Electronic Colloquium on Computational Complexity, Report No. 15 (2013)

New lower bounds for privacy in communication protocols

Iordanis Kerenidis* Mathieu Laurière† David Xiao ‡

January 15, 2013

Abstract Communication complexity is a central model of computation introduced by Yao [Yao79], where two players, Alice and Bob, receive inputs x and y respectively and want to compute $f(x, y)$ for some fixed function f with the least amount of communication. Recently people have revisited the question of the privacy of such protocols: is it possible for Alice and Bob to compute $f(x, y)$ without revealing too much information about their inputs? There are two types of privacy for communication protocols that have been proposed: first, an information theoretic definition ([BYCKO93, K04]), which for Boolean functions is equivalent to the notion of information cost introduced by [CSWY01] and that has since found many important applications; second, a combinatorial definition introduced by [FJS10] and further developed by [ACCFKP12].

We provide new results for both notions of privacy, as well as the relation between them. Our new lower bound techniques both for the combinatorial and the information-theoretic definitions enable us to give tight bounds for the privacy of several functions, including Equality, Disjointness, Inner Product, Greater Than. In the process we also prove tight bounds (up to 1 or 2 additive bits) for the external information complexity of these functions.

We also extend the definitions of privacy to bounded-error randomized protocols and provide a relation between the two notions and the communication complexity. Again, we are able to prove tight bounds for the above-mentioned functions as well as the Vector in Subspace and Gap Hamming Distance problems.

*CNRS, LIAFA, Universite Paris 7 and CQT, NUS Singapore. jkeren@liafa.univ-paris-diderot.fr
†ENS Cachan, LIAFA, Universite Paris 7. mathieu.lauriere@ens-cachan.fr
‡CNRS, LIAFA, Universite Paris 7. dxiao@liafa.univ-paris-diderot.fr

ISSN 1433-8092

1 Introduction

Communication complexity is a central model of computation, first defined by Yao, [Yao79], that has found applications in many areas of theoretical computer science. In the 2-*party communication complexity setting*, we consider two players, Alice and Bob with unlimited computational power. Each of them receives an input, say $x \in \mathcal{X}$ for Alice and $y \in \mathcal{Y}$ for Bob, and their goal is to compute $f(x, y) \in \mathcal{Z}$ for some fixed function f with the minimum amount of communication.

Imagine now that Alice and Bob still want to collaboratively compute $f(x, y)$, while retaining privacy of their input. More precisely, the loss of privacy measures how much information about (x, y) is leaked to an eavesdropper who has only access to the transcript (*extrenal* privacy), or how much information about one party's input is leaked through the transcript to the other pary (*internal* privacy). A perfectly private protocol will reveal no information about x and y, other than what can be inferred from the value of $f(x, y)$.

For example, if Alice and Bob both want to output the minimum of $x, y \in \{0, 1\}^n$, then the optimal deterministic communication protocol is the trivial protocol of complexity $2n$. In fact one can show that any deterministic protocol that has optimal communication complexity is not private at all against an eavesdropper since basically both players have to send the input to the other one. However a perfectly private deterministic protocol exists, alas with much worse communication complexity: the two parties initiate a counter $i = 0$ and in each round $i = 0$ to $2^n - 1$, Alice announces "Yes" if $x = i$, otherwise "No"; Bob announces "Yes" if $y = i$, otherwise "No". If neither party says "Yes" then they increment i, otherwise the protocol ends when someone says "Yes". It is clear that from the transcript, one only learns what can be inferred from the value of the function and nothing more.

In order to quantify the notion of privacy, Bar-Yehuda *et al.* [BYCKO93] provided a definition of *internal privacy* of a function f according to an input distribution μ, a variation of which has been subsequently referred to as *internal information cost* ($\mathrm{IC}^{int}_{\mu}(f)$). In high level, it measures the amount of information Alice learns about Bob's input from the transcript and vice versa. A second type of information cost, called *external information cost* ($\mathrm{IC}^{ext}_{\mu}(f)$) was defined in [CSWY01] and measures the amount of information that is learned by an external observer about Alice and Bob's inputs given the messages they exchanged during the protocol. The notion of internal and external information cost has recently found many important applications in communication complexity [CSWY01, BJKS02, BBCR10, Bra11].

Klauck [K04] also defined an information theoretic notion of privacy, which we denote here by $\mathrm{PRIV}^{int}_{\mu}(f)$, which is closely related to the internal information cost (the only difference being that it measures the amount of information Alice learns about Bob's input from the transcript conditioned on knowing the value of the function, and vice versa). In fact, the two notions are basically equivalent for boolean functions and all our results about PRIV can be translated to results about information cost. These definitions have the advantage to be easily related to other tools in information theory, but are not easily seen in a combinatorial way.

Feigenbaum *et al.* [FJS10] gave a combinatorial definition of privacy, called *objective privacy-approximation ratio* (that we will refer to as external privacy-approximation ratio), that is equal to the expected value over the inputs (x, y) drawn from some distribution μ of the following ratio: the number of inputs that are mapped to the same value by f (that are indistinguishable of (x, y) by looking only at the function's output) over the number of inputs giving rise to the same transcript as the one of (x, y) (that are indistinguishable of (x, y) by looking only at the protocol's transcript). They also defined a notion of *subjective (or internal) privacy-approximation ratio*, that captures how much more one player learns about the input of the other one through the transcript than through the value of the function, and equals the ratio of the number of Alice's possible inputs x that are indistinguishable by looking only at Bob's input y and the output of the function, over the number of x's that are instiguishable by looking at y and the full transript. Last, they studied several functions and computed lower bounds for their privacy-approximation ratio, however restricting themselves to the case of uniformly distributed inputs.

More recently, Ada *et al.* in [ACCFKP12] have modified the definition of privacy-approximation ratio, which we denote as $\mathrm{PAR}_\mu^{ext}(f)$ and $\mathrm{PAR}_\mu^{int}(f)$, so that it measures the size of subsets of $\mathcal{X} \times \mathcal{Y}$ not just by counting the number of elements, but relative to the inputs' distribution μ. They showed that the logarithm of this new definition of internal PAR can be lower bounded by the zero-error internal information cost (which nevertheless can be arbitrarily smaller for certain functions with large output range). They also proved a tradeoff between privacy and communication complexity for a specific function (`Vickrey-auction`) and the uniform distribution of inputs.

We note that in [FJS10] and [ACCFKP12] only deterministic protocols were considered. Moreover, the relation between the two measures was not very well understood.

1.1 Our results

We provide new results for the two notions of privacy, PRIV and PAR, both external and internal, as well as about their relation. These results enable us to give tight bounds for the privacy of several functions. We also extend the definitions of PRIV and PAR to bounded-error randomized protocols, provide a relation between the two notions and the communication complexity and prove tight bounds for several functions.

New lower bounds for external PAR First, we present a general lower bound technique for $\mathrm{PAR}_\mu^{ext}(f)$ via linear programming. We relate it to two other well known lower bound techniques for communication complexity (see [JK10]): the *rectangle* bound ($\mathrm{rec}(f)$) and the *partition* bound ($\mathrm{prt}(f)$). This linear program can be written as a weighted sum of rectangle bounds $\mathrm{rec}^z(f)$, where the weight is equal to the weight of the inputs (x, y) according to μ that are mapped to z by f. It is, hence, easy to compute for many functions:

Theorem: For all functions f, $\mathrm{PAR}_\mu^{ext}(f) \geq \sum_{z \in \mathcal{Z}} \left|f^{-1}(z)\right|_\mu \cdot \mathrm{rec}^z(f)$.

Moreover, for the case of boolean functions we present two more lower bound techniques. First, as a weighted sum of the size of the 0- and 1-fooling sets of M_f ($|F_0|$ and $|F_1|$ respectively); and second, by the rank of M_f:

Theorem: For boolean functions f, $\mathrm{PAR}_\mu^{ext}(f) \geq \left|f^{-1}(0)\right|_\mu \cdot |F_0| + \left|f^{-1}(1)\right|_\mu \cdot |F_1|$.

Theorem: For boolean functions f, $\mathrm{PAR}_\mu^{ext}(f) \geq \min\{\mathrm{rank}\left(\mathcal{M}_f\right), \mathrm{rank}\left(\mathcal{M}_{\mathrm{not}f}\right)\} \geq \mathrm{rank}\left(\mathcal{M}_f\right) - 1$.

In addition, we prove that external PAR is larger than internal PAR and that external PAR and *communication complexity* are polynomially related, provided that the *log-rank conjecture* holds.

New lower bound techniques for external IC **and** PRIV We prove a new lower bound on the external zero-error information cost which using the equivalence between IC and PRIV given in Theorem 2.8 will in turn give new lower bounds on $\mathrm{PRIV}_\mu^{ext}(f)$.

Theorem: Fix a function f. Suppose there exists $\delta > 0$ and a distribution μ over the inputs of f whose support is a rectangle, such that for all monochromatic rectangles R of f, $\mu(R) \leq \delta$. Then it holds for every protocol P that computes f with zero error that $\mathrm{IC}_\mu^{ext}(P) \geq \log(1/\delta)$.

We remark that our theorem allows us to prove exact bounds for zero-error IC up to an *additive* constant term (with a small constant, between 1 and 2).

Theorem: For each of $f = \mathrm{EQ}, \mathrm{GT}, \mathrm{DISJ}$, there exists μ such that $\mathrm{IC}_\mu^{ext}(f) \geq n$. Also, there exists μ such that $\mathrm{IC}_\mu^{ext}(\mathrm{IP}) \geq n - 1 - o(1)$.

These are much sharper than typical lower bounds on IC, which work in the bounded-error case and incur multiplicative constants [BJKS02, Bra11, BW12, KLLRX12]. The only other such sharp lower bounds we are aware of are due to Braverman *et al.* [BGPW12] who study the AND and DISJ functions. However they

Problem	$\mathbf{PAR}^{\mathbf{ext}}_{\mu}$		$\mathbf{PRIV}^{\mathbf{ext}}_{\mu}$ (for some μ)	$\mathbf{PAR}^{\mathbf{int}}_{\mu,\epsilon}$ (for some μ)
	[FJS10] (for uniform μ)	**Our contribution** (for any μ)		
`Equality`	-	2^n	$n-1$	$\Theta(1)$
`Disjointness`	$\left(\frac{3}{2}\right)^n$	2^n-1	$n-1$	$2^{\Theta(n)}$
`Inner Product`	-	2^n-1	$n-2-o(1)$	$2^{\Theta(n)}$
`Greater Than`	$2^n+\frac{1}{2^{n+1}}-\frac{1}{2}$	2^n-1	$n-1$	$2^{\Theta(\log n)}$

Table 1: Lower bounds for specific functions

prove sharp bounds for the internal IC of DISJ, not for the external IC as we study here.

Our bound can be used to prove an *optimal* lower bound on the zero-error information complexity of certain functions (*i.e.* without even an additive constant loss). For example, for the single bit AND function, our theorem implies that there exists μ such that $\mathrm{IC}^{ext}_{\mu}(P) \geq \log_2 3$. This matches a recent upper bound of [BGPW12] (they also proved that their upper bound is tight via different techniques).

Applications We exhibit the power of these new lower bound techniques for PAR and PRIV by proving optimal lower bounds on most of the examples of functions left open in [FJS10] and more: `Equality`, `Disjointness`, `Inner Product`, `Greater Than` (`Millionaire's problem`).

Privacy for bounded-error randomized protocols We extend the definition of PAR and PRIV to bounded-error randomized protocols and show that for any protocol, external PRIV is a lower bound on external PAR and the same for the internal notions.

Theorem: $\mathrm{PRIV}^{ext}_{\mu}(P) \leq \log \mathrm{PAR}^{ext}_{\mu}(P)$ and $\mathrm{PRIV}^{int}_{\mu}(P) \leq 2 \cdot \log\left(\mathrm{PAR}^{int}_{\mu}(P)\right)$.

Since PRIV is lower bounded by IC, which was shown in [KLLRX12] to subsume almost all known lower bounds for communication complexity, i.e. smooth rectangle, γ_2-norm bound, discrepancy, etc., the two notions of privacy, for the bounded-error case, are in fact both equal to the communication complexity for all boolean functions for which we have a tight bound on their communication complexity. Interestingly, the notion of PAR sits between information and communication complexity, and it is an important open question whether these two notions are equal (which would also make PAR equal to them).

Comparison between the two notions of privacy As we have said, for the case of bounded-error protocols, the two notions of privacy seem to be practically equal for most functions. However, for the zero-error case, they can diverge for certain functions. In order to understand the differences between the notions, we study their robustness when we change slightly the input distribution and we show that the information theoretic notion of privacy is more robust to such changes. Moreover, we show that while PRIV is always less than the expected communication complexity of the protocol, the same is not true for PAR. We also discuss an error in the appendix of [FJS10] where they claim that PRIV is not as robust as PAR for distinguishing the privacy of two specific protocols they exhibited.

2 Preliminaries

We consider three non empty sets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ and a function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$. μ denotes a distribution over $\mathcal{X} \times \mathcal{Y}$, and for any set $E \subseteq \mathcal{X} \times \mathcal{Y}$, $|E|_{\mu} := \sum_{(x,y)\in E} \mu(x,y)$. $\mathcal{M}_f$ is the matrix of f: $\mathcal{M}_f[x,y] := f(x,y)$.

We let P denote a two-party communication protocol. Protocols may use both public and private random coins. For any protocol P, let T_P denote its transcript. For randomized protocols, $T_P(x,y)$ is a random

variable comprised of all messages as well as the public coins, but not the private coins. We will simply write T if the protocol is clear from context. Given a protocol P and a transcript T, for any input $(x,y) \in \mathcal{X} \times \mathcal{Y}$, $P(x,y)$ is the value output by Alice and Bob upon running the protocol, and $T(x,y)$ is the whole transcript (these are random variables if the protocol is randomized). Let $\mathbf{CC}(P)$ be the maximum number of bits communicated by P over all choices of inputs and random coins. Let $\mathbf{D}(f) = \min_P \mathbf{CC}(P)$ where P ranges over all *deterministic* protocols computing f. Let $\mathbf{R}^\epsilon(f) = \min_P \mathbf{CC}(P)$ where P ranges over all randomized protocols computing f with error at most ϵ.

For any input $(x,y) \in \mathcal{X} \times \mathcal{Y}$, the monochromatic f-region of (x,y) is defined as $\mathrm{D}^f_{x,y} := f^{-1}(f(x,y))$, and is equal to the monochromatic P-region $\mathrm{D}^P_{x,y}$ of (x,y) if P is deterministic or 0-error. The monochromatic P-rectangle of (x,y) is defined as $\mathrm{D}^{T_P}_{x,y} := T_P^{-1}(T_P(x,y))$. For any output $z \in \mathcal{Z}$, the monochromatic f-region of z is: $f^{-1}(z) := f^{-1}(\{z\})$, which is equal to the monochromatic P-region of z, $P^{-1}(z)$, if P is deterministic or 0-error. Let $\mathcal{R}^P_z$ be the set of P-rectangles covering $P^{-1}(z)$, that is: $\mathcal{R}^P_z := \{\mathrm{D}^{T_P}_{x,y} | (x,y) : P(x,y) = z\}$. Let $\mathcal{R}^P = \cup_{z \in \mathcal{Z}} \mathcal{R}^P_z = \left\{\mathrm{D}^{T_P}_{x,y} \middle| (x,y) \in \mathcal{X} \times \mathcal{Y}\right\}$ be the set of all P-rectangles. If P is a deterministic or zero-error protocol for f, for each $z \in \mathcal{Z}$, $\mathrm{cut}_P(z)$ is the number of P-rectangles in $f^{-1}(z)$; $\mathcal{R}(\mathcal{X} \times \mathcal{Y})$ is the set of all rectangles in $\mathcal{X} \times \mathcal{Y}$.

For three random variables A, B, C the conditional mutual information is defined as $\mathbf{I}(A;B|C) := \mathbf{H}(A|C) - \mathbf{H}(A|BC)$, where $\mathbf{H}$ denotes Shannon entropy: if X and Y are two random variables $\mathbf{H}(X) = \sum_x \mathbb{P}\{X = x\} \log(1/\mathbb{P}\{X = x\})$ and $\mathbf{H}(X|Y) = \mathbb{E}[-\log(\mathbb{P}(X|Y)]$. We recall some simple facts about information and entropy (more details about information theory can be found in the textbook of Cover and Thomas [CT06].) For any random variables X, Y, Z, W, the Chain Rule says that $\mathbf{H}(X,Y) = \mathbf{H}(X) + \mathbf{H}(Y|X)$ and $\mathbf{I}(X,Z;Y) = \mathbf{I}(X;Y) + \mathbf{I}(Z;Y|X)$. Another easy fact (see for example [ACCFKP12]) is that:

$$|\mathbf{I}(X;Y|W) - \mathbf{I}(X;Y|W,Z)| \leq \mathbf{H}(Z) \tag{1}$$

2.1 Definitions of privacy for deterministic or zero-error protocols

Information complexiy: We define the external and internal information complexity, notions that have recently found many applications in communication complexity [CSWY01, BJKS02, BBCR10, Bra11]. The external information complexity measures the amount of information that is learned from someone who looks at the messages exchanged between Alice and Bob during the protocol about their inputs. The internal information complexity measures the amount of information that Alice learns about Bob's input and vice versa.

Definition 2.1. *We define the external information complexity of P as $\mathrm{IC}^{ext}_\mu(P) := \mathbf{I}(T_P; X, Y)$. The external information complexity of f is $\mathrm{IC}^{ext}_\mu(f) := \min_P \mathrm{IC}^{ext}_\mu(P)$ where the minimum is over all protocols P computing f.*

Definition 2.2. *We define the internal information complexity of P as $\mathrm{IC}^{int}_\mu(P) := \mathbf{I}(T_P; X|Y) + \mathbf{I}(T_P; Y|X)$. The internal information complexity of f is $\mathrm{IC}^{int}_\mu(f) := \min_P \mathrm{IC}^{int}_\mu(P)$ where the minimum is over all protocols P computing f.*

Information-theoretic privacy: In [BYCKO93], the definition of privacy ($\mathcal{I}^{\mathrm{det}}_{\mathrm{c-i}}$ in their notations) is basically the same as what we now call $\mathrm{IC}^{int}_\mu(P)$ (they used the max instead of the sum of the two terms). A related notion of privacy has been defined by Klauck in [K04], where again he takes the max of the two terms instead of the sum. We give a distribution-dependent version of his definition. Roughly, it represents how much an observer learns about the inputs conditioned on knowing the output of the function. We also define an internal version of the definition.

Definition 2.3. *The external privacy of P is defined as $\mathrm{PRIV}^{ext}_\mu(f,P) := \mathbf{I}(T_P(X,Y); X, Y | f(X,Y))$. The external privacy of a function f is defined as $\mathrm{PRIV}^{ext}_\mu(f) := \min_P \mathrm{PRIV}^{ext}_\mu(f,P)$ where the minimum is taken over all protocols P for f.*

Definition 2.4. *The internal privacy of P is defined as* $\mathrm{PRIV}^{int}_{\mu}(f,P) := \mathbf{I}(T_P(X,Y);X|Y,f(X,Y)) + \mathbf{I}(T_P(X,Y);Y|X,f(X,Y))$ *The internal privacy of a function f is defined as* $\mathrm{PRIV}^{int}_{\mu}(f) := \min_P \mathrm{PRIV}^{int}_{\mu}(f,P)$ *where the minimum is taken over all protocols P for f.*

Combinatorial privacy PAR**:** We present here the definition of PAR given by [ACCFKP12], which modified the original definition in [FJS10] in order to measure the size of regions relative to the inputs' distribution.

Definition 2.5. *The external privacy-approximation ratio of a deterministic protocol P computing f is defined as:*

$$\mathrm{PAR}^{ext}_{\mu}(f,P) := \mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}^{f}_{x,y}\right|_{\mu}}{\left|\mathrm{D}^{T_P}_{x,y}\right|_{\mu}}\right] = \mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}^{P}_{x,y}\right|_{\mu}}{\left|\mathrm{D}^{T_P}_{x,y}\right|_{\mu}}\right]$$

(where the equality holds because P has zero error). The external privacy-approximation ratio of a function f is defined as: $\mathrm{PAR}^{ext}_{\mu}(f) := \min_P \mathrm{PAR}^{ext}_{\mu}(f,P)$.

Definition 2.6. *The internal privacy-approximation ratio of a protocol P computing f is defined as:*

$$\mathrm{PAR}^{int}_{\mu}(f,P) := \mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}^{f}_{x,y}\cap\mathcal{X}\times\{y\}\right|_{\mu}}{\left|\mathrm{D}^{T_P}_{x,y}\cap\mathcal{X}\times\{y\}\right|_{\mu}}\right] + \mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}^{f}_{x,y}\cap\{x\}\times\mathcal{Y}\right|_{\mu}}{\left|\mathrm{D}^{T_P}_{x,y}\cap\{x\}\times\mathcal{Y}\right|_{\mu}}\right]$$

The internal privacy-approximation ratio of a function f is defined as: $\mathrm{PAR}^{int}_{\mu}(f) := \min_P \mathrm{PAR}^{int}_{\mu}(f,P)$.

Basic relations and bounds: The relation between internal IC and internal PRIV was explained in [ACCFKP12]. It is possible to improve the lower bound and to show the same relationship for external notions.

Theorem 2.7. *For any protocol P and any distribution μ,*

$$\mathrm{PRIV}^{int}_{\mu}(f,P) \le \mathrm{IC}^{int}_{\mu}(P) \le \mathrm{PRIV}^{int}_{\mu}(f,P) + 2\log(|\mathcal{Z}|)$$

and similarly:

Theorem 2.8. *For any protocol P and any distribution μ,*

$$\mathrm{PRIV}^{ext}_{\mu}(f,P) \le \mathrm{IC}^{ext}_{\mu}(P) \le \mathrm{PRIV}^{ext}_{\mu}(f,P) + \log(|\mathcal{Z}|)$$

Proof of Theorem 2.7 *and Theorem* 2.8. Let us first prove the upper bounds. By definition of IC and PRIV we have, respectively for the external and the internal notions:

$$\begin{aligned}
\mathrm{IC}^{int}_{\mu}(P) - \mathrm{PRIV}^{int}_{\mu}(f,P) &= \mathbf{I}(T_P(X,Y);X|Y) - \mathbf{I}(T_P(X,Y);X|Y,f(X,Y)) \\
&\quad + \mathbf{I}(T_P(X,Y);Y|X) - \mathbf{I}(T_P(X,Y);Y|X,f(X,Y)) \\
&\le 2\log(|\mathcal{Z}|), \\
\mathrm{IC}^{ext}_{\mu}(P) - \mathrm{PRIV}^{ext}_{\mu}(f,P) &= \mathbf{I}(T_P(X,Y);X,Y) - \mathbf{I}(T_P(X,Y);X,Y|f(X,Y)) \\
&\le \log(|\mathcal{Z}|).
\end{aligned}$$

The inequalities come from Inequality (1). Let us prove now the lower bound. Since the output of function is completely determined by the input,

$$\begin{aligned}
\mathrm{PRIV}^{ext}_{\mu}(f,P) = \mathbf{I}(T_P(X,Y);X,Y|f(X,Y)) &= \mathbf{H}(T_P(X,Y)|f(X,Y)) - \mathbf{H}(T_P(X,Y)|X,Y,f(X,Y)) \\
&= \mathbf{H}(T_P(X,Y)|f(X,Y)) - \mathbf{H}(T_P(X,Y)|X,Y) \\
&\le \mathbf{H}(T_P(X,Y)) - \mathbf{H}(T_P(X,Y)|X,Y) \\
&= \mathrm{IC}^{ext}_{\mu}(P).
\end{aligned}$$

The same arguments work also for the internal notions. □

We can show that internal PAR is smaller than external one:

Theorem 2.9. *For any deterministic protocol P computing f:*

$$\mathrm{PAR}_\mu^{int}(f,P) \leq 2 \cdot \mathrm{PAR}_\mu^{ext}(f,P).$$

Proof. Recall that: $\mathrm{PAR}_\mu^{ext}(f,P) = \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_\mu \cdot n_z$, where $n_z = \#\mathcal{R}_z^P$. If we denote by $n_{z,y}$ and $n_{z,x}$ respectively $|\{R \cap \mathcal{X}\times\{y\} | R \in \mathcal{R}_z^P\}|$ and $|\{R \cap \{x\}\times\mathcal{Y} | R \in \mathcal{R}_z^P\}|$, then:

$$\begin{aligned}
\mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}_{x,y}^f \cap \mathcal{X}\times\{y\}\right|_\mu}{\left|\mathrm{D}_{x,y}^{T_P} \cap \mathcal{X}\times\{y\}\right|_\mu}\right] &= \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \mu(x,y) \cdot \frac{\left|\mathrm{D}_{x,y}^f \cap \mathcal{X}\times\{y\}\right|_\mu}{\left|\mathrm{D}_{x,y}^{T_P} \cap \mathcal{X}\times\{y\}\right|_\mu} \\
&= \sum_{z\in\mathcal{Z}}\sum_{y\in\mathcal{Y}} \left|f^{-1}(z)\cap\mathcal{X}\times\{y\}\right|_\mu \sum_{R\in\mathcal{R}_z^P} \frac{1}{|R\cap\mathcal{X}\times\{y\}|_\mu} \underbrace{\sum_{x:(x,y)\in R} \mu(x,y)}_{=|R\cap\mathcal{X}\times\{y\}|_\mu} \\
&= \sum_{z\in\mathcal{Z}}\sum_{y\in\mathcal{Y}} \left|f^{-1}(z)\cap\mathcal{X}\times\{y\}\right|_\mu \cdot n_{z,y}.
\end{aligned}$$

Similarly we prove:

$$\mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}_{x,y}^f \cap \{x\}\times\mathcal{Y}\right|_\mu}{\left|\mathrm{D}_{x,y}^{T_P} \cap \{x\}\times\mathcal{Y}\right|_\mu}\right] = \sum_{z\in\mathcal{Z}}\sum_{x\in\mathcal{X}} \left|f^{-1}(z)\cap\{x\}\times\mathcal{Y}\right|_\mu \cdot n_{z,x}.$$

Hence:

$$\begin{aligned}
\mathrm{PAR}_\mu^{int}(f,P) &= \sum_{z\in\mathcal{Z}} \left(\sum_{y\in\mathcal{Y}} \left|f^{-1}(z)\cap\mathcal{X}\times\{y\}\right|_\mu \cdot n_{z,y} + \sum_{x\in\mathcal{X}} \left|f^{-1}(z)\cap\{x\}\times\mathcal{Y}\right|_\mu \cdot n_{z,x}\right) \\
&\leq \sum_{z\in\mathcal{Z}} \left(2 \cdot \left|f^{-1}(z)\right|_\mu \cdot n_z\right) = 2 \cdot \mathrm{PAR}_\mu^{ext}(f,P)
\end{aligned}$$

the inequality follows from the fact that for each z, for any x and y, $n_{z,x}$ and $n_{z,y}$ are smaller than n_z, and $\sum_{y\in\mathcal{Y}} \left|f^{-1}(z)\cap\mathcal{X}\times\{y\}\right|_\mu = \sum_{x\in\mathcal{X}} \left|f^{-1}(z)\cap\{x\}\times\mathcal{Y}\right|_\mu = \left|f^{-1}(z)\right|_\mu$. □

It is also easy to see that

Theorem 2.10. $\mathrm{PRIV}_\mu^{int}(f,P) \leq \mathrm{PRIV}_\mu^{ext}(f,P) + \log(|\mathcal{Z}|)$.

Proof. Braverman [Bra11] proved that: $\mathrm{IC}_\mu^{int}(P) \leq \mathrm{IC}_\mu^{ext}(P)$. Hence, with Theorem 2.8:

$$\mathrm{PRIV}_\mu^{int}(f,P) \leq \mathrm{IC}_\mu^{int}(P) \leq \mathrm{IC}_\mu^{ext}(P) \leq \mathrm{PRIV}_\mu^{ext}(f,P) + \log(|\mathcal{Z}|).$$

□

We can also relate PAR and PRIV:

Theorem 2.11. *(Theorem 19 in [ACCFKP12]) For any deterministic protocol P and any distribution μ:* $\mathrm{PRIV}_\mu^{int}(f,P) \leq 2 \cdot \log\left(\mathrm{PAR}_\mu^{int}(f,P)\right)$.

Similarly, for external privacy we can prove:

Theorem 2.12. *For any deterministic protocol P and any distribution μ,* $\mathrm{PRIV}^{ext}_{\mu}(f,P) \leq \log\left(\mathrm{PAR}^{ext}_{\mu}(f,P)\right)$.

The proof follows from a more general result that we prove in section 4.

The external PAR equals a weighted sum of the number of rectangles tiling each f-monochromatic region.

Theorem 2.13 ([ACCFKP12])**.** *For any deterministic protocol P, we have:*

$$\mathrm{PAR}^{ext}_{\mu}(f,P) = \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_{\mu} \cdot \mathrm{cut}_P(f^{-1}(z)).$$

This result was stated in [ACCFKP12] but for completeness we present a proof:

Proof. We successively write:

$$\begin{aligned}
\mathrm{PAR}^{ext}_{\mu}(f,P) &= \mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}^f_{x,y}\right|_{\mu}}{\left|\mathrm{D}^{T_P}_{x,y}\right|_{\mu}}\right] = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \mu(x,y)\cdot\frac{\left|\mathrm{D}^f_{x,y}\right|_{\mu}}{\left|\mathrm{D}^{T_P}_{x,y}\right|_{\mu}} \\
&= \sum_{z\in\mathcal{Z}}\sum_{R\in\mathcal{R}^P_z} \frac{\left|f^{-1}(z)\right|_{\mu}}{|R|_{\mu}} \underbrace{\sum_{(x,y)\in R}\mu(x,y)}_{=|R|_{\mu}} \\
&= \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_{\mu}\cdot \#\mathcal{R}^P_z = \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_{\mu}\cdot \mathrm{cut}_P(f^{-1}(z)).
\end{aligned}$$

The third equality holds because each rectangle generated by the transcript is included in the f-region that corresponds to its output (since P is a zero-error protocol). □

Finally, if $\mathrm{Dist}_{\mu,\epsilon}$ for $\epsilon \geq 0$ represents the *expected* distributional complexity of a randomized ϵ-error protocol with respect to some input distribution μ, we have:

Theorem 2.14. *For any randomized ϵ-error protocol and any input distribution,* $\mathrm{Dist}_{\mu,\epsilon}(P) \geq \mathrm{IC}^{ext}_{\mu,\epsilon}(P)$.

Proof. We consider the code $T_P(X,Y)$ where the codewords are all the possible transcripts over the alphabet $\{0,1\}$, for a randomized ϵ-error protocol P over the input X,Y distributed according to μ. This code is uniquely decodable, and hence we can use the Kraft inequality of theorem 5.5.1 in [CT06]: the expected length of $T_P(X,Y)$ is greater than its entropy. Hence, for any input distribution, $\mathrm{Dist}_{\mu,\epsilon}(P) \geq H(T_P)$. The conclusion follows from: $H(T_P) \geq H(T_P) - H(T_P|X,Y) = \mathbf{I}(T_P;X,Y) = \mathrm{IC}^{ext}_{\mu,\epsilon}(P)$. □

Note that, since $\mathrm{IC}^{ext}_{\mu,\epsilon}(P) \geq \mathrm{IC}^{int}_{\mu,\epsilon}(P)$, we also have: $\mathrm{Dist}_{\mu,\epsilon}(P) \geq \mathrm{IC}^{int}_{\mu,\epsilon}(P)$.

We can summarize the relations between the notions of communication or information complexity and privacy by the diagrams of figure 2.1. In the diagram, an arrow $A \leftarrow B$ indicates that $A \leq B$ (up to constants). The quantities indicate *worst-case* complexity except for Dist, clarified below. Relations between:

- $\mathbf{D}$ (resp. $\mathbf{R}^{\epsilon}$) and PAR is given by Theorem 4.4;
- $\mathrm{PAR}^{\mathrm{ext}}$ and $\mathrm{PAR}^{\mathrm{int}}$ comes from Theorem 2.9;
- PAR and PRIV for the deterministic 0-error setting are given in Theorem 2.11 (internal) and Theorem 2.12 (external), while Theorem 4.5 give these relations in the bounded-error setting (internal and external);

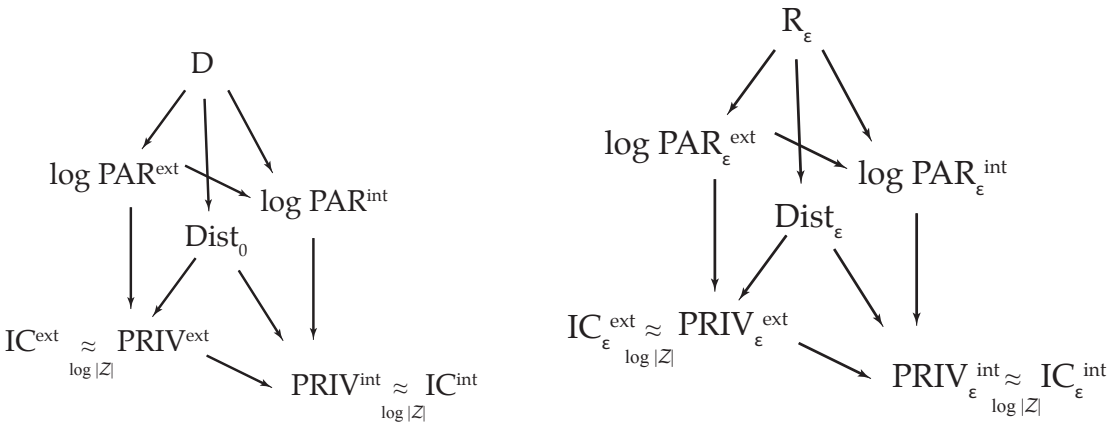

Figure 1: Lower bounds diagrams for deterministic and bounded error cases

- $\mathrm{PRIV}^{\mathrm{ext}}$ and $\mathrm{PRIV}^{\mathrm{int}}$ (both for deterministic 0-error and bounded error) comes from Theorem 2.10;
- IC and PRIV (both for deterministic 0-error and bounded error) are in Theorem 2.7 (internal) and Theorem 2.8 (external).
- The expected distributional complexity and IC (or PRIV) for every possible input distribution is given in Theorem 2.14

3 New lower bound techniques for PAR and PRIV of deterministic protocols

3.1 External PAR and the rectangle bound

We start by defining a linear program and prove that its optimal value is a lower bound for external PAR. Then we relate it to two other well known bounds (the *rectangle* and the *partition* bound) that often provide tight lower bounds for communication complexity.

Definition 3.1. *Let* $\widetilde{\mathrm{PAR}}_\mu(f)$ *be the value of the following linear program:*

$$\min_{w_{z,R}} \sum_{z,R} w_{z,R} \cdot \left|f^{-1}(z)\right|_\mu \quad \textit{s.t.} \qquad \forall\,(x,y) \in f^{-1}(\mathcal{Z}) : \sum_{R:R\ni(x,y)} w_{f(x,y),R} = 1 \tag{2}$$

$$\forall\,(x,y) \in f^{-1}(\mathcal{Z}) : \sum_{R:R\ni(x,y)} \sum_{z} w_{z,R} = 1 \tag{3}$$

$$\forall\, z,\ \forall\, R : \qquad w_{z,R} \geq 0. \tag{4}$$

where the z's and the R's are always taken respectively in $\mathcal{Z}$ and in $\mathcal{R}(\mathcal{X}\times\mathcal{Y})$.

Intuitively, conditions (3) and (4) say that we can interprete $w_{z,R}$ as a probability distribution. In fact, $w_{z,R}$ can be seen as the probability distribution that picks R and outputs z on (x,y). This is because condition (2) forces the probability of outputting $f(x,y)$ on (x,y) to be 1. We prove the following theorem that the optimal value of this linear program is a lower bound for PAR:

Theorem 3.2. $\mathrm{PAR}^{ext}_\mu(f) \geq \widetilde{\mathrm{PAR}}_\mu(f)$.

Proof. Let P be a deterministic protocol for f and T its transcript. Since P is zero-error, its rectangles are included in f-monochromatic regions. Given a rectangle R of the transcript, let D^f_R be the unique f-monochromatic region containing R. Then:

$$\mathrm{PAR}^{ext}_\mu(f,P) = \mathbb{E}_{(x,y)\sim\mu}\left[\frac{\left|\mathrm{D}^f_{x,y}\right|_\mu}{\left|\mathrm{D}^T_{x,y}\right|_\mu}\right] = \sum_{(x,y)} \mu(x,y)\cdot\frac{\left|\mathrm{D}^f_{x,y}\right|_\mu}{\left|\mathrm{D}^T_{x,y}\right|_\mu} = \sum_{R\in\mathcal{R}^P}\frac{\left|\mathrm{D}^f_R\right|_\mu}{|R|_\mu}\sum_{(x,y)\in R}\mu(x,y) = \sum_{R\in\mathcal{R}^P}\left|\mathrm{D}^f_R\right|_\mu$$

Letting $w_{z,R} := \mathbb{1}_{R\in\mathcal{R}^P_z}$ ($=1$ if P outputs z on R, 0 otherwise), we can write:

$$\mathrm{PAR}^{ext}_\mu(f,P) = \sum_{R\in\mathcal{R}(\mathcal{X}\times\mathcal{Y})}\sum_{z\in\mathcal{Z}} w_{z,R}\cdot\left|\mathrm{D}^f_R\right|_\mu = \sum_{z\in\mathcal{Z},R\in\mathcal{R}(\mathcal{X}\times\mathcal{Y})} w_{z,R}\cdot\left|f^{-1}(z)\right|_\mu.$$

Moreover it is easy to see that this family $w_{z,R}$ satisfies the constraints of Definition 3.1 because P is zero-error. Hence it is a valid solution, whose corresponding objective value is at least the optimal value. □

3.1.1 Relation with *rectangle* linear program:

We use the formulation of the rectangle bound given by [JK10] as the following linear program:

Definition 3.3. $\mathrm{rec}^z(f)$ *is the optimal value of the following linear program:*

$$\begin{aligned}\min_{w_R}\sum_R w_R \quad \textit{s.t.} \qquad & \forall (x,y)\in f^{-1}(z): \sum_{R:R\ni(x,y)} w_R = 1\\ & \forall (x,y)\in \mathcal{X}\times\mathcal{Y}\setminus f^{-1}(z): \sum_{R:R\ni(x,y)} w_R = 0\\ & \forall R: \qquad w_R\geq 0.\end{aligned}$$

where the R's are taken in $\mathcal{R}(\mathcal{X}\times\mathcal{Y})$.

Theorem 3.4. $\widetilde{\mathrm{PAR}}_\mu(f) \geq \sum_{z\in\mathcal{Z}}\left|f^{-1}(z)\right|_\mu\cdot\mathrm{rec}^z(f)$.

Proof. Note first that, thanks to the first and the third conditions in the definition of $\mathrm{rec}^z(f)$, we can replace the second condition by $\forall(x,y)\in f^{-1}(\mathcal{Z}) : \sum_{R:R\ni(x,y)} w_R = 1$. Now, let us rewrite Definition 3.1 in the following way, that exhibits the independance of constraints for each f-monochromatic region $D\in\mathcal{D}^f$:

$$\begin{aligned}\min_{w_{z,R}}\sum_z\left|f^{-1}(z)\right|_\mu\sum_R w_{z,R} \quad \text{s.t.} \qquad & \forall\, z,\ \forall\,(x,y)\in f^{-1}(z): \sum_{R:R\ni(x,y)} w_{f(x,y),R} = 1\\ & \forall\, z,\ \forall\,(x,y)\in f^{-1}(z): \sum_{R:R\ni(x,y)}\sum_{z'} w_{z',R} = 1\\ & \forall\, z,\ \forall\, R: \qquad w_{z,R}\geq 0,\end{aligned}$$

where $z\in\mathcal{Z}$ and R denotes a rectangle in $\mathcal{R}(\mathcal{X}\times\mathcal{Y})$. The set of three constraints are independent from one region $f^{-1}(z)$ to another. Furthermore, since the $\left|f^{-1}(z)\right|_\mu$ are non-negative, we can replace the objective function by: $\sum_z\left|f^{-1}(z)\right|_\mu\cdot\min_{w_R}\sum_R w_R$, without changing either the set of solutions nor the optimal value. We finally get the formula (equality of the optimal values) as a consequence of LP formulations' equality. □

3.2 External PAR and the partition bound

Following Jain and Klauck [JK10], we define the partition lower bound in the following way:

Definition 3.5. *We define* $\mathrm{prt}(f)$ *as the optimal value of the following linear program:*

$$\begin{aligned} \min_{w_{z,R}} \ & \sum_z \sum_R w_{z,R} \\ \text{s.t.} \quad & \forall (x,y) \in \mathcal{X} \times \mathcal{Y} : \sum_{R : R \ni (x,y)} w_{f(x,y),R} = 1 \\ & \forall (x,y) \in \mathcal{X} \times \mathcal{Y} : \sum_{R : R \ni (x,y)} \sum_z w_{z,R} = 1 \\ & \forall z,\ \forall R : \quad w_{z,R} \geq 0. \end{aligned}$$

where $z \in \mathcal{Z}$ *and* R *denotes a rectangle in* $\mathcal{R}(\mathcal{X} \times \mathcal{Y})$.

From [JK10], we know that $\log \mathrm{prt}(f)$ is a lower bound on the communication complexity of f, that is even tighter than $\log \mathrm{rec}(f)$:

Theorem 3.6. *1.* $\mathbf{D}(f) \geq \log \mathrm{prt}(f)$.
2. $\mathrm{prt}(f) \geq \mathrm{rec}^z(f)$ *for every* $z \in \mathcal{Z}$.

We can relate this known bound to PAR in the case of a uniform output distribution:

Theorem 3.7. *Assume the distribution* μ *of the input is such that the distribution of the outputs is uniform, i.e.:* $\forall\, z \in \mathcal{Z} \qquad \left|f^{-1}(z)\right|_\mu = \mathbb{P}_{(x,y)\sim\mu}\left\{f(x,y) = z\right\} = \frac{1}{|\mathcal{Z}|}$, *then:*

$$\mathrm{PAR}^{ext}_\mu(f) \geq \widetilde{\mathrm{PAR}}_\mu(f) = \frac{1}{|\mathcal{Z}|} \cdot \mathrm{prt}(f).$$

Proof. It follows directly from PAR and partition linear program formulations. □

3.3 External PAR of boolean functions

We now turn our attention to the case of boolean functions: we prove lower bounds on PAR with respect to the number of monochromatic rectangles of the communication matrix and the communication matrix rank. We also prove that if the log-rank conjecture holds, PAR and communication complexity are polynomially related.

Let f be a *boolean* function (i.e. $|\mathcal{Z}| = 2$, so that we identify $\mathcal{Z}$ with $\{0,1\}$), P a deterministic protocol for f and T its transcript. Let n_0 and n_1 be the number of P-rectangles where the output is respectively 0 and 1 ($n_0 = |\mathcal{R}_0^P|, n_1 = |\mathcal{R}_1^P|$).

Theorem 3.8. $\mathrm{PAR}^{ext}_\mu(f) \geq \min\{\mathrm{rank}\left(\mathcal{M}_f\right), \mathrm{rank}\left(\mathcal{M}_{\mathrm{not}f}\right)\} \geq \mathrm{rank}\left(\mathcal{M}_f\right) - 1$

Proof. We can rewrite the PAR of any deterministic protocol for f as:

$$\mathrm{PAR}^{ext}_\mu(f,P) = \sum_{z \in \mathcal{Z}} \left|f^{-1}(z)\right|_\mu \cdot \mathrm{cut}_P(f^{-1}(z)) = \left|f^{-1}(0)\right|_\mu \cdot n_0 + \left|f^{-1}(1)\right|_\mu \cdot n_1 \tag{5}$$

where the first equality comes from Theorem 2.13. Then, $\mathrm{PAR}^{ext}_\mu(f,P) \geq \min(n_0, n_1)$ since $\left|f^{-1}(0)\right|_\mu + \left|f^{-1}(1)\right|_\mu = 1$. The theorem follows since for any protocol $\mathrm{rank}\left(\mathcal{M}_f\right) \leq n_1$ and $\mathrm{rank}\left(\mathcal{M}_{\mathrm{not}f}\right) \leq n_0$ (see for example [KN97]).

Observe that $\operatorname{rank}(\mathcal{M}_f)$ and $\operatorname{rank}(\mathcal{M}_{\text{not}f})$ differ by at most 1 ($\mathcal{M}_{\text{not}f} = J - \mathcal{M}_f$ where J is the all-ones matrix), hence we deduce the second statement. □

Relation with log-rank conjecture: The *log-rank conjecture* states that there exists γ such that $\mathbf{D}(f) \leq \log^{\gamma} \operatorname{rank}(\mathcal{M}_f)$.

Theorem 3.9. *The log-rank conjecture implies* $\mathrm{PAR}_{\mu}^{ext}(f)$ and $\mathbf{D}(f)$ are polynomially related for boolean f.

Proof. It suffices to note that $\mathbf{D}(f) + 1 \geq \log\left(\mathrm{PAR}_{\mu}^{ext}(f) + 1\right) \geq \log \operatorname{rank}(\mathcal{M}_f) \geq (\mathbf{D}(f))^{1/\gamma}$, where the first equality comes from Theorem 4.4 applied to deterministic protocols. □

Moreover, we show that Theorem 3.8 and Theorem 3.9 fail for non-boolean functions.

Rank argument fails for non-Boolean functions Theorem 3.8 and Theorem 3.9 are not true in general for non-boolean functions. For instance, consider the following function that take three values: let $\mathrm{EQ}' : \{1, \ldots, m\}^2 \to \{0, 1, 2\}$ be the function defined by:

$$\mathrm{EQ}'(x,y) = \begin{cases} 0 & \text{if } x \neq y \text{ and } x < m \text{ or } y < m \\ 1 & \text{if } x = y \text{ and } x < m \text{ or } y < m \\ 2 & \text{otherwise } (x = m \text{ or } y = m). \end{cases} \quad \text{whose matrix is:} \quad \mathcal{M}_{\mathrm{EQ}'} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 2 \\ 0 & 1 & \cdots & 0 & 2 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & 2 \\ 2 & 2 & \cdots & 2 & 2 \end{pmatrix}.$$

Then, for any (zero-error) protocol P solving EQ', the number of 0-rectangles and the number of 1-rectangles are at least the minimum number of such rectangles for EQ_{m-1}:

$$\mathrm{EQ}_{m-1} : \{1, \ldots, m-1\}^2 \to \{0, 1\},\ (x, y) \mapsto 1 \text{ iff } x = y.$$

But the number of 2-rectangles can be only 2. Now, if we pick a distribution μ and δ satisfying $\left|\mathrm{EQ}'^{-1}(0)\right|_{\mu} = \left|\mathrm{EQ}'^{-1}(1)\right|_{\mu} = \delta/2 < 2^{-(2m-2)}$ and $\left|\mathrm{EQ}'^{-1}(2)\right|_{\mu} = 1 - \delta$, then:

$$\begin{aligned} \mathrm{PAR}_{\mu}^{ext}(\mathrm{EQ}') &= \left|\mathrm{EQ}'^{-1}(0)\right|_{\mu} \cdot n_0 + \left|\mathrm{EQ}'^{-1}(1)\right|_{\mu} \cdot n_1 + \left|\mathrm{EQ}'^{-1}(2)\right|_{\mu} \cdot n_2 = \frac{\delta}{2} \cdot (n_0 + n_1) + 2 \cdot (1 - \delta) \\ &\leq 2^{-(2m-2)} \cdot 2^{2m-2} + 2 \cdot (1 - \delta) \qquad \text{(since } n_0 + n_1 \leq \#\{(x,y) \in \{1, \ldots, m-1\}^2\}) \\ &\leq 3. \end{aligned}$$

Hence for this function EQ' and this distribution μ:

$$\mathrm{PAR}_{\mu}^{ext}(\mathrm{EQ}', P) \leq 3$$
$$\text{whereas} \qquad \operatorname{rank}\left(\mathcal{M}_{\mathrm{EQ}'}\right) \geq \operatorname{rank}\left(\mathcal{M}_{\mathrm{EQ}_{n-1}}\right) = 2^{n-1}.$$

As a consequence, we see that Theorem 3.8 and Theorem 3.9 are not true in general for non-boolean functions.

Relation between PAR and fooling sets: Recall that a z-fooling set ($z \in \mathcal{Z}$) for $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ is a subset $F_z \subseteq f^{-1}(z)$ such that: $\forall\, (x,y) \in F_z,\ f(x,y) = z$ and $\forall\, (x_1,y_1), (x_2,y_2) \in F_z,\ (x_1,y_1) \neq (x_2,y_2)$ it holds that $f(x_1,y_2) \neq z$ or $f(x_2,y_1) \neq z$. Recall the following theorem about fooling sets and monochromatic rectangles:

Theorem 3.10 ([KN97])**.** *If F_z is a z-fooling set for f, then any covering of $f^{-1}(z)$ by monochromatic rectangles has at least $|F_z|$ rectangles.*

As a corollary of Theorem 2.13 and Theorem 3.10:

Corollary 3.11. *For any function f and any set of z-fooling sets $\{F_z\}_{z\in\mathcal{Z}}$,*

$$\mathrm{PAR}_\mu^{ext}(f) \geq \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_\mu \cdot |F_z|.$$

3.4 New lower bound techniques for external IC

We show lower bounds on the external zero-error information complexity, which using Theorem 2.8 will in turn give new lower bounds on information-theoretic privacy.

Theorem 3.12. *Fix a function f. Suppose there exists $\delta > 0$ and a distribution μ over the inputs of f whose support is a rectangle, such that for all monochromatic rectangles R of f, $\mu(R) \leq \delta$. Then it holds for every P that computes f with zero error that $\mathrm{IC}_\mu^{ext}(P) \geq \log(1/\delta)$.*

Proof. Suppose t is a transcript of the protocol P, let $\mathrm{supp}_{XY}(t)$ denote the support of inputs for transcript t, namely all x, y such that $\Pr[T_P(x,y) = t] > 0$. Since P has no error, $\mathrm{supp}_{XY}(t)$ is monochromatic. Also, since $\mathrm{supp}(\mu)$ is a rectangle, this implies that $\mathrm{supp}_{XY}(t)$ is a monochromatic rectangle. It holds that:

$$\begin{aligned}
\mathbf{I}(T_P(X,Y);(X,Y)) &= \mathbb{E}_{(x,y,t)\sim(X,Y,T_P(X,Y))} \log \tfrac{\mathbb{P}[(X,Y)=(x,y)|T_P(X,Y)=t]}{\mathbb{P}[(X,Y)=(x,y)]} \\
&= \mathbb{E}_{t\sim T_P} \sum_{(x,y)\in\mathrm{supp}_{XY}(t)} \mathbb{P}[(X,Y) = (x,y \mid T_P(X,Y) = t] \log \tfrac{\mathbb{P}[(X,Y)=(x,y)|T_P(X,Y)=t]}{\mathbb{P}[(X,Y)=(x,y)]} \\
&\geq \mathbb{E}_{t\sim T_P} \log \tfrac{1}{\sum_{(x,y)\in\mathrm{supp}_{XY}(t)} \mathbb{P}[(X,Y)=(x,y)]} \;=\; \mathbb{E}_{t\sim T_P} \log \tfrac{1}{\mathbb{P}[(X,Y)\in\mathrm{supp}_{XY}(t)]} \geq \log(1/\delta)
\end{aligned}$$

The first inequality is a consequence of the log-sum inequality and the second inequality follows because $\mathrm{supp}_{XY}(t)$ is a monochromatic rectangle and by the hypothesis that all monochromatic rectangles have mass at most δ. □

We prove the following corollary for fooling sets.

Corollary 3.13. *For any function f with a fooling set S of size $|S| = k$, there exists a distribution μ such that for all protocols P that compute f with zero error over μ, it holds that $\mathrm{IC}_\mu^{ext}(P) \geq \log k$.*

Proof. Let $S = \{(x_i, y_i)\}_{i=1,\ldots,k}$. Let z be the element such that $f(x_i, y_i) = z$ for all $(x_i, y_i) \in S$.

Set $\gamma = 2^{-k}$. Construct μ as follows: with probability $1-\gamma$ pick a random element of S, with probability γ choose (x_i, y_j) where $i \neq j$ are chosen uniformly from $[k]$.

Observe that the z-rectangles in the support of μ are exactly the singleton elements (x_i, y_i), and $\mu(x_i, y_i) = (1-\gamma)2^{-k} \leq 2^{-k}$. On the other hand, the total mass on all elements not labeled z is γ, so in particular the mass of any z'-rectangles for $z' \neq z$ is at most $\gamma = 2^{-k}$. Finally observe that the support of μ is a rectangle.

Therefore μ satisfies the hypotheses of Theorem 3.12 with $\delta = 2^{-k}$, and therefore $\mathrm{IC}_\mu^{ext}(P) \geq \log k$. □

Remark 3.14. *Theorem 3.12 can be used to prove an optimal lower bound on the zero-error information complexity of certain functions. For example, single bit AND function, the hard distribution μ is uniform over $(0,1),(1,0),(1,1)$, and our theorem implies that $\mathrm{IC}_\mu^{ext}(P) \geq \log_2 3$. This matches a recent upper bound by Braverman* et al. *[BGPW12]. We note that it is not possible to extend Theorem 3.12 for the internal information complexity, as the AND function (on single bits) has strictly smaller internal information complexity than external [BGPW12].*

3.5 Applications: tight bounds on external PAR and PRIV for specific functions

Our applications as described in Table 1 follow from the lower bounds techniques that we have seen and applying well known facts about the rank or the size of the fooling sets of the communication matrix of the functions in question. An advantage of our techniques is that they give bounds for *any* distribution of input μ, and not only for a uniform distribution as what was done in [FJS10]. Since any of these problems can be solved by sending Alice's entire input (n bits), the communication complexity is always upper-bounded by n, hence so PAR is always upper-bounded by 2^n.

Theorem 3.15. *For any distribution μ:*

1. $\mathrm{PAR}_{\mu}^{ext}(\mathrm{EQ}) = 2^n$
2. $\mathrm{PAR}_{\mu}^{ext}(\mathrm{DISJ}) = 2^n - 1$
3. $\mathrm{PAR}_{\mu}^{ext}(\mathrm{IP}) = 2^n - 1$
4. $\mathrm{PAR}_{\mu}^{ext}(\mathrm{GT}) \geq 2^n - 1$. *If $\mu = \mathcal{U}$ is uniform, then:* $\mathrm{PAR}_{\mathcal{U}}^{ext}(\mathrm{GT}) \geq 2^n + \frac{1}{2^{n+1}} - \frac{1}{2}$.

Proof. For each of the functions, the lower bound is attained by computing the rank of the matrix and applying Theorem 3.8. Since $\mathrm{rank}\,(\mathcal{M}_{not\mathrm{EQ}}) = 2^n$ too, we have $\min(n_0, n_1) = 2^n$ in the particular case of EQ. For GT and the uniform distribution, we count the 0- and 1-rectangles and apply equation (5) □

For PRIV we have:

Theorem 3.16. *For each of the following, there exists an appropriate distribution μ such that:*

1. $\mathrm{IC}_{\mu}^{ext}(\mathrm{EQ}) \geq n$, *and* $\mathrm{PRIV}_{\mu}^{ext}(\mathrm{EQ}) \geq n - 1$
2. $\mathrm{IC}_{\mu}^{ext}(\mathrm{DISJ}) \geq n$, *and* $\mathrm{PRIV}_{\mu}^{ext}(\mathrm{DISJ}) \geq n - 1$
3. $\mathrm{IC}_{\mu}^{ext}(\mathrm{IP}) \geq n - 1 - o(1)$, *and* $\mathrm{PRIV}_{\mu}^{ext}(\mathrm{IP}) \geq n - 2 - o(1)$
4. $\mathrm{IC}_{\mu}^{ext}(\mathrm{GT}) \geq n$, *and* $\mathrm{PRIV}_{\mu}^{ext}(\mathrm{GT}) \geq n - 1$. *If $\mu = \mathcal{U}$ is uniform, then:* $\mathrm{PRIV}_{\mathcal{U}}^{ext}(\mathrm{GT}) = O(1)$.

Proof. For the lower bounds for $\mathrm{EQ}, \mathrm{DISJ}, \mathrm{GT}$, we can apply Corollary 3.13 using an appropriate fooling set, followed by the relationship between IC and PRIV given in Theorem 2.8.

For IP we use the well-known fact that all 0-monochromatic rectangles of the IP function contain at most 2^n elements. Construct the distribution μ that with probability $\gamma = 2^{-n}$ picks a uniformly random x, y and with probability $1 - \gamma$ uniformly random non-zero x, y such that $\mathrm{IP}(x, y) = 0$. Clearly μ has support over a $(2^n - 1) \times (2^n - 1)$ rectangle, and it is balanced. Since the combined weight of all 1-inputs is at most 2^{-n}, the weight of all 1-monochromatic rectangles is at most 2^{-n}. Since each 0-monochromatic rectangle has at most 2^n elements and each of these receives weight $(1-\gamma)2/(2^n-1)^2$, the weight of each 0-monochromatic rectangle is at most $\frac{2}{2^n-1}$ and therefore $\log \frac{2^n-1}{2} > n - 1 - o(1)$. Applying Theorem 2.8 and the observation gives a lower bound of $n - 2 - o(1)$.

The fact that $\mathrm{PRIV}_{\mathcal{U}}^{ext}(\mathrm{GT}) = O(1)$ can be observed by considering the deterministic protocol that sends the inputs of both parties bit-by-bit, stopping when they reach a disagreeing bit. In expectation they exchange 2 bits before finding a disagreement, and therefore it holds that $\mathrm{PRIV}_{\mathcal{U}}^{ext}(\mathrm{GT}) \leq \mathrm{IC}_{\mathcal{U}}^{ext}(\mathrm{GT}) \leq \mathrm{Dist}_{\mathcal{U},0}(\mathrm{GT}) \leq O(1)$. □

Remark 3.17. *We note that in Theorem* 3.16 *it is possible to improve the lower bounds on* PRIV *for* $\mathrm{EQ}, \mathrm{DISJ}, \mathrm{GT}$ *from* $n - 1$ *to* $n - o(1)$ *and for* IP *from* $n - 2 - o(1)$ *to* $n - 1 - o(1)$. *This is achieved by noticing that the loss incurred by going from* IC *to* PRIV *can be bounded by the entropy of the output of the function, and for our choice of distributions this entropy is* $o(1)$.

4 Privacy for bounded-error protocols

4.1 Considerations for defining privacy for bounded-error protocols

We start by extending the definitions of PRIV and PAR to the bounded-error randomized model. We assume that the output of the protocol depends only on the transcript (*i.e.*$P(x,y)$ is a deterministic function of $T(x,y)$). Our definitions measures the amount of information learned by an observer conditioned on knowing the output of the *protocol*, rather than conditioned on knowing the output of the *function* as in Klauck's definition [K04].

At a high level our definitions attempt to measure how much additional information an observer (either exterior or one of the players) learns about the players' inputs above and beyond what he is supposed to learn.

The main choice in taking this approach that we must make is what constitutes "what is permissible for the observers to learn". The two most reasonable choices are the output of the protocol and the output of the function. (Observe that this dilemma does not exist for deterministic or zero-error protocols, since in that case the output of the protocol and the output of the function must be identical, but when there is non-zero error these can differ.)

We choose to define all of our measures in terms of the output of the protocol: our measures capture how much information is leaked above and beyond the output of the protocol. We believe this is the more natural choice for several reasons.[1] First it cleanly separates the issues of correctness and privacy: the privacy of a protocol does not vary depending on what "target" function it is trying to compute. Second, an observer learns the output of the protocol but she *does not know the output of the function*: she is only able to observe the transcript of the protocol and this does not necessarily reveal the value of the function if the protocol has error. In this sense, the output of the protocol corresponds more closely to "what is permissible for the players to learn by running the protocol".

Finally, on a technical level our definition is better-behaved: suppose that instead we used an alternative definition of external PAR that conditioned on the output of the function instead of the value of the protocol, for instance:

$$\mathrm{PAR}^{altext}_{\mu,\epsilon}(f) := \inf_{P\ \epsilon\text{-computes } f} \mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}_{X,Y}(f(X,Y)=z_t)}{\mathbb{P}_{X,Y,T}(T(X,Y)=t)}\right]$$

For any $\epsilon > 0$, consider some f such that $\Pr[f(X,Y)=1] = 1-\epsilon$. Then the protocol P that always outputs 1 without communication computes f with error ϵ and therefore $\mathrm{PAR}^{altext}_{\mathcal{U},\epsilon}(f) \leq 1-2\epsilon$. This is non-sensical, since the PAR should always be at least 1.

Definition 4.1. *We define:*

- *The external privacy of a randomized protocol P as:* $\mathrm{PRIV}^{ext}_{\mu}(P) := \mathbf{I}(T_P(X,Y); X,Y|P(X,Y))$. *For $\epsilon \geq 0$, the external ϵ-error privacy of f is defined as the following, where the infimum is taken over all protocols P computing f with error at most ϵ:* $\mathrm{PRIV}^{ext}_{\mu,\epsilon}(f) := \inf_P \mathrm{PRIV}^{ext}_{\mu}(P)$.
- *The internal privacy of a randomized protocol P:*

$$\mathrm{PRIV}^{int}_{\mu}(P) := \mathbf{I}(T_P(X,Y); X|P(X,Y),Y) + \mathbf{I}(T_P(X,Y); Y|P(X,Y),X)$$

  For $\epsilon \geq 0$, the internal ϵ-error privacy of f is defined as the following, where the infimum is taken over all protocols P computing f with error at most ϵ: $\mathrm{PRIV}^{int}_{\mu,\epsilon}(f) := \inf_P \mathrm{PRIV}^{int}_{\mu}(P)$.

In the following, the expectations are taken over inputs x, y, transcripts t resulting from computing $P(x,y)$ (where P may be randomized), and z_t which is the output contained in t.

[1] We observe that Klauck [K04] studied the other notion, conditioning on the output of the function.

Definition 4.2. *We define:*

- *The external* PAR *of a randomized protocol P as:* $\mathrm{PAR}^{ext}_{\mu}(P) := \mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}_{X,Y,P}(P(X,Y)=z_t)}{\mathbb{P}_{X,Y,T}(T(X,Y)=t)}\right]$. *For $\epsilon \geq 0$, the external ϵ-error* PAR *of f is defined as the following, where the infimum is taken over all protocols P computing f with error at most ϵ:* $\mathrm{PAR}^{ext}_{\mu,\epsilon}(f) := \inf_P \mathrm{PAR}^{ext}_{\mu}(P)$.
- *The internal* PAR *of a randomized protocol P as:*

$$\mathrm{PAR}^{int}_{\mu}(P) := \mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}_{X,Y,P}(P(X,Y)=z_t \wedge Y=y)}{\mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge Y=y)}\right] + \mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}_{X,Y,P}(P(X,Y)=z_t \wedge X=x)}{\mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge X=x)}\right].$$

  For $\epsilon \geq 0$, the external ϵ-error PAR *of f is defined as the following, where the infimum is taken over all protocols P computing f with error at most ϵ:* $\mathrm{PAR}^{int}_{\mu,\epsilon}(f) := \inf_P \mathrm{PAR}^{int}_{\mu}(P)$.

We can show that internal PAR is smaller than external one in the randomized bounded error setting (see Theorem 2.9 for 0-error deterministic case):

Theorem 4.3. *For any randomized bounded error protocol P computing f:*

$$\mathrm{PAR}^{int}_{\mu}(f,P) \leq 2 \cdot \mathrm{PAR}^{ext}_{\mu}(f,P).$$

Proof. Let ν denote the distribution of x, y, t. Then we have:

$$\begin{aligned}
\mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}_{X,Y,P}(P(X,Y)=z_t \wedge Y=y)}{\mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge Y=y)}\right] &= \sum_{x,y,t} \nu(x,y,t) \cdot \frac{\mathbb{P}_{X,Y,P}(P(X,Y)=z_t \wedge Y=y)}{\mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge Y=y)} \\
&= \sum_{y,z} \mathbb{P}_{X,Y,P}(P(X,Y)=z \wedge Y=y) \cdot \sum_{t:z_t=z} \frac{1}{\mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge Y=y)} \sum_{x} \nu(x,y,t) \\
&= \sum_{y,z} \mathbb{P}_{X,Y,P}(P(X,Y)=z \wedge Y=y) \cdot |\{t \in T_{(\cdot,y)} : z_t = z\}|,
\end{aligned}$$

since $\nu(x,y,t) = \mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge Y=y \wedge X=x)$ and $\sum_x \nu(x,y,t) = \mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge Y=y)$, where $T_{(\cdot,y)}$ is the set of all possible transcripts provided that $Y=y$. Similarly we prove:

$$\mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}_{X,Y,P}(P(X,Y)=z_t \wedge X=x)}{\mathbb{P}_{X,Y,T}(T(X,Y)=t \wedge X=x)}\right] = \sum_{x,z} \mathbb{P}_{X,Y,P}(P(X,Y)=z \wedge X=x) \cdot |\{t \in T_{(x,\cdot)} : z_t = z\}|.$$

Hence:

$$\begin{aligned}
\mathrm{PAR}^{int}_{\mu}(f,P) &= \sum_z |\{t \in T_{(x,\cdot)} : z_t = z\}| \cdot \sum_x \mathbb{P}_{X,Y,P}(P(X,Y)=z \wedge X=x) \\
&\quad + \sum_z |\{t \in T_{(\cdot,y)} : z_t = z\}| \cdot \sum_y \mathbb{P}_{X,Y,P}(P(X,Y)=z \wedge Y=y) \\
&\leq 2 \cdot \sum_z |\{t : z_t = z\}| \cdot \sum_x \mathbb{P}_{X,Y,P}(P(X,Y)=z) = 2 \cdot \mathrm{PAR}^{ext}_{\mu}(f,P),
\end{aligned}$$

since $\mathrm{PAR}^{ext}_{\mu}(f,P) = \sum_z \mathbb{P}(P(X,Y)=z) \sum_{t:z_t=z} \frac{\sum_{x,y}\nu(x,y,t)}{\mathbb{P}(T(X,Y)=t)} = \sum_z \mathbb{P}(P(X,Y)=z)|t : z_t = z|$. □

4.2 Bounds

We first observe that several bounds from the deterministic case carry over immediately: for example the proof Theorem 2.8 carries over immediately to the bounded-error case. Here we also show that for any protocol (deterministic or randomized), the external privacy-approximation ratio is at most exponential in the communication of the protocol. We also prove that for bounded-error randomized protocols PRIV is a lower bound for PAR both for the external and the internal case.

Theorem 4.4. *For any protocol P,* $\mathrm{PAR}_{\mu}^{ext}(P) \leq 2^{\mathbf{CC}(P)}$.

Proof. Fix a protocol P.

$$\begin{aligned}
\mathrm{PAR}_{\mu}^{ext}(P) &= \sum_{x,y,t} \mathbb{P}[(X,Y,T_P(X,Y)) = (x,y,t)] \cdot \mathbb{P}[P(X,Y) = z_t]/\mathbb{P}[T_P(X,Y) = t] \\
&= \sum_{x,y,t} \mathbb{P}[T_P(X,Y) = t] \cdot \mathbb{P}[(X,Y) = (x,y) | T_P(X,Y) = t] \cdot \mathbb{P}[P(X,Y) = z_t]/\mathbb{P}[T_P(X,Y) = t] \\
&\leq \sum_{t} \sum_{x,y} \mathbb{P}[(X,Y) = (x,y) | T_P(X,Y) = t] \\
&\leq |\{ \text{ possible } t\}| \\
&\leq 2^{\mathbf{CC}(P)}.
\end{aligned}$$

□

Theorem 4.5. *For any input distribution μ and any randomized protocol P, it holds that* $\mathrm{PRIV}_{\mu}^{ext}(P) \leq \log\left(\mathrm{PAR}_{\mu}^{ext}(P)\right)$ *and* $\mathrm{PRIV}_{\mu}^{int}(P) \leq 2 \cdot \log\left(\mathrm{PAR}_{\mu}^{int}(P)\right)$. *As a consequence, $\forall \mu, f, \epsilon$ it holds that* $\mathrm{PRIV}_{\mu,\epsilon}^{ext}(f) \leq \log\left(\mathrm{PAR}_{\mu,\epsilon}^{ext}(f)\right)$ *and* $\mathrm{PRIV}_{\mu,\epsilon}^{int}(f) \leq \log\left(\mathrm{PAR}_{\mu,\epsilon}^{int}(f)\right)$.

Proof. **External:** notice that, by the definition of conditional probability

$$\mathbb{P}\left\{(X,Y) = (x,y) | T(X,Y) = t\right\} = \frac{\mathbb{P}\{(X,Y)=(x,y) \wedge T(X,Y)=t\}}{\mathbb{P}\{T(X,Y)=t\}},$$

$$\mathbb{P}\left\{(X,Y) = (x,y) | P(X,Y) = z_t\right\} = \frac{\mathbb{P}\{(X,Y)=(x,y) \wedge P(X,Y)=z_t\}}{\mathbb{P}\{P(X,Y)=z_t\}}.$$

Using the fact that the transcript determines the output of the protocol and that for any two variables X, Y, $\mathbf{H}(X|Y) = \mathbb{E}[-\log(\mathbb{P}(X|Y)]$, we have

$$\begin{aligned}
\mathrm{PRIV}_{\mu}^{ext}(P) &= \mathbb{E}\left[\log \frac{\mathbb{P}\{(X,Y)=(x,y) \wedge T(X,Y)=t\}}{\mathbb{P}\{T(X,Y)=t\}} \cdot \frac{\mathbb{P}\{P(X,Y)=z_t\}}{\mathbb{P}\{(X,Y)=(x,y) \wedge P(X,Y)=z_t\}}\right] \\
&= \mathbb{E}\left[\log \frac{\mathbb{P}\{(X,Y)=(x,y) \wedge T(X,Y)=t\}}{\mathbb{P}\{(X,Y)=(x,y) \wedge P(X,Y)=z_t\}} \cdot \frac{\mathbb{P}\{P(X,Y)=z_t\}}{\mathbb{P}\{T(X,Y)=t\}}\right] \\
&\leq \mathbb{E}\left[\log 1 \cdot \frac{\mathbb{P}\{P(X,Y)=z_t\}}{\mathbb{P}\{T(X,Y)=t\}}\right] \;\leq\; \log \mathbb{E}\left[\frac{\mathbb{P}\{P(X,Y)=z_t\}}{\mathbb{P}\{T(X,Y)=t\}}\right] \;=\; \log \mathrm{PAR}_{\mu}^{ext}(P),
\end{aligned}$$

where the the second inequality follows from the concavity of $\log$ and the first inequality comes from the fact that: $\{(X,Y) = (x,y) \wedge P(X,Y) = z_t\} \supseteq \{(X,Y) = (x,y) \wedge T(X,Y) = t\}$. **Internal:** let us prove that:

$$\mathbb{E}\left[\log \frac{\mathbb{P}\left\{X = x | T(X,Y) = t \wedge P(X,Y) = z_t \wedge Y = y\right\}}{\mathbb{P}\left\{X = x | P(X,Y) = z_t \wedge Y = y\right\}}\right] \leq \log \mathbb{E}\left[\frac{\mathbb{P}(P(X,Y) = z_t \wedge Y = y)}{\mathbb{P}(T(X,Y) = t \wedge Y = y)}\right].$$

Notice that, by the definition of conditional probability and the fact that the transcript determines the output of the protocol:

$$\mathbb{P}\left\{X = x | T(X,Y) = t \wedge Y = y\right\} = \frac{\mathbb{P}\left\{X = x \wedge T(X,Y) = t \wedge Y = y\right\}}{\mathbb{P}\left\{T(X,Y) = t \wedge Y = y\right\}}$$

and similarly:

$$\mathbb{P}\left\{X = x | P(X,Y) = z_t \wedge Y = y\right\} = \frac{\mathbb{P}\left\{X = x \wedge P(X,Y) = z_t \wedge Y = y\right\}}{\mathbb{P}\left\{P(X,Y) = z_t \wedge Y = y\right\}}$$

Hence:

$$\begin{aligned}
&\mathbb{E}\left[\log \frac{\mathbb{P}\left\{X=x|T(X,Y)=t\wedge P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{X=x|P(X,Y)=z_t\wedge Y=y\right\}}\right]\\
&=\mathbb{E}\left[\log \frac{\mathbb{P}\left\{X=x\wedge T(X,Y)=t\wedge Y=y\right\}}{\mathbb{P}\left\{T(X,Y)=t\wedge Y=y\right\}}\cdot\frac{\mathbb{P}\left\{P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{X=x\wedge P(X,Y)=z_t\wedge Y=y\right\}}\right]\\
&=\mathbb{E}\left[\log \frac{\mathbb{P}\left\{X=x\wedge T(X,Y)=t\wedge Y=y\right\}}{\mathbb{P}\left\{X=x\wedge P(X,Y)=z_t\wedge Y=y\right\}}\cdot\frac{\mathbb{P}\left\{P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{T(X,Y)=t\wedge Y=y\right\}}\right]\\
&\leq\mathbb{E}\left[\log 1\cdot\frac{\mathbb{P}\left\{P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{T(X,Y)=t\wedge Y=y\right\}}\right]\\
&\leq\log\mathbb{E}\left[\frac{\mathbb{P}\left\{P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{T(X,Y)=t\wedge Y=y\right\}}\right],
\end{aligned}$$

where the first inequality comes from the fact that:

$$\{(X,Y)=(x,y)\wedge P(X,Y)=z_t\}\supseteq\{(X,Y)=(x,y)\wedge T(X,Y)=t\}$$

and the second inequality follows from the concavity of $\log$.
By symmetry in x and y, we also have:

$$\mathbb{E}\left[\log \frac{\mathbb{P}\left\{Y=y|T(X,Y)=t\wedge P(X,Y)=z_t\wedge X=x\right\}}{\mathbb{P}\left\{Y=y|P(X,Y)=z_t\wedge X=x\right\}}\right]\leq\log\mathbb{E}\left[\frac{\mathbb{P}(P(X,Y)=z_t\wedge X=x)}{\mathbb{P}(T(X,Y)=t\wedge X=x)}\right].$$

So we have:

$$\mathrm{PRIV}_\mu^{int}(P)\leq\log\mathbb{E}\left[\frac{\mathbb{P}\left\{P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{T(X,Y)=t\wedge Y=y\right\}}\right]+\log\mathbb{E}\left[\frac{\mathbb{P}(P(X,Y)=z_t\wedge X=x)}{\mathbb{P}(T(X,Y)=t\wedge X=x)}\right].$$

Hence:

$$\begin{aligned}
\text{either:}\qquad & 2^{\frac{1}{2}\cdot\mathrm{PRIV}_\mu^{int}(P)}\leq\mathbb{E}\left[\frac{\mathbb{P}\left\{P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{T(X,Y)=t\wedge Y=y\right\}}\right]\\
\text{or:}\qquad & 2^{\frac{1}{2}\cdot\mathrm{PRIV}_\mu^{int}(P)}\leq\mathbb{E}\left[\frac{\mathbb{P}(P(X,Y)=z_t\wedge X=x)}{\mathbb{P}(T(X,Y)=t\wedge X=x)}\right].
\end{aligned}$$

Finally:

$$\begin{aligned}
\frac{1}{2}\cdot\mathrm{PRIV}_\mu^{int}(P)&\leq\log\left(\mathbb{E}\left[\frac{\mathbb{P}\left\{P(X,Y)=z_t\wedge Y=y\right\}}{\mathbb{P}\left\{T(X,Y)=t\wedge Y=y\right\}}\right]+\mathbb{E}\left[\frac{\mathbb{P}(P(X,Y)=z_t\wedge X=x)}{\mathbb{P}(T(X,Y)=t\wedge X=x)}\right]\right)\\
&=\log\mathrm{PAR}_\mu^{int}(P)
\end{aligned}$$

□

4.3 Applications: tight bounds on PAR and PRIV for specific functions

Note that internal PRIV is lower bounded by Information Complexity, which was shown in [KLLRX12] to subsume almost all known lower bounds for communication complexity, i.e. smooth rectangle, γ_2-norm bound, discrepancy, etc. Hence, for the bounded-error case, the two notions of privacy are in fact both equal to the communication complexity for all boolean functions for which we have a tight bound on their communication complexity. Interestingly, the notion of PAR sits between information and communication complexity, and it is an important open question whether these two notions are equal (which would also make PAR equal to them). For the bounds in Table 1, the results follow immediately from known lower bounds on the IC of these functions: for EQ the lower bound is trivial, for DISJ one can look at IC directly [BJKS02, Bra11], while for EQ, IP, GT one can look at their discrepancies [BW12]. Then, using Theorem 2.8 and Theorem 4.5 we obtain bounds on internal PAR. Note that the bounds also hold for external PRIV and PAR (since internal is always at most external, see Theorem 2.10). Moreover, we can also get similar lower bounds for the functions Vector in Subspace and Gap-Hamming distance by the results in [KLLRX12].

5 Quality of the two definitions

5.1 Privacy for deterministic protocols:

For deterministic protocols, the two definitions of privacy, PRIV and PAR, can be arbitrarily different for the same distribution. In high level, PRIV captures the expected privacy loss of a protocol, while PAR captures a more "risk-averse" notion of privacy, where a protocol is penalized heaviliy for high-privacy-loss events, even if they occur with small probability.

We show that this difference makes PRIV a much more robust definition: an ϵ change in the input distribution causes at most an ϵn change in PRIV, so PRIV is "smooth". Furthermore, PRIV always remains less than the expected communication of the protocol, which we believe to be another natural property. We prove that this is not the case for PAR: sometimes an ϵ change in the input distribution can cause PAR to change exponentially, and PAR can grow arbitrarily larger than the expected communication. Finally we also point out an error in the appendix of [FJS10] and show that for the example they gave, in fact PRIV is just as good as PAR at distinguishing two protocols in their example.

5.1.1 Robustness over the input distribution

We show that PAR is not robust over the input distribution μ. More precisely, we give an example of a function and of two distributions with exponentially small statistical distance, but whose privacy-approximation ratio is constant for one and exponential for the other.

Proposition 5.1. *There exists a function f and two input distributions μ_1, μ_2 satisfying $|\mu_1 - \mu_2| \leq 2^{-n/2}$ in statistical distance, and yet such that $\mathrm{PAR}^{ext}_{\mu_1}(f) = \Theta(1)$ and $\mathrm{PAR}^{ext}_{\mu_2}(f) = \Omega(2^{n/2})$.*

Proof. Let $m = 2^n$ and $f : \{0, \ldots, m\}^2 \to \{0, 1, 2\}$ be the function defined by:

$$f(x,y) = \begin{cases} 0 & \text{if } x \neq y \text{ and } x \neq m \text{ and } y \neq m \\ 1 & \text{if } x = y \text{ and } x \neq m \text{ and } y \neq m \\ 2 & \text{otherwise } (x = m \text{ or } y = m). \end{cases} \quad \text{whose matrix is:} \quad \mathcal{M}_f = \begin{pmatrix} 1 & 0 & \cdots & 0 & 2 \\ 0 & 1 & \cdots & 0 & 2 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & 2 \\ 2 & 2 & \cdots & 2 & 2 \end{pmatrix}.$$

Let μ_1 be the following distribution: with probability 2^{-n} pick a random element of $f^{-1}(0) \cup f^{-1}(1)$, and with probability $1 - 2^{-n}$ pick a random element of $f^{-1}(2)$.

Set $\epsilon = 2^{-n/2}$ and let μ_2 be the following distribution: with probability $2^{-n} + \epsilon$ pick a random element of $f^{-1}(0) \cup f^{-1}(1)$, and with probability $1 - 2^{-n} - \epsilon$ pick a random element of $f^{-1}(2)$.

Consider now the protocol P, where first Alice and Bob exchange a single bit to check whether $x = m$ or $y = m$ and if they are both different than m, Alice and Bob solve Equality (by having Alice send her entire input to Bob).

Then we have:

$$\begin{aligned} \mathrm{PAR}^{ext}_{\mu_1}(f) &\leq \mathrm{PAR}^{ext}_{\mu_1}(f, P) = \left|f^{-1}(0)\right|_\mu \cdot n_0 + \left|f^{-1}(1)\right|_\mu \cdot n_1 + \left|f^{-1}(2)\right|_\mu \cdot n_2 \\ &\leq (\left|f^{-1}(0)\right|_\mu + \left|f^{-1}(1)\right|_\mu) \cdot 2^n + \left|f^{-1}(2)\right|_\mu \cdot 3 = \Theta(1) \end{aligned}$$

On the other hand,any protocol for this function must solve Equality so n_0 and n_1 must be at least 2^n,

since they have to be larger than the rank of the matrix. Consider the optimal protocol P for f

$$\begin{aligned}\mathrm{PAR}^{ext}_{\mu_2}(f) &= \mathrm{PAR}^{ext}_{\mu_2}(f,P) = \left|f^{-1}(0)\right|_{\mu}\cdot n_0 + \left|f^{-1}(1)\right|_{\mu}\cdot n_1 + \left|f^{-1}(2)\right|_{\mu}\cdot n_2\\ &\geq (\left|f^{-1}(0)\right|_{\mu} + \left|f^{-1}(1)\right|_{\mu})\cdot 2^n\\ &= (\frac{1}{2^n}+\epsilon)\cdot 2^n = \Omega(2^{n/2}).\end{aligned}$$

One can finally verify that $|\mu_1 - \mu_2| = \epsilon = 2^{-n/2}$. □

In fact, the right way to look at the robustness of PAR is to talk about $\log \mathrm{PAR}^{ext}_{\mu}(f)$. Even in this case, we see that an exponentially small change to the input distribution can change the $\log \mathrm{PAR}^{ext}_{\mu}(f)$ from constant to $\Omega(n)$.

On the other hand, we can prove that when the statistical distance of the input distributions is ϵ, then the PRIV changes by at most $O(\epsilon n)$. This implies, that in our previous example, PRIV changes only by an exponentially small amount.

Theorem 5.2. *For any protocol P and any two input distributions μ, μ' with statistical distance $|\mu - \mu'| \leq \epsilon$, it holds that*

$$\begin{aligned}|\mathrm{PRIV}^{ext}_{\mu}(P) - \mathrm{PRIV}^{ext}_{\mu'}(P)| &\leq O(\epsilon n)\\ |\mathrm{PRIV}^{int}_{\mu}(P) - \mathrm{PRIV}^{int}_{\mu'}(P)| &\leq O(\epsilon n)\end{aligned}$$

Proof. The proof is a consequence of the fact that two statistically close joint distributions must have similar mutual information. To prove this formally we use the following lemma:

Lemma 5.3 (Lemma 3.15 of [MX12])**.** *For any random variables $XY, X'Y'$ such that $|XY - X'Y'| \leq \epsilon$ and where X, X' take value in $\{0,1\}^n$, it holds that*

$$|H(X \mid Y) - H(X' \mid Y')| \leq 4(H(\epsilon) + \epsilon n)$$

Let $XY \sim \mu$ and T be the output of the transcript of P applied to XY and Z be the value output by T. Similarly define $X'Y'T'Z'$ with respect to μ'. Because $|\mu - \mu'| \leq \epsilon$ it also holds that $|XYTZ - X'Y'T'Z'| \leq \epsilon$. Therefore, it holds that:

$$\begin{aligned}|\mathrm{PRIV}^{ext}_{\mu}(P) - \mathrm{PRIV}^{ext}_{\mu'}(P)| &= |I(XY;T \mid Z) - I(X'Y';T' \mid Z')|\\ &= |H(XY \mid Z) - H(X'Y' \mid Z') + H(X'Y' \mid T'Z') - H(XY \mid TZ)|\\ &\leq |H(XY \mid Z) - H(X'Y' \mid Z')| + |H(X'Y' \mid T'Z') - H(XY \mid TZ)|\\ &\leq 8(H(\epsilon) + \epsilon n)\end{aligned}$$

A similar derivation holds for the internal privacy cost. □

5.1.2 Relationship between communication and privacy

A natural methodology for studying privacy is to measure the amount of information revealed by the transcript above and beyond what is supposed to be revealed. We believe that both PRIV and PAR were designed with this methodology in mind.

One intuitive bound that "natural" measures of information should satisfy is the following: a transcript of length c can reveal at most c bits of information. As a consequence, the privacy loss should also be bounded by the communication (appropriately normalized of course: for example in the case of PAR, one would compare $\log \mathrm{PAR}$ to communication).

When taking an expectation over randomized protocols, as one does for instance when measuring the complexity of zero-error randomized protocols, one would therefore also expect that the privacy loss revealed should be bounded by the expected communication. While PRIV does indeed satisfy this property, we observe that PAR does not:

Remark 5.4. *For the* `Greater Than` *function* GT *under the uniform input distribution $\mathcal{U}$, the following holds:*

1. *For all zero-error protocols P solving* GT, $\mathrm{PAR}^{ext}_{\mathcal{U}}(P) \geq 2^n - 1$.
2. *There exist a zero-error protocol for* GT *where the expected communication is constant.*

The first point was proved in Theorem 3.15. The second point follows from the trivial protocol that exchanges their inputs bit-by-bit starting with the highest order bits until the players find a difference, at which point they terminate because they know which player has the greater value. Then clearly under uniform inputs, for each $i \geq 1$ the probability of terminating after $2i$ bits is $1-2^{-i}$, and so the expected communication is $2\sum_{i=1}^{\infty} i \cdot 2^{-i} = 4$ regardless of the size of the inputs.

Thus, the above remark shows that PAR can tend to infinity even though the expected communication is constant, which violates the "natural" property that c bits of communication can reveal at most c bits of information.

On the other hand, one could argue that PAR captures a "risk-averse" notion of privacy, where one does not want the expected privacy loss but rather the privacy loss with higher weights assigned to high-privacy-loss events. In this case one may also want to look at worst-case choices of inputs and random coins; worst-case inputs were defined in [FJS10, ACCFKP12], although they did not study worst-case random coins since they focused on deterministic protocols.

5.2 Bounded-error case:

As we explained in section 4.3, in the case of bounded-error randomized protocols, the two notions of privacy are in fact both equal to the communication complexity for all boolean functions for which we have a tight bound on their communication complexity. Moreover, for functions with large output, we still do not have any example where PRIV and PAR are different when we are allowed bounded error.

5.3 Error in appendix of [FJS10]

An example was given in the appendix of [FJS10] that claimed to exhibit a function f and two protocols P, Q such that $\mathrm{PAR}^{ext}_{\mathcal{U}}(P) = O(1)$ and $\mathrm{PAR}^{ext}_{\mathcal{U}}(Q) = 2^{\Omega(n)}$, whereas it was claimed that $\mathrm{PRIV}^{ext}_{\mathcal{U}}(P) = \mathrm{PRIV}^{ext}_{\mathcal{U}}(Q) = \Theta(n)$. This was interpreted to mean that PRIV was not sufficiently precise enough to capture the difference between these two protocols.

However the second claim is incorrect as a simple calculation reveals that $\mathrm{PRIV}^{ext}_{\mathcal{U}}(P) = O(1)$ and so PRIV does indeed distinguish between the two protocols. The flaw in their argument was in using the geometric interpretation of PRIV: the characterization of [BYCKO93] that they use only applies to the *worst* distribution for a function (which for the function they give is *not* uniform), whereas they explicitly want to study the uniform distribution. For the worst distribution μ it is indeed the case that $\mathrm{PRIV}^{ext}_{\mu}(P) = \Theta(n)$, but not for the uniform distribution. Therefore, for their example, PRIV is actually just as capable as PAR in distinguishing the two protocols P, Q.

Conclusion

We proved new lower bound techniques both for PAR and PRIV (or IC), which enabled us to give tight bounds for the privacy of several functions. In fact, for boolean functions we believe that our techniques give tight bounds for almost all interesting functions. We also extended the definitions of privacy to bounded-error randomized protocols and showed that PRIV is a lower bound on PAR, which in turn is less than the randomized communication complexity. Since PRIV (in fact IC) subsumes most of the known lower bound techniques for communication complexity, we get tight lower bounds for a large number of boolean functions. For functions with large output, our techniques do not provide any strong bounds, which in fact is justified, since some of these functions (for example, Vickrey Auction) admit perfectly private protocols. Ultimately, we would like to understand the tradeoff between the communication complexity of a protocol and its privacy.

References

[ACCFKP12] A. Ada, A. Chattopadhyay, S. Cook, L. Fontes, M. Koucký, T. Pitassi. *The Hardness of Being Private*. CCC 2012.

[BBCR10] B. Barak, M. Braverman, X. Chen, and A. Rao. *How to compress interactive communication*. In Proc. 42nd STOC, pages 67–76, 2010.

[BGPW12] M. Braverman, A. Garg, D. Pankratov, and O. Weinstein. *Private communication.*

[BJKS02] Z. Bar-Yossef, T. Jayram, R. Kumar, and D. Sivakumar. *An information statistics approach to data stream and communication complexity*. In Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, pages 209–218, 2002.

[Bra11] M. Braverman. *Interactive information complexity*. ECCC, report No. 123 (2011); STOC'12

[BYCKO93] R. Bar-Yehuda, B. Chor, E. Kushilevitz, and A. Orlitsky. *Privacy, additional information and communication*. IEEE Transactions on Information Theory 39(6)(1993).

[CT06] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, Second Edition, ISBN: 0-471-24195-4, Hardcover, 776 pages, July 2006.

[CSWY01] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao, *Informational complexity and the direct sum problem for simultaneous message complexity*. 42nd IEEE FOCS, pp. 270–278, 2001.

[FJS10] J. Feigenbaum, A. D. Jaggard, and M. Schapira. *Approximate Privacy: Foundations and Quantification*. Proceedings of the 11th Conference on Electronic Commerce (EC), ACM Press, New York, 2010, pp. 167 – 178.

[JK10] R. Jain, H. Klauck. *The Partition Bound for Classical Communication Complexity and Query Complexity*. 25th IEEE Conference on Computational Complexity, 2010.

[K04] H. Klauck, *On quantum and approximate privacy*. Proc. STACS, 2002.

[KLLRX12] I. Kerenidis, S. Laplante, V. Lerays, J. Roland, and D. Xiao. *Lower bounds on information complexity via zero-communication protocols and applications*. FOCS 2012, pages 500-509, 2012.

[KN97] E. Kushilevitz and N. Nisan. *Communication complexity*. Cambridge University Press, 1997.

[L90] L. Lovasz. *Communication Complexity: A Survey*. In Paths, Flows, and VLSI Layout, edited by B. H. Korte, Springer, 1990.

[MX12] M. Mahmoody and D. Xiao. *Languages with efficient Zero Knowledge **PCP**s are in **SZK***. ECCC Technical Report TR2012-052, 2012.

[Yao79] A. C-C. Yao. *Some complexity questions related to distributive computing*. In Proceedings of the 11th ACM Symposium on Theory of Computing (STOC), pages 209–213, 1979.

[BW12] M. Braverman and O. Weinstein. *A Discrepancy Lower Bound for Information Complexity.* In Proc. APPROX-RANDOM 2012, p. 459-470, 2012.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de