Anonymous Quantum Communication

Gilles Brassard[1], Anne Broadbent[1], Joseph Fitzsimons[2], Sébastien Gambs[1], and Alain Tapp[1]

[1] Université de Montréal
Département d'informatique et de recherche opérationnelle
C.P. 6128, Succ. Centre-Ville, Montréal (Québec), H3C 3J7 Canada
{brassard,broadbea,gambsseb,tappa}@iro.umontreal.ca
[2] University of Oxford
Department of Materials
Parks Road, Oxford, OX1 3PH United Kingdom
joe.fitzsimons@materials.ox.ac.uk

Abstract. We present the first protocol for the anonymous transmission of a quantum state that is information-theoretically secure against an active adversary, without any assumption on the number of corrupt participants. The anonymity of the sender and receiver, as well as the privacy of the quantum state, are perfectly protected except with exponentially small probability. Even though a single corrupt participant can cause the protocol to abort, the quantum state can only be destroyed with exponentially small probability: if the protocol succeeds, the state is transferred to the receiver and otherwise it remains in the hands of the sender (provided the receiver is honest).

Keywords: quantum cryptography, multiparty computation, anonymity, dining cryptographers.

1 Introduction

In David Chaum's classic dining cryptographers scenario [Cha88], a group of cryptographers is having dinner at a restaurant and it is the case that either one of them has anonymously paid the dinner bill or the NSA has paid. The task that the cryptographers wish to accomplish is to find out which of the two cases occurred, without revealing any additional information. The security of Chaum's protocol does not rely on any computational assumption, but only on the cryptographers having access to pairwise private channels and to a broadcast channel. A simple extension to this protocol allows a single participant, say Alice, to broadcast a message to all the other participants in such a way that Alice's identity is information-theoretically protected.

But what if Alice wishes to send a private message to Bob (who is also sitting at the dinner table), while ensuring the anonymity of both herself and of Bob? This task is called *anonymous message transmission*. As an instance of multiparty secure computation, such a protocol can be accomplished, assuming pairwise private channels and a broadcast channel, as long as a majority of participants are honest [RB89]. Recently, two of us [BT07] have given a protocol that requires pairwise private channels and a broadcast channel, and accomplishes anonymous message transmission *without* any assumption on the number of honest participants. The protocol, however, allows even a single corrupt participant to cause an abort.

Our main contribution is to give the first information-theoretically secure protocol for *quantum* anonymous transmission that tolerates any number of corrupt participants. That is, our protocol allows Alice to send a quantum message to Bob such that both Alice and Bob remain anonymous (no participant learns the identity of Alice—even if Bob is corrupt—and the identity of Bob remains known only to Alice), and the quantum message remains private (nothing about it leaks to participants other than Bob, unless of course Bob is corrupt). The anonymity of the sender and receiver, as well as the privacy of the quantum message, are perfect except with exponentially small probability, regardless of the behaviour of cheating parties, with no need to rely on any assumptions other than the availability of a classical broadcast channel as well as private authenticated quantum channels between each pair of participants. Our protocol has features similar to the anonymous (classical) message transmission protocol mentioned above: we can tolerate an arbitrary number of corrupt participants, but any single corrupt participant can cause the protocol to abort. However, no private information can be obtained by making the protocol abort.

Since Alice sends quantum information, we need to address a concern that did not exist in the context of classical anonymous message transmission: the state to be transmitted should never be destroyed *even if the protocol aborts* (unless the receiver is corrupt, since in that case he can follow honestly the protocol until the very end, and then destroy the successfully transmitted message!). Because of the no-cloning theorem [WZ82], the sender cannot generally keep a backup copy of the message before entering the protocol. Nevertheless, we accomplish this safeguard as part of the main protocol with a simple and novel notion called *fail-safe teleportation*. This notion ensures that if something went wrong with the transmission of the state, its integrity is never at stake because the receiver can always teleport it back to the sender in a way that does not compromise anonymity.

1.1 Anonymity

Anonymity is a basic cryptographic concept whose goal is to hide the identity of the sender or receiver of a message (or both). It is different from, but often complementary to *privacy*, which ensures the confidentiality of a message. Examples of anonymous tasks include sending an anonymous letter to one's love, using an email account with a pseudonym, accessing a web page through a trusted identity proxy server or blind reviewing of a conference paper. Three approaches to classical anonymity are generally considered. The first one requires the help of a trusted third party that forwards messages between participants without revealing the identity of the senders. Anonymizers [Boy97,GGK+99] belong to this class. The second approach uses chains of untrusted servers that randomize the ordering of messages. This reordering prevents an outside observer from linking the sender and the receiver of a particular message. The privacy of messages is generally assured by a public-key cryptosystem. Chaum's MixNets [Cha81] are an instance of techniques using this approach. The third and last approach offers information-theoretic security, assuming resources such as a broadcast channel and pairwise private channels. Chaum's dining cryptographers protocol [Cha88] is the archetypical example of a protocol in this category.

1.2 Model

In our model, we suppose that each pair of participants shares a *private authenticated quantum channel*, which means that a participant can send an authenticated private message (quantum or classical) to any other participant. Such a channel can be implemented if the participants share pairwise quantum channels as well as classical secret keys. An extra tool is given to the participants under the form of a (classical) *broadcast channel*. This channel guarantees that all participants receive the same message from a publicly known sender, and that the message is not modified while in transit.

Two security models are generally considered in secure multiparty computation: *honest-but-curious* and *malicious*. In the honest-but-curious model (also called *semi-honest*), the participants are assumed to follow the protocol (thus being honest) but at the same time record all the information they have seen during its execution (thus being curious). In this model, a protocol is said to be secure against a *collusion* of participants if, by pooling their data, these participants cannot learn more information than from their inputs and the output of the protocol alone. In the malicious model, participants may actively cheat and deviate from the original prescription of the protocol. Cheaters can for instance try to learn information about the input of honest participants or tamper with the output of the protocol. Formal definitions can be found in Chapter 7 of [Gol04]. Both these models are neatly encapsulated by considering a central entity called an *adversary*, which controls some of the participants, rendering them *corrupt*. The adversary is *passive* if the corrupt participants are honest-but-curious, and *active* if the corrupt participants are malicious. In this paper, we consider the case of an active adversary that chooses the set of corrupt participants before the execution of the protocol.

In the scenario that we consider, within a group of n participants, the anonymous sender communicates a private quantum message to an anonymous receiver. The sender is unknown to all participants and the receiver is unknown to all participants except to the sender. We give the following definitions:

Definition 1 (Sender Anonymity). *A protocol achieves* sender anonymity *if it does not reveal any information concerning the identity of the sender to any adversary. An exception concerns the receiver (or the adversary, if the receiver is corrupt), who may legitimately learn something about the identity of the sender by virtue of the contents of the transmitted message.*

Note that in particular, if the sender is corrupt, a protocol vacuously achieves sender anonymity, and that sender anonymity requires that no adversary can learn the identity of the sender, *even if the receiver is corrupt.*

Definition 2 (Receiver Anonymity). *A protocol achieves* receiver anonymity *if it does not reveal any information concerning the identity of the receiver to any adversary beyond what could be legitimately learned by knowing for each corrupt participant whether or not he is the receiver.*

Note that in particular, if the sender or receiver is corrupt, a protocol vacuously achieves receiver anonymity.

Definition 3 (Full Anonymity). *A protocol achieves* full anonymity *if it does not reveal any information about the relation between the identity of the sender and receiver to any adversary beyond what could be legitimately learned by knowing for each corrupt participant whether or not he is the receiver.*

Note that full anonymity implies sender and receiver anonymity and that if the sender is corrupt, a protocol vacuously achieves full anonymity.
Remark. *The asymmetry between the definitions of sender and receiver anonymity stems from the fact that, contrary to the sender, the receiver does not know at the onset of the protocol that such a role will be imparted upon him.*

In what follows, we are only interested in protocols that are unconditionally secure in the information-theoretic sense for the purpose of achieving full anonymity. We place no limit on the number of corrupt participants. However, our protocol could abort if even a single corrupt participant deviates from the prescribed protocol. Even if the protocol aborts, full anonymity as well as message privacy are never compromised, except with exponentially small probability. Note that if we had some sort of guarantee that a strict majority of participants is honest, then anonymous quantum message transmission could be implemented as a special case of quantum secure multiparty computation [BCG$^+$06].

1.3 Anonymity in the Quantum World

The first protocol based on quantum mechanics that allows the anonymous communication of *classical* information was proposed by P. Oscar Boykin [Boy02]. In the case of a *quantum* message, Matthias Christandl and Stephanie Wehner were first to define the concept of *anonymous quantum message transmission* and to give an explicit protocol for solving this task [Weh04,CW05], but under the *deus ex machina* assumption that the n participants share ahead of time entangled state $|+_n\rangle = \frac{1}{\sqrt{2}}|0^n\rangle + \frac{1}{\sqrt{2}}|1^n\rangle$. (No mechanism is proposed to verify the validity of that state.) Under that assumption, their protocol is information-theoretically secure in terms of full anonymity, but malicious participants can alter the transmitted state in a way that will not be detected by the honest participants.

One key notion introduced in the paper of Christandl and Wehner is that of *anonymous entanglement*. Starting with the assumed n-party entangled state $|+_n\rangle$, the sender and the receiver end up sharing a two-party entangled state $|+_2\rangle$, better known as Bell State $|\Phi^+\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$, provided the other parties follow the protocol honestly. This entanglement is *anonymous* because the sender has chosen with which other party (the receiver) he shares it, but the receiver has no information concerning the party with which he is entangled. Moreover, the other parties have no information concerning who are the two entangled parties (assuming the entangled parties are not corrupt).

A first attempt to accomplish quantum message transmission in the presence of an unlimited number of corrupt participants *without* assuming that a trusted state $|+_n\rangle$ is shared between the participants before the onset of the protocol was made by Jan Bouda and Josef Šprojcar [BŠ07], but in a public-receiver model (the sender is anonymous but the receiver is public). The creation and distribution of a $|+_n\rangle$ state is an important part of their protocol. From there, they attempt to establish semi-anonymous entanglement (the identity of one of the entangled parties, the receiver, is public). However, careful analysis reveals that an active adversary can proceed in such a way that the probability that the protocol aborts becomes correlated with the identity of the sender, thus compromising his anonymity. If the protocol requires the receiver to stay quiet in order not to reveal whether or not the protocol has succeeded, it is true that the anonymity of the sender is preserved. However, this is very different from the model usually considered in secure multiparty computation, in which all the participants learn at the end of the protocol whether or not it has succeeded. More importantly, this approach makes it impossible to preserve the identity of the sender whenever the receiver is corrupt. Indeed, if we wanted to cope with a corrupt receiver and still preserve sender anonymity, this would require the need to hide from the receiver himself whether or not the protocol has succeeded. But if it were the case that the message itself (if received) did not convey any information on the success of the protocol, then it would mean that it is no more useful than a totally random state. Then, why bother send it at all?

Our own protocol is also based on the establishment of anonymous entanglement between the sender and the receiver. However, compared to the protocol of Christandl and Wehner, we do not need to assume an *a priori* shared $|+_n\rangle$ state and no malicious attempt at corrupting the intended final $|\Phi^+\rangle$ state between the sender and the receiver can succeed (except with exponentially small probability) without causing an abort. It follows that the intended state will be transmitted faithfully unless the protocol aborts, in which case it will end up intact at the sender's by virtue of fail-safe teleportation (unless the receiver is corrupt). Compared with the protocol of Bouda and Šprojcar, our receiver is anonymous and the identity of the sender and the receiver cannot be correlated with the probability that the protocol aborts, allowing us to achieve full anonymity according to Definition 3.

2 Toolbox

We now survey the classical and quantum tools that are used in our main protocol. Two of us recently developed several classical secure multiparty protocols [BT07]; we present below some of the relevant results, which will be used in the next section. All protocols assume pairwise authentic private classical channels and a broadcast channel. They offer information-theoretic security and have polynomial complexity in the number of participants as well as in a security parameter and, in the case of Theorem 4, in the number of bits in the transmitted message. In all cases, the expression "exponentially close to 1" or "exponentially small" means "exponentially in the security parameter". We also review a key result from [BCG$^+$02].

Theorem 1 (Logical OR–[BT07]). *There exists a secure multiparty protocol to compute the* logical OR *of the participants' input bits (one bit per participant). If all participants are honest, the correct answer is computed with probability exponentially close to 1. Misbehaving participants cannot cause the protocol to abort. (Any refusal to participate when expected will cause the output to be 1.) The only information an active adversary can learn through the protocol is if at least one honest participant has input 1. No information about the number of such participants or their identity is revealed.*

Theorem 2 (Collision Detection–[BT07]). *There exists a* collision detection *protocol in which each participant inputs a bit. Let r denote the number of 1s among these input bits. The protocol has three possible outcomes corresponding to whether $r = 0$, $r = 1$ or $r \geq 2$. If all participants are honest, the correct value is computed with probability exponentially close to 1. No participant can make the protocol abort, and an adversary cannot learn more than it could have learned by assigning to all corrupt participants the input 0 and letting them follow the protocol faithfully. A single corrupt participant can cause the output corresponding to $r \geq 2$ regardless of the other inputs (even if all the other inputs are 0). Also, it is possible for a corrupt participant to set his input to 0 if all other participants have input 0 (producing an $r = 0$ output) and to 1 otherwise (producing an $r \geq 2$ output). No other form of cheating is possible.*

Although the collision detection protocol outlined above may look rather imperfect, it is actually just as useful as the ideal protocol for our purpose.

Theorem 3 (Notification–[BT07]). *There exists a* notification *protocol in which participants can notify other participants of their choosing. Each player's output is one private bit specifying if he has been notified at least once; this value is correctly computed with probability exponentially close to 1. This is the only information accessible through the protocol even in the case of an active adversary.*

According to [BT07], it is possible in general to invoke the notification protocol even if multiple senders want to notify several receivers. However, in the specific context of our use of this protocol for the purpose of anonymous quantum message transmission, we forbid any honest participant to engage in the above notification protocol without having previously caused output "$r = 1$" in the collision detection protocol (Theorem 2). Similarly, no honest participant S will ever engage in the anonymous message transmission protocol below unless he has initially caused output "$r = 1$" in the collision detection protocol *and* has notified a single other participant R.

Theorem 4 (Anonymous Message Transmission–[BT07]). *There exists an anonymous message transmission protocol in which a sender can transmit a classical message to a receiver such that even in the presence of an active adversary, full anonymity is achieved and the privacy of the message is perfect. If all participants are honest then the message is transmitted perfectly. Any attempt by a corrupt participant to modify the message will cause the protocol to abort, except with exponentially small probability.*

In 2002, Howard Barnum, Claude Crépeau, Daniel Gottesman and Alain Tapp presented a non-interactive scheme for the authentication of quantum messages [BCG+02]. The protocol also encrypts the quantum state to be transmitted and is information-theoretically secure.

Theorem 5 (Quantum Authentication–[BCG+02]). *There exists an information-theoretically secure quantum authentication scheme to authenticate an arbitrary quantum message $|\psi\rangle$ of length m with an encoding circuit (called authenticate) and a decoding circuit (called decode) of size polynomial in m, which uses a random private key of length $2m + 2s + 1$ and has authenticated message of length $m + s$. Let p the probability that the message is accepted. If the message is accepted then let q be the probability of obtaining outcome $|\psi\rangle$ when measuring in a basis containing $|\psi\rangle$. If the authenticated message is not modified, then $p = q = 1$. Otherwise, $pq + (1-p) > 1 - \frac{m+s}{s(2^s+1)}$. The protocol also perfectly preserves the privacy of the transmitted message.*

3 Protocol for Anonymous Quantum Message Transmission

In this section, we describe and analyse our protocol for anonymous quantum message transmission. Our protocol allows an anonymous sender S to transmit an m-qubit message $|\psi\rangle$ to an anonymous receiver R. We assume a broadcast channel as well as an information-theoretically secure private and authenticated quantum channel between each pair of participants (which can also be used, of course, to transmit classical information). Our protocol achieves full anonymity and message privacy, except with exponentially small probability. The security proof for the protocol makes no assumption on the number of corrupt participants, but a single corrupt participant can make the protocol abort. However, if the sender and the receiver are honest, the quantum message to be transmitted will only be lost with exponentially small probability.

Here is an informal description of the protocol. In the first step, the purely classical collision detection protocol of Theorem 2 is performed to establish that exactly one participant wants to send an anonymous quantum message. If this is not the case, the protocol aborts. In case it is found that more than one participant wants to speak, one might imagine alternative scenarios such as asking each one of them to decide at random whether or not to skip their turn and trying again the collision detection protocol until a single-sender occurrence occurs. This will reveal information on the number of honest would-be senders and may take too many trials if there are too many of them, so that more sophisticated solutions might need to be considered. (Further elaboration on this issue would go beyond the scope of this paper.)

In the next two steps, the participants collaborate to establish multiple instances of a shared state $|+_n\rangle = \frac{1}{\sqrt{2}}|0^n\rangle + \frac{1}{\sqrt{2}}|1^n\rangle$. Then, the sender designates a receiver by use of the notification protocol (Theorem 3).

If honest, the receiver will act differently from the other participants, but in a way that is indistinguishable, so that his anonymity is preserved. The shared instances of $|+_n\rangle$ are then used to create anonymous entanglement between the sender and the receiver. However, the anonymous entanglement could be imperfect if other participants misbehave. For this reason, the sender then creates a sufficient number of instances of Bell state $|\Phi^+\rangle$. The possibly imperfect anonymous entanglement is used to teleport [BBC$^+$93] an authenticated version of half of each $|\Phi^+\rangle$. If this first teleportation is successful, the sender uses this newly established perfect anonymous entanglement to teleport the quantum message itself. Our fail-safe quantum teleportation protocol ensures that unless the receiver is corrupt, the quantum message is never destroyed, except with exponentially small probability: either it is safely transmitted to the receiver, or it comes back intact at the sender's.

In more detail, all classical communication from the sender to the receiver is performed anonymously using the anonymous message transmission protocol (Theorem 4). To create anonymous entanglement, all participants must be involved. One participant (who is chosen arbitrarily, for instance the first participant in lexicographic order) creates a state $|+_n\rangle$ and distributes one qubit to each participant, keeping one for himself. Of course, this participant could be corrupt, so that there is no guarantee that a proper $|+_n\rangle$ has been distributed. Moreover, a corrupt distributor could send different states to different honest participants, in the hope that the future evolution of the protocol may depend on who is the sender and who is the receiver. Foiling this threat constitutes a key contribution of our protocol. For this reason, all participants *verify* this state *without* destroying it in the next step. If the verification succeeds, the state shared amongst all participants is guaranteed to be invariant under permutation of the honest participants (Lemma 1), even though it could still not be a genuine $|+_n\rangle$ state. This ensures full anonymity. Furthermore, the behaviour of the state $|+_n\rangle$, when measured by all but two parties in the Hadamard basis, ensures correctness (unless it aborts) as shown in Theorems 6 and 8.

The full protocol is given as **Protocol 1**, where we denote by P the *conditional phase change* defined by $P|0\rangle = |0\rangle$ and $P|1\rangle = -|1\rangle$. Note that if two participants (such as the sender and the receiver) share an instance of Bell state $|\Phi^-\rangle = \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle$, a single participant (such as the sender) can convert this to a $|\Phi^+\rangle$ by locally applying the P operation. Note also that such a local operation (performed by the sender) has no detectable effect that could be measured by the other participants (in particular the receiver), which ensures that the anonymity of the sender is not compromised. It is easy to see that **Protocol 1** has polynomial complexity in n (the number of participants), s (the security parameter) and m (the length of the message).

Theorem 6 (Correctness). *Assume all participants are honest in* ***Protocol 1****. If more than one of them wishes to be a sender, this will be detected with probability exponentially close to 1 in the first step. Otherwise, the message is transmitted perfectly with probability exponentially close to 1, and the protocol can abort only with exponentially small probability.*

Proof. Even if all participants are honest, it is possible for collision detection or notification to produce an incorrect output (the notification protocol may also abort); however, this happens with exponentially small probability.

To ensure correctness of the protocol, we only have to verify that S and R share a sufficient number of proper Bell states $|\Phi^+\rangle$ at the end of step 5. It is clear

---

Protocol 1 Anonymous Quantum Message Transmission

Let s be the security parameter and m be the length of quantum message $|\psi\rangle$. All quantum communication is performed using the private authenticated quantum channels.

1. **Multiple Sender Detection**
   - 1.1 The collision detection protocol (Theorem 2) is used to determine if one and only one participant wants to be the sender. If not, the protocol aborts.
2. **Entanglement Distribution**
   - 2.1 One arbitrarily designated participant creates $2m+s$ instances of the state $|+_n\rangle$ and sends one qubit of each instance to each participant, keeping one qubit of each instance for himself.
3. **Entanglement Verification**
   For each of the $2m+s$ instances:
   - 3.1 Each participant makes $n-1$ *pseudo-copies* of his qubit by applying a control-not with it as the source and a qubit initialized to $|0\rangle$ as the target. One such pseudo-copy is sent to every other participant.
   - 3.2 Each participant verifies that all the n qubits in his possession are in the subspace spanned by $\{|0^n\rangle, |1^n\rangle\}$.
   - 3.3 Each participant broadcasts the outcome of the previous step. If any outcome is negative, the protocol aborts.
   - 3.4 Each participant *resets* $n-1$ of his qubits to $|0\rangle$ by performing $n-1$ control-not operations. These qubits are discarded and the one remaining is back to the state distributed at step 2.

Protocol 1 Anonymous Quantum Message Transmission (continued)

4. Receiver Notification

4.1 The participants execute the notification protocol (Theorem 3) in which only S notifies a single R.

5. Anonymous Entanglement Generation

For each of the $2m + s$ instances:

5.1 All participants except S and R measure in the Hadamard basis the qubit that remains from step 3.

5.2 Each participant broadcasts the result of his measurement (S and R broadcast two random dummy bits).

5.3 S computes the parity of all the bits received during the previous step (except his own and that of R).

5.4 If the parity is odd, S applies P, the conditional phase change, to his remaining qubit (the two qubits shared by S and R are now in Bell state $|\Phi^+\rangle$).

6. Perfect Anonymous Entanglement

6.1 S creates $2m$ instances of Bell state $|\Phi^+\rangle$. He keeps the first qubit of each pair; let ρ be the rest of the pairs.

6.2 S creates a random classical key k of length $4m + 2s + 1$, and computes $\rho' = \mathsf{authenticate}(\rho, k)$.

6.3 S performs a teleportation measurement on ρ' using the anonymous $|\Phi^+\rangle$ states generated during steps 2–5.

6.4 S uses the anonymous message transmission protocol (Theorem 4) to send k and the teleportation bits to R.

6.5 R completes the teleportation and computes $\rho = \mathsf{decode}(\rho', k)$. If the decoding is successful, S and R share perfect anonymous entanglement (they share $2m$ instances of $|\Phi^+\rangle$).

6.6 A logical OR is computed (Theorem 1): all players input 0 except R, who inputs 1 if the authentication failed and 0 otherwise. If the outcome is 1, the protocol aborts.

7. Fail-Safe Teleportation

7.1 S teleports the state $|\psi\rangle$ to R using the first m pairs generated in the previous step. The teleportation bits are anonymously transmitted to R (Theorem 4). If the communication succeeds, R terminates the teleportation.

7.2 A logical OR is performed (Theorem 1): all players input 0 except R, who inputs 1 if the communication of the teleportation bits failed. If the outcome is 0, the protocol succeeds. Otherwise, S and R do the following:

7.2.1 R performs a teleportation measurement using the remaining perfect anonymous entanglement to teleport back to S the quantum state resulting from partially failed step 7.1.

7.2.2 All participants broadcast $2m$ random bits, except R who broadcasts the teleportation bits from above. The protocol continues even if one of the participants refuses to broadcast.

7.2.3 S reconstructs $|\psi\rangle$ from his own teleportation bits from step 7.1 and R's teleportation bits received from the broadcast. The protocol aborts.

that at the end of step 3, the participants share proper instances of state $|+_n\rangle$ (since we are assuming in this theorem that they are honest). When S computes the parity of the measurement outcomes in step 5, this corresponds to the parity of the measurement results in the Hadamard basis of the state $|+_n\rangle$, where all but two qubits are measured. If the parity is even, S and R share $|\Phi^+\rangle$ and otherwise $|\Phi^-\rangle$, which is corrected by the sender by the application of the conditional phase change P. □

The following Lemma is necessary in the proof of anonymity and privacy (Theorem 7).

Lemma 1 (Invariance Under Permutation of Honest Participants). *In* ***Protocol 1****, if step 3 succeeds, then the state of the system at the end of the step is:*

$$\alpha|00\ldots0\rangle_H|\psi_0\rangle_C + \beta|11\ldots1\rangle_H|\psi_1\rangle_C\,, \tag{1}$$

where H denotes the honest participants' subsystem, C denotes the corrupt participants' subsystem, and $\alpha, \beta \in \mathbb{C}$ are such that $|\alpha|^2 + |\beta|^2 = 1$.

Proof. In the entanglement verification step, each honest participant sends a pseudo-copy of his state to every other honest participant. Therefore, after a single honest participant verifies that his qubits are in the subspace spanned by $\{|0^n\rangle, |1^n\rangle\}$, we are already ensured that if the entanglement verification succeeds, the state will be of the form given above. Note that the corrupt participants' subsystem C could span more than t qubits since they can bring arbitrary ancillas into their cheating strategy. □

Theorem 7 (Anonymity and Privacy). *Regardless of the number of corrupt participants and except with exponentially small probability,* ***Protocol 1*** *achieves full anonymity and privacy of the transmitted message $|\psi\rangle$.*

Proof. We analyse the protocol step by step in order to prove the statement.

By virtue of Theorem 2, step 1 does not compromise the identity of the sender, and it involves neither the receiver nor the quantum state to be transmitted. Steps 2 and 3 are done without any reference to S or R and thus cannot compromise their anonymity either. Furthermore, the state obtained at the end of step 3 (if it does not abort) cannot be specifically correlated with any honest participant even if some other participants are corrupt. More precisely, by Lemma 1, the state is *invariant under any permutation of the honest participants*. This is crucial for the anonymity and privacy of the rest of the protocol. In particular, it guarantees that the probability that the protocol aborts does not depend on the identity of S or R, or any relationship between them. We prove this below in the analysis of step 6.

The security of step 4 follows directly from the unconditional security of the notification protocol (Theorem 3). However, if S fails to notify R in step 4 (this happens with exponentially small probability), an adversary can surreptitiously take over the role of the honest receiver in the rest of the protocol without being detected. In that case, the adversary will violate the secrecy of the transmitted state, yet without compromising the sender and receiver anonymity beyond what can be learned by inspecting the illegitimately received state.

In step 5, anonymous entanglement is generated. No information is revealed to the adversary in this step since all communication is done by honest participants broadcasting random bits.

For step 6, all communication is done using the anonymous message transmission protocol, which is secure according to Theorem 4, except in logical OR computation at the end, which reveals the success or failure of the authentication part of the protocol. We now show that this last substep cannot reveal any information on the identity of S or R. This is because the success or failure of the authentication step is uncorrelated to the identity of S and R: by Lemma 1, as far as the qubits are concerned, all honest participants are identical under permutation. Thus the adversary has no strategy that would allow him to determine any information about the identity of S or R, or even about any relation between them.

During step 7, all the bits sent from S to R are randomly and uniformly distributed because they are the classical bits resulting from the teleportation protocol, therefore they do not reveal any information about the identity of S. A similar observation about the bits broadcast by R in the case that the very last part of the protocol is executed ensures that R and S remain anonymous.

The privacy of the state $|\psi\rangle$ in the case that S successfully notified R in step 4 (which happens with probability exponentially close to 1) is guaranteed by the basic properties of teleportation. □

Theorem 8 (Integrity). *At the end of* ***Protocol 1****, if R is honest then the state $|\psi\rangle$ is either in the possession of S or R, except with exponentially small probability. Furthermore, $|\psi\rangle$ can only stay with S if the protocol has aborted.*

Proof. If all participants are honest, then by Theorem 6, the state is in the possession of R except with exponentially small probability. Otherwise, the protocol might abort before step 7, in which case S still has $|\psi\rangle$. If the protocol reaches step 7, due to the quantum authentication of step 6, S and R share $2m$ perfect Bell states $|\Phi^+\rangle$ (with probability exponentially close to 1), which are used for teleportation in step 7. If the first step of the fail-safe teleportation fails, then S no longer has $|\psi\rangle$; however, the last three substeps of the protocol will always succeed and S will reconstruct $|\psi\rangle$ (provided R is honest). Furthermore, it follows from the virtues of teleportation that if the protocol does not abort, the state is no longer with S. □

The reason why we specify in Theorem 8 that R must be honest is that a corrupt R can destroy $|\psi\rangle$ by simply discarding it after having faithfully followed the entire protocol. There remains one subtlety to mention: a corrupt R could behave honestly until the last step. Then, he would input 1 in the logical OR computation to force S to accept the teleportation back of the state. At that point, the corrupt R could teleport back to S a fake state. As a result, S would be fooled into thinking he still has custody of the original quantum state when, in fact, that state is in the hands of R. (In general, there will be no way for S to know that this has happened.)

4 Conclusion and Discussion

We have presented the first information-theoretically secure protocol for quantum communication between an anonymous sender and an anonymous receiver that tolerates an arbitrary number of corrupt participants. In particular, this means that no adversary can learn any information that will break the anonymity of the sender or receiver. Our protocol also provides perfect privacy for the quantum message and ensures that the quantum message is never destroyed, except with exponentially small probability. The drawback of our protocol is that any participant can disrupt the protocol and make it abort.

5 Acknowledgements

We are grateful to Patrick Hayden and Flavien Serge Mani Onana for insightful discussions. G. B. is supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Research Chair program and the Canadian Institute for Advanced Research (CIFAR). A. B. is supported in part by scholarships from the Canadian Federation of University Women and the Fonds Québecois de Recherche sur la Nature et les Technologies (FQRNT). J. F. is supported by a Helmore Award. S. G. is supported by NSERC. A. T. is supported in part by CIFAR, FQRNT, the Mathematics of Information Technology and Complex Systems Network, and NSERC. Furthermore, we acknowledge the support of the Quantum*Works* Network and of the Institut transdisciplinaire d'informatique quantique (INTRIQ).

References

[BBC+93] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70:1895–1899, 1993.

[BCG+02] H. Barnum, C. Crépeau, D. Gottesman, A. Smith, and A. Tapp. Authentication of quantum messages. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS02)*, page 449, 2002.

[BCG+06] M. Ben-Or, C. Crépeau, D. Gottesman, A. Hassidim, and A. Smith. Secure multiparty quantum computation with (only) a strict honest majority. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006)*, pages 249–260, 2006.

[Boy97] J. Boyan. The Anonymizer: protecting user privacy on the Web. *Computer-Mediated Communication Magazine*, 4(9), 1997.

[Boy02] P. O. Boykin. *Information security and quantum mechanics: security of quantum protocols*. PhD thesis, University of California, Los Angeles, 2002.

[BŠ07] J. Bouda and J. Šprojcar. Anonymous transmission of quantum information. In *Proceedings of the First International Conference on Quantum, Nano, and Micro Technologies (ICQNM'07)*, 2007.

[BT07] A. Broadbent and A. Tapp. Information-theoretic security without an honest majority. In these *ASIACRYPT Proceedings*, 2007.

[Cha81] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24:84–88, 1981.

[Cha88] D. Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1:65–75, 1988.

[CW05] M. Christandl and S. Wehner. Quantum anonymous transmissions. In *Proceedings of the 11th International Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT 2005)*, pages 217–235, 2005.

[GGK+99] E. Gabber, P. B. Gibbons, D. M. Kristol, Y. Matias, and A. J. Mayer. Consistent, yet anonymous, Web access with LPWA. *Communications of the ACM*, 42(2):42–47, 1999.

[Gol04] O. Goldreich. *The Foundations of Cryptography — volume 2*. Cambridge University Press, 2004.

[RB89] T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority. In *Proceedings of the twenty-first annual ACM Symposium on Theory of Computing (STOC)*, pages 73–85, 1989.

[Weh04] S. Wehner. Quantum computation and privacy. Master's thesis, CWI Amsterdam, 2004.

[WZ82] W. K. Wootters and W. H. Żurek. A single quantum cannot be cloned. *Nature*, 299:802–803, 1982.