi i}}_{\infty,\mathsf{id}}$ is contained between $\mathsf{awae}^{f}_{\infty,\mathsf{id}}$ and $\mathsf{awae}^{f}_{\infty,\mathsf{id}} + \mathrm{H}_\infty(\pi_{\Phi i})$.

in order to measure and control both the distortion that f' and π_Φ introduce, and the privacy gain that it offers over using f for SMC instead. We therefore develop now the formalism needed for studying the inherent trade-off between the accuracy of the output and the privacy of supplied inputs. We also recall that Assumption 1 ensures that any of the parties can perform the methods we introduce next and compute optimal virtual distributions. When replacing the computation of a function f by an approximation f', the output accuracy is directly influenced by the choice of f'. A function f' that is the constant 0 function, e.g., would not reveal anything about the inputs, but be very inaccurate.

Assumption 3: In the remaining paper, we will focus on additive approximations f^+ of f.

This is a natural assumption to make: it simplifies our problem, as shown in (24), and enables us to characterize optimal virtual distributions. We will also propose some practical methods from optimization for discovering virtual distributions that realize this trade-off in an optimal manner. We now make an assumption on the adversarial power.

Assumption 4: In this section, we will consider that the parties in $\mathbb{A}$ are passive *attackers.*

In particular, we assume that the attackers $\mathbb{A}$ will not be deceitful and will provide their honest and intended inputs to the SMC protocol. This is a rational assumption: in this section, we are studying the trade-off between privacy of inputs and accuracy of outputs. However, if the attackers are deceitful, their inputs $\mathbf{x}_\mathbb{A}$ might be far different from their intended inputs, which might lead the output to be devoid of any significance. In this case, it would be hopeless to control the accuracy of such an output. The approach and the results presented in this section are thus only sensible in the presence of passive attackers. However, all the theoretical results that have been developed in the previous sections

are still relevant in the presence of passive and deceitful adversaries. Indeed, the functions $\mathsf{awae}^{f}_{\alpha,g}$ and $\mathsf{awae}^{f',\pi_\Phi}_{\alpha,g}$ convey the amount of information that the attackers get when inputting a value $\mathbf{x}_\mathbb{A}$ regardless of their adversarial power. The study of these functions thus still informs us of the influence that a deceitful attacker may have in a computation. Finally, Assumption 4 will be critical for introducing the objective function of Definition 7.

A. Maximal and Optimal Distortion

We want to contain the distortion introduced by computing f^+ instead of f. Formally, for a given virtual input φ with distribution π_Φ, we will tolerate a certain distortion threshold Δ in $\mathbb{N}_{>0}$ that serves as upper bound for the maximal absolute difference $\xi(f, f^+)$ between the output of f and f^+, i.e. $\xi(f, f^+) \leq \Delta$ where:

$$\xi(f, f^+) \coloneqq \max_{\substack{\varphi \in \mathrm{supp}(\pi_\Phi) \\ \mathbf{x} \in D}} |f(\mathbf{x}) - f^+(\mathbf{x}, \varphi)| \tag{23}$$

and where $\mathrm{supp}(\pi_\Phi) \coloneqq \{\varphi \in D_\Phi \mid \pi_\Phi(\varphi) > 0\}$ denotes the support of π_Φ. For the additive approximation f^+ of f, we can see that $\xi(f, f^+)$ equals $\max\{|\varphi| \mid \varphi \in \mathrm{supp}(\pi_\Phi)\}$. Thus, we have:

$$\xi(f, f^+) \leq \Delta \iff \mathrm{supp}(\pi_\Phi) \subseteq [\![-\Delta, +\Delta]\!] \tag{24}$$

For such additive approximation f^+, our examples suggested that different distributions π_Φ for a virtual input φ can yield different privacy gains for the targeted inputs. We are thus interested in studying the influence of the distribution π_Φ of the virtual input on the behavior of the leakage measure $\mathsf{awae}^{f^+,\pi_\Phi}_{\alpha,g}$. To that end, we first want to evaluate how much privacy is being protected by f^+ and π_Φ within a distortion threshold Δ. We can do this through a metric, our objective function for optimization, which we define next. In accordance with Assumption 4, we consider that the attackers $\mathbb{A}$ are passive, and their inputs $\mathbf{x}_\mathbb{A}$ will thus be considered as a random variable following distribution $\pi_\mathbb{A}$ introduced in Assumption 1. Our objective function will thus be the weighted average of $\mathsf{awae}^{f^+,\pi_\Phi}_{\alpha,g}$ over all $\mathbf{x}_\mathbb{A}$.

Definition 7: Let α be in $\mathbb{R}_{>0} \cup \{\infty\}$. The function $\mathrm{obj}_{\alpha,g}\colon \Omega(D_\Phi) \to \mathbb{R}_{\geq 0}$ is defined, for all π_Φ in $\Omega(D_\Phi)$, as:

$$\mathrm{obj}_{\alpha,g}(\pi_\Phi) = \sum_{\mathbf{x}_\mathbb{A} \in D_\mathbb{A}} p(\mathbf{x}_\mathbb{A}) \cdot \mathsf{awae}^{f^+,\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \tag{25}$$

We observe that the information gained by a deceitful attacker would have been better conveyed by $\min_{\mathbf{x}_\mathbb{A}} \mathsf{awae}^{f^+,\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A})$ since such attackers could select the most informative $\mathbf{x}_\mathbb{A}$, but Assumption 4 ensures that the attackers $\mathbb{A}$ are passive.

The targeted parties in $\mathbb{T}$ — and perhaps others — now want to find a distribution π_Φ that will be optimal for this given metric, under the constraint that the distortion should remain below the threshold Δ. Entropies, as mathematical functions, are such that the larger their output is, the less do we actually know. Therefore, we mean to find a *global maximum* of the above metric, subject to the distortion-bound constraint. This ensures that an attacker has, on average, the least information gain for this from all possible virtual distributions. Using the equivalence in Equation (24), this naturally leads us to the following optimization problem.

Definition 8: Let Δ be in $\mathbb{N}_{>0}$, let α be in $\mathbb{R}_{>0} \cup \{\infty\}$. Then we denote by $\mathsf{OP}_{\alpha,g}(\Delta)$ the optimization problem:

$$\underset{\pi_\Phi \in \Omega([\![-\Delta,+\Delta]\!])}{\text{maximise}} \ \mathrm{obj}_{\alpha,g}(\pi_\Phi) \tag{26}$$

We write $\omega_{\alpha,g}$ for the optimal objective value in (26).

Note that this optimization problem can equivalently be expressed as optimizing the $2\Delta + 1$ values of distribution π_Φ:

$$\begin{aligned} \underset{(\pi_\Phi(i))_{-\Delta \leq i \leq +\Delta}}{\text{maximise}} \ & \mathrm{obj}_{\alpha,g}(\pi_\Phi) \\ \text{subject to} \ & \sum_{i \in [\![-\Delta,+\Delta]\!]} \pi_\Phi(i) = 1 \\ & \text{and } \forall i \in [\![-\Delta, +\Delta]\!] \colon 0 \leq \pi_\Phi(i) \leq 1 \end{aligned} \tag{27}$$

B. Computing Optimal Virtual Distributions

We now discuss methods for solving this optimization problem and computing optimal virtual distributions, where we distinguish between the cases in which α is ∞ or greater or equal to 1.

Optimal Virtual Input Randomization When $1 \leq \alpha < \infty$: For a gain function $g\colon \mathcal{W} \times D_\mathbb{T} \to [0, 1]$, let us study the objective function of $\mathsf{OP}_{\alpha,g}(\Delta)$. We recall that for all $\mathbf{x}_\mathbb{A}$ in $D_\mathbb{A}$ and for $\mathrm{V}_{\alpha,g}$ as defined in (9), we have:

$$\mathsf{awae}^{f^+,\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A}) = \frac{\alpha}{1-\alpha} \cdot \log\big(\mathrm{V}_{\alpha,g}(X_\mathbb{T} \mid O', \mathbf{x}_\mathbb{A})\big)$$

and where, for all $\mathbf{x}_\mathbb{T}$ in $D_\mathbb{T}$, the term $p(o' \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A})$ is a linear function of π_Φ, namely:

$$p(o' \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A}) = \sum_{\substack{\mathbf{x}_\mathbb{S}, \varphi \\ f^+(\mathbf{x}_\mathbb{A}, \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{S}, \varphi) = o'}} p(\mathbf{x}_\mathbb{S}) \cdot p(\varphi)$$

Below, we may write $p(o', \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A})[\pi_\Phi]$ for $p(o', \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A})$ in order to make this linear dependency on π_Φ explicit.

We thus have a non-linear and non-convex optimization problem with linear constraints and where the objective function is twice continuously differentiable almost everywhere. Sequential Quadratic Programming (SQP) [52], [53] would thus seem like an adequate and simple solution for finding a local optimum for our optimization problem. However, SQP requires the constraints and the objective function to be twice continuously differentiable, which is not the case of our objective function: for all $\alpha > 1$ and all integer $n > 1$, the function $y \mapsto \|y\|_\alpha$ is not differentiable at the origin even when restricted to $(\mathbb{R}_{\geq 0})^n \to \mathbb{R}_{\geq 0}$. Consequently, our objective function is not differentiable at the points $\pi_{\Phi 0}$ in $\Omega(D_\Phi)$ such that $\pi_{\Phi 0}$ makes $p(o', \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A})$ be 0 but where $p(o', \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A})$ is not always 0, i.e., when:

$$\begin{aligned} &\big(p(o', \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A})[\pi_{\Phi 0}] = 0\big) \\ &\qquad \wedge \big(\exists \pi_{\Phi 1} \in \Omega(D_\Phi) \colon p(o', \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A})[\pi_{\Phi 1}] > 0\big) \end{aligned}$$

We will address this by smoothening the objective function in (25) through a non-zero offset vector $\boldsymbol{\delta}$ in $(\mathbb{R}_{\geq 0})^{|D_\mathbb{T}|}$ that

is added to the argument of the α-norm — the expression in (9) with O' instead of O. This approximation is then twice continuously differentiable everywhere. We introduce some definitions for formalizing this:

Definition 9: Let α be in $]1,\infty[$. Let $\boldsymbol{\delta} \neq 0$ be in $(\mathbb{R}_{\geq 0})^{|D_{\mathbb{T}}|}$.

1) *Let π_Φ be in $\Omega(D_\Phi)$. For all $\mathbf{x}_\mathbb{A}$ in $D_\mathbb{A}$, we define:*

$$\mathrm{V}^{\boldsymbol{\delta}}_{\alpha,g}(X_\mathbb{T} \mid O', \mathbf{x}_\mathbb{A}) \coloneqq \sum_{o'} \left\| \boldsymbol{\delta} + \left\langle \sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o' \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A}) \cdot g(w, \mathbf{x}_\mathbb{T}) \right\rangle_w \right\|_\alpha \tag{28}$$

$$\mathsf{awae}^{f^+,\pi_\Phi,\boldsymbol{\delta}}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \coloneqq \frac{\alpha}{1-\alpha} \cdot \log\left(\mathrm{V}^{\boldsymbol{\delta}}_{\alpha,g}(X_\mathbb{T} \mid O', \mathbf{x}_\mathbb{A})\right) \tag{29}$$

2) *We define the function $\mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g} \colon \Omega(D_\Phi) \to \mathbb{R}_{\geq 0}$ for all π_Φ in $\Omega(D_\Phi)$ as:*

$$\mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}(\pi_\Phi) \coloneqq \sum_{\mathbf{x}_\mathbb{A} \in D_\mathbb{A}} p(\mathbf{x}_\mathbb{A}) \cdot \mathsf{awae}^{f^+,\pi_\Phi,\boldsymbol{\delta}}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \tag{30}$$

For Δ in $\mathbb{N}_{>0}$, we define $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$ as the following optimization problem:

$$\underset{\pi_\Phi \in \Omega([\![-\Delta,+\Delta]\!])}{\text{maximise}} \quad \mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}(\pi_\Phi) \tag{31}$$

We write $\omega^{\boldsymbol{\delta}}_{\alpha,g}$ for the global maximum of the optimization problem $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$.

Using the above optimization problems, we are now able to approximate the result of the original problem in (26) with an arbitrary accuracy by choosing the value of $\boldsymbol{\delta}$. We formalize this next:

Theorem 3: Let α be in $]1,\infty[$. Let g be a β-positive gain function (as defined in Definition 1). Let Δ be in $\mathbb{N}_{>0}$ and let $\boldsymbol{\delta}$ be the vector in $(\mathbb{R}_{>0})^{|D_\mathbb{T}|}$ whose $|D_\mathbb{T}|$ components all equal δ in $\mathbb{R}_{>0}$. Then, for all ε in $\mathbb{R}_{>0}$, we have:

$$\left(\delta \leq (1-\frac{1}{\alpha}) \cdot \frac{\varepsilon \cdot \beta \cdot \ln(2)}{|D_{O'}| \cdot |\mathcal{W}|}\right) \implies \left(|\omega_{\alpha,g} - \omega^{\boldsymbol{\delta}}_{\alpha,g}| \leq \varepsilon\right) \tag{32}$$

where $\ln(2)$ *refers to the natural logarithm of* 2.

Proof: Let π_Φ be in $\Omega(D_\Phi)$, let o' be in $D_{O'}$, and let $\mathbf{x}_\mathbb{A}$ be in $D_\mathbb{A}$ For sake of convenience, let us define the vector:

$$W^{o'}_{\mathbf{x}_\mathbb{A}} \coloneqq \left\langle \sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o' \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A}) \cdot g(w, \mathbf{x}_\mathbb{T}) \right\rangle_w$$

First, as all the components of the vectors are non-negative, we have:

$$\left\| W^{o'}_{\mathbf{x}_\mathbb{A}} + \boldsymbol{\delta} \right\|_\alpha \geq \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha$$

Since α is greater than 1, we know that $\frac{\alpha}{1-\alpha}$ is negative, and thus:

$$\mathsf{awae}^{f^+,\pi_\Phi,\boldsymbol{\delta}}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \leq \mathsf{awae}^{f^+,\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A})$$

Moreover, application of the triangular inequality yields:

$$\sum_{o'} \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} + \boldsymbol{\delta} \right\|_\alpha \leq \sum_{o'} \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha + \sum_{o'} \|\boldsymbol{\delta}\|_\alpha$$

Applying logarithm and multiplying by $\frac{\alpha}{1-\alpha}$ on both sides, we obtain:

$$\mathsf{awae}^{f^+,\pi_\Phi,\boldsymbol{\delta}}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \geq \frac{\alpha}{1-\alpha} \cdot \log\left(\sum_{o'} \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha + \sum_{o'} \|\boldsymbol{\delta}\|_\alpha\right)$$

However, for all a and b in $\mathbb{R}_{>0}$, we have $\log(a+b) = \log(a) + \log(1+\frac{b}{a})$. Therefore, we conclude that:

$$\mathsf{awae}^{f^+,\pi_\Phi,\boldsymbol{\delta}}_{\alpha,g}(\mathbf{x}_\mathbb{A}) \geq \mathsf{awae}^{f^+,\pi_\Phi}_{\alpha,g}(\mathbf{x}_\mathbb{A}) + \frac{\alpha}{1-\alpha} \cdot \log\left(1 + \frac{\sum_{o'} \|\boldsymbol{\delta}\|_\alpha}{\sum_{o'} \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha}\right)$$

Rearranging the terms and summing over $\mathbf{x}_\mathbb{A}$ gives us:

$$\mathrm{obj}_{\alpha,g}(\pi_\Phi) - \mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}(\pi_\Phi) \leq \sum_{\mathbf{x}_\mathbb{A}} p(\mathbf{x}_\mathbb{A}) \cdot \frac{\alpha}{\alpha-1} \cdot \log\left(1 + \frac{\sum_{o'} \|\boldsymbol{\delta}\|_\alpha}{\sum_{o'} \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha}\right)$$

Moreover, for all x in $\mathbb{R}_{>0}$, we know that $\log(1+x) \leq x/\ln(2)$. Thus, we infer:

$$\mathrm{obj}_{\alpha,g}(\pi_\Phi) - \mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}(\pi_\Phi) \leq \sum_{\mathbf{x}_\mathbb{A}} p(\mathbf{x}_\mathbb{A}) \cdot \frac{\alpha}{\alpha-1} \cdot \frac{\sum_{o'} \|\boldsymbol{\delta}\|_\alpha}{\ln(2) \cdot \sum_{o'} \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha} \tag{33}$$

Furthermore, for all p in $[1,\infty]$ and n in $\mathbb{N}_{>0}$, we get from the topological equivalence of the norms in finite dimension that for all x in $\mathbb{R}^n$, we have $\|x\|_p \geq \|x\|_1 \cdot n^{\frac{1}{p}-1}$. Therefore:

$$\left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha \geq \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_1 \cdot |\mathcal{W}|^{\frac{1}{\alpha}-1}$$

Now, we know that:

$$\left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_1 = \sum_w \sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o' \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A}) \cdot g(w, \mathbf{x}_\mathbb{T}) = \sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T}) \cdot p(o' \mid \mathbf{x}_\mathbb{T}, \mathbf{x}_\mathbb{A}) \cdot \left(\sum_w g(w, \mathbf{x}_\mathbb{T})\right)$$

Since g is β-positive, we obtain:

$$\sum_{o'} \left\| W^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha \geq \beta \cdot |\mathcal{W}|^{\frac{1}{\alpha}-1}$$

On the other hand, by definition of $\boldsymbol{\delta}$ we have:

$$\sum_{o'} \|\boldsymbol{\delta}\|_\alpha = \delta \cdot |D_{O'}| \cdot |\mathcal{W}|^{\frac{1}{\alpha}}$$

and thus Equation (33) becomes:

$$\mathrm{obj}_{\alpha,g}(\pi_\Phi) - \mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}(\pi_\Phi) \leq \sum_{\mathbf{x}_\mathbb{A}} p(\mathbf{x}_\mathbb{A}) \cdot \frac{\alpha}{\alpha-1} \cdot \frac{\delta \cdot |D_{O'}| \cdot |\mathcal{W}|^{\frac{1}{\alpha}}}{\ln(2) \cdot \beta \cdot |\mathcal{W}|^{\frac{1}{\alpha}-1}} \leq \frac{\alpha}{\alpha-1} \cdot \frac{\delta \cdot |D_{O'}| \cdot |\mathcal{W}|}{\ln 2 \cdot \beta} \tag{34}$$

Consider now any ε in $\mathbb{R}_{>0}$. In order for $\mathrm{obj}_{[\alpha,g]}(\pi_\Phi) - \mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}(\pi_\Phi)$ not to exceed ε, Equation (34) ensures that it suffices to have:

$$\delta \leq (1-\frac{1}{\alpha}) \cdot \frac{\varepsilon \cdot \beta \cdot \ln(2)}{|D_{O'}| \cdot |\mathcal{W}|} \tag{35}$$

Finally, the reverse triangle inequality applied on functions $\mathrm{obj}_{\alpha,g}$ and $\mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}$ with the uniform norm yields:

$$|\omega_{\alpha,g} - \omega^{\boldsymbol{\delta}}_{\alpha,g}| \leq \max_{\pi_\Phi} |\,\mathrm{obj}_{\alpha,g}(\pi_\Phi) - \mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}(\pi_\Phi)|$$

and thus the condition in (35) implies $|\omega_{\alpha,g} - \omega^{\boldsymbol{\delta}}_{\alpha,g}| < \varepsilon$. □

The last theorem states that, if we are able to solve the optimization problem $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$ for any non-zero offset vector $\boldsymbol{\delta}$ in $\mathbb{R}^{|D_\mathbb{T}|}_{\geq 0}$, then we are able to approximate the optimal outcome of the original optimization problem $\mathsf{OP}_{\alpha,g}(\Delta)$ with arbitrary precision. We now present a method for solving the approximate optimization problems $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$.

Method 1: Let us consider the optimization problem $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$ *of (31) where* α *is in* $]1,\infty[$ *The objective function* $\mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}$ *is twice differentiable and the constraints are linear. Thus, we may apply SQP [52], [53] to find a local optimum for* $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$. *However, as the objective function* $\mathrm{obj}^{\boldsymbol{\delta}}_{\alpha,g}$ *is non-convex, we will use a globalization technique known as the basin-hopping algorithm [54]. In order to respect the linear constraints of this problem, the starting points of this algorithm will be drawn from a symmetric Dirichlet distribution.*

This computational method lets us solve optimization problems of the form $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$. Consequently, Theorem 3 enables us to build a method for solving our original optimization problem $\mathsf{OP}_{\alpha,g}(\Delta)$ with an arbitrary precision ε, which we formalize in the next method:

Method 2: We seek a solution of the optimization problem $\mathsf{OP}_{\alpha,g}(\Delta)$ *where* α *is in* $[1,\infty[$ *and* g *is a* β*-positive gain function.*

Case $\alpha > 1$*: We will solve* $\mathsf{OP}_{\alpha,g}(\Delta)$ *with a given accuracy* $\varepsilon > 0$*. In other words, a solution* π_Φ *should satisfy* $|\,\mathrm{obj}_{\alpha,g}(\pi_\Phi) - \omega_{\alpha,g}| \leq \varepsilon$*. First, let us choose* δ *in* $\mathbb{R}_{>0}$ *such that:*

$$\delta \leq (1 - \frac{1}{\alpha}) \cdot \frac{\varepsilon \cdot \beta}{|D_{O'}| \cdot |\mathcal{W}|}$$

Let $\boldsymbol{\delta}$ *be the vector in* $\mathbb{R}^{|D_\mathbb{T}|}$ *whose components all equal* δ*. We apply Method 1 in order to solve the optimization problem* $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$ *. Let* π_Φ *be the solution output by Method 1. By virtue of Theorem 3, we have* $|\,\mathrm{obj}_{\alpha,g}(\pi_\Phi) - \omega_{\alpha,g}| \leq \varepsilon$.

Case $\alpha = 1$*: Let* g *be unitary. We can solve* $\mathsf{OP}_{1,g}(\Delta)$ *using the same procedure as that of Method 1 since the objective function* $\mathrm{obj}_{1,g}$ *and the constraints of the problem are twice continuously differentiable.*

Now that we are able to solve the optimization problem $\mathsf{OP}_{\alpha,g}(\Delta)$ when $\alpha \geq 1$ is finite, we now turn our attention to the case of $\alpha = \infty$. In the same way as Method 2 builds on Method 1 to approximate a solution, our next idea will be to approximate the optimal result of $\mathsf{OP}_{\infty,g}(\Delta)$ with a multiple of that of $\mathsf{OP}_{\alpha,g}(\Delta)$ for a sufficiently large α.

First, we introduce $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$, a slightly modified version of $\mathsf{OP}_{\alpha,g}(\Delta)$ whose objective function is a multiple of $\mathrm{obj}_{\alpha,g}$. Then, we prove that the solutions of $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$ converge towards a solution of $\mathsf{OP}_{\infty,g}(\Delta)$. Moreover, we make the convergence rate explicit for computational purposes. We define $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$ next:

Definition 10: Let α *be in* $]1,\infty[$.

1) *We define the function* $\overline{\mathrm{obj}}_{\alpha,g} \colon \Omega(D_\Phi) \to \mathbb{R}_{\geq 0}$ *for all* π_Φ *in* $\Omega(D_\Phi)$ *as:*

$$\overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi) \coloneqq \frac{\alpha - 1}{\alpha} \cdot \mathrm{obj}_{\alpha,g}(\pi_\Phi) \tag{36}$$

2) *For* Δ *in* $\mathbb{N}_{>0}$*, the optimization problem* $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$ *is:*

$$\underset{\pi_\Phi \in \Omega([\![-\Delta,+\Delta]\!])}{\text{maximise}} \ \overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi) \tag{37}$$

We write $\overline{\omega}_{\alpha,g}$ *denote the optimal objective value for* $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$.

From this definition it is clear that $\mathsf{OP}_{\alpha,g}(\Delta)$ and $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$ are equivalent optimization problems, in that:

$$\overline{\omega}_{\alpha,g} = \frac{\alpha - 1}{\alpha} \cdot \omega_{\alpha,g} \tag{38}$$

More precisely, the optimal values of $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$ under-approximate that of $\mathsf{OP}_{\infty,g}(\Delta)$, with an error rate dominated by $\frac{1}{\alpha}$:

Theorem 4: Let the functions $\tau, \theta \colon]1,+\infty[\to \mathbb{R}_{\geq 0}$ *be defined as* $\tau(\alpha) = \overline{\omega}_{\alpha,g}$ *and* $\theta(\alpha) = |\omega_{\infty,g} - \overline{\omega}_{\alpha,g}|$*. Then, for all* $\alpha > 1$*, we have* $\tau(\alpha) \leq \omega_{\infty,g}$*,* $\lim_{\alpha\to\infty} \tau(\alpha) = \omega_{\infty,g}$*, and* $\theta(\alpha) = \mathcal{O}(\frac{1}{\alpha})$.

Proof: Let π_Φ be in $\Omega(\mathbb{Z})$. To simplify notation, we define the vector $Y^{o'}_{\mathbf{x}_\mathbb{A}}$ for all $\mathbf{x}_\mathbb{A}$ in $D_\mathbb{A}$ and o' in $D_{O'}$ as:

$$Y^{o'}_{\mathbf{x}_\mathbb{A}} \coloneqq \left\langle \sum_{\mathbf{x}_\mathbb{T}} p(\mathbf{x}_\mathbb{T} \mid o', \mathbf{x}_\mathbb{A}) \cdot g(w, \mathbf{x}_\mathbb{T}) \right\rangle_w$$

For all α in $]1,\infty]$ and π_Φ in D_Φ, we have by definition that $\overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi)$ equals:

$$\overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi) = -\sum_{\mathbf{x}_\mathbb{A}} p(\mathbf{x}_\mathbb{A}) \cdot \log\left(\sum_{o'} p(o' \mid \mathbf{x}_\mathbb{A}) \cdot \left\| Y^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha \right)$$

We know that in finite dimension, all the norms are topologically equivalent. In particular, for all n in $\mathbb{N}_{>0}$, x in $\mathbb{R}^n$, and p in $]1,\infty[$, we have:

$$\|x\|_\infty \leq \|x\|_p \leq \|x\|_\infty \cdot n^{\frac{1}{p}}$$

Let α be in $]1,\infty[$. We thus have:

$$\left\| Y^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\infty \leq \left\| Y^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\alpha \leq \left\| Y^{o'}_{\mathbf{x}_\mathbb{A}} \right\|_\infty \cdot |\mathcal{W}|^{\frac{1}{\alpha}}$$

and thus:

$$\mathrm{obj}_{\infty,g}(\pi_\Phi) - \frac{1}{\alpha} \cdot \log(|\mathcal{W}|) \leq \overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi) \leq \mathrm{obj}_{\infty,g}(\pi_\Phi)$$

From this inequality, we can see that $\tau(\alpha) \leq \omega_{\infty,g}$ for all $\alpha > 1$. Moreover:

$$\begin{aligned} \theta(\alpha) &\leq \mathrm{obj}_{\infty,g}(\pi_\Phi) - \left(\mathrm{obj}_{\infty,g}(\pi_\Phi) - \frac{1}{\alpha} \cdot \log(|\mathcal{W}|) \right) \\ &\leq \frac{1}{\alpha} \cdot \log(|\mathcal{W}|) \end{aligned}$$

and thus $\theta(\alpha) = \mathcal{O}(\frac{1}{\alpha})$. In particular, θ converges to 0 as α goes to infinity, which ensures that τ converges at infinity such that $\lim_{\tau\to\infty} \tau(\alpha) = \omega_{\infty,g}$. Moreover, for any ε in $\mathbb{R}_{>0}$, in order to have $\theta(\alpha) \leq \varepsilon$, it suffices to have:

$$\alpha \geq \frac{1}{\varepsilon} \log(|\mathcal{W}|)$$

□

From this theorem, we can build a method for solving the optimization problems of the form $\mathsf{OP}_{\infty,g}(\Delta)$. Indeed, even though the objective function $\mathrm{obj}_{\infty,g}$ is not twice differentiable, we can approximate the solution of that optimization problem with that of $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$ for a sufficiently large α. We recall that, by (38), the optimal value of the latter problem is a multiple of that of $\mathsf{OP}_{\alpha,g}(\Delta)$, which we can solve with Method 2. However, Method 2 also requires a non-zero accuracy threshold. Thus, for a given ε in $\mathbb{R}_{>0}$, we will invoke Method 2 in order to solve $\mathsf{OP}_{\alpha,g}(\Delta)$ with accuracy $\frac{\alpha}{\alpha-1}\cdot\frac{\varepsilon}{2}$, and we will take advantage of Theorem 4 to ensure that the output of our method indeed approximates the optimal objective value with precision ε. We formalize this idea next:

Method 3: Let ε be in $\mathbb{R}_{>0}$ and let us assume that we wish to solve $\mathsf{OP}_{\infty,g}(\Delta)$ with accuracy ε, i.e. that the solution π_Φ we get satisfies $|\,\mathrm{obj}_{\infty,g}(\pi_\Phi)-\omega_{\infty,g}|\leq\varepsilon$. First, we take some $\alpha>1$ which satisfies:

$$\alpha \geq \frac{2}{\varepsilon}\cdot\log(|\mathcal{W}|) \tag{39}$$

Then, we invoke Method 2 in order to solve $\mathsf{OP}_{\alpha,g}(\Delta)$ with accuracy $\frac{\alpha}{\alpha-1}\cdot\frac{\varepsilon}{2}$. Let π_Φ be an optimal solution for this produced by Method 2. Then, π_Φ is an optimal solution of $\mathsf{OP}_{\infty,g}(\Delta)$ with accuracy ε.

Proof: As π_Φ is the output of Method 2, we know that $|\,\mathrm{obj}_{\alpha,g}(\pi_\Phi)-\omega_{\alpha,g}|\leq\frac{\alpha}{\alpha-1}\cdot\frac{\varepsilon}{2}$. Multiplying both sides by $\frac{\alpha-1}{\alpha}$ yields:

$$|\overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi)-\overline{\omega}_{\alpha,g}|\leq\frac{\varepsilon}{2}$$

Moreover, by virtue of Theorem 4 and as we have Equation (39), we know that $|\overline{\omega}_{\alpha,g}-\omega_{\alpha,g}|\leq\frac{\varepsilon}{2}$. Finally, we know that:

$$\overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi)\leq\mathrm{obj}_{\infty,g}(\pi_\Phi)\leq\omega_{\infty,g}$$

Appealing to the triangular inequality, we then get:

$$\begin{aligned}|\,\mathrm{obj}_{\infty,g}(\pi_\Phi)-\omega_{\infty,g}| &\leq |\overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi)-\omega_{\infty,g}|\\ &\leq |\overline{\mathrm{obj}}_{\alpha,g}(\pi_\Phi)-\overline{\omega}_{\alpha,g}|\\ &\quad+|\overline{\omega}_{\alpha,g}-\omega_{\infty,g}|\\ &\leq\varepsilon\end{aligned}$$

□

The following example illustrates how the solution of $\mathsf{OP}_{\infty,g}(\Delta)$ is approximated by the successive solutions of $\overline{\mathsf{OP}}_{\alpha,g}(\Delta)$ for different values of α. It is worth noting that the calculation of α-norms involves the exponentiation of real numbers ranged in $[0,1]$ which can quickly be rounded to 0 for large values of α. In order to mitigate against the effects of such numerical errors, results reports in this paper rely on use of the `mpmath` Python library, which enables us to perform arbitrary-precision floating-point arithmetic [55].

Example 4: Let us consider 3 parties X, Y and Z with respective inputs x, y and z, and where $\mathbb{A}=\{X\}$ is attacking $\mathbb{T}=\{Y\}$ under spectator $\mathbb{S}=\{Z\}$. Let $D_\mathbb{A}=D_\mathbb{T}=D_\mathbb{S}=\{1,2\}$. Let $\pi_\mathbb{T}$ and $\pi_\mathbb{S}$ be linear distributions over their domains and let $\pi_\mathbb{A}=\{1\colon 1\}$ be a point-mass distribution centred in 1. Function $f:\mathbb{Z}^3\to\mathbb{Z}$ is defined by $f(x,y,z)=5xy-2yz$.

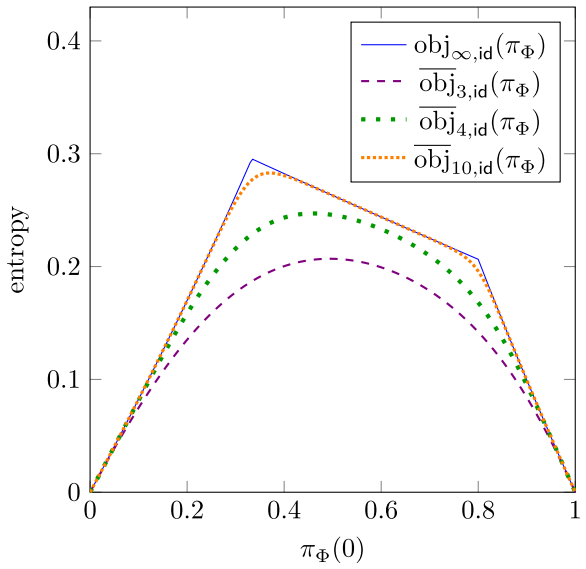

Fig. 6. Approximation of $\mathrm{obj}_{\infty,\mathsf{id}}$ by $\overline{\mathrm{obj}}_{p,\mathsf{id}}$ for p in $\{3,4,10\}$ while computing $f(x,y,z)=5xy-2yz$ with linear distributions over $\{1,2\}$ for $\pi_\mathbb{T}$ and $\pi_\mathbb{S}$ and $\pi_\mathbb{A}=\{1\colon 1\}$.

We study the influence of distributions π_Φ for virtual inputs over $\Omega(\{0,1\})$ on $\mathrm{obj}_{\infty,\mathsf{id}}$ produced by the output randomization f^+. Such two-dimensional distributions π_Φ will be represented by a single real r in $[0,1]$, which fully characterizes π_Φ as $\{0\colon r, 1\colon 1-r\}$. We evenly discretize the interval $[0,1]$ into 201 values for r, and we plot the values of $\mathrm{obj}_{\infty,\mathsf{id}}$ in Figure 6. In order to see the influence of our smoothing method, we also plot the values of $\overline{\mathrm{obj}}_{\alpha,\mathsf{id}}$ for different values of α. We can notice that, as suggested by our previous discussion and by Theorem 4, the approximations $\overline{\mathrm{obj}}_{\alpha,\mathsf{id}}$ of $\mathrm{obj}_{\infty,\mathsf{id}}$ are functions that are twice differentiable and that also under-approximate $\mathrm{obj}_{\infty,\mathsf{id}}$. Moreover, larger values of α produce more accurate approximations of the original objective function.

Let us now illustrate how the methods we developed here help us to find virtual distributions that protect the inputs' privacy optimally, given some accuracy bound on the distorted output.

Example 5: Let us consider 3 parties X, Y and Z with respective inputs x, y and z, and where $\mathbb{A}=\{X\}$ is attacking $\mathbb{T}=\{Y\}$ under spectator $\mathbb{S}=\{Z\}$. Let $D_\mathbb{A}=D_\mathbb{T}=D_\mathbb{S}=[\![1,30]\!]$. Let $\pi_\mathbb{T}$ and $\pi_\mathbb{S}$ be linear distributions over their domains and for the sake of the example, let $\pi_\mathbb{A}=\{5\colon 1\}$ be a point-mass distribution centred in 5. Let us consider the function $f:\mathbb{Z}^3\to\mathbb{Z}$ defined by $f(x,y,z)=x(3y-5z)+2z$. Let $\mathcal{W}=\{0,1\}$ be a set of allowable guesses and let $g:\mathcal{W}\times D_\mathbb{T}\to[0,1]$ be the gain function defined for all w in $\mathcal{W}$ and $\mathbf{x}_\mathbb{T}$ in $D_\mathbb{T}$ as:

$$g(w,\mathbf{x}_\mathbb{T})=\begin{cases}1 & \text{if } w\equiv\mathbf{x}_\mathbb{T} \mod 2\\ 0 & \text{otherwise}\end{cases}$$

In other words, this gain function g measures the information that an attacker has on the least significant bit of the secret $\mathbf{x}_\mathbb{T}$. More generally, we can consider other gain functions that could gauge the information that an attacker

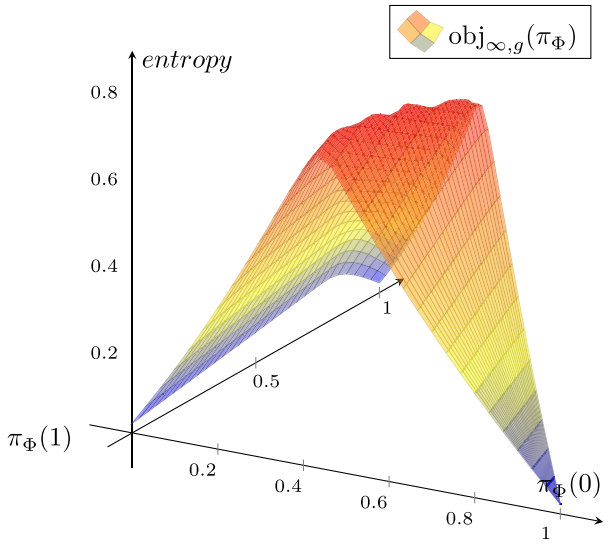

Fig. 7. Influence of π_Φ in $\Omega([\![-1,+1]\!])$ in the optimization problem $\mathsf{OP}_{\infty,g}(1)$ in the computation of $f(x,y,z) = x(3y-5z)+2z$ with linear distributions over $[\![1,30]\!]$ for $\pi_{\mathbb{T}}$ and $\pi_{\mathbb{S}}$ and $\pi_{\mathbb{A}} = \{5\colon 1\}$.

learns on a particular property of a secret. We note that g is β-positive with $\beta = 1$.

In comparison to Example 4, a distribution π_Φ in $\Omega([\![-1,1]\!])$ will now be characterized by two variables $\pi_\Phi(0)$ and $\pi_\Phi(1)$ since then $\pi_\Phi(-1) = 1-\pi_\Phi(0)-\pi_\Phi(1)$ will be fixed. The first variable $\pi_\Phi(0)$ will take its values in $[0,1]$ while the second one $\pi_\Phi(1)$ will take its values in $[0, 1-\pi_\Phi(0)]$. We discretize the interval $[0,1]$ into 101 values so that $\pi_\Phi(0)$ was assigned these values consecutively. For each of these values of $\pi_\Phi(0)$, the interval $[0, 1-\pi_\Phi(0)]$ is furthermore discretized into 101 values that $\pi_\Phi(1)$ took consecutively. For each pair $(\pi_\Phi(0), \pi_\Phi(1))$, we compute the value of $\text{obj}^g_\infty(\pi_\Phi)$ for the corresponding π_Φ, and we plot the resulting graph in Figure 7.

Let us now solve the optimization problem $\mathsf{OP}_{\infty,g}$ with accuracy $\varepsilon = 10^{-2}$ through Method 3. Here, $|\mathcal{W}|$ equals 2. Let us then take $\alpha = \frac{2}{\varepsilon}\cdot\log(|\mathcal{W}|) = 200$. We then invoke Method 2 to solve $\mathsf{OP}_{\alpha,g}(\Delta)$ with accuracy $\varepsilon' = \frac{\alpha}{\alpha-1}\cdot\frac{\varepsilon}{2} = 5.0\cdot 10^{-3}$. Moreover, a combinatorial calculation gives us $|D_{O'}| = 5656$. We thus let:

$$\beta = (1-\frac{1}{\alpha})\frac{\varepsilon'\mu\cdot\ln 2}{|D_{O'}|\cdot|\mathcal{W}|} = 3.0\cdot 10^{-7}$$

and we let $\boldsymbol{\delta}$ be the vector in $\mathbb{R}^2_{>0}$ whose components are all equal to β. Finally, we invoke Method 1 to solve $\mathsf{OP}^{\boldsymbol{\delta}}_{\alpha,g}(\Delta)$. This produces a nearly optimal solution $\pi_{\Phi o} = \{-1\colon 0.30, 0\colon 0.49, 1\colon 0.21\}$ for which $\text{obj}_{\infty,g}(\pi_{\Phi o})$ equals 0.77. This ensures that $\omega_{\infty,g}$ is in $[0.77, 0.78]$ while a uniform distribution $\pi_{\Phi u}$ over $\{-1,0,1\}$ would have only $\text{obj}_{\infty,g}(\pi_{\Phi u}) = 0.56$.

X. Discussion and Future Works

In this work, we proposed an approach for quantifying the information that attackers can retrieve about private inputs from public outputs in black-box computations of a public function. We also developed concepts and methods for mitigating against such information leakage, by distorting the public function with virtual, private inputs: we introduced some methods for maximizing the posterior entropy of the targeted inputs, and developed non-linear optimization techniques that can compute virtual inputs that optimally trade off the privacy protection stemming from virtual inputs and the accuracy of the distorted output in comparison with the un-distorted output.

Our approach is generic in that, depending on the nature of the inputs and on the use context of the secure computation, the participants can agree on a particular type of entropy to maximize before entering the optimization protocol. Participants may also want their inputs to be protected with respect to different kinds of entropy, and this could lead us to study multi-objective optimization and Pareto optimality — a topic for future work. The quantities and distortions that our approach can compute may also inform the risk management of using SMC for the same function repeatedly, with potentially different but related inputs — such as the logging of daily health data.

In a practical secure computation, once an optimal virtual distribution π_Φ has been computed by our methods for a given type of entropy, the participants of the SMC would have to securely produce a virtual input drawn from distribution π_Φ. For example, the parties may enter an SMC protocol in order to produce a value φ that is secret to all the participants, and that follows distribution π_Φ. To that end, parties may generate locally shares of a virtual input such that the value obtained by the combination of these shares follows the specified distribution. Alternatively, it may also be practical to let a central authority compute the virtual inputs — and these virtual inputs could then be fed into SMC protocols in addition to the x_i as seen on the right of Figure 3. For example, if parties are health insurance providers, then the computation of virtual inputs by a central authority does not require any proof of compliance with health and data regulations, since the insurance providers would not share sensitive health data with that central authority. Designing such secure protocols is subject to future works.

Our work considered the prior beliefs on the inputs to be public, constant, and part of the common knowledge. In SMC, this would enable participants to come to a consensus in order to agree on a common optimal virtual distribution π_Φ and to securely compute the output of f'. In comparison to the setting of SMC which assumes that participants have agreed on an actively or passively secure protocol to use, our setting assumes that participants will agree on an approximate function f' and a virtual distribution π_Φ that protects the targets' privacy. In the case of outsourced computations, those public distributions could simply be used by a trusted third party in order to produce a virtual input drawn from π_Φ and randomize the computation of f.

On the other hand, it would be of interest to relax these assumptions. In particular, computing an optimal virtual distribution π_Φ requires having a prior belief $\pi_{\mathbb{A}}$ on the attackers' input. Distribution π_Φ would then maximize the targeted inputs' privacy given the prior belief $\pi_{\mathbb{A}}$. But as the computation of π_Φ can be performed offline by any of the parties, this could enable an attacker to substitute his input

accordingly. This could thus update the belief $\pi_{\mathbb{A}}$ and we would require another computation of π_{Φ}. The setting where two attackers would try to learn information about each other's input could also lead to interesting game-theoretic situations to be studied in future work.

We also assumed that the partition of the participants into attackers, targets and spectators was given, but it would be of interest to develop techniques that can protect the participants' inputs when the set of potential attackers is not known. Moreover, we would like to further generalize our approach to the secure computation of vector-valued functions, i.e. of functions that compute several outputs, and where each of the outputs can be opened to different sets of parties. Finally, scaling our approach to large input spaces is also one of our future research objectives.

XI. Conclusion

Although efficient SMC protocols have been designed, information flow of outputs is inevitable, and has recently been rigorously formalized and quantified [15]. In this work, we first proposed a generalized notion of entropy that makes our approach compatible with various widely used measures of information flow. We then introduced the concepts of function substitution and virtual input that aim at randomizing the output of SMC computations in order to impede the influence of deceitful attackers wishing to use input substitution to gain maximal information about private inputs from opened outputs. We have established some theoretical bounds for the privacy gain that approximations and close approximations provide. We then focused on additive approximations and formalized an optimization problem that aims at maximizing participants' privacy while controlling the distortion introduced on the output in the presence of passive adversaries. We proposed different methods for solving such optimization problems in practice and we experimentally showed that additive approximations give rise to significant privacy gains under specified distortion bounds.

Acknowledgements

The authors would like to thank anonymous reviewers for their insightful comments which considerably improved our paper.

References

[1] A. C.-C. Yao, "How to generate and exchange secrets," in *Proc. 27th Annu. Symp. Found. Comput. Sci.*, Oct. 1986, pp. 162–167.

[2] A. C. Yao, "Protocols for secure computations," in *Proc. 23rd Annu. Symp. Found. Comput. Sci. (SFCS)*, Nov. 1982, pp. 160–164.

[3] A. Shamir, "How to share a secret," *Commun. ACM*, vol. 22, no. 11, pp. 612–613, Nov. 1979.

[4] T. Rabin and M. Ben-Or, "Verifiable secret sharing and multiparty protocols with honest majority," in *Proc. 21st Annu. ACM Symp. Theory Comput.*, 1989, pp. 73–85.

[5] M. Ben-Or, S. Goldwasser, and A. Wigderson, "Completeness theorems for non-cryptographic fault-tolerant distributed computation," in *Proc. 20th Annu. ACM Symp. Theory Comput.*, 1988, pp. 1–10.

[6] D. Chaum and C. Crépeau, and I. Damgard, "Multiparty unconditionally secure protocols," in *Proc. 20th Annu. ACM Symp. Theory Comput.*, 1988, pp. 11–19.

[7] V. Kolesnikov and T. Schneider, "Improved garbled circuit: Free XOR gates and applications," in *International Colloquium on Automata, Languages, and Programming* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2008, pp. 486–498.

[8] Y. Lindell, B. Pinkas, N. P. Smart, and A. Yanai, "Efficient constant round multi-party computation combining BMR and SPDZ," in *Advances in Cryptology*. Berlin, Germany: Springer, 2015, pp. 319–338.

[9] Y. Lindell and B. Pinkas, "Secure two-party computation via cut-and-choose oblivious transfer," *J. Cryptol.*, vol. 25, no. 4, pp. 680–722, 2012.

[10] T. Araki, J. Furukawa, Y. Lindell, A. Nof, and K. Ohara, "High-throughput semi-honest secure three-party computation with an honest majority," in *Proc. SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 805–817.

[11] Y. Lindell and B. Pinkas, "Secure multiparty computation for privacy-preserving data mining," *J. Privacy Confidentiality*, vol. 1, no. 1, p. 5, 2009.

[12] C. Orlandi, "Is multiparty computation any good in practice?" in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 5848–5851.

[13] R. Cramer, I. B. Damgård, and J. B. Nielsen, *Secure Multiparty Computation and Secret Sharing*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[14] Y. Aumann and Y. Lindell, "Security against covert adversaries: Efficient protocols for realistic adversaries," in *Theory of Cryptography*. Berlin, Germany: Springer, 2007, pp. 137–156.

[15] P. Ah-Fat and M. Huth, "Secure multi-party computation: Information flow of outputs and game theory," in *Proc. Int. Conf. Princ. Secur. Trust*. Berlin, Germany: Springer, 2017, pp. 71–92.

[16] D. E. Denning, "A lattice model of secure information flow," *Commun. ACM*, vol. 19, no. 5, pp. 236–243, 1976.

[17] D. Volpano, C. Irvine, and G. Smith, "A sound type system for secure flow analysis," *J. Comput. Secur.*, vol. 4, nos. 2–3, pp. 167–187, 1996.

[18] R. Joshi and K. R. M. Leino, "A semantic approach to secure information flow," *Sci. Comput. Program.*, vol. 37, nos. 1–3, pp. 113–138, 2000.

[19] G. Smith, "Principles of secure information flow analysis," in *Malware Detection*. Berlin, Germany: Springer, 2007, pp. 291–307.

[20] C. Dima, C. Enea, and R. Gramatovici, "Nondeterministic nointerference and deducible information flow," Univ. Paris, Paris, France, LACL, Tech. Rep. 2006-01, 2006.

[21] H. Yasuoka and T. Terauchi, "Quantitative information flow as safety and liveness hyperproperties," *Theor. Comput. Sci.*, vol. 538, pp. 167–182, Jun. 2014.

[22] D. Clark, S. Hunt, and P. Malacaria, "A static analysis for quantifying information flow in a simple imperative language," *J. Comput. Secur.*, vol. 15, no. 3, pp. 321–371, 2007.

[23] M. R. Clarkson, A. C. Myers, and F. B. Schneider, "Quantifying information flow with beliefs," *J. Comput. Security*, vol. 17, no. 5, pp. 655–701, 2009.

[24] Q.-S. Phan, P. Malacaria, C. S. Păsăreanu, and D. Marcelo, "Amorim, "Quantifying information leaks using reliability analysis," in *Proc. Int. SPIN Symp. Model Checking Softw.*, 2014, pp. 105–108.

[25] P. Malacaria, "Algebraic foundations for quantitative information flow," *Math. Struct. Comput. Sci.*, vol. 25, no. 2, pp. 404–428, 2015.

[26] G. Smith, "Quantifying information flow using min-entropy," in *Proc. 8th Int. Conf. Quant. Eval. Syst. (QEST)*, Sep. 2011, pp. 159–167.

[27] A. McIver and C. Morgan, "A probabilistic approach to information hiding," in *Programming Methodology*. Berlin, Germany: Springer, 2003, pp. 441–460.

[28] G. Smith, "On the foundations of quantitative information flow," in *Proc. Int. Conf. Found. Softw. Sci. Comput. Struct.* Berlin, Germany: Springer, 2009, pp. 288–302.

[29] C. Dwork, "Differential privacy: A survey of results," in *Proc. Int. Conf. Theory Appl. Models Comput.* Berlin, Germany: Springer, 2008, pp. 1–19.

[30] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, nos. 3–4, pp. 211–407, 2014.

[31] M. S. Alvim, A. E. Andrés, K. Chatzikokolakis, P. Degano, and C. Palamidessi, "On the information leakage of differentially-private mechanisms," *J. Comput. Secur.*, vol. 23, no. 4, pp. 427–469, 2015.

[32] D. G. Nair, V. P. Binu, and G. S. Kumar. (2015). "An improved E-voting scheme using secret sharing based secure multi-party computation." [Online]. Available: https://arxiv.org/abs/1502.07469

[33] P. Bogetoft *et al.*, "Secure multiparty computation goes live," in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer 2009, pp. 325–343.

[34] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. Theory Cryptogr. Conf.* Berlin, Germany: Springer, 2006, pp. 265–284.
[35] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proc. IEEE 48th Annu. Symp. Found. Comput. Sci. (FOCS)*, Oct. 2007, pp. 94–103.
[36] A. Beimel, K. Nissim, and E. Omri, "Distributed private data analysis: Simultaneously solving how and what," in *Proc. Annu. Int. Cryptol. Conf.* Berlin, Germany: Springer, 2008, pp. 451–468.
[37] C. H. Bennett, G. Brassard, and J.-M. Robert, "Privacy amplification by public discussion," *SIAM J. Comput.*, vol. 17, no. 2, pp. 210–229, Apr. 1988.
[38] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, 1995.
[39] C. Cachin, "Smooth entropy and Rényi entropy," in *Proc. Int. Conf. Theory Appl. Cryptograph. Techn.* Berlin, Germany: Springer, 1997, pp. 193–208.
[40] C. Cachin, "Entropy measures and unconditional security in cryptography," Ph.D. dissertation, Swiss Federal Inst. Technol. Zürich, Zürich, Switzerland, 1997.
[41] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
[42] G. Brassard and C. Crépeau, "Oblivious transfers and privacy amplification," in *Proc. Int. Conf. Theory Appl. Cryptograph. Techn.* Berlin, Germany: Springer, 1997, pp. 334–347.
[43] J. HÅstad, R. Impagliazzo, L. A. Levin, and M. Luby, "A pseudorandom generator from any one-way function," *SIAM J. Comput.*, vol. 28, no. 4, pp. 1364–1396, 1999.
[44] M. H. R. Khouzani and P. Malacaria, "Relative perfect secrecy: Universally optimal strategies and channel design," in *Proc. IEEE 29th Comput. Secur. Found. Symp. (CSF)*, Jun./Jul. 2016, pp. 61–76.
[45] M. S. Alvim, K. Chatzikokolakis, A. McIver, C. Morgan, C. Palamidessi, and G. Smith, "Axioms for information leakage," in *Proc. IEEE 29th Comput. Secur. Found. Symp. (CSF)*, Jun./Jul. 2016, pp. 77–92.
[46] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.
[47] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *Proc. Int. Conf. Theory Appl. Cryptograph. Techn.* Berlin, Germany: Springer, 2004, pp. 523–540.
[48] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Symp. Math. Statist. Probab.* Berkeley, CA, USA: Univ. California, 1961.
[49] S. Fehr and S. Berens, "On the conditional Rényi entropy," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6801–6810, Nov. 2014.
[50] S. Arimoto, "Information measures and capacity of order α for discrete memoryless channels," in *Topics in Information Theory (Colloquia Mathematica Societatis Janos Bolyai)*, vol. 16. Amsterdam, The Netherlands: North-Holland, 1977, pp. 41–52.
[51] M. S. Alvim, K. Chatzikokolakis, C. Palamidessi, and G. Smith, "Measuring information leakage using generalized gain functions," in *Proc. IEEE 25th Comput. Secur. Found. Symp. (CSF)*, Jun. 2012, pp. 265–279.
[52] J. Nocedal and S. J. Wright, *Sequential Quadratic Programming*. Berlin, Germany: Springer, 2006.
[53] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta Numer.*, vol. 4, pp. 1–51, Jan. 1995.
[54] D. J. Wales and J. P. Doye, "Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms," *J. Phys. Chem. A*, vol. 101, no. 28, pp. 5111–5116, 1997.
[55] F. Johansson *et al.* (Dec. 2013). *Mpmath: A Python Library for Arbitraryprecision Floating-Point Arithmetic (Version 0.18)*. [Online]. Available: http://mpmath.org/

Patrick Ah-Fat is a PhD student in the Department of Computing at Imperial College London under the supervision of Michael Huth. In 2012, he started studying at Enseeiht, Toulouse, France, from where he received a Masters in Computing and Applied Mathematics. He then received a Masters in Computing in 2015 from Imperial College London, United Kingdom. His research interests include information theory, information flow and secure multi-party computation.

Michael Huth is Professor of Computer Science in the Department of Computer Science at Imperial College London. He is a Diplom-Mathematiker (TU Darmstadt, Germany), obtained his PhD in 1991 (Tulane University of Louisiana, USA), was an Assistant Professor at Kansas State University from 1996-2001, and completed several postdoctoral studies in the US, Germany, and the UK on programming language semantics and design, formal verification, and probabilistic modeling. His present research focuses on cybersecurity, especially modeling and reasoning about the interplay of trust, security, risk, and economics. Currently funded projects of his include work on blockchain technology for intelligent transportation systems and for machine learning and optimization applied to cybersecurity. He is the Technical Lead of the Theme Harnessing Economic Value in the UK PETRAS IoT Research Hub and on the editorial board of the International Journal on Software Tools for Technology Transfer. Professor Huth is also CTO of XAIN AG, an AI/Cybersecurity startup, in Berlin and active as research and product advisor in the London and Cybersecurity startup scene.