Receiver-Device-Independent Quantum Key Distribution

Marie Ioannou, Maria Ana Pereira, Davide Rusca, Fadri Grünenfelder, Alberto Boaron, Matthieu Perrenoud, Alastair A. Abbott, Pavel Sekatski, Jean-Daniel Bancal, Nicolas Maring, et al.

Receiver-Device-Independent Quantum Key Distribution

Marie Ioannou[1], Maria Ana Pereira[1], Davide Rusca[1], Fadri Grünenfelder[1], Alberto Boaron[1], Matthieu Perrenoud[1], Alastair A. Abbott[1,2], Pavel Sekatski[1], Jean-Daniel Bancal[1,3], Nicolas Maring[1], Hugo Zbinden[1], and Nicolas Brunner[1]

[1]Department of Applied Physics University of Geneva, 1211 Geneva, Switzerland
[2]Univ. Grenoble Alpes, Inria, 38000 Grenoble, France
[3]Université Paris-Saclay, CEA, CNRS, Institut de physique théorique, 91191, Gif-sur-Yvette, France

We present protocols for quantum key distribution in a prepare-and-measure setup with an asymmetric level of trust. While the device of the sender (Alice) is partially characterized, the receiver's (Bob's) device is treated as a black-box. The security of the protocols is based on the assumption that Alice's prepared states have limited overlaps, but no explicit bound on the Hilbert space dimension is required. The protocols are immune to attacks on the receiver's device, such as blinding attacks. The users can establish a secret key while continuously monitoring the correct functioning of their devices through observed statistics. We report a proof-of-principle demonstration, involving mostly off-the-shelf equipment, as well as a high-efficiency superconducting nanowire detector. The possibility to establish a secret key is demonstrated over a 4.8 km low-loss optical fiber with a finite-size analysis. The prospects of implementing these protocols over longer distances is discussed.

1 Introduction

Quantum communication has witnessed an extremely fast evolution over the last two decades [1, 2, 3]. On the practical level, record-distance implementations of quantum key distribution (QKD) have been reported [4, 5, 6], and first commercial systems have been developed [7]. On the more fundamental level, significant progress has been achieved as well, notably through the development of the concept of device-independent (DI) QKD [8, 9, 10, 11]. The observation of strong nonlocal quantum correlations allows distant users to establish a secret in a black-box setting, i.e., without relying on a detailed characterization of their cryptographic devices. This represents the strongest form of security for QKD [12].

From a practical point of view, the concept of DI QKD has also generated interest, in particular as a potential solution for countering experimentally demonstrated hacking attacks [13, 14]. While first proof-of-principle experiments have just been reported using state-of-the-art setups [15, 16, 17], any practical implementation of DI QKD is still arguably far out of reach.

This motivates research on more general scenarios for quantum communication where trust is relaxed on some of the observers or devices. The most well-known approach is that of Measurement-Device-Independent (MDI) QKD [18, 19, 1] where the honest parties (Alice and Bob) both send a quantum system to an intermediate third party (Charlie) performing a joint measurement. Security can be demonstrated without any assumption on Charlie's device, the protocol being in this sense MDI. However, the devices of both Alice and Bob must be characterized and cannot be treated as black boxes.

Another relevant scenario is the one where trust is relaxed on one of the honest parties. Consider for instance that Alice's device is trusted while Bob's device is viewed as a black-box. In practice, such an asymmetric scenario can be well-motivated, considering for example quantum communication between a large company and some end-user. On the one hand, the company has access to advanced technology and can verify the correct operation of its setup. On the other hand, the end user has only very limited resources and no possibility to verify the correct functioning of their cryptographic device.

The above scenario, referred to as one-sided DI,

has been introduced in [20], and key rates have been derived considering the effect of noise and finite-size data [21, 22]. However, the effect of losses has not been considered in these works. Instead, a fair-sampling type assumption is made, which opens the door to the detection loophole and to attacks such as blinding [13, 14], which impose severe requirements in terms of transmission and detection efficiency [23]. In practice, where losses are unavoidable, such an approach can thus no longer be considered one-sided DI. An alternative approach was followed in Ref. [24], considering an entanglement-based one-sided DI setup, establishing a connection to quantum steering. The implementation of such a setup is however challenging, as it requires a similar level of complexity compared to a fully DI protocol (notably in terms of detector efficiency) which explains why it has not been experimentally demonstrated so far.

In this work we discuss QKD protocols in a prepare-and-measure scenario where the sender's device is (partially) trusted while the receiver's device can be treated as a black-box. We term these protocols "receiver-DI". Our approach is based on the assumption that the prepared states have limited overlaps (i.e., we assume a bound on how distinguishable the states are from each other), inspired by recent developments in quantum randomness generation [25, 26, 27] and quantum correlations in prepare-and-measure scenarios [28]. Our approach can be classified as semi-DI [29] and one-sided DI [20], but differs from previous proposals since (i) we do not need an explicit bound on the Hilbert space dimension of the quantum systems (as in Refs [29, 30]), and (ii) we do not rely on any type of fair-sampling assumption (as in Refs [21, 22]).

In our protocols, the users can establish a secret key while continuously monitoring the quantum channel (as in any QKD protocol), but also continuously verifying (or self-testing) the correct operation of their devices. For Bob's device (and the communication channel) this verification procedure is performed based only on the observed statistics, similarly to the full DI model. Our protocols are therefore black-box on Bob's side, and do not rely on any physical model of the detector. This is a sensitive point in standard QKD protocols, because, contrary to the devices used for state preparation, the input of the detectors can be directly controlled by the eavesdropper Eve, as demonstrated in practice via so-called blinding attacks [13, 14]. In contrast, our protocols are immune to attacks on Bob's device.

Alice's device requires a partial characterisation via the bounds on the overlap of the emitted states. We argue that this assumption can be rather well justified in practice, and our implementation features a monitoring module allowing Alice to ensure the validity of the bounds on the overlaps. In practice, the introduction of this additional assumption considerably simplifies the implementation: a prepare-and-measure setup can be used, and low detection efficiencies can be tolerated (contrary to full DI [9, 31, 32, 33, 34, 35] or entanglement-based one-sided DI protocols [24]).

After presenting our protocols and their security analysis, we report on a proof-of-principle experiment. The setup involves mostly off-the-shelf equipment, with the addition of a high-efficiency superconducting nano-wire single-photon detector (SNSPD). For our simplest two-state protocol, an expected secret key rate of order $10^{-2} \sim 10^{-3}$ per round is demonstrated over a 4.8 km low-loss optical fiber, taking finite-size statistics into account. We illustrate the self-testing feature of the protocol, by showing that an artificial decrease of Bob's detector efficiency (as, e.g., in a blinding attack) is immediately detected by the users. Finally, we show that a three-state protocol can tolerate more loss than the two-state variation, and discuss more generally the prospects of implementations over longer distances.

2 Scenario

We consider a prepare-and-measure setup, where Alice sends quantum systems to Bob who per-

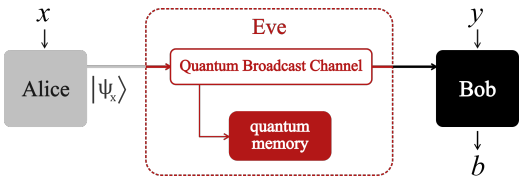

Figure 1: Scenario: Based on the observed data $p(b|x,y)$, and the assumption that Alice's preparations $|\psi_x\rangle$ have bounded overlap (see text), Alice and Bob can establish a secret key. Eve controls the quantum channel, but can also have full knowledge of the functioning of the devices of Alice and Bob.

forms measurements (see Fig. 1). In each round, Alice prepares her system in one of the n possible states $\{|\psi_x\rangle\}_{x=0}^{n-1}$ and sends it to Bob (note that the restriction to pure states is not necessary, see below). Bob then performs one of n possible measurements, labelled by $y = 0, ..., n-1$, and obtains a binary outcome $b = 0, 1$. After sufficiently many repetitions, Alice and Bob can estimate the conditional probability distribution $p(b|x, y)$ [1]. The security of the protocols, as given by a lower bound on the raw secret key rate (obtained after a post-processing step described below), is guaranteed based on the observed statistics $p(b|x, y)$, given that the setup complies with several assumptions that we now specify.

We begin with general assumptions common to all QKD protocols, including the DI case. (i) The choices of state preparation x and measurement setting y are made independently from Eve, i.e., she can not predict these values better than at random. (ii) No information about x and y leaks to Eve, except via the quantum and classical communication specified in the protocol at the given round. (iii) We assume the validity of quantum physics (note, that some DI protocols do not require this).

The central assumption specific to our protocols concerns the relation between the various states prepared by Alice. More precisely, (iv) we assume that their respective pairwise (possibly complex) overlaps $\gamma_{ij} = \langle\psi_i|\psi_j\rangle$ are bounded. One can think of the states $|\psi_x\rangle$ as describing the quantum system prepared by Alice's device and sent through the communication channel. More generally, they describe the states of all systems outside of Alice's lab conditioned on her applying the preparation sequence labeled by x. Note that if the states prepared by Alice are mixed, we require that they admit purifications that satisfy the overlap bounds – the purifying system can be attributed to Alice's lab by assumption and does not compromise the security. The bound on the overlaps makes the no-leakage assumptions for x in (ii) redundant; indeed it forbids the existence of a side-channel leaking any additional knowledge of x to Eve. Nevertheless if such a side-channel is possible, it can be accounted for by adapting the overlaps γ_{ij}.

Note that we do not require to explicitly specify the relevant degrees of freedom or the Hilbert space dimension of the any quantum system. Loosely speaking, only the relative distinguishability of the states matters. Thus, Alice's device is partially characterized, but prone to unavoidable errors due to technical imperfections.

Concerning the receiver (Bob), no characterisation of their device is required and no fair-sampling type assumption is used. In particular, our protocols are robust to attacks where Eve controls Bob's device [13], which can compromise the security of standard QKD protocols. This strong security comes however at a certain price, namely that the protocol is sensitive to losses. Importantly, this is not a particular weakness of our protocol, but a general feature of any QKD protocol that is device-independent on Bob's side (even considering a fully trusted Alice). Indeed, the possibility for Eve to perform a blinding attack sets severe bounds on the allowed transmission of the channel η. Specifically, no secret key can be obtained when $\eta \leq 1/n$ [23]; note that in order to overcome this limit, one may add the fair-sampling assumption, as in Ref. [21]. As we will see below, our protocols can reach this limit (i.e. provide a positive key rate when $\eta \to 1/n$), and are therefore optimally robust to loss in the receiver-device-independent scenario.

Finally, note that in the present analysis we restrict the eavesdropper to collective attacks. That is, Eve interacts with each communication round independently of previous rounds and stores her systems in a quantum memory. We believe that our security analysis could be lifted to general attacks with established reduction techniques [36, 37].

3 Protocols

We now present our simplest protocol (where Alice can prepare $n = 2$ different states), which is similar to the B92 protocol [38]. While the presentation below fits the implementation reported later, a more abstract and general presentation of these protocols is given in the companion paper [39].

Given a key bit $k = 0, 1$, Alice prepares one of two possible states, simply setting $x = k$. She uses a coherent state $|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$

[1] In practice to estimate the probability they reveal their outcomes on a small sample of the rounds chosen randomly, and use the results of the remaining rounds to distill the secret key.

with two possible polarization states $|\phi_x\rangle = \cos(\theta/2)\,|H\rangle + \sin(\theta/2)e^{i\pi x}\,|V\rangle$, which we write

$$|\psi_x\rangle = |\alpha\cos(\theta/2)\rangle_H \left|\alpha\sin(\theta/2)e^{i\pi x}\right\rangle_V, \quad (1)$$

where H and V denote the two orthogonal polarization modes. The overlap between Alice's preparations is given by $\langle\psi_1|\psi_0\rangle = e^{-2|\alpha|^2\sin(\theta)^2}$. The main assumption of our protocol is then

$$\gamma = \langle\psi_1|\psi_0\rangle \geq C, \quad (2)$$

where C is a parameter chosen by the users.

Bob performs measurements of the polarization, using two possible bases. Specifically, for $y = 0$, Bob projects the incoming signal onto a polarization $\left|\phi_0^\perp\right\rangle$, i.e., orthogonal to the polarization of Alice's first preparation. Similarly, for $y = 1$, he projects onto a polarization $\left|\phi_1^\perp\right\rangle$. For both measurements, if Bob gets a click (i.e., detects some light in the orthogonal polarization mode), then the round is conclusive and he outputs $b = 0$; otherwise $b = 1$, and the round will be discarded during sifting.

In the case of an ideal channel (loss and noise free), Alice and Bob will observe the statistics

$$p(b=0|x,y) = 1 - e^{-|\alpha|^2\sin(\theta)^2\sin\left(\frac{\pi(x-y)}{2}\right)^2}. \quad (3)$$

Note that $p(b = 0|x, y) > 0$ only when $x \neq y$. Hence, to establish the sifted key, Bob announces which rounds are successful, i.e. when $b = 0$. In this case, Bob infers his raw key bit to be $k' = y \oplus 1$. For an ideal channel, Alice and Bob obtain a perfectly correlated sifted key, i.e. $k = k'$.

This protocol can be generalized to the case where Alice can prepare $n > 2$ different states $|\psi_x\rangle$. In order to encode the raw key bit, Alice chooses now a pair of states, $\mathbf{r} = (r_0, r_1)$ with $0 \leq r_0 < r_1 \leq n-1$, among $\binom{n}{2}$ possible pairs. Then, for a key bit k, Alice sets $x = r_k$. Note that every state $|\psi_x\rangle$ can now encode either bit value, 0 or 1. Bob has n possible measurements, corresponding to projections onto the polarizations orthogonal to the states that Alice can prepare. Each measurement has two outputs $b = 0, 1$; $b = 0$ corresponds to the projection onto $\left|\phi_y^\perp\right\rangle$ while $b = 1$ corresponds to the projection onto $|\phi_y\rangle$. The sifting is now slightly more complicated. Alice announces $\mathbf{r}$, i.e., which pair of states she used. If Bob chose measurement $y = r_0$ or $y = r_1$ and got a conclusive outcome $b = 0$, he announces that the round is successful; if not, the round is discarded.

For our implementation, we consider again polarized coherent states, similarly to Eq. (1), with n possible polarizations $|\phi_x\rangle = \cos(\theta/2)\,|H\rangle + \sin(\theta/2)e^{ix2\pi/n}\,|V\rangle$, resulting in overlaps $\gamma_{ij} = \langle\psi_i|\psi_j\rangle = e^{-|\alpha|^2\sin(\theta/2)^2(1-e^{i2\pi/n(j-i)})}$. The main assumption in Eq. (2) is now replaced by an assumption on these complex overlaps, i.e. the entries of the Gram matrix G with elements $G_{ij} = \gamma_{ij}$. In fact, we can weaken this assumption by assuming only that the overlaps are in the vicinity of some ideal values γ_{ij}, i.e. $|\mathrm{Re}(\gamma_{ij}) - \mathrm{Re}(\langle\psi_i|\psi_j\rangle)| \leq \epsilon$ and $|\mathrm{Im}(\gamma_{ij}) - \mathrm{Im}(\langle\psi_i|\psi_j\rangle)| \leq \epsilon$ for a fixed ϵ.

4 Security analysis

Alice and Bob must bound Eve's knowledge of the key in order to perform the final privacy amplification for distilling the secret key. In our protocol, this estimation can be performed based solely on the observed data $p(b|x, y)$, given the overlap assumption of Eq. (2) is satisfied.

The channel, controlled by Eve, is viewed as a quantum broadcast channel (see Fig. 1), where part of the information reaches Bob's lab, while the remainder is held by Eve. Given Alice sent the state $|\psi_{r_k}\rangle$, the state at the output of the channel is denoted $\rho_{r_k}^{BE}$. Bob's measurements are denoted by a set of operators $M_{b|y}$. Eve performs a measurement on her subsystem, possibly after sifting, which is denoted $E_{e|z}$. As we do not impose any bound on the Hilbert space dimension, the measurements $M_{b|y}$ and $E_{e|z}$ can be taken to be projective (via Naimark dilation).

Since the min-entropy lower-bounds the von Neuman entropy as $H_{\min}(A|E, \mathrm{succ}) \leq H(A|E, \mathrm{succ})$, the asymptotic key rate (per round) is lower-bounded by [40]

$$R = \left(H_{\min}(A|E, \mathrm{succ}) - H_2[\mathrm{QBER}]\right) p(\mathrm{succ}), \quad (4)$$

where $p(\mathrm{succ})$ denotes the probability that a round of the protocol is conclusive (hence used to generate the key). The second term captures the error correction cost (H_2 is the binary entropy and QBER is the quantum bit error rate, which can be estimated from the data $p(b|x, y)$). The first term captures the privacy amplification efficiency and is given by Eve's conditional min entropy $H_{\min}(A|E, \mathrm{succ}) = -\log_2\left(p_g(e = k|\mathrm{succ})\right)$,

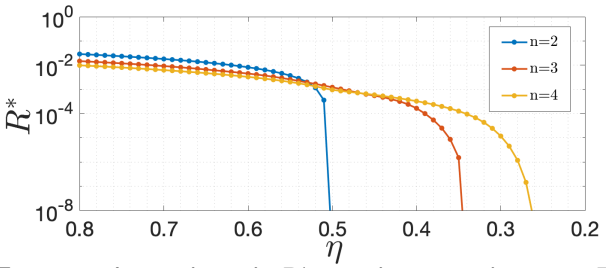

Figure 2: Lower bound, R^*, on the secret key rate R as a function of the transmission η. A positive key rate is obtained for transmissions down to $\eta = 1/n$, when considering a protocol where Alice prepares n states of the form in Eq. (1) (fixing the polarizations to $\theta = 0.2$ and optimizing over the coherent state α).

where $p_g(e = k|\text{succ})$ is the probability of correctly guessing the secret bit k given that the round is conclusive. We note that the expression R provides a lower bound on the key rate, and the bound is in general not tight. In order to obtain tight bounds, one should ~~consider a different quantity, namely~~ directly consider the conditional von Neumann entropy, see e.g. Ref. [40, 37, 41].

The main challenge is now to bound Eve's guessing probability $p_g := p_g(e = k|\text{succ}) = \frac{p(e=k,\text{succ})}{p(\text{succ})}$, under the constraint that the observed data is given by $p(b|x, y)$ and that the states prepared by Alice have given overlaps. More precisely, we want to upper bound the quantity:

$$p_g = \frac{\sum_{\mathbf{r}} \sum_k \text{tr}\left(\rho^{BE}_{r_k}(M_{0|r_0} + M_{0|r_1}) \otimes E_{k|\mathbf{r}}\right)}{\sum_{\mathbf{r}} \sum_k \text{tr}\left(\rho^{BE}_{r_k}(M_{0|r_0} + M_{0|r_1}) \otimes \mathbb{1}_E\right)} \quad (5)$$

over any possible output state $\rho^{BE}_{r_k}$ and measurements for Bob and Eve that are compatible with the data and the overlap assumption. We assume that Alice and Bob choose their respective inputs (k, $\mathbf{r}$ and y) uniformly at random. It turns out that this problem can be relaxed to a hierarchy of semi-definite programs (SDPs) by taking advantage of the method introduced in Ref. [28] (see Appendix A).

In Fig. 2, we show how the corresponding bound on the key rate R behaves as a function of the transmission η of the channel, for protocols involving $n = 2, 3, 4$ states. As noted before, any QKD protocol in the receiver-DI model must have $R \to 0$ when $\eta \to 1/n$. Nevertheless, we see from the plot that the security approaches this threshold, so that lower transmissions can be reached by increasing the number of states n in the protocol.

5 Experimental realization

The main challenge for the experimental implementation of our protocols lies in the limited loss budget (< 3 dB for the 2-states protocol) of the quantum channel (QC), including Bob's measurement setup. For this feasibility experiment we used, for simplicity, weak coherent states, encoded and measured with fibered piezo-electric polarization controllers. Although these devices have low transmission loss, they encompass a limit in the operation speed (~ 1 kHz) (a much faster low-loss setup would be possible, though requiring a considerably higher degree of complexity). To minimize loss, on the detection side, we used a high efficiency SNSPD with 90% detection efficiency at telecom wavelength and a dark count rate of 200 Hz (ID Quantique).

The experimental setup is shown on Fig. 3. On Alice's side, a distributed feedback laser, triggered at a 1 MHz rate, generates pulses at 1559 nm with 90 ps FWHM duration. The power fluctuation of the laser output signal is monitored every second using a powermeter. The rate of the optical signal is then reduced to 1 kHz using an EOM. The polarization states $|\phi_x\rangle$ are encoded via a polarization controller (General Photonics' PolaRITE) comprised of 4 piezoelectric fiber squeezers. The polarization at the input of the controller is aligned so that two piezoelectric squeezers control the angles θ and ϕ respectively, allowing for the preparation of any desired state $|\phi_x\rangle = \cos(\theta/2)|H\rangle + \sin(\theta/2)\, e^{-i\phi_x}|V\rangle$. In order to monitor the stability of the state preparation, part of the signal is sent to a moni-

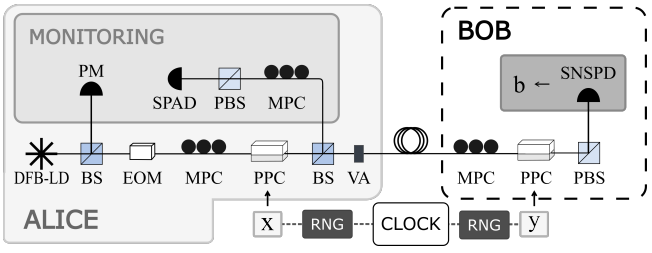

Figure 3: Experimental setup. DFB-LD: Distributed feedback laser diode; BS: Beam splitter; EOM: Electro-optic modulator; MPC: Manual polarization controller; PPC: Piezo-electric polarization controller; VA: Variable attenuator; PBS: Polarizing beam splitter; SNSPD: Superconducting nanowire single-photon detector; SPAD: Single-photon avalanche diode; PM: Powermeter; RNG: Random number generator.

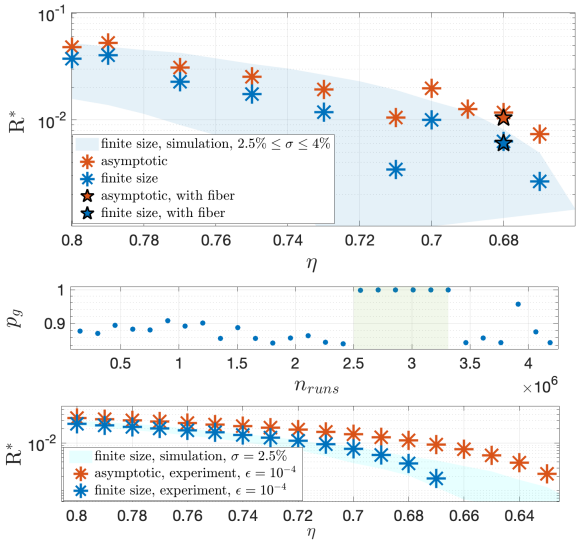

Figure 4: Experimental results. (Top) Key rate R as a function of the transmission η for the protocol with $n = 2$ states. Each point represents a run over half an hour, with finite-size bound on the guessing probability (blue), see Appendix A.1 for details, and in the asymptotic regime (red). Data taken with 4.8 km fiber corresponds to blue and red stars. Data is consistent with Monte Carlo simulations with polarization fluctuations $2.5\% \leq \sigma \leq 4\%$ (blue region, estimated from data). (Middle) Illustration of the self-testing feature of the protocol. After 2.5 hours of operation, we artificially lower the detection efficiency of the SNSPD (shaded region), resulting in a guessing probability p_g for Eve that jumps to one, hence $R = 0$. Later, the SNSPD's efficiency is brought back to normal, hence $p_g < 1$ again and $R > 0$. (Bottom) Key rate R vs transmission η for the protocol with $n = 3$ states, showing enhanced robustness to losses.

toring setup that includes a manual polarization controller, a PBS, and a SPAD (ID Quantique ID210).

Bob's measurement is implemented via a polarization controller (identical to Alice's), allowing for an active choice of the measurement basis, followed by a PBS and the SNSPD.

The detection timestamps are recorded within a detection window of 1 ns, using a time-to-digital converter (ID Quantique ID800). All time driven components of the setup are triggered with an external clock (Silicon Labs SI5341). The inputs for Alice (k, $\mathbf{r}$) and Bob (y) are generated by a personal computer and transferred to the PPCs via Teensy micro controllers.

At the output of Alice, a VA controls the average photon number per pulse, $|\alpha|^2$, and when desired, introduces additional transmission loss. The QC, including Bob's setup and the detection efficiency, has a total transmission of $\eta = 80.3\,\%\,(-0.953\,\mathrm{dB})$. Adding 4815 m of ultra-low loss telecom fiber (Corning SMF-28 ULL), the total transmission is $\eta = 68\,\%\,(-1.674\,\mathrm{dB})$.

We first implemented the simplest protocol with $n = 2$ states, investigating the performance as a function of the transmission of the QC, η. The two polarization states are prepared as in Eq. (1) with $\theta = 0.6$ rad. The coherent state α is optimized depending on η (see Appendix C.1). The monitoring stage allows one to continuously measure the polarization and intensity of the light, in order to ensure the validity of the overlap assumption of Eq. (2).

To establish a secret key, we run each experiment for 0.5 hours, collecting blocks of $N = 1.8 \times 10^6$ events. For each round, a random key bit $k = x$ is generated for Alice, and a random measurement setting y for Bob. The detection (non-detection) at the SNSPD sets the outcome to $b = 0$ ($b = 1$). Once a block of data is collected, sifting is performed followed by a finite-size analysis to determine the statistics $p(b|x, y)$ (see Appendix A). Given the overlap assumption and the statistics, we then compute a lower-bound on the key rate R as in Eq. (4) with the semi-definite programming (SDP) relaxation.

The results are shown in Fig. 4 (Top); we also provide a bound on R in the asymptotic regime, i.e. omitting finite-size corrections. To model the experiment, we perform a finite-size Monte Carlo simulation by ranging the estimated misalignment error σ (see Appendix B) between 2.5% and 4% (estimated from data). Key rates in the order of 10^{-3} were obtained for a transmission of $\eta = 67\%$. A similar rate is obtained over the 4.8 km fiber.

In Fig. 4 (Middle), we illustrate the self-testing feature of the protocol. After 2.5×10^6 runs, the detector efficiency was intentionally reduced from 90 % to 42 % by lowering the bias current of the SNSPD. This sudden change is immediately detected in the post-processing, resulting in an immediate increase of Eve's guessing probability to one. Hence, no secret key can be distilled and the users become aware of the setup's malfunction.

Lastly, we implemented the protocol with $n = 3$ states (see Appendix C.3 for details). To lower misalignment errors (down to $\sigma = 2.5\%$), due mostly to patterning effects in the polar-

ization controllers, we implemented this protocol omitting the random choice of settings. Results, shown in Fig. 4 (Bottom), demonstrate increased robustness to loss, with positive key rates at $\eta = 63\%$ in the asymptotic regime.

6 Conclusion

We presented protocols for receiver-DI QKD in a prepare-and-measure setup. Such a scenario is relevant in practice when Alice and Bob have a different level of trust in their devices; for instance Alice being a large company and Bob an end-user. Note that the MDI approach cannot provide security in the receiver-DI scenario, as some level of trust will always be required for both honest parties. We note that a detailed comparison of our approach with previous works is given in the companion paper [39], as well as a more detailed discussion of how to justify the overlap bounds in practice.

From the observed statistics, the users can establish a secret key, while monitoring in real-time the correct operation of their devices, and thus immediately detect potential failure due, for instance, to a malfunctioning device or an attack by the eavesdropper (e.g., blinding). Our protocols do not rely on any type of fair-sampling assumption, contrary to Ref. [21].

The main limitation of our protocols are their loss sensitivity, which is in fact a fundamental limit for any QKD protocol in the receiver-DI scenario. Still, the requirements for transmission and efficiency are considerably relaxed compared to the full DI model (or the one-sided DI approach of Ref. [24]), with the additional advantage that no source of entanglement is required, as well as significantly reducing the required efficiency on Bob's side (in principle, the efficiency can be made arbitrarily low, as we show in the companion article [39]).

We reported here a proof-of-principle implementation achieving security over a distance of few kilometers. We expect notable improvements in terms of rates and transmission for a high-speed polarization encoding setup as demonstrated in previous works [42]. Indeed, by multiplying the clock-rate by a factor of 10^3, we would be much less limited by the block size and could increase the number of states of the protocol. For instance (following the results presented in Fig. 3) an implementation of the four-state protocol with weak coherent states encoded at a 1 GHz rate, would lead to key rates of $\sim$ 100 kbps at a 30 km distance. Alternatively an implementation based on polarization-encoded single photons would also be more robust to loss [39].

To conclude, we believe that our work opens a new interesting approach in the intermediate regime between standard "device-dependent" QKD and the fully device-independent (DI) model, which is amenable to experiments.

Acknowledgements.—We thank Jonathan Brask and Denis Rosset for discussions, and ID Quantique for supplying the SNSPD. We acknowledge financial support from the EU Quantum Flagship project QRANGE, and the Swiss National Science Foundation (BRIDGE, project 2000021_192244/1). This work was supported as a part of NCCR QSIT, a National Centre of Competence (or Excellence) in Research, funded by the Swiss National Science Foundation (grant number $51NF40-185902$).

References

[1] Eleni Diamanti, Hoi-Kwong Lo, Bing Qi, and Zhiliang Yuan. "Practical challenges in quantum key distribution". npj Quantum Inf. **2**, 16025 (2016).

[2] Feihu Xu, Xiongfeng Ma, Qiang Zhang, Hoi-Kwong Lo, and Jian-Wei Pan. "Secure quantum key distribution with realistic devices". Rev. Mod. Phys. **92**, 025002 (2020).

[3] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, and P. Wallden. "Advances in quantum cryptography". Adv. Opt. Photon. **12**, 1012–1236 (2020).

[4] Alberto Boaron, Gianluca Boso, Davide Rusca, Cédric Vulliez, Claire Autebert, Misael Caloz, Matthieu Perrenoud, Gaëtan Gras, Félix Bussières, Ming-Jun Li, Daniel Nolan, Anthony Martin, and Hugo Zbinden. "Secure quantum key distribution over 421 km of optical fiber". Phys. Rev. Lett. **121**, 190502 (2018).

[5] Juan Yin, Yu-Huai Li, Sheng-Kai Liao, Meng Yang, Yuan Cao, Liang Zhang, Ji-Gang Ren, Wen-Qi Cai, Wei-Yue Liu, Shuang-Lin Li, Rong Shu, Yong-Mei Huang, Lei Deng, Li Li, Qiang Zhang, Nai-Le Liu, Yu-Ao Chen, Chao-Yang Lu, Xiang-Bin Wang, Feihu Xu, Jian-Yu Wang, Cheng-Zhi Peng, Artur K. Ekert, and Jian-Wei Pan. "Entanglement-based secure quantum cryptography over 1,120 kilometres". Nature **582**, 501–505 (2020).

[6] Jiu-Peng Chen, Chi Zhang, Yang Liu, Cong Jiang, Weijun Zhang, Xiao-Long Hu, Jian-Yu Guan, Zong-Wen Yu, Hai Xu, Jin Lin, Ming-Jun Li, Hao Chen, Hao Li, Lixing You, Zhen Wang, Xiang-Bin Wang, Qiang Zhang, and Jian-Wei Pan. "Sending-or-not-sending with independent lasers: Secure twin-field quantum key distribution over 509 km". Phys. Rev. Lett. **124**, 070501 (2020).

[7] ID Quantique SA Switzerland.

[8] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. "Device-independent security of quantum cryptography against collective attacks". Phys. Rev. Lett. **98**, 230501 (2007).

[9] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. "Device-independent quantum key distribution secure against collective attacks". New J. Phys. **11**, 045021 (2009).

[10] Umesh Vazirani and Thomas Vidick. "Fully device-independent quantum key distribution". Phys. Rev. Lett. **113**, 140501 (2014).

[11] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. "Practical device-independent quantum cryptography via entropy accumulation". Nature Commun. **9**, 459 (2018).

[12] Artur Ekert and Renato Renner. "The ultimate physical limits of privacy". Nature **507**, 443–447 (2014).

[13] Lars Lydersen, Carlos Wiechers, Christoffer Wittmann, Dominique Elser, Johannes Skaar, and Vadim Makarov. "Hacking commercial quantum cryptography systems by tailored bright illumination". Nature Photonics **4**, 686–689 (2010).

[14] Ilja Gerhardt, Qin Liu, Antía Lamas-Linares, Johannes Skaar, Christian Kurtsiefer, and Vadim Makarov. "Full-field implementation of a perfect eavesdropper on a quantum cryptography system". Nature Commun. **2**, 349 (2011).

[15] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M. Lucas, C. J. Ballance, K. Ivanov, E. Y-Z. Tan, P. Sekatski, R. L. Urbanke, R. Renner, N. Sangouard, and J-D. Bancal. "Device-independent quantum key distribution" (2021). arXiv:2109.14600.

[16] Wei Zhang, Tim van Leent, Kai Redeker, Robert Garthoff, Rene Schwonnek, Florian Fertig, Sebastian Eppelt, Valerio Scarani, Charles C. W. Lim, and Harald Weinfurter. "Experimental device-independent quantum key distribution between distant users" (2021). arXiv:2110.00575.

[17] Wen-Zhao Liu, Yu-Zhe Zhang, Yi-Zheng Zhen, Ming-Han Li, Yang Liu, Jingyun Fan, Feihu Xu, Qiang Zhang, and Jian-Wei Pan. "High-speed device-independent quantum key distribution against collective attacks" (2021). arXiv:2110.01480.

[18] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. "Measurement-device-independent quantum key distribution". Phys. Rev. Lett. **108**, 130503 (2012).

[19] Samuel L. Braunstein and Stefano Pirandola. "Side-channel-free quantum key distribution". Phys. Rev. Lett. **108**, 130502 (2012).

[20] Marco Tomamichel and Renato Renner. "Uncertainty relation for smooth entropies". Phys. Rev. Lett. **106**, 110506 (2011).

[21] Marco Tomamichel, Charles Ci Wen Lim, Nicolas Gisin, and Renato Renner. "Tight finite-key analysis for quantum cryptography". Nature Communications **3**, 634 (2012).

[22] Marco Tomamichel and Anthony Leverrier. "A largely self-contained and complete security proof for quantum key distribution". Quantum **1**, 14 (2017).

[23] Antonio Acín, Daniel Cavalcanti, Elsa Passaro, Stefano Pironio, and Paul Skrzypczyk. "Necessary detection efficiencies for secure quantum key distribution and bound randomness". Phys. Rev. A **93**, 012319 (2016).

[24] Cyril Branciard, Eric G. Cavalcanti, Stephen P. Walborn, Valerio Scarani, and Howard M. Wiseman. "One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering". Phys. Rev. A **85**, 010301 (2012).

[25] J B Brask, A Martin, W Esposito, R Houlmann, J Bowles, H Zbinden, and N Brunner. "Megahertz-rate semi-device-independent quantum random number generators based on unambiguous state discrimination". Phys. Rev. Applied **7**, 054018 (2017).

[26] T Van Himbeeck, E Woodhead, N J Cerf, R García-Patrón, and S Pironio. "Semi-device-independent framework based on natural physical assumptions". Quantum **1**, 33 (2017).

[27] D Rusca, T van Himbeeck, A Martin, J B Brask, W Shi, S Pironio, N Brunner, and H Zbinden. "Practical self-testing quantum random number generator based on a energy bound". Phys. Rev. A **100**, 062338 (2019).

[28] Yukun Wang, Ignatius William Primaatmaja, Emilien Lavie, Antonios Varvitsiotis, and Charles Ci Wen Lim. "Characterising the correlations of prepare-and-measure quantum networks". npj Quantum Inf. **5**, 17 (2019).

[29] M Pawłowski and N Brunner. "Semi-device-independent security of one-way quantum key distribution". Phys. Rev. A **84**, 010302(R) (2011).

[30] Erik Woodhead and Stefano Pironio. "Secrecy in prepare-and-measure clauser-horne-shimony-holt tests with a qubit bound". Phys. Rev. Lett. **115**, 150501 (2015).

[31] X. Ma and N. Lutkenhaus. "Improved data post-processing in quantum key distribution and application to loss thresholds in device independent qkd". Quantum Inf. Comput. **12**, 3 (2011).

[32] M. Ho, P. Sekatski, E. Y.-Z. Tan, R. Renner, J.-D. Bancal, and N. Sangouard. "Noisy preprocessing facilitates a photonic realization of device-independent quantum key distribution". Phys. Rev. Lett. **124**, 230502 (2020).

[33] Rene Schwonnek, Koon Tong Goh, Ignatius W. Primaatmaja, Ernest Y.-Z. Tan, Ramona Wolf, Valerio Scarani, and Charles C.-W. Lim. "Device-independent quantum key distribution with random key basis". Nature Commun. **12**, 2880 (2021).

[34] Erik Woodhead, Antonio Acín, and Stefano Pironio. "Device-independent quantum key distribution with asymmetric CHSH inequalities". Quantum **5**, 443 (2021).

[35] Pavel Sekatski, Jean-Daniel Bancal, Xavier Valcarce, Ernest Y.-Z. Tan, Renato Renner, and Nicolas Sangouard. "Device-independent quantum key distribution from generalized CHSH inequalities". Quantum **5**, 444 (2021).

[36] Renato Renner. "Symmetry of large physical systems implies independence of subsystems". Nature Physics **3**, 645–649 (2007).

[37] Frederic Dupuis, Omar Fawzi, and Renato Renner. "Entropy accumulation". Commn. Math. Phys. **379**, 867–913 (2020).

[38] Charles H. Bennett. "Quantum cryptography using any two nonorthogonal states". Phys. Rev. Lett. **68**, 3121–3124 (1992).

[39] Marie Ioannou, Pavel Sekatski, Alastair A. Abbott, Denis Rosset, Jean-Daniel Bancal, and Nicolas Brunner. "Receiver-device-independent quantum key distribution protocols" (2021). arXiv:2110.00575.

[40] Igor Devetak and Andreas Winter. "Distillation of secret key and entanglement from quantum states". Proc. Roy. Soc. **A461**, 207–235 (2005).

[41] Hamza Fawzi, James Saunderson, and Pablo A. Parrilo. "Semidefinite approximations of the matrix logarithm". Found. Comput. Math. **19**, 259 (2018).

[42] Fadri Grünenfelder, Alberto Boaron, Davide Rusca, Anthony Martin, and Hugo Zbinden. "Performance and security of 5 ghz repetition rate polarization-based quantum key distribution". Appl. Phys. Lett. **117**, 144003 (2020).

[43] Abraham Charnes and William W Cooper. "Programming with linear fractional functionals". Naval Research logistics quarterly **9**, 181–186 (1962).

[44] Pei-Sheng Lin, Denis Rosset, Yanbao Zhang, Jean-Daniel Bancal, and Yeong-Cherng Liang. "Device-independent point estimation from finite data and its application to device-independent property estimation". Phys. Rev. A **97**, 032309 (2018).

[45] Jean-Daniel Bancal, Kai Redeker, Pavel Sekatski, Wenjamin Rosenfeld, and Nicolas Sangouard. "Self-testing with finite statistics enabling the certification of a quantum network link". Quantum **5**, 401 (2021).

A Bounding Eve's information

In this appendix, we detail the procedure for obtaining upper bounds on Eve's guessing probability, which in turn leads to a lower bound on the key rate R.

The main assumption of the protocol is that the overlaps of Alice's prepared states $|\psi_x\rangle$ are limited. This is compactly expressed in terms of the Gram matrix

$$G = \sum_{i,j=0}^{n-1} \gamma_{ij} \, |i\rangle\langle j| \, . \tag{6}$$

Moreover, the observed statistics $p(b|x,y)$ are estimated by Alice and Bob. A lower bound on the asymptotic keyrate R is given by

$$R = (H_{\min}(p_g(e=k|\text{succ})) - H_2[\text{QBER}])p(\text{succ}) \tag{7}$$

where $p_g(e=k|\text{succ})$ is the probability that Eve guesses correctly the secret bit k given that a round is conclusive, $H_{\min}$ is the min-entropy, H_2 is the binary entropy function, QBER is the quantum bit error rate and $p(\text{succ})$ is the average probability to generate key. The QBER and $p(\text{succ})$ can be extracted from the observed statistics $p(b|x,y)$ while the guessing probability $p_g(e=k|\text{succ})$ needs to be upper bounded. The guessing probability is given by

$$\begin{aligned} p_g(e=k|\text{succ}) &= \frac{p(e=k,\text{succ})}{p(\text{succ})} \\ &= \frac{\sum_{r=0}^{\binom{n}{2}-1} p_R(r) \sum_{k=0}^{1} p_K(k) \sum_{y=0}^{n-1} p_Y(y) \, \text{tr}(\rho_{r_k}^{BE} M_{0|y} E_{k|r})(\delta_{y,r_0} + \delta_{y,r_1})}{\sum_{r=0}^{\binom{n}{2}-1} p_R(r) \sum_{k=0}^{1} p_K(k) \sum_{y=0}^{n-1} p_Y(y) \, \text{tr}\,(\rho_{r_k}^{BE} M_{0|y} \mathbb{1})(\delta_{y,r_0} + \delta_{y,r_1})}, \end{aligned} \tag{8}$$

where $M_{b|y}$ are Bob's measurement operators with $b = 0,1$ and $y = 0,...,n-1$ and $E_{k|z}$ are Eve's measurement operators with $k = 0,1$ and $z = 0,...,\binom{n}{2}-1$. $p_R(r)$, $p_Y(y)$ and $p_K(k)$ are the probabilities of choosing the inputs r, y and k, satisfying $\sum_r p_R(r) = \sum_k p_K(k) = \sum_y p_Y(y) = 1$, $p_K(k) \geq 0 \ \forall k$, $p_Y(y) \geq 0 \ \forall y$ and $p_R(r) \geq 0 \ \forall r$. Here, we take the input probabilities to be uniformly random for all inputs. As we impose no limit on the Hilbert space dimension, we can, without loss of generality (using Naimark's dilation theorem), assert that Bob's and Eve's measurements are projectors satisfying the following properties: (i) $M_{b|y}M_{b'|y} = \delta_{b,b'}M_{b|y} \quad \forall y,b,b'$, (ii) $\sum_b M_{b|y} = \mathbb{1} \quad \forall y$, (iii) $E_{e|z}E_{e'|z} = \delta_{e,e'}E_{e|z} \quad \forall z,e,e'$, (iv) $\sum_e E_{e|z} = \mathbb{1} \quad \forall z$ and (v) $[M_{b|y}, E_{e|z}] = 0 \ \forall b,e,y,z$. The last property comes from the fact that Bob and Eve act on two different Hilbert spaces.

To upper bound $p_g(e=k|\text{succ})$, we will use the numerical method presented in [28]. The numerical method consists of an SDP hierarchy providing increasingly tight outer approximations of the set of quantum correlations in a discrete prepare-and-measure (PM) scenario given the Gram matrix of the set of quantum states. This hierarchy provides a computationally tractable method to bounding $p_g(e=k|\text{succ})$ in the absence of any upper bound on the Hilbert space dimension. Let us define the moment matrix Γ of size $nq \times nq$:

$$\Gamma = \sum_{i,j=1}^{n} \Gamma_{ij} \otimes |\hat{e}_i\rangle\langle\hat{e}_j| \tag{9}$$

where $\{|\hat{e}_i\rangle\}_{i=1,...,n}$ is an orthonormal basis in $\mathbb{R}^n$ and recall that n is the number of states prepared by Alice. The sub-blocks Γ_{ij} are defined as

$$\Gamma_{ij} = \sum_{k,l=1}^{r} \langle\psi_i| \, S_k^\dagger S_l \, |\psi_j\rangle \otimes |\hat{f}_k\rangle\langle\hat{f}_l| \tag{10}$$

where $\{|\hat{f}_i\rangle\}_{i=1,...,q}$ is an orthonormal basis for $\mathbb{R}^q$ and $\{S_i\}_{i=1,...,q}$ is a set of products of measurement operators $B_{b|y}$ and $E_{e|z}$. We note that the moment matrix Γ is positive semi-definite and all the correlations $p(b,e|x,y,z)$ appear as elements in Γ. One can choose the set of operators arbitrarily but the aim is to have as many independent operators as possible in the moment matrix. We can organize the operators into levels of the hierarchy. The first two levels are given by the two following sets of operators

$$\begin{aligned} &\mathbb{S}_\mathbb{1} = \{\mathbb{1}, B_{b|y}, E_{e|z}\}, \\ &\mathbb{S}_2 = \mathbb{S}_\mathbb{1} \cup \{B_{b|y}B_{b'|y'}, E_{e|z}E_{e'|z'}, B_{b|y}E_{e|z}\}. \end{aligned} \tag{11}$$

One can go to higher levels by including increasingly long products of measurement operators. In Ref. [28], it has been demonstrated that by going to increasingly large levels the hierarchy converges to the quantum set.

The SDP maximizing the guessing probability is given by

$$\begin{aligned} \max_{\Gamma} \quad & p_g(e=k|\text{succ}) = \frac{\text{tr}(\Gamma A)}{\text{tr}(\Gamma B)} \\ \text{s.t.} \quad & \text{tr}(\Gamma C_{bxy}) = p(b|x,y) \qquad \forall b,y,x \\ & \text{tr}(\Gamma D_{jk}) = G_{jk} \qquad \forall j,k \\ & \text{tr}(\Gamma F_k) = f_k \\ & \Gamma \geq 0 \end{aligned} \tag{12}$$

where A and B are constant matrices selecting the term of the moment matrix necessary to compute the guessing probability Eq. (5). The fixed matrices C_{bxy} set the appropriate entries of Γ equal to the observed statistics $p(b|x,y)$, and D_{jk} apply the inner-product constraint of the encoding states to Γ (recall that G is the Gram matrix (6) specifying these overlaps). Finally, F_k's set the linear constraints arising from the operators $B_{b|y}$ and $E_{e|z}$. As written above, the optimization has the form of a semidefinite-fractional program, i.e., a semidefinite program with an objective function consisting of the fraction between two linear functions. Such a program can in general be transformed into a regular semidefinite program by applying a Charnes-Cooper transform as described in the theory of linear-fractional programming [43].

However, in our case we notice that the denominator $\text{tr}(\Gamma B) = p(\text{succ})$ only involves directly observable quantities, i.e., terms of the Γ matrix which are fixed by the first constraint. Therefore, it is sufficient to simply optimize

$$\begin{aligned} \max_{\Gamma} \quad & p_g(e=k,\text{succ}) = \text{tr}(\Gamma A) \\ \text{s.t.} \quad & \text{tr}(\Gamma C_{bxy}) = p(b|x,y) \qquad \forall b,y,x \\ & \text{tr}(\Gamma D_{jk}) = G_{jk} \qquad \forall j,k \\ & \text{tr}(\Gamma F_k) = f_k \\ & \Gamma \geq 0. \end{aligned} \tag{13}$$

The SDP (13) assumes the overlap assumption is satisfied exactly, but this can easily be relaxed as follows. For $n = 2$ we can, without loss of generality, consider real overlaps and hence enforce simply a lower bound on the overlaps, i.e., $\gamma = \langle\psi_0|\psi_1\rangle \geq C$. For $n > 2$, we assume that the the real and imaginary part of the overlap are in the vicinity of the ideal values γ_{ij} by upper and lower bounding the real and imaginary parts of γ_{ij}

$$\begin{aligned} |\text{Re}(\gamma_{ij}) - \text{Re}(\langle\psi_i|\psi_j\rangle)| \leq \epsilon \\ |\text{Im}(\gamma_{ij}) - \text{Im}(\langle\psi_i|\psi_j\rangle)| \leq \epsilon. \end{aligned} \tag{14}$$

A.1 Finite-size statistics

The primal SDP (13) gives optimal bounds on the joint guessing probability given a Gram matrix and an asymptotic probability distribution, but it is not practical when we consider finite-size statistical effects. Indeed, finite experimental data does not describe, in general, asymptotically valid distributions, and it is not clear *a priori* how finite-size effects propagate in the primal SDP. A general way to deal with the first issue in NPA-type SDP hierarchies was proposed in [44]. In our case, since we are simply interested in obtaining a valid bound on the keyrate which applies to the finite statistics, we rely only on the symmetrized observed statistics $p(b|x,y)$. We then analyse the effect of finite statistical fluctuations with the help of the dual of the SDP. Namely, we will first show that the dual objective function upper bounds the primal objective function, and then we will upper-bound the objective function of the dual by taking into account the finite-size statistical effects. Finally, we will lower-bound $p(\text{succ})$.

The primal SDP has the generic form

$$\max_{\Gamma} \text{tr}(\Gamma A) \tag{15a}$$

$$\text{s.t.} \quad \text{tr}(\Gamma B) = b \tag{15b}$$

$$\Gamma \geq 0. \tag{15c}$$

To derive the dual, we introduce for each constraint a Lagrangian multiplier β, and H. The Lagrangian function of the primal SDP is given by

$$\mathcal{L} = \text{tr}(\Gamma A) + \beta(b - \text{tr}(\Gamma B)) + H\Gamma. \tag{16}$$

We define $\mathcal{S}$ to be the supremum of the Lagrangian over the primal SDP variable in its domain:

$$\mathcal{S} = \sup_{\Gamma} \mathcal{L}. \tag{17}$$

Examining this quantity, we notice that the contributions coming from the second term of the Lagrangian (16) vanish because of the constraint

(15b). Imposing $H \geq 0$ implies that the second term of the Lagrangian is positive. Since the first term of the Lagrangian is equal to the objective function of the primal SDP (15a), $\mathcal{S}(\beta, H)$ is an upper bound on the objective function of the primal (15a) whenever $H \geq 0$, i.e. $\mathcal{S} \geq p_g^{\text{primal}}$. Let us rewrite $\mathcal{S}$ by grouping the primal SDP variables

$$\mathcal{S} = \sup_{\Gamma} \operatorname{tr}(\Gamma(A - \beta B + H)) + \beta b. \tag{18}$$

To obtain an optimal upper bound on p_g^{primal} from $\mathcal{S}$, we minimize $\mathcal{S}$ over the Lagrange multipliers. Since $\Gamma \geq 0$ and $H \geq 0$ the supremum may be unbounded if the first term doesn't vanish. Hence, we impose $A - \beta B + H = 0$. This leads to the following SDP

$$p_g^{\text{dual}} = \min_{\beta} \quad \beta b \tag{19a}$$

$$\text{s.t.} \quad A - \beta B \leq 0. \tag{19b}$$

By construction, we thus showed that $p_g \leq p_g^{\text{primal}} \leq p_g^{\text{dual}}$.

With the above considerations it is easy to show that the objective function of the dual is given by

$$p_g \leq K + \sum_{x,y} \nu_{x,y}^{b=0}\, p(b=0|y,x), \tag{20}$$

where K takes into account all the terms that do not contain any finite-statistical effects and $\nu_{x,y}^{b=1}$ is the associated set of dual variables corresponding to the constraints imposed by the probability distribution. Let us upper bound the second term of Eq. (20). We first write

$$\sum_{x,y} \nu_{x,y}^{b=0} p(b=0|x,y) = \sum_{x,y} \frac{\nu_{x,y}^{b=0}}{p(x)p(y)} \delta_{b,0}^{x,y} p(b,x,y) = \sum_{x,y} g_{x,y}^{b=0} \mathbb{E}(\delta_{b,0}^{x,y}).$$

Here, $\delta_{b,0}^{x,y}$ can be interpreted as a binary game which is won if $b = 0$ and lost if $b = 1$. If the game is won, we score $g_{x,y}^{b=0} := \frac{\nu_{x,y}^{b=0}}{p(x)p(y)}$. Then, using Theorem B.2 of Ref. [45], we obtain the following bound on $\mathbb{E}(\delta_{b,0}^{x,y})$ with confidence $1 - \alpha_1$: $P(\mathbb{E}(\delta_{b,0}^{x,y}) \leq \hat{q}_{x,y}) \geq 1 - \alpha_1$ for $\hat{q}_{x,y} = 1 - I_{\alpha_1}^{-1}(f(b=0,x,y)N, N(1 - f(b=0,x,y)) + 1)$ where $I_{\alpha}^{-1}(a,b)$ denotes the inverse regularized Beta function and $\alpha_1 = 10^{-9}$. We thus obtain the following bound on Eq. (**??**) with confidence $1 - \alpha_1$:

$$\sum_{x,y} g_{x,y}^{b=0} \mathbb{E}(\delta_{b,0}^{x,y}) \leq \sum_{x,y} g_{x,y}^{b=0} \hat{q}_{x,y}. \tag{21}$$

Using Eq. (20) this translates into a bound on p_g with the same confidence level $1 - \alpha_1$:

$$p_g \leq K + \sum_{x,y} g_{x,y}^{b=0} \hat{q}_{x,y}. \tag{22}$$

We hence upper-bounded the joint guessing probability (13). It remains to lower-bound $p(\text{succ}) \geq p_2^*$.

The probability of success is given by

$$\begin{aligned} p(\text{succ}) &= \sum_{r,k,y} \frac{(\delta_{y,r_0} + \delta_{y,r_1})}{p_R(r)p_K(k)p_Y(y)} p(b=0, r_k, y) \\ &= \sum_{r,k,y} g_{r_k,y}^{b=0} (\delta_{y,r_0} + \delta_{y,r_1}) \delta_{b,0} p(b, r_k, y) \\ &= \sum_{r,k,y} g_{r_k,y}^{b=0} \mathbb{E}(\chi(b, r_k, y)) \\ &\geq \sum_{r,k,y} g_{r_k,y}^{b=0} \hat{q}_{r_k,y}. \end{aligned} \tag{23}$$

Similarly as before, $\chi(b, r_k, y) := (\delta_{y,r_0} + \delta_{y,r_1})\delta_{b,0}$ can be interpreted as a binary game which is won if $b = 0$ and $y = r_1$ or $y = r_0$ and lost if $b = 1$. If the game is won, we now score $g_{r_k,y}^{b=0} = \frac{1}{p_R(r)p_K(k)p_Y(y)}$. We obtain the following lower bound on $\mathbb{E}(\chi(b, y, r_k))$ with confidence $1 - \alpha_2$: $P(\mathbb{E}(\chi(b, r_k, y) \geq \hat{q}_{r_k,y}) \geq 1 - \alpha_2$ for $\hat{q}_{r_k,y} = I_{\alpha_1}^{-1}(f(b=0,r_k,y)N, N(1 - f(b=0,r_k,y)) + 1)$.

Since $p(e = k, \text{succ}) \leq p_1^*$ with prob $1 - \alpha_1$ and $p(\text{succ}) \geq p_2^*$ with prob $1 - \alpha_2$, we deduce that $p(e = k|\text{succ}) \leq \frac{p_1^*}{p_2^*}$ with confidence at least $1 - \alpha$, for $\alpha = \alpha_1 + \alpha_2$.

With the upper bound on the guessing probability $p(e = k|\text{succ}) \leq \frac{p_1^*}{p_2^*}$ taking into account finite statistical effects, we estimate the key rate in Eq. (4) as

$$R^* \geq \left(H_{\min}\left(\frac{p_1^*}{p_2^*}\right) - H_2[\text{QBER}] \right) p(\text{succ}). \tag{24}$$

B Noise model

In this section we are going to consider a channel model that allows us to estimate the amount of noise in our experiment due to misalignment between the states prepared by Alice and the bases chosen by Bob. These errors can be divided in two different categories: systematic errors and stochastic ones.

In order to model the former it is sufficient to add a constant term of misalignment for both the angle θ and the state and basis choice x and y.

If we define these terms as Δ_θ and $\Delta_{x,y}$ we can rewrite Eq. (3) as:

$$p(b=0|x,y) = 1 - e^{-|\alpha|^2 \sin(\theta+\Delta_\theta)^2 \sin\left(\frac{\pi(x-y+\Delta_{x,y})}{2}\right)^2}. \tag{25}$$

If we now consider instead possible stochastic errors we have to assume a certain noise distribution of the polarization inside the fiber. In this scenario, we analyze the case of polarization fluctuation, which average state is aligned with the prepared one. In order to do so we can divide our channel in different steps. First, we consider our source to generate a coherent state with mean photon number $\mu = |\alpha|^2$ in the polarization H:

$$|\psi\rangle = |\alpha\rangle_H |0\rangle_V . \tag{26}$$

Alice then turns this state into the prepared one shown in Eq. (1) by a unitary transformation $A_{\theta,x}$ that turns the polarization into the desired one. The state propagates into the channel that can be divided into a transformation that represents the loss, i.e., C_η and one that represents the polarization fluctuations that we want to model in this section, i.e., $C_{\theta,\phi}$ where θ and ϕ are the random variables representing respectively the polar and azimuthal rotation on the Poincaré sphere. Bob chooses the measurement basis by an analogous transformation of Alice, i.e. $B_{\theta,y}$. Finally, the two polarization modes are separated by a polarizing beam splitter in two paths (as shown in Fig. 3). One of the two paths corresponds to the projection needed and leads to a single photon detector. The resulting state before the PBS has the form:

$$A_{\theta,x} C_{\theta,\phi} C_\eta B_{\theta,y} |\alpha\rangle_H |0\rangle_V = \left|\beta_{0|\theta,x,y,\theta,\phi}\right\rangle_{H'} \left|\beta_{1|\theta,x,y,\theta,\phi}\right\rangle_{V'}, \tag{27}$$

where H' and V' are the proper polarization modes of Bob's PBS and $\beta_{1|\theta,x,y,\theta,\phi}$ is the coherent state amplitude arriving on the single photon detector whilst $\beta_{2|\theta,x,y,\theta,\phi}$ is the amplitude that correspond to the other output port of the PBS.

The transformation we are interested in is $C_{\theta,\phi}$, since this corresponds to our source of misalignment. Without loss of generality, since the rotation acts randomly and independently on each state, we can consider the effect of this transformation already in the input state of the system giving us a new input state

$$|\psi'\rangle = |\alpha\cos(\theta/2)\rangle_H \left|\alpha\sin(\theta/2)e^{i\phi}\right\rangle_V . \tag{28}$$

For simplicity we consider θ to be a random variable with a normal Gaussian distribution, i.e. $P(\theta) = \frac{1}{\sigma_\theta\sqrt{2\pi}} e^{-\frac{\theta^2}{2}\sigma_\theta^2}$ with null mean and standard deviation σ_θ, and ϕ to be uniformly distributed in the interval $[0, 2\pi]$. In order to show the effect of this fluctuation on the detection statistics, it is sufficient to focus on the probability of having a conclusive event (corresponding to $b = 0$). This probability has the form $p(b=0|x,y) = 1 - e^{-|\beta_{0|\theta,x,y}|^2}$, where $|\beta_{0|\theta,x,y}|^2$ is the mean photon number of the coherent state arriving to the detector in the configuration where θ is fixed and Alice and Bob chose their inputs as x and y respectively.

By trivial calculation using $|\psi'\rangle$ as initial state, we can express the value of $|\beta_0|^2$ depending on the specific values of θ and ϕ as follows:

$$\begin{aligned}
&|\beta_{0|\theta,x,y,\theta,\phi}|^2 = \\
&[\cos^2(\theta/2)(r^2 + t^2 + 2rt\cos(x-y)) \\
&+ \sin^2(\theta/2) 2rt(1 - \cos(x-y)) \\
&- \cos(\theta/2)\sin(\theta/2)[(t-r)2\cos(\phi) \\
&- 2\cos(\phi + x - y) + 4r\cos(\phi)\cos(x-y)]]|\alpha^2|,
\end{aligned} \tag{29}$$

where $r = \sin(\theta/2)^2$ and $t = \cos(\theta/2)^2$.

Finally, it is sufficient to take the average of this quantity with respect to the probability density function of θ and Φ in order to obtain the final mean photon number $|\beta_{0|\theta,x,y}|^2$:

$$\begin{aligned}
|\beta_{0|\theta,x,y}|^2 &= \int_0^{2pi} \int_{-\infty}^{+\infty} P(\theta) U(\phi) |\beta_{0|\theta,x,y,\theta,\phi}|^2 d\phi d\theta \\
&= [\frac{1}{2}(1 + e^{-2\sigma_\theta^2})(r^2 + t^2 + 2rt\cos(x-y)) \\
&\quad + \frac{1}{2}(1 - e^{-2\sigma_\theta^2}) 2rt(1 - \cos(x-y))]|\alpha|^2.
\end{aligned} \tag{30}$$

In conclusion, this simple model allows us to evaluate the stability and the fluctuation of our channel just by a simple figure of merit given by the standard deviation σ_θ. It is important to stress that this calculation shows just a possible channel model between Alice and Bob for simulation purposes and are not needed to grant the security of the protocol.

C Experimental implementation

C.1 Mean photon number

Here we discuss the question of how to choose the polarizations and coherent state amplitude α

for the implementation. For a given transmission η, we optimized numerically the key rate R over α and θ (note that the second angle ϕ is already defined, and depends on the number of states n). To do so, we also take into account the probability of having a dark count (estimated to $p_{dc} = 3.24 \cdot 10^{-7}$) and the misalignment error (pessimistically estimated to $\sigma = 4\%$; see Appendix B). For practical convenience, we then decided to fix θ in order to be able to swipe automatically over α while running the experiment and not having to realign manually the polarization states for each transmission. We choose θ so that we can get key for a large span of transmissions η. For every η and fixed $\theta = 0.6$ rad, we extract the optimal α. In Fig. 5, we show the numerical optimization for $n = 2$.

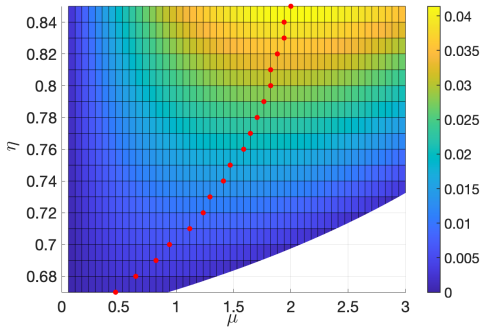

Figure 5: $n = 2$. Key rate R as a function of η and $\mu = |\alpha|^2$ for $\theta = 0.6$ rad, $\sigma = 4\%$ and $p_{dc} = 3.24 \cdot 10^{-7}$. The red dots indicate the optimal value of $|\alpha|^2$ for each η.

C.2 Illustration of the self-testing feature

Fig. 4 (Middle) shows the evolution of the guessing probability $p_g(e = k|\text{succ})$ throughout the measurement's duration for the two-state protocol implementation with the 4.8 km optical fiber. After 2.5×10^6 runs, the detector efficiency was intentionally reduced from 90 % to 42 % by lowering the bias current of the SNSPD. This was done to mimic a malfunction or tampering by Eve of the detector. The bias current was kept low until 3.5×10^6 runs had elapsed. During this time, the guessing probability rises up to $p_g = 1$, hence no key can be extracted. The parties become aware of a deviation from the correct operation regime, and can then abort the protocol and re-calibrate their devices.

C.3 Implementation of the protocol with $n = 3$ states

For the 3-state protocol we set the polarizations to $\theta = 0.7$ rad; and $\phi = \frac{2}{3}\pi$ as defined in the presentation of the protocol. Because of the limiting patterning effect of the polarization controller in this scenario, the measurement was performed without the random basis choice, i.e., using fixed states and measurement basis. The detection rate was kept at 1 kHz but the states were switched after 50 ms, following a fixed sequence. As a result, we observe a reduction of the the misalignment error down to $\sigma = 2.5\%$. The measurement was run with an intensity $|\alpha|^2 = 0.647$ and transmission $\eta = 65\%$ (setting the variable attenuator at the output of Alice's device), integrating over a period of 2 hours. As the variable attenuator sets both the intensity of the prepared states and the transmission of the channel, given by the product $|\alpha|^2\eta = c$, the data can in fact be analyzed in different ways. Specifically, we can consider a value of η between 80% and 60%. The corresponding value of the intensity is then given by $|\alpha|^2 = c/\eta$. We observed a positive secret key rate for η as low as 63% and 67% for the asymptotic and finite-size analysis, respectively (see Fig. 4 (Bottom)). To check the consistency of the data, we performed a Monte Carlo simulation, with finite-size analysis (assuming errors $\sigma = 2.5\%$). The expected range is given by the blue area, and all data points are inside.