Certified Randomness in Tight Space

Andreas Fyrillas,[1,†] Boris Bourdoncle,[1,*,†] Alexandre Maïnos,[1,2] Pierre-Emmanuel Emeriau,[1]
Kayleigh Start,[1] Nico Margaria,[1] Martina Morassi,[3] Aristide Lemaître,[3] Isabelle Sagnes,[3]
Petr Stepanov,[1] Thi Huong Au,[1] Sébastien Boissier,[1] Niccolo Somaschi,[1] Nicolas Maring,[1]
Nadia Belabas,[3,‡] and Shane Mansfield[1,‡]

[1] *Quandela, 7 Rue Léonard de Vinci, Massy 91300, France*
[2] *Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS81FD, United Kingdom*
[3] *Université Paris-Saclay, CNRS, Centre de Nanosciences et de Nanotechnologies, Palaiseau 91120, France*

(Received 22 September 2023; revised 2 February 2024; accepted 30 April 2024; published 30 May 2024)

Reliable randomness is a core ingredient in algorithms and applications ranging from numerical simulations to statistical sampling and cryptography. The outcomes of measurements on entangled quantum states can violate Bell inequalities, thus guaranteeing their intrinsic randomness. This constitutes the basis for certified randomness generation. However, this certification requires spacelike separated devices, making it unfit for a compact apparatus. Here we provide a general method for certified randomness generation on a small-scale application-ready device and perform an integrated photonic demonstration combining a solid-state emitter and a glass chip. In contrast to most existing certification protocols, which in the absence of spacelike separation are vulnerable to loopholes inherent to realistic devices, the protocol we implement accounts for information leakage and is thus compatible with emerging compact scalable devices. We demonstrate a two-qubit photonic device that achieves the highest standard in randomness, yet is cut out for real-world applications. The full 94.5-h-long stabilized process harnesses a bright and stable single-photon quantum-dot-based source, feeding into a reconfigurable photonic chip, with stability in the milliradian range on the implemented phases and consistent indistinguishability of the entangled photons above 93%. Using the contextuality framework, we certify private randomness generation and achieve a rate compatible with randomness expansion secure against quantum adversaries.

DOI: 10.1103/PRXQuantum.5.020348

I. INTRODUCTION

The strictest requirements on randomness sources are typically destined to cryptographic applications. There, randomness should ideally be both unpredictable and private, so that no information about the generated sequence can be gained by an eavesdropper either prior to or immediately after its generation. Quantum sources admit certification of these properties, by exploiting links between the unpredictability of quantum behavior and the violation of Bell inequalities. A guarantee that numbers have been sampled from empirical data exhibiting Bell nonlocality [1] or, more generally, contextuality [2–5] can suffice to certify unpredictability and privacy [6,7].

Randomness certification and other Bell-inequality-based protocols offer attainable practical advantages for quantum information processing with relatively low numbers of qubits, but they are nevertheless susceptible to loopholes [8]. One way to close the locality or, more generally, the compatibility loophole is to ensure spacelike separation between the players of the nonlocal game [9–13]. However, that is not an option for a practical compact device. Merely asserting that the relevant parts of the device are shielded [14] is unsatisfactory for users who would like to protect themselves against a device deteriorating with time. For such a device, the compatibility loophole must be carefully addressed, because crosstalk can lead to detrimental information flow between components. This compromises theoretical analyses and security proofs even outside of adversarial scenarios. More broadly, all future on-chip quantum information processing will be susceptible to such effects, for which reason it is essential

[*] Corresponding author: boris.bourdoncle@quandela.com
[†] These authors contributed equally to this work.
[‡] These authors jointly supervised this work.

that they be taken into account in protocols and algorithms at the information processing level.

In this work, we introduce novel theoretical tools to address the locality loophole, which we demonstrate in a randomness certification protocol performed on a compact two-qubit photonic processor. Idealized analyses typically lead to relations between relevant figures of merit (such as fidelities, rates, and guessing probabilities) on the one hand, and Bell violations or more general contextuality measures [15] on the other hand. Here, however, we provide relations suited to realistic devices, which allow the evaluation of the relevant figures of merit in terms of both beneficial contextuality and detrimental crosstalk. Moreover, we introduce a method to upper bound the amount of crosstalk by computing how far the device's observed behavior is from the set of quantum correlations approximated by the Navascués-Pironio-Acín (NPA) hierarchy [16]. This enables detection of adversarial manipulation of the device that may seek to exploit the locality loophole to spoof certification.

We propose a certification method that is secure against quantum side information, meeting the highest security standards. This method requires acquiring large statistics while maintaining high photon purity and indistinguishability [17,18], which places knock-on constraints on hardware efficiency and stability. Our theoretical contribution bridging the gap between ideal situations and realistic implementations, combined with finely controlled and robust hardware, allows us to implement the first on-chip certified quantum random number generation protocol with a full security proof.

II. MAXIMAL SCORES IN REALISTIC CONTEXTUAL GAMES

In our protocol, two parties play the Clauser-Horne-Shimony-Holt (CHSH) game [19,20] to guarantee the generation of randomness. The protocol is divided between test rounds that assess a Bell inequality violation and generation rounds that produce random numbers. During test rounds, Alice and Bob each perform one of two measurements $x \in \{0, 1\}$ for Alice and $y \in \{0, 1\}$ for Bob. This yields one of two outputs $a \in \{0, 1\}$ for Alice and $b \in \{0, 1\}$ for Bob. Following the terminology of contextuality, we call a set of measurement choices (or set of "inputs") a "context." Our corresponding implementation is sketched in Fig. 1(a) and detailed in Fig. 2 below.

The winning condition for the CHSH game for a context (x, y) and joint results (a, b) is $a \oplus b = x \cdot y$. The probability of obtaining (a, b) when measuring (x, y) is denoted $p(ab|xy)$, and the set of the four conditional distributions $\{p(a, b|x, y)\}$ forms the behavior of our device, denoted p. Its CHSH score $S_{\mathrm{CHSH}}(p)$ and its CHSH inequality

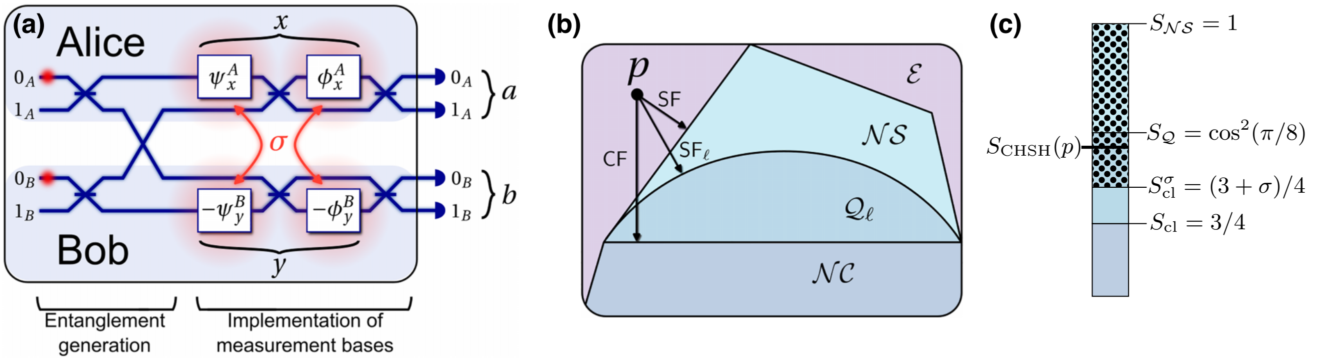

FIG. 1. Bell test on a compact device. (a) CHSH game on a photonic chip. The glass chip uses two dual-rail encoded qubits: modes 0_A and 1_A correspond to Alice's qubit, modes 0_B and 1_B to Bob's qubit. The thick lines represent waveguides; crosses with horizontal lines represent symmetric beam splitters. The first part of the chip generates the Bell state $(|0_A 1_B\rangle + |1_A 0_B\rangle)/\sqrt{2}$ after postselection on detection events to project into the dual-rail qubit basis (see Appendix D for details). The second part enables the agents to select the inputs (x, y) of the Bell test, i.e., the measurement bases, by applying phases via thermo-optic phase shifters. Here σ represents the crosstalk that can occur between Alice and Bob. The outputs (a, b) correspond to where the photons are detected at the exit of the chip. The distributions of the outputs conditioned on the input form the behavior, denoted p. (b) Sketch of the set $\mathcal{E}$ of all possible behaviors and its subsets. The behavior p underlying the statistics of our implementation lies outside the no-signaling polytope $\mathcal{NS}$ because σ is nonzero. We quantify by SF, CF, and SF_ℓ how far p is respectively from the no-signaling set $\mathcal{NS}$, the noncontextual set $\mathcal{NC}$, and the approximation of the quantum set defined by the ℓ th level of the Navascues-Pironio-Acín hierarchy $\mathcal{Q}_l$. (c) Maximal scores for the CHSH game. Thanks to our bound relating contextuality and signaling, we know that a score above S^{σ}_{cl} (dotted region) cannot be achieved with noncontextual behavior even when an amount σ of signaling is allowed: above that threshold, the behavior p has to be contextual and can be used to certify randomness. We denote by S_{cl}, $S_{\mathcal{Q}}$, and $S_{\mathcal{NS}}$ the maximal scores over the noncontextual, quantum, and no-signaling sets, respectively.

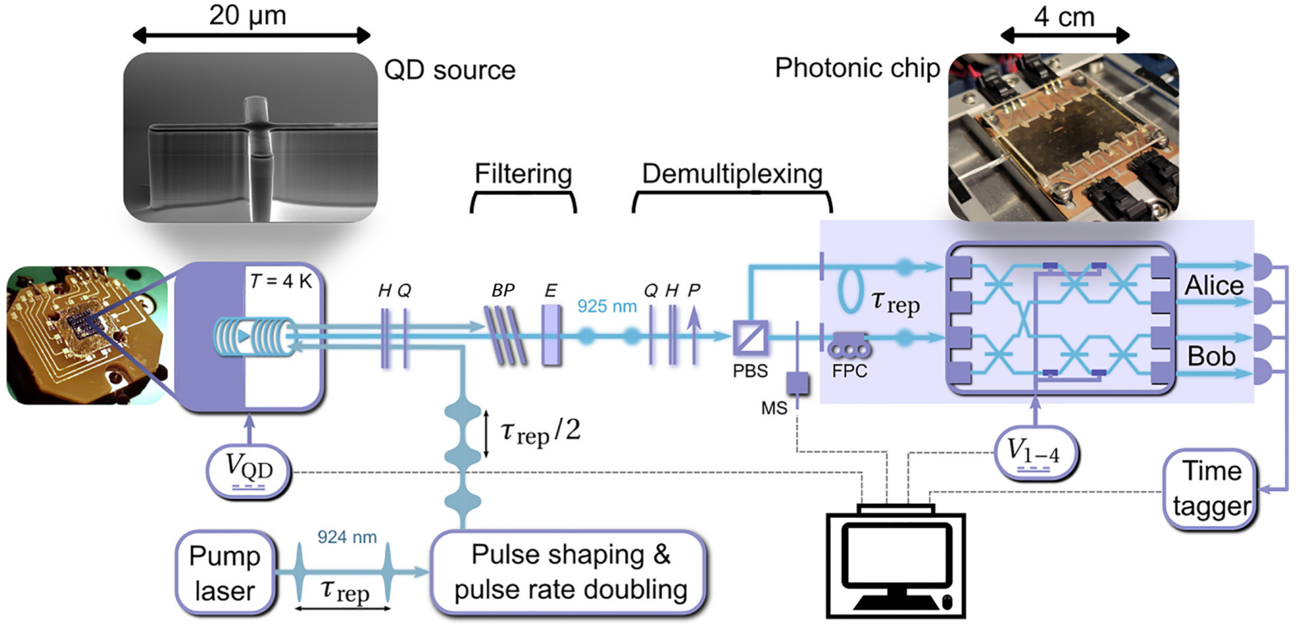

FIG. 2. Compact implementation of a certified quantum random number generator. The quantum dot (QD) photon emitter generates photons at 925 nm via a phonon-assisted excitation scheme (see Sec. IX below). H, Q, half- and quarter-wave plates; BP, bandpass filter; E, etalon; P, polarizer. After a demultiplexing stage (see Sec. IX), the outputs of a polarizing beam splitter (PBS) are collected with collimators. The setup is entirely fibered or waveguided in the blue area. A fibered delay τ_{rep} allows us to synchronize pairs of photons sent into the photonic chip. A motorized shutter (MS) enables chip voltage calibration. The fibered polarization controller (FPC) ensures that both photons enter the photonic chip with the same polarization. Dashed gray lines indicate that elements of the setup are automated to implement the randomness generation protocol, by adapting the voltage on the photon source for optimal brightness and periodic calibrations of the thermo-optic phase-shifter voltages. Here $V_{1\text{–}4}$ control the phases on chip and hence measurement bases of Alice and Bob. The V_{QD} feedback loop ensures that the QD emission remains bright and the emitted photons indistinguishable.

violation [19] $I_{\text{CHSH}}(p)$ are defined as

$$S_{\text{CHSH}}(p) = \sum_{a,b,x,y} p(x,y)V(a,b,x,y)p(ab|xy), \qquad (1)$$

$$I_{\text{CHSH}}(p) = \langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle, \qquad (2)$$

where $p(x,y)$ is the distribution of the measurement choices, V is the scoring function for the game, defined by $V(a,b,x,y) = 1$ if $a \oplus b = xy$ and $V(a,b,x,y) = 0$ otherwise, and where $\langle A_xB_y\rangle = p(a = b|x,y) - p(a \neq b|x,y)$. When the input distribution is uniform, $S_{\text{CHSH}} = (I_{\text{CHSH}} + 4)/8$. For a nonsignaling behavior, a value $S_{\text{CHSH}}(p) > 0.75$ [$I_{\text{CHSH}}(p) > 2$] is a signature that p is contextual. For a given experimental setup, the set of all possible behaviors is often represented within a geometric space [1], in which behaviors with relevant properties typically form polytopes or convex subsets [Fig. 1(b)].

In Bell-based quantum information processing, the key element to examine the intrinsic properties of a behavior is its decomposition into hidden-variable models (HVMs). Operationally, one can think of an HVM as a description of the devices held by an eavesdropper that reproduces the observed behavior, thus giving the eavesdropper more predictive power while being indistinguishable from the original behavior for the users. A behavior is noncontextual if it admits a factorizable HVM [3], and we quantify the departure from the set of noncontextual behaviors with the contextual fraction CF [15]. When no signaling is guaranteed, a behavior is noncontextual if and only if it can be decomposed as a mixture of deterministic models [21]. Conversely, this means that the underlying behavior is inherently nondeterministic, and randomness can be certified from it, when one observes a score $S(p)$ higher than the noncontextual (or "classical") maximum S_{cl}, defined as

$$S_{\text{cl}} := \max_{p}\{S(p) \mid \mathsf{CF}(p) = 0\}. \qquad (3)$$

Note that the score $S(p)$ we define here is not limited to the CHSH game: this definition applies more generally to any contextual game, which can have more than two measurement choices, more than two measurement results, more than two players, and any scoring function mapping the input-output tuples to $\{0, 1\}$. One can also define S_{cl}^{ξ}, the maximal score over HVMs with contextual fraction at most ξ, as

$$S_{\text{cl}}^{\xi} := \max_{p}\{S(p) \mid \mathsf{CF}(p) \leq \xi\}, \qquad (4)$$

and relate it to the consistency of the game [22].

For every context, the scoring function of a game defines a logical formula on the outputs that has to be satisfied to win the game [for the CHSH game, $a \oplus b = 1$ for inputs $(1, 1)$ and $a \oplus b = 0$ for the other inputs]. A game is said to be k-consistent if at most k formulae can be simultaneously satisfied by an assignment of the outputs. For the CHSH game, $a = b = 0$ satisfies the formulae defined by all inputs except $(1, 1)$, and that is the best one can do: the CHSH game is 3-consistent.

Proposition 1 (Theorem 4 of Refs. [15,23]). Let S_{cl}^{ξ} be the maximal score for a contextual game with m measurement choices that is k-consistent. Then

$$S_{\mathrm{cl}}^{\xi} \leq \frac{k + \xi(m - k)}{m}. \tag{5}$$

Unfortunately, the measurements performed in practical scenarios and on devices without spacelike separation are not perfectly compatible [24], which implies that the no-signaling conditions [25] are not met and the tools that were developed to certify randomness [6,7,26] do not apply in a straightforward way. To overcome this limitation, the set of admissible HVMs needs to be extended. One can, for instance, admit more general measurements [27] or more general behaviors [28]. Here, we allow a limited amount of signaling at the hidden-variable level, and we compute new maximal scores achievable over these models. More precisely, we define the signaling fraction SF (simultaneously introduced in Ref. [29]) to quantify how far a behavior is from the no-signaling set. Like the contextual fraction, SF is the solution of a linear program and can thus be computed efficiently. Its formal definition can be found in Appendix A. We then define $S_{x,y}^{\sigma}$, the maximal score over HVMs with signaling fraction at most σ, and the further requirement that the HVMs be deterministic on a specific context (x, y), called the distinguished context:

$$\begin{aligned} S_{x,y}^{\sigma} := \max_{p} \{ S(p) \mid \text{ there exists } (a, b) \\ \text{such that } p(ab|xy) = 1 \text{ and } \mathsf{SF}(p) \leq \sigma \}. \end{aligned} \tag{6}$$

The score with a distinguished context is the relevant quantity to certify a lower bound on randomness against quantum side information using the techniques introduced by Miller and Shi [26]. We extend the score definition to the case of nonzero signaling ($\sigma > 0$), in order to design a protocol for realistic devices, and we relate $S_{x,y}^{\sigma}$ to S_{cl}^{ξ} with the following proposition.

Proposition 2. For a contextual bipartite game with binary input choices for both players, let (x, y) be a distinguished context. Then

$$S_{x,y}^{\sigma} \leq S_{\mathrm{cl}}^{\xi=\sigma}. \tag{7}$$

The proof can be found in Appendix B, where the bound is formulated in the more general case of n-partite games with binary inputs. This relation can be interpreted as follows: $S_{x,y}^{\sigma}$ is an extension of the distinguished context score to behaviors with signaling fraction at most σ; S_{cl}^{ξ} is an extension of the classical score to behaviors with contextual fraction at most ξ; the latter is an upper bound on the former, with $\xi = \sigma$. In other words, the extent to which an amount of signaling σ can improve the distinguished context score is bounded by the extent to which an amount $\xi = \sigma$ of contextuality can improve the classical score. For our protocol to generate randomness, one has to observe a score above $S_{x,y}^{\sigma}$. In the CHSH case, combining Propositions 1 and 2, this means that $S_{\mathrm{CHSH}}(p) > (3 + \sigma)/4$ is a sufficient condition, as depicted in Fig. 1(c). Compared to Um *et al.* [28], who also used the spot-checking security proof of Miller and Shi [26] and took into account imperfect compatibility between the measurements, our analysis is more general, as it applies to any n-partite nonlocal game with binary inputs.

III. ESTIMATING PHYSICAL CROSSTALK σ

In order to define the set of admissible HVMs associated with our implementation, we allow a limited amount of information between the subsystems of Alice and Bob, due to crosstalk between the components (see Fig. 1). This means that the admissible HVMs can have a positive signaling fraction σ, for some σ that can be characterized either by an upstream partial characterization of the devices or through an on-the-fly estimation based on the statistics of the inputs and outputs collected during the protocol. The first approach requires a device-dependent physical analysis of the setup, while the second approach is semi-device-independent: it derives a cross-talk estimate only from the observed input-output correlations, and hence its device-independent characterization, and then relates it to the underlying crosstalk via an assumption on the device. We follow the second one here and assume that σ is related to the amount of signaling observed empirically. This assumption is well founded if the devices were fabricated by an honest provider, i.e., were not programmed to act maliciously in order to function with a high level of crosstalk while keeping the empirically observable signaling low. In that case, an eavesdropper can only take advantage of flaws in the implementation and deterioration of the devices with time to try and predict the outputs.

More precisely, we assume that our implementation obeys the laws of quantum mechanics, and we thus take $\sigma = \mathsf{SF}_{\ell}$, where SF_{ℓ} is the extension of SF obtained by replacing the nonsignaling set with the set of quantum correlations approximated at the ℓ th level of the NPA hierarchy [16,30]. The reasoning for that choice is the following: if we instead took $\sigma = \mathsf{SF} \leq \mathsf{SF}_{\ell}$, an adversarial quantum strategy consisting in preparing a distribution p_{adv}

such that $S(p_{\text{adv}})$ is high, $\mathsf{SF}(p_{\text{adv}})$ is low and the output distribution for the distinguished context is biased for the benefit of the adversary would not be discarded by our analysis, while preparing such a behavior quantumly requires more than σ signaling.

We give the formal definition of SF_ℓ in Appendix A, along with some properties of this measure and a comparison to the crosstalk measure introduced by Silman *et al.* [27].

IV. PROTOCOL FOR CERTIFIED RANDOMNESS GENERATION

We build upon Miller and Shi's spot-checking random number generation protocol [26] to compute a lower bound on the min-entropy of the total string of bits obtained during the execution of our protocol. This protocol is valid for general contextual games and is secure in the presence of the most general kind of information accessible to an eavesdropper, i.e., quantum side information. The complete description of this protocol is given in Appendix C, along with the assumptions required for the min-entropy bound to be valid. In that appendix, we also describe how our idea can be combined with a protocol for randomness certification against classical side information such as those introduced in Refs. [31–33], because in the trusted provider scenario this restriction on the side information is reasonable.

When the protocol succeeds, the min-entropy of the output sequence **AB** conditioned on the input sequence **XY** and all information potentially available to an eavesdropper E [34] satisfies

$$H^\delta_{\min}(\mathbf{AB}|\mathbf{XY},E) \geq N[\pi(\chi) - \Delta]; \tag{8}$$

$H^\delta_{\min}$ quantifies how many bits can be extracted from the output sequence of N bits, such that these sifted bits are uniformly random and uncorrelated to the quantum side information held by the eavesdropper. Here χ is the score threshold fixed prior to the execution of the protocol, π is the rate curve, and Δ is the correction term, whose expression is given in Appendix C.

V. DESCRIPTION OF OUR SETUP

A scheme of the setup we use to implement our certified randomness generation is provided in Fig. 2: an electrically controlled semiconductor quantum dot in a 2-μm-diameter micropilar cavity generates single photons that are sent to a reconfigurable glass chip, implementing the CHSH game by varying the measurement context via optical phases and measuring output coincidences.

By a periodical calibration during the experiment, we maintain a high precision over the implemented measurement bases to limit signaling between the two agents, quantified by SF_ℓ. We use a bright and stable Quandela semiconductor quantum-dot- (QD) based single-photon source [35] that delivers indistinguishable single photons, allowing us to obtain high Bell inequality violations.

The polarized fibered-device brightness of our QD-based single-photon source, i.e., the probability to detect after the filtering stage with a polarizer an emitted photon following an excitation pulse, is 8.3% ± 0.8% (all error bars represent one standard deviation).

We quantify the purity of the single photons, i.e., the proportion of $|1\rangle$ Fock states compared to $|2\rangle$, with the second-order normalized correlation function $g^{(2)}(0) \approx 2.31\% \pm 0.03\%$ [36] and their indistinguishability with the Hong-Ou-Mandel (HOM) visibility $V_{\text{HOM}} = 93.09\% \pm 0.04\%$ [37]. The train of emitted photons is converted with a passive demultiplexing stage (see Sec. IX below) into pairs of photons entering simultaneously the photonic chip.

On exiting the chip, the photons are detected by high-efficiency single-photon detectors and time tagged. The overall transmission of the setup, i.e., the probability that a pump pulse results in a detected photon, is 2.7%.

The details about the source, the chip, and the selection of the measurement bases can be found in Sec. IX below.

VI. RELATION BETWEEN PHOTON DISTINGUISHABILITY AND CHSH VIOLATION

Achieving certified randomness requires witnessing correlations that violate the CHSH inequality. This places requirements on the purity and indistinguishability of the photons emitted by the single-photon source. For the ideal case of a source emitting only pure photons in a lossless optical circuit, we derive the relation

$$I_{\text{CHSH}} = \sqrt{2}(V_{\text{HOM}} + 1), \tag{9}$$

where V_{HOM} is the HOM visibility of the photons. We derive this relation in Appendix D. We simulate the relationship between I_{CHSH} and V_{HOM}, taking into account photon impurity and circuit losses with Perceval, a software platform specialized in simulations of photonic circuits in the discrete variable paradigm [38]. In our case, the main causes of photon distinguishability arise from polarization fluctuations in the optical fibers and charge noise around the QD [39].

The comparison of the CHSH value obtained with our experimental setup and the simulated one is given in Fig. 3. We acquire I_{CHSH} experimentally as a function of V_{HOM}. The photon distinguishability is adjusted by manually shifting the polarization of one of the two photons entering the photonic chip with a fibered polarization controller (see the FPC in Fig. 2). The HOM visibility V_{HOM} is measured independently from I_{CHSH} by setting Alice's on-chip interferometer phases $\psi^A = 0$ and $\phi^A = -\pi/2$ in Fig. 2. Her interferometer then implements the following unitary matrix on the first two spatial modes (up to a global

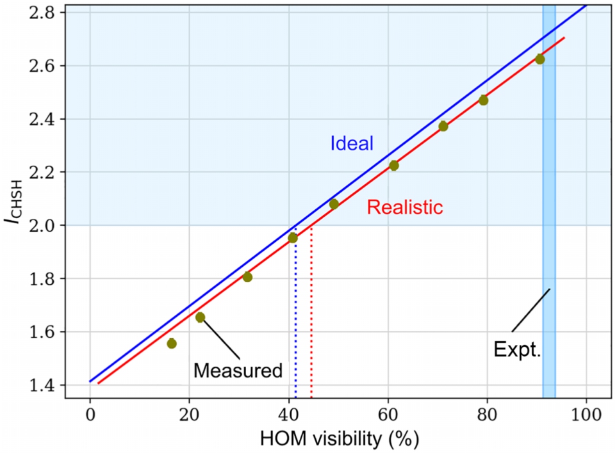

FIG. 3. CHSH expression value $I_{\rm CHSH}$ as a function of HOM visibility. Blue line represents the analytical dependence derived for a photon source emitting only pure single photons $g^{(2)}(0) = 0$ in a lossless circuit $T = 100\%$. Red line represents the realistic curve simulated with the Perceval package for $g^{(2)}(0) = 0.023$ and an optical circuit transmission of $T = 2.7\%$, which are the experimental values. The blue and red dotted lines indicate the minimum HOM visibility required to certify quantum correlations for the ideal and realistic cases, respectively. The minimum values are respectively approximately 41.4% and approximately 44.5%. Green dots denote measured data points. The HOM visibility is decreased by changing the polarization of one of the two photons entering the photonic chip. The error bars extending over ±0.02 representing one standard deviation are contained in the plot markers. Light blue area indicates CHSH violations certifying quantum correlations in the no-signaling case. Blue vertical bar indicates the range of HOM visibility values measured during our 94.5-h main experiment.

phase; see Sec. IX C below):

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}. \tag{10}$$

This is a 50:50 beam-splitter matrix like in the canonical Hong-Ou-Mandel experiment. The HOM visibility is then $V_{\rm HOM} = 1 - 2p_{\rm coinc}$, where $p_{\rm coinc}$ is the probability of simultaneously detecting a photon on both of Alice's detectors in this configuration. The close match between simulated and measured data points validates the performance of our setup.

VII. IMPLEMENTATION OF THE PROTOCOL AND EXPERIMENTAL RESULTS

The full protocol is described in Fig. 4. Before each round of the experiment, Alice and Bob choose the measurement contexts (x, y) via phases (ψ_x^A, ϕ_x^A) and (ψ_y^B, ϕ_y^B) indicated in Fig. 1(a). The corresponding values of the phases are given in Table I in Sec. IX below. Alice and Bob also agree that $(x = 0, y = 0)$ is the generation context, i.e., the measurement context in which the random number will be generated. We choose $\phi_0^A = -\pi/2$, so that $(x = 0, y = 0)$ allows us to acquire the HOM visibility in parallel to the generation rounds because Alice's interferometer acts as a symmetric beam splitter in that configuration. For our experiment, $\psi_x^A = \psi_y^B = 0$. To implement in practice on the chip a set of phases, we solve the phase-voltage matrix equation written in Appendix E using standard optimization procedures. Alice (Bob) maps her (his) result to "0" when a photon is detected in mode 0_A (0_B) and to "1" when a photon is detected in mode 1_A (1_B), according to the mode labeling of Fig. 1(a).

The protocol alternates between generation rounds and test rounds. A round consists in the measurement of a coincidence between Alice and Bob, whose result is stored in the form 00, 01, 10, or 11, where the first bit describes Alice's result and the second one Bob's result. For the generation rounds, Alice and Bob set their phase to the generation context, and for the test rounds, they randomly choose a measurement context among the four. The protocol outputs three binary sequences: generation round results, test round results, and test round contexts. The first stores the coincidence results for the generation rounds and represents the sequence of random bits before randomness extraction. The test round results and contexts sequences are used to determine the experiment's behavior p. From it, we can compute $S_{\rm CHSH}(p)$ and $\mathsf{SF}(p)$. We assume that the detected photons are representative of the whole optical setup behavior, i.e., that the sampling is fair.

The predicted coincidence rate between Alice and Bob, which corresponds to the round processing rate is computed as

$$\begin{aligned}
&158 \times 10^6 && \text{(pump laser pulse rate)}\\
&\times (0.0266)^2 && \text{(overall photon transmission)}\\
&\times 1/4 && \text{(passive demultiplexing)}\\
&\times 1/2 && \text{(state postselection)}\\
&= 14\,000 \pm 3000 \text{ s}^{-1}. &&
\end{aligned} \tag{11}$$

The measured coincidence rate is about 14 200 ± 600 s^{-1}. Considering the thermalization waiting time of 250 ms after each measurement context switch, the expected protocol round processing rate from the measured coincidence rate is around 8300 ± 300 s^{-1}, which is close to the measured rate of 7300 s^{-1}. The disparity indicates an opportunity for improvement in the programming implementation of the randomness generation protocol.

The polarization of photons reaching the polarization-sensitive detectors exhibits temporal fluctuations, leading to asymmetric count rate variations on Alice's and Bob's pair of detectors. This causes a drift of the order of 0.25 mrad/h on the measured phases ϕ^A and ϕ^B of Alice's and Bob's interferometers. To compensate this effect and

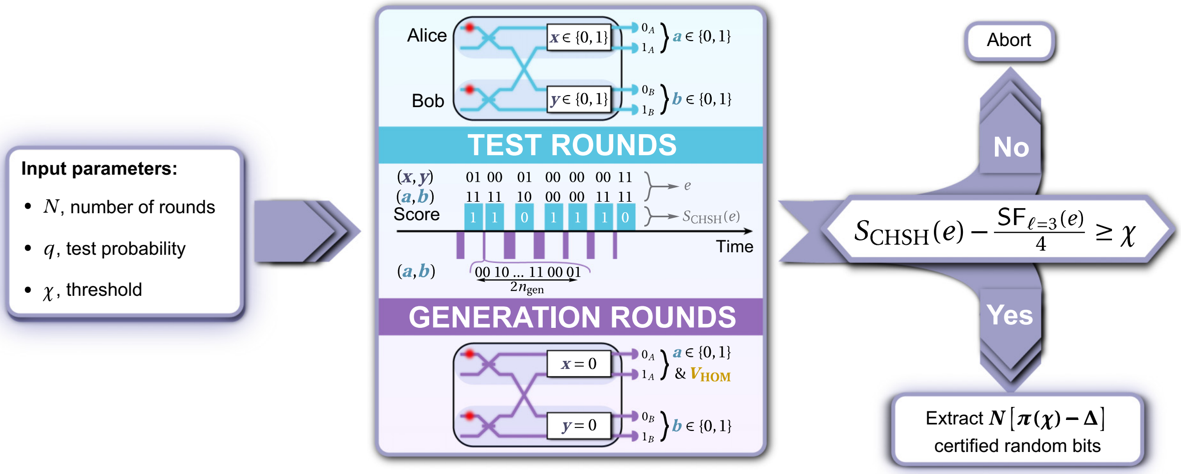

FIG. 4. Certification protocol. The protocol certifies privacy and unpredictability of the output sequence. The main loop consists of two steps: generation rounds (purple area) followed by a test round (blue area). The number of generation rounds n_{gen} to acquire in each iteration of the main loop is determined by a geometric probability distribution of parameter q. Each generation round consists of measuring a coincidence (a,b) in the generation context $(x=0, y=0)$. After n_{gen} generation rounds, a measurement context (x,y) is chosen randomly among $(0,0), (0,1), (1,0)$, and $(1,1)$, and a coincidence (a,b) is measured in a test round. The test rounds are used to compute the CHSH game score $S_{\mathrm{CHSH}}(e) = c/Nq$, where e is the behavior constructed from the test round measurements and c is the sum of all individual test round scores. When the total number of rounds carried out reaches N, the protocol exits the main loop. If $S_{\mathrm{CHSH}}(e) - \mathrm{SF}_{l=3} < \chi$, the protocol aborts; otherwise, we perform randomness extraction on the generation rounds to yield $N[\pi(\chi) - \Delta]$ certified random bits. In parallel to the generation rounds, we monitor the photon indistinguishability via the HOM visibility V_{HOM}.

additional phase errors, the voltages used to implement the corresponding phases for each measurement context with the heating resistors are calibrated at the beginning of the protocol and then after every 6 h of operation (see Appendix F). As a result, all phases stayed confined within an interval of 3 mrad around the target phases during an almost 100-h-long run.

Before data acquisition we measured a violation of the CHSH inequality $I_{\mathrm{CHSH}} = 2.68$, which we used to fix the optimal parameters for the protocol. The total number of rounds $N = 2.4 \times 10^9$ was determined by the desired duration of the acquisition together with the expected rate. We optimized the test probability q (cf. Fig. 4) to maximize the min-entropy bound given by Eq. (8) and fixed $q = 1.34 \times 10^{-4}$. As a result, the number of test rounds carried out to construct the behavior of our device and estimate the CHSH violation is 3.2×10^5. We obtained $I_{\mathrm{CHSH}} = 2.685$ on the test rounds (or $S_{\mathrm{CHSH}} = 0.8356$ in the CHSH game picture) and $\mathrm{SF}_3(e) = 0.005$. In Fig. 5, we present a few key figures recorded during the acquisition, showcasing the stability and precision of our setup. According to Eq. (8), these values certify that the δ-smooth min-entropy $H^{\delta}_{\min}(\mathbf{AB}|\mathbf{XY}, E)$ of the obtained sequence of results $\mathbf{AB}$, conditioned on the sequence of input measurement choices $\mathbf{XY}$ and the quantum side information held by any potential eavesdropper E, is at least

$$H^{\delta}_{\min}(\mathbf{AB}|\mathbf{XY}, E) \geq 7.21 \times 10^6, \tag{12}$$

where we chose as security parameter $\delta = 10^{-10}$. We can hence extract 7.21×10^6 random bits with a Toeplitz matrix hashing randomness extractor [40].

Note that this amount of certified randomness is compatible with randomness expansion, in the sense that if we were to use the interval algorithm to generate our strongly biased input bits from a small number of uniform bits [41,42], the input randomness required for our implementation would be

$$\mathrm{rand}_{\mathrm{in}} = N[h(q) + 2q] = 5,24 \times 10^6, \tag{13}$$

where h is the binary Shannon entropy, which is smaller than the amount of randomness we generate at the output.

VIII. DISCUSSION AND OUTLOOK

This work focuses on certifying randomness with a high standard of security and a compact device, as is needed for real-world applications. Previous implementations of quantum random number generation on a chip focused on the device-dependent framework [43] or did not provide a complete security proof [44], while we report here the first on-chip Bell-based randomness generation protocol with a complete security proof. Improvements, both theoretical and experimental, would allow us to maintain this level of security at higher rates.

On the theoretical side, the Miller-Shi protocol is valid for generic nonlocal and contextual games, but it is

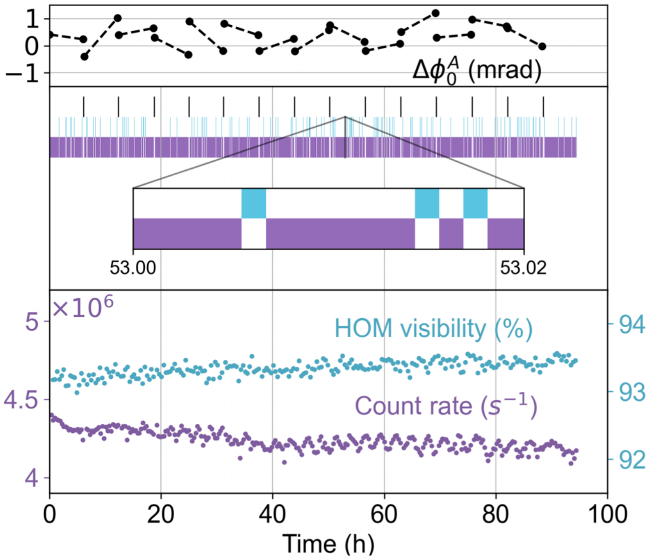

FIG. 5. Stability, brightness, and indistinguishability of photons emitted by the source during a 94.5-h-long experiment. Top plot: difference $\Delta\phi_0^A = \phi_0^A - (-\pi/2)$ between Alice's measured and target interferometer phases in context $(0, 0)$. Phase is measured using the interferometer splitting before and after each context voltage calibration. Middle plot: typical sequence of calibration (black), test (blue) and generation (purple) sequences generated randomly using parameter values of the experiment. Bottom plot: total count rate and HOM visibility at the output of the chip.

not optimal. Recent results on how entropy accumulates [45,46] have been used to obtain optimal rates for randomness expansion with the CHSH game [47] and better rates for arbitrary nonlocal games [48]. The general framework of Brown *et al.* [48] in combination with new techniques for efficiently lower bounding the conditional von Neumann entropy [49] and with our randomness-versus-signaling trade-off can lead to order-of-magnitude improvements in min-entropy rates. This will require extending the aforementioned results to contextuality scenarios with limited but nonzero signaling. We could also use a similar setup to certify randomness against classical side information only. In that case, biasing the input distribution is useful in two cases: when an output distribution for some choice of inputs contains more randomness than others, and when one aims to do randomness expansion (that is the goal of spot checking). When one wishes to simply generate private randomness, and when the randomness-bounding functions are identical for all input choices because of symmetries of the underlying behavior, it is more favourable to generate randomness from all inputs. However, using a biased input distribution is still beneficial to avoid a lengthy recalibration time between each round. By implementing a protocol against classical side information on a silicon nitride chip, where thermo-optic phase shifters react faster than in glass, and with an optimized choice of bias for the inputs, we could certify bits against classical side information at a rate of kbit/s after only a few minutes of acquisition.

On the hardware side, the efficiency of deterministic single-photon sources does not have fundamental limits and can be increased while keeping single-photon purity high. The main limitations are optical losses within the setup and the reconfiguration time of the photonic circuit. We estimate the experiment duration with short-term hardware improvements. With active photon demultiplexing, the photon coincidence rate can be doubled, and using a photonic chip featuring electro-optic [50] or mechanical [51] phase shifters, the round processing rate equals the coincidence rate because it removes the need for thermalization waiting times. Commercially available single-photon detectors can reach 90% detection efficiency. For the single-photon source, these characteristics are currently achieved in the lab: 50% polarized first lens brightness, $g^{(2)}(0) = 0.9\%$, and $V_{\mathrm{HOM}} = 94\%$. As a consequence, the etalon filter is not needed. Factoring in all of these improvements, the total setup transmission increases from 2.66% to 10.5%, and the CHSH inequality value increases from 2.66 to 2.73. Repeating our randomness generation experiment with the same number of rounds to process, that is, $N = 2.4 \times 10^9$, it would take only 1.3 h and generate 9×10^6 certified random bits, resulting in a kbit/s rate.

IX. METHODS

A. Single-photon generation

We illustrate the optical setup in Fig. 2. Single photons at 925.16 nm are generated by a Quandela single-photon source relying on an InAs/GaAs quantum dot embedded in a cavity [35]. A voltage of the order of −1.5 V is applied on the dot, such that the emission line is in resonance with the cavity and to reduce charge noise. The source is pumped using the longitudinal-acoustic phonon-assisted excitation scheme [52–55] at around 924.24 nm and with a pump spectral FWHM $\Delta\omega \approx 0.1$ nm, corresponding to a pulse duration of the order of 12 ps. The pump is a mode-locked femtosecond laser with a repetition rate of 79.08 MHz, corresponding to a duration $\tau_{\mathrm{rep}} \approx 12,6$ ns between two consecutive pulses.

The laser pulses are subsequently temporally shaped using a filtering setup based on a $4f$ line principle, which includes a grating splitting incoming wavelengths into different directions, a slit for wavelength selection, and another grating recombining the light into a single Gaussian beam, ensuring optimal pumping of the source. To increase the random number generation rate of the experiment, the pulse rate is doubled using a fibered Mach-Zehnder interferometer with an approximately $\tau_{\mathrm{rep}}/2$ delay line on one arm.

The excitation pulses are then sent to the photon source. Emitted single photons and reflected pump light are sent to a filtering stage consisting of three bandpass filters (10^{-3} transmission at 924 nm, 805 pm FWHM) and a Fabry-Pérot etalon [FSR 204 pm and finesse 14 at 925 nm, $59\% \pm 1\%$ (error bar is one standard deviation) single-photon transmission].

The first lens brightness of our single-photon source (number of photons collected per excitation pulse at the level of the source [35]) amounts to $39\% \pm 3\%$ and the polarized fibered brightness (number of photons collected per excitation pulse after the filtering stage, including the polarizer; see Fig. 2) is $8.3\% \pm 0.8\%$, corresponding to a polarized photon output rate of $(13.0 \pm 0.1) \times 10^6$ s $^{-1}$.

The purity and indistinguishability of the generated photons are increased by inserting a polarizer and a Fabry-Pérot etalon (see Fig. 2). The purity with the etalon is $g^{(2)}(0) \approx 2.31\% \pm 0.03\%$. The HOM visibility is $93.09\% \pm 0.04\%$. We can deduce from these measurements the photon mean wave-packet overlap $M_s = 97.65\% \pm 0.06\%$ [37].

B. Single-photon manipulation

A passive demultiplexing stage (20% insertion loss) converts the photon stream into pairs of photons arriving simultaneously at the photonic chip input. The demultiplexer consists of a polarizer set in the diagonal position, such that the subsequent polarizing beam splitter in Fig. 2 acts like a symmetric beam splitter. The wave plates preceding the polarizer are set to maximize the number of transmitted photons through the polarizer. One of the outputs of the beam splitter leads to a fibered delay loop. A pair of photons is successfully demultiplexed when the first photon of the pair takes the long path via the fibered loop and the second photon the short path. Hence, only 1/4 of the photon pairs are successfully demultiplexed.

The silica glass chip features laser-written waveguides and four configurable thermo-optic phase shifters (see Appendices E and F for details and operation). The optical transmission of the chip is $58\% \pm 1\%$ (averaged over the two inputs used). Its output is sent to a superconducting nanowire single-photon detector (70% detection efficiency). Photon times of arrival are processed by a time tagger module. We measure an overall setup transmission, i.e., the probability of a photon being detected after an excitation pulse arriving on the source, of 2.7% by using the photon count rate on the detectors.

C. Selection of the measurement bases

We compute the on-chip phases that should be applied in order to maximally violate the CHSH inequality and explain how to calibrate the voltages accordingly. Alice and Bob each control an interferometer [see Fig. 1(a)] consisting of a phase ψ, followed by a symmetric beam splitter, a phase ϕ, and a second symmetric beam splitter.

TABLE I. Phase shifts for maximal CHSH inequality violation. Alice's measurement bases are labeled $x = 0$ and $x = 1$, and Bob's bases are labeled $y = 0$ and $y = 1$. A measurement context is a pair (x, y). From the measurement bases, we compute the phases, and hence the voltages, that should be applied to Alice's and Bob's phase shifters.

Measurement context (x, y)	ψ^A	ϕ^A	ψ^B	ϕ^B
$(0, 0)$	0	$-\pi/2$	0	$-\pi/4$
$(0, 1)$	0	$-\pi/2$	0	$\pi/4$
$(1, 0)$	0	0	0	$-\pi/4$
$(1, 1)$	0	0	0	$\pi/4$

Their interferometer is described by the unitary matrix

$$\widehat{U}(\psi,\phi) = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}\begin{bmatrix}e^{i\phi} & 0\\ 0 & 1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & i\\ i & 1\end{bmatrix}\begin{bmatrix}e^{i\psi} & 0\\ 0 & 1\end{bmatrix}$$
$$= ie^{i\frac{\phi}{2}}\begin{bmatrix}\sin(\phi/2)e^{i\psi} & \cos(\phi/2)\\ \cos(\phi/2)e^{i\psi} & -\sin(\phi/2)\end{bmatrix}. \qquad (14)$$

A possible choice of measurement bases that maximally violates the CHSH inequality is described in Table I. With this choice, when Alice's measurement basis is $x = 0$, we can measure the HOM visibility of the photons by recording the photon coincidences on her outputs. Indeed, for that basis, her interferometer behaves like a symmetric beam splitter.

D. CHSH inequality value computation

A behavior for a CHSH Bell test can be described by a table containing four columns, one for each measurement context (x, y), and the observed probability of each coincidence result (a, b) on the corresponding row. For $V_{\mathrm{HOM}} = 93\,\%$, which characterizes our single-photon source, the expected behavior is as follows (see Appendix D).

a	b	00	01	10	11
0	0	0.414	0.086	0.073	0.073
0	1	0.086	0.414	0.427	0.427
1	0	0.086	0.414	0.427	0.427
1	1	0.414	0.086	0.073	0.073

From the 3.2×10^5 test rounds of our 94.5-h-long randomness generation experiment, we construct the following observed behavior.

a	b	00	01	10	11
0	0	0.424	0.090	0.085	0.077
0	1	0.087	0.416	0.418	0.429
1	0	0.084	0.410	0.418	0.423
1	1	0.405	0.084	0.079	0.071

The uncertainty due to finite statistics is of the order of 0.001 for each cell. Discrepancies between the two tables can be attributed to multiphoton emissions combined with optical losses, the reflectivities of the on-chip directional couplers estimated at 51%, errors on phase shifter phases, and dark counts.

For the theoretical behavior, we get, as expected, $S_{\mathrm{CHSH}} \approx 0.84$ ($I_{\mathrm{CHSH}} \approx 2.73$) and, from our experimentally observed behavior, $S_{\mathrm{CHSH}} \approx 0.835$ ($I_{\mathrm{CHSH}} \approx 2.685$).

The data that support the findings of this study are available from the corresponding authors upon request. The code for the randomness extraction step is available from the corresponding authors upon request.

ACKNOWLEDGMENTS

We thank Pascale Senellart for valuable discussions and feedback, Mathias Pont for experimental help and the multiphoton model of the Perceval package used in our simulations, Damian Markham and Ivan Šupić for discussions on quantum correlations and signaling, and Swabian Instruments for helpful discussions.

This work has received funding by the European Union's Horizon 2020 research and innovation programme QUDOT-TECH under the Marie Skłodowska Curie Grant Agreement No. 861097 and from BPI France Concours Innovation PIA3 projects DOS0148634/00 and DOS0148633/00—Reconfigurable Optical Quantum Computing.

B.B., P-E.E., and S.M. conceptualized the work. A.F., B.B., and N.B. were responsible for the methodology. A.F., A.M., N.B., and N.Maring conducted the experimental investigation. I.S., N.Margaria, S.B., and T.H.A. performed the QD device processing. A.L. and M.M. were responsible for the QD sample growth. P.S. carried out the QD device characterization. K.S. performed the randomness extraction. A.F. and N.B. conducted the data analysis. B.B., N.B., P-E.E., and S.M. performed the formal analysis. A.F., B.B., N.B., and S.M. were responsible for visualization and writing. N.B., N.Maring, N.S., and S.M. supervised the study.

A patent has been filed listing B.B., P.-E.E., and S.M. as inventors.

APPENDIX A: CONTEXTUAL FRACTION AND SIGNALING FRACTION

For a contextual n-partite game, we denote by p an associated empirical behavior, i.e., a set of probability distributions on the outputs conditioned on the inputs. We call a given choice of inputs a context, denoted C. In an ideal contextual game, the measurements in a context are compatible, so the marginal distributions computed from two different distributions on the intersection of their contexts are equal, i.e., they obey the generalized no-signaling condition. In a real implementation however the measurements of a context might not be perfectly compatible because of physical crosstalk. We nonetheless define the same contexts as in the ideal description of the game, which means that we can observe behaviors that do no satisfy the no-signaling conditions. We call $\mathcal{NS}$ the space of behaviors that does satisfy no-signaling and $\mathcal{E}$ the bigger space of all behaviors, i.e., the space of sets of real numbers for each context that satisfies the usual normalization and positiveness conditions.

A no-signaling behavior is noncontextual if the distribution for each context $\{p_C\}$ can be obtained as the marginal of a single distribution on global assignments [3]. The contextual fraction CF quantifies how contextual a behavior is and is the solution of a linear program (LP) [15]. It generalizes the nonlocal fraction [56]. For all behaviors p, it holds that $\mathsf{CF}(p) \in [0,1]$, that p is noncontextual if and only if $\mathsf{CF}(p) = 0$, and that p is maximally (or strongly) contextual if and only if $\mathsf{CF}(p) = 1$.

Similarly, the signaling fraction SF that we introduce in this work (simultaneously introduced in Ref. [29]) quantifies how far a behavior is from the no-signaling set. For an empirical behavior p, we consider affine decompositions of the form $p = sp' + (1-s)p''$, for which p' is a nonsignaling behavior. If s^* is the maximum weight that can be assigned to a nonsignaling p' in such a decomposition then we define the nonsignaling fraction as $\mathsf{NSF}(p) := s^*$. Then we define the signaling fraction $\mathsf{SF}(p) := (1-s^*)$, as this is the irreducibly signaling "fraction" of p. In this way it quantifies how far from $\mathcal{NS}$ a behavior is. We can thus decompose p as

$$p = \mathsf{NSF}(p)\, p^{\mathrm{NS}} + \mathsf{SF}(p)\, p^{\mathrm{SS}}, \tag{A1}$$

where p^{NS} is a no-signaling behavior and p^{SS} is a strongly signaling behavior, i.e., $\mathsf{SF}(p^{\mathrm{SS}}) = 1$. For all behaviors p, it holds that $\mathsf{SF}(p) \in [0,1]$, and that p is nonsignaling if and only if $\mathsf{SF}(p) = 0$. The signaling and nonsignaling fractions $\mathsf{SF}(p)$ can be computed by a linear program, similarly to $\mathsf{CF}(p)$.

One can study the potential decompositions of behaviors into HVMs. An HVM for a measurement scenario is defined by a set of hidden variables Λ, a distribution $h(\lambda)$ over Λ, and, for each λ, a behavior h^{λ} on the same measurement scenario. A behavior e is said to be realized by an HVM $\{\Lambda, h(\lambda), h^{\lambda}\}$ if it arises as the weighted average

$$p = \sum_{\lambda \in \Lambda} h(\lambda) h^{\lambda}. \tag{A2}$$

In this way HVMs provide a framework to reason about deeper or more fine-grained descriptions of hypothetical underlying processes that may be giving rise to an observed empirical behavior, and that in principle an adversary may seek to exploit. Note that, for finite measurement scenarios, signaling, nonsignaling, and noncontextual behaviors each arise as convex combinations of a finite number of vertex behaviors. These provide canonical hidden variable spaces, and when we wish to consider these properties, it thus suffices to consider averages over finite hidden-variable spaces (for extensions to the continuum, see Ref. [57]).

In the device-independent approach, it is thus crucial to examine all acceptable HVMs, i.e., all HVMs that are compatible with the underlying theory and the description of the experimental setup we implement, to bound the power of the eavesdropper. In particular, in the case of certified randomness, an HVM decomposition describes the eavesdropper's ability to predict the outputs of an experiment. In the case of a multipartite game implemented in a spacelike separated manner, all the behaviors of the HVMs must, for instance, be no signaling (the term "parameter independent" is sometimes used instead, when talking about hidden-variable properties [29]). Under these conditions, noncontextuality is equivalent to perfect predictability, because any noncontextual behavior can be decomposed into no-signaling deterministic behaviors [21].

The constraint that the elements of the HVM are no signaling can be replaced by an upper bound on their signaling fraction, if one has reasons to believe that the physical setup underlying the examined behavior allows some flow of information between the physical components implementing the measurements of each context, but only in a limited amount. Indeed, crosstalk between the components, which happens at the physical level, is related to signaling at the level of the behaviors. We assume here that the amount of signaling allowed at the hidden-variable level σ is not greater than the signaling SF_ℓ observed in the estimated behavior, which we define in the following way: we look for decompositions of the form $p = sp' + (1-s)p''$, for which p' is contained in $\mathsf{NPA}_\ell \subseteq \mathcal{NS}$, the ℓ th level of the NPA hierarchy [16,30], which approximates the set of behaviors achievable by performing quantum measurements on quantum states. The set NPA_ℓ is a strict subset of the no-signaling space. We project into NPA_ℓ rather than into $\mathcal{NS}$ because a no-signaling but supraquantum behavior can also be achieved by a signaling and quantum behavior. Similarly to Eq. (A1), we can thus decompose p as

$$p = [1 - \mathsf{SF}_\ell(p)]\, p^{\mathsf{NPA}_\ell} + \mathsf{SF}_\ell(p)\, p''. \tag{A3}$$

The signaling fraction SF_ℓ satisfies the following properties: SF_ℓ increases with ℓ, and, by definition, $\sigma \geq \mathsf{SF}$ and $\sigma \geq \mathsf{SF}_\ell$ for all ℓ if we assume that quantum mechanics is valid.

Note that the security of the protocol described in the next section relies only on Propositions 1 and 2, which do not rely on any assumption on the signaling σ. It can thus be straightforwardly adapted to another choice of σ, based on a semi-device-dependent characterization of the devices or related to different cryptographic assumptions, by replacing SF_ℓ step 6 of the protocol in Appendix C by the appropriate choice.

Another approach to derive a lower bound on the physical crosstalk from the observed behavior was proposed by Silman *et al.* [27]. It puts a constraint at the level of the measurements rather than on the behaviors [see Eq. (6) of Ref. [27]]. This problem being computationally intractable, the authors also proposed using the NPA hierarchy to obtain a lower bound. The cross-talk measure χ is smaller than SF_ℓ because the optimization space for SF_ℓ is a subspace of the optimization space for the former. Indeed, the second constraint of Eq. (6) of Ref. [27] is equivalent to

$$-\chi \leq p(ab|xy) - p^{\mathsf{NPA}_\ell}(ab|xy) \leq \chi \tag{A4}$$

with p^{NPA_ℓ} in NPA_ℓ, where, without loss of generality, we took the same state for both behaviors, while Eq. (A3) is equivalent to

$$p - p^{\mathsf{NPA}_\ell} = \mathsf{SF}_\ell(p)(p'' - p^{\mathsf{NPA}_\ell}), \tag{A5}$$

which implies Eq. (A4) for $\chi = \mathsf{SF}_\ell(p)$, as $-1 \leq p'' - p^{\mathsf{NPA}_\ell} \leq 1$.

Both metrics provide valid lower bounds on the crosstalk, as long as they are used coherently (i.e., one should compute the maximal score in the presence of crosstalk using χ if one uses χ to estimate crosstalk). The χ metrics are based on a statistical distance, while the SF metrics look for a convex decomposition. The advantages of our chosen approach are that it integrates well with convex optimization techniques used to compute maximal scores, Bell inequality violations, and guessing probabilities, and it allows us to leverage all duality properties that relate these quantities to Bell inequalities [15,58,59].

APPENDIX B: RELATION BETWEEN THE SCORE WITH BOUNDED CONTEXTUALITY AND THE SCORE WITH BOUNDED SIGNALING

We reformulate Proposition 2 in the more general case of an n-partite contextual game.

Proposition 3. For a contextual n-partite game with binary input choices for all players, let C be a distinguished context. Then

$$S_C^\sigma \leq S_{\mathrm{cl}}^{\xi=\sigma}. \tag{B1}$$

Proof. Let p^* be a solution for the optimization problem defining S_C^σ. We can decompose p^* as

$$p^* = (1-\tau)p' + \tau p'' \tag{B2}$$

with $\tau \in [0,\sigma]$ and p' a no-signaling behavior. Moreover, by definition, we must have $p'_C = p''_C = 1$. In the case of binary inputs, using the same construction as in Appendix D of Ref. [26], we can find a local HVM for any no-signaling behavior with binary inputs that is deterministic on a context, which implies that $\mathsf{CF}(p') = 0$. Equation (B2) is then a feasible point for the LP defining the contextual fraction of p^*, which in turn implies that

$$\mathsf{CF}(p^*) \leq \tau \leq \sigma, \tag{B3}$$

and p^* is thus a feasible point for the optimization problem corresponding to $S_{\text{cl}}^{\xi=\sigma}$. ■

Combining Propositions 1 and 3, we see that, for randomness to be certified, the observed score c/qN (see Fig. 6) must satisfy $c/qN > [\sigma(2^n - k) + k]/2^n$, which in turn implies that $\sigma < [2^n(c/qN) - k]/(2^n - k)$. This gives the upper bound on the amount of crosstalk that our analysis can tolerate while certifying randomness.

Arguments:

- G: an k-consistent n-party contextual game with binary inputs for each player and a distinguished context C
- N: the output length (a positive integer)
- q: the test probability (a real number in $[0,1]$)
- χ: the score threshold (a real number in $[0,1]$)
- ℓ: the level of the NPA hierarchy (a positive integer)

Variables:

- c: the number of wins in test rounds (a positive integer that we set to 0)
- $\hat{p}$: the estimated behavior (a table indexed by the input and output choices that we set to 0)

Protocol:

1. Choose a bit $t \in \{0,1\}$ according to the Bernoulli distribution $(1-q, q)$.
2. If $t = 1$ ("test round"), play G, record the input and output in $\hat{p}$, add score to c.
3. If $t = 0$ ("generation round"), input C and record the output.
4. Steps 2–4 are repeated $(N-1)$ more times.
5. Normalize $\hat{p}$ and compute $\mathsf{SF}_\ell(\hat{p})$.
6. If $c/(qN) - \mathsf{SF}_\ell(\hat{p})(2^n - k)/2^n < \chi$ then the protocol aborts. Otherwise, it succeeds.

FIG. 6. Protocol for the randomness generation protocol with nonzero crosstalk. The protocol is based on an n-player k-consistent nonlocal game with binary inputs, and it is secure provided that the signaling at the HV level σ is not greater than SF_ℓ. Compared to previous protocols that ignored the effect of crosstalk, the observed average score c/qN has to be greater than a fixed threshold plus the correction due to signaling for the protocol to succeed.

APPENDIX C: PROTOCOLS FOR RANDOMNESS GENERATION AND EXPANSION AGAINST CLASSICAL AND QUANTUM SIDE INFORMATION

To certify randomness against quantum side information, we use the protocol described in Fig. 6. We call **C** and **S** the sequence of context choices and outputs produced when the protocol is implemented, and E the side information accessible to an eavesdropper. Then, when the protocol succeeds, the δ-smooth min-entropy of the outputs **S** is at least

$$H_{\min}^{\delta}(\mathbf{S}|\mathbf{C},E) \geq N[\pi(\chi) - \Delta] \tag{C1}$$

with

$$\pi(\chi) = 2\frac{\log(e)(\chi - S_{\text{cl}})^2}{nd-1} \tag{C2}$$

and

$$\Delta = \frac{\log(2/\delta^2)}{N\varepsilon} + 2ndq + \frac{\varepsilon}{q}\frac{8\log(e)(\chi - S_{\text{cl}})^2}{(nd-1)^2} + \left(\frac{\varepsilon}{q}\right)^2\frac{32\log(e)(\chi - S_{\text{cl}})^3}{3(nd-1)^3}2^{(\varepsilon/q)4\log(e)(\chi - S_{\text{cl}})/(nd-1)} \tag{C3}$$

for any $\varepsilon \in]0,1]$, where $\log(e)$ is the base-2 logarithm of the exponential and d is the total number of outputs.

Bound (C1) is derived from the security proof introduced in Ref. [26]. It is valid because the compatibility assumption in Ref. [26], which translates to the no-signaling assumption at the behavior's level, is required only to bound the distinguished score for the game. We derived a new relaxed distinguished score with signaling, S_C^σ, and can then use the tools of Ref. [26] to lower bound the min-entropy with that new score. This score is, according to Propositions 1 and 3, the classical score in the absence of signaling S_{cl} increased by $\sigma(2^n - k)/2^n$. An alternative way to present our protocol and the associated bound would be to instead test $c/(qN) < \chi$ in step 6 and to replace S_{cl} by S_C^σ in Eqs. (C2) and (C3). The effect of signaling would then be visible in the expression of the bound rather than, as it is now, in the description of the protocol itself through the requirement on χ. These two approaches are equivalent because what matters in the min-entropy bound is the difference $\chi - S_{\text{cl}}$.

The expressions for π and Δ given by Eqs. (C2) and (C3) are obtained in the following way. The rate curve $\pi(\chi)$ is described by Theorem 5.8 of Ref. [26], where, for the CHSH game, $r = 4$ (the size of the output alphabet; nd in our notation) and $W_{G,\bar{a}} = 3/4$ [the score with distinguished input, which, for binary inputs, is equal to the classical score (S_{cl} in our notation), as proved in Appendix D of Ref. [26]]. The term $\log\big(2/\delta^2\big)/N\epsilon$ in Eq. (C3) follows from Theorem 3.2 of Ref. [26], rewriting it as $1 + 2\log(1/\delta) = \log\big(2/\delta^2\big)$. The other terms are expressed in big-O notation in Ref. [26], where the authors were interested in the asymptotic case, and we replace them by actual bounds on a finite number of rounds using a derivation similar to that used in Appendix G of Ref. [28]. More precisely, in Ref. [26], the term $O(q)$ in Proposition 6.8 comes from an induction on Propositions B.2 and B.3 together with the fact that $(1 - x)^\alpha \geq 1 - \alpha x$, to bound the term $\sum_x \langle \rho_{\bar{a}}^x \rangle_{1+\epsilon} / \langle \rho \rangle_{1+\epsilon}$ in Eq. (6.24), and can thus be replaced by $2ndq$ in our Eq. (C3). The term $O(\epsilon/q)$ comes from the Taylor expansion at order 3 of $x \mapsto 2^x$ around 0 in the proof of Proposition 6.3, and can thus be replaced by the last two terms of our Eq. (C3). The bound given by our Eq. (8) in the main text was optimized on ε, which led to taking $\varepsilon = 6 \times 10^{-5}$.

For the min-entropy bound to hold, the following assumptions have to be satisfied.

(1) The user implements the protocol in a secure lab from which information leakage can be prevented.
(2) This lab can be partitioned into two sites, corresponding to Alice and Bob, and the information transfer between these two sites can be upper bounded.
(3) Quantum mechanics is correct.
(4) The user has access to a trusted classical computer.
(5) The user has access to a source of private random numbers or of public random numbers that are independent of the state of the devices used to implement the protocol.

The first assumption has to be satisfied for any cryptographic applications. The second assumption enables us to address the locality loophole by introducing a measure of signaling, and the third assumption allows us to quantify it via the distance to the set of quantum correlations. The fourth assumption is required for the processing of the data output by the protocol. The last assumption is needed to achieve random expansion (first case) or private randomness generation (second case). No additional assumption, in particular on the inner working of the device, is required.

In order to derive a bound on the min-entropy against classical side information, one can use the protocols proposed in Refs. [31–33]. To adapt them to nonzero signaling and our measure of crosstalk, the adequate modified version of the guessing probability problem [58–60] with a fixed Bell inequality β is

$$\begin{aligned} G_C(I, \sigma, \ell) = \max_p \ \max_s \ & e_C(p) \\ \text{such that} \quad & p \in \mathcal{E}, \\ & \beta(p) = I, \\ & \mathsf{SF}_\ell(p) \leq \sigma, \end{aligned}$$

which is similar to that introduced by Silman *et al.* [27], but where the amount of signaling is bounded at the behavior level. It would be interesting to study how G_C varies as a function of σ and to compare it to the $P^*_{xy}(I, \chi)$ introduced in Eq. (2) of Ref. [13]. We leave this as future work.

APPENDIX D: RELATION BETWEEN CHSH VIOLATION AND PHOTON INDISTINGUISHABILITY

In this appendix, we derive the relation between the measured CHSH expression I_{CHSH} in our setup and the photon indistinguishability characterized by the HOM visibility V_{HOM}:

$$I_{\mathrm{CHSH}} = \sqrt{2}(V_{\mathrm{HOM}} + 1). \tag{D1}$$

This allows us to set the requirements on the single-photon sources used to violate Bell inequalities. We use the formalism of quantum creation operators to predict the behavior of two partially distinguishable photons in the photonic chip used in our optical setup. We calculate the expression of the output state, and use it to compute the coincidence probabilities. We then compute the expected behavior as a function of V_{HOM}, which yields the relation between I_{CHSH} and V_{HOM}.

We initialize the photonic chip by injecting one photon in the spatial mode 0_A and a second one in mode 0_B (see Fig. 1 in the main text). We assume in addition that the two injected photons are not necessarily identical. They could, for instance, not arrive exactly at the same time in the chip, have slightly different wavelengths, or not share the same polarization. The information about these degrees of freedom is encoded in configuration wave functions $|\alpha\rangle$ for the first photon and $|\beta\rangle$ for the second one. A photon is thus completely described by its mode and configuration, e.g., $|0_B : \alpha\rangle$ for a photon in mode 0_B and configuration $|\alpha\rangle$. The wave-function overlap is then $\langle\alpha|\beta\rangle$. If the photons are completely distinguishable, $\langle\alpha|\beta\rangle = 0$, and, on the contrary, if they are identical, $\langle\alpha|\beta\rangle = 1$ (up to a phase).

We denote by $a^\dagger_{0,\alpha}$, $a^\dagger_{1,\alpha}$, $b^\dagger_{0,\beta}$, and $b^\dagger_{0,\beta}$ and creation operators respectively associated with the states $|0_A : \alpha\rangle$, $|1_A : \alpha\rangle$, $|0_B : \beta\rangle$, and $|1_B : \beta\rangle$. We use the notation $|\mathrm{vac}\rangle$ for the vacuum state. The chip input state is then $|\psi_0\rangle = |0_A : \alpha, 0_B : \beta\rangle = a^\dagger_{0,\alpha} b^\dagger_{0,\beta} |\mathrm{vac}\rangle$. The first part of the chip is dedicated to generating a Bell state. We can write the full quantum state after the first column of symmetric beam splitters and the swap operation, before postselection, as

$$|\psi_1\rangle = \left(\frac{1}{2}a^\dagger_{0,\alpha}a^\dagger_{1,\beta} + \frac{i}{2}a^\dagger_{0,\alpha}b^\dagger_{1,\beta} + \frac{i}{2}b^\dagger_{0,\alpha}a^\dagger_{1,\beta} - \frac{1}{2}b^\dagger_{0,\alpha}b^\dagger_{1,\beta}\right)|\text{vac}\rangle. \tag{D2}$$

The Bell state is generated from this state by postselection, by keeping the results where Alice and Bob simultaneously measure a photon, which yields

$$|\psi_{\text{ent}}\rangle = \frac{1}{\sqrt{2}}a^\dagger_{0,\alpha}b^\dagger_{1,\beta}|\text{vac}\rangle + \frac{1}{\sqrt{2}}b^\dagger_{0,\alpha}a^\dagger_{1,\beta}|\text{vac}\rangle, \tag{D3}$$

where we discarded the global phase factor i.

When we set the phase shift $\psi = 0$, as is the case in our experiment, the unitary matrix associated with Alice and Bob's interferometer is [see Eq. (14)]

$$\widehat{U}(\phi, \psi = 0) = \begin{bmatrix} \sin(\phi/2) & \cos(\phi/2) \\ \cos(\phi/2) & -\sin(\phi/2) \end{bmatrix}, \tag{D4}$$

where we have discarded the global phase factor, because the interferometer outputs are connected to detectors. The quantum state at the exit of the chip is then

$$\begin{aligned} |\psi_{\text{out}}\rangle = &\frac{1}{\sqrt{2}}\left[\sin\left(\frac{\phi^A}{2}\right)a^\dagger_{0,\alpha} + \cos\left(\frac{\phi^A}{2}\right)a^\dagger_{1,\alpha}\right] \\ &\times\left[\cos\left(\frac{\phi^B}{2}\right)b^\dagger_{0,\beta} - \sin\left(\frac{\phi^B}{2}\right)b^\dagger_{1,\beta}\right]|\text{vac}\rangle \\ &+\frac{1}{\sqrt{2}}\left[\sin\left(\frac{\phi^B}{2}\right)b^\dagger_{0,\alpha} + \cos\left(\frac{\phi^B}{2}\right)b^\dagger_{1,\alpha}\right] \\ &\times\left[\cos\left(\frac{\phi^A}{2}\right)a^\dagger_{0,\beta} - \sin\left(\frac{\phi^A}{2}\right)a^\dagger_{1,\beta}\right]|\text{vac}\rangle. \end{aligned} \tag{D5}$$

Writing $|\beta\rangle = c_\parallel|\alpha\rangle + c_\perp|\alpha_\perp\rangle$ with $|\alpha_\perp\rangle$ a unit vector such that $\langle\alpha|\alpha_\perp\rangle = 0$, and $|c_\parallel|^2 + |c_\perp|^2 = 1$, we have $a^\dagger_\beta = c_\parallel a^\dagger_\alpha + c_\perp a^\dagger_{\alpha_\perp}$ and the output state is equal to

$$\begin{aligned} |\psi_{\text{out}}\rangle = &\frac{1}{\sqrt{2}}\left[\sin\left(\frac{\phi^A}{2}\right)a^\dagger_{0,\alpha} + \cos\left(\frac{\phi^A}{2}\right)a^\dagger_{1,\alpha}\right] \\ &\times\left[c_\parallel\cos\left(\frac{\phi^B}{2}\right)b^\dagger_{0,\alpha} + c_\perp\cos\left(\frac{\phi^B}{2}\right)b^\dagger_{0,\alpha_\perp}\right. \\ &\left.- c_\parallel\sin\left(\frac{\phi^B}{2}\right)b^\dagger_{1,\alpha} - c_\perp\sin\left(\frac{\phi^B}{2}\right)b^\dagger_{1,\alpha_\perp}\right]|\text{vac}\rangle \\ &+\frac{1}{\sqrt{2}}\left[\sin\left(\frac{\phi^B}{2}\right)b^\dagger_{0,\alpha} + \cos\left(\frac{\phi^B}{2}\right)b^\dagger_{1,\alpha}\right] \\ &\times\left[c_\parallel\cos\left(\frac{\phi^A}{2}\right)a^\dagger_{0,\alpha} + c_\perp\cos\left(\frac{\phi^A}{2}\right)a^\dagger_{0,\alpha_\perp}\right. \\ &\left.- c_\parallel\sin\left(\frac{\phi^A}{2}\right)a^\dagger_{1,\alpha} - c_\perp\sin\left(\frac{\phi^A}{2}\right)a^\dagger_{1,\alpha_\perp}\right]|\text{vac}\rangle. \end{aligned} \tag{D6}$$

In the protocol, the chip modes 0_A and 1_A correspond to Alice's results 0 and 1, respectively, and modes 0_B and 1_B correspond to Bob's results 0 and 1. Let $p(a,b|\phi^A,\phi^B)$ be the probability that Alice measures a and Bob measures b at the same time, knowing that their interferometer phases are ϕ^A and ϕ^B, respectively. We can compute it from the expression of $|\psi_{\text{out}}\rangle$ by pairing the creation operators according to (a,b) and summing the modulus square of the coefficients. For instance, to compute $p(0,0|\phi^A,\phi^B)$, we identify the pairs $a^\dagger_{0,\alpha}b^\dagger_{0,\alpha}$, $a^\dagger_{0,\alpha}b^\dagger_{0,\alpha_\perp}$, $b^\dagger_{0,\alpha}a^\dagger_{0,\alpha}$, and $b^\dagger_{0,\alpha}a^\dagger_{0,\alpha_\perp}$ in $|\psi_{\text{out}}\rangle$. Each of these pairs of creation operators, when applied on $|\text{vac}\rangle$, yields a state where Alice and Bob both measure the result 0. Note that $a^\dagger_{0,\alpha}b^\dagger_{0,\alpha} = b^\dagger_{0,\alpha}a^\dagger_{0,\alpha}$. We thus obtain

$$\begin{aligned} p(0,0|\phi^A,\phi^B) = &\left|\frac{c_\parallel}{\sqrt{2}}\left[\sin\left(\frac{\phi^A}{2}\right)\cos\left(\frac{\phi^B}{2}\right)\right.\right. \\ &\left.\left.+\sin\left(\frac{\phi^B}{2}\right)\cos\left(\frac{\phi^A}{2}\right)\right]\right|^2 \\ &+\left|\frac{c_\perp}{\sqrt{2}}\sin\left(\frac{\phi^A}{2}\right)\cos\left(\frac{\phi^B}{2}\right)\right|^2 \\ &+\left|\frac{c_\perp}{\sqrt{2}}\sin\left(\frac{\phi^B}{2}\right)\cos\left(\frac{\phi^A}{2}\right)\right|^2, \end{aligned} \tag{D7}$$

which amounts to

$$\begin{aligned} p(0,0|\phi^A,\phi^B) &= p(1,1|\phi^A,\phi^B) \\ &= \frac{|c_\parallel|^2-1}{8}\cos(\phi^A-\phi^B) \\ &\quad-\frac{|c_\parallel|^2+1}{8}\cos(\phi^A+\phi^B)+\frac{1}{4}, \end{aligned} \tag{D8}$$

$$\begin{aligned} p(0,1|\phi^A,\phi^B) &= p(1,0|\phi^A,\phi^B) \\ &= -\frac{|c_\parallel|^2-1}{8}\cos(\phi^A-\phi^B) \\ &\quad+\frac{|c_\parallel|^2+1}{8}\cos(\phi^A+\phi^B)+\frac{1}{4}. \end{aligned} \tag{D9}$$

Photon indistinguishability is commonly quantified with the HOM visibility. Consider a symmetric beam splitter. We inject two photons, one in each beam-splitter input,

TABLE II. Measured behavior as a function of single-photon HOM visibility V_{HOM}.

a	b	00	01	10	11
0	0	$\frac{1}{4}+\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}-\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}-\sqrt{2}/8$	$\frac{1}{4}-\sqrt{2}/8$
0	1	$\frac{1}{4}-\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}+\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}+\sqrt{2}/8$	$\frac{1}{4}+\sqrt{2}/8$
1	0	$\frac{1}{4}-\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}+\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}+\sqrt{2}/8$	$\frac{1}{4}+\sqrt{2}/8$
1	1	$\frac{1}{4}+\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}-\sqrt{2}V_{\text{HOM}}/8$	$\frac{1}{4}-\sqrt{2}/8$	$\frac{1}{4}-\sqrt{2}/8$

and measure the probability p_{coinc} of measuring a coincidence, that is, two photons going out of the beam splitter on different outputs. The HOM visibility is then defined as $V_{\text{HOM}} = 1 - 2p_{\text{coinc}}$. Indistinguishable photons will in this case always both come out of the same beam-splitter output, and no coincidences will be measured, yielding $V_{\text{HOM}} = 1$. For perfectly distinguishable photons, $p_{\text{coinc}} = 1/2$ and $V_{\text{HOM}} = 0$.

To relate $c_{\parallel}$ to V_{HOM}, we use the same mathematical treatment as above, i.e., we write the input state with creation operators, propagate the state in a symmetric beam splitter, and extract the probability of measuring one photon on each beam-splitter output, yielding $V_{\text{HOM}} = |\langle\alpha|\beta\rangle|^2 = |c_{\parallel}|^2$.

As a function of V_{HOM}, Eqs. (D8) and (D9) become

$$\begin{aligned} p(0,0|\phi^A,\phi^B) &= p(1,1|\phi^A,\phi^B) \\ &= \frac{V_{\text{HOM}}-1}{8}\cos(\phi^A-\phi^B) \\ &\quad - \frac{V_{\text{HOM}}+1}{8}\cos(\phi^A+\phi^B)+\frac{1}{4}, \end{aligned} \tag{D10}$$

$$\begin{aligned} p(0,1|\phi^A,\phi^B) &= p(1,0|\phi^A,\phi^B) \\ &= -\frac{V_{\text{HOM}}-1}{8}\cos(\phi^A-\phi^B) \\ &\quad + \frac{V_{\text{HOM}}+1}{8}\cos(\phi^A+\phi^B)+\frac{1}{4}. \end{aligned} \tag{D11}$$

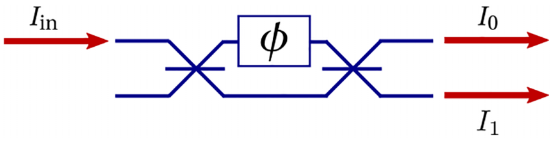

FIG. 7. Mach-Zehnder interferometer (MZI) in the configuration used to calibrate the voltages. Here I_{in} is the input power and I_0, I_1 the output ones. The MZI splitting $n_0 = I_0/(I_0 + I_1)$ is related to the MZI's phase ϕ by $n_0(\phi) = \sin^2(\phi/2)$ according to Eq. (14).

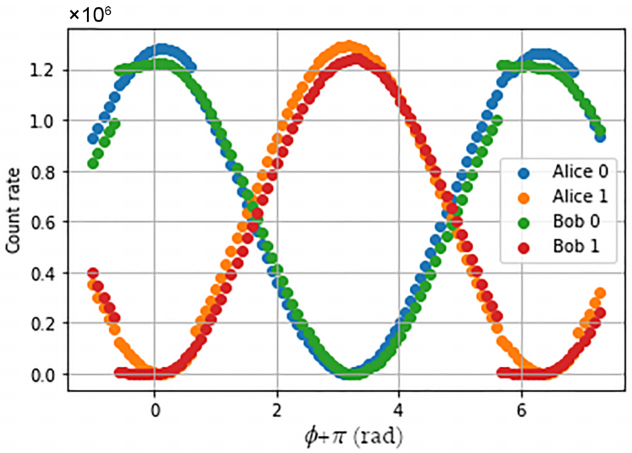

FIG. 8. Measured count rates on Alice's and Bob's outputs while sweeping ϕ^A and ϕ^B. Note that the heating resistors do not achieve a full 2π sweep.

The corresponding behavior is displayed in Table II, and its associated CHSH value is

$$I_{\text{CHSH}} = \sqrt{2}(V_{\text{HOM}} + 1). \tag{D12}$$

APPENDIX E: PHOTONIC CHIP CROSS-TALK MATRICES

Heat generated by Alice's and Bob's heating resistors propagates in the chip. We denote by V_1, V_2, V_3, and V_4 the voltages respectively applied on the phase shifters ψ_x^A, ψ_y^B, ϕ_x^A, and ϕ_x^B [see Fig. 1(a)]. The full characterization of the phases implemented by the phase shifters as a function of applied voltages is given with a typical error of the order of 0.1 rad by

$$\begin{aligned} \begin{bmatrix} \phi_Z^{(A)} \\ \phi_Z^{(B)} \end{bmatrix} &= \begin{bmatrix} 1.2890 & -0.0785 \\ 0.0988 & -1.2777 \end{bmatrix} \begin{bmatrix} V_1^2 \\ V_2^2 \end{bmatrix} \\ &\quad - \begin{bmatrix} 0.0192 & -0.0009 \\ 0.0012 & -0.0203 \end{bmatrix} \begin{bmatrix} V_1^4 \\ V_2^4 \end{bmatrix}, \end{aligned} \tag{E1}$$

TABLE III. Target phases and MZI splittings for each measurement context.

Measurement context	Target ϕ^A	Target n_0^A	Target ϕ^B	Target n_0^B
$(x=0, y=0)$	$-\pi/2$	0.5	$-\pi/4$	0.146
$(x=0, y=1)$	$-\pi/2$	0.5	$\pi/4$	0.146
$(x=1, y=0)$	0	0	$-\pi/4$	0.146
$(x=1, y=1)$	0	0	$\pi/4$	0.146

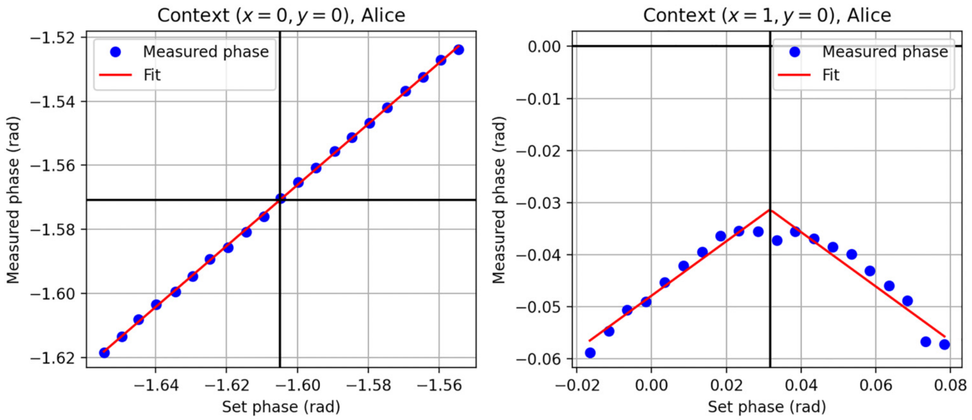

FIG. 9. For each measurement context, the input phase of Alice’s and Bob’s MZI is swept around the target phase and we record the measured phase from the MZI splitting. Data are fitted with a line or a triangle depending on the measurement context. Here we show the result for Alice in contexts 00 and 01. The horizontal black line indicates the target phase for each context, and the vertical one represents the input phase that should be used to implement the context, knowing that, in general, the input phase is not equal to the measured phase. Note that, for Alice, in contexts $(x = 1, y = 0)$ and $(x = 1, y = 1)$, the measured phase should be π at the triangle’s peak. This is due to dark counts, which prevent the measured interferometer balance from going to 0, and thus we rely on a triangular fit of the sweep.

$$\begin{bmatrix}\phi_Y^{(A)}\\ \phi_Y^{(B)}\end{bmatrix} = \begin{bmatrix}0.2703\\ -0.2799\end{bmatrix} + \begin{bmatrix}1.4693 & -0.1111\\ 0.1120 & -1.4776\end{bmatrix}\begin{bmatrix}V_3^2\\ V_4^2\end{bmatrix} - \begin{bmatrix}0.0351 & -0.0026\\ 0.0027 & -0.0343\end{bmatrix}\begin{bmatrix}V_3^4\\ V_4^4\end{bmatrix}, \tag{E2}$$

where the phases are expressed in radians as a function of the voltages in volts applied on each heating resistor. Injecting a continuous diode laser in the chip and measuring the outputs with photodiodes confirms that there is measurable thermal crosstalk only between resistors belonging to the same column in Fig. 1(a) in the main text.

APPENDIX F: PHASE CALIBRATION PROTOCOL

The goal of the voltage calibration protocol is to compute the voltages to apply on Alice’s and Bob’s Mach-Zehnder interferometer (MZI) phases, such that the measurement contexts presented in Table I can be implemented

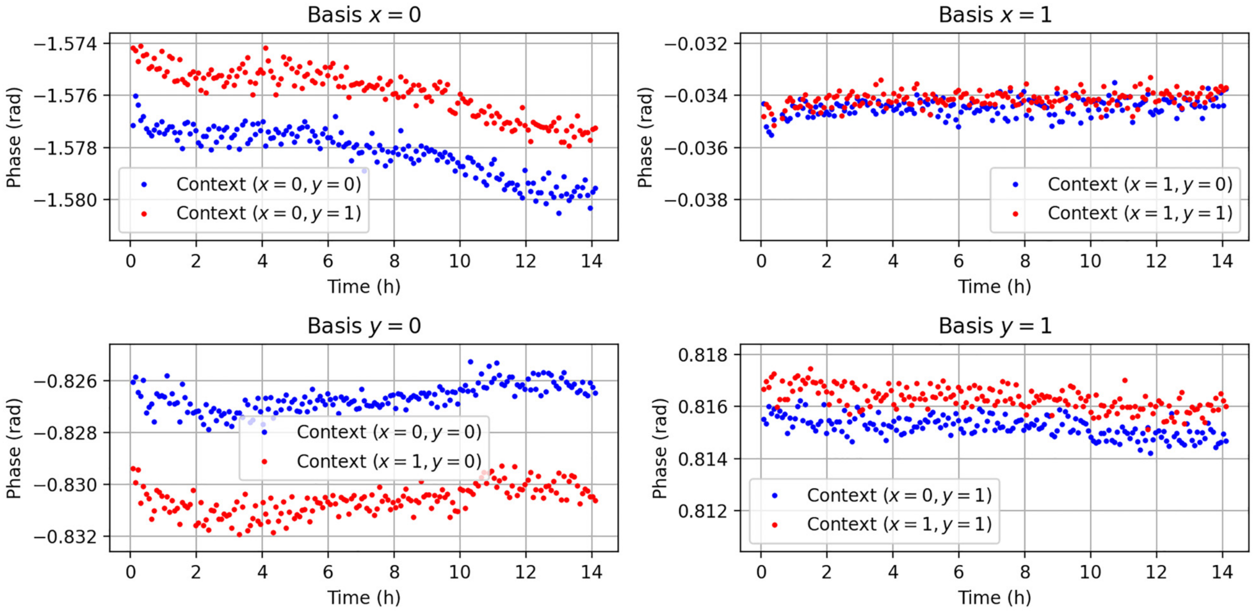

FIG. 10. For each measurement context, the current MZI phase is measured using the MZI’s balance at Alice’s and Bob’s outputs. For each plot, the vertical scale is 2 mrad per division.

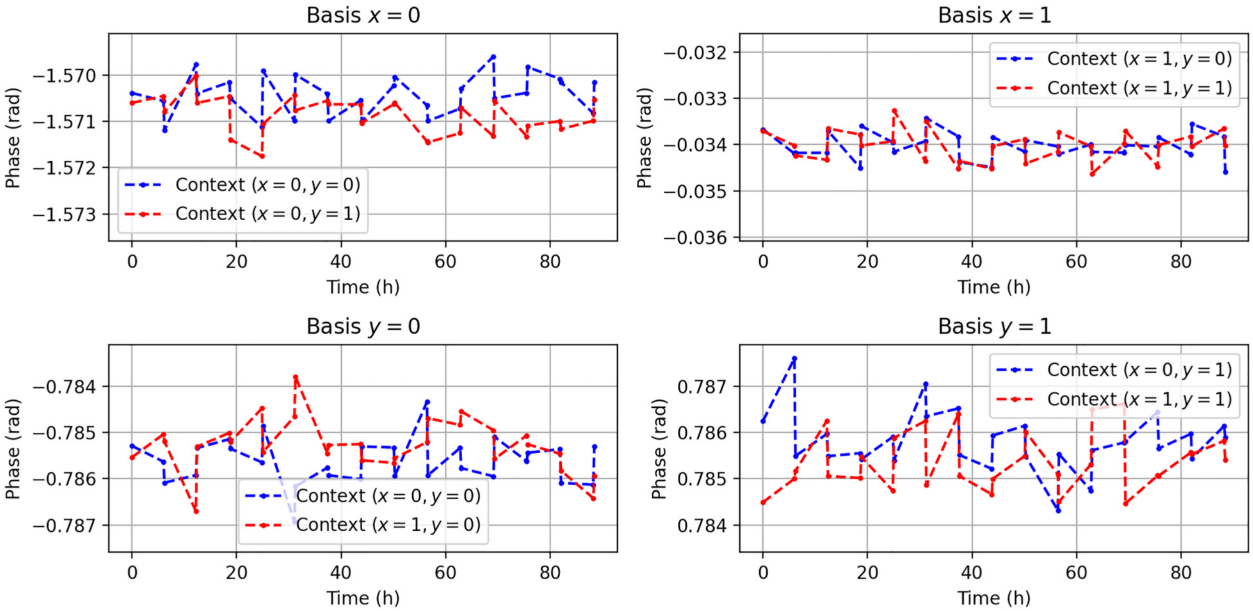

FIG. 11. For each measurement context, the current MZI phase is measured every 6 h before and after a voltage calibration using the MZI's balance at Alice's and Bob's outputs. For each plot, the vertical scale is 1 mrad per division.

with high precision. The voltages are regularly calibrated during the execution of the protocol because the polarization of photons in the fibers between the chip and the detectors fluctuate, causing detection efficiency fluctuations on the detectors that are sensitive to polarization. We compensate for these by shifting the MZI phase. Because we cannot rely on the phases computed from the cross-talk matrix relations [Eqs. (E1) and (E2)] to accurately apply a phase on the interferometer, we use Alice's and Bob's MZI splitting as a measure of the implemented phase (see Fig. 7). To do so, the motorized shutter in Fig. 2 is closed, such that only the upper input modes of Alice's and Bob's MZI provide photons.

1. Relative detector efficiency measurement

First, ϕ^A and ϕ^B are swept simultaneously while recording the count rate on Alice's and Bob's outputs, which produces the data presented in Fig. 8.

We call $N^A_{0,\max}$ the maximum count rate recorded during the sweep by Alice on her 0 output, and similarly for $N^A_{1,\max}$, $N^B_{0,\max}$, and $N^B_{1,\max}$. Instead of using the raw detector count rates N^A_0, we worked with the normalized count rates $\tilde{N}^A_0 = N^A_0/N^A_{0,\max}$ to compensate for the different detector efficiencies. The corrected MZI splittings are $\tilde{n}^A_0 = \tilde{N}^A_0/\left(\tilde{N}^A_0 + \tilde{N}^A_1\right), \tilde{n}^B_0 = \tilde{N}^B_0/\left(\tilde{N}^B_0 + \tilde{N}^B_1\right)$ [61].

2. Local phase sweeps

For each measurement context, ϕ^A and ϕ^B are swept simultaneously around the target phase while recording the corrected MZI splittings. The target phases and MZI splittings for each context are displayed in Table III. From the measured MZI splittings, we deduce the measured interferometer phase: $\phi^A_{\text{measured}} = 2\arcsin\left(\sqrt{\tilde{n}^A_0}\right)$, $\phi^B_{\text{measured}} = 2\arcsin\left(\sqrt{\tilde{n}^B_0}\right)$. The plots of the measured phases as a function of the input phases are displayed in Fig. 9. The data are processed with a linear or triangular fit depending on the context, and the input phase that should be applied to the chip is retrieved from the vertical lines on the plots.

3. Phase drifts

We measured the phases of Alice's and Bob's MZIs for 14 h, always using the same voltages to implement the four measurement contexts, and by cycling through them in the order $(x = 0, y = 0)$, $(x = 0, y = 1)$, $(x = 1, y = 0)$, and $(x = 1, x = 1)$. The results are summarized in Fig. 10. Overall, we observe a typical drift of the implemented phase of the order of 0.25 mrad/h.

4. Phase stabilization

In Fig. 11 we show Alice's and Bob's MZI phases measured every 6 h before and after calibration during our main experiment. As a result of these frequent calibrations, the implemented phases stayed confined in an interval of 3 mrad around the targets.

[1] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[2] S. Kochen and E. P. Specker, in *The Logico-Algebraic Approach to Quantum Mechanics*, The University of Western Ontario Series in Philosophy of Science (Springer, Dordrecht, 1975), p. 293.

[3] S. Abramsky and A. Brandenburger, The sheaf-theoretic structure of non-locality and contextuality, New J. Phys. **13**, 113036 (2011).

[4] A. Cabello, S. Severini, and A. Winter, Graph-theoretic approach to quantum correlations, Phys. Rev. Lett. **112**, 040401 (2014).

[5] A. Acín, T. Fritz, A. Leverrier, and A. B. Sainz, A combinatorial approach to nonlocality and contextuality, Commun. Math. Phys. **334**, 533 (2015).

[6] R. Colbeck, Quantum and relativistic protocols for secure multi-party computation, Ph.D. thesis, University of Cambridge, 2007.

[7] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).

[8] J.-Å. Larsson, Loopholes in Bell inequality tests of local realism, J. Phys. A: Math. Theor. **47**, 424003 (2014).

[9] P. Bierhorst, E. Knill, S. Glancy, Y. Zhang, A. Mink, S. Jordan, A. Rommal, Y.-K. Liu, B. Christensen, S. W. Nam, M. J. Stevens, and L. K. Shalm, Experimentally generated randomness certified by the impossibility of superluminal signals, Nature **556**, 223 (2018).

[10] Y. Liu, Q. Zhao, M.-H. Li, J.-Y. Guan, Y. Zhang, B. Bai, W. Zhang, W.-Z. Liu, C. Wu, X. Yuan, H. Li, W. J. Munro, Z. Wang, L. You, J. Zhang, X. Ma, J. Fan, Q. Zhang, and J.-W. Pan, Device-independent quantum random-number generation, Nature **562**, 548 (2018).

[11] Y. Zhang, L. K. Shalm, J. C. Bienfang, M. J. Stevens, M. D. Mazurek, S. W. Nam, C. Abellán, W. Amaya, M. W. Mitchell, H. Fu, C. A. Miller, A. Mink, and E. Knill, Experimental low-latency device-independent quantum randomness, Phys. Rev. Lett. **124**, 010505 (2020).

[12] M.-H. Li, X. Zhang, W.-Z. Liu, S.-R. Zhao, B. Bai, Y. Liu, Q. Zhao, Y. Peng, J. Zhang, Y. Zhang, W. J. Munro, X. Ma, Q. Zhang, J. Fan, and J.-W. Pan, Experimental realization of device-independent quantum randomness expansion, Phys. Rev. Lett. **126**, 050503 (2021).

[13] L. K. Shalm, Y. Zhang, J. C. Bienfang, C. Schlager, M. J. Stevens, M. D. Mazurek, C. Abellán, W. Amaya, M. W. Mitchell, M. A. Alhejji, H. Fu, J. Ornstein, R. P. Mirin, S. W. Nam, and E. Knill, Device-independent randomness expansion with entangled photons, Nat. Phys. **17**, 452 (2021).

[14] W.-Z. Liu, M.-H. Li, S. Ragy, S.-R. Zhao, B. Bai, Y. Liu, P. J. Brown, J. Zhang, R. Colbeck, J. Fan, Q. Zhang, and J.-W. Pan, Device-independent randomness expansion against quantum side information, Nat. Phys. **17**, 448 (2021).

[15] S. Abramsky, R. S. Barbosa, and S. Mansfield, Contextual fraction as a measure of contextuality, Phys. Rev. Lett. **119**, 050504 (2017).

[16] M. Navascués, S. Pironio, and A. Acín, Bounding the set of quantum correlations, Phys. Rev. Lett. **98**, 010401 (2007).

[17] I. Marcikic, H. de Riedmatten, W. Tittel, H. Zbinden, M. Legré, and N. Gisin, Distribution of time-bin entangled qubits over 50 km of optical fiber, Phys. Rev. Lett. **93**, 180502 (2004).

[18] E. M. González-Ruiz, S. K. Das, P. Lodahl, and A. S. Sørensen, Violation of Bell's inequality with quantum-dot single-photon sources, Phys. Rev. A **106**, 012222 (2022).

[19] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, Phys. Rev. Lett. **23**, 880 (1969).

[20] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, Consequences and limits of nonlocal strategies, ArXiv:0404076.

[21] A. Fine, Hidden variables, joint probability, and the Bell inequalities, Phys. Rev. Lett. **48**, 291 (1982).

[22] S. Abramsky and L. Hardy, Logical Bell inequalities, Phys. Rev. A **85**, 062114 (2012).

[23] Note that there is a typo in the statement of Theorem 4 of Ref. [15]: the denominator on the right-hand side should be n, not k, which is equal to m in our case.

[24] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, Incompatible measurements in quantum information science, ArXiv:2112.06784.

[25] G. C. Ghirardi, A. Rimini, and T. Weber, A general argument against superluminal transmission through the quantum mechanical measurement process, Lett. Nuovo Cimento **27**, 293 (1980).

[26] C. A. Miller and Y. Shi, Universal security for randomness expansion from the spot-checking protocol, SIAM J. Comput. **46**, 1304 (2017).

[27] J. Silman, S. Pironio, and S. Massar, Device-independent randomness generation in the presence of weak cross-talk, Phys. Rev. Lett. **110**, 100504 (2013).

[28] M. Um, Q. Zhao, J. Zhang, P. Wang, Y. Wang, M. Qiao, H. Zhou, X. Ma, and K. Kim, Randomness expansion secured by quantum contextuality, Phys. Rev. Appl. **13**, 034077 (2020).

[29] K. Vallée, P.-E. Emeriau, B. Bourdoncle, A. Sohbi, S. Mansfield, and D. Markham, Corrected Bell and noncontextuality inequalities for realistic experiments, ArXiv:2310.19383.

[30] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, New J. Phys. **10**, 073013 (2008).

[31] S. Pironio and S. Massar, Security of practical private randomness generation, Phys. Rev. A **87**, 012336 (2013).

[32] S. Fehr, R. Gelles, and C. Schaffner, Security and composability of randomness expansion from Bell inequalities, Phys. Rev. A **87**, 012335 (2013).

[33] O. Nieto-Silleras, C. Bamps, J. Silman, and S. Pironio, Device-independent randomness generation from several Bell estimators, New J. Phys. **20**, 023049 (2018).

[34] R. Renner, Security of quantum key distrubtion, Ph.D. thesis, ETH Zurich, 2005.

[35] N. Somaschi, V. Giesz, L. De Santis, J. C. Loredo, M. P. Almeida, G. Hornecker, S. L. Portalupi, T. Grange, C. Antón, J. Demory, C. Gómez, I. Sagnes, N. D. Lanzillotti-Kimura, A. Lemaítre, A. Auffeves, A. G. White, L. Lanco, and P. Senellart, Near-optimal single-photon sources in the solid state, Nat. Photonics **10**, 340 (2016).

[36] R. Loudon, *The Quantum Theory of Light* (Oxford University Press, Oxford, 2000), p. 107, 227.

[37] H. Ollivier, S. E. Thomas, S. C. Wein, I. M. de Buy Wenniger, N. Coste, J. C. Loredo, N. Somaschi, A. Harouri,

A. Lemaitre, I. Sagnes, L. Lanco, C. Simon, C. Anton, O. Krebs, and P. Senellart, Hong-Ou-Mandel interference with imperfect single photon sources, Phys. Rev. Lett. **126,** 063602 (2021).

[38] N. Heurtel, A. Fyrillas, G. D. Gliniasty, R. Le Bihan, S. Malherbe, M. Pailhas, E. Bertasi, B. Bourdoncle, P.-E. Emeriau, R. Mezher, L. Music, N. Belabas, B. Valiron, P. Senellart, S. Mansfield, and J. Senellart, Perceval: A software platform for discrete variable photonic quantum computing, Quantum **7,** 931 (2023).

[39] A. V. Kuhlmann, J. Houel, A. Ludwig, L. Greuter, D. Reuter, A. D. Wieck, M. Poggio, and R. J. Warburton, Charge noise and spin noise in a semiconductor quantum device, Nat. Phys. **9,** 570 (2013).

[40] H. Krawczyk, in *Advances in Cryptology—EUROCRYPT '95*, edited by L. C. Guillou and J.-J. Quisquater (Springer Berlin Heidelberg, Berlin, Heidelberg, 1995), p. 301.

[41] D. Knuth and A. Yao, in *Algorithms and Complexity: New Directions and Recent Results* (Academic Press, New York, 1976), p. 357.

[42] T. S. Hao and M. Hoshi, Interval algorithm for random number generation, IEEE Trans. Inf. Theory **43**, 599 (1997).

[43] B. Bai, J. Huang, G.-R. Qiao, Y.-Q. Nie, W. Tang, T. Chu, J. Zhang, and J.-W. Pan, 18.8 Gbps real-time quantum random number generator with a photonic integrated chip, Appl. Phys. Lett. **118,** 264001 (2021).

[44] J. Wang, S. Paesani, Y. Ding, R. Santagati, P. Skrzypczyk, A. Salavrakos, J. Tura, R. Augusiak, L. Mančinska, D. Bacco, D. Bonneau, J. W. Silverstone, Q. Gong, A. Acín, K. Rottwitt, L. K. Oxenløwe, J. L. O'Brien, A. Laing, and M. G. Thompson, Multidimensional quantum entanglement with large-scale integrated optics, Science **360,** 285 (2018).

[45] F. Dupuis, O. Fawzi, and R. Renner, Entropy accumulation, Commun. Math. Phys. **379,** 867 (2020).

[46] F. Dupuis and O. Fawzi, Entropy accumulation with improved second-order term, IEEE Trans. Inf. Theory **65,** 7596 (2019).

[47] R. Arnon-Friedman, R. Renner, and T. Vidick, Simple and tight device-independent security proofs, SIAM J. Comput. **48,** 181 (2019).

[48] P. J. Brown, S. Ragy, and R. Colbeck, A framework for quantum-secure device-independent randomness expansion, IEEE Trans. Inf. Theory **66,** 2964 (2020).

[49] P. Brown, H. Fawzi, and O. Fawzi, Device-independent lower bounds on the conditional von Neumann entropy, ArXiv:2106.13692.

[50] D. Janner, D. Tulli, M. García-Granda, M. Belmonte, and V. Pruneri, Micro-structured integrated electro-optic $LiNbO_3$ modulators, Laser Photonics Rev. **3,** 301 (2009).

[51] N. Quack, H. Sattari, A. Y. Takabayashi, Y. Zhang, P. Verheyen, W. Bogaerts, P. Edinger, C. Errando-Herranz, and K. B. Gylfason, MEMS-enabled silicon photonic integrated devices and circuits, IEEE J. Quantum Electron. **56,** 1 (2020).

[52] A. M. Barth, S. Lüker, A. Vagov, D. E. Reiter, T. Kuhn, and V. M. Axt, Fast and selective phonon-assisted state preparation of a quantum dot by adiabatic undressing, Phys. Rev. B **94,** 045306 (2016).

[53] M. Cosacchi, F. Ungar, M. Cygorek, A. Vagov, and V. M. Axt, Emission-frequency separated high quality single-photon sources enabled by phonons, Phys. Rev. Lett. **123,** 017403 (2019).

[54] C. Gustin and S. Hughes, Efficient pulse-excitation techniques for single photon sources from quantum dots in optical cavities, Adv. Quantum Technol. **3,** 1900073 (2020).

[55] S. E. Thomas, M. Billard, N. Coste, S. C. Wein, Priya, H. Ollivier, O. Krebs, L. Tazaïrt, A. Harouri, A. Lemaitre, I. Sagnes, C. Anton, L. Lanco, N. Somaschi, J. C. Loredo, and P. Senellart, Bright polarized single-photon source based on a linear dipole, Phys. Rev. Lett. **126,** 233601 (2021).

[56] A. C. Elitzur, S. Popescu, and D. Rohrlich, Quantum nonlocality for each pair in an ensemble, Phys. Lett. A **162,** 25 (1992).

[57] R. S. Barbosa, T. Douce, P.-E. Emeriau, E. Kashefi, and S. Mansfield, Continuous-variable nonlocality and contextuality, Commun. Math. Phys. **391,** 1047 (2022).

[58] J.-D. Bancal, L. Sheridan, and V. Scarani, More randomness from the same data, New J. Phys. **16,** 033011 (2014).

[59] O. Nieto-Silleras, S. Pironio, and J. Silman, Using complete measurement statistics for optimal device-independent randomness evaluation, New J. Phys. **16,** 013035 (2014).

[60] A. Acín, S. Massar, and S. Pironio, Randomness versus nonlocality and entanglement, Phys. Rev. Lett. **108,** 100402 (2012).

[61] It was realized after the data acquisition that the detector efficiency should in fact not be accounted for, because correcting for that efficiency introduces biases in the measured empirical table that reduce the CHSH violation. In practice, the detector efficiencies differ only at the third significant digit, so the impact on the obtained results is negligible: correcting for these biases diminishes the maximum obtainable CHSH violation by only 10^{-4} (simulation using the Perceval package in PYTHON). In addition, in principle, these biases do not increase the signaling fraction. Therefore, we can still exploit the results from our experiment.