Insecurity of detector-device-independent quantum key distribution

Anqi Huang,[1,2] Shihan Sajeed,[1,2] Shihai Sun,[3] Feihu Xu,[4] Vadim Makarov,[1,5,2] and Marcos Curty[6]

[1]*Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[2]*Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[3]*College of Science, National University of Defense Technology, Changsha 410073, China*
[4]*Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*
[5]*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[6]*EI Telecomunicación, Department of Signal Theory and Communications, University of Vigo, Vigo E-36310, Spain*

Introduction.—It is time to close the gap between theory and practice in quantum key distribution (QKD) [1], a technique to distribute a secret random bit string between two separated parties (Alice and Bob). In theory, QKD provides information-theoretic security based on the laws of physics. In practice, however, it does not, as QKD realisations cannot typically fulfill the demands imposed by the theory.

To bridge this gap, measurement-device-independent QKD (mdiQKD) [2, 3] currently appears to be the most promising approach so far, both in terms of feasibility and performance. It removes all side-channels from the measurement unit, which is arguably the most critical element of QKD systems [4]. Its security is based on post-selected entanglement [2]. Given that Alice and Bob know their state preparation processes, mdiQKD can deliver a key rate that is comparable to that of standard entanglement-based QKD schemes [3]. Also, its practicality has been already confirmed [5–7]. A limitation of this technique is, however, its experimental complexity, as it requires high-visibility two-photon interference between two independent light sources. Another limitation is that current finite-key security bounds for mdiQKD against general attacks [3] require larger post-processing data block sizes than those of standard QKD.

To overcome these two main limitations, a novel approach, called detector-device-independent QKD (ddiQKD), has been introduced very recently [8–11]. By assuming that Alice and Bob can trust their state preparation processes, this technique avoids the problem of interfering photons from independent sources by using the concept of a single-photon Bell state measurement (BSM) [12]. This makes its experimental implementation simpler than mdiQKD. Also, its classical post-processing data block sizes and its finite secret key rate are expected to be similar to those of standard prepare-and-measure QKD schemes. Despite this promising performance, its robustness against detector side-channel attacks has not been rigorously proven, and but only conjectured [8–11]. The security of ddiQKD was assumed to be based on post-selected entanglement, similar to mdi-QKD [8–11]. Nonetheless, these security assumptions can not be guaranteed in practice [13].

The main contributions of this work are twofold. First, we show that, in contrast to mdiQKD, the security of ddiQKD cannot be based on post-selected entanglement alone, as has been initially thought in [8–11]. Second, we argue that ddiQKD is actually insecure under detector side-channel attacks. For this, we present various eavesdropping strategies, all of them exploiting the detector blinding attack [4], that can provide Eve with full information about the distributed secret key without introducing any error.

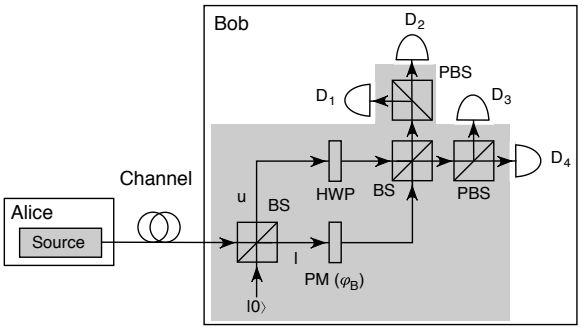

FIG. 1: Schematic representation of a possible ddiQKD realisation [9]. HWP, half-wave plate; and PM, phase modulator. A click in D_1 indicates a projection into the Bell state $|\Psi^+\rangle = (|H\rangle|l\rangle + |V\rangle|u\rangle)/\sqrt{2}$; a click in D_2 indicates a projection into the Bell state $|\Phi^+\rangle = (|H\rangle|u\rangle + |V\rangle|l\rangle)/\sqrt{2}$; a click in D_3 indicates a projection into the Bell state $|\Psi^-\rangle = (|H\rangle|l\rangle - |V\rangle|u\rangle)/\sqrt{2}$; and a click in D_4 indicates a projection into the Bell state $|\Phi^-\rangle = (|H\rangle|u\rangle - |V\rangle|l\rangle)/\sqrt{2}$; The grey area denotes devices that need to be characterised and trusted. Also, Alice's and Bob's laboratories need to be protected from any information leakage to the outside.

The security of ddiQKD cannot be based on post-selected entanglement.— To see this, we consider a slightly simplified version of the ddiQKD scheme illustrated in Fig. 1. In particular, we assume that Bob's receiver has only one active detector, say for instance detector D_1, while the detectors D_2, D_3 and D_4 are disabled. That is, the BSM at Bob's side projects the incoming photons into only one Bell state. If the security of ddiQKD is based on post-selected entanglement alone, this modification should not affect its security (only its final secret key rate would be reduced by a factor of four) because a projection on a single Bell state should be sufficient to guarantee security [2]. Below we show that ddiQKD is insecure in this

scenario, which confirms that its security cannot be based on the same principles as mdiQKD.

For this, we elaborate that a blinding attack [4] allows Eve to obtain full information about the secret key without introducing any error. In particular, suppose that Eve shines bright light onto Bob's detector D_1 to make it enter linear-mode operation. This way, the detector is no longer sensitive to single-photon pulses, but it can only detect strong light. More precisely, we will assume that D_1 can only produce a "click" when the intensity μ of the incoming light is above a certain threshold value $\mu_{\rm th}$, and otherwise produces a "no click" event. Now, Eve performs an intercept-resend attack: she measures out every signal emitted by Alice in one of the two BB84 bases (which Eve selects at random for every signal), and then sends a new signal, depending on the result obtained, to Bob. Intercept-resend attacks correspond to entanglement-breaking channels and, therefore, they cannot lead to a secure key [14]. In particular, we will assume that the signals that Eve sends to Bob are coherent states of the form $|\sqrt{2}\alpha\rangle$, with creation operator $a^\dagger = (a_{\rm H}^\dagger + e^{i\phi_{\rm E}} a_{\rm V}^\dagger)/\sqrt{2}$, where $a_{\rm H}^\dagger$ ($a_{\rm V}^\dagger$) denotes the creation operator for horizontally (vertically) polarised photons. The value of the angle $\phi_{\rm E} \in \{0, \pi/2, \pi, 3\pi/2\}$ depends on Eve's measurement result. That is, Eve sends Bob coherent states prepared in the polarisation states output by her measurement at each given time. It can be shown that the state at the input ports of Bob's detectors is of the form

$$|\psi\rangle = |\frac{\alpha}{2}\big(e^{i\phi_{\rm E}} + e^{i\varphi_{\rm B}}\big)\rangle_{\rm D_1} \otimes |\frac{\alpha}{2}\big(1 + e^{i(\phi_{\rm E}+\varphi_{\rm B})}\big)\rangle_{\rm D_2} \otimes |\frac{\alpha}{2}\big(e^{i\phi_{\rm E}} - e^{i\varphi_{\rm B}}\big)\rangle_{\rm D_3} \otimes |\frac{\alpha}{2}\big(1 - e^{i(\phi_{\rm E}+\varphi_{\rm B})}\big)\rangle_{\rm D_4}. \quad (1)$$

This is illustrated in Table I, where we show the intensity of the incoming light at the input ports of Bob's detectors for all possible combinations of $\phi_{\rm E}$ and $\varphi_{\rm B}$. Here we use the notation $\mu = |\alpha|^2$. Also, we assume that Eve selects the intensity μ such that $\mu/2 < \mu_{\rm th} < \mu$. Importantly, note that Bob's detector D_1 only produces a "click" when Bob uses the same measurement basis as Eve (*i.e.*, $\varphi_{\rm B}, \phi_{\rm E} \in \{0, \pi\}$ or $\varphi_{\rm B}, \phi_{\rm E} \in \{\pi/2, 3\pi/2\}$), and therefore no errors are generated. This means that whenever the system loss is above 3 dB the ddiQKD illustrated in Fig. 1 with one detector is insecure. The same conclusion applies as well to the ddiQKD schemes introduced in [8, 10, 11]

Insecurity of ddiQKD under detector side-channel attacks.— The eavesdropping strategy described above has one main drawback: when Bob uses two or more detectors D_i, it can produce double-clicks [13]. Indeed, suppose the original ddiQKD scheme with four detectors. From Table I one can see that whenever Bob uses the same measurement basis as Eve, there are two detectors that will always "click". For instance, when $\varphi_{\rm B} = \phi_{\rm E} = 0$ then both detectors D_1 and D_2 will "click", and similar for the other cases. However, below we show that by exploiting other side-channels, Eve is able to eliminate double clicks.

$\varphi_{\rm B}$	D_1	D_2	D_3	D_4
0	μ	μ	0	0
$\frac{\pi}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$
π	0	0	μ	μ
$\frac{3\pi}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$

(a) Angle $\phi_{\rm E} = 0$

$\varphi_{\rm B}$	D_1	D_2	D_3	D_4
0	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$
$\frac{\pi}{2}$	μ	0	0	μ
π	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$
$\frac{3\pi}{2}$	0	μ	μ	0

(b) Angle $\phi_{\rm E} = \frac{\pi}{2}$

$\varphi_{\rm B}$	D_1	D_2	D_3	D_4
0	0	0	μ	μ
$\frac{\pi}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$
π	μ	μ	0	0
$\frac{3\pi}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$

(c) Angle $\phi_{\rm E} = \pi$

$\varphi_{\rm B}$	D_1	D_2	D_3	D_4
0	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$
$\frac{\pi}{2}$	0	μ	μ	0
π	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$
$\frac{3\pi}{2}$	μ	0	0	μ

(d) Angle $\phi_{\rm E} = \frac{3\pi}{2}$

TABLE I: Intensity of the input light to Bob's detectors as a function of the angles $\phi_{\rm E}$ and $\varphi_{\rm B}$. Here, the parameter $\mu = |\alpha|^2$.

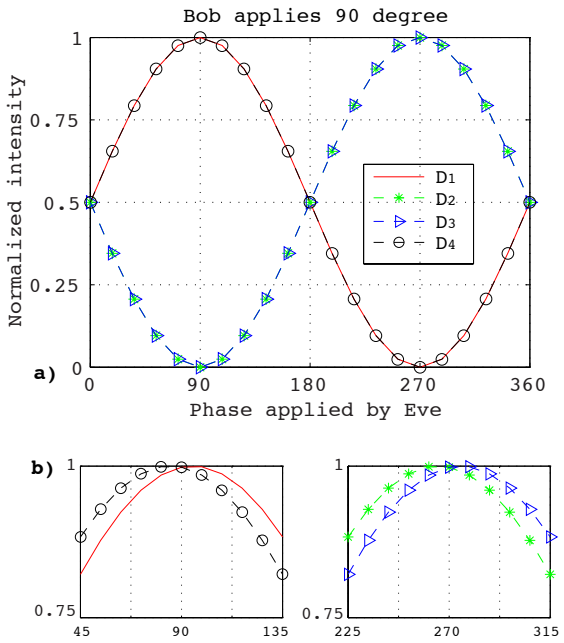

FIG. 2: Normalized intensities at the four detectors as a function of $\phi_{\rm E}$.

Firstly, if Bob's phase modulation is not perfect, as normally is the case in practical implementations, double clicks can be eliminated due to the distinguishable intensity distribution. If in Eq. 1, we add a modulation error of $\pm\Delta\varphi_{\rm B}$ to $\varphi_{\rm B}$, we find that the intensity distributions shift with respect to each other as highlighted in Fig. 2b. In this simulation we have assumed a modulation error of 5° which is not uncommon in practical systems [15]. Eve can exploit this asymmetry. Instead

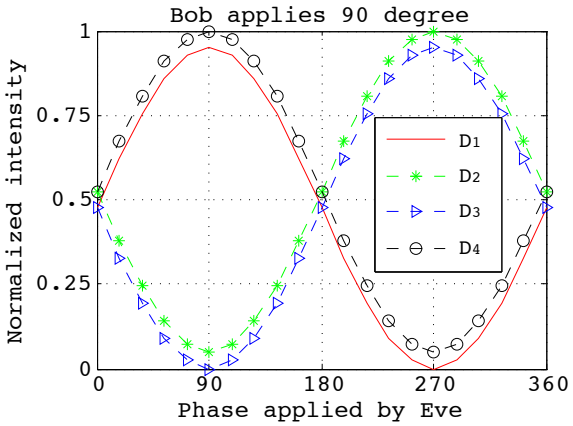

FIG. 3: Normalized intensities at the four detectors for Bob's first BS with ratio 64:36

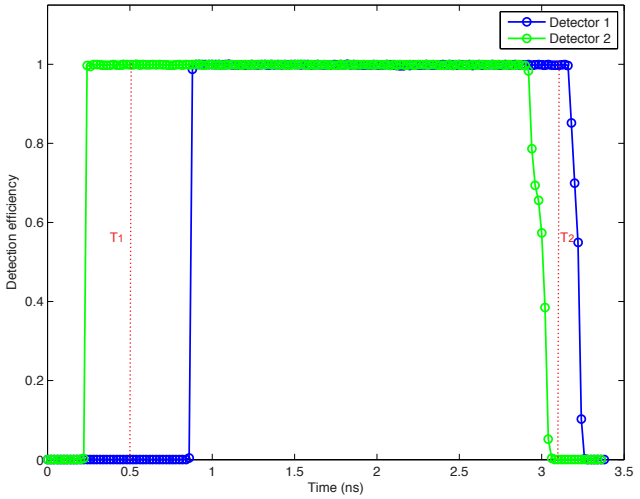

FIG. 4: Experimental result of efficiency mismatch when detectors are under blinding attack.

of selecting pre-defined values of $\phi_{\rm E} \in \{0, \pi/2, \pi, 3\pi/2\}$, she can use, say $\phi_{\rm E} \pm 10^{\circ}$ with a carefully chosen intensity to make sure only one of the detectors click. In this way, double clicks will be eliminated and the protocol is insecure against blinding attacks also in an scenario with four detectors.

Secondly, an imperfection in Bob's beam splitter (BS) also opens a side-channel to avoid double clicks. Beam splitters are wavelength-sensitive. A 50 : 50 BS designed for 1550 nm wavelength will have a nearly perfect splitting ratio at that wavelength. However, for a different wavelength this output ratio can be quite different [16]. For example, let's assume that instead of having a splitting ratio of 50 : 50, the first beam splitter in Bob (see Fig. 1) has a splitting ratio of 64 : 36 for a certain wavelength used by Eve. Eve also performs intercept-resend attack as described above. When $\phi_{\rm E} = 0$, she sends a coherent state $|\sqrt{2}\alpha\rangle$ with creation operator $a^{\dagger} = (0.8a_{\rm H}^{\dagger} + 0.6e^{i\phi_{\rm E}}a_{\rm V}^{\dagger})$ to Bob's blinded detectors. When Bob applies $\varphi_{\rm B}$, the intensities at the four detectors are shown in Fig. 3. By adjusting the intensities of H and V in the faked state, Eve is able to control either detector D_1 or D_4 clicks to avoid double clicks.

Another strategy is exploiting efficiency mismatch to perform a time-shift attack [17]. To investigate the detection efficiencies of blinded detectors, we test two detectors' detection probabilities under blinding attack on another non DDI-QKD system. Fig. 4 shows our experimental result; it comfirms that the efficiency mismatch is still possible among the blinded detectors, which means that a time-shift attack can be utilized to control which detector clicks. For example, when the signal pulse arrives at T_1, only Detector 2 clicks, while, this temporal point is out of the response region of Detector 1. However, if Eve delays the arrival time to T_2, only Detector 1 clicks, while Detector 2 keeps silent. If this efficiency mismatch exists in ddiQKD system as well, Eve can also avoid double clicks successfully by using this method.

Conclusion—In this work we have shown that, in contrast to mdiQKD, the security principle of ddiQKD cannot be based on post-selected entanglement. Also, we have presented several potential eavesdropping strategies, which exploit detector blinding attacks in combination with other side-channels from other device imperfections, which would render ddiQKD insecure.

[1] V. Scarani, H. Bechmann-Pasquinucci, N. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, Rev. Mod. Phys. **81**, 1301-1350 (2009).
[2] H.-K. Lo, M. Curty, and B. Qi, Phys. Rev. Lett. **108**, 130503 (2012).
[3] M. Curty *et al.*, Nat. Commun. **5**, 3732 (2014).
[4] L. Lydersen *et al.*, Nat. Photonics **4**, 686-689 (2010).
[5] Z. Tang, Z. Liao, F. Xu, B. Qi, L. Qian, and H.-K. Lo, Phys. Rev. Lett. **112**, 190503 (2014).
[6] Y.-L. Tang *et al.*, IEEE J. Sel. Topics Quantum Elect. **21**, 6600407 (2015).
[7] L. C. Comandar *et al.*, Nat. Photonics **10**, 312-315 (2016).
[8] P. Gonzalez *et al.*, Phys. Rev. A **92**, 022337 (2015).
[9] C. C. W. Lim, B. Korzh, A. Martin, F. Bussieres, R. Thew, and H. Zbinden, Appl. Phys. Lett. **105**, 221112 (2014).
[10] W.-F. Cao *et al.*, preprint arXiv:1410.2928 (2014).
[11] W.-Y. Liang *et al.*, Phys. Rev. A **92**, 012319 (2015).
[12] Y.-H. Kim, Phys. Rev. A **67**, 040301(R) (2003).
[13] Qi, B, Phys. Rev. A **91(2)**, 020303(R) (2015).
[14] M. Curty, M. Lewenstein, and N. Lütkenhaus, Phys. Rev. Lett. **92**, 217903 (2004).
[15] F. Xu *et al.*, Phys. Rev. A **92**, 032305 (2015).
[16] H.-W. Li *et al.*, Phys. Rev. A **84**, 062308 (2011).
[17] B. Qi *et al.*, Quant. Inf. Comp. **7**, 73-82 (2007).