



OPEN

Bright-light detector control emulates the local bounds of Bell-type inequalities

Shihan Sajeed^{1,2,3,4✉}, Nigar Sultana^{1,3}, Charles Ci Wen Lim^{5,6} & Vadim Makarov^{7,8,9}

It is well-known that no local model—in theory—can simulate the outcome statistics of a Bell-type experiment as long as the detection efficiency is higher than a threshold value. For the Clauser–Horne–Shimony–Holt (CHSH) Bell inequality this theoretical threshold value is $\eta_T = 2(\sqrt{2} - 1) \approx 0.8284$. On the other hand, Phys. Rev. Lett. 107, 170404 (2011) outlined an explicit practical model that can fake the CHSH inequality for a detection efficiency of up to 0.5. In this work, we close this gap. More specifically, we propose a method to emulate a Bell inequality at the threshold detection efficiency using existing optical detector control techniques. For a Clauser–Horne–Shimony–Holt inequality, it emulates the CHSH violation predicted by quantum mechanics up to η_T . For the Garg–Mermin inequality—re-calibrated by incorporating non-detection events—our method emulates its exact local bound at any efficiency above the threshold. This confirms that attacks on secure quantum communication protocols based on Bell violation is a real threat if the detection efficiency loophole is not closed.

More than 50 years ago, John Stewart Bell showed that any physical theory based on the assumptions of locality (i.e., nothing can communicate faster than light) and realism (i.e., physical properties of an object are fixed and pre-defined) must satisfy a set of statistical criteria called Bell inequalities¹. That is, if a Bell-type experiment is performed and the results show a violation of a Bell inequality, then the underlying physical process cannot be explained by a local theory. This kind of tests are called Bell tests and the violation of the inequality is called Bell violation. Since the earlier demonstrations utilizing cascade decays in atoms^{2–5}, Bell violations have been observed in tests utilizing nonlinear optical processes^{6–9}, ions¹⁰, neutral atoms¹¹, Josephson junction¹² and solid state qubits¹³. The implications of the Bell test not only change our understanding of nature, but also find application in device independent (DI) quantum communications^{14–16}, randomness generation and amplification^{17–19}, DI-verified quantum computation^{20,21}, certifying quantum devices^{18,22,23} and DI bit commitment²⁴. Entanglement, a necessary precondition for unconditional security^{25,26} in quantum key distribution, can also be certified from the violation of a Bell inequality, independently of the underlying implementation details. This paves the way for the device-independent tests of security^{27,28}. However, for the observed Bell violation to be conclusive, it is important that the Bell test is loophole-free.

More specifically, a loophole-free Bell test is an entanglement experiment that requires multiple implementation loopholes such as the detection, locality, and measurement-independent loopholes to be closed simultaneously. Here, we focus on the detection loophole, and defer the rest to Ref.²⁹. In general, the detection loophole is a scenario in which the observed Bell violation (a test statistics) is no longer reliable as the measurement sample and may not be a true representative of the population (i.e., the entire measurement statistics). Crucially, this situation commonly happens in practice as practical detectors have finite detection efficiencies and hence one could

¹Institute for Quantum Computing, University of Waterloo, Waterloo, ON N2L 3G1, Canada. ²Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada. ³Department of Electrical and Computer Engineering, University of Waterloo, N2L 3G1 Waterloo, ON, Canada. ⁴Department of Electrical and Computer Engineering, University of Toronto, M5S-3G4 Toronto, ON, Canada. ⁵Centre for Quantum Technologies, National University of Singapore, Singapore, Singapore. ⁶Department of Electrical & Computer Engineering, National University of Singapore, Singapore, Singapore. ⁷Russian Quantum Center, Skolkovo, Moscow 121205, Russia. ⁸Shanghai Branch, National Laboratory for Physical Sciences at Microscale and CAS Center for Excellence in Quantum Information, University of Science and Technology of China, Shanghai 201315, People's Republic of China. ⁹NTI Center for Quantum Communications, National University of Science and Technology MISiS, Moscow 119049, Russia. ✉email: shihan.sajeed@gmail.com

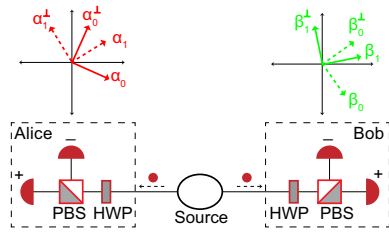


Figure 1. Setup for a CHSH test. The measurement angles shown are arbitrary. Here, for each party, the set of possible outcomes is given by $\{+, -, ?\}$, where “+” (“-”) are assigned when only the lower (upper) detector clicks and the other detector is silent. “?” is assigned when none of the detectors registers a detection.

end up with samples that are non-representative. While the detection loophole is not an issue for non-adversarial settings, the same is not true for the case of quantum cryptography since an adversary can take advantage of it to come up with a local model to *fake* Bell violations³⁰. For this reason, much effort has been devoted to closing the detection loophole in practice.

How a local model can theoretically simulate non-local correlations—taking advantage of the detection loophole—has already been reported in the literature^{31,32}. However, methods of experimentally implementing such correlations using practical means have been rarely discussed, despite its importance in practical quantum cryptography. The state-of-the-art method is arguably that of Ref.³⁰, where the authors demonstrated how an adversary could implement a local model using existing optical detector control methods to violate a Bell inequality for active basis choice schemes. However, their local model is effective only for a detector efficiency of up to $\eta = 0.5$, while theoretically it is possible to fake the inequality for a threshold efficiency of up to $\eta_T = 2(\sqrt{2} - 1) \approx 0.8284$ (here, efficiency η refers to the probability that one party observes a conclusive outcome given a measurement is made). In this article, we discuss how to experimentally close this gap and fake the violation at higher efficiencies. More specifically, we show how existing optical detector control methods^{33–35} can be exploited to both fake the violation of the standard Clauser–Horne–Shimony–Holt (CHSH) Bell test all the way up to its threshold efficiency and simulate the local bound of the more general Garg–Mermin Bell test. Our results point out once again that when Bell tests are performed for certifying randomness, guaranteeing security in quantum communications, or detecting non-locality, they should either be performed with an efficiency at which the test is robust against detection loopholes, or should use the bound given by more general inequalities (for example, Eq. (3) below). Otherwise, existing optical detection control methods may allow to implement a local model to simulate the results of the test.

The article is organized as follows. First we outline the assumptions and methodology of the Bell test that we consider in this article. Then we present several local models that allow an adversary to implement a practical setup to fake the Bell test or emulate the local bounds given by the inequalities. Then we make our conclusion.

Assumptions for Bell test

The experimental setup of the CHSH Bell test for two parties with binary inputs and outputs³⁶ is shown in Fig. 1. The test assumes that a source of entangled photon pairs sends each member of the pairs to two legitimate parties, Alice and Bob. Alice randomly measures the polarizations along direction α_0 or α_1 and Bob randomly along β_0 or β_1 as shown in Fig. 1. The measurement along a particular direction is performed with the help of a rotatable half wave plate (HWP) followed by a polarization beamsplitter (PBS) and two single photon detectors. This type of analyzer is called an active basis choice analyzer. The possible polarization measurement outcomes expected at Alice and Bob are $P_A \in \{\alpha_0, \alpha_0^\perp, \alpha_1, \alpha_1^\perp\}$ and $P_B \in \{\beta_0, \beta_0^\perp, \beta_1, \beta_1^\perp\}$, and they are mapped into outcomes $\{+, -, ?\} \times \{+, -, ?\}$; see Fig. 1 for outcome assignments.

We assume that Alice and Bob are situated far apart, so that the locality loophole is closed. However, due to the finite efficiency of the detectors and optical losses in the setup, it is not possible to measure the polarization of all the photons. So, the final statistics are calculated using only the detected events, i.e., events in which photons have been detected on both sides. In this case, for each pair of measurement settings $\{\alpha_i, \beta_j\}$ with $ij \in \{00, 01, 10, 11\}$ chosen by Alice and Bob, the correlation function $E(\alpha, \beta)$ is given by

$$E(\alpha, \beta) = \frac{N_{\alpha,\beta}(++) + N_{\alpha,\beta}(--) - N_{\alpha,\beta}(+-) - N_{\alpha,\beta}(-+)}{N_{\alpha,\beta}(++) + N_{\alpha,\beta}(--) + N_{\alpha,\beta}(+-) + N_{\alpha,\beta}(-+)}, \tag{1}$$

where $N_{\alpha,\beta}(i, j)$ represents the number of coincidences with successfully detected outcome $\{i, j\} \in \{++, +-, -+, --\}$ for a particular setting (α, β) . The associated CHSH Bell inequality is then

$$S_{\text{CHSH}} = E(\alpha_0, \beta_0) + E(\alpha_1, \beta_0) + E(\alpha_1, \beta_1) - E(\alpha_0, \beta_1) \leq 2. \tag{2}$$

Quantum mechanics predicts a maximum violation of $S = 2\sqrt{2}$ for the setting choice $\alpha_0 = -78.75^\circ, \alpha_1 = 56.25^\circ, \beta_0 = 11.25^\circ, \beta_1 = -33.75^\circ$ ³⁷, and even stronger correlations are algebraically possible in theory leading to $S \leq 4$ ³⁸. However, as long as the efficiency of a measurement is $\eta = 1$, all local models must necessarily satisfy Eq. (2). Unfortunately, this is not true for $\eta < 1$. In particular, when η is less than some threshold η_T , it is possible to devise local models that violate Eq. (2). For the CHSH test described here, $\eta_T = 2(\sqrt{2} - 1) \approx 82.84\%$ ³⁹. In order to avoid this, these tests are performed in the region $\eta > \eta_T$. Note that the CHSH test is not the most robust Bell test as one

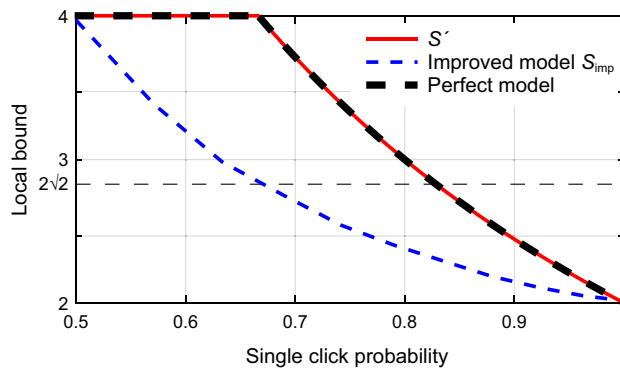


Figure 2. Local bounds for recalibrated inequality S' [Eq. (3)], improved faking model [Eq. (6)], and perfect faking model. The quantum mechanical bound $2\sqrt{2}$ is also shown. The improved faking model achieves this bound at $\eta \approx 0.6678$ and the perfect model at $\eta = 2(\sqrt{2} - 1)$. The perfect model can fully emulate Eq. (3) for efficiency range $2/3 \leq \eta \leq 1$.

can further reduce the detection threshold by looking at marginal correlations (or singles statistics). This is given by the Eberhard Bell inequality³⁷, which has a detection threshold of $\eta_T = 2/3 \approx 66.67\%$. Alternatively, one can include the ‘efficiency’ in the inequality and recalibrate it as a function of η as proposed by Garg and Mermin³⁹

$$\begin{aligned} S'(\eta) &= E(\alpha_0, \beta_0) + E(\alpha_1, \beta_0) + E(\alpha_1, \beta_1) - E(\alpha_0, \beta_1) \\ &\leq \frac{4}{\eta} - 2. \end{aligned} \quad (3)$$

The recalibrated CHSH Bell inequality gives the local bound of S' as a function of η , i.e., how much violation is required to certify non-locality for a given efficiency. This is shown by the solid (red) curve in Fig. 2. Note that, when $\eta = 1$, Eq. (3) becomes Eq. (2) since the post-selected correlation set becomes the entire measurement set. Also, when $\eta \leq 2/3$, one can always set the bound to be 4, which is the maximum value attainable by the sum of four correlation functions. Thus, a local model that can simulate Eq. (3) for efficiency range $2/3 \leq \eta \leq 1$ would be the optimum model to exploit detection loopholes in a Bell test. We present it in the next section.

Faking Bell inequality with improved efficiency

For ease of understanding, we will go step by step. First, we review an existing local model that can fake Eq. (2) for $\eta \leq 1/2$ ³⁰ and point out its limitations. Then we propose a modification to this model that enables it to fake Eq. (2) up to $\eta \leq 2/3$. We then present our perfect model that can not only fake Eq. (2) for $\eta \leq 2(\sqrt{2} - 1)$ but also emulate the local bounds given by Eq. (3). Since all three models exploit an existing detector control method—bright-light detector control^{33–35}—we first recap it.

Bright-light detector control: Single-photon detectors used in a Bell test may become insensitive to single photons when exposed to bright light^{33,40}. Even in this mode, they can produce a detection event (‘click’) when additionally exposed to a light pulse of intensity I equal to or higher than a threshold level I_{th} . This allows an adversary Eve to have control over the detectors by tailoring I . For example, if the measurement basis matches that of the incoming light pulse, then all of it is incident on a single detector with intensity $I \geq I_{th}$ and results in a detection event. However, in case of basis mismatch, the incoming light is split between two detectors with intensity $I/2 < I_{th}$ (assuming a conjugate basis) and none of the detectors click. This is how the adversary can have control over detection outcomes. The feasibility of bright-light control has been confirmed numerous times, with both detectors based on avalanche photodiodes^{33,34,41–47} and superconducting nanowires^{40,48,49}. Next, we show how an adversary can exploit it to implement a local-realistic model.

Conditions for violation: Let us assume that an arbitrary value of $|S| \leq 4$ needs to be simulated by the local model. Assuming symmetry for each setting combination (α, β) , this implies $|E| = S/4$. Assuming $N_{\alpha,\beta}(++) = N_{\alpha,\beta}(--) = N_{sim}$ and $N_{\alpha,\beta}(+-) = N_{\alpha,\beta}(-+) = N_{dif}$, (where $2N_{sim} + 2N_{dif} = 1$), Eq. (1) can be written as

$$\frac{N_{sim}}{N_{dif}} = \frac{1 + E}{1 - E}. \quad (4)$$

This implies that under the assumptions specified above, an arbitrary correlation value E requires the ratio of similar to different outcomes to follow Eq. (4). For example, the quantum mechanical prediction of $S = 2\sqrt{2}$, which corresponds to $E = \pm 1/\sqrt{2}$, requires

$$N_{sim} = (3 \pm 2\sqrt{2})N_{dif}. \quad (5)$$

Below we describe several techniques by which an active attacker can satisfy this condition.

Existing model: A straightforward approach to force the outcomes to follow Eq. (4) is to generate polarization combinations at the source with desired statistics and then force deterministic outcomes during the measurement,

Towards Alice	Towards Bob			
	β_0	β_0^\perp	β_1	β_1^\perp
α_0	$N_{\text{sim}}/4$	$N_{\text{dif}}/4$	$N_{\text{dif}}/4$	$N_{\text{sim}}/4$
α_0^\perp	$N_{\text{dif}}/4$	$N_{\text{sim}}/4$	$N_{\text{sim}}/4$	$N_{\text{dif}}/4$
α_1	$N_{\text{sim}}/4$	$N_{\text{dif}}/4$	$N_{\text{sim}}/4$	$N_{\text{dif}}/4$
α_1^\perp	$N_{\text{dif}}/4$	$N_{\text{sim}}/4$	$N_{\text{dif}}/4$	$N_{\text{sim}}/4$

Table 1. Probability of each polarization combination generated by the source in the existing faking model³⁰. They are normalized to maintain $2N_{\text{sim}} + 2N_{\text{dif}} = 1$.

as done in Ref.³⁰. We assume each polarization combination is generated according to the probabilities given in Table 1, where N_{sim} and N_{dif} obey Eq. (4). We assume the intensity is tailored to bring the bright-light control method into play, i.e., matched (mismatched) bases lead to deterministic outcome with unity probability (no detection). Let's consider the case when the source generates polarization combination $\alpha_0\beta_0$ ($\alpha_0\beta_1$) with probability $N_{\text{sim}}/4$ ($N_{\text{dif}}/4$). They result in coincidences only for the setting $\alpha_0\beta_0$ ($\alpha_0\beta_1$) and lead to deterministic similar (similar) outcomes with unity probability. For the remaining three setting choices, no coincidence happens and the outcomes have no effect on the correlation. This is true for all the polarization combinations in Table 1. In this way, it is possible to generate outcomes to match Eq. (4) for any desired value of E and achieve any value of S up to $S_1 = 4$. A problem with this method, however, is that half of the time the measurement basis does not match the preparation basis and results in no detection. Thus the efficiency at each side $\eta_1 = 0.5$. This is a limitation in Ref.³⁰. Next, we outline how to implement an improved local realistic model with a higher detection efficiency.

Improvement to existing model: Above we have recapped the existing first method that leads to CHSH parameter $S_1 = 4$ with an efficiency $\eta_1 = 0.5$. We now generate a second method that leads to CHSH parameter $S_2 = 2$ with efficiency $\eta_2 = 1$. For this, let's assume that the source always sends polarization α (β) to Alice (Bob), where α (β) is polarized at an angle that is midway between α_0 and α_1 (β_0 and β_1). In this case, irrespective of the measurement settings, the input intensity I is split at a ratio of $\cos^2(\phi_A) : \sin^2(\phi_A)$ between the two detectors in Alice and at $\cos^2(\phi_B) : \sin^2(\phi_B)$ in Bob. Here, $\phi_A = |\alpha_1 - \alpha_0|/2$ and $\phi_B = |\beta_1 - \beta_0|/2$. Tailoring the intensity to satisfy $I \cos^2(\phi) \geq I_{\text{th}}$ and $I \sin^2(\phi) < I_{\text{th}}$ at the respective sides ensures that only one of the detectors clicks (with outcome +), irrespective of the basis choice, and efficiency stays 1. This will result in $E = +1$ for each measurement setting and lead to a CHSH parameter $S_2 = 2$ with an efficiency $\eta_2 = 1$. Note that this method (presented here for its ease of explanation) results in only ++ outcomes. It can be symmetricized to produce all four outcomes ++, +-, -+, --, which we omit for brevity.

Thus, we have outlined two independent approaches to control S : the first one leads to $S_1 = 4$ with an efficiency $\eta_1 = 0.5$, while the second one leads to $S_2 = 2$ with efficiency $\eta_2 = 1$. An adversary can then use a probabilistic mixture of these two approaches to increase the faking efficiency of the Bell test. With probability p_1 ($p_2 = 1 - p_1$) she uses the first (second) method. The input intensity needs to be tailored to $2I_{\text{th}} > I \geq I_{\text{th}}/\cos^2(\phi)$ to ensure that the first (second) method leads to detection efficiency of $\eta_1 = 0.5$ ($\eta_2 = 1$) and results in $S_1 = 4$ ($S_2 = 2$). The resultant efficiency as seen by Alice and Bob will be $\eta = \sqrt{p_1\eta_1^2 + p_2\eta_2^2}$ and the improved CHSH parameter will be

$$S_{\text{imp}} = \frac{p_1 S_1 \eta_1^2 + p_2 S_2 \eta_2^2}{\eta^2}. \quad (6)$$

The variation of S_{imp} with η is shown in Fig. 2. The left-most point $(\eta, S_{\text{imp}}) = (0.5, 4)$ corresponds to the first method with $p_2 = 0$. As p_2 is increased, S_{imp} becomes smaller with increasing efficiency and eventually becomes $(\eta, S_{\text{imp}}) = (1, 2)$ at the rightmost point with $p_2 = 1$. Quantum mechanical prediction $S = 2\sqrt{2}$ is obtained at $p_2 \approx 0.2612$ and the corresponding efficiency is $\eta \approx 0.6678$. This is still lower than the threshold efficiency limit $\eta_T = 2(\sqrt{2} - 1) \approx 0.8284$ for CHSH inequality. To achieve higher local bounds, one more degree of freedom needs to be introduced, as discussed in our next model.

Perfect local model: Now we present a perfect local model that can not only fake a violation of inequality (2) for $\eta \leq 2(\sqrt{2} - 1)$ but also emulate the local bounds given by Eq. (3) for $2/3 \leq \eta \leq 1$. For this model, we make three assumptions:

1. The adversary at the source always generates one of the two polarization combinations $\alpha_0\beta_0$ and $\alpha_1\beta_1$ with equal probability of 1/2 each.
2. The adversary tailors the light intensity towards Bob in such a way that they result in a deterministic outcome with unity probability. For the ease of analysis we will assume that at Bob, the polarization β_0 (β_1) leads—with unity efficiency—to deterministic outcome “+” (“+”) when measured along β_0 and “-” (“+”) when measured along β_1 (however, any other outcomes will also do as long as they are deterministic and have unity probability).
3. At Alice, whenever the measurement basis matches (does not match) that of the incoming light, a deterministic “+” (random) outcome is produced with probability a (b), and no-detection outcome with probability $1 - a$ ($1 - b$).

Polarization emitted from source	Measurement outcome	Joint probability at measurement setting			
		$\alpha_0\beta_0$	$\alpha_1\beta_0$	$\alpha_0\beta_1$	$\alpha_1\beta_1$
$\alpha_0\beta_0$	++	a	$b/2$	0	0
	+-	0	0	a	$b/2$
	-+	0	$b/2$	0	0
	--	0	0	0	$b/2$
	?+	$1-a$	$1-b$	0	0
	?-	0	0	$1-a$	$1-b$
$\alpha_1\beta_1$	++	$b/2$	a	$b/2$	a
	+-	0	0	0	0
	-+	$b/2$	0	$b/2$	0
	--	0	0	0	0
	?+	$1-b$	$1-a$	$1-b$	$1-a$
	?-	0	0	0	0

Table 2. Possible outcomes and the corresponding probabilities for different measurement settings in the perfect model. Outcome $ij \in \{+, -, ?\} \times \{+, -, ?\}$ represents i at Alice and j at Bob. It can be verified that the conditional probability distributions are no-signalling⁵⁰.

For details on how the above assumptions can be met in practice, please see the Methods. For each setting, the possible outcomes at Alice and Bob and the corresponding coincidence probabilities are shown in Table 2. For any measurement setting $\{\alpha, \beta\}$, the correlation function E is related to a and b as

$$|E| = \frac{\frac{a}{2} + \frac{b}{4} - \frac{b}{4}}{\frac{a}{2} + 2\frac{b}{4}} = \frac{a}{a+b}, \quad (7)$$

and the coincidence probability is

$$\frac{a}{2} + \frac{b}{2} = \eta^2. \quad (8)$$

Solving Eqs. (7) and (8), we get,

$$\begin{aligned} a &= 2E\eta^2 \\ b &= 2(1-E)\eta^2 \end{aligned} \quad (9)$$

Thus, to emulate the local bounds in an actual experiment having detector efficiency η , an adversary can use Eq. (3) to calculate the maximum correlation value E corresponding to that η , and then use Eq. (9) to set the values of a, b . As long as Eq. (9) is maintained, the single click probability during the test is equal to η and the CHSH value is equal to the bound as shown by the thick black dashed line in Fig. 2. For example, for a Bell test done with detector efficiency $\eta = 2(\sqrt{2} - 1)$, the local bound is $S' = 2\sqrt{2}$ according to Eq. (3). This can be attained—according to Eq. (9)—if $a = 12\sqrt{2} - 16 = 0.97$ and $b = 40 - 28\sqrt{2} = 0.40$, which leads to $|E| = 1/\sqrt{2}$. Similarly, the local bound of S' [Eq. (3)] can be achieved for any $2/3 \leq \eta \leq 1$. For $\eta \leq 2/3$, the local bound reaches the algebraic maximum $S' = 4$ [Eq. (3)]. Note that this method leads to asymmetric detection efficiency, as Bob's efficiency is always higher than Alice's. However, this can be avoided by reversing the roles of Alice and Bob half of the time. This concludes our local model that can emulate the local bounds given by Eq. (3) for every value of $2/3 \leq \eta \leq 1$.

Conclusion

Although it is a known fact that a local theory can violate a Bell inequality up to a threshold detection efficiency, it is rarely addressed in the literature how an adversary can actually implement it. In this work, we have shown that the existing detector control method can be exploited to implement a local model that can fake the CHSH Bell inequality [Eq. (2)] up to the threshold efficiency. Our model can also simulate the local bound of the Garg–Mermin Bell inequality [Eq. (3)] for efficiency over $2/3$. Our results point out that whenever Bell violations are used for testing less-conventional theories, implementing device-independent quantum secure communication²⁷, certifying randomness⁴⁸ and nonlocality, loophole-free Bell tests^{51–53} should be performed. We would like to point out that there are Bell inequalities that use non-maximally entangled states with threshold efficiency $\eta_T = 2/3^{37}$; however, whether our model is effective against those would be a task for future study.

Methods

Strategies for controlling a and b . Here we show that regardless of the value of α_0 and α_1 an adversary can satisfy the assumption that whenever the Alice's basis matches (does not match) that of the incoming light, a deterministic (random) outcome is produced with probability a (b). For simplicity, let us assume the case when

Probability	Intensity	Polarization	Outcome when basis		Required value of I
			Matches	Mismatches	
$a - b$	I	α_0	α_0	x	$I \geq I_{th}, I \cos^2(2\phi_0) < I_{th}, I \sin^2(2\phi_0) < I_{th}$
$b/2$	I	$\alpha_0 + \phi_0$	α_0	α_1	$I \sin^2(\phi_0) < I_{th} \leq I \cos^2(\phi_0)$
$b/2$	I	$\alpha_0 - \phi_1$	α_0	α_1^\perp	$I \sin^2(\phi_1) < I_{th} \leq I \cos^2(\phi_1)$
$1 - a$	Vacuum		x	x	

Table 3. Strategy to practically simulate deterministic (random) outcome with efficiency a (b). Here, $\phi_0 = |\alpha_0 - \alpha_1|/2$, $\phi_1 = |\alpha_0 - \alpha_1^\perp|/2$, and ‘x’ represents no detection.

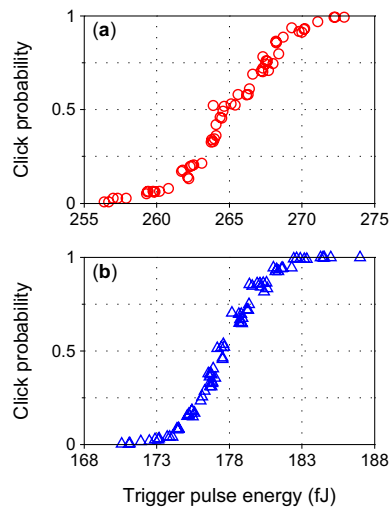


Figure 3. Control characteristics of a detector in commercial quantum key distribution system Clavis2^{45, 54}, responding to a short trigger pulse atop continuous-wave blinding power of (a) 740 μ W and (b) 367 μ W. Wavelength of light was $\sim 1.55 \mu$ m.

the adversary sends a light polarized at angle α_0 towards Alice (strategies for the other polarizations are similar). Then, with probability $(a - b)$, she sends light polarized at angle α_0 which, when measured in the same (different) basis, results in detection (no detection) if intensity is tailored properly (see Table 3). With probability $b/2$, she sends the light at an angle midway between α_0 and α_1 (α_1^\perp) at angle $\alpha_0 + \phi_1$ ($\alpha_0 - \phi_1^\perp$). Here, $\phi_1 = |\alpha_0 - \alpha_1|/2$, $\phi_1^\perp = |\alpha_0 - \alpha_1^\perp|/2$. As a result, when the basis matches, for both the cases, outcome is α_0 while for basis mismatch the outcome is α_1 and α_1^\perp with probability $b/2$. The condition for this is $I \sin^2 \phi < I_{th} < I \cos^2(\phi)$ for $\phi \in \{\phi_1, \phi_1^\perp\}$ as shown in Table 3. For the remaining times (with probability $1 - a$), the adversary sends vacuum. Overall, from Table 3, it can be seen that when the basis matches that of the incoming light, it results in a deterministic outcome with probability a ; while when the basis mismatches, it results in a random outcome with probability b . This supports the practicality of our assumption. Note that this method leads to asymmetric detection efficiency, as Bob’s efficiency is always higher than Alice’s. However, this can be avoided by reversing the roles of Alice and Bob half of the time.

We have so far assumed that the blinded detector is controllable as a step function: for $I < I_{th}$ the click probability is 0, and for $I \geq I_{th}$ it is 1. This is of course a simplification^{33–35, 40–47}. Real detectors have noise, which leads to them having two thresholds $I_{never} < I_{always}$, with click probability 0 for $I \leq I_{never}$ and 1 for $I \geq I_{always}$. In the range $I_{never} < I < I_{always}$, the click probability gradually increases from 0 to 1. These thresholds depend on the blinding power and regime. Furthermore, no two detector samples are identical, and require tweaking the faked states to achieve perfect or near-perfect control^{34, 35, 44}. Generally, if the ratio I_{always}/I_{never} can be made sufficiently small, perfect control can be achieved. These issues are device-specific and should be treated at the implementation stage. However, the ability to obtain an arbitrary click probability by adjusting I may allow an alternative method of controlling a and b , as we show below.

Practical detectors, when blinded, gradually increase their click probability from 0 to 1 in a certain range of trigger intensity^{33–35, 40–47}. This can be used to obtain probabilistic detections. To illustrate this, we have measured control characteristics of one avalanche photodiode detector in a commercial QKD system Clavis2^{45, 54}. At a particular continuous-wave blinding power, we varied the trigger pulse energy and recorded the corresponding click probability as shown in Fig. 3. The result shows that it is in principle possible for an adversary to select a value of trigger pulse intensity I (without varying the polarization by $\pm\phi$) that in a matching basis leads to click probability 1 in one detector, and when halved owing to basis mismatch, leads to a random click in either

detector with probability ~ 0.40 . However, some double clicks (i.e., simultaneous clicks in both detectors) will happen in this strategy. Their handling in a Bell test will need to be considered.

Received: 24 March 2020; Accepted: 15 July 2020

Published online: 06 August 2020

References

- Bell, J. S. On the Einstein Podolsky Rosen paradox. *Physics* **1**, 195–200 (1964).
- Freedman, S. J. & Clauser, J. F. Experimental test of local hidden-variable theories. *Phys. Rev. Lett.* **28**, 938–941 (1972).
- Aspect, A., Grangier, P. & Roger, G. Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.* **47**, 460–463 (1981).
- Aspect, A., Dalibard, J. & Roger, G. Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.* **49**, 1804–1807 (1982).
- Aspect, A., Grangier, P. & Roger, G. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities. *Phys. Rev. Lett.* **49**, 91–94 (1982).
- Weih, G., Jennewein, T., Simon, C., Weinfurter, H. & Zeilinger, A. Violation of Bell's inequality under strict Einstein locality conditions. *Phys. Rev. Lett.* **81**, 5039–5043 (1998).
- Giustina, M. *et al.* Bell violation using entangled photons without the fair-sampling assumption. *Nature* **497**, 227–230 (2013).
- Christensen, B. G. *et al.* Detection-loophole-free test of quantum nonlocality, and applications. *Phys. Rev. Lett.* **111**, 130406 (2013).
- Poh, H. S., Joshi, S. K., Cerè, A., Cabello, A. & Kurtsiefer, C. Approaching Tsirelson's bound in a photon pair experiment. *Phys. Rev. Lett.* **115**, 180408 (2015).
- Rowe, M. A. *et al.* Experimental violation of a Bell's inequality with efficient detection. *Nature* **409**, 791 (2001).
- Hofmann, J. *et al.* Heralded entanglement between widely separated atoms. *Science* **337**, 72–75 (2012).
- Ansmann, M. *et al.* Violation of Bell's inequality in Josephson phase qubits. *Nature* **461**, 504–506 (2009).
- Pfaff, W. *et al.* Demonstration of entanglement-by-measurement of solid-state qubits. *Nat. Phys.* **9**, 29–33 (2012).
- Ekert, A. K. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* **67**, 661–663 (1991).
- Mayers, D. & Yao, A. Quantum cryptography with imperfect apparatus. In *Proc. 39th Annual Symposium on Foundations of Computer Science*, 503–509 (IEEE, 1998).
- Barrett, J., Hardy, L. & Kent, A. No signaling and quantum key distribution. *Phys. Rev. Lett.* **95**, 010503 (2005).
- Colbeck, R. *Quantum and relativistic protocols for secure multi-party computation*. Ph.D. thesis, University of Cambridge (2006).
- Pironio, S. *et al.* Random numbers certified by Bell's theorem. *Nature* **464**, 1021–1024 (2010).
- Colbeck, R. & Renner, R. Free randomness can be amplified. *Nat. Phys.* **8**, 450–453 (2012).
- Gheorghiu, A., Kashefi, E. & Wallden, P. Robustness and device independence of verifiable blind quantum computing. *New J. Phys.* **17**, 083040 (2015).
- Hajdušek, M., Pérez-Delgado, C. A. & Fitzsimons, J. F. Device-independent verifiable blind quantum computation. *Preprint at arXiv:1502.02563* (2015).
- Mayers, D. & Yao, A. Self testing quantum apparatus. *Quant. Inf. Comput.* **4**, 273–286 (2004).
- McKague, M. Quantum information processing with adversarial devices. *Preprint at arXiv:1006.2352* (2004).
- Aharon, N., Massar, S., Pironio, S. & Silman, J. Device-independent bit commitment based on the CHSH inequality. *New J. Phys.* **18**, 025014 (2016).
- Lo, H.-K. & Chau, H. F. Unconditional security of quantum key distribution over arbitrarily long distances. *Science* **283**, 2050–2056 (1999).
- Curry, M., Lewenstein, M. & Lütkenhaus, N. Entanglement as a precondition for secure quantum key distribution. *Phys. Rev. Lett.* **92**, 217903 (2004).
- Acín, A. *et al.* Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.* **98**, 230501 (2007).
- Acín, A., Gisin, N. & Masanes, L. From Bell's theorem to secure quantum key distribution. *Phys. Rev. Lett.* **97**, 120405 (2006).
- Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. & Wehner, S. Bell nonlocality. *Rev. Mod. Phys.* **86**, 419–478 (2014).
- Gerhardt, I. *et al.* Experimentally faking the violation of Bell's inequalities. *Phys. Rev. Lett.* **107**, 170404 (2011).
- Gisin, N. & Gisin, B. A local hidden variable model of quantum correlation exploiting the detection loophole. *Phys. Lett. A* **260**, 323–327 (1999).
- Larsson, J.-Å. Modeling the singlet state with local variables. *Phys. Lett. A* **256**, 245–252 (1999).
- Lydersen, L. *et al.* Hacking commercial quantum cryptography systems by tailored bright illumination. *Nat. Photonics* **4**, 686–689 (2010).
- Gerhardt, I. *et al.* Full-field implementation of a perfect eavesdropper on a quantum cryptography system. *Nat. Commun.* **2**, 349 (2011).
- Liu, Q. *et al.* A universal setup for active control of a single-photon detector. *Rev. Sci. Instrum.* **85**, 013108 (2014).
- Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed experiment to test local hidden variable theories. *Phys. Rev. Lett.* **23**, 880–884 (1969).
- Eberhard, P. H. Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment. *Phys. Rev. A* **47**, R747–R750 (1993).
- Popescu, S. & Rohrlich, D. Quantum nonlocality as an axiom. *Found. Phys.* **24**, 379–385 (1994).
- Garg, A. & Mermin, N. D. Detector inefficiencies in the Einstein-Podolsky-Rosen experiment. *Phys. Rev. D* **35**, 3831–3835 (1987).
- Lydersen, L., Akhlaghi, M. K., Majedi, A. H., Skaar, J. & Makarov, V. Controlling a superconducting nanowire single-photon detector using tailored bright illumination. *New J. Phys.* **13**, 113042 (2011).
- Lydersen, L. *et al.* Thermal blinding of gated detectors in quantum cryptography. *Opt. Express* **18**, 27938–27954 (2010).
- Wiechers, C. *et al.* After-gate attack on a quantum cryptosystem. *New J. Phys.* **13**, 013043 (2011).
- Sauge, S., Lydersen, L., Anisimov, A., Skaar, J. & Makarov, V. Controlling an actively-quenched single photon detector with bright light. *Opt. Express* **19**, 23590–23600 (2011).
- Jogenfors, J., Elhassan, A. M., Ahrens, J., Bourennane, M. & Larsson, J.-Å. Hacking the Bell test using classical light in energy-time entanglement-based quantum key distribution. *Sci. Adv.* **1**, e1500793 (2015).
- Huang, A. *et al.* Testing random-detector-efficiency countermeasure in a commercial system reveals a breakable unrealistic assumption. *IEEE J. Quant. Electron.* **52**, 8000211 (2016).
- Chistiakov, V., Huang, A., Egorov, V. & Makarov, V. Controlling single-photon detector ID210 with bright light. *Opt. Express* **27**, 32253–32262 (2019).
- Gras, G. *et al.* Optical control of single-photon negative-feedback avalanche diode detector. *J. Appl. Phys.* **127**, 094502 (2020).
- Tanner, M. G., Makarov, V. & Hadfield, R. H. Optimised quantum hacking of superconducting nanowire single-photon detectors. *Opt. Express* **22**, 6734–6748 (2014).
- Elezov, M., Ozhegov, R., Goltsman, G. & Makarov, V. Countermeasure against bright-light attack on superconducting nanowire single-photon detector in quantum key distribution. *Opt. Express* **27**, 30979 (2019).

50. Branciard, C. Detection loophole in Bell experiments: How postselection modifies the requirements to observe nonlocality. *Phys. Rev. A* **83**, 032123 (2011).
51. Hensen, B. *et al.* Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526**, 682–686 (2015).
52. Giustina, M. *et al.* Significant-loophole-free test of Bell's theorem with entangled photons. *Phys. Rev. Lett.* **115**, 250401 (2015).
53. Shalm, L. K. *et al.* Strong loophole-free test of local realism. *Phys. Rev. Lett.* **115**, 250402 (2015).
54. Clavis2 specification sheet, <http://marketing.idquantique.com/acton/attachment/11868/f-00a0/1/-/-/-/Clavis%20QKD%20DataSheet.pdf>, visited 10 May 2018.

Acknowledgements

We thank D. A. Graft, G. Adenier, H.-K. Lo, and Y. Zhang for discussions. This work was funded by Industry Canada, CFI, NSERC (programs Discovery and CryptoWorks21), Ontario MRI, US Office of Naval Research, and the Ministry of Education and Science of Russia (program NTI center for quantum communications).

Author contributions

S.S. developed the perfect local model, measured detector characteristics, and wrote the manuscript with input from all authors. N.S. developed the improvement to existing model. C.C.W.L. contributed to the study of the perfect local model and V.M. supervised the study.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020