Authentication of Quantum Messages

Howard Barnum *, Claude Crépeau †, Daniel Gottesman ‡, Adam Smith §, and Alain Tapp ¶

Abstract

Authentication is a well-studied area of classical cryptography: a sender $\mathcal{A}$ and a receiver $\mathcal{B}$ sharing a classical secret key want to exchange a classical message with the guarantee that the message has not been modified or replaced by a dishonest party with control of the communication line. In this paper we study the authentication of messages composed of quantum states.

We give a formal definition of authentication in the quantum setting. Assuming $\mathcal{A}$ and $\mathcal{B}$ have access to an insecure quantum channel and share a secret, classical random key, we provide a non-interactive *scheme that enables $\mathcal{A}$ to both encrypt and authenticate an m qubit message by encoding it into $m + s$ qubits, where the error probability decreases exponentially in the security parameter s. The scheme requires a secret key of size $2m + O(s)$. To achieve this, we give a highly efficient protocol for testing the purity of shared EPR pairs.*

It has long been known that learning information about a general quantum state will necessarily disturb it. We refine this result to show that such a disturbance can be done with few side effects, allowing it to circumvent cryptographic protections. Consequently, any scheme to authenticate quantum messages must also encrypt them. In contrast, no such constraint exists classically.

This reasoning has two important consequences: It allows us to give a lower bound of $2m$ key bits for authenticating m qubits, which makes our protocol asymptotically optimal. Moreover, we use it to show that digitally signing quantum states is impossible.

Keywords. Authentication, quantum information.

* CCS-3 Group, Los Alamos National Laboratories, Los Alamos, New Mexico 87554 USA. e-mail: `barnum@lanl.gov`.

† School of Computer Science, McGill University, Montréal (Qc), Canada H3A 2A7. e-mail: `crepeau@cs.mcgill.ca`.

‡ UC Berkeley, EECS: Computer Science Division, Berkeley, CA 94720, USA. e-mail: `gottesma@eecs.berkeley.edu`. Supported by the Clay Mathematics Institute.

§ M.I.T., Lab. for Computer Science, Cambridge MA 02139, USA. e-mail: `asmith@theory.lcs.mit.edu`. Supported in part by U.S. Army Research Office Grant DAAD19-00-1-0177.

¶ Département IRO, Université de Montréal, Montréal (Qc), Canada H3C 3J7. e-mail: `tappa@iro.umontreal.ca`.

1. Introduction

Until recently, the expression "quantum cryptography" referred mostly to quantum key distribution protocols [3]. However, these words now refer to a larger set of problems. While QKD and many other quantum protocols attempt to provide improved security for tasks involving classical information, an emerging area of quantum cryptography attempts instead to create secure protocols for tasks involving *quantum* information. One standard cryptographic task is the *authentication* of messages: $\mathcal{A}$ transmits some information to $\mathcal{B}$ over an insecure channel, and they wish to be sure that it has not been tampered with *en route*. When the message is classical, and $\mathcal{A}$ and $\mathcal{B}$ share a random secret key, this problem can be solved by, for instance, the Wegman-Carter scheme [7]. In this paper, we discuss the analogous question for quantum messages.

If we assume $\mathcal{A}$ and $\mathcal{B}$ share a secret *quantum* key in the form of m EPR pairs, and a secret classical key, there is a straightforward solution to the problem: $\mathcal{A}$ teleports her message to $\mathcal{B}$, authenticating the $2m$ classical bits transmitted in the quantum teleportation protocol [4]. If $\mathcal{A}$ and $\mathcal{B}$ initially share only a classical key, however, the task is more difficult. We start with a simple approach: first distribute EPR pairs (which might get corrupted in transit), and then use entanglement purification [5] to establish clean pairs for teleportation. However, we do not need entanglement *purification*, which produces good EPR pairs even if the channel is noisy; instead we only need a *purity testing protocol*, which checks that EPR pairs are correct, but does not repair erroneous ones.

Unfortunately, any such protocol will have to be interactive, since $\mathcal{A}$ must first send some qubits to $\mathcal{B}$ and then wait for confirmation of receipt before completing the transmission. This is unsuitable for situations in which a message is stored and must be checked for authenticity at a later time. Also, this interactive protocol achieves something stronger than what is required of a quantum authentication scheme: at the end of the purity-testing based scheme, *both* $\mathcal{A}$ and $\mathcal{B}$ know that the transmission was successful, whereas for authentication, we only require that $\mathcal{B}$ knows.

Contributions In this paper we study *non-interactive* quantum authentication schemes with classical keys. Our primary contributions are:

Formal definition of authentication for quantum states – In classical authentication, one simply limits the probability that the adversary can make *any* change to the state without detection. This condition is too stringent for quantum information, where we only require high fidelity to the original state. We state our definition in terms of the transmission of pure states (section 3). In the full version, we also show that the same definition implies security for mixed or entangled states.

Construction of efficient purity testing protocols – We show how to create purity-testing protocols using families of quantum error-correcting codes with a particular covering property, namely that any Pauli error is detected by most of the codes in the family. We construct an efficient such family based on projective geometry, yielding a purity-testing protocol requiring only $O(s)$ (classical) bits of communication, where s is the security parameter (section 4).

Purity-testing codes have not explicitly appeared before in the literature, but have been present implicitly in earlier work, for instance [10, 12]. To prove our purity-testing protocols secure, we use a "quantum-to-classical" reduction, due to Lo and Chau [10]. Subsequently to our work, Ambainis, Smith, and Yang [2] used our construction in a study of more general entanglement extraction procedures.

*Construction of non-interactive quantum authentication schemes (*QAS*)* – We show that a secure non-interactive QAS can be constructed from any purity-testing protocol derived, as above, from QECCs (section 5). In particular, for our family of codes, we obtain an authentication scheme which requires sending $m + s$ qubits, and consuming $2m + O(s)$ bits of classical key for a message of m qubits. The proof techniques in the Shor and Preskill paper [12] serve as inspiration for the transformation from an interactive purity-testing protocol to a non-interactive QAS.

A relation between encryption and authentication – One feature of our authentication protocol is that it completely encrypts the quantum message being sent. We show that this is necessary in *any* QAS(section 6), in striking contrast to the situation for classical information, where common authentication schemes leave the message completely intelligible. It therefore follows that any authentication protocol for an m-qubit message must use nearly $2m$ bits of classical key, enough to encrypt the message. The protocol we present approaches this bound asymptotically.

Impossibility of digitally signing quantum states – Since authentication requires encryption, it is impossible to create digital signature schemes for quantum messages: any protocol which allows one recipient to read a message also allows him or her to modify it without risk of detection, and therefore all potential recipients of an authenticated message must be trustworthy (section 7). This conclusion holds true even if we require only computationally secure digital signatures. Note that this does not in any way preclude the possibility of signing *classical* messages with or without quantum states [9].

Why should we prefer a scheme with classical keys to a scheme with entangled quantum keys? The task of authenticating quantum data is only useful in a scenario where quantum information can be reliably stored, manipulated, and transmitted over communication lines, so it would not be unreasonable to assume quantum keys. However, many manipulations are easier with classical keys. Certainly, the technology for storing and manipulating them is already available, but there are additional advantages. Consider, for example, public key cryptography; it is possible to sign and encrypt classical key bits with public key systems, but signing a general quantum state is impossible. Thus, quantum keys would be unsuitable for an asymmetric quantum authentication scheme such as the one we describe in section 7.1.

2. Preliminaries

Classical Authentication In the classical setting, an authentication scheme is defined by a pair of functions $A : \mathcal{K} \times M \to C$ and $B : \mathcal{K} \times C \to M \times \{\text{valid, invalid}\}$ such that for any message $\mu \in M$ and key $k \in \mathcal{K}$ we have *completeness*

$$B_k(A_k(\mu)) = \langle \mu, \text{valid} \rangle$$

and that for any opponent O, we have *soundness*

$$\text{Prob}\,\{B_k(O(A_k(\mu))) \in \{\langle \mu', \text{valid}\rangle, \mu' \neq \mu\}\} \geq 1 - 2^{-\Omega(t)}$$

where $t = \lg \#C - \lg \#M$ is the security parameter creating the tradeoff between the expansion of the messages and the security level. Note that we only consider information-theoretically secure schemes, not schemes that are based on computational assumptions.

Wegman and Carter [7] introduced several constructions for such schemes; their most efficient uses keys of size only $4(t + \lg \lg m) \lg m$ and achieves security $1 - 2^{-t+2}$. This compares rather well to the known lower bound of $t + \lg m - \lg t$ for such a result [7]. The same work also introduced a technique to re-use an authentication function several times by using one-time-pad encryption on the tag, so that an opponent cannot learn *anything* about the particular key being used by $\mathcal{A}$ and $\mathcal{B}$. Thus, at a marginal cost of only t secret key bits per authentication, the confidentiality of the authentication function h is guaranteed and thus may be re-used (a polynomial number of times).

For the remainder of this paper, we assume the reader is familiar with the basic notions and notation of quantum computing (see textbooks such as [11]).

Quantum Stabilizer Codes A quantum error-correcting code (QECC) is a way of encoding quantum data (say m qubits) into n qubits ($m < n$). Usually the goal in the construction of codes is give it the ability to correct errors on as many qubits as possible. However, in this paper, we use the theory developed for those purposes to construct families of codes with a slightly different type of property. For now, we review the necessary theory on a very general class of codes known as *stabilizer codes*.

A basis for the set of all operators on a qubit is the "Pauli" error basis, defined via $E_{ab} = X^a Z^b$, where $\langle i|X|j\rangle = \delta_{i,j+1}$, $\langle i|Z|j\rangle = (-1)^i \delta_{i,j}$. We define the error group E to be the tensor product of n copies of the single-qubit error group $\{(-1)^c E_{ab}\}$. Each element of E corresponds to a $2n$-dimensional vector, and the vectors $x = (\mathbf{a}|\mathbf{b})$, $y = (\mathbf{a}'|\mathbf{b}')$ come from commuting operators iff their symplectic inner product is 0 in $\mathbb{Z}_2$:

$$E_x E_y = E_y E_x \quad \iff \quad B(x, y) = \mathbf{a}'\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{b}' = 0. \tag{1}$$

A *stabilizer code* is a QECC given by an Abelian subgroup S of E which does not contain any multiples of the identity other than I itself. S can be described by the set of $2n$-dimensional vectors x such that $E_x \in S$. This will be a subspace of $\mathbb{Z}_2^{2n}$. Moreover, it will be *totally isotropic*, i.e. $B(x, y) = 0$ for all x, y in the subspace. If we take a set of generators for S, we can divide Hilbert space into a set of equidimensional orthogonal subspaces. Each such space T consists of common eigenvectors of all operators of S having a fixed pattern of eigenvalues, unique to T. The space with all eigenvalues $+1$ is the "code space," its elements are "codewords," and the orthogonal spaces are labelled by "syndromes."

Note that one can also view $B(\cdot,\cdot)$ as a symplectic form over $GF(2^{2n})$, by choosing a set of generators for $GF(2^{2n})$ as a vector space over $\mathbb{Z}_2$. By choosing different sets of generators for $GF(2^{2n})$ as a vector space over $\mathbb{Z}_2$, we can get different symplectic forms $B(\cdot,\cdot)$ over this finite vector space. By judicious choice of the generators, one can make $B(\cdot,\cdot)$ correspond to *any* non-degenerate symplectic form over $GF(2^{2n})$.

Undetectable errors We can classify errors which lie in E into three categories: The errors corresponding to elements of Q are not truly errors—they leave the codewords unchanged. Errors which fail to commute with some element of Q move codewords into a subspace orthogonal to the code, so can be detected by the QECC. The remaining errors, those which commute with all elements in S but are not themselves in S, are the undetectable errors of the code. Thus, if $Q^\perp$ is the space of vectors y for which $B(x, y) = 0$ for all $x \in Q$, the set of undetectable errors is just $Q^\perp - Q$.

Purification and purity testing Quantum error-correcting codes may be used for *entanglement purification* ([5]). In this setting, $\mathcal{A}$ and $\mathcal{B}$ share some Bell states (say $|\Phi^+\rangle = |00\rangle + |11\rangle$) which have been corrupted by transmission through a noisy quantum channel. They want a protocol which processes these imperfect EPR pairs and produces a smaller number of higher-quality pairs. We assume that $\mathcal{A}$ and $\mathcal{B}$ have access to an authenticated, public classical channel. At the end of the protocol, they either accept or reject based on any inconsistencies they have observed. As long as $\mathcal{A}$ and $\mathcal{B}$ have a noticeable probability of accepting, then conditioned on accepting, the state they share should have fidelity almost 1 to the pure state $|\Phi^+\rangle^{\otimes m}$. Moreover, small amounts of noise in their initial shared state should not cause failure of the protocol.

Stabilizer codes can be particularly useful for purification because of the following observation: for any stabilizer code Q, if we measure the syndrome of one half of a set of Bell states $|\Phi^+\rangle^{\otimes n}$ and obtain the result y, then the result is the state $|\Phi^+\rangle^{\otimes m}$, with each of its two halves encoded in the coset with syndrome y. (Moreover, in this case the distribution on y is uniform.) If the original state is erroneous, $\mathcal{A}$ and $\mathcal{B}$ will likely find different syndromes, which will differ by the syndrome associated with the actual error.

Most purification protocols based on stabilizer codes require efficient error correction; we measure the syndrome, and use that information to efficiently restore the encoded state. However, one can imagine a weaker task in which $\mathcal{A}$ and $\mathcal{B}$ only want to *test* their EPR pairs for purity, i.e. they want a guarantee that if their pairs pass the test, their shared state will probably be close to $|\Phi^+\rangle^{\otimes m}$. In that case, we can use the code for error detection, not correction, and need only be able to encode and decode efficiently from the space Q.

Encryption of Quantum Messages A useful ingredient for much recent work in quantum cryptography is the concept of quantum teleportation, put forward by Bennett et al [4]. After $\mathcal{A}$ and $\mathcal{B}$ have shared a singlet state, $\mathcal{A}$ can later secretly send a single qubit in an arbitrary quantum state ρ to $\mathcal{B}$ by measuring her half of the singlet state together with her state ρ in the Bell basis to get two classical bits b_0, b_1. As a result, $\mathcal{B}$'s half of the singlet state will become one of four possibilities $\rho' := \sigma_z^{b_0}\sigma_x^{b_1}\rho\sigma_x^{b_1}\sigma_z^{b_0}$. If $\mathcal{A}$ sends b_0, b_1, then $\mathcal{B}$ can easily recover ρ.

Now without the bits b_0, b_1, the state ρ' reveals no information about ρ. Thus, one can turn this into an encryption scheme which uses only a classical key: after $\mathcal{A}$ and $\mathcal{B}$ have secretly shared two classical bits b_0, b_1, $\mathcal{A}$ can later secretly send a single qubit in an arbitrary quantum state ρ to $\mathcal{B}$ by sending a qubit in state ρ' as above. This is called a quantum one-time pad (QOTP). This scheme is optimal [1]: any quantum encryption must use 2 bits of classical key for every transmitted qubit.

3. Quantum Authentication

At an intuitive level, a quantum authentication scheme is a keyed system which allows $\mathcal{A}$ to send a state ρ to $\mathcal{B}$ with a guarantee: if $\mathcal{B}$ accepts the received state as "good", the fidelity of that state to ρ is almost 1. Moreover, if the adversary makes no changes, $\mathcal{B}$ should always accept, and the fidelity should be exactly 1.

Of course, this informal definition is impossible to attain. The adversary might always replace $\mathcal{A}$'s transmitted message with a completely mixed state. There would be a small probability that $\mathcal{B}$ would accept, and in that case the fidelity of the received state to $\mathcal{A}$'s initial state would be very low.

The problem comes from conditioning on $\mathcal{B}$'s acceptance of the received state; this causes trouble if the adversary's a priori chances of cheating are high. A more reasonable definition requires a tradeoff between $\mathcal{B}$'s chances of accepting, and the expected fidelity of the received system to $\mathcal{A}$'s initial state given his acceptance: as $\mathcal{B}$'s chance of accepting increases, so should the expected fidelity.

There is no reason to use the languages of both probability and fidelity here: for classical tests, fidelity and probability of acceptance coincide. With this in mind we first define what constitutes a quantum authentication scheme, and then give a definition of security:

Definition 1 *A* quantum authentication scheme *(*QAS*) is a pair of polynomial time quantum algorithms A and B together with a set of* classical *keys $\mathcal{K}$ such that:*

- *A takes as input an m-qubit message system M and a key $k \in \mathcal{K}$ and outputs a transmitted system T of $m+t$ qubits.*
- *B takes as input the (possibly altered) transmitted system T' and a classical key $k \in \mathcal{K}$ and outputs two systems: a m-qubit message state M, and a single qubit V which indicates acceptance or rejection. The classical basis states of V are called $|\text{ACC}\rangle, |\text{REJ}\rangle$ by convention.*

For any fixed key k, we denote the corresponding super-operators by A_k and B_k.

Note that $\mathcal{B}$ may well have measured the qubit V to see whether or not the transmission was accepted or rejected. Nonetheless, we think of V as a qubit rather than a classical bit since it will allow us to describe the joint state of the two systems M, V with a density matrix.

There are two conditions which should be met by a quantum authentication protocol. On the one hand, in the absence of intervention, the received state should be the same as the initial state and $\mathcal{B}$ should accept.

On the other hand, we want that when the adversary does intervene, $\mathcal{B}$'s output systems have high fidelity to the statement "either $\mathcal{B}$rejects or his received state is the same as that sent by $\mathcal{A}$". One difficulty with this is that it is not clear what is meant by "the same state" when $\mathcal{A}$'s input is a mixed state. It turns out that it is sufficient to define security in terms of pure states; one can deduce an appropriate statement about fidelity of mixed states.

Given a pure state $|\psi\rangle \in \mathcal{H}_M$, consider the following test on the joint system M, V: output a 1 if the first m qubits are in state $|\psi\rangle$ *or* if the last qubit is in state $|\text{REJ}\rangle$ (otherwise, output a 0). The projectors corresponding to this measurement are

$$\begin{aligned} P_1^{|\psi\rangle} &= |\psi\rangle\langle\psi| \otimes I_V \,+\, I_M \otimes |\text{REJ}\rangle\langle\text{REJ}| \\ &\qquad - |\psi\rangle\langle\psi| \otimes |\text{REJ}\rangle\langle\text{REJ}| \\ P_0^{|\psi\rangle} &= (I_M - |\psi\rangle\langle\psi|) \otimes (|\text{ACC}\rangle\langle\text{ACC}|) \end{aligned}$$

We want that for all possible input states $|\psi\rangle$ and for all possible interventions by the adversary, the expected fidelity of $\mathcal{B}$'s output to the space defined by $P_1^{|\psi\rangle}$ is high. This is captured in the following definition of security.

Definition 2 *A* QAS *is secure with error ϵ for a state $|\psi\rangle$ if it satisfies:*

Completeness: *For all keys $k \in \mathcal{K}$: $B_k(A_k(|\psi\rangle\langle\psi|)) = |\psi\rangle\langle\psi| \otimes |\text{ACC}\rangle\langle\text{ACC}|$*

Soundness: *For all super-operators $\mathcal{O}$, let ρ_{Bob} be the state output by $\mathcal{B}$ when the adversary's intervention[1] is characterized by $\mathcal{O}$, that is:*

$$\rho_{Bob} = \mathbb{E}_k\Big[B_k(\mathcal{O}(A_k(|\psi\rangle\langle\psi|)))\Big] = \frac{1}{|\mathcal{K}|}\sum_k B_k(\mathcal{O}(A_k(|\psi\rangle\langle\psi|)))$$

[1] We make no assumptions on the running time of the adversary.

where " $\mathbb{E}_k$*" means the expectation when* k *is chosen uniformly at random from* $\mathcal{K}$*. The* QAS *has soundness error* ϵ *for* $|\psi\rangle$ *if:*

$$\mathrm{Tr}\left(P_1^{|\psi\rangle}\rho_{Bob}\right) \geq 1-\epsilon$$

A QAS *is secure with error* ϵ *if it is secure with error* ϵ *for all states* $|\psi\rangle$*.*

Note that our definition of completeness assumes that the channel connecting $\mathcal{A}$ to $\mathcal{B}$ is noiseless in the absence of the adversary's intervention. This is in fact not a significant problem, as we can simulate a noiseless channel using standard quantum error correction.

Interactive protocols In the previous section, we dealt only with non-interactive quantum authentication schemes, since that is both the most natural notion, and the one we achieve in this paper. However, there is no reason to rule out interactive protocols. The definitions of completeness and soundness extend naturally: as before, $\mathcal{B}$'s final output is a pair of systems M, V, where the state space of V is spanned by $|\text{ACC}\rangle, |\text{REJ}\rangle$. In that case ρ_{Bob} is $\mathcal{B}$'s density matrix at the end of the protocol, averaged over all possible choices of shared secret key and executions of the protocol. The soundness error is ϵ, where $\mathrm{Tr}\left(P_1^{|\psi\rangle}\rho_{Bob}\right) \geq 1-\epsilon$.

4. Purity Testing Codes

An important tool in our proof is the notion of a *purity testing code*, which is a way for $\mathcal{A}$ and $\mathcal{B}$ to ensure that they share (almost) perfect EPR pairs. We shall concentrate on purity testing codes based on stabilizer QECCs.

Definition 3 *A stabilizer purity testing code (SPTC) with error* ϵ *is a set of stabilizer codes* $\{Q_k\}$*, for* $k \in \mathcal{K}$*, such that* $\forall\, E_x \in E$ *with* $x \neq 0$*,* $\#\{k | x \in Q_k^\perp - Q_k\} \leq (\#\mathcal{K})\epsilon$*.*

That is, for any error x in the error group, if k is chosen later at random, the probability that the code Q_k detects x is at least $1-\epsilon$.

Definition 4 *A purity testing protocol with error* ϵ *is a superoperator* $\mathcal{T}$ *which can be implemented with local operations and classical communication, and which maps* $2n$ *qubits (half held by* $\mathcal{A}$ *and half held by* $\mathcal{B}$*) to* $2m+1$ *qubits and satisfies the following two conditions:*

Completeness: $\mathcal{T}(|\Phi^+\rangle^{\otimes n}) = |\Phi^+\rangle^{\otimes m} \otimes |\text{ACC}\rangle$

Soundness: *Let* P *be the projection on the subspace spanned by* $|\Phi^+\rangle^{\otimes m} \otimes |\text{ACC}\rangle$ *and* $|\psi\rangle \otimes |\text{REJ}\rangle$ *for all* $|\psi\rangle$*. Then* $\mathcal{T}$ *satisfies the soundness condition if for all* ρ*,* $\mathrm{Tr}\left(P\mathcal{T}(\rho)\right) \geq 1-\epsilon$*.*

The obvious way of constructing a purity testing protocol $\mathcal{T}$ is to start with a purity testing code $\{Q_k\}$. When $\mathcal{A}$ and $\mathcal{B}$ are given the state ρ, $\mathcal{A}$ chooses a random $k \in \mathcal{K}$ and tells it to $\mathcal{B}$. They both measure the syndrome of Q_k and compare. If the syndromes are the same, they accept and perform the decoding procedure for Q_k; otherwise they reject.

Proposition 1 *If the purity testing code* $\{Q_k\}$ *has error* ϵ*, then* $\mathcal{T}$ *is a purity testing protocol with error* ϵ*.*

Proof: If $\mathcal{A}$ and $\mathcal{B}$ are given n EPR pairs, this procedure will always accept, and the output will always be $|\Phi^+\rangle^{\otimes m}$. Thus, $\mathcal{T}$ satisfies the completeness condition.

Suppose for the moment that the input state is $(E_x \otimes I)|\Phi^+\rangle^{\otimes n}$, for $E_x \in E$, $x \neq 0$. Then when k is chosen at random, there is only probability ϵ that $x \in Q_k^\perp - Q_k$. If $x \notin Q_k^\perp$, then $\mathcal{A}$ and $\mathcal{B}$ will find different error syndromes, and therefore reject the state. If $x \in Q_k^\perp$, then $\mathcal{A}$ and $\mathcal{B}$ will accept the state, but if $x \in Q_k$, then the output state will be $|\Phi^+\rangle^{\otimes m}$. Thus, the probability that $\mathcal{A}$ and $\mathcal{B}$ accept an incorrect state is at most ϵ.

To prove the soundness condition, we can use this fact and a technique of Lo and Chau [10]. The states $(E_x \otimes I)|\Phi^+\rangle^{\otimes n}$ form the Bell basis for the Hilbert space of $\mathcal{A}$ and $\mathcal{B}$. Suppose a nonlocal third party first measured the input state ρ in the Bell basis; call this measurement B. Then the argument of the previous paragraph would apply to show the soundness condition. In fact, it would be sufficient if $\mathcal{A}$ and $\mathcal{B}$ used the nonlocal measurement $Q_k \otimes Q_k$ which compares the Q_k-syndromes for $\mathcal{A}$ and $\mathcal{B}$without measuring them precisely. This is a submeasurement of the Bell measurement B — that is, it gives no additional information about the state. Therefore it commutes with B, so the sequence B followed by $Q_k \otimes Q_k$ is the same as $Q_k \otimes Q_k$ followed by B, which gives probability at least $1-\epsilon$ of success for general input states ρ. But if the state after $Q_k \otimes Q_k$ gives, from a Bell measurement, $|\Phi^+\rangle^{\otimes m}$ or $|\text{REJ}\rangle$ with probability $1-\epsilon$, then the state itself must have fidelity $1-\epsilon$ to the projection P. Thus, the measurement $Q_k \otimes Q_k$ without B satisfies the soundness condition. Moreover, $\mathcal{A}$ and $\mathcal{B}$'s actual procedure $\mathcal{T}$ is a refinement of $Q_k \otimes Q_k$—that is, it gathers strictly more information. Thus, it also satisfies the soundness condition, and $\mathcal{T}$ is a purity testing protocol with error ϵ. □

An Efficient Purity Testing Code Now we give an example of a particularly efficient purity testing code. To do so, we apply the techniques of stabilizer quantum error-correcting codes. We build a set of codes Q_k each encoding $m = (r-1)s$ qubits in $n = rs$ qubits, and show that the Q_k form a SPTC. (Note that the construction below works for registers with dimension equal to any prime power.) Using qubits in groups of s allows us to view our field $GF(2^{2rs})$ as both a $2r$-dimensional vector space over $GF(2^s)$ and a $2rs$-dimensional binary vector space. We need a symplectic form that is compatible with this decomposition. One possibility is

$$B(x, y) := \mathrm{Tr}\,(xy^{2^{rs}}), \tag{2}$$

where $\mathrm{Tr}(z) = \sum_{i=0}^{2rs-1} z^{2^i}$ is the standard trace function, which maps $GF(2^{2rs})$ onto $GF(2)$.

We consider a *normal rational curve* in $PG(2r-1, 2^s)$ (the projective geometry whose points are the 1-d subspaces of the $2r$-dimensional vector space over $GF(2^s)$). (See, e.g., the excellent introductory text [6].) Such a curve is given by:

$$\Upsilon = \{[1 : y : y^2 : \cdots : y^{2r-1}] : y \in GF(2^s)\} \\ \cup \{[0 : 0 : 0 : \cdots : 1]\}$$

Thus, there are $2^s + 1$ points on the curve.

Since each "point" of this curve is actually a one-dimensional subspace over $GF(2^s)$, it can also be considered as an s-dimensional binary subspace Q_k in a vector space of dimension $2rs = 2n$. We show that Q_k is totally isotropic with respect to the symplectic inner product (2), and encodes $m = n - s$ qubits in n qubits.

Theorem 2 *The set of codes Q_k form a SPTC with error*

$$\epsilon = \frac{2r}{2^s + 1}\ . \tag{3}$$

Each Q_k encodes $m = (r-1)s$ qubits in $n = rs$ qubits.

Proof: We must show (a) that Q_k is totally isotropic, and (b) that the error probability is at most ϵ.

(a) Follows from a straightforward calculation.

(b) We must find, for an arbitrary error E_x (which can be described via a $2n$-dimensional $GF(q)$ vector x), an upper bound on the number of $Q_k^\perp - Q_k$ it can belong to. It will be sufficient to bound the number of $Q_k^\perp$ the error can belong to, since $|Q_k|$ is small compared to $|Q_k^\perp|$ in our context. $x \in Q_k^\perp$ means $B(x, y) = 0$ for all $y \in Q_k$. By choice of s linearly independent $y \in Q_k$ this imposes s linearly independent linear equations on x. We will show below that if we take any $2r$ codes Q_k defined by points on Υ, and take s independent vectors from each, the resulting set of $2rs$ vectors is linearly independent. Thus if E_x is undetectable in $2r$ such codes, this imposes the dimension's worth ($2rs$) of linearly independent equations on x. Consequently, E_x must be detectable in all the remaining codes, i.e., E_x can satisfy $x \in Q_k^\perp$ for *at most* $2r$ values of k, when Q_k are chosen among the $q^s + 1$ available s-dimensional spaces corresponding to points on Υ. Thus, the $\{Q_k\}$ form a purity testing code with error

$$\epsilon \leq \frac{2r}{q^s + 1}\ . \tag{4}$$

We now show the claimed property of codes defined by Υ. A set of points in a projective geometry of dimension $d - 1$ are said to be in general position if any d (= dimension of the underlying vector space, when, as in our case, such exists) of them are linearly independent. The points on the normal rational curve Υ are in general position, and indeed a maximal set of such points. (To verify that they are in general position one shows that for any $2r$ points on the curve, the determinant of the matrix of their coordinates is nonzero; these are Vandermonde determinants.) That is, any $2r$ points on Υ are linearly independent. Each point k on Υ corresponds to an s-dimensional code Q_k, consisting of $2rs$-dimensional vectors. Let z be any nonzero element of Q_k. As α ranges over $GF(q^s)$, αz ranges over all vectors in Q_k. Thus, if any vector from Q_k is a linear combination of vectors from other codes $\{Q_j\}$, than all of Q_k is also a linear combination of vectors from $\{Q_j\}$, and k is linearly dependent on the points $\{j\}$ of Υ. So if we take any $2r$ codes Q_k, and take s independent vectors from each, the resulting set of $2rs$ vectors is linearly independent. □

Note that the above efficient SPTC also yields a corresponding efficient QKD protocol.

5. Protocols

In this section we describe a secure non-interactive quantum authentication scheme (Protocol 5.2) which satisfies the definition of section 3.

In order to prove our scheme secure, we begin with a purity testing protocol as per Section 4 (summarized as Protocol 5.1). The security of this protocol follows from Prop. 1. We then perform several transformations to the protocol that strictly preserve its security and goals but which remove the interaction, replacing it with a shared secret key. We thus obtain two less interactive intermediate protocols and a final protocol (Protocol 5.2), which is completely non-interactive. The transformations are similar in flavor to those of Shor and Preskill [12], who use the technique to obtain a simple proof of the security of a completely different task, namely the BB84 [3] quantum key exchange scheme.

Following the notation of Section 4, let P be the projector onto the subspace described by "either $\mathcal{B}$ has aborted or the joint state held by $\mathcal{A}$ and $\mathcal{B}$ is $|\Phi^+\rangle^{\otimes m}$ ".

Protocol 5.1 (Purity Testing Based Protocol)

1: *$\mathcal{A}$ and $\mathcal{B}$ agree on some SPTC $\{Q_k\}$.*

2: *$\mathcal{A}$ generates $2n$ qubits in state $|\Phi^+\rangle^{\otimes n}$. $\mathcal{A}$ sends the first half of each $|\Phi^+\rangle$ state to $\mathcal{B}$.*

3: *$\mathcal{B}$ announces that he has received the n qubits.*

4: *$\mathcal{A}$ picks a random $k \in \mathcal{K}$, and announces it to $\mathcal{B}$.*

5: *$\mathcal{A}$ and $\mathcal{B}$ measure the syndrome of the stabilizer code Q_k. $\mathcal{A}$ announces her results to $\mathcal{B}$ who compares them to his own results. If any error is detected, $\mathcal{B}$ aborts.*

6: *$\mathcal{A}$ and $\mathcal{B}$ decode their n-qubit words according to Q_k. Each is left with m qubits, which together should be nearly in state $|\Phi^+\rangle^{\otimes m}$.*

7: *$\mathcal{A}$ uses her half of $|\Phi^+\rangle^{\otimes m}$ to teleport an arbitrary m-qubit state ρ to $\mathcal{B}$.*

Let ρ_{AB} be the joint density matrix of $\mathcal{A}$ and $\mathcal{B}$'s systems. Then Prop. 1 states that at the end of step 6, $\mathrm{Tr}(P\rho_{AB})$ is exponentially close to 1 in n. The soundness of our first protocol follows immediately:

Corollary 3 *If $\mathcal{A}$ and $\mathcal{B}$ are connected by an authenticated classical channel, then Protocol 5.1 is a secure* interactive *quantum authentication protocol, with soundness error exponentially small in n.*

Protocol 5.2 (Non-interactive authentication)

1: *Preprocessing: $\mathcal{A}$ and $\mathcal{B}$ agree on some SPTC $\{Q_k\}$ and some secret and random binary strings k, x, and y.*

2: *$\mathcal{A}$ q-encrypts ρ as τ using key x. $\mathcal{A}$ encodes τ according to Q_k for the code Q_k with syndrome y to produce σ. $\mathcal{A}$ sends the result to $\mathcal{B}$.*

3: *$\mathcal{B}$ receives the n qubits. Denote the received state by σ'. $\mathcal{B}$ measures the syndrome y' of the code Q_k on his qubits. $\mathcal{B}$ compares y to y', and aborts if any error is detected. $\mathcal{B}$ decodes his n-qubit word according to Q_k, obtaining τ'. $\mathcal{B}$ q-decrypts τ' using x and obtains ρ'.*

Theorem 4 *When the purity testing code $\{Q_k\}$ has error ϵ, the protocol 5.2 is a secure quantum authentication scheme with key length $O(n + \log_2(\#\mathcal{K}))$ and soundness error ϵ. In particular, for the purity testing code described in section 4, the authentication scheme has key length $2m + s + \log_2(2^s + 1) \leq 2n + 1$ and soundness error $2n/[s(2^s + 1)]$, where m is the length of the message in qubits, s is the security parameter, and $\mathcal{A}$ sends a total of $n = m + s$ qubits.*

Proof: From Corollary 3 we have that Protocol 5.1 is a secure interactive authentication protocol. We show that Protocol 5.2 is equivalent to Protocol 5.1, in the sense that any attack on Protocol 5.2 implies an equally successful attack on Protocol 5.1. To do so, we proceed by a series of reductions through two intermediate protocols.

PROTOCOL 5.1 $\rightarrow$ 1$^{\text{st}}$ INT. PROTOCOL: We obtain our first intermediate protocol by observing that in protocol 5.1, $\mathcal{A}$ can perform all of her operations (except for the transmissions) *before* she actually sends anything to $\mathcal{B}$, since these actions do not depend on $\mathcal{B}$'s feedback. This will not change any of the states transmitted in the protocol or computed by $\mathcal{B}$, and so both completeness and soundness will remain the same.

1$^{\text{st}}$ INT. PROTOCOL $\rightarrow$ 2$^{\text{nd}}$ INT. PROTOCOL: There are two changes between the first and second intermediate protocols. First, note that measuring the first qubit of $|\Phi^+\rangle$ and obtaining a random bit c_i is equivalent to choosing c_i at random and preparing the pure state $|c_i\rangle \otimes |c_i\rangle$. Therefore, instead of preparing $|\Phi^+\rangle^{\otimes n}$ and measuring the syndrome of half of it, $\mathcal{A}$ may as well choose the syndromes s at random and encode both halves of $|\Phi^+\rangle^{\otimes m}$ using code Q_k and the syndrome s.

Second, rather than teleporting state ρ to $\mathcal{B}$ using the EPR halves which were encoded in Q_{s_1,s_2}, $\mathcal{A}$ encrypts ρ using a quantum one-time pad and sends it to $\mathcal{B}$ directly, further encoded in Q_k. These behaviors are equivalent since either way, the encoded state is $\sigma_x^{\vec{t}_1}\sigma_z^{\vec{t}_2}\rho\sigma_z^{\vec{t}_2}\sigma_x^{\vec{t}_1}$, where $\vec{t}_1$ and $\vec{t}_2$ are random n-bit vectors.

2$^{\text{nd}}$ INT. PROTOCOL $\rightarrow$ PROTOCOL 5.2: In Protocol 5.2, all the random choices of $\mathcal{A}$ are replaced with the bits taken from a secret random key shared only by her and $\mathcal{B}$. This eliminates the need for an authenticated classical channel, and for any interaction in the protocol. This transformation can only increase the security of the protocol as it removes the adversary's ability to jam the classical communication. $\square$

If the quantum channel is noisy we can modify our scheme to benefit from the error-correction capability of the quantum code. If $\mathcal{B}$ rejects only when the number of observed errors is too large then error correction will fix natural noise or tampering of small amplitude.

6. Authentication Implies Encryption

One notable feature of any protocol derived using Theorem 4 is that the information being authenticated is also completely encrypted. For classical information, authentication and encryption can be considered completely separately, but in this section we will show that quantum information is different. While quantum states can be encrypted without any form of authentication, the converse is not true: any scheme which guarantees authenticity must also encrypt the quantum state almost perfectly. The difference can be understood as a complementarity feature of quantum mechanics: authenticating a message in one basis requires encrypting it in the complementary Fourier-transformed basis. This is essentially another realization of the principle that measuring data in one basis disturbs it in any complementary basis. For classical messages, therefore, encryption is not required: only one basis is relevant. In contrast, for quantum messages, we require authentication in all bases and therefore we must also require encryption in all bases.

To show this, let us consider any fixed authentication scheme. Denote by $\rho_{|\psi\rangle}$ the density matrix transmitted in this scheme when $\mathcal{A}$'s input is $|\psi\rangle$. Let $\rho^{(k)}_{|\psi\rangle}$ denote the density matrix for key k.

Definition 5 *An encryption scheme with error ϵ for quantum states hides information so that if ρ_0 and ρ_1 are any two distinct encrypted states, then the trace distance $D(\rho_0, \rho_1) = \frac{1}{2}\mathrm{Tr}\,|\rho_0 - \rho_1| \le \epsilon$.*

We claim that any good QAS must necessarily also be a good encryption scheme. That is:

Theorem 5 (Main Lower Bound) *A QAS with error ϵ is an encryption scheme with error at most $4\epsilon^{1/6}$.*

Corollary 6 *A QAS with error ϵ requires at least $2m(1 - poly(\epsilon))$ classical key bits.*

The proof of the corollary is quite similar to the proof that $2m$ key bits are needed to perfectly encrypt a quantum message. Instead, we concentrate on Theorem 5.

The intuition behind the proof of this main theorem is that measurement disturbs quantum states, so if the adversary can learn information about the state, she can change the state. More precisely, if the adversary can distinguish between two states $|0\rangle$ and $|1\rangle$, she can change the state $|0\rangle+|1\rangle$ to $|0\rangle-|1\rangle$. An extreme version of this fact is contained in the following proposition:

Proposition 7 *Suppose that there are two states $|0\rangle, |1\rangle$ whose corresponding density matrices $\rho_{|0\rangle}, \rho_{|1\rangle}$ are perfectly distinguishable. Then the scheme is not an ϵ-secure QAS for any $\epsilon < 1$.*

Proof: Since $\rho_{|0\rangle}$, $\rho_{|1\rangle}$ can be distinguished, they must have orthogonal support, say on subspaces V_0, V_1. So consider an adversary who applies a phase shift of -1 conditioned on being in V_1. Then for all k, $\rho^{(k)}_{|0\rangle+|1\rangle}$ becomes $\rho^{(k)}_{|0\rangle-|1\rangle}$. Thus, $\mathcal{B}$ will decode the (orthogonal) state $|0\rangle - |1\rangle$. □

However, in general, the adversary cannot exactly distinguish two states, so we must allow some probability of failure. Note that it is sufficient in general to consider two encoded pure states, since any two mixed states can be written as ensembles of pure states, and the mixed states are distinguishable only if some pair of pure states are. Furthermore, we might as well let the two pure states be orthogonal, since if two nonorthogonal states $|\psi_0\rangle$ and $|\psi_1\rangle$ are distinguishable, two basis states $|0\rangle$ and $|1\rangle$ for the space spanned by $|\psi_0\rangle$ and $|\psi_1\rangle$ are at least as distinguishable.

Due to space limitations, we outline the proof with a sequence of lemmas, only proving the last and most difficult.

We first consider the case when $|0\rangle$ and $|1\rangle$ can *almost* perfectly be distinguished. In that case, the adversary can change $|0\rangle + |1\rangle$ to $|0\rangle - |1\rangle$ with high (but not perfect) fidelity (stated formally in Lemma 8). When $|0\rangle$ and $|1\rangle$ are more similar, we first magnify the difference between them by repeatedly encoding the same state in multiple copies of the authentication scheme, then apply the above argument.

Lemma 8 *Suppose that there are two states $|0\rangle, |1\rangle$ such that $D(\rho_{|0\rangle}, \rho_{|1\rangle}) \ge 1 - \eta$. Then the scheme is not ϵ-secure for $|\psi\rangle = |0\rangle + |1\rangle$ for any $\epsilon < 1 - 2\eta$.*

When two states can be distinguished, but only just barely, the above lemma is not sufficient. Instead, we must magnify the distinguishability of the states $|0\rangle$ and $|1\rangle$ by repeating them by considering the tensor product of many copies of the same state. The probability of distinguishing then goes to 1 exponentially fast in the number of copies:

Lemma 9 *Let ρ_0, ρ_1 be density matrices with $D(\rho_0, \rho_1) = \delta$. Then $D(\rho_0^{\otimes t}, \rho_1^{\otimes t}) \ge 1 - 2\exp(-t\delta^2/2)$.*

We create these repeated states by encoding them in an iterated QAS consisting of t copies of the original QAS (with independent values of the key for each copy).

Lemma 10 *Suppose we iterate the scheme t times. Let $|\psi\rangle = \frac{1}{\sqrt{2}}(|000...0\rangle + |111...1\rangle)$. If $(A, B, \mathcal{K})$ is an ϵ-secure QAS, then the iterated scheme is $10t^3\epsilon$-secure for the state $|\psi\rangle$.*

Note that the proof of this lemma goes through the following crucial claim, which follows from a simple hybrid argument.

Claim 11 (Product states) *The iterated scheme is $t\epsilon$-secure for any product state.*

Proof: For simplicity we prove the claim for the state $|000...0\rangle$. The same proof works for any product pure state (and in fact for separable states in general).

Intuitively, an adversary who modifies the state $|000...0\rangle$ must change some component of the state. We can formalize this by rewriting the projector $P_0^{|000...0\rangle}$ in terms of the individual projectors $P_0^{|0\rangle_i}$.

For the case $t = 2$, $\mathcal{B}$ accepts only if he finds the verification qubits for both schemes in the accept state.

$$\begin{aligned}
&P_0^{|00\rangle} \\
&= (I_{m_1m_2} - |00\rangle\langle 00|) \otimes |\text{ACC}_1\rangle\langle\text{ACC}_1| \\
&\quad \otimes |\text{ACC}_2\rangle\langle\text{ACC}_2| \\
&= \Big((I_{m_1} - |0\rangle\langle 0|) \otimes I_{m_2} + I_{m_1} \otimes (I_{m_2} - |0\rangle\langle 0|) \\
&\quad -(I_{m_1} - |0\rangle\langle 0|) \otimes (I_{m_2} - |0\rangle\langle 0|)\Big) \\
&\quad \otimes |\text{ACC}_1\rangle\langle\text{ACC}_1| \otimes |\text{ACC}_2\rangle\langle\text{ACC}_2| \\
&= P_0^{|0\rangle_1} \otimes |\text{ACC}_2\rangle\langle\text{ACC}_2| + P_0^{|0\rangle_2} \otimes |\text{ACC}_1\rangle\langle\text{ACC}_1| \\
&\quad -P_0^{|0\rangle_1} \otimes P_0^{|0\rangle_2}
\end{aligned}$$

Since $P_0^{|0\rangle_1} \otimes P_0^{|0\rangle_2}$ is positive, for all ρ, we have

$$\text{Tr}(P_0^{|00\rangle}\rho) \le \text{Tr}(P_0^{|0\rangle_1}\rho) + \text{Tr}(P_0^{|0\rangle_2}\rho) \le 2\epsilon$$

Similarly, for larger values of t we have

$$\text{Tr}(P_0^{|000...0\rangle}\rho) \le \sum_{i=1}^{t} \text{Tr}(P_0^{|0\rangle_i}\rho) \le t\epsilon$$

Thus the iterated scheme is $t\epsilon$-secure for $|000...0\rangle$ (and in fact for all product states). □

Proof (of Lemma 10): Consider the net superoperator due to encoding, decoding, and the adversary's intervention, i.e. $\mathcal{O}_{net} = \frac{1}{|\mathcal{K}|}\sum_k B_k\mathcal{O}_{adv}A_k$. By introducing an ancilla system R, we can extend this superoperator to a linear transformation on the joint system $M \otimes R \otimes V$ (where M is the message system and V is $\mathcal{B}$'s verification qubit). For a pure state $|\psi\rangle$, write its image as

$$|\psi\rangle|\gamma_{|\psi\rangle}\rangle|\text{ACC}\rangle + |\beta_{|\psi\rangle}\rangle|\text{REJ}\rangle + |\delta_{|\psi\rangle}\rangle|\text{ACC}\rangle$$

where $|\delta_{|\psi\rangle}\rangle$ is a joint state of MR which is orthogonal to the subspace $|\psi\rangle \otimes R$.

Now consider the family of states $|\psi_i\rangle = |\underbrace{000...0}_{i}\underbrace{111...1}_{t-i}\rangle$, and let $|\gamma_i\rangle = |\gamma_{|\psi_i\rangle}\rangle$ and $|\delta_i\rangle = |\delta_{|\psi_i\rangle}\rangle$.

Claim 12 *For all $i = 0, ..., t-1$, we have $\|\frac{1}{2}(|\gamma_{i+1}\rangle - |\gamma_i\rangle)\| \le (1+\sqrt{2})\sqrt{t\epsilon}$*

Proof: Fix i. Note that $|\psi_+\rangle = \frac{1}{\sqrt{2}}(|\psi_{i+1}\rangle + |\psi_i\rangle)$ is a product state (with $H|0\rangle$ in one position), as is $|\psi_-\rangle = \frac{1}{\sqrt{2}}(|\psi_{i+1}\rangle - |\psi_i\rangle)$. The image of $|\psi_+\rangle$ can be written

$$\begin{aligned}
&\frac{1}{\sqrt{2}}\Big((|\psi_{i+1}\rangle|\gamma_{i+1}\rangle + |\psi_i\rangle|\gamma_i\rangle)|\text{ACC}\rangle \\
&\quad + (|\delta_{i+1}\rangle + |\delta_i\rangle)|\text{ACC}\rangle + (|\beta_{i+1}\rangle + |\beta_i\rangle)|\text{REJ}\rangle\Big) \\
&= \Big(|\psi_+\rangle\frac{1}{2}(|\gamma_{i+1}\rangle + |\gamma_i\rangle) + |\psi_-\rangle\frac{1}{2}(|\gamma_{i+1}\rangle - |\gamma_i\rangle) \\
&+\frac{1}{\sqrt{2}}(|\delta_{i+1}\rangle+|\delta_i\rangle)\Big)|\text{ACC}\rangle + \frac{1}{\sqrt{2}}(|\beta_{i+1}\rangle+|\beta_i\rangle)|\text{REJ}\rangle
\end{aligned}$$

Now we know that $\||\delta_i\rangle\|^2 \le t\epsilon$ for all i (since $|\gamma_i\rangle$ is a product state). Thus, $\|\frac{1}{\sqrt{2}}(|\delta_{i+1}\rangle + |\delta_i\rangle)\| \le \sqrt{2t\epsilon}$.

Moreover, $|\psi_+\rangle$ is a product state and so we have

$$\||\psi_-\rangle\frac{1}{2}(|\gamma_{i+1}\rangle - |\gamma_i\rangle) + \frac{1}{\sqrt{2}}(|\delta_{i+1}\rangle + |\delta_i\rangle)\| \le \sqrt{t\epsilon}$$

Thus, $\||\psi_-\rangle\frac{1}{2}(|\gamma_{i+1}\rangle - |\gamma_i\rangle)\| = \|\frac{1}{2}(|\gamma_{i+1}\rangle - |\gamma_i\rangle)\| \le (1+\sqrt{2})\sqrt{t\epsilon}$. □

Then by the triangle inequality, $\|\frac{1}{2}(|\gamma_t\rangle - |\gamma_0\rangle)\| \le (1+\sqrt{2})t\sqrt{t\epsilon}$. Let $|\Psi_\pm\rangle = \frac{1}{\sqrt{2}}(|\psi_t\rangle \pm |\psi_0\rangle)$. The image of $|\Psi_+\rangle = \frac{1}{\sqrt{2}}(|000...0\rangle + |111...1\rangle)$ is:

$$\begin{aligned}
&\Big(|\Psi_+\rangle\frac{1}{2}(|\gamma_t\rangle + |\gamma_0\rangle) + |\Psi_-\rangle\frac{1}{2}(|\gamma_t\rangle - |\gamma_0\rangle) \\
&+\frac{1}{\sqrt{2}}(|\delta_t\rangle + |\delta_0\rangle)\Big)|\text{ACC}\rangle + \frac{1}{\sqrt{2}}(|\beta_t\rangle + |\beta_0\rangle)|\text{REJ}\rangle
\end{aligned}$$

Now the trace of this state with $P_0^{|\Psi_+\rangle}$ is the square of

$$\begin{aligned}
&\||\Psi_-\rangle\frac{1}{2}(|\gamma_t\rangle - |\gamma_0\rangle) + \frac{1}{\sqrt{2}}(|\delta_t\rangle + |\delta_0\rangle)\| \\
&\le \||\Psi_-\rangle\frac{1}{2}(|\gamma_t\rangle - |\gamma_0\rangle)\| + \|\frac{1}{\sqrt{2}}(|\delta_t\rangle + |\delta_0\rangle)\| \\
&\le (1+\sqrt{2})t\sqrt{t\epsilon} + \sqrt{2t\epsilon} \\
&\le \sqrt{10t^3\epsilon},
\end{aligned}$$

where in the last line, we have assumed $t \ge 2$. That is, the iterated scheme is $10t^3\epsilon$-secure for $|\Psi_+\rangle$. □

Putting the various lemmas together, we find that, given two states $|0\rangle$ and $|1\rangle$ which are slightly distinguishable by the adversary, so $D(\rho_0, \rho_1) \ge \delta$, then in the iterated scheme, $|000...0\rangle$ and $|111...1\rangle$ are more distinguishable: $D(\rho_{|000...0\rangle}, \rho_{|111...1\rangle}) \ge 1 - \eta$, where $\eta \le 2\exp(-t\delta^2/2)$. Since the iterated scheme is $10t^3\epsilon$-secure for the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|000...0\rangle + |111...1\rangle)$, then by Lemma 8,

$$10t^3\epsilon > 1 - 2\eta \ge 1 - 4\exp(-t\delta^2/2)$$

Choosing $t = 1/\sqrt[3]{20\epsilon}$, we get $\delta \le 4\epsilon^{1/6}$.

7. Quantum Signatures

One consequence of the previous theorem is that digitally signing quantum messages is impossible. One can imagine more than one way of defining this task, but any reasonable definition must allow a recipient—who should not be able to alter signed messages—to learn something about the contents of the message. However, this is precisely what is forbidden by the previous theorem: in an information-theoretic setting, any adversary who can gain a non-trivial amount of information must be able to modify the authenticated state with non-negligible success.

If we consider computationally secure schemes, a somewhat narrower definition of digitally signing quantum states remains impossible to realize. If we assume a quantum digital signature protocol should allow any recipient to efficiently extract the original message, then a simple argument shows that he can also efficiently change it without being detected, contradicting the security of the scheme. Namely: Assume that there is transformation U with a small circuit which extracts the original message ρ, leaving auxiliary state $|\varphi\rangle$ (which may not all be held by $\mathcal{B}$). In order to preserve any entanglement between ρ and a reference system, the auxiliary state $|\varphi\rangle$ must be independent of ρ. Therefore, $\mathcal{B}$ can replace ρ with any other state ρ' and then perform $U^\dagger$ on ρ' and his portion of $|\varphi\rangle$, producing a valid signature for ρ'. This is an efficient procedure: the circuit for $U^\dagger$ is just the circuit for U executed backwards.

7.1. Public Key Quantum Authentication

As an alternative to digital signature we introduce the weaker notion of public key quantum authentication.

Let E_b, D_b be $\mathcal{B}$'s public and private keyed algorithms to a PKC and S_a, V_a be $\mathcal{A}$'s private and public keyed algorithms to a digital signature scheme, both resistant to quantum computers' attacks. To perform authentication, $\mathcal{A}$ picks secret and random binary strings k, x, and y, and uses them as keys to q-authenticate ρ as ρ'. $\mathcal{A}$ encrypts and signs the key as $\sigma := S_a(E_b(k|x|y))$. $\mathcal{A}$ sends $\langle\rho', \sigma\rangle$ to $\mathcal{B}$. To verify a state, $\mathcal{B}$ verifies $\mathcal{A}$'s signature on σ using V_a and then discovers the key k, x and y using his private decryption function D_b. $\mathcal{B}$ checks that ρ' is a valid q-authenticated message according to k, x, y, and recovers ρ.

Public Key Quantum Authentication is the basis of a scheme for Uncloneable Encryption [8] allowing transmission of quantum or classical messages in a way that if the authentication succeeds then despite the computational assumption the confidentiality of the message is forever guaranteed unconditionally.

Acknowledgments

We would like to thank H. Bernstein, A. Blokhuis, H.F. Chau, D. DiVincenzo, M. Knill, D. Leung, M. Mosca, E. Rains, and R. de Wolf for helpful discussions or comments.

References

[1] A. Ambainis, M. Mosca, A. Tapp and R. de Wolf “Private quantum channels”, *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pp. 547 – 553, IEEE Computer Society Press, 2000.

[2] A. Ambainis, A. Smith and K. Yang, “Extracting Quantum Entanglement (General Entanglement Purification Protocols)”, *Proceedings of CCC 2002*, May 2002.

[3] C. H. Bennett and G. Brassard. “Quantum cryptography: Public-key distribution and coin-tossing.” In *Proc. of IEEE Conference on Computers, Systems and Signal Processing*, 1984, pp. 175 – 179.

[4] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. Wootters, “Teleporting an unknown quantum state via dual classical and EPR channels”, *Phys. Rev. Lett.*, pp. 1895 – 1899, March 1993.

[5] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, W. K. Wootters, “Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels”, *Phys. Rev. Lett.*, vol. 76, 1996, pp. 722 – 725.

[6] A. Beutelspacher and U. Rosenbaum, *Projective Geometry*, Cambridge University Press, Cambridge, 1998.

[7] J. L. Carter and M. N. Wegman, “New hash functions and their use in authentication and set equality”, *Journal of Computer and System Sciences*, vol. 22, 1981, pp. 265 – 279.

[8] D. Gottesman, “Uncloneable Encryption with Quantum States”, to appear in *Proceedings of QCMC ’02*, Rinton Press, Dec. 2002.

[9] D. Gottesman and I. L. Chuang, “Quantum digital signatures”, to appear in *Proceedings of QCMC ’02*, Rinton Press, Dec. 2002. e-print quant-ph/0105032.

[10] H.-K. Lo and H. F. Chau, “Unconditional Security Of Quantum Key Distribution Over Arbitrarily Long Distances”, *Science*, vol. 283, 1999, pp. 2050 – 2056.

[11] M. A. Nielsen and I. L. Chuang, “Quantum Computation and Quantum Information”, *Cambridge University Press*, 676 pages, 2000.

[12] P. W. Shor and J. Preskill, “Simple Proof of Security of the BB84 Quantum Key Distribution Protocol”, *Phys. Rev. Lett.*, vol. 85, 2000, pp. 441 – 444.