Quantum Information and Computation, Vol. 16, No. 3&4 (2016) 0181–0196


# INFORMATION COST OF QUANTUM COMMUNICATION PROTOCOLS

IORDANIS KERENIDIS
*LIAFA, CNRS, Université Paris Diderot*
*8 place Aurélie Nemours, Paris, 75013, France.*

MATHIEU LAURIÈRE
*LIAFA, Université Paris Diderot*
*8 place Aurélie Nemours, Paris, 75013, France.*

FRANÇOIS LE GALL
*Department of Computer Science, The University of Tokyo,*
*7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan.*

MATHYS RENNELA
*Institute for Computing and Information Sciences, Radboud University,*
*Toernooiveld 212, Nijmegen, 6525 EC, The Netherlands.*

Received July 3, 2015
Revised November 19, 2015

In two-party quantum communication complexity, Alice and Bob receive some classical inputs and wish to compute some function that depends on both these inputs, while minimizing the communication. This model has found numerous applications in many areas of computer science. One notion that has received a lot of attention recently is the information cost of the protocol, namely how much information the players reveal about their inputs when they run the protocol. In the quantum world, it is not straightforward to define a notion of quantum information cost. We study two different notions and analyze their relation. We also provide new quantum protocols for the Inner Product function and for Private Information Retrieval, and show that protocols for Private Information Retrieval whose classical or quantum information cost for the user is zero can have exponentially different information cost for the server.

*Keywords*:
*Communicated by*: R Cleve & R de Wolf

## 1 Introduction

In two-party communication complexity [1], Alice and Bob receive inputs and wish to compute some function that depends on both these inputs, while minimizing the communication. This model has found numerous applications in many areas of computer science. One question that has received a lot of attention recently is whether it is possible to perform such protocols without leaking much information.

In classical communication protocols, the information cost (or privacy loss) is defined as the information that the transcript reveals to each player about the input of the other one. In this model, one is interested in the information cost of a specific protocol and hence we

only consider the case where the players honestly follow the protocol and not how they can increase the information by deviating from it.

In quantum communication protocols [2], Alice and Bob receive classical inputs $x$ and $y$ and wish to compute $f(x, y)$ but they are allowed to use quantum communication and quantum workspaces. We consider three quantum registers $A, M, B$ that correspond to Alice's workspace, the message qubits, and Bob's workspace. At each round of the protocol, one player applies a unitary operation that depends on her input on her workspace and the message qubits, and sends the message qubits to the other player, who continues the protocol. We always assume that Alice starts the protocol. Since the message qubits can be reused throughout the protocol and copying of the quantum states may be impossible, there is no clear notion of a transcript. Hence, we know of no way to define notions of quantum information cost, other than by a *round-by-round* definition.

We can see, by a chain rule argument, that the classical definition of information cost is equivalent to a *round-by-round* definition where for each round, we calculate the information that the current message reveals about the sender's input to the receiver, who knows her input and has kept a copy of all previous messages in her workspace.

Again, this definition is not readily applicable to quantum protocols, since the players may not be able to copy the messages and continue the protocol at the same time. Nevertheless, they have a quantum workspace, where, depending on the protocol, they may keep information about previous messages. We would like to calculate how much information every new message reveals to them, given that they already know their own input and have kept some information in their quantum workspace according to the protocol.

There is one additional important issue to consider. Each player has a register where the input is written in the beginning of the protocol. In the setting we consider here, this input is always a classical one. One natural possibility is therefore to consider that the input register is a classical register, meaning it stays unentangled with the workspaces and the message space. The second possibility is to consider that the input is written in a quantum register, which could be entangled with the players' workspaces or even with the environment. We discuss below in more details these two possibilities (formal definitions and more complete discussion are given in Section 3).

*Information cost with classical input registers.* This is in fact the definition that has been mostly used in the past [3, 4, 5]. In high level, it is defined in the following way: Alice and Bob in the beginning of the protocol receive the classical inputs $x, y$ according to some input distribution in two classical registers $X$ and $Y$. Then, they take turns applying some unitaries on the quantum registers $A, M, B$ controlled by the classical strings $x, y$. At each round, we measure how much information is revealed by the message $M$ about the sender's input register, given the receiver's working space. This definition has also been used as a measure of privacy of quantum protocols. We will denote Alice's and Bob's classical input information cost for a protocol $\pi$ by $CIC_A(\pi)$ and $CIC_B(\pi)$, respectively. Note that we split the information cost into two separate informaiton costs, one for each player.

*Information cost with quantum input registers.* A new definition of quantum information cost for quantum protocols, denoted in this paper by $QIC_A(\pi)$ and $QIC_B(\pi)$, was proposed by Touchette [6]. This definition has nice properties, e.g., it is equal to the amortized quantum communication complexity [6]. In this case, the input registers are initially purified through

an external register (an environment, not accessible to the players). Then, the information that is measured is with respect to this environment register and not the input registers of the players.

*Information cost as privacy.* The information cost of a protocol measures the amount of information about the inputs that is leaked during the protocol and hence can be seen as a measure of privacy loss of the protocol. Note that we are not in a cryptographic scenario with cheating players, rather we want to compute the privacy loss of the specific protocol that computes the function (the players cannot deviate from the prescribed protocol). In the case of quantum protocols, one could possibly define different notions of privacy for the different notions of information cost. We argue that the most relevant definition of privacy is the information cost with classical input registers, since it is the only one that measures the privacy loss of the prescribed protocol. This notion of privacy with classical input registers has previously been discussed for some classes of quantum protocols, for example in [4, 5].

Moreover, with the definition of $QIC$ we do not measure the information revealed about each player's input but about the environment register $R$. Nevertheless, we will see that the quantum information cost is a stronger technique for providing lower bounds on communication complexity.

**Our results** The main goal of this paper is to investigate quantum communication protocols under these two variants of information cost, and in particular study the differences between information cost with classical input registers and information cost with quantum input registers. In the present work, we first prove the following inequalities between these definitions.

**Result 1**: $CIC_A(\pi) \leq QIC_A(\pi)$ and $CIC_B(\pi) \leq QIC_B(\pi)$ for any protocol $\pi$.

The details are given in Theorem 1 of Section 3. We then show that for some protocols, the quantum information cost for one player can be arbitrarily higher than his classical information cost, while the two notions of information cost for the other player are equal. This is done by considering the Inner Product function. In addition, for Private Information Retrieval we show that protocols whose classical or quantum information cost for the user is zero can have exponentially different information cost for the server. In order to obtain these gaps, we construct new quantum protocols for these tasks and analyze their informaiton cost.

We describe below in more details our results for the Inner Product function and for Private Information Retrieval.

*The information cost of Inner Product.* We provide a protocol for Inner Product that shows a gap between the two notions of information cost (see Theorem 2 in Section 4).

**Result 2**: There exists a quantum protocol for Inner Product over $n$-bit strings, which is perfectly private for Bob and where Alice's classical input information cost is only $n/2 + 1/2$.

We also show that for the protocol we contruct the quantum information cost is basically $n/2$ for both parties, hence providing a gap between these notions.

*The information cost of Private Information Retrieval.* Private Information Retrieval has been extensively studied so as to find the minimum communication necessary between the user and one or more servers, while keeping the perfect privacy of the user. Here we consider

the one-server setting: the server has for input a database $x = x_1 \cdots x_n \in \{0,1\}^n$, the user has for input an index $i \in \{1, \ldots, n\}$, and the goal is for the user to output $x_i$. It is well known that any classical protocol perfectly private for the user (i.e., in which the server obtains no information about the index $i$) requires $\Omega(n)$ bits of communication [7]. Moreover, the quantum communication complexity, as well as the quantum information costs are also $\Omega(n)$ [8].

Recently, Le Gall [5] showed that there exists a quantum protocol for this task, perfectly private for the user (according to Definition 1 in Section 3), with communication complexity $O(\sqrt{n})$. This upper bound has then been improved to $O(n^{1/3})$ by Ruben Brokkelkamp [9]. Here we ask the question: Can these upper bounds be further improved? Or, more specifically, how much information does a single server have to leak about the database in any protocol which is perfectly private for the user (with regard to the classical input information cost)? We show the following surprising result (see Corollary 2 in Section 5).

**Result 3**: There exists a quantum protocol for Private Information Retrieval whose classical input information cost is zero for the user and polylogarithmic on the size of the database for the server.

The protocol is constructed explicitly and its communication cost is also polylogarithmic in $n$, which is an exponential improvement over the prior works [5, 9] mentioned above. Moreover, since any scheme whose quantum information cost is zero for the user must have quantum information cost linear on the size of the database for the server, our result provides the first exponential separation between the two notions of information cost of quantum protocols (namely, classical input information cost versus quantum information cost).

The proof has two steps: first, we show how to take any $\ell$-server classical PIR scheme and translate it into a quantum one-server scheme, such that the index remains perfectly private. Then, we use a classical PIR scheme with a logarithmic number of servers and polylogarithmic communication [7], which implies that the information cost of the database is polylogarithmic, since it is always less than the communication.

Finally, we improve the above upper bounds when the user and the server share prior entanglement (see Theorem 4 in Section 6):

**Result 4**: There exists a quantum protocol for Private Information Retrieval using shared prior entanglement whose classical input information cost is zero for the user and logarithmic on the size of the database for the server.

The protocol is again constructed explicitly and its communication cost is $O(\log n)$, which is optimal since, even with prior entanglement, the quantum communication complexity of the Index Function is $\Omega(\log n)$.

## 2 Preliminaries

In this paper we write, for a positive integer $p$, $[p] := \{1, 2, \ldots, p\}$ and, for two positive integer $p < q$, write $[p, q] := \{p, p+1, \ldots, q\}$.

In two-party communication complexity, Alice and Bob receive inputs $x$ and $y$ respectively and wish to compute some function $f(x, y)$ that depends on both these inputs, while minimizing the *communication cost*, i.e., the number of exchanged bits. The *communication*

*complexity* of a function is the least amount of communication possible in a protocol computing $f$. We refer to [10] for details about classical communication complexity, and to [11] for asymmetric communication complexity.

In two-party quantum communication complexity, the players are now allowed to exchange quantum bits. The standard model consists of three quantum registers: A, M and B. Here A and B are private workspaces of Alice and Bob respectively, while M is used to communicate qubits and is sent from one player to the other one. Additionally, Alice and Bob hold a register (classical or quantum), say $X$ and $Y$ respectively, where they store their respective input. At every round, one player applies a unitary operation on their workspace and the message qubits (that also depends on their input) and sends the message qubits to the other player who continues the protocol. Note that in general protocols, the unitary operation could modify the player's input register. However in this paper we will be interested only in protocols where the players never modify their input registers. In other words, we consider that the players always keep a safe copy of their inputs, which seems reasonable since they want to compute a certain function of these inputs.

In the above setting, the Inner Product (IP) problem consists in computing $f(x, y) = x \cdot y := \sum_i x_i \cdot y_i$. In [12], it is proved as a particular case of the bounded error setting, that computing classically and perfectly IP requires a communication of $n$, and the same holds for quantum protocols [13].

Another well studied problem is Private Information Retrieval (PIR): a user, whose input is an index $i \in [n]$, interacts with a server holding a database $x = (x_j)_{j\in[n]} \in \{0,1\}^n$. The goal for the user is to learn $x_i$ in such a way that the server does not learn his index $i$. In [14, 7], it is shown that the communication complexity of this problem is $\Omega(n)$. The same paper also shows that it is possible to improve the communication complexity if the user can interact with several independent servers: in this setting it is possible to obtain a communication polylogarithmic in $n$. In [5], the author gives a quantum protocol using a single server and only $O(\sqrt{n})$ qubits of communication, which yields a quadratic improvement over what is possible classically. This upper bound has then been improved to $O(n^{1/3})$ by Ruben Brokkelkamp [9]. In both cases the protocol is perfectly private for the user (as long as the server follows exactly the prescribed scheme). If we allow the players to create superpositions of inputs or act as specious adversaries, then it is known that the communication from the server must be linear [8, 15].

The information cost will be analyzed with information theoretical tools. More precisely, $S(X)$ will denote the entropy of $X$, that is $S(X)$ is equal to $-\sum_x p_x \log(p_x)$ if $X$ is a classical random variable taking value $x$ with probability $p_x$, or to $-\operatorname{Tr}(\rho_X \log(\rho_X))$ if $X$ is a quantum register whose state is denoted by $\rho_X$. If $A, B$ and $C$ are either classical random variables or quantum registers, the mutual information between $A$ and $B$ (resp. the mutual information between $A$ and $B$ conditioned on $C$) is defined by $I(A : B) = S(A) + S(B) - S(AB)$ (resp. $I(A : B|C) = I(AC : B) - I(C : B) = S(AC) + S(BC) - S(C) - S(ABC)$, which can be interpreted as the knowledge that $A$ gives about $B$ provided that we already knew $C$).

## 3 Definitions of Information Cost for Quantum Protocols and their Relation

In classical communication protocols, the information cost of a protocol is defined as the information that the transcript $\Pi$ of the communication reveals to each player about the

input of the other one. Using a chain rule argument, it is not hard to see that the classical definition of information cost is equivalent to a *round-by-round* definition where for every round $k$, we calculate the information that the message at round $k$ reveals about each player's input to the other player, who already knows his input and has kept a copy of all previous messages in his workspace:

$$I(\Pi : X|Y) = \sum_{k:\text{ odd}} I(M_k : X|Y, M_1, \ldots, M_{k-1}), \tag{1}$$

$$I(\Pi : Y|X) = \sum_{k:\text{ even}} I(M_k : Y|X, M_1, \ldots, M_{k-1}). \tag{2}$$

Note that we define the information cost of each player separately, since their input sizes or their privacy considerations can be different. For example, for Private Information Retrieval, we will look at protocols which are perfectly secure for the user (whose input has size $\log n$) and leak a logarithmic amount of information about the database (whose size is $n$).

In quantum communication protocols, since there is no notion of transcript, we define notions of privacy or quantum information cost by a *round-by-round* definition. As we said, we will also differentiate between the case where the input registers contain classical or quantum inputs.

We will try to provide a unified view of the two definitions of information cost. Let us denote the input distribution by $\mu$. We start by assuming that a third party creates the following state

$$\sum_{x,y} \sqrt{\mu(x,y)}|x\rangle_{X_R}|y\rangle_{Y_R}|x\rangle_X|y\rangle_Y. \tag{3}$$

He keeps the registers $X_R Y_R = R$ to himself and sends the registers $X$ and $Y$ to Alice and Bob respectively. Hence, Alice and Bob receive in their registers classical inputs $x, y$ with probability $\mu(x, y)$, while $R$ holds a purification of the inputs. Then, they start the protocol and using the inputs as control bits, they apply the prescribed unitaries, so that at round $k$ the joint state can be written as

$$\sum_{x,y} \sqrt{\mu(x,y)}|x\rangle_{X_R}|y\rangle_{Y_R}|x\rangle_X|y\rangle_Y|\phi^k_{xy}\rangle_{AMB}. \tag{4}$$

Using this state, we can now define the two notions of information cost.

### *3.1 Classical input information cost*

In this definition the register $R$ does not appear and hence without loss of generality, we can assume that it has been measured. This is why the registers $X$ and $Y$ become now classical registers. We can now provide the definition of classical input information cost

**Definition 1** *For a protocol $\pi$, the Classical input Information Cost of Alice and Bob are defined as*

$$CIC_A(\pi) = \sum_{k\,:\,odd} I(M_k : X|Y, B_k) \text{ and } CIC_B(\pi) = \sum_{k\,:\,even} I(M_k : Y|X, A_k),$$

*where according to equation* (4)*, $X, Y$ are the input registers, and $M_k, A_k, B_k$ are quantum registers that correspond to the message and Alice's and Bob's workspaces at round $k$.*

**Remark 1** *Note that in the model of quantum communication, where only unitary operations are allowed during the protocol, there is no obvious equivalent of a classical protocol $\pi$ when $\pi$ uses random coins (as is usual in a classical communication protocols). It is possible to define a quantum communication protocol $\pi'$ where the players create and keep in their workspace a uniform superposition of coins, use them as control bits to simulate $\pi$ in superposition, and measure them at the end of the protocol $\pi'$. As far as communication is concerned, this construction is fine, since the communication of $\pi'$ is the same as the one of $\pi$. However, when dealing with information, things are more subtle. In particular the CIC of $\pi'$ could be larger than the (classical) information cost of $\pi$. This is because using coins in superposition might reveal more information than just using classical coins. We will not discuss further this issue here, since in the present work we focus on the difference between the two notions of informaiton cost for quantum protocols and not how they relate to the informaiton cost of classical protocols.*

### *3.2 Quantum information cost*

Recently, another definition of quantum information cost for protocols with entanglement was proposed by Touchette [6]. This definition has very nice properties: for example, it is equal to the amortized quantum communication complexity. In this case, the information cost is measured with respect to the register $R$ and not the input registers $X, Y$.

**Definition 2** *For a protocol $\pi$, the Quantum Information Cost of Alice and Bob are defined as*

$$QIC_A(\pi) = \sum_{k\,:\,odd} I(M_k : R|Y, B_k) \text{ and } QIC_B(\pi) = \sum_{k\,:\,even} I(M_k : R|X, A_k), \tag{5}$$

*where according to Equation 4, $X, Y$ are the input registers, $R = (X_R, Y_R)$ holds their purifications and $M_k, A_k, B_k$ are quantum registers that correspond to the message qubits and Alice's and Bob's workspaces at round $k$.*

### *3.3 Relation between the two definitions of information cost*

We have seen two different definitions which measure in some way the information transmitted during the protocol. We now prove a general inequality between these notions.

**Theorem 1** *For any protocol $\pi$ we have*

$$CIC_A(\pi) \leq QIC_A(\pi) \text{ and } CIC_B(\pi) \leq QIC_B(\pi).$$

**Proof :** We have that at round $k$,

$$I(M_k : X|YB_k) = I(M_k : X_R|YB_k) = I(M_k : X_RY_R|YB_k) - I(M_k : Y_R|X_RYB_k) \tag{6}$$
$$\leq I(M_k : X_RY_R|YB_k). \tag{7}$$

Summing over odd $k$ we obtain $CIC_A(\pi) \leq QIC_A(\pi)$. Similar for Bob. □

## 4 Information Cost for Inner Product

In this section we describe a quantum protocol for Inner Product and compute all different information cost quantities for it. Alice and Bob have input $x \in \{0,1\}^n, y \in \{0,1\}^n$ respectively and want to compute $x \cdot y$.

The protocol is given in Fig. 1. Here we assume that only Alice needs to learn the value of the function (then she could communicate it to Bob, leaking at most one bit of information about her input).

**Protocol** $\Pi_{IP}$ :

1. Alice creates and sends to Bob $|\phi^1_{xy}\rangle := \frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\rangle_Q |r\cdot x\rangle_T$
2. Bob applies the unitary $V_y : |r\rangle \mapsto |r\oplus y\rangle$ to register $Q$ and sends back to Alice the state $|\phi^2_{xy}\rangle := \frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\oplus y\rangle_Q |r\cdot x\rangle_T$

Fig. 1. Quantum protocol for inner product.

Let us prove the correctness of the protocol. Observe that

$$|\phi^2_{xy}\rangle = \frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\oplus y\rangle_Q |r\cdot x\rangle_T = \frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\rangle_Q |(r\oplus y)\cdot x\rangle_T. \tag{8}$$

At the end of the protocol, Alice, by applying the unitary $U_x : |r\rangle|b\rangle \mapsto |r\rangle|b\oplus r\cdot x\rangle$, can transform $|\phi^2_{xy}\rangle$ to the state

$$|\phi^3_{xy}\rangle := \frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\rangle_Q |((r\oplus y)\cdot x)\oplus(r\cdot x)\rangle_T = \left(\frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\rangle_Q\right) |x\cdot y\rangle_T. \tag{9}$$

By measuring Register $T$, Alice obtains the bit $x\cdot y$.

Assuming the inputs are distributed uniformly, we then evaluate the information cost of this protocol.

**Theorem 2** *For the above protocol* $\Pi_{IP}$ *under uniform distribution of inputs, we have (up to exponentially small terms)*

$$CIC_A(\Pi_{IP}) = n/2 + 1/2 \quad , \quad CIC_B(\Pi_{IP}) = 1.$$
$$QIC_A(\Pi_{IP}) = n/2 + 1/2 \quad , \quad QIC_B(\Pi_{IP}) = n/2 + 3/2.$$

The proof of Theorem 2 follows from the following claims.

**Claim 1** *Bob gets* $n/2+1/2$ *bits of information (up to exponentially small terms) from the first message and Alice one bit from the second message. More precisely:* $CIC_A(\Pi_{IP}) = n/2+1/2$ *and* $CIC_B(\Pi_{IP}) = 1$.

**Proof :** After receiving the first message, the information that Bob has about Alice's input is, by definition:

$$I(M_1 : X|Y) = S(M_1Y) - S(Y) - S(XYM_1) + S(XY) = S(M_1), \tag{10}$$

since $S(XYM_1) = S(XY) = 2n$ and $M_1$ is independent of $Y$. It remains to calculate $S(M_1)$. Define

$$M_1^x = |\phi^1_{xy}\rangle\langle\phi^1_{xy}| = \frac{1}{2^n} \sum_{r,r'} |r\rangle|r\cdot x\rangle\langle r'|\langle r'\cdot x|. \tag{11}$$

Then

$$M_1 = \sum_{x\in\{0,1\}^n} \frac{1}{2^n} M_1^x = \frac{1}{2^{2n}} \sum_{r,r',i,j} c(r,r',i,j)|r\rangle|i\rangle\langle r'|\langle j|, \tag{12}$$

where the coefficient $c(r,r',i,j)$ is defined on $\{0,1\}^n \times \{0,1\}^n$ for $i,j \in \{0,1\}$ as:

$$c(r,r',i,j) := \#\Big\{x \in \{0,1\}^n : r\cdot x = i, r'\cdot x = j\Big\} = \begin{cases} 2^n & \text{if } r = r' = i = j = 0 \\ 0 & \text{if } r = r', i \neq j \\ & \text{or } r = 0, i = 1 \text{ or } r' = 0, j = 1 \\ 2^{n-1} & \text{if } r = r' \neq 0, i = j \\ & \text{or } r = 0 \neq r', i = 0 \\ & \text{or } r' = 0 \neq r, j = 0 \\ 2^{n-2} & \text{otherwise.} \end{cases} \tag{13}$$

We can show by computing the matrix and its eigenvalues that $S(M_1) = n/2 + 1/2$ (up to exponentially small terms).

Alice receives only one message from Bob, and after this message she has the state $\rho^2_{x,y} = \frac{1}{2^{2n}} \sum_{x,y} |x\rangle\langle x| \otimes |\phi^2_{xy}\rangle\langle\phi^2_{xy}|$ with

$$|\phi^2_{xy}\rangle = \frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\rangle|(r \oplus y)\cdot x\rangle. \tag{14}$$

We have

$$\begin{aligned} I(M_2 : Y|X) &= S(M_2X) - S(X) - S(M_2XY) + S(XY) && (15)\\ &= (n+1) - n + 2n - 2n && (16)\\ &= 1, && (17) \end{aligned}$$

where we used the fact that the state in the registers $M_2X$ has the same entropy as the following state (since there is a unitary on $M_2X$ that turns one into the other): $\rho^3_{x,y} = \frac{1}{2^{2n}} \sum_{x,y} |x\rangle\langle x| \otimes |\phi^3_{xy}\rangle\langle\phi^3_{xy}|$ with

$$|\phi^3_{xy}\rangle = \left(\frac{1}{\sqrt{2^n}} \sum_{r\in\{0,1\}^n} |r\rangle\right) |x\cdot y\rangle. \qquad \square \tag{18}$$

**Claim 2** *For the Quantum Information Cost of the protocol, we have (up to exponentially small terms)*

$$QIC_A(\Pi_{IP}) = n/2 + 1/2 \qquad , \qquad QIC_B(\Pi_{IP}) = n/2 + 3/2.$$

**Proof :** Let us compute $QIC_A(\Pi_{IP})$. For the first round, the state is

$$|\phi^1\rangle = \frac{1}{2^{3n/2}} \sum_{xy} |x\rangle_X |y\rangle_Y |xy\rangle_T \sum_{r\in\{0,1\}^n} |r\rangle|r\cdot x\rangle. \tag{19}$$

Then, we have

$$I(M_1 : R|Y) = S(M_1Y) - S(Y) + S(YR) - S(YM_1R) = S(M_1) = n/2 + 1/2. \quad (20)$$

We used here the fact that $S(YM_1R) = S(YR) = n$, the fact that $M_1$ is independent of $Y$, and the equality $S(M_1) = n/2 + 1/2$ we have already proven when analyzing the privacy loss in the proof of Claim 1.

We finally compute $QIC_B(\Pi_{IP})$. For the second round, we have

$$|\phi^2\rangle = \frac{1}{2^{3n/2}} \sum_{xy} |x\rangle_X |y\rangle_Y |xy\rangle_T \sum_{r\in\{0,1\}^n} |r\rangle|(r \oplus y) \cdot x\rangle, \quad (21)$$

and thus

$$\begin{aligned} I(M_2 : R|X) &= S(M_2X) - S(X) - S(XM_2R) + S(XR) && (22)\\ &= (n+1) - n - n + (n + n/2 + 1/2) && (23)\\ &= n/2 + 3/2. \quad \square && (24) \end{aligned}$$

Note that, since Alice must output $x \cdot y$, the quantity $CIC_B$ is at least one for any protocol computing Inner Product, which means that our protocol is optimal with respect to this quantity. Also note that the lower bound of Cleve et al. [13] on the quantum communication complexity of Inner Product shows that the sum of the information cost of both players is at least $n/2$.

**Remark 2** *We can easily describe a family of protocols $\Pi^t_{IP}, t \in [n]$ that provide a tradeoff between the CIC of Alice and Bob: Alice and Bob apply Protocol $\Pi_{IP}$ for the first $t$ bits of their inputs. Then, for the remaining $n - t$ bits they switch roles and in the end Bob sends the outcome to Alice. This new protocol is correct, since the inner product of $x, y$ is the XOR of the inner products of the smaller strings. Alice leaks at most $t/2 + 1/2$ bits from the first invocation and 1 from the second one. Bob leaks at most 1 bit from the first one and $(n - t)/2 + 1/2$ from the second one and hence $n/2 + 3$ in total.*

## 5 The Information Cost of Private Information Retrieval

In this section we construct a quantum protocol for Private Information Retrieval with polylogarithmic $CIC_S$ and polylogarithmic communication cost, by describing a general method to convert a classical scheme for Private Information Retrieval with $\ell > 1$ servers into a quantum scheme with a single server.

**Simulation of an $\ell$-server classical scheme by a 1-server quantum scheme** Consider a two-round classical scheme $\Pi_{PIR}$, where a user interacts with $\ell > 1$ servers that each possess a copy of the $n$-bit database and are not allowed to interact with each other. We can describe such a scheme as in Fig. 2, where $m_q, m_a, R \in \mathbf{N}$. We assume that the distribution of queries that each server receives is uniform, and hence do not reveal any information about the user's input. This assumption is true for essentially all known classical protocols for (information-theoretic) Private Information Retrieval, including the protocols described in [14, 7] that we will later use.

**Protocol** $\Pi_{PIR}$

1. The user picks uniformly at random $r \in [R]$ that corresponds to a $\ell$-tuple of queries $\{q_1^r, ..., q_\ell^r\}$ and asks query $q_i^r \in \{0,1\}^{m_q}$ to server $i \in [\ell]$.
2. Each server $i$, who received $q_i^r$, sends his answer $a_i^r \in \{0,1\}^{m_a}$, to the user.

Fig. 2. General form of a 2-round $\ell$-server classical protocol for Private Information Retrieval.

Let us now describe a quantum protocol $Q_{PIR}$ that simulates the classical protocol $\Pi_{PIR}$, but with a single server. The server and the user use $\ell$ query registers $Q_1, \ldots, Q_\ell$ of size $m_q$ each and $\ell$ answer registers $\mathrm{Ans}_1, \ldots, \mathrm{Ans}_\ell$ of size $m_a$ each. Moreover the user also holds a private register $Q$ of size $\ell \cdot m_q$ to keep a copy of the queries. The protocol is given in Fig. 3, where we use the notations $|a^r_{[i-1]}\rangle_{\mathrm{Ans}_{[i-1]}} := |a^r_1\rangle_{\mathrm{Ans}_1} \ldots |a^r_{i-1}\rangle_{\mathrm{Ans}_{i-1}}$ and $|0\rangle_{\mathrm{Ans}_{[i,\ell]}} := |0\rangle_{\mathrm{Ans}_i} \ldots |0\rangle_{\mathrm{Ans}_\ell}$.

**Protocol** $Q_{PIR}$

1. The user prepares the pure state
$$|\phi^1\rangle := \frac{1}{\sqrt{R}} \sum_r |q_1^r \ldots q_\ell^r\rangle_Q |q_1^r\rangle_{Q_1} \ldots |q_\ell^r\rangle_{Q_\ell} |0\rangle_{\mathrm{Ans}_1} \ldots |0\rangle_{\mathrm{Ans}_k}.$$
2. The user and the server iterate for $i = 1$ to $\ell$ :
   - at each odd round $2i-1$ the user holds the whole (pure) state
   $$|\phi^{2i-1}\rangle := \frac{1}{\sqrt{R}} \sum_r |q_1^r \ldots q_\ell^r\rangle_Q |q_1^r\rangle_{Q_1} \ldots |q_\ell^r\rangle_{Q_\ell} |a^r_{[i-1]}\rangle_{\mathrm{Ans}_{[i-1]}} |0\rangle_{\mathrm{Ans}_{[i,\ell]}}$$
   and sends registers $(Q_i, \mathrm{Ans}_i)$ to the server ;
   - at each round $2i$, the server holds $(Q_i, \mathrm{Ans}_i)$. He reads the query, writes the answer in the $\mathrm{Ans}_i$ register and sends the two registers to the user.

Fig. 3. Quantum protocol simulating $\Pi_{PIR}$ with one server.

This protocol indeed simulates the classical protocol $\Pi_{PIR}$: at the end of the protocol, the user holds
$$|\phi^{2k}\rangle = \frac{1}{\sqrt{R}} \sum_r |q_1^r \ldots q_\ell^r\rangle_Q |q_1^r\rangle_{Q_1} \ldots |q_\ell^r\rangle_{Q_\ell} |a^r_{[\ell]}\rangle_{\mathrm{Ans}_{[\ell]}} \tag{25}$$
and by measuring in the computational basis, he gets a uniformly random $\ell$-tuple of queries and their answers, hence he has the same success probability as the user in the classical scheme. The communication cost of Protocol $Q_{PIR}$ is $2\ell(m_a + m_q)$ qubits. We now describe its information cost.

**Theorem 3** *For the above protocol $\Pi_{PIR}$ under uniform distribution of inputs, we have*

$$CIC_S(\Pi_{PIR}) = O(\ell(m_a + m_q)) \quad , \quad CIC_U(\Pi_{PIR}) = 0,$$
$$QIC_S(\Pi_{PIR}) = O(\ell(m_a + m_q)) \quad , \quad QIC_U(\Pi_{PIR}) = \Omega(\log(n)),$$

*where the queries and answers are in $\{0,1\}^{m_q}$ and $\{0,1\}^{m_a}$.*

**Proof :** The first statement is obvious since in $\Pi_{PIR}$ the total communication is $2\ell(m_a+m_q)$ and hence $CIC_S(\Pi_{PIR}), QIC_S(\Pi_{PIR})$ are $O(\ell(m_a + m_q))$.

As for the $CIC_U$, note that each message independently does not leak any information about the user's input, since the quantum message is exactly the same distribution over classical queries that each server receives in the classical scheme, which we know is perfectly private.

Now, for $QIC_U(\Pi_{PIR})$, we know from Theorem 3.2 in [8] that, if the user leaks at most $b$ bits about his input, then the server has to leak at least $\Omega(n/2^{O(b)})$ bits about his $n$-bit database, or equivalently, if the server leaks at most $t$ bits about the database, then the user must leak at least $\Omega(\log(n/t))$ bits about his input. Since in our scheme the communication is bounded by polylog$(n)$, we obtain that the user has to leak at least $\Omega(\log(n))$ about his input. □

**Application: a quantum protocol for PIR with polylogarithmic privacy loss** We consider the classical scheme proposed in [14, 7].

**Lemma 1** [See Theorem 3 in [14]] *There is a (classical) private information retrieval scheme for $\frac{1}{2} \cdot (\log n + \log\log n) + 1$ servers, each holding $n$ bits of data, where the server sends to each server a query of $O(\log(n) \cdot \log\log(n))$ bits, and receives from each server an answer of $O(\log\log(n))$ bits (so that the total communication cost is $O(\log^2 n \cdot \log\log(n))$ bits).*

By converting the classical protocol of Lemma 1 into a one-server quantum protocol by the above construction, and applying Theorem 3, we obtain the following result.

**Corollary 1** *There exists a one-server quantum protocol for Private Information Retrieval, with total communication cost $O(\log^2(n) \cdot \log\log(n))$, such that:*

- *it is perfectly private for the user according to CIC ;*
- *the server leaks $O\big(\text{polylog}(n)\big)$ information.*

By Corollary 1 and Theorem 3.2 in [8] we obtain the following separation between the different notions of information cost.

**Corollary 2** *There exists a quantum protocol which is perfectly private for the user according to CIC and has information cost (CIC and QIC) for the server only polylogarithmic on the size of the database. On the other hand, any quantum protocol which is perfectly private for the user according to QIC has information cost (CIC and QIC) for the server linear on the size of the database.*

## 6 Logarithmic Scheme for PIR with Prior Entanglement

We now study one-server quantum private information retrieval in the same setting as in the previous section, but allowing prior entanglement between the server and the user, and construct a protocol with $CIC_S$ and communication cost $O(\log(n))$. For simplicity we will

assume in this section that $n = 2^\ell$, and write the user's input using its binary representation as $i = i_1 i_2 \dots i_\ell$, where $i_1, \dots, i_\ell$ are bits such that $i = 1 + \sum_{k=1}^{\ell} i_k 2^{\ell-k}$. The case where $n$ is not a power of two can be dealt in a similar way, or simply by adding zeros to the database in order to obtain a size that is a power of two.

For convenience we introduce the following notation.

**Definition 3** *Let $s$ be any positive integer, and $z$ be any binary string of length $2^s$. Define $z[0]$ and $z[1]$ as the first and second halves of the string $z$, respectively. For any $k \in \{2, \dots, s\}$ and any $k$ bits $j_1, \dots, j_k$, let $z[j_1, \dots, j_k]$ be the binary string of length $2^{s-k}$ defined by the recurrence relation $z[j_1, \dots, j_k] = (z[j_1, \dots, j_{k-1}])[j_k]$.*

Let us consider an example to illustrate this definition: if $s = 3$ and $z = 10100110$, then $z[0] = 1010$, $z[1] = 0110$, $z[0,0] = 10$, $z[0,1] = 10$, $z[1,0] = 01$, $z[1,1] = 10$ and, for instance, $z[0,0,0] = 1$ or $z[1,0,1] = 1$. Note that, with these definitions, the bit $x_i$ that the user wants to output in a protocol for Private Information Retrieval is $x[i_1, \dots, i_\ell]$.

Our protocol will use, besides the two registers containing the inputs, the following quantum registers:

- $\ell$ quantum registers $R_1, \dots, R_\ell$ where $R_k$ is a register of $2^{\ell-k}$ qubits for $k \in \{1, \dots, \ell\}$;
- $\ell$ quantum registers $R'_1, \dots, R'_\ell$ where $R'_k$ is a register of $2^{\ell-k}$ qubits for $k \in \{1, \dots, \ell\}$;
- two one-qubit quantum registers $Q_0$ and $Q_1$.

Define the unitary operator $V_1$ acting on $(R_1, Q_0, Q_1)$ as follows:

$$V_1 \left( |z\rangle_{R_1} |a\rangle_{Q_0} |b\rangle_{Q_1} \right) = |z\rangle_{R_1} |a \oplus z \cdot x[0]\rangle_{Q_0} |b \oplus z \cdot x[1]\rangle_{Q_1} \tag{26}$$

for any string $z \in \{0,1\}^{2^{\ell-1}}$ and any bits $a, b \in \{0,1\}$. For any integer $k \in \{2, \dots, \ell\}$, we define the unitary operator $V_k$ acting on $(R_{k-1}, R_k, Q_0, Q_1)$ as follows:

$$V_k \left( |y\rangle_{R_{k-1}} |z\rangle_{R_k} |a\rangle_{Q_0} |b\rangle_{Q_1} \right) = |y\rangle_{R_{k-1}} |z\rangle_{R_k} |a \oplus z \cdot y[0]\rangle_{Q_0} |b \oplus z \cdot y[1]\rangle_{Q_1} \tag{27}$$

for any strings $y \in \{0,1\}^{2^{\ell-k+1}}$, $z \in \{0,1\}^{2^{\ell-k}}$ and any bits $a, b \in \{0,1\}$.

For any integer $k \in \{1, \dots, \ell\}$, define the state

$$|\Phi_k\rangle_{(R_k, R'_k)} = \frac{1}{\sqrt{2^{2^{\ell-k}}}} \sum_{z \in \{0,1\}^{2^{\ell-k}}} |z\rangle_{R_k} |z\rangle_{R'_k}. \tag{28}$$

We assume that the server and the user initially share the quantum state

$$|\Phi_1\rangle_{(R_1, R'_1)} \otimes |\Phi_2\rangle_{(R_2, R'_2)} \otimes \dots \otimes |\Phi_\ell\rangle_{(R_\ell, R'_\ell)} \otimes |0\rangle_{Q_0} |0\rangle_{Q_1}, \tag{29}$$

where $R_1, \dots, R_\ell, Q_0, Q_1$ are owned by the server and $R'_1, \dots, R'_\ell$ are owned by the user. Our quantum protocol is given in Fig. 4.

We analyze the correctness, the communication cost and the information cost of Protocol $\mathcal{P}_{PIR}$, and prove the following theorem.

**Theorem 4** *The protocol $\mathcal{P}_{PIR}$ for input size $n = 2^\ell$ has communication cost $4\ell + 1$ qubits and correctly computes the index function. Moreover, under uniform distribution of inputs, we have*

$$\begin{aligned} CIC_S(\mathcal{P}_{PIR}) &\leq 2\ell + 1 \quad , \quad CIC_U(\mathcal{P}_{PIR}) = 0, \\ QIC_S(\mathcal{P}_{PIR}) &\leq 2\ell + 1 \quad , \quad QIC_U(\mathcal{P}_{PIR}) = \Omega(\ell). \end{aligned}$$

**Protocol $\mathcal{P}_{PIR}$**

1. For $k$ from 1 to $\ell$, the server and the user do the following:
   (a) The server applies $V_k$, and then sends Registers $Q_0$ and $Q_1$ to the user;
   (b) The user applies the Pauli gate $Z$ over Register $Q_{i_k}$ and sends back Registers $Q_0$ and $Q_1$ to the server;
   (c) The server applies $V_k$, and applies a Hadamard transform on each of the $2^{\ell-k}$ qubits in Register $R_k$;
   (d) The user applies a Hadamard transform on each of the $2^{\ell-k}$ qubits in Register $R'_k$.
2. The server sends Register $R_\ell$ to the user.

Fig. 4. Quantum protocol for private information retrieval with prior entanglement.

The following lemma, which can be easily shown by recursion on $k$, will be convenient to analyze Protocol $\mathcal{P}_{PIR}$.

**Lemma 2** *Assume that Protocol $\mathcal{P}_{PIR}$ is applied when the server's input is $x \in \{0,1\}^{2^N}$ and the user's input is $i \in \{0,1\}^N$. Then, at the end of the $k$-th iteration of the loop in Step 1, the state of the quantum system is (omitting a global normalization factor)*

$$\left[\sum_{y^1,\ldots,y^k}\left(|y^1\rangle_{R_1}|x[i_1]\oplus y^1\rangle_{R'_1}\bigotimes_{j=2}^{k}|y^j\rangle_{R_j}|y^{j-1}[i_j]\oplus y^j\rangle_{R'_j}\right)\right]\otimes|0\rangle_{Q_0}|0\rangle_{Q_1}\otimes\left[\bigotimes_{j=k+1}^{N}|\Phi_j\rangle_{(R_j,R'_j)}\right],$$

*where the sum is over all strings $y^1 \in \{0,1\}^{2^{N-1}}, \ldots, y^k \in \{0,1\}^{2^{N-k}}$.*

**Proof of Theorem 4:** Since each iteration of the loop in Step 1 uses four qubits of communication, and one additional qubit is used at Step 2, the overall communication cost is $4N+1$.

Next, we show that this protocol correctly computes the index function, i.e., the user can output $x_i$. From Lemma 2, the state of the quantum system at the end of Protocol $\mathcal{P}_{PIR}$ is (omitting a global normalization factor)

$$\sum_{y^1,\ldots,y^N}|y^1\rangle_{R_1}|x[i_1]\oplus y^1\rangle_{R'_1}|y^2\rangle_{R_2}|y^1[i_2]\oplus y^2\rangle_{R'_2}\cdots|y^N\rangle_{R_N}|y^{N-1}[i_N]\oplus y^N\rangle_{R'_N}|0\rangle_{Q_0}|0\rangle_{Q_1}, \quad (30)$$

where the server owns Registers $R_1,\ldots,R_{N-1},Q_0,Q_1$, and the user owns Register $R_N$ and Registers $R'_1,\ldots,R'_N$. If the server and the user measure all their registers, the user obtains strings $a^N, b^1, \ldots, b^N$ such that

$$\begin{cases} a^N &= y^N, \\ b^1 &= x[i_1]\oplus y^1, \\ b^2 &= y^1[i_2]\oplus y^2, \\ \vdots &\vdots \quad \vdots \\ b^N &= y^{N-1}[i_N]\oplus y^N, \end{cases}$$

for some strings $y^1, \ldots, y^N$ corresponding to the server's measurement outcomes. Note that

$$x[i_1, i_2, \ldots, i_N] = b^1[i_2, \ldots, i_N] \oplus b^2[i_3, \ldots, i_N] \oplus \cdots \oplus b^{N-1}[i_N] \oplus b^N \oplus a^N, \qquad (31)$$

which means that the user can recover $x_i = x[i_1, i_2, \ldots, i_N]$ from his measurement outcomes.

The upper bounds on $CIC_S(\mathcal{P}_{PIR})$ and $QIC_S(\mathcal{P}_{PIR})$ follow from the observation that the total length of the messages received by the user is $2N+1$. The lower bounds on $QIC_U(\mathcal{P}_{PIR})$ follow from the same argument (based on [8]) as in Theorem 3.

Finally, let us prove that $CIC_U(\mathcal{P}_{PIR}) = 0$, by showing that the server's state just after receiving the message from the user during the $k$-th iteration of Step 1 of Protocol $\mathcal{P}_{PIR}$ is independent of $i$, for each $k \in \{1, \ldots, N\}$. In the case $k = 1$, the state of the registers owned by the server just after receiving the message from the user is, omitting a global normalization factor,

$$\left[ \sum_{z \in \{0,1\}^{2^{N-1}}} |\Psi(z)\rangle\langle\Psi(z)| \right] \otimes \left[ \bigotimes_{j=2}^{N} \sum_{z \in \{0,1\}^{2^{N-j}}} |z\rangle_{R_j}\langle z|_{R_j} \right], \qquad (32)$$

where $|\Psi(z)\rangle = (-1)^{x[i_1]\cdot z}|z\rangle_{R_1}|x[0]\cdot z\rangle_{Q_0}|x[1]\cdot z\rangle_{Q_1}$. Since

$$|\Psi(z)\rangle\langle\Psi(z)| = |z\rangle_{R_1}|x[0]\cdot z\rangle_{Q_0}|x[1]\cdot z\rangle_{Q_1}\langle z|_{R_1}\langle x[0]\cdot z|_{Q_0}\langle x[1]\cdot z|_{Q_1} \qquad (33)$$

is independent of $i$, the above state is also independent of $i$. For the case $k \geq 2$, by using Lemma 2. the state of the registers owned by the server just after receiving the $k$-th message from the user is, omitting a global normalization factor,

$$\left[ \sum_{y^1, \ldots, y^{k-1}} \sum_{z \in \{0,1\}^{2^{N-k}}} |\Psi_x(y^1, \ldots, y^{k-1}, z)\rangle\langle\Psi_x(y^1, \ldots, y^{k-1}, z)| \right] \otimes \left[ \bigotimes_{j=k+1}^{N} \sum_{z \in \{0,1\}^{2^{N-j}}} |z\rangle_{R_j}\langle z|_{R_j} \right], \qquad (34)$$

where

$$|\Psi_x(y^1, \ldots, y^{k-1}, z)\rangle = (-1)^{y^{k-1}[i_k]\cdot z}|y^1\rangle_{R_1} \cdots |y^{k-1}\rangle_{R_{k-1}}|z\rangle_{R_k}|y^{k-1}[0]\cdot z\rangle_{Q_0}|y^{k-1}[1]\cdot z\rangle_{Q_1}, \qquad (35)$$

and is again independent of $i$. □

**Acknowledgments.** The authors are grateful to Ronald de Wolf for helpful comments about this work, and for pointing out Ref. [9]. The authors are also grateful to Rahul Jain for helpful discussion, and to the anonymous reviewers for helpful comments. Iordanis Kerenidis and Mathieu Laurière have been supported by the ERC grant QCC and the EU grant QAlgo. François Le Gall has been supported by the Grant-in-Aid for Scientific Research (A) No. 24240001 of the Japan Society for the Promotion of Science and the Grant-in-Aid for Scientific Research on Innovative Areas No. 24106009 of the Ministry of Education, Culture, Sports, Science and Technology in Japan. Mathys Rennela has been supported by the ERC grant QCLS.

## References

1. A. C-C. Yao (1979). Some complexity questions related to distributive computing. *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pp. 209-213.
2. A. C-C. Yao (1993). Quantum circuit complexity. *Proceedings of the 34th Annual Symposium on Foundations of Computer Science*, pp. 352-361.
3. R. Jain, J. Radhakrishnan and P. Sen (2003). A lower bound for bounded round quantum communication complexity of set disjointness. *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pp. 220-229.
4. H. Klauck. On quantum and approximate privacy (2002). In *Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science*, Lecture Notes in Computer Science, Vol. 2285, pp. 335-346.
5. F. Le Gall (2012). Quantum private information retrieval with sublinear communication complexity. *Theory of Computing*, Vol. 8, pp. 369-374.
6. D. Touchette (2015). A new, fully quantum notion of information complexity, and an application to direct sum for bounded round quantum communication complexity. STOC, 2015.
7. B. Chor, O. Goldreich, E. Kushilevitz and M. Sudan (1998). Private information retrieval. *Journal of the ACM*, Vol. 45(6), pp. 965-981.
8. R. Jain, J. Radhakrishnan and P. Sen (2009). A new information-theoretic property about quantum states with an application to privacy in quantum communication. *Journal of the ACM*, Vol. 56(6), Article 33.
9. K. Ruben Brokkelkamp (2013). *Quantum private information retrieval.* Bachelor Thesis, University of Amsterdam.
10. E. Kushilevitz and N. Nisan (1997). *Communication Complexity.* Cambridge University Press.
11. P. Bro Miltersen, N. Nisan, S. Safra and A. Wigderson (1998). On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, Vol. 57(1), pp. 37-49.
12. B. Chor and O. Goldreich (1988), Unbiased bits from weak sources of randomness and probabilistic communication complexity, *SIAM Journal on Computing*, Vol. 17(2), pp. 230-261.
13. R. Cleve, W. van Dam, M. Nielsen and A. Tapp (1999). Quantum entanglement and the communication complexity of the inner product function. *Proceedings of the First NASA International Conference on Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science, Vol. 1509, pp. 61-74.
14. B. Chor, O. Goldreich, E. Kushilevitz and M. Sudan (1995). Private information retrieval. *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, pp. 41-50.
15. A. Baumeler and A. Broadbent (2014), Quantum Private Information Retrieval has linear communication complexity, *Journal of Cryptology*, to appear. Also arXiv.org e-Print archive, arXiv:1304.5490v2, 2014.