

Preparing a commercial quantum key distribution system for certification against implementation loopholes

Vadim Makarov,^{1,2,*} Alexey Abrikosov,^{1,2} Poompong Chaiwongkhot,^{3,4,5,6} Aleksey K. Fedorov,^{1,7} Anqi Huang,⁸ Evgeny Kiktenko,^{1,2,9} Mikhail Petrov,^{1,2,10,11} Anastasiya Ponosova,^{1,2} Daria Ruzhitskaya,^{1,2} Andrey Tayduganov,^{2,7} Daniil Trefilov,^{1,2,10,11,12,13} and Konstantin Zaitsev^{1,2,10,11,12}

¹*Russian Quantum Center, Skolkovo, Moscow 121205, Russia*

²*NTI Center for Quantum Communications, National University of Science and Technology MISiS, Moscow 119049, Russia*

³*Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*

⁴*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*

⁵*Department of Physics, Faculty of Science, Mahidol University, Bangkok, 10400 Thailand*

⁶*Quantum technology foundation (Thailand), Bangkok, 10110 Thailand*

⁷*QRate, Skolkovo, Moscow 143025, Russia*

⁸*Institute for Quantum Information & State Key Laboratory of High Performance Computing, College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, People's Republic of China*

⁹*Steklov Mathematical Institute, Russian Academy of Sciences, Moscow 119991, Russia*

¹⁰*Vigo Quantum Communication Center, University of Vigo, Vigo E-36310, Spain*

¹¹*atlanTTic Research Center, University of Vigo, Vigo E-36310, Spain*

¹²*School of Telecommunication Engineering, Department of Signal*

Theory and Communications, University of Vigo, Vigo E-36310, Spain

¹³*National Research University Higher School of Economics, Moscow 101000, Russia*

(Dated: October 31, 2023)

A commercial quantum key distribution (QKD) system needs to be formally certified to enable its wide deployment. The certification should include the system's robustness against known implementation loopholes and attacks that exploit them. Here we ready a fiber-optic QKD system for this procedure. The system has a prepare-and-measure scheme with decoy-state BB84 protocol, polarisation encoding, qubit source rate of 312.5 MHz, and is manufactured by QRate in Russia. We detail its hardware and post-processing. We analyse the hardware for any possible implementation loopholes and discuss countermeasures. We then amend the system design to address the highest-risk loopholes identified. We also work out technical requirements on the certification lab and outline its possible structure.

I. INTRODUCTION

Over the past three decades, quantum key distribution (QKD) has progressed from a proof-of-principle tabletop demonstration [1] to commercial deployment in fiber networks in many countries [2–4]. Cryptographic systems must ensure reliable and secure operation, and therefore undergo a formal certification procedure [5, 6]. This involves analysing the system's robustness against known vulnerabilities that exploit the imperfections in its hardware [7–12]. While both national and international certification standards for QKD are being developed [13, 14], the full certification ecosystem for it is not yet established.

Preparing a QKD system for certification involves (i) documenting the system in sufficient detail for it to be analysed, (ii) analysing it, (iii) patching the security loopholes found [15], and (iv) proposing the requirements for future certification tests. Here we perform these four steps for a commercial system from QRate, utilising the latest developments in vulnerabilities, countermeasures, and security proofs. These steps are to be followed by (v) the actual implementation of certification, however

in Russia this last step is classified, thus our paper probably constitutes all we can publicly disclose about this system's preparation to it.

The paper is organised as follows. In Section II, we define a risk factor that tells the manufacturer whether a given vulnerability is easily exploitable and thus must be closed by a countermeasure before the system is passed to the formal certification. In Section III, we decide how to combine existing security proofs for systems with imperfections. We describe the QKD system under evaluation in Sec. IV, including a fairly detailed disclosure of its optical scheme and post-processing protocol. We discuss every potential vulnerability in this system and possible countermeasures to them in Sec. V and summarise this initial analysis in Sec. VI. Section VII reports how the manufacturer has subsequently addressed the high-risk vulnerabilities. We outline the test capabilities the certification lab should have in Sec. VIII and conclude in Sec. IX.

II. RISK EVALUATION SCALE

The company should prioritise patching security issues that are more easily exploitable in practice [15]. We thus need to score each issue identified. The cryptography

* makarov@vad1.com

community commonly ranks attacks by their likelihood of success, time and other resources needed to execute them. The proposed ISO standard for QKD attempts to follow this practice [13]. It uses a set of factors to evaluate an attack potential that follows a standard evaluation method for security products. Unfortunately, we find that the possible values currently suggested of each factor are not suitable for QKD yet. All the vulnerabilities we discuss in our paper score as either “highly resistant to attack” or “beyond-high” in the ISO scale. I.e., developing a working exploit for a vulnerability regarded as easy today in the QKD community still requires a multiple-experts team, longer than six months of work, and bespoke equipment (of which a good example is [16]). It is then difficult to differentiate between the vulnerabilities, as they tend to be off that scale. The values of each factor in the ISO standard clearly need to be adjusted before they become applicable to QKD.

Meanwhile, we temporarily adopt an alternative risk evaluation scale that essentially spans the difficulties of exploit higher than the ISO scale. This allows us to compare the risk of vulnerabilities. Our empirical scale is the following. If the security issue has been eliminated or addressed sufficiently well such that it no longer presents a security risk, its overall risk factor is set to ‘solved’. For those issues where this is not the case, we first evaluate the severity of the issue by three parameters.

- *Loophole likelihood:* How likely is it that the particular loophole exists in the system, according to our present knowledge? If its existence has been confirmed or suspected to be likely, this parameter has value 1. If the loophole is considered to be possible in principle but not very likely (and we have not tested the system yet to find out for sure if the particular imperfection exists), the value is 0. The idea here is that security problems known to be more likely to exist should be more important for the manufacturer to address.
- *Future or current technology:* If the loophole may be exploited with today’s technology (i.e., all the components for building a full security exploit can be purchased or easily developed), the value is 1. If it would need future technology that does not exist yet, the value is 0. For example, if the exploit requires an adversary Eve to use 95% efficient single-photon detectors in her setup, these are available commercially today. If the exploit however requires Eve to use a lossless optical communication line, it is of course possible in principle (physics doesn’t prohibit lossless lines) but not available today. In the latter case, Eve would need to wait until optical fiber with much lower loss than exists today is produced [17], or until high-quality quantum repeaters are built so that she can use them to implement lossless quantum teleportation over today’s lossy optical fibers. Both are long shots for Eve in practice, this she would not be able to build the exploit

today and the security problem is less urgent for the company to address.

The value is also 0 if it is presently not known how to construct an attack.

- *Amount of key leakage:* If the attack provides Eve full or nearly full information about the secret key, the value is 1. If the attack can only provide Eve a minor partial information about the secret key, the value is 0. For example, most intercept-resend type attacks give Eve 100% (or close to that) information about the secret key [16]. It would then be relatively easy for her to attack a classical cryptographic algorithm that subsequently uses this compromised key. However if the attack only results in the leakage of partial key information, this presents Eve two additional practical challenges. First, she would need to construct her exploit apparatus very carefully such that it works almost perfectly and does not introduce side effects (such as additional errors in the key) that would make Eve’s eavesdropped key information zero. Second, she needs to solve a non-trivial classical cryptanalytic task when attacking the classical cryptographic scheme with only partial key information. Although these problems have not been explored, we feel that vulnerabilities that deliver Eve full or nearly full key information should be more urgent for the company to address.

We add up the values of the three parameters and evaluate the overall risk factor. If the sum is 0 or 1, the overall risk is low (**L**); 2, medium (**M**); 3, high (**H**). As the reader will see, these three rough risk grades are evenly distributed across today’s QKD vulnerabilities.

III. SECURING THE SYSTEM IN THE ABSENCE OF A UNIFIED SECURITY PROOF

At least five attacks in this report require updating the key rate formula according to available security proofs taking each individual attack into account. However, there is no unified security proof that takes into account all these attacks simultaneously and offers a general key rate formula simultaneously accounting for the effects of several attacks. Having said that, we have to mention recent attempts in this direction. In order to take into account various source flaws and side channels, the so-called loss-tolerant QKD protocol was proposed by Tamaki and his coworkers [18]. The three-state ($|0_Z\rangle, |1_Z\rangle, |0_X\rangle$) loss-tolerant protocol with imperfect state preparation is studied together with either intensity fluctuations [19, 20] or Trojan-horse attacks [21, 22], and can simultaneously account for correlations among the source pulses [20, 22]. Similarly, the four-state ($|0_Z\rangle, |1_Z\rangle, |0_X\rangle, |1_X\rangle$) loss-tolerant protocol with imperfect state preparation has been investigated

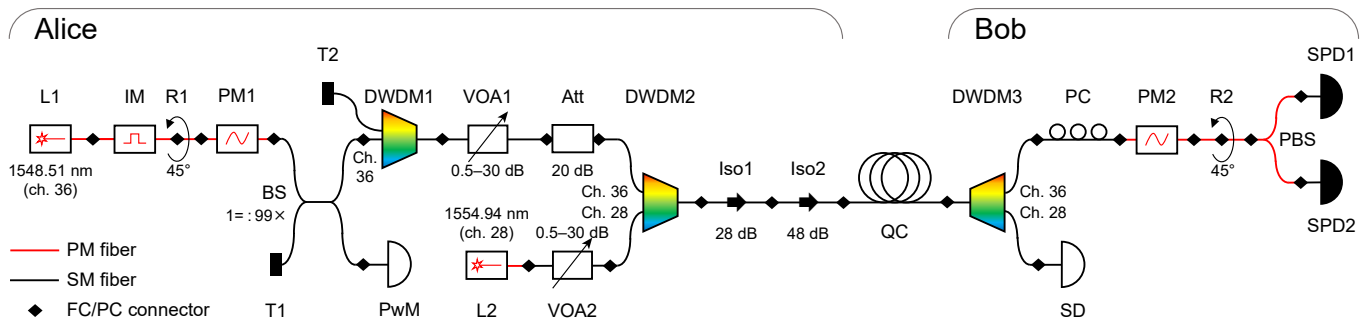


FIG. 1. Optical scheme of the QKD system under evaluation. L, lasers [L1: Nolatech DFB-1550-5PM; L2: Shengshi Optical SWLD-1554.94-FC/PC-05-PM(DFB)]; IM, intensity modulator (iXblue MX-LN-10); R, FC/PC connector with 45° rotation (custom-made by QRate based on bulkhead adapter Opneti AD-FC/SM-SP04); PM, phase modulator (iXblue MPZ-LN-10); T, optical terminator; BS, beamsplitter (with its splitting ratio noted; Opneti CP-S-P-1x2-1550-1/99-900-1-0.3-FC-3x54); PwM, power meter (Thorlabs PM101 with S154C sensor); DWDM, dense-wavelength-division multiplexer (DWDM1 and DWDM3: Opneti DWDM-1-100-36-900-1-0.3-FC; DWDM2: Opneti DWDM-1-100-28-900-1-0.3-FC); VOA, variable optical attenuator (Opneti SVOA-B-1550-30-5.2250-1-1-FC); Att, fixed attenuator (Opneti FOA-P-1-20-FC); Iso, polarisation-independent isolator (Iso1: Opneti IS-S-P-1550-900-1-0.3-FC-5.5x35; Iso2: Opneti IS-D-P-1550-900-1-0.3-FC-5.5x35); QC, quantum channel; SD, synchronisation detector (Fujitsu FRM5W232BS); PC, polarisation controller (General Photonics MPC-4X-7-P-FC/PC); PBS, polarising beamsplitter (Opneti PBS-1x2-P-1550-900-1-0.8-FC); SPD, single-photon detector. Note that the components used in the system at the time of the initial analysis may be replaced with other similar models before the final certification, especially because some of the original components may no longer be available in Russia.

as well, combined with the Trojan-horse attacks [23, 24] and correlations among the source pulses [23]. One has to point out that only a single-photon source is considered in [21–24]. The “standard” decoy-state BB84 protocol with four encoding states and three intensities is also studied in [25, 26], where the Trojan-horse attack is considered along with the vulnerabilities of detector backflash [25] or detector efficiency mismatch [26]. However, several new attacks, such as light injection and induced photorefractive, seem to be not included in the security proofs, which is a subject for future research. Thus, we conclude that no complete security proof currently exists that takes into account all the potential imperfections and side channels we list in Sec. V. Deriving such security proof is an open academic question, and a very non-trivial one.

Without this theoretical treatment, we are in the realm of guessing. We still, however, need to make a practical decision how to treat these vulnerabilities in QRate’s system. Our first idea was to sum algebraically the key rate corrections owing to the different vulnerabilities. We discussed this idea with theoreticians [27] and, while they conceded it might turn out to be approximately correct, no one really liked it.

Our second idea is to use hardware countermeasures (filters, isolators, etc.) to minimise the key rate reduction of *each and every vulnerability considered alone* to a negligible level. This means that for every individual vulnerability for which a security proof is available, the hardware is characterised, then reinforced and improved until the proof gives a very small correction to the key rate and maximum transmission distance comparing to the case of a perfect hardware. An incidental advantage of this is that the key rate formula for the perfect hardware can be used in the system.

We hope that the latter approach turns out to be robust, and suggest to use it to claim the QRate’s system is secure. Again, there is no strict proof of that, but this is a reasonable best-practice approach we can currently do. The rest of this report adopts this approach.

IV. SYSTEM UNDER EVALUATION

The QKD system we study is an industrial prototype under development at QRate. It has a prepare-and-measure scheme and uses a decoy-state Bennett-Brassard 1984 (BB84) protocol with polarisation-encoded states at approximately 1550 nm wavelength and 312.5 MHz clock rate. The optical scheme is shown in Fig. 1 and photos in Fig. 2. Further details can be found in a Russian-language Ph.D. thesis [28].

The system manufacturer has also provided us a Design specification sheet of the overall scheme (dated 2018-11-07) that contains a high-level description of the hardware and software structure, as well as later documents on extensive changes and updates made by the end of 2021. We have received further oral information and written notes on various aspects of design and manufacturing from the company engineers. At this evaluation stage, we have not yet tested the system hardware for most vulnerabilities (with a few exceptions that will be noted through the text).

The system software uses the post-processing procedure containing the following standard steps.

1. *Sifting*. Bob announces the positions of registered pulses and their measurement bases. Alice announces whether her basis matches and they dis-

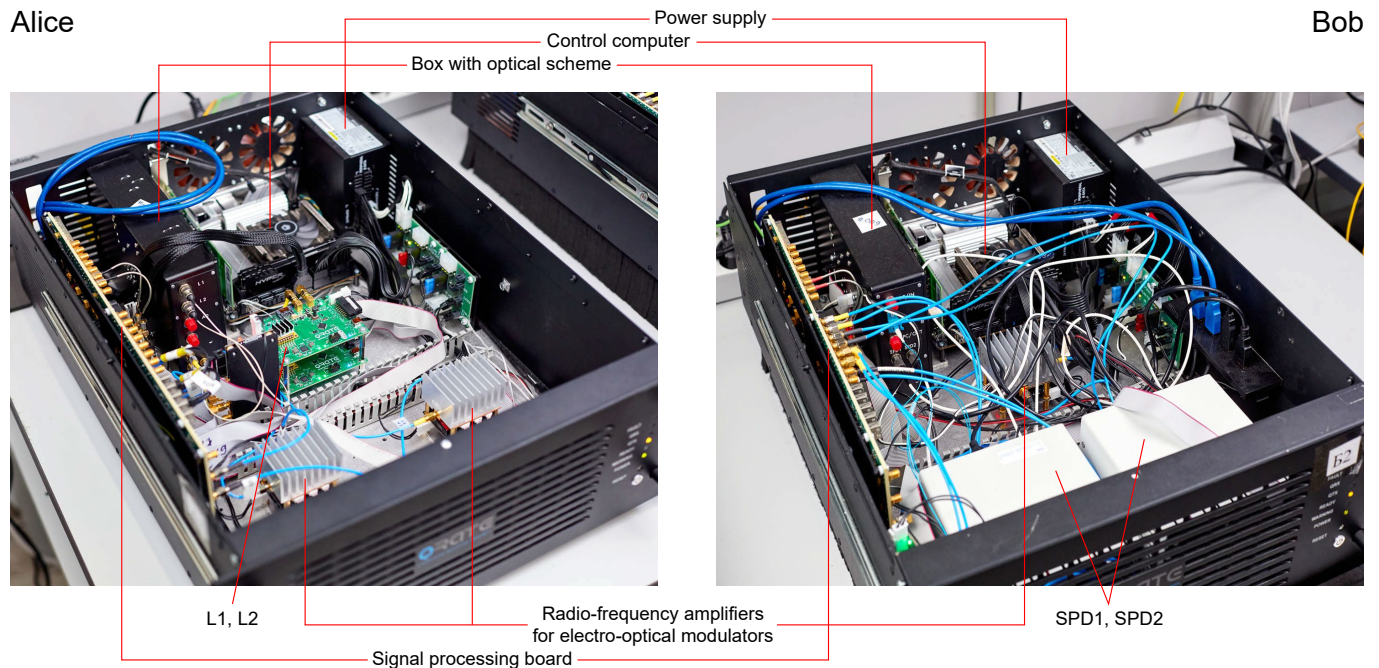


FIG. 2. Quantum key distribution system under evaluation (a prototype built in 2021), with covers removed from Alice and Bob.

card all events with incompatible bases. For matching bases, Alice also announces the type of each pulse (signal or decoy).

2. *Information reconciliation.* Alice's sifted key is taken as a reference, while Bob attempts to find and eliminate the discrepancies (errors) between the keys. For this purpose the low-density parity-check (LDPC) codes are used [29, 30].
3. *Verification and parameter estimation.* The identity of error-corrected keys is verified using a modified PolyP32 hash function [31, 32].
4. *Estimation of the level of eavesdropping.* Bob estimates the total amount of information leaked to Eve during the quantum phase and the previous post-processing steps, and computes the secret key length ℓ_{sec} (A12) according to the model from [33] taking into account the finite-key-size effects. If $\ell_{\text{sec}} \leq 0$, Alice and Bob abort the protocol and proceed to the next generated raw key block.
5. *Privacy amplification.* In order to get rid of Eve's residual information about the verified key, it is compressed using a 2-universal hash function from the Toeplitz family [34, 35]. As a result, Alice and Bob obtain a common shorter key of length ℓ_{sec} , Eve's information about which is now negligible.

A more detailed description of the post-processing is given in Appendix A.

V. POTENTIAL VULNERABILITIES

In order to simplify the task of security evaluation, and ease of understanding, we have subdivided the system implementation into several layers according to the hierarchical order of information flow [15] (recapped in Appendix B). In this work, we perform a complete security analysis of the bottom four layers (Q1–Q4) that correspond to optics, analog electronics, driver and calibration algorithms, and operation cycle of the system. For these layers, we aim to examine all suspected implementation security issues according to the current knowledge. For higher layers Q5 and up (from QKD protocol post-processing and up), we cannot perform a complete security evaluation as they lay outside the expertise of most of the authors. Nevertheless, we point out a few issues in the layer Q5 and we include its fairly detailed description in Appendix A to aid any independent analysis.

Based on the information received about the system, we have identified a number of potential security issues that might be exploitable by Eve. A summary of them is given in Table I. Note that QRate has subsequently addressed all the high-risk issues, as detailed later in Sec. VII. We now explain the identified issues.

A. Choice of QKD protocol

The choice of the QKD protocol and scheme is one of the most important decisions a designer makes. It affects the product through its lifetime.

TABLE I. **Summary of potential security issues in QRate 312.5 MHz QKD system found at its initial evaluation (completed in January 2022).** Q , system implementation layers involved (see [15] or Appendix B).

Potential security issue	Q	Target component	Action recommended to the company	Risk evaluation
Choice of QKD protocol	Q5	Protocol	None.	Solved
Superlinear detector control	Q1–5,7	SPDs	The development of the photocurrent-measurement countermeasure should continue at the company. It should be tested in our lab.	H ^a
Detector efficiency mismatch	Q1–5	SPDs, Bob’s PM	Update the key rate equation. Spectrally characterise Bob’s components. Discuss countermeasures to timing attacks.	H ^a
Detector deadtime	Q1,2,5	SPDs	Supplement the hardware simultaneous deadtime with implementing it in post-processing.	H ^a
Trojan-horse	Q1,2	Alice’s optics	Characterise Alice’s components in a wide spectral range. Install additional isolators and, possibly, spectral filters.	L
Laser seeding	Q1,2	Laser	None.	Solved
Light injection into Alice’s power meter	Q1–3	IM	Characterise Alice’s components in a wide spectral range. Install additional isolators and, possibly, spectral filters.	L
Induced photorefraction ^b	Q1–3	Alice’s IM and PM	Characterise Alice’s components in a wide spectral range. Optical measurements should be done in our lab.	M
Laser damage	Q1	Alice’s & Bob’s optics	Install an additional sacrificial isolator at Alice’s exit.	M
APD backflash	Q1,2	SPDs	Characterise Bob’s components in a wide spectral range. Measure backflash photon emission probability of the SPD.	M
Intersymbol interference	Q1–3	Alice’s active components	Optical measurements should be done in our lab.	L
Imperfect state preparation	Q1–3,5	Alice’s optics	Optical measurements should be done in our lab.	L
Calibration via channel Alice–Bob	Q1–5	SPDs, IM, PM	The analysis team did not know how to solve this and proposed to discuss with QRate. QRate has subsequently found solutions acceptable for the manufacturing process, see Sec. VII.	H ^a
Quantum random number generator	Q5	Protocol	Implement the quantum random number generator and integrate it into the system.	L
Compromised supply chain	All	Any	Learn mitigation strategies from the national cryptography licensing authority.	M

^a All the high-risk issues identified have been addressed by QRate before publication of this report, see Sec. VII.

^b Issue added in mid-2022 when we learned about a recent study [36].

QRate has chosen the best understood and most widely studied scheme and protocol: the prepare-and-measure (one-way) scheme and BB84 protocol with decoy states of three intensities (vacuum, decoy, signal). This choice has the advantage that complete general security proofs are available that have been widely scrutinised for correctness. An additional advantage crucial for our analysis is that modifications of these security proofs that take into account various hardware imperfections are often also available. We cite these through this report.

We remark that not every company has made the same choices. Sometimes the motivation of developing its own intellectual property prevails and a less studied proto-

col that lacks the general security proof is chosen. This often raises questions. For instance, the excellent current commercial QKD system by ID Quantique (Switzerland) [37, 38] implements a coherent-one-way protocol [39]. This protocol lacked the general proof at the time of its initial commercialisation in 2014. Subsequently, quantum attacks on the original coherent-one-way protocol have been discovered that severely limit the key rate and communication distance [40]. Although the latest version of the system [37] uses a modified protocol, the general security proof for it is also not available. In other examples, the subcarrier-wave QKD system being commercialised by Quantum Communications Ltd. (Russia)

[15] still has its security proof in development [41]. For the system with a geometrically-uniform-coherent-states QKD protocol [42] commercialised by Infotecs (Russia) [43], the integrity of its security proof is being debated in the scientific community [44]. None of the available partial security proofs for these systems incorporate device imperfections, which may further hinder their analysis.

Risk evaluation: Solved.

Further suggestions: None. Decoy-state BB84 protocol is the safest available choice for the prepare-and-measure QKD scheme.

B. Superlinear detector control

Superlinear detector control attacks are based on three phenomena. First, most single-photon detectors (SPDs) are threshold detectors, which means that they cannot resolve the number of photons in a pulse. When they produce a detection event, called a click, they do not distinguish whether it has been caused by one or multiple photons. Second, the SPD's detection efficiency of multiphoton pulses may exhibit a so-called superlinearity effect [45]. SPDs are usually characterised by their quantum efficiency η , which is the probability to detect a single photon ($\eta \sim 10\%$ for QRate's SPD). For a multiphoton pulse the detection probability can be estimated as

$$p_{\text{det}}(n) = 1 - (1 - \eta)^n, \quad (1)$$

where n is the number of photons in the pulse. An SPD whose multiphoton detection probability is higher than Eq. (1) exhibits superlinear behavior. The third phenomenon is a threshold level shift, which is the ability of the detector to reduce its quantum efficiency partially or completely to zero. Engineers exploit the latter effect in a gated regime to decrease the detector's dark count rate [46]. The reverse-bias voltage at an avalanche photodiode is lowered between the gates, so that the detector is insensitive to single photons ($\eta = 0$) in between the gates. It then behaves as a normal photodiode and may only respond to bright light pulses at this time, with a classical threshold on the pulse energy [47].

Several attacks that exploit these and other phenomena in the SPDs have been developed and multiple countermeasures to them have been proposed. This is arguably the most difficult group of vulnerabilities in today's QKD. For readers not familiar with these developments, we survey them in Appendix C.

Features of the QKD system under analysis: The detection system is developed by QRate with the use of the avalanche photodiode (APD) PGA-025u-1550TF based on InGaAs/InP structure from Princeton Lightwave. From our discussion with QRate's engineers, we have found that no measures have been taken to prevent the superlinearity detector control attacks. As our preliminary detector tests show, the detector is blinded with continuous-wave (cw) light of 3 μW (-25 dBm)

power. It allows total control at 250 μW (-6 dBm) blinding power and trigger pulse energies $E_{\text{never}} = 12$ fJ and $E_{\text{always}} \gtrsim 22$ fJ (see Appendix D).

In QRate implementation shown in Fig. 1, synchronisation detector (SD) can be used as a watchdog (see Countermeasure 4 in Appendix C). However we think this would be a bad idea, for the following reasons. First of all, the SD is not sufficiently sensitive, its threshold starting at a few microwatt (-20 to -30 dBm) level. The presence of demultiplexer (DWDM3) adds about 35 dB to this level (at the particular wavelength of 1548.5 nm). Secondly, the SD's sensitivity can be controllably reduced by the laser damage attack [48, 49]. Thirdly, putting extra functionality on the SD would complicate synchronisation routines that are already far from perfect (see Sec. VC).

A more promising approach is to add a photocurrent measurement to the SPDs (see Countermeasure 3 in Appendix C). QRate has implemented this measurement at a stand-alone sinusoidally-gated SPD, but haven't integrated it as a countermeasure into the system yet. Our preliminary tests of this implementation in a setup from Appendix D show a countermeasure readout (roughly proportional to a logarithm of averaged APD photocurrent) of 400–1200 arbitrary units under single-photon pulses, depending on the count rate. Under the blinding attack, the readout is 2100–2400 arbitrary units. There is a clear separation between the normal operation and blinding, which is encouraging. However this countermeasure needs to be tested with a pulsed blinding [50–52] and be fully integrated into the QKD system. We treat this problem further in [53, 54].

The after-gate attack is probably possible in the current implementation, especially given that Eve may control the timing synchronisation inside Bob and that Bob registers clicks with a coarse 3.2-ns resolution corresponding to one bit period (Sec. VC). One possible countermeasure would be to make the phase modulator pulse shorter than the detector gate, i.e., shorter than 400 to 800 ps. This however can be difficult to implement and may lead to less accurate state preparation (see Sec. VL). Another possible countermeasure is a precise click time measurement, however the detector jitter and timing drift may make this difficult to implement.

Risk evaluation: **H** (1 vulnerability is likely exploitable, 1 with current technology, 1 might give Eve high key information).

Further suggestions: We suggest the company to finish the implementation of the photocurrent-measurement countermeasure (which has been built and tested preliminarily). Our lab will test it in a stand-alone detector against the blinding, after-gate, and falling-edge attacks. We will then possibly repeat the tests in the complete QKD system. This should, at least, allow the company to claim that the system is protected against the detector blinding attack.

If the detector's vulnerability to the after-gate and falling-edge attacks is experimentally confirmed, counter-

measures against them would require a discussion with QRate engineers. We are unsure what solutions are practical given the high time precision and calibration requirements on the PM pulse.

Developing a measurement-device-independent or twin-field commercial system [55–57] is a radical alternative that may be considered, as this would remove all the detector vulnerabilities. However this is a major business decision influenced by many factors.

We remark that the work currently progresses according to the above suggestions, as detailed in Secs. VII and VIII. This note applies to every potential vulnerability from here on.

C. Detector efficiency mismatch

In a theoretical security proof it is assumed that Bob’s SPDs are identical [58]. For real-world SPDs that are not identical, there are three possible mismatches that have to be included in the security proof.

1. Static efficiency mismatch. The average photon detection probability in the SPDs is 10% [28]. If we assume that one SPD has 9% efficiency and another 11%, the probability ratio of bits detected would be 45 : 55 instead of 50 : 50, with no Eve’s influence. This asymmetry gives Eve some *a priori* information about the raw key. While the current QRate’s firmware ignores this issue and assumes the equal 50 : 50 probabilities, the company plans to update the key rate equation to one that takes into account unequal static probabilities, according to the security proof [59, 60].

A simpler solution that does not require the modification of Eq. (A12) is the “four-state measurement” scheme, originally proposed as a countermeasure against the time-shift attack [61]. Bob randomly chooses not only his basis but also bit-0 and bit-1 assignment of his detectors. In such setup, even if Eve has the information about which detector clicks, she still does not know Bob’s bit value since she is not aware about which detector corresponds to bit-0. During the sifting communication rounds Bob announces the bit positions when the detectors were “swapped”, and Alice performs a bit-flip in respective positions on her side. In this way, the distribution of zeros and ones becomes uniform. The potential loophole of the four-state measurement method is that Eve may try to read out Bob’s detector assignments by injecting a strong pulse like in the Trojan-horse attack [62–64].

2. Time mismatch. The detectors are sinusoidally-gated and are sensitive to single photons for about 800 ps out of the 3.2 ns gate period. Any gated detectors are likely vulnerable to time-shift attacks (TSA) [65, 66].

3. Wavelength mismatch. Characteristics of all optical components depend on the wavelength, which often leads to loopholes. On Bob’s side an attack is in principle possible using wavelength dependence of the polarising beamsplitter (PBS) and SPDs [67, 68]. A combination of this attack with other attacks should also be considered. Spectral characterisation of Bob’s components that is necessary for further study of this attack is discussed in Appendix E.

Risk evaluation: **H** (1 vulnerability is likely exploitable, 1 with current technology, 1 might give Eve high key information).

Further suggestions: Although the security proofs [59, 60] derive the key rate equation that accounts for the static efficiency mismatch, they are not applicable to the mismatch in the time and wavelength domain that Eve can dynamically control. This leaves us with the only realistic option to solve this problem by implementing a four-state Bob (i.e., Bob who randomly swaps or not swaps his detectors’ assignment to bit values 0 and 1 by applying or not applying an additional π phase shift at his PM [61]). This eliminates all the efficiency mismatches and corresponding corrections to the key rate equation. However, Bob then needs to additionally guarantee a certain amount of isolation against the Trojan-horse attack on him, which becomes necessary because the detectors’ assignment has to remain secret. A security proof that estimates the required amount of the latter isolation is not available in the literature. It needs to be developed and Eq. (A18) amended by including a Trojan-horse leakage term.

We remind that measurement-device-independent and twin-field QKD systems do not suffer from the detector vulnerabilities. They may be considered as an alternative solution.

D. Detector deadtime attack

The security proof requires that both Bob’s detectors are sensitive to photons when Bob registers a click. If one detector remains sensitive and clicks from it are accepted as valid while the other detector is having a deadtime, an attack becomes possible [69].

In QRate’s system, whenever one detector clicks, a simultaneous deadtime of about 4.5 μ s is introduced to both detectors, via electrical cross-links between the detector units. Figure 3 shows the effect of the simultaneous deadtime on cross-correlation between the detectors’ clicks. While for the detector that has clicked the deadtime begins instantly, the other detector starts it a few 3.2-ns gating periods later, owing to the delay in the electrical cross-link. We thus see a few cross-clicks early in the deadtime, in which only one detector remains sensitive to single photons. This would present a loophole if these clicks are accepted into the raw key [65]. The gradual recovery from the deadtime is also uneven between

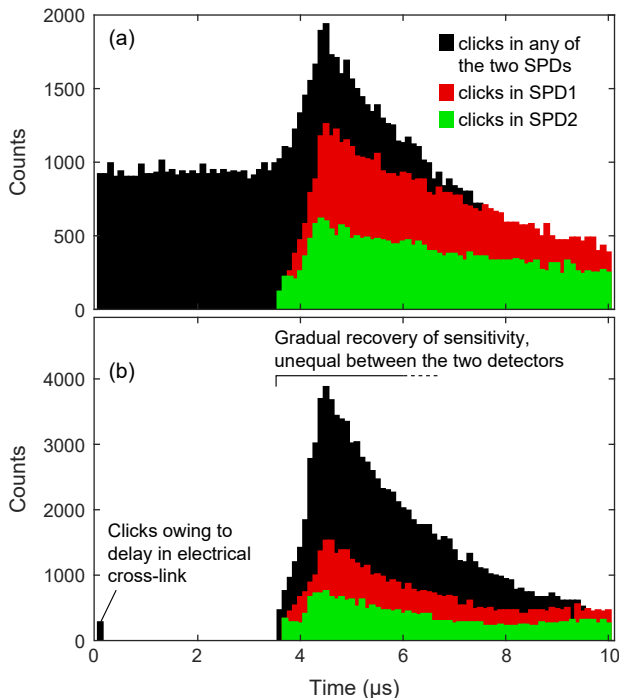


FIG. 3. Test of simultaneous deadtime of Bob’s two SPDs (performed by QRate). The histograms show click rate versus time after a click (a) in independent detectors without simultaneous deadtime and (b) in interlinked detectors with simultaneous deadtime. The detectors are illuminated with Alice’s light typical for QKD operation.

the detector units, with a significant efficiency mismatch visible in the time range starting at $3.4 \mu\text{s}$ and extending roughly to $6\text{--}9 \mu\text{s}$. Similarly, this may leave Eve possibilities to construct attacks.

Risk evaluation: **H** (1 vulnerability is likely exploitable, 1 with current technology, 1 might give Eve high key information).

Further suggestions: Implementing the simultaneous deadtime precisely in post-processing should be sufficient to close this vulnerability. In the case of QRate system, this would supplement the hardware deadtime that already prevents the majority of the unusable clicks, thus reducing their impact on the key rate. The software should then discard all clicks that occur fewer than a fixed number of gates after any click in either detector (corresponding to at least $6 \mu\text{s}$, exact time to be determined by a more accurate cross-correlation measurement). Note that if a click is being discarded, it also renews the discard time period. This should close this vulnerability.

Here we assume that the system does not implement the four-state Bob countermeasure (Sec. V C). If it does, the detector deadtime attack should be re-evaluated.

E. Trojan-horse attack

In this section, we consider the Trojan-horse attack (THA) on Alice’s phase and intensity modulators [62, 63, 70]. In this attack, Eve attempts to read Alice’s IM and PM settings by injecting light, called Trojan photons, into her apparatus. The outbound photons that have passed Alice’s PM and IM will thus contain the secret information about the phase and intensity encoded into them. There are several security proofs for the decoy-state BB84 protocol that take this information leakage into account [71–73]. Here we use the latest proof to calculate the required isolation values in the finite-key regime [73]. For this, we need to upper-bound the intensity (conventionally called ‘intensity’ in QKD but actually meaning energy) of the leaked signals

$$I_{\text{max}} = 10^{-\frac{\alpha_A}{10}} I_{\text{in}}, \quad (2)$$

where α_A is the total loss of the Trojan photons in Alice (in decibel) and

$$I_{\text{in}} = \frac{W_{\text{in}} \lambda}{f_p h c} = \frac{100 \text{ W}}{312.5 \text{ MHz}} \times \frac{1550 \text{ nm}}{1.99 \times 10^{-25} \text{ J}\cdot\text{m}} \quad (3)$$

$$= 2.5 \times 10^{12} \text{ photons per pulse},$$

where f_p is the qubit repetition rate and W_{in} the maximum optical power that can be transmitted through the standard telecommunication optical fiber (assumed here to be 100 W). To estimate α_A we need to know the losses inside Alice; her component parameters are given in Table II. Taking into account that the Trojan photons pass each component twice, we obtain

$$\alpha_A = 2 (\alpha_{\text{IM}} + \alpha_{\text{PM1}} + \alpha_{\text{BS}} + \alpha_{\text{DWDM1}} + \alpha_{\text{VOA1}} + \alpha_{\text{Att}} + \alpha_{\text{DWDM2}}) + \alpha_{\text{Iso1rev}} + \alpha_{\text{Iso2rev}} + \alpha_{\text{Iso1forw}} + \alpha_{\text{Iso2forw}}. \quad (4)$$

This formula incorporates the following assumptions.

1. Here we assume that Eve’s Trojan photons have 1548.51 nm (i.e., channel 36) wavelength. Thus both DWDMs have 1 dB insertion loss and the insertion loss values of the other components can be taken from their data sheets. This allows us to make a quick estimate but is in no way sufficient to treat this vulnerability [64, 74]. Eve is, of course, not limited to this wavelength. She may use any other wavelength if the combined loss at it is lower. None of the components in the QRate system have been characterised in a sufficiently wide spectral range. This data is never available from the component manufacturers, because it is not needed for normal applications, not measured, and not guaranteed. We must thus perform a wide spectral characterisation of all the components ourselves in $\sim 350\text{--}2400 \text{ nm}$ range (see Appendix E), then find the minimum of α_A over this entire spectral range.

TABLE II. **Optical insertion loss α of system components**, in the quantum signal path (L1-QC-SPDs), at the system operating wavelength of 1548.51 nm. The values are taken from component data sheets. The values at other wavelengths are not specified and may differ considerably. Connector loss (typically 0.3 dB) is neglected.

Alice's component	α (dB)	Bob's component	α (dB)
IM	2.7	DWDM3	1
PM1	2.5	PC	0.05
BS	20	PM2	2.5
DWDM1	1	PBS	0.5
VOA	0.5–30		
Att	20		
DWDM2	1		
Iso1 reverse / forward	28 / 0.35		
Iso2 reverse / forward	48 / 0.4		

- The Trojan pulses experience losses and reflections from different surfaces *behind* the IM. However, Eve might manipulate the phase and delay of each consecutive pulse such that the reflections from each of those surfaces arrive at IM in-phase at the same time [14, 75]. Those pulses will interfere constructively, resulting in the total photon number passing through the IM being much higher than a mere sum of individual reflections. This effective reflectance depends on the number of reflective surfaces Eve could exploit. Although measuring individual reflections that are widely spaced apart is possible [63, 70, 71, 76], a general characterisation technique that takes into account closely spaced reflections is complex and not yet proven [14]. Also the individual reflections might be wavelength-dependent, which further adds to the challenge. It is much easier and safer to adopt a conservative assumption that all the photons behind the IM are reflected back [14, 75]. Thus all the losses behind the IM are neglected.
- The variable optical attenuator (VOA) can be set anywhere in the range 0.5–30 dB. It might be used at the lower-attenuation end of the range during QKD, according to QRate. We thus assume here the worst case with the minimum attenuation of 0.5 dB.
- We neglect the loss in FC/PC connectors, which can typically be 0.3 dB per connection.
- Eve can attempt to change the attenuation characteristics of the last isolator (Iso2) by the laser-damage attack (see Sec. VI). Here we do not consider this.

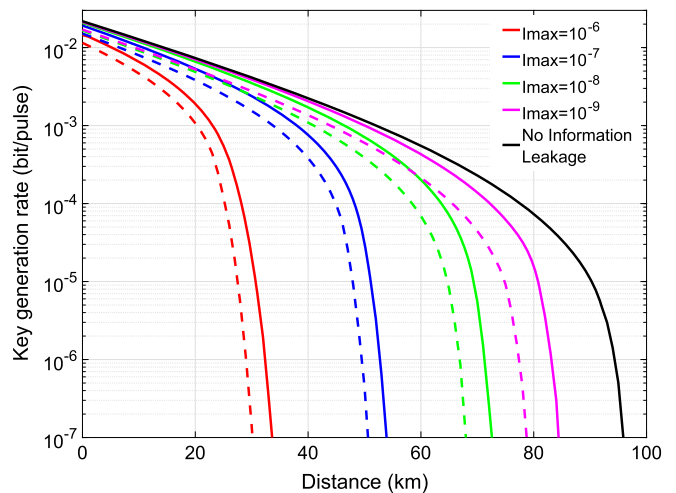


FIG. 4. Secret key rates for different leaked intensities, for a typical QKD system (not QRate's), reprinted from [73]. The total number of transmitted pulses $N = 10^{12}$.

Combining Eqs. (2) to (4) with the data from Table II we obtain $\alpha_A \approx 172$ dB, $I_{\max} \approx 1.5 \times 10^{-5}$ photons per pulse. We can quickly estimate an expected key rate by looking at the plots from [73] calculated for a typical QKD system with slightly different parameters in the finite-key-size regime (Fig. 4). Our high value of I_{\max} is not in the plots and leads to zero key rate at most distances. We can roughly estimate that an additional isolation of more than 40 dB is needed to approach the ideal case with no information leakage (i.e., $I_{\max} \lesssim 10^{-9}$).

Risk evaluation: **L** (0 loophole is unlikely to be present taking into account the very conservative assumptions in our calculation, 0 requires research and possibly future technology to exploit because a complete low-reflectance Trojan-horse attack on the source has not been demonstrated, 0 probably gives Eve low key information).

Further suggestions: First, it is necessary to characterise Alice's optical components in the wide spectral range (Appendix E), to determine the minimum total loss α_A over the entire wavelength range accessible to Eve. Then, a more accurate calculation of the key rate [73] should be done with our actual system parameters and the finite key size. Given our strategy of reducing the information leakage under each individual imperfection to a negligible value (Sec. III), an acceptable key rate reduction threshold should be set arbitrarily (e.g., no less than 0.9 of the ideal-case rate) and additional isolators and, possibly, spectral filters should be installed in Alice to guarantee it. Note that the key rate curves always diverge near the maximum transmission distance (Fig. 4); this means that a restriction should be implemented in the system software to prevent operation close to the distance limit.

F. Laser-seeding attack

From the quantum channel, Eve might be able to inject light into Alice’s laser diode (L1) and modify its emission characteristics, e.g., phase, intensity, and wavelength [77–80]. According to previous research, the injection power reaching the connector of Alice’s laser should be in the milliwatt range (assuming the laser has a built-in optical isolator) [78] or nanowatt range (assuming the laser without the built-in isolator) [80]. Similarly to Sec. V E, we assume the laser power entering Alice is 100 W. Then the loss in Alice for the laser-seeding attack

$$\alpha_{As} = \alpha_{IM} + \alpha_{PM1} + \alpha_{BS} + \alpha_{DWDM1} + \alpha_{VOA1} + \alpha_{Att} + \alpha_{DWDM2} + \alpha_{Iso1rev} + \alpha_{Iso2rev} = 123.7 \text{ dB.} \quad (5)$$

The continuous-wave power reaching L1 $W_{L1} = 10^{-\alpha_{As}/10} W_{in} \approx 40 \text{ pW}$, which is already orders of magnitude lower than the power needed for the successful hacking. Note that additional isolation will be added to Alice to protect her against the Trojan-horse attack. With this large margin, we consider this vulnerability to be eliminated.

Risk evaluation: Solved.

Further suggestions: None.

G. Light injection into Alice’s power meter

The Alice’s internal power meter (PwM, see Fig. 1) is used to maintain the working point of her intensity modulator (IM; iXblue MX-LN-10). This intensity modulator internally consists of a Mach-Zehnder interferometer with a fast-modulation section and bias section in its arms. The zero point of the interferometer drifts over time and requires compensation by applying a static voltage at the bias section. The power meter indirectly measures the deviation from the zero point, by measuring an average power of the mix of vacuum, decoy, and signal states emitted by Alice in the normal QKD operation. If the zero point drifts, this power deviates from a factory-preset value, which lies in the range of 2–5 μW . The difference acts on the bias voltage via a slow negative-feedback loop implemented in the system software. Currently the PwM is implemented with Thorlabs PM101 power meter with S154C sensor. The company plans to replace it with a discrete photodiode and their own current measurement circuit.

Injecting additional light into the PwM externally would thus cause the IM’s working point to be set improperly. This would change the intensities of vacuum, signal, and decoy states, as well as their ratios. Notably the intensity of the vacuum state would be increased. This may lower the actual secure key rate below that calculated by the system.

Let’s roughly estimate how much power Eve might inject in the current system at its operating wavelength of

1548.51 nm, similarly to Sec. V E. The loss in Alice

$$\alpha_{Ap} = \alpha_{DWDM1} + \alpha_{VOA1} + \alpha_{Att} + \alpha_{DWDM2} + \alpha_{Iso1rev} + \alpha_{Iso2rev} = 98.5 \text{ dB.} \quad (6)$$

Here we conservatively assume the injected light totally reflects at the BS (the actual reflection coefficient is tricky to calculate owing to possible interference effects Eve might exploit). The upper bound on the power reaching PwM $W_{PwM} = 10^{-\alpha_{Ap}/10} W_{in} \approx 14 \text{ nW}$, which is a fraction of the power it measures in the normal operation. This leaves a small risk Eve might manage to tamper with the operation of PwM and the state intensities emitted.

Risk evaluation: **L** (1 vulnerability is known to exist in principle, 0 requires significant research and possibly future technology to exploit, 0 probably gives Eve low key information). Note added in 2023: an explicit attack on measurement-device-independent QKD that exploits this vulnerability has been published [81], slightly raising the risk.

Further suggestions: Re-evaluate W_{PwM} after additional isolation is added to Alice to protect her against the Trojan-horse attack. This will likely solve this vulnerability as well.

H. Induced-photorefractive attack

Recently, a new light-injection attack based on photorefractive effect in modulators has been proposed [36, 81, 82]. A demonstration has been made of Eve’s shifting the bias point of Bob’s lithium-niobate device by illuminating it using 405 nm laser emission with power of just 3 nW. This might open security vulnerabilities and, in particular, in the case of variable optical attenuators, enables Eve to steal a secret key being undetected by the legitimate users. It is also claimed that the photorefractive effect is effective over a wide range of wavelengths (from ultraviolet to even 1549 nm [83]).

In the QRate system, lithium niobate devices in Alice, namely IM and PM1, prepare the quantum state. They both might be affected by this attack. In the case of the phase modulator, a shift of its working point can have effects similar to those considered in Secs. V L and V M. In the case of the intensity modulator, the effect will be similar to that in Sec. V G. I.e., the induced-photorefractive attack on the modulators is a potential vulnerability.

Similarly to the light-injection attacks considered in Secs. V E to V G, we really need a wide spectral characterisation of the system components to treat this vulnerability. The photorefractive effect in lithium niobate modulators is most easily produced by short-wavelength illumination of blue to green color [36, 81, 82], thus we need to consider primarily the short-wavelength end of the spectrum. But, as the first step, let’s calculate how much Eve’s power at 1548.51 nm might reach Alice’s modulators. The loss in Alice before the PM1 and IM

is

$$\alpha_{\text{Apm1}} = \alpha_{\text{BS}} + \alpha_{\text{DWDM1}} + \alpha_{\text{VOA1}} + \alpha_{\text{Att}} + \alpha_{\text{DWDM2}} + \alpha_{\text{Iso1rev}} + \alpha_{\text{Iso2rev}} = 118.5 \text{ dB}, \quad (7)$$

$$\alpha_{\text{Aim}} = \alpha_{\text{PM1}} + \alpha_{\text{Apm1}} = 121 \text{ dB}. \quad (8)$$

Assuming the laser power entering Alice is 100 W, the power reaching the PM1 and IM is about 141 and 79 pW. Owing to the low efficiency of the photorefractive effect at the long wavelengths [36], the existing isolation in the system will prevent this attack at the operating wavelength. However, we stress that this attack should be characterised in the ultra-wide spectral range.

Risk evaluation: **M** (0 vulnerability is not likely to exist, 1 is exploitable with today’s technology, 1 potentially gives Eve high key information).

Further suggestions: Test IM and PM1 for sensitivity to induced photorefractive at short wavelengths, similarly to [36, 81, 82]. This will establish the isolation required. It is also necessary to characterise Alice’s optical components in the wide spectral range (Appendix E), to determine the minimum total loss α_{Apm1} and α_{Aim} over the entire wavelength range accessible to Eve.

I. Laser damage

High-power laser radiation may cause temporary or permanent changes of properties of both absorbing media (for example, via heating and vaporisation) and transparent media (for example, via nonlinear effects [84]). This potentially affects many optical and optoelectronic components. Laser-damage attacks have been demonstrated on various QKD systems by targeting optical attenuators [85, 86], isolators [87, 88], a photodiode [49], and an avalanche single-photon detector [48]. Let’s consider the laser-damage attack on the QRate system’s Alice and Bob.

1. In earlier QKD systems, the optical attenuator was the last component in Alice before the quantum channel. However, its attenuation might be significantly decreased during the laser-damage attack, what leads to an increased Alice’s output mean photon number and thus leakage of the secret key [85, 86]. To mitigate this known risk, QRate’s system has two optical isolators in series at its output (Fig. 1). According to our more recent experimental results [87, 88], placing an additional sacrificial fiber-optic isolator or circulator at Alice’s exit might be required to complete the countermeasure against the laser-damage attack, at least by a 1550-nm continuous-wave laser.

We have tested three models of fiber-optic circulators and four models of fiber-optic isolators, including the isolator previously used in the QRate QKD system (QRate has recently replaced the exit isolator Iso2 with another model Opneti D-P-1550-900-1-0.3-FC-5.5x35 that we have not tested) [87, 88].

The samples tested exhibit a temporary reduction of isolation by about 15–35 dB achieved at a certain cw laser power specific to each sample. In the current system configuration with two isolators, this reduction of isolation may open loopholes for the Trojan-horse attack, laser-seeding attack, and power-meter-injection attack (Secs. V E to V G). However, attempts to reduce the isolation further under a higher illumination power result in the sample’s catastrophic failure. The latter manifests in an extremely large insertion loss and isolation, safely and permanently interrupting key generation.

Almost all the samples tested had a residual isolation (before the catastrophic failure) of more than 17 dB. This is sufficient to protect the next isolator behind it and the remaining system components from the laser damage, because the residual power reaching them never exceeds their specified maximum operating power. The isolator previously used by QRate (Thorlabs IO-G-1550APC; ISO PM 2 in [87]) exhibited maximum isolation reduction from 37 dB to about 17 dB residual value at 3.37 W laser power. Therefore, this isolator may itself be a good passive countermeasure, when an extra copy of it is added at the channel interface. We stress that the current system configuration with untested isolator models is already unsafe against the Trojan-horse attack because of insufficient isolation (Sec. V E) and might be further impaired by the laser-damage attack.

We have only tested the isolators under cw illumination at 1550 nm. However, damage mechanisms depend on illumination regime and wavelength. Continuous lasers and pulsed lasers with pulse duration longer than 1 ns typically cause damage via thermal effects; short and ultrashort laser pulses often strip electrons from the lattice structure of optical material before causing thermal damage [89]. The damage thresholds strongly depend on wavelength. It is thus important to test the front-end components against damage by a short-pulsed laser and lasers at different wavelengths [88].

Furthermore, the isolation properties of fiber-optic isolators often strongly depend on the wavelength. For instance, one model of isolator (not the one in QRate’s system) has the minimum of 11 dB isolation at 1150 nm (Fig. 5). Therefore, the laser-damage attack at this wavelength might bypass the isolators with enough power to affect the subsequent components. We discuss this problem further in Appendix E.

2. Bob’s setup is not protected against the laser damage. Theoretically, each component might be affected by the high-power laser. Let’s consider them one by one.

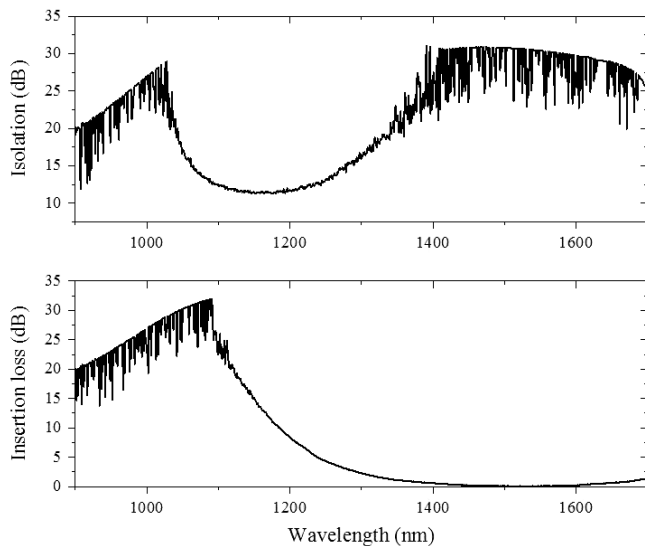


FIG. 5. Typical isolation and insertion loss of a fiber-optic isolator (FOCI M-II-2-15-S-C-C-E-1-FC/FC; not the one in QRate’s system) over a wider wavelength range. While the isolation is ~ 60 dB at the operating wavelength of 1550 nm, it drops to 11 dB at 1150 nm.

SD: the laser-damage attack might reduce sensitivity of a photodiode [49]. This does not compromise the security of QKD. However, SD shall not be employed as a countermeasure against the blinding attack.

DWDM: even if it is damaged by a high-power laser, this does not impact the QKD security. However, it shall not be employed as a security component against wavelength-dependent attacks.

PC and PM: their damage is unhelpful for Eve. We do not see how changing a detection basis setting could be exploited.

PBS: although changing the polarising beamsplitter’s properties may assist the detector-efficiency-mismatch attacks (Sec. V C), we consider this unlikely.

SDPs: Bob’s detectors are unprotected against the high-power illumination. It is known that a single brief application of high power can cause a permanent switching of the single-photon detector into a linear regime, which is equivalent to its blinding [48]. However, photocurrent monitoring should be an effective countermeasure against this.

Risk evaluation: **M** (in Alice: 1 vulnerability exists, 0 requires research and possibly future technology to exploit similarly to the Trojan-horse attack in Sec. V E, 1 potentially gives Eve higher key information than the Trojan-horse attack without laser damage; in Bob: 0 vulnerability is not likely to exist, 1 is exploitable with today’s technology, 1 may give Eve high key information).

Further suggestions: We recommend QRate to add the already tested isolator Thorlabs IO-G-1550APC as an additional sacrificial component at the exit of Alice, i.e., between the channel and the rest of Alice’s setup. This component’s only function is protecting the rest of the setup from damage, thus its own isolation should not be included into the isolation estimation of the source [87]. This isolator model should be further tested in a large range of laser powers and wavelengths, in continuous-wave and pulsed illumination regimes. We plan to test it under a 1064 nm, sub-nanosecond pulsed laser [88]. Also, all the components in Alice including this isolator should be characterised in a wide spectral range (Appendix E).

J. APD backflash

It has been shown that avalanching APDs emit photons that are coupled back to the quantum channel [90–92]. This emission has a broad spectrum. Although the state of each photon might not be correlated to the photon that caused the detection, these backflash photons pass through optical components and carry information about the originating detector, thus leaking information about the key to Eve. In the QRate system, Bob’s setup contains the PBS that splits the incoming photons into the two detectors. We assume this PBS encodes different polarisations into the backflash photons from different SPDs, allowing Eve to distinguish them and learn which detector has clicked, thus learning this bit of the raw key. The information leakage due to backflash is proportional to the probability of such events; the latter can be up to 10% [90]. A modified formula for the key rate should be used in the system, reducing the key generation rate, considering the worst-case assumption that Eve can distinguish all the backflash photons and map them to the raw key of Alice and Bob. The secret key rate bound is estimated in [91, 92] but only for a perfect single-photon source. A modified decoy-state security analysis for a realistic photon source that takes into account several side channels, including the APD backflash, is attempted in [25].

Given our strategy of reducing the information leakage under each individual imperfection to a negligible value (Sec. III), the emission probability from our SPDs needs to be measured and its transmission through Bob’s scheme into the quantum channel calculated. The latter depends on the spectrum of the backflash and spectral properties of Bob’s components, such as DWDM3. If necessary, additional spectral filters and isolators can then be added to Bob’s scheme, to reduce the emission probability into the channel to a specified level.

While the spectral characterisation of Bob’s passive components is straightforward (Appendix E), the spectral measurement of the broadband backflash emission of the SPDs is more challenging, owing to the single-photon sensitivity required [90, 91, 93, 94]. The availability of single-photon detectors and their noise level

restrict the wavelength range and spectral resolution of this measurement. We will probably need to make a relatively broadband integral measurement [90, 91, 94] and a crude bandpass measurement around the DWDM3's working wavelength [90, 93], then make some reasonable assumptions about the SPD's true backflash spectrum to upper-bound the emission probability into the channel.

Risk evaluation: **M** (1 vulnerability is known to exist, 1 exploitable with today's technology, 0 gives Eve low key information).

Further suggestions: First, DWDM3 should be characterised in a wide spectral range in the reverse direction of light propagation; we may conservatively assume the other Bob's components to be transparent. Then, the probability of backflash emission from one of Bob's SPDs should be measured as outlined above. (Alternatively we may skip the latter measurement and assume the emission spectral density of our SPD to be equal to that of different devices measured in [90, 92].) Based on the results, additional spectral filters and isolators may need to be installed in Bob to reduce the backflash emission to a negligible level.

K. Intersymbol interference

The security proof of the decoy-state BB84 protocol incorporates the assumption that all the intensity states (signal, weak and vacuum decoy) and the signal states are prepared independently of each other. However, a realistic frequency response of the intensity modulator that has a finite bandwidth might break this assumption and introduce correlations in the shape of the adjacent pulses [95]. The intensity of the pulse deviates from the set value, depending on the state of the preceding pulse, exhibiting a so-called pattern effect. The same study describes an experimental approach needed to quantify the intersymbol interference and suggests the additional post-processing procedures that effectively restore non-correlated pulses assumption. In a later study [96], the correlation between adjacent pulses in both the intensity and phase modulators was quantified in a high-repetition rate polarisation-based QKD system.

At the same time, an intersymbol interference can occur even in the phase of the adjacent laser pulses. This effect becomes stronger as the QKD system's repetition rate increases. While, in principle, the decoy-state method relies on the assumption that the phase of the emitting weak coherent pulses is independently random in the range $[0, 2\pi)$, this might not be the case for most commercial QKD systems with gain-switching laser sources. The residual electromagnetic field after the emitted light pulse in the cavity of the laser can affect the next light pulse and their phases can be correlated [96, 97].

Risk evaluation: **L** (1 vulnerability likely exists, 0 requires research and possibly future technology to exploit, 0 probably gives Eve low key information).

Further suggestions: Our preliminary measurements on the QRate system have shown correlations between adjacent pulses in the electrical signals feeding the phase and intensity modulators [98]. This indirectly indicates the optical pulses have correlations as well. Optical measurements are planned to quantify the phase, intensity, and polarization correlations in the optical pulses. Once quantified, they can be incorporated in the security proof [22, 23, 99–101]. A possible challenge here is that the available proof gives a zero or low key rate even for small correlations. In case our measurements result in an unsatisfactory key rate, we may additionally consider a software-based countermeasure in the post-processing [95] or replacement of the existing electro-optical modulators and/or their electronics with ones that have a higher bandwidth.

L. Imperfect individual state preparation

Most security proofs imply that the quantum states are prepared perfectly in any parameters like amplitude, relative phase, etc., which is generally not the case in reality. Regardless of the above-mentioned effects of intersymbol interference in phase and intensity of the emitted qubits, deviations of these parameters from the ideal can be considered for both average and individual qubits. Such deviations have been studied experimentally and theoretically for a loss-tolerant protocol [20, 21, 102, 103]. To our knowledge, the analysis is yet to be developed for BB84.

Risk evaluation: **L** (1 imperfection certainly exists, 0 research is required for exploitation, 0 probably gives Eve low key information).

Further suggestions: The optical measurements of intersymbol correlations that we plan in our lab will also yield data on imperfect state preparation, including average deviation and its statistical distribution. Based on that, we may attempt to apply the loss-tolerant protocol to incorporate these flaws [20, 21, 102, 103]. Further theory development is needed for a full understanding of this imperfection.

M. Calibration performed via channel Alice–Bob

The current system implementation conducts several calibration routines through the channel Alice–Bob. During the calibration, Bob sends low-level commands to Alice via the classical channel, Alice transmits signals via the quantum channel, Bob receives them using his SPDs and collects photon click statistics. This calibration sets multiple vital parameters such as signal timing and working points of the modulators. It is always performed at the system power-up and repeated as necessary whenever the system fails to generate keys for a relatively long time (about 1 h). The following parameters are determined by this calibration.

- Precise timing of Bob’s detector gate to maximise the count rate and correctly identify Alice’s bit number (i.e., not register clicks in an adjacent bit), separately for each of the two detectors. The scanning range is 6.4 ns, which spans two adjacent bit slots. The scanning is done with 100 ps step.
- Zero point of Alice’s IM, set by applying voltage at its bias section. While it is maintained by a feedback from PwM during QKD (Sec. V G), the initial setting is calibrated using Bob’s SPDs.
- Precise timing of electrical signal applied at Bob’s PM to correctly modulate the light pulse received from Alice. The scanning is done with 400 ps step.
- Precise timing of electrical signal applied at Alice’s PM to correctly modulate the laser pulse being sent. The scanning is done with 400 ps step.
- Precise timing of electrical signal applied at Alice’s IM to correctly modulate the laser pulse being sent. The scanning is done with 400 ps step.
- Initial setting of Bob’s PC to minimise QBER. (The PC is then adjusted in realtime during QKD to keep the QBER low.)

During QKD, three realtime adjustments are being performed continuously: Alice maintains her IM’s zero point (Sec. V G), Bob adjusts his PC to maintain low QBER, and Bob’s master clock generator is adjusted ten times per second.

Since the initial calibration is performed via the quantum channel, it is totally exposed to Eve’s tampering [104, 105]. We have to assume, and it is likely the case, that Eve can set any values of the parameters being calibrated at her discretion. Additionally she may interfere with Bob’s realtime clock adjustment and whatever timing parameters this affects. Additionally we have to assume Eve may issue low-level commands to Alice (unless this communication is strongly authenticated).

This has fairly horrible consequences. Several attacks become possible.

- Eve can induce a large time-efficiency mismatch between Bob’s detectors (Sec. V C), which has been demonstrated experimentally in other QKD systems [104, 105]. An additional possibility is that, since the system scans both click acceptance window positions over the time range that spans more than one bit slot, Eve may diverge them. I.e., she may set them such that a click resulting from a qubit detection at SPD1 is registered as one key bit while a click from the same qubit at SPD2 is registered as another key bit. This generally makes any security proof for BB84 inapplicable, because they all implicitly assume this situation is impossible. Besides, we can think of a practical attack that combines this diverged click registration with the simultaneous deadtime (Sec. V D) and allows

Eve to suppress clicks in the detector that gets registered in the later bit slot and, possibly, exploit an asymmetric click discarding in one bit slot at the end of the deadtime.

- Eve can shift Alice’s PM signal in time such that it modulates the light pulse at the transition between the modulation levels. The quantum states prepared are then less separated in phase and a phase-remapping attack may become possible [106]. Additionally, any careful characterisation of state preparation imperfections we perform (Secs. V K and V L) becomes meaningless. In particular, intersymbol correlations may be amplified.

If Eve can arbitrarily control Alice’s and Bob’s PM phase shifts, an extreme attack becomes possible. Eve sets Alice to use an identical pair of phase shifts in both her bases, and she tricks Bob to do the same as well. Then she performs a quantum intercept-resend attack in this one basis setting, while Alice and Bob think they are using different choices of bases. This attack does not increase the QBER and gives Eve the complete key. It is important to note that Alice and Bob can detect the presence of Eve in the discarded cases when their bases don’t match since the measurement outcome for Bob will no longer be random. Nevertheless, there are no such verification procedures in the classical BB84 protocol and its version used in the QRate’s system. The potential existence of this extreme attack hints that we have a problem that has to be treated by a security proof even in a milder case when Eve has a limited control over the PM settings. The most complete security proof for the BB84 protocol with imperfect state preparation [23] cannot account for the most extreme case of this attack, where four quantum states in two bases in reality degenerate into two quantum states in one basis.

- Eve can shift Alice’s IM signal in time and/or make the IM operate with an incorrectly set zero point. This would have similar effects on the intensity states being prepared by Alice. The security proof of the decoy-state protocol [107] becomes inapplicable when the actual intensities are unknown and are chosen by Eve. Characterisation of state preparation imperfections becomes meaningless.

To summarise, the public exposure of the calibration routines presents multiple security issues. The analysis team did not know how to solve this and suspected that a significant redesign of the system hardware might be required. This is a decision that should be taken by the system manufacturer.

Risk evaluation: **H** (1 vulnerabilities are likely exploitable, 1 with current technology, 1 might give Eve high key information).

Further suggestions: An extensive discussion with QRate was needed. QRate has subsequently found so-

lutions acceptable for the manufacturing process, see Sec. VII. Regarding the four-state Bob countermeasure to the time mismatch vulnerability suggested in Sec. VC and presently implemented by QRate, we are still not sure it is sufficient, given that Eve might be able to set all the time parameters and diverge them between bits.

N. Quantum random number generator

We remind the manufacturer that, in order to be compliant with the security proof, a real quantum random number generator must provide all the state, basis, and intensity choices in Alice and Bob, as well as random bit-value assignment in the event of a double click and other random values needed in the protocol. A mathematical ‘random’ number generator (used currently) or randomness expansion are, strictly speaking, insufficient.

Risk evaluation: **L** (1 vulnerability is known to exist in principle, 0 requires significant research and possibly future technology to exploit, 0 probably gives Eve low key information).

Further suggestions: The company should implement a full-bandwidth quantum random number generator, without resorting to the randomness expansion, and integrate it into the system.

O. Compromised chain of supply

Like most manufacturers of cryptographic hardware, QRate buys the constituent parts of its products from a multitude of external suppliers. It is the nature of cryptography that many of these parts may subvert the security of the product if the part’s supplier, or a third party, modifies it in a malicious way before it is installed into the product. The modification may be a covert change of characteristics that enables an attack, a change of the part’s behaviour that does the same, or a hidden transmitter (either radio-frequency or optical) that communicates the secret information outside the device. The modification will, of course, be difficult to detect: it will not reveal itself in the standard factory assembly and testing procedures, neither will it hinder the normal operation of the product. In the QKD system, the parts that may be compromised include optical, electrooptical and electronic components, third-party electronic modules, and even integrated circuits.

This problem is general to all cryptography hardware. We can also think of attacks and information leakage tactics specific to QKD that might be enabled in such a way.

A significant drawback of these attacks from Eve’s point of view is the need to plan them well in advance. She must initiate them before the equipment is assembled and deployed for protection of the asset of interest.

Risk evaluation: **M** (0 it is unlikely that any player will spend significant resources on preventively attacking

a niche product that is not yet being deployed for protection of high-value information assets, 1 can certainly be done today, 1 can be arranged to leak the entire key).

Further suggestions: The company should learn suitable mitigation strategies from the national cryptography licensing authority.

VI. SUMMARY OF INITIAL SECURITY ANALYSIS

At the end of our initial security analysis concluded in January 2022, we identify more than ten potential implementation security issues in QRate 312.5 MHz QKD system and rank them by their practical risk (see Table I). The vulnerabilities in Bob’s single-photon detection subsystem related to detector controllability and timing calibration are of a high concern (Secs. VB and VC). We are not sure if it is possible to construct sufficient countermeasures and stop all detector-related attacks that are implementable with today’s technology. From this point of view, the measurement-device-independent and twin-field QKD schemes are an attractive alternative. The accessibility of the calibration routines for Eve’s tampering (Sec. VM) is another difficult problem that needs to be discussed.

Actions needed to address the remaining vulnerabilities are mostly clear. Most optical components in the scheme need to be spectrally characterised in a wide spectral range ($\sim 350\text{--}2400$ nm, see Appendix E). Optical measurements of imperfections in the state preparation in Alice and light emission from Bob’s APDs need to be performed. Several inexpensive additional passive components, such as isolators and spectral filters, should be added to the scheme. QRate should make minor improvements in the post-processing algorithms and update the key rate equation.

Finally, we ask QRate to provide the complete QKD system to our testing lab on a permanent basis. Further security analysis requires a level of familiarity with the system implementation that cannot be gained by reading technical documentation and can only be obtained via extensive hands-on experience during experiments. This sample of the system should be reserved for the hacking experiments and serve no other purposes.

VII. ADDRESSING HIGH-RISK SECURITY ISSUES

After the delivery of our initial analysis report, actions have taken place during the year 2022. The four high-risk security issues (marked **H** in Table I) have been prioritised and QRate has implemented countermeasures to all of them. Meanwhile, we hope that a formal certification methodology that is being designed covers all, or most of, the security issues identified by us. QRate has also provided us the QKD system for testing.

The photocurrent-monitoring countermeasure against detector blinding (Sec. VB) has been implemented by QRate and tested in our lab [53, 54]. It reliably protects against cw blinding. However, pulsed blinding and control remain possible, owing to the photocurrent measured being averaged over a relatively long time [53]. A higher-bandwidth photocurrent registration scheme has subsequently been implemented in the sinusoidally-gated detector for this QKD system, in order to close this issue. Its testing on an automated testbench [54] is in progress. Also, testing this detector for the after-gate and falling-edge attacks is in progress [108].

The four-state Bob has been implemented by QRate as a countermeasure against the detector efficiency mismatch (Sec. VC) and timing calibration vulnerability (Sec. VM).

The software component of the simultaneous deadtime has been implemented, to complete the countermeasure against the deadtime attack (Sec. VD).

In order to address the calibration vulnerabilities (Sec. VM), QRate has eliminated the calibration of Alice’s IM and PM via the channel. The intensity modulator is now instead always calibrated via Alice’s PwM. The phase modulator is now only calibrated at the factory once, then its settings remain fixed during the lifetime of the system. With these changes and the four-state Bob in place, we hope that the existing calibration of Bob’s PM and timing of his detectors via the channel no longer constitute a vulnerability and may remain unchanged.

The above modifications to the system and additional tests planned should close all the high-risk vulnerabilities from Table I. This protects the system from the attacks known to be readily implementable today.

VIII. PROPOSAL FOR CERTIFICATION

To perform a complete set of measurements and tests for certifying implementation security of the “quantum” part of the system (i.e., to cover all the potential issues identified in this report), five testbenches are needed.

1. Wideband spectral characterisation of components, as detailed in Appendix E.
2. Characterisation of detector controllability, deadtime, efficiency mismatch, and Bob’s calibration routines. This includes testing the efficiency of any countermeasures to these issues. The testbench design is sketched in Appendix D and [53].
3. Characterisation of state preparation imperfections in Alice. The testbench design can be based on [96, 109].
4. Characterisation of light emission from the detectors, as detailed in [90].

5. Laser damage, based on [87]. Although different lasers may be used by Eve, we propose to initially implement the basic testing under 1550 nm continuous-wave laser.

A formal certification methodology for QKD is currently under development, in coordination with the Russian national cryptography licensing authority. This report is one of the inputs to this process. Traditionally, Russian national certification standards for cryptographic systems are classified. The actual domestic certification procedures being implemented thus cannot be disclosed.

IX. CONCLUSION

We have performed security analysis of the commercial QKD prototype system from QRate. Since this system uses a fairly standard prepare-and-measure BB84 scheme, this analysis should be partially applicable to other systems of the same type. Out of several potential vulnerabilities identified (Table I), four are deemed high-risk (**H**), because attacks exploiting them are likely implementable today. These four security issues are addressed first. QRate has implemented countermeasures to each of them (Sec. VII). The remaining security issues might be addressed routinely in the course of the formal certification that is being developed (Sec. VIII). We hope that this work contributes to the establishment of a Russian domestic certification lab and national certification standard for implementation imperfections in QKD.

ACKNOWLEDGMENTS

This manuscript has been reviewed by QRate prior to its publication. We thank QRate’s engineering division, theory group, and management for discussions and support. The measurement in Fig. 3 was performed by Ilya Gerasimov and Nikita Rudavin. We also thank Marcos Curty, Anton Trushechkin, Aleksei Reutov, Daniil Menskoy, and Nikolay Borisov for comments and discussions. This work was funded by the Ministry of Education and Science of Russia (program NTI center for quantum communications) and Russian Science Foundation (grant 21-42-00040).

Author contributions: All authors except A.K.F., E.K., and A.T. analysed the system documentation and different vulnerabilities, reviewed the general approach and conclusions. A.K.F., E.K., and A.T. described the post-processing and security proofs used in the system under analysis. All authors contributed to writing the manuscript.

Appendix A: Post-processing in QRate's system

The concept of the QKD is that two legitimate users (Alice and Bob) generate “long” symmetric keys by using a classical and a quantum channels together with a “short” pre-shared key. The pre-shared key is used for authentication of a classical communication only and can be discarded, or even publicly announced, after the end of the first run (round) of the QKD protocol. In the next round, a piece of the previous quantum-generated key can be used for the authentication purposes. In this way, the QKD have to be considered as a *quantum key growing*.

The core of the QKD protocol is in preparing quantum states and encoding information on Alice's side, and measuring the states on Bob's side. In the BB84 protocol [110], Alice and Bob use four qubit states that form two orthogonal bases in two-dimensional Hilbert space, $Z : \{|0_Z\rangle, |1_Z\rangle\}$ and $X : \{|0_X\rangle, |1_X\rangle\}$, where 0 and 1 indicate a classical bit encoded by the corresponding basis vector. The basis vectors are related as

$$|0_X\rangle = \frac{|0_Z\rangle + |1_Z\rangle}{\sqrt{2}}, \quad |1_X\rangle = \frac{|0_Z\rangle - |1_Z\rangle}{\sqrt{2}}. \quad (\text{A1})$$

If the information is encoded into polarization of a single photon, then $|0\rangle_Z$ and $|1\rangle_Z$ can correspond to the horizontal and vertical polarizations. In this case, $|0\rangle_X$ and $|1\rangle_X$ represent two diagonal polarizations, rotated by 45° and 135° relative to the horizontal direction. This polarization encoding is used to illustrate the idea, but, in fact, there is no restriction on the method of information encoding. Formally, $|0\rangle_Z, |1\rangle_Z, |0\rangle_X,$ and $|1\rangle_X$ are just vectors in the Hilbert space, and one can use any encoding scheme that fulfills Eq. (A1). The equivalence of the polarization and phase encodings is explained in detail in [111].

Importantly, as can be seen from Eq. (A1), when measuring a qubit in a basis different from the preparation one, the result is a completely random value. This is the consequence of the well-known fact that two non-orthogonal quantum states cannot be perfectly distinguished. On the contrary, if the preparation and measurement bases coincide, the result perfectly correlates with the initial qubit state (in the absence of errors in the channel, measuring devices, etc.). In this way, if Eve does not know the preparation basis, due to the no-cloning theorem [112, 113] she has to employ imperfect copying techniques that induce errors on Bob's side.

In practice, however, true single-photon states are very difficult to generate, and weak coherent states with a phase randomization are used instead. In the case of the polarization encoding, the state preparation takes the form

$$\begin{aligned} |0_Z\rangle &\rightarrow \rho_H(\alpha) \otimes \rho_V(0), & |1_Z\rangle &\rightarrow \rho_H(0) \otimes \rho_V(\alpha), \\ |0_X\rangle &\rightarrow \rho_D(\alpha) \otimes \rho_A(0), & |1_X\rangle &\rightarrow \rho_D(0) \otimes \rho_A(\alpha), \end{aligned} \quad (\text{A2})$$

where $\rho_M(\beta)$ stands for a phase-randomized coherent

state in mode M with mean photon number β :

$$\rho_M(\beta) = \sum_{n=0}^{\infty} \frac{e^{-\beta} \beta^n}{n!} |n\rangle_M \langle n|, \quad (\text{A3})$$

$|n\rangle_M$ denotes an n -photon state in mode M , H and V (D and A) indicate horizontal and vertical (diagonal and antidiagonal) modes with corresponding annihilation operators satisfying

$$\hat{a}_D = \frac{1}{\sqrt{2}}(\hat{a}_H + \hat{a}_V), \quad \hat{a}_A = \frac{1}{\sqrt{2}}(\hat{a}_H - \hat{a}_V). \quad (\text{A4})$$

The chosen photon number α in Eq. (A2) is specified by the protocol. The projection of considered states on a single-photon subspace results in four states

$$\begin{aligned} |0_Z\rangle &\rightarrow |1\rangle_H |0\rangle_V, & |0_X\rangle &\rightarrow \frac{1}{\sqrt{2}}(|1\rangle_H |0\rangle_V + |0\rangle_H |1\rangle_V), \\ |1_Z\rangle &\rightarrow |0\rangle_H |1\rangle_V, & |1_X\rangle &\rightarrow \frac{1}{\sqrt{2}}(|1\rangle_H |0\rangle_V - |0\rangle_H |1\rangle_V) \end{aligned} \quad (\text{A5})$$

that are suitable for the BB84 protocol [cf. Eq. (A1)]. Unfortunately, multiphoton components of states (A2) are vulnerable to a photon-number-splitting (PNS) attack and can not be used for secure key generation. Therefore, an estimation of the number of detections on Bob's side that resulted from single-photons states generated on Alice's side is required.

For the past decades, the BB84 protocol has been theoretically studied in detail. The first security proofs [114–116] were made for ideal version of the protocol with perfect single-photon source, and then generalised for realistic photon source [107]. In order to eliminate the vulnerability against the PNS attack and increase the secure communication distance, the decoy-state technique was developed [117] and combined with the entanglement distillation approach from [107]. As a result, an improved secret key rate formula was obtained [118, 119]. The security against not only the PNS attack but all possible general attacks is usually considered by the community as “obvious”, and until recently the complete mathematical proof has not been available in the literature. The formal security proof is summarised and presented in [111].

In the QKD system under evaluation, the practical realisation of the decoy-state BB84 protocol contains the following steps.

1. *State preparation and measurement.* Alice randomly with equal probabilities chooses a basis from the set $\{Z, X\}$ and an information bit from $\{0, 1\}$.

In order to counteract the PNS attack, the widely applied decoy-state technique is used. Alice randomly chooses the laser pulse intensity [α in Eq. (A2)] from the set $\{\mu, \nu_1, \nu_2\}$ with corresponding probabilities $\{p_\mu, p_{\nu_1}, p_{\nu_2}\}$. Here μ corresponds to the signal-type state, ν_1 and ν_2 ($\nu_1 + \nu_2 < \mu$, $\nu_2 < \nu_1$) correspond to the weak and vacuum

decoy-type states respectively. The optimal intensities and probabilities are determined for a given communication distance and experimental setup from the numerical maximization of the simulated secret key rate.

Then, the photon pulse is prepared in the corresponding quantum state and is transmitted through the quantum channel. It is important that Alice's laser emits each pulse with a random phase, owing to it being internally seeded with spontaneous emission. Thus the photon number statistics of Alice's pulses is Poissonian [Eq. (A3)], as required in the decoy-state technique.

Bob randomly and independently of Alice chooses a measurement basis from $\{Z, X\}$ and measures the qubit state in the selected basis. In case of a double-click of Bob's detectors, he randomly chooses the bit value.

The above steps are repeated many times until a sufficient number of quantum states are detected. More specifically, Alice sends pulses in so-called "trains" of fixed size ($\sim 10^6$ pulses per train). Owing to the time synchronisation with Alice, Bob knows the train number and the position of each detected pulse in the train.

2. *Sifting.* When Bob accumulates enough statistics (~ 1900 clicks), he announces the train number and position of each registered pulse together with its measurement basis. Alice in turn compares her preparation basis with it and announces the positions with matching bases and their corresponding pulse types (signal, weak or vacuum decoy).

After that, Alice and Bob select the signal-type bits with matching bases and form two bit strings, called *sifted keys*. Ideally, they should be identical, but due to natural noise in the channel or adversary actions they do not match 100%. Moreover, Eve may have partial information about them.

3. *Statistics estimation.* For practical reasons, the sifted keys are assembled into post-processing blocks of equal fixed size. In order to minimise the effect of statistical fluctuations on the final secret key length and have a reasonable block generation time, it is chosen to be $\ell_{\text{block}} = 1.36 \times 10^6$ bits.

For each block, Alice counts the corresponding total numbers of transmitted (N_α) and detected (M_α) pulses of each intensity $\alpha \in \{\mu, \nu_1, \nu_2\}$ regardless of their preparation and measurement basis. Then Alice estimates a *gain* Q_α – the probability that a pulse of intensity α is detected by Bob,

$$\hat{Q}_\alpha = \frac{M_\alpha}{N_\alpha}, \quad \alpha \in \{\mu, \nu_1, \nu_2\}, \quad (\text{A6})$$

and sends all three sets $\{N_\alpha, \hat{Q}_\alpha\}$ to Bob. Here and below, Q_α denotes a true probability value of

binomial distribution $M_\alpha \sim \text{Bi}(N_\alpha, Q_\alpha)$, while \hat{Q}_α denotes its statistical estimate (i.e., a random variable). I.e., Alice computes this statistics before sifting, in order to maximise the statistical sample size.

4. *Information reconciliation.* Alice's key is considered to be a reference one, while Bob attempts to eliminate the discrepancies between the keys caused by errors. In order to correct them, low-density parity-check (LDPC) codes are commonly used. Since ℓ_{block} is too large for high-speed and efficient LDPC-based algorithms, the block is split into 50 subblocks of length $\ell_{\text{subblock}} = 27\,200$ bits and the error correction is performed on each subblock separately. If the correction of a subblock fails, it is discarded from the block by both sides. As a result, Alice and Bob obtain the corrected keys K_{cor}^A and K_{cor}^B of length $\ell_{\text{cor}} = n_{\text{cor}} \ell_{\text{subblock}} \leq \ell_{\text{block}}$, where n_{cor} is the number of corrected subblocks. For a more detailed description of the symmetric blind information reconciliation scheme used, see [29, 30].

For each successfully corrected subblock, Bob computes the signal QBER

$$E_\mu^{(i)} = \frac{\text{number of errors in } i^{\text{th}} \text{ subblock}}{\ell_{\text{subblock}}}. \quad (\text{A7})$$

5. *Verification and parameter estimation.* The identity of obtained K_{cor}^A and K_{cor}^B is checked using an ε -universal polynomial hash function `PolyHash`, computed according to a modified PolyP32 algorithm [31, 32]. First, Alice generates a random number $k \in \{0, 1, \dots, q-1\}$ where q is a prime number, chosen to be $q = 2^{50} - 27$. Then she computes the hash-tag of her key and sends it together with k to Bob to compare with his hash-tag. If $\text{PolyHash}(k, K_{\text{cor}}^A) = \text{PolyHash}(k, K_{\text{cor}}^B)$, the verification is considered successful and the protocol proceeds to the next step. Otherwise, Alice and Bob start comparing the hash-tags of every single subblock until all the corrupted subblocks are found and discarded. In this way, the legitimate users obtain identical *verified keys* K_{ver}^A and K_{ver}^B of length $\ell_{\text{ver}} = n_{\text{ver}} \ell_{\text{subblock}} \leq \ell_{\text{cor}}$, where n_{ver} is the number of verified subblocks.

The probability of remaining errors in the verified keys can be estimated as

$$\varepsilon_{\text{ver}} \leq \varepsilon_{\text{col}}(\ell_{\text{cor}}) \quad (\text{A8})$$

if $\text{PolyHash}(k, K_{\text{cor}}^A) = \text{PolyHash}(k, K_{\text{cor}}^B)$ or

$$\varepsilon_{\text{ver}} \leq 1 - [1 - \varepsilon_{\text{col}}(\ell_{\text{subblock}})]^{n_{\text{ver}}} \quad (\text{A9})$$

otherwise. Here the probability of a hash collision, i.e., $\text{PolyHash}(k, K^A) = \text{PolyHash}(k, K^B)$ when $K^A \neq K^B$, is evaluated as [32]

$$\varepsilon_{\text{col}}(\ell_K) = \frac{\lceil \ell_K / \lfloor \log_2 q \rfloor \rceil - 1}{q}. \quad (\text{A10})$$

At the end of this step Bob computes the overall average QBER

$$E_\mu = \frac{1}{n_{\text{ver}}} \sum_{i \in \mathcal{V}} E_\mu^{(i)}, \quad (\text{A11})$$

where the summation is performed over the ensemble of successfully corrected and verified sub-blocks \mathcal{V} .

6. *Estimation of the level of eavesdropping.* After the successful error correction and verification, Bob estimates the final secret key length [33]

$$\ell_{\text{sec}} = m_1^l [1 - h_2(E_1^u)] - \text{leak} - \log_2 \varepsilon_{\text{pa}}^{-5}, \quad (\text{A12})$$

where the first and last terms represent the privacy amplification step and are determined by m_1^l —the lower bound on the number of bits in the verified key, obtained from signal single-photon pulses, E_1^u —the upper bound on the single-photon QBER, and $\varepsilon_{\text{pa}} = 10^{-12}$ —the tolerable failure probability for the privacy amplification step. The h_2 -function is the standard Shannon binary entropy. The second term in Eq. (A12) is the amount of information about the key leaked to Eve during the error correction and verification steps

$$\text{leak} = \sum_{i \in \mathcal{V}} [\ell_{\text{synd}} - p + d_i] + \xi \ell_{\text{hash}}, \quad (\text{A13})$$

where ℓ_{synd} is the syndrome length, p is the initial number of punctured bits, d_i is the total number of disclosed punctured bits in additional rounds (ℓ_{synd} , p and d_i depend on the LDPC code rate and *a priori* QBER estimation, see [29, 30]), $\ell_{\text{hash}} = \lceil \log_2 q \rceil = 50$ is the hash-tag length, $\xi = 1$ if $\text{PolyHash}(k, K_{\text{cor}}^A) = \text{PolyHash}(k, K_{\text{cor}}^B)$ and $\xi = n_{\text{cor}} + 1$ otherwise.

Using the decoy-state technique, the lower bound on the single-photon gain Q_1 is estimated as [33, 120]

$$Q_1^l = \frac{\mu^2 e^{-\mu}}{(\nu_1 - \nu_2)(\mu - \nu_1 - \nu_2)} \left[Q_{\nu_1}^l e^{\nu_1} - Q_{\nu_2}^u e^{\nu_2} - \frac{\nu_1^2 - \nu_2^2}{\mu^2} (Q_\mu^u e^\mu - Y_0^l) \right], \quad (\text{A14})$$

$$Y_0^l = \max \left\{ \frac{\nu_1 Q_{\nu_2}^l e^{\nu_2} - \nu_2 Q_{\nu_1}^u e^{\nu_1}}{\nu_1 - \nu_2}, 0 \right\}. \quad (\text{A15})$$

The finite key effect and statistical fluctuations are taken into account in our analysis. According to the central limit theorem, the binomial distributions of $M_\alpha \sim \text{Bi}(N_\alpha, Q_\alpha)$ and $m_1 \sim \text{Bi}(\ell_{\text{ver}}, Q_1/Q_\mu)$ can be well approximated by the normal distribution. The upper and lower bounds on Q_α and m_1 are [33]

$$Q_\alpha^{u,l} = \hat{Q}_\alpha \pm z \sqrt{\frac{\hat{Q}_\alpha(1 - \hat{Q}_\alpha)}{N_\alpha}}, \quad (\text{A16})$$

$$m_1^l = \ell_{\text{ver}} \frac{Q_1^l}{Q_\mu^u} - z \sqrt{\ell_{\text{ver}} \frac{Q_1^l}{Q_\mu^u} \left(1 - \frac{Q_1^l}{Q_\mu^u} \right)}, \quad (\text{A17})$$

where z is the normal distribution quantile, and the bounds on the true value of Q_α are evaluated via the Wald confidence interval.

In general, one cannot use binomial distribution for QBER since if Eve performs a coherent attack, the errors in different positions in the key cannot be treated as independent events. Therefore, E_1 is estimated in a different way [33]:

$$E_1^u = \frac{\ell_{\text{ver}} E_\mu - \bar{m}_0^l}{m_1^l}, \quad (\text{A18})$$

where the lower bound on the number of bit errors in the verified key, obtained from the 0-photon pulses due to background events [$\bar{m}_0 \sim \text{Bi}(N_\mu, e^{-\mu} Y_0/4)$], is given by

$$\bar{m}_0^l = N_\mu \frac{e^{-\mu} Y_0^l}{4} - z \sqrt{N_\mu \frac{e^{-\mu} Y_0^l}{4} \left(1 - \frac{e^{-\mu} Y_0^l}{4} \right)}. \quad (\text{A19})$$

One can see that 7 confidence bounds in total are required to compute ℓ_{sec} . Therefore, in order to have the estimation Eq. (A12) satisfied with probability not less than $1 - \varepsilon_{\text{decoy}}$, one has to define the quantile as

$$z = \Phi^{-1} \left(1 - \frac{\varepsilon_{\text{decoy}}}{7} \right). \quad (\text{A20})$$

If $\ell_{\text{sec}} \leq 0$, Eve is assumed to have more information about Alice's string than Bob. Hence, the key block is considered insecure and is discarded by both sides. The protocol is aborted, and Alice and Bob proceed to the next accumulated sifted block.

7. *Privacy amplification.* If $\ell_{\text{sec}} > 0$, Alice and Bob proceed to the privacy amplification step, aimed to shorten the verified key K_{ver} even further and destroy Eve's potential knowledge about the key. This procedure is performed using a hash function from the Toeplitz family of 2-universal hash functions [34, 35]. Bob generates a random string S of length $\ell_S = \ell_{\text{ver}} + \ell_{\text{sec}} - 1$ and sends it to Alice [121]. Alice computes $\ell_{\text{sec}} = \ell_S - \ell_{\text{ver}} + 1$. Then both sides symmetrically generate a Toeplitz matrix T_S of dimension $\ell_{\text{sec}} \times \ell_{\text{ver}}$ using S and compute the final key $K_{\text{sec}} = T_S K_{\text{ver}}$. As a result, Alice and Bob obtain a common shorter *secret key* of length ℓ_{sec} , Eve's information about which is now negligible. The security of privacy amplification is based on the leftover hash lemma [122].

One can notice that all the post-processing steps require Alice and Bob to communicate via the classical

channel. In order to verify both data integrity and authenticity of each message, a hash-based message authentication code and a secret key taken from the common quantum key, is used. The message authentication code uses a Russian national standard hash function Streebog-512 (GOST R 34.11-2012) [123]. The authentication failure probability (i.e., the probability that Eve will guess the secret key from the hash-tag of initial message) is considered to be much less than 10^{-12} and hence is neglected in Eq. (A21). There is also an option to supply the system with a certified hardware authentication device (“Continent” manufactured by the Russian company Security Code LLC) that replaces Streebog-512.

Theoretically the QKD security level is expressed in terms of the trace distance between the real classical-quantum state (in which the classical subsystem corresponds to the key, and the quantum one belongs to Eve) and the respective ideal state. The latter is characterised by a uniform distribution of the key and the absence of correlations between the key and Eve’s quantum subsystem. If the trace distance does not exceed ε , the key is called ε -secure (see, e.g., [124]). This overall (in)security parameter of the entire QKD system has several contributions

$$\begin{aligned} \varepsilon &= \varepsilon_{\text{decoy}} + \varepsilon_{\text{ver}} + \varepsilon_{\text{pa}} \\ &= 10^{-12} + 2.5 \times 10^{-11} + 10^{-12} < 3 \times 10^{-11}, \end{aligned} \quad (\text{A21})$$

where $\varepsilon_{\text{decoy}}$ is a failure probability of the single-photon gain and QBER estimation, ε_{ver} and ε_{pa} are the failure probabilities of key verification and privacy amplification. If the authentication at QKD round $r = 2, 3, \dots$ is realized by using a part of a key generated at the $(r - 1)^{\text{th}}$ round, then the security parameter for the r^{th} round is given by $\varepsilon^{(r)} = r\varepsilon$.

Appendix B: Implementation layers in a quantum communication system

For convenience, we reprint the description of layers from [15] here:

Layer	Description
Q7. Installation and maintenance	Manual management procedures done by the manufacturer, network operator, and end users.
Q6. Application interface	Handles the communication between the quantum communication protocol and the (classical) application that has asked for the service. For example, for QKD this layer may transfer the generated key to an encryption device or key distribution network.

Q5. Post-processing	Handles the post-processing of the raw data. For QKD it involves preparation and storage of raw key data, sifting, error correction, privacy amplification, authentication, and the communication over a classical public channel involved in these steps.
Q4. Operation cycle	State machine that decides when to run subsystems in different regimes, at any given time, alternating between qubit transmission, calibration and other service procedures.
Q3. Driver and calibration algorithms	Firmware/software routines that control low-level operation of analog electronics and electro-optical devices in different regimes.
Q2. Analog electronics interface	Electronic signal processing and conditioning between firmware/software and electro-optical devices. This includes for example current-to-voltage conversion, signal amplification, mixing, frequency filtering, limiting, sampling, timing-to-digital and analog-to-digital conversions.
Q1. Optics	Generation, modulation, transmission and detection of optical signals, implemented with optical and electro-optical components. This includes both quantum states and service optical signals for synchronisation and calibration. For example, in a decoy-state BB84 QKD protocol this layer may include generation of weak coherent pulses with different polarisation and intensity, their transmission, polarisation splitting and detection.

Appendix C: Attacks exploiting superlinear detector control

Many commercially available gated SPDs exhibit superlinearity at the edge of the gate [45, 125, 126]. This is an unwanted SPD behavior that creates a loophole in QKD security. It may be exploited, for instance, in the following intercept-resend attack on the BB84 [110] family of protocols. Eve uses a random basis to measure quantum states sent by Alice and resends her measurement results as multiphoton pulses, which are split into four (with a passive basis choice) or two (with an active basis choice) detectors at Bob. If the basis and bit value of the detector coincides with Eve’s basis and bit value, it will absorb twice as many photons as each detector in the opposite basis to Eve’s. Due to the superlinearity of the SPD, the probability of detection for Bob in the basis matching Eve’s is higher than in the opposite basis. This contradicts the assumptions on Bob’s measurement in the BB84 security proof. If the superlinearity is strong

enough, the quantum bit error rate (QBER) under attack falls below 11%. However, a constraint of this regime is that Bob doesn't always detect Eve's multiphoton pulse, even in the basis matching Eve's. She can compensate for this efficiency loss by making her intercept setup more efficient and placing it close to Alice (thus excluding line loss), which may make her attack successful depending on the setup parameters [45].

A step for Eve to improve her control of Bob would be to make his detection probability unity. If Bob uses gated detectors, she can achieve this by sending her multiphoton pulse in between the gates [47]. This is a so-called "after-gate attack". However Eve's pulse, typically of hundreds of fJ energy, creates afterpulses in Bob's SPDs in the following gates. They contribute to QBER, together with Bob's normal dark count rate.

The next step for Eve would be to take Bob's detectors under a complete control, by eliminating his dark counts. She can completely blind them to single photons and make the dark count rate zero. This is usually achieved by illuminating Bob with a continuous-wave laser of power ranging from nW to W depending on the type of SPD [127–129]. The blinding is caused by either constant photocurrent through the avalanche photodiode [127], its raised temperature [128], or even its permanent damage from a brief one-time application of a high-power laser [48]. There are versions of this attack that use pulsed blinding illumination [128, 130, 131]. Eve causes Bob's blinded detectors to click controllably typically by adding a bright pulse with appropriate timing and energy ranging from hundreds aJ to dozens fJ [127–129, 131], similarly to the after-gate attack. For some SPDs, she can make clicks by introducing gaps in her blinding illumination [130, 132]. The blinding attack often allows a total detector control, with unity probability and no artefacts like afterpulses or dark counts.

In summary, Eve has two passive ways to use superlinearity in Bob's detectors—find it at the edges of [45, 125, 126] or between the gates [47]—and four active ways to induce it—influence electronics by constant light (creating photocurrent) [127], cause heating by constant light [128], influence electronics by blinding pulses [128, 131], and change the properties of the SPD by laser damage [48]. This gives her several ways to attack the SPD and makes it tricky to develop reliable countermeasures [133]. This is arguably the most difficult vulnerability in today's QKD.

Let's consider if the detector control attacks can be revealed by statistical means, e.g., by analysing attack's signature and any possible artefacts in QBER, dark count rate, key rate, and other parameters. To begin discussing this we define two energy levels, E_{never} and E_{always} [127]. The former is the highest pulse energy that the SPD does not respond to with a click and the latter is the lowest energy that always causes a click. I.e.,

$$\begin{aligned} P(E_{\text{never}}) &= 0, \\ P(E_{\text{always}}) &= 1, \end{aligned} \tag{C1}$$

where $P(E)$ is the probability of the SPD to respond to the light pulse with energy E . If E_{never} is much higher than the single-photon energy (which means SPD works as a classical power meter), Eve can send the pulse with energy $4E_{\text{never}}$ for passive basis choice or $2E_{\text{never}}$ for active basis choice. In the former case energy $2E_{\text{never}}$ would always impinge on the detector decoding her state and energy E_{never} would impinge on each of the two detectors in the opposite basis. In the latter case either energy $2E_{\text{never}}$ would impinge on the appropriate detector (if Bob's and Eve's bases match) or this energy would be split equally with E_{never} impinging on both detectors (if the bases don't match). While this control method doesn't introduce any QBER, $P(2E_{\text{never}})$ can be much less than one, reducing the key generation rate. As discussed above, Eve needs to consider this constraint carefully [45]. The situation becomes easier for Eve when $2E_{\text{never}} \geq E_{\text{always}}$, thus $P(2E_{\text{never}}) = P(E_{\text{always}}) = 1$. Under such condition Eve can always get her resent state detected by Bob in case of the passive basis choice or half the time in case of the active basis choice [127]. This is generally enough to maintain the same key rate as before the attack.

Countermeasures: Most SPDs suffer from superlinearity. In the ten years following the discovery of this vulnerability, many countermeasures have been proposed. Let's group and review them.

1. "Too good to be true". Many detector-control attacks ironically improve the system performance: they decrease the QBER, decrease the dark count rate and increase the detection rate. However, monitoring for improved performance cannot be a reliable countermeasure, because Eve can always throttle the rate and intentionally introduce random errors in order to simulate the normal system performance [16].
2. "Change the paradigm". Measurement-device-independent (MDI) and twin-field QKD protocols [55–57] exclude the detectors from the secure environment. This solution totally removes all the detector vulnerabilities. Unfortunately implementing one of these protocols in a commercial system requires a complete redesign of the system and makes it more expensive and slow. So far, only laboratory demonstrations and prototypes have been made (notably one by Toshiba [134]), but no commercial product.
3. "Observe the observer". When SPDs are pushed into the superlinear regime, they manifest some artefacts unusual for normal workflow. The countermeasure can be watching the parameters of detectors. For the blinding attack, it could be measuring photocurrent [50–52, 135] (for further reading, see [136]; for probabilistic blinding attack model and security proofs of the photocurrent measurement, see [137]). The after-gate attack can be caught by

exact time measurement and observing afterpulse effects [138, 139]. The thermal blinding can be observed by temperature measurement. Such countermeasures are usually very effective against the specific type of attack but close one loophole only and make the system more complex and expensive. However, they cannot eliminate attacks that cause small changes of physical parameters. For example, the attack at the falling edge of the gate uses a small amount of energy and small time delay [126].

4. “Add a watchdog”. Adding a beamsplitter and a separate monitoring detector at the entrance of the receiver allows in principle to monitor for bright blinding light [127, 140]. However a hack-proof construction of this detector is a separate challenge and it may miss attacks that use a small amount of energy. Practical implementations of such monitoring detector have not been reported in the literature.
5. “Check double clicks”. The basic detector-control attack produces too frequent double clicks in pairs of SPDs, which can be the basis of a countermeasure [141]. This countermeasure needs further experiments to check if an improved attack that circumvents it can be constructed.
6. “Test the detectors”. Placing a calibrated light source inside the receiver and activating it at random times allows to test the detector response during a QKD session, e.g., check that it is not blinded [142]. When this countermeasure is integrated into a security proof, this imposes tight conditions on the equipment [143]. It is not clear if these conditions are sufficiently practical to implement.
7. “Detector decoy”. Ideas similar to the well-known decoy-state protocol [119] but implemented by varying the detector sensitivity between two levels were suggested as a countermeasure against the detector-control attacks. Distinguishing between weak and strong avalanches in a self-differencing detector allows to detect its blinding [144]. Another implementation places variable attenuators in front of each detector that randomly introduce 3 dB loss [145]. QBER and qubit rate for both 0 and 3 dB loss settings are measured. Without the attack, QBER is expected to be below 11% at both loss settings and the rate with 3 dB is expected to be half that with 0 dB. This countermeasure is promising but needs further experiments to check its security.
8. “Shake the box”. To catch unexpected superlinear regime Bob can decrease sensitivity or even turn off his SPDs for some time. So he wouldn’t expect any qubits from Alice would be measured. Whatever is measured is either noise or Eve’s attack. Bob can randomly turn off his gate (to catch the blinding or after-gate attacks [146]) or shift the gate time (to

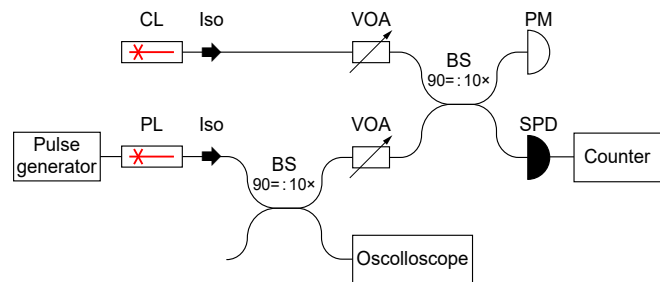


FIG. 6. Setup for testing detector control by bright light. CL, cw laser (1552 nm, 40 mW, Thorlabs SFL1550P); PL, pulsed laser (1552 nm, Gooch & Housego AA1406); Iso, optical isolator; VOA, variable optical attenuator; BS, fiber beamsplitter; PM, power meter (Thorlabs PM400 with S155C head); SPD, single-photon detector under test. The pulse generator (Highland Technology P400) drives PL directly and can induce relaxation-limited short laser pulses. The counter (Stanford Research Systems SR620) was typically accumulating clicks over 100 s for each data point. The oscilloscope (LeCroy 816Zi with OE555 optical-to-electrical converter) was used to observe the laser pulse shape.

catch the after-gate attack and even attack at the falling edge [147]). This can be effective against the basic attacks [45, 127, 129] but can be hacked by some modification of the basic attack [148]. Note that this countermeasure requires individual control over each detector gate, which may be impossible to implement in sinusoidally-gated and self-differencing detector schemes.

9. “Kill the superlinearity”. Eliminating the superlinearity would achieve perfect security against detector control attacks. Optical limiters may be investigated for this purpose [149–152]. Unfortunately, they start nonlinear behavior at power much higher (from dozens milliwatt) than used for blinding attacks (microwatts to milliwatts) and need a sufficient time to react (milliseconds) [152].

Appendix D: Test of QRate’s detector for bright-light control

Detector control by bright light was first proposed in [130] and later demonstrated in a number of SPDs, making systems that use them insecure [16, 127–129, 131, 132, 140, 153]. We have subjected the SPD from QRate’s system (detector serial number 20PD010013G “Gleb”) to the same test, using a standard experimental setup shown in Fig. 6 [127]. The scheme allows application of cw and pulsed light of controllable power to the SPD. We observe the SPD becoming blind (i.e., stopping producing output pulses) at cw power of 3 μ W.

We then add bright trigger pulses that should produce a controlled click response. This SPD works in a sinusoidally-gated regime at 312.5 MHz. In this test, we

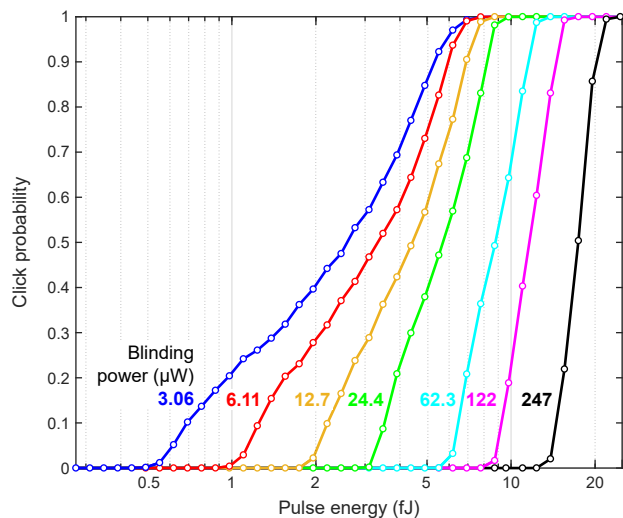


FIG. 7. Detector control characteristics in the asynchronous regime at different cw blinding power levels. The trigger pulse FWHM was 0.4 ns.

apply our trigger pulses at 100 kHz and do not synchronise them to the detector gates. They thus impinge on the SPD at random times relative to the detector gate. One would expect the sensitivity of the blinded SPD to the trigger pulses to vary through the detector gate [126, 148]. Thus our asynchronous regime represents a worst-case condition for Eve. The measured control characteristics are shown in Fig. 7. While the pulse response at the minimum blinding power of 3 μW is somewhat unstable, from 6 μW on we observe a transition from 0 to exactly 100% click probability. A close-to-perfect detector control, manifested in the click probability rising from 0 to $> 99\%$ at 3 dB increase of the trigger pulse energy [127], is achieved in our SPD at blinding power $\geq 62 \mu\text{W}$.

This SPD is well controllable even in the asynchronous regime. The QKD system is thus certainly vulnerable and needs a countermeasure. A further treatment of this problem is given in [53, 54].

Appendix E: Wide spectral testing

Most of the known attacks and countermeasures for them traditionally considered Eve’s access in Alice’s and Bob’s setups at around the QKD system operating wavelength (~ 1550 nm). However, the transmission channel has much wider bandwidth (for quartz fiber it is ~ 350 – 2400 nm [154]) and gives Eve a potential to vary her light wavelength at which the attacks can be made. While the countermeasures implemented to protect from the attacks often work well at the QKD system operating wavelength, they may be completely unsafe in case of attacks in another spectral region [74].

As example, a standard approach to protect QKD systems from several attacks is their optical isolation with

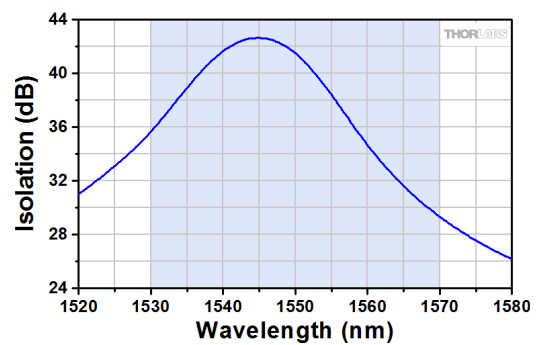


FIG. 8. Isolation of Thorlabs IO-H-1550 fiber-optic isolator, from its specification sheet [155].

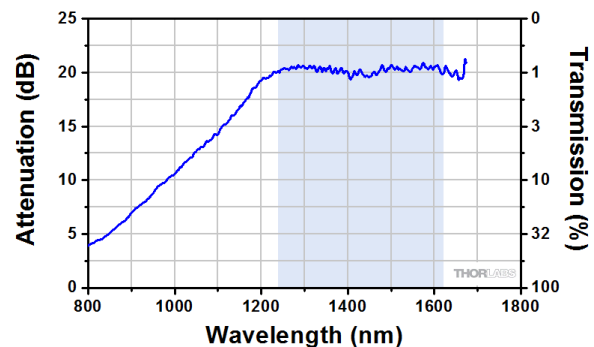


FIG. 9. Attenuation of Thorlabs FA20T fiber-optic attenuator, from its specification sheet [156].

attenuators and optical isolators. Spectral characteristics of isolators (Iso1, Opneti IS-S-P-1550-900-1-0.3-FC-5.5x35; Iso2, Opneti IS-D-P-1550-900-1-0.3-FC-5.5x35) and attenuator (Att, Opneti FOA-P-1-20-FC), used for this purpose in QKD QRate system, are not provided by the manufacturer. But to illustrate the broadband attack principle let’s consider similar devices from Thorlabs. Spectral attenuation of isolators (Thorlabs IO-H-1550) and attenuators (Thorlabs FA20T) are shown in Figs. 8 and 9 (from specification sheets). For IO-H-1550 isolation at the operating wavelength is about 43 dB, but as the wavelength shift only to 1580 nm it becomes noticeably less, about 26 dB. In the case of FA20T, its attenuation when shifting from the operating wavelength to 800 nm decreases from 20 dB to 4 dB. Such a significant reduction of isolation may make the countermeasure ineffective. Note that the manufacturer’s data shows their spectral properties only in a narrow range (800–1700 nm and 1520–1580 nm, respectively) while the quantum channel bandwidth is much wider (~ 350 – 2400 nm). It is possible that in the rest of the spectrum their attenuation is even less.

In general, every QKD system has this problem. Most attacks can be optimised by Eve via varying the attack wavelength. Thus, the transmission and response to light of all the optical components involved in a particular attack must be characterised in the wide spectral range.

Then, an optimum wavelength should be found at which each attack is the most efficient. The information leakage is quantified at this wavelength. To ease requirements on the dynamic range of the characterisation testbench, the components in the attack path are characterised individually, then their measured transmission characteristics are multiplied together.

We next list the attacks from this report that benefit from the wavelength-dependent component properties, and specify which components should be characterised. We then propose a testbench for the characterisation of fiber-optic components.

1. Wavelength-dependent attacks

Superlinear detector control (Sec. V B).

In the current implementation, the QRate QKD system is vulnerable to the detector control attacks at wavelengths over the entire sensitivity range of the SPD. In principle, SD (Fig. 1) might be used as a countermeasure to identify this attack's presence. But in this scheme SD is located after DWDM3 (Opneti DWDM-1-100-36-900-1-0.3-FC [157]) in the 1554.94 nm port (channel 28). DWDM non-adjacent channel isolation is > 35 dB [157]. This means that if Eve attacks outside the 1553.33–1556.55 nm range, the SD may not detect this. Thus, if a watchdog photodetector is used to monitor for the detector control attacks, it should be placed either in front of the DWDM3 or after it in the main signal path, via a beamsplitter not selective by wavelength. In order to exclude Eve's attempts to vary the attack wavelength, the watchdog photodetector's spectral sensitivity range should be wider than the SPDs have. Otherwise Eve can choose a wavelength outside the photodetector's range, and blind and control Bob's SPDs unnoticed.

Components whose insertion loss or splitting ratio should be spectrally characterised: DWDM3.

Components whose sensitivity should be spectrally characterised: watchdog photodetector, SPD.

Detector efficiency mismatch (Sec. V C).

Choosing the wavelength benefits Eve. The difference in the spectral and spatio-temporal properties of the beamsplitter (PBS) and photodetectors (SPD1, SPD2) mentioned above allows Eve to distinguish the photodetectors and activate them selectively, gaining the ability to steal the key. Eve can in addition select the attack wavelength at which the differences in the time and amplitude of the photodetector responses are maximized. To determine the optimal attack wavelength, one should measure the PBS splitting ratio over the wavelength, check each photodetector's spectral and time sensitivity separately, combine these measurements and determine the wavelength when the photodetectors' efficiency mismatch is maximized. However, if the four-state Bob is implemented as we suggest in Sec. V C, this characterisation is not necessary.

Components whose insertion loss or splitting ratio should be spectrally characterised: PBS.

Components whose sensitivity should be spectrally characterised: SPD1, SPD2.

Detector deadline attack (Sec. V D).

The deadline loophole should be closed algorithmically as we suggest in Sec. V D. For completeness we remark that if it is not closed, the wavelength-dependent properties that affect the efficiency mismatch attack would also help Eve to select which of the two SPDs enters the deadline.

Trojan-horse attack (Sec. V E).

As discussed in Sec. V E, if Eve uses $I_{\text{in}} = 2.5 \times 10^{12}$ photons at the operating wavelength ~ 1550 nm, the mean Trojan photon number $I_{\text{max}} \approx 1.5 \times 10^{-5}$ exiting Alice leads to significant information leakage, owing to 172 dB attenuation by Alice's components.

We noted in Sec. V E, that to determine the maximum level of vulnerability of the QKD system by a Trojan-horse attack, one should find out the minimum total level of losses introduced by the entire system throughout whole range of quantum channel transparency and calculate corresponding maximum value of leaked signals I_{max} . Unfortunately, the manufacturer does not specify the spectral characteristics for QRate system components. These should be measured separately. Here, just to illustrate how crucial the choice of Trojan-horse attack wavelength is, we consider the spectral data of the Thorlabs devices discussed at the beginning of this section (Appendix E), taking them as analogue to Iso2 and Att.

We roughly estimate Trojan-horse photon attenuation α_A (Eq. (4)) in spectral region where Thorlabs elements (Iso2 and Att) are more transparent and determine the corresponding value of I_{max} . As mentioned above, manufacturer shows components' spectral data (Figs. 8 and 9) only in a narrow band nearby operation wavelength, but even from these submitted short spectral range data it is obvious that, in principle, it is possible to select a spectral part in which attenuation becomes noticeably lower. From Figs. 8 and 9 we conservatively assume $\alpha_{\text{Iso2rev}} = 26$ dB, $\alpha_{\text{Att}} = 4$ dB. We assume that reverse loss of Iso1 decreases proportionally to that Iso2 and $\alpha_{\text{Iso1rev}} = 17$ dB. We guess that the attenuation of DWDM1 (Opneti DWDM-1-100-36-900-1-0.3-FC) and DWDM2 (Opneti DWDM-1-100-28-900-1-0.3-FC) for a non-adjacent channel is > 35 dB [157], but of course, losses outside the operating range certainly require experimental verification. We assume that the losses of VOA1, BS, PM1, and IM do not change significantly. From Eq. (4) with these data we obtain $\alpha_A \approx 243$ dB and $I_{\text{max}} \approx 1.25 \times 10^{-12}$. From [73] and Fig. 4 we can estimate that in case of Eve's Trojan-horse attack at DWDM non-adjacent channel wavelength information leakage is low ($I_{\text{max}} \ll 10^{-9}$). It should be emphasised once again that the key value of DWDM loss outside its design spectral range requires experimental verification.

Components whose insertion loss or splitting ratio

should be spectrally characterised: Iso2, Iso1, DWDM2, Att, VOA1, DWDM1, BS, PM1, IM.

Laser-seeding attack (Sec. V F).

As discussed in Sec. V F Eve might be able to inject light into Alice's L1 laser diode and modify its emission characteristics, e.g., phase, intensity, and wavelength. But at operation wavelength attenuation inside Alice till laser L1 $\alpha_{As} \approx 124$ dB, making this attack unsuccessful. We estimate Alice's entry path losses in more transparent spectral region. With the assumptions made in Sec. V F and components' attenuation values discussed in the previous item, from Eq. (5) we get $\alpha_{As} \approx 142$ dB. The power reaching L1 is $W_{L1} = 10^{-\alpha_{As}/10} W_{in} \approx 0.63$ pW. This shows that even in the optimal wavelength case the seeding attack is still impossible [78]. Furthermore, L1 sensitivity for seeding outside the close vicinity of its emission wavelength is, in principle, significantly lower. This means that choosing the best wavelength is unlikely to give any benefit for this type of attack even if W_{seed} reaches the values much higher than 0.63 pW. Note that it is only a rough estimation and spectral minimum loss value in Alice requires additional broadband testing.

Components whose insertion loss or splitting ratio should be spectrally characterised: Iso2, Iso1, DWDM2, Att, VOA1, DWDM1, BS, PM1, IM.

Components whose sensitivity should be spectrally characterised: L1.

Light injection into power meter (Sec. V G).

As explained in Sec. V G, attacking Alice's power meter PwM by injecting additional light into it, Eve tries to force Alice to disbalance her intensity modulator's zero point. This leads to a change in the intensities of the vacuum, decoy, and signal states, as well as their ratios. This would reduce the real secure key rate below that calculated by the system.

It was shown in Sec. V G that, if Eve attacks at the QKD operating wavelength, her additional power reaching PwM is less than 14 nW, which is negligible in comparison with the power it measures in the normal operation. But, Eve can try to attack in the Alice's components maximum transparency spectral region. With all the Sec. V G assumptions and applying the above-mentioned component attenuation values in transparency region, by Eq. (6) we get the Alice's losses up to PwM $\alpha_{Ap} \approx 117$ dB and seeding power reaching power meter $W_{PwM} = 10^{-\alpha_{Ap}/10} W_{in} \approx 0.2$ nW. This value is minor and even noticeably smaller that Eve can get with 1548.51 nm attack. Note that it is only a rough estimation and spectral minimum loss value in Alice requires additional broadband testing.

Components whose insertion loss or splitting ratio should be spectrally characterised: Iso2, Iso1, DWDM2, Att, VOA1, DWDM1.

Induced-photorefractive attack (Sec. V H).

If Eve injects light into the QKD device this can change the photorefractive properties of Alice's and Bob's active lithium niobate elements and allow Eve to perform an induced photorefractive attack as described in Sec. V H.

Obviously, the magnitude of the photorefractive effect, and hence the effectiveness of the attack depends on radiation intensity reaching the active elements. Selection of injected light wavelength, in principle, can make it possible to use the spectral region with maximum optical channel transparency and increase the level of radiation power at PM and IM. We estimate radiation power at Alice's active elements in the case when the attack wavelength does not coincide with the QKD operating wavelength. Elements loss values are the same as previously assumed in this section. Using Eqs. (7) and (8), estimated power reaching PM1 and IM are as follows: $\alpha_{Apm1} \approx 137.5$ dB, $W_{Apm1} \approx 1.8$ pW; $\alpha_{Aim} \approx 140$ dB, $W_{Aim} \approx 1$ pW. These power values turn out to be less than at the operating wavelength. This is due to the fact that two DWDMs introduce strong additional losses, which are essentially spectral filters that strictly pass only the operating wavelength. In this case, the photorefractive attack is completely ineffective.

It has already been discussed in this section, but needs to be emphasised again, that the DWDM insertion loss outside the working channel is conservatively assumed to be about 35 dB, but we guess that outside the range of all working DWDM channels the transparency of this element may be significantly higher. This may cause much more power to reach the phase and amplitude modulators and the photorefractive attack may become possible. To evaluate this correctly, broadband spectral characterisation of the DWDM is needed.

Components whose insertion loss or splitting ratio should be spectrally characterised: Iso2, Iso1, DWDM2, Att, VOA1, DWDM1, BS, PM1.

Laser-damage attack (Sec. VI).

As has been shown above in this Appendix and in Sec. VI, the Bob's side of QRate QKD is unprotected from all known types of receiver-side attacks (detector control, deadtime, mismatch). In this QKD implementation, there is no special isolation component at Bob's input to Eve's damage attack at the operating wavelength. In this way, Eve might try to reach inside Bob and implement the laser-damage attack on PBS, trying to change its polarisation splitting ratio. The DWDM3 installed at Bob's input limits Eve's ability to get inside Bob at other wavelengths to attack the PBS. But, if Eve applies a laser-damage attack to DWDM3, she might change its spectral properties and make DWDM3 transparent not only at a channel 36 wavelength, but also in other regions of the spectrum. After that, varying the attack wavelength Eve can choose the optimal one, at which changes in PBS under the laser light happen the most efficiently and as much as possible unbalance the PBS splitting ratio. Such PBS damaging may improve the detector deadtime and mismatch attacks.

Unlike Bob, Alice has protective components (Iso1, Iso2) that hamper Eve's attacks. Section VI recommends placing an additional isolator between Iso2 and the channel, as a countermeasure to the laser-damage attack. With such attack, under the action of laser radia-

tion (~ 1550 nm, ~ 3.4 W), the isolator either completely breaks and becomes practically opaque, or retains a residual isolation level of about 17 dB. This limits the power of the attacking radiation passing through it and makes it impossible to defeat Iso2 and the other components beyond it.

However, Eve may benefit from the choice of wavelength. Due to different mechanisms of action, the impact of powerful attacking radiation with different wavelengths and illumination regimes may lead to different spectral changes in the losses introduced by the attacked isolator. It may be possible to choose such parameters of the attacking radiation that in some regions of the spectrum it will be possible to bleach the isolator and significantly reduce the losses introduced by it, which make it ineffective as a countermeasure and weaken protection against the other types of attacks. Furthermore, during the laser-damage attack the properties of optical components might not change uniformly in the entire spectral range. In principle, Eve might choose such a laser-damage regime when properties at the operating wavelength do not change significantly (i.e., from the point of view of Alice and Bob, everything is okay), but change strongly in a different spectral range. This potentially gives Eve additional opportunities to implement “invisible” attacks. These properties of isolators under the laser-damage attack in different regimes require a separate detailed study.

Components whose insertion loss or splitting ratio should be spectrally characterised: DWDM3, Iso2, Iso1, and PBS—all after the laser damage-attack.

APD backflash (Sec. V J).

Bob’s components attenuate the spectrally broadband detector backflash. Their transmission in the backwards direction is needed for the calculation of emission probability into the channel, as explained in Sec. V J.

Components whose insertion loss or splitting ratio should be spectrally characterised: DWDM3 in the reverse direction and eventually any additional components for Bob’s setup.

2. Ultra-wideband spectral testbench

Since component manufacturers never provide the spectral characteristics in the full range we need (~ 350 – 2400 nm), these need to be measured.

At least two different implementations for measuring fiber-optic elements attenuation spectra are possible. One uses a spectrum analyser [Fig. 10(a)], another a monochromator before device-under-test (DUT) and photodetector [Fig. 10(b)]. In both schemes, the spectrum is first scanned without the DUT (to characterise the instrument response), then with the DUT in place. The two spectral curves are then subtracted from one another, to obtain the DUT transmission curve. The testbench using a monochromator has the advantage that DUT is not exposed to high power, eliminating the pos-

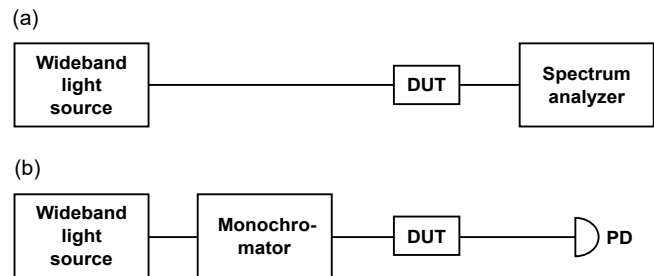


FIG. 10. Measuring fiber-optic component attenuation spectra (a) with a spectrum analyser and (b) with a monochromator and photodetector. DUT, device under test; PD, photodetector.

sibility of heating it and potentially changing its characteristics. But, since the spectral range of measurements is wide and the required sensitivity is at least 60 dB, as we need to characterize transmission of high-absorption fiber-optic components, in practice the setup with monochromator and photodetector is more difficult to implement instead of using a purchased spectrum analyser. The setup with spectrum analyser is also preferable due to its convenience of use and ergonomics. The difficulties encountered in the technical implementation of a setup using a monochromator are well illustrated in [76], where a single-photon detector is employed as PD. They report 10 nm spectral resolution, spectral range of 1100–1800 nm, and dynamic range of insertion loss measurement of about 70 dB (as visible in the plots in [76]). In comparison, our proposed testbench that uses an off-the-shelf spectrum analyser allows spectral resolution of 0.05 nm, spectral range of 350–2400 nm, and the dynamic range of measurements comparable to that demonstrated in [76]. We expect our testbench to be easier to align and operate. Next we will consider only measurements using the purchased spectrum analyser.

The key devices in the attenuation measurement spectrum setup are a spectrometer (spectrum analyser) and a light source.

In today’s optical instrumentation market there are spectrum analysers that completely cover the quantum channel transition wavelength range and have necessary sensitivity. Yokogawa here is the established leader. The advantage of these devices is high quality, user-friendly interface, a wide range of measured wavelengths, high sensitivity and dynamic range. The models line of Yokogawa spectrum analysers lacks a single device that completely covers the wavelength range necessary for our purposes. But a set of two devices does. These are the models AQ6374 (350–1700 nm) and AQ6375B (1200–2400 nm). Their spectral ranges overlap, allowing a more accurate “stitching” of data obtained in the two different wavelength ranges.

Broadband light sources can be fundamentally divided into two large groups, according to their physical principle: incandescent lamps and laser (supercontinuum) sources. A practical disadvantage of incandescent-lamp-

based sources is their low output power when coupled into a single-mode fiber. The supercontinuum (laser) sources are free from this shortcoming, their main disadvantage being a high cost. Supercontinuum sources that satisfy our requirements for power density and spectral range are commercially available.

The recognised world leader in the production of supercontinuum sources is NKT Photonics. This company offers a wide range of supercontinuum sources that differ in the range of generated wavelengths, average power and power spectral density in different parts of the spectrum. One of the optimal devices for our application is the SuperK Extreme/Fianium FIU-15 model. The range of light generation is from about 350 to 2500 nm. The integrated optical power is about 4.5 W, the spectral power density varies through the wavelength range and averages about 3–4 mW/nm. The output radiation power can be manually and programmatically adjusted. The source is characterised by high stability of the output optical power in the entire range (<0.5%). The polarisation of the output light is random.

Most of the fiber-optic components to be tested have standard single-mode fiber pigtailed (9.5/125 μm) with FC type connectors. The advantage of Yokogawa spectrum analysers is that they have an input for a single-mode fiber with this type of connector. For connecting the supercontinuum source to the FC connectorised single-mode fiber, an accessory optical coupler from NKT Photonics has to be used. To cover our entire range, we can use two of them: 350–1200 nm model (SuperK Connect FD7) and 1200–2400 nm model (SuperK Connect FD6).

One of the possible implementations of the testbench with these instruments is shown in Fig. 11. The setup consists of three parts:

- supercontinuum source SuperK Extreme/Fianium FIU-15;
- path for measurements in 350–1200 nm range: FD7 connector, DUT, AQ6374 spectrum analyser;
- path for measurements in 1200–2400 nm range: FD6 connector, DUT, AQ6375B spectrum analyser.

The device under test is measured in both paths and the results are combined to obtain its complete transmission spectrum.

To eliminate the possibility of influence on DUT by the source’s radiation, which can potentially lead to its characteristics changing, we estimated the power impinging on DUT during the measurements. The maximum FIU-15 power density after single mode fiber coupler FD6 is approximately 1 mW/nm. We hope that significant changes to spectral characteristics of fiber-optic components under such power density is unlikely. However, this should be experimentally verified for every type of DUT. This can be done by carrying out several successive measurements with different radiation power of FIU-15 light source. If DUT’s characteristics change under higher

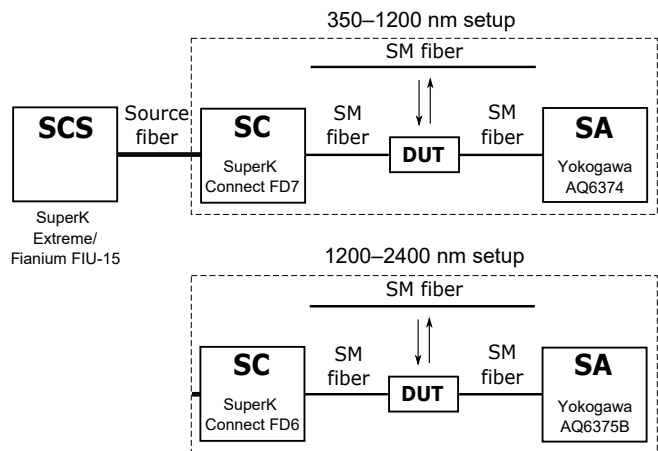


FIG. 11. Setup implementation for measuring fiber-optic component transmission spectra with NKT Photonics & Yokogawa devices. SCS, supercontinuum source; SC, single-mode fiber coupler; SA, spectrum analyser.

power, the results of these measurements will not match. Changing the integral output power of FIU-15 source is possible by varying its pulse repetition rate, which does not affect its spectrum.

We have assembled this testbench and are currently refining its usage methodology [158]. Meanwhile, the testbench with monochromator and single-photon detector is implemented by SFB Lab in Moscow [76].

As discussed in the previous section (Appendix E1), not only the fiber-optic components’ spectral characteristics need to be measured, but the photodetectors’ spectral sensitivity as well. These PD spectral characteristics are critical in the risk evaluation of several attacks: superlinear detector control, detector efficiency mismatch, and light injection into calibration photodetector.

The characterisation of photodetectors’ spectral sensitivity is not as demanding task as for high-absorbing fiber-optic components, for the following reasons. Since QKD systems use sensitive photodetectors, there is no need to characterise them using a bright light source with high power spectral density. The spectral sensitivity range of photodetectors is limited (about 900–1700 nm), which allows the use of non-ultra-wideband light sources. Thus, a simple to implement and relatively cheap setup can be used as a light source, shown in Fig. 10(b) (without the DUT), with the incandescent lamp (e.g., Thorlabs SLS201L/M stabilised fiber coupled light source) and a narrow-band monochromator with a fiber-optic output (e.g., Zolix Omni- λ 305i).

3. Consistency of broadband spectral properties

Even if all broadband spectral tests are performed properly for all the elements of a particular QKD system and its safety is fully proven, it is not possible to say with complete certainty, without additional assumptions, that

another sample of the same QKD system is safe without performing exactly the same extensive tests. This can be due to a sample-to-sample variation of the system elements. Most manufacturers guarantee parameters of the elements only in a very narrow spectral region, close to a specific wavelength. In one manufacturing batch, the spectral properties of elements from the same manufacturer may coincide. But another batch of these elements may be made with a slight change of the manufacturing technology, maintaining the declared properties in the narrow spectral range. Such changes in technology can however lead to uncontrollable spectral transparency loopholes outside of the narrow range. To prevent this, three approaches are possible.

- Perform the spectral testing of all elements of the system, several samples for each of them. Assume that the properties of all the other samples will be the same as in those tested. There is a risk that this is not the case.
- Spectrally test all elements for each particular QKD system. This is very expensive.
- Install a spectral filtering system at Alice's output and Bob's input of each QKD system that will block the entire spectrum except the communication wavelength. Only these filtering components will be subject to the spectral testing of their every sample.

-
- [1] C. H. Bennett, F. Bessette, L. Salvail, G. Brassard, and J. Smolin, "Experimental quantum cryptography," *J. Cryptology* **5**, 3–28 (1992).
- [2] Yu-Ao Chen, Qiang Zhang, Teng-Yun Chen, Wen-Qi Cai, Sheng-Kai Liao, Jun Zhang, Kai Chen, Juan Yin, Ji-Gang Ren, Zhu Chen, Sheng-Long Han, Qing Yu, Ken Liang, Fei Zhou, Xiao Yuan, Mei-Sheng Zhao, Tian-Yin Wang, Xiao Jiang, Liang Zhang, Wei-Yue Liu, Yang Li, Qi Shen, Yuan Cao, Chao-Yang Lu, Rong Shu, Jian-Yu Wang, Li Li, Nai-Le Liu, Feihu Xu, Xiang-Bin Wang, Cheng-Zhi Peng, and Jian-Wei Pan, "An integrated space-to-ground quantum communication network over 4,600 kilometres," *Nature* **589**, 214 (2021).
- [3] Francesco Vedovato, Paolo Villoresi, Giuseppe Vallone, Florian Kutschera, Victor Lopez, Vicente Martin, Jose Luis Rosales Bejarano, Rydlichowski Piotr, and Marc Geitz, "OPENQKD deliverable D8.3. Report on testbed replicability and performance," <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5da0754d3&appId=PPGMS>, visited 13 Feb 2023.
- [4] Miralem Mehic, Marcin Niemiec, Stefan Rass, Jiajun Ma, Momtchil Peev, Alejandro Aguado, Vicente Martin, Stefan Schauer, Andreas Poppe, Christoph Pacher, and Miroslav Voznak, "Quantum key distribution: a networking perspective," *ACM Comp. Surv.* **53**, 96 (2020).
- [5] Thomas Länger and Gaby Lenhart, "Standardization of quantum key distribution and the ETSI standardization initiative ISG-QKD," *New J. Phys.* **11**, 055051 (2009).
- [6] Romain Alléaume, Thomas E. Chapuran, Christopher J. Chunnillall, Ivo P. Degiovanni, Norbert Lütkenhaus, Vincente Martin, Alan Mink, Momtchil Peev, Marco Lucamarini, Martin Ward, and Andrew Shields, "Worldwide standardization activity for quantum key distribution," in *Proc. IEEE Globecom Workshop 2014* (IEEE Press, 2014) pp. 656–661.
- [7] Hoi-Kwong Lo, Marcos Curty, and Kiyoshi Tamaki, "Secure quantum key distribution," *Nat. Photonics* **8**, 595–604 (2014).
- [8] A. R. Dixon, J. F. Dynes, M. Lucamarini, B. Fröhlich, A. W. Sharpe, A. Plews, W. Tam, Z. L. Yuan, Y. Tanizawa, H. Sato, S. Kawamura, M. Fujiwara, M. Sasaki, and A. J. Shields, "Quantum key distribution with hacking countermeasures and long term field trial," *Sci. Rep.* **7**, 1978 (2017).
- [9] "ETSI white paper no. 27. Implementation security of quantum cryptography," https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp27_qkd_imp_sec_FINAL.pdf, visited 13 Feb 2023.
- [10] Akihisa Tomita, "Implementation security certification of decoy-BB84 quantum key distribution systems," *Adv. Quantum Technol.* **2**, 1900005 (2019).
- [11] 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).
- [12] Shihai Sun and Anqi Huang, "A review of security evaluation of practical quantum key distribution system," *Entropy* **24**, 260 (2022).
- [13] "ISO/IEC DIS 23837-2(en). Information technology security techniques — Security requirements, test and evaluation methods for quantum key distribution — Part 2: Evaluation and testing methods," <https://www.iso.org/obp/ui/#iso:std:iso-iec:23837:-2:dis:ed-1:v1:en>, visited 13 Feb 2023.
- [14] "Draft ETSI GS QKD 010 V0.4.1 (2021-06). Quantum key distribution (QKD); Implementation security: protection against Trojan horse attacks," https://docbox.etsi.org/ISG/QKD/Open/GS-QKD-0010_IStrojan_v0.4.1_OpenArea.pdf, visited 13 Oct 2023.
- [15] Shihan Sajeed, Poompong Chaiwongkhot, Anqi Huang, Hao Qin, Vladimir Egorov, Anton Kozubov, Andrei Gaidash, Vladimir Chistiakov, Artur Vasiliev, Artur Gleim, and Vadim Makarov, "An approach for security evaluation and certification of a complete quantum communication system," *Sci. Rep.* **11**, 5110 (2021).
- [16] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, "Full-field implementation of a perfect eavesdropper on a quantum cryptography system," *Nat. Commun.* **2**, 349 (2011).
- [17] Eric Numkam Fokou, Seyed Abokhamis Mousavi, Gregory T. Jasion, David J. Richardson, and Francesco

- Poletti, “Loss in hollow-core optical fibers: mechanisms, scaling rules, and limits,” *Adv. Opt. Photonics* **15**, 1 (2023).
- [18] Kiyoshi Tamaki, Marcos Curty, Go Kato, Hoi-Kwong Lo, and Koji Azuma, “Loss-tolerant quantum cryptography with imperfect sources,” *Phys. Rev. A* **90**, 052314 (2014).
- [19] Akihiro Mizutani, Marcos Curty, Charles Ci Wen Lim, Nobuyuki Imoto, and Kiyoshi Tamaki, “Finite-key security analysis of quantum key distribution with imperfect light sources,” *New J. Phys.* **17**, 093011 (2015).
- [20] Akihiro Mizutani, Go Kato, Koji Azuma, Marcos Curty, Rikizo Ikuta, Takashi Yamamoto, Nobuyuki Imoto, Hoi-Kwong Lo, and Kiyoshi Tamaki, “Quantum key distribution with setting-choice-independently correlated light sources,” *npj Quantum Inf.* **5**, 8 (2019).
- [21] Margarida Pereira, Marcos Curty, and Kiyoshi Tamaki, “Quantum key distribution with flawed and leaky sources,” *npj Quantum Inf.* **5**, 62 (2019).
- [22] Margarida Pereira, Go Kato, Akihiro Mizutani, Marcos Curty, and Kiyoshi Tamaki, “Quantum key distribution with correlated sources,” *Sci. Adv.* **6**, eaaz4487 (2020).
- [23] Margarida Pereira, Guillermo Currás-Lorenzo, Álvaro Navarrete, Akihiro Mizutani, Go Kato, Marcos Curty, and Kiyoshi Tamaki, “Modified BB84 quantum key distribution protocol robust to source imperfections,” arXiv:2210.11754 [quant-ph].
- [24] Guillermo Currás-Lorenzo, Margarida Pereira, Go Kato, Marcos Curty, and Kiyoshi Tamaki, “A security framework for quantum key distribution implementations,” arXiv:2305.05930 [quant-ph].
- [25] S.N. Molotkov, “Trojan horse attacks, decoy state method, and side channels of information leakage in quantum cryptography,” *J. Exp. Theor. Phys.* **130**, 809–832 (2020).
- [26] S.N. Molotkov, “Side channels of information leakage in quantum cryptography: nonstrictly single-photon states, different quantum efficiencies of detectors, and finite transmitted sequences,” *J. Exp. Theor. Phys.* **133**, 272–304 (2021).
- [27] Private discussions at *Security proofs in QKD online workshop, 15–17 July 2020*.
- [28] Alexander Duplinskiy, *Quantum key distribution with high-rate polarization encoding*, Ph.D. thesis, Moscow Institute of Physics and Technology (2019).
- [29] E. O. Kiktenko, A. S. Trushechkin, C. C. W. Lim, Y. V. Kurochkin, and A. K. Fedorov, “Symmetric blind information reconciliation for quantum key distribution,” *Phys. Rev. Appl.* **8**, 044017 (2017).
- [30] Nikolay Borisov, Ivan Petrov, and Andrey Tayduganov, “Asymmetric adaptive LDPC-based information reconciliation for industrial quantum key distribution,” *Entropy* **25**, 31 (2023).
- [31] Ted Krovetz and Phillip Rogaway, “Fast universal hashing with small keys and no preprocessing: The PolyR construction,” *Lect. Notes Comp. Sci.* **2015**, 73–89 (2001).
- [32] A. K. Fedorov, E. O. Kiktenko, and A. S. Trushechkin, “Symmetric blind information reconciliation and hash-function-based verification for quantum key distribution,” *Lobachevskii J. Math.* **39**, 992–996 (2018).
- [33] A. S. Trushechkin, E. O. Kiktenko, and A. K. Fedorov, “Practical issues in decoy-state quantum key distribution based on the central limit theorem,” *Phys. Rev. A* **96**, 022316 (2017).
- [34] Hugo Krawczyk, “LFSR-based hashing and authentication,” *Lect. Notes Comp. Sci.* **839**, 129–139 (1994).
- [35] Hugo Krawczyk, “New hash functions for message authentication,” *Lect. Notes Comp. Sci.* **921**, 301–310 (1995).
- [36] Peng Ye, Wei Chen, Guo-Wei Zhang, Feng-Yu Lu, Fang-Xiang Wang, Guan-Zhong Huang, Shuang Wang, De-Yong He, Zhen-Qiang Yin, Guang-Can Guo, and Zheng-Fu Han, “Induced-photorefractive attack against quantum key distribution,” *Phys. Rev. Appl.* **19**, 054052 (2023).
- [37] “Cerberis XG QKD system,” <https://www.idquantique.com/quantum-safe-security/products/cerberis-xg-qkd-system/>, visited 14 Feb 2023.
- [38] N. Walenta, A. Burg, D. Caselunghe, J. Constantin, N. Gisin, O. Guinnard, R. Houlmann, P. Junod, B. Korzh, N. Kulesza, M. Legré, C. W. Lim, T. Lunghi, L. Monat, C. Portmann, M. Soucarros, R. T. Thew, P. Trinkler, G. Trollet, F. Vannel, and H. Zbinden, “A fast and versatile quantum key distribution system with hardware key distillation and wavelength multiplexing,” *New J. Phys.* **16**, 013047 (2014).
- [39] D. Stucki, N. Brunner, N. Gisin, V. Scarani, and H. Zbinden, “Fast and simple one-way quantum key distribution,” *Appl. Phys. Lett.* **87**, 194108 (2005).
- [40] Róbert Trényi and Marcos Curty, “Zero-error attack against coherent-one-way quantum key distribution,” *New J. Phys.* **23**, 093005 (2021).
- [41] Anton Kozubov, Andrei Gaidash, and George Miroshnichenko, “Finite-key security for quantum key distribution systems utilizing weak coherent states,” arXiv:1903.04371 [quant-ph].
- [42] S. P. Kulik, S. N. Molotkov, and A. P. Makhaveev, “Combined phase-time encoding method in quantum cryptography,” *JETP Lett.* **85**, 297 (2007).
- [43] “ViPNet Quandor,” <https://quantum-crypto.ru/projects/vipnet-quandor/>, visited 15 Feb 2023.
- [44] D. A. Kronberg, “Vulnerability of quantum cryptography with phase-time coding under attenuation conditions,” *Theor. Math. Phys.* **214**, 121 (2023).
- [45] L. Lydersen, N. Jain, C. Wittmann, Ø. Marøy, J. Skaar, C. Marquardt, V. Makarov, and G. Leuchs, “Super-linear threshold detectors in quantum cryptography,” *Phys. Rev. A* **84**, 032320 (2011).
- [46] S. Cova, M. Ghioni, A. Lotito, I. Rech, and F. Zappa, “Evolution and prospects for single-photon avalanche diodes and quenching circuits,” *J. Mod. Opt.* **51**, 1267–1288 (2004).
- [47] C. Wiechers, L. Lydersen, C. Wittmann, D. Elser, J. Skaar, C. Marquardt, V. Makarov, and G. Leuchs, “After-gate attack on a quantum cryptosystem,” *New J. Phys.* **13**, 013043 (2011).
- [48] Audun Nystad Bugge, Sebastien Sauge, Aina Mardhiyah M. Ghazali, Johannes Skaar, Lars Lydersen, and Vadim Makarov, “Laser damage helps the eavesdropper in quantum cryptography,” *Phys. Rev. Lett.* **112**, 070503 (2014).
- [49] Vadim Makarov, Jean-Philippe Bourgoin, Poompong Chaiwongkhot, Mathieu Gagné, Thomas Jennewein, Sarah Kaiser, Raman Kashyap, Matthieu Legré, Carter Minshull, and Shihan Sajeed, “Creation of backdoors

- in quantum communications via laser damage,” *Phys. Rev. A* **94**, 030302 (2016).
- [50] Gaëtan Gras, Nigar Sultana, Anqi Huang, Thomas Jennewein, Félix Bussi eres, Vadim Makarov, and Hugo Zbinden, “Optical control of single-photon negative-feedback avalanche diode detector,” *J. Appl. Phys.* **127**, 094502 (2020).
- [51] Zhihao Wu, Anqi Huang, Huan Chen, Shi-Hai Sun, Jiangfang Ding, Xiaogang Qiang, Xiang Fu, Ping Xu, and Junjie Wu, “Hacking single-photon avalanche detectors in quantum key distribution via pulse illumination,” *Opt. Express* **28**, 25574 (2020).
- [52] Bulavkin D.S., Sushchev I.S., Bugai K.E., Bogdanov S.A., and Dvoretzkiy D.A., “Study of a single-photon detector blinding attack with modulated bright light,” *Proc. SPIE* **12323**, 123230E (2022).
- [53] Polina Acheva, Konstantin Zaitsev, Vladimir Zavadilenko, Anton Losev, Anqi Huang, and Vadim Makarov, “Automated verification of countermeasure against detector-control attack in quantum key distribution,” *EPJ Quantum Technol.* **10**, 22 (2023).
- [54] Mikhail Kuzmin, *Resistance of an avalanche photon detector to bright-light attacks in quantum key distribution*, Master’s thesis, Moscow Technical University of Communication and Informatics (2023).
- [55] H.-K. Lo, M. Curty, and B. Qi, “Measurement-device-independent quantum key distribution,” *Phys. Rev. Lett.* **108**, 130503 (2012).
- [56] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, “Overcoming the rate–distance limit of quantum key distribution without quantum repeaters,” *Nature* **557**, 400 (2018).
- [57] Xiang-Bin Wang, Zong-Wen Yu, and Xiao-Long Hu, “Twin-field quantum key distribution with large misalignment error,” *Phys. Rev. A* **98**, 062323 (2018).
- [58] P. Rice and J. Harrington, “Numerical analysis of decoy state quantum key distribution protocols,” arXiv:0901.0013v2 [quant-ph].
- [59] M. K. Bochkov and A. S. Trushechkin, “Security of quantum key distribution with detection-efficiency mismatch in the single-photon case: Tight bounds,” *Phys. Rev. A* **99**, 032308 (2019).
- [60] Anton Trushechkin, “Security of quantum key distribution with detection-efficiency mismatch in the multiphoton case,” *Quantum* **6**, 771 (2022).
- [61] Bing Qi, Chi-Hang Fred Fung, Hoi-Kwong Lo, and Xiongfeng Ma, “Time-shift attack in practical quantum cryptosystems,” *Quantum Inf. Comput.* **7**, 73–82 (2007).
- [62] A. Vakhitov, V. Makarov, and D. R. Hjelle, “Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography,” *J. Mod. Opt.* **48**, 2023–2038 (2001).
- [63] Nitin Jain, Elena Anisimova, Imran Khan, Vadim Makarov, Christoph Marquardt, and Gerd Leuchs, “Trojan-horse attacks threaten the security of practical quantum cryptography,” *New J. Phys.* **16**, 123030 (2014).
- [64] Shihan Sajeed, Carter Minshull, Nitin Jain, and Vadim Makarov, “Invisible Trojan-horse attack,” *Sci. Rep.* **7**, 8403 (2017).
- [65] V. Makarov, A. Anisimov, and J. Skaar, “Effects of detector efficiency mismatch on security of quantum cryptosystems,” *Phys. Rev. A* **74**, 022313 (2006), erratum *ibid.* **78**, 019905 (2008).
- [66] Yi Zhao, Chi-Hang Fred Fung, Bing Qi, Christine Chen, and Hoi-Kwong Lo, “Quantum hacking: Experimental demonstration of time-shift attack against practical quantum-key-distribution systems,” *Phys. Rev. A* **78**, 042333 (2008).
- [67] Hong-Wei Li, Shuang Wang, Jing-Zheng Huang, Wei Chen, Zhen-Qiang Yin, Fang-Yi Li, Zheng Zhou, Dong Liu, Yang Zhang, Guang-Can Guo, Wan-Su Bao, and Zheng-Fu Han, “Attacking a practical quantum-key-distribution system with wavelength-dependent beam-splitter and multiwavelength sources,” *Phys. Rev. A* **84**, 062308 (2011).
- [68] Jing-Zheng Huang, Christian Weedbrook, Zhen-Qiang Yin, Shuang Wang, Hong-Wei Li, Wei Chen, Guang-Can Guo, and Zheng-Fu Han, “Quantum hacking of a continuous-variable quantum-key-distribution system using a wavelength attack,” *Phys. Rev. A* **87**, 062329 (2013).
- [69] H. Weier, H. Krauss, M. Rau, M. F urst, S. Nauerth, and H. Weinfurter, “Quantum eavesdropping without interception: an attack exploiting the dead time of single-photon detectors,” *New J. Phys.* **13**, 073024 (2011).
- [70] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, “Trojan-horse attacks on quantum-key-distribution systems,” *Phys. Rev. A* **73**, 022320 (2006).
- [71] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. L. Yuan, and A. J. Shields, “Practical security bounds against the Trojan-horse attack in quantum key distribution,” *Phys. Rev. X* **5**, 031030 (2015).
- [72] Kiyoshi Tamaki, Marcos Curty, and Marco Lucamarini, “Decoy-state quantum key distribution with a leaky source,” *New J. Phys.* **18**, 065008 (2016).
- [73] Weilong Wang, Kiyoshi Tamaki, and Marcos Curty, “Finite-key security analysis for quantum key distribution with leaky sources,” *New J. Phys.* **20**, 083027 (2018).
- [74] Nitin Jain, Birgit Stiller, Imran Khan, Vadim Makarov, Christoph Marquardt, and Gerd Leuch, “Risk analysis of Trojan-horse attacks on practical quantum key distribution systems,” *IEEE J. Sel. Top. Quantum Electron.* **21**, 6600710 (2015).
- [75] Poompong Chaiwongkhot, Hao Qin, Iris Choi, Norbert L utkenhaus, Martin B. Ward, and Vadim Makarov, “The role of interference in Trojan-horse attack on quantum cryptography,” (2018), unpublished internal report for ETSI ISG-QKD.
- [76] Ivan S. Sushchev, Diana M. Guzairova, Andrey N. Klimov, Dmitriy A. Dvoretzkiy, Sergey A. Bogdanov, Klim D. Bondar, and Anton P. Naumenko, “Practical security analysis against the Trojan-horse attacks on fiber-based phase-coding QKD system in the wide spectral range,” *Proc. SPIE* **11868**, 118680H (2021).
- [77] Shi-Hai Sun, Feihu Xu, Mu-Sheng Jiang, Xiang-Chun Ma, Hoi-Kwong Lo, and Lin-Mei Liang, “Effect of source tampering in the security of quantum cryptography,” *Phys. Rev. A* **92**, 022304 (2015).
- [78] Anqi Huang,  lvaro Navarrete, Shi-Hai Sun, Poompong Chaiwongkhot, Marcos Curty, and Vadim Makarov, “Laser-seeding attack in quantum key distribution,” *Phys. Rev. Appl.* **12**, 064043 (2019).
- [79] Xiao-Ling Pang, Ai-Lin Yang, Chao-Ni Zhang, Jian-

- Peng Dou, Hang Li, Jun Gao, and Xian-Min Jin, “Hacking quantum key distribution via injection locking,” *Phys. Rev. Appl.* **13**, 034008 (2020).
- [80] V. Lovic, D.G. Marangon, P.R. Smith, R.I. Woodward, and A.J. Shields, “Quantified effects of the laser-seeding attack in quantum key distribution,” *Phys. Rev. Appl.* **20**, 044005 (2023).
- [81] Feng-Yu Lu, Peng Ye, Ze-Hao Wang, Shuang Wang, Zhen-Qiang Yin, Rong Wang, Xiao-Juan Huang, Wei Chen, De-Yong He, Guan-Jie Fan-Yuan, Guang-Can Guo, and Zheng-Fu Han, “Hacking measurement-device-independent quantum key distribution,” *Optica* **10**, 520 (2023).
- [82] Liying Han, Yang Li, Hao Tan, Weiyang Zhang, Wenqi Cai, Juan Yin, Jigang Ren, Feihu Xu, Shengkai Liao, and Chengzhi Peng, “Effect of light injection on the security of practical quantum key distribution,” *Phys. Rev. Appl.* **20**, 044013 (2023).
- [83] S.M. Kostritskii, “Photorefractive effect in LiNbO₃-based integrated-optical circuits at wavelengths of third telecom window,” *Appl. Phys. B* **95**, 421–428 (2009).
- [84] S. R. Friberg, A. M. Weiner, Y. Silberberg, B. G. Sfez, and P. S. Smith, “Femtosecond switching in a dual-core-fiber nonlinear coupler,” *Opt. Lett.* **13**, 904–906 (1988).
- [85] Anqi Huang, Ruoping Li, Vladimir Egorov, Serguei Tchouragoulov, Krtin Kumar, and Vadim Makarov, “Laser-damage attack against optical attenuators in quantum key distribution,” *Phys. Rev. Appl.* **13**, 034017 (2020).
- [86] Bugai K.E., Zyzykin A.P., Bulavkin D.S., Bogdanov S.A., Sushchev I.S., and Dvoretzkiy D.A., “Laser damage attack on a simple optical attenuator widely used in fiber-based QKD systems,” in *Proc. 2022 International Conference Laser Optics (ICLO)* (IEEE, 2022) p. 393.
- [87] Anastasiya Ponosova, Daria Ruzhitskaya, Poompong Chaiwongkhot, Vladimir Egorov, Vadim Makarov, and Anqi Huang, “Protecting fiber-optic quantum key distribution sources against light-injection attacks,” *PRX Quantum* **3**, 040307 (2022).
- [88] Daria D. Ruzhitskaya, Irina V. Zhluktova, Mikhail A. Petrov, Konstantin A. Zaitsev, Polina P. Acheva, Nikolay A. Zunikov, Aleksei V. Shilko, Djeylan Akta, Friederike Johlinger, Daniil O. Trefilov, Anastasiya A. Ponosova, Vladimir A. Kamynin, and Vadim V. Makarov, “Vulnerabilities in the quantum key distribution system induced under a pulsed laser attack,” *Sci. Tech. J. Inf. Technol. Mech. Opt.* **21**, 837–847 (2021), in Russian.
- [89] Roger M. Wood, *Laser-Induced Damage of Optical Materials*, 1st ed. (CRC Press, 2003).
- [90] Alice Meda, Ivo Pietro Degiovanni, Alberto Tosi, Zhiliang Yuan, Giorgio Brida, and Marco Genovese, “Quantifying backflash radiation to prevent zero-error attacks in quantum key distribution,” *Light Sci. Appl.* **6**, e16261 (2017).
- [91] Paulo Vinicius Pereira Pinheiro, Poompong Chaiwongkhot, Shihan Sajeed, Rolf T. Horn, Jean-Philippe Bourgoin, Thomas Jennewein, Norbert Lütkenhaus, and Vadim Makarov, “Eavesdropping and countermeasures for backflash side channel in quantum cryptography,” *Opt. Express* **26**, 21020–21032 (2018).
- [92] A. Koehler-Sidki, J.F. Dynes, T.K. Paraíso, M. Lucamarini, A. W. Sharpe, Z.L. Yuan, and A.J. Shields, “Backflashes from fast-gated avalanche photodiodes in quantum key distribution,” *Appl. Phys. Lett.* **116**, 154001 (2020).
- [93] Yicheng Shi, Janet Zheng Jie Lim, Hou Shun Poh, Peng Kian Tan, Peiyu Amelia Tan, Alexander Ling, and Christian Kurtsiefer, “Breakdown flash at telecom wavelengths in InGaAs avalanche photodiodes,” *Opt. Express* **25**, 30388 (2017).
- [94] Sergey A. Bogdanov, Ivan S. Sushchev, Andrey N. Klimov, Kirill E. Bugai, Daniil S. Bulavkin, and Dmitriy A. Dvoretzkiy, “Influence of QKD apparatus parameters on the “backflash” attack,” *Proc. SPIE* **12133**, 121330G (2022).
- [95] Ken-ichiro Yoshino, Mikio Fujiwara, Kensuke Nakata, Tatsuya Sumiya, Toshihiko Sasaki, Masahiro Takeoka, Masahide Sasaki, Akio Tajima, Masato Koashi, and Akihisa Tomita, “Quantum key distribution with an efficient countermeasure against correlated intensity fluctuations in optical pulses,” *npj Quantum Inf.* **4**, 8 (2018).
- [96] 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).
- [97] Toshiya Kobayashi, Akihisa Tomita, and Atsushi Okamoto, “Evaluation of the phase randomness of a light source in quantum-key-distribution systems with an attenuated laser,” *Phys. Rev. A* **90**, 032320 (2014).
- [98] Daniil Trefilov, *Imperfect state preparation in quantum key distribution*, Master’s thesis, Higher School of Economics (2021).
- [99] Víctor Zapatero, Álvaro Navarrete, Kiyoshi Tamaki, and Marcos Curty, “Security of quantum key distribution with intensity correlations,” *Quantum* **5**, 602 (2021).
- [100] Guillermo Currás-Lorenzo, Kiyoshi Tamaki, and Marcos Curty, “Security of decoy-state quantum key distribution with imperfect phase randomization,” arXiv:2210.08183v1 [quant-ph].
- [101] Xoel Sixto, Víctor Zapatero, and Marcos Curty, “Security of decoy-state quantum key distribution with correlated intensity fluctuations,” *Phys. Rev. Appl.* **18**, 044069 (2022).
- [102] Feihu Xu, Kejin Wei, Shihan Sajeed, Sarah Kaiser, Shihai Sun, Zhiyuan Tang, Li Qian, Vadim Makarov, and Hoi-Kwong Lo, “Experimental quantum key distribution with source flaws,” *Phys. Rev. A* **92**, 032305 (2015).
- [103] Guillermo Currás-Lorenzo, Álvaro Navarrete, Margarida Pereira, and Kiyoshi Tamaki, “Finite-key analysis of loss-tolerant quantum key distribution based on random sampling theory,” *Phys. Rev. A* **104**, 012406 (2021).
- [104] Nitin Jain, Christoffer Wittmann, Lars Lydersen, Carlos Wiechers, Dominique Elser, Christoph Marquardt, Vadim Makarov, and Gerd Leuchs, “Device calibration impacts security of quantum key distribution,” *Phys. Rev. Lett.* **107**, 110501 (2011).
- [105] Yang-Yang Fei, Xiang-Dong Meng, Ming Gao, Hong Wang, and Zhi Ma, “Quantum man-in-the-middle attack on the calibration process of quantum key distribution,” *Sci. Rep.* **8**, 4283 (2018).
- [106] F. Xu, B. Qi, and H.-K. Lo, “Experimental demonstration of phase-remapping attack in a practical quan-

- tum key distribution system,” *New J. Phys.* **12**, 113026 (2010).
- [107] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, “Security of quantum key distribution with imperfect devices,” *Quantum Inf. Comput.* **4**, 325–360 (2004).
- [108] Dmitriy Kuzmin, *Resistance of a single-photon detector to after-gate attack in quantum key distribution*, Master’s thesis, Moscow Technical University of Communication and Informatics (2023).
- [109] Anqi Huang, Akihiro Mizutani, Hoi-Kwong Lo, Vadim Makarov, and Kiyoshi Tamaki, “Characterization of state-preparation uncertainty in quantum key distribution,” *Phys. Rev. Appl.* **19**, 014048 (2023).
- [110] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” in *Proc. International Conference on Computers, Systems, and Signal Processing* (IEEE Press, New York, Bangalore, India, 1984) pp. 175–179.
- [111] A.S. Trushechkin, E.O. Kiktenko, D.A. Kronberg, and A.K. Fedorov, “Security of the decoy state method for quantum key distribution,” *Phys. Uspekhi* **64**, 88 (2021).
- [112] D. Dieks, “Communication by EPR devices,” *Phys. Lett. A* **92**, 271–272 (1982).
- [113] W. K. Wootters and W. H. Zurek, “A single quantum cannot be cloned,” *Nature* **299**, 802–803 (1982).
- [114] Dominic Mayers, “Quantum key distribution and string oblivious transfer in noisy channels,” *Lect. Notes Comp. Sci.* **1109**, 343–357 (1996).
- [115] Hoi-Kwong Lo and H. F. Chau, “Unconditional security of quantum key distribution over arbitrarily long distances,” *Science* **283**, 2050–2056 (1999).
- [116] Peter W. Shor and John Preskill, “Simple proof of security of the BB84 quantum key distribution protocol,” *Phys. Rev. Lett.* **85**, 441–444 (2000).
- [117] Won-Young Hwang, “Quantum key distribution with high loss: Toward global secure communication,” *Phys. Rev. Lett.* **91**, 057901 (2003).
- [118] Hoi-Kwong Lo, Xiongfeng Ma, and Kai Chen, “Decoy state quantum key distribution,” *Phys. Rev. Lett.* **94**, 230504 (2005).
- [119] Xiongfeng Ma, Bing Qi, Yi Zhao, and Hoi-Kwong Lo, “Practical decoy state for quantum key distribution,” *Phys. Rev. A* **72**, 012326 (2005).
- [120] Zhen Zhang, Qi Zhao, Mohsen Razavi, and Xiongfeng Ma, “Improved key-rate bounds for practical decoy-state quantum-key-distribution systems,” *Phys. Rev. A* **95**, 012333 (2017).
- [121] Evgeny Kiktenko, Anton Trushechkin, Yury Kurochkin, and Aleksey Fedorov, “Post-processing procedure for industrial quantum key distribution systems,” *J. Phys. Conf. Ser.* **741**, 012081 (2016).
- [122] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner, “Leftover hashing against quantum side information,” *IEEE Trans. Inf. Theory* **57**, 5524–5535 (2011).
- [123] V. Dolmatov, Ed. and A. Degtyarev, “GOST R 34.11-2012: Hash function,” RFC **6986** (2013).
- [124] Anton Trushechkin, “On the operational meaning and practical aspects of using the security parameter in quantum key distribution,” *Quantum Electron.* **50**, 426–439 (2020).
- [125] Z. L. Yuan, J. F. Dynes, and A. J. Shields, “Response to ‘Comment on ‘Resilience of gated avalanche photodiodes against bright illumination attacks in quantum cryptography’” [Appl. Phys. Lett. 99, 196101 (2011)],” *Appl. Phys. Lett.* **99**, 196102 (2011).
- [126] Yong-Jun Qian, De-Yong He, Shuang Wang, Wei Chen, Zhen-Qiang Yin, Guang-Can Guo, and Zheng-Fu Han, “Hacking the quantum key distribution system by exploiting the avalanche-transition region of single-photon detectors,” *Phys. Rev. Appl.* **10**, 064062 (2018).
- [127] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, “Hacking commercial quantum cryptography systems by tailored bright illumination,” *Nat. Photonics* **4**, 686–689 (2010).
- [128] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, “Thermal blinding of gated detectors in quantum cryptography,” *Opt. Express* **18**, 27938–27954 (2010).
- [129] L. Lydersen, M. K. Akhlaghi, A. H. Majedi, J. Skaar, and V. Makarov, “Controlling a superconducting nanowire single-photon detector using tailored bright illumination,” *New J. Phys.* **13**, 113042 (2011).
- [130] V. Makarov, “Controlling passively quenched single photon detectors by bright light,” *New J. Phys.* **11**, 065003 (2009).
- [131] S. Sauge, L. Lydersen, A. Anisimov, J. Skaar, and V. Makarov, “Controlling an actively-quenched single photon detector with bright light,” *Opt. Express* **19**, 23590–23600 (2011).
- [132] Michael G. Tanner, Vadim Makarov, and Robert H. Hadfield, “Optimised quantum hacking of superconducting nanowire single-photon detectors,” *Opt. Express* **22**, 6734–6748 (2014).
- [133] Aleksey Fedorov, Ilja Gerhardt, Anqi Huang, Jonathan Jogenfors, Yury Kurochkin, Antía Lamas-Linares, Jan-Åke Larsson, Gerd Leuchs, Lars Lydersen, Vadim Makarov, and Johannes Skaar, “Comment on ‘Inherent security of phase coding quantum key distribution systems against detector blinding attacks’ (2018 Laser Phys. Lett. 15 095203),” *Laser Phys. Lett.* **16**, 019401 (2019).
- [134] R. I. Woodward, Y. S. Lo, M. Pittaluga, M. Minder, T. K. Paraíso, M. Lucamarini, Z. L. Yuan, and A. J. Shields, “Gigahertz measurement-device-independent quantum key distribution using directly modulated lasers,” *npj Quantum Inf.* **7**, 58 (2021).
- [135] Z. L. Yuan, J. F. Dynes, and A. J. Shields, “Resilience of gated avalanche photodiodes against bright illumination attacks in quantum cryptography,” *Appl. Phys. Lett.* **98**, 231104 (2011).
- [136] A. Koehler-Sidki, J. F. Dynes, M. Lucamarini, G. L. Roberts, A. W. Sharpe, Z. L. Yuan, and A. J. Shields, “Best-practice criteria for practical security of self-differencing avalanche photodiode detectors in quantum key distribution,” *Phys. Rev. Appl.* **9**, 044027 (2018).
- [137] Yong-Jun Qian, Hong-Wei Li, De-Yong He, Zhen-Qiang Yin, Chun-Mei Zhang, Wei Chen, Shuang Wang, and Zheng-Fu Han, “Countermeasure against probabilistic blinding attack in practical quantum key distribution systems,” *Chin. Phys. B* **24**, 090305 (2015).
- [138] T. F. da Silva, G. B. Xavier, G. P. Temporão, and J. P. von der Weid, “Real-time monitoring of single-photon detectors against eavesdropping in quantum key distribution systems,” *Opt. Express* **20**, 18911–18924 (2012).
- [139] A. Koehler-Sidki, J. F. Dynes, A. Martinez, M. Luca-

- marini, G. L. Roberts, A. W. Sharpe, Z. L. Yuan, and A. J. Shields, “Intrinsic mitigation of the after-gate attack in quantum key distribution through fast-gated delayed detection,” *Phys. Rev. Appl.* **12**, 024050 (2019).
- [140] Vladimir Chistiakov, Anqi Huang, Vladimir Egorov, and Vadim Makarov, “Controlling single-photon detector ID210 with bright light,” *Opt. Express* **27**, 32253 (2019).
- [141] Muataz Alhoussein, Kyo Inoue, and Toshimori Honjo, “Monitoring coincident clicks in differential-quadrature-phase shift QKD to reveal detector blinding and control attacks,” *Jpn. J. Appl. Phys.* **58**, 012006 (2019).
- [142] L. Lydersen, V. Makarov, and J. Skaar, “Secure gated detection scheme for quantum cryptography,” *Phys. Rev. A* **83**, 032306 (2011).
- [143] Øystein Marøy, Vadim Makarov, and Johannes Skaar, “Secure detection in quantum key distribution by real-time calibration of receiver,” *Quantum Sci. Technol.* **2**, 044013 (2017).
- [144] Min Soo Lee, Byung Kwon Park, Min Ki Woo, Chang Hoon Park, Yong-Su Kim, Sang-Wook Han, and Sung Moon, “Countermeasure against blinding attacks on low-noise detectors with a background-noise-cancellation scheme,” *Phys. Rev. A* **94**, 062321 (2016).
- [145] Yong-Jun Qian, De-Yong He, Shuang Wang, Wei Chen, Zhen-Qiang Yin, Guang-Can Guo, and Zheng-Fu Han, “Robust countermeasure against detector control attack in a practical quantum key distribution system,” *Optica* **6**, 1178–1184 (2019).
- [146] Charles Ci Wen Lim, Nino Walenta, Matthieu Legré, Nicolas Gisin, and Hugo Zbinden, “Random variation of detector efficiency: a countermeasure against detector blinding attacks for quantum key distribution,” *IEEE J. Sel. Top. Quantum Electron.* **21**, 6601305 (2015).
- [147] Thiago Ferreira da Silva, Gustavo C. do Amaral, Guilherme B. Xavier, Guilherme P. Temporão, and Jean Pierre von der Weid, “Safeguarding quantum key distribution through detection randomization,” *IEEE J. Sel. Top. Quantum Electron.* **21**, 6600309 (2015).
- [148] Anqi Huang, Shihan Sajeed, Poompong Chaiwongkhot, Mathilde Soucarros, Matthieu Legré, and Vadim Makarov, “Testing random-detector-efficiency countermeasure in a commercial system reveals a breakable unrealistic assumption,” *IEEE J. Quantum Electron.* **52**, 8000211 (2016).
- [149] Lee W. Tutt and Thomas F. Boggess, “A review of optical limiting mechanisms and devices using organics, fullerenes, semiconductors and other materials,” *Prog. Quantum Electron.* **17**, 299–338 (1993).
- [150] Michael E. DeRosa and Stephan L. Logunov, “Fiber-optic power limiter based on photothermal defocusing in an optical polymer,” *Appl. Opt.* **42**, 2683 (2003).
- [151] Ivan Martinec and Dusan Pudis, “Fiber-optical power limiter and cut-off switch based on thermo-optical effect,” *IEEE Photon. Technol. Lett.* **24**, 297–299 (2012).
- [152] Gong Zhang, Ignatius William Primaatmaja, Jing Yan Haw, Xiao Gong, Chao Wang, and Charles Ci Wen Lim, “Securing practical quantum communication systems with optical power limiters,” *PRX Quantum* **2**, 030304 (2021).
- [153] Jonathan Jogenfors, Ashraf Mohamed Elhassan, Johan Ahrens, Mohamed Bourennane, and Jan-Åke Larsson, “Hacking the Bell test using classical light in energy-time entanglement-based quantum key distribution,” *Sci. Adv.* **1**, e1500793 (2015).
- [154] The quartz fiber transparency range is estimated approximately, since the boundaries of this range significantly depend on the fiber type and the technology of its production. The long-wavelength cutoff is determined by the radius of curvature of installed fiber, which is device-specific. We are not aware of reliable data in the literature.
- [155] Thorlabs, IO-H-1550 fiber-optic isolator isolation plot, https://www.thorlabs.com/images/tabImages/IO-H-1550_780.gif, visited 16 April 2022.
- [156] Thorlabs, FA20T fiber-optic attenuator attenuation plot, https://www.thorlabs.com/images/TabImages/FA20T_780.gif, visited 16 April 2022.
- [157] Opneti, 100Ghz Dense Wavelength Division Multiplexer data sheet, <http://www.opneti.com/uploadfile/20101208/20101208120646747.pdf>, visited 16 April 2022.
- [158] Hao Tan, Mikhail Petrov, *et al.*, (2023), unpublished.