Laser Damage Helps the Eavesdropper in Quantum Cryptography

Audun Nystad Bugge,[1] Sebastien Sauge,[2] Aina Mardhiyah M. Ghazali,[3] Johannes Skaar,[1] Lars Lydersen,[1] and Vadim Makarov[4,*]

[1]*Department of Electronics and Telecommunications, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*

[2]*School of Information and Communication Technology, Royal Institute of Technology (KTH), Electrum 229, SE-16440 Kista, Sweden*

[3]*Department of Science in Engineering, Faculty of Engineering, International Islamic University Malaysia, P.O. Box 10, 50728 Kuala Lumpur, Malaysia*

[4]*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

(Received 31 October 2013; published 18 February 2014)

We propose a class of attacks on quantum key distribution (QKD) systems where an eavesdropper actively engineers new loopholes by using damaging laser illumination to permanently change properties of system components. This can turn a perfect QKD system into a completely insecure system. A proof-of-principle experiment performed on an avalanche photodiode-based detector shows that laser damage can be used to create loopholes. After ~1 W illumination, the detectors' dark count rate reduces 2–5 times, permanently improving single-photon counting performance. After ~1.5 W, the detectors switch permanently into the linear photodetection mode and become completely insecure for QKD applications.

DOI: 10.1103/PhysRevLett.112.070503 PACS numbers: 03.67.Dd, 42.62.Cf, 61.80.Ba, 85.60.Dw

Quantum key distribution (QKD) enables two remote parties to grow a secret key [1]. The security relies on the laws of physics, provided the components and the system behave according to the models in the security proof [2–4]. Practical implementations contain imperfections, however, which may enable so-called quantum hacking attacks [5–8]. Work is now in progress to restore security, by modifying the implementations to avoid large loopholes [9–13], and generalizing the security proofs [2,14–17] to take the remaining, unavoidable imperfections into account [18]. From these promising directions of research, it may seem that quantum key distribution systems will become nearly perfect in the future, in the sense that all imperfections are either eliminated, or accounted for by additional privacy amplification as quantified by security proofs.

In other words, the eavesdropper Eve in QKD seems to have a sad destiny. She initially had two tools in her suitcase: attacking perfect QKD systems with optimal quantum attacks, and quantum hacking attacks exploiting imperfections. The security proofs eliminated the first tool, while the recent developments in implementations and practical security proofs are about to eliminate the second. However, in this Letter, we demonstrate a third tool in her suitcase. Eve may intentionally damage the system, to actively engineer exploitable imperfections. In this way, even an initially perfect setup can become totally insecure, without raising any alarms. This clearly demonstrates the fact that it is not sufficient to have well-characterized components and systems. Eve may totally change their behavior at some later point. The results ultimately question if communication security is physically attainable at all, in principle.

Changes in characteristics of most optical components inside a QKD system can lead to loopholes being created. QKD schemes rely on known characteristics of, for example, attenuators, beam splitters, modulators, polarization control components, spatial and spectral filters, optical connectors, lenses, mirrors, light sources, and detectors. For a proof-of-principle demonstration of the new class of attacks, we needed to pick a target component and a target type of QKD system. A natural choice was an avalanche photodiode (APD) in a free-space system, for the following reasons. A high-power laser beam is experimentally easier to apply through free-space optics. The APD absorbs most of the incoming light in a small area, which makes it likely to suffer damage at lower power than other optical components. We decided to investigate a widely used Si APD (PerkinElmer C30902SH), employed in single-photon detectors in several QKD experiments [19–25]. For this component, we have demonstrated permanent laser damage useful for eavesdropping, as detailed below.

Initial tests showed that useful laser damage could be achieved. For thorough characterization of effects, we subsequently built an automated setup (Fig. 1) that applied damaging light in small increments and fully characterized the APD in between the exposures [26]. The setup tests a stand-alone APD; however, the results are applicable to a complete QKD system as discussed through this Letter. High-power continuous-wave (cw) illumination is produced by electrically controlled 807 nm laser diode, pigtailed with multimode (MM) fiber of 200 μm core diameter. The beam exiting the MM fiber is collimated, passes through a 50:50 nonpolarizing beam splitter (diverting half of the power into a power meter), a mechanical shutter, and is focused at the

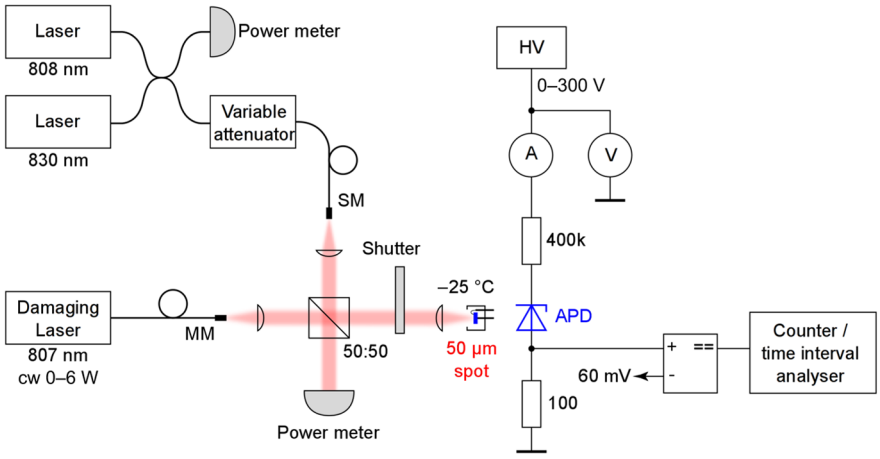

FIG. 1 (color online). Experimental setup for APD damage and characterization. See text for details.

APD in 50 μm full width at half-maximum (FWHM) diameter spot. In an actual attack, the wavelength of the damaging laser would have to be close to that of signal photons, because all free-space QKD systems employ a narrow-band interference filter at the entrance to cut background light in the channel. Many of these systems operate in 770–850 nm wavelength range [19–24], not far from the damaging laser wavelength in our test.

In addition to the high-power laser, our setup has two single-mode (SM) fiber pigtailed lasers and a variable attenuator. These provide calibrated pulsed and cw illumination for characterizing the APD. The APD is connected into a standard passively quenched single-photon detector scheme, and thermoelectrically chilled to −25 °C [27,28]. The detector output is connected to a counter and time interval analyzer. Bias is applied to the APD from a programmable voltage source (HV), allowing measurement of I, V curves and several electrical and optical characteristics [26]. We measured detector dark count rate and photon detection efficiency with APD biased 15 V above its initial (undamaged) breakdown voltage value $V_{\mathrm{br\,orig.}}$, APD breakdown voltage V_{br}, dark current when biased in linear amplification mode 5 V below $V_{\mathrm{br\,orig.}}$, and photoconversion quantum efficiency when biased at 0 V.

Most of our tests proceeded by applying a cycle of cw illumination for 60 s, then characterizing the APD. The power level was increased in small increments between the cycles. The software paused the experiment and alerted the operator if any characterized parameter deviated significantly from its initial value. Then, the operator would either continue the test to higher powers and further destruction of the sample, or terminate the test to check for sample's long-term stability. Results of the tests are plotted in Fig. 2.

We tested ten samples of PerkinElmer C30902SH APD in total, numbered APD-1 through APD-10 in this Letter. The samples came from different production batches manufactured in 2009–2010. The changes observed in all samples after high-power illumination were generally consistent between the samples and permanent, unless noted otherwise. As we applied illumination of increasing power, we observed seven distinct effects denoted by vertical bands a–f in Fig. 2 and explained below.

a. After illumination of less than 0.25 W power, dark count rate of the APDs rose by several times. This is the only nonpermanent effect, dissipating after the APD is left in darkness for several hours. (Also the only known, noted in the APD data sheet.)

b. In 0.3 to 0.45 W range, four out of eight APDs tested in this range exhibited rise of their V_{br} by 2.3–2.5 V. This was accompanied by a reduction in their photon detection efficiency by a factor of 0.83–0.90. Hypothesized mechanism of the efficiency reduction is that while the APDs are biased at a constant voltage in the detector, the rise of V_{br} lowered overvoltage (the difference between the bias voltage and V_{br}), leading to lower detection efficiency [28,29]. When attacking a complete Bob, Eve could thus reduce sensitivity of a selected APD. This is because individual APDs are addressable by varying polarization or other parameters of the damaging light at Bob's input. This would create a permanent efficiency mismatch between Bob's detectors [5]. This efficiency mismatch can potentially increase Eve's knowledge of the key, if Bob does not recharacterize his detectors or accounts for such imperfections in the postprocessing procedure.

c. In 0.5 to 0.8 W range, all APD parameters returned to normal, with the exception of the dark count rate that remarkably fell 1.7–5.4 times comparing to the original dark count rate measured before starting the treatment. The dark count rate reduction was observed in all eight samples tested in this power range, and the change has been verified to be permanent. This is to our knowledge the first demonstration that Eve can improve legitimate user's equipment. The default treatment of all errors in QKD is that they resulted from eavesdropping, regardless of the actual error source. Detector dark counts therefore limit the maximum transmission distance of a given system, raising

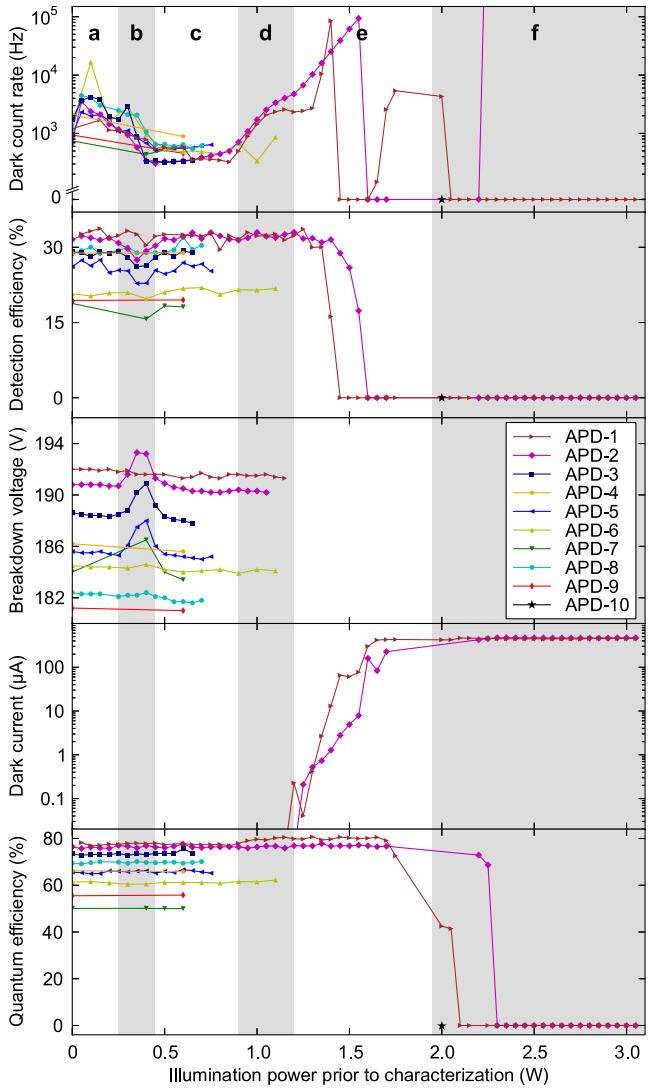

FIG. 2 (color online). Results of applying high-power illumination to ten APD samples. The data points show APD and detector parameters measured after each successive application of illumination of increasing peak power. The leftmost point on each trace is the initial value of parameter prior to illumination.

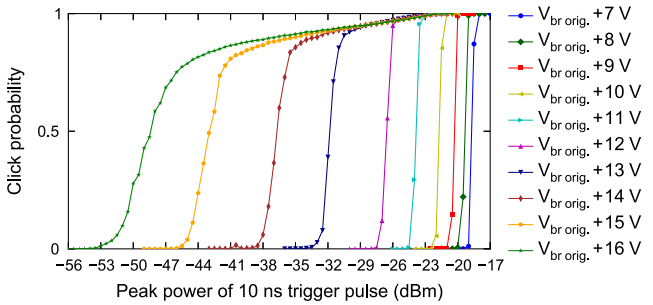

FIG. 3 (color online). Detector control characteristics of a permanently blinded APD-1, at different overvoltage values. Note that trigger pulse power needs to change less than 3 dB (i.e., less than 2 times) for a change of click probability from 0 to >0.5, at typical operating overvoltages of the APD in 10–15 V range. This would allow a perfect or near-perfect faked-state attack on a QKD system [7,32,33]. Note that perfect deterministic 0-or-1 click probability control, as evident at overvoltages ≤ 11 V, is not required for a successful attack. Even probabilistic control at larger overvoltages should suffice to break security in most, if not all, practical settings [34].

the quantum bit error rate (QBER) beyond the secure limit as the photon transmission probability drops. With some extra assumptions or complications in the detection setup, it is possible to improve QKD performance beyond this limit [30,31]. Similarly, it is tempting to simply subtract a calibrated dark count rate from the QBER. Our result clearly shows that this can be dangerous; all errors in the raw key must be treated as caused by eavesdropping [31].

d. In 0.9 to 1.2 W range, the dark count rate permanently rose to large values.

e. In 1.2 to 1.7 W range, the APDs developed a large dark current. This led to blinding of the passively quenched detector, dropping the photon detection efficiency and dark count rate to zero in both samples tested in this range. The blinding mechanism is that excessive current drawn by the APD from the bias circuit (in our case, from 400 kΩ ballast resistor) leads to the voltage supplied by the circuit dropping below V_{br}, as previously demonstrated by weaker cw illumination [35]. The difference here is that the laser damage blinding is permanent and does not require continuing illumination. Under the blinded condition, the detector remains photosensitive to moderately bright light and is either perfectly controllable or well controllable (depending on overvoltage operating setpoint) by 10 ns wide light pulses, see Fig. 3. This renders it insecure for QKD applications [7,8,33].

f. At ≥ 2 W, catastrophic structural damage took place. We tested three samples to this power range. In one of them (APD-10, single experimental point at 2 W in Fig. 2), the bonding wires melted off, leaving the device an open circuit. The other two reduced then completely lost all photosensitivity, with the device becoming a resistor in 10–100 kΩ range. If this APD were employed for a watchdog power meter as in one countermeasure proposal [7], the countermeasure would be defeated.

Later stages of damage result in visible changes to the APD chip (Fig. 4). The first visible change is disfiguring of the gold electrode, possibly resulting from Si-Au alloy formation at $>370\,^{\circ}\mathrm{C}$ [36]. In the last stage of damage, the laser beam always produces a hole in Si chip.

The permanent reduction of dark count rate is an interesting effect. We tested most of our samples illuminated with 50 μm focused, 60 s square pulses of successively increasing power levels, and kept the detector high-voltage source at $V_{\mathrm{br\,orig.}} + 15$ V through the test. However, we have also tested with a single 60 s square pulse applied to a fresh sample (APD-7), with illumination slowly linearly ramped up in 900 s, kept constant for 60 s then linearly ramped down to zero in 900 s (APD-4), with illumination defocused such

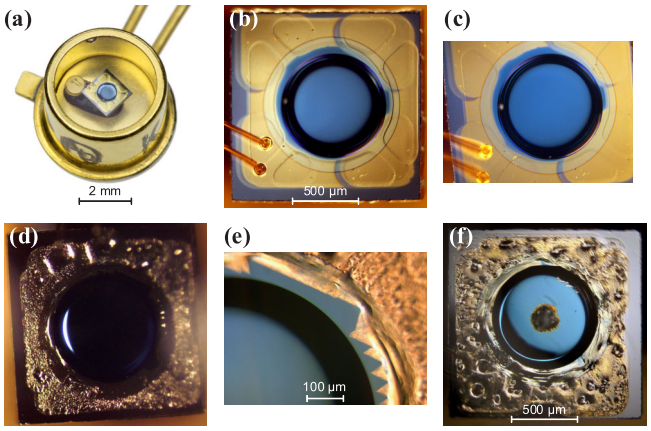

FIG. 4 (color online). Microscope images of APDs at various stages of damage. (a) APD package with Si chip behind a glass window; the other images show chip close-ups. (b) Untreated APD-3. (c) APD-3 after 0.65 W illumination, which has reduced the dark count rate and produced no visible damage. (d)–(e) APD-1 after 2 W illumination, showing remelted gold electrode and gold flowing into clear area along Si crystal lattice planes; this sample has a large dark current but unchanged quantum efficiency. (f) APD-2 after 3 W illumination with a hole blown in the middle through the entire thickness of the silicon chip; it has zero photosensitivity, resistorlike state. Damaging illumination in all cases was applied for 60 s. Images (b)–(e) were taken with bright-field illumination, (d) chip surface intentionally tilted, (f) dark-field illumination.

that the spot became larger than the APD photosensitive area (APD-6), and, finally, with the high-voltage source switched off for the duration of laser treatment (APD-8). In all cases, we observed permanent reduction of dark count rate. It appears that the main cause of it is heating the APD chip to a certain peak temperature. A similar effect has previously been observed and attributed to localized annealing when APD junction was heated by electrical current [36].

The results of testing this component clearly support that Eve may, in general, alter the system characteristics by altering characteristics of its optical components. Then, the system no longer complies with the security proof. Then, even with a sufficiently general security proof, and with a QKD implementation that is precharacterized to comply with the security proof, security cannot be guaranteed. The countermeasure can be to characterize the system more frequently to ensure the validity of the characterization. One could imagine doing this whenever an unusual event was detected, as the bright power of the damaging laser surely has a temporary signature on the system. Meanwhile, it is difficult to exhaustively list all events that should trigger a recharacterization. Eve could, for instance, wait for a power outage, and perform the damage when the system is unpowered.

It is therefore advisable to monitor the characteristics of the system directly during QKD, or at least such that the characteristics are bounded during QKD with a sufficiently high probability. Thus, the security proofs should minimize the number of necessary characteristics about the system. One example is the Bennett-Brassard-Mermin 1992 (BBM92) scheme where the source of entangled photons does not need to be characterized [37]. Another example is the proof for measurement-device-independent QKD systems [11] that has no necessary characterized parameters for the Bell-state analyzer including the detectors. Yet another example is the security proof in Ref. [16], where the secure key generation rate is only dependent on one imperfection parameter at Bob's side, namely, the minimum detection efficiency of a nonvacuum state incident to Bob.

On the implementation side, it turns slightly into a cat-and-mouse game, where Alice and Bob must ensure that the in-field characterization during QKD is reliable and untampered by Eve. Optical power limiters is a well-studied technology that may be applied against tampering at the entrances of Alice and Bob [38], and using a watchdog power meter has been proposed [7,39]. However, our results clearly show that Eve might tamper with these countermeasures. In a more narrow example, detectors can be tested for single-photon sensitivity at random times to bound the minimum detection efficiency [9]. Again, to do this in field is not trivial, and the security then again relies on the precharacterization of the single-photon source and path used for testing. A reliable in-field scheme to characterize crucial equipment parameters during operation can be a future study.

Finally, our study shows the practical challenge of physically securing a QKD system from all side channels. This is one of the most fundamental assumptions in most security proofs (even in the device-independent security proofs [40]), and possibly, the hardest to fully characterize. For example, one can envision a situation where Eve damages the detectors or other crucial components, not by using the fiber, but rather by focusing high-power x-ray radiation onto the components from outside of the system. Another, probably future way to gain access could be a nanorobot burrowing through the fiber core.

We thank E. Anisimova for valuable assistance during the experiments; C. Kurtsiefer, Y.-S. Kim, and Q. Liu for sharing electronics and mechanical design used in parts of the experimental setup. This work was supported by the Research Council of Norway (Grant No. 180439/V30), University Graduate Center in Kjeller, and Industry Canada.

*makarov@vad1.com

[1] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing* (IEEE Press, New York, 1984), p. 175.
[2] D. Mayers, in *Proceedings of Crypto'96*, edited by N. Koblitz (Springer, New York, 1996), Vol. **1109**, p. 343.

[3] H.-K. Lo and H. F. Chau, Science **283**, 2050 (1999).
[4] P. W. Shor and J. Preskill, Phys. Rev. Lett. **85**, 441 (2000).
[5] V. Makarov, A. Anisimov, and J. Skaar, Phys. Rev. A **74**, 022313 (2006); **78**, 019905(E) (2008).
[6] Y. Zhao, C.-H. F. Fung, B. Qi, C. Chen, and H.-K. Lo, Phys. Rev. A **78**, 042333 (2008).
[7] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nat. Photonics **4**, 686 (2010).
[8] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Nat. Commun. **2**, 349 (2011).
[9] L. Lydersen, V. Makarov, and J. Skaar, Phys. Rev. A **83**, 032306 (2011).
[10] Z. L. Yuan, J. F. Dynes, and A. J. Shields, Appl. Phys. Lett. **98**, 231104 (2011).
[11] H.-K. Lo, M. Curty, and B. Qi, Phys. Rev. Lett. **108**, 130503 (2012).
[12] S. L. Braunstein and S. Pirandola, Phys. Rev. Lett. **108**, 130502 (2012).
[13] A. Rubenok, J. A. Slater, P. Chan, I. Lucio-Martinez, and W. Tittel, arXiv:1204.0738v2.
[14] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill, Quantum Inf. Comput. **4**, 325 (2004).
[15] M. Koashi, New J. Phys. **11**, 045018 (2009).
[16] Ø. Marøy, L. Lydersen, and J. Skaar, Phys. Rev. A **82**, 032337 (2010).
[17] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, Nat. Commun. **3**, 634 (2012).
[18] It is still necessary for the manufacturer to show that the actual system is within the more general models of these security proofs. This is typically obtained by characterizing the system in a controlled environment. Note also, however, that the manufacturers could in turn rely on the specifications of the components, thereby anchoring the security in the hands of other external companies.
[19] C. Kurtsiefer, P. Zarda, M. Halder, H. Weinfurter, P. M. Gorman, P. R. Tapster, and J. G. Rarity, Nature (London) **419**, 450 (2002).
[20] T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity, A. Zeilinger, and H. Weinfurter, Phys. Rev. Lett. **98**, 010504 (2007).
[21] M. P. Peloso, I. Gerhardt, C. Ho, A. Lamas-Linares, and C. Kurtsiefer, New J. Phys. **11**, 045007 (2009).
[22] R. J. Hughes, J. E. Nordholt, D. Derkacs, and C. G. Peterson, New J. Phys. **4**, 43 (2002).
[23] A. Poppe, A. Fedrizzi, R. Ursin, H. R. Böhm, T. Loünser, O. Maurhardt, M. Peev, M. Suda, C. Kurtsiefer, H. Weinfurter, T. Jennewein, and A. Zeilinger, Opt. Express **12**, 3865 (2004).
[24] V. L. Kurochkin, I. I. Ryabtsev, and I. G. Neizvestny, Tech. Phys. **50**, 727 (2005).
[25] J. L. Duligall, M. S. Godfrey, K. A. Harrison, W. J. Munro, and J. G. Rarity, New J. Phys. **8**, 249 (2006).
[26] A. M. M. Ghazali, A. N. Bugge, S. Sauge, and V. Makarov, IIUM Eng. J. **12**, 97 (2011).
[27] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, Appl. Opt. **35**, 1956 (1996).
[28] Y.-S. Kim, Y.-C. Jeong, S. Sauge, V. Makarov, and Y.-H. Kim, Rev. Sci. Instrum. **82**, 093110 (2011).
[29] H. Dautet, P. Deschamps, B. Dion, A. D. MacGregor, D. MacSween, R. J. McIntyre, C. Trottier, and P. P. Webb, Appl. Opt. **32**, 3894 (1993).
[30] B. Qi, Y. Zhao, X. Ma, H.-K. Lo, and L. Qian, Phys. Rev. A **75**, 052304 (2007).
[31] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301 (2009).
[32] V. Makarov, A. Anisimov, and S. Sauge, arXiv:0809.3408v1.
[33] L. Lydersen, J. Skaar, and V. Makarov, J. Mod. Opt. **58**, 680 (2011).
[34] L. Lydersen, N. Jain, C. Wittmann, Ø. Marøy, J. Skaar, C. Marquardt, V. Makarov, and G. Leuchs, Phys. Rev. A **84**, 032320 (2011).
[35] V. Makarov, New J. Phys. **11**, 065003 (2009).
[36] R. H. Haitz, J. Appl. Phys. **36**, 3123 (1965).
[37] C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).
[38] L. W. Tutt and T. F. Boggess, Prog. Quantum Electron. **17**, 299 (1993).
[39] V. Makarov and D. R. Hjelme, J. Mod. Opt. **52**, 691 (2005).
[40] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).