Experimentally Certified Transmission of a Quantum Message through an Untrusted and Lossy Quantum Channel via Bell's Theorem

Simon Neves,[1, 2] Laura dos Santos Martins,[1] Verena Yacoub,[1] Pascal Lefebvre,[1] Ivan Šupić,[1] Damian Markham,[1] and Eleni Diamanti[1]

[1]*Sorbonne Université, CNRS, LIP6, 4 Place Jussieu, Paris F-75005, France*
[2]*Université Marie et Louis Pasteur, CNRS, Institut FEMTO-ST, F-25000 Besançon, France*

(Dated: August 26, 2025)

Quantum transmission links are central elements in essentially all protocols involving the exchange of quantum messages. Emerging progress in quantum technologies involving such links needs to be accompanied by appropriate certification tools. In adversarial scenarios, a certification method can be vulnerable to attacks if too much trust is placed on the underlying system. Here, we propose a protocol in a device independent framework, which allows for the certification of practical quantum transmission links in scenarios where minimal assumptions are made about the functioning of the certification setup. In particular, we take unavoidable transmission losses into account by modeling the link as a completely-positive trace-decreasing map. We also, crucially, remove the assumption of independent and identically distributed samples, which is known to be incompatible with adversarial settings. Particular emphasis is put on a one-sided device-independent scenario, in which the sender possesses trusted resources. Finally, in view of the use of the certified transmitted states for follow-up applications, our protocol moves beyond certification of the channel to allow us to estimate the quality of the transmitted quantum message itself. To illustrate the practical relevance and the feasibility of our protocol with currently available technology, we provide an experimental implementation in the one-sided device-independent setting, based on a state-of-the-art polarization entangled photon pair source in a Sagnac configuration and analyze its robustness for realistic losses and errors.

Introduction

The ability to send and receive quantum information is at the heart of the rapidly developing quantum technologies. Transmitting quantum information over quantum networks promises unparalleled efficiency and security [1], as well as new functionalities such as the delegation of quantum computation [2] and quantum sensing [3]. Within quantum computers themselves we will need to input, share and distribute quantum information to different parts, particularly important for architectures relying on multiple quantum processors [4, 5]. The reliable transmission of quantum information is thus an essential building block for future quantum technologies, and, as such, we must be very sure of its working. When the physical devices used to test and use these quantum channels are trusted, this question can be answered by standard quantum channel authentication [6], and there are various approaches to this end, from those requiring incredibly expensive entangled resources [6–8], to those more achievable, but at cost to security scaling [9–12]. In this work, we consider a much stronger requirement, where some or all devices used are not trusted, in a so-called device independent setting. This will be a crucial step for testing the transmission through quantum channels for future applications.

Device independence uses Bell-like correlations to imply correct behaviour of quantum hardware, without the need to understand or trust their inner workings [13, 14], that is, independently of the physical device used. It is motivated by the inevitable situation where the user of a quantum technology is not necessarily the one who built all the hardware and does not necessarily want to trust it to behave as specified. It has first been applied in quantum information to prove security in quantum key distribution devices, thus making them secure against potential hardware hacks. It has then expanded in many directions, including random number generation [15], verification of quantum computation [16], and more [17, 18]. The application to quantum channels is relatively recent [19] (but see also [20]), however there are some important missing elements in order to obtain useful certification. Measurement-device-independent approaches have been successfully demonstrated [21–23], but these do not directly quantify the channel quality and its ability to faithfully transmit arbitrary quantum information.

Here, we address the main remaining obstacles to certify the transmission of quantum information in the device independent framework. First, in our approach we explicitly take into account loss. This is particularly important in optical implementations (which is the most natural choice for quantum channels). It is not addressed in current schemes[19, 20], which effectively assume that any loss is innocent; this is somewhat against the goals of device independence and opens a security loophole if the loss is controlled by malicious parties. Second, we remove the assumption that each time a channel is used, it is done so in an independent, uncorrelated way, known as identical independent distribution (IID). This assumption similarly makes us vulnerable in terms of security so should be avoided in general. Third, we certify the transmission of quantum information itself. Previous works assume IID, and they certify that the channel that was used during the test was good but without a statement on actual transmitted quantum information [19, 24]. We develop the treatment of loss as a non trace preserving channel, bounding the diamond fidelity between an untrusted channel and an ideal one. We use this to build protocols certifying a transmitted quantum message using this channel. Our protocols are secure in the one-sided device independent setting (where the sender's devices are fully trusted, but not the receiver's), and also in the fully device independent setting when IID is assumed on the source; in both cases no IID needs to be assumed on the uses of the channel.

We also demonstrate the feasibility of our protocol and experimentally validate the main elements of one-sided device independent certified transmission with an implementation exploiting a high-quality entangled photon source with polarization encoding obtained in a Sagnac configuration. This allows us to explore the behavior of the minimum fidelity that we can certify for realistic losses in honest channels and confirm the robustness of the protocol against simulated errors introduced by dishonest channels.

Results

Certification protocol. In our framework, a player Alice wishes to send a qubit state from Hilbert space $\mathcal{H}_i$ to Bob, through a local unitary quantum channel $\mathcal{E}_0$. This *quantum message* is possibly entangled with another system of Hilbert space $\mathcal{S}$ of arbitrary dimension, so the global state reads $\rho_i \in \mathcal{L}(\mathcal{H}_i \otimes \mathcal{S})$. The channel takes any qubit from $\mathcal{L}(\mathcal{H}_i)$ to another qubit from $\mathcal{L}(\mathcal{H}_o)$, the output qubit space, with output global state $\rho_o = (\mathcal{E}_0 \otimes \mathbb{I})[\rho_i] = (U \otimes \mathbb{I})\rho_i(U^\dagger \otimes \mathbb{I})$, where U is a local unitary and $\mathbb{I}$ is the identity. This model describes a perfect unitary gate in a quantum computer, quantum transmission link (carried on through quantum teleportation or a simple optical fiber) or quantum memory. Without loss of generality, we take $U = \mathbb{I}$ and $(\mathcal{E}_0 \otimes \mathbb{I})[\rho_i] = \rho_i$, as this case encompasses all unitaries in a device independent scenario [19]. This channel is called the *reference channel*.

In real world situations, the channel would be lossy, noisy, or even operated by a malicious party Eve. Also, Alice and Bob normally do not have access to isolated qubit spaces, but operate with physical systems such as photons or atoms, displaying other degrees of freedom. This way, without further assumptions, Alice and Bob have access to a completely positive trace-decreasing (CPTD) map $\mathcal{E}$, *i.e.* a probabilistic channel, that sends density operators from an input Hilbert space $\mathcal{H}_{\mathcal{A}_1}$ to positive operators of trace smaller than 1 on an output Hilbert space $\mathcal{H}_{\mathcal{B}}$. This channel is called the *physical channel*. Alice also possesses a source of bipartite states Φ_i shared between $\mathcal{H}_{\mathcal{A}_1}$ and a secondary Hilbert space $\mathcal{H}_{\mathcal{A}_2}$, that we call the *probe* input state. She can send one part of Φ_i through the channel $\mathcal{E}$, resulting in the probe output state Φ_o, shared with Bob:

$$\Phi_o = (\mathcal{E} \otimes \mathbb{I})[\Phi_i]/t(\mathcal{E}|\Phi_i), \tag{1}$$

where $t(\mathcal{E}|\Phi_i) = \mathrm{Tr}(\mathcal{E} \otimes \mathbb{I})[\Phi_i]$ is the *transmissivity* of $\mathcal{E}$ which *a priori* depends on the input state, as it does in polarizing channels for instance. Note that the transmissivity can be defined for any given quantum channel, and thus no assumption is made to formalize it. For more details on this relatively new notion, the reader can refer to SUPP. MAT. A [25]. Finally, the players can measure states with 2-outcome positive operator-valued measures (POVMs) $\{M^{\mathcal{P}}_{l|q}\}_{l=0,1}$ where $\mathcal{P} = \mathcal{A}_1, \mathcal{A}_2$ or $\mathcal{B}$ indicating the Hilbert space on which the measurement is acting, and q indicates which POVM is measured, see Eqs. (9) to (12) below. Fig. 1 illustrates our setting.

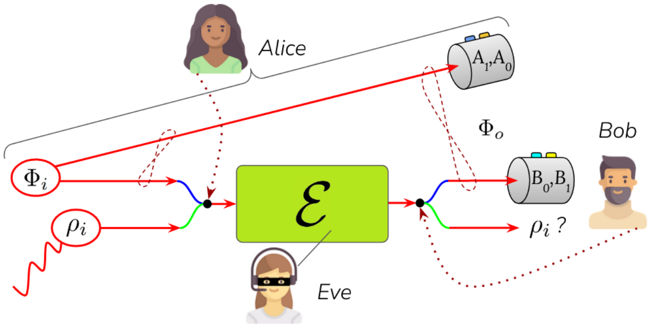

FIG. 1: Sketch of the problem. Alice's goal is to send a qubit, potentially part of a larger system, in state ρ_i, through an untrusted quantum channel $\mathcal{E}$ (green path). To do so, she sometimes tests the channel by sending half an entangled state (blue path). Alice and Bob can then measure the output state Φ_o, to assess how close the action of the physical channel $\mathcal{E}$ is to an ideal reference channel $\mathcal{E}_0$ on the transmitted state ρ_i.

In an adversarial scenario, Alice and Bob wish to draw device independent conclusions, meaning they make no assumption whatsoever on the states or the measurements. In particular, physical Hilbert spaces are of arbitrarily big dimensions, which include all degrees of freedom of the physical systems and possible entanglement with the rest of the universe. In this way, players can only certify objects up to local isometries, which associate finite-dimension qubit spaces $\mathcal{H}_i$ and $\mathcal{H}_o$, to these infinite-dimension physical spaces $\mathcal{H}_{\mathcal{A}_1}$, $\mathcal{H}_{\mathcal{A}_2}$, $\mathcal{H}_{\mathcal{B}}$. As a device independent procedure, self-testing is actually "blind" to local isometries such that it does not certify a single state, but a whole equivalence class of quantum states mutually related by locally isometric transformations. As shown in [19], similar conclusions can be drawn in order to device-independently test the equivalence between the physical channel $\mathcal{E} \otimes \mathbb{I}$ and the reference operation $\mathcal{E}_0 \otimes \mathbb{I}$. Note, however, that as a quantum channel is associated to two Hilbert spaces (one in input and the other in output), two isometries are involved in order to extract a qubit-to-qubit channel from a physical channel. This way, the input isometry brings a qubit input state to a physical state that can be fed into the physical channel, while the output isometry extracts a qubit state from the physical channel's output state. However, this formalism, in principle, only applies to completely positive trace-preserving (CPTP) maps. In our case, a trace-decreasing physical channel only returns a state with a certain probability, such that it can only be compared to the reference channel multiplied by a constant $t \leq 1$. Then, one can only make a statement about equivalence between the physical and reference channels, when considering rounds in which the transmission was successful. We capture this intuition with the following definition.

Definition 1 (Self-testing of a CPTD map)**.** *Let us consider a physical channel $\mathcal{E} : \mathcal{H}_{\mathcal{A}_1} \longrightarrow \mathcal{H}_{\mathcal{B}}$. With two local isometries $\Gamma_i : \mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_i \longrightarrow \mathcal{H}_{\mathcal{A}_1} \otimes \mathcal{H}_i^{ext}$ (encoding map) and $\Gamma_o : \mathcal{H}_{\mathcal{B}} \longrightarrow \mathcal{H}_o \otimes \mathcal{H}_o^{ext}$ (decoding map), and an ancillary state $\rho_{\mathcal{A}_1} \in \mathcal{L}(\mathcal{H}_{\mathcal{A}_1})$, we can define an extracted qubit channel $\mathcal{E}_{i,o}$ as:*

$$\mathcal{E}_{i,o} : \rho \in \mathcal{L}(\mathcal{H}_i) \longrightarrow Tr_{ext}\big((\Gamma_o \circ \mathcal{E} \circ \Gamma_i)[\rho_{\mathcal{A}_1} \otimes \rho\,]\big), \quad (2)$$

where the trace is taken over $\mathcal{H}_i^{ext}$ and $\mathcal{H}_o^{ext}$ [26]. The self-testing equivalence between a probabilistic channel $\mathcal{E}$ and the reference channel $\mathcal{E}_0$ is established if there exists $t\ \in]0;1]$ giving:

$$\mathcal{E}_{i,o} = t\mathcal{E}_0. \quad (3)$$

Note that we exclude the value $t = 0$, otherwise the null quantum channel, that never outputs any quantum state whatever the input, would be equivalent to all quantum channels. In addition, we highlight that the transmissivity is invariable on the encoding and decoding maps chosen in the definition. The reader can refer to SUPP. MAT. A 2 for more details on the lossy channels' equivalence classes. In experiments, we can never perfectly certify $\mathcal{E}$, therefore we quantify the ability of this probabilistic channel to implement the deterministic channel $\mathcal{E}_0$ by generalizing the diamond fidelity to probabilistic quantum channels:

$$\begin{aligned}\mathcal{F}_\diamond^{\Gamma_{i,o}}(\mathcal{E},\mathcal{E}_0) &= \mathcal{F}_\diamond(\mathcal{E}_{i,o},\mathcal{E}_0)\\ &= \inf_{|\phi\rangle} F((\mathcal{E}_{i,o}\otimes\mathbb{I})[\phi]/t(\mathcal{E}_{i,o}|\phi),(\mathcal{E}_0\otimes\mathbb{I})[\phi]),\end{aligned} \quad (4)$$

where $F(\rho,\sigma) = \mathrm{Tr}\big(\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\big)^2$ is the Ulhmann fidelity for quantum states, and the lower bound is taken over all pure states $|\phi\rangle$ from $\mathcal{H}_i^{\otimes 2}$ such that $t(\mathcal{E}_1|\phi) \neq 0$ and $t(\mathcal{E}_2|\phi) \neq 0$. Note that the left state is normalized by the transmissivity. Consequently, contrary to CPTP maps fidelities, $\mathcal{F}_\diamond(\mathcal{E}_{i,o},\mathcal{E}_0) = 1$ does not imply $\mathcal{E}_{i,o} = \mathcal{E}_0$, but only that there exists $t \in]0,1]$ such that $\mathcal{E}_{i,o} = t\mathcal{E}_0$, meaning that the channels are equivalent in the sense of our definition. Physically speaking, these two channels output the same states, under the condition those were not lost. The diamond fidelity is particularly useful here, as it can be interpreted as the minimum probability that $\mathcal{E}\otimes\mathbb{I}$ successfully implements the operation $\mathcal{E}_0\otimes\mathbb{I}$ on any state, under the condition that a state successfully passes through the channel. The main goal of our protocol is therefore to certify that fidelity.

For that purpose, let us consider the situation where Alice can certify the probe input state Φ_i up to two local isometries $\Gamma^{\mathcal{A}_1/\mathcal{A}_2} : \mathcal{H}_{\mathcal{A}_1/\mathcal{A}_2} \longrightarrow \mathcal{H}_{\mathcal{A}_1/\mathcal{A}_2} \otimes \mathcal{H}_i$ with the following fidelity to a maximally entangled state:

$$F^i = F\big((\Lambda^{\mathcal{A}_1}\otimes\Lambda^{\mathcal{A}_2})[\Phi_i],\Phi_+\big), \quad (5)$$

where Φ_+ is a maximally-entangled state (for instance $|\Phi_+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$) and $\Lambda^j[\cdot] = \mathrm{Tr}_j(\Gamma^j[\cdot])$. We next consider the situation that Alice and Bob are able to certify the probe output state Φ_o up to local isometries $\Gamma^{\mathcal{A}_2}$ and $\Gamma^{\mathcal{B}} : \mathcal{H}_{\mathcal{B}} \longrightarrow \mathcal{H}_{\mathcal{B}} \otimes \mathcal{H}_o$ with the following fidelity:

$$F^o = F\big((\Lambda^{\mathcal{B}}\otimes\Lambda^{\mathcal{A}_2})[(\mathcal{E}\otimes\mathbb{I})[\Phi_i]]/t(\mathcal{E}|\Phi_i),(\mathcal{E}_0\otimes\mathbb{I})[\Phi_+]\big). \quad (6)$$

Given Eqs. (5) and (6), we show in SUPP. MAT. D 2 that there exist isometries Γ_i, Γ_o such that Alice and Bob are able to lower bound the diamond fidelity on the corresponding extracted channel $\mathcal{E}_{i,o}$:

$$\mathcal{F}_\diamond(\mathcal{E}_{i,o},\mathcal{E}_0) \geq 1-4\sin^2\Big(\arcsin\big(C^i/t(\mathcal{E}|\Phi_i)\big)+\arcsin C^o\Big), \quad (7)$$

where $C^j = \sqrt{1-F^j}$ are sine distances associated to their corresponding fidelities [27]. In this way, checking the input and output fidelities allows us to assess the fidelity of the channel itself. This bound generalizes what is shown in [19] to probabilistic channels. It also uses the diamond fidelity, which informs on the behavior of the channel on any state, instead of the Choi-Jamiołkowski fidelity, which only informs on the behavior of the channel on a maximally entangled state.

This bound gives the direction for estimating the fidelity of a quantum channel. The idea is to evaluate the fidelity F^i of the probe input state to a Bell state, then send one part of that probe state through the channel Alice wishes to send ρ_i through, and finally evaluate the fidelity F^o of the corresponding output state to the same Bell state. Such procedure is possible using recent self-testing results [28], but requires a very large number of experimental rounds in the absence of the IID assumption, as both input and output probe states require certification. We significantly decrease that number by making the IID assumption on the probe state, or by leaving its full characterization to Alice's responsibility. Still, as we make no IID assumption on the channel, optimal security cannot be reached by first testing that channel, and only then using it to send the message state ρ_i, as Eve may change the channel's expression in the last moment. Our protocol works around this problem by allowing Alice to hide the message ρ_i among a large number of probe states, at a random position unknown to Eve. In that case, we show in SUPP. MAT. D 5 that the bound (7) holds for the average channel $\bar{\mathcal{E}}_{i,o}$ over the whole protocol. Then the *transmission fidelity* between the output quantum message $\bar{\rho}_o = (\bar{\mathcal{E}}_{i,o}\otimes\mathbb{I})[\rho_i]/t(\bar{\mathcal{E}}|\rho_i)$ and the input quantum message ρ_i is certified:

$$F(\rho_i,\bar{\rho}_o) \geq \mathcal{F}_\diamond(\bar{\mathcal{E}}_{i,o},\mathbb{I}). \quad (8)$$

As long as the message's position among the probe states remains hidden, we can use $\bar{\rho}_o$ to describe accurately any statistics that would occur when processing the output state of the protocol, and estimate the quality of an actual transmitted state, instead of a verification of a channel only (see SUPP. MAT. D 1 for more details).

In SUPP. MAT. C we give detailed protocols where we apply these ideas to test a transmitted quantum message under the device independent (DI) and one-sided device independent (1sDI) scenarios. In Table I, we provide a detailed comparison of the level of trust put in the 1sDI and full-DI scenarios.

Protocol Elements	1sDI	DI
Quantum Channel	Untrusted, Non-IID	Untrusted, Non-IID
Alice's Source	Trusted, IID, Characterized	Untrusted, IID
Alice's Measurement Apparatus	Trusted, Characterized	Uncharacterized
Bob's Measurement Apparatus	Uncharacterized	Uncharacterized
Classical Communication Channel	Trusted, Private	Trusted, Private

TABLE I: Summary of the assumptions made on the elements of the protocol, in the ideal 1sDI and DI scenarii. Additional assumptions are made in the experimental implementations, which are summarized in Methods. Fair-sampling assumptions are made on the measurement apparatuses which are detailes in SUPP. MAT. E 3.

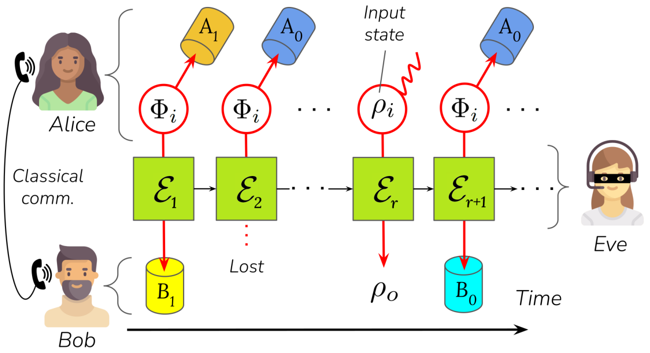

FIG. 2: Protocol sketch in a one-sided device independent scenario: Alice prepares N copies of the probe state Φ_i, and sends them through the untrusted channel $\mathcal{E}$ that varies with time, as well as ρ_i at a random secret position r. Some states are lost such that Bob only receives a fraction of them. Alice tells Bob the value of r. If ρ_i was lost, then the protocol aborts. Otherwise, Bob stores ρ_i and, together with Alice, tests the violation of the steering inequality with the output probe states. They deduce the average channel's quality over the protocol, which informs on the probability that the message ρ_i was accurately transmitted to Bob, up to isometries.

For the purpose of our demonstration, we focus on an one-sided device independent scenario. A summary of the protocol in this case is given in Fig. 2 (for a detailed recipe, the reader can refer to the Supplementary Material). Here, Alice's measurement setup is trusted, such that her Hilbert spaces are qubit spaces $\mathcal{H}_{\mathcal{A}_1} = \mathcal{H}_{\mathcal{A}_2} = \mathcal{H}_i$, her isometries are trivial $\Gamma_i = \Gamma^{\mathcal{A}_1} = \Gamma^{\mathcal{A}_2} = \mathbb{I}$, and she performs measurements in the Pauli X and Z bases:

$$A_0 = M^{\mathcal{A}_2}_{0|0} - M^{\mathcal{A}_2}_{1|0} = Z, \tag{9}$$

$$A_1 = M^{\mathcal{A}_2}_{0|1} - M^{\mathcal{A}_2}_{1|1} = X. \tag{10}$$

This fits a variety of scenarios where Alice is a powerful server, trying to provide states to a weaker client, Bob, whose measurement apparatus is still untrusted. For that reason, Bob's observables, defined as:

$$B_0 = M^{\mathcal{B}}_{0|0} - M^{\mathcal{B}}_{1|0}, \tag{11}$$

$$B_1 = M^{\mathcal{B}}_{0|1} - M^{\mathcal{B}}_{1|1}, \tag{12}$$

are *a priori* unknown. We refer to such measurements as *uncharacterized*. In order to bound F^o, Alice and Bob use self-testing through steering [29]. Namely, the maximal violation of the *steering* inequality [30]:

$$\beta = |\langle A_0 B_0\rangle + \langle A_1 B_1\rangle| \leq \sqrt{2}, \tag{13}$$

self-tests the maximally entangled pair of qubits. We then combine recent self-testing results [28] with further finite statistics methods in a non-IID setting and with a lossy channel, in order to estimate F^o in bound (7) with high confidence, when a close-to-maximal violation $\beta = 2 - \epsilon$ is measured:

$$F^o \geq 1 - \alpha f(\epsilon, K) \simeq 1 - \alpha\epsilon, \tag{14}$$

with f a function of ϵ and the number K of states measured by Alice and Bob during the protocol (see Eq. (27) in Methods), and $\alpha = 1.26$ [28]. This outlines the protocol: by sending N characterized probe states through the channel, Alice and Bob estimate F_o and thus the diamond fidelity between the extracted channel and the identity channel, and therefore the transmission fidelity of an unknown state ρ_i, as a function of N, ϵ, and the number K of transmitted states.

Experimental implementation. In order to test the feasibility of our protocol, we perform a proof-of-principle experiment based on photon pairs, emitted at telecom wavelength via type-II spontaneous parametric down-conversion (SPDC) in a periodically-poled KTP crystal (ppKTP). Photons are entangled in polarization thanks to a Sagnac interferometer [31], encoding in this way a close-to-maximally entangled pair of qubits. Details of the setup are given in Fig. 3. [32]

The states emitted by the source are characterized at each iteration of the protocol via quantum state tomography [33], without inserting any untrusted quantum channel (green box in Fig. 3). Polarization analyzers (PA) are trusted for that task, as it is performed by Alice. This way we measured a fidelity of the probe's polarization state to a Bell state of $\overline{F^i} = 99.19\% \pm 0.03\%$ on average over all protocol attempts, with a maximum reached fidelity of $F^i = 99.43\% \pm 0.05\%$. We then send the probe states through an untrusted quantum channel. For this first demonstration we use a variable optical attenuator (VOA) in order to simulate a lossy but honest

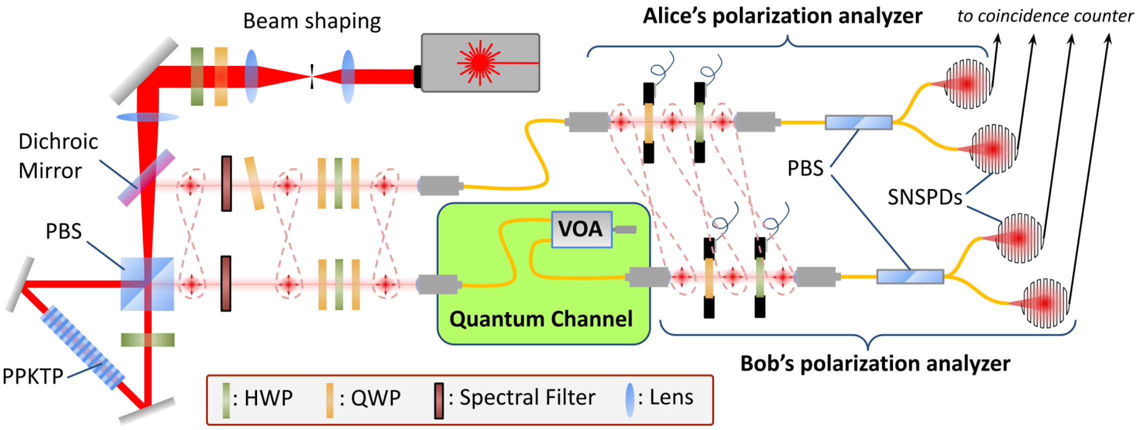

FIG. 3: Experimental setup for photonic certified quantum communication through an unstrusted channel. Photon pairs are generated via type-II SPDC, in a ppKTP crystal ($30\,\mathrm{mm}$-long, $46.2\,\mu\mathrm{m}$ poling period), and entangled in polarization in a Sagnac interferometer. The source is pumped with a $770\,\mathrm{nm}$ continuous laser. Signal and idler photons are emitted around $1540\,\mathrm{nm}$, separated from the pump by a dichroic mirror, and from each other by the polarizing beam splitter (PBS) of the interferometer. They are then coupled into single-mode fibers, and sent to the different players. The idler photon is both used as Alice's part of the maximally-entangled pair and to herald the probe state. The signal photon is sent to Bob through the untrusted lossy channel. A variable optical attenuator (VOA) allows to simulate an honest channel with a tunable amount of loss. The biphoton state is measured with polarization analyzers, each made of two waveplates (WPs), a fibered PBS, and $> 80\%$-efficiency Superconducting Nanowire Single-Photon Detectors (SNSPDs). The WPs are mounted on motorized stages, allowing to both regularly randomize the measurement basis and implement dishonest channels. Detection events are then sent to a fast coincidence counter which gathers all the data required in order to evaluate the quantum correlations and channel's transmissivity.

channel that requires certification. Detecting an idler photon in Alice's PA heralds a signal photon being sent through the quantum channel, which is then detected in Bob's PA. In each protocol attempt, the transmissivity is identified as the probability that Bob detects a state, knowing Alice heralded that state, and is also known as the heralding efficiency η_s:

$$t(\mathcal{E}|\Phi_i) \simeq \eta_s = R_{si}/R_i, \tag{15}$$

where R_{si} is the pair detection rate and R_i the idler detection rate. We measure the pairs in random bases A_0B_0 or A_1B_1, and evaluate a close-to-maximum violation of steering inequality $\beta = 2-\epsilon$, with an average deviation $\bar{\epsilon} = 1.42 \cdot 10^{-2}$ over all protocols, a minimum deviation measured in a protocol attempt $\epsilon_{\text{min}} = 1.34 \cdot 10^{-2}$, and a maximum deviation $\epsilon_{\text{max}} = 1.48 \cdot 10^{-2}$.

For each protocol attempt we set a different transmissivity of the VOA, such that η_s ranges from 21.9% to 47.3%, the maximum value corresponding to the replacement of the VOA by a simple fiber connector. Following the 1sDI setting, Alice trusts her devices, so we are allowed to take losses originating from her equipment as trusted. However, the experimental set up makes it difficult to distinguish between the source of losses. To allow for all cases we consider that a certain fraction of the losses is not induced by the channel itself, but by other components which are characterized by Alice, as part of the source. These act as an unbiased filter, which losses are considered homogeneous and trusted, so the overall channel reads

$$\mathcal{E} = (1-\lambda_c)\mathcal{E}', \tag{16}$$

with λ_c the amount of losses that is trusted and state-independent, and $\mathcal{E}'$ a quantum channel that is strictly equivalent to $\mathcal{E}$ by definition, and therefore returns the same output states; see Fig. 4. In that case we can certify $\mathcal{E}'$ instead of $\mathcal{E}$, and evaluate the transmissivity in bound (7) as

$$t(\mathcal{E}'|\Phi_i) = t(\mathcal{E}|\Phi_i)/(1-\lambda_c) = \eta_s/(1-\lambda_c). \tag{17}$$

This tightens the bound compared to the naive approach where all losses are attributed to the channel. Adopting this interpretation is quite realistic, considering that Alice preforms a full characterization of the probe states, which potentially includes a lower bound on the coupling losses. In the most paranoid scenario, we can always set $\lambda_c = 0$ we attribute all loss (including Alice's coupling and detection losses) to the quantum channel.

We show the results of our implementations in Fig. 5. Thanks to our close-to-maximum violation of steering inequality and relatively high coupling efficiency, we are able to certify the transmission of an unknown qubit state through the untrusted channel, with a non-trivial transmission fidelity $F(\rho_i, \rho_o) > 50\%$. This is true even when Alice attributes all losses to the channel, *i.e.* $\lambda_c = 0$, for channels with the highest transmissivities. The certified fidelity increases as Alice trusts a larger amount of homogeneous losses λ_c, reaching

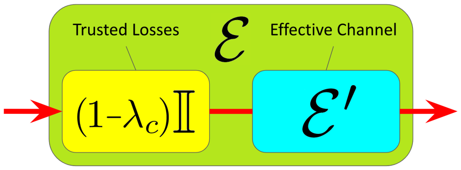

FIG. 4: Schematic decomposition of the untrusted channel $\mathcal{E}$, into an equivalent channel $\mathcal{E}'$ that the protocol effectively certifies, and a trusted channel, corresponding to the characterized and homogeneous losses λ_c trusted by Alice.

$F(\rho_i, \rho_o) \geq 77.1\% \pm 0.6\%$ when she assumes a maximum value $\lambda_c = 0.526$ and the channel is close to lossless. In any case, the certified fidelity decreases as the channel gets more lossy, as a direct consequence of bound (7), highlighting the difficulties of certifying lossy channels. This gives further motivation to assume that a fraction of the losses is trusted, in order to certify, for example, long-distance quantum communications. In our implementation, assuming maximum trusted losses $\lambda_c = 0.526$, we could certify a non-trivial transmission fidelity $F(\rho_i, \rho_o) > 50\%$, for total transmissivities as low as $t(\mathcal{E}|\Phi_i) = \eta_s \simeq 0.263$, while such certification was possible only for $\eta_s \gtrsim 0.44$ with no trusted losses $\lambda_c = 0$.

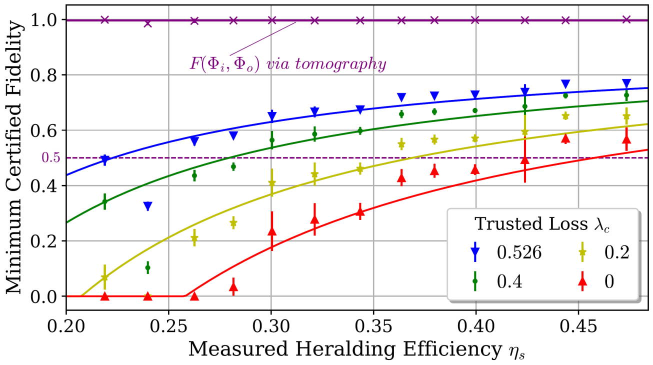

FIG. 5: Minimum fidelity $F(\rho_i, \rho_o)$ certified via our protocol as a function of the measured heralding efficiency, tuned with a VOA, and for different trusted losses λ_c (colored curves). The curves are plotted by taking the average fidelity of the probe state to a Bell state $\overline{F_i}$, and the average of the deviation from maximum violation ϵ, over all protocol attempts. Experimental results deviate from these curves, as F^i and ϵ vary between experiments. Errors induced by the finite statistics are directly subtracted from the certified fidelity, as detailed in Methods (see Eqs. (28) and (29) in particular). Error bars include effects induced by the unbalance in detectors' efficiency and the propagation of errors on F^i. We also display the fidelity $F(\rho_i, \rho_o)$ measured via quantum state tomography, for $\rho_i = \Phi_i$.

In order to fully demonstrate the protocol, one should send a single quantum message ρ_i through the channel, hidden among the probe states. The value of that state does not matter in our implementation as we do not use it in a later protocol, so we choose $\rho_i = \Phi_i$ and consider that a random copy of the probe state is actually the quantum message. To check the correctness of our protocol, we fully characterize the corresponding transmitted message ρ_o after the channel, by quantum state tomography. Note that this procedure is done after the protocol ends, and in the same experimental conditions, though placing full trust in our laboratory. Hence this could not be performed inside an untrusted quantum network. The full trust placed in our measurement apparatus enables a more precise evaluation of the transmission fidelity, which we thus estimate at $F(\rho_i, \rho_o) = 99.64\% \pm 0.08\%$ on average over all protocol attempts, with a minimum value of $F(\rho_i, \rho_o) = 98.6\% \pm 0.6\%$. As displayed on Fig. 5, this is far higher than the values certified by our protocol (due to the trust added during the quantum state tomography), which shows the state was indeed properly transmitted. Note that, in this case, the channel and measurement stations are trusted during the tomography of ρ_o, as it is performed outside of the protocol. This allows us to measure numerous copies of ρ_o, which is necessary for a full characterization of the state. In order to show that the correctness of our certification protocol would hold for other quantum messages ρ_i, we perform a full-process tomography of the quantum channel [34], and lower-bound the fidelity between the physical channel and the identity $\mathcal{F}_\diamond(\mathcal{E}, \mathbb{I}) \geq 94\% \pm 3\%$. We expect this bound to be far from tight, as it is evaluated using the equivalence between diamond and Choi-Jamiołkowski distances [35] (see Lemma 2 in Methods). Still, the fidelity is greatly above the values certified by our protocol, showing the certification procedure is indeed valid for any quantum message ρ_i.

The resilience of the protocol is further shown by experimentally simulating examples of dishonest channels. Let us first recall that the operator of the channel has no information on the position of the quantum message ρ_i before the end of the protocol. This way, a typical attack consists in applying a disruptive transformation with small probability, hoping it will be applied to ρ_i and stay undetected by Alice and Bob. Here we consider such a transformation to be a bit flip and/or a phase flip. For this experimental demonstration, we remove the VOA and consider that all losses are trusted. Note that performing a phase flip is equivalent to turning Bob's first measurement B_0 into $-B_0$:

$$B_0 = M^{\mathcal{B}}_{0|0} - M^{\mathcal{B}}_{1|0} \longrightarrow -B_0 = M^{\mathcal{B}}_{1|0} - M^{\mathcal{B}}_{0|0}. \tag{18}$$

Similarly, a bit flip is equivalent to turning Bob's second measurement B_1 into $-B_1$. Thus, we perform these flips in practice by randomly changing the waveplate angles in order to get the opposite measurement bases. This simulates dishonest channels of the form:

$$\begin{aligned} \mathcal{E}_{p,q}[\rho] = (1-p)(1-q)\rho + p(1-q)X\rho X \\ + pqY\rho Y + (1-p)qZ\rho Z, \end{aligned} \tag{19}$$

with p the bit flip probability and q the phase flip probability.

The certification results are displayed in Fig. 6, for different bit and phase flip probabilities. These show that

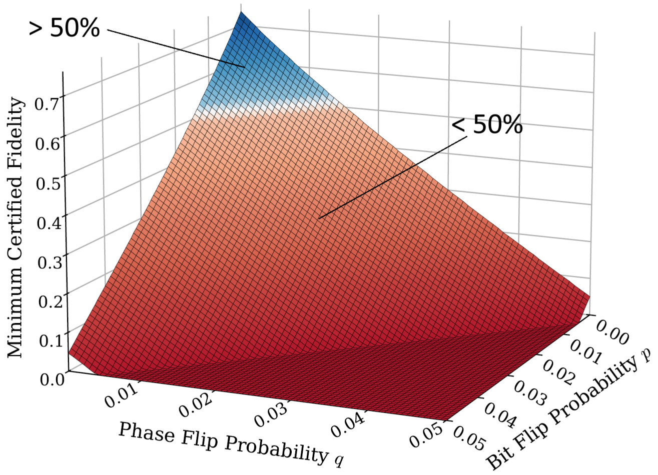

FIG. 6: Minimum fidelity $F(\rho_i,\rho_o)$ certified via our protocol, for malicious channels $\mathcal{E}_{p,q}$, where p is the probability of applying gate X and q is the probability of applying gate Z. Here we measured a probe state fidelity to a Bell state of $F^i = 99.12\% \pm 0.1\%$, and we trust a maximum amount of losses $\lambda_c = 0.526$.

our implementation is quite sensitive to these attacks, such that a flip probability of 0.01 induces a collapse of 16% of the certified fidelity, and we only certify $F(\rho_i,\rho_o) \geq 58\%$. The certified fidelity falls below the trivial value 50% for flip probabilities as low as 0.016. In this way, any attempt of Eve to disrupt the input state ρ_i with such a method can only succeed with very small probabilities $p, q < 0.02$, or it will be detected by Alice and Bob.

Discussion

In this work, we have provided a protocol to certify the transmission of a qubit through an untrusted and lossy quantum channel, by probing the latter with close-to-maximally entangled states and witnessing non-classical correlations at its output. In the DI case these are Bell correlations, in the 1sDI they are steering correlations. Our theoretical investigations rely only on assumptions made on the probe state's source and the sender's measurement apparatus (in the case of 1sDI), while relaxing assumptions made on the quantum channel and the receiver's measurement apparatus. This setting proves to be an interesting trade-off between realistic experimental conditions and reasonable cryptographic requirements. It also embodies a practical scenario in which a strong server provides a weaker receiver with a quantum bit.

Compared to previously proposed verification procedures, our protocol not only certifies the probed channels, but also an unmeasured channel through which a single unknown state can be sent. As quantum measurements deteriorate the quantum states, this task can only be performed at the price of measuring a large amount of probe states, which limits the repeatability of the protocol. With our experimental parameters, certifying the transmission of a single qubit thus takes 1 h to 2 h, depending on the channel's transmissivity. In these conditions the protocol can still serve as a practical primitive for other single-shot protocols that require a single quantum state, such as the recently demonstrated quantum weak coin-flipping [36, 37]. Additionally, we show in SUPP. MAT. E 5 that the duration of the protocol can be reduced to a few seconds, by increasing the probe states' quality and emission rate within attainable performances for current technologies. Further scientific advancements such as the development of multiplexed photon-pair sources [32, 38–41], or the emergence of bright and deterministic single-photon sources [42–44], could improve our protocol's repeatability by drastically increasing the entangled photon-pair generation rate.

Our proof-of-principle implementation shows the correctness of this certification procedure, and its feasibility with current technology. This way we could certify non-trivial transmission fidelities for a wide range of losses induced by the channel, by making some mild but realistic assumptions, such as the characterization of a fraction of trusted losses, induced for instance by the coupling of probe states inside optical fibers. By implementing random bit and phase flips, we could show that even a small probability attempt to disrupt the quantum information degrades the certified transmission fidelity, and is therefore detected by the players.

Future developments could demonstrate the feasibility of a fully device independent version of our protocol, in which Alice's measurement or even the probe states' source are not trusted. Such a protocol could be achieved by linking the probe state quality to that of the corresponding output state, or by making the IID assumption on the probe state's source. Also, more investigation on quantum-memory-based attacks could give a sharper idea on the possibilities of deceiving the certification procedure.

Our work opens the way to certification of a wide variety of more sophisticated lossy quantum channels. In particular, the rapid improvements of quantum technologies could soon provide possible applications of this protocol to the authentication of quantum teleportation, memories or repeaters. The development of the latter could also temper the losses in large-scale quantum transmission links, making our protocol particularly suited for the certification of long-distance quantum communications [45–47].

Methods

Two Useful Lemmas. The proof of bound (7) relies on two lemmas, which give fundamental results on lossy quantum channels, and that we provide here.

Lemma 1 (Extended Processing Inequality)**.** *For any probabilistic channel $\mathcal{E}$ (CPTD), and any input states ρ_i and σ_i, the following inequality holds for the sine distance $C(\rho,\sigma) = \sqrt{1-F(\rho,\sigma)}$:*

$$C(\rho_i,\sigma_i) \geq t \cdot C(\rho_o,\sigma_o), \tag{20}$$

where $\rho_o = \mathcal{E}[\rho_i]/t(\mathcal{E}|\rho_i)$ and $\sigma_o = \mathcal{E}[\sigma_i]/t(\mathcal{E}|\sigma_i)$ are the output states of the channel, and $t = t(\mathcal{E}|\rho_i)$ or $t = t(\mathcal{E}|\sigma_i)$.

This first lemma generalizes to CPTD maps the well-known fidelity processing inequality $F(\rho,\sigma) \leq F(\mathcal{E}[\rho],\mathcal{E}[\sigma])$, which holds for any CPTP map $\mathcal{E}$.

Lemma 2 (Channel's Metrics Equivalence)**.** *For any probabilistic channel $\mathcal{E}_1$, and any $\mathcal{E}_2$ that is proportional to a deterministic channel (CPTP map), both acting on $\mathcal{L}(\mathcal{H}_i)$, we have the following inequalities:*

$$\mathcal{C}_J(\mathcal{E}_1,\mathcal{E}_2) \leq \mathcal{C}_\diamond(\mathcal{E}_1,\mathcal{E}_2) \leq \dim\mathcal{H}_i \times \mathcal{C}_J(\mathcal{E}_1,\mathcal{E}_2), \tag{21}$$

where the $\mathcal{C}_J$, resp. $\mathcal{C}_\diamond$, are the Choi-Jamiołkowski, resp. diamond, sine distances of probabilistic quantum channels:

$$\mathcal{C}_J(\mathcal{E}_1,\mathcal{E}_2) = C\Big(\frac{(\mathcal{E}_1\otimes\mathbb{I})[\Phi_+]}{t(\mathcal{E}_1|\Phi_+)},(\mathcal{E}_2\otimes\mathbb{I})[\Phi_+]\Big), \tag{22}$$

$$\mathcal{C}_\diamond(\mathcal{E}_1,\mathcal{E}_2) = \sup_{|\phi\rangle} C\Big(\frac{(\mathcal{E}_1\otimes\mathbb{I})[\phi]}{t(\mathcal{E}_1|\phi)},(\mathcal{E}_2\otimes\mathbb{I})[\phi]\Big). \tag{23}$$

This lemma shows the equivalence between Choi-Jamiołkowski and diamond distances, which is fundamental when trying to link the behaviour of the channel on a maximally-entangled state, to its behaviour on any quantum state. We also use this lemma in order to bound the diamond fidelity after performing a full process tomography of the channel, by evaluating the more straightforward Choi-Jamiołkowski fidelity.

Note that both these lemmas also apply to the trace distance $D(\rho,\sigma) = \frac{1}{2}\mathrm{Tr}|\rho-\sigma|$, and are proven in SUPP. MAT. B 1 and B 2.

Protocol Security. In our protocol, the quantum channel is allowed to evolve through time, with some potential memory of the experiment's past history. This way we define the channel $\mathcal{E}_{k|[k-1]}$, where $[k-1] = k-1, k-2, ..., 1$, that operates on the k-th state sent by Alice through the protocol. In particular, Alice sends the quantum message ρ_i at a random position r through channel $\mathcal{E}_{r|[r-1]}$. We then define the expected channel over the protocol:

$$\bar{\mathcal{E}} = \frac{1}{N+1}\sum_{k=1}^{N+1}\mathcal{E}_{k|[k-1]}. \tag{24}$$

As ρ_i is sent at a random position that stayed concealed from the channel's operator, the expected transmitted message is $\bar{\rho}_o = (\bar{\mathcal{E}}\otimes\mathbb{I})[\rho_i]/t(\bar{\mathcal{E}}|\rho_i)$. As long as r stays hidden and random, any measurement performed on the transmitted message later after the protocol would follow the same statistics as if it was performed on $\bar{\rho}_o$ (see SUPP. MAT. D 1 for more details). This way, we derive the protocol security by applying bound (7) to the average channel $\bar{\mathcal{E}}$, in order to bound the fidelity of $\bar{\rho}_o$ to ρ_i, up to isometry. In particular, the output probe state fidelity to a maximally entangled state now reads

$$F^o = F\big((\Lambda^{\mathcal{B}}\otimes\Lambda^{\mathcal{A}_2})[(\mathcal{E}\otimes\mathbb{I})[\Phi_i]]/t(\bar{\mathcal{E}}|\Phi_i),(\mathcal{E}_0\otimes\mathbb{I})[\Phi_+]\big). \tag{25}$$

Using recent self-testing results in a non-IID setting [28] applied to the output probe state, we show in SUPP. MAT. D that for any $x > 0$, $C^o = \sqrt{1-F^o}$ can be bounded by two terms, with confidence of at least $c_x = (1-e^{-x})\cdot(1-2e^{-x})^2$:

$$\arcsin C^o \leq \arcsin\sqrt{\alpha f_x(\epsilon,K)} + \Delta_x(\eta_s,K), \tag{26}$$

where K is the number of pairs measured by Alice and Bob, η_s is the measured heralding efficiency, $\Delta_x(\eta_s,K)$ is an error function that goes to 0 for high values of K, αf_x gives self-testing bound on the output state, in a non-IID regime, with

$$f_x(\epsilon,K) = 8\sqrt{\frac{x}{K}} + \frac{\epsilon}{2} + \frac{\epsilon+8/K}{2+1/K} \xrightarrow[K\to+\infty]{} \epsilon, \tag{27}$$

and $\alpha = 1.26$. We choose $x = 7$ to get a confidence $c_x > 99.5\%$, and measure $K \simeq 10^9$ copies of the probe state, in order to reach the asymptotic values, which takes from 1 to 3 hours in our experiments depending on the channel transmissivity. Note that the error function is due to both the non-IID regime and the lack of information on channels that do not output any state. A similar error occurs when we evaluate the transmissivity as the measured heralding efficiency:

$$t(\bar{\mathcal{E}}|\Phi_i) \gtrsim \tau_x(\eta_s,K), \tag{28}$$

where $\tau_x(\eta_s,K) \simeq \eta_s$ for high values of K. This way, the actual bound on the fidelity between the input and output state reads, with confidence c_x,

$$\begin{aligned} F(\bar{\rho}_o,\rho_i) \geq 1 - 4\cdot\sin^2\Big(&\arcsin\big(C^i/\tau_x\big) + \\ &\arcsin\sqrt{\alpha f_x(\epsilon,K)} + \Delta_x\Big), \end{aligned} \tag{29}$$

which includes additive error terms compared to bound (7). In the analysis of our data, we include these terms that are minimized thanks to the large number K of states measured for each implementation. Note that the expressions for all the mentioned functions are detailed in SUPP. MAT. D 4.

Assumptions. For clarity we highlight the assumptions made in our security analysis.

First, we assume Alice and Bob can communicate via a trusted private classical channel. It allows the players to agree on their measurement settings, Alice to send Bob the position r of the quantum message ρ_i, and Bob to tell Alice if the states were properly received. Players can hence perform measurements on the fly. Alternatively, Bob could use quantum memories to store all the states sent through the quantum channel, and then only perform the measurements in agreement with Alice. This would however require a large number of trusted quantum memories, which would raise important security issues, as well as practical realization problems.

Secondly, the fair sampling assumption is required on the measurement apparatus for the self-testing procedure, as we allow a large amount of losses to be induced by the quantum channel. Alice's measurement apparatus is completely trusted and characterized, according to the one-sided device independent scenario. On Bob's side, we assume the efficiency of the measurement apparatus to be independent of the measurement setting B_0 or B_1. If the efficiency depends on the state measured, then we consider that dependence to be part of the quantum channel. A slight unbalance of efficiency is allowed between the two different measurement outcomes, and we show in the SUPP. MAT. E 3 that the error induced by this unbalance is negligible.

Finally, in keeping with the 1sDI setting, we make the IID assumption on the probe state source, during each attempt of the protocol. To show the legitimacy of this assumption in our implementation, we performed a series of quantum state tomography measurements, during 8 hours, in order to characterize the fluctuation of the probe state with time. This characterization shows the probe states are stable at the scale of one protocol (see SUPP. MAT. E 1 for the detailed results).

Loopholes. Depending on the scenario, important loopholes may remain open when self-testing the quantum states involved in the protocol. Most notably, the freedom-of-choice loophole is open in our experimental implementation, as the measurement basis is randomized every $1\,\mathrm{s}$ only, by using the simulated randomness of our classical computer. Closing this loophole would require the use of certified quantum random number generators [48], which goes out of the scope of this study. For this first demonstration, we thus assume the results are not affected by the loophole.

The detection loophole also remains open due to the channels' losses, whatever the implementation. Fair-sampling assumptions are thus required, as mentioned earlier.

Finally, the locality loophole is irrelevant in 1sDI and thus the experimental implementation, as Alice's apparatus is trusted and does not communicate with Bob's. In a full-DI setting, closing the loophole requires Alice to send her states over long distances to ensure spacelike-separated detection events during certification procedures. More quantum channels are thus required to carry these states, which have to be assumed trusted in order to perform the protocol.

Source and Detection. Probe states are generated via type-II SPDC in a ppKTP crystal combined with a Sagnac interferometer. We maximized the heralding efficiency $\eta_s = R_{si}/R_i$, with R_i the idler photon detection rate and R_{si} the pair detection rate, following the method proposed in [49, 50]. For that purpose, the pump's spatial mode and focus as well as the pair's collection modes, were tuned carefully when coupling to single-mode fibers, and losses on the signal photon path were minimized. This way the pump is in a collimated mode at the scale of the crystal, close to a gaussian mode of waist $w_p \simeq 315\,\mu\mathrm{m}$, which maximizes the heralding efficiency [50, 51]. The signal photon's coupling mode has a waist $w_s \simeq 190\,\mu\mathrm{m}$, and the idler photon's is $w_i \simeq 218\,\mu\mathrm{m}$. We also used high-efficiency SNSPDs to detect the photons. Losses on the idler photon were not limiting, so we selected the best components and detectors for the signal photon. All detection events were recorded by a time tagger, and dated with picosecond precision. Two detection events were considered simultaneous when measured within the same $500\,\mathrm{ps}$ coincidence window. In this way, we detect idler photons in Alice's detectors with a rate $R_i = 600 \pm 40\,\mathrm{kHz}$ (varying from one protocol attempt to another), for a brilliance of $\simeq 670 \pm 50\,\mathrm{kHz\,W^{-1}\,nm^{-1}}$. SNSPDs display dark count rates of $\leq 500\,\mathrm{Hz}$, such that the probability of falsely heralding a probe state is negligible. Finally, $1\,\mathrm{nm}$-bandwidth spectral filters were used to limit the spectrum spread that would otherwise degrade the polarization state because of birefringence and dispersion in optical fibers.

Quantum State Tomography. We perform quantum state tomographies via linear regression estimation [52] and fast maximum likelihood estimation [53]. Photon counts are corrected by measuring relative efficiencies of the detectors. We use this method in order to reconstruct the probe state Φ_i, and to calculate the probe state fidelity to a maximally entangled state F^i. For this calculation, we maximize the fidelity

$$F^i_U = F\big((\mathbb{I}\otimes U)\Phi_i(\mathbb{I}\otimes U^\dagger), \Phi_+\big) \tag{30}$$

on a local unitary U, to evaluate the maximum fidelity up to isometries, as defined in Eq. (5).

The estimation of uncertainties on the reconstructed states takes into account different factors (see SUPP. MAT. E 4). We use the Monte Carlo method to evaluate the influence of photon counting statistics and systematic errors on measurement bases [54]. Calibration of the measurement stations allows to estimate the maximum deviation from the ideal measurement. Then, we generate 1000 new data samples from simulated experiments, including random deviation on measurement bases within the characterized interval, and photon-counting Poissonian statistics. We thus reconstruct 1000 new density matrices from which we compute the fidelity to the target state. The standard deviation on the resulting fidelities gives an estimation of the uncertainty. In addition, the resulting average fidelity a priori deviates from F^i, so we take this deviation as another uncertainty. Finally, slow thermal fluctuation also induce some uncertainty on the fidelity, as our experiment lasts for a relatively long period of time. By continuously performing quantum state tomographies for 8 hours, we are able to evaluate the fluctuations in the quantum state on time spans of the order of a protocol duration. This way, we measure an additional 0.02% error on the quantum state fidelities to Bell states, due to thermal fluctuations. The reader can refer to SUPP. MAT. E 1 for more details on the evaluation of these thermal fluctuations and the drift of the quantum state through time. Note that throughout the manuscript, all error bars on fidelities are enlarged by a factor 2, giving a $95\,\%$ confidence level.

Steering measurement. When testing the violation of steering inequality, players should in principle pick a random measurement basis between A_0B_0 and A_1B_1 for each new photon pair. However, because of technical limitations of our motorized waveplate stages, we only operate this randomization at a limited rate of $1\,\mathrm{Hz}$. A fully secure protocol would therefore require faster electronics and active optical components.

For the implementation of malicious channels, we perform a 7-hours measurement run. From this single run we generate the data that could be acquired in the certification procedure of a variety of channels $\mathcal{E}_{p,q}$, as defined in Eq. (19). For this run, we randomize the measurement basis, with equal probabilities between A_0B_0, A_1B_1 (the channel chooses to act honestly), and $-A_0B_0$, $-A_1B_1$ (the channel chooses to disrupt the state). In order to simulate a larger variety of data samples, we perform that randomization at a 5 Hz-rate. We then generate the data for the certification of channel $\mathcal{E}_{p,q}$, by picking a random set of samples, with the following proportions:

- $q/2$ in basis $-A_0B_0$,
- $p/2$ in basis $-A_1B_1$,
- $(1-q)/2$ in basis A_0B_0,
- $(1-p)/2$ in basis A_1B_1.

The data acquired in basis $-A_0B_0$ and $-A_1B_1$ is treated as if it was acquired in basis A_0B_0 and A_1B_1, respectively, when calculating the average violation of steering inequality $\beta = |\langle A_0B_0\rangle + \langle A_1B_1\rangle|$.

Note added. While finishing this manuscript we became aware of a related work by Bock *et al.* [55].

Acknowledgments

We acknowledge useful discussions with Anupama Unnikrishnan and Massimiliano Smania on self-testing techniques and assistance from ID Quantique with the single-photon detectors. We also acknowledge financial support from the European Research Council project QUSCO (E.D.), the ANR VanQuTe project ANR-17-CE24-0035, the PEPR integrated projects EPiQ ANR-22-PETQ-0007, DI-QKD ANR-22-PETQ-0009, and QCommTestbed ANR-22-PETQ-0011, which are part of Plan France 2030, and the Marie Skłodowska-Curie grant agreement No 956071 (AppQInfo).

Author contributions

S.N. developed the theoretical protocols and proofs, together with I.S., D.M. and E.D. S.N. and E.D. conceived the experimental setup, S.N. developed it, and S.N., L.M., V.Y. and P.L. performed the protocol implementation. S.N. and L.M. processed the data. All authors discussed the analysis of the data, and contributed to writing or proofreading the manuscript. D.M. and E.D. supervised the project.

[1] S. Wehner, D. Elkouss, and R. Hanson, Science **362**, eaam9288 (2018).
[2] J. F. Fitzsimons, npj Quantum Information **3**, 26 (2017).
[3] N. Shettell and D. Markham, Phys. Rev. A **106**, 052427 (2022).
[4] D. Awschalom and et al., PRX Quantum **2**, 017002 (2021).
[5] Z. H. Saleem, T. Tomesh, M. A. Perlin, P. Gokhale, and M. Suchara, arXiv:2107.07532 [quant-ph] (2021).
[6] H. Barnum, C. Crépeau, D. Gottesman, A. Smith, and A. Tapp, in *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.* (IEEE, 2002) pp. 449–458.
[7] F. Dupuis, J. B. Nielsen, and L. Salvail, in *Advances in Cryptology–CRYPTO 2012: 32nd Annual Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2012. Proceedings* (Springer, 2012) pp. 794–811.
[8] A. Broadbent, G. Gutoski, and D. Stebila, in *Advances in Cryptology–CRYPTO 2013: 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part II* (Springer, 2013) pp. 344–360.
[9] D. Markham and A. Marin, in *Information Theoretic Security: 8th International Conference, ICITS 2015, Lugano, Switzerland, May 2-5, 2015. Proceedings 8* (Springer, 2015) pp. 1–14.
[10] D. Markham and A. Krause, Cryptography **4**, 3 (2020).
[11] H. Zhu and M. Hayashi, Physical Review A **100**, 062335 (2019).
[12] Y. Takeuchi, A. Mantri, T. Morimae, A. Mizutani, and J. F. Fitzsimons, npj Quantum Information **5**, 27 (2019).
[13] R. Colbeck, arXiv:0911.3814 [quant-ph] (2011).
[14] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).
[15] S. Pironio, A. Acín, S. Massar, A. B. d. l. Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature **464**, 1021 (2010).
[16] B. W. Reichardt, F. Unger, and U. Vazirani, Nature **496**, 456 (2013).
[17] F. Baccari, R. Augusiak, I. Šupić, and A. Acín, Phys. Rev. Lett. **125**, 260507 (2020).
[18] I. Šupić and N. Brunner, arXiv:2203.13171 [quant-ph] (2022).
[19] P. Sekatski, J.-D. Bancal, S. Wagner, and N. Sangouard, Phys. Rev. Lett. **121**, 180505 (2018).
[20] F. Magniez, D. Mayers, M. Mosca, and H. Ollivier, arXiv:quant-ph/0512111 (2005).
[21] F. Graffitti, A. Pickston, P. Barrow, M. Proietti, D. Kundys, D. Rosset, M. Ringbauer, and A. Fedrizzi, Physical Review Letters **124**, 010503 (2020), arXiv: 1906.11130.
[22] Y. Mao, Y.-Z. Zhen, H. Liu, M. Zou, Q.-J. Tang, S.-J. Zhang, J. Wang, H. Liang, W. Zhang, H. Li, L. You, Z. Wang, L. Li, N.-L. Liu, K. Chen, T.-Y. Chen, and J.-W. Pan, Physical Review Letters **124**, 010502 (2020).
[23] Y. Yu, P.-F. Sun, Y.-Z. Zhang, B. Bai, Y.-Q. Fang, X.-Y. Luo, Z.-Y. An, J. Li, J. Zhang, F. Xu, X.-H. Bao, and J.-W. Pan, Physical Review Letters **127**, 160502 (2021).
[24] P. Sekatski, J.-D. Bancal, M. Ioannou, M. Afzelius, and N. Brunner, Physical Review Letters **131**, 170802 (2023).
[25] See Supplemental Material [url] for more details, which includes Refs. [56–65].
[26] The identity channel on $\mathcal{H}_i^{ext}$ is omitted in (2) for more clarity.
[27] A. E. Rastegin, arXiv:quant-ph/0602112 (2006).
[28] A. Unnikrishnan and D. Markham, Phys. Rev. A **100**, 032314 (2019).
[29] I. Šupić and M. J. Hoban, New J. Phys. **18**, 075006 (2016).
[30] E. G. Cavalcanti, S. J. Jones, H. M. Wiseman, and M. D. Reid, Phys. Rev. A **80**, 032112 (2009).

[31] A. Fedrizzi, T. Herbst, A. Poppe, T. Jennewein, and A. Zeilinger, Opt. Express **15**, 15377 (2007).

[32] Y.-H. Li, Z.-Y. Zhou, Z.-H. Xu, L.-X. Xu, B.-S. Shi, and G.-C. Guo, Physical Review A **94**, 043810 (2016).

[33] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).

[34] I. Bongioanni, L. Sansoni, F. Sciarrino, G. Vallone, and P. Mataloni, Phys. Rev. A **82**, 042307 (2010).

[35] M.-D. Choi, Linear Algebra and its Applications **10**, 285 (1975).

[36] S. Neves, V. Yacoub, U. Chabaud, M. Bozzio, I. Kerenidis, and E. Diamanti, Nature Communications **14**, 1855 (2023).

[37] M. Bozzio, U. Chabaud, I. Kerenidis, and E. Diamanti, Phys. Rev. A **102**, 022414 (2020), arXiv: 2002.09005.

[38] H. E. Guilbert, *Efficient Entangled Biphoton Production and Manipulation for Quantum Applications*, Ph.D. thesis, Duke University (2015).

[39] M. Smania, *Photonic multipartite entanglement: Generation, measurement and applications*, Ph.D. thesis, Department of Physics, Stockholm University (2020).

[40] E. Meyer-Scott, C. Silberhorn, and A. Migdall, Review of Scientific Instruments **91** (2020).

[41] L. d. S. Martins, N. Laurent-Puig, P. Lefebvre, S. Neves, and E. Diamanti, arXiv preprint arXiv:2407.00802 (2024).

[42] H. Wang, Y.-M. He, T.-H. Chung, H. Hu, Y. Yu, S. Chen, X. Ding, M.-C. Chen, J. Qin, X. Yang, *et al.*, Nature Photonics **13**, 770 (2019).

[43] Y.-M. He, H. Wang, C. Wang, M.-C. Chen, X. Ding, J. Qin, Z.-C. Duan, S. Chen, J.-P. Li, R.-Z. Liu, *et al.*, Nature Physics **15**, 941 (2019).

[44] S. Thomas, M. Billard, N. Coste, S. Wein, Priya, H. Ollivier, O. Krebs, L. Tazaïrt, A. Harouri, A. Lemaitre, *et al.*, Physical review letters **126**, 233601 (2021).

[45] M. Cao, F. Hoffet, S. Qiu, A. S. Sheremet, and J. Laurat, Optica **7**, 1440 (2020).

[46] Y.-M. Xie, B.-H. Li, Y.-S. Lu, X.-Y. Cao, W.-B. Liu, H.-L. Yin, and Z.-B. Chen, Optics Letters **46**, 1632 (2021).

[47] C.-L. Li, H.-L. Yin, and Z.-B. Chen, Reports on Progress in Physics **87**, 127901 (2024).

[48] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, npj Quantum Information **2**, 1 (2016).

[49] R. S. Bennink, Phys. Rev. A **81**, 053805 (2010).

[50] N. Bruno, A. Martin, T. Guerreiro, B. Sanguinetti, and R. T. Thew, Optics Express **22**, 17246 (2014).

[51] T. Guerreiro, A. Martin, B. Sanguinetti, N. Bruno, H. Zbinden, and R. Thew, Opt. Express **21**, 27641 (2013).

[52] B. Qi, Z. Hou, L. Li, D. Dong, G. Xiang, and G. Guo, Sci. Rep. **3**, 1 (2013).

[53] J. A. Smolin, J. M. Gambetta, and G. Smith, Phys. Rev. Lett. **108**, 070502 (2012).

[54] J. B. Altepeter, E. R. Jeffrey, and P. G. Kwiat, Advances in Atomic, Molecular, and Optical Physics **52**, 105 (2005).

[55] M. Bock, P. Sekatski, J.-D. Bancal, S. Kucera, T. Bauer, N. Sangouard, C. Becher, and J. Eschner, npj Quantum Information **10**, 63 (2024).

[56] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2011).

[57] N. Lütkenhaus, J. Calsamiglia, and K.-A. Suominen, Phys. Rev. A **59**, 3295 (1999).

[58] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Phys. Rev. A **71**, 062310 (2005).

[59] A. Unnikrishnan, I. J. MacFarlane, R. Yi, E. Diamanti, D. Markham, and I. Kerenidis, Physical Review Letters **122**, 240501 (2019), arXiv: 1811.04729.

[60] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[61] A. Unnikrishnan, *Enforcing trust in quantum networks*, Ph.D. thesis, University of Oxford, University of Oxford (2019).

[62] A. Gočanin, I. Šupić, and B. Dakić, PRX Quantum **3**, 010317 (2022).

[63] I. Šupić and J. Bowles, Quantum **4**, 337 (2020).

[64] D. Orsucci, J.-D. Bancal, N. Sangouard, and P. Sekatski, Quantum **4**, 238 (2020).

[65] “Data sheet for the "broadband polarization-entangled photon source", from oz optics,” https://www.ozoptics.com/ALLNEW_PDF/DTS0198.pdf.