Supplementary Notes for the paper 'Practical quantum tokens without quantum memories and experimental tests'

Adrian Kent,[1,2] David Lowndes,[3] Damián Pitalúa-García,[1,*] and John Rarity[3]

[1] *Centre for Quantum Information and Foundations, DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, U.K.*

[2] *Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada*

[3] *Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Bristol, U.K.*

I. SUMMARY

These supplementary notes provide the mathematical proofs of the lemmas and theorems given in the main text, as well as some known mathematical results, and new lemmas and a new theorem derived here to prove these results. Supplementary Note II states some well known mathematical results that are used along this text. The rest of this text is divided in three parts. The first part comprises Supplementary Notes III and IV. Supplementary Note III provides Lemmas 6 and 7, which are used in other supplementary notes to prove various results. In Supplementary Note IV, Lemma 1 is proved from Lemma 6. Thus, the first part, along with Supplementary Note II is all that the reader requires for the security proof in the case of two presentation points (the case $M = 1$) in the ideal case where there are not any errors or losses, i.e. Lemma 1. The second part comprises Supplementary Notes V, VI and VII, which give mathematical details for the case of two presentation points in the general case that there are errors, losses and other experimental imperfections. The third part comprises Supplementary Note VIII, which gives mathematical details for the case of 2^M presentation points, for any integer $M \geq 1$, in the general case that there are errors, losses and other experimental imperfections.

In the second part, Lemmas 2, 3 and 4, which indicate the robustness, correctness and privacy for the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ given in the main text, are proved in Supplementary Note V. Supplementary Note VI proves Lemma 5, showing that reporting strategies 1 and 2 with the photonic setup of Fig. 3 guarantee perfect protection against arbitrary multi-photon attacks [1], given assumptions E and F of Table 6. Using Lemma 6, Supplementary Note VII proves Theorem 1, which corresponds to unforgeability for the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ given in the main text.

In the third part, Theorem 2 given in the main text is proved in Supplementary Note VIII in the following way. First, the quantum token scheme $\mathcal{QT}_a$ is extended to a quantum token scheme $\mathcal{QT}_a^M$ for the case of 2^M presentation points, for any integer $M \geq 1$, and for $a \in \{1, 2\}$. Then, Lemmas 8, 9 and 10 and Theorem 3, which respectively state the robustness, correctness, privacy and unforgeability for the token schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$, are given and proved.

II. MATHEMATICAL PRELIMINARIES

The following results are well known in the literature. We state them here for completeness. In the following, $\|O\|$ denotes the Schatten $\infty-$norm of the linear operator O, which equals the greatest eigenvalue of O, if O is a positive semi definite operator acting on a finite dimensional Hilbert space.

Proposition 1. *Let $X_1, X_2, \ldots, X_N$ be independent random variables taking values $X_k \in \{0, 1\}$, for $k \in [N]$. Let $X = \sum_{k=1}^{N} X_k$, and let $E(X)$ be the average value of X. Two Chernoff bounds state that [2]*

$$\begin{aligned} Pr[X \leq (1-\epsilon)E(X)] &\leq e^{-\frac{E(X)}{2}\epsilon^2}, \\ Pr[X \geq (1+\epsilon)E(X)] &\leq e^{-\frac{E(X)}{3}\epsilon^2}, \end{aligned} \tag{1}$$

for $0 < \epsilon < 1$.

Proposition 2. *For any quantum density matrix ξ and any positive semi definite operator O acting on a finite dimensional Hilbert space $\mathcal{H}$, we have*

$$Tr(O\xi) \leq \|O\|. \tag{2}$$

Proof. Since O acts on a finite dimensional Hilbert space $\mathcal{H}$ and it is positive semi definite, it is also Hermitian, hence, from the spectral theorem there exists an orthonormal basis $\{|e_i\rangle\}_i$ of $\mathcal{H}$ which is an eigenbasis of O, with real eigenvalues $\{\mu_i\}_i$. Suppose that $\xi = |\xi\rangle\langle\xi|$ is pure, then we express it in the eigenbasis $\{|e_i\rangle\}_i$ of O. We have $|\xi\rangle = \sum_i \alpha_i |e_i\rangle$, where $\sum_i |\alpha_i|^2 = 1$, hence, $\mathrm{Tr}(O\xi) = \sum_i \mu_i |\alpha_i|^2 \leq \|O\|$. If ξ is not pure, it can be written as the convex combination of pure states, hence, applying the inequality to each of these pure states, the result follows. □

Proposition 3. *For any finite set of positive semi definite operators $\{D_a\}_{a\in\Omega}$ acting on a finite dimensional Hilbert space $\mathcal{H}_A$ and any projective measurement $\{\Pi_a\}_{a\in\Omega}$ acting on a finite dimensional Hilbert space $\mathcal{H}_B$, it holds that*

$$\left\| \sum_{a\in\Omega} (D_a)_A \otimes (\Pi_a)_B \right\| = \max_{a\in\Omega} \|D_a\|. \tag{3}$$

* D.Pitalua-Garcia@damtp.cam.ac.uk

Proof. Let $|\psi\rangle$ be the eigenstate of $O = \sum_{a\in\Omega}(D_a)_A \otimes (\Pi_a)_B$ with the greatest eigenvalue, hence, $\|O\| = \langle\psi|O|\psi\rangle$. We can write $|\psi\rangle = \sum_{a\in\Omega}\alpha_a|\omega_a\rangle$, where $\sum_{a\in\Omega}|\alpha_a|^2 = 1$, and where $\big(\mathbb{1}_A \otimes (\Pi_a)_B\big)|\omega_b\rangle = \delta_{a,b}|\omega_b\rangle$, for $a, b \in \Omega$. Thus, we obtain

$$\|O\| = \sum_{a\in\Omega}|\alpha_a|^2\langle\omega_a|\big((D_a)_A \otimes \mathbb{1}_B\big)|\omega_a\rangle. \tag{4}$$

We can then write $|\omega_a\rangle = \sum_{i,j}\beta^a_{i,j}|e^a_i\rangle\otimes|j\rangle$, where $\{|e^a_i\rangle\}_i$ is the eigenbasis of D_a, with eigenvalues $\{\mu^a_i\}_i$, $\{|j\rangle\}_j$ is an orthonormal basis of $\mathcal{H}_B$, and where $\sum_{i,j}|\beta^a_{i,j}|^2 = 1$, for $a \in \Omega$. From (4), we have

$$\begin{aligned}\|O\| &= \sum_{a\in\Omega}|\alpha_a|^2\sum_{i,j}|\beta^a_{i,j}|^2\mu^a_i \\ &\le \sum_{a\in\Omega}|\alpha_a|^2\sum_{i,j}|\beta^a_{i,j}|^2\|D_a\| \\ &= \sum_{a\in\Omega}|\alpha_a|^2\|D_a\| \\ &\le \max_{a\in\Omega}\|D_a\|, \end{aligned}\tag{5}$$

where in the second line we used that $\|D_a\|$ is the greatest of the eigenvalues $\{\mu^a_i\}_i$. Now let $|\tau_b\rangle \in \mathcal{H}_A$ be an eigenstate of D_b whose corresponding eigenvalue is $\|D_b\|$, i.e. the greatest eigenvalue of D_b, and where $\|D_b\| = \max_{a\in\Omega}\|D_a\|$, for some $b \in \Omega$. Let $|\chi_b\rangle \in \mathcal{H}_B$ be a pure state satisfying $\Pi_a|\chi_b\rangle = \delta_{a,b}|\chi_b\rangle$, for $a \in \Omega$. It can easily be verified that $|\tau_b\rangle \otimes |\chi_b\rangle$ is an eigenstate of O with eigenvalue $\|D_b\| = \max_{a\in\Omega}\|D_a\|$, i.e. there is an eigenvalue of O equal to $\max_{a\in\Omega}\|D_a\|$. Thus, since O is positive semi definite, $\|O\|$ is the greatest eigenvalue of O, hence, the result follows from (5). □

III. USEFUL MATHEMATICAL RESULTS

In this supplementary note,we state and prove Lemmas 6 and 7. Lemma 6 extends results of Ref. [3], for example in allowing a small deviation from the random distribution, as characterized by the parameters $\beta_{\rm PS} > 0$ and $\beta_{\rm PB} > 0$. Lemma 6 is a central mathematical result that we use to prove Lemma 1 and Theorem 1 in Supplementary Notes IV and VII, respectively. Lemma 7 states an upper bound on the maximum eigenvalue of a particular qubit density matrix. It will be useful in the proof of Theorem 1 in Supplementary Note VII.

Lemma 6. *For $r, r', s \in \{0,1\}$ and $k \in [N]$, and for some $O \in \big[\frac{1}{\sqrt{2}}, 1\big)$, let $|\phi^k_{r,s}\rangle$ be qubit pure states satisfying $\langle\phi^k_{0,s}|\phi^k_{1,s}\rangle = 0$ and $\big|\langle\phi^k_{r,0}|\phi^k_{r',1}\rangle\big| \le O$. Let $\mathbf{h} = (h_1, \ldots, h_N)$ be a $N-$bit string. For any $N-$bit string $\mathbf{x} = (x_1, \ldots, x_N)$ and for $i \in \{0,1\}$, let $\mathbf{x}_i$ denote the restriction of $\mathbf{x}$ to $S^{\mathbf{h}}_i = \big\{k \in [N] | s_k = h_k \oplus i\big\}$. Let $P_{\mathbf{s}} = \prod_{k=1}^N P^k_{s_k}$ be the probability distribution for $\mathbf{s} = (s_1, \ldots, s_N) \in \{0,1\}^N$, where $\{P^k_0, P^k_1\}$ is a binary probability distribution satisfying $P^{LB}_0 \le P^k_0 \le P^{UB}_0$, for $k \in [N]$. Let $d(\cdot,\cdot)$ denote the Hamming distance, and let $\big(\phi_{\mathbf{r},\mathbf{s}}\big)_B = \bigotimes_{k=1}^N\big(|\phi^k_{r_k,s_k}\rangle\langle\phi^k_{r_k,s_k}|\big)_{B_k}$, where $B = B_1B_2\cdots B_N$ denotes a quantum system of N qubits. For $\mathbf{a}, \mathbf{b} \in \{0,1\}^N$, we define*

$$\big(D_{\mathbf{a},\mathbf{b}}\big)_B = \sum_{\mathbf{s}\in\{0,1\}^N} P_{\mathbf{s}} \sum_{\substack{\mathbf{r}\in\{0,1\}^N \\ d(\mathbf{a}_0,\mathbf{r}_0)+d(\mathbf{b}_1,\mathbf{r}_1)\le N\gamma_{err}}} \big(\phi_{\mathbf{r},\mathbf{s}}\big)_B, \tag{6}$$

for some $\gamma_{err} \ge 0$. Let λ be a lower bound on the minimum of the function $B(P_0, O)$ evaluated over the range $P_0 \in [P^{LB}_0, P^{UB}_0]$, with

$$B(P_0, O) = \frac{1 - \sqrt{1 - 4(1-O^2)P_0(1-P_0)}}{2}. \tag{7}$$

In particular, if $P^{LB}_0 = \frac{1}{2} - \beta_{PB}$ and $P^{UB}_0 = \frac{1}{2} + \beta_{PB}$ for some $\beta_{PB} \ge 0$ then

$$\lambda = \frac{1}{2}\Big(1 - \sqrt{1 - (1-O^2)(1-4\beta^2_{PB})}\Big). \tag{8}$$

If $\gamma_{err} = 0$ then it holds that

$$\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N}\|D_{\mathbf{a},\mathbf{b}}\| \le \big(1-\lambda\big)^N. \tag{9}$$

If $0 < \gamma_{err} < \lambda$ then it holds that

$$\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N}\|D_{\mathbf{a},\mathbf{b}}\| \le e^{-\frac{N\lambda}{2}\left(1-\frac{\gamma_{err}}{\lambda}\right)^2}. \tag{10}$$

Proof. For $\mathbf{a}, \mathbf{b}, \mathbf{s} \in \{0,1\}^N$, we define $\mathbf{u} \in \{0,1\}^N$ satisfying $\mathbf{u}_0 = \mathbf{a}_0$ and $\mathbf{u}_1 = \mathbf{b}_1$. Then, in (6), we change variables $\mathbf{r} = \mathbf{x} \oplus \mathbf{u}$ and we sum over $\mathbf{x} \in \{0,1\}^N$, where '$\oplus$' denotes bit-wise sum modulo 2. We obtain

$$D_{\mathbf{a},\mathbf{b}} = \sum_{\mathbf{s}\in\{0,1\}^N} P_{\mathbf{s}} \sum_{\substack{\mathbf{x}\in\{0,1\}^N \\ w(\mathbf{x})\le N\gamma_{\rm err}}} \big(\phi_{\mathbf{x}\oplus\mathbf{u},\mathbf{s}}\big)_B, \tag{11}$$

where $w(\mathbf{x})$ denotes the Hamming weight of $\mathbf{x}$.

We note that $D_{\mathbf{a},\mathbf{b}}$ is a positive semi definite operator, hence, $\|D_{\mathbf{a},\mathbf{b}}\|$ corresponds to the greatest eigenvalue of $D_{\mathbf{a},\mathbf{b}}$, for $\mathbf{a}, \mathbf{b} \in \{0,1\}^N$. For given $\mathbf{a}, \mathbf{b} \in \{0,1\}^N$, in order to compute $\big\|D_{\mathbf{a},\mathbf{b}}\big\|$, we first evaluate the sum over $\mathbf{s} \in \{0,1\}^N$ in (11). We obtain

$$\begin{aligned}\sum_{\mathbf{s}\in\{0,1\}^N} P_{\mathbf{s}}\phi_{\mathbf{x}\oplus\mathbf{u},\mathbf{s}} &= \bigotimes_{k=1}^N\Big(\sum_{s_k=0}^1 P^k_{s_k}|\phi^k_{x_k\oplus u_k,s_k}\rangle\langle\phi^k_{x_k\oplus u_k,s_k}|\Big) \\ &= \bigotimes_{k=1}^N \rho^k_{x_k\oplus a_k, x_k\oplus b_k}, \end{aligned}\tag{12}$$

where

$$\rho^k_{b,c} = \Big(P^k_{h_k}|\phi^k_{b,h_k}\rangle\langle\phi^k_{b,h_k}| + P^k_{h_k\oplus1}|\phi^k_{c,h_k\oplus1}\rangle\langle\phi^k_{c,h_k\oplus1}|\Big), \tag{13}$$

for $b, c \in \{0,1\}$ and $k \in [N]$. We note that $\rho^k_{b,c} + \rho^k_{b\oplus1,c\oplus1} = \mathbb{1}$ since $\{|\phi^k_{r,s}\rangle\}_{r\in\{0,1\}}$ is a qubit orthonormal

basis for $s \in \{0,1\}$ and since $\{P_0^k, P_1^k\}$ is a probability distribution, for $k \in [N]$. Thus, $\rho_{b,c}^k$ and $\rho_{b\oplus 1,c\oplus 1}^k$ are diagonal in the same basis, for $b,c \in \{0,1\}$ and $k \in [N]$. Therefore, without loss of generality, we can write

$$\rho_{b,c}^k = \sum_{t=0}^{1} \lambda_{t\oplus b,b\oplus c}^k \big|\mu_{t,b\oplus c}^k\big\rangle\big\langle\mu_{t,b\oplus c}^k\big|, \tag{14}$$

where $\big\{\big|\mu_{t,b\oplus c}^k\big\rangle\big\}_{t=0}^1$ is the eigenbasis of $\rho_{b,c}^k$ with real non-negative eigenvalues $\big\{\lambda_{t\oplus b,b\oplus c}^k\big\}_{t=0}^1$, and where

$$\lambda_{0,c}^k + \lambda_{1,c}^k = 1, \tag{15}$$

for $b,c \in \{0,1\}$ and $j \in [N]$. Thus, we have

$$\left(\bigotimes_{k=1}^N \rho_{x_k\oplus a_k,x_k\oplus b_k}^k\right)\left(\bigotimes_{k=1}^N \big|\mu_{t_k,a_k\oplus b_k}^k\big\rangle\right) = \left(\prod_{k=1}^N \lambda_{t_k\oplus x_k\oplus a_k,a_k\oplus b_k}^k\right)\left(\bigotimes_{k=1}^N \big|\mu_{t_k,a_k\oplus b_k}^k\big\rangle\right), \tag{16}$$

for $\mathbf{t} \in \{0,1\}^N$. Importantly, we see from (16) that the eigenbasis of $\bigotimes_{k=1}^N \rho_{x_k\oplus a_k,x_k\oplus b_k}^k$ is the same for all $\mathbf{x} \in \{0,1\}^N$. Thus, from (11), (12) and (16), we see that the eigenbasis of $D_{\mathbf{a},\mathbf{b}}$ is $\left\{\bigotimes_{k=1}^N \big|\mu_{t_k,a_k\oplus b_k}^k\big\rangle\right\}_{\mathbf{t}\in\{0,1\}^N}$, with eigenvalues

$$\sum_{\substack{\mathbf{x}\in\{0,1\}^N \\ w(\mathbf{x})\le N\gamma_{\rm err}}} \left(\prod_{k=1}^N \lambda_{t_k\oplus x_k\oplus a_k,a_k\oplus b_k}^k\right),$$

for $\mathbf{t} \in \{0,1\}^N$. It follows that

$$\begin{aligned}
&\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N} \|D_{\mathbf{a},\mathbf{b}}\| \\
&\quad= \max_{\mathbf{a},\mathbf{b},\mathbf{t}\in\{0,1\}^N} \sum_{\substack{\mathbf{x}\in\{0,1\}^N \\ w(\mathbf{x})\le N\gamma_{\rm err}}} \left(\prod_{k=1}^N \lambda_{t_k\oplus x_k\oplus a_k,a_k\oplus b_k}^k\right) \\
&\quad= \max_{\alpha,\beta\in\{0,1\}^N} \sum_{\substack{\mathbf{x}\in\{0,1\}^N \\ w(\mathbf{x})\le N\gamma_{\rm err}}} \left(\prod_{k=1}^N \lambda_{x_k\oplus\beta_k,\alpha_k}^k\right),
\end{aligned} \tag{17}$$

by taking the change of variables $\alpha_k = a_k \oplus b_k$ and $\beta_k = a_k \oplus t_k$, for $k \in [N]$.

Below we compute the maximum given by the second line of (17). We consider two cases separately, the case $\gamma_{\rm err} = 0$, and the case $0 < \gamma_{\rm err} < \lambda$. Within the second case we consider the subcases $0 < \gamma_{\rm err} < \frac{1}{N}$ and $\gamma_{\rm err} \ge \frac{1}{N}$. We use the following definitions:

$$\begin{aligned}
\lambda_0^k &= \max\{\lambda_{\beta,\alpha}^k\}_{\alpha,\beta\in\{0,1\}} \\
\lambda_1^k &= 1 - \lambda_0^k,
\end{aligned} \tag{18}$$

where in the second line we used (15), for $k \in [N]$. We also define parameters $\lambda_0 \le 1$ and λ_1 that satisfy

$$\begin{aligned}
\lambda_0^k &\le \lambda_0, \\
\lambda_1 &= 1 - \lambda_0,
\end{aligned} \tag{19}$$

for $k \in [N]$.

In the case $\gamma_{\rm err} = 0$, we have from (17) that

$$\begin{aligned}
\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N} \|D_{\mathbf{a},\mathbf{b}}\| &= \max_{\alpha,\beta\in\{0,1\}^N} \prod_{k=1}^N \lambda_{\beta_k,\alpha_k}^k \\
&= \prod_{k=1}^N \lambda_0^k \\
&\le (1-\lambda_1)^N \\
&= (1-\lambda)^N,
\end{aligned} \tag{20}$$

where in the second line we used (18), in the third line we used (19), and in the last line we used that $\lambda_1 = \lambda$, which is shown below. The bound (9) follows from (20).

In the case $0 < \gamma_{\rm err} < \frac{1}{N}$, we note that since $\lambda \in (0,1)$, we have $\ln(1-\lambda) \le -\lambda < -\frac{\lambda}{2}$. It follows that

$$(1-\lambda)^N < e^{-\frac{N\lambda}{2}\left(1-\frac{\gamma_{\rm err}}{\lambda}\right)^2}. \tag{21}$$

Thus, from (20) and (21), it follows that in the case that the conditions $0 < \gamma_{\rm err} < \lambda$ and $0 < \gamma_{\rm err} < \frac{1}{N}$ hold, the bound (10) is satisfied.

We show below that the bound (10) is satisfied too in the case that $\frac{1}{N} \le \gamma_{\rm err} < \lambda$ holds. It follows that (10) holds if $0 < \gamma_{\rm err} < \lambda$, as stated in the lemma.

We consider $\frac{1}{N} \le \gamma_{\rm err} < \lambda$. For any $\alpha,\beta \in \{0,1\}^N$ and for any $l \in [N]$, we define $\tilde{\mathbf{x}}_l = (x_1, x_2, \ldots, x_{l-1}, x_{l+1}, x_{l+2}, \ldots, x_N)$, and we can write

$$\begin{aligned}
&\sum_{\substack{\mathbf{x}\in\{0,1\}^N \\ w(\mathbf{x})\le N\gamma_{\rm err}}} \prod_{k=1}^N \lambda_{x_k\oplus\beta_k,\alpha_k}^k = \\
&\quad \lambda_{\beta_l\oplus 1,\alpha_l}^l \left(\sum_{\substack{\tilde{\mathbf{x}}_l\in\{0,1\}^{N-1} \\ w(\tilde{\mathbf{x}}_l)\le N\gamma_{\rm err}-1}} \prod_{\substack{k=1 \\ k\ne l}}^N \lambda_{x_k\oplus\beta_k,\alpha_k}^k\right) \\
&\quad + \lambda_{\beta_l,\alpha_l}^l \left(\sum_{\substack{\tilde{\mathbf{x}}_l\in\{0,1\}^{N-1} \\ w(\tilde{\mathbf{x}}_l)\le N\gamma_{\rm err}}} \prod_{\substack{k=1 \\ k\ne l}}^N \lambda_{x_k\oplus\beta_k,\alpha_k}^k\right).
\end{aligned} \tag{22}$$

We see that the term inside the second bracket cannot be smaller than the term inside the first one. Since this holds for any choice of $l \in [N]$, in order to maximize the quantity on the left-hand side, we need to maximize $\lambda_{\beta_l,\alpha_l}^l$ for $l \in [N]$. Thus, we obtain from (17), (18) and (22) that

$$\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N} \|D_{\mathbf{a},\mathbf{b}}\| = \sum_{\substack{\mathbf{x}\in\{0,1\}^N \\ w(\mathbf{x})\le N\gamma_{\rm err}}} \left(\prod_{k=1}^N \lambda_{x_k}^k\right). \tag{23}$$

Similarly, reasoning as in the previous lines, it is straightforward to obtain from (19) and (23) that

$$\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N}\|D_{\mathbf{a},\mathbf{b}}\| \leq \sum_{\substack{\mathbf{x}\in\{0,1\}^N \\ w(\mathbf{x})\leq N\gamma_{\rm err}}} \left(\prod_{k=1}^{N}\lambda_{x_k}\right) = \sum_{n=0}^{\lfloor N\gamma_{\rm err}\rfloor} \tbinom{N}{n}(\lambda_0)^{N-n}(\lambda_1)^n. \quad (24)$$

We upper bound the right-hand side of (24) using the Chernoff bound given by Proposition 1. Let X_k be a random variable taking value $X_k = i$ with probability λ_i, for $i \in \{0,1\}$ and $k \in [N]$. Let $X = \sum_{k=1}^{N} X_k$, whose average value is $E(X) = N\lambda_1$. We have

$$\sum_{n=0}^{\lfloor N\gamma_{\rm err}\rfloor} \tbinom{N}{n}(\lambda_0)^{N-n}(\lambda_1)^n \leq \Pr[X \leq N\gamma_{\rm err}] \leq e^{-\frac{N\lambda_1}{2}\left(1-\frac{\gamma_{\rm err}}{\lambda_1}\right)^2}, \quad (25)$$

for $0 < \gamma_{\rm err} < \lambda_1$, where in the second line we used the Chernoff bound (1), by taking $\epsilon = 1 - \frac{\gamma_{\rm err}}{\lambda_1}$. By taking $\lambda_1 = \lambda$, it follows from (24) and (25) that

$$\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N}\|D_{\mathbf{a},\mathbf{b}}\| \leq e^{-\frac{N\lambda}{2}\left(1-\frac{\gamma_{\rm err}}{\lambda}\right)^2}, \quad (26)$$

for $0 < \gamma_{\rm err} < \lambda$, as claimed.

We complete the proof below by showing that $\lambda_1 = \lambda$ satisfies the condition (19). We write $\langle\phi^k_{b,h_k}|\phi^k_{c,h_k\oplus 1}\rangle = \omega^k_{b,c}e^{i\chi^k_{b,c}}$, with $\omega^k_{b,c} = \left|\langle\phi^k_{b,h_k}|\phi^k_{c,h_k\oplus 1}\rangle\right|$, for $b,c \in \{0,1\}$ and $k \in [N]$. We define

$$R^k_{\pm,b,c} = \frac{P^k_{h_k\oplus 1} - P^k_{h_k} \pm \sqrt{(P^k_1 - P^k_0)^2 + 4(\omega^k_{b,c})^2 P^k_0 P^k_1}}{2\omega^k_{b,c}P^k_{h_k}}, \quad (27)$$

for $b,c \in \{0,1\}$ and $k \in [N]$. It is straightforward to verify that the density matrix $\rho^k_{b,c}$ given by (13) has eigenstates

$$\left|e^k_{\pm,b,c}\right\rangle = \frac{1}{\sqrt{1+\left(R^k_{\pm,b,c}\right)^2}}\left(\left|\phi^k_{b,h_k}\right\rangle + R_{\pm,b,c}e^{-i\chi^k_{b,c}}\left|\phi^k_{c,h_k\oplus 1}\right\rangle\right) \quad (28)$$

with eigenvalues

$$\lambda^k_{\pm,b,c} = P^k_{h_k}\left(1 + \omega^k_{b,c}R^k_{\pm,b,c}\right), \quad (29)$$

for $b,c \in \{0,1\}$ and $k \in [N]$. Thus, from the definition (18) and from (29), we obtain

$$2\lambda^k_0 = P^k_0 + P^k_1 + \sqrt{\left(P^k_1 - P^k_0\right)^2 + 4P^k_0P^k_1 \max_{b,c\in\{0,1\}}\{(\omega^k_{b,c})^2\}}, \quad (30)$$

for $k \in [N]$. Since by assumption of the lemma, $\left|\langle\phi^k_{b,0}|\phi^k_{c,1}\rangle\right| \leq O$ for some $O \in \left[\frac{1}{\sqrt{2}},1\right)$, using $P^k_1 = 1-P^k_0$ and $\omega^k_{b,c} = \left|\langle\phi^k_{b,h_k}|\phi^k_{c,h_k\oplus 1}\rangle\right|$, we have from (30) that

$$\lambda^k_0 \leq A(P^k_0, O), \quad (31)$$

for $k \in [N]$, where

$$A(P_0,O) = \frac{1+\sqrt{1-4[1-O^2]P_0(1-P_0)}}{2}. \quad (32)$$

Thus, since $P^k_0 \in [P^{\rm LB}_0, P^{\rm UB}_0]$ for $k \in [N]$, by defining λ_0 as an upper bound on the maximum of the function $A(P_0,O)$ evaluated over the range $P_0 \in [P^{\rm LB}_0, P^{\rm UB}_0]$, with $\lambda_0 < 1-\gamma_{\rm err}$, and by defining $\lambda_1 = 1-\lambda_0$, the conditions given by (19) hold. We see from (32) that the function $B(P_0,O)$ given by (7) satisfies $B(P_0,O) = 1 - A(P_0,O)$. Thus, we define λ_1 as a lower bound on the minimum of the function $B(P_0,O)$ evaluated over the range $P_0 \in [P^{\rm LB}_0, P^{\rm UB}_0]$, and we define $\lambda = \lambda_1$. In the case that $P^{\rm LB}_0 = \frac{1}{2} - \beta_{\rm PB}$ and $P^{\rm UB}_0 = \frac{1}{2} + \beta_{\rm PB}$ for $\beta_{\rm PB} \geq 0$, we define λ_1 as the minimum of the function $B(P_0,O)$ evaluated over the range $P_0 \in [P^{\rm LB}_0, P^{\rm UB}_0]$, and we define $\lambda = \lambda_1$. It is straightforward to see from (32) that in this case $\lambda_1 = \frac{1}{2}\left(1 - \sqrt{1-[1-O^2](1-4\beta^2_{\rm PB})}\right)$. The result follows by noting that, as stated in the lemma, $\lambda = \frac{1}{2}\left(1 - \sqrt{1-[1-O^2](1-4\beta^2_{\rm PB})}\right)$. □

Lemma 7. *Let ρ be a qubit density matrix given by*

$$\rho = \sum_{u=0}^{1}\sum_{t=0}^{1} P_{PB}(u)P_{PS}(t)|\phi_{tu}\rangle\langle\phi_{tu}|, \quad (33)$$

where $\{|\phi_{tu}\rangle\}_{t,u\in\{0,1\}}$ is a set of qubit states satisfying $\langle\phi_{0u}|\phi_{1u}\rangle = 0$ for $u \in \{0,1\}$, where the qubit orthonormal basis $\mathcal{D}_u = \{|\phi_{tu}\rangle\}^1_{t=0}$ is the computational (Hadamard) basis within an uncertainty angle $\theta \in \left(0, \frac{\pi}{4}\right)$ on the Bloch sphere if $u = 0$ ($u = 1$); and where the binary probability distributions $\{P_{PB}(u)\}^1_{u=0}$ and $\{P_{PS}(t)\}^1_{t=0}$ satisfy $\frac{1}{2} - \beta_{PB} \leq P_{PB}(u) \leq \frac{1}{2} + \beta_{PB}$ and $\frac{1}{2} - \beta_{PS} \leq P_{PS}(u) \leq \frac{1}{2} + \beta_{PS}$ for $u \in \{0,1\}$, and for given parameters $\beta_{PB}, \beta_{PS} \in \left(0, \frac{1}{2}\right)$. Let μ_+ be the greatest eignevalues of ρ. It holds that

$$\mu_+ \leq \frac{1}{2}\left(1 + h(\beta_{PS}, \beta_{PB}, \theta)\right), \quad (34)$$

where

$$h(\beta_{PS}, \beta_{PB}, \theta) = 2\beta_{PS}\sqrt{\frac{1}{2} + 2\beta^2_{PB} + \left(\frac{1}{2} - 2\beta^2_{PB}\right)\sin(2\theta)}. \quad (35)$$

Proof. To simplify notation, we define $P = P_{\rm PB}(0)$, $1 - P = P_{\rm PB}(1)$, $R = P_{\rm PS}(0)$, and $1 - R = P_{\rm PS}(1)$. Since applying a unitary operation U on ρ does not change its eigenvalues, we define $\rho' = U\rho U^\dagger$ and we compute an upperbound on the greatest eigenvalue of ρ'. Since $\{|\phi_{tu}\rangle\}^1_{t=0}$ is a qubit orthonormal basis, for $u \in \{0,1\}$, we can choose U such that $U|\phi_{t0}\rangle = |t\rangle$ and $U|\phi_{t1}\rangle =$

$|\tilde{t}\rangle$, where $\{|0\rangle, |1\rangle\}$ is the computational basis, which has Bloch vectors in the z axis, and where $\{|\tilde{0}\rangle, |\tilde{1}\rangle\}$ is another orthonormal basis with Bloch vectors on the $z-x$ plane. Thus, we see that from the statement of the lemma, we can choose U such that $|\tilde{0}\rangle$ has a Bloch vector with angle ξ above the x axis, towards the z axis; hence, $|\tilde{1}\rangle$ has a Bloch vector with angle ξ below the $-x$ axis, towards the $-z$ axis; where $\xi \in [-2\theta, 2\theta]$ for some $\theta \in \big(0, \frac{\pi}{4}\big)$.

Thus, using this notation, from (33), we obtain

$$\rho' = P\rho_0 + (1-P)\rho_1, \tag{36}$$

where

$$\begin{aligned} \rho_0 &= R|0\rangle\langle 0| + (1-R)|1\rangle\langle 1|, \\ \rho_1 &= R|\tilde{0}\rangle\langle\tilde{0}| + (1-R)|\tilde{1}\rangle\langle\tilde{1}|. \end{aligned} \tag{37}$$

The Bloch vector of ρ_0 is $(2R-1)\hat{z}$ and the Bloch vector of ρ_1 is $(2R-1)\big(\cos\xi\hat{x} + \sin\xi\hat{z}\big)$, where $\hat{x}$ and $\hat{z}$ are unit vectors pointing along the x and z axes, respectively. Thus, the Bloch vector of ρ' is

$$\vec{r} = (2R-1)\big[\big(P + (1-P)\sin\xi\big)\hat{z} + (1-P)\cos\xi\hat{x}\big]. \tag{38}$$

The eigenvalues of ρ', hence also of ρ, are given by $\mu_\pm = \frac{1}{2}\big(1 \pm |\vec{r}|\big)$. Thus, from (38), the greatest eigenvalue is given by

$$\mu_+ = \frac{1}{2}\big(1 + |\vec{r}|\big), \tag{39}$$

where

$$|\vec{r}| = |2R-1|\sqrt{\big(P + (1-P)\sin\xi\big)^2 + (1-P)^2\cos^2\xi}. \tag{40}$$

Since from the statement of the lemma we have $\frac{1}{2} - \beta_{\rm PS} \leq R \leq \frac{1}{2} + \beta_{\rm PS}$, we see from (39) and (40) that for fixed values of P and ξ, μ_+ achieves its maximum if $R = \frac{1}{2} \pm \beta_{\rm PS}$. Thus, it holds that

$$\mu_+ \leq \frac{1}{2}\big(1 + 2\beta_{\rm PS}\sqrt{g(P,\xi)}\big), \tag{41}$$

where

$$g(P,\xi) = \big(P + (1-P)\sin\xi\big)^2 + (1-P)^2\cos^2\xi. \tag{42}$$

We write $P = \frac{1}{2} + d$. From the statement of the lemma, we have $d \in [-\beta_{\rm PB}, \beta_{\rm PB}]$ for some $\beta_{\rm PB} \in \big(0, \frac{1}{2}\big)$. It follows from (42) that

$$g(P,\xi) = \frac{1}{2}(1+\sin\xi) + 2(1-\sin\xi)d^2. \tag{43}$$

Since $\xi \in [-2\theta, 2\theta]$ with $0 < \theta < \frac{\pi}{4}$, we have $(1-\sin\xi) > 0$. Thus, $g(P,\xi)$ is maximum when d^2 is maximum. Since $d \in [-\beta_{\rm PB}, \beta_{\rm PB}]$, it follows that

$$\begin{aligned} g(P,\xi) &\leq \frac{1}{2}(1+\sin\xi) + 2(1-\sin\xi)\beta_{\rm PB}^2 \\ &= \frac{1}{2} + 2\beta_{\rm PB}^2 + \Big(\frac{1}{2} - 2\beta_{\rm PB}^2\Big)\sin\xi. \end{aligned} \tag{44}$$

Since $\beta_{\rm PB} \in \big(0, \frac{1}{2}\big)$, we have $\frac{1}{2} - 2\beta_{\rm PB}^2 > 0$. Thus, since $\xi \in [-2\theta, 2\theta]$ with $0 < \theta < \frac{\pi}{4}$, the second line of (44) is maximum when $\xi = 2\theta$. It follows that

$$g(P,\xi) \leq \frac{1}{2} + 2\beta_{\rm PB}^2 + \Big(\frac{1}{2} - 2\beta_{\rm PB}^2\Big)\sin(2\theta). \tag{45}$$

Thus, from (41) and (45), we obtain (34):

$$\mu_+ \leq \frac{1}{2}\big(1 + h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)\big), \tag{46}$$

where $h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)$ is given by (35), as claimed. □

IV. PROOF OF LEMMA 1

Lemma 1. *The quantum token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$ are $\epsilon_{unf}-$unforgeable with*

$$\epsilon_{unf} = \Big(\frac{1}{2} + \frac{1}{2\sqrt{2}}\Big)^N. \tag{47}$$

Proof. In summary, the proof comprises reducing a general cheating strategy by Alice in the token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$ to the task of producing the $N-$bit strings $\mathbf{a}$ and $\mathbf{b}$ given in Lemma 6, with $\gamma_{\rm err} = 0$, $\beta_{\rm PB} = 0$, and $\theta = 0$, i.e. $O(\theta) \equiv O = \frac{1}{\sqrt{2}}$. As we show, then Alice's success probability is upper bounded by the quantity $\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N}\|D_{\mathbf{a},\mathbf{b}}\|$, which for these parameters is upper bounded by $\big(\frac{1}{2} + \frac{1}{2\sqrt{2}}\big)^N$.

Consider the token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$. In these token schemes Alice gives Bob a $N-$bit strings $\mathbf{d} = (d_1, \ldots, d_N)$ and a bit c. Using this information, honest Bob computes the $N-$bit string $\tilde{\mathbf{d}}_i = (\tilde{d}_{i,1}, \ldots, \tilde{d}_{i,N})$ in the causal past of the presentation point Q_i, where $\tilde{d}_{i,k} = d_k \oplus c \oplus i$, for $k \in [N]$ and $i \in \{0,1\}$. In a general cheating strategy $\mathcal{S}$, Alice applies a joint projective measurement on the quantum system A of $N-$qubits in the state $|\phi_{\mathbf{tu}}\rangle_A = \bigotimes_{k=1}^N |\phi_{t_k u_k}\rangle_{A_k}$ received from Bob and an ancilla E of arbitrary finite Hilbert space dimension in a quantum state $|\chi\rangle_E$, and obtains the classical message $x = (\mathbf{d}, c)$ of $N+1$ bits that she gives Bob within the causal pasts of Q_0 and Q_1 and two $N-$bit (token) strings $\mathbf{a} = (a_1 \ldots, a_N)$ and $\mathbf{b} = (b_1, \ldots, b_N)$ that she gives to Bob at Q_0 and Q_1, respectively. Alice succeeds in making Bob validate these token strings at Q_0 and Q_1 if $\mathbf{a}_0 = \mathbf{t}_0$ and $\mathbf{b}_1 = \mathbf{t}_1$, where $\mathbf{x}_i$ is a restriction of the string $\mathbf{x} \in \{\mathbf{a}, \mathbf{b}, \mathbf{t}\}$ to the bit entries $x_k \in \Delta_i$, where $\Delta_i = \{k \in [N] | \tilde{d}_{i,k} = u_k\}$, for $i \in \{0,1\}$. Since $\tilde{d}_{i,k} = d_k \oplus c \oplus i$, for $k \in [N]$, we have that $\Delta_i = \{k \in [N] | u_k = d_k \oplus c \oplus i\}$, for $i \in \{0,1\}$.

Now consider the task of Lemma 6 with the following parameters. The states $|\phi^k_{r,s}\rangle$ are the BB84 states: $|\phi^k_{0,0}\rangle \equiv |\phi_{00}\rangle = |0\rangle$, $|\phi^k_{1,0}\rangle \equiv |\phi_{10}\rangle = |1\rangle$, $|\phi^k_{0,1}\rangle \equiv |\phi_{01}\rangle = |+\rangle$, $|\phi^k_{1,1}\rangle \equiv |\phi_{11}\rangle = |-\rangle$, for $k \in [N]$. It follows that $O = \frac{1}{\sqrt{2}}$. We also consider that $P_{\mathbf{s}} = \big(\frac{1}{2}\big)^N$ for

$\mathbf{s} \in \{0,1\}^N$, i.e. $\beta_{\rm PB} = 0$. It follows from Lemma 6 that $\lambda = \frac{1}{2} - \frac{1}{2\sqrt{2}}$ and that

$$\max_{\mathbf{a},\mathbf{b}\in\{0,1\}^N} \|D_{\mathbf{a},\mathbf{b}}\| \leq \Big(\frac{1}{2} + \frac{1}{2\sqrt{2}}\Big)^N. \tag{48}$$

We define the $N-$bit strings $\mathbf{r} = (r_1, \ldots, r_N)$, $\mathbf{s} = (s_1, \ldots, s_N)$ and $\mathbf{h} = (h_1, \ldots, h_N)$ in terms of the strings $\mathbf{t}$, $\mathbf{u}$ and $\mathbf{d}$ and of the bit c of the token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$ as follows: $r_k = t_k$, $s_k = u_k$ and $h_k = d_k \oplus c$, for $k \in [N]$. Thus, the set $S_i^{\mathbf{h}}$ in Lemma 6 is the set Δ_i in the token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$: $S_i^{\mathbf{h}} = \Delta_i$, for $i \in \{0,1\}$. It follows that the operator $D_{\mathbf{a},\mathbf{b}}$ in Lemma 6 can be associated to Alice's cheating strategy in the token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$. We deduce this connection below.

We consider an entanglement-based version of the token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$. Bob prepares a pair of qubits B_kA_k in the Bell state $|\Phi^+\rangle_{B_kA_k} = \frac{1}{\sqrt{2}}\big(|0\rangle|0\rangle + |1\rangle|1\rangle\big)_{B_kA_k}$, sends the qubit A_k to Alice, chooses $u_k \in \{0,1\}$ with probability $\frac{1}{2}$ and then measures the qubit B_k in the basis $\mathcal{D}_{u_k} = \{|\phi_{tu_k}\rangle\}_{t=0}^1$, obtaining the outcome $|\phi_{t_ku_k}\rangle$ randomly, with Alice's qubit A_k projecting into the same state, for $t_k \in \{0,1\}$. In a general cheating strategy $\mathcal{S}$, Alice introduces an ancillary quantum system E of arbitrary finite Hilbert space dimension in a pure state $|\chi\rangle_E$ and then applies a projective measurement on AE, with projector operators $\Pi_{x\mathbf{ab}}$, where the measurement outcomes $x = (\mathbf{d}, c) \in \{0,1\}^{N+1}$ and $\mathbf{a}, \mathbf{b} \in \{0,1\}^N$ correspond to the classical messages that Alice gives Bob, and where $A = A_1 \cdots A_N$. The probability that Alice obtains outcomes $x, \mathbf{a}$ and $\mathbf{b}$ following her strategy $\mathcal{S}$, for given values of $\mathbf{u}$ and $\mathbf{t}$, is given by

$$P_{\mathcal{S}}[x\mathbf{ab}|\mathbf{tu}] = \mathrm{Tr}\Big[\Phi_{\mathbf{tu}}\Pi_{x\mathbf{ab}}\Big], \tag{49}$$

where $\Phi_{\mathbf{tu}} = \big(|\phi_{\mathbf{tu}}\rangle\langle\phi_{\mathbf{tu}}|\big)_A \otimes \big(|\chi\rangle\langle\chi|\big)_E$. We define the sets

$$\Gamma^x_{\mathbf{abu}} = \big\{(\mathbf{a},\mathbf{b}) \in \{0,1\}^N \times \{0,1\}^N | \mathbf{a}_0 = \mathbf{t}_0, \mathbf{b}_1 = \mathbf{t}_1\big\}, \tag{50}$$

$$\xi^x_{\mathbf{abu}} = \big\{\mathbf{t} \in \{0,1\}^N | \mathbf{a}_0 = \mathbf{t}_0, \mathbf{b}_1 = \mathbf{t}_1\big\}. \tag{51}$$

It follows that Alice's success probability $P_{\mathcal{S}}$ satisfies

$$\begin{aligned} P_{\mathcal{S}} &= \Big(\frac{1}{4}\Big)^N \sum_{x,\mathbf{u},\mathbf{t}} \sum_{(\mathbf{a},\mathbf{b})\in\Gamma^x_{\mathbf{tu}}} P_{\mathcal{S}}[x\mathbf{ab}|\mathbf{tu}] \\ &= \Big(\frac{1}{4}\Big)^N \sum_{\mathbf{a},\mathbf{b},x,\mathbf{u}} \sum_{\mathbf{t}\in\xi^x_{\mathbf{abu}}} P_{\mathcal{S}}[x\mathbf{ab}|\mathbf{tu}], \end{aligned} \tag{52}$$

where in the first line we used (49) and (50); and where in the second line we used (50) and (51), and the fact that the string $\mathbf{z}_i$ has bit entries with labels from the set Δ_i satisfying $\Delta_0 \cap \Delta_1 = \emptyset$ and $\Delta_0 \cup \Delta_1 = [N]$, for $i \in \{0,1\}$ and $\mathbf{z} \in \{\mathbf{a}, \mathbf{b}, \mathbf{t}\}$.

We define the quantum state

$$\rho = \big(\Phi^+\big)_{BA} \otimes \big(|\chi\rangle\langle\chi|\big)_E, \tag{53}$$

where B denotes the system held by Bob and where $\big(\Phi^+\big)_{BA} = \bigotimes_{k\in[N]}\big(|\Phi^+\rangle\langle\Phi^+|\big)_{B_kA_k}$. We define the positive semi definite (and Hermitian) operators

$$D_{x\mathbf{ab}} = \Big(\frac{1}{2}\Big)^N \sum_{\mathbf{u}} \sum_{\mathbf{t}\in\xi^x_{\mathbf{abu}}} \big(\phi_{\mathbf{tu}}\big)_B, \tag{54}$$

$$\tilde{P} = \sum_{x,\mathbf{a},\mathbf{b}} \big(D_{x\mathbf{ab}}\big)_B \otimes \big(\Pi_{x\mathbf{ab}}\big)_{AE}, \tag{55}$$

where $\big(\phi_{\mathbf{tu}}\big)_B = \bigotimes_{k\in[N]}\big(|\phi_{t_ku_k}\rangle\langle\phi_{t_ku_k}|\big)_{B_k}$ and where $\mathbf{u}$ runs over $\{0,1\}^N$, x runs over $\{0,1\}^{N+1}$, and $\mathbf{a}$ and $\mathbf{b}$ run over $\{0,1\}^N$. It follows straightforwardly from (49) – (55) that

$$\begin{aligned} P_{\mathcal{S}} &= \mathrm{Tr}\big(\tilde{P}\rho\big) \\ &\leq \|\tilde{P}\| \\ &= \max_{x,\mathbf{a},\mathbf{b}} \big\|D_{x\mathbf{ab}}\big\|, \end{aligned} \tag{56}$$

where in the second line we used Proposition 2; and where in the third line we used (55) and Proposition 3, since $\{\Pi_{x\mathbf{ab}}\}_{x,\mathbf{a},\mathbf{b}}$ is a projective measurement acting on a finite dimensional Hilbert space and $\{D_{x\mathbf{ab}}\}_{x,\mathbf{a},\mathbf{b}}$ is a finite set of positive semi definite operators acting on a finite dimensional Hilbert space. We note that the operator $D_{x\mathbf{ab}}$ defined by (54) equals the operator $D_{\mathbf{ab}}$ given in Lemma 6, for the parameters $\gamma_{\rm err} = 0$, $O = \frac{1}{\sqrt{2}}$ and $\beta_{\rm PB} = 0$ that we are considering here. Thus, from (48) and (56), and because this bound does not depend on x, we obtain

$$P_{\mathcal{S}} \leq \Big(\frac{1}{2} + \frac{1}{2\sqrt{2}}\Big)^N. \tag{57}$$

Thus, the quantum token schemes $\mathcal{IQOT}_1$ and $\mathcal{IQOT}_2$ are $\epsilon_{\rm unf}-$unforgeable with $\epsilon_{\rm unf} = \big(\frac{1}{2}+\frac{1}{2\sqrt{2}}\big)^N$, as claimed. □

V. PROOFS OF LEMMAS 2, 3 AND 4

We recall that Lemmas 2, 3 and 4 consider parameters $\gamma_{\rm det}, \gamma_{\rm err} \in (0,1)$, allow for the experimental imperfections of Table 5 and make the assumptions of Table 6.

A. Proof of Lemma 2

Lemma 2. *If*

$$0 < \gamma_{det} < P_{det}, \tag{58}$$

then $\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{rob}-$robust with

$$\epsilon_{rob} = e^{-\frac{P_{det}N}{2}\left(1-\frac{\gamma_{det}}{P_{det}}\right)^2}. \tag{59}$$

We note that the condition (58) is necessary to guarantee robustness, as in the limit $N \to \infty$ the number n of quantum states $|\psi_k\rangle$ reported by Alice as being successfully measured tends to its expectation value

$E(n) = P_{\rm det} N$ with probability tending to unity. Thus, if $P_{\rm det} < \gamma_{\rm det}$ then $n < \gamma_{\rm det} N$ and Bob aborts with probability tending to unity for $N \to \infty$.

Proof of Lemma 2. Let $P_{\rm abort}$ be the probability that Bob aborts the token scheme if Alice and Bob follow the token scheme honestly. By definition of the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$, we have

$$P_{\rm abort} = \Pr[n < \gamma_{\rm det} N]. \tag{60}$$

We note that the expectation value of n is $E(n) = N P_{\rm det}$. From (58), we have that $0 < 1 - \frac{\gamma_{\rm det}}{P_{\rm det}} < 1$. Thus, we obtain from a Chernoff bound of Proposition 1 that

$$\Pr[n < \gamma_{\rm det} N] < e^{-\frac{P_{\rm det} N}{2}\left(1 - \frac{\gamma_{\rm det}}{P_{\rm det}}\right)^2}. \tag{61}$$

It follows from (60) and (61) that

$$P_{\rm abort} < \epsilon_{\rm rob}, \tag{62}$$

with $\epsilon_{\rm rob}$ given by (59). It follows from (62) that the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{\rm rob}$-robust with $\epsilon_{\rm rob}$ given by (59).

□

B. Proof of Lemma 3

Lemma 3. *If*

$$\begin{aligned} &0 < \frac{\gamma_{err}}{2} < E < \gamma_{err}, \\ &0 < \nu_{cor} < \frac{P_{det}(1 - 2\beta_{PB})}{2}, \end{aligned} \tag{63}$$

then $\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{cor}-$correct with

$$\epsilon_{cor} = e^{-\frac{P_{det}(1-2\beta_{PB})N}{4}\left(1 - \frac{2\nu_{cor}}{P_{det}(1-2\beta_{PB})}\right)^2} + e^{-\frac{E\nu_{cor}N}{3}\left(\frac{\gamma_{err}}{E} - 1\right)^2}. \tag{64}$$

We recall that $E = \max_{t,u}\{E_{tu}\}$, where E_{tu} is Alice's error rate when Bob prepares states $|\phi^k_{tu}\rangle$ and Alice measures in the basis of preparation by Bob, for $t, u \in \{0, 1\}$. The condition

$$E_{\rm min} < \gamma_{\rm err}, \tag{65}$$

with $E_{\rm min} = \min_{t,u}\{E_{tu}\}$, is necessary to guarantee correctness. To see this, suppose that $E_{\rm min} > \gamma_{\rm err}$. In the limit $N \to \infty$, we have $|\Delta_b| \to \infty$, in which case the number of error outcomes $n_{\rm errors}$ when Alice measures in the same basis of preparation by Bob satisfies $n_{\rm errors} \geq E_{\rm min}|\Delta_b| > \gamma_{\rm err}|\Delta_b|$ with probability tending to unity. Thus, with probability tending to unity, Bob does not accept Alice's token as valid, for $N \to \infty$, if $E_{\rm min} > \gamma_{\rm err}$.

Proof of Lemma 3. Let $P_{\rm error}$ be the probability that Bob does not accept Alice's token as valid if Alice and Bob follow the token scheme honestly. By definition of the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$, we have

$$P_{\rm error} = \sum_{|\Delta_b|=0}^{N} P_{\rm error}(|\Delta_b|)\Pr(|\Delta_b|), \tag{66}$$

where

$$P_{\rm error}(|\Delta_b|) = \Pr\Big[n_{\rm errors} > |\Delta_b|\gamma_{\rm err}\big||\Delta_b|\Big], \tag{67}$$

and where $n_{\rm errors}$ is the number of bit errors in the substring $\mathbf{x}_b$ of the the token $\mathbf{x}$ that Alice presents to Bob at Q_b, compared to the bits of the substring $\mathbf{r}_b$ of $\mathbf{r}$ encoded by Bob. From (66), we have

$$\begin{aligned} P_{\rm error} = &\sum_{|\Delta_b| < \nu_{\rm cor} N} P_{\rm error}(|\Delta_b|)\Pr(|\Delta_b|) \\ &+ \sum_{|\Delta_b| \geq \nu_{\rm cor} N} P_{\rm error}(|\Delta_b|)\Pr(|\Delta_b|) \\ \leq &\Pr\big[|\Delta_b| < \nu_{\rm cor} N\big] \\ &+ \sum_{|\Delta_b| \geq \nu_{\rm cor} N} P_{\rm error}(|\Delta_b|)\Pr(|\Delta_b|). \end{aligned} \tag{68}$$

We show below that

$$P_{\rm error}(|\Delta_b|) < e^{-\frac{E|\Delta_b|}{3}\left(\frac{\gamma_{\rm err}}{E} - 1\right)^2}, \tag{69}$$

and that

$$\Pr[|\Delta_b| < \nu_{\rm cor} N] \leq e^{-\frac{P_{\rm det}(1-2\beta_{\rm PB})N}{4}\left(1 - \frac{2\nu_{\rm cor}}{P_{\rm det}(1-2\beta_{\rm PB})}\right)^2}. \tag{70}$$

From (68) – (70), and noting that $e^{-\frac{E|\Delta_b|}{3}\left(\frac{\gamma_{\rm err}}{E} - 1\right)^2}$ decreases with increasing $|\Delta_b|$, we obtain

$$P_{\rm error} < \epsilon_{\rm cor}, \tag{71}$$

with $\epsilon_{\rm cor}$ given by (64). Thus, the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{\rm cor}-$correct with $\epsilon_{\rm cor}$ given by (64), as claimed.

We show (69). Let us assume for now that $E_{tu} = E$ for $t, u \in \{0, 1\}$. Given $|\Delta_b|$, we note that the expectation value of $n_{\rm error}$ equals $E|\Delta_b|$. From (63), we have $0 < \frac{\gamma_{\rm err}}{E} - 1 < 1$. Thus, from a Chernoff bound of Proposition 1, we have

$$\Pr\big[n_{\rm errors} > |\Delta_b|\gamma_{\rm err}\big||\Delta_b|\big] < e^{-\frac{E|\Delta_b|}{3}\left(\frac{\gamma_{\rm err}}{E} - 1\right)^2}. \tag{72}$$

The function $f(E) = E\big(\frac{\gamma_{\rm err}}{E} - 1\big)^2$ is decreasing with increasing E, because from (63) we have that $f'(E) = 1 - (\frac{\gamma_{\rm error}}{E})^2 < 0$. Let $E_{\rm max} \geq E$. Thus, from (72), we have

$$\Pr\big[n_{\rm errors} > |\Delta_b|\gamma_{\rm err}\big||\Delta_b|\big] < e^{-\frac{E_{\rm max}|\Delta_b|}{3}\left(\frac{\gamma_{\rm err}}{E_{\rm max}} - 1\right)^2}. \tag{73}$$

It follows from (67) and (73) that

$$P_{\rm error}(|\Delta_b|) < e^{-\frac{E_{\rm max}|\Delta_b|}{3}\left(\frac{\gamma_{\rm err}}{E_{\rm max}} - 1\right)^2}. \tag{74}$$

Since in general we have $E_{tu} \leq E$, for $t, u \in \{0,1\}$, we can replace $E_{\max}$ by E in (74) and obtain (69).

We show (70). Since for the quantum state $|\psi_k\rangle$, with $g(k) = j$, for $k \in \Lambda$ and $j \in [n]$, $\mathcal{B}$ encodes the bit $t_k = r_j$ in the basis labelled by $u_k = s_j$, with u_k chosen with probability $P^k_{\rm PB}(u_k)$ satisfying $\frac{1}{2} - \beta_{\rm PB} \leq P^k_{\rm PB}(u_k) \leq \frac{1}{2} + \beta_{\rm PB}$ for $t_k, u_k \in \{0,1\}$, the expectation value $E(|\Delta_b|)$ of $|\Delta_b|$ satisfies

$$E(|\Delta_b|) \geq P_{\rm det} N\Big(\frac{1}{2} - \beta_{\rm PB}\Big). \tag{75}$$

This is easily seen as follows. By the definition of Δ_b given in the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$, we see that $|\Delta_b|$ corresponds to the number of labels $k \in \Lambda$ satisfying $g(k) = j \in [n]$ for which it holds that $y_j = s_j$, where we recall y_j and s_j are the bits labelling the qubit measurement basis by Alice and the preparation basis by Bob, respectively. Thus, $E(|\Delta_b|) = P_{\rm det} N \Pr[y_j = s_j] = P_{\rm det} N \sum_{a=0}^{1} \Pr[s_j = a] \Pr[y_j = a] \geq P_{\rm det} N\Big(\frac{1}{2} - \beta_{\rm PB}\Big)$, as claimed. We define

$$\epsilon = 1 - \frac{2\nu_{\rm cor}}{P_{\rm det}(1 - 2\beta_{\rm PB})}. \tag{76}$$

From the condition (63), we have $0 < \epsilon < 1$. It follows that

$$\nu_{\rm cor} N = (1-\epsilon) P_{\rm det} N\Big(\frac{1}{2} - \beta_{\rm PB}\Big) = (1-\epsilon') E(|\Delta_b|), \tag{77}$$

for some ϵ' satisfying $0 < \epsilon \leq \epsilon' < 1$. Thus, from the Chernoff bound of Proposition 1, we have

$$\begin{aligned} \Pr[|\Delta_b| < \nu_{\rm cor} N] &= \Pr[|\Delta_b| < (1-\epsilon') E(|\Delta_b|)] \\ &\leq e^{-\frac{E(|\Delta_b|)}{2}\epsilon'^2} \\ &\leq e^{-\frac{P_{\rm det} N (1-2\beta_{\rm PB})}{4}\epsilon^2} \\ &= e^{-\frac{P_{\rm det}(1-2\beta_{\rm PB})N}{4}\left(1 - \frac{2\nu_{\rm cor}}{P_{\rm det}(1-2\beta_{\rm PB})}\right)^2}, \end{aligned} \tag{78}$$

where in the first line we used (77); in the second line we used the Chernoff bound of Proposition 1; in the third line we used (75) and $0 < \epsilon \leq \epsilon'$; and in the last line we used (76).

□

C. Proof of Lemma 4

Lemma 4. *$\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{priv}-$private with*

$$\epsilon_{priv} = \beta_E. \tag{79}$$

Proof. From assumption C (see Table 6), the set Λ of labels transmitted to $\mathcal{B}$ in step 2 of $\mathcal{QT}_1$ and $\mathcal{QT}_1$ gives $\mathcal{B}$ no information about the string W and the bit z. Furthermore, from assumption E (see Table 6), $\mathcal{B}$ cannot use degrees of freedom not previously agreed for the transmission of the quantum states to affect, or obtain information about, the statistics of the quantum measurement devices of $\mathcal{A}$. Moreover, in our setting, we assume that Alice's laboratories are secure and that communication among Alice's agents is made through secure and authenticated classical channels. It follows from these assumptions that the only way in which Bob can obtain information about Alice's bit b before she presents the token is via the message $c = z \oplus b$.

In order to prove our result, let us assume that Bob knows Alice's probability distributions $P_{\rm E}(z)$. Since this cannot make it more difficult for Bob to guess Alice's bit b, we can assume this without loss of generality. In the ideal case that the probability distribution $P_{\rm E}(z)$ is totally random Bob cannot obtain any information about b. However, as stated by our allowed experimental imperfection 7 (see Table 5), this probability distribution is only close to random:

$$\frac{1}{2} - \beta_{\rm E} \leq P_{\rm E}(z) \leq \frac{1}{2} + \beta_{\rm E}, \tag{80}$$

for a small parameters $\beta_{\rm E} > 0$, for $z \in \{0,1\}$. Thus, Bob can guess b with some probability greater than $\frac{1}{2}$.

Let $P^{(i)}_{\rm bit}(c)$ be the probability distribution for the bit c that $\mathcal{A}$ sends $\mathcal{B}$, when $b = i$, for $i, c \in \{0,1\}$. Since $c = b \oplus z$, we have

$$P^{(i)}_{\rm bit}(c) = P_{\rm E}(z = i \oplus c), \tag{81}$$

for $c, i \in \{0,1\}$.

For any two probability distributions $P(x)$ and $Q(x)$ over a set of values $x \in \mathcal{X}$, the maximum probability $P_{\max}$ to distinguish them is given by $P_{\max} = \frac{1}{2} + \frac{1}{2}\|P - Q\|$, where $\|P - Q\|$ is their variational distance. Thus, Bob's probability $P_{\rm Bob}$ to guess Alice's bit b is upper bounded by

$$\begin{aligned} P_{\rm Bob} &\leq \frac{1}{2} + \frac{1}{2}\big\|P^{(0)}_{\rm bit} - P^{(1)}_{\rm bit}\big\| \\ &= \frac{1}{2} + \frac{1}{4}\sum_{c=0}^{1}\big|P^{(0)}_{\rm bit}(c) - P^{(1)}_{\rm bit}(c)\big| \\ &= \frac{1}{2} + \frac{1}{4}\sum_{c=0}^{1}\big|P_{\rm E}(z = c) - P_{\rm E}(z = c \oplus 1)\big| \\ &\leq \frac{1}{2} + \frac{1}{4}\sum_{c=0}^{1}\big|2\beta_{\rm E}\big| \\ &= \frac{1}{2} + \beta_{\rm E}, \end{aligned} \tag{82}$$

where in the second line we used the definition of the variational distance; in the third line we used (81); and in the fourth line we used (80).

It follows from (82) that the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{\rm priv}$-private, with $\epsilon_{\rm priv}$ given by (79), as claimed. □

VI. PROOF OF LEMMA 5

Lemma 5. *Suppose that Bob sends Alice N photon pulses, labelled by $k \in [N]$. Let the kth pulse have L_k photons. Let ρ be an arbitrary quantum state prepared by Bob in the polarization degrees of freedom of the photons sent to Alice, which can be arbitrarily entangled among all photons in all pulses and can also be arbitrarily entangled with an ancilla held by Bob. Let $\mathcal{D}_0$ and $\mathcal{D}_1$ be two arbitrary qubit orthogonal bases. Suppose that either Alice uses the setup of Fig. 3 with reporting strategy 1 to implement the quantum token scheme $\mathcal{QT}_1$ (see Table 2), or Alice uses the setup of Fig. 3 with reporting strategy 2 to implement the quantum token scheme $\mathcal{QT}_2$ (see Table 3). Suppose also that assumptions E and F (see Table 6) hold. For $k \in [N]$, let $m_k = 1$ if Alice assigns a successful measurement to the kth pulse and $m_k = 0$ otherwise; let $w_k = 0$ ($w_k = 1$) if Alice assigns a measurement basis to the kth pulse in the basis $\mathcal{D}_0$ ($\mathcal{D}_1$). If Alice uses the setup of Fig. 3 and reporting strategy 1 to implement the scheme $\mathcal{QT}_1$, without loss of generality, suppose also that Alice sets $w_k = 0$ with unit probability, if $m_k = 0$, for $k \in [N]$. Let $m = (m_1, \ldots, m_N)$, $w = (w_1, \ldots, w_N)$ and $L = (L_1, \ldots, L_N)$.*

If Alice uses the setup of Fig. 3 with reporting strategy 1 to implement the scheme $\mathcal{QT}_1$, then the probability that Alice reports the string m to Bob and assigns the string of measurement bases w, given ρ and L, is

$$P^{(1)}_{rep}(m, w|\rho, L) = \prod_{k=1}^{N} G^{(1)}_{m_k, w_k}(d_0, d_1, \eta, L_k), \qquad (83)$$

where

$$\begin{aligned} G^{(1)}_{1,b}(d_0, d_1, \eta, a) &= (1-d_0)(1-d_1)\Big(1-\frac{\eta}{2}\Big)^a \\ &\quad -(1-d_0)^2(1-d_1)^2(1-\eta)^a, \\ G^{(1)}_{0,0}(d_0, d_1, \eta, a) &= 1 - 2G^{(1)}_{1,0}(d_0, d_1, \eta, a), \\ G^{(1)}_{0,1}(d_0, d_1, \eta, a) &= 0, \end{aligned} \qquad (84)$$

for $b \in \{0,1\}$, $m, w \in \{0,1\}^N$ and $a, L_1, \ldots, L_N \in \{0, 1, 2, \ldots\}$. Furthermore, the probability $P_{MB}(w_k)$ that Alice assigns a measurement in the basis $\mathcal{D}_{w_k}$, conditioned on the value $m_k = 1$, for the kth pulse, satisfies

$$P_{MB}(w_k) = \frac{1}{2}, \qquad (85)$$

for $w_k \in \{0,1\}$ and $k \in [N]$.

If Alice uses the setup of Fig. 3 with reporting strategy 2 to implement the scheme $\mathcal{QT}_2$, then the probability that Alice reports the string m to Bob, given ρ, w and L, is

$$P^{(2)}_{rep}(m|w, \rho, L) = \prod_{k=1}^{N} G^{(2)}_{m_k}(d_0, d_1, \eta, L_k), \qquad (86)$$

where

$$\begin{aligned} G^{(2)}_0(d_0, d_1, \eta, a) &= (1-d_0)(1-d_1)(1-\eta)^a, \\ G^{(2)}_1(d_0, d_1, \eta, a) &= 1 - (1-d_0)(1-d_1)(1-\eta)^a, \end{aligned} \qquad (87)$$

for $m, w \in \{0,1\}^N$ and $a, L_1, \ldots, L_N \in \{0, 1, 2, \ldots\}$.

In any of the two cases, the message m gives Bob no information about the bit entries w_k for which $m_k = 1$. Equivalently, the set $\Lambda \subset [N]$ of labels transmitted to Bob in step 2 of $\mathcal{QT}_1$ and $\mathcal{QT}_2$ gives Bob no information about the string W and the bit z.

Proof. We note from (83) and (84) that if Alice uses the setup of Fig. 3 with reporting strategy 1 to implement the scheme $\mathcal{QT}_1$, then Alice's probability $P^{(1)}_{\rm rep}(m, w|\rho, L)$ to report the message m to Bob and assign measurement basis with string of labels w is the same for all strings w satisfying that $w_k = 0$ if $m_k = 0$, for arbitrary fixed given values of m, ρ and L. It follows from this and from assumption E (see Table 6) that the message m gives Bob no information about the bit entries w_k for which $m_k = 1$.

Similarly, we note from (86) and (87) that if Alice uses the setup of Fig. 3 with reporting strategy 2 to implement the scheme $\mathcal{QT}_2$, then Alice's probability $P^{(2)}_{\rm rep}(m|w, \rho, L)$ to report the message m to Bob, given that she applied quantum measurements with string of labels w, is the same for all strings w, for arbitrary fixed given values of m, ρ and L. It follows from this and from assumption E (see Table 6) that the message m gives Bob no information about any bit entries w_k of w. We note that in the scheme $\mathcal{QT}_2$, $w_k = z$ for $k \in [N]$, and for some bit z chosen by Alice. Thus, it follows that the message m gives Bob no information about the bit z.

Alice sending the message m to Bob is equivalent to Alice sending the set Λ of labels $k \in [N]$ for which $m_k = 1$, i.e. the labels of pulses that were successfully measured by Alice. Furthermore, the string W is defined on the set of labels Λ, and has entries equal to w_k, for $k \in \Lambda$, i.e. for $k \in [N]$ satisfying $m_k = 1$. It follows that the set $\Lambda \subset [N]$ of labels transmitted to Bob in step 2 of $\mathcal{QT}_1$ and $\mathcal{QT}_2$ gives Bob no information about the string W and the bit z.

We prove (83). We suppose that Alice uses the setup of Fig. 3 with reporting strategy 1 to implement the scheme $\mathcal{QT}_1$, and that assumptions E and F of Table 6 hold. It is straightforward to obtain

$$P^{(1)}_{\rm rep}(m, w|\rho, L) = \prod_{k=1}^{N} P^{(1,k)}_{\rm rep}(m_k, w_k|\rho, L, \tau_k), \qquad (88)$$

where $\tau_k = (m_0, w_0, \ldots, m_{k-1}, w_{k-1})$, $P^{(1,k)}_{\rm rep}(m_k, w_k|\rho, L, \tau_k)$ is the probability that Alice reports the bit message m_k to Bob and assigns a measurement in the basis $\mathcal{D}_{w_k}$ for the kth pulse, given ρ, L, and τ_k, for $k \in [N]$; and where without loss of generality we define $m_0 = w_0 = 1$.

From Lemma 12 of Ref. [1] and the definition (84), after measuring the first pulse received from Bob, Alice reports the message $m_1 = 1$ to Bob and assigns a measurement outcome in the basis $\mathcal{D}_{w_1}$ with probability

$$P^{(1,1)}_{\rm rep}(1, w_1|\rho, L, \tau_1) = G^{(1)}_{1,w_1}(d_0, d_1, \eta, L_1), \qquad (89)$$

for $w_1 \in \{0,1\}$, which only depends on the dark count probabilities d_0 and d_1, on the detector efficiency η, and on the number of photons L_1 of the first pulse, for an arbitrary quantum state ρ, which can be arbitrarily entangled with an ancilla held by Bob. We can consider this ancilla to include the pulses labelled by $2, 3, \ldots, N$.

Similarly, for $k \in \{2, 3, \ldots, N\}$, after measuring the pulses with labels $1, 2, \ldots, k-1$, Alice obtains a value τ_k according to her obtained detection statistics for these pulses, and the joint quantum state of the pulses with labels $k, k+1, \ldots, N$ and any ancilla held by Bob changes to some quantum state ρ_k. Then, from Lemma 12 of Ref. [1] and the definition (84), after measuring the kth pulse, Alice reports the message $m_k = 1$ to Bob and assigns a measurement outcome in the basis $\mathcal{D}_{w_k}$ with probability

$$P_{\text{rep}}^{(1,k)}(1, w_k|\rho, L, \tau_k) = G_{1,w_k}^{(1)}(d_0, d_1, \eta, L_k), \tag{90}$$

for $w_k \in \{0,1\}$, which only depends on the dark count probabilities d_0 and d_1, on the detector efficiency η, and on the number of photons L_k of the kth pulse; in particular, $P_{\text{rep}}^{(1,k)}(1, w_k|\rho, L, \tau_k)$ does not depend on the quantum state ρ_k, which can be arbitrarily entangled with an ancilla held by Bob. In this case, we can consider this ancilla to include the pulses labelled by $k+1, k+2, \ldots, N$. We note that since $P_{\text{rep}}^{(1,k)}(1, w_k|\rho, L, \tau_k)$ does not depend on ρ_k, it does not depend on ρ, apart from the number of photons L_k of the kth pulse, and it does not depend on τ_k either.

By definition of Alice's reporting strategy 1, we have

$$\begin{aligned} P_{\text{rep}}^{(1,k)}(0, 0|\rho, L, \tau_k) &= 1 - P_{\text{rep}}^{(1,k)}(1, 0|\rho, L, \tau_k) \\ &\quad - P_{\text{rep}}^{(1,k)}(1, 1|\rho, L, \tau_k) \\ &= 1 - 2G_{1,0}^{(1)}(d_0, d_1, \eta, L_k) \\ &= G_{0,0}^{(1)}(d_0, d_1, \eta, L_k), \end{aligned} \tag{91}$$

for $k \in [N]$, where in the second line we used (89) and (90), and the definition (84); and in the third line we used (84) again. Similarly, by definition of Alice's reporting strategy and from the definition (84), we have

$$P_{\text{rep}}^{(1,k)}(0, 1|\rho, L, \tau_k) = G_{0,1}^{(1)}(d_0, d_1, \eta, L_k), \tag{92}$$

for $k \in [N]$. Thus, the claimed result (83) follows straightforwardly from (88) – (91).

We prove (85). Let $P_{\text{MB}}^{(1,k)}(w_k|m_k = 1, \rho, L, \tau_k)$ be the probability that Alice assigns a measurement in the basis $\mathcal{D}_{w_k}$, conditioned on the value $m_k = 1$, for the kth pulse, given the values of ρ, L and τ_k, for $k \in [N]$. We have

$$\begin{aligned} &P_{\text{MB}}^{(1,k)}(w_k|m_k = 1, \rho, L, \tau_k) \\ &\quad = \frac{P_{\text{rep}}^{(1,k)}(1, w_k|\rho, L, \tau_k)}{P_{\text{rep}}^{(1,k)}(1, 0|\rho, L, \tau_k) + P_{\text{rep}}^{(1,k)}(1, 1|\rho, L, \tau_k)} \\ &\quad = \frac{G_{1,w_k}^{(1)}(d_0, d_1, \eta, L_k)}{2G_{1,w_k}^{(1)}(d_0, d_1, \eta, L_k)} \\ &\quad = \frac{1}{2}, \end{aligned} \tag{93}$$

for $w_k \in \{0,1\}$ and $k \in [N]$; where in the second line we used (89) and (90), and the definition (84); and in the third line we used that $G_{1,w_k}^{(1)}(d_0, d_1, \eta, L_k) > 0$ from (84) and from the fact that $d_0, d_1, \eta \in (0,1)$, as stated in assumption F (see Table 6). From (93), since $P_{\text{MB}}^{(1,k)}(w_k|m_k = 1, \rho, L, \tau_k)$ does not depend on k, ρ, L or τ_k, we have that

$$P_{\text{MB}}^{(1,k)}(w_k|m_k = 1, \rho, L, \tau_k) = P_{\text{MB}}(w_k), \tag{94}$$

for $w_k \in \{0,1\}$ and $k \in [N]$, and the claimed result (85) follows.

We prove (86). We suppose that Alice uses the setup of Fig. 3 with reporting strategy 2 to implement the scheme $\mathcal{QT}_2$, and that assumptions E and F of Table 6 hold. We note that in the scheme $\mathcal{QT}_2$, the string w has bit entries $w_k = z$, for a bit z chosen by Alice, and for $k \in [N]$. However, the analysis below is more general, and works for arbitrary $w \in \{0,1\}^N$. It is straightforward to obtain

$$P_{\text{rep}}^{(2)}(m|w, \rho, L) = \prod_{k=1}^{N} P_{\text{rep}}^{(2,k)}(m_k|w, \rho, L, \tilde{\tau}_k), \tag{95}$$

where $\tilde{\tau}_k = (m_0, \ldots, m_{k-1})$ and $P_{\text{rep}}^{(2,k)}(m_k|w, \rho, L, \tilde{\tau}_k)$ is the probability that Alice reports the bit message m_k to Bob for the kth pulse, given ρ, L, w and $\tilde{\tau}_k$, for $k \in [N]$; and where without loss of generality we define $m_0 = 1$.

From Lemma 1 of Ref. [1] and the definition (87), after measuring the first pulse received from Bob in the basis $\mathcal{D}_{w_1}$, Alice reports the message $m_1 = 1$ to Bob with probability

$$P_{\text{rep}}^{(2,1)}(1|w, \rho, L, \tilde{\tau}_1) = G_1^{(2)}(d_0, d_1, \eta, L_1), \tag{96}$$

which only depends on the dark count probabilities d_0 and d_1, on the detector efficiency η, and on the number of photons L_1 of the first pulse, for an arbitrary quantum state ρ, which can be arbitrarily entangled with an ancilla held by Bob. We can consider this ancilla to include the pulses labelled by $2, 3, \ldots, N$.

Similarly, for $k \in \{2, 3, \ldots, N\}$, after measuring the pulses with labels $1, 2, \ldots, k-1$ in the bases $\mathcal{D}_{w_1}, \ldots, \mathcal{D}_{w_{k-1}}$, Alice obtains a value $\tilde{\tau}_k$ according to her obtained detection statistics for these pulses, and the joint quantum state of the pulses with labels $k, k+$

$1,\ldots,N$ and any ancilla held by Bob changes to some quantum state ρ_k. Then, from Lemma 1 of Ref. [1] and the definition (87), after measuring the kth pulse in the basis $\mathcal{D}_{w_k}$, Alice reports the message $m_k = 1$ to Bob with probability

$$P_{\text{rep}}^{(2,k)}(1|w,\rho,L,\tilde{\tau}_k) = G_1^{(2)}(d_0,d_1,\eta,L_k), \tag{97}$$

which only depends on the dark count probabilities d_0 and d_1, on the detector efficiency η, and on the number of photons L_k of the kth pulse; in particular, $P_{\text{rep}}^{(2,k)}(1|w,\rho,L,\tilde{\tau}_k)$ does not depend on the quantum state ρ_k, which can be arbitrarily entangled with an ancilla held by Bob. In this case, we can consider this ancilla to include the pulses labelled by $k+1,k+2,\ldots,N$. We note that since $P_{\text{rep}}^{(2,k)}(1|w,\rho,L,\tilde{\tau}_k)$ does not depend on ρ_k, it does not depend on ρ, apart from the number of photons L_k of the kth pulse, and it does not depend on $\tilde{\tau}_k$ either.

By definition of Alice's reporting strategy 2, we have

$$\begin{aligned} &P_{\text{rep}}^{(2,k)}(0|w,\rho,L,\tilde{\tau}_k) \\ &\quad = 1 - P_{\text{rep}}^{(2,k)}(1|w,\rho,L,\tilde{\tau}_k) \\ &\quad = 1 - G_1^{(2)}(d_0,d_1,\eta,L_k) \\ &\quad = G_0^{(2)}(d_0,d_1,\eta,L_k), \end{aligned} \tag{98}$$

for $k \in [N]$, where in the second line we used (96) and (97), and in the third line we used the definition (87). Thus, the claimed result (86) follows straightforwardly from (95) – (98).

□

VII. PROOF OF THEOREM 1

Theorem 1. *Consider the constraints*

$$\begin{aligned} &0 < \gamma_{err} < \lambda(\theta,\beta_{PB}), \\ &0 < P_{noqub} < \nu_{unf} < \min\left\{2P_{noqub}, \gamma_{det}\left(1-\frac{\gamma_{err}}{\lambda(\theta,\beta_{PB})}\right)\right\}, \\ &0 < \beta_{PS} < \frac{1}{2}\left[e^{\frac{\lambda(\theta,\beta_{PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{PB})}\right)^2} - 1\right]. \end{aligned} \tag{99}$$

We define the function

$$\begin{aligned} &f(\gamma_{err},\beta_{PS},\beta_{PB},\theta,\nu_{unf},\gamma_{det}) \\ &\quad = (\gamma_{det}-\nu_{unf})\left[\frac{\lambda(\theta,\beta_{PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{PB})}\right)^2 - \ln(1+2\beta_{PS})\right] \\ &\qquad -\big(1-(\gamma_{det}-\nu_{unf})\big)\ln\big[1+h(\beta_{PS},\beta_{PB},\theta)\big], \end{aligned} \tag{100}$$

where

$$\begin{aligned} h(\beta_{PS},\beta_{PB},\theta) &= 2\beta_{PS}\sqrt{\frac{1}{2}+2\beta_{PB}^2+\left(\frac{1}{2}-2\beta_{PB}^2\right)\sin(2\theta)}, \\ \delta &= \frac{\gamma_{det}\gamma_{err}}{\gamma_{det}-\nu_{unf}}. \end{aligned} \tag{101}$$

There exist parameters satisfying the constraints (99), for which $f(\gamma_{err},\beta_{PS},\beta_{PB},\theta,\nu_{unf},\gamma_{det}) > 0$. For these parameters, $\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{unf}-$unforgeable with

$$\epsilon_{unf} = e^{-\frac{P_{noqub}N}{3}\left(\frac{\nu_{unf}}{P_{noqub}}-1\right)^2} + e^{-Nf(\gamma_{err},\beta_{PS},\beta_{PB},\theta,\nu_{unf},\gamma_{det})}. \tag{102}$$

We recall that Theorem 1 considers parameters $\gamma_{\text{det}},\gamma_{\text{err}} \in (0,1)$, allows for the experimental imperfections of Table 5 and makes the assumptions of Table 6.

A. Summary of the Proof

We allow Alice to have arbitrarily advanced quantum technology. In particular, we assume that Alice knows the set Ω_{qub} and Ω_{noqub} and that she receives all quantum systems A_k, for $k \in [N]$. Let $|\psi\rangle_A = \otimes_{k\in[N]}|\psi_k\rangle_{A_k}$ be the quantum state that Alice receives from Bob, where A is the global quantum system received from Bob.

Alice's more general cheating strategy in the token scheme $\mathcal{QT}_1$ is as follows. In the intersection of the causal pasts of Q_0 and Q_1, Alice adds an ancillary system E of arbitrary finite Hilbert space dimension in a quantum state $|\chi\rangle_E$ and applies an arbitrary projective measurement on AE, which may depend on Ω_{qub} and Ω_{noqub}, and obtains an outcome that includes $\Lambda, g, \mathbf{d}$ and c satisfying the required constraints, as well as respective tokens $\mathbf{a}$ and $\mathbf{b}$ to give at the presentation points Q_0 and Q_1. Alice sends Λ, g, $\mathbf{d}$, and c to Bob, as required by the task. Alice gives Bob tokens $\mathbf{a}$ at Q_0 and $\mathbf{b}$ at Q_1. Alice's more general cheating strategy in the token scheme $\mathcal{QT}_2$ is equivalent, with the only difference that the string $\mathbf{d}$ is not required in Alice's measurement outcome. We recall that the jth entry of $\mathbf{d}$ in $\mathcal{QT}_1$ is $d_j = y_j \oplus z$, for $j \in [n]$. On the other hand, in $\mathcal{QT}_2$ we have $y_j = z$, for $j \in [n]$. Thus, without loss of generality, in Alice's general cheating strategy in the token scheme $\mathcal{QT}_2$ we simply set $\mathbf{d}$ to be a fixed string with all bit entries being zero.

We note that if Q_1 is in the causal future of Q_0 Alice's agent $\mathcal{A}_0$ can send a signal to $\mathcal{A}_1$ indicating whether her token $\mathbf{a}$ was validated or not at Q_0, and $\mathcal{A}_1$ can in principle use this information to adapt her strategy at Q_1. However, this possibility cannot increase Alice's probability to have tokens validated at both Q_0 and Q_1. If $\mathcal{A}_0$ fails in having $\mathbf{a}$ validated at Q_0 then Alice fails in her attempt to have Bob validating her tokens at both Q_0 and Q_1. Thus, without loss of generality we assume that the signal sent from $\mathcal{A}_0$ to $\mathcal{A}_1$ indicates that $\mathbf{a}$ was successfully validated at Q_0, which is equivalent to $\mathcal{A}_0$ not sending any signal to $\mathcal{A}_1$ and $\mathcal{A}_1$ always acting as if $\mathbf{a}$ were successfully validated at Q_0. The same reasoning applies if Q_0 is in the causal future of Q_1. Furthermore, if Q_0 and Q_1 are spacelike separated $\mathcal{A}_1$ cannot receive any signals informing her whether the token $\mathbf{a}$ was successfully validated at Q_0 before $\mathcal{A}_1$ presents the token $\mathbf{b}$ at Q_1. This means that the strategy outlined in the previous paragraph can be considered as the most general one.

If $|\Omega_{\rm noqub}| > \nu_{\rm unf} N$ then we assume that Alice can succeed in giving valid tokens $\mathbf{a}$ and $\mathbf{b}$ at Q_0 and Q_1. This is because, for example, if $|\Omega_{\rm noqub}|$ is too large, then there is not any constraint on the quantum states $|\psi_k\rangle$ for a large number of labels k, i.e for $k \in \Omega_{\rm noqub}$. For example, if Bob's quantum state source is a Poissonian photon source (e.g. weak coherent) with average photon number $\mu << 1$ then we associate pulses with two or more photons to have labels from the set $\Omega_{\rm noqub}$, giving $P_{\rm noqub} = 1 - (1+\mu)e^{-\mu}$. In this case, the states $|\psi_k\rangle$ with $k \in \Omega_{\rm noqub}$ consist of two or more copies of quantum states $|\phi^k_{t_k u_k}\rangle$ and Alice can measure each copy in the corresponding basis, being able to present correct bit outcomes at Q_0 and Q_1, for bit entries with labels $k \in \Omega_{\rm noqub}$. But, since the probability $P_{\rm noqub}$ that Bob prepares states $|\psi_k\rangle$ with labels $k \in \Omega_{\rm noqub}$ is bounded, the probability that $|\Omega_{\rm noqub}| > \nu_{\rm unf} N$ is bounded by the first term of $\epsilon_{\rm unf}$ in (102).

On the other hand, we show that there exist parameters satisfying the constraints (99) for which $f(\gamma_{\rm err}, \beta_{\rm PS}, \beta_{\rm PB}, \theta, \nu_{\rm unf}, \gamma_{\rm det}) > 0$. We show that, for these parameters, and for the case $|\Omega_{\rm noqub}| \leq \nu_{\rm unf} N$, the probability that Alice gives valid tokens $\mathbf{a}$ and $\mathbf{b}$ at Q_0 and Q_1 is upper bounded by $e^{-N f(\gamma_{\rm err}, \beta_{\rm PS}, \beta_{\rm PB}, \theta, \nu_{\rm unf}, \gamma_{\rm det})}$. This analysis gives the second term of $\epsilon_{\rm unf}$ in (102).

If $|\Omega_{\rm noqub}| \leq \nu_{\rm unf} N$, from the conditions (99) it follows that Alice must obtain the correct bits in two token strings $\mathbf{a}$ and $\mathbf{b}$, given respectively at Q_0 and Q_1, for a sufficiently large number of entries $j \in \Delta_0$ with $j = g(k)$ for some $k \in \Omega_{\rm qub}$, for $\mathbf{a}$, and $j \in \Delta_1$ with $j = g(k)$ for some $k \in \Omega_{\rm qub}$, for $\mathbf{b}$. That is, Alice could in principle learn perfectly the bit entries of both strings $\mathbf{a}_0$ and $\mathbf{b}_1$ with labels $k \in \Omega_{\rm noqub}$, but the number of these entries is small and thus Alice must be able to obtain the correct bit entries of the strings $\mathbf{a}_0$ and $\mathbf{b}_1$ for a sufficiently large number of labels $k \in \Omega_{\rm qub}$ for which such bits are encoded in single qubits, and not in multiple copies of the same quantum states. The probability that Alice succeeds in this task is upper bounded using Lemma 6, which considers the ideal situation in which Bob encodes each bit in a single qubit.

We see that the $n-$bit strings $\tilde{\mathbf{d}}_0 = (\tilde{d}_{0,1}, \ldots, \tilde{d}_{0,n})$ and $\tilde{\mathbf{d}}_1 = (\tilde{d}_{1,1}, \ldots, \tilde{d}_{1,n})$ are the complement of each other. In the scheme $\mathcal{QT}_1$, we have $\tilde{d}_{1,j} = d_j \oplus c \oplus 1 = \tilde{d}_{0,j} \oplus 1$, for $j \in [n]$. Similarly, in the scheme $\mathcal{QT}_2$, we have $\tilde{d}_{1,j} = c \oplus 1 = \tilde{d}_{0,j} \oplus 1$, for $j \in [n]$. Thus, in both schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$, we define the sets of labels $\Delta_0 = \{j \in [n] | \tilde{d}_{0,j} = s_j\}$ and $\Delta_1 = \{j \in [n] | \tilde{d}_{1,j} = s_j\}$, which do not intersect, implying that $|\Delta_0| + |\Delta_1| = n$. We define the sets $\underline{\Delta}_0$ and $\underline{\Delta}_1$ as the sets of labels $j \in \Delta_0$ and $j \in \Delta_1$ with $j = g(k)$ for some $k \in \Omega_{\rm qub}$, respectively. We also define the strings $\underline{\mathbf{a}}_0$ and $\underline{\mathbf{b}}_1$ as the restrictions of the strings $\mathbf{a}_0$ and $\mathbf{b}_1$ to bit entries with labels $j \in \underline{\Delta}_0$ and $j \in \underline{\Delta}_1$, respectively. The strings $\underline{\mathbf{r}}_0$ and $\underline{\mathbf{r}}_1$ are defined similarly. Then, we can upper bound Alice's success probability by the probability that the string $\underline{\mathbf{a}}_0$ and the string $\underline{\mathbf{b}}_1$ satisfy $d(\underline{\mathbf{a}}_0, \underline{\mathbf{r}}_0) + d(\underline{\mathbf{b}}_1, \underline{\mathbf{r}}_1) \leq (|\underline{\Delta}_0| + |\underline{\Delta}_1|)\delta$, where $0 < \tilde{\delta} \leq \delta < \lambda(\theta, \beta_{\rm PB})$, and where $\tilde{\delta} = \frac{(|\Delta_0| + |\Delta_1|)\gamma_{\rm err}}{(|\underline{\Delta}_0| + |\underline{\Delta}_1|)}$. Finally, this probability is upper bounded using Lemma 6.

B. Preliminaries

We consider that Alice performs an arbitrary cheating strategy $\mathcal{S}$ trying to have Bob validating tokens at Q_0 and Q_1. Let $P_{\mathcal{S}}$ be Alice's success probability. We show an upper bound on $P_{\mathcal{S}}$, for any strategy $\mathcal{S}$. We assume that Alice has arbitrarily advanced quantum technology. In particular, we assume that Alice knows the set $\Omega_{\rm qub}$, hence also the set $\Omega_{\rm noqub}$, and that she receives all quantum systems A_k transmitted by Bob, for $k \in [N]$.

Let $P_{\mathcal{S}}^{\Omega_{\rm qub}}$ be Alice's success probability following the strategy $\mathcal{S}$ given a set $\Omega_{\rm qub}$, and let $P_{\Omega_{\rm qub}}$ be the probability that the set $\Omega_{\rm qub}$ is generated. We have

$$\begin{aligned} P_{\mathcal{S}} &= \sum_{m=0}^{N} \sum_{\Omega_{\rm qub}:|\Omega_{\rm noqub}|=m} P_{\mathcal{S}}^{\Omega_{\rm qub}} P_{\Omega_{\rm qub}} \\ &= \sum_{m \leq \nu_{\rm unf} N} \sum_{\Omega_{\rm qub}:|\Omega_{\rm noqub}|=m} P_{\mathcal{S}}^{\Omega_{\rm qub}} P_{\Omega_{\rm qub}} \\ &\quad + \sum_{m > \nu_{\rm unf} N} \sum_{\Omega_{\rm qub}:|\Omega_{\rm noqub}|=m} P_{\mathcal{S}}^{\Omega_{\rm qub}} P_{\Omega_{\rm qub}} \\ &\leq \sum_{m \leq \nu_{\rm unf} N} \sum_{\Omega_{\rm qub}:|\Omega_{\rm noqub}|=m} P_{\mathcal{S}}^{\Omega_{\rm qub}} P_{\Omega_{\rm qub}} \\ &\quad + \Pr[|\Omega_{\rm noqub}| > \nu_{\rm unf} N], \end{aligned} \tag{103}$$

where in the third line we used $\Pr[|\Omega_{\rm noqub}| > \nu_{\rm unf} N] = \sum_{m > \nu_{\rm unf} N} \sum_{\Omega_{\rm qub}:|\Omega_{\rm noqub}|=m} P_{\Omega_{\rm qub}}$ and the tivial bound $P_{\mathcal{S}}^{\Omega_{\rm qub}} \leq 1$ for $|\Omega_{\rm noqub}| > \nu_{\rm unf} N$.

Since for each element $k \in [N]$, Bob's agent $\mathcal{B}$ asigns it to be an element of the set $\Omega_{\rm noqub}$ with probability $P_{\rm noqub}$, the probability that $\Omega_{\rm noqub}$ has m elements is

$$\sum_{\Omega_{\rm qub}:|\Omega_{\rm noqub}|=m} P_{\Omega_{\rm qub}} = \binom{N}{m} (P_{\rm noqub})^m (1 - P_{\rm noqub})^{N-m}, \tag{104}$$

for $m \in \{0, 1, \ldots, N\}$. To simplify notation, below we write $m = |\Omega_{\rm noqub}|$. We use the Chernoff bound of Proposition 1 to show below that

$$\Pr[m > \nu_{\rm unf} N] < e^{-\frac{P_{\rm noqub} N}{3} \left(\frac{\nu_{\rm unf}}{P_{\rm noqub}} - 1 \right)^2}. \tag{105}$$

Let $Z_1, Z_2, \ldots, Z_N$ be independent random variables, where $Z_k \in \{0, 1\}$ and $\Pr[Z_k = 1] = P_{\rm noqub}$ for $k \in [N]$. We can then write $m = \sum_{k=1}^{N} Z_k$. The expectation value of m is $E(m) = P_{\rm noqub} N$. Let $\epsilon = \frac{\nu_{\rm unf}}{P_{\rm noqub}} - 1$. It follows from (99) that $0 < \epsilon < 1$. Thus, we have

$$\begin{aligned} \Pr[m > \nu_{\rm unf} N] &= \Pr[m > (1+\epsilon) E(m)] \\ &< e^{-\frac{P_{\rm noqub} N}{3} \left(\frac{\nu_{\rm unf}}{P_{\rm noqub}} - 1 \right)^2}, \end{aligned} \tag{106}$$

as claimed, where in the second line we used $E(m) = P_{\rm noqub}N$, $0 < \epsilon = \frac{\nu_{\rm unf}}{P_{\rm noqub}} - 1 < 1$ and the Chernoff bound of Proposition 1.

From (103) and (105), we have

$$P_{\mathcal{S}} < e^{-\frac{P_{\rm noqub}N}{3}\left(\frac{\nu_{\rm unf}}{P_{\rm noqub}}-1\right)^2} + \sum_{m\leq \nu_{\rm unf}N}\ \sum_{\Omega_{\rm qub}:|\Omega_{\rm noqub}|=m} P_{\mathcal{S}}^{\Omega_{\rm qub}} P_{\Omega_{\rm qub}}. \quad (107)$$

Let

$$|\Psi_{\mathbf{tu}}^{\Omega_{\rm qub}}\rangle_A = \bigotimes_{k\in\Omega_{\rm qub}} |\phi_{t_k u_k}^k\rangle_{A_k} \bigotimes_{k\in\Omega_{\rm noqub}} |\Phi_{t_k u_k}^k\rangle_{A_k} \quad (108)$$

be the quantum state that Alice's agent $\mathcal{A}$ receives from Bob's agent $\mathcal{B}$, where the global quantum system received by $\mathcal{A}$ is $A = A_1 \cdots A_N$, and where $\mathbf{t} = (t_1, \ldots, t_N)$ and $\mathbf{u} = (u_1, \ldots, u_N)$. We recall that $\Omega_{\rm noqub} = [N] \setminus \Omega_{\rm qub}$. For a given set $\Omega_{\rm qub} \subseteq [N]$, let $\underline{\mathbf{x}}$ and $\overline{\mathbf{x}}$ denote the substrings of the $N-$bit string $\mathbf{x}$ with bit entries $k \in \Omega_{\rm qub}$ and $k \in \Omega_{\rm noqub}$, respectively, for $\mathbf{x} \in \{\mathbf{u}, \mathbf{t}\}$. Thus, from (108), we can write

$$|\Psi_{\mathbf{tu}}^{\Omega_{\rm qub}}\rangle_A = |\phi_{\underline{\mathbf{tu}}}\rangle_{\underline{A}} \otimes |\Phi_{\overline{\mathbf{tu}}}\rangle_{\overline{A}}, \quad (109)$$

where

$$|\phi_{\underline{\mathbf{tu}}}\rangle_{\underline{A}} = \bigotimes_{k\in\Omega_{\rm qub}} |\phi_{t_k u_k}^k\rangle_{A_k},$$
$$|\Phi_{\overline{\mathbf{tu}}}\rangle_{\overline{A}} = \bigotimes_{k\in\Omega_{\rm noqub}} |\Phi_{t_k u_k}^k\rangle_{A_k}, \quad (110)$$

and where $\underline{A} = \bigotimes_{k\in\Omega_{\rm qub}} A_k$ and $\overline{A} = \bigotimes_{k\in\Omega_{\rm noqub}} A_k$. Let $P_{\mathbf{tu}}$ be the probability distribution for the variables $(\mathbf{t}, \mathbf{u}) \in \{0,1\}^N \times \{0,1\}^N$. By the statement of the theorem, we have

$$P_{\mathbf{tu}} = \prod_{k=1}^{N} P_{\rm PS}^k(t_k) P_{\rm PB}^k(u_k), \quad (111)$$

where $\left\{P_{\rm PB}^k(0), P_{\rm PB}^k(1)\right\}$ and $\left\{P_{\rm PS}^k(0), P_{\rm PS}^k(1)\right\}$ are binary probability distributions, for $k \in [N]$. Thus, for any sets $F_0 \subseteq [N]$ and $F_1 = [N] \backslash F_0$ with $\mathbf{u}_0$ and $\mathbf{u}_1$ being substrings of $\mathbf{u}$, and with $\mathbf{t}_0$ and $\mathbf{t}_1$ being substrings of $\mathbf{t}$, with bit entries with labels from the sets F_0 and F_1, respectively, we use the notation $P_{\mathbf{tu}} = P_{\mathbf{t}_0\mathbf{t}_1\mathbf{u}_0\mathbf{u}_1} = P_{\mathbf{t}_0\mathbf{t}_1} P_{\mathbf{u}_0\mathbf{u}_1}$. Thus, we can express $P_{\mathcal{S}}^{\Omega_{\rm qub}}$ by

$$P_{\mathcal{S}}^{\Omega_{\rm qub}} = \sum_{\overline{\mathbf{t}},\overline{\mathbf{u}}} P_{\overline{\mathbf{tu}}} P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{tu}}}, \quad (112)$$

where $P_{\overline{\mathbf{tu}}}$ is the probability distribution for the variables $\overline{\mathbf{t}}, \overline{\mathbf{u}}$; and where $P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{tu}}}$ is the probability that Alice succeeds in giving Bob valid tokens at the presentation points Q_0 and Q_1 by following the strategy $\mathcal{S}$, given a set $\Omega_{\rm qub}$ and given variables $\overline{\mathbf{t}}$ and $\overline{\mathbf{u}}$.

We show below that there exist parameters satisfying the constraints (99) and satisfying

$$f(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det}) > 0, \quad (113)$$

where $f(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det})$ is given by (100). We show below that, for the parameters satisfying (99) and (113), it holds that

$$P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{tu}}} \leq e^{-Nf(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det})}, \quad (114)$$

for any finite dimensional quantum state $|\Phi_{\overline{\mathbf{tu}}}\rangle_{\overline{A}}$ and for any set $\Omega_{\rm qub}$ satisfying $m = |\Omega_{\rm noqub}| \leq \nu_{\rm unf}N$. It follows from (107) and from (112) – (114) that Alice's probability to succeed in her cheating strategy is not greater than $\epsilon_{\rm unf}$, with $\epsilon_{\rm unf}$ given by (102). Thus, $\mathcal{QT}_1$ and $\mathcal{QT}_2$ are $\epsilon_{\rm unf}-$unforgeable, with $\epsilon_{\rm unf}$ given by (102), as claimed.

We show that there exist parameters satisfying the constraints (99) for which (113) holds. Consider parameters satisfying (99) and the following constraint:

$$0 < \beta_{\rm PS} < \frac{1}{2}\left[e^{\frac{(\gamma_{\rm det}-\nu_{\rm unf})\lambda(\theta,\beta_{\rm PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{\rm PB})}\right)^2} - 1\right], \quad (115)$$

which is equivalent to

$$0 < \ln\left(1+2\beta_{\rm PS}\right) < \frac{(\gamma_{\rm det}-\nu_{\rm unf})\lambda(\theta,\beta_{\rm PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{\rm PB})}\right)^2. \quad (116)$$

From (99), we have

$$0 < \gamma_{\rm det} - \nu_{\rm unf} < 1. \quad (117)$$

Thus, from (115) and (117), we have

$$0 < \beta_{\rm PS} < \frac{1}{2}\left[e^{\frac{\lambda(\theta,\beta_{\rm PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{\rm PB})}\right)^2} - 1\right], \quad (118)$$

as required by (99). Then, since we have $0 < \theta < \frac{\pi}{4}$, $0 < \beta_{\rm PB} < \frac{1}{2}$ and $0 < \beta_{\rm PS} < \frac{1}{2}$, we obtain from the definition of $h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)$ in (101) that

$$0 < h(\beta_{\rm PS}, \beta_{\rm PB}, \theta) < 2\beta_{\rm PS}. \quad (119)$$

From (119), we have

$$0 < \ln\left[1 + h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)\right] < \ln\left(1 + 2\beta_{\rm PS}\right). \quad (120)$$

From (100), we have

$$\begin{aligned} &f(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det}) \\ &= \frac{(\gamma_{\rm det}-\nu_{\rm unf})\lambda(\theta,\beta_{\rm PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{\rm PB})}\right)^2 - \ln(1+2\beta_{\rm PS}). \\ &\quad + \left(1-(\gamma_{\rm det}-\nu_{\rm unf})\right)\left[\ln\left(1+2\beta_{\rm PB}\right) - \ln\left[1+h(\beta_{\rm PS},\beta_{\rm PB},\theta)\right]\right] \\ &> \frac{(\gamma_{\rm det}-\nu_{\rm unf})\lambda(\theta,\beta_{\rm PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{\rm PB})}\right)^2 - \ln(1+2\beta_{\rm PS}). \\ &> 0, \end{aligned} \quad (121)$$

where in the second line we used (117) and (120), and in the third line we used (116).

A general cheating strategy $\mathcal{S}$ by Alice is as follows. Alice receives the quantum system A from Bob in the quantum state $|\Psi_{\mathbf{tu}}^{\Omega_{\rm qub}}\rangle_A$ given by (109), she adds an ancillary system E of arbitrary finite Hilbert space dimension in a quantum state $|\chi\rangle_E$ and applies an arbitrary projective measurement $\{\Pi_{x\mathbf{ab}}\}_{x,\mathbf{a},\mathbf{b}}$ on AE, which may depend on $\Omega_{\rm qub}$, and obtains an outcome $(x,\mathbf{a},\mathbf{b})$, where $x=(\Lambda,g,\mathbf{d},c,\zeta)$, with $\Lambda,g,\mathbf{d}$ and c comprising the information that Alice must send Bob before token presentation satisfying the required constraints, $\mathbf{a},\mathbf{b}\in\{0,1\}^n$ are token strings to present at the respective presentation points Q_0 and Q_1, and ζ is any other classical variable obtained by Alice's measurement. This means that $\Lambda\subseteq[N]$ with $|\Lambda|=n$, for some integer $n\geq\gamma_{\rm det}N$ and for a predetermined $\gamma_{\rm det}\in(0,1)$, g is a one-to-one function of the form $g(k)=j$ for $k\in\Lambda$ and $j\in[n]$, $\mathbf{d}\in\{0,1\}^n$ and $c\in\{0,1\}$. In the intersection of the causal pasts of Q_0 and Q_1, Alice sends Λ,g,c and $\mathbf{d}$ to Bob, as required by the task. Bob does not abort as he receives the set Λ satisfying the required condition, as well as the $g,\mathbf{d}$ and c of the required form. Alice sends the tokens to her respective agents who present them at the corresponding presentation points.

Below, we show the bound (114) on the probability $P_{\mathcal{S}}^{\Omega_{\rm qub}\bar{\mathbf{t}}\bar{\mathbf{u}}}$ that Alice succeeds in giving Bob valid tokens at the presentation points Q_0 and Q_1 by following the strategy $\mathcal{S}$, given a set $\Omega_{\rm qub}$ and given variables $\bar{\mathbf{t}}$ and $\bar{\mathbf{u}}$. Thus, we consider that Alice gives tokens $\mathbf{a}\in\{0,1\}^n$ at Q_0 and $\mathbf{b}\in\{0,1\}^n$ at Q_1.

C. Notation and useful relations

We recall notation. Using the set Λ and the function $g:\Lambda\to[n]$, we have the following relations between $\mathbf{t},\mathbf{u}\in\{0,1\}^N$ and $\mathbf{r},\mathbf{s}\in\{0,1\}^n$. We have $r_j=t_k$, and $s_j=u_k$, where $j=g(k)$, for $j\in[n]$ and $k\in\Lambda$. We define $\mathbf{r}=(r_1,\ldots,r_n)$ and $\mathbf{s}=(s_1,\ldots,s_n)$. In the token scheme $\mathcal{QT}_1$, we have defined

$$\Delta_i=\{j\in[n]|\tilde{d}_{i,j}=s_j\},\tag{122}$$

and $\mathbf{a}_i$ as the restriction of a string $\mathbf{a}$ to entries a_j with $j\in\Delta_i$, for $i\in\{0,1\}$. These variables are defined similarly in the token scheme $\mathcal{QT}_2$ by simply setting $\mathbf{d}$ as a string whose bit entries are only zero.

By definition of Λ, we have

$$\Lambda\subseteq[N],\qquad|\Lambda|=n\leq N.\tag{123}$$

By definition of $\Omega_{\rm qub}$ and $\Omega_{\rm noqub}$, we have

$$\Omega_{\rm qub}\cap\Omega_{\rm noqub}=\emptyset,\qquad\Omega_{\rm qub}\cup\Omega_{\rm noqub}=[N].\tag{124}$$

We note that the sets Δ_i depend on $x=(\Lambda,g,\mathbf{d},c,\zeta)$, for $i\in\{0,1\}$, but we do not write this dependence explicitly, in order to simplify the notation. Since $\tilde{d}_{1,j}=\tilde{d}_{0,j}\oplus 1$, for $j\in[n]$, we have

$$\Delta_0\cap\Delta_1=\emptyset,\qquad\Delta_0\cup\Delta_1=[n].\tag{125}$$

We define

$$\underline{\Lambda}=\Lambda\cap\Omega_{\rm qub},\qquad\overline{\Lambda}=\Lambda\cap\Omega_{\rm noqub}.\tag{126}$$

It follows that

$$\underline{\Lambda}\cap\overline{\Lambda}=\emptyset,\qquad\underline{\Lambda}\cup\overline{\Lambda}=\Lambda.\tag{127}$$

Similarly, we define

$$\begin{aligned}\underline{\Delta}&=\{j\in[n]|\ \exists k\in\underline{\Lambda}\text{ s. t. }g(k)=j\},\\ \overline{\Delta}&=\{j\in[n]|\ \exists k\in\overline{\Lambda}\text{ s. t. }g(k)=j\}.\end{aligned}\tag{128}$$

Since the function $g:\Lambda\to[n]$ is one-to-one, there is a one-to-one correspondence between the elements of $\underline{\Lambda}$ ($\overline{\Lambda}$) and the elements of $\underline{\Delta}$ ($\overline{\Delta}$). Thus, from (127) and (128), we have

$$\underline{\Delta}\cap\overline{\Delta}=\emptyset,\qquad\underline{\Delta}\cup\overline{\Delta}=[n].\tag{129}$$

We define

$$\underline{\Delta}_i=\Delta_i\cap\underline{\Delta},\tag{130}$$

for $i\in\{0,1\}$. It follows that $\underline{\Delta}_i\subseteq\Delta_i$, for $i\in\{0,1\}$. From the definitions of $\underline{\Delta}_0$, $\underline{\Delta}_1$ and $\underline{\Delta}$, we have

$$\underline{\Delta}_0\cap\underline{\Delta}_1=\emptyset,\qquad\underline{\Delta}_0\cup\underline{\Delta}_1=\underline{\Delta}.\tag{131}$$

We define

$$\begin{aligned}\Lambda_i&=\{k\in\Lambda|g(k)\in\Delta_i\},\\ \underline{\Lambda}_i&=\{k\in\Lambda|g(k)\in\underline{\Delta}_i\},\end{aligned}\tag{132}$$

for $i\in\{0,1\}$. Since the function $g:\Lambda\to[n]$ is one-to-one, we have from (125), (128), (131) and (132) that

$$\Lambda_0\cap\Lambda_1=\emptyset,\qquad\Lambda_0\cup\Lambda_1=\Lambda,\tag{133}$$

$$\underline{\Lambda}_0\cap\underline{\Lambda}_1=\emptyset,\qquad\underline{\Lambda}_0\cup\underline{\Lambda}_1=\underline{\Lambda}.\tag{134}$$

We define $\underline{\mathbf{e}}$, $\underline{\overline{\mathbf{e}}}$, $\underline{\mathbf{e}}_0$ and $\underline{\mathbf{e}}_1$ to be the restrictions of the string $\mathbf{e}\in\{0,1\}^n$ to the bit entries e_j with labels $j\in\underline{\Delta}$, $j\in\overline{\Delta}$, $j\in\underline{\Delta}_0$ and $j\in\underline{\Delta}_1$, respectively, for $\mathbf{e}\in\{\mathbf{a},\mathbf{b},\mathbf{r},\mathbf{s}\}$. We have chosen the notation $\underline{\mathbf{e}}$ and $\underline{\overline{\mathbf{e}}}$ instead of the more obvious $\underline{\mathbf{e}}$ and $\overline{\mathbf{e}}$ for consistency with the notation chosen below, which simplifies the notation in various equations that follow. From (130), we have $\underline{\Delta}_i\subseteq\Delta_i$, for $i\in\{0,1\}$. Thus, and since $\mathbf{e}_0$ and $\mathbf{e}_1$ are restrictions of the string $\mathbf{e}$ to the bit entries e_j with labels $j\in\Delta_0$ and $j\in\Delta_1$, respectively, we have that $\underline{\mathbf{e}}_0$ and $\underline{\mathbf{e}}_1$ are substrings of the strings $\mathbf{e}_0$ and $\mathbf{e}_1$, respectively, for $\mathbf{e}\in\{\mathbf{a},\mathbf{b},\mathbf{r},\mathbf{s}\}$. From (129) and (131), we have

$$\mathbf{e}=(\underline{\mathbf{e}},\underline{\overline{\mathbf{e}}})=(\underline{\mathbf{e}}_0,\underline{\mathbf{e}}_1,\underline{\overline{\mathbf{e}}}),\qquad\underline{\mathbf{e}}=(\underline{\mathbf{e}}_0,\underline{\mathbf{e}}_1),\tag{135}$$

where $\underline{\mathbf{e}}\in\{0,1\}^{\underline{\Delta}}$, $\underline{\overline{\mathbf{e}}}\in\{0,1\}^{\overline{\Delta}}$, $\underline{\mathbf{e}}_0\in\{0,1\}^{\underline{\Delta}_0}$ and $\underline{\mathbf{e}}_1\in\{0,1\}^{\underline{\Delta}_1}$, for $\mathbf{e}\in\{\mathbf{a},\mathbf{b},\mathbf{r},\mathbf{s}\}$.

We define $\underline{\mathbf{e}}_i$ to be the restrictions of the string $\mathbf{e}\in\{0,1\}^N$ to the bit entries e_k with labels $k\in\underline{\Lambda}_i$, for $i\in\{0,1\}$ and $\mathbf{e}\in\{\mathbf{u},\mathbf{t}\}$. Since the function $g:\Lambda\to[n]$ is one to one, there is a one to one correspondence between

$\underline{\mathbf{t}}_i$ and $\underline{\mathbf{r}}_i$, and between $\underline{\mathbf{u}}_i$ and $\underline{\mathbf{s}}_i$, for $i \in \{0,1\}$. We define the string $\underline{\underline{\mathbf{e}}}$ to be the restriction of the string $\mathbf{e} \in \{0,1\}^N$ to the bit entries e_k with labels $k \in \underline{\underline{\Lambda}}$, for $\mathbf{e} \in \{\mathbf{u}, \mathbf{t}\}$. We define $\underline{\mathbf{e}}$ to be the restriction of $\mathbf{e} \in \{0,1\}^N$ to the bit entries e_k with labels $k \in \Omega_{\rm qub}$, for $\mathbf{e} \in \{\mathbf{u}, \mathbf{t}\}$. Thus, from (126), $\underline{\underline{\mathbf{e}}}$ is a sub-string of $\underline{\mathbf{e}}$, for $\mathbf{e} \in \{\mathbf{u}, \mathbf{t}\}$. From (134), we can write

$$\underline{\underline{\mathbf{e}}} = (\underline{\mathbf{e}}_0, \underline{\mathbf{e}}_1), \tag{136}$$

where $\underline{\underline{\mathbf{e}}} \in \{0,1\}^{\underline{\underline{\Lambda}}}$, $\underline{\mathbf{e}}_0 \in \{0,1\}^{\underline{\Lambda}_0}$ and $\underline{\mathbf{e}}_1 \in \{0,1\}^{\underline{\Lambda}_1}$, for $\mathbf{e} \in \{\mathbf{t}, \mathbf{u}\}$. We define $\underline{\mathbf{e}}'$ as the restriction of the string $\mathbf{e} \in \{0,1\}^N$ to the bit entries e_k with labels $k \in \Omega_{\rm qub} \setminus \underline{\Lambda}$, for $\mathbf{e} \in \{\mathbf{u}, \mathbf{t}\}$. It follows that we can write

$$\underline{\mathbf{e}} = (\underline{\underline{\mathbf{e}}}, \underline{\mathbf{e}}') = (\underline{\mathbf{e}}_0, \underline{\mathbf{e}}_1, \underline{\mathbf{e}}'), \tag{137}$$

where $\underline{\mathbf{e}} \in \{0,1\}^{\Omega_{\rm qub}}$, $\underline{\underline{\mathbf{e}}} \in \{0,1\}^{\underline{\Lambda}}$, $\underline{\mathbf{e}}' \in \{0,1\}^{\Omega_{\rm qub} \setminus \underline{\Lambda}}$, $\underline{\mathbf{e}}_0 \in \{0,1\}^{\underline{\Lambda}_0}$ and $\underline{\mathbf{e}}_1 \in \{0,1\}^{\underline{\Lambda}_1}$, for $\mathbf{e} \in \{\mathbf{t}, \mathbf{u}\}$.

As mentioned above, given our notation, there is a one-to-one correspondence between $\underline{\mathbf{t}}_i$ and $\underline{\mathbf{r}}_i$, and between $\underline{\mathbf{u}}_i$ and $\underline{\mathbf{s}}_i$, for $i \in \{0,1\}$. Similarly, there is a one-to-one correspondence between $\underline{\underline{\mathbf{t}}}$ and $\underline{\underline{\mathbf{r}}}$, and between $\underline{\underline{\mathbf{u}}}$ and $\underline{\underline{\mathbf{s}}}$. We express these, and previously mentioned, one-to-one correspondences as follows

$$\begin{aligned}
\Lambda &\leftrightarrow [n],\\
\underline{\Lambda} &\leftrightarrow \underline{\Delta},\\
\overline{\Lambda} &\leftrightarrow \overline{\Delta},\\
\Lambda_i &\leftrightarrow \Delta_i,\\
\underline{\Lambda}_i &\leftrightarrow \underline{\Delta}_i,\\
\underline{\underline{\mathbf{u}}} &\leftrightarrow \underline{\underline{\mathbf{s}}},\\
\underline{\underline{\mathbf{t}}} &\leftrightarrow \underline{\underline{\mathbf{r}}},\\
\underline{\mathbf{u}}_i &\leftrightarrow \underline{\mathbf{s}}_i,\\
\underline{\mathbf{t}}_i &\leftrightarrow \underline{\mathbf{r}}_i,
\end{aligned} \tag{138}$$

for $i \in \{0,1\}$.

In the rest of this proof we assume

$$m = |\Omega_{\rm noqub}| \leq \nu_{\rm unf} N. \tag{139}$$

We consider the $n-$bit strings $\tilde{\mathbf{d}}_i = (\tilde{d}_{i,1}, \ldots, \tilde{d}_{i,n})$ for $i \in \{0,1\}$. By definition of the token scheme, we have $\tilde{d}_{i,j} = d_j \oplus i \oplus c$, for $j \in [n]$ and $i \in \{0,1\}$. Thus, we have $\tilde{d}_{1,j} = \tilde{d}_{0,j} \oplus 1$, for $j \in [n]$, from which follows that the $n-$bit strings $\tilde{\mathbf{d}}_0$ and $\tilde{\mathbf{d}}_1$ are the complement of each other. We have

$$\begin{aligned}
|\underline{\Delta}_0 \cup \underline{\Delta}_1| &= |\underline{\underline{\Delta}}|\\
&= |\underline{\underline{\Lambda}}|\\
&\geq |\Lambda| - |\Omega_{\rm noqub}|\\
&\geq n - \nu_{\rm unf} N,
\end{aligned} \tag{140}$$

where in the first line we used (131); in the second line we used (128) and the fact that $g : \Lambda \to [n]$ is a one-to-one function; in the third line we used (124) and (126); and in the last line we used (123) and (139).

For a given possible outcome x, we define the sets

$$\begin{aligned}
\Gamma^x_{\mathbf{tu}} &= \{(\mathbf{a}, \mathbf{b}) \in \{0,1\}^n \times \{0,1\}^n | d(\mathbf{a}_0, \mathbf{r}_0)\\
&\qquad \leq |\Delta_0|\gamma_{\rm err}, d(\mathbf{b}_1, \mathbf{r}_1) \leq |\Delta_1|\gamma_{\rm err}\}, \quad (141)\\
\underline{\Gamma}^x_{\mathbf{tu}} &= \{(\mathbf{a}, \mathbf{b}) \in \{0,1\}^n \times \{0,1\}^n | d(\underline{\mathbf{a}}_0, \underline{\mathbf{r}}_0)\\
&\qquad \leq |\Delta_0|\gamma_{\rm err}, d(\underline{\mathbf{b}}_1, \underline{\mathbf{r}}_1) \leq |\Delta_1|\gamma_{\rm err}\}, \quad (142)\\
\tilde{\underline{\Gamma}}^x_{\mathbf{tu}} &= \{(\mathbf{a}, \mathbf{b}) \in \{0,1\}^n \times \{0,1\}^n | d(\underline{\mathbf{a}}_0, \underline{\mathbf{r}}_0)\\
&\qquad + d(\underline{\mathbf{b}}_1, \underline{\mathbf{r}}_1) \leq \big(|\Delta_0| + |\Delta_1|\big)\gamma_{\rm err}\}, \quad (143)\\
\overline{\underline{\Gamma}}^x_{\mathbf{tu}} &= \{(\mathbf{a}, \mathbf{b}) \in \{0,1\}^n \times \{0,1\}^n | d(\underline{\mathbf{a}}_0, \underline{\mathbf{r}}_0)\\
&\qquad + d(\underline{\mathbf{b}}_1, \underline{\mathbf{r}}_1) \leq \big(|\underline{\Delta}_0| + |\underline{\Delta}_1|\big)\delta\}, \quad (144)\\
\xi^x_{\mathbf{ab}\underline{\mathbf{u}}} &= \{\underline{\mathbf{t}} \in \{0,1\}^{\Omega_{\rm qub}} | d(\underline{\mathbf{a}}_0, \underline{\mathbf{r}}_0) + d(\underline{\mathbf{b}}_1, \underline{\mathbf{r}}_1)\\
&\qquad \leq \big(|\underline{\Delta}_0| + |\underline{\Delta}_1|\big)\delta\}, \quad (145)
\end{aligned}$$

where

$$\tilde{\delta} = \frac{\big(|\Delta_0| + |\Delta_1|\big)\gamma_{\rm err}}{|\underline{\Delta}_0| + |\underline{\Delta}_1|}, \tag{146}$$

and where δ is given by (101). As this is useful below, we have clarified with the chosen notation that $\xi^x_{\mathbf{ab}\underline{\mathbf{u}}}$ does not depend on $\mathbf{t}\overline{\mathbf{u}}$. We show below that

$$0 < \tilde{\delta} \leq \delta < \lambda(\theta, \beta_{\rm PB}). \tag{147}$$

It follows straightforwardly that

$$\Gamma^x_{\mathbf{tu}} \subseteq \underline{\Gamma}^x_{\mathbf{tu}} \subseteq \tilde{\underline{\Gamma}}^x_{\mathbf{tu}} \subseteq \overline{\underline{\Gamma}}^x_{\mathbf{tu}}. \tag{148}$$

We show (147). From the condition $n \geq \gamma_{\rm det} N$ for Bob not aborting and from (99), we have

$$n - \nu_{\rm unf} N \geq (\gamma_{\rm det} - \nu_{\rm unf})N > 0. \tag{149}$$

Thus, from $\gamma_{\rm err} > 0$, (125), (131), (140), (146) and (149), we have $\tilde{\delta} > 0$. From (125), (131), (140) and (146), we have

$$\begin{aligned}
\tilde{\delta} &\leq \frac{n\gamma_{\rm err}}{n - \nu_{\rm unf} N}\\
&= \gamma_{\rm err}\Big(1 + \frac{\nu_{\rm unf} N}{n - \nu_{\rm unf} N}\Big)\\
&\leq \gamma_{\rm err}\Big(1 + \frac{\nu_{\rm unf}}{\gamma_{\rm det} - \nu_{\rm unf}}\Big)\\
&< \lambda(\theta, \beta_{\rm PB}),
\end{aligned} \tag{150}$$

where in the third line we used the condition $n \geq \gamma_{\rm det} N$ for Bob not aborting, and in the last line we used (99). Since $\delta = \frac{\gamma_{\rm err}\gamma_{\rm det}}{\gamma_{\rm det} - \nu_{\rm unf}}$, as defined by (101), (147) follows from (150).

The probability that Alice obtains outcomes $x, \mathbf{a}$ and $\mathbf{b}$ following her strategy $\mathcal{S}$ for given values of $\Omega_{\rm qub}$, $\mathbf{u}$ and $\mathbf{t}$ is given by

$$P^{\Omega_{\rm qub}\mathbf{t}\overline{\mathbf{u}}}_{\mathcal{S}}[x\mathbf{ab}|\underline{\mathbf{t}}\underline{\mathbf{u}}] = {\rm Tr}\Big[\Big(\big(\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}\big)_{\underline{A}} \otimes \big(\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\big)_{\overline{A}E}\Big)\Pi_{x\mathbf{ab}}\Big], \tag{151}$$

where

$$\begin{aligned}
\big(\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}\big)_{\underline{A}} &= \big(|\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}\rangle\langle\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}|\big)_{\underline{A}},\\
\big(\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\big)_{\overline{A}E} &= \big(|\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\rangle\langle\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}|\big)_{\overline{A}} \otimes \big(|\chi\rangle\langle\chi|\big)_E,
\end{aligned} \tag{152}$$

and where $|\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}\rangle_{\underline{A}}$ and $|\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\rangle_{\overline{A}}$ are defined by (110). Thus, Alice's success probability $P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}}$ satisfies

$$\begin{aligned}
P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}} &= \sum_{\underline{\mathbf{t}},\underline{\mathbf{u}},x} \sum_{(\mathbf{a},\mathbf{b})\in\Gamma^x_{\underline{\mathbf{t}}\underline{\mathbf{u}}}} P_{\underline{\mathbf{t}}\underline{\mathbf{u}}} P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}}[x\mathbf{a}\mathbf{b}|\underline{\mathbf{t}}\underline{\mathbf{u}}] \\
&\leq \sum_{\underline{\mathbf{t}},\underline{\mathbf{u}},x} \sum_{(\mathbf{a},\mathbf{b})\in\overline{\Gamma}^x_{\underline{\mathbf{t}}\underline{\mathbf{u}}}} P_{\underline{\mathbf{t}}\underline{\mathbf{u}}} P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}}[x\mathbf{a}\mathbf{b}|\underline{\mathbf{t}}\underline{\mathbf{u}}] \\
&= \sum_{\underline{\mathbf{u}},x} \sum_{\substack{\underline{\mathbf{t}}_0,\underline{\mathbf{t}}_1,\underline{\mathbf{t}}' \\ \underline{\mathbf{b}}_0,\underline{\mathbf{a}}_1,\overline{\mathbf{a}},\overline{\mathbf{b}}}} \sum_{(\underline{\mathbf{a}}_0,\underline{\mathbf{b}}_1):C} P_{\underline{\mathbf{t}}\underline{\mathbf{u}}} P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}}[x\mathbf{a}\mathbf{b}|\underline{\mathbf{t}}\underline{\mathbf{u}}] \\
&= \sum_{\underline{\mathbf{u}},x} \sum_{\substack{\underline{\mathbf{a}}_0,\underline{\mathbf{b}}_1,\underline{\mathbf{t}}' \\ \underline{\mathbf{b}}_0,\underline{\mathbf{a}}_1,\overline{\mathbf{a}},\overline{\mathbf{b}}}} \sum_{(\underline{\mathbf{t}}_0,\underline{\mathbf{t}}_1):C} P_{\underline{\mathbf{t}}\underline{\mathbf{u}}} P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}}[x\mathbf{a}\mathbf{b}|\underline{\mathbf{t}}\underline{\mathbf{u}}] \\
&= \sum_{\mathbf{a},\mathbf{b},x,\underline{\mathbf{u}}} \sum_{\underline{\mathbf{t}}\in\xi^x_{\mathbf{a}\mathbf{b}\underline{\mathbf{u}}}} P_{\underline{\mathbf{t}}\underline{\mathbf{u}}} P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}}[x\mathbf{a}\mathbf{b}|\underline{\mathbf{t}}\underline{\mathbf{u}}], \qquad (153)
\end{aligned}$$

where C denotes the constraint

$$d(\underline{\mathbf{a}}_0,\underline{\mathbf{r}}_0)+d(\underline{\mathbf{b}}_1,\underline{\mathbf{r}}_1)\leq(|\underline{\Delta}_0|+|\underline{\Delta}_1|)\delta; \qquad (154)$$

where in the first line we used (141) and (151); in the second line we used (148); in the third line we used (135), (137), and (144); in the fourth line we used that $\underline{\mathbf{t}}_i$ is in one to one correspondence with $\underline{\mathbf{r}}_i$, for $i\in\{0,1\}$; and in the last line we used (135), (137) and (145).

D. Entanglement-based version

We use an entanglement-based version of the task to re-write the last line of (153). For $k\in\Omega_{\rm qub}$, Bob first prepares a pair of qubits B_kA_k in the Bell state $|\Phi^+\rangle_{B_kA_k}=\frac{1}{\sqrt{2}}\big(|0\rangle|0\rangle+|1\rangle|1\rangle\big)_{B_kA_k}$, sends the qubit A_k to Alice, chooses $u_k\in\{0,1\}$ with probability $P^k_{\rm PB}(u_k)$ and then measures the qubit B_k in the the basis $\mathcal{D}^k_{u_k}=\{|\phi^k_{tu_k}\rangle\}^1_{t=0}$, obtaining the outcome $|\phi^k_{t_ku_k}\rangle$ randomly, with Alice's qubit A_k projecting into the same state, for $t_k\in\{0,1\}$. We note that in order to deal with the fact that the probability $P_{\underline{\mathbf{t}}}$ does not necessarily correspond to the random distribution, for $\underline{\mathbf{t}}\in\{0,1\}^{\Omega_{\rm qub}}$, we will need to introduce a factor of $P_{\underline{\mathbf{t}}}2^{|\Omega_{\rm qub}|}$. For $k\in\Omega_{\rm noqub}$, Bob generates the bits t_k and u_k in such a way that the strings $\overline{\mathbf{t}}$ and $\overline{\mathbf{u}}$ are generated with the probability distribution $P_{\overline{\mathbf{t}}\overline{\mathbf{u}}}$. Given the obtained values for $\overline{\mathbf{t}}$ and $\overline{\mathbf{u}}$, Bob prepares a finite dimensional quantum state $|\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\rangle_{\overline{A}}$ and sends it to Alice. Alice introduces an ancillary quantum system E of arbitrary finite Hilbert space dimension in a pure state $|\chi\rangle_E$ and then applies a projective measurement on AE, with projector operators $\Pi_{x\mathbf{a}\mathbf{b}}$, where the possible measurement outcomes $x=(\Lambda,g,\mathbf{d},c,\zeta)$ and $\mathbf{a},\mathbf{b}\in\{0,1\}^n$, run over the set of values satisfying the constraints, and where $A=A_1\cdots A_N=\underline{A}\overline{A}$. We define the quantum state

$$\rho_{\overline{\mathbf{t}}\overline{\mathbf{u}}}=\big(\Phi^+\big)_{\underline{B}\underline{A}}\otimes\big(\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\big)_{\overline{A}E}, \qquad (155)$$

where $\underline{B}$ denotes the system held by Bob, $\big(\Phi^+\big)_{\underline{B}\underline{A}}=\bigotimes_{k\in\Omega_{\rm qub}}\big(|\Phi^+\rangle\langle\Phi^+|\big)_{B_kA_k}$, and where the state $\big(\Phi_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\big)_{\overline{A}E}$is defined by (152). We define the positive semi definite (and Hermitian) operators

$$D_{x\mathbf{a}\mathbf{b}}=\sum_{\underline{\mathbf{u}}}P_{\underline{\mathbf{u}}}\sum_{\underline{\mathbf{t}}\in\xi^x_{\mathbf{a}\mathbf{b}\underline{\mathbf{u}}}}P_{\underline{\mathbf{t}}}2^{|\Omega_{\rm qub}|}\big(\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}\big)_{\underline{B}}, \qquad (156)$$

and

$$\tilde{P}=\sum_{x,\mathbf{a},\mathbf{b}}\big(D_{x\mathbf{a}\mathbf{b}}\big)_{\underline{B}}\otimes\big(\Pi_{x\mathbf{a}\mathbf{b}}\big)_{AE}, \qquad (157)$$

where $\big(\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}\big)_{\underline{B}}$ is given by (152), replacing $\underline{A}$ by $\underline{B}$, i.e. $\big(\phi_{\underline{\mathbf{t}}\underline{\mathbf{u}}}\big)_{\underline{B}}=\bigotimes_{k\in\Omega_{\rm qub}}\big(|\phi^k_{t_ku_k}\rangle\langle\phi^k_{t_ku_k}|\big)_{B_k}$; and where $\underline{\mathbf{u}}$ runs over $\{0,1\}^{\Omega_{\rm qub}}$, x runs over its set of possible values, and $\mathbf{a}$ and $\mathbf{b}$ run over $\{0,1\}^n$. It follows straightforwardly from (151) – (157), and from $P_{\underline{\mathbf{t}}\underline{\mathbf{u}}}=P_{\underline{\mathbf{t}}}P_{\underline{\mathbf{u}}}$, which follows from (111), that

$$\begin{aligned}
P_{\mathcal{S}}^{\Omega_{\rm qub}\overline{\mathbf{t}}\overline{\mathbf{u}}} &\leq \mathrm{Tr}\big(\tilde{P}\rho_{\overline{\mathbf{t}}\overline{\mathbf{u}}}\big) \\
&\leq \|\tilde{P}\| \\
&= \max_{x,\mathbf{a},\mathbf{b}}\big\|D_{x\mathbf{a}\mathbf{b}}\big\|, \qquad (158)
\end{aligned}$$

where in the second line we used Proposition 2; and where in the third line we used (157) and Proposition 3, since $\{\Pi_{x\mathbf{a}\mathbf{b}}\}_{x,\mathbf{a},\mathbf{b}}$ is a projective measurement acting on a finite dimensional Hilbert space, and $\{D_{x\mathbf{a}\mathbf{b}}\}_{x,\mathbf{a},\mathbf{b}}$ is a finite set of positive semi definite operators acting on a finite dimensional Hilbert space. We show below that

$$\max_{x,\mathbf{a},\mathbf{b}}\|D_{x\mathbf{a}\mathbf{b}}\|\leq e^{-Nf(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det})}, \qquad (159)$$

with $f(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det})$ given by (100). Thus, the claimed bound (114) follows from (158) and (159).

E. Using Lemma 6 to prove (159)

We compute an upper bound on $\max_{x,\mathbf{a},\mathbf{b}}\big\|D_{x\mathbf{a}\mathbf{b}}\big\|$. First, we define the set

$$\begin{aligned}
\underline{\xi}^x_{\mathbf{a}\mathbf{b}\underline{\mathbf{u}}_0\underline{\mathbf{u}}_1} = \big\{&(\underline{\mathbf{t}}_0,\underline{\mathbf{t}}_1)\in\{0,1\}^{\underline{\Lambda}_0}\times\{0,1\}^{\underline{\Lambda}_1}|d(\underline{\mathbf{a}}_0,\underline{\mathbf{r}}_0) \\
&+d(\underline{\mathbf{b}}_1,\underline{\mathbf{r}}_1)\leq\big(|\underline{\Delta}_0|+|\underline{\Delta}_1|\big)\delta\big\}. \qquad (160)
\end{aligned}$$

As explicitly stated by the notation, we note that the dependence of $\underline{\xi}^x_{\mathbf{a}\mathbf{b}\underline{\mathbf{u}}_0\underline{\mathbf{u}}_1}$ on $\mathbf{u}$ is only via the sub-strings $\underline{\mathbf{u}}_0$ and $\underline{\mathbf{u}}_1$. This follow in particular because the constant δ does not depend on $\mathbf{u}$, as follows from (101). Thus, from (137), (145), (156) and (160), we have

$$D_{x\mathbf{a}\mathbf{b}}=\tilde{D}_{x\mathbf{a}\mathbf{b}}\otimes\tilde{\phi}_x, \qquad (161)$$

where

$$\tilde{D}_{x\mathbf{a}\mathbf{b}}=\sum_{\underline{\mathbf{u}}_0,\underline{\mathbf{u}}_1}P_{\underline{\mathbf{u}}_0\underline{\mathbf{u}}_1}\sum_{(\underline{\mathbf{t}}_0,\underline{\mathbf{t}}_1)\in\underline{\xi}^x_{\mathbf{a}\mathbf{b}\underline{\mathbf{u}}_0\underline{\mathbf{u}}_1}}P_{\underline{\mathbf{t}}_0\underline{\mathbf{t}}_1}\big(\phi_{\underline{\mathbf{t}}_0\underline{\mathbf{t}}_1\underline{\mathbf{u}}_0\underline{\mathbf{u}}_1}\big)_{\underline{B}_0\underline{B}_1}, \qquad (162)$$

$$\tilde{\phi}_x=2^{|\Omega_{\rm qub}|}\sum_{\underline{\mathbf{t}}',\underline{\mathbf{u}}'}P_{\underline{\mathbf{u}}'\underline{\mathbf{t}}'}\big(\phi_{\underline{\mathbf{t}}'\underline{\mathbf{u}}'}\big)_{\underline{B}'}, \qquad (163)$$

and where $\underline{B} = \underline{B}_0 \underline{B}_1 \underline{B}'$. It follows from (161) that

$$\|D_{x\mathbf{ab}}\| = \|\tilde{D}_{x\mathbf{ab}}\| \, \|\tilde{\phi}_x\|. \tag{164}$$

We deduce an upper bound on $\max_{\mathbf{a},\mathbf{b}} \|\tilde{D}_{x\mathbf{ab}}\|$. For given values of $x, \mathbf{a}$ and $\mathbf{b}$, we define the operator

$$\underline{\tilde{D}}_{x\mathbf{ab}} = \sum_{\underline{\mathbf{u}}_0, \underline{\mathbf{u}}_1} P_{\underline{\mathbf{u}}_0 \underline{\mathbf{u}}_1} \sum_{(\underline{\mathbf{t}}_0, \underline{\mathbf{t}}_1) \in \xi^x_{\mathbf{a}\underline{\mathbf{u}}_0 \underline{\mathbf{u}}_1}} \big(\phi_{\underline{\mathbf{t}}_0 \underline{\mathbf{t}}_1 \underline{\mathbf{u}}_0 \underline{\mathbf{u}}_1}\big)_{\underline{B}_0 \underline{B}_1}. \tag{165}$$

For a given x, it follows from (147), (160) and (165), and from Lemma 6 that

$$\max_{\mathbf{a},\mathbf{b}} \|\underline{\tilde{D}}_{x\mathbf{ab}}\| \leq e^{-\big(|\underline{\Delta}_0| + |\underline{\Delta}_1|\big) \frac{\lambda(\theta, \beta_{\rm PB})}{2} \Big(1 - \frac{\delta}{\lambda(\theta, \beta_{\rm PB})}\Big)^2}, \tag{166}$$

where

$$\lambda(\theta, \beta_{\rm PB}) = \frac{1}{2}\Big(1 - \sqrt{1 - [1 - (O(\theta))^2](1 - 4\beta_{\rm PB}^2)}\Big). \tag{167}$$

To see this more clearly, recall the following facts. First, $\underline{\mathbf{e}}_i$ is a bit string with bit entries e_j with labels $j \in \underline{\Delta}_i$, for $i \in \{0,1\}$ and $\mathbf{e} \in \{\mathbf{a}, \mathbf{b}, \mathbf{r}, \mathbf{s}\}$. Second, $\underline{\mathbf{e}}_i$ is a bit string with bit entries e_k with labels $k \in \underline{\Lambda}_i$, for $i \in \{0,1\}$ and $\mathbf{e} \in \{\mathbf{u}, \mathbf{t}\}$. Third, there is a one-to-one correspondence between the sets $\underline{\Lambda}_i$ and $\underline{\Delta}_i$ via the one-to-one function g, i.e $k \in \underline{\Lambda}_i$ iff $g(k) = j \in \underline{\Delta}_i$, for $i \in \{0,1\}$. Fourth, there is a one-to-one correspondence between $\underline{\mathbf{u}}_i$ and $\underline{\mathbf{s}}_i$, and between $\underline{\mathbf{t}}_i$ and $\underline{\mathbf{r}}_i$, i.e. $u_k = s_j$ and $t_k = r_j$ for $j = g(k)$ and $k \in \underline{\Lambda}_i$, and for $i \in \{0,1\}$. Fifth, the sets $\underline{\Lambda}_0$ and $\underline{\Lambda}_1$ do not intersect, hence the sets $\underline{\Delta}_0$ and $\underline{\Delta}_1$ do not intersect either, which implies that $|\underline{\Lambda}_0 \cup \underline{\Lambda}_1| = |\underline{\Delta}_0 \cup \underline{\Delta}_1| = |\underline{\Delta}_0| + |\underline{\Delta}_1|$. Sixth, the bit entries $\tilde{d}_{0,j} = d_j \oplus c$ and $\tilde{d}_{1,j} = d_j \oplus 1 \oplus c$ of the respective strings $\tilde{\mathbf{d}}_0$ and $\tilde{\mathbf{d}}_1$ defined in the token scheme $\mathcal{QT}_1$ are different: $\tilde{d}_{0,j} = \tilde{d}_{1,j} \oplus 1$; which are also different in the token scheme $\mathcal{QT}_2$ by setting $d_j = 0$ for $j \in [n]$. Seventh, $\underline{\Delta}_i \subseteq \Delta_i$ for $i \in \{0,1\}$, where $\Delta_i = \{j \in [n] | \tilde{d}_{i,j} = s_j\}$, for $i \in \{0,1\}$. From the previous observations, we can associate the parameter N and the $N-$bit string $\mathbf{h}$ in Lemma 6 with $|\underline{\Delta}_0| + |\underline{\Delta}_1|$ and with a string of bit entries $h_j = \tilde{d}_{0,j}$ for $j \in \underline{\Delta}_0 \cup \underline{\Delta}_1$, respectively. From (147) and (167), we associate the parameters $\gamma_{\rm err}$, λ and O in Lemma 6 with the parameters δ, $\lambda(\theta, \beta_{\rm PB})$ and $O(\theta)$ here, respectively. Thus, since $\tilde{d}_{0,j} = \tilde{d}_{1,j} \oplus 1$, for $j \in \underline{\Delta}_0 \cup \underline{\Delta}_1$, the set $S_i^{\mathbf{h}}$ in Lemma 6 corresponds to the set $\underline{\Delta}_i$ here, for $i \in \{0,1\}$. Thus, from (160) and (165), we can associate the operator $D_{\mathbf{a},\mathbf{b}}$ in Lemma 6 with the operator $\underline{\tilde{D}}_{x\mathbf{ab}}$ here. Therefore, the bound (166) follows from Lemma 6.

Since as follows from (111), $P_{\underline{\mathbf{t}}} = \prod_{k \in \Omega_{\rm qub}} P^k_{\rm PS}(t_k)$ with $\frac{1}{2} - \beta_{\rm PS} \leq P^k_{\rm PS}(t_k) \leq \frac{1}{2} + \beta_{\rm PS}$ for $t_k \in \{0,1\}$ and for $k \in \Omega_{\rm qub}$, we have from (162) and (165) that

$$\tilde{D}_{x\mathbf{ab}} \leq \Big(\frac{1}{2} + \beta_{\rm PS}\Big)^{(|\underline{\Delta}_0| + |\underline{\Delta}_1|)} \underline{\tilde{D}}_{x\mathbf{ab}}, \tag{168}$$

where we used that $|\underline{\Lambda}_i| = |\underline{\Delta}_i|$, for $i \in \{0,1\}$. Thus, from (166) and (168), we have

$$\begin{aligned}
&\max_{\mathbf{a},\mathbf{b}} \|\tilde{D}_{x\mathbf{ab}}\| \\
&\quad \leq e^{-\big(|\underline{\Delta}_0| + |\underline{\Delta}_1|\big) \Big[\frac{\lambda(\theta, \beta_{\rm PB})}{2} \big(1 - \frac{\delta}{\lambda(\theta, \beta_{\rm PB})}\big)^2 - \ln(1 + 2\beta_{\rm PS})\Big]} \times \\
&\qquad \times 2^{-\big(|\underline{\Delta}_0| + |\underline{\Delta}_1|\big)}.
\end{aligned} \tag{169}$$

We derive an upper bound on $\|\tilde{\phi}_x\|$. From (137) and (163), we have

$$\tilde{\phi}_x = 2^{|\Omega_{\rm qub}|} \bigotimes_{k \in \Omega_{\rm qub} \setminus \underline{\Lambda}} \big(\rho^k\big)_{B_k}, \tag{170}$$

where ρ^k is a qubit density matrix given by

$$\rho^k = \sum_{u=0}^{1} \sum_{t=0}^{1} P^k_{\rm PB}(u) P^k_{\rm PS}(t) |\phi^k_{tu}\rangle\langle\phi^k_{tu}|, \tag{171}$$

for $k \in \Omega_{\rm qub} \setminus \underline{\Lambda}$. Let $\mu^k_{\pm}$ be the eigenvalues of ρ^k, satisfying $\mu^k_- \leq \mu^k_+$, for $k \in \Omega_{\rm qub} \setminus \underline{\Lambda}$. It follows from Lemma 7 that

$$\mu^k_+ \leq \frac{1}{2}\Big(1 + h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)\Big), \tag{172}$$

where $h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)$ is given by (101), for $k \in \Omega_{\rm qub} \setminus \underline{\Lambda}$. It follows that

$$\begin{aligned}
\|\tilde{\phi}_x\| &= 2^{|\Omega_{\rm qub}|} \prod_{k \in \Omega_{\rm qub} \setminus \underline{\Lambda}} \mu^k_+ \\
&\leq 2^{(|\underline{\Delta}_0| + |\underline{\Delta}_1|)} \big(1 + h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)\big)^{|\Omega_{\rm qub}| - |\underline{\Delta}_0| - |\underline{\Delta}_1|},
\end{aligned} \tag{173}$$

where in the first line we used (170) and (171); and in the second line we used (172), the fact that $\Omega_{\rm qub} = \underline{\Lambda}_0 \cup \underline{\Lambda}_1 \cup \{\Omega_{\rm qub} \setminus \underline{\Lambda}\}$, that the sets $\underline{\Lambda}_0$ and $\underline{\Lambda}_1$ do not intersect, that $\underline{\Lambda} = \underline{\Lambda}_0 \cup \underline{\Lambda}_1$, and that $|\underline{\Delta}_i| = |\underline{\Lambda}_i|$ for $i \in \{0,1\}$.

It follows from (164), (169) and (173) that

$$\begin{aligned}
&\max_{\mathbf{a},\mathbf{b}} \|D_{x\mathbf{ab}}\| \\
&\quad \leq e^{-(|\underline{\Delta}_0| + |\underline{\Delta}_1|) \Big[\frac{\lambda(\theta, \beta_{\rm PB})}{2} \big(1 - \frac{\delta}{\lambda(\theta, \beta_{\rm PB})}\big)^2 - \ln(1 + 2\beta_{\rm PS})\Big]} \times \\
&\qquad \times e^{\ln\big(1 + h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)\big) \big(|\Omega_{\rm qub}| - |\underline{\Delta}_0| - |\underline{\Delta}_1|\big)} \\
&\quad \leq e^{-N f(\gamma_{\rm err}, \beta_{\rm PS}, \beta_{\rm PB}, \theta, \nu_{\rm unf}, \gamma_{\rm det})},
\end{aligned} \tag{174}$$

where in the second line we used (100) and

$$\begin{aligned}
&N f(\gamma_{\rm err}, \beta_{\rm PS}, \beta_{\rm PB}, \theta, \nu_{\rm unf}, \gamma_{\rm det}) \\
&\leq (|\underline{\Delta}_0| + |\underline{\Delta}_1|)\Big[\frac{\lambda(\theta, \beta_{\rm PB})}{2}\Big(1 - \frac{\delta}{\lambda(\theta, \beta_{\rm PB})}\Big)^2 - \ln(1 + 2\beta_{\rm PS})\Big] \\
&\quad - \ln\big(1 + h(\beta_{\rm PS}, \beta_{\rm PB}, \theta)\big)\big(|\Omega_{\rm qub}| - |\underline{\Delta}_0| - |\underline{\Delta}_1|\big).
\end{aligned} \tag{175}$$

As we show below, (175) holds for all possible values of x. Thus, (174) holds for all possible values of x. It follows that

$$\max_{x,\mathbf{a},\mathbf{b}} \|D_{x\mathbf{ab}}\| \leq e^{-N f(\gamma_{\rm err}, \beta_{\rm PS}, \beta_{\rm PB}, \theta, \nu_{\rm unf}, \gamma_{\rm det})}, \tag{176}$$

with $f(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det})$ given by (100), which is the claimed bound (159).

We show that (175) holds for all possible values of x. First, we note from (99) that

$$\frac{\lambda(\theta,\beta_{\rm PB})}{2}\Big(1-\frac{\delta}{\lambda(\theta,\beta_{\rm PB})}\Big)^2-\ln(1+2\beta_{\rm PS})>0. \quad (177)$$

Second, from $\beta_{\rm PS}>0$ and from the definition (101) we have $h(\beta_{\rm PS},\beta_{\rm PB},\theta)>0$. Thus, we have

$$\ln\big(1+h(\beta_{\rm PS},\beta_{\rm PB},\theta)\big)>0. \quad (178)$$

Third, from (131) and (140), and from the condition $n\geq\gamma_{\rm det}N$ for Bob not aborting, we have

$$|\underline{\Delta}_0|+|\underline{\Delta}_1|\geq N(\gamma_{\rm det}-\nu_{\rm unf}). \quad (179)$$

Fourth, it follows from (99) that

$$\gamma_{\rm det}-\nu_{\rm unf}>0. \quad (180)$$

Thus, since $|\Omega_{\rm qub}|\leq N$, (175) follows from (177) – (180) and from the definition of $f(\gamma_{\rm err},\beta_{\rm PS},\beta_{\rm PB},\theta,\nu_{\rm unf},\gamma_{\rm det})$ given by (100).

VIII. THE CASE OF 2^M PRESENTATION POINTS

The proof of Theorem 2 follows straightforwardly from Lemmas 8 – 10 and from Theorem 3 at the end of this section.

The quantum token schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ presented below extend the quantum token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ of Tables 2 and Table 3 to the case of 2^M presentation points, for arbitrary integer $M\geq 1$. Broadly speaking, the schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ generate the classical inputs and outputs of the schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ as subroutines, M times in parallel, with a few differences arising due to the fact that $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ have 2^M presentation points instead of two. Similarly to $\mathcal{QT}_1$ and $\mathcal{QT}_2$, $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ can be implemented in practice with the photonic setups of Fig. 3, respectively.

In the schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$, there are two presentation points Q_0 and Q_1, Alice has agents $\mathcal{A},\mathcal{A}_0,\mathcal{A}_1$ and Bob has agents $\mathcal{B},\mathcal{B}_0,\mathcal{B}_1$. From Tables 2 and Table 3, we see that in $\mathcal{QT}_1$ and $\mathcal{QT}_2$, $\mathcal{B}$ obtains $\mathbf{t},\mathbf{u}\in\{0,1\}^N$, $\Omega_{\rm noqub}\subseteq[N]$, and $\mathbf{r},\mathbf{s}\in\{0,1\}^n$ in the intersection of the causal pasts of the presentation points; $\mathcal{A}$ obtains $\Lambda\subseteq[N]$, $n=|\Lambda|$, $W\in\{0,1\}^\Lambda$, $g:\Lambda\to[n]$, $\mathbf{y},\mathbf{x},\mathbf{d}\in\{0,1\}^n$, and $b,c,z\in\{0,1\}$ in the intersection of the causal pasts of the presentation points; $\mathcal{B}_i$ obtains $\mathbf{d}_i,\tilde{\mathbf{d}}_i\in\{0,1\}^n$ in the causal past of Q_i, for $i\in\{0,1\}$; $\mathcal{A}_b$ presents $\mathbf{x}$ to $\mathcal{B}_b$ in Q_b; and $\mathcal{B}_b$ obtains $\mathbf{x}_b,\mathbf{r}_b\in\{0,1\}^n$ and Δ_b in Q_b.

On the other hand, in the schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$, there are 2^M presentation points Q_i, Alice has agents $\mathcal{A},\mathcal{A}_i$, and Bob has agents $\mathcal{B},\mathcal{B}_i$, where $i=(i^1,\ldots,i^M)\in\{0,1\}^M$. In these schemes, Alice's and Bob's agents obtain the inputs and outputs of the schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ in the corresponding spacetime regions, as mentioned above, in M independent rounds. For the lth round, we label the inputs and outputs mentioned above by a superscript l. In the schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$, the experimental imperfections of Table 5 and the assumptions of Table 6 apply independently to each of the M rounds.

The schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ have a new step (step 12 of $\mathcal{QT}_1^M$) compared to $\mathcal{QT}_1$ and $\mathcal{QT}_2$, in which $\mathcal{B}_i$ sends a signal to agents $\mathcal{B}_{i'}$ with $Q_{i'}$ in the causal future of Q_i indicating whether a token was presented at Q_i by $\mathcal{A}_i$. This extra step allows us to reduce the proof of unforgeability to the case of spacelike separated presentation points. We note that instant validation is still satisfied, as no extra delays for token validation due to cross-checking are required.

The schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are presented precisely below.

A. Quantum token scheme $\mathcal{QT}_1^M$ for 2^M presentation points

Steps 1 to 10 below are repeated in M independent rounds, labelled by $l\in[M]$. Steps 1 to 9 take place within the intersection of the causal pasts of the presentation points.

1. Preparation stage

0. Alice and Bob agree on a reference frame, on presentation points Q_i in the agreed frame, for $i\in\{0,1\}^M$, and on parameters $N\in\mathbb{N}$, $\beta_{\rm PB}\in\big(0,\frac{1}{2}\big)$, $\gamma_{\rm det}\in(0,1)$ and $\gamma_{\rm err}\in(0,1)$.

2. Stage I

1. For $k\in[N]$, $\mathcal{B}$ prepares bits t_k^l and u_k^l with respective probability distributions $P_{\rm PS}^{k,l}(t_k^l)$ and $P_{\rm PB}^{k,l}(u_k^l)$, satisfying $\frac{1}{2}-\beta_{\rm X}\leq P_{\rm X}^{k,l}(t)\leq\frac{1}{2}+\beta_{\rm X}$, where $\beta_{\rm X}\in\big(0,\frac{1}{2}\big)$ is a small parameter, for ${\rm X}\in\{{\rm PS},{\rm PB}\}$, $t\in\{0,1\}$ and $k\in[N]$. We define $\mathbf{t}^l=(t_1^l,\ldots,t_N^l)$ and $\mathbf{u}^l=(u_1^l,\ldots,u_N^l)$. For $k\in[N]$, $\mathcal{B}$ prepares a quantum system A_k^l in a quantum state $|\psi_k^l\rangle$ and sends it to $\mathcal{A}$ with its label (k,l). $\mathcal{B}$ chooses $(k,l)\in\Omega_{\rm noqub}^l$ with probability $P_{\rm noqub}>0$ or $(k,l)\in\Omega_{\rm qub}^l$ with probability $1-P_{\rm noqub}$. For $(k,l)\in\Omega_{\rm qub}$, $|\psi_k^l\rangle=|\phi^{k,l}_{t_k^l u_k^l}\rangle$ is a qubit state, where $\langle\phi_{0u}^{k,l}|\phi_{1u}^{k,l}\rangle=0$ for $u\in\{0,1\}$, where the qubit orthonormal basis $\mathcal{D}_u^{k,l}=\{|\phi_{tu}^{k,l}\rangle\}_{t=0}^1$ is the computational (Hadamard) basis up to an uncertainty angle θ on the Bloch sphere if $u=0$ ($u=1$). For

$(k,l) \in \Omega^l_{\text{noqub}}$, $|\psi^l_k\rangle = |\Phi^{l,k}_{t^l_k u^l_k}\rangle$ is a quantum state of arbitrary finite Hilbert space dimension greater than two. In photonic implementations, a vacuum or one-photon pulse has label $(k,l) \in \Omega^l_{\text{qub}}$, with a one-photon pulse encoding a qubit state, while a multi-photon pulse has label $(k,l) \in \Omega^l_{\text{noqub}}$ and encodes a quantum state of finite Hilbert space dimension greater than two.

2. For $k \in [N]$, $\mathcal{A}$ measures A^l_k in the qubit orthonormal basis $\mathcal{D}_{w^l_k}$, for $w^l_k \in \{0,1\}$. Due to losses, $\mathcal{A}$ only successfully measures quantum states $|\psi^l_k\rangle$ with labels (k,l) from a proper subset Λ^l of $[N]$. Let W^l be the string of bit entries w^l_k for $(k,l) \in \Lambda^l$ and let $n^l = |\Lambda^l|$. Conditioned on $(k,l) \in \Lambda^l$, the probability that $\mathcal{A}$ measures A^l_k in the basis $\mathcal{D}_{w^l_k}$ satisfies $P_{\text{MB}}(w^l_k) = \frac{1}{2}$, for $w^l_k \in \{0,1\}$ and $k \in [N]$. $\mathcal{A}$ reports to $\mathcal{B}$ the set Λ^l with its label l. $\mathcal{B}$ does not abort if and only if $n^l \geq \gamma_{\text{det}} N$.

3. $\mathcal{A}$ chooses a one-to-one function $g^l : \Lambda^l \to [n]$, for example the numerical ordering, and sends it to $\mathcal{B}$ with its label l. Let $y^l_j \in \{0,1\}$ indicate the basis $\mathcal{D}_{y^l_j}$ on which the quantum state $|\psi^l_k\rangle$ is measured by $\mathcal{A}$ and let $x^l_j \in \{0,1\}$ be the measurement outcome, where $j = g^l(k)$, for $k \in \Lambda^l$ and $j \in [n]$. Let $\mathbf{y}^l = (y^l_1, \ldots, y^l_n) \in \{0,1\}^n$ and $\mathbf{x}^l = (x^l_1, \ldots, x^l_n) \in \{0,1\}^n$ denote the strings of Alice's measurement bases and outcomes, respectively.

4. $\mathcal{A}$ sends $\mathbf{x}^l$ to $\mathcal{A}_i$ with its label l, for $i \in \{0,1\}^M$.

5. $\mathcal{A}$ chooses a bit $z^l \in \{0,1\}$ with probability $P^l_{\text{E}}(z^l)$ that satisfies $\frac{1}{2} - \beta_{\text{E}} \leq P^l_{\text{E}}(z^l) \leq \frac{1}{2} + \beta_{\text{E}}$, for $z^l \in \{0,1\}$, and for a small parameter $\beta_{\text{E}} \in \left(0, \frac{1}{2}\right)$. $\mathcal{A}$ computes the string $\mathbf{d}^l \in \{0,1\}^n$ with bit entries $d^l_j = y^l_j \oplus z^l$, for $j \in [n]$. $\mathcal{A}$ sends $\mathbf{d}^l$ to $\mathcal{B}$ with its label l.

6. For $i = (i^1, \ldots, i^M) \in \{0,1\}^M$, $\mathcal{B}$ sends $\mathbf{d}^l$ to $\mathcal{B}_i$ with its label l, and $\mathcal{B}_i$ computes the string $\mathbf{d}^l_{i^l} \in \{0,1\}^n$ with bit entries $d^l_{i^l,j} = d^l_j \oplus i^l$, for $j \in [n]$.

7. $\mathcal{B}$ uses $\mathbf{t}^l, \mathbf{u}^l, \Lambda^l$ and g^l to compute the strings $\mathbf{s}^l, \mathbf{r}^l \in \{0,1\}^n$, as follows. We define $r^l_j = t^l_k$, and $s^l_j = u^l_k$, where $j = g^l(k)$, for $j \in [n]$ and $k \in \Lambda^l$. We define $\mathbf{r}^l$ and $\mathbf{s}^l$ as the strings with bit entries r^l_j and s^l_j, for $j \in [n]$, respectively. For $\mathcal{B}$ sends $\mathbf{s}^l$ and $\mathbf{r}^l$ to $\mathcal{B}_i$ with its label l, for $i \in \{0,1\}^M$.

3. Stage II

8. $\mathcal{A}$ chooses the lth entry $b^l \in \{0,1\}$ for the bit string $b = (b^1, \ldots, b^M) \in \{0,1\}^M$ that labels the presentation point Q_b where to present the token. $\mathcal{A}$ computes the bit $c^l = b^l \oplus z^l$ and sends it to $\mathcal{B}$ with its label l.

9. $\mathcal{B}$ sends c^l with its label l to $\mathcal{B}_i$, for $i \in \{0,1\}^M$.

10. For $i \in \{0,1\}^M$, in the causal past of Q_i, $\mathcal{B}_i$ computes the string $\tilde{\mathbf{d}}^l_{i^l} \in \{0,1\}^n$ with bit entries $\tilde{d}^l_{i^l,j} = d^l_{i^l,j} \oplus c^l$, for $j \in [n]$.

11. $\mathcal{A}$ sends a signal to $\mathcal{A}_b$ indicating to present the token at Q_b, and $\mathcal{A}_b$ presents the token $\mathbf{x} = (\mathbf{x}^1, \ldots, \mathbf{x}^M)$ to $\mathcal{B}_b$ in Q_b.

12. For all $i \in \{0,1\}^M$, if $\mathcal{B}_i$ receives a token from $\mathcal{A}_i$ at Q_i, $\mathcal{B}_i$ sends a signal to $\mathcal{B}_{i'}$ indicating so, for all $i' \in \{0,1\}^M$ such that $Q_{i'}$ is in the causal future of Q_i.

13. $\mathcal{B}_b$ validates the token $\mathbf{x}$ received in Q_b if two conditions hold: 1) $\mathcal{B}_b$ does not receive signals from Bob's agent $\mathcal{B}_i$ indicating that a token has been presented by Alice at Q_i, for any $i \in \{0,1\}^M$ such that Q_i is in the causal past of Q_b; and 2) for all $l \in [M]$, the Hamming distance between the strings $\mathbf{x}^l_{b^l}$ and $\mathbf{r}^l_{b^l}$ satisfies $d(\mathbf{x}^l_{b^l}, \mathbf{r}^l_{b^l}) \leq |\Delta^l_{b^l}|\gamma_{\text{err}}$, where $\Delta^l_v = \{j \in [n] | \tilde{d}^l_{v,j} = s^l_j\}$, and where $\mathbf{a}^l_v$ is the restriction of a string $\mathbf{a}^l \in \{\mathbf{x}^l, \mathbf{r}^l\}$ to entries a^l_j with $j \in \Delta^l_v$, for $v \in \{0,1\}$.

B. Quantum token scheme $\mathcal{QT}^M_2$ for 2^M presentation points

Steps 1 to 7 below are repeated in M independent rounds, labelled by $l \in [M]$. Steps 1 to 6 take place within the intersection of the causal pasts of the presentation points.

1. Preparation stage

0. As step 0 of $\mathcal{QT}^M_1$.

2. Stage I

1. As step 1 of $\mathcal{QT}^M_1$.

2. The step 2 of $\mathcal{QT}^M_1$ is replaced by the following. $\mathcal{A}$ chooses a bit z^l with probability $P^l_{\text{E}}(z^l)$ satisfying $\frac{1}{2} - \beta_{\text{E}} \leq P^l_{\text{E}}(z^l) \leq \frac{1}{2} + \beta_{\text{E}}$, for $z^l \in \{0,1\}$, and for a small parameter $\beta_{\text{E}} \in \left(0, \frac{1}{2}\right)$. $\mathcal{A}$ measures A^l_k in the qubit orthonormal basis $\mathcal{D}_{z^l}$, for $k \in [N]$. Due to losses, $\mathcal{A}$ only successfully measures quantum states $|\psi^l_k\rangle$ with labels (l,k) from a proper subset Λ^l of $[N]$. $\mathcal{A}$ reports to $\mathcal{B}$ the set Λ^l with its label l. Let $n^l = |\Lambda^l|$. $\mathcal{B}$ does not abort if and only if $n^l \geq \gamma_{\text{det}} N$.

3. As step 3 of $\mathcal{QT}_1^M$. The string $\mathbf{y}^l \in \{0,1\}^n$ of Alice's measurement bases has bit entries $y_j^l = z^l$, for $j \in [n]$.

4. As step 4 of $\mathcal{QT}_1^M$. The steps 5 and 6 of $\mathcal{QT}_1^M$ are discarded.

5. As step 7 of $\mathcal{QT}_1^M$.

3. Stage II

6. As steps 8 and 9 of $\mathcal{QT}_1^M$.

7. The step 10 of $\mathcal{QT}_1^M$ is replaced by the following. For $i = (i^1, \ldots, i^M) \in \{0,1\}^M$, in the causal past of Q_i, $\mathcal{B}_i$ computes the string $\tilde{\mathbf{d}}_{i^l}^l \in \{0,1\}^n$ with bit entries $\tilde{d}_{i^l,j}^l = i^l \oplus c^l$, for $j \in [n]$.

8. As steps 11, 12 and 13 of $\mathcal{QT}_1^M$.

C. Comments

We note that steps 1 to 11 and 1 to 7 of the token schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are straightforward extensions of the corresponding steps in the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$, respectively. As we discussed above, this basically comprises applying the corresponding steps of $\mathcal{QT}_1$ and $\mathcal{QT}_2$ in M parallel and independent rounds. However, step 12 of $\mathcal{QT}_1^M$ is a new step, and steps 13 and 8 of $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ modify steps 12 and 8 of $\mathcal{QT}_1$ and $\mathcal{QT}_2$, respectively, to account for the new step.

We note from steps 12 and 13 of $\mathcal{QT}_1^M$, and from step 8 of $\mathcal{QT}_2^M$, that a token received by Bob's agent $\mathcal{B}_b$ from Alice's agent $\mathcal{A}_b$ at a presentation point Q_b can be validated by $\mathcal{B}_b$ nearly instantly at Q_b. In particular, $\mathcal{B}_b$ does not need to wait for any signals coming from agents $\mathcal{B}_i$ who have possibly received tokens from Alice's agents at presentation points Q_i that are not in the causal past of Q_b.

For $M > 1$, steps 12 and 13 of $\mathcal{QT}_1^M$, and step 8 of $\mathcal{QT}_2^M$, allow us to guarantee unforgeability, as discussed below. In the case $M = 1$, steps 12 and 13 of $\mathcal{QT}_1^M$, and step 8 of $\mathcal{QT}_2^M$, can simply be replaced by step 12 of $\mathcal{QT}_1$, and by step 8 of $\mathcal{QT}_2$, respectively.

Every observation made previously about the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$ also applies to the token schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$. In particular, the schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ also allow for the experimental imperfections of Table 5 and make the assumptions of Table 6, for the lth round and for $l \in [M]$. Stage I includes the quantum communication, which can take place between adjacent laboratories, and must be implemented within the intersection of the causal pasts of all the presentation points. This stage can take an arbitrarily long time and can be completed arbitrarily in the past of the presentation points, which is very helpful for practical implementations. Stage II comprises only classical processing and communication, and must usually be completed within a very short time. We note that Alice chooses her presentation point in stage II, meaning in particular that it can take place after her quantum measurements have been completed, which gives Alice great flexibility in spacetime to choose her presentation point. The token schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ can be modified in various ways, as discussed previously for the $\mathcal{QT}_1$ and $\mathcal{QT}_2$ schemes.

D. Robustness, correctness, privacy and unforgeability

As discussed for the token schemes $\mathcal{QT}_1$ and $\mathcal{QT}_2$, in the token schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ we define P_{det} as the probability that a quantum state $|\psi_k^l\rangle$ transmitted by Bob is reported by Alice as being successfully measured, with label $(l,k) \in \Lambda^l$, for $k \in [N]$ and $l \in [M]$. We define E as the probability that Alice obtains a wrong measurement outcome when she measures a quantum state $|\psi_k^l\rangle$ in the basis of preparation by Bob; if the error rates E_{tu} are different for different prepared states, labelled by t, and for different measurement bases, labelled by u, we simply take $E = \max_{t,u}\{E_{tu}\}$.

The robustness, correctness, privacy and unforgeability of $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are stated by the following lemmas and theorem. These lemmas and theorem consider parameters $\gamma_{\text{det}}, \gamma_{\text{err}} \in (0,1)$, allow for the experimental imperfections of Table 5 and make the assumptions of Table 6, for each of the M rounds labelled by $l \in [M]$, as discussed above.

Lemma 8. *If*

$$0 < \gamma_{det} < P_{det}, \tag{181}$$

then $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are ϵ_{rob}^M-robust with

$$\epsilon_{rob}^M = 1 - (1 - \epsilon_{rob})^M \leq M\epsilon_{rob}, \tag{182}$$

where

$$\epsilon_{rob} = e^{-\frac{P_{det}N}{2}\left(1-\frac{\gamma_{det}}{P_{det}}\right)^2}. \tag{183}$$

Proof. Suppose that (181) holds and that Alice and Bob follow the scheme $\mathcal{QT}_a^M$ honestly, for $a \in \{0,1\}$. From Lemma 2, the probability P_{abort}^1 that Bob aborts in the round label by $l = 1$ satisfies $P_{\text{abort}}^1 \leq \epsilon_{\text{rob}}$, with ϵ_{rob} given by (183). Since steps 1 to 10 of $\mathcal{QT}_1^M$ and steps 1 to 7 of $\mathcal{QT}_2^M$ are implemented in M independent rounds, we also have from Lemma 2 that the probability P_{abort}^l that Bob aborts in the lth round, given that he does not abort in the rounds $1, 2, \ldots, l-1$, satisfies $P_{\text{abort}}^l \leq \epsilon_{\text{rob}}$, with ϵ_{rob} given by (183), for $l \in \{2, \ldots, M\}$. Thus, the probability P_{abort} that Bob aborts in the scheme satisfies

$$P_{\text{abort}} \leq 1 - (1 - \epsilon_{\text{rob}})^M. \tag{184}$$

Thus, the schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are $\epsilon_{\rm rob}^M-$robust with $\epsilon_{\rm rob}^M$ given by (182), as claimed. The inequality in (182) follows from Bernoulli's inequality. □

Lemma 9. *If*

$$0 < \frac{\gamma_{err}}{2} < E < \gamma_{err},$$
$$0 < \nu_{cor} < \frac{P_{det}(1-2\beta_{PB})}{2}, \tag{185}$$

then $\mathcal{QT}_1^M$ *and* $\mathcal{QT}_2^M$ *are* ϵ_{cor}^M-*correct with*

$$\epsilon_{cor}^M = 1-(1-\epsilon_{cor})^M \leq M\epsilon_{cor}, \tag{186}$$

where

$$\epsilon_{cor} = e^{-\frac{P_{det}(1-2\beta_{PB})N}{4}\left(1-\frac{2\nu_{cor}}{P_{det}(1-2\beta_{PB})}\right)^2} + e^{-\frac{E\nu_{cor}N}{3}\left(\frac{\gamma_{err}}{E}-1\right)^2}. \tag{187}$$

Proof. Suppose that (185) holds and that Alice and Bob follow the scheme $\mathcal{QT}_a^M$ honestly, for $a \in \{0,1\}$. We see from step 13 of $\mathcal{QT}_1^M$ and step 8 of $\mathcal{QT}_2^M$ that $\mathcal{B}_b$ validates Alice's token $\mathbf{x} = (\mathbf{x}^1, \ldots, \mathbf{x}^M)$ at the presentation point Q_b if the condition $d(\mathbf{x}_{b^l}^l, \mathbf{r}_{b^l}^l) \leq |\Delta_{b^l}^l|\gamma_{\rm err}$ is satisfied for all $l \in [M]$. Since steps 1 to 10 of $\mathcal{QT}_1^M$ and steps 1 to 7 of $\mathcal{QT}_2^M$ are implemented in M independent rounds, we see that the probability that each of these conditions is satisfied is independent of whether the other conditions are satisfied. We see that steps 1 to 10 (1 to 7) of the lth round in $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$) and the lth condition for token validation in step 13 (8) of $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$) are equivalent to the corresponding steps and the condition for token validation in $\mathcal{QT}_1$ ($\mathcal{QT}_2$), for $l \in [M]$. Thus, we have from Lemma 3 that the probability $P_{\rm fail}^l$ that the lth condition for token validation in $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$) is not passed satisfies $P_{\rm fail}^l \leq \epsilon_{\rm cor}$, with $\epsilon_{\rm cor}$ given by (187), for $l \in [M]$. Thus, the probability $P_{\rm fail}$ that the token $\mathbf{x}$ is not validated by $\mathcal{B}_b$ in Q_b in either scheme $\mathcal{QT}_1^M$ or $\mathcal{QT}_2^M$ satisfies

$$P_{\rm fail} \leq 1-(1-\epsilon_{\rm cor})^M. \tag{188}$$

Thus, the schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are $\epsilon_{\rm cor}^M-$correct with $\epsilon_{\rm cor}^M$ given by (186), as claimed. The inequality in (186) follows from Bernoulli's inequality. □

Lemma 10. $\mathcal{QT}_1^M$ *and* $\mathcal{QT}_2^M$ *are* ϵ_{priv}^M-*private with*

$$\epsilon_{priv}^M = \frac{1}{2^M}\left[(1+2\epsilon_{priv})^M - 1\right], \tag{189}$$

with

$$\epsilon_{priv} = \beta_E. \tag{190}$$

Proof. Suppose that Alice follows the scheme $\mathcal{QT}_a^M$ honestly, for $a \in \{0,1\}$. From assumption C (see Table 6), the set Λ^l of labels transmitted to $\mathcal{B}$ in step 2 of $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ in the lth round gives $\mathcal{B}$ no information about the string W^l and the bit z^l, for $l \in [M]$. Furthermore, from assumption E (see Table 6), $\mathcal{B}$ cannot use degrees of freedom not previously agreed for the transmission of the quantum states to affect, or obtain information about, the statistics of the quantum measurement devices of $\mathcal{A}$. Moreover, in our setting, we assume that Alice's laboratories are secure and that communication among Alice's agents is made through secure and authenticated classical channels. It follows from these assumptions that the only way in which Bob can obtain information about Alice's bit string $b = (b^1, \ldots, b^M)$ before she presents the token is via the bit messages $c^l = z^l \oplus b^l$, for $l \in [M]$. Since Alice prepares the bits z^l in independent rounds, for $l \in [M]$, the probability that Bob guesses Alice's bit string b is given by

$$P_{\rm Bob} = \prod_{l=1}^{M} P_{\rm Bob}^l, \tag{191}$$

where $P_{\rm Bob}^l$ is the probability that Bob guesses Alice's bit b^l, for $l \in [M]$.

We see that the steps 1 to 10 (1 to 7) of the lth round in $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$) are equivalent to the corresponding steps in $\mathcal{QT}_1$ ($\mathcal{QT}_2$), for $l \in [M]$. Thus, from Lemma 4, we have

$$P_{\rm Bob}^l \leq \frac{1}{2} + \epsilon_{\rm priv}, \tag{192}$$

for $l \in [M]$, with $\epsilon_{\rm priv}$ given by (190). From (191) and (192), we have that

$$P_{\rm Bob} \leq \left(\frac{1}{2} + \epsilon_{\rm priv}\right)^M = \frac{1}{2^M} + \epsilon_{\rm priv}^M, \tag{193}$$

with $\epsilon_{\rm priv}^M$ given by (189). Thus, the schemes $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are $\epsilon_{\rm priv}^M-$private with $\epsilon_{\rm priv}^M$ given by (189), as claimed. □

Theorem 3. *Consider the constraints*

$$0 < \gamma_{err} < \lambda(\theta, \beta_{PB}),$$
$$0 < P_{noqub} < \nu_{unf} < \min\left\{2P_{noqub}, \gamma_{det}\left(1-\frac{\gamma_{err}}{\lambda(\theta,\beta_{PB})}\right)\right\},$$
$$0 < \beta_{PS} < \frac{1}{2}\left[e^{\frac{\lambda(\theta,\beta_{PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{PB})}\right)^2} - 1\right]. \tag{194}$$

We define the function

$$f(\gamma_{err},\beta_{PS},\beta_{PB},\theta,\nu_{unf},\gamma_{det}) = (\gamma_{det}-\nu_{unf})\left[\frac{\lambda(\theta,\beta_{PB})}{2}\left(1-\frac{\delta}{\lambda(\theta,\beta_{PB})}\right)^2 - \ln(1+2\beta_{PS})\right] - \left(1-(\gamma_{det}-\nu_{unf})\right)\ln\left[1+h(\beta_{PS},\beta_{PB},\theta)\right], \tag{195}$$

where

$$h(\beta_{PS},\beta_{PB},\theta) = 2\beta_{PS}\sqrt{\frac{1}{2}+2\beta_{PB}^2+\left(\frac{1}{2}-2\beta_{PB}^2\right)\sin(2\theta)},$$
$$\delta = \frac{\gamma_{det}\gamma_{err}}{\gamma_{det}-\nu_{unf}}. \tag{196}$$

Let $L \leq 2^M$ be the number of pair-wise spacelike separated presentation points among the 2^M presentation points and let $C = \frac{L(L-1)}{2}$ be the number of pairs of spacelike separated presentation points, if $L \geq 2$, and $C = 0$ if $L = 0$. For any $M \geq 1$, there exist parameters satisfying the constraints (194), for which $f(\gamma_{err},\beta_{PS},\beta_{PB},\theta,\nu_{unf},\gamma_{det}) > 0$. For these parameters, $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are ϵ_{unf}^M-unforgeable with

$$\epsilon_{unf}^M = C\epsilon_{unf}, \tag{197}$$

with

$$\epsilon_{unf} = e^{-\frac{P_{noqub}N}{3}\left(\frac{\nu_{unf}}{P_{noqub}}-1\right)^2} + e^{-Nf(\gamma_{err},\beta_{PS},\beta_{PB},\theta,\nu_{unf},\gamma_{det})}. \tag{198}$$

Proof. From Theorem 1, there exist parameters satisfying the constraints (194), for which $f(\gamma_{\text{err}},\beta_{\text{PS}},\beta_{\text{PB}},\theta,\nu_{\text{unf}},\gamma_{\text{det}}) > 0$. This holds for arbtrary $M \geq 1$ because the constraints (194) and the function $f(\gamma_{\text{err}},\beta_{\text{PS}},\beta_{\text{PB}},\theta,\nu_{\text{unf}},\gamma_{\text{det}})$ are independent of M.

Suppose that the constraints (194) hold and that $f(\gamma_{\text{err}},\beta_{\text{PS}},\beta_{\text{PB}},\theta,\nu_{\text{unf}},\gamma_{\text{det}}) > 0$. Suppose that Bob follows the scheme $\mathcal{QT}_a^M$ honestly and Alice follows an arbitrary cheating strategy $\mathcal{S}$, for $a \in \{0,1\}$. From step 12 (8) of $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$), Alice cannot succeed in making Bob validate tokens at timelike separated presentation points. For this reason, we consider without loss of generality that Alice tries to make Bob validate tokens at spacelike separated presentation points.

Let $L \leq 2^M$ be the number of spacelike separated presentation points. Alice's general cheating strategy $\mathcal{S}$ comprises using the received classical information and quantum sates from Bob to output classical data to give Bob as required by the scheme $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$) in steps 1 to 8 (1 to 6), and to obtain a token to give Bob at each of the L spacelike separated presentation points. We note that in our schemes there are no penalties for Alice if Bob catches her cheating. Thus, it does not affect Alice to present tokens at all spacelike separated presentation points, even if this can in principle increase the probability that Bob catches her cheating. Moreover, by presenting tokens at all spacelike separated presentation points, Alice has a greater probability to make Bob validate tokens at any two or more different presentation points. For example, if by giving tokens at $K < L$ spacelike separated presentation points Alice can make Bob validate tokens at any two or more different presentation points with probability P, Alice can additionally give a random token at another spacelike separated presentation point and in this way increase the probability to some value $P' > P$.

Let $\mathcal{P}_{\text{spacelike}}$ be the set of labels $i = (i^1,\ldots,i^M) \in \{0,1\}^M$ for the spacetime presentation points Q_i that are spacelike separated. Let $v,w \in \mathcal{P}_{\text{spacelike}}$ with $v \neq w$. Let $\mathbf{a} = (\mathbf{a}^1,\ldots,\mathbf{a}^M)$ and $\mathbf{b} = (\mathbf{b}^1,\ldots,\mathbf{b}^M)$ be the tokens that Alice gives Bob at Q_v and Q_w, respectively. Let $P_{vw}^{\mathcal{S}}$ be the probability that Bob validates the token $\mathbf{a}$ at Q_v and the token $\mathbf{b}$ at Q_w. Let $P^{\mathcal{S}}$ be the probability that Bob validates tokens at any two or more different presentation points. We have

$$P^{\mathcal{S}} \leq \sum_{\substack{v,w\in\mathcal{P}_{\text{spacelike}}\\ v\neq w}} P_{vw}^{\mathcal{S}}. \tag{199}$$

We show below that

$$P_{vw}^{\mathcal{S}} \leq \epsilon_{\text{unf}}, \tag{200}$$

for any $v,w \in \mathcal{P}_{\text{spacelike}}$ with $v \neq w$, and for any cheating strategy $\mathcal{S}$ by Alice, where ϵ_{unf} is given by (198). By noticing that by definition, $L = |\mathcal{P}_{\text{spacelike}}|$ is the number of spacelike separated presentation points and C is the number of pairs of spacelike separated presentation points, it follows from (199) and (200) that $\mathcal{QT}_1^M$ and $\mathcal{QT}_2^M$ are ϵ_{unf}^M unforgeable, with ϵ_{unf}^M given by (197), as claimed.

We show (200). Let $v,w \in \mathcal{P}_{\text{spacelike}}$ with $v \neq w$. Let $v = (v^1,\ldots,v^M)$ and $w = (w^1,\ldots,w^M)$, where $v^l, w^l \in \{0,1\}$, for $l \in [M]$. Since $v \neq w$, there exists $l' \in [M]$ such that

$$v^{l'} = w^{l'} \oplus 1. \tag{201}$$

Thus, without loss of generality, let

$$v^{l'} = 0 \qquad \text{and} \qquad w^{l'} = 1. \tag{202}$$

Bob validating the token $\mathbf{a}$ at Q_v and the token $\mathbf{b}$ at Q_w requires satisfaction of the conditions $d(\mathbf{a}_{v^l}^l, \mathbf{r}_{v^l}^l) \leq |\Delta_{v^l}^l|\gamma_{\text{err}}$ at Q_v and $d(\mathbf{b}_{w^l}^l, \mathbf{r}_{w^l}^l) \leq |\Delta_{w^l}^l|\gamma_{\text{err}}$ at Q_w, for all $l \in [M]$. Thus, it requires in particular satisfaction of the conditions

$$\begin{aligned} d(\mathbf{a}_0^{l'}, \mathbf{r}_0^{l'}) &\leq |\Delta_0^{l'}|\gamma_{\text{err}}, \\ d(\mathbf{b}_1^{l'}, \mathbf{r}_1^{l'}) &\leq |\Delta_1^{l'}|\gamma_{\text{err}}, \end{aligned} \tag{203}$$

at Q_v and Q_w, respectively, where we used (202).

Since Bob follows the scheme honestly, he follows the steps 1 to 10 (1 to 7) of the l'th round in $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$) independently of rounds with label $l \neq l'$. we see that Bob's steps 1 to 10 (1 to 7) of the l'th round in $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$) and his l'th conditions for token validation at Q_v and Q_w given by (203) are equivalent to Bob's corresponding steps and conditions for token validation at the two presentation points in $\mathcal{QT}_1$ ($\mathcal{QT}_2$). Thus, from Theorem 1, the probability that both conditions (203) are satisfied is upper bounded by ϵ_{unf}, with ϵ_{unf} given by (198). It follows that

$$P_{vw}^{\mathcal{S}} \leq \epsilon_{\text{unf}}, \tag{204}$$

for any $v,w \in \mathcal{P}_{\text{spacelike}}$ with $v \neq w$, and for any cheating strategy $\mathcal{S}$ by Alice.

□

We note that Lemmas 8, 9 and 10 reduce to Lemmas 2, 3 and 4 in the case $M = 1$, respectively. Similarly, Theorem 3 reduces to Theorem 1 in the case $M = 1$ if the presentation points are spacelike separated. This follows straightforwardly from the fact that $\mathcal{QT}_a^M$ reduces to $\mathcal{QT}_a$ for the case $M = 1$, for $a \in \{1, 2\}$, except for steps 12 and 13 of $\mathcal{QT}_1^M$ and step 8 of $\mathcal{QT}_2^M$, which as mentioned above can simply be replaced by step 12 of $\mathcal{QT}_1$ and step 8 of $\mathcal{QT}_2$ in this case, respectively. In the case $M = 1$ with timelike separated presentation points, differently to Theorem 1, Theorem 3 sates that the probability that Bob validates tokens at both presentation points is zero. This is due to the extra step 12 (8) in $\mathcal{QT}_1^M$ ($\mathcal{QT}_2^M$). In any case, Theorem 3 is consistent with Theorem 1 in the case $M = 1$, if the presentation points are timelike or spacelike separated.

[1] M. Bozzio, A. Cavaillès, E. Diamanti, A. Kent, and D. Pitalúa-García, Multiphoton and side-channel attacks in mistrustful quantum cryptography, PRX Quantum **2**, 030338 (2021).

[2] M. Mitzenmacher and E. Upfal, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis* (Cambridge University Press, Cambridge, UK, 2005).

[3] T. Lunghi *et al.*, Experimental bit commitment based on quantum communication and special relativity, Phys. Rev. Lett. **111**, 180504 (2013).