Privacy in Quantum Communication Complexity

Iordanis Kerenidis* CNRS, Université Paris 7
Mathieu Laurière† Université Paris 7
François Le Gall‡ University of Tokyo
Mathys Rennela§ University of Nijmegen

Technical version available at `http://arxiv.org/abs/1409.8488`

In two-party communication complexity, Alice and Bob receive inputs x and y 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 in a private way.

In classical communication protocols, the privacy loss (or information cost) 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 privacy loss 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 the protocol.

In quantum communication protocols, Alice and Bob again receive classical inputs x and y and wish to compute a function $f(x, y)$. Here, 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 on his workspace and the message qubits, and sends the message qubits to the other player who continues the protocol. Since the message qubits can be reused in different rounds of the protocol and copying of the quantum states may be impossible, we cannot define a transcript. Hence, we know of no way to define notions of privacy or quantum information cost, other than by a *round-by-round* definition.

By a chain rule argument, the classical definition of privacy loss is equivalent to a *round-by-round* definition where for every round i, we calculate the information that the message at round i reveals about the sender's input to the receiver, who knows his input and has kept a copy of all previous messages in his 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. What we would like to calculate then, is 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 more issue to discuss before we provide a definition of privacy loss. Each player has a register where the input is written in the beginning of the protocol. This input is of course a classical input. One natural possibility is therefore to consider that the input register is a classical register, meaning it cannot be entangled with the workspaces and the message space. The second possibility, studied for instance in [1, 4, 8], 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. One of the main goals of this paper is to investigate quantum communication complexity under the former definition of privacy. In particular, we show how to construct private protocols for several fundamental tasks, such as Private Information Retrieval and the Inner Product function. The second main goal of the paper is to investigate the differences between these two concepts of privacy (classical input registers versus quantum input registers). We now detail the definitions.

Privacy for quantum protocols with classical input registers Privacy with classical input registers has previously been discussed for some classes of quantum protocols, for example in [5, 6]. Klauck defined a notion of honest behaviour of the players, according to which, at every round of the protocol, the state of the message qubits sent must be equal to the one prescribed by the protocol. He, then, considered the privacy of Disjointness by looking at a quantum protocol with pure state messages. However, for protocols with mixed state messages, his definition allows for a player to change the execution of the protocol considerably. Hence, we propose the following definition of honest execution of a protocol that takes into consideration protocols in which the players can send mixed states, and is equivalent to Klauck's definition for pure state protocols.

*jkeren@liafa.univ-paris-diderot.fr
†mathieu.laurere@univ-paris-diderot.fr
‡legall@is.s.u-tokyo.ac.jp
§m.rennela@cs.ru.nl

Definition 1. *Let π be a quantum protocol where, at each round i, the quantum registers corresponding to the message sent, Alice's workspace and Bob's workspace are denoted M_i, A_i and B_i, respectively. An honest execution of the protocol is such that for all i, the joint state in the registers A_i, M_i, B_i is equal to the state described by the protocol, up to a possible local operation on A_i and a possible local operation on B_i.*

We can now provide the definition of privacy loss.

Definition 2. *For a protocol π, the privacy loss of Alice and Bob are defined respectively by*

$$L_A(\pi) = \sum\nolimits_i I(M_i : X|Y, B_i) \text{ and } L_B(\pi) = \sum\nolimits_i I(M_i : Y|X, A_i),$$

where X, Y are classical registers that hold the inputs, and M_i, A_i, B_i are quantum registers that correspond, respectively, to the message qubits, Alice's workspace and Bob's workspace at round i.

It is easy to see that the privacy loss for any honest execution of the protocol is the same, hence we only need to consider the states described by the protocol itself. Also, if according to the protocol Alice holds a pure state at some round, then an honest Bob can not entangle his workspace with Alice's state. Last if π computes some function f, then any honest execution also computes f. This is important, since as in the classical case, it makes sense to consider only the privacy of protocols that actually compute the function f.

Let us remark at this point that the above definition is not appropriate for classical protocols. Consider the protocol, where Alice sends a random string r, Bob returns the string r and Alice replies with the string $r + x$. If Bob honestly follows the protocol, he gets no information about x from the first message and no information from the second message, since he had not kept anything on his workspace. This is why in the round-by-round definition of classical privacy, we condition on the previous messages, to avoid that the parties forget what they already know. For this example, Definition 2 is actually still meaningful since we consider in this work protocols that only allow the players to perform unitary operations, as standard in quantum communication complexity (see, e.g., [8]). This way, 'sending a random string r' implies that the player has kept a quantum purification of the string, meaning that the other player cannot copy it anymore without changing the execution of the protocol. It is an interesting question to define a notion of privacy for protocols where the players are allowed to both apply unitary operations and send classical messages at the same time.

Information cost of quantum protocols with quantum input registers In the previous definition X and Y are classical registers. There are definitions of quantum information cost where the registers X and Y are quantum registers that can hold superpositions of inputs and be entangled with other registers (e.g., [1, 4, 8]).

This is the case for example with the "superposed" information cost for a protocol π, variants of which have been used in [1, 4]. We denote it by $SIC_A(\pi)$ and $SIC_B(\pi)$ respectively, where when one considers the privacy loss for Alice, Bob is allowed to have in his quantum register Y a superposition of all his possible inputs instead of the fixed classical input y he received, and for each input in the superposition he honestly executes π. While this is certainly a strategy that Bob can follow in order to acquire more information about Alice's input, as we said before, we are not in a cryptographic scenario with cheating players, but we want to compute the privacy loss of the specific protocol that computes the function. Also note that if Bob starts with a superposition of all possible inputs, Alice and Bob are not able to compute the value of the function f on the received inputs (x, y), but only on a point (x, y') for a random y'. Similar definitions of the superposed information cost have been given in [1, 4].

Recently, another definition of quantum information cost was proposed by Touchette [8]. In this case, the registers X and Y are initially entangled with an external register R (an environment, not accessible to the players) so that the initial input state of the players for a distribution μ is $\sum_{x,y} \mu(x, y)|x, y\rangle_R|x\rangle_X|y\rangle_Y$.

For a protocol π, the Quantum Information Cost of Alice and Bob are defined respectively as follows [8]:

$$QIC_A(\pi) = \sum\nolimits_i I(M_i : R|Y, B_i) \text{ and } QIC_B(\pi) = \sum\nolimits_i I(M_i : R|X, A_i).$$

This definition has nice properties, eg. it is equal to the amortized quantum communication complexity [8].

We believe that the notions of SIC and QIC are more suited as tools to lower bound the communication complexity than as tools to measure the privacy of the protocol (the parties can indeed, as mentioned above, considerably deviate from the protocol and may not even compute the function f on the received inputs). In comparison, privacy with classical input registers (Definition 2) appears to be more suited as a tool to discuss the privacy of a protocol computing a function with classical inputs. Notice also that in the classical case,

if we allow Alice to run the protocol with a random input instead of the input x she received, then private protocols, like for the IdMinimum function (where the output is $\min(x, y)$ with the identity of the player who has this value) are rendered not private. In the present work, we first prove the following inequalities between these definitions and show that in some cases the gaps can be exponentially large.

Theorem 1. *$L_A(\pi) \leq SIC_A(\pi) \leq QIC_A(\pi)$ and $L_B(\pi) \leq SIC_B(\pi) \leq QIC_B(\pi)$.*

We then analyze the privacy of new quantum protocols (with classical input registers) for computing the Inner Product function and for Private Information Retrieval, as described below.

The privacy of Inner Product We start by proving a simple gap between quantum communication complexity and both privacy loss and information cost for the Inner Product function.

Theorem 2. *There exists a quantum protocol for Inner Product, which is perfectly private for Bob and where Alice's privacy loss is $n/2$.*

We use the following protocol where the players exchange only pure states as messages (here $x \in \{0,1\}^n$ denotes Alice's input and $y \in \{0,1\}^n$ denotes Bob's input).

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

The protocol is correct, since Alice can apply the unitary $U_x : |r\rangle|b\rangle \mapsto |r\rangle|b \oplus r \cdot x\rangle$, to get $x \cdot y$. For the privacy, we can calculate Bob's information from the first message and show that it is $n/2$ bits of information. Moreover, Alice gets no information about Bob's input apart from the value $x \cdot y$, since the pure state she receives in the second round is locally equivalent to a tensor product of two states, the first independent of x, y and the second equal to $|x \cdot y\rangle$. Note also that we can actually split the input into two parts and exchange the roles of Alice and Bob in the protocol in order to provide a tradeoff between the respective privacy losses, where for any t, we have that one player's privacy loss is $t/2$ and the other's $(n-t)/2$.

The privacy of Private Information Retrieval PIR 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 $x \in \{0,1\}^n$, the user has an index $y \in \{1, \ldots, n\}$, and the goal is for the user to output x_y. Any classical protocol perfectly private for the user (i.e., in which the server obtains no information about y) requires $\Omega(n)$ bits of communication [2]. Recently, Le Gall [6] showed that there exists a quantum protocol for this task, perfectly private for the user (according to Definition 2), with communication complexity $O(\sqrt{n})$. Here we ask: Can this quadratic upper bound be improved? More generally, how much information does a single server have to leak about the database in any protocol which is perfectly private for the user? We show the following result.

Theorem 3. *There exists a quantum protocol for Private Information Retrieval, which is perfectly private for the user and in which the server's privacy loss is polylogarithmic on the size of the database. Moreover its communication complexity is also polylogarithmic on the size of the database.*

Note that, again, we are not discussing how parties can deviate from the protocol in order to gain more information. The proof has two steps: first, we show how to take any k-server classical PIR scheme and translate it into a quantum one-server scheme, such that the index remains perfectly private. In high level, let the user in the classical k-server scheme use some uniform randomness r in order to pick the queries $(q_1^r, \ldots, q_k^r)$ and let the servers reply with $(a_1^r, \ldots, a_k^r)$. In our quantum scheme, the user creates the state

$$\frac{1}{\sqrt{|\sharp r|}} \sum_r |q_1^r \ldots q_k^r\rangle_Q |q_1^r\rangle_{Q_1} \ldots |q_k^r\rangle_{Q_k} |0\rangle_{\mathrm{Ans}_1} \ldots |0\rangle_{\mathrm{Ans}_k}$$

and at each round i of the protocol, he sends the two registers Q_i and Ans_i to the single server who writes the answer on the Ans_i register and returns both of them. It is easy to see that the protocol is correct and we can show that it is perfectly private for the user (according to Definition 2). Then, we use a classical PIR scheme with a logarithmic number of servers and polylogarithmic communication [2], which implies that the privacy loss about the database is polylogarithmic, since it is always less than the communication.

Finally, we improve the above upper bounds in the case where the user and the server share prior entanglement: we construct a new quantum protocol for Private Information Retrieval, perfectly private for the user (again according to Definition 2), where the server's privacy loss is $O(\log n)$ bits. The communication complexity is $O(\log n)$ qubits, which is optimal since, even with prior entanglement, the quantum communication complexity of the Index Function is $\Omega(\log n)$.

References

[1] A. Chailloux and G. Scarpa, Parallel repetition of entangled games with exponential decay via the superposed information cost. *Proceedings of the 41st International Colloquium on Automata, Languages, and Programming*, Lecture Notes in Computer Science, Vol. 8572, pp. 296-307, 2014.

[2] B. Chor, O. Goldreich, E. Kushilevitz and M. Sudan. Private information retrieval. *Journal of the ACM*, Vol. 45(6), pp. 965-981, 1998.

[3] R. Cleve, W. van Dam, M. Nielsen and A. Tapp. 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, 1999.

[4] R. Jain, J. Radhakrishnan and P. Sen. 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, 2003.

[5] H. Klauck. On quantum and approximate privacy. In *Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science*, Lecture Notes in Computer Science, Vol. 2285, pp. 335-346, 2002.

[6] F. Le Gall. Quantum private information retrieval with sublinear communication complexity. *Theory of Computing*, Vol. 8, pp. 369-374, 2012.

[7] P. Bro Miltersen, N. Nisan, S. Safra and A. Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, Vol. 57(1), pp. 37-49, 1998.

[8] D. Touchette. Quantum information complexity and amortized communication. arXiv.org e-Print archive, arXiv:1404.3733, 2014.