New Lower Bounds for Privacy in Communication Protocols

Iordanis Kerenidis[1,2], Mathieu Laurière[3(✉)], and David Xiao[1]

[1] CNRS, LIAFA, Université Paris 7, Paris, France
{jkeren,dxiao}@liafa.univ-paris-diderot.fr
[2] CQT, NUS Singapore, Singapore, Singapore
[3] LIAFA, Université Paris 7, Paris, France
lauriere@liafa.univ-paris-diderot.fr

Abstract. Communication complexity is a central model of computation introduced by Yao [22], 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 ([9,15]), which for Boolean functions is equivalent to the notion of information cost introduced by [12] and that has since found many important applications; second, a combinatorial definition introduced by [13] and further developed by [1].

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.

Keywords: Communication complexity · Information complexity · Lower bound · Privacy

1 Introduction

Communication complexity is a central model of computation, first defined by Yao, [22], that has found applications in many areas of theoretical computer

C. Padró (Ed.): ICITS 2013, LNCS 8317, pp. 69–89, 2014.
DOI: 10.1007/978-3-319-04268-8_5,

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. The loss of privacy measures how much information about (x, y) is leaked to an eavesdropper who has only access to the transcript (*external* 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 $x, y \in \{0, 1\}^n$ and the identity of the person holding it, then the deterministic communication protocol with optimal communication is the trivial protocol of complexity $2n$. In fact one can show that any deterministic protocol that has optimal communication 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 privacy, Bar-Yehuda *et al.* [9] 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* ($\mathrm{IC}_\mu^{int}(f)$). At a 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* ($\mathrm{IC}_\mu^{ext}(f)$) was defined in [12] 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, including better communication lower bounds for important functions, direct sum theorems and new compression schemes [2,3,5–8,12,20].

Klauck [15] also defined an information theoretic notion of privacy, which we denote here by $\mathrm{PRIV}_\mu^{int}(f)$, which is closely related to the internal information cost (the only difference being that we subtract the information that the function reveals about the inputs, which the players are allowed to learn). 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 of being easily related to other tools in information theory, but are not easily seen in a combinatorial way.

Feigenbaum *et al.* [13] gave a combinatorial definition of privacy for the uniform distribution over inputs that was extended by Ada *et al.* [1] to any

distribution μ, called *average case objective privacy-approximation ratio*, that we will refer to simply as *external privacy-approximation ratio* (we only study this average-case notion, and not a related worst-case notion also defined in that work), and we denote this by $\mathrm{PAR}_{\mu}^{ext}(f, P)$. It 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 from (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 *(average) subjective (or internal) privacy-approximation ratio* (here again we will omit "average") which we denote $\mathrm{PAR}_{\mu}^{int}(f, P)$, which captures how much more one player learns about the input of the other one by the transcript than by 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 transcript plus the symmetric ratio for Bob. Last, they computed lower bounds for the privacy-approximation ratio of several functions, however restricting themselves to the case of uniformly distributed inputs.

More recently, Ada *et al.* in [1] 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 [13] and [1] only deterministic protocols were considered. Moreover, the relation between PRIV and PAR was not very well understood.

Our Results: We prove new relationships between PRIV, PAR and communication complexity, as well as providing new lower bound techniques for the two notions of privacy, PRIV and PAR, both external and internal, enabling us to give tight bounds for the privacy of various functions in the case of deterministic protocols. We also extend the definitions of PRIV and PAR to bounded-error randomized protocols, and derive linear lower bounds for various functions.

New lower bounds for external PAR *of deterministic protocols for boolean functions:* For boolean functions we give new lower bounds techniques, relating it to the rank of the function and the deterministic complexity.

Theorem 1. *For boolean f, for any distribution μ with full support,* $\mathrm{PAR}_{\mu}^{ext}(f) \geq \mathrm{rank}\,(\mathcal{M}_f) - 1$.

Theorem 2. *For boolean f, for any distribution μ with full support,* $\log \mathrm{PAR}_{\mu}^{ext}(f) \geq \sqrt{\mathbf{D}(f)}$.

Observe that this implies that $\log \mathrm{PAR}_{\mu}^{ext}(f)$ is in fact *polynomially* related to the deterministic communication complexity. Notably, it therefore holds that

the *only* boolean functions with low privacy loss (as measured using PAR^{ext}) are functions that have low communication complexity (this is not the case with non-boolean functions as was already observed by [13]).

New lower bounds for external PAR *of deterministic protocols for non-boolean functions:* For simplicity we restrict ourselves to full support distibutions μ, but it is possible to extend the results to general distributions by considering summations over only the rectangles whose intersection with μ's support is not empty. First, we present a general lower bound technique for $\mathrm{PAR}^{ext}_{\mu}(f)$ via linear programming. We relate it to two other well known lower bound techniques for communication complexity (see [14]): the *rectangle* bound ($\mathrm{rec}(f)$) and the *partition* bound ($\mathrm{prt}(f)$). This linear program, whose optimal value is denoted by $\widetilde{\mathrm{PAR}}_{\mu}(f)$, 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 3. *For all f, for any distribution μ with full support,* $\mathrm{PAR}^{ext}_{\mu}(f) \geq \widetilde{\mathrm{PAR}}_{\mu}(f)$.

Theorem 4. *For all f, for any distribution μ with full support,* $\widetilde{\mathrm{PAR}}_{\mu}(f) \geq \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_{\mu} \cdot \mathrm{rec}^z(f)$.

Moreover, we bound external PAR as a weighted sum of the size of the z-fooling sets F_z of $\mathcal{M}_f$:

Theorem 5. *For all f, for any distribution μ with full support,* $\mathrm{PAR}^{ext}_{\mu}(f) \geq \sum_z \left|f^{-1}(z)\right|_{\mu} \cdot |F_z|$.

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 12, will in turn give new lower bounds on $\mathrm{PRIV}^{ext}_{\mu}(f)$.

Theorem 6. *Fix a function f. Suppose there exists $\delta > 0$ and a distribution μ over the inputs of f, such that for all monochromatic rectangles R of f, $\mu(R) \leq \delta$. Then it holds for every protocol P that computes f without error on any input that* $\mathrm{IC}^{ext}_{\mu}(P) \geq \log(1/\delta)$.

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

These are much sharper than typical lower bounds on IC, which work in the bounded-error case and incur multiplicative constants [6,7,10,16]. The only other such sharp lower bounds we are aware of are due to Braverman *et al.* [4] who study the AND and DISJ functions. However they prove sharp bounds for the internal IC of DISJ, not for the external IC as we study here.

Our bound proves an *optimal* lower bound on the zero-error information cost of certain functions (i.e. without any additive constant loss). For the one bit AND, we show that there exists μ with $\mathrm{IC}_{\mu}^{ext}(AND) \geq \log_2 3$. This matches the bound of [4] (they also proved optimality via different techniques).

Privacy for bounded-error randomized protocols: We define for the first time PAR and PRIV for bounded-error protocols. Such protocols can be much more efficient than deterministic ones and it is important to see whether they remain private or not. These definitions capture again how much more information is leaked by the protocol than by the output of the function, where now we consider randomized protocols that compute the function with some bounded error. We show that for any protocol, PRIV is a lower bound on PAR, both for the external and internal notions.

Theorem 8. $\forall \mu, f, \epsilon, \quad \mathrm{PRIV}_{\mu,\epsilon}^{ext}(f) \leq \log \mathrm{PAR}_{\mu,\epsilon}^{ext}(f)$ *and* $\mathrm{PRIV}_{\mu,\epsilon}^{int}(f) \leq 2 \log \mathrm{PAR}_{\mu,\epsilon}^{int}(f) - 2$.

Internal PRIV is lower bounded by internal IC, which was shown in [16] to subsume almost all known lower bounds for communication complexity, i.e. smooth rectangle, γ_2-norm, discrepancy, etc. Hence,

Corollary 1 (Informal). *In the bounded error setting, for all boolean f whose internal information complexity equals communication complexity, all notions of privacy loss (PRIV, PAR, external, internal) are equivalent to each other and to the communication complexity.*

Interestingly, PAR sits between information and communication complexity, and it is an important open question whether these two notions are equal for all functions (and hence make PAR equal to them).

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 [13] and more: `Equality`, `Disjointness`, `Inner Product`, `Greater Than` (`Millionaire's problem`).

Table 1. Lower bounds for specific functions, zero error.

Problem	$\mathbf{PAR}_{\mu}^{\mathbf{ext}}$		$\mathbf{PRIV}_{\mu}^{\mathbf{ext}}$
	[13] (for uniform μ)	**Our contribution** (for μ with full support)	(for some μ)
`Equality`	-	2^n	$n-1$
`Disjointness`	$\left(\frac{3}{2}\right)^n$	$2^n - 1$	$n-1$
`Inner Product`	-	$2^n - 1$	$n - 2 - o(1)$
`Greater Than`	$2^n + \frac{1}{2^{n+1}} - \frac{1}{2}$	$2^n - 1$	$n-1$

Comparison between the two notions of privacy: 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

Table 2. Lower bounds for specific functions, with bounded error

	$\mathbf{PRIV}^{\mathbf{int}}_{\mu,\epsilon}, \mathbf{PRIV}^{\mathbf{ext}}_{\mu,\epsilon}$ (for some μ)	$\mathbf{PAR}^{\mathbf{int}}_{\mu,\epsilon}, \mathbf{PAR}^{\mathbf{ext}}_{\mu,\epsilon}$ (for some μ)
`Equality`	$\Theta(1)$	$\Theta(1)$
`Disjointness`	$\Theta(n)$	$2^{\Theta(n)}$
`Inner Product`	$\Theta(n)$	$2^{\Theta(n)}$
`Greater Than`	$\Theta(\log n)$	$2^{\Theta(\log n)}$

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 [13] where they claim that PRIV is not as robust as PAR.

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: $\mathcal{M}_f[x,y] := f(x,y)$. A *rectangle* of $\mathcal{X} \times \mathcal{Y}$ is a product set $A \times B$ where $A \subseteq \mathcal{X}, B \subseteq \mathcal{Y}$.

We let P denote a two-party communication protocol. Protocols may use both public and private random coins. We let r denote the ensemble of all random coins (public and private) a protocol may use; we let R denote a random variable of all these coins, and R_{pub} denote just the public coins. Given a (possibly randomized) protocol P, for any input $(x,y) \in \mathcal{X} \times \mathcal{Y}$ and random coins r, $P(x,y,r)$ is the value output by Alice and Bob upon running the protocol, and $T_P(x,y,r)$ is the transcript, comprising all messages and public coins. We omit r in the previous if P is deterministic. 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 on each input.

In the following paragraph we let P be a *deterministic* protocol that perfectly computes a function f. 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). The monochromatic P-rectangle of (x,y) is defined as $\mathrm{D}^{T_P}_{x,y} := T_P^{-1}(T_P(x,y))$ (the fact that this is a rectangle and not an arbitrary subset is a well-known consequence of P being a communication protocol). 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, $P^{-1}(z)$. 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. For each

$z \in \mathcal{Z}$, $\mathrm{cut}_P(f^{-1}(z))$ 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 [11].) 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 [1]) is that:

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

We let D_{KL} denote the KL-divergence, $\mathsf{D}_{\mathsf{KL}}(X \parallel Y) = \mathbb{E}_{x \sim X} \log \frac{\Pr[X=x]}{\Pr[Y=x]}$. It is easy to see that $I(X; Y) = \mathsf{D}_{\mathsf{KL}}(XY \parallel X'Y)$ where X' is an independent copy of X. We will also use the following data processing inequality for KL-divergence (we include a proof in the appendix for the sake of completeness):

Lemma 1. *For any X, Y and any deterministic function L, the following holds:*

$$\mathsf{D}_{\mathsf{KL}}(X \parallel Y) \geq \mathsf{D}_{\mathsf{KL}}(L(X) \parallel L(Y)) \tag{2}$$

2.1 Definitions of Privacy

In the following, (X, Y) denotes a pair of random variables, distributed according to μ, and P denotes a (possibly randomized) protocol.

Information Cost: We define the external and internal information cost, notions that have recently found many applications in communication complexity [2,6,7,12]. The external information cost 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 cost measures the amount of information that Alice learns about Bob's input and vice versa.

Definition 1. *The external information cost of P is defined as $\mathrm{IC}^{ext}_{\mu}(P) := \mathbf{I}(X, Y; T_P(X, Y, R))$. The external information cost of f is $\mathrm{IC}^{ext}_{\mu,\epsilon}(f) := \inf_P \mathrm{IC}^{ext}_{\mu}(P)$ where the minimum is over all protocols P computing f with distributional error ϵ.*

Definition 2. *We define the internal information cost of P as $\mathrm{IC}^{int}_{\mu}(P) := \mathbf{I}(X; T_P(X, Y, R)|Y) + \mathbf{I}(Y; T_P(X, Y, R)|X)$. The internal information cost of f is $\mathrm{IC}^{int}_{\mu,\epsilon}(f) := \inf_P \mathrm{IC}^{int}_{\mu}(P)$ where the minimum is over all protocols P computing f with distributional error ϵ.*

Information-theoretic privacy: In [9], 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 [15]. We give a distribution-dependent version of his definition. At a high level, it quantifies how much *more* an observer learns about the inputs from the transcript than from the value of the function. We also define an internal version of the definition. We assume that the output of a randomized protocol depends only on the transcript (i.e. $P(x, y, r)$ is a deterministic function of $T(x, y, r)$).

Definition 3. *The external privacy of P is defined as* $\mathrm{PRIV}^{ext}_{\mu}(f, P) := \mathbf{I}(X, Y; T_P(X, Y, R)) - \mathbf{I}(X, Y; f(X, Y))$. *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 distributional error at most ϵ:* $\mathrm{PRIV}^{ext}_{\mu,\epsilon}(f) := \inf_P \mathrm{PRIV}^{ext}_{\mu}(f, P)$. *We let* $\mathrm{PRIV}^{ext}_{\mu}(f) := \mathrm{PRIV}^{ext}_{\mu,0}(f)$.

Definition 4. *The internal privacy of P is defined as* $\mathrm{PRIV}^{int}_{\mu}(f, P) := \mathbf{I}(X; T_P(X, Y, R)|Y) - \mathbf{I}(X; f(X, Y)|Y) + \mathbf{I}(Y; T_P(X, Y, R)|X) - \mathbf{I}(Y; f(X, Y)|X)$. *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 distributional error at most ϵ:* $\mathrm{PRIV}^{int}_{\mu,\epsilon}(f) := \inf_P \mathrm{PRIV}^{int}_{\mu}(f, P)$. *We let* $\mathrm{PRIV}^{int}_{\mu}(f) := \mathrm{PRIV}^{int}_{\mu,0}(f)$.

It is easy to see that our definition is equivalent to the one in [15] for deterministic or zero-error protocols.

Combinatorial privacy PAR: We present here the definition of PAR for *deterministic* protocols given by [1], which modified the original definition in [13] in order to measure the size of regions relative to the inputs' distribution.

Definition 5. *The external privacy-approximation ratio of a deterministic protocol P for f is defined as:* $\mathrm{PAR}^{ext}_{\mu}(f, P) := \mathbb{E}_{(x,y)\sim\mu}\left[\frac{|\mathrm{D}^f_{x,y}|_\mu}{|\mathrm{D}^{T_P}_{x,y}|_\mu}\right] = \mathbb{E}_{(x,y)\sim\mu}\left[\frac{|\mathrm{D}^P_{x,y}|_\mu}{|\mathrm{D}^{T_P}_{x,y}|_\mu}\right]$ *(where the equality holds because P has zero error). The external privacy-approximation ratio of a function f is defined as:* $\mathrm{PAR}^{ext}_{\mu}(f) := \inf_P \mathrm{PAR}^{ext}_{\mu}(f, P)$ *where the infimum is over all deterministic P computing f with zero error.*

Definition 6. *The internal privacy-approximation ratio of a deterministic protocol P for f is defined as:* $\mathrm{PAR}^{int}_{\mu}(f, P) := \mathbb{E}_{(x,y)\sim\mu}\left[\frac{|\mathrm{D}^f_{x,y}\cap\mathcal{X}\times\{y\}|_\mu}{|\mathrm{D}^{T_P}_{x,y}\cap\mathcal{X}\times\{y\}|_\mu}\right] + \mathbb{E}_{(x,y)\sim\mu}\left[\frac{|\mathrm{D}^f_{x,y}\cap\{x\}\times\mathcal{Y}|_\mu}{|\mathrm{D}^{T_P}_{x,y}\cap\{x\}\times\mathcal{Y}|_\mu}\right]$. *The internal privacy-approximation ratio of a function f is defined as:* $\mathrm{PAR}^{int}_{\mu}(f) := \inf_P \mathrm{PAR}^{int}_{\mu}(f, P)$ *where the infimum is over all deterministic P computing f with zero error.*

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

Theorem 9 ([1]). *For any deterministic protocol P, we have:* $\mathrm{PAR}_\mu^{ext}(f,P) = \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_\mu \cdot \mathrm{cut}_P(f^{-1}(z))$.

This result was stated in [1] but for completeness we present a proof in the appendix.

We now extend the definition to randomized protocols. In the following, the expectations are taken over inputs x, y and random coins r. A simple calculation shows that the following definition coincides with the definition of [1,13] in the case of deterministic zero-error protocols.

Definition 7. *We define:*

- *The external* PAR *of a randomized protocol P as:*

$$\mathrm{PAR}_\mu^{ext}(f,P) := \mathbb{E}_{x,y,r}\left[\frac{\mathbb{P}_{X,Y,R}((X,Y)=(x,y)\mid T_P(X,Y,R)=T_P(x,y,r))}{\mathbb{P}_{X,Y}((X,Y)=(x,y)\mid f(X,Y)=f(x,y))}\right].$$

  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:* $\mathrm{PAR}_{\mu,\epsilon}^{ext}(f) := \inf_P \mathrm{PAR}_\mu^{ext}(f,P)$.
- *The internal* PAR *of a randomized protocol P as:*

$$\mathrm{PAR}_\mu^{int}(f,P) := \mathbb{E}_{x,y,r}\left[\frac{\mathbb{P}_{X,Y,R}(Y=y\mid T_P(X,Y,R)=T_P(x,y,r)\wedge X=x)}{\mathbb{P}_{X,Y}(Y=y\mid f(X,Y)=f(x,y)\wedge X=x)}\right] + \mathbb{E}_{x,y,r}\left[\frac{\mathbb{P}_{X,Y,R}(X=x\mid T_P(X,Y,R)=T_P(x,y,r)\wedge Y=y)}{\mathbb{P}_{X,Y}(X=x\mid f(X,Y)=f(x,y)\wedge Y=y)}\right].$$

  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:* $\mathrm{PAR}_{\mu,\epsilon}^{int}(f) := \inf_P \mathrm{PAR}_\mu^{int}(f,P)$.

Remark 1. There is another way to generalize the definition of PAR for 0-error protocols. This alternative definition is deferred to the appendix.

3 Relations Between Privacy Notions and Communication

We prove a number of relations between the different notions of privacy, communication complexity and information cost both for deterministic and randomized protocols. We summarize them in Fig. 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 (see Theorem 13). Relations between:

- PAR and PRIV are given in Theorem 8 (which was proved in [1] only for the *deterministic 0-error internal* case);
- **D** (resp. $\mathbf{R}^\epsilon$) and PAR is given by Theorem 11;
- IC and PRIV are given in Theorem 12 (which was proved in [1] only for the *deterministic 0-error internal* case);
- The expected distributional complexity and IC (or PRIV) for every possible input distribution is given in Theorem 13;
- $\mathrm{PRIV}^{\mathrm{ext}}$ and $\mathrm{PRIV}^{\mathrm{int}}$ is given in Theorem 14;
- $\mathrm{PAR}^{\mathrm{ext}}$ and $\mathrm{PAR}^{\mathrm{int}}$ comes from Theorem 15 (for the deterministic case).

We start by proving that PRIV provides a lower bound for the log of PAR:

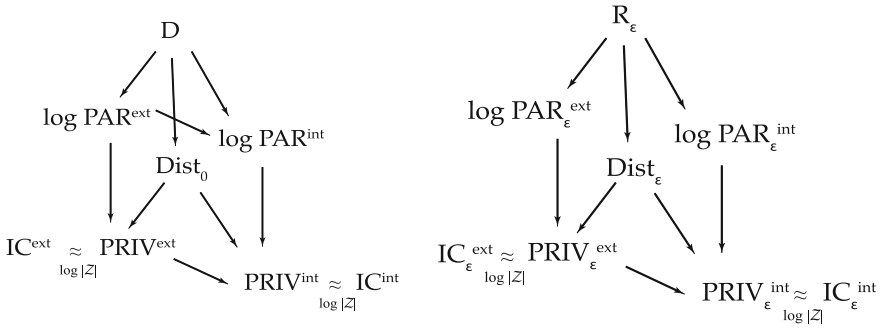

Fig. 1. Lower bounds diagrams for deterministic and bounded error cases

3.1 Relations Between the Different Notions of Privacy and Communication Complexity

We provide below the proof of Theorem 8. The other theorems are proven in the appendix.

Theorem 8 [restated]. For any input distribution μ and any (deterministic or randomized) protocol P, it holds that $\mathrm{PRIV}_{\mu}^{ext}(f,P) \leq \log\left(\mathrm{PAR}_{\mu}^{ext}(f,P)\right)$ *and* $\mathrm{PRIV}_{\mu}^{int}(f,P) \leq 2 \cdot \log\left(\mathrm{PAR}_{\mu}^{int}(f,P)\right) - 2$. 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* $\mathrm{PRIV}_{\mu,\epsilon}^{int}(f) \leq 2 \cdot \log\left(\mathrm{PAR}_{\mu,\epsilon}^{int}(f)\right) - 2$.

Proof. For external privacy, this is a consequence of Bayes rule, and for internal privacy, this is a consequence of Bayes rule and an argument about the worse of the two terms comprising internal PRIV and PAR. The details of this proof will appear in the full version of the article ([17]).

Remark 2. Note that when the protocol is externally (resp. internally) perfectly private, the inequality is tight since $\mathrm{PRIV}_{\mu}^{ext}(f,P) = 0$ and $\mathrm{PAR}_{\mu}^{ext}(f,P) = 1$ (resp. $\mathrm{PRIV}_{\mu}^{int}(f,P) = 0$ and $\mathrm{PAR}_{\mu}^{int}(f,P) = 2$).

3.2 Applications: Tight Bounds on PAR and PRIV for Specific Functions

For boolean functions, PRIV is essentially lower bounded by Information Cost, which subsumes almost all known lower bounds for communication complexity, i.e. smooth rectangle, γ_2-norm, discrepancy, etc [16]. Hence, Theorem 8 implies

Corollary 1 [restated]. For all f, μ, ϵ such that $\mathbf{R}^{\epsilon}(f) = O(\mathrm{IC}_{\mu,\epsilon}^{int}(f))$, it holds that $\log \mathrm{PAR}_{\mu,\epsilon}(f) = \mathrm{PRIV}_{\mu,\epsilon}(f) = \mathbf{R}^{\epsilon}(f)$ up to constant factors (for both internal and external notions).

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 2, 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 [6,7], while for EQ, IP, GT one can look at their discrepancies [10]. Then, using Theorem 12 and 8 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 14). Moreover, we can also get similar lower bounds for the functions Vector in Subspace and Gap-Hamming distance by the results in [16].

4 New Lower Bound Techniques for PAR and PRIV of Deterministic Protocols

In Subsects. 4.1 and 4.2, we assume μ to be full-support for simplicity. By restricting the summations to the rectangles that intersect the support of μ, it is possible to get similar results for a general distribution.

4.1 External PAR of Boolean Functions

Let f be a *boolean* function, P a deterministic protocol for f and T its transcript. Let n_0 and n_1 be the number of P-rectangles with output 0 and 1 ($n_0 = |\mathcal{R}_0^P|, n_1 = |\mathcal{R}_1^P|$). We lower bound PAR by the communication matrix rank.

Theorem 1 [restated]. *For boolean f, for any distribution μ with full support,*

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

The proof will appear in the full version of the article ([17]) and uses Theorem 9 in [1]. Moreover, we are going to use the following result of Yannakakis, which restated in our notation says that

Lemma 2 (Lemma 1 in [21]). *For boolean f and any deterministic protocol P, $\log \min(n_0, n_1) \geq \sqrt{\mathbf{D}(f)}$.*

In fact, Yannakakis proves only that $\log n_1 \geq \sqrt{\mathbf{D}(f)}$, but it is easy to verify that the proof is independent of the value of the monochromatic rectangles, so it similarly follows for the 0-rectangle case. Using in addition the fact that $\mathrm{PAR}_\mu^{ext}(f) \geq \min(n_0, n_1)$, we have

Theorem 2 [restated]. For boolean f, for any distribution μ with full support, $\log \mathrm{PAR}_\mu^{ext}(f) \geq \sqrt{\mathbf{D}(f)}$.

Note that Theorem 1 is not true in general for non-boolean functions (see Appendix).

4.2 External PAR for Non-boolean Functions

Definition 8. *Let* $\widetilde{\mathrm{PAR}}_\mu(f)$ *be the value of the following linear program:*

$$\min_{w_{z,R}} \sum_{z,R} w_{z,R} \cdot \left|f^{-1}(z)\right|_\mu \quad s.t. \quad \forall\,(x,y) \in f^{-1}(\mathcal{Z}) : \sum_{R:R\ni(x,y)} w_{f(x,y),R} = 1 \quad (3)$$

$$\forall\,(x,y) \in f^{-1}(\mathcal{Z}) : \sum_{R:R\ni(x,y)} \sum_{z} w_{z,R} = 1 \quad (4)$$

$$\forall\, z,\ \forall\, R : \quad w_{z,R} \geq 0. \quad (5)$$

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, from conditions (4) and (5), we can interpret $w_{z,R}$ as a probability distribution. In fact, $w_{z,R}$ is the probability to pick R and outputs z on (x,y). This is because condition (3) forces the probability of outputting $f(x,y)$ on (x,y) to be 1.

Theorem 3 [restated]. For all f, for any distribution μ with full support, $\mathrm{PAR}^{ext}_\mu(f) \geq \widetilde{\mathrm{PAR}}_\mu(f)$.

Proof. Let P be a deterministic protocol for f and T its transcript. We can show that $w_{z,R} := \mathbf{1}_{R\in\mathcal{R}^P_z}$ satisfies the conditions of Definition 8 and deduce the lower bound. The details of this proof will appear in the full version of the article ([17]).

Relation with rectangle linear program: We relate this linear program to the rectangle bound defined in [14]. For uniform output distribution, we can generalize this relation to the partition bound (see Appendix).

Definition 9. $\mathrm{rec}^z(f)$ *is the optimal value of the following linear program, where* R *is taken in* $\mathcal{R}(\mathcal{X}\times\mathcal{Y})$*:*

$$\min_{w_R} \sum_{R} w_R \quad s.t. \quad \forall (x,y) \in f^{-1}(z) : \sum_{R:R\ni(x,y)} w_R = 1 \quad (6)$$

$$\forall (x,y) \in \mathcal{X}\times\mathcal{Y} \setminus f^{-1}(z) : \sum_{R:R\ni(x,y)} w_R = 0 \quad (7)$$

$$\forall R : \quad w_R \geq 0. \quad (8)$$

Theorem 4 [restated]. For all f, for any distribution μ with full support, $\widetilde{\mathrm{PAR}}_\mu(f) \geq \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_\mu \cdot \mathrm{rec}^z(f)$.

The proof will appear in the full version of the article ([17]).

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 $f(x_2,y_1) \neq z$. By Theorem 9 in [1] and the following theorem we lower bound PAR by fooling sets.

Theorem 10 ([18]). *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.*

Theorem 5 [restated]. For all f and any set of z-fooling sets $\{F_z\}_{z\in\mathcal{Z}}$, for any distribution μ with full support, $\mathrm{PAR}^{ext}_{\mu}(f) \geq \sum_{z\in\mathcal{Z}} \left|f^{-1}(z)\right|_{\mu} \cdot |F_z|$.

4.3 New Lower Bound Techniques for External IC

We show lower bounds on the external information complexity, which using Theorem 12 will in turn give new lower bounds on information-theoretic privacy. Our lower bounds hold for zero-error randomized protocols, which of course imply the same bounds for deterministic protocols.

Theorem 6 [restated]. Fix a function f. Suppose there exists $\delta > 0$ and a distribution μ over the inputs of f, such that for all monochromatic rectangles R of f, $\mu(R) \leq \delta$. Then it holds for every P that computes f without error on any input (i.e. even on pairs of inputs lying outside μ's support) that $\mathrm{IC}^{ext}_{\mu}(P) \geq \log(1/\delta)$.

The proof will appear in the full version of the article ([17]).

Corollary 2. *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 $\mathrm{IC}^{ext}_{\mu}(P) \geq \log k$.*

The proof of this corollary will appear in the full version of the article ([17]). Note that Theorem 6 can be used to prove an optimal lower bound on the zero-error information complexity of certain functions. For example, for one bit AND, the hard distribution μ is uniform over (0, 1), (1, 0), (1, 1), and our theorem implies that $\mathrm{IC}^{ext}_{\mu}(P) \geq \log_2 3$. This matches a recent exact bound (which is in particular an upper bound) by Braverman *et al.* [4].

4.4 Applications: Tight Bounds on External PAR and PRIV for Specific Functions

Our applications 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.

The proofs will appear in the full version of the article ([17]).

5 Quality of the Two Definitions

5.1 Privacy for Deterministic Protocols

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 heavily for high-privacy-loss events, even if they occur with small probability.

In the appendix, 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 [13] and show that for the example they gave, in fact PRIV is just as good as PAR at distinguishing two protocols in their example.

Bounded-error case: As we explained in Sect. 3.2, 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.

Acknowledgements. We would like to thank Salil Vadhan for useful comments regarding our definition for bounded error and for observing that the original proof of Theorem 6 could be greatly simplified. We would also like to thank Omri Weinstein and Lila Fontes for useful discussions.

This work was partially supported by the ANR Blanc project ANR-12-BS02-005 (RDAM) and ANR Jeune Chercheur project CRYQ, ANR Blanc project QRAC (ANR-08-EMER-012), and EU ANR Chist-ERA project DIQIP.

A Appendix

A.1 Complements to Sect. 2

Omitted Proofs. The proofs of Lemma 1 and Theorem 9 will appear in the full version of the article [17].

Discussion About the Definition of PAR. In Sect. 2, definition 7, we have defined PAR for randomized bounded-error protocols relatively to the transcript and the output value of the function. This definition is consistent with the one for deterministic protocols. However it is also possible to extend the definition of PAR by taking the output of the protocol instead of the output of the function:

Definition 10. – *An alternative definition for the external* PAR *of a randomized protocol P is:* $\mathrm{PAR}_{\mu}^{ext,alt}(P) := \mathbb{E}_{x,y,r}\left[\frac{\mathbb{P}_{X,Y,R}((X,Y)=(x,y)\,|\,T_P(X,Y,R)=T_P(x,y,r))}{\mathbb{P}_{X,Y}((X,Y)=(x,y)\,|\,P(X,Y)=P(x,y))}\right]$. *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:* $\mathrm{PAR}_{\mu,\epsilon}^{ext,alt}(f) := \inf_P \mathrm{PAR}_{\mu}^{ext,alt}(P)$.

– *An alternative definition for the internal* PAR *of a randomized protocol P is:*

$$\mathrm{PAR}_{\mu}^{int,alt}(P) := \mathbb{E}_{x,y,r}\left[\frac{\mathbb{P}_{X,Y,R}(Y=y \mid T_P(X,Y,R)=T_P(x,y,r)\wedge X=x)}{\mathbb{P}_{X,Y}(Y=y \mid P(X,Y)=P(x,y)\wedge X=x)}\right] + \mathbb{E}_{x,y,r}\left[\frac{\mathbb{P}_{X,Y,R}(X=x \mid T_P(X,Y,R)=T_P(x,y,r)\wedge Y=y)}{\mathbb{P}_{X,Y}(X=x \mid P(X,Y)=P(x,y)\wedge Y=y)}\right].$$

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:* $\mathrm{PAR}_{\mu,\epsilon}^{int,alt}(f) := \inf_P \mathrm{PAR}_{\mu}^{int,alt}(P)$.

A.2 Omitted Roofs from Sect. 3

We have proven Theorem 8 in Sect. 3. We prove here the other theorems stated in this section.

Relations Between the Different Notions of Privacy and Communication Complexity. Firstly we show that for any protocol (deterministic or randomized), the external privacy-approximation ratio is at most exponential in the communication of the protocol.

Theorem 11. *For any protocol P,* $\mathrm{PAR}_{\mu}^{ext}(f,P) \leq 2^{\mathbf{CC}(P)}$.

The proof will appear in the full version of the article [17].

The relation between internal IC and internal PRIV for deterministic protocols was explained in [1]. It is possible to improve the lower bound and to show the same relationship for external notions and any (deterministic or randomized) protocol.

Theorem 12. *For any protocol P and any distribution μ,*

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

Proof. By definition of IC and PRIV we have, respectively for the external and the internal notions:

$$\mathrm{IC}_{\mu}^{int}(P) - \mathrm{PRIV}_{\mu}^{int}(f,P) = \mathbf{I}(X; f(X,Y)|Y) + \mathbf{I}(Y; f(X,Y)|X) \leq 2\log(|\mathcal{Z}|),$$
$$\mathrm{IC}_{\mu}^{ext}(P) - \mathrm{PRIV}_{\mu}^{ext}(f,P) = \mathbf{I}(X,Y; f(X,Y)) \leq \log(|\mathcal{Z}|).$$

For the lower bounds, note that mutual information is always positive.

Moreover, 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 13 ([11]). *For any randomized ϵ-error protocol and any input distribution,* $\mathrm{Dist}_{\mu,\epsilon}(P) \geq \mathrm{IC}_{\mu,\epsilon}^{ext}(P)$.

The proof of this well-known fact can be found in [11] for example.

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

Relation Between Internal and External Privacy. We first study the case of PRIV and then focus on PAR.

Theorem 14. $\mathrm{PRIV}^{int}_{\mu}(f,P) \leq \mathrm{PRIV}^{ext}_{\mu}(f,P) + \log(|\mathcal{Z}|)$.

Proof. Braverman [7] proved that: $\mathrm{IC}^{int}_{\mu}(P) \leq \mathrm{IC}^{ext}_{\mu}(P)$. Hence, with 12:

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

Moreover, we show that internal PAR is smaller than external one for deterministic protocols:

Theorem 15. *For any deterministic protocol P computing f:*

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

The proof will appear in the full version of the article [17].

However, Theorem 15 does not hold in general for ϵ-error randomized protocols. For instance, consider that Alice receives an s-bit string x, and Bob receives x plus an n-bit string y, such that x and y are independent, and they want to compute the function that reveals x: $f(x,y) = x$. The protocol they use, where only Bob sends messages, is the following: if $x = 0^s$ then Bob sends y, otherwise he sends a random n-bit string (independent of x and y). Then:

$$\begin{aligned}
\mathrm{PAR}^{int}_{\mu}(f,P) &= \mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}(XY = xy | T = t, X = x)}{\mathbb{P}(XY = xy | X = x)}\right] + 1 \\
&= \sum_{x,y,t} \mathbb{P}(X = x, Y = y, T = t) \frac{\mathbb{P}(Y = y | T = t, X = x)}{\mathbb{P}(Y = y | X = x)} + 1 \\
&= 2^n \sum_{x,y,t} \mathbb{P}(X = x, Y = y, T = t) \mathbb{P}(Y = y | X = x, T = t) + 1 \\
&= 2^n \left(\sum_{x \neq 0, y, t} \frac{1}{2^{2n+s}} \frac{1}{2^n} + \sum_{x=0, y=t} \frac{1}{2^{n+s}} \cdot 1 \right) + 1 = 2^{n-s} + o(1)
\end{aligned}$$

and:

$$\begin{aligned}
\mathrm{PAR}^{ext}_{\mu}(f,P) &= \mathbb{E}_{x,y,t}\left[\frac{\mathbb{P}(X = x, XY = xy | T = t)}{\mathbb{P}(X = x, XY = xy | f(X,Y) = f(x,y))}\right] \\
&= \sum_{x,y,t} \mathbb{P}(X = x, Y = y, T = t) \frac{\mathbb{P}(X = x, Y = y | T = t)}{\mathbb{P}(Y = y)} \\
&\qquad\qquad (\text{since} f(x,y) = x) \\
&= 2^n \sum_{x,y,t} \mathbb{P}(X = x, Y = y, T = t) \mathbb{P}(X = x, Y = y | T = t) \\
&= 2^n \left(\sum_{x \neq 0, y, t} \frac{1}{2^{2n+s}} \frac{1}{2^{n+s}} + \sum_{x=0, y=t} \frac{1}{2^{n+s}} \frac{1}{2^s} \right) = 2^n + o(1)
\end{aligned}$$

Hence, if x is of length $s = n/2$, then $\mathrm{PAR}^{int}_{\mu}(f,P) = 2^{n/2} + o(1)$ is exponentially bigger than $\mathrm{PAR}^{ext}_{\mu}(f,P) = o(1)$.

A.3 Omitted Proofs for Sect. 4

Relation with Partition Linear Program. It is also possible to lower bound $\mathrm{PAR}_{\mu}^{ext}(f)$ by $\frac{1}{|\mathcal{Z}|} \cdot \mathrm{prt}(f)$, where $\mathrm{prt}(f)$ is defined in [14]. The details of this fact wille appear in the full version of the article [17].

Rank Argument Fails 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:

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

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 one can see that $\mathrm{PAR}_{\mu}^{ext}(\mathrm{EQ}') \leq 3$. Hence for this function EQ' and this distribution μ: $\mathrm{PAR}_{\mu}^{ext}(\mathrm{EQ}', P) \leq 3$ whereas : $\mathrm{rank}\left(\mathcal{M}_{\mathrm{EQ}'}\right) \geq \mathrm{rank}\left(\mathcal{M}_{\mathrm{EQ}_{n-1}}\right) = 2^{n-1}$.

Proofs of Applications. An advantage of our techniques is that they give bounds for *any* distribution of input μ, and not only for a uniform distribution as in [13]. 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. The lower bounds stated in Table 1 can be proved using Theorem 1.

Now we explain briefly how to obtain the results of Theorem 7 (see the full version of the article ([17]) for the details). For the lower bounds for EQ, DISJ, GT, we can apply Corollary 2 using an appropriate fooling set, followed by the relationship between IC and PRIV given in Theorem 12. For IP it is possible to use the well-known fact that all 0-monochromatic rectangles of the IP function contain at most 2^n elements.

A.4 Privacy for Deterministic Protocols

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 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:

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

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:

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

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

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

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 16. *For any protocol P and any two input distributions μ, μ' with statistical distance $|\mu - \mu'| \leq \epsilon$, it holds that :* $|\mathrm{PRIV}^{ext}_{\mu}(P) - \mathrm{PRIV}^{ext}_{\mu'}(P)| \leq O(\epsilon n)$ *and* $|\mathrm{PRIV}^{int}_{\mu}(P) - \mathrm{PRIV}^{int}_{\mu'}(P)| \leq O(\epsilon n)$.

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 3 (Lemma 3.15 of [19]). *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*

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

The details of this proof will appear in the full version of the article [17].

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 3. 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 1. 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 [1,13], although they did not study worst-case random coins since they focused on deterministic protocols.

Error in Appendix of [13]. An example was given in the appendix of [13] 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 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 [9] 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.

References

1. Ada, A., Chattopadhyay, A., Cook, S., Fontes, L., Koucký, M., Pitassi, T.: The Hardness of Being Private. In: 27th Annual IEEE Conference on Computational Complexity, CCC'12, pp. 192–202 (2012)
2. Barak, B., Braverman, M., Chen, X., Rao, A.: How to compress interactive communication. In: Proceedings of the 42nd STOC, pp. 67–76 (2010)
3. Brody, J., Buhrman, H., Koucky, M., Loff, B., Speelman, F., Vereshchagin, N.: Towards a reverse Newman's theorem in interactive information complexity, CCC (2013)
4. Braverman, M., Garg, A., Pankratov, D., Weinstein, O.: From information to exact communication, In: STOC, pp. 151–160 (2013)
5. Braverman, M., Garg, A., Pankratov, D., Weinstein, O.: Information lower bounds via self-reducibility. In: Bulatov, A.A., Shur, A.M. (eds.) CSR 2013. LNCS, vol. 7913, pp. 183–194. Springer, Heidelberg (2013)
6. Bar-Yossef, Z., Jayram, T., Kumar, R., Sivakumar, D.: An information statistics approach to data stream and communication complexity. In: Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, pp. 209–218 (2002)
7. Braverman, M.: Interactive information complexity. ECCC, report No. 123, STOC'12 (2011)
8. Braverman, M., Moitra, A.: An information complexity approach to extended formulations. In: STOC'13 (2013)
9. Bar-Yehuda, R., Chor, B., Kushilevitz, E., Orlitsky, A.: Privacy, additional information and communication. IEEE Trans. Inf. Theory **39**(6), 1930–1943 (1993)

10. Braverman, M., Weinstein, O.: A discrepancy lower bound for information complexity. In: Proceedings of the APPROX-RANDOM 2012, pp. 459–470 (2012)
11. Cover, T.M., Thomas, J.A.: Elements of Information Theory, 2nd, Hardcover, New York, pp. 776 2006 ISBN: 0-471-24195-4
12. Chakrabarti, A., Shi, Y., Wirth, A., Yao, A.: Informational complexity and the direct sum problem for simultaneous message complexity. In: 42nd IEEE FOCS, pp. 270–278 (2001)
13. Feigenbaum, J., Jaggard, A.D., Schapira, M.: Approximate privacy: foundations and quantification. In: Proceedings of the 11th Conference on Electronic Commerce (EC)., ACM Press, New York, pp. 167–178 (2010)
14. Jain, R., Klauck, H.: The partition bound for classical communication complexity and query complexity. In: 25th IEEE Conference on Computational Complexity (2010)
15. Klauck, H.: On quantum and approximate privacy. In: Proceedings STACS (2002)
16. Kerenidis, I., Laplante, S., Lerays, V., Roland, J., Xiao, D.: Lower bounds on information complexity via zero-communication protocols and applications. FOCS **2012**, 500–509 (2012)
17. Kerenidis, I., Laurière, M., Xiao, D.: New lower bounds for privacy in communication protocols, http://eccc.hpi-web.de/report/2013/015/ (full version, 2013)
18. Kushilevitz, E., Nisan, N.: Communication Complexity. Cambridge University Press, New York (1997)
19. Mahmoody, M., Xiao, D.: Languages with efficient zero Knowledge PCPs are in SZK. ECCC technical report TR2012-052 (2012)
20. Jain, R.: New strong direct product results in communication complexity. J. ACM (2013)
21. Yannakakis, M.: Expressing combinatorial optimization problems by linear programs. J. Comput. Syst. Sci. **43**, 441–466 (1991)
22. Yao, A.C-C.: Some complexity questions related to distributive computing. In: Proceedings of the 11th ACM Symposium on Theory of Computing (STOC), pp. 209–213 (1979)