Experimentally Faking the Violation of Bell's Inequalities

Ilja Gerhardt,[1,2] Qin Liu,[3] Antía Lamas-Linares,[1] Johannes Skaar,[3,4] Valerio Scarani,[2,5]
Vadim Makarov,[3,4,*] and Christian Kurtsiefer[1,5,†]

[1]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*
[2]*Chemistry Department, Low Temperature Group, University of British Columbia,*
Vancouver, British Columbia V6T 1Z1, Canada
[3]*Department of Electronics and Telecommunications, Norwegian University of Science and Technology,*
NO-7491 Trondheim, Norway
[4]*University Graduate Center, NO-2027 Kjeller, Norway*
[5]*Department of Physics, National University of Singapore, 3 Science Drive 2, Singapore 117543*
(Received 9 August 2011; published 20 October 2011)

Entanglement witnesses such as Bell inequalities are frequently used to prove the nonclassicality of a
light source and its suitability for further tasks. By demonstrating Bell inequality violations using classical
light in common experimental arrangements, we highlight why strict locality and efficiency conditions are
not optional, particularly in security-related scenarios.

Introduction.—Experimental demonstrations of theoretical ideas require a reliable translation between the mathematical objects used in the theory and a specific setup of devices. In some instances trust may not be warranted due to the unnoticed failure of some measurement devices. In the context of quantum cryptography, a warning was raised recently: the possibility of so-called *faked states* [1]. In this scenario, honest scientists are recording real measurement outcomes, performed by devices that seemingly work as they should. What is exploited is the fact that a physical device, even in its presumably normal state, may be sensitive to other degrees of freedom than the ones that are thought to be relevant. For instance, it is customary to encode qubits in (e.g., polarization) states of the light field, but the field is much more than a qubit and the "rest" may trigger a detector as well.

A parallel theoretical development, triggered by quantum cryptography in conjunction with foundational studies, proposes an unexpectedly powerful solution to the problem of trust using Bell's inequalities [2]. These purely statistical criteria can be checked without any knowledge of the degree of freedom that is studied and of the measurements that are performed—in fact, they are even independent of quantum physics. Therefore, whatever physical device leads to a violation of Bell's inequalities can immediately be labeled as nonclassical. This intuition can be refined to provide device-independent criteria for secure cryptography [3,4], trusted randomness [5], entangled states [6,7], and measurements [8]. These works, together with the parallel approach called self testing [9,10], show that some quantum devices can be certified solely from the observed statistics. There is a warning, though: trust should be given if and only if the violation of Bell's inequalities is loophole free. In the context of Bell's inequalities, a loophole means that the observed statistics may in fact be due to

classical communication ("locality loophole") or to shared randomness ("detection loophole"). If a loophole is not closed, the violation of Bell's inequalities can be produced by trivial classical means, for instance two suitably programmed and possibly causally connected computers.

It is widely accepted that the loopholes must be closed if the devices are not fully characterized. Experimentalists, however, assume to have an *a priori* knowledge that their data are not generated by a pair of conspiring computers, and may be tempted to adopt a more relaxed stance. A warning against this complacency was issued when an unplanned postselection of data led to a violation of Bell's inequalities larger than theoretically possible [11]. Bell experiments, in which one photon is sent through an optical amplifier before detection [12] are another instance where nonclosed loopholes can lead to misinterpretation of results. Here we discuss another scenario: using faked-state techniques [1,13–15], we show a fake violation of Bell's inequalities in the same conditions in which ordinary tests with genuine entangled states are performed, by exploiting the physics of single-photon detectors. In particular, two of our experimental demonstrations feature no postselection in a passive-choice scheme, and a third demonstration features active choices, with a postselection that keeps half of the data (more than any Bell experiment reported to date). The takeaway message is that, when dealing with a malicious adversary, it is necessary to close all the loopholes, even if the measurement devices are known to the user and seem to behave normally.

Experimental Bell test.—With a typical setup to generate entangled states we carry out polarization correlation measurements (Fig. 1). A source of polarization-entangled photon pairs based on parametric down conversion [16] feeds each member of the pair to two legitimate parties, Alice and Bob, who measure the polarization in one of two

possible bases using a beam splitter as a random basis choice, and polarizers followed by avalanche single-photon detectors (APDs).

The test measures photon polarization correlations between two parties in two bases,

$$S = E_{AB} + E_{A'B} + E_{AB'} - E_{A'B'}, \qquad (1)$$

where A, A' and B, B' correspond to the measurements at Alice's and Bob's side, respectively, and E_{AB} is the correlation function

$$E = \frac{P(1,1) - P(1,0) - P(0,1) + P(0,0)}{P(1,1) + P(1,0) + P(0,1) + P(0,0)}, \qquad (2)$$

where P is the probability of a joint outcome. The Clauser-Horne-Shimony-Holt (CHSH) inequality [17] states that any classical state will have $|S| \le 2$. Quantum mechanical states can produce up to $|S| \le 2\sqrt{2}$, and the mathematical maximum value is $|S| = 4$. Note that this expression already postselects only cases where pairs of detection events were seen on both sides—it excludes the cases where one of the photons was lost due to inefficiencies in the detectors, source, or transmission path.

Bell's inequality with an eavesdropper.—In between the source of entangled photons and Bob, we set up another observer, Eve. She has a measuring apparatus identical to Bob's, and a "faked-state generator" (FSG), which takes advantage of the detailed physical mechanism behind APDs to manipulate their output. When in normal operation, the arrival of a photon on an APD creates an electron-hole pair, which gets separated by an applied voltage and in the process creates an avalanche resulting in a macroscopic current; when the current exceeds a certain threshold, a "click" is registered and interpreted as the arrival of a photon. By sending tailored strong pulses of light into an APD, it is possible to manipulate the output such that "clicks" appear at Eve's will [13,15]: Beyond some optical power level, the detector does not have time to recharge and is constantly producing small avalanches that do not fulfill the threshold condition, and can thus be blinded or saturated. At even higher powers, the number of electron-hole pairs produced is such that the current can exceed the threshold condition set by the electronics without the need for an avalanche. By tailoring the polarization and power of the optical pulses sent into Bob's apparatus, it is possible to force him to obtain any desired measurement result.

An incoherent mixture of laser beams of two different polarizations allows the FSG [Fig. 1(a)] to address any of the four detectors. One of the beams is circularly polarized (σ), and thus gets distributed evenly among all detectors at a blinding intensity level. The other beam is linearly polarized in the same direction as the desired measurement result. The beam splitter and polarization optics in the measurement setup will direct a larger fraction of the latter beam to the targeted APD than to any of the other APDs, driving only the targeted APD into the faking regime and
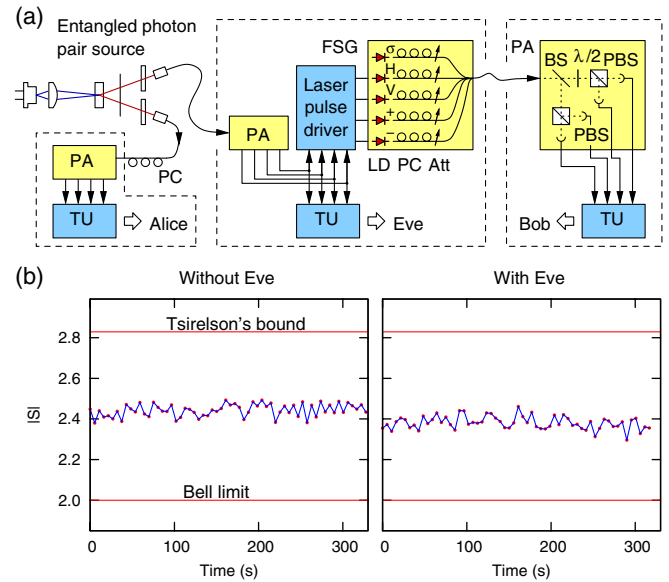

FIG. 1 (color online). Testing a Bell inequality in an intercept-resend scenario. (a) Experimental setup. Pairs of polarization-entangled photons are generated via spontaneous parametric down-conversion and sent for polarization analysis (PA) in two conjugate linearly polarized bases on both sides using polarizing beam splitters (PBS), half-wave plates ($\lambda/2$), and nonpolarizing beam splitters (BS); details are described in Ref. [24]. An intercept-resend system (Eve) is inserted in the fiber line that carries one of the photons. Eve consists of a PA and faked-state generator (FSG), and generates measurement results that are copied into the receiver Bob. Key elements of the FSG are laser diodes (LD), polarization controllers (PC), and attenuators (Att); see details in Ref. [15]. Detection events are recorded with time-stamp units (TU) for later correlations. (b) From the correlations between the measurement results, Alice and Bob can test a CHSH Bell inequality. With the photodetectors used for this experiment, they obtain the same result (i.e., a violation of a Bell inequality) with and without the presence of Eve, and would conclude that they witnessed entanglement in a pair of photons.

forcing a false photon detection (see [15] for experimental details). The two parties testing the inequality would be registering clicks at the respective detectors and assembling joint probability distributions of the possible simultaneous results. While Alice is measuring genuine photons, Bob is looking at a classical pulse of light. A test of the CHSH inequality is shown in Fig. 1(b). As this is an "intercept-resend" configuration, the Bell value registered will be a reflection of the entanglement quality of the original source. In this case the visibility in the $\pm$ basis is around 92%, and results in a value of $S = 2.381 \pm 0.036$. This is well above the classical limit, highlights the possible breakdown of the Ekert quantum key distribution protocol [18], and the need for strictly fulfilling the assumptions associated with the Bell theorem in a cryptography scenario.

Preprogrammed Bell violation.—Since we can fake any result, there is no need for the entanglement source, so we

(a) Alice ... Bob
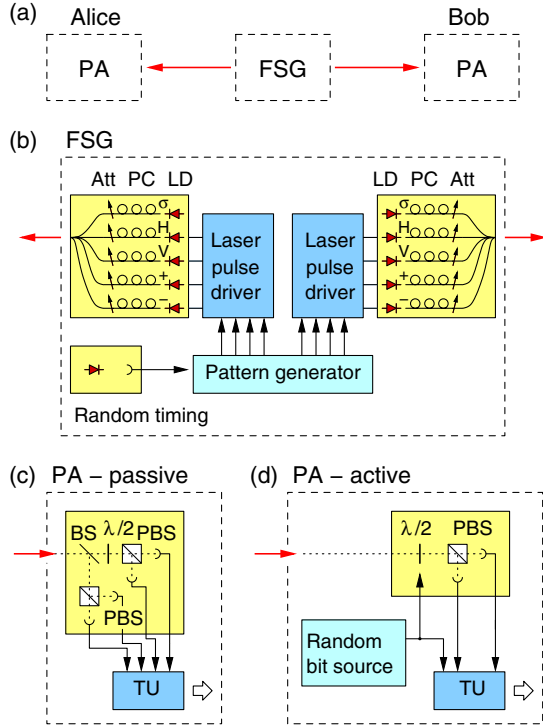
(b) FSG

(c) PA – passive   (d) PA – active

FIG. 2 (color online). Bell test with the source replaced by a twin FSG. (a) Optical outputs of the FSG are connected to two polarization analyzers at Alice and Bob. (b) Implementation of the twin FSG: A preprogrammed list of correlation events is sent to two single FSGs at random times defined by a photocounting detector illuminated with a constant intensity. (c) Polarization analyzer with a passive choice of measurement basis, and (d) with an active choice, where a half-wave plate ($\lambda/2$) is rotated by a motor according to the value of an independent random bit source.

replace it with a pair of FSGs [Fig. 2(a)]. The polarizations expected by Alice are denoted (H, V, $+$, $-$); those expected by Bob are rotated by 22.5° compared to Alice's and are denoted ($\tilde{H}$, $\tilde{V}$, $\tilde{+}$, $\tilde{-}$). The frequencies $p_{P_A P_B}$ with which the FSG sends out polarization P_A to Alice and P_B to Bob are programmed to be

$$\begin{pmatrix} p_{H\tilde{H}} & p_{H\tilde{V}} & p_{H\tilde{+}} & p_{H\tilde{-}} \\ p_{V\tilde{H}} & p_{V\tilde{V}} & p_{V\tilde{+}} & p_{V\tilde{-}} \\ p_{+\tilde{H}} & p_{+\tilde{V}} & p_{+\tilde{+}} & p_{+\tilde{-}} \\ p_{-\tilde{H}} & p_{-\tilde{V}} & p_{-\tilde{+}} & p_{-\tilde{-}} \end{pmatrix} = \begin{pmatrix} x & y & x & y \\ y & x & y & x \\ x & y & y & x \\ y & x & x & y \end{pmatrix}, \quad (3)$$

where $x = (1 + q)/16$, $y = (1 - q)/16$, and q varies from -1 to $+1$. Such correlations lead to a CHSH violation of $S = 4q$: in particular, Tsirelson's bound [19] is recovered for $q = 1/\sqrt{2}$, while the choice $q = \pm 1$ simulates the unphysical Popescu-Rohrlich box [20].

For each value of q, a sequence of bright faked-state pairs is generated by a random number generator and the twin FSG [Fig. 2(b)]. Pairs are sent out upon photodetection events of an independent illuminated photodetector, to mimic the detection times of light emitted from the sponta-
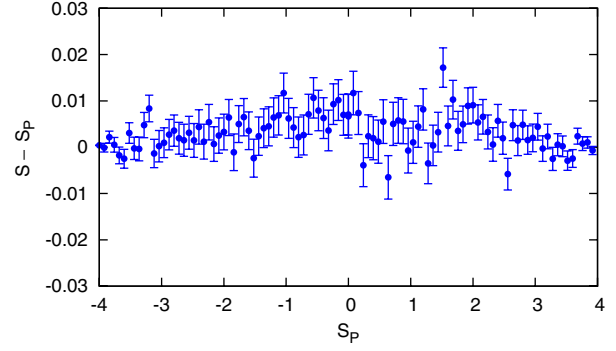

FIG. 3 (color online). Difference between observed correlation function S and expected S_p in a CHSH Bell test, as a function of the programmed value $S_p = S(q) = 4q$. The error bars were assigned through propagated Poissonian counting statistics.

neous parametric down-conversion source. The experimental errors evaluated assuming Poisson counting statistics are on the order of 4×10^{-3} when averaging over about 700 thousand events per q value (Fig. 3). Moreover, because the pulses of light can be made bright enough to overcome any channel losses, this configuration allows faking of correlations with efficiencies of 100%, meaning that there are no "unpaired" events: every time Alice registers a click, so does Bob.

The fixed nature of the analysis optics means that it is possible for the eavesdropper to use the same degree of freedom that is used to guide the genuine photons to the right detector (i.e., polarization) in a malicious way to target her control pulses. Note that, although the random choice of measurement basis for an arriving photon is a genuine one [21], its unchanging status results in an attack vector. So, is the problem solved just by having an active choice of basis?

Active basis choice.—We now come to the third experiment, where the measurement basis is chosen based on a random setting causally connected to neither the source nor the measurement apparatus. For this, we replace the BS by a half-wave plate that is rotated by a motor between two positions, determined by the parity of photodetection events from an illuminated, independent detector [see Fig. 2(d)]. The faking scheme will still work, once intensity levels are adjusted, but the immediate consequence is a decrease in efficiency from 100% to 50%, since a basis choice not matched to the faked state results in an absence of the detection event. Efficiency here means number of paired detection events over total detected events, on each side. Although this reduction is dramatic, we are not aware of any experiment with entangled photons that has demonstrated an efficiency this high (50%) so far. A correlation matrix between the prepared faked-state pairs and the combination of detection events and measurement bases on both sides is shown in Fig. 4. Tiny imperfections in reproducing programmed coincidences (mean of unwanted clicks of 0.07%) result from using a simplified version of
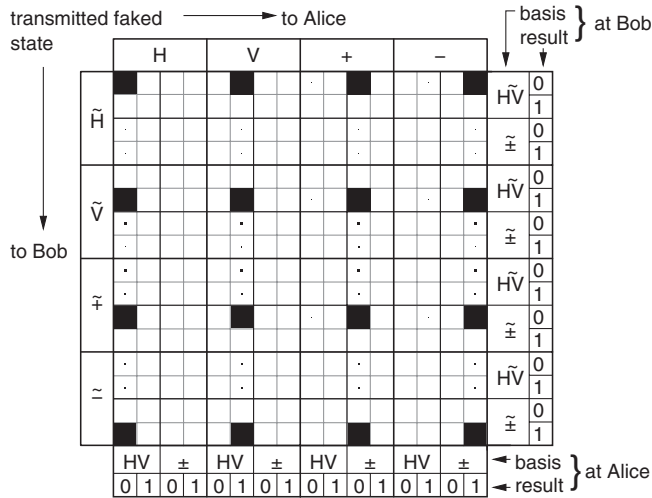
FIG. 4. Experimentally obtained coincidences between pairs of detection events (0 or 1) at receivers Alice and Bob for their different measurement bases (HV or $\pm$) as a function of different faked-state pairs of bright pulses sent to both sides. Each faked-state pair is made up by a combination of two pulses for horizontal (H), vertical (V), and $\pm 45°$ linear polarization ($+$ and $-$). Since bases for Alice and Bob are rotated in a Bell test by $22.5°$, we denote the faked-state polarizations and measurement bases on Bob's side with a tilde. Each cell's area is filled proportionally to the number of coincidence events; empty cell corresponds to 0 coincidence events, while a completely filled one corresponds to 40 229 coincidences. The level of unwanted clicks is always below 1.4% (≤ 566 coincidences).

the twin FSG, with a single bright-light source per polarization [15]. Again, this matrix demonstrates full controllability of pair correlations even in an active basis choice scenario. Using this with a randomized distribution of faked coincidences corresponding to the Bell correlations, we obtain a value of $S = 2.7971 \pm 0.0005$ in a run with 4561 basis settings (motor positions) and about 7.25×10^6 coincidences. In principle, all values up to $|S| = 4$ can be generated at will. Thus, an active choice by itself is still not sufficient to prevent a Bell inequality violation unless it is paired with an efficiency high enough to avoid the need for the fair sampling assumption.

Conclusion.—The strength of Bell inequalities as entanglement witnesses is their formulation independently of the underlying physical model and measurement processes. Once we have an operational definition of what constitutes a measurement, such as the click mentioned repeatedly, there is a prescription to determine whether correlations between measurement results could be explained by a classical model. The results in this Letter highlight how important the underlying assumptions in the theorem are to the interpretation of results. We have shown that ignoring the need for active choice and causal separation might result in an apparent violation with 100% efficiency. Introducing an active choice scenario without consideration for the possible statistical bias is also not sufficient,

and we achieve again an apparent violation beyond what has been achieved with genuine photonic entangled states. Our experiment shows that it is not legitimate to assume that a sampling of detection events is fair. Both loopholes need to be closed simultaneously for interpreting the violation of a Bell inequality as a refutation of local variables.

While the detector control method used in this work could be defeated, e.g., by monitoring optical power at the analyzer entrance [15,22], such additional monitoring is neither a part of the Bell theorem nor of the security proofs for quantum cryptography. Furthermore, by trading off some detection efficiency, a control method requiring much less light (tens of pW) can be used [13], which would demand close to single-photon sensitivity of the power meter. The feasibility of implementing a countermeasure once the method of attack is explicit, does not detract from the conclusions, but rather highlights the need to carefully scrutinize any assumptions that are not part of the protocol.

We have focused here on Bell's inequalities, but it is evident that implementation of ''self-testing'' [9,10] and even standard entanglement witnessing [23] needs some critical rethinking along the same lines.

*makarov@vad1.com
†christian.kurtsiefer@gmail.com

[1] V. Makarov and D. R. Hjelme, J. Mod. Opt. **52**, 691 (2005).
[2] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[3] A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[4] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).
[5] S. Pironio *et al.*, Nature (London) **464**, 1021 (2010).
[6] C.-E. Bardyn, T. C. H. Liew, S. Massar, M. McKague, and V. Scarani, Phys. Rev. A **80**, 062327 (2009).
[7] J.-D. Bancal, N. Gisin, Y.-C. Liang, and S. Pironio, Phys. Rev. Lett. **106**, 250404 (2011).
[8] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner, and V. Scarani, Phys. Rev. Lett. **107**, 050502 (2011).
[9] D. Mayers and A. Yao, Quantum Inf. Comput. **4**, 273 (2004).
[10] M. McKague, arXiv:1006.2352.
[11] D. S. Tasca, S. P. Walborn, F. Toscano, and P. H. Souto Ribeiro, Phys. Rev. A **80**, 030101(R) (2009).
[12] E. Pomarico, B. Sanguinetti, P. Sekatski, H. Zbinden, and N. Gisin, New J. Phys. **13**, 063031 (2011).
[13] V. Makarov, New J. Phys. **11**, 065003 (2009).
[14] L. Lydersen, M. K. Akhlaghi, A. H. Majedi, J. Skaar, and V. Makarov, arXiv:1106.2396.

[15] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, Nature Commun. **2**, 349 (2011).

[16] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, Phys. Rev. Lett. **75**, 4337 (1995).

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

[18] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[19] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).

[20] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[21] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger, Rev. Sci. Instrum. **71**, 1675 (2000).

[22] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, Nat. Photon. **4**, 686 (2010).

[23] P. Skwara, H. Kampermann, M. Kleinmann, and D. Bruß, Phys. Rev. A **76**, 012312 (2007).

[24] I. Marcikic, A. Lamas-Linares, and C. Kurtsiefer, Appl. Phys. Lett. **89**, 101122 (2006).