Cryptanalysis of four arbitrated quantum signature schemes

Pierre-Alain Jacqmin*[a] and Jean Liénardy†[a]

[a] *Department of Mathematics, Royal Military Academy, Brussels, Belgium*

March 20, 2026

Abstract

Arbitrated quantum signature (AQS) schemes aim at ensuring the authenticity of a message with the help of an arbitrator. Moreover, they aim at preventing repudiation, both from a sender that denies the origin of a message, and from a receiver who disavows its reception. Such protocols use quantum communication and are often designed to protect quantum messages. In this paper, we study four recently submitted AQS schemes and propose attacks on their security.

Firstly, we look at Zhang, Sun, Zhang and Jia's AQS scheme which aims at signing quantum messages with chained CNOT encryption. We show that the sender can repudiate her messages and make false allegation of reception. Moreover, we show that a dishonest receiver can forge signatures.

Secondly, we analyse Ding, Xin, Yang and Sang's AQS protocol to sign classical messages based on GHZ states. We show that both the sender and the receiver have simple repudiation strategies.

Thirdly, we study Lu, Li, Yu and Han's AQS scheme that uses controlled teleportation to protect quantum messages. We expose forgeries, false allegation attacks and the possibility of repudiation by both parties.

Fourthly, we focus on the AQS scheme by Zhang, Xin, Sun, Li and Li designed to sign classical messages without entangled states. We show that one can disavow the reception of messages, and that information-theoretic security is not achieved for other security goals.

Keywords: arbitrated quantum signature, cryptanalysis, quantum message, authentication, non-repudiation, forgery.

Contents

1 Introduction **2**

2 Cryptanalysis of a scheme based on chained CNOT **5**
2.1 Description of the protocol 5
2.2 Analysis of the security claims 7

3 Cryptanalysis of a scheme based on GHZ states **10**
3.1 Description of the protocol 10
3.2 Analysis of the security claims 12

4 Cryptanalysis of a scheme based on controlled teleportation **13**
4.1 Description of the protocol 13
4.2 Analysis of the security claims 16

*Email: pierre-alain.jacqmin@mil.be; ORCID: 0000-0001-9834-1433
†Email: jean.lienardy@mil.be; ORCID: 0009-0002-6377-0734

5 Cryptanalysis of a scheme without entangled states **17**
5.1 Description of the protocol 17
5.2 Analysis of the security claims 19

6 Conclusion **21**

References **22**

List of Protocols

AQS Protocol 1, of Zhang, Sun, Zhang and Jia 6
AQS Protocol 2, of Ding, Xin, Yang and Sang 11
AQS Protocol 3, of Lu, Li, Yu and Han 14
AQS Protocol 4, of Zhang, Xin, Sun, Li and Li 18

1 Introduction

Signatures are one of the most important primitives of cryptography, as they allow binding a message to its originator while ensuring integrity and non-repudiation. Classical schemes achieve this with public-key primitives, but their security typically rests on computational assumptions that may be threatened by quantum algorithms. Although post-quantum cryptographic signatures have been proposed as a solution to this problem, this has, in parallel, motivated signature mechanisms whose security derives from physics rather than hardness assumptions.

It is well known that public-key signatures of quantum messages with unconditional security are impossible [1]. The seminal work of Zeng and Keitel [23] paved the way for the field of arbitrated quantum signatures (AQS), which aims to overcome this limitation by means of an arbitrator, in addition to the sender and the receiver.

The goal of AQS is often to sign a generic quantum state opposed, say, to signing a classical message by quantum means, as done in quantum digital signatures [8]. This is performed by involving an arbitrator in the protocol, who participates to the signing process and whose role is to resolve disputes. There are multiple models of arbitrators: it can be fully trusted, honest-but-curious (following the protocol but trying to read the message) or adversarial (trying to forge a signature).

The security goals of AQS are unforgeability and non-repudiation. The former is the impossibility for anyone other than the signer to create a legitimate signature on their name, while the later, also referred to as impossibility to disavow, splits in non-repudiation of origin (the signer cannot claim that they did not sign the message) and non-repudiation of receipt (the receiver cannot claim that they did not receive it). With classical public-key digital signatures, usually only unforgeability and non-repudiation of origin are aimed. Non-repudiation of receipt is not a goal of classical signatures, and it is usually not achieved by signatures alone. There exist classical protocols aiming to provide both non-repudiation of origin and receipt, typically referred to as fair non-repudiation protocols. They usually require additional assumptions such as trusted third parties, synchronization, or constraints on computing power [10]. In order to perform an AQS protocol, some secret is usually pre-shared between parties, such as a classical bit string or in the form of an entangled quantum state. Note that multiple copies of the quantum message are often required to perform the signature.

After the seminal work of Zeng and Keitel [23, 5, 22], in which the three parties pre-share a GHZ state, numerous modifications and improvements have been proposed. These include relying on Bell states instead of GHZ triplets [12], removing the need for entanglement [31], modifying the encryption algorithm [25, 26, 32, 24, 9], preventing denial-of-service attacks using Bell states [20], using chaos-based encryption [21], introducing decoy states [14], and chaining CNOT operations for encryption [11, 27].

In parallel, a substantial body of cryptanalysis has highlighted flaws in many published schemes. For example, [31] showed that the initial schemes [23, 12] are vulnerable to disavowal by Bob, who can claim that a legitimate signature is invalid. The works [4, 7]

further demonstrated that the use of the qubit-wise quantum one-time pad (QOTP) makes the transmitted message malleable, enabling existential forgery in earlier protocols. The authors of [7] also noted that modification by Alice of messages between Bob and the arbitrator can lead to repudiation by Alice. This line of cryptanalysis continued with [19, 15, 20, 17, 28], each targeting one or multiple schemes and paving the way towards more secure designs.

Contributions. This article continues that cryptanalytic agenda by focussing on four recent schemes [27, 6, 16, 29] that, to our knowledge, have up to now only been partially analysed.

First, we focus on the scheme of [27] built upon the idea of key-controlled chained CNOT encryption, whose goal is to thwart forgery attacks caused by the malleability of the QOTP. It uses an encryption method that involves and mixes multiple registers of the quantum message as proposed in [26].

Second, we analyse the scheme from [6] that relies on shared GHZ states and which aims at unforgeability and non-repudiation. A particularity of this scheme is that it does not rely on quantum teleportation, even though entangled states are shared.

In third place, we consider an AQS protocol based on controlled teleportation, designed by Lu et al. [16]. There, a 5-qubit quantum state shared between the parties serves as a medium to perform quantum teleportation.

Finally, we address an AQS scheme that signs classical messages without relying on entangled states [29], thereby aiming at better performance and simpler implementation.

The outline of this article is straightforward: in each of Sections 2–5, a scheme is first described. If there are any ambiguities in the protocol description, they are raised and discussed. Security claims of the protocol are then stated, analysed, and challenged. For each protocol, we show that most security claims are not met, either partially or completely. In the last section, we draw short conclusions. In the rest of the present section, we present the general structure of an AQS protocol, which is shared by all the schemes reviewed in this article, and we fix some notation that we will later use.

Generic description of an arbitrated quantum signature scheme. AQS protocols are usually divided into three procedures, although we consider two additional procedures in this paper for better precision. These procedures involve a sender Alice ($\mathcal{A}$), a receiver Bob ($\mathcal{B}$) and an arbitrator Trent ($\mathcal{T}$). A first initialisation step, that we note Init, is usually required to exchange secret bit strings or pre-share quantum states if required by the protocol. This step is assumed to be repeated for each new message to be signed, except if stated explicitly. The actual signature of the message is done during the Sign procedure, which takes as input one or multiple copies of a (classical or quantum) message, as well as the data established during the Init procedure, in order to produce a quantum signature. This procedure often ends with the signature being sent to the receiver Bob. The third step of the protocol is the verification of the signature, noted Verify, during which Bob gets confirmation that the message indeed comes from Alice and has not been modified. Two additional procedures are needed for dispute resolution. Although they are not part of the signature protocol itself, they are nevertheless required in the case where parties disagree at some point after the initial protocol. One of them is the proof of origin, noted Proof-of-Origin, in which Bob proves to the arbitrator that Alice has indeed sent him a certain message. The second one is the proof of receipt, noted Proof-of-Receipt, in which Alice proves to the arbitrator that Bob has indeed received a certain message from her. These additional procedures must be precisely formulated, as the security claims (non-repudiation of origin and non-repudiation of receipt) will depend on how they are performed. We stress that for certain protocols, these procedures are not supposed to be run several times for the same message to sign. For instance, their second run might even be badly defined because a quantum state has been sent or measured.

The communication during the procedures Init, Proof-of-Origin and Proof-of-Receipt is assumed to be made with authenticity of the originator, message integrity, and guaranteed reception, contrarily to Sign and Verify. Moreover, the communication during Init is assumed to be secret. Some protocols require a public board, that we note $\mathcal{P}_b$, on which parties can write public classical messages. It is usually assumed that communications with the public board are authenticated and that the public board is always available.

As far as security properties are concerned, we require that an arbitrator gives a judgement in favour of the complainant Bob for a Proof-of-Origin procedure on a certain message if and only if Alice indeed sent this message. If Bob manages that Trent gives a judgement in his favour when Alice did not sent it, we call it a *forgery* attack; while if Alice manages that Trent gives a judgement in her favour when she has indeed sent the message, it is a *repudiation of origin* attack. In addition, we require that an arbitrator gives a judgement in favour of the complainant Alice in a Proof-of-Receipt procedure on a certain message if and only if Bob indeed received this message. If Bob manages to make Trent judge in his favour when his has actually received the message, we call it a *repudiation of receipt* attack or a *disavowal of reception*; while if Alice makes Trent say that Bob received a message that he did not, we call it a *false allegation* attack. Avoiding false allegations from Alice is often not considered as a security goal. However, it should not be omitted; otherwise, a trivial Proof-of-Receipt procedure could consist of Trent always being in favour of Alice. This would avoid any repudiation from Bob, but would not give any meaningful guarantees.

Notation. We denote a classical bit string by a lowercase letter (e.g., keys k), and its single bits by subscripts $k = k_1 \dots k_n$. Its bit-length is noted with $|\cdot|$ (e.g., $|k| = n$). We use the standard bra-ket notation for quantum states. The tensor product of states is often implicit, that is, we write $|X\rangle |Y\rangle$ for $|X\rangle \otimes |Y\rangle$. We even use the notation $|X, Y\rangle$ when no confusion is possible. Note that the respective lengths of the registers of $|X\rangle$ and $|Y\rangle$ are assumed known by legitimate parties. Qubits are two-dimensional quantum states that are represented by unit vectors in $\mathbb{C}^2$. Given a state defined as a tensor product of n qubits, we call n its length.

The four Pauli matrices are $\{\sigma_\alpha\}_{\alpha=0}^3 = \{I, X, Y, Z\}$ and the Hadamard operator is noted H. We recall that each of these operators is an involution $(\sigma_\alpha)^2 = I = H^2$ for any α, that different non-identity Pauli matrices anticommute ($XY = -YX$, $XZ = -ZX$ and $YZ = -ZY$) and that Pauli and Hadamard operators satisfy:

$$HX = ZH, \quad HZ = XH, \quad HY = -YH.$$

Hadamard and Pauli gates act qubit-wise: given an n-qubit register $|M\rangle = \sum_j c_j \bigotimes_{i=1}^n |M_i^j\rangle$ and a bit string y of length greater than or equal to n, $|S\rangle = H^y |M\rangle$ means that $|S\rangle = \sum_j c_j \bigotimes_{i=1}^n |S_i^j\rangle$ where $|S_i^j\rangle = H^{y_i} |M_i^j\rangle$ and analogously for Pauli gates. We use the standard computational basis $\{|0\rangle, |1\rangle\}$ composed of eigenvectors of Z and the diagonal basis $\{|+\rangle, |-\rangle\}$, defined as $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. Given a classical message m, we note $|m\rangle$ its encoding in the computational basis: $|m\rangle = \bigotimes_i |m_i\rangle$ with $|m_i\rangle = X^{m_i} |0\rangle$. In other words, we apply the canonical mapping $m_i \mapsto |m_i\rangle$ to $m_i \in \{0, 1\}$.

When a party transmits a message, this is indicated with the symbol $\rightarrow$. The notation $\mathcal{A} \leftrightarrow \mathcal{B} : m$, used during Init procedures, indicates that the classical message m has been exchanged between $\mathcal{A}$ and $\mathcal{B}$.

In a description of a protocol, the phrase "$\mathcal{X}$: recover $|\cdot\rangle$" is meant to apply the appropriate operations on a register in $\mathcal{X}$'s possession using information available to $\mathcal{X}$ in order to recover a given state, for example, as in standard quantum teleportation or to decrypt an encrypted message.

In protocols, we sometimes use a semicolon to separate instructions. We also use the notation $X = Y \ ? \ P_T : P_F$ to indicate that an equality test is performed between X and Y. If the test succeeds, that is, if one evaluates $X = Y$, then the instructions P_T are executed, else the instructions P_F are executed. Many of these tests are of the form $E(Z) = Y \ ? \ P_T : P_F$, that is, with $X = E(Z)$ being the result of an invertible function E. When this is the case, we assume that if the test is passed, the party implicitly recomputes the input states Z and Y. Note that if X and Y are quantum states, one usually needs several copies of them to trustfully perform such an equality test by using SWAP tests [3].

Some instructions are written in bold and indicate that a party aborts: the party does not continue the procedure. When such an instruction is reached, it is assumed that all other parties are aware of it. Specifically, in a Verify procedure, the instruction **abort** simply indicates that the party aborts the procedure, **accept** means that the receiver considers the signature as valid and ends the procedure, **reject** indicates that the receiver considers the signature invalid and aborts. In a phase of dispute resolution, the instructions **favour** $\mathcal{A}$ and

favour $\mathcal{B}$ respectively mean that the arbitrator gives a judgement in favour of Alice or Bob, and the procedure is stopped.

Acknowledgment. Both authors were funded by the Defense Funded Research (DFR) study DAP/23-01.

2 Cryptanalysis of a scheme based on chained CNOT

In this section, we analyse the AQS scheme proposed by Zhang et al. [27], aimed at the signature of quantum messages and based on the idea of chained CNOT encryption [11, 17]. We first describe the scheme and its building blocks before analysing its security claims.

2.1 Description of the protocol

The main building block for the security of this scheme is a multi-layer encryption technique aimed at diffusing the information in each qubit of the message across multiple qubits, rather than merely performing a qubit-wise operation. The entire protocol is given as Protocol 1.

The first encryption layer is performed by using multiple CNOT gates in a chain-like fashion. We recall that the gate $\mathrm{CNOT}_{i,j}$ (with $1 \leqslant i, j \leqslant n$, $i \neq j$) applied on an n-qubit state $|P\rangle = \bigotimes_{\ell=1}^{n} |P_\ell\rangle$ with $P_1, \ldots, P_n \in \{0,1\}$ acts on the i^{th} and j^{th} tensor register, mapping $|P_i\rangle\,|P_j\rangle$ given in the computational basis to $|P_i\rangle\,|P_i \oplus P_j\rangle$ (and this definition is extended linearly). Furthermore, $\mathrm{CNOT}_{i,i}$ is the identity operator. The chained CNOT encryption of an n-qubit message $|P\rangle$ is noted $E_{k^1}^{\mathrm{CNOT}}\,|P\rangle$ and requires a key $k^1 = (k^1_{(1)}, \ldots, k^1_{(n)})$ which is a permutation of $(1, \ldots, n)$ where each $k^1_{(i)}$ is the image of i. The permutation can be encoded in $\lceil \log_2(n!) \rceil$ bits, which we consider as the key length. The chained CNOT encryption is given by

$$E_{k^1}^{\mathrm{CNOT}}\,|P\rangle = \mathrm{CNOT}_{n,k^1_{(n)}} \cdots \mathrm{CNOT}_{2,k^1_{(2)}}\,\mathrm{CNOT}_{1,k^1_{(1)}}\,|P\rangle\,.$$

The second encryption layer is H^{k^2}, which consists in applying H conditionally to the key bits of $k^2 \in \{0,1\}^n$.

The third and final step for the encryption is a permutation that depends on the key $k^3 \in \{0,1\}^n$, that we denote by $E_{k^3}^{\mathrm{perm}}$. The authors describe two such key-controlled permutations that can be applied to the qubits of a state.

The first one is a controlled pairwise transposition of the qubits i and $n+1-i$. For each $i \in \{1, \ldots, \lfloor \frac{n}{2} \rfloor\}$, let $\tau_i = k^3_i \oplus k^3_{n+1-i}$. Then, $S^{\tau_i}_{i,n+1-i}$ is a linear operator that acts on $|P\rangle = \bigotimes_{\ell=1}^{n} |P_\ell\rangle$ with $P_1, \ldots, P_n \in \{0,1\}$ as the identity if $\tau_i = 0$ and by swapping the qubits $|P_i\rangle$ and $|P_{n+1-i}\rangle$ if $\tau_i = 1$. This action is extended on generic states by linearity. One then defines $E_{k^3}^{\mathrm{transp}} = S^{\tau_{\lfloor \frac{n}{2} \rfloor}}_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil + 1} \circ \cdots \circ S^{\tau_2}_{2,n-1} \circ S^{\tau_1}_{1,n}$.

The second controlled permutation, denoted $E_{k^3}^{\mathrm{rot}} = S_n^{(\tau)}$, is a rotation of the n qubits. It rotates the qubits by τ positions where $\tau = (\sum_i k^3_i) \pmod n$:

$$S_n^{(\tau)}\,|P_1, P_2, \cdots, P_n\rangle = |P_{\tau+1}, \cdots, P_n, P_1, \cdots, P_\tau\rangle$$

for $P_1, \ldots, P_n \in \{0,1\}$, extended linearly.

The full encryption of a state $|P\rangle$ of length n is denoted $E_k^{\mathrm{KCCC}}\,|P\rangle$. It requires a key $k = k^1 \| k^2 \| k^3$ with k^1 a permutation of $(1, \ldots, n)$ and $k^2, k^3 \in \{0,1\}^n$:

$$E^{\mathrm{KCCC}}_{k^1||k^2||k^3}\,|P\rangle = E^{\mathrm{perm}}_{k^3} \circ H^{k^2} \circ E^{\mathrm{CNOT}}_{k^1}\,|P\rangle\,,$$

where "perm" is either "transp" or "rot", depending on the chosen controlled permutation, which is a public information chosen before the start of the protocol. The protocol is given[1] in Protocol 1. We note that it does not use any pre-shared quantum states, such as Bell states or GHZ triplets.

[1] with a small correction to the protocol given in [27]. Indeed, the condition "$V_T = 0$" in (V2) appears to be an error. Consistency with the subsequent step requires $V_T = 1$. We adopted this corrected condition in our analysis.

Protocol 1 AQS protocol of [27] (the encryption of $|M\rangle$ to $|P\rangle$ is chosen as E^{KCCC}).

procedure INIT

1. $\mathcal{A} \leftrightarrow \mathcal{B}$: k_{AB}.
2. $\mathcal{A} \leftrightarrow \mathcal{T}$: k_{AT}.
3. $\mathcal{B} \leftrightarrow \mathcal{T}$: k_{BT}.

end procedure

procedure SIGN (on 3 copies $|M^0\rangle = |M^1\rangle = |M^2\rangle$ of $|M\rangle = \bigotimes_{i=1}^{n} |M_i\rangle$)

1. $\mathcal{A}$: choose random r; $|P^i\rangle = E_r^{\mathrm{KCCC}} |M^i\rangle$ for $i \in \{0,1,2\}$.
2. $\mathcal{A}$: $|P\rangle = |P^0\rangle$; $\quad |R_{AB}\rangle = E_{k_{AB}}^{\mathrm{KCCC}} |P^1\rangle$; $\quad |S_A\rangle = E_{k_{AT}}^{\mathrm{KCCC}} |P^2\rangle$.
3. $\mathcal{A}$: $E_{k_{AB}}^{\mathrm{KCCC}}\big(|P\rangle, |R_{AB}\rangle, |S_A\rangle\big) \to \mathcal{B}$.

end procedure

procedure VERIFY

1. $\mathcal{B}$: recover $|P\rangle, |R_{AB}\rangle, |S_A\rangle$.
2. $\mathcal{B}$: $|Y_{BT}\rangle = E_{k_{BT}}^{\mathrm{KCCC}}\big(|P\rangle, |S_A\rangle\big) \to \mathcal{T}$.
3. $\mathcal{T}$: recover $|P\rangle, |S_A\rangle$.
4. $\mathcal{T}$: $E_{k_{AT}}^{\mathrm{KCCC}} |P\rangle = |S_A\rangle$? $v_T = 1 : v_T = 0$; $v_T \to \mathcal{P}_b$.
5. $\mathcal{B}$: $v_T = 0$? **reject** : continue.
6. $\mathcal{T}$: $|Y_{TB}\rangle = E_{k_{BT}}^{\mathrm{KCCC}}\big(|P\rangle, |S_A\rangle\big) \to \mathcal{B}$.
7. $\mathcal{B}$: recover $|P\rangle, |S_A\rangle$.
8. $\mathcal{B}$: $E_{k_{AB}}^{\mathrm{KCCC}} |P\rangle = |R_{AB}\rangle$? $v_B = 1 : v_B = 0$; $v_B \to \mathcal{P}_b$.
9. $\mathcal{B}$: $v_B = 0$? **reject** : continue.
10. $\mathcal{A}$: $v_T = v_B = 1$? $r \to \mathcal{P}_b$: **abort**.
11. $\mathcal{B}$: recover $|M^0\rangle$ from $|P\rangle$ and r.
12. $\mathcal{B}$: store signature $(|M^0\rangle, |S_A\rangle, r)$; **accept**.

end procedure

procedure PROOF-OF-ORIGIN (on input $|\overline{M}\rangle = \bigotimes_{i=1}^{n} |\overline{M}_i\rangle$)

1. $\mathcal{T}$: $v_T = v_B = 1$? continue : **favour** $\mathcal{A}$.
2. $\mathcal{B}$: $|S_A\rangle \to \mathcal{T}$.
3. $\mathcal{T}$: $|S_A\rangle = E_{k_{AT}}^{\mathrm{KCCC}} E_r^{\mathrm{KCCC}} |\overline{M}\rangle$? **favour** $\mathcal{B}$: **favour** $\mathcal{A}$.

end procedure

procedure PROOF-OF-RECEIPT (on input $|\overline{M}\rangle = \bigotimes_{i=1}^{n} |\overline{M}_i\rangle$)

1. $\mathcal{T}$: $v_T = v_B = 1$? **favour** $\mathcal{A}$: **favour** $\mathcal{B}$.

end procedure

Ambiguities. We note that the description of [27] contains some ambiguities. We list and explain how we deal with these ambiguities in our analysis of the scheme.

A first ambiguity concerns the subkey sizes. Given a key $k = k^1 \| k^2 \| k^3$ used to encrypt a state of n qubits, the key k^3 (noted L in [27]) is obtained from k^2 through a hash function: $k^3 = f(k^2)$. However, the hash function has input size $2n$, while k^2 is of length n. We instead treat k^3 as independent from k^2, both of length n. This choice clears the ambiguity and furthermore removes the need for a hash function, whose presence could affect the ITS nature of the whole scheme, and which was only marginally useful, as the key length is dominated by that of k^1 for large n. Moreover, this does not change the success probability of our attacks if the hash function is modelled as a random oracle.

Still regarding the subkey size, we note a suboptimal usage of the key k^3. If the permutation chosen is E^{transp}, then only $\lfloor \frac{n}{2} \rfloor$ bits are required to specify the transposition. These are the bits denoted by τ_i above. For E^{rot}, an integer $\tau \in [0, n-1]$ is sufficient, hence requiring $\lceil \log_2(n) \rceil$ bits. In both cases, some of the n bits of k^3 are superfluous. Moreover, while the distribution of the τ_i used by E^{transp} is uniform if the k_i^3 are independent and uniformly distributed, the distribution of $\tau = (\sum_i k_i^3) \mod n$, used for E^{rot}, is not uniform for a uniform distribution of k_i^3. While we do not know if this behaviour was desired, we decided to keep, in the following, the initial construction using the n bits of k^3. We stress that our attacks would still hold (although with smaller or equal probability) if τ was uniformly selected in $[0, n-1]$.

A second imprecision regards the size of the whole keys and their re-use. The authors mention a single key length, which is equal to n. However, as stated above, a key that encrypts a state of n qubits must have a size of at least $2n + \lceil \log_2(n!) \rceil$ (or still $n + \lceil \log_2(n!) \rceil$ if we keep the hash function, and $n + \lceil \log_2(n!) \rceil + \lfloor \frac{n}{2} \rfloor$ or $n + \lceil \log_2(n!) \rceil + \lceil \log_2(n) \rceil$ if we had chosen optimal subkeys for E^{perm}). Furthermore, while the key k_{AT} does encrypt and decrypt states of size n, the key k_{BT} is used for states of size $2n$ (see steps 2 and 6 in the verifying procedure), hence it is of minimal length $4n + \lceil \log_2((2n)!) \rceil$. As for the key k_{AB}, it is used to encrypt both a state of size n and a state of size $3n$ in steps 2 and 3 of the signing procedure, respectively. It is unclear what the impact of reusing a key is. In the following, we suppose that two different keys k_{AB} are used to prevent any interference. The size of the pre-shared secret between Alice and Bob is thus $8n + \lceil \log_2((3n)!) \rceil + \lceil \log_2(n!) \rceil$.

A third ambiguity is the transformation of $|M\rangle$ into $|P\rangle$ using r, which is not explicitly stated. In the following, we assume that this is done using the same layered encryption E^{KCCC}, using a classical key r of appropriate size.

A fourth ambiguity concerns the procedures Proof-of-Origin and Proof-of-Receipt since they are not explicitly described in the paper [27]. The Proof-of-Origin, where Bob wants to prove that Alice sent him a certain message $|\overline{M}\rangle$, is quite intuitive, and the one we propose in Protocol 1 is probably what the authors have in mind. However, the procedure Proof-of-Receipt, in which Alice wants to prove that Bob received a certain message $|\overline{M}\rangle$ from her, is far less straightforward. According to our understanding of Section 4.2 of [27], it works as follows: on input $|\overline{M}\rangle$ from Alice, Trent gives a judgement in favour of Alice if and only if $v_T = v_B = 1$. In this case, Alice has a trivial false allegation attack: as the procedure does not depend on $|\overline{M}\rangle$, Alice can pretend to have sent any message to Bob. Although being trivial, we choose to describe this option in Protocol 1. We thought of another option, in which Trent asks Bob to send him the state $|S_A\rangle$ from the verification procedure to make an additional equality test, as in the Proof-of-Origin procedure. Since Bob could modify $|S_A\rangle$ as he wishes, he can trivially deny having received the message $|\overline{M}\rangle$ from Alice.

A fifth issue concerns the set of messages Alice can sign. In the paper [27], it is mentioned that she can only sign n-qubit messages of the factorised form $|M\rangle = \bigotimes_{i=1}^{n} |M_i\rangle$, although there exists n-qubit messages which are not of this form. The utility of this restriction is unclear to us as the scheme could easily be extended to any entangled state. Moreover, as no checks on format are made by the participants, we do not know what happens if an non-factorised message is used. Nevertheless, our attacks are not impacted by this restriction.

2.2 Analysis of the security claims

The authors of [27] base their security analysis on the fact that an error in the ciphertext may propagate with the decryption of E^{KCCC} through different qubits in the plaintext. In this

subsection, we show that despite the enhanced encryption used, there is still the possibility for Alice to repudiate a message and to make false allegations, and for Bob to do a forgery. This contradicts the claims of [27].

Repudiation by Alice. Suppose that there are two distinct quantum messages $|M\rangle$ and $|M'\rangle$, with their corresponding states, as defined at step 2 of the SIGN procedure, being $|P\rangle$, $|S_A\rangle$, $|R_{AB}\rangle$ and $|P'\rangle$, $|S'_A\rangle$, $|R'_{AB}\rangle$. In the following, we suppose that Alice has sent $E^{\mathrm{KCCC}}_{k_{AB}}(|P'\rangle, |R_{AB}\rangle, |S'_A\rangle)$ to Bob during the last step of the SIGN procedure. This message is considered valid by the arbitrator during the VERIFY procedure and $v_T = 1$. Alice wishes to intercept and modify the message $|Y_{TB}\rangle$ on its way back to Bob in step 6 so that he obtains $|P\rangle |S'_A\rangle$ instead of $|P'\rangle |S'_A\rangle$. If she succeeds, Bob's checks will pass, and he will consider the signature $(|M\rangle, |S'_A\rangle, r)$ as valid for the message $|M\rangle$. If Bob wants to prove that Alice sent him this message, he will show $|M\rangle$ to Trent as input to the procedure PROOF-OF-ORIGIN and send $|S'_A\rangle$ in step 2. Since $|M\rangle \neq |M'\rangle$, the arbitrator will agree with Alice, who will have succeeded in her disavowal.

For a message $|M\rangle$ to be repudiated, we consider $|P\rangle = E^{\mathrm{KCCC}}_r |M\rangle$, and $|P'\rangle = X_1 |P\rangle$ with X_1 being the gate X applied on the first register. The state $|M'\rangle = (E^{\mathrm{KCCC}}_r)^{-1} |P'\rangle$ is of the factorised form $\bigotimes_{i=1}^n |M'_i\rangle$ if $|M\rangle$ is. This observation follows from the commutation relations between the gates CNOT and X:

$$X_j \mathrm{CNOT}_{i,j} = \mathrm{CNOT}_{i,j} X_j \quad \text{and} \quad X_i \mathrm{CNOT}_{i,j} = \mathrm{CNOT}_{i,j} X_i X_j \quad \text{for } i \neq j.$$

These relations imply that conjugating a product of Pauli matrices by $E^{\mathrm{CNOT}}_{r^1}$ yields, up to a global phase, another tensor product of Pauli matrices. The same observation also holds for the last two steps of encryption and therefore for E^{KCCC}_r.

Given $|M\rangle$, Alice chooses r so that $|M'\rangle$ is distinct from $|M\rangle$. Writing the key k_{BT} simply as $k = k^1 \| k^2 \| k^3$, the state sent back to Bob by the arbitrator is

$$|Y_{TB}\rangle = E^{\mathrm{KCCC}}_k (|P'\rangle |S'_A\rangle) = E^{\mathrm{perm}}_{k^3} \circ H^{k^2} \circ E^{\mathrm{CNOT}}_{k^1} (|P'\rangle |S'_A\rangle).$$

The first layer of encryption, $E^{\mathrm{CNOT}}_{k^1}$, depends on the permutation $k^1 = (k^1_{(1)}, \ldots, k^1_{(2n)})$, with the image of i written $k^1_{(i)}$. We consider the event that $k^1_{(1)} = 1$, that is, the permutation has 1 as fixed point. If that event occurs, one has $X_1 E^{\mathrm{CNOT}}_{k^1} = E^{\mathrm{CNOT}}_{k^1} X_1$.

In the case where E^{transp} is used (i.e., "perm" is "transp"), we consider the events that $k^2_1 = 0$ and $\tau_1 \coloneqq k^3_1 \oplus k^3_{2n} = 0$. These events are independent and each of them occurs with probability $\frac{1}{2}$. In this case, Alice's strategy is to apply the operator X_1 to the state sent by Trent to Bob in step 6 of the verification procedure. If $k^1_{(1)} = 1$, $k^2_1 = 0$ and $\tau_1 = 0$ occur, $E^{\mathrm{transp}}_{k^3}$ and H^{k^2} do not act on the first register and one has

$$\begin{aligned} X_1 |Y_{TB}\rangle &= X_1 E^{\mathrm{KCCC}}_k (|P'\rangle |S'_A\rangle) \\ &= X_1 E^{\mathrm{transp}}_{k^3} H^{k^2} E^{\mathrm{CNOT}}_{k^1} (|P'\rangle |S'_A\rangle) \\ &= E^{\mathrm{transp}}_{k^3} H^{k^2} E^{\mathrm{CNOT}}_{k^1} X_1 (|P'\rangle |S'_A\rangle) \\ &= E^{\mathrm{KCCC}}_k (|P\rangle |S'_A\rangle). \end{aligned}$$

The state that Bob obtains by decryption is $|P\rangle |S'_A\rangle$. It will pass Bob checks (as he has both $|P\rangle$ and $|R_{AB}\rangle$) and Bob will accept the message $|M\rangle$ as legitimate. However, it can be later repudiated by Alice as Bob has $|S'_A\rangle$ and not $|S_A\rangle$. The overall probability of such a successful repudiation is lower-bounded by

$$\Pr(\text{repudiation}) \geqslant \Pr\left(k^1_{(1)} = 1\right) \cdot \Pr\left(k^2_1 = 0\right) \cdot \Pr(\tau_1 = 0) = \frac{1}{8n}.$$

In the case where E^{rot} is used (i.e., "perm" is "rot"), we now consider the events $k^2_1 = 0$ and $\tau \coloneqq (\sum_{i=1}^{2n} k^3_i) \pmod{2n} = n$. These events are independent of each other, and one has $\Pr(k^2_1 = 0) = \frac{1}{2}$ and $\Pr(\tau = n) = 2^{-2n} \cdot \binom{2n}{n} \geqslant \dfrac{1}{2\sqrt{n}}$ (using Lemma 7 of Chapter 10 of [18]). In this case, Alice's strategy is to apply the operator X_{n+1}, defined as the gate X applied to

the $(n+1)^{\text{th}}$ register, to the state sent by Trent to Bob in step 6 of the verification procedure. If $k^1_{(1)} = 1$, $k^2_1 = 0$ and $\tau = n$ occur, one has

$$\begin{aligned} X_{n+1} |Y_{TB}\rangle &= X_{n+1} E_k^{\text{KCCC}} \left(|P'\rangle |S'_A\rangle \right) \\ &= X_{n+1} E_{k^3}^{\text{rot}} H^{k^2} E_{k^1}^{\text{CNOT}} \left(|P'\rangle |S'_A\rangle \right) \\ &= E_{k^3}^{\text{rot}} X_1 H^{k^2} E_{k^1}^{\text{CNOT}} \left(|P'\rangle |S'_A\rangle \right) \\ &= E_{k^3}^{\text{rot}} H^{k^2} E_{k^1}^{\text{CNOT}} X_1 \left(|P'\rangle |S'_A\rangle \right) \\ &= E_k^{\text{KCCC}} \left(|P\rangle |S'_A\rangle \right). \end{aligned}$$

In this case, similarly as above, Alice succeeds in repudiating her message. The success probability is lower-bounded by

$$\Pr\left(\text{repudiation}\right) \geqslant \Pr\left(k^1_{(1)} = 1\right) \cdot \Pr\left(k^2_1 = 0\right) \cdot \Pr\left(\tau = n\right) \geqslant \frac{1}{8n\sqrt{n}}.$$

Although the authors of [27] did not bound the probability of a disavowal, the claim that it is "*obviously impossible*" and that Alice cannot disavow is unfounded. In any case, such a probability should decrease exponentially (or at least as a negligible function, with respect to some security parameter), which is not the case here.

Even though the attack uses the fact that the permutation k^1 has a fixed point, we stress that there exist similar attacks even if we restrict k^1 to permutations without any fixed points. In this article, we have chosen the simplest attack that highlights the flaws in the design.

False allegation by Alice. As already mentioned in Subsection 2.1, Alice has a trivial false allegation strategy consisting in requiring a Proof-of-Receipt procedure for any message $|\overline{M}\rangle$ of her choice. As this procedure does not depend on the input, it succeeds with probability 1.

Forgery by Bob. Suppose that Alice trustfully signs the message $|M\rangle$ by sending $E_{k_{AB}}^{\text{KCCC}}\big(|P\rangle, |R_{AB}\rangle, |S_A\rangle \big)$ to Bob. We denote the key k_{AT} as $k = k^1 \| k^2 \| k^3$.

If one chooses the controlled pairwise transposition for the final step of the encryption (i.e., "perm" is "transp"), we consider the three independent events $k^1_{(1)} = 1$ (i.e., k^1 fixes 1), $k^2_1 = 0$ and $\tau_1 \coloneqq k^3_1 \oplus k^3_n = 0$ which occur respectively with probability $\frac{1}{n}$, $\frac{1}{2}$ and $\frac{1}{2}$ (or 1 if $n = 1$). In that case, Bob's strategy consists in applying X_1 to both $|P\rangle$ and $|S_A\rangle$ in the beginning of the verification procedure to obtain $|P'\rangle$ and $|S'_A\rangle$. If $k^1_{(1)} = 1$, $k^2_1 = 0$ and $\tau_1 = 0$ hold, one has

$$\begin{aligned} |S'_A\rangle &= X_1 |S_A\rangle \\ &= X_1 E_k^{\text{KCCC}} |P\rangle \\ &= X_1 E_{k^3}^{\text{transp}} H^{k^2} E_{k^1}^{\text{CNOT}} |P\rangle \\ &= E_{k^3}^{\text{transp}} H^{k^2} E_{k^1}^{\text{CNOT}} X_1 |P\rangle \\ &= E_k^{\text{KCCC}} |P'\rangle. \end{aligned}$$

Therefore, in that case, if Bob sends $|Y'_{BT}\rangle = E_{k_{BT}}^{\text{KCCC}}\big(|P'\rangle, |S'_A\rangle \big)$ to Trent, the later will write $v_T = 1$ on the public board. Then, Bob can write $v_B = 1$ at step 8 of the Verify procedure. At the end, he will save $(|M'\rangle, |S'_A\rangle, r)$ where $|M'\rangle = \left(E_r^{\text{KCCC}}\right)^{-1} |P'\rangle$. During a Proof-of-Origin procedure, he will thus be able to prove that Alice sent him the message $|M'\rangle$, which is not true if $|M\rangle \neq |M'\rangle$, that is if $|P\rangle \neq |P'\rangle$ (up to a global phase). Since $|P'\rangle = X_1 |P\rangle$, this last condition occurs exactly when $|P\rangle$ is neither of the form $|+\rangle |Q\rangle$ nor of the form $|-\rangle |Q\rangle$. Fixing all bits of $r = r^1 \| r^2 \| r^3$ except that of r^2 influencing the first qubit of $|P\rangle$, this happens for at least one of the two choices for that bit, leading to $\Pr(|P\rangle \neq |P'\rangle) \geqslant \frac{1}{2}$. The overall success probability of this forgery attack is lower-bounded by

$$\Pr\left(\text{forgery}\right) \geqslant \Pr\left(k^1_{(1)} = 1\right) \cdot \Pr\left(k^2_1 = 0\right) \cdot \Pr\left(\tau_1 = 0\right) \cdot \Pr\left(|P\rangle \neq |P'\rangle\right) \geqslant \frac{1}{8n}$$

which is not negligible.

If one chooses the rotation as controlled permutation (i.e., "perm" is "rot"), we consider the three independent events $k^1_{(1)} = 1$, $k^2_1 = 0$ and $\tau \coloneqq (\sum_{i=1}^n k^3_i) \pmod n = \lfloor \frac{n}{2} \rfloor$. Using Lemma 7 of Chapter 10 of [18], the third event occurs with probability $\Pr(\tau = \lfloor \frac{n}{2} \rfloor) = 2^{-n} \cdot \binom{n}{\lfloor \frac{n}{2} \rfloor} \geqslant \frac{1}{2\sqrt{n}}$ if $n > 1$ (or probability 1 if $n = 1$). Bob can use a strategy similar to that above, applying the gate X to the first qubit of $|P\rangle$ and, if $n > 1$, to the $(n - \lfloor \frac{n}{2} \rfloor + 1)^{\text{th}}$ qubit of $|S_A\rangle$ (or its first qubit if $n = 1$) to get $|P'\rangle$ and $|S'_A\rangle$ respectively. The overall success probability of this forgery attack is lower-bounded by

$$\Pr(\text{forgery}) \geqslant \Pr\left(k^1_{(1)} = 1\right) \cdot \Pr\left(k^2_1 = 0\right) \cdot \Pr\left(\tau = \left\lfloor \frac{n}{2} \right\rfloor\right) \cdot \Pr(|P\rangle \neq |P'\rangle) \geqslant \frac{1}{8n\sqrt{n}}$$

which is again not negligible.

For the same reason given in Alice's repudiation, it is not hard to see that, in both cases, the message $|M'\rangle$ that Bob pretends Alice sent is of the factorised form $\bigotimes_{i=1}^n |M'_i\rangle$ if $|M\rangle$ is.

3 Cryptanalysis of a scheme based on GHZ states

In this section, we analyse the scheme proposed in [6]. Its security is based on GHZ-like states shared between parties, but which are not used in a teleportation protocol. Instead, the parties verify that the three registers of the GHZ states have a predetermined XOR-sum in the computational basis. The scheme builds upon the protocol of [30], which is shown by the authors of [6] to be trivially insecure: Bob can apply X gates to the message and its signature to forge a valid signed message. The improved protocol aims to address this insecurity.

3.1 Description of the protocol

The AQS protocol of [6], described in Protocol 2, is designed to sign classical messages.

The protocol uses GHZ-like states that are shared between parties during initialisation. Specifically, given a GHZ state $|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$, the state

$$H^{\otimes 3} |GHZ\rangle = \frac{1}{\sqrt{2}}(|+++\rangle + |---\rangle) = \frac{1}{2}(|000\rangle + |011\rangle + |101\rangle + |110\rangle)$$

is created and shared by sending one of the three registers to each involved party. To sign a classical message of length n, n such states are produced and distributed. We denote the n qubit state received by Alice (respectively by Bob, by Trent) as $|A\rangle$ (respectively $|B\rangle$, $|T\rangle$). This step is assumed to be done securely by the arbitrator. The scheme relies on two one-way hash functions, $G_1, G_2 \colon \{0,1\}^* \to \{0,1\}^n$, which are considered secure.

By abuse of notation, for two n qubit states $|P\rangle$ and $|Q\rangle$, we write $\text{CNOT}(|P\rangle, |Q\rangle)$ simply by $|P\rangle |P \oplus Q\rangle$ where CNOT is defined for $P_1, \ldots, P_n, Q_1, \ldots, Q_n \in \{0,1\}$ by

$$\text{CNOT}\left(\bigotimes_{i=1}^n |P_i\rangle, \bigotimes_{i=1}^n |Q_i\rangle\right) = \left(\bigotimes_{i=1}^n |P_i\rangle\right) \otimes \left(\bigotimes_{i=1}^n |P_i \oplus Q_i\rangle\right)$$

and extended linearly. We stress that, following [6], we use throughout this section a notation in which possibly entangled states are written as factorised kets. Thus, expressions such as $|A\rangle |B\rangle |T\rangle$ or $|P\rangle |P \oplus Q\rangle$ should not be interpreted as tensor products of independent pure states, but as a convenient description of the relevant registers of a possibly entangled state. As an example, during step 3 of the VERIFY procedure, the state $\text{CNOT}(|T\rangle |B \oplus A \oplus P^1_A\rangle)$ may be entangled, but we still denote it by $|T\rangle |P'_A\rangle$ and refer to operations on the second register as acting on $|P'_A\rangle$.

Ambiguities. The proposed protocol contains several ambiguities. We list them below and state the conventions we adopt to resolve them.

In step 4 of the SIGN procedure and in step 2 of the VERIFY procedure, quantum messages are sent, and these communications are secured "by the eavesdropping detection technology [13]", which is based on decoy states. The authors do not make precise statement

Protocol 2 AQS protocol of [6] based on GHZ states.

procedure INIT

1. $\mathcal{T}$: prepare and distribute n GHZ states $H^{\otimes 3}\left|GHZ\right\rangle_{ABT}$.
2. $\mathcal{A} \leftrightarrow \mathcal{T}$: $k_A \in \{0,1\}^n$.

end procedure

procedure SIGN (on a classical message $m \in \{0,1\}^n$)

1. $\mathcal{A}$: $t = G_1(m\|k_A)$ and $s = G_2(m\|k_A)$.
2. $\mathcal{A}$: $\left|P_A^0\right\rangle = \left|P_A^1\right\rangle = H^t X^s \left|m\right\rangle$.
3. $\mathcal{A}$: $\mathrm{CNOT}(\left|A\right\rangle \left|P_A^i\right\rangle) = \left|A\right\rangle \left|A \oplus P_A^i\right\rangle$ for each $i \in \{0,1\}$.
4. $\mathcal{A}$: $\left(\left|A \oplus P_A^0\right\rangle, \left|A \oplus P_A^1\right\rangle, m\right) \to \mathcal{B}$ // secured using decoy states.
5. $\mathcal{B}$: store $\left(\left|B\right\rangle, \left|A \oplus P_A^0\right\rangle, m\right)$.

end procedure

procedure VERIFY

1. $\mathcal{B}$: $\mathrm{CNOT}(\left|B\right\rangle \left|A \oplus P_A^1\right\rangle) = \left|B\right\rangle \left|B \oplus A \oplus P_A^1\right\rangle$.
2. $\mathcal{B}$: $\left(\left|B \oplus A \oplus P_A^1\right\rangle, \left|m\right\rangle\right) \to \mathcal{T}$ // secured using decoy states.
3. $\mathcal{T}$: $\mathrm{CNOT}\left(\left|T\right\rangle \left|B \oplus A \oplus P_A^1\right\rangle\right) = \left|T\right\rangle \left|T \oplus B \oplus A \oplus P_A^1\right\rangle =: \left|T\right\rangle \left|P_A'\right\rangle$.
4. $\mathcal{T}$: m_T = measure of $\left|m\right\rangle$ along Z basis.
5. $\mathcal{T}$: $t' = G_1(m_T\|k_A)$ and $s' = G_2(m_T\|k_A)$.
6. $\mathcal{T}$: m' = measure of $X^{s'}H^{t'}\left|P_A'\right\rangle$ along Z basis.
7. $\mathcal{T}$: $m' = m_T$? store $(\left|T\right\rangle, \left|P_A'\right\rangle, m_T)$ and **accept** : store $(m_T = \perp)$ and **reject**.

end procedure

procedure PROOF-OF-ORIGIN (on input $\overline{m} \in \{0,1\}^n$)

1. $\mathcal{T}$: $m_T = \perp$? **favour** $\mathcal{A}$: continue.
2. $\mathcal{B}$: $\mathrm{CNOT}(\left|B\right\rangle \left|A \oplus P_A^0\right\rangle) = \left|B\right\rangle \left|B \oplus A \oplus P_A^0\right\rangle$.
3. $\mathcal{B}$: $\left|B \oplus A \oplus P_A^0\right\rangle \to \mathcal{T}$.
4. $\mathcal{T}$: $\mathrm{CNOT}\left(\left|T\right\rangle \left|B \oplus A \oplus P_A^0\right\rangle\right) = \left|T\right\rangle \left|T \oplus B \oplus A \oplus P_A^0\right\rangle =: \left|T\right\rangle \left|P_A''\right\rangle$.
5. $\mathcal{T}$: $\bar{t} = G_1(\overline{m}\|k_A)$ and $\bar{s} = G_2(\overline{m}\|k_A)$.
6. $\mathcal{T}$: m'' = measure of $X^{\bar{s}}H^{\bar{t}}\left|P_A''\right\rangle$ along Z basis.
7. $\mathcal{T}$: $m'' = \overline{m}$? **favour** $\mathcal{B}$: **favour** $\mathcal{A}$.

end procedure

procedure PROOF-OF-RECEIPT (on input $\overline{m} \in \{0,1\}^n$)

1. $\mathcal{T}$: $m_T = \perp$? **favour** $\mathcal{B}$: continue.
2. $\mathcal{T}$: $\bar{t} = G_1(\overline{m}\|k_A)$ and $\bar{s} = G_2(\overline{m}\|k_A)$.
3. $\mathcal{T}$: m' = measure of $X^{\bar{s}}H^{\bar{t}}\left|P_A'\right\rangle$ along Z basis.
4. $\mathcal{T}$: $m' = \overline{m}$? **favour** $\mathcal{A}$: **favour** $\mathcal{B}$.

end procedure

on this matter (type of decoys, number...). As our attacks do not rely on that, we do not make it more explicit in Protocol 2. It is implicit that the protocol is aborted and the signature rejected if the decoy checks fail. Additionally, we note that the message m is sent classically to Bob during step 4 of the SIGN procedure, and then as a quantum state during the VERIFY one. Besides the fact that we do not see the added value of this discrepancy, not protecting the integrity of m in the SIGN procedure enlarges the attack surface.

Regarding the registers and messages that are stored by the parties, we assume that both Bob and the arbitrator keep their GHZ register: Bob stores $(|B\rangle, |A \oplus P_A^0\rangle, m)$ and Trent keeps $(|T\rangle |P'_A\rangle, m_T)$ in case he accepts. We note that Trent must reconstruct the state $|P'_A\rangle$ as $H^{t'} X^{s'} |m'\rangle$ after its measurement during the step 6 of the VERIFY procedure. The protocol does not describe what happens if he rejects the signature (at the end of the VERIFY procedure or during decoy verification). For the sake of simplicity, we assume he stores a rejection token $\perp$ used in the later procedures.

The paper [6] does not specify the procedures PROOF-OF-ORIGIN and PROOF-OF-RECEIPT. Several descriptions are possible here, the simplest being that, since Trent saved m_T at the end of a successful verification procedure, he checks whether m_T is equal to the input $\overline{m}$ of the procedure PROOF-OF-ORIGIN or PROOF-OF-RECEIPT, and gives a judgement in favour of the complainant if and only if $m_T = \overline{m}$. However, this would make useless the storage of quantum states by Bob and Trent at the end of the SIGN and VERIFY procedures, respectively. Based on the statements in [6] that "Bob stores $\{|A \oplus P_A\rangle, m\}$ as Alice's signature", "If Bob refuses that he has ever received the quantum signature, Trent can judge the fact by checking the sequence $\{|P'_A\rangle, m\}$" and "Trent can decode $|P'_A\rangle$ [...] by performing the signature verification", we opt for more complex procedures PROOF-OF-ORIGIN and PROOF-OF-RECEIPT in Protocol 2 that do not compare m_T with $\overline{m}$ (hence that do not require Trent to keep a transcript of the messages that have been signed).

3.2 Analysis of the security claims

The authors of [6] claim that their scheme offers resistance to forgery and non-repudiation. While we have found no attack regarding the forgery, we note that the scheme is probably not information-theoretically secure (ITS) as its security relies on the security of hash functions. As there exist classical schemes to sign classical messages with computational security, this scheme may not be the most efficient solution.

In the following, we focus on security claims regarding disavowal by the sender or the receiver, as well as a false allegation strategy for Alice if equipped with unbounded resources.

Repudiation by Alice. In step 4 of the SIGN procedure, suppose that Alice sends $\big(|\text{gibberish}\rangle, |A \oplus P_A^1\rangle, m\big)$. In that case, the message will first be deemed valid by the arbitrator during the VERIFY procedure. If later Alice repudiates her message and that Bob confronts her with the signature during a PROOF-OF-ORIGIN procedure on input m, the arbitrator will (most probably) reject the signature Bob kept since $|P''_A\rangle$ is not related to m anymore. Therefore, Alice will have successfully disavowed. However, note that this attack does not work if the simpler PROOF-OF-ORIGIN procedure consisting in the single step "$\mathcal{T}$: $\overline{m} = m_T$? **favour** $\mathcal{B}$: **favour** $\mathcal{A}$" is chosen instead, but in which case the quantum information stored by Bob is unused.

False allegation by Alice with unbounded resources. We consider the case of a malicious Alice with unbounded resources. She wishes to find two distinct messages $m, \overline{m}$ yielding the same $|P_A^0\rangle$. For convenience, we introduce the function $G'_2(m,k) = m \oplus G_2(m\|k)$, and recall that $|P_A^0\rangle$ is given by

$$|P_A^0\rangle = H^{G_1(m\|k_A)} X^{G_2(m\|k_A)} |m\rangle = H^{G_1(m\|k_A)} X^{G'_2(m,k_A)} |0\rangle .$$

To succeed in her false allegation attempt, Alice uses her unbounded resources to find two distinct messages $m, \overline{m}$ such that $G_1(m\|k_A) = G_1(\overline{m}\|k_A)$ and $G'_2(m,k_A) = G'_2(\overline{m},k_A)$. Alice can then claim she has sent $\overline{m}$ while she sent m. For random functions G_1 and G'_2, the probability that there exists such messages is non-negligible. To evaluate this, we consider the function

$$F_{k_A} : \{0,1\}^n \to \{0,1\}^{2n}, \quad m \mapsto (G_1(m\|k_A), G'_2(m,k_A)),$$

and we apply the standard birthday argument. The probability that we are evaluating is the collision probability of F_{k_A}:

$$\Pr[\exists\, m, \overline{m} | F_{k_A}(m) = F_{k_A}(\overline{m})] = 1 - \prod_{i=0}^{2^n-1} \left(1 - \frac{i}{2^{2n}}\right) \geqslant 1 - e^{-\frac{2^n(2^n-1)}{2^{2n+1}}}.$$

This probability tends toward $1 - e^{-1/2} \simeq 39\%$ for large n. We note that if Alice has some control on k_A (which is usually the case in standard QKD protocols such as BB84), then she can find such a collision with overwhelming probability.

As for the above repudiation attack, we note that this attack does not work if the simpler Proof-of-Receipt procedure consisting in the single step "$\mathcal{T}$: $\overline{m} = m_T$? **favour** $\mathcal{A}$: **favour** $\mathcal{B}$" is chosen instead.

Disavowal by Bob. During the course of the protocol, the receiver has access to the complete cleartext message before the arbitrator performs his part of the verification. Therefore, if the message is not to his liking, Bob can decide not to forward the message and the quantum states to Trent. This simple disavowal strategy is not considered by the authors as the protocol has not been terminated. However, this is a winning strategy that we consider valid. Note that in the most recent AQS schemes, this issue is resolved by delaying the full disclosure of the message to the recipient at the last step of the verification (see e.g., [31, 11, 30] or Protocol 1).

4 Cryptanalysis of a scheme based on controlled teleportation

In this section, we review the AQS protocol of Lu et al. [16] described in Protocol 3, which is based on a five-qubit entangled system, and we discuss its security claims.

4.1 Description of the protocol

Two encryptions. This scheme involves two encryption methods. The first one is the standard quantum one-time pad [2] defined as

$$E_k |M\rangle = \bigotimes_{i=1}^{n} X^{k_{2i-1}} Z^{k_{2i}} |M_i\rangle$$

for a key $k \in \{0,1\}^{2n}$ and an n-qubit state $|M\rangle = \bigotimes_{i=1}^{n} |M_i\rangle$ (and extended linearly). The other encryption method used in Protocol 3 is that of [31], defined as

$$R_k |M\rangle = \bigotimes_{i=1}^{n} X^{k_i} Z^{k_i \oplus 1} |M_i\rangle$$

for a key $k \in \{0,1\}^{n}$ and an n-qubit state $|M\rangle = \bigotimes_{i=1}^{n} |M_i\rangle$ (and extended linearly).

Two bases and a five-qubit state. The Bell state basis is denoted by $\{|\Psi_\alpha\rangle\}_{\alpha=0}^{3}$ and defined as

$$|\Psi_0\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle), \quad |\Psi_\alpha\rangle = \left(Z^{b_1} \otimes X^{b_2}\right) |\Psi_0\rangle,$$

where $b_1 \| b_2$ is the binary representation of α. One also needs the von Neumann basis $\{|\chi_\beta\rangle\}_{\beta=0}^{7}$ defined by

$$|\chi_0\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle), \quad |\chi_\beta\rangle = \left(X^{b_1} \otimes X^{b_2} \otimes Z^{b_3}\right) |\chi_0\rangle,$$

where $b_1 \| b_2 \| b_3$ is the binary representation of β (the states $(\chi_0, \chi_1, \chi_2, \chi_3, \chi_4, \chi_5, \chi_6, \chi_7)$ are denoted in [16] as $(\chi^1, \chi^2, \chi^5, \chi^6, \chi^7, \chi^8, \chi^3, -\chi^4)$).

Protocol 3 AQS protocol of [16], based on controlled teleportation.

procedure INIT

1. $\mathcal{T}$: $k_A \to \mathcal{A}$.
2. $\mathcal{T}$: $k_B \to \mathcal{B}$.

end procedure

procedure SIGN (on 4 copies $|M^0\rangle = |M^1\rangle = |M^2\rangle = |M^3\rangle$ of $|M\rangle = \bigotimes_{i=1}^{n} |M_i\rangle$)

1. $\mathcal{A}$: $|R_A\rangle = R_{k_A} |M^0\rangle$.
2. $\mathcal{A}$: $|\Omega\rangle = \bigotimes_{i=1}^{n} |\Omega_i\rangle$ with $|\Omega_i\rangle = |M_i^1\rangle \otimes |\xi\rangle$.
3. $\mathcal{A}$: share $|\Omega\rangle_{ATAATB} = \bigotimes_{i=1}^{n} |\Omega_i\rangle_{ATAATB}$ into $|\Omega_A\rangle, |\Omega_B\rangle$ and $|\Omega_T\rangle$.
4. $\mathcal{A}$: $|\Omega_T\rangle \to \mathcal{T}$, $|\Omega_B\rangle \to \mathcal{B}$ // secured using decoy states.
5. $\mathcal{A}$: $|\omega_A\rangle = |\Omega_A\rangle$ after measurement according to the basis $\{|\chi_\beta\rangle\}_{\beta=0}^{7}$.
6. $\mathcal{A}$: $|S\rangle = E_{k_A}(|R_A\rangle, |\omega_A\rangle)$.
7. $\mathcal{A}$: $(|M^2\rangle, |M^3\rangle, |S\rangle) \to \mathcal{B}$.

end procedure

procedure VERIFY

1. $\mathcal{B}$: $|Y_{BT}\rangle = E_{k_B}(|S\rangle, |M^2\rangle) \to \mathcal{T}$ // without decoy.
2. $\mathcal{T}$: recover $|S\rangle, |M^2\rangle$; from $|S\rangle$, recover $|R_A\rangle, |\omega_A\rangle$.
3. $\mathcal{T}$: $|\omega_T\rangle = |\Omega_T\rangle$ after measurement according to Bell states basis $\{|\Psi_\alpha\rangle\}_{\alpha=0}^{3}$.
4. $\mathcal{T}$: $|R_A\rangle = R_{k_A} |M^2\rangle$? $\theta = 1 : \theta = 0$ // $\mathcal{T}$ has now 2 copies of $|R_A\rangle$.
5. $\mathcal{T}$: $|S^0\rangle = |S^1\rangle = E_{k_A}(|R_A\rangle, |\omega_A\rangle)$ // $\mathcal{T}$ uses his 2 copies of $|R_A\rangle$.
6. $\mathcal{T}$: choose a hash function H and calculate $\mathrm{H}(|S^1\rangle)$.
7. $\mathcal{T}$: $|Y_{TB}\rangle = E_{k_B}(|S^0\rangle, \mathrm{H}(|S^1\rangle), |\omega_A\rangle, |\omega_T\rangle, |\theta\rangle) \to \mathcal{B}$ // without decoy.
8. $\mathcal{T}$: store (ω_A, θ).
9. $\mathcal{B}$: recover $|S^0\rangle$, $\mathrm{H}(|S^1\rangle)$, $|\omega_A\rangle$, $|\omega_T\rangle$, θ; $\theta = 0$? **reject** : continue.
10. $\mathcal{B}$: recover $|M^1\rangle$ from $|\Omega_B\rangle$ using $|\omega_A\rangle$ and $|\omega_T\rangle$.
11. $\mathcal{B}$: $|M^1\rangle = |M^3\rangle$? continue : **reject**.
12. $\mathcal{B}$: $\mathrm{H}(|S^0\rangle) = \mathrm{H}(|S^1\rangle)$? store $(|M^1\rangle, |S^0\rangle, |S^1\rangle)$ and **accept** : **reject**.

end procedure

procedure PROOF-OF-ORIGIN (on input $|\overline{M}\rangle = \bigotimes_{i=1}^{n} |\overline{M_i}\rangle$)

1. $\mathcal{T}$: $\theta = 1$? continue : **favour** $\mathcal{A}$.
2. $\mathcal{B}$: $|S^0\rangle \to \mathcal{T}$.
3. $\mathcal{T}$: $|\overline{R_A}\rangle = R_{k_A} |\overline{M}\rangle$.
4. $\mathcal{T}$: $E_{k_A}\left(|\overline{R_A}\rangle, |\omega_A\rangle\right) = |S^0\rangle$? **favour** $\mathcal{B}$: **favour** $\mathcal{A}$.

end procedure

procedure PROOF-OF-RECEIPT (on input $|\overline{M}\rangle = \bigotimes_{i=1}^{n} |\overline{M_i}\rangle$)

1. $\mathcal{T}$: $\theta = 1$? continue : **favour** $\mathcal{B}$.
2. $\mathcal{B}$: $|S^1\rangle \to \mathcal{T}$.
3. $\mathcal{T}$: $|\overline{R_A}\rangle = R_{k_A} |\overline{M}\rangle$.
4. $\mathcal{T}$: $E_{k_A}\left(|\overline{R_A}\rangle, |\omega_A\rangle\right) = |S^1\rangle$? **favour** $\mathcal{A}$: **favour** $\mathcal{B}$.

end procedure

The five-qubit state $|\xi\rangle$ used in the protocol is

$$|\xi\rangle = \frac{1}{2}\left(|100\rangle|\Psi_0\rangle + |111\rangle|\Psi_1\rangle + |001\rangle|\Psi_2\rangle + |010\rangle|\Psi_3\rangle\right).$$

This state is used in a teleportation-like protocol of a qubit $|\varphi\rangle$. First the joint state $|\Omega\rangle = |\varphi\rangle \otimes |\xi\rangle$ is created. It is composed of six entangled states which are shared between the parties. We note $|\Omega\rangle_{ATAATB}$ to indicate this sharing. This notation means that Alice has the first, the third and the fourth registers composing $|\Omega\rangle$, and similarly for Trent and Bob. The resulting states are noted $|\Omega_A\rangle, |\Omega_T\rangle$ and $|\Omega_B\rangle$. Second, Alice and the arbitrator both perform a measurement of their state and note the respective collapsed states $|\omega_A\rangle$ and $|\omega_T\rangle$. Finally, after the measurement of the first five registers, Bob can recover the initial message $|\varphi\rangle$ by applying the correct transformations (depending on $|\omega_A\rangle$ and $|\omega_T\rangle$) on the sixth register that he owns, $|\Omega_B\rangle$.

As in the previous protocol, some quantum communications are secured using decoy states. If the decoy checks fail upon reception of a quantum message, the signature is rejected and Trent stores $\theta = 0$ before aborting. As our attack does not exploit a weakness related to these decoys, we only briefly mention them in the protocol description. In particular, we do not describe the exchange of polynomials in the INIT procedure aimed at selecting the state of these decoys. Note that only the states that share the teleportation state $|\Omega\rangle$ are protected with decoys, and not all quantum messages.

Ambiguities. We note that there are multiple ambiguities in [16]. The first one concerns the number of copies of the initial message required. The article mentions that three copies are needed. However, as can be seen from Protocol 3, one copy of the message is used to construct $|R_A\rangle$, one copy is entangled with the five-qubit states, and two are sent to Bob at the end of the signing procedure. Moreover, the protocol does not take into account that multiple copies should be used for the equality tests in steps 4, 11 and 12 of the VERIFY procedure.

A second ambiguity concerns the hash function H and its usage. The protocol requires the arbitrator to use a hash function on the signature $|S^1\rangle$. This hash is claimed to ensure the integrity of the signature. However, the nature of H is not specified in [16]. We consider two cases: either $\mathrm{H}(|S\rangle)$ is a classical bit string computed from $|S\rangle$, or H is a quantum unitary operation applied on $|S\rangle$ such that $|h\rangle = \mathrm{H}(|S\rangle)$ is a quantum state. In the first case, Bob would be unable to send $|S^0\rangle$ to Trent in a PROOF-OF-ORIGIN procedure. Another problem is that such a classical string should be obtained after a measurement of $|S\rangle$. As this state is unknown to Bob, the result of H would be random, and likely leading to a rejection in step 12 of the SIGN procedure. For these reasons, we reject this first option and consider the second one, where H is a quantum unitary operation. Note that in that case, Trent does not hold a copy of $\mathrm{H}(|S^1\rangle)$ after the VERIFY procedure, which causes issues for Trent to judge whether Bob has indeed received the message as explained below. Furthermore, since H is unitary, it is easily invertible, which is anyway mandatory for Bob to re-obtain $|S^0\rangle$ as a signature at the last step of the VERIFY procedure. In addition, the scheme specifies that the arbitrator unilaterally chooses the hash function. We assume that this choice is publicly announced to Bob after the reception of $|Y_{TB}\rangle$. The length of the hash output is not given. As a unitary, it cannot be smaller than the one of the input and we therefore suppose it is length-preserving.

It is unclear whether classical information (e.g., the bit θ) is sent classically or as a quantum state ($|\theta\rangle$). We opt for the second option, as the authors use the quantum one-time pad which is defined for quantum states, and assume that the recipient implicitly measures the quantum state in the appropriate basis to recover classical information. Note that this choice does not influence our attacks.

A fourth ambiguity concerns the keys k_A and k_B. In the description of the scheme in [16], their length is not given explicitly. In Appendix A of that paper, it appears in an example that $k_A, k_B \in \{0,1\}^{2n}$. However, in step 6 of the SIGN procedure, Alice computes $E_{k_A}(|R_A\rangle, |\omega_A\rangle)$ where $|R_A\rangle$ is n-qubit long and $|\omega_A\rangle$ is $3n$-qubit long. This requires k_A to be at least $8n$-bit long. Moreover, in step 7 of the VERIFY procedure, Trent computes $E_{k_B}(|S^0\rangle, \mathrm{H}(|S^1\rangle), |\omega_A\rangle, |\omega_T\rangle, |\theta\rangle)$ where $|S^0\rangle$ and $\mathrm{H}(|S^1\rangle)$ are both $4n$-qubit long, $|\omega_A\rangle$ and $|\omega_T\rangle$ have respective length $3n$ and $2n$ qubits, and θ is a single bit encoded in a qubit. One thus needs that k_B is at least $(26n+2)$-bit long. In addition, the same key k_A is used for two

different encryptions, $|R_A\rangle$ and $|S\rangle$ and, similarly, k_B is used in $|Y_{BT}\rangle$ and $|Y_{TB}\rangle$. All these encryptions require a different key length (i.e., n and $8n$ for k_A, and $10n$ and $26n+2$ bits for k_B). In order to use independent keys and increase security, we assume that one chooses $k_A \in \{0,1\}^{9n}$ and $k_B \in \{0,1\}^{36n+2}$, where their first part is used for the first encryption and the last part for the second one.

The authors of [16] do not describe the procedures Proof-of-Origin and Proof-of-Receipt. We make a proposition for both of them in Protocol 3. This requires that Trent saves θ, k_A and the measurement results $|\omega_A\rangle$, and that Bob keeps $|S^0\rangle$, $|S^1\rangle$ and $|M\rangle$. Our proposition for the Proof-of-Origin procedure is quite intuitive, but we did not find a good solution for the Proof-of-Receipt procedure. Indeed, similarly to Protocol 1, we could have required Trent to check only $\theta = 1$, but this would have led to a trivial false allegation attack by Alice who could have pretended that Bob received any message since this procedure does not depend on $|\overline{M}\rangle$. We rather opt for another solution, which is similar to the Proof-of-Origin procedure, but which is trivially deniable by Bob as he could send a garbage state in step 2.

Another issue concerns the space of messages that Alice can sign. Similarly to Protocol 1, Alice cannot sign all n-qubit messages, but only those of the factorised form $|M\rangle = \bigotimes_{i=1}^{n} |M_i\rangle$. This restriction is due to the quantum teleportation which is realised qubit-wise. However, it is not clear to us if the protocol is claimed to be protected against adversaries that do not follow this requirement, for example, if Alice starts the Sign procedure with an entangled message. Although we think these are legitimate attacks, we make sure to use factorised messages in our attacks.

4.2 Analysis of the security claims

The authors of [16] claim that their protocol achieves unforgeability and non-repudiation for both the sender and the receiver. In this section, we show that these claims are not valid.

Repudiation and false allegation by Alice. In the beginning of the Sign procedure, Alice chooses $|M^0\rangle = |M^2\rangle \neq |M^1\rangle = |M^3\rangle$. In that case, Bob gets two different messages $|M^2\rangle$ and $|M^3\rangle$ at the end of the Sign procedure, but all checks succeed in the Verify procedure. Since $|S^0\rangle$ is computed from $|M^2\rangle$, Bob will be able to prove in a Proof-of-Origin procedure that he received it, but not that he received $|M^3\rangle$, which is the copy of the message that he could used. This strategy also allows Alice to prove that Bob received $|M^2\rangle$ in a Proof-of-Receipt procedure, although he believes he received $|M^3\rangle$.

Disavowal by Bob. Bob has two trivial disavowal strategies. The first is similar to that of Protocol 2. As Bob receives the full message in the last step of the Sign procedure, he can decide to stop the protocol. The solution to this issue is to design a protocol for which the delivery of the intended message is delayed until the end of the protocol, as in [31].

For the second one, Bob simply sends another state than $|S^1\rangle$ in step 2 of the Proof-of-Receipt procedure.

Forgery by Bob. We show here a forgery attack that Bob can apply to any qubit of the message $|M\rangle$. Let us suppose that Bob attacks the first qubit. At the end of the Sign procedure, Bob receives from Alice two copies $|M^2\rangle = |M^3\rangle$ of $|M\rangle$, as well as $|S\rangle = E_{k_A}(|R_A\rangle, |\omega_A\rangle)$ where $|R_A\rangle = R_{k_A}|M^0\rangle$. Let a be the first bit of the part of k_A used to compute $|R_A\rangle$, and b and c be the first and second bits of the part of k_A used to compute $|S\rangle$. In view of our assumptions we made to resolve the ambiguities, one has $a = (k_A)_1$, $b = (k_A)_{n+1}$ and $c = (k_A)_{n+2}$. The first qubit of $|R_A\rangle$ is $X^a Z^{a\oplus 1}|M_1^0\rangle$, and the first qubit of $|S\rangle$ is $X^b Z^c X^a Z^{a\oplus 1}|M_1^0\rangle$.

We suppose that, at the start of the Verify procedure, Bob applies an X gate to the first qubit of $|M^2\rangle$ and to the first qubit of $|S\rangle$. They thus become $X|M_1^2\rangle$ and $X^{b\oplus 1} Z^c X^a Z^{a\oplus 1}|M_1^0\rangle$, respectively, while the other qubits are left unchanged. When Trent

decrypts $|S\rangle$ to get $|R_A\rangle$, he obtains the correct $|R_A\rangle$ except for the first qubit which is now

$$\begin{aligned} Z^c X^b X^{b\oplus 1} Z^c X^a Z^{a\oplus 1} |M_1^0\rangle &= Z^c X Z^c X^a Z^{a\oplus 1} |M_1^0\rangle \\ &= (-1)^c X Z^c Z^c X^a Z^{a\oplus 1} |M_1^0\rangle \\ &= (-1)^c X^{a\oplus 1} Z^{a\oplus 1} |M_1^0\rangle . \end{aligned}$$

In step 4 of the verification procedure, when Trent computes $R_{k_A} |M^2\rangle$, he obtains the correct state except for the first qubit which is now

$$X^a Z^{a\oplus 1} X |M_1^2\rangle = (-1)^{a\oplus 1} X^{a\oplus 1} Z^{a\oplus 1} |M_1^2\rangle .$$

Therefore, as the prefactors only act as an unimportant global phase, the test in step 4 will be passed and give $\theta = 1$. The states $|S^0\rangle = |S^1\rangle$ computed by Trent will correspond to the modified state $|R_A\rangle$ and Bob will accept the signature. Then, in a Proof-of-Origin procedure, Bob can prove that Alice sent him the message $|\overline{M}\rangle = X_1 |M\rangle$ (i.e., the message $|M\rangle$ to which an X gate has been applied to the first qubit), which constitutes a forgery. Indeed, $\theta = 1$ and the test $E_{k_A}(|\overline{R_A}\rangle, |\omega_A\rangle) = |S^0\rangle$ in step 4 will pass for the same reason as above. Bob therefore has a forgery attack that works with probability 1. This strategy easily generalises: for any operator U that applies a (possibly different) Pauli gate on each qubit, $U = \bigotimes_{i=1}^n \sigma_{\alpha_i}$, Bob can obtain a successful forgery by applying U on $|M^2\rangle$ and on the first n qubits of $|S\rangle$.

5 Cryptanalysis of a scheme without entangled states

In this section, we review the claims of the AQS scheme of Zhang et al. [29], which aims at providing quantum signatures of classical messages without the use of entangled states.

After having recalled the scheme, we review some of the claims made by the authors, among which the claimed impossibility of forgery by Bob, as well as the information-theoretical security of the scheme.

5.1 Description of the protocol

In this section, we recall in a synthetic way the protocol of [29]. One important feature of this protocol is the use of a so-called quantum one way function, denoted F_k for a key $k \in \{0,1\}^\ell$, and which maps classical bit strings to quantum states. It is based on three functions $f, g, h \colon \{0,1\}^* \to \{0,1\}^n$, which are used to create n-bit strings $f_k(m) = f(k\|m)$, $g_k(m) = g(k\|m)$ and $h_k(m) = h(k\|m)$. The n-qubit quantum state created is

$$F_k(m) = H^{g_k(m)} \widetilde{Y}^{h_k(m)} |f_k(m)\rangle$$

where $\widetilde{Y} = iY = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Due to the classical nature of $f_k(m)$, the effect of $\widetilde{Y}$ is, up to a factorisable phase, the same as X. Therefore, we may write, omitting the unimportant global phase $(-1)^{\sum_i h_i(k\|m)\cdot(1+f_i(k\|m))}$,

$$F_k(m) = H^{g_k(m)} X^{h_k(m)} |f_k(m)\rangle = H^{g_k(m)} X^{h_k(m)} X^{f_k(m)} |0\rangle = H^{g_k(m)} X^{\widetilde{f}_k(m)} |0\rangle ,$$

where $\widetilde{f} \colon \{0,1\}^* \to \{0,1\}^n$ is defined by the bit-wise XOR $\widetilde{f}_k(m) = \widetilde{f}(k\|m) = f_k(m) \oplus h_k(m)$. Similarly, one may also write, omitting a global phase factor $(-1)^{\sum_i f_i(k\|m)}$,

$$F_k(m) = H^{g_k(m)} \widetilde{Y}^{h_k(m)} \widetilde{Y}^{f_k(m)} Z^{f_k(m)} |0\rangle = H^{g_k(m)} \widetilde{Y}^{\widetilde{f}_k(m)} |0\rangle .$$

We will use both rewrites during the analysis of the security claims. The protocol is described in Protocol 4.

Ambiguities. A first ambiguity in [29] concerns the use of decoy states. It is mentioned that "sufficient decoy particles" should be used, without specifying how many. It is also mentioned that these decoy particles are inserted randomly in the states, but nothing is said about the distribution of the positions. Moreover, the authors allow the decoy checks to

Protocol 4 AQS protocol of [29] that does not require entangled states.

procedure INIT

1. $\mathcal{A} \leftrightarrow \mathcal{T}$: $k \in \{0,1\}^{\ell}$ (where $\ell \gg n$).
2. $\mathcal{A} \leftrightarrow \mathcal{B}$: $x, y \in \{0,1\}^{n}$.

end procedure

procedure SIGN (on a classical message $m \in \{0,1\}^{*}$)

1. $\mathcal{A}$: $|M\rangle = F_k(m)$.
2. $\mathcal{A}$: $|S\rangle = \widetilde{Y}^{x} H^{y} |M\rangle$.
3. $\mathcal{A}$: $(m, |S\rangle) \rightarrow \mathcal{B}$ // secured using decoy states for $|S\rangle$.

end procedure

procedure VERIFY

1. $\mathcal{B}$: $|M'\rangle = H^{y}(-\widetilde{Y})^{x} |S\rangle$.
2. $\mathcal{B}$: $(m, |M'\rangle) \rightarrow \mathcal{T}$ // secured using decoy states for $|M'\rangle$.
3. $\mathcal{T}$: $|f'\rangle = (-\widetilde{Y})^{h_k(m)} H^{g_k(m)} |M'\rangle$.
4. $\mathcal{T}$: f' = measure of $|f'\rangle$ according to $\{|0\rangle, |1\rangle\}$.
5. $\mathcal{T}$: $f' = f(k\|m)$? $v_T = 1$ and stores (m, f') : $v_T = 0$; $v_T \rightarrow \mathcal{P}_b$.
6. $\mathcal{B}$: $v_T = 1$? **accept** : **reject**.

end procedure

procedure PROOF-OF-ORIGIN (on input $\overline{m} \in \{0,1\}^{*}$)

1. $\mathcal{T}$: $v_T = 1$? continue : **favour** $\mathcal{A}$.
2. $\mathcal{T}$: $(m, f') = (\overline{m}, f(k\|\overline{m}))$? **favour** $\mathcal{B}$: **favour** $\mathcal{A}$.

end procedure

procedure PROOF-OF-RECEIPT (on input $\overline{m} \in \{0,1\}^{*}$)

1. $\mathcal{T}$: $v_T = 1$? continue : **favour** $\mathcal{B}$.
2. $\mathcal{T}$: $m = \overline{m}$? **favour** $\mathcal{A}$: **favour** $\mathcal{B}$.

end procedure

fail up to a certain threshold, but do not explain how to determine it. Also, it is not made explicit what happens if the decoy verification do not succeed. In the following, we assume that in the case of a rejection due to a failure of the decoy checks, the parties discard the message and abort, while Trent further stores $v_T = 0$.

Another ambiguity concerns the procedures Proof-of-Origin and Proof-of-Receipt. No precise descriptions of them were given in [29], but we have inferred them from Section 4.2.3 therein. Another possibility for the Proof-of-Receipt procedure would be to check $(m, f') = (\overline{m}, f(k\|\overline{m}))$ in step 2, which would not change the validity of our attacks.

5.2 Analysis of the security claims

The authors of [29] claim that the protocol is secure against existential forgeries and non-repudiation by the sender and the receiver. It is not clear to us whether these security properties are claimed to hold computationally (i.e., for a bounded adversary), or are claimed information-theoretically secure (ITS, i.e., for an unbounded adversary). For instance, Theorem 6 in [29] involves a random oracle whereas Theorem 3 states that "the proposed quantum signature has the information-theoretical security".

Nevertheless, as the scheme can only be used to sign classical messages, its utility would be questionable if it did not provide information-theoretic security. Indeed, classical public-key cryptography (in particular post-quantum cryptography) provides computationally-secure ways of obtaining authenticity and non-repudiation of origin for classical messages that are much easier to implement. We expose here attacks on the non-repudiation and forgery that an adversary with unbounded resources can mount, showing that the scheme is not ITS.

Choice of hash functions and security lemmas. The protocol does not indicate any restriction on the choice of f, g and h, except that they should be one-way hash functions with uniform output. However, one-wayness is only relevant for computational security. Without further indication, some of the stated results are incorrect, one of which being the Theorem 4 of [29], which states that for any message m, the density operator of $F_K(m)$ is uniform (where K is a uniform random variable on $\{0,1\}^\ell$ representing the key k), that is, under proper normalisation, it equals the identity matrix[2]. As a simple example, let us analyse the choice $f = h$, which could be made for ease of implementation. With such a choice, one has $\widetilde{f} = 0$ and the density operator of $F_K(m)$ is $\rho = \bigotimes_{i=1}^{n} \rho_i$ where

$$\rho_i = \frac{1}{2} \sum_{g_i \in \{0,1\}} H^{g_i} |0\rangle \langle 0| H^{g_i} = \frac{1}{4} \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix},$$

which differs from the identity operator. In the following, we do not expose further the attacks that originate from such degenerate choices, but we stress that one should take care of interactions between primitives when no restrictions or conditions are imposed between each other.

Repudiation and false allegation by Alice with unbounded resources. If Alice wants to send the message m and wants to be able to repudiate it afterwards, she uses her unbounded resources to find another $m' \in \{0,1\}^*$ such that $g_k(m) = g_k(m')$ and $\widetilde{f}_k(m) = \widetilde{f}_k(m')$ (i.e., $f_k(m) \oplus h_k(m) = f_k(m') \oplus h_k(m')$). Such a second preimage exists almost certainly as the message space is infinite ($m \in \{0,1\}^*$).

Then, during step 2 of the Verify procedure, she intercepts the message $(m, |M'\rangle)$ and replaces it with $(m', |M'\rangle)$. Note that this does not interfere with the decoys as Alice only changes the classical part of the communication. Trent's check $f' = f(k\|m')$ will then pass, and he will store (m', f'). When Bob wants to prove that he received the message m from Alice, the second check in the Proof-of-Origin procedure will fail and Alice will succeed her repudiation strategy. Moreover, Alice can prove that Bob received m' via a Proof-of-Receipt procedure, which is a false allegation.

[2]While the statement of Theorem 4 is that the output of $F_K(\cdot)$ have same density operator for different input, this is not the property of interest (e.g., a trivial $F_K(\cdot) = 0$ would have this feature). Rather, the uniformity of the density operator is what is actually used in the following Theorem 5, hence we discuss it here.

This attack can be more sophisticated. For instance, suppose that Alice finds an m' such that the equalities $g_k(m) = g_k(m')$ and $\widetilde{f}_k(m) = \widetilde{f}_k(m')$ hold for every bit except the first one. She can suppose that the first qubit of $|M'\rangle$ is not a decoy state (which has a reasonable success rate, depending on the number of decoy particles). Then, she can substitute m by m' and the first qubit of $|M'\rangle$ by the first qubit of $F_k(m')$ so that Trent's check will pass and $(m', f(k\|m'))$ will be stored, making Alice succeed in her attack.

Notice that, according to [29], the key k is shared between Alice and Trent via a quantum key distribution protocol, allowing Alice to have some control over the key k, thus making the task of finding a (quasi-)collision (m, m') easier.

Forgery by Bob with a single signature, for specific hashes. The first forgery we present uses the supposition that the functions f, g and h are of the form

$$f(k\|m) = f_a(k) \oplus f_b(m), \quad g(k\|m) = g_a(k) \oplus g_b(m), \quad h(k\|m) = h_a(k) \oplus h_b(m),$$

with f_a, f_b, g_a, g_b, h_a and h_b chosen independently and at random in the set of functions $\{0,1\}^* \to \{0,1\}^n$. Although this assumption might be considered as a very specific choice of hash functions, such choice could be made for performance reasons[3]. For example, the parties could use the function $f(k\|m) = \mathrm{SHA3}(0\|k) \oplus \mathrm{SHA3}(1\|m)$ in order to pre-compute the hash of a very long key.

We now detail the forgery in the case where the hashes f, g and h (and therefore $\widetilde{f}$) are of the aforementioned form. Given a message m, one has

$$\begin{aligned} |M\rangle &= H^{g(k\|m)} \widetilde{Y}^{\widetilde{f}(k\|m)} |0\rangle \\ &= H^{g_a(k)\oplus g_b(m)} \widetilde{Y}^{\widetilde{f}_a(k)\oplus \widetilde{f}_b(m)} |0\rangle . \end{aligned}$$

To forge a valid signature on the message $\overline{m}$, Bob computes $g_b(m) \oplus g_b(\overline{m})$ and $\widetilde{f}_b(m) \oplus \widetilde{f}_b(\overline{m})$ and modifies the state $|M'\rangle$ computed in step 1 of the Verify procedure to

$$\begin{aligned} |\overline{M}\rangle &= H^{g_b(m)\oplus g_b(\overline{m})} \widetilde{Y}^{\widetilde{f}_b(m)\oplus \widetilde{f}_b(\overline{m})} |M'\rangle \\ &= H^{g_b(m)\oplus g_b(\overline{m})} \widetilde{Y}^{\widetilde{f}_b(m)\oplus \widetilde{f}_b(\overline{m})} H^{g_a(k)\oplus g_b(m)} \widetilde{Y}^{\widetilde{f}_a(k)\oplus \widetilde{f}_b(m)} |0\rangle \\ &= \pm H^{g_a(k)\oplus g_b(\overline{m})} \widetilde{Y}^{\widetilde{f}_a(k)\oplus \widetilde{f}_b(\overline{m})} |0\rangle , \end{aligned}$$

where the ± 1 prefactor is the result of the anticommutation relation between $\widetilde{Y}$ and H, $\widetilde{Y}H = -H\widetilde{Y}$, but that only acts as a global phase, hence without any measurable effect on the result. In step 2 of the Verify procedure, Bob sends $(\overline{m}, |\overline{M}\rangle)$ to Trent. This will pass all the arbitrator's verifications, hence resulting in a successful universal forgery.

Forgery by Bob with polynomially many signatures and unbounded resources. This final forgery requires adversarial Bob to ask a polynomial number of signatures to the signing party under a reused key k, and assumes that Bob is allowed to launch a Verify procedure for each of them. We note that key reuse is explicitly allowed by the authors of [29] in their formulation of the security of the scheme against forgery. As in the example provided in [29], we fix the parameter $\ell = 2n \gg n$.

Bob's strategy consists in modifying the first qubit of $|M'\rangle$ when he sends it to the arbitrator in step 2 of the Verify procedure. Depending on the statistics of the acceptance or rejection of the signature, Bob learns the first bit of the hash functions with high probability. Repeating that process for each qubit, he can learn the whole strings $g_k(m)$ and $\widetilde{f}_k(m) = f_k(m) \oplus h_k(m)$. Using his unbounded resources, Bob can find the key k and forge new messages.

Let us describe this attack in detail for $F_k(m) = |M\rangle = \bigotimes_{i=1}^n |M_i\rangle$. By the construction of F_k, one has $|M_i\rangle = H^{g_i} X^{\widetilde{f}_i} |0\rangle$, where g_i denotes the i^{th} bit of $g_k(m)$, and similarly for $\widetilde{f}_i$. Upon reception of $(m, |S\rangle)$ and recovery of $|M\rangle$, Bob replaces the first qubit of $|M\rangle$ by $|0\rangle$, and thus sends to the arbitrator the message $(m, |0\rangle \otimes \bigotimes_{i=2}^n |M_i\rangle)$. If $g_1 = 0$, Trent will accept or reject with probability 1, depending on $\widetilde{f}_1$ being 0 or 1, respectively. If $g_1 = 1$, Trent

[3] Additionally, we add that a security claim should not depend on unwritten assumptions on the primitives that are used.

$(g_i, \widetilde{f}_i)$	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
Bob sends $(\bigotimes_{j=1}^{i-1} \lvert M_j\rangle) \otimes \lvert 0\rangle \otimes (\bigotimes_{j=i+1}^{n} \lvert M_j\rangle)$	1	0	1/2	1/2
Bob sends $(\bigotimes_{j=1}^{i-1} \lvert M_j\rangle) \otimes \lvert +\rangle \otimes (\bigotimes_{j=i+1}^{n} \lvert M_j\rangle)$	1/2	1/2	1	0

Table 1: Probability of acceptance by the arbitrator of the modified $|M\rangle$ when Bob modifies the i^{th} qubit of $|M\rangle$ by $|0\rangle$ or $|+\rangle$, depending on the bits g_i and $\widetilde{f}_i$.

will accept or reject with equiprobability. Bob repeats this process ν times, each requiring a new signature from Alice of the same message. Afterwards, Bob redoes the same scenario but by changing the first qubit by $|+\rangle$. In that case, for $g_1 = 1$, the arbitrator accepts or rejects with certainty if $\widetilde{f}_1 = 0$ or $\widetilde{f}_1 = 1$ respectively, and his result is unpredictable if $g_1 = 0$. Table 1 summarises the various cases.

After 2ν messages, there is a probability $\frac{1}{2^{\nu-1}}$ that Bob cannot exactly determine the bits g_1 and $\widetilde{f}_1$. Indeed, the only way Bob will not be able to certainly determine g_1 and $\widetilde{f}_1$ is when the events with $1/2$ probability of acceptance all give the same output. This occurs with probability $2^{-(\nu-1)}$, and in this case, Bob can guess the correct values of g_1 and $\widetilde{f}_1$ with $\frac{1}{2}$ probability. In other words, Bob knows the correct bits with probability $1 - \frac{1}{2^{\nu}}$. By doing this strategy for all qubits, Bob learns the values of $g(k\|m)$ and $\widetilde{f}(k\|m)$ with probability $(1 - 2^{-\nu})^n$. Choosing $\nu = n$, this probability tends to 1 for large messages, which is the case studied by the authors ($n = 128$). This adversarial strategy only requires a polynomial amount of messages as it requires at most $2\nu \cdot n = 2n^2$ signatures.

With the knowledge of $g(k\|m)$ and $\widetilde{f}(k\|m)$, the adversarial Bob, with unbounded computing capacities, can obtain the key k with large probability. Indeed, by uniformity, there are $\frac{2^{\ell+|m|}}{2^n}$ preimages of $g(k\|m)$, among which only about $2^{\ell-n}$ end with the correct message m. Each of these preimages for g gives the correct $\widetilde{f}(k\|m)$ with probability 2^{-n}. Hence, with unbounded resources, Bob finds the key k with large probability (as $\ell = 2n$), and can mount any forgery of his choice[4].

This attack also works for longer keys whose size is polynomial in the message length. Indeed, Bob extracts $2n$ bits of information in $2n^2$ signatures of a fixed message, and can iterate this process with different messages, hence obtaining the key in a polynomial amount of queries (with an exponentially low probability of failure).

Disavowal by Bob. If Bob receives a pair $(m, |S\rangle)$ that he dislikes, he can simply discard it (or retain it for later). This issue, due to the message m being fully available to Bob at the end of the Sign procedure, has already been discussed in Subsections 3.2 and 4.2. Note that many of the recent AQS schemes delay the full disclosure of the message to the recipient until the last step of verification, after Bob has communicated with the arbitrator (e.g., [31, 11, 30]).

Even if the trivial disavowal strategy is forbidden, a corollary of the forgery detailed above is that a malicious Bob with unbounded resources can disavow messages. Indeed, once he has obtained the key k, Bob can modify the message to his liking and forge its signature so that Trent publishes $v_T = 1$. A Proof-of-receipt started by Alice with her message will fail, and Bob will have succeeded in his repudiation.

6 Conclusion

In this article, we have shown that several recent AQS schemes, designed to provide secure signatures with an arbitrator, suffer from security flaws. In many cases, the core AQS security features (unforgeability and the impossibility of disavowal by either sender or receiver) are not met. Additionally, we have often identified specification ambiguities that render the claimed

[4]To phrase it differently, both $g(k\|m)$ and $\widetilde{f}(k\|m)$ give Bob n bits of information on the key, which is of size $\ell = 2n$.

guarantees difficult to verify. These gaps in the specification may even lead to unforeseen attacks. This enlightens one more time the necessity of precise descriptions of algorithms and protocols, precise security definitions and comprehensive proofs thereof.

References

[1] H. Barnum, C. Crépeau, D. Gottesman, A.D. Smith and A. Tapp, Authentication of quantum messages, *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, 449–458 (2002).

[2] P.O. Boykin and V. Roychowdhury, Optimal encryption of quantum bits, *Physical Review A* **67**, 042317 (2003).

[3] H. Buhrman, R. Cleve, J. Watrous and R. de Wolf, Quantum fingerprinting, *Physical Review Letters* **87**, 167902 (2001).

[4] J.W. Choi, K.-Y. Chang and D. Hong, Security problem on arbitrated quantum signature schemes, *Physical Review A* **84**, 062330 (2011).

[5] M. Curty and N. Lütkenhaus, Comment on "Arbitrated quantum-signature scheme", *Physical Review A* **77**, 046301 (2008).

[6] L. Ding, X. Xin, Q. Yang and Y. Sang, Security analysis and improvements of XOR arbitrated quantum signature-based GHZ state, *Modern Physics Letters A* **37**, 2250008 (2022).

[7] F. Gao, S.-J. Qin, F.-Z. Guo and Q.-Y. Wen, Cryptanalysis of the arbitrated quantum signature protocols, *Physical Review A* **84**, 022344 (2011).

[8] D. Gottesman and I. Chuang, Quantum digital signatures, *arXiv:quant-ph/0105032* (2001).

[9] T. Kim, J.W. Choi, N.-S. Jho and S. Lee, Quantum messages with signatures forgeable in arbitrated quantum signature schemes, *Physica Scripta* **90**, 025101 (2015).

[10] S. Kremer, O. Markowitch and J. Zhou, An intensive survey of fair non-repudiation protocols, *Computer Communications* **25**, 1606–1621 (2002).

[11] F.-G. Li and J.-H. Shi, An arbitrated quantum signature protocol based on the chained CNOT operations encryption, *Quantum Information Processing* **14**, 2171–2181 (2015).

[12] Q. Li, W.H. Chan and D.-Y. Long, Arbitrated quantum signature scheme using Bell states, *Physical Review A* **79**, 054307 (2009).

[13] X.-H. Li, C.-Y. Li, F.-G. Deng, P. Zhou, Y.-J. Liang and H.-Y. Zhou, Quantum secure direct communication with quantum encryption based on pure entangled states, *Chinese Physics* **16**, 2149 (2007).

[14] F. Liu, S.-J. Qin and Q. Su, An arbitrated quantum signature scheme with fast signing and verifying, *Quantum Information Processing* **13**, 491–502 (2014).

[15] F. Liu, K. Zhang and T. Cao, Security weaknesses in arbitrated quantum signature protocols, *International Journal of Theoretical Physics* **53**, 277–288 (2014).

[16] D. Lu, Z. Li, J. Yu and Z. Han, A verifiable arbitrated quantum signature scheme based on controlled quantum teleportation, *Entropy* **24**, 111 (2022).

[17] Y.-P. Luo and T. Hwang, Comment on "An arbitrated quantum signature protocol based on the chained CNOT operations encryption", *arXiv:1512.00711* (2015).

[18] F.J. MacWilliams and N.J.A. Sloane, The theory of error-correcting codes (3^{rd} edition), *North-Holland Publishing Company* (1981).

[19] Z. Sun, R. Du, B. Wang, Q. Li and D. Long, Improvements on the security of arbitrated quantum signature protocols, *arXiv:1107.2459v3* (2011).

[20] C. Wang, J.-W. Liu and T. Shang, Enhanced arbitrated quantum signature scheme using Bell states, *Chinese Physics B* **23**, 060309 (2014).

[21] Y. Wang, K. Xu and Y. Guo, A chaos-based arbitrated quantum signature scheme in quantum crypotosystem, *International Journal of Theoretical Physics* **53**, 28–38 (2014).

[22] G. Zeng, Reply to "Comment on 'Arbitrated quantum-signature scheme'", *Physical Review A* **78**, 016301 (2008).

[23] G. Zeng and C.H. Keitel, Arbitrated quantum-signature scheme, *Physical review A* **65**, 042312 (2002).

[24] K.-J. Zhang, D. Li and Q. Su, Security of the arbitrated quantum signature protocols revisited, *Physica Scripta* **89**, 015102 (2014).

[25] K.-J. Zhang, S.-J. Qin, Y. Sun, T.-T. Song and Q. Su, Reexamination of arbitrated quantum signature: the impossible and the possible, *Quantum Information Processing* **12**, 3127–3141 (2013).

[26] K.-J. Zhang, W.-W. Zhang and D. Li, Improving the security of arbitrated quantum signature against the forgery attack, *Quantum Information Processing* **12**, 2655–2669 (2013).

[27] L. Zhang, H.-W. Sun, K.-J. Zhang and H.-Y. Jia, An improved arbitrated quantum signature protocol based on the key-controlled chained CNOT encryption, *Quantum Information Processing* **16**, 70 (2017).

[28] L. Zhang, H.-W. Sun, K.-J. Zhang, Q.-L. Wang and X.-Q. Cai, The security problems in some novel arbitrated quantum signature protocols, *International Journal of Theoretical Physics* **56**, 2433–2444 (2017).

[29] T. Zhang, X. Xin, L. Sun, C. Li and F. Li, Secure quantum signature scheme without entangled state, *Quantum Information Processing* **23**, 49 (2024).

[30] X.-Y. Zheng and C. Kuang, Arbitration quantum signature protocol based on XOR encryption, *International Journal of Quantum Information* **18**, 2050025 (2020).

[31] X. Zou and D. Qiu, Security analysis and improvements of arbitrated quantum signature schemes, *Physical Review A* **82**, 042325 (2010).

[32] X. Zou and D. Qiu, Arbitrated quantum signature schemes: Attacks and security, *FAW-AAIM 2013, Springer Lecture Notes in Computer Science* **7924**, 48–59 (2013).