Quantum Hacking in the Age of Measurement-Device-Independent Quantum Cryptography

by

Anqi Huang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering (Quantum Information)

Waterloo, Ontario, Canada, 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Gilles Brassard Professor Dept. of Computer Science and Operations Research Universit de Montral
Supervisors:	Vadim Makarov Research Assistant Professor Dept. of Physics and Astronomy
	Christopher Wilson Professor Dept. of Electrical and Computer Engineering
Internal Members:	Michal Bajcsy Assistant Professor Dept. of Electrical and Computer Engineering
	Guo-Xing Miao Assistant Professor Dept. of Electrical and Computer Engineering
Internal-External Member:	Michele Mosca Professor Dept. of Combinatorics and Optimization

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Cryptography is essential for secure communication in the digital era. Today, public-key cryptography is widely employed, and has provided an efficient method for encrypting content and ensuring both confidentiality and authenticity of electronic communications. However, the security of these systems is based on assumptions of computational hardness within the constraints of current computing capability. Thus, as quantum computing becomes a reality, public-key algorithms will be genuinely vulnerable to attack. By contrast, quantum cryptography, which is based on quantum physics instead of mathematical assumptions, is able to achieve information-theoretic security.

Advances in practical quantum cryptographic systems have not kept pace with theory, where an eavesdropper can relatively easily exploit loopholes in practical implementations to compromise theory-proved security. Bridging the gap between perfect theory and imperfect practice has become a priority for the growing field of quantum key distribution (QKD), which has strived to strengthen the practical security of QKD systems. Among all the countermeasures against quantum hacking, the measurement-device-independent (MDI) QKD protocol is promising because it is immune to all side-channel attacks on measurement devices. However, the MDI QKD protocol has some limitations that critically restrict its practical usefulness. Technically, the MDI scheme is not compatible with existing QKD systems, and produces a low key rate. In addition, the theory underlying MDI QKD security is based on the use of trusted source stations. Thus, this protocol is not a universal solution. This thesis further investigates the practical security of quantum cryptography in and beyond MDI quantum cryptography.

To overcome the technical limitations of MDI QKD, we first evaluate two other countermeasures against imperfect detections. The first is an industrial patch based on random detection efficiency, recently implemented by ID Quantique in the commercial Clavis2 QKD system. While powerful, experimental testing shows that this countermeasure is not sufficient to defeat the detector blinding attack. The second countermeasure aims to achieve a higher key rate than MDI QKD while maintaining the same security properties. However, our research shows that detector-device-independent (DDI) QKD security is not equivalent to that of MDI QKD and, further, that DDI QKD is insecure against detector side-channel attacks.

While this initial work points to the superior performance of MDI QKD systems, core challenges remain. The fundamental security assumption adopted for MDI QKD systems, regarding the exclusive use of trustable source stations, cannot always be satisfied in practice. Our study revealed several side channels of source devices. The first is disclosed

from the implementation of a decoy-state protocol, which is widely used in QKD systems with weak coherent sources. The pump-current-modulated intensities result in a timing mismatch between the signal and decoy states, violating the key assumption in the decoy-state QKD protocol. Moreover, an active Eve can break the basic assumption about photon numbers in the QKD system. In this work, we experimentally demonstrate a laser seeding attack on the laser source, which shows that Eve can increase the emission power of the laser diode. Furthermore, by shining a high-power laser into an optical attenuator, Eve can decrease the attenuation values. The increase in laser emission power and the decrease in attenuation leads to an increase in mean photon numbers.

In summary, MDI QKD is a milestone in quantum cryptography. However, this thesis indicates the importance of continued investigations into the practical security of MDI QKD. The analysis of practical security should be extended to other countermeasures against side-channel attacks and the source stations in MDI QKD systems. Practical quantum hacking and security analysis promote the development of quantum cryptographic systems, which will eventually achieve the unconditional security claimed in theory.

Acknowledgements

I would like to thank various people who made my Ph.D. study at the Institute for Quantum Computing in University of Waterloo exciting and enjoyable. First of all, I would like to express my gratitude for my supervisor, Dr. Vadim Makarov, who has supported me with his knowledge and kindness during my study. He guided me towards the world of quantum hacking and let me be interested in my research. I enjoyed every moment at which I worked with him. I am thankful for many opportunities and resources he provided me to present my work and communicate with excellent researchers around the world. It is my fortune to work with him. He is always my supervisor in my heart.

I would also like to thank Prof. Christopher Wilson for being my supervisor to support my study going smoothly. Special thanks are extended to Prof. Michele Mosca, Prof. Michal Bajcsy, and Prof. Guo-Xing Miao for being my committee members. Especially, I am very grateful to Prof. Gilles Brassard for his time in studying my work and for being my thesis examiner. It is my honor to have one of the founders of QKD examining my thesis. I am also thankful to Prof. Norbert Lütkenhaus and Prof. Thomas Jennewein for teaching me the knowledge about QKD and providing help when I needed. Being supported and guided by all of these outstanding professors made my Ph.D. study amazing.

Then it is a pleasure to thank all of the people I had the chance to work with. Thank members from our group: Shihan Sajeed, Poompong Chaiwongkhot, Hao Qin, Ruoping Li, and Jan Gulla. Thank you for all the discussion, assistance, and thesis proofreading. I will always remember the time we spent in the quantum hacking lab. There are other IQC members that I would like to thanks for the help: Katanya Kuntz, Nigar Sultana, Yanbao Zhang, Chang Liu, Guiyang Han, Le Han, Honghao Fu, Jie Lin, and Elena Anisimova. You made my life in IQC interesting. Besides the people in IQC, I thanks for the privilege to collaborate with Shihai Sun, Marcos Curty, Feihu Xu, M. Soucarros, M. Legré, Stefanie Barz, Erika Andersson, Hoi-Kwong Lo, Akihiro Mizutani, Kiyoshi Tamaki, Álvaro Navarrete, Vladimir Chisriakov, Vladimir Egorov, and Gaëtan Gras.

I would like to thank all the friends I made in Waterloo. Especially, I am fortunate to meet Yu Huang, Peyton Shi, Liang Dong, Tianyu Yang, Rodney Ding, Jiaqi Xu, Demin Yin, and Bingyao Tan. Thank you for giving me a feeling like a family in Waterloo. I enjoyed the time with you. I am also grateful to my parents and parents-in-law for their endless encouragement and support.

Most importantly, I deeply thank my husband Zhihong Liu for being with me, giving me the love, and supporting my every step. Thank you for giving me the happiness in my life and giving me the energy to fulfill my study. Even though we had to be apart from each other for more than two years, your love is always with me. Thank you for all!

Dedication

This is dedicated to dear Zhihong Liu, my husband and my love.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Quantum cryptography (QC)

1.1.1 Why quantum cryptography?

Information is invisible but valuable property. Information security guarantees that only authorized users have access to precise and integrated information when requested [1]. Locally preserving information security is easy and feasible by setting a protection zone. However, maintaining security properties during information transmission is challenging, as it is difficult to guard the entire transmitting channel, particularly when it is long. Cryptography is therefore applied to protect information from unauthorized disclosure in transit.

Cryptology is the art of code-making and code-breaking. Once an encryption code is made, code breakers attempt to defeat it. Subsequently, code makers propose new types of code to replace broken ones. This centuries-old battle to maintain the security of confidential information, on the one hand, and exploit vulnerabilities in that security. On several occasions, this competition changed the course of human history. An often-cited example was the breaking of Axis' codes, including the famous German Enigma machine, which was instrumental to ending World War II two years earlier than predicted [156]. Though that war is over, the war between code makers and code breakers continues. For code makers, inventing a truely unbreakable cryptography is the ultimate goal to achieving information security. Given the unprecedented power of quantum computing, this is a colossal task.

Today, cyber technology is one of the fastest growing industries worldwide, and the Internet has spread to a global scale. With pervasive adoption of modern information and communication technologies, cybersecurity is essential. Cryptography is a vital tool for achieving information security in a cyber environment, particularly when untrusted network channels are used. Public-key cryptography currently satisfies the requirements of security in our digital economy. The ease of creating and distributing key pairs using public-key algorithms makes them a convenient solution for ensuring secure authentication and nonrepudiation [130].

The security of public-key cryptographic systems relies on *unproven* computational assumptions, yet it is reasonable to believe that the currently-used public-key systems are unlikely to be cracked by current computing capability. However, a mathematical assumption is not a foundation that can be relied upon forever. For example, RSA (Rivest-Shamir-Adleman), a widely used public-key algorithm, relies on the difficulty of factoring: it is easy to multiply two large prime numbers, but it is assumed to be difficult to factor them. The difficulty is based on today's computational and algorithmic capability. The advent of quantum computation poses a significant threat to this assumption. Quantum computers will have the computational power to run Shor's algorithm [154] to factor large numbers in a reasonable period of time, rendering the RAS protocol vulnerable.

Quantum computing will usher in a new era for computation and information security. Once the universal quantum computer is available, the cryptographic systems based on mathematical assumptions will become vulnerable. This may become a reality in just a few decades. Should we start considering the capability and threat of quantum computers even now? Michele Mosca's theorem [129] answers this question. Let us assume that we need to secure our secret information for x years, and it will take y years to re-tool the existing cryptographic infrastructure with a large-scale quantum-safe solution. Building a large and universal quantum computer will take z years. The theorem says that if $x+y > z$, we should worry about the threat from quantum computing right now.

To counter the threat from quantum computers, we must replace our cryptographic infrastructure with quantum-safe technology. There are currently two "quantum-safe" candidates with potential to provide security against quantum attacks [2]. One is post-quantum cryptography, a conventional cryptographic system that, while based on mathematical assumptions, is believed to be secure against quantum attacks. However, this statement may need more solid proof. The second candidate is quantum cryptography, which is based on quantum mechanics, and is not limited by computational assumptions. Quantum cryptography can be immune to quantum attacks and, therefore, can achieve information-theoretic security.

1.1.2 What is quantum cryptography?

Quantum cryptography is the science and art of exploiting quantum mechanics to fulfill cryptographic objectives. There are many quantum cryptography primitives, including (but not limited to): quantum key distribution (QKD) [35], quantum digital signature (QDS) [69], quantum random number generator (QRNG) [81], quantum secret sharing (QSS) [49], quantum coin tossing (QCT) [35], quantum secure direct communication (QSDC) [55], quantum oblivious transfer (QOT) [37], and blind quantum computing (BQC) [44]. With a strong foundation in quantum physics, including Heisenberg's uncertainty principle, no-cloning theorem, and quantum entanglement, quantum cryptography has potential to achieve information-theoretic, or unconditional, security.

As the best-known application of quantum information for cryptography, QKD [35] (invented in 1983) took the spotlight and has seen rapid development in the past two decades. Remarkably, several QKD protocols have been proved to be unconditionally secure (in theory) by rigorous security proofs [70,109,155], and has even been implemented over long distances in free space [152] and optical fibers [106,144]. In the security model, an eavesdropper is allowed to have unlimited computational power, even a quantum computer, but is restricted by the laws of physics. The security of QKD also follows Kerckhoffs's principle [89], which states that Eve knows the entire protocol and system, except for the shared key.

To introduce the working mechanism of QKD, we take the well-known Bennett-Brassard 1984 (BB84) protocol [35] as an example. The convention in cryptography is to refer to the information sender as Alice, the receiver as Bob, and the eavesdropper as Eve. The phases of the BB84 protocol are the following:

- State preparation. Alice prepares a string of quantum states, which are randomly selected from X basis or Z basis. For instance, she chooses the polarization of photons as the encoding bases: horizontal ($|H\rangle$) and vertical ($|V\rangle$) polarizations as X basis and diagonal ($|+\rangle$), and anti-diagonal ($|-\rangle$) polarizations as Z basis. Among these four quantum states, $|H\rangle$ and $|+\rangle$ correspond to bit 0; $|V\rangle$ and $|-\rangle$ project to bit 1. Thus, Alice randomly and independently picks one of four quantum states for every slot in the bit string.

- Transmission. Alice transmits the prepared quantum states via a quantum channel, which is an optical fiber or free space.

- State measurement. Once the states reach Bob's side, he also randomly selects X basis or Z basis to measure them. Quantum mechanics shows that when Alice and

Bob select the same basis, they obtain the same bit value from the quantum state with certainty, whereas if they choose different bases, they may get opposite bit values with 50 % probability.

- Sifting. To exclude the cases of bases mismatch, they communicate via a classical authenticated public channel to announce their basis choices. Then they keep only the slots in which they choose the same bases, and discard the rest. The kept part is called the 'sifted key'.

- Parameter estimation. Alice and Bob take a small part of the sifted key and announce their exact bit values. Due to channel noise or Eve's disturbance, the bit values may not match perfectly, which is treated as an error. By calculating the error rate among these disclosed bits, Alice and Bob can regard the value of the error rate as a reference for the rest of the secret bits. If the error rate is higher than a preset threshold, the protocol aborts. Otherwise, Alice and Bob continue to the next phase.

- Error correction. To correct the small fraction of errors, Alice and Bob apply an error correction algorithm in this phase. After error correction, ideally, Alice and Bob obtain the identical bit string at each end. During error correction, some information about the secret key is disclosed over the public channel.

- Privacy amplification. Eve eavesdrops the key during the transmission, and some secret information is announced in the last step. Thus, Alice and Bob apply privacy amplification to squeeze out the information that Eve may know. After this phase, Alice and Bob finally share the secret key.

Please note: The QKD protocol requires an authenticated channel as the public channel. Thus it ensures that Alice communicates with a real Bob, but not Eve, to avoid man-in-the-middle attacks. Once the secret key is established by QKD, it can be applied to any encryption algorithm. However, unconditional security can only be achieved if the secret key is used in a one-time pad algorithm to transmit the message. This is true as long as the secret key is never reused.

1.2 Practical quantum key distribution (QKD)

1.2.1 Experimental implementations

The first QKD system, which ran the BB84 protocol, was realized in 1989 [36]. Over the past 29 years, various types of QKD protocols have been implemented in the lab and even

in the field.

In practice, optical fibers and free space are used as quantum channels to transmit quantum states. Quantum states can be defined in several different degrees of freedom of a photon. As previously mentioned, polarization is a typical degree of freedom that is used in free-space QKD systems [34,46,99,100,152]. For fibre-based QKD, one often chooses phase encoding [66,159,163], frequency encoding [41], and time-bin encoding [107,170]. Polarization encoding is subject to birefringence in the fibers, which affects its stability [174].

Weak coherent sources are commonly used in practical QKD systems. To satisfy the requirement of a single photon, an attenuator usually follows the weak coherent source. A decoy-state BB84 system with this type of weak coherent source can reach transmission distance of about 140-200 km [106, 152]. It is remarkable that, in recent years, demonstrations based on measurement-device-independent (MDI) QKD can reach distances up to 300-400 km over optical fibers [169, 182]. If Alice and Bob wish to extend to a longer distance, another option is entanglement-based QKD protocols [58, 119, 137], which can tolerate around 70 dB loss. In this protocol, an untrusted entanglement source is applied to distribute entangled states between Alice and Bob. They measure the entangled states to obtain the shared information. However, this approach generates only relatively low secure key rates.

Other solutions that can generate a high key rate in a short distance (below 100 km) are distributed-phase-reference QKD protocols, which include the differential-phase-shift protocol [78,163] and the coherent-one-way protocol [158,160]. In these protocols, information is encoded in the coherence between adjacent pulses, instead of individual pulses. For the receiver, Bob uses a Mach-Zehnder interferometer to check the coherent information. In the coherent-one-way protocol, Alice also changes the intensities of the pulses. Bob detects different intensities to get bit information.

The protocols described above are discrete variable QKD protocols, which require single-photon detectors to detect every individual photon. InGaAs/InP avalanche photodiodes have been introduced and are widely used, but they demonstrate low detection efficiency (about $10-20\%$) [66,92], thus limiting the secure key rate. Continuous-variable QKD was proposed to overcome the limitations of single-photon detectors [71,86,140]. This type of protocol now employs standard telecom p-i-n photo diodes and homodyne detection to measure light-field quadratures. As these protocols require only standard telecom components and do not need single-photon detectors, they are suited for experimental realization.

1.2.2 Commercialization and globalization

As QKD implementation continues to mature, several companies have commercialized QKD systems and related products [3].

ID Quantique (IDQ), a Swiss company with a significant presence in the quantum cryptography market, has been producing commercial QKD systems since 2007. Their flagship QKD implementation was used to encrypt the election data in Geneva in 2007 [4]. This application worked reliably. Moreover, the company provides quantum-safe cryptography solutions [5–7] to customers. For example, IDQ offers a hybrid encryption solution, merging state-of-the-art 10 Gigabit Ethernet encryption with QKD [8]. The 10-Gigabit Ethernet encryptors secure the backbone link between a company's headquarters and a database center, with one more layer of security provided by the QKD system.

In China, there are two other QKD companies, Qasky and QuantumCTek. Qasky offers a decoy BB84 QKD system that employs the Faraday-Michelson (F-M) phase encoding scheme [9], which is immune to channel disturbance and thus can operate stably. The system is available to work at 20 MHz [10], 50 MHz [10] or 1 GHz repetition frequency [11]. Qasky also offers several services: point-to-point secure transmission [12], QKD network based on an optical switch [13], and a real-time full-access network [14]. Qasky customers include power-grid industries, banks, and government.

QuantumCTek is rapidly developing in the field of commercialized quantum information technology, and is now the world's largest manufacturer and provider of quantum-safe products and services. QuantumCTek is committed to providing competitive quantum-safe solutions in telecom infrastructure, enterprise networks, cloud computing, big data technology, and its services [15]. The solutions, products, and services are provided to government, financial institutions, and energy industries [16]. They are also the product supplier of the Beijing-Shanghai QKD backbone link [17]. So far, more than one thousand QuantumCTek quantum secure products have been manufactured and are running online, securing communication links longer than 4000 km [18].

As point-to-point QKD can transmit a secret key only in the range of several hundreds of kilometers, a QKD network is one solution to securely extend the distance. Significant efforts all over the world have been put into QKD networks and advance QKD globalization.

In 2003, the European Union launched a project, the Development of a Global Network for SEcure COmmunication based on Quantum Cryptography (SECOQC) [134], which aimed to enhance the practical applications of QKD. Forty-one research and industrial organizations participated in this project. An eight-point network was formed based on the trusted-node scheme [134]. The point-to-point links were built by various types of

QKD technology: plug-and-play systems, a one-way weak pulse system, a coherent-one-way system, an entangled photons system, a continuous-variables system and a free-space system.

In the UK, a quantum technology hub for quantum communications [19] was established in 2014, linking 8 UK universities and several companies, including BT, Toshiba Research Europe Ltd, and the National Physical Laboratory. The Hub partnership mainly focuses on three areas to provide technology prototypes: the UK's first quantum network; chip-scale integration of QKD modules; and short-range QKD to guarantee secure communication between low-cost personal devices and services [20]. The hub aims at the commercialization and affordability of QKD [21].

The European Commission recently announced an ambitious flagship programme to be launched in 2018, in which €1 billion will be granted [62]. The initiative is expected to turn Europe's research in key areas such as quantum communication, quantum sensing, quantum simulation, and quantum computing into real technological opportunities, which can be taken up by industry [22]. For quantum communication, this programme aims to address long-distance communication in ways that will enable QKD networks in metro areas and ranges larger than 1000 km [143].

In Korea, the company, SK Telecom, has developed a QKD network based on trusted repeaters, reaching a distance of 112 km [23]. In 2016, the company applied quantum cryptography technologies to a commercial LTE network in Sejong city [24]. In February 2017, SK Telecom started a collaboration with Nokia and Deutsche Telekom to develop secure communication using quantum cryptography [24]. SK Telecom is currently planning to build a quantum network to share a quantum key between Seoul and Busan (about 460 km) [24]. Five trusted repeaters will be employed.

In Japan, a live demonstration of the Tokyo QKD network was delivered in 2010 [149]. A real-time encrypted video conference, including the detection of an eavesdropper and a quantum link switch, was shown [149]. This project involved both domestic partners (NICT, NEC, NTT, Mitsubishi Electric) and international participants (Toshiba Research Europe, ID Quantique, University of Vienna, etc.) [25]. A three-layer structure was applied based on trusted nodes: a quantum layer, a key management layer, and a communication layer [149].

To communicate over longer distances than that is currently possible through fiber-based QKD networks, satellites can be introduced as middle nodes in the network. This idea was first realized by China, which is becoming a leader in quantum communication. China launched a quantum satellite named Micius in 2016 [26, 84]. The satellite has successfully completed three tasks: 1200-km entanglement distribution [186], quantum

teleportation [142], and QKD [99]. The quantum satellite has also established the key crossing continents between China and Austria [98]. This shared key supported a video call between Beijing and Vienna [98].

In April 2017, the Government of Canada announced an $80.9 million grant over five years to support space-related research [27]. One of the major projects is a demonstration of applications of quantum technologies in space, which involves the Institute for Quantum Computing (IQC) at University of Waterloo [27]. This project aims to position Canada as a leading player in quantum encryption. The goals of this project are to achieve more secure communications, more reliable government services and greater protection of Canadians' privacy.

As QKD enters the initial phases of commercialization and globalization, standardization has become a growing concern. The Telecommunications Standards Institute (ETSI) is aware of this demand and has organized an Industry Specification Group (ISG) to work on a QKD international standard [28]. ISG-ETSI brings together important parties from science, industry, and commence to address standardization issues in quantum cryptography, and quantum technology in general. The ESTI white paper [2] about quantum safe cryptography and security was published, providing a broad view of the task.

Today, governments and industries pay increasing attention to quantum-safe cryptography in order to mitigate impending threats from quantum computing. Moreover, countries all over the world are contributing to QKD technology development, with an emphasis on improved performance and commercialization, but also international cooperation and standardization.

1.3 Practical security of QKD

1.3.1 Quantum hacking

The unconditional security of QKD has been proven in theory [108]. QKD security proofs are based on three factors, beginning with a solid foundation in the laws of quantum physics. The proofs are trustable because they are logical and restrict. Importantly, the proofs also incorporate assumptions and models of practical devices. Although, it should be noted that these assumptions and models cannot precisely match real-life devices. For example, some assumptions may not be satisfied in practice [114], or models might not describe the overall characteristics of the practical QKD devices [117]. As we believe the

theory is perfect, the mismatch between theory and implementation is due to practical imperfections.

Practical implementation presents unique challenges. Imperfect equipment may disclose some loopholes, providing Eve opportunities steal the secret key without being noticed. This is called quantum hacking. Quantum hackers attempt to exploit practical imperfections to compromise the security of QKD. The possibility of quantum hacking has been shown in research experiments and even commercial QKD systems [45, 61, 80, 117, 122, 146, 161, 168, 179]. However, quantum hacking remains at the scientific stage of study, where scientific researchers publish analyses about practical imperfections and vulnerabilities to attacks (as opposed to real hacking). This work helps the entire research community to better understand practical QKD systems. Most importantly, scrutinizing practical security is a necessary phase during the battle-testing period. This flaw analysis gradually enhances the practical security of QKD. From this point of view, quantum hacking evaluates and verifies QKD implementations, especially for commercial systems.

In the standard prepare-and-measure QKD scheme, it is assumed that Eve only has access to the quantum channel, but Alice (source of state preparation) and Bob (state measurement) are in protected laboratories. However, Eve could try to break these assumptions by exploiting flaws from the source or the measurement devices. For the source, imperfect state preparation may leak information about the secret key [164, 165, 168, 181]. For instance, the QKD protocol assumes that quantum states are indistinguishable in the non-encoded degrees of freedom. However, imperfect encoding methods result in side channels from which encoded states are partially distinguishable [132]. Secondly, it is also assumed that Alice prepares the required quantum states correctly. Unfortunately, practical preparation may introduce some errors due to imperfect devices or Eve's disturbance [162]. To steal the information about the states, Eve can also actively perform the Trojan-horse attack [63, 79, 172] on intensity modulators and phase modulators.

The measurement party is usually more vulnerable to quantum hacking than the source. Since Eve sends everything in the same direction as Alice, it is difficult to protect measurement devices from attacks by simply using an optical isolator. So far, a significant number of attacks focus on single-photon detectors [45, 61, 115–118, 176]. For this reason, single-photon detectors are regarded as the "Achilles heel" of QKD [111]. In fact, the vulnerabilities of single-photon detectors are the result of their complex working mechanism: the detection is affected by incoming light and the control circuit. Therefore, Eve can manipulate the intensity [117], the time [121, 139], or the wavelength [96] of incoming light to control the response of single-photon detectors. The detection also really depends on the electronic control and processing, which could have loopholes that can be exploited by Eve [176]. Due to the complex mechanism of the single-photon detector, it is challenging

to fully characterize them in a security model, which can result in the serious attacks on detectors.

1.3.2 Countermeasures

There are four main approaches to bridge the gap between perfect theory and imperfect practice.

The first idea is to precisely characterize and describe the practical devices in mathematical models. Then the models can be included in the security proof to estimate the real secure key rate based on an imperfect setup [180]. While this approach seems straightforward, developing models to fully match the behavior of QKD devices is rarely possible because the components are complex. This approach is also limited by our understanding of the devices. It is remarkable that even though it is hard to fully characterize all the QKD components, there are ongoing efforts to consider as many imperfections as possible into the security models as shown in Ref. [113, 165, 180].

Instead of considering all the characteristics of all components, an effective solution to close the known loopholes is to patch them. More specifically, once one discovers a new type of attack, a corresponding countermeasure against this attack can be proposed and realized in an existing QKD system [103, 146]. This approach usually only requires modifying the software or the hardware (sometimes both) of a current system. Thus, patching loopholes is more feasible and more practical than the previous method. However, the main drawback of patches is that they only prevent the known attacks. For potential and unknown attacks, the countermeasures may fail [122, 148]. Furthermore, the patched countermeasures themselves might open other loopholes, introducing one more layer of security risk [146, 148].

The third method is to apply device-independent QKD (DI QKD) [31, 127]. In this protocol, Alice and Bob are not required to know how their devices work, but simply treat their devices as "black boxes". However, this protocol still incorporates some assumptions. For example, true random number generators, trusted classical post-processing, and an authenticated classical channel are needed. The key assumption is that there is no information leakage from the devices. With these assumptions, the security of DI QKD has been proven based on the violation of the Bell inequality [173]. This idea is exciting as it can remove all known and potential side channels. Unfortunately, this protocol is hard to realize with off-the-shelf technology, because it needs almost perfect single-photon detectors [64]. Furthermore, even if it is realized, DI QKD can only achieve an extremely low key rate [64]. An exciting news is that researchers realized the Bell inequality test that closes

the non-locality loophole and the detection loophole in the same experiment [65, 72, 153], which is a milestone of implementing DI QKD. Advanced technology in the future might make DI QKD more practical.

The fourth solution is to eliminate as many security assumptions about the devices as possible, but still keep some devices under protection. One of the most promising candidates in this approach is measurement-device-independent QKD (MDI QKD; see Sec. 2.4) [110]. MDI QKD removes all the assumptions about the measurement station, which can even be held by Eve. In this way, the protocol is immune to all the side channels of measurement devices [110]; these devices are treated as the weakest part of a QKD system. The security of MDI QKD is equivalent to the Einstein-Podolsky-Rosen (EPR) based QKD protocol [77], as it is a time-reversed version of EPR-based QKD [110]. Remarkably, MDI QKD is also highly practical, and it can be realized by current technology [54, 107, 136, 145, 166, 167, 169, 170, 182]. Additionally, the source stations of MDI QKD still need to be in protected laboratories [110].

1.4 Motivation

The practical security of quantum cryptography has drawn much attention in the past decade. Various types of quantum attacks were disclosed, and researchers proposed solutions to tolerate such imperfections in quantum cryptography, especially in QKD. As most of the discovered loopholes are from detectors, the MDI scheme is treated here as a viable solution to get rid of the threat to vulnerable detectors (see Sec. 2.4).

However, is MDI the end of the story? Once we have a MDI QKD protocol as a tool, can a QKD system reach the security level guaranteed by its theory? The obvious answer is no, because MDI QKD technology is limited in critical ways. The MDI QKD protocol is not compatible with other QKD protocols, which means it cannot be implemented in existing QKD systems. Another technical bottleneck is the relatively low key rate compared to standard prepare-and-measure QKD systems, because MDI QKD requires interference from two individual sources and coincidence clicks as detection events. Most importantly, MDI QKD is still based on an essential assumption: the source stations must be trustable. This assumption may not hold in practice, which compromises the security of MDI QKD.

To further scrutinize the practical security of general quantum cryptographic systems, several questions must be addressed (even after MDI quantum cryptography has been proposed). (i) Except for MDI QKD, what countermeasure(s) can commercial companies deploy in their current QKD systems to avoid detector loopholes? Is this countermeasure

robust enough? (ii) How could we tackle the technical limitations of MDI QKD? Is there any other semi-DI protocol that achieves the same security performance as MDI QKD, but is easier to realize than MDI QKD? (iii) How secure and trustable is the source? Is there any imperfection in the source? Are there more possible attacks on the source? If new types of attacks are discovered, how will they affect the security of MDI QKD?

This thesis is motivated by the above questions and explores their answers. By collaborating with ID Quantique, we first evaluate its countermeasure against a detector blinding attack on the commercial Clavis2 plug-and-play QKD system. The countermeasure is a patch that randomly changes the detection efficiency. We show that the countermeasure is not sufficient to protect the system from the modified detector blinding attack. Then, the second project shifts to verify the security of detector-device-independent QKD (DDI QKD). DDI QKD claimed that its security is equivalent to MDI QKD, and this protocol is easier to realize than MDI QKD with a high key rate. However, our research proves that DDI QKD is insecure for detector-control attacks, fundamentally violating its security statement. For the source, we investigate the implementation of decoy-state QKD and show experimentally that the source is imperfect. The partially distinguishable signal state and decoy state compromise the security of QKD. Apart from the protocol's inherent imperfection, we experimentally show two active attacks: laser seeding attack on laser diodes and laser damage attack on optical attenuators, in which Eve injects different amounts of light into the source to create loopholes.

1.5 Outline

The remaining chapters of the thesis are organized as follows. Chapter 2 provides related background, including theoretical attacks, practical attacks based on imperfect equipment, and MDI QC as a countermeasure. Chapter 3 presents the counterattack on the random-detector-efficiency countermeasure against the detector blinding attack. In Chapter 4, DDI QKD security is directly evaluated and compared against the MDI QKD protocol. Several practical hacking strategies to crack the security of DDI QKD are presented. In Chapter 5, the imperfections in the implementations of the decoy-state protocol are investigated, and a modified security model to tolerate these imperfections is given. In Chapter 6, a laser seeding attack is presented as a tool for Eve to increase the mean photon number of the source. In Chapter 7, a laser damage attack on various attenuators is presented. The experimental results show that this attack is likely to decrease the value of attenuation, which also breaks the assumption about the mean photon number in the QKD system. Chapter 8 briefly introduces two other projects I participated in. In Chapter 9, the thesis

is concluded with an outlook on potential future work.

1.6 List of contributions

1. **Counterattack on the random-detector-efficiency countermeasure**.
   To protect the deployed QKD system from the detector-control attacks, especially the detector blinding attack, a countermeasure that randomizes the detection efficiency was proposed. The first project tested and evaluated this countermeasure implemented in the ID Quantique Clavis2 commercial QKD system. I performed the original and modified detector blinding attacks with help from Shihan Sajeed and Poompong Chaiwongkhot. The experimental results show that this countermeasure is not sufficient to defeat the detector blinding attack. I scrutinized the processes in the gated detector supervised by Vadim Makarov. I also theoretically analyzed the general conditions for a successful blinding attack with assistance from Vadim Makarov. The detailed description is provided in Chapter 3, and the results are published in Ref. [73].

2. **Insecurity of detector-device-independent quantum key distribution**.
   DDI QKD was proposed and claimed that it is secure against all detector side-channel attacks. However, in contrast to its security statement, I, along with Shihan Sajeed, developed several attack strategies that showed DDI QKD is, in fact, insecure against detector side-channel attacks. I tested detector efficiency mismatch under detector blinding attack and proposed the attack of exploiting an imperfect beam splitter. Shihan Sajeed conceptualized the attacks exploiting trigger-pulse-energy-threshold difference under different blinding conditions and imperfect phase modulation. Shihan Sajeed and I contributed equally to this work. The details are provided in Chapter 4, and the results are published in Ref. [147].

3. **Decoy-state QKD with an imperfect source.**
   The decoy-state protocol is now widely used in QKD systems as a countermeasure against the photon-number-splitting (PNS) attack. However, the implementations of the decoy-state protocol may have flaws. In this project, I investigated the possible imperfections in practice. I performed the measurement of two decoy-state implementations: pump-current modulation and external modulation of the laser's intensity. The experiments show a timing side channel. Then, I modeled a PNS attack with support from Zhihong Liu. I participated Shihai Sun in modifying the security proof to tolerate a general imperfect source with distinguishable signal and

decoy states. I applied the security model to the measured results and shown the real secure key rate under the practical imperfections. The details of this project are provided in Chapter 5.

4. **Laser seeding attack on the source.** For a QKD system that employs a weak coherent laser source, a value of mean photon number is designed and fixed during QKD operation. However, this designed mean photon number may be manipulated by Eve. In the fourth project, I performed a laser seeding attack with assistance from Shihai Sun and Poompong Chaiwongkhot. It experimentally showed that Eve could increase the output power of the weak coherent laser in Alice. I conducted the initial and basic security analysis of a decoy-state BB84 QKD system under this attack. More specific description of this project is in Chapter 6.

5. **Laser damage attack on the source.** In Alice's apparatus, an optical attenuator is usually the last component the signal passes through before being transmitted via a quantum channel. The attenuator is used to attenuate light intensity to a single-photon level. However, I found that Eve can shine a high-power laser to modify the attenuation. I first analyzed the optical-power handling capacity of a single-mode fiber. Then I, along with Ruoping Li, tested three types of attenuators: a fixed attenuator, a microelectromechanical system-based (MEMS-based) attenuator, and a manual variable attenuator. The experimental results show that Eve has a chance to decrease the value of attenuation, which leaks more photons from Alice. To further investigate the MEMS attenuator, I supervised Ruoping Li and participated in the testing of actual attenuation boards in a QKD system. The results still show the possibility of decreasing attenuation. The details of this project are described in Chapter 7.

1.7 List of publications related to this thesis

1. **A. Huang**, S. Sajeed, P. Chaiwongkhot, M. Soucarros, M. Legré, and V. Makarov, Testing random-detector-efficiency countermeasure in a commercial system reveals a breakable unrealistic assumption, IEEE J. Quantum Electron. **52**, 8000211 (2016).

2. S. Sajeed, **A. Huang**, S. Sun, F. Xu, V. Makarov, and M. Curty, Insecurity of detector-device-independent quantum key distribution, Phys. Rev. Lett. **117**, 250505 (2016). Contributed-equally authors with S. Sajeed.

3. **A. Huang**, S.-H. Sun, Z. Liu, and V. Makarov, Decoy state quantum key distribution with an imperfect source, arXiv:1711.00597.

More publications are in preparation. In summary, 4 scientific articles on these research projects are in preparation (items No. 1 - No. 4 listed below); a survey on the practical security of quantum communication is being written (item No. 5 listed below); 3 confidential reports on implementation security of industrial systems from 3 different manufacturers have been delivered (items No. 6 - No. 8 listed below).

1. **A. Huang**, R. Li, S. Tchouragoulov, and V. Makarov, Laser damage attacks against optical attenuators in QKD systems.
2. **A. Huang**, Á. Navarrete, S.-H. Sun, P. Chaiwongkhot, M. Curty, and V. Makarov, Laser seeding attack in quantum key distribution.
3. H. Qin, **A. Huang**, and V. Makarov, Short pulse attack on continuous-variable quantum key distribution system.
4. P. Chaiwongkhot, K. Kuntz, Y. Zhang, **A. Huang**, J.-P. Bourgoin, S. Sajeed, N. Lütkenhaus, T. Jennewein, and V. Makarov, Effect of atmospheric turbulence on spatial-mode detector efficiency mismatch.

1.8 List of conference presentations

1.8.1 Presented by me

1. **A. Huang**, S. Sajeed, P. Chaiwongkhot, M. Soucarros, M. Legré, and V. Makarov, Gap between industrial and academic solutions to implementation loopholes: testing random-gate-removal countermeasure in commercial QKD system (poster),
presented at QCrypt 2015, Tokyo, Japan, September 28 - October 2, 2015.
2. **A. Huang**, S. Sajeed, P. Chaiwongkhot, M. Soucarros, M. Legré, and V. Makarov, Gaps between industrial and academic solutions to implementation loopholes in QKD: testing random-detector-efficiency countermeasure in a commercial system (poster),
presented at Trustworthy Quantum Information (TyQI) Workshop, Shanghai, China, June 27 - 30, 2016.

3. **A. Huang**, S. Sajeed, S. Sun, F. Xu, V. Makarov, and M. Curty, Insecurity of detector-device-independent quantum key distribution (contributed talk),
presented at QCrypt 2016, Washington DC, USA, September 12 - 16, 2016.

4. **A. Huang**, S. Sajeed, P. Chaiwongkhot, M. Soucarros, M. Legré, and V. Makarov, An advanced Eve of QKD: breaking a security assumption and hacking a black box (poster),
presented at QCrypt 2016, Washington DC, USA, September 12 - 16, 2016.

5. **A. Huang**, S. Sajeed, S. Sun, F. Xu, V. Makarov, and M. Curty, Insecurity of detector-device-independent quantum key distribution (poster),
presented at 4th ETSI/IQC Workshop on Quantum-Safe Cryptography, Toronto ON, Canada, September 19 - 21, 2016.

6. **A. Huang**, S. Sajeed, P. Chaiwongkhot, M. Soucarros, M. Legré, and V. Makarov, An advanced Eve of QKD: breaking a security assumption and hacking a black box (poster),
presented at 4th ETSI/IQC Workshop on Quantum-Safe Cryptography, Toronto ON, Canada, September 19 - 21, 2016.

7. **A. Huang**, Can quantum physics break cryptography's curse? (invited talk)
presented at CAQCR 2017, Nanchang Jiangxi, China, July 27 - 28, 2017.

8. **A. Huang**, S.-H. Sun, Z. Liu, and V. Makarov, Decoy state quantum key distribution with imperfect source (poster),
presented at QCrypt 2017, Cambridge, UK, September 18 - 22, 2017.

1.8.2 Presented by my coauthors (the first person in each author list)

1. S. Sajeed, **A. Huang**, S. Sun, F. Xu, V. Makarov, and M. Curty, Insecurity of detector-device-independent quantum key distribution (3rd prize for best poster),
presented at ICNFP 2017, Crete, Greece, August 17 - 29, 2017;

2. P. Chaiwongkhot, K. B. Kuntz, **A. Huang**, J.-P. Bourgoin, S. Sajeed, N. Lütkenhaus, T. Jennewein, and V. Makarov, Effect of atmospheric turbulence on spatial-mode detector-efficiency mismatch (poster),
presented at QCrypt 2017, Cambridge, UK, September 18 - 22, 2017.

3. H. Qin, **A. Huang**, and V. Makarov, Short pulse attack on continuous-variable quantum key distribution system (poster),
presented at QCrypt 2017, Cambridge, UK, September 18 - 22, 2017.

Chapter 2

Background of quantum hacking

Cryptanalysis studies the capability of cryptography focusing on investigating the weaknesses therein. In quantum cryptography, a specific term "quantum hacking" is widely used to represent the cryptanalysis in which Eve's ability is allowed by quantum mechanics. In this chapter, theoretical and practical attacks on quantum cryptography are introduced. As a remarkable countermeasure against detector side channels, the idea of MDI QKD is also presented.

2.1 Individual attack, collective attack, and coherent attack

In theory, Eve's attack strategies are classified into three families. The first family is called individual attack [60]. In this type of attack, Eve is limited to probing each qubit independently and measuring it one after the other. The individual attack also requires that Eve must perform her measurement before post-processing. This attack is the simplest one because it does not need any quantum memory. An important sub-family is the intercept-resend attack that is applied in many practical attacks. I will explain this subfamily in detail in the next section.

Another family of attack is collective attack [40]. In this attack, Eve still probes each qubit independently, but she can keep her probes in the quantum memory. Later, after post-processing, or whenever it is convenient for her, Eve performs an optimal measurement to obtain the maximum amount of information. Please note that in the collective

attack, Eve is allowed to measure several probes coherently, which constitutes collective measurement.

The last family of attack is coherent attack [39], in which Eve can probe and measure qubits coherently. Generally speaking, Eve is allowed to perform any attack allowed, only restricted by the laws of quantum mechanics. The hacking strategies are various, e.g., Eve can entangle with qubits transmitted from Alice to Bob, or she can adjust her attack strategy according to intermediate measurement results. By optimizing the hacking strategy, she can obtain as much information as possible.

2.2 Intercept-resend attack

As aforementioned, the intercept-resend attack [34] is a typical and simple individual attack. In this attack, Eve first intercepts the qubits individually and measures each of them in her interception basis. She then prepares new states according to her measurement results and resends them to Bob. Being that Eve chooses measurement bases independently without any information about Bob's bases choices, she indeed introduces some errors into the sifted key. The simplest attack for Eve is to measure the qubit in a basis selected randomly between the bases used by Alice. For example, Eve can perform this attack in the BB84 protocol [35]. After the raw key exchange and the sifting, Alice and Bob keep the slots if they choose the same bases, while Eve can only correctly guess half of them. The correct guesses allow Eve to precisely measure Alice's states and resend them to Bob, which introduces no error. For the other half of the slots, Eve chooses different bases from Alice's, so the measurement results are inconclusive. The resent states result in random detections at Bob's side: half of them are correct, but the other half of them are wrong. Therefore, Eve introduces 25% error overall in the sifted key, if there are no other sources of error and noise. This intercept-resend attack makes secure key generation become impossible in principle [51, 52].

It is wise to mention a fake-state attack here [117, 123], as it is regarded as the modified version of the intercept-resend attack. In particular, the intercept phase of this attack is the same as that of the intercept-resend attack. However, instead of reconstructing and resending original states to Bob, Eve can resend tailored states, e.g., stronger intensity, different wavelengths, or different triggering time. In this way, Eve aims to control Bob's detection results by exploiting detectors' flaws. This hacking strategy is usually combined with other practical attacks to achieve better performance. I will show several specific examples od practical attacks later in the next section.

2.3 Attacks based on practical imperfections

The flaws in implementations can help Eve discover the secret key. More importantly, by exploiting the practical imperfections, Eve's attacks can be hidden from Alice and Bob. Hence, they are not aware of the attacks. Several representative attacks based on imperfect implementations will be explained in this section, since they are relevant to my Ph.D. research.

2.3.1 Photon-number-splitting attack

In the ideal QKD protocol, a single-photon source is assumed. Unfortunately, implementing the single-photon source is challenging. Instead, weak coherent sources are usually utilized in practical QKD systems. The optical pulses are attenuated to a single-photon level with a mean photon number $\mu < 1$. Please note that the photon number distribution of a weak coherent source follows a Poisson distribution, so there must be a portion of pulses containing multiple photons. These multi-photon pulses leak the information about prepared states to Eve. An attack named photon-number-splitting (PNS) attack [34, 43, 75, 114] exploits this flaw.

The specific steps of this attack are as follows. For every pulse sent from Alice, Eve first performs quantum non-demolition measurement (QND) to know the photon number in the pulse. Upon finding a multi-photon pulse, Eve splits one photon from the pulse and keeps this photon in her quantum memory. The rest of the photons in this pulse are forwarded to Bob. This operation does not disturb the states prepared by Alice, so it introduces no error. For the pulses only containing a single photon, Eve blocks them. Thus, Bob has no information about them at all. The split pulses have a weaker intensity, which is equivalent to the loss during transmission. This loss can be compensated by replacing the lossy channel with a lossless channel between Eve and Bob. In this way, Alice and Bob do not notice the attack. After Bob measures each pulse and announces the basis he selects, Eve measures the photon in her quantum memory by the same basis.

2.3.2 Detector-efficiency mismatch attacks

To measure two different bit values, there are usually at least two detectors in a QKD system. To detect a single photon, InGaAs/InP avalanche photodiodes (APDs) are widely used. This type of APD often works in a gated mode, which means that APDs are only

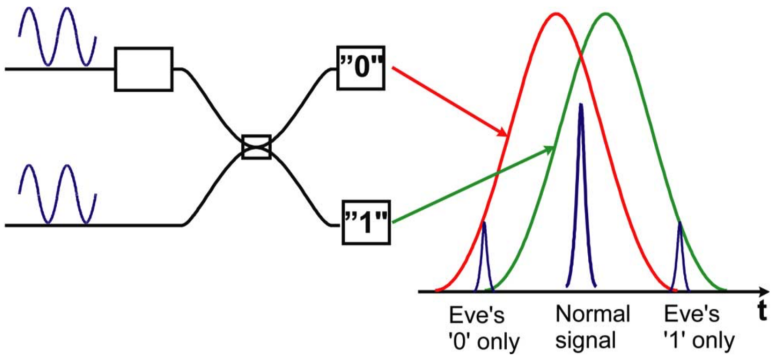

Figure 2.1: Mismatch of detector efficiencies. Reprinted from [121].

sensitive to a single photon during the gated time. As the detection efficiency is time-dependent, it is difficult to perfectly match the same detection efficiency from two different detectors. This is so because of the inherent manufacturing difference and different optical lengths coupled to detectors. The misalignment of the gate time causes the detector-efficiency mismatch [121, 139] as shown in Fig. 2.1.

The loophole of detector-efficiency mismatch allows Eve to perform a fake-state attack to control Bob's detection results. Instead of resending photons to Bob's detectors in the middle of a gate, shown by the position of the normal signal in Fig. 2.1, Eve shifts the arriving time of the photons to the efficiency mismatch areas, in which one detector is more sensitive than the other. For example, if Eve sends photons earlier than normal, she has a higher chance of triggering Detector "0". Conversely, the photons arriving at Bob later than normal may trigger only Detector "1", as it is shown in Fig. 2.1. Thus, by controlling the arrival time of photons, Eve can control the detection results. Besides the time domain, an efficiency mismatch could also exist in other degrees of freedom, e.g., the spectral and spatial domains [132, 146].

A time-shift attack also exploits the detector-efficiency mismatch, as was proposed in Ref. [139] and experimentally verified in Ref. [189]. It is remarkable that this was the first experimental attack on a commercial QKD system. Its principle is similar to that of the attack described above. The main difference is that, instead of performing a faked-state attack, Eve does not intercept the states, but just randomly shifts the arrival time of the incoming photon by changing the length of the transmission fiber. This operation achieves the same result as the fake-state attack described above. Thus, by manipulating the arrival time, Eve can also control the detection results.

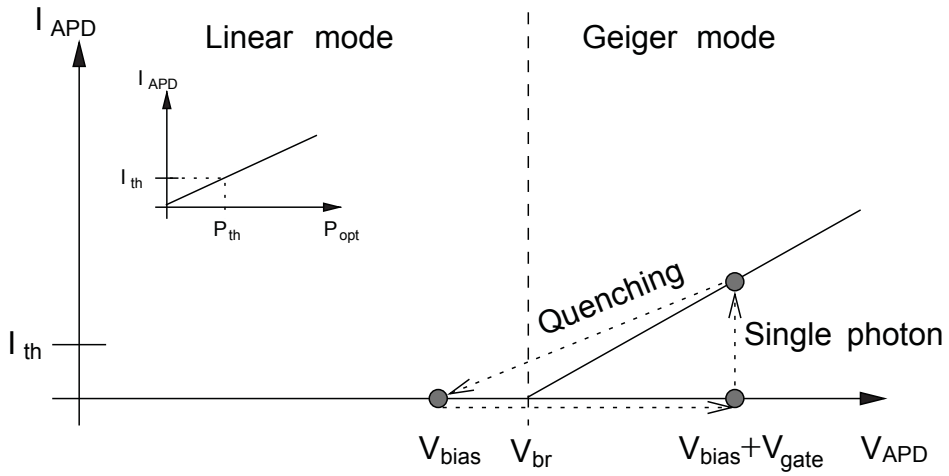

Figure 2.2: Linear-mode and Geiger-mode APD operation (reprinted from [117]).

2.3.3 Wavelength-dependent attack

In the polarization-encoding QKD systems, as described in Sec. 1.1.2, a passive selection of measurement bases is commonly used at Bob's side [68, 74, 106, 152]. In this scheme, a 50:50 beam splitter (BS) is applied to randomly pass the incoming photons to the X basis (from output 1) or the Z basis (from output 2). This is an ideal case. However, in practice, researchers find that the output splitting ratio of BS is not always 50:50 [96]. The splitting ratio is however dependent on the wavelength. It has been shown in an extreme case that the splitting ratio of the BS made by fused biconical taper technology is 98.6:1.4 at 1290 nm, while it becomes 0.3:99.7 at 1470 nm [96].

This changeable ratio can help Eve manipulate the detection results during the intercept-resend attack. If Eve obtains a measurement result in the X basis, she resends the measured results at 1290 nm to Bob. Thus, the resent states likely go through output 1 of BS and are detected by Bob in the X basis as well. Conversely, if Eve measured in the Z basis, she resends the states at 1470 nm. This results in a much higher chance of Bob's detection at the Z basis. The wavelengths are the tools for Eve to control the detection results at Bob's side actively. Please note that the detection efficiency is also wavelength-dependent. To compensate for lower detection efficiencies at 1290 nm and 1470 nm, Eve resends the states with higher intensities to obtain a similar detection efficiency at 1550 nm.

2.3.4 Detector control attack

Most available single-photon detectors are InGaAs/InP APDs operating in a Geiger mode, in which they are sensitive to a single photon [50]. The working principle of this type of APDs is shown in Fig. 2.2. When the APD is reverse-biased above its breakdown voltage V_{br}, a single photon can cause a large current I_{APD} to flow. If this current exceeds the threshold I_{th}, then the electronics register this as photon detection (a 'click'). After that, an external circuit quenches the avalanche by lowering the bias voltage V_{APD} below V_{br}, and then the APD goes into a linear mode. In the linear mode, I_{APD} is proportional to the incident bright optical power P_{opt}. The current threshold I_{th} then becomes a threshold on the incident optical power P_{th} that makes a click [117].

If Eve sends a bright continuous-wave (c.w.) illumination to the gated detectors, then the bright light makes the APD generate a significant photocurrent that monotonically increases with P_{opt}. This large current reduces the voltage across the APD V_{APD} [117]. If we apply enough illumination power, V_{APD} will be less than V_{br} even inside the gate, and the APD then always stays in the linear mode. Consequently, the detector becomes blinded to single photons. After blinding Bob's detectors, Eve can conduct a faked-state attack. Eve first intercepts all photons sent by Alice. Whenever Eve detects a photon, she sends the same state to Bob via a bright-light pulse with specific energy, superimposed on her blinding illumination. Only if Bob chooses the same measurement basis as Eve and applies the gate, does one of Bob's detectors click, and he will get the same bit value as Eve. Otherwise, there is no click at Bob's side. During the sifting procedure, Alice and Bob keep the bit values when they have chosen the same basis, and so does Eve. Therefore, Eve has identical bit values to Bob, introduces no extra QBER, and does not increase the alarm counter. Eve then listens to the public communication between Alice and Bob and performs the same error correction and privacy amplification procedures as them to obtain the identical copy of their secret key [117].

2.3.5 Laser damage attack

In QKD theory, there is no limitation on the power of Eve's light. In other words, Eve can send as much light as she wants to QKD systems. This strong light may change the characteristics of some optical components employed in QKD systems, which is called the laser damage attack [45, 122]. The effect of a high-power laser was first tested for stand-alone APDs [45]. The testing shows that 1.2-1.7 W laser can permanently blind the APDs, making them insensitive to single photons. Nevertheless, the blinded APDs are still

sensitive to a certain amount of bright light. Thus, the detector control attack becomes possible again.

To investigate the capability of laser damage attack on the practical QKD systems further, a fiber-based plug-and-play QKD system and a free-space QKD system were tested [122]. For the fiber-based system, the first component that the high-power laser destroys at Alice's side is a monitoring detector that monitors the energy of incoming light [148]. The power above 1 W reduces the photosensitivity of the InGaAs p-i-n photodiodes. The power of 1.7 W can entirely damage the photodiodes, which means they lose the capability to monitor the injected light. A Trojan-horse attack [63, 79, 172] then becomes possible. For the free-space QKD system, the first component destroyed is a spatial filter, a pinhole. The 3.6 W laser enlarges the size of the pinhole, so it cannot restrict the area and beam size of the incoming light. Without the protection of the pinhole, Eve can exploit the efficiency mismatch in the spatial domain to hack the QKD system.

2.4 Countermeasure against detector control attacks: measurement-device-independent QC (MDI QC)

2.4.1 Idea of MDI

The idea of measurement device independence is inspired by an Einstein-Podolsky-Rosen (EPR) based QKD protocol [38, 58]. In this protocol, Alice and Bob individually prepare an EPR pair at each side and send one photon from each pair to an untrusted party, Charles. Charles then performs a Bell state measurement (BSM) to swap entanglements. The measurement result is announced. Once the BSM is finished, Alice and Bob measure the other photon of the EPR pairs locally by randomly choosing between the X and Z bases. Comparing a subset of their measurement results allows Alice and Bob to know whether Charles is honest.

Importantly, the EPR protocol can also work in a "time-reversal" version [38], in which the order of measurements is reversed. Thus, Alice and Bob can measure their local photons first, instead of waiting for Charles' measurement results. This order of preparation and measurement is equivalent to that of the prepare-and-measurement QKD scheme in which Alice and Bob prepare BB84 states [35] and send them to Charles to perform the BSM. After that, the Charles' honesty can still be checked by comparing a part of Alice's and Bob's results. This time-reversal EPR protocol is the main concept behind MDI. The

advantage of MDI QC is that it removes all detector side channels because the third-party Charles who performs the measurement can be fully untrusted.

Moreover, the idea of MDI can be extended to a multi-party scenario [59]. Instead of using the EPR state, a Greenberger-Horne-Zeilinger (GHZ) entangled state can be established among multiple parties. Similar to the EPR version, the security of multi-party quantum cryptography is based on post-selected GHZ states. The correlation in the GHZ entangled state guarantees the security. This security is also independent of the measurement station, which can be untrusted as well. Most importantly, this post-selected entanglement scheme is applicable not to only QKD, but also to other quantum cryptographic protocols, such as quantum digital signature (QDS) [185] and quantum cryptographic conferencing (QCC) [48].

2.4.2 MDI QKD protocol

To understand the MDI QC further, I take the MDI QKD protocol [110] as an example and explain it in detail. The MDI QKD protocol can be divided into several phases:

Phase I: Alice and Bob randomly and individually prepare one of four BB84 states [35]. They then send the states to an untrusted party, Charles.

Phase II: An honest Charles performs a BSM that makes Alice's and Bob's states interfere with each other, generating a Bell state. The quantum communication phases are completed at this point. The rest of the phases use only the classical public channel.

Phase III: Whether Charles is honest or not, he announces the outcome of BSM when he obtains a successful measurement.

Phase IV: Post-processing. Alice and Bob keep the data that corresponds to Charles' successful measurement events and discard the rest. Next, similar to the sifting in BB84 protocol, Alice and Bob announce their basis choices for sifting the events and keep the events using same bases. Based on Charles' measurement result, Alice flips part of her bits to guarantee the correct correlation with those of Bob.

MDI QKD protocol is very practical and can be implemented by current technology. A typical implementation scheme based on polarization encoding is shown in Fig. 2.3. In practice, Alice and Bob prepare phase-randomized weak coherent pulses (WCP) first. Individual BB84 polarization states are independently and randomly modulated by a polarization modulator (pol-M). To protect the weak coherent source from the PNS attack, a decoy state method [112] is applied with the help of an intensity modulator (decoy-IM) [110]. Alice and Bob simultaneously transmit the prepared photons to Charles, and

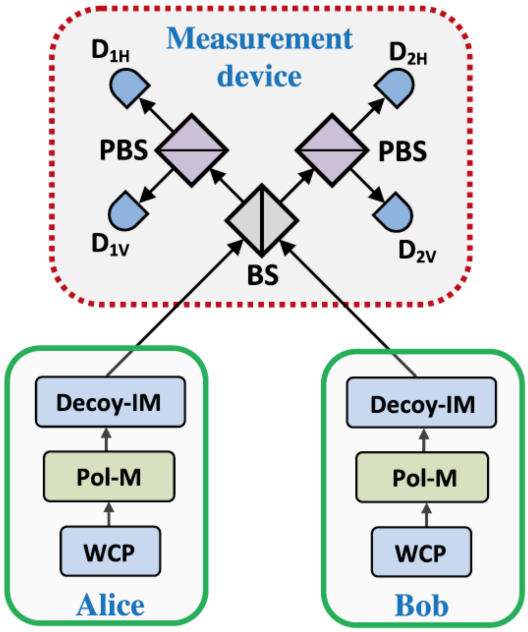

Figure 2.3: Basic schematics of MDI QKD implementation. Reprinted from [110].

the photons interfere at a 50:50 beamsplitter (BS) at Charles' station. Each output of the BS is followed by a polarization beam splitter (PBS) that projects the incoming photons to H or V polarization. Four single-photon detectors ($D_{1\mathrm{H}}$, $D_{1\mathrm{V}}$, $D_{2\mathrm{H}}$, and $D_{2\mathrm{V}}$) are employed to detect the photons at each output of PBS.

In this case, a successful BSM corresponds to two detectors being triggered simultaneously. The detection combination of D_{1H} and D_{1V}, or D_{2H} and D_{2V}, means a projection into the Bell state $|\psi^+\rangle = 1/\sqrt{2}(|HV\rangle + |VH\rangle)$, while the combination of D_{1H} and D_{2V}, or D_{1V} and D_{2H}, indicates a projection into the Bell state $|\psi^-\rangle = 1/\sqrt{2}(|HV\rangle - |VH\rangle)$. Alice flip her bits to be identical to Bob's. Only when Alice and Bob choose the diagonal basis and the measured result is $|\psi^+\rangle$, does Alice not flip her bits.

2.4.3 Limitations of MDI QKD

MDI QKD is indeed a milestone in the progress of quantum cryptography. It eliminates the threat of imperfect detections. However, MDI QKD protocol still has some limitations. One major concern from industry is how to apply it to commercial products. MDI QKD is a new protocol that is not compatible with the existing QKD systems. Replacing all

the deployed systems with MDI QKD systems is costly. Thus, instead of using MDI QKD scheme, the industry should consider patching detection loopholes to strengthen the existing QKD systems.

A technical drawback of MDI QKD is that it requires high-visibility two-photon interference between two independent sources. This requirement makes its implementation more demanding than that of conventional QKD schemes, as it is hard to maintain the identical characteristics, e.g., wavelength, arrival time and pulse width, for two individual sources. In addition, the current finite-key security analysis against general attacks [53] requires larger post-processing data block sizes of MDI QKD than those of standard prepare-and-measure QKD, even though recent proposals [190] significantly improve the performance of MDI QKD in the finite-key regime. These two technical obstacles limit the secure key rate of MDI QKD.

There is no doubt that the MDI QKD scheme can remove all security assumptions about the detection devices. Thus, it eliminates all detector side channels from QKD implementations, which are regarded as significant imperfections in the QKD system. Please note that an essential assumption in MDI QKD is that the source stations are trusted. That is, Alice and Bob are believed to be located at secure laboratories under protection and fully know the prepared states. However, this assumption might not be satisfied in a realistic scenario. Instead, Eve might exploit side channels of the sources, like imperfect state preparation [132] and the active Trojan-horse attack [63, 79, 172], to compromise the security of MDI QKD. The sources may become the "Achilles' heel" of MDI QKD.

Chapter 3

Counterattack on random-detector-efficiency countermeasure

As one of the limitations of MDI QKD mentioned in the past chapter, it is hard to adapt the deployed QKD systems to the MDI scheme. Alternatively, the solutions from industry are attempting to patch the existing systems against specific attacks. Importantly, the security of these patches should be verified. In this chapter, an example of testing the security of an implemented countermeasure is given. I examine ID Quantique's attempted countermeasure to earlier discovered bright-light detector control attacks [117, 118, 177]. It shows that the countermeasure can be counterattacked. This work is published in Ref. 73.

3.1 From loophole discovery to countermeasure implementation

In 2009, the vulnerability of the commercial QKD system Clavis2 [29] to detector blinding attacks was identified and a confidential report was submitted to ID Quantique (the work was published shortly afterwards [117]). After this, ID Quantique has been trying to figure out an experimental countermeasure against these attacks. The timeline of this security problem is shown in Fig. 3.1. In 2010, ID Quantique proposed a countermeasure that randomizes the efficiency of a gated avalanche photodiode (APD) by randomly choosing one out of two different gate voltages, and filed this idea for a patent [30]. In this way, an

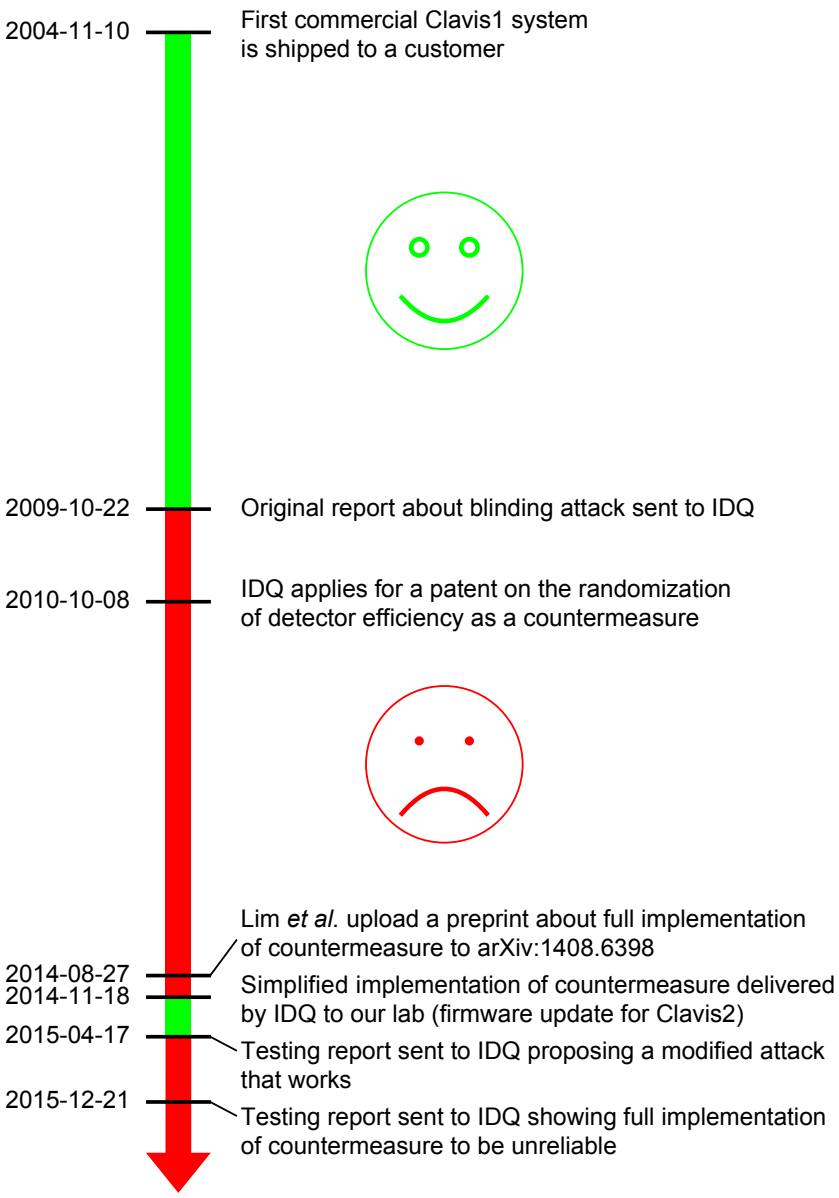

Figure 3.1: Timeline of hacking-countermeasure-hacking for the bright-light detector control class of attacks.

eavesdropper Eve does not know the exact efficiency of Bob in every gated slot and thus cannot maintain his detection statistics. At the sifting phase, if the observed detection rates differ from the expected values, Alice and Bob would be aware of Eve's presence and discard their raw keys.

In 2014, Lim *et al.* proposed a specific protocol to realize this countermeasure [103], which analyses the security mathematically for blinding attacks that obey a certain assumption on their behaviour. In the protocol, Bob randomly applies two non-zero detection efficiencies $\eta_1 > \eta_2 > 0$, and measures detection rates R_1 and R_2 conditioned on these efficiencies. The effect of detector blinding attack is accounted via the factor $(\eta_1 R_2 - \eta_2 R_1) / (\eta_1 - \eta_2)$. Without the blinding attack, the detection rate is proportional to the efficiency, making this factor zero. The analysis makes a crucial assumption that the detection rate under blinding attack $R_1 = R_2$, i.e., it will be independent of Bob's choice of $\eta_{1,2}$. Under attack the factor then will be greater than zero, and reduces the secure key rate. This solution intends to introduce an information gap between Eve and Bob, for Eve has no information about Bob's random efficiency choice.

Later in 2014, ID Quantique implemented the countermeasure as a firmware patch. The hardware in Clavis2 is not capable of generating two nonzero efficiency levels that switch randomly between adjacent detector gates. As a result, implementation is in a simple form by suppressing gates randomly with 2% probability. The suppressed gates represent zero efficiency $\eta_2 = 0$, while the rest of the gates represent calibrated efficiency $\eta_1 = \eta$. Ideally, in the updated system, there should be no click in the absence of the gate. In practice, transient electromagnetic interference may extremely infrequently lead to a click without a gate. Therefore, an alarm counter is used with the system lifetime limit of 15 clicks in the absence of the gate. If this limit is reached, it triggers the firmware to brick the system and requires factory maintenance. This implementation assumes that under blinding attack [117], click probability should not depend on the gate voltage and the attack should, therefore, cause clicks at the slots of gate absence.

3.2 Counterattack on the countermeasure

I demonstrate that the countermeasure presently implemented by ID Quantique is effective against the original blinding attack [117], but not sufficient against the general class of attacks attempting to take control of Bob's single-photon detectors.

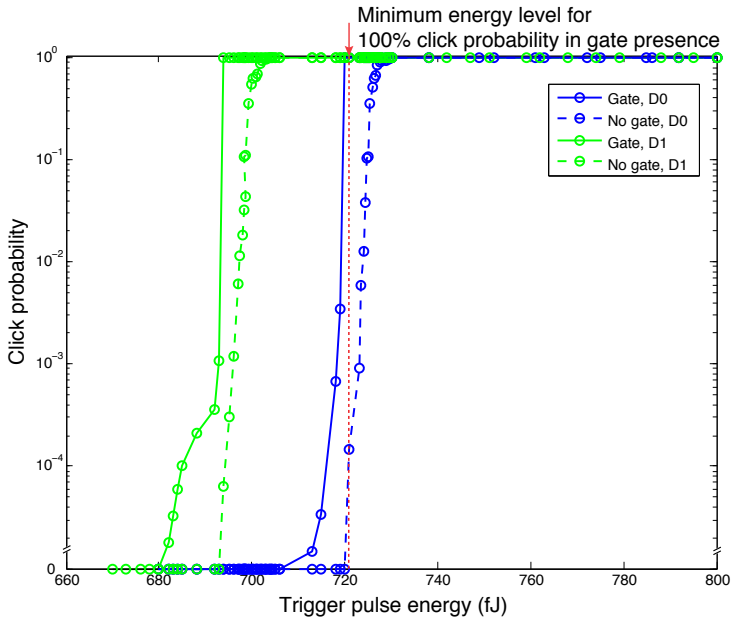

Figure 3.2: Click probability under original blinding attack [117] versus energy of trigger pulse.

3.2.1 Hack by the original blinding attack

Let me briefly remind the reader how Clavis2 and the original blinding attack against it work. Clavis2 is a bidirectional phase-encoding QKD system [29, 159]. After Bob sends multi-photon bright pulses to Alice, Alice randomly modulates one of the four BB84 phase states [35], attenuates the pulses and sends them back to Bob. Bob randomly chooses one out of two measurement bases. Interference happens between pulses from longer and shorter paths of an interferometer at Bob's side, and the outcomes of interference depend on the phase difference between Alice's and Bob's modulation [131]. However, Eve is able to control the outcomes by the following strategy. She shines a bright light to blind the detectors, and then intercepts Alice's states [117]. According to Eve's interception results, she re-sends faked states by multi-photon pulses to Bob's blinded detectors. If Bob chooses the same measurement basis as Eve's, the pulses interfere at Bob's interferometer, so that all power of the pulse goes to one detector to trigger a click. If the measurement bases chosen by Bob and Eve are mismatched, there is no interference, and the power of the pulse is split equally between Bob's two detectors. In this case, neither detector clicks. In this attack, Eve can fully control Bob's detectors and obtain the whole key tracelessly [117].

For the original blinding attack, Eve sends bright-light continuous-wave (c.w.) laser light to blind Bob's detectors. A trigger pulse then is sent slightly after the gate to make a click. I repeat this attack for an improved Clavis2 system and test the amount of energy to trigger a click which is shown in Fig. 3.2. In this testing, the blinding power is 1.08 mW, as the same as the power used in the published original attack [117]. The timing of trigger pulse is 0.7 ns long, 3 ns after the center of the gate signal, which should roughly reproduce the original attack [117]. From Fig. 3.2, we can see the trigger pulse energy for gate presence (solid curves) is lower than that for gate absence (dashed curves), because minute electrical fluctuations of APD voltage following the gate signal lower the click threshold slightly.

However, if Eve tries to trigger a click with 100% probability when the gate is applied, this amount of trigger pulse energy (marked by a dotted vertical line in Fig. 3.2) also might trigger a click with non-zero probability when the gate is suppressed, which is monitored and results in an alarm. Therefore, Eve cannot hack the system with full controllability. To avoid clicks in slots of gate suppression, Eve could, in theory, decrease the level of trigger pulse energy to trigger a click sometimes with gate presence, but never with gate absence. This also satisfies a necessary condition of a successful attack which we will discuss in Sec. 3.3 later. Unfortunately, in practice, the testing result shows the amount of trigger pulse energy required to trigger D0 without the gate is about 710 fJ, which is only 1.5% less than the amount of energy for 100% click (720 fJ) when the gate is present. The 1.5% difference of these two energy levels is likely not big enough to achieve a reliable attack operation that avoids triggering the countermeasure. Also, D1 will always trigger at these energy levels, revealing the attack. Eve could target D1 using a slightly lower energy level, but the relative precision required is similar there. Routine fluctuations of temperature and other equipment parameters may lead to some instability of these trigger pulse energy levels, causing a risk for Eve to trigger a few clicks in the gate absence and brick the system being attacked. From this point of view, we think this first implementation of countermeasure is effective against the original blinding attack.

3.2.2 Hack by the modified blinding attack

I slightly modify the blinding attack to break the security of this countermeasure. Similarly to the original blinding attack, Bob's detectors are blinded by a bright-light laser first. Then, instead of sending a trigger pulse slightly after the gate as in the original attacks [117], I send a 0.7 ns long trigger pulse on top of the c.w. illumination *during the detector gate*, as shown in Fig. 3.3. The relative time between the gate voltage transitions and the optical pulse is approximate. The c.w. signal is generated by a 1536 nm laser diode; the trigger pulse signal is obtained by modulating pump current of a separate 1551 nm

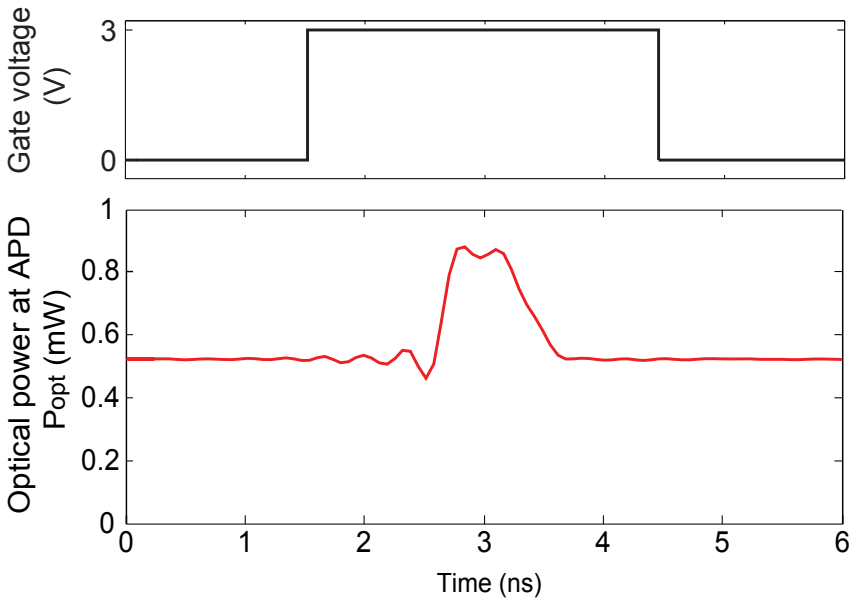

Figure 3.3: Idealized APD gate signal and real oscillogram of optical trigger pulse.

laser diode, using an electrical pulse generator [117]. This trigger pulse produces a click in one of Bob's two detectors only if Bob applies the gate and his basis choice matches that of Eve; otherwise there is no click.

To explain why this modified attack succeeds, let me remind the reader the normal operation of an avalanche photodiode (APD). The detectors in Clavis2 are gated APDs. When the gate signal is applied, the voltage across the APD V_{APD} is greater than its breakdown voltage V_{br}. If a single photon comes during the gated time, an avalanche happens and causes a large current. This current is converted into a voltage by the detector electronic circuit. If the peak voltage is larger than a threshold $V_{\mathrm{th}} = 70$ mV, the detector registers a photon detection (a 'click'). Fig. 3.4(a) and (b) show the cases of no photon coming and a photon introducing an avalanche.

A bright laser is able to blind the APDs. Under c.w. illumination, the APD produces constant photocurrent that overloads the high-voltage supply and lowers V_{APD}. Then, even when the gate signal is applied, V_{APD} does not exceed V_{br} and the APD remains in the linear mode as a classical photodetector that is no longer sensitive to single photons. This means the detectors become blinded.

Under the blinding attack with 0.56 mW c.w. illumination, Fig. 3.4(c–e) shows the detector voltages in different cases: when (c) no trigger pulse is applied and when 0.32 pJ

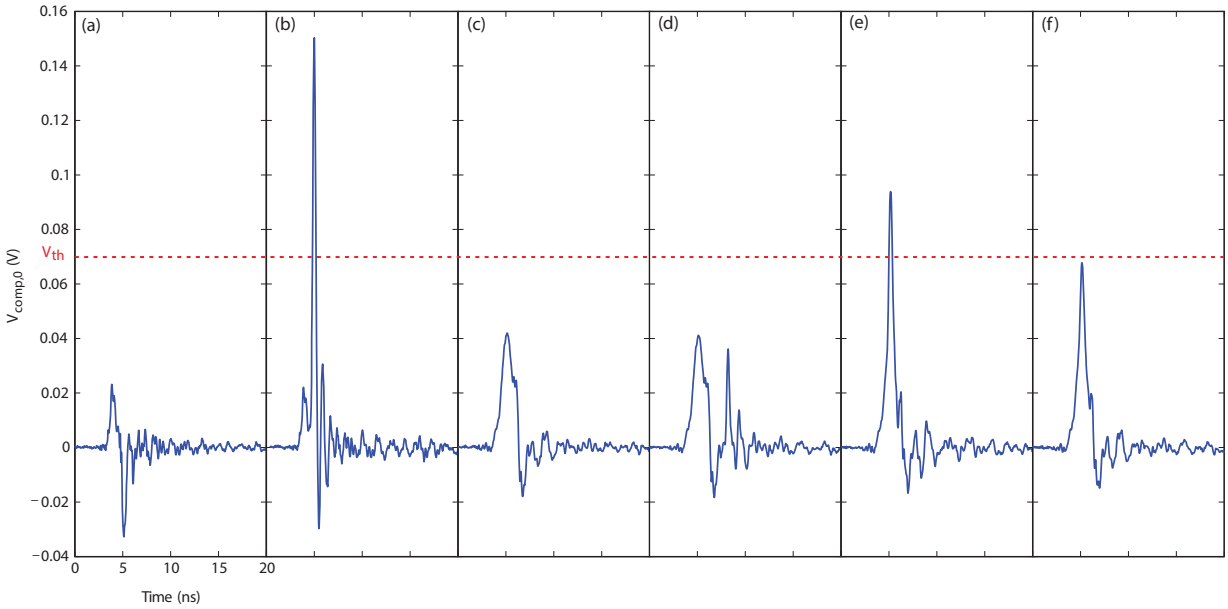

Figure 3.4: Oscillograms at comparator input in the detector circuit, proportional to APD current.

trigger pulse is applied either after (case (d)) or in the gate (case (e)). Since in the linear mode the gain factor of secondary electron-hole pairs generation in the APD depends on the voltage across it, the 3 V gate applied to the APD increases the gain factor. This larger gain during the gated time assists the APD in generating a larger photocurrent than the photocurrent outside the gate. Therefore the gate signal causes a positive pulse as shown in Fig. 3.4(c). The trigger pulse applied after the gate produces a second pulse, but the peak voltages of neither pulses exceed V_{th} [Fig. 3.4(d)]. However, when the trigger pulse is shifted inside the gate, the two pulse amplitudes add up, reach V_{th} and produce a detector click [Fig. 3.4(e)]. If Bob chooses a different measurement basis than Eve, only half of the trigger pulse energy (0.16 pJ) arrives at each detector [117]. In this case, the peak voltage does not reach V_{th} [Fig. 3.4(f)]. Overall, only when the trigger pulse is applied during the gate time and Bob chooses the same basis as Eve, the detector under the blinding attack clicks. As a result, Eve can control Bob's detectors to make Bob obtain the same measurement result as her, and does not introduce extra errors [117].

Contrary to most of previously demonstrated attacks attempting to take control of single-photon detectors [115, 117, 118], in the present demonstration the timing of the trigger pulse has to be aligned with the gate. Besides timing alignment, another important

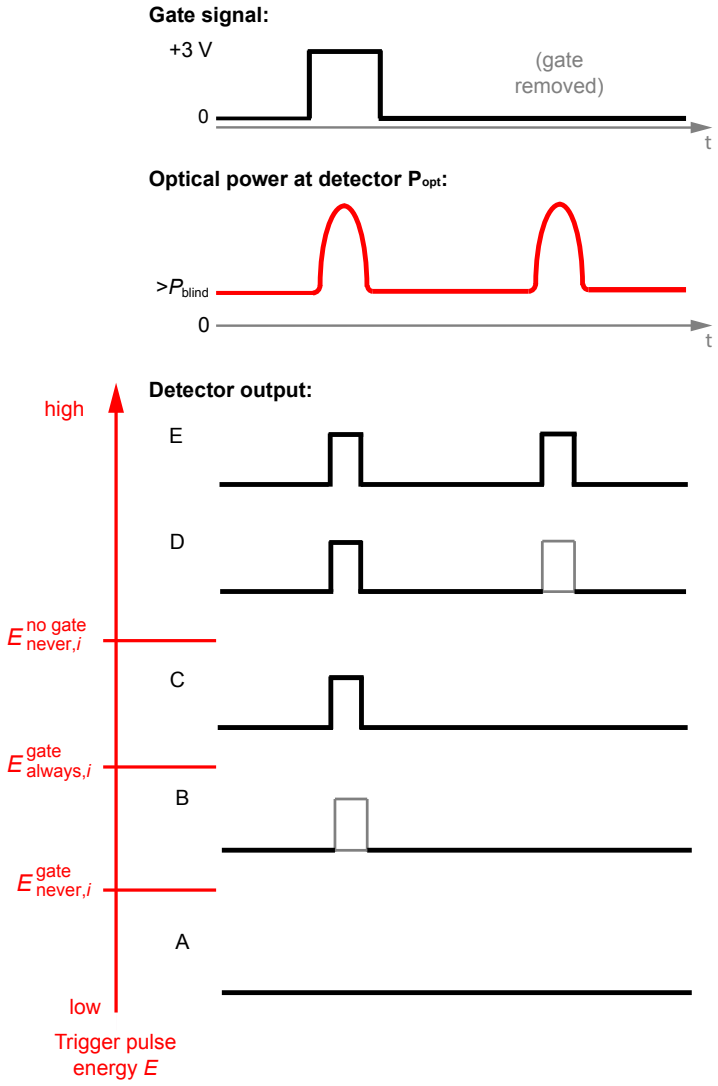

Figure 3.5: Output of a blinded detector in Clavis2 under control of trigger pulses of different energy.

factor of the attack is the trigger pulse energy E. To test the effect of different trigger pulse energies, I gradually increase it and observe the detection outcomes. Figure 3.5 shows schematically in which order clicks appear in Clavis2 as E is increased. The top graph shows a gate applied at the first slot, but suppressed at the second slot. However, an optical trigger pulse is sent to the detector in both slots. Graphs A–E show detector output versus trigger pulse energy E. In graph A, the energy is insufficient to produce a click. As the energy is increased above $E^{\text{gate}}_{\text{never},i}$, clicks intermittently appear in the presence of the gate, as shown in graph B. At the energy level above $E^{\text{gate}}_{\text{always},i}$, the gate always has a click, as shown in graph C. However, there is never a click when there is no gate. At a higher energy level above $E^{\text{no gate}}_{\text{never},i}$, clicks in the gate absence appear intermittently (graph D) or always (graph E). I observe three thresholds.

- If $E \leq E^{\text{gate}}_{\text{never},i}$ (where $i \in \{0, 1\}$ is detector number), the detector never clicks when the gate is applied.
- If $E \geq E^{\text{gate}}_{\text{always},i}$, the detector always clicks when the gate is applied.
- If $E \leq E^{\text{no gate}}_{\text{never},i}$, the detector never clicks when the gate is suppressed.

Figure 3.6 shows these detection thresholds measured for a range of c.w. blinding powers. All the thresholds rise with the blinding power, because higher blinding power leads to a larger photocurrent and lower V_{APD}. The decreased V_{APD} leads to smaller gain and thus lower sensitivity to the trigger pulse. As can be seen, for any given blinding power, $E^{\text{no gate}}_{\text{never},i}$ is much higher than the other click thresholds. This easily allows the original detector control attack [117] to proceed undetected by the countermeasure. A more formal analysis will be stated in the next section.

3.3 Conditions of a successful attack

Experimental result of the previous section shows that the attack of Ref. 117 is possible in Clavis2. However, general conditions for a successful attack should be analyzed theoretically. In this section, we first consider *strong conditions* for a perfect attack, in which Eve induces a click in Bob with 100% probability if their bases match and the gate is applied, and 0% probability otherwise. These conditions are definitely sufficient for a successful attack [117]. However, as we remark later in this section, even if these strong conditions are not satisfied, an attack may still be possible.

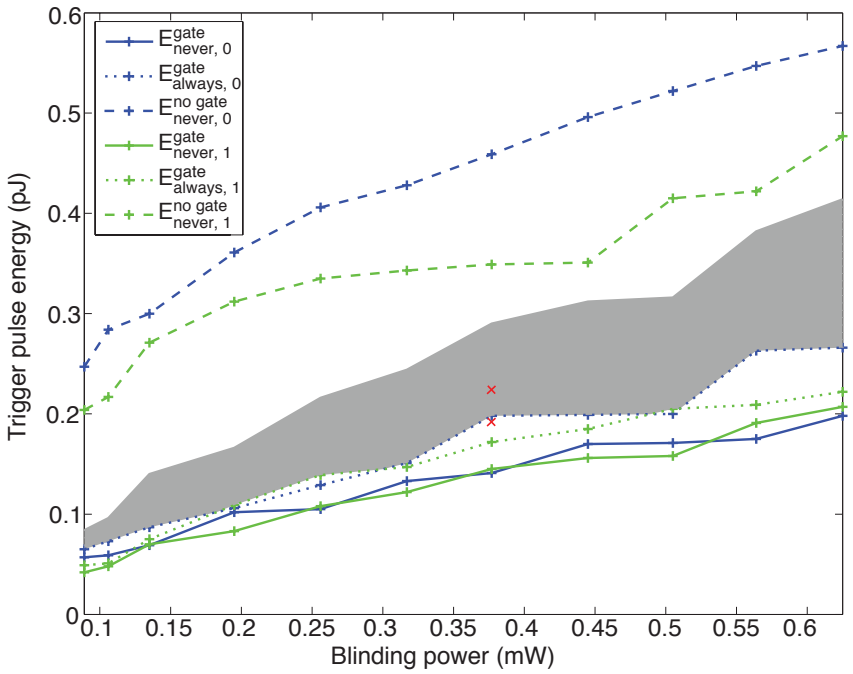

Figure 3.6: Energy thresholds of trigger pulse versus c.w. blinding power. Shaded area shows the range of trigger pulse energies of the perfect attack. The red × will be explained in Sec. 3.4.

Strong conditions. If the detection outcome varies as Fig. 3.5 with the increase of trigger pulse energy, the order of the three thresholds is:

$$E^{\text{no gate}}_{\text{never},i} > E^{\text{gate}}_{\text{always},i} > E^{\text{gate}}_{\text{never},i}. \tag{3.1}$$

If Eve and Bob select opposite bases, half of the energy of trigger pulse goes to each Bob's detector. In this case, none of the detectors should click despite the gate presence. This is achieved if [117]

$$\frac{1}{2}\max_i \left\{E^{\text{gate}}_{\text{always},i}\right\} < \left(\min_i \left\{E^{\text{gate}}_{\text{never},i}\right\}\right). \tag{3.2}$$

The random gate suppression imposes additional conditions. In case of basis mismatch, half of the trigger pulse energy is arriving at each detector. It should induce a click in neither detector when the gate signal is absent. For the target detector i, there is no click once Eq. (3.1) is satisfied. For the other detector $i\oplus 1$, no click is achieved when half of the trigger pulse energy is still lower than the detection threshold in the no-gate case. That is,

$$\frac{1}{2}E^{\text{gate}}_{\text{always},i} < E^{\text{no gate}}_{\text{never},i\oplus 1}. \tag{3.3}$$

If the bases match, we need to make sure there is no click when the gate is suppressed, but always a click in the expected detector in the gate presence. This is achieved if $E^{\text{gate}}_{\text{always},i} < E^{\text{no gate}}_{\text{never},i}$, which is already included in inequality (3.1). Although inequality (3.3) has a physical meaning, it mathematically follows from inequalities (3.1) and (3.2). Thus satisfying inequalities (3.1) and (3.2) represents the strong attack conditions and guarantees the same performance as in Ref. 117. The shaded area in Fig. 3.6 indicates a range of the trigger pulse energies Eve can apply for the perfect attack. The range is sufficiently wide to allow for a robust implementation, only requiring Eve to set correct energy with about $\pm 15\%$ precision.

Necessary condition. An attack may still be possible even if Eve's trigger pulse does not always cause a click in Bob when their bases match, and/or sometimes causes a click when their bases do not match [116]. The latter introduces some additional QBER but as long as it's below the protocol abort threshold, Alice and Bob may still produce a key. The random gate removal countermeasure imposes the condition

$$E^{\text{no gate}}_{\text{never},i} > E^{\text{gate}}_{\text{never},i}, \tag{3.4}$$

which means Eve should be able to at least sometimes cause a click in the gate while never causing a click without the gate (lest the alarm counter is increased). This is a necessary condition for an attack. As the present paper details, there are strong engineering reasons why this condition is likely to be satisfied in a detector. Additional conditions will depend on exact system characteristics [116].

3.4 Will a full implementation of the countermeasure be robust?

I have proved so far that the current countermeasure with gate suppression cannot defeat the detector blinding attack. However, the paper of Lim *et al.* [103] claims that the full version of countermeasure with two non-zero detection efficiencies is effective against a large class of detector side-channel attacks including the blinding attack [117]. Even though this full countermeasure has not been implemented by ID Quantique, I have tested some properties of the detectors in Clavis2 to show two possible methods to hack the full countermeasure, based on certain assumptions about a future implementation.

Bob could choose randomly between $P/2$ and P detection efficiency by changing either gate voltage amplitude V_{gate} or high-voltage supply V_{bias} [103]. Since in Clavis2 hardware V_{gate} is fixed, we assume an engineer will change V_{bias} to achieve different non-zero detection efficiencies. To achieve half of original detection efficiency, we lower V_{bias} manually. When $V_{\text{bias},0}$ of D0 drops from -55.26 V to -54.86 V, the detection efficiency P_0 reduces from 22.6% to 12.8%. Similarly, we decrease $V_{\text{bias},1}$ of D1 from -54.70 V to -54.40 V, leading to the detection efficiency P_1 reduction from 18.9% to 9.7%. After that, we test Eve's controllability of these two detectors.

First, I blind the detectors and then measure the relationship between the energy of trigger pulse and probability to cause a click. The position of trigger pulse is fixed in the middle of the gate signal. Figure 3.7 shows the testing result which indicates there is a transition range between 0% and 100% click probability. Solid curves show the energy of trigger pulse for original V_{bias}, while dashed curves for reduced V_{bias} lowering photon detection efficiency by about a factor of 2. The blinding power is 0.38 mW and the timing of trigger pulse is aligned in the middle of the gate by minimizing its energy required to make a click.

From the measurement result, Eve can randomly select different levels of trigger pulse energy (shown as dotted lines in Fig. 3.7) to attack the full version of countermeasure. As we know, only when Bob chooses the same measurement basis as Eve, all the energy of trigger pulse arrives at the targeted detector and achieves a click. For target D0, if trigger pulse energy E_1 is chosen, D0 always clicks, while at E_2, the detector only clicks if higher V_{bias} is applied. When E_1 and E_2 are chosen randomly with the same probability $P_0/2$, the detection probability for higher V_{bias} is P_0 and the detection probability for lower V_{bias} is only $P_0/2$. Therefore, the attack reproduces correct detection probabilities as the protocol requires. Similarly, for target D1, Eve can choose E_3 to trigger click always and choose E_4 to get a click only if higher V_{bias} is applied. This reproduces correct detection probabilities,

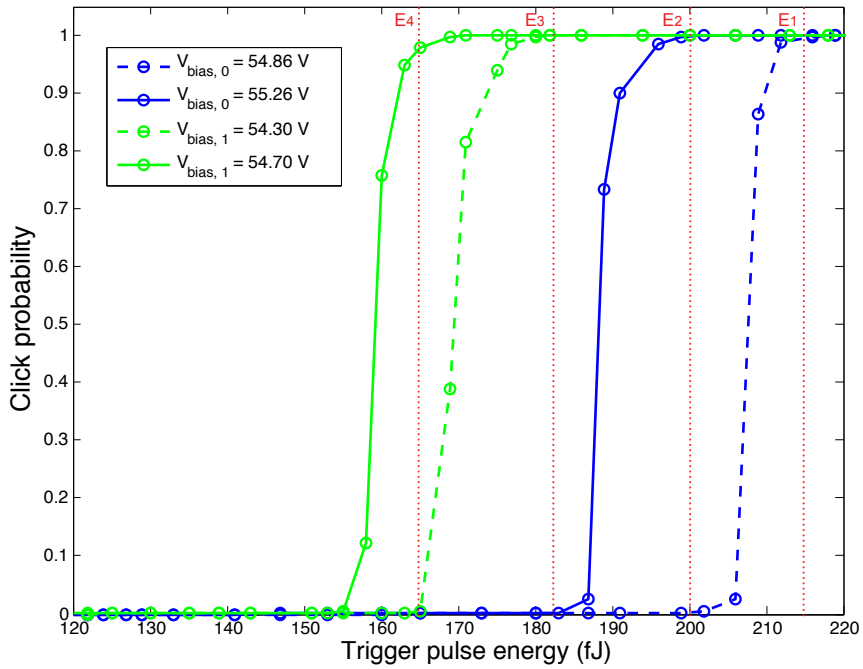

Figure 3.7: Click probabilities under blinding attack versus energy of trigger pulse.

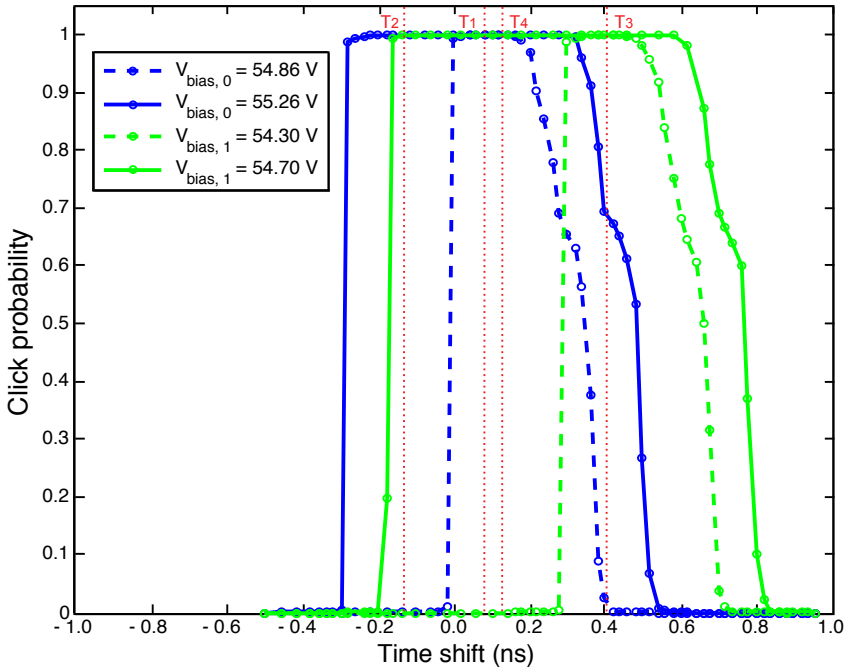

Figure 3.8: Click probabilities under blinding attack versus relative time shift of trigger pulse.

$P_1/2$ and P_1. At the same time, E_1 and E_3 remain safely below $E^{\text{no gate}}_{\text{never},0,1}$ shown in Fig. 3.6, so clicks are never produced in the absence of the gate and alarm is not triggered. This allows Eve to hack the countermeasure tracelessly.

Second, I test the correlation between time shift of trigger pulse and click probability of blinded detector. The trigger pulse energy I use in this test for D1 is slightly lower than that of D0, but both levels of energy are above $E^{\text{gate}}_{\text{always},0,1}$ in Fig. 3.6 marked as red $\times$. The measurement result is shown in Fig. 3.8. Solid curves give the detection probability at the original V_{bias}, and dashed curves give the detection probability at lower V_{bias}. Note that the latter extends over a relatively narrower time window. The blinding power is 0.38 mW. The energy of trigger pulse for D0 is 0.22 pJ and for D1 is 0.19 pJ. These energy levels are marked as red $\times$ in Fig. 3.6.

This testing result illustrates another method to attack the countermeasure: randomly adjusting the time shift of the trigger pulse. For D0, after fixing the suitable energy level of the trigger pulse, Eve can always trigger a click by choosing time shift T_1, but only trigger a click at higher V_{bias} by choosing T_2. Similarly, if target detector is D1, the detector always

clicks at T_3, but only clicks at higher V_{bias} at T_4. Then, when Eve sends the trigger pulse to control D0, she randomly selects T_1 and T_2 with equal probability $P_0/2$ to reproduce the correct detection efficiencies of D0. Eve utilizes the same strategy for D1 to achieve correct detection probabilities, $P_1/2$ and P_1. In this way, Eve also hacks Clavis2 system tracelessly.

Generally, a finite set of decoy detection efficiency levels $\eta_1 < \eta_2 < \eta_3 < ... < \eta_n$ can be hacked by properly setting probabilities of different attacking energy levels or time-shifts. We take energy levels of trigger pulse as an example. According to the result in Fig. 3.7, it is reasonable to extrapolate that we can find n distinct levels of trigger pulse energy $E_1 > E_2 > E_3 > ... > E_n$ in this situation. Then Eve can apply E_k $(k = 1, ..., n)$ with probability q_k to satisfy $\eta_k = \sum_{i=1}^{k} q_i$. This would reproduce every expected value of η_k and hack the system. We have so far assumed that applying energy level E_k causes zero click probability for decoy levels up to η_{k-1}, and 100% click probability for η_k and above. However this is not a necessary condition. More generally, under energy E_k, the click probability for efficiency level η_i is $\beta_{\eta_i}^{E_k}$. To reproduce the expected efficiencies, we need to satisfy the following set of equations:

$$\begin{aligned}
&q_1\beta_{\eta_1}^{E_1} + q_2\beta_{\eta_1}^{E_2} + ... + q_n\beta_{\eta_1}^{E_n} = \eta_1 \\
&q_1\beta_{\eta_2}^{E_1} + q_2\beta_{\eta_2}^{E_2} + ... + q_n\beta_{\eta_2}^{E_n} = \eta_2 \\
&..... \\
&q_1\beta_{\eta_n}^{E_1} + q_2\beta_{\eta_n}^{E_2} + ... + q_n\beta_{\eta_n}^{E_n} = \eta_n.
\end{aligned} \tag{3.5}$$

We might solve these equations to get values $0 \leq q_k < 1$. A worse case would be if Eve cannot find values of all q_k, which means she may only have a partial control of Bob's η_k. However, it still breaks the assumption in the security proof [103] that Eve cannot form faked states with click probability conditional on Bob's randomly chosen efficiency. For quantitative analysis, an updated security proof would be needed first.

From the above testing and analysis of the implementation that changes V_{bias}, we can guess that an alternative implementation that changes V_{gate} [103] or adds an intensity modulator in front of the detectors [128], may leave a similar loophole. If we apply the intensity modulator, the energy of the trigger pulse arriving at the detector is not constant but depends on the modulation. However, this case is similar to gate voltage modulation, as we only consider the total energy from the gate signal and trigger pulse. Therefore, we will get similar results as Figs. 3.7 and 3.8, but the amount of trigger pulse energy and time shift might be different.

The reason for this practical loophole is a wrong assumption made by Lim and his colleagues [103]. They assume Eve cannot generate faked states that trigger detections

with probabilities that are *proportional* to the original photon detection efficiency. Here I have proved this is in fact possible. Therefore, the model of a practical detector should be more precise in security analysis, if one wishes to close the detector control loophole without resorting to measurement-device-independent QKD.

3.5 Conclusion

I have tested the first implementation of the countermeasure against the blinding attack in the commercial QKD system Clavis2. Our testing result demonstrates that presently implemented countermeasure is effective against the original blinding attack but not effective against a modified blinding attack. The modified attack fully controls Bob's single-photon detectors but does not trigger the security alarm. The modified attack is similar to the original detector blinding attack [117] with the only difference that the trigger pulses are time-aligned to coincide with the detector gates, instead of following it. We argue that this attack should be implementable in practice against an installed QKD communication line where Eve does not have physical access to characterizing Alice and Bob. However such full demonstration has not yet been done, to our knowledge.

I have also tested the full proposed implementation of countermeasure with two non-zero efficiency levels, and found its security to be unreliable despite predictions of the theory proposal [103]. From the current testing results, bright-pulse triggering probabilities of the blinded detectors depend on several factors including V_{bias}, timing and energy of the trigger pulse (see Sec. 3.4). This in principle allows Eve to compromise the full countermeasure implementation.

According to the testing result, this countermeasure is not as reliable as would be expected in a high-security environment of QKD. Although an ideal industrial countermeasure has not been achieved, everybody now has a clearer concept about the detector loopholes. This procedure emphasizes the necessity of security testing every time practical QKD systems are developed or updated. We only can reach the final practical security of any QKD system after several iterations of implementation development and testing verification. Our countermeasure testing also illustrates that patching a loophole is still time-consuming and difficult. However, addressing practical vulnerabilities at the design stage of a QKD system is both cheaper and less messy than trying to retrofit patches on an existing deployed solution.

Chapter 4

Insecurity of detector-device-independent QKD (DDI QKD)

As mentioned before, the security of MDI QKD is based on post-selected entanglement, and it is able to eliminate all detector side channels from QKD implementations, which are major security loopholes [61,85,95,117,121,139,146,176,179]. However, a limitation of MDI QKD is that it requires high-visibility interference between two independent sources, which makes its implementation more demanding than that of conventional prepare-and-measure QKD. Moreover, current finite-key security proof [53] requires larger post-processing data block sizes than those of standard QKD schemes.

To overcome the limitations above, detector-device-independent QKD (DDI QKD), has been proposed recently [47, 67, 97, 102]. It avoids the challenge of interference from individual sources by applying the concept of a single-photon Bell state measurement (BSM) [90]. Consequently, its post-processing data block sizes are similar to these of prepare-and-measure QKD schemes [101]. It is claimed by the inventors of DDI QKD protocol that DDI QKD has the same security performance as that of MDI QKD. However, its security against detector side-channel attacks has not been rigorously proven yet.

In this chapter, the security of DDI QKD is investigated. It shows that, in contrast to the claimed statement [47, 67, 97, 102], the security of DDI QKD cannot rely on the same principles as MDI QKD (*i.e.,* post-selected entanglement). Importantly, I demonstrate that DDI QKD is actually vulnerable to detector side-channel attacks and to other attacks that exploit imperfections of Bob's linear optical network. These attacks are effective even

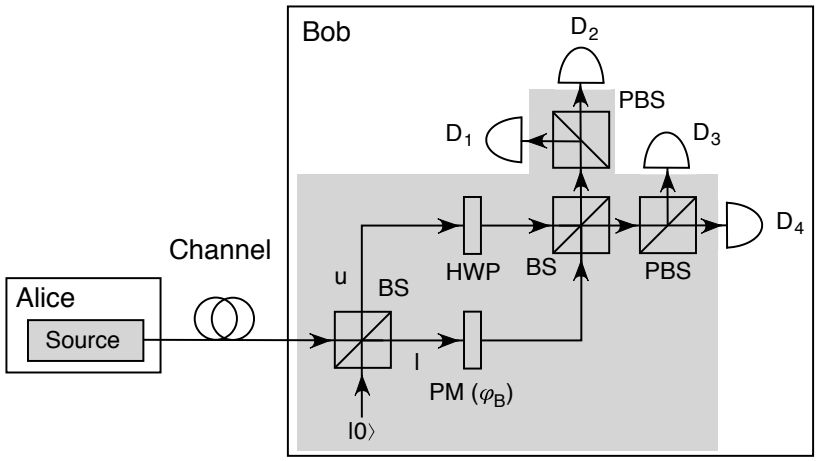

Figure 4.1: Possible implementations of detector-device-independent QKD with linear optics. HWP, half-wave plate; and PM, phase modulator. One single click in the detector D_1, D_2, D_3, or D_4 corresponds to a projection into the Bell state $|\Psi^+\rangle$, $|\Phi^+\rangle$, $|\Psi^-\rangle$, or $|\Phi^-\rangle$ respectively (see main text for further details).

when all the devices, except for detectors, are fully characterized and trusted, which is an essential assumption in DDI QKD.

4.1 Principles of DDI QKD

DDI QKD [47,67,97,102] attempts to follow the same idea of MDI QKD. The key difference is to replace the two-photon BSM with a two-qubit single-photon BSM [90]. This means that Alice and Bob use two different degrees of freedom of the same single photons to encode their secret information. In this scheme, there is no need for interfering photons from two independent sources. An example of a possible implementation is shown in Fig. 4.1 [102] (Ref. 47, 67, 97 propose similar schemes). In Fig. 4.1, Alice sends Bob one of BB84 polarization states: $(|\mathrm{H}\rangle + e^{i\theta_\mathrm{A}}|\mathrm{V}\rangle)/\sqrt{2}$, where $|\mathrm{H}\rangle$ ($|\mathrm{V}\rangle$) denotes the horizontal (vertical) polarization, and the phase $\theta_\mathrm{A} \in \{0, \pi/2, \pi, 3\pi/2\}$. Once Bob receives the photons from the quantum channel, he encodes his bit information in the spatial degree of freedom. This is completed by a $50:50$ beamsplitter (BS) to randomly choose either upper path, denoted as state $|u\rangle$, or lower path, denoted as state $|l\rangle$. A phase modulator (PM) at lower path randomly applies a phase $\varphi_\mathrm{B} \in \{0, \pi/2, \pi, 3\pi/2\}$ to each incoming signal (see Fig. 4.1). Finally, Bob performs a BSM that projects each encoded photon into a Bell state: $|\Phi^\pm\rangle =$

$(|\mathrm{H}\rangle|\mathrm{u}\rangle \pm |\mathrm{V}\rangle|\mathrm{l}\rangle)/\sqrt{2}$ and $|\Psi^{\pm}\rangle = (|\mathrm{H}\rangle|\mathrm{l}\rangle \pm |\mathrm{V}\rangle|\mathrm{u}\rangle)/\sqrt{2}$. A detection event in each detector D_i corresponds to a projection on a specific Bell state.

Similarly to MDI QKD, DDI QKD requires that Alice's and Bob's state preparation devices are characterized and trusted. This requirement is indicated by the grey areas shown in Fig. 4.1. In DDI QKD, BS, PM, and half-wave plate (HWP) inside Bob's grey area are a linear optical network. The rest of trusted elements in Bob belong to the BSM. Most importantly, the detectors D_i do not need to be characterized, but only need to be trusted that no information leaks to the outside.

4.2 The security of DDI QKD is not based on post-selected entanglement

At first glance, it seems that the security of DDI QKD follows that of MDI QKD, but only changes two-photon BSM to the two-qubit single-photon BSM. Both MDI QKD and DDI QKD assumes that Alice's and Bob's state preparation processes are trusted [47, 67, 97, 102, 110]. If it is true, that means the security of DDI QKD also relies on post-selected entanglement between Alice and Bob after BSM. Ref. 138 first indicates a confrontation of this DDI QKD idea. In that article, it was shown that DDI QKD is, in fact, insecure if Eve replaces Bob's detectors with measurement devices that leak detection results to the public channel [138]. Even though this result posed the possible insecurity of DDI QKD, it violates one of the important assumptions in DDI QKD. That is, Bob's detectors have to be built by a trusted party to avoid information leakage to the outside [67], but they are not fully characterized. This section shows that even following the security assumptions, the security of DDI QKD cannot be based on post-selected entanglement, in contrast to MDI QKD.

The security analysis is based on the DDI QKD scheme shown in Fig. 4.1. A modification is made for a clear explanation. Particularly, it is assumed that Bob's receiver contains only one detector, for example, the detector D_1, but the other three detectors are disconnected. It means that now Bob's BSM can only project the encoded photons into the Bell state $|\Psi^{+}\rangle$. Logically, if the security of DDI QKD is based on post-selected entanglement, this slight modification should not compromise its security, but reduces the secret key rate to one fourth. The projection into a single Bell state in MDI QKD is sufficient to guarantee security [110]. However, the following analysis shows that a detector blinding attack [61, 117] breaks the security of DDI QKD in this scenario.

First, it is supposed that Eve sends continuous-wave (c.w.) bright light to Bob's detector

D_1, forcing it to be in the linear mode [61, 117]. In this mode, the detector operates as a classical detector. It means that is not sensitive to single-photon pulses anymore, but is only able to detect strong pulses. It is reasonable to assume that when blinded D_1 receives a bright pulse with mean photon number μ it always gets a click, while if the mean photon number reduces to $\mu/2$, it never obtains a click. This performance has been confirmed in experiments for different detector types [61, 73, 83, 115, 117, 118, 150, 177].

Eve then performs an intercept-resend attack based on the blinded D_1. She first measures every signal sent by Alice in one of the two BB84 bases that are randomly selected. According to the measured result, she prepares a new signal and send it to Bob. Assume that the faked states that Eve sends to Bob are coherent states, $\left|\sqrt{2\mu}\right\rangle$, with creation operator $a^\dagger = (a_{\rm H}^\dagger + e^{i\phi_{\rm E}} a_{\rm V}^\dagger)/\sqrt{2}$. Here, $a_{\rm H}^\dagger$ ($a_{\rm V}^\dagger$) is the creation operator for horizontal (vertical) polarization, and the phase $\phi_{\rm E} \in \{0, \pi/2, \pi, 3\pi/2\}$ is the same as Eve's measurement result. Following this strategy, it can be proved that the state at the inputs of Bob's detectors is a coherent state

$$\begin{aligned} |\psi\rangle \;=\; & \left|\frac{\sqrt{\mu}}{2}\left(e^{i\phi_{\rm E}} + e^{i\varphi_{\rm B}}\right)\right\rangle_{D_1} \otimes \left|\frac{\sqrt{\mu}}{2}\left(1 + e^{i(\phi_{\rm E}+\varphi_{\rm B})}\right)\right\rangle_{D_2} \\ \otimes \; & \left|\frac{\sqrt{\mu}}{2}\left(e^{i\phi_{\rm E}} - e^{i\varphi_{\rm B}}\right)\right\rangle_{D_3} \otimes \left|\frac{\sqrt{\mu}}{2}\left(1 - e^{i(\phi_{\rm E}+\varphi_{\rm B})}\right)\right\rangle_{D_4}. \end{aligned} \tag{4.1}$$

The specific results under different cases are listed in Table 4.1, where it shows the mean photon number reaching each of Bob's detectors for all combinations of $\phi_{\rm E}$ and $\varphi_{\rm B}$. From this table, please note that if D_1 is the only one detector, Bob only can obtain detections when he chooses the same measurement basis as Eve's. For example, when $\varphi_{\rm B}, \phi_{\rm E} \in \{0, \pi\}$ or $\varphi_{\rm B}, \phi_{\rm E} \in \{\pi/2, 3\pi/2\}$), and $\varphi_{\rm B} = \phi_{\rm E}$, D_1 always receive μ photons, which triggers a click, but does not introduce any error. This attack indicates that the DDI QKD scheme illustrated in Fig. 4.1 with only one detector D_1 is indeed insecure against the detector blinding attack, which is just as the same as the standard QKD schemes. This shows that the DDI QKD scheme cannot defeat the detector side-channel attack. That is, the security of DDI QKD cannot be based on post-selected entanglement. The same conclusion can be applied to other DDI QKD implementations in Refs. 47, 67, 97.

Table 4.1: Mean photon number of the input light to Bob's detectors as a function of the phases ϕ_{E} and φ_{B}.

(a) $\phi_{\mathrm{E}} = 0$

φ_{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
$\frac{3\pi}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$

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

φ_{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

(c) $\phi_{\mathrm{E}} = \pi$

φ_{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
$\frac{3\pi}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$	$\frac{\mu}{2}$

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

φ_{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

4.3 Insecurity of DDI QKD against side-channel attacks

4.3.1 Side-channel attacks against Bob's detectors

Following the analysis in the previous section, let's consider a full DDI QKD scheme with four detectors as shown in Fig. 4.1. In the four-detector scheme, there is one main drawback for the detector blinding attack: it triggers double-clicks [138]. From Table 4.1, it is obvious that whenever Bob selects the same measurement basis as Eve, it always makes two detectors click. For instance, when $\varphi_{\mathrm{B}} = \phi_{\mathrm{E}} = 0$, the detectors D_1 and D_2 always click. Similar results happen for the other cases. These double clicks allow Alice and Bob to notice the presence of Eve. So, the security question becomes that whether or not four active detectors can guarantee the claimed performance of DDI QKD again. As shown below, the answer is "no". In this section, two possible hacking strategies exploiting imperfections of Bob's detectors are presented to avoid double-clicks.

The first strategy that allows Eve to avoid double-clicks is based on a time-shift attack [121, 139], which exploits the detection efficiency mismatch between Bob's detectors. In this type of attack, Eve shifts the arrival time of each signal that she sends to Bob

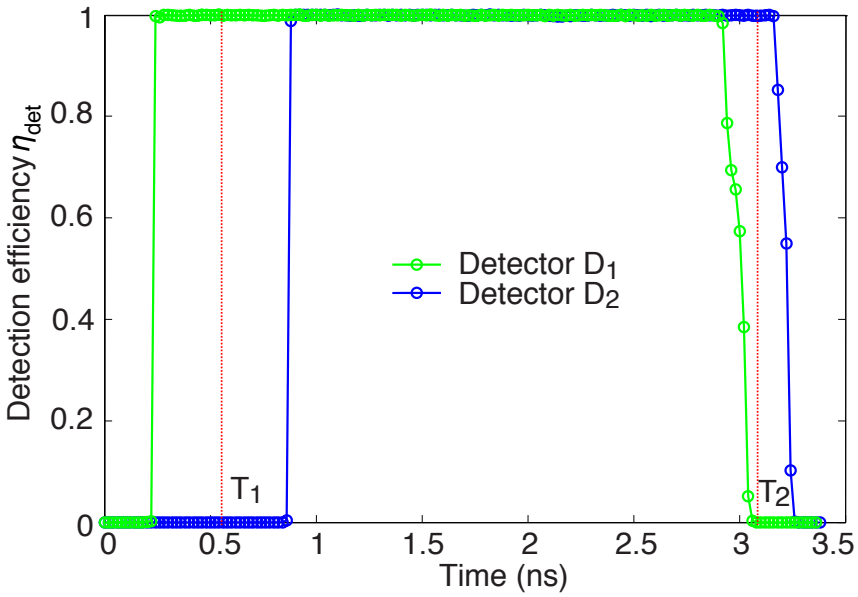

Figure 4.2: Measured detection efficiency mismatch in bright-light blinded regime in commercial QKD system Clavis2 at $P_{\mathrm{B}} = 0.32$ mW, $E_{\mathrm{T}} = 0.24$ pJ, and 0.7 ns wide trigger pulse (see main text for further details).

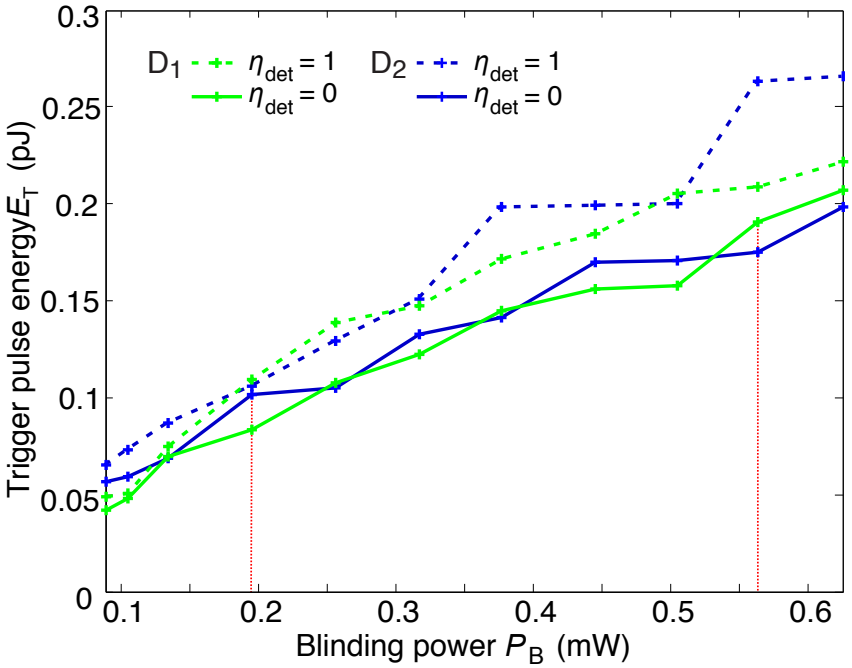

Figure 4.3: Detector click trigger thresholds versus blinding power P_{B} for two different single-photon detectors D_1 and D_2 under the blinding attack in commercial QKD system Clavis2.

such that only one detector can produce a click each given time. Here, we have confirmed experimentally that this type of attack is also possible with *blinded* detectors. For this, we blinded two single-photon detectors from the commercial QKD system Clavis2 [29] and we measured their detection efficiency mismatch. The experimental results are shown in Fig. 4.2. These results are applicable to the DDI QKD scheme in Fig. 4.1. Let us consider again the case where $\varphi_{\mathrm{B}} = \phi_{\mathrm{E}} = 0$. Suppose that Eve would like to force a click only on detector D_1, and no click on detector D_2. Then, to achieve this goal, she can simply shift the arriving time of signals to T_1. We find that only the detector D_1 can produce a click because this instance is outside of the response region of the detector D_2. Similarly, sending the signals at T_2 only triggers the detector D_2. That is, by combining the time-shift attack with the blinding attack introduced in the previous section, Eve could again break the security of DDI QKD without introducing errors nor double-clicks.

Another eavesdropping strategy is using the fact that single-photon detectors respond differently to the same blinding power P_{B}. The testing results have been presented in Chapter 3. Reprinted Fig. 4.3 from Ref. [73] shows again the response of two single-photon

detectors in a commercial QKD system Clavis2 [29] to varying blinding power P_{B}. The maximum and minimum value of the trigger pulse energy E_{T} for which the click probabilities are 0 and 1 respectively during the gate time. By choosing different blinding power and trigger pulse energy, one is able to avoid double-clicks as well. For example, as shown in Fig. 4.3, the values $P_{\mathrm{B}} \approx 0.2$ mW and $E_{\mathrm{T}} \approx 0.1$ pJ can only trigger the detector D_1. Similarly, $P_{\mathrm{B}} \approx 0.56$ mW and $E_{\mathrm{T}} \approx 0.19$ pJ can only make the detector D_2 register a click.

The attacks described above show that if Bob's detectors are uncharacterized, as assumed in DDI QKD, this type of schemes are indeed insecure against detector side-channel attacks. That is, Eve could learn the whole secret key without producing any error nor a double-click.

4.3.2 Side-channel attacks against Bob's linear optics network

One main assumption of DDI QKD is that Bob's linear optics network (*i.e.,* the grey area within Bob's receiver in Fig. 4.1) is fully characterized and trusted. Note that, however, this does not mean that its devices need to be perfect, as this would be impossible to achieve in practice. In this section, we show that Eve could also exploit various typical imperfections of Bob's linear optics to avoid double clicks when performing the blinding attack described in **??**.

For example, we consider the situation where Eve exploits the fact that Bob's BSes are not perfect to avoid double-clicks. Although a $50:50$ BS designed to operate at a certain wavelength (say, for example, at 1550 nm) can achieve nearly perfect splitting ratio at that wavelength, its splitting ratio can vary significantly at a different wavelength. For instance, a custom-made beamsplitter sample studied in Ref. [96] exhibited an extreme behaviour with splitting ratio of $98.6:1.4$ ($0.3:99.7$) at 1470 nm (1290 nm). While commercial beamsplitter models may exhibit less variation, Eve in general can to some extent control the splitting ratio by simply changing the wavelength of the signals [96], and this could be used to avoid double-clicks.

In particular, suppose that Eve's signals are in a wavelength such that the splitting ratio of Bob's first (second) BS is $t_1 : 1 - t_1$ ($t_2 : 1 - t_2$). In addition, suppose that the creation operator of Eve's coherent states $\left|\sqrt{2\mu}\right\rangle$ is now given by $a^\dagger = (\sqrt{\gamma} a_{\mathrm{H}}^\dagger + e^{i\phi_{\mathrm{E}}}\sqrt{1-\gamma} a_{\mathrm{V}}^\dagger)$, where the parameter γ is chosen by Eve. In this scenario, it can be shown that the state

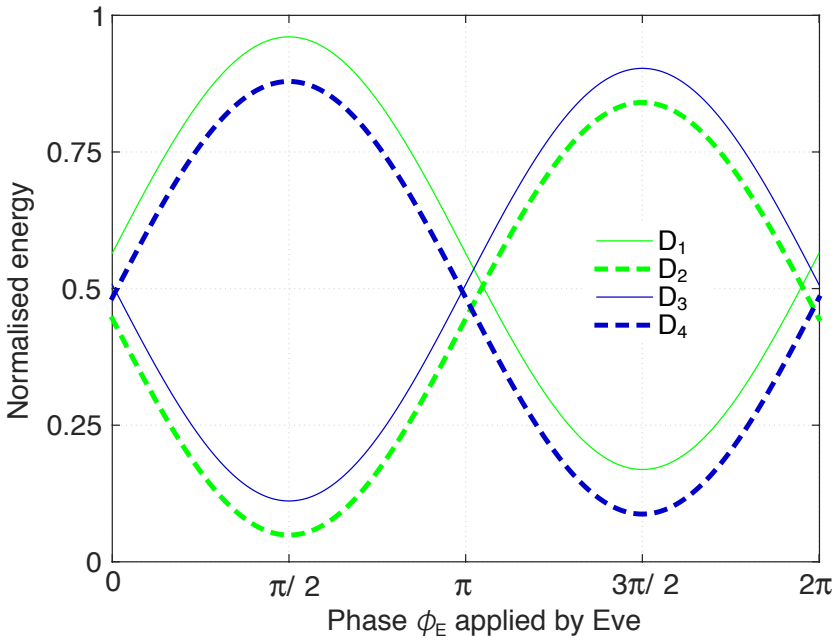

Figure 4.4: Normalised energy at the input ports of Bob's detectors D_i as a function of ϕ_E, when $\varphi_B = \pi/2$.

at the input ports of Bob's detectors D_i is a coherent state of the form

$$\begin{aligned} |\psi\rangle \;=\; & \left|\alpha\big(\sqrt{\hat{t}_1\hat{t}_2\hat{\gamma}}e^{i\phi_E} + \sqrt{t_1t_2\gamma}e^{i\varphi_B}\big)\right\rangle_{D_1} \\ \otimes\; & \left|\alpha\big(\sqrt{\hat{t}_1\hat{t}_2\gamma} + \sqrt{t_1t_2\hat{\gamma}}e^{i(\phi_E+\varphi_B)}\big)\right\rangle_{D_2} \\ \otimes\; & \left|\alpha\big(\sqrt{\hat{t}_1t_2\hat{\gamma}}e^{i\phi_E} - \sqrt{t_1\hat{t}_2\gamma}e^{i\varphi_B}\big)\right\rangle_{D_3} \\ \otimes\; & \left|\alpha\big(\sqrt{\hat{t}_1t_2\gamma} - \sqrt{t_1\hat{t}_2\hat{\gamma}}e^{i(\phi_E+\varphi_B)}\big)\right\rangle_{D_4}, \end{aligned} \tag{4.2}$$

where $\hat{x} = 1 - x$, and $\alpha = \sqrt{2\mu}$. Note that when $t_1 = t_2 = \gamma = 1/2$ we obtain Eq. 4.1.

This means that, in principle, Eve might select the parameter γ and the wavelength of her signals such that the resulting splitting ratios t_1 and t_2 make the input energies at Bob's detectors asymmetric. In so doing, and following a similar argumentation to the one introduced in the previous eavesdropping strategy, Eve can guarantee that when she and

Bob choose the same basis, only one detector clicks. This situation is illustrated in Fig. 4.4 for a particular example where $\varphi_{\mathrm{B}} = \pi/2$, $t_1 = 0.44$, $t_2 = 0.46$, and $\gamma = 0.2$. In this scenario, we assume that the splitting ratio of Bob's first (second) BS is $44:56$ $(46:54)$, and Eve's state parameter $\gamma = 0.2$. We find that the maximum normalized energy at the input ports of Bob's detectors D_1 and D_4 when Eve selects $\phi_{\mathrm{E}} = \pi/2$ is, respectively, 0.96 and 0.87. Similarly, when she chooses $\phi_{\mathrm{E}} = 3\pi/2$ the maximum normalized energy at the detectors D_3 and D_2 is, respectively, 0.9 and 0.84. Therefore, Eve can choose the energy of her signals such that only the detector D_1 (D_3) clicks when $\phi_{\mathrm{E}} = \pi/2$ $(\phi_{\mathrm{E}} = 3\pi/2)$. That is, by changing the values of the parameters t_1, t_2, and γ, Eve can guarantee that only one detector clicks each given time.

4.4 Conclusion

In this chapter, the security of the DDI QKD has been analyzed. In the beginning, DDI QKD promised to be secure against detector side-channel attacks. However, it has been shown that its security cannot be based on post-selected entanglement, which was claimed. Most importantly, we have presented several types of attacks to show that DDI QKD is actually vulnerable to detector side-channel attacks and other side-channel attacks exploiting imperfections of Bob's linear optical network. These attacks are effective even under the assumptions of the DDI QKD. That is, Alice's and Bob's state preparation devices are fully characterized and trusted, and Bob's detectors are only trusted but not characterized. From this study, it is clear that the security of DDI QKD scheme is actually the same as that of a standard prepare-and-measure scheme. It means that DDI QKD scheme does not provide any security advantage. Most importantly, the DDI QKD cannot be treated as a simple version of the MDI QKD. The MDI QKD still seems to be the only practical solution to defeat all the detector side channels.

Chapter 5

Decoy state QKD with imperfect source

The security of the measurement station is not required for the MDI QKD. However, even in MDI QKD, the source still needs to be protected. Unfortunately, in practice, the requirements for the source may not be satisfied. It is hence important to investigate the practical imperfections in the source station. This chapter is based on a reprint (arXiv:1711.00597).

5.1 Motivation

In practical quantum cryptographic systems, a weak coherent source (WCS) is widely used to replace the single photon source. One inherent imperfection in WCS is the emission of multi-photon pulses, which gives Eve more than one copy of Alice's quantum states. Eve could then perform the photon-number-splitting (PNS) attack [43, 75], in which she blocks all single-photon pulses, and keeps one photon from the multi-photon state. She could then get all the final key after Alice and Bob announce their basis choices. Note that a modified PNS attack based on a beam splitter has been demonstrated [105]. Thus, the danger of PNS attack is not only theoretical but also practical. Fortunately, decoy state protocols [76, 112, 175] were proposed to beat such attack, which has been implemented in many QKD systems [106, 144, 152, 166, 182]. It has also been employed in other quantum cryptographic systems [42, 59, 183, 185] to guarantee their security.

Generally speaking, in the decoy state protocol, signal and decoy states only have different mean photon numbers. Decoy states are used to estimate the detection gain and

error rate of single-photon pulses in signal states. If Eve could not distinguish the signal and decoy states, then she would change the photon number in both signal and decoy pulses during the PNS attack [43, 75]. Thus, she would disturb the yield and error rate of decoy states, which affects the estimation of single-photon detection gain and error rate in signal states. It results in the decrease of the secure key rate [112].

However, the essential assumption - the indistinguishability of the signal and decoy states - may not be guaranteed in practice. In fact, Eve might exploit practical imperfections to find a side channel which allows her to distinguish the signal and decoy states. She could then perform different hacking strategies to keep the normal statistic distributions, while spying on some secret information silently without being discovered. Several types of source imperfections and corresponding attacks have been shown in different QKD systems [82, 132, 165, 168]. Importantly, the first quantum satellite also employs one of such imperfect sources [99]. The security of the decoy state QKD with a leaky source has been considered by Tamaki and his co-workers [165], in which the imperfection of signal and decoy states is taken into account. However, a linear program problem should be solved to estimate the contribution of the single photon pulse in their security proof, which is complex.

In this chapter, I study several types of implementations for the decoy state protocol to show some side channels, the corresponding attack, and solutions.

5.2 Decoy state protocol

As a fundamental theory of our research, we recap the decoy state protocol first in this section. Here we take the weak + vacuum decoy state protocol [120] as an example to explain the basic idea of the decoy state protocol. This simple one weak + vacuum decoy state protocol is commonly used in Bennett-Brassard 1984 (BB84) QKD system [35], as it provides the optimal key rate in the case of only two decoy states [120]. The security analysis in Secs. 5.4 to 5.6 also follows this decoy state model.

According to the analysis of Gottesman-Lo-Lütkenhaus-Preskill (GLLP) [70], the key rate of QKD with the WCS can be written as

$$R \geq q\{-Q_\mu H_2(E_\mu) f(E_\mu) + P_1^\mu Y_1^\mu [1 - H_2(e_1^\mu)]\}. \tag{5.1}$$

Here $q = 1/2$ for BB84 protocol (if one uses the efficient BB84 protocol [109], $q \approx 1$), the subscript μ means the intensity of a signal state, Q_μ (E_μ) is the total gain (the error rate) of the signal state, Y_1^μ and e_1^μ are the yield and the error rate of single-photon pulses, P_1^μ is

the probability of single-photon pulses, $f(x)$ is the bidirectional error correction efficiency, normally $f(x) \geq 1$ with Shannon limit $f(x) = 1$, and $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary Shannon information entropy.

In Eq. (5.1), Q_μ and E_μ are directly obtained in an experiment, and P_1^μ is known for a given source. Thus, the major task of the decoy state is to tightly estimate the lower bound of Y_1^μ and upper bound of e_1^μ. Note the fact that, if the phase of the WCS is randomized from 0 to 2π (the phase randomization assumption), the density matrix of the WCS can be written as

$$\rho_\omega = \sum_{n=0}^{\infty} P_n^\omega |n\rangle\langle n|, \tag{5.2}$$

where $\omega = \{\mu, \nu, 0\}$ represents the average intensity of pulse signal state μ, decoy state ν, and vacuum state that is always 0. P_n^ω is the probability distribution of n-photon number from the source with the intensity ω. For the WCS, $P_n^\omega = e^{-\omega}\omega^n/n!$. Without loss of generality, we assume $\mu > \nu$. Thus, the total gain and the error rate can be written as

$$\begin{aligned} Q_\omega &= \sum_{n=0}^{\infty} P_n^\omega Y_n^\omega, \\ Q_\omega E_\omega &= \sum_{n=0}^{\infty} P_n^\omega Y_n^\omega e_n^\omega. \end{aligned} \tag{5.3}$$

Here Y_n^ω (or e_n^ω) is the yield (or the error rate) given that Alice sends a n-photon pulse from the source with intensity ω. Obviously, if Eve does not have any prior information about the intensity of Alice's pulse, we can assume that

$$\begin{aligned} Y_n^\mu &= Y_n^\nu = Y_n, \\ e_n^\mu &= e_n^\nu = e_n. \end{aligned} \tag{5.4}$$

The lower bound of Y_1 and upper bound of e_1 can be then estimated by solving the linear Eqs. (5.3) with weak + vacuum decoy states [120].

5.3 Intensity modulation test

To evaluate the realization of the weak + vacuum decoy state protocol, we test two intensity modulation methods. The first method under testing is the pump-current modulation, similar to Refs. 100, 184. For the signal and weak decoy states, different intensities are

produced by applying different pulses of pump current to a laser diode. Thus, the laser diode directly emits optical pulses with different intensities. The vacuum state is generated by turning off the pump current. An optical attenuator then applies a fixed attenuation to all the optical pulses, to reach single-photon level. The second method under testing is an external intensity modulator, similar to Refs. 57, 144, 188. Optical pulses could be produced with a constant intensity from a laser diode first, and then the different intensities of signal and decoy states are modulated by an intensity modulator (IM). Similar to the former method, a fixed attenuator provides attenuation to the single-photon level.

Our intensity measurement of the optical pulses is taken before the fixed attenuation is applied. The optical pulses are measured by a photodetector (40 GHz bandwidth) and an oscilloscope (33 GHz bandwidth), averaging $\gtrsim 5000$ pulses. We obtain the normalized probability distribution of emitting photons over time which is shown in Figs. 5.1(a) and 5.2. Although we measure the intensity of classical optical pulses, the probability of emitting single photon should follow the same distribution, because constant attenuation is applied.

For the case of pump-current modulation, Fig. 5.1(a) clearly shows that the probability distributions of emitting the signal state and the decoy state do not totally overlap. The main peaks of these two distributions are mismatched. The signal state emits earlier than the decoy state with high probability and has a secondary peak from 662 to 937 ps. Over the same time interval, the probability distribution of the decoy state drops to low values. The timing mismatch of the signal state and the decoy state clearly violates the basic assumption of indistinguishability in the decoy state protocol. As we show numerically in Sec. 5.4, this can be exploited by Eve to bypass the protection of the decoy state protocol. However, the measured result of external intensity modulation in Fig. 5.2 does not show a measurable timing mismatch between signal and decoy states. This is expected because the pulse generation and intensity modulation in this type of source are physically decoupled and performed by separate devices. As long as there is no electrical crosstalk between the laser diode driver and intensity modulator driver, no correlation is expected. This is the case, as Fig. 5.2 shows.

To investigate the reason for timing mismatch in the case of pump-current modulation, we measure the current flowing through the laser diode (Agilecom WSLS-940010C4123). A differential probe with 30 GHz bandwidth (Agilent N5445A) is used to measure the differential voltage V across the laser diode and its built-in serial resistor $R_s = 20\ \Omega$. Since the laser diode forward voltage $V_d = 1.23$ V is known from its test sheet, we can calculate the pump current $I = (V - V_d)/R_s$. This calculated current is shown in Fig. 5.1(b). The lasing threshold, 14 mA, is shown as a line.

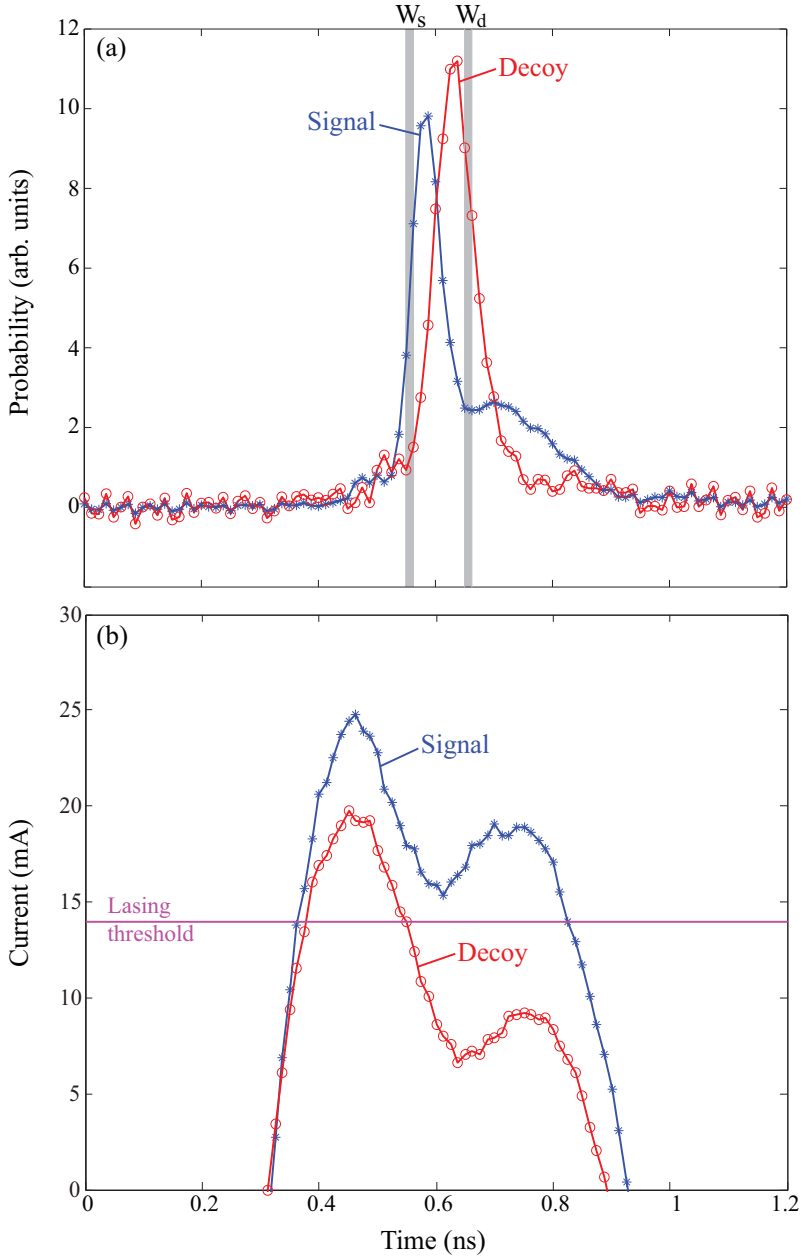

Figure 5.1: Pump-current modulation. (a) Normalized intensity distribution of the signal state and the decoy state measured in the time domain. For ease of comparison, the pulses are normalized to have the same area. (b) Laser-diode's pump current. The relative time alignment between (a) and (b) is a guess.

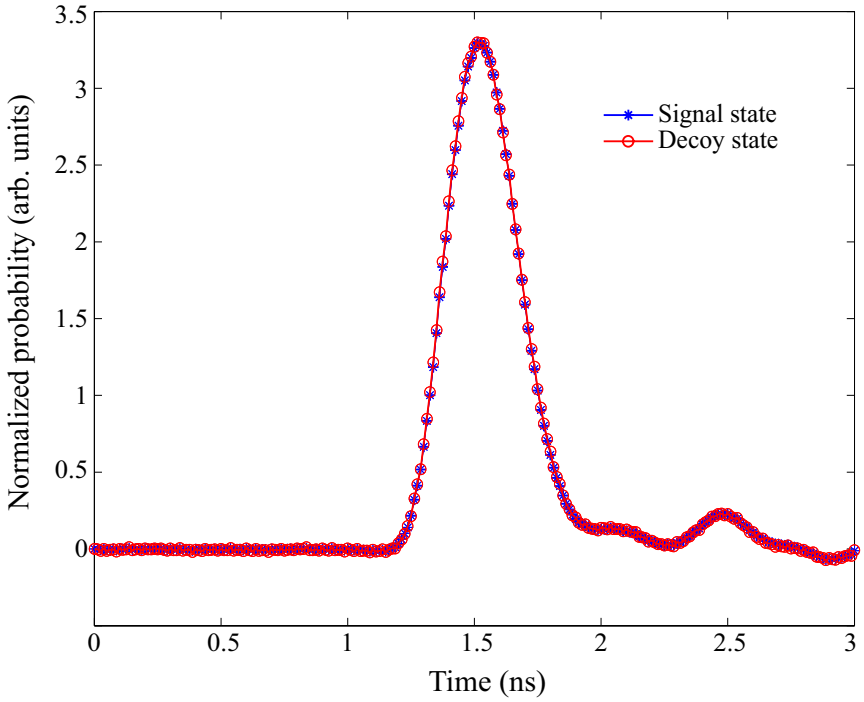

Figure 5.2: External intensity modulation. Normalized intensity distribution of the signal state and the decoy state measured in the time domain.

If the laser diode was pumped by a constant current, any current above the lasing threshold $I_{\text{th}} = 14$ mA (shown in the figure) would result in continuous-wave (c.w.) laser emission. However, when the current is initially zero then rapidly increased above I_{th}, the diode does not begin to lase immediately [33]. First, a certain number of carriers should be injected into the p-n junction before the diode reaches population inversion, and that takes time (the higher the current, the less time). Once the population inversion is reached and the diode attains light amplification condition, the few spontaneously emitted photons present in the optical cavity need time to amplify into the strong coherent light. This results in a fraction-of-nanosecond delay between the application of current and the start of strong light emission. In this process, the population inversion and emitted light power briefly overshoot the steady-state. They then undergo a few oscillations with ~ 100 ps period and eventually settle at the steady-state c.w. level if the pump current continues [33]. However, if the pump current is interrupted, as is the case with our device under test, the lasing stops. As can be seen in Fig. 5.1, the signal state is produced by a higher peak current pulse, the laser begins emitting light earlier and has time to emit two light pulses (i.e., light power oscillations) before the current stops. When the decoy state is produced by a lower peak current pulse, light emission begins later and the laser only has time to emit one light pulse. Although this physics of laser diode operation is well-known to the manufacturers of pulsed laser diodes (e.g., PicoQuant), it is a somewhat obscure topic for many electronics engineers. The engineers who have selected this modulation method for the QKD system under test may not have been aware of its implications on the timing and shape of emitted pulses.

5.4 PNS attack

In the case of pump-current modulation, being that the signal state and the decoy state are partially distinguishable in the time domain, the PNS attack becomes possible again. Here we consider a special PNS attack summarized in Table 5.1. Eve selects time windows W_s and W_d to observe states sent by Alice. By properly setting the intervals of W_s and W_d, Eve treats all the states observed in W_s (W_d) as the signal state (the decoy state). She then performs the PNS attack. For single-photon states, Eve blocks or forwards those that are in the observation windows, while she blocks all of those that are out of the observation windows. Once the states contain two or more photons, Eve keeps one photon and either blocks or forwards the rest of the photons to Bob in the observation windows but forwards all the photons to Bob when the states are out of the windows. If Eve obtains photons in both W_s and W_d, she randomly keeps photons in only one window and forwards the rest

Table 5.1: Hacking strategy and corresponding yields.

PNS attack	In the time windows	Outside the time windows
Single-photon states	Forward or block	Block
Multiphoton states	Keep one photon and forward or block others	Forward
Yield $Y_n^{\omega_{\text{Eve}}}$	Z_n^{ω}	0 $(n=1)$ Y_n^{ω} $(n \geq 2)$

of photons to Bob.

By following the criteria of a successful attack proposed in Ref. 168, the success of the above attack could be analyzed. A successful attack lets Eve know partial information about the final secret key. In other words, Alice and Bob's key remains partially insecure after post-processing. To show this, a lower bound of the key rate under Alice and Bob's estimation, R^l, and an upper bound of the key rate under Eve's attack, R^u, are compared. If

$$R^l > R^u, \tag{5.5}$$

the shared final key must be partially insecure, and then Eve knows some amount of information. This is the result that Eve's attack would like to achieve.

The lower bound of the key rate is the one used in the decoy state protocol [120]:

$$R^l = -Q_\mu H_2(E_\mu) f(E_\mu) + Y_1^\mu \mu e^{-\mu}[1 - H_2(e_1^\mu)], \tag{5.6}$$

which is consistent with Eq. (5.1) when we consider the efficient BB84 protocol [109], $q = 1$. Here Y_1^μ and e_1^μ are the single-photon yield and the error rate in the normal decoy state protocol. It is the secure key rate from Alice and Bob's point of view under the attack. Since Alice and Bob do not know about Eve's attack, the estimation of the lower bound of Y_1^μ and the upper bound of e_1^μ still follows the weak + vacuum decoy protocol [120] with the assumption of indistinguishability. The actual upper bound of the key rate under the PNS attack [168] is

$$R^u = Y_1^{\mu_{\text{Eve}}} \mu e^{-\mu}, \tag{5.7}$$

where $Y_1^{\mu_{\text{Eve}}}$ is the real overall yield of single-photon states under Eve's attack. Apparently, the goal of our attack is to minimize the upper bound in Eq. (6.2) to satisfy inequality (5.5), while matching the value of Q_ω and even reaching lower QBER than $Q_\omega E_\omega$. The attack will then remain unnoticed.

Based on the measurement result in Fig. 5.1(a), Eve can only partially distinguish the signal state and the decoy state. Hence, within a certain observation window, we define the following guessing probability. The conditional probability $P(i|j)$ is defined as Eve guesses the state is i given Alice actually sending the j state. Here $i, j \in [s, d]$, which means i or j is either the signal state, s, or the decoy state, d. Thus, $P(s|s)$ and $P(d|d)$ are the probabilities of correct guess in W_s and W_d respectively, while $P(s|d)$ and $P(d|s)$ are the probabilities of wrong guess in the same windows.

As mentioned in the hacking strategy, once Eve observes multiphoton states in W_s or W_d, she keeps a single photon and might forward or block the remaining photons to Bob. To maintain the statistics of Q_ω and $Q_\omega E_\omega$, Eve has to manipulate detection yield in the observation windows from Y_n^ω to Z_n^ω as shown in Table 5.1. In the time window W_s (W_d), the yield is denoted as Z_n^μ (Z_n^ν). Please note that Eve allows to use a lower-loss, or even lossless, channel, which means Z_n^ω could be greater than Y_n^ω. At the phase of decoy announcement in QKD protocol, Bob classifies detection slots according to Alice's signal and decoy information. Thus, under Eve's attack, the yields $Y_n^{\omega_{\text{Eve}}}$ actually should be recalculated as follows. For the single-photon states, Eve fully controls the yields, since the single-photon states out of time windows are blocked. Thus, $Y_1^{\omega_{\text{Eve}}}$ are given by

$$\begin{aligned} Y_1^{\mu_{\text{Eve}}} &= P(s|s)Z_1^\mu + P(d|s)Z_1^\nu, \\ Y_1^{\nu_{\text{Eve}}} &= P(s|d)Z_1^\mu + P(d|d)Z_1^\nu. \end{aligned} \tag{5.8}$$

For multiphoton states ($n \geq 2$), Eve forwards the states to Bob when these states are out of observation windows, so $Y_n^{\omega_{\text{Eve}}}$ are given by

$$\begin{aligned} & Y_n^{\mu_{\text{Eve}}} \\ =& P(s|s)Z_n^\mu + P(d|s)Z_n^\nu + [1 - P(s|s) - P(d|s)]Y_n^\mu, \\ & Y_n^{\nu_{\text{Eve}}} \\ =& P(s|d)Z_n^\mu + P(d|d)Z_n^\nu + [1 - P(s|d) - P(d|d)]Y_n^\nu. \end{aligned} \tag{5.9}$$

Correspondingly, the overall gains of the signal state and the decoy state are

$$\begin{aligned} Q_{\mu_{\text{Eve}}} &= Y_0^{\mu_{\text{Eve}}} e^{-\mu} + \sum_{n=1}^{\infty} Y_n^{\mu_{\text{Eve}}} e^{-\mu} \frac{\mu^n}{n!}, \\ Q_{\nu_{\text{Eve}}} &= Y_0^{\nu_{\text{Eve}}} e^{-\nu} + \sum_{n=1}^{\infty} Y_n^{\nu_{\text{Eve}}} e^{-\nu} \frac{\nu^n}{n!}, \end{aligned} \tag{5.10}$$

where $Y_0^{\omega_{\text{Eve}}}$ are dark count rates under the attack. The overall QBER are given by

$$\begin{aligned} E_{\mu_{\text{Eve}}} Q_{\mu_{\text{Eve}}} &= \frac{1}{2} Y_0^{\mu_{\text{Eve}}} e^{-\mu} + \sum_{n=1}^{\infty} \frac{1}{2} P(d|s) Z_n^{\nu} e^{-\mu} \frac{\mu^n}{n!}, \\ E_{\nu_{\text{Eve}}} Q_{\nu_{\text{Eve}}} &= \frac{1}{2} Y_0^{\nu_{\text{Eve}}} e^{-\nu} + \sum_{n=1}^{\infty} \frac{1}{2} P(s|d) Z_n^{\mu} e^{-\nu} \frac{\nu^n}{n!}. \end{aligned} \tag{5.11}$$

Here we consider an extreme case. A dark count introduces error half the time. There is no error if signal and decoy states are correctly distinguished by Eve or states are out of the windows W_s and W_d. However, a wrong guess in W_s and W_d results in random clicks, which introduces error half the time.

According to the standard decoy state protocol [120], the normal overall gains should be

$$Q_{\omega} = Y_0 + 1 - e^{-\eta\omega}, \tag{5.12}$$

where Y_0 is the dark count rate, and η is the total transmittance of the QKD system. η is given by

$$\eta = \eta_{\text{Bob}} 10^{-\alpha L/10}, \tag{5.13}$$

where η_{Bob} is the transmittance of Bob's optical device, including detector efficiency, and α is the transmittance of channel between Alice and Bob. Typically, $\alpha = 0.21$ dB/km for the commercial fibre at 1550 nm. L is the length of channel. The normal overall QBERs should be

$$E_{\omega} Q_{\omega} = \frac{1}{2} Y_0 + e_{\text{detector}} (1 - e^{-\eta\omega}), \tag{5.14}$$

where e_{detector} is the probability that a photon goes to erroneous detector, characterizing the alignment and stability of a QKD system.

To achieve a successful attack, the upper bound of the key rate R^u should be minimized, which is equivalent to minimizing $Y_1^{\mu_{\text{Eve}}}$ in Eq. (5.8). Meanwhile, to achieve a traceless attack, the attack has to follow the same detection statistics by optimizing Z_n^{μ}, Z_n^{ν}, $P(s|s)$, $P(d|s)$, $P(s|d)$ and $P(d|d)$ for every distance value. Therefore, it becomes an optimization problem under certain constraints:

$$\min_{Z_n^{\mu}, Z_n^{\nu}; P(s|s), P(d|s)} Y_1^{\mu_{\text{Eve}}} \tag{5.15}$$

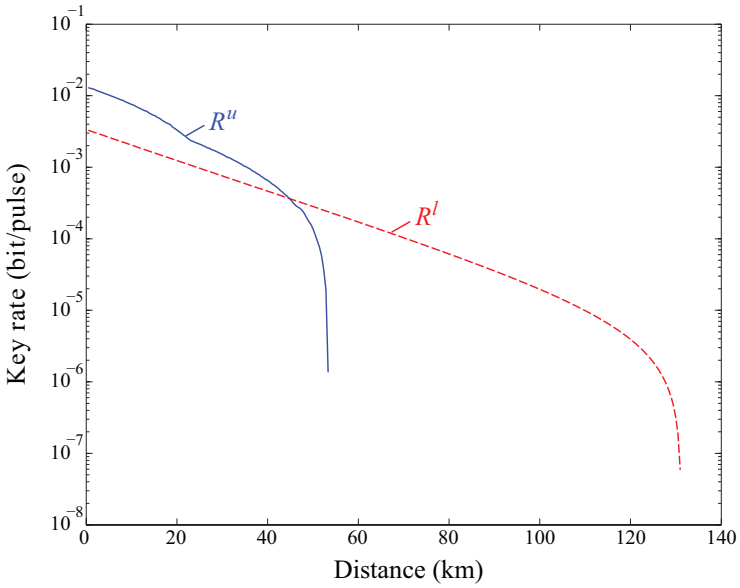

Figure 5.3: The lower bound R^l and optimized upper bound R^u of the key rate under our simulated attack.

subject to

$$\begin{aligned} Q_\mu &= Q_{\mu_{\text{Eve}}}, \\ Q_\nu &= Q_{\nu_{\text{Eve}}}, \\ E_\mu Q_\mu &\geqslant E_{\mu_{\text{Eve}}} Q_{\mu_{\text{Eve}}}, \\ E_\nu Q_\nu &\geqslant E_{\nu_{\text{Eve}}} Q_{\nu_{\text{Eve}}}, \\ Z_n^\mu, Z_n^\nu &\in [0,1], \\ P(s|s), P(d|s), P(s|d), P(d|d) &\in [0,1]. \end{aligned} \tag{5.16}$$

Ideally, the detection efficiency could be 100%, so the yield Z_n^ω could reach 1. We also remark that the probabilities $P(i\big|j)$ are taken from the measured probability distribution of the states sent by Alice in Fig. 5.1(a). $P(s\big|s)$ and $P(s\big|d)$ should be taken from the time window W_s; $P(d\big|s)$ and $P(d\big|d)$ should be taken from the time window W_d. Importantly, since every time window could contain several timing intervals, any observation probabilities mentioned above should be the summary of all the possibilities in the time windows.

The simulation result is shown in Fig. 5.3. To follow the initial analysis of weak + vacuum decoy state protocol in Ref. 120, we also use the detection parameters from Gobby-

Yuan-Shields (GYS) experiment [66] in our attack simulation. The dark count rate $Y_0 = 1.7 \times 10^{-6}$, the transmission in Bob's apparatus $\eta_{\text{Bob}} = 4.5\%$, the misalignment error rate $e_{\text{detector}} = 3.3\%$ and the error correction efficiency $f(E_\mu) = 1.22$. However, we assume the source has characteristics as in Fig. 5.1(a); this source actually comes from a different QKD system with the mean photon number $\mu = 0.6$ for the signal state and $\nu = 0.2$ for the weak decoy state. According to inequality (5.5), once the optimized upper bound starts becoming smaller than the lower bound, Eve can successfully execute the PNS attack. Figure 5.3 shows that Eve is able to hack it successfully and eavesdrop some of the secret keys when the distance between Alice and Bob is longer than 47 km. The attack windows W_s and W_d are optimized for every distance point to get the lowest R^u at this distance. For example, when the distance between Alice and Bob is 49 km, the optimized W_s and W_d are shown as grey zones in Fig. 5.1(a).

5.5 Tightened the secure key rate with an imperfect source

The previous section shows the effect of partial distinguishability between signal and decoy states in the time domain. However, the side channel that partially distinguishes signal and decoy states could be more general. For example, generating signal and decoy states by individual laser diodes is widely employed in QKD systems [106, 135, 187], even in the first quantum satellite [99]. Unfortunately, this type of state preparation might leak the modulation information in the time and frequency (spectral) domains, which was shown in the previous research [132]. For another preparation method of one laser diode with an IM in a plug-and-play system, Eve shifts the arriving time of pulses to the rising edge of intensity modulation, obtaining a side channel in the frequency domain in the plug-and-play system [82]. Moreover, the modulation information of IM might be read out by an active Trojan-horse attack [165, 172]. Even if the intensity modulation is perfect, the laser pulses with non-random phases give Eve a chance to distinguish signal and decoy states [168]. Therefore, it is important to build a general security model that tolerates such side channels. In this section, we modify the model of the decoy state protocol to consider such imperfect sources and derive two tight analytic formulas that estimate the contribution of single-photon pulses.

5.5.1 Model

The following analysis is based on the weak + vacuum decoy state protocol with intensities $\omega = \{\mu, \nu, \nu_1\}$. Without loss of generality, it is assumed $\mu > \nu > \nu_1$ and $\mu > \nu + \nu_1$. When the imperfection of source is considered in the security model, the density matrix of Alice's states [Eq. (5.2)] becomes

$$\rho'_\omega = \rho_\omega \otimes \rho_\omega(\lambda) = \sum_{n=0}^{\infty} \sum_{\lambda} P_n^\omega f_\omega(\lambda) |n, \lambda\rangle\langle n, \lambda|. \tag{5.17}$$

Here $\rho_\omega(\lambda)$ is Eve's quantum state applied to tell the signal state and the decoy state for each pulse. We remark that $\rho_\omega(\lambda)$ can be an extra quantum state, or any additional dimension of Alice's pulses. λ denotes the dimension, like the time, frequency etc., used by Eve, which is measured to distinguish between the signal state and the decoy state. $f_\omega(\lambda)$ is the normalized probability distribution of λ ($\sum_\lambda f_\omega(\lambda) = 1$), which is also depends on the intensities of Alice's pulse ω. Apparently, when $\rho_\omega(\lambda)$ is not correlated to the intensities of Alice's pulse, which means $\rho_\mu(\lambda) = \rho_\nu(\lambda) = \rho_{\nu_1}(\lambda) \equiv \rho(\lambda)$. Thus, the general decoy state method in Eq. (5.2) is able to estimate the bound of the yield and the error rate for the single photon pulses [120].

By integrating Eq. (5.17) into Eqs. (5.3), the total gain and the error rate of Alice's states can be rewritten as

$$\begin{aligned} Q_\omega &= \sum_{n=0}^{\infty} P_n^\omega Y_n^\omega = \sum_{n=0}^{\infty} P_n^\omega \sum_{\lambda} f_\omega(\lambda) Y_n(\lambda), \\ Q_\omega E_\omega &= \sum_{n=0}^{\infty} P_n^\omega Y_n^\omega e_n^\omega = \sum_{n=0}^{\infty} P_n^\omega \sum_{\lambda} f_\omega(\lambda) Y_n(\lambda) e_n(\lambda), \end{aligned} \tag{5.18}$$

where $Y_n(\lambda)$ and $e_n(\lambda)$ are the yield and the error rate when Alice sends a n-photon pulse and meanwhile Eve obtains λ from her measurement. Hence, $Y_n(\lambda)$ and $e_n(\lambda)$ rely on the parameter λ, but are not dependent on ω.

We can characterize the imperfection of source by the distance between $\rho_\omega(\lambda)$ and $\rho_{\omega'}(\lambda)$ as follows.

$$D_{\omega\omega'} = \frac{1}{2} tr|\rho_\omega - \rho_{\omega'}| = \frac{1}{2} \sum_{\lambda} |f_\omega(\lambda) - f_{\omega'}(\lambda)|, \tag{5.19}$$

where $tr|x|$ is defined as the trace distance of quantum state.

Here, we choose the case of slight timing difference between the signal state (μ) and the decoy state (ν) to show the imperfection of source $D_{\omega\omega'}$. Without loss of generality, it is

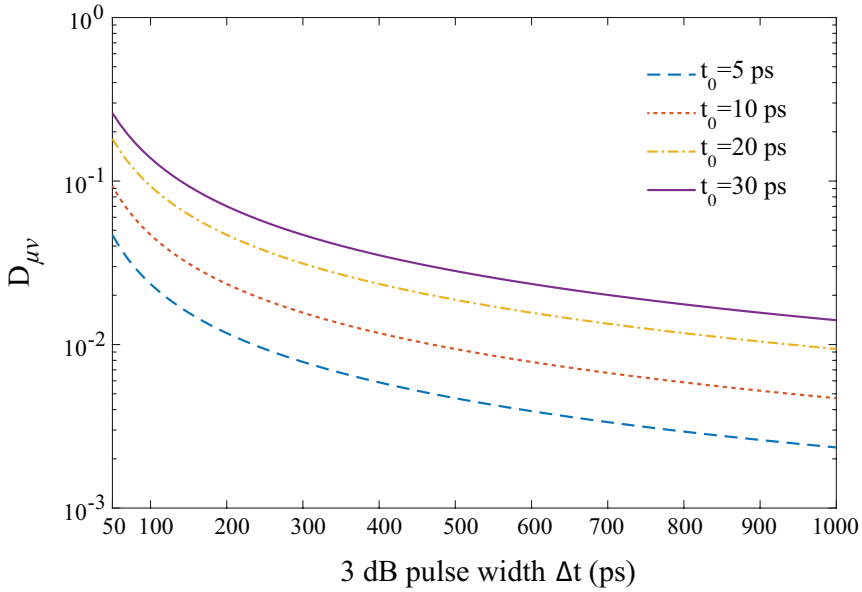

Figure 5.4: The imperfection of the signal state and the decoy state, $D_{\mu\nu}$, for different widths of pulses.

assumed that the shape of the signal pulse and the decoy pulse follow Gaussian distribution. After attenuated to be single-photon level, the probability distributions of the signal state and the decoy state are

$$\begin{aligned} f_\mu(t) &= \frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{t^2}{2\sigma^2}], \\ f_\nu(t) &= \frac{1}{\sqrt{2\pi}\sigma}\exp[-\frac{(t-t_0)^2}{2\sigma^2}]. \end{aligned} \tag{5.20}$$

Here t_0 is the centering shift between the signal state and the decoy state; σ is the standard deviation of Gaussian distribution. It is assumed the signal and decoy states have the same standard deviation. If the 3 dB pulse width is Δt, it is easy to know that $\sigma = \Delta t/\sqrt{8\ln 2}$. Submitting Eq. (5.20) into Eq. (5.19), the distance of the signal state and the decoy state can be then

$$D_{\mu\nu} = \frac{1}{2}\operatorname{erf}\frac{t_0}{2\sqrt{2}\sigma} = \frac{1}{2}\operatorname{erf}\frac{\sqrt{\ln(2)}t_0}{\Delta t}. \tag{5.21}$$

Here $\operatorname{erf} x = \int_0^x \frac{2}{\sqrt{\pi}}e^{-x^2}dx$ is the error correction function. The result in Fig. 5.4 clearly illustrates that even a small timing difference t_0 will leak secret information to Eve. In

a practical QKD system, it may be challenging for Alice to fully match the central time of the signal state and decoy state (i.e., $t_0 = 0$). However, Fig. 5.4 shows that Alice is able to reduce $D_{\mu\nu}$ by broadening her pulse. For instance, in the case that $t_0 = 10$ ps, $D_{\mu\nu} = 0.0931$ when $\Delta t = 100$ ps, but it decreases to $D_{\mu\nu} = 0.0094$ when $\Delta t = 1$ ns. Furthermore, it is remarkable that even though Fig. 5.4 illustrates the mismatch between the signal state and the decoy state in time, the method of characterization is effective for any other dimensions, for example, frequency, spatial mode in free space, and so on.

5.5.2 Lower bound of Y_1^μ

Now we calculate the lower bound of Y_1^μ from the model given above. According to Eq. (5.19), it is easy to obtain inequalities

$$\begin{aligned} |Y_n^\omega - Y_n^{\omega'}| &\leq 2D_{\omega\omega'}, \\ |Y_n^\omega e_n^\omega - Y_n^{\omega'} e_n^{\omega'}| &\leq 2D_{\omega\omega'}, \end{aligned} \tag{5.22}$$

where $\omega, \omega' = \mu, \nu, \nu_1$ and $1 > \mu > \nu > \nu_1 > 0$. The lower bound of Y_1^μ can be then estimated by following the procedure in Ref 120. We assume that $Y_0^\mu = Y_0^\nu = Y_0^{\nu_1} = Y_0$, since the vacuum pulse are the same and cannot provide any information to Eve. Hence, the lower bound of the background rate Y_0 can be estimated as follows.

$$\begin{aligned} &\nu e^{\nu_1} Q_{\nu_1} - \nu_1 e^{\nu} Q_\nu = \sum_{n=0}^{\infty} \frac{1}{n!} (\nu \nu_1^n Y_n^{\nu_1} - \nu_1 \nu^n Y_n^\nu) \\ &= (\nu - \nu_1) Y_0 + \sum_{n=1}^{\infty} \frac{\nu\nu_1}{n!} (\nu_1^{n-1} Y_n^{\nu_1} - \nu^{n-1} Y_n^\nu) \\ &\leq (\nu - \nu_1) Y_0 + \sum_{n=1}^{\infty} \frac{\nu\nu_1}{n!} [\nu_1^{n-1} (Y_n^\mu + 2D_{\mu\nu_1}) - \nu^{n-1} (Y_n^\mu - 2D_{\mu\nu})] \\ &= (\nu - \nu_1) Y_0 + \sum_{n=1}^{\infty} [2D_{\mu\nu_1} \frac{\nu\nu_1}{n!} \nu_1^{n-1} + 2D_{\mu\nu} \frac{\nu\nu_1}{n!} \nu^{n-1} + \frac{\nu\nu_1}{n!} (\nu_1^{n-1} - \nu^{n-1}) Y_n^\mu] \\ &= (\nu - \nu_1) Y_0 + 2D_{\mu\nu_1} \nu (e^{\nu_1} - 1) + 2D_{\mu\nu} \nu_1 (e^\nu - 1) - \nu\nu_1 \sum_{n=1}^{\infty} \frac{\nu^{n-1} - \nu_1^{n-1}}{n!} Y_n^\mu \\ &\leq (\nu - \nu_1) Y_0 + 2D_{\mu\nu_1} \nu (e^{\nu_1} - 1) + 2D_{\mu\nu} \nu_1 (e^\nu - 1) \\ &\equiv (\nu - \nu_1) Y_0 + g'(\mu, \nu, \nu_1). \end{aligned} \tag{5.23}$$

Note that $\nu^{n-1} - \nu_1^{n-1} \geq 0$ is applied above. From inequality (5.23), the lower bound of Y_0 can be then obtained

$$Y_0 \geq Y_0^L = \max\{\frac{1}{\nu - \nu_1}[\nu e^{\nu_1} Q_{\nu_1} - \nu_1 e^{\nu} Q_{\nu} - g'(\mu, \nu, \nu_1)], 0\}. \tag{5.24}$$

We can estimate the lower bound of Y_1 as shown below.

$$\begin{aligned}
& e^{\nu} Q_{\nu} - e^{\nu_1} Q_{\nu_1} = \sum_{n=0}^{\infty} \frac{\nu^n}{n!} Y_n^{\nu} - \sum_{n=0}^{\infty} \frac{\nu_1^n}{n!} Y_n^{\nu_1} \\
&= Y_0^{\nu} - Y_0^{\nu_1} + \nu Y_1^{\nu} - \nu_1 Y_1^{\nu_1} + \sum_{n=2}^{\infty} \frac{1}{n!} (\nu^n Y_n^{\nu} - \nu_1^n Y_n^{\nu_1}) \\
&\leq \nu(Y_1^{\mu} + 2D_{\mu\nu}) - \nu_1(Y_1^{\mu} - 2D_{\mu\nu_1}) + \sum_{n=2}^{\infty} \frac{1}{n!} [\nu^n (Y_n^{\mu} + 2D_{\mu\nu}) - \nu_1^n (Y_n^{\mu} - 2D_{\mu\nu_1})] \\
&= (\nu - \nu_1) Y_1^{\mu} + 2\nu D_{\mu\nu} + 2\nu_1 D_{\mu\nu_1} + \sum_{n=2}^{\infty} \frac{1}{n!} [(\nu^n - \nu_1^n) Y_n^{\mu} + 2\nu^n D_{\mu\nu} + 2\nu_1^n D_{\mu\nu_1}] \\
&= (\nu - \nu_1) Y_1^{\mu} + \sum_{n=2}^{\infty} \frac{1}{n!} (\nu^n - \nu_1^n) Y_n^{\mu} + 2D_{\mu\nu}(e^{\nu} - 1) + 2D_{\mu\nu_1}(e^{\nu_1} - 1) \\
&\leq (\nu - \nu_1) Y_1^{\mu} + \frac{\nu^2 - \nu_1^2}{\mu^2} (e^{\mu} Q_{\mu} - Y_0^L - \mu Y_1^{\mu}) + \frac{\mu(\nu - \nu_1) - (\nu^2 - \nu_1^2)}{\mu} g(\mu, \nu, \nu_1) \\
&= \frac{\mu(\nu - \nu_1) - (\nu^2 - \nu_1^2)}{\mu} Y_1^{\mu} + \frac{\nu^2 - \nu_1^2}{\mu^2} (e^{\mu} Q_{\mu} - Y_0^L) + \frac{\mu(\nu - \nu_1) - (\nu^2 - \nu_1^2)}{\mu} g(\mu, \nu, \nu_1).
\end{aligned} \tag{5.25}$$

Here, the inequality $\frac{\nu^n - \nu_1^n}{\mu^n} \leq \frac{\nu^2 - \nu_1^2}{\mu^2}$ is used for all $n \geq 2$, and Y_0^L is given by Eq. (5.24). In Eq. (5.25),

$$g(\mu, \nu, \nu_1) \equiv \frac{2\mu[D_{\mu\nu}(e^{\nu} - 1) + D_{\mu\nu_1}(e^{\nu_1} - 1)]}{\mu(\nu - \nu_1) - (\nu^2 - \nu_1^2)}. \tag{5.26}$$

It is assumed $\mu \geq \nu + \nu_1$. By deriving inequality (5.25), the lower bound of Y_1^{μ} is the following.

$$\begin{aligned}
Y_1^{\mu} &\geq \frac{\mu[e^{\nu} Q_{\nu} - e^{\nu_1} Q_{\nu_1} - \frac{\nu^2 - \nu_1^2}{\mu^2}(e^{\mu} Q_{\mu} - Y_0^L)]}{\mu(\nu - \nu_1) - (\nu^2 - \nu_1^2)} - g(\mu, \nu, \nu_1) \\
&\equiv G(\mu, \nu, \nu_1) - g(\mu, \nu, \nu_1).
\end{aligned} \tag{5.27}$$

$G(\mu,\nu,\nu_1)$, the same as Eq. (21) in Ref. 120, represents the yield of the single photon pulse with a perfect source. Thus, $g(\mu,\nu,\nu_1)$ means the leaky information because of the imperfection of source.

In general, weak + vacuum state protocol is applied in experiments, which means $\nu_1 = 0$. Hence, the lower bound of Y0 can be obtained from Eq. (5.24): $Y_0^L = Q_{\nu_1=0} = Q_{\text{vac}}$. Equation (5.27) can be then rewritten as

$$Y_1^\mu \geq \frac{\mu}{\mu\nu-\nu^2}[e^\nu Q_\nu - \frac{\nu^2}{\mu^2}e^\mu Q_\mu - \frac{\mu^2-\nu^2}{\mu^2}Q_{\text{vac}} - 2D_{\mu\nu}(e^\nu-1)]. \tag{5.28}$$

5.5.3 Upper bound of e_1^μ

Now we calculate the upper bound of e_1^μ. Since

$$e^\omega Q_\omega E_\omega = \sum_{n=0}^{\infty}\frac{\omega^n}{n!}Y_n^\omega e_n^\omega \geq e_0 Y_0 + \omega Y_1^\omega e_1^\omega, \tag{5.29}$$

we can have

$$e_1^\mu \leq \frac{e^\mu Q_\mu E_\mu - e_0 Y_0^L}{\mu Y_1^\mu} \equiv K^\mu, \tag{5.30a}$$

$$e_1^\mu \leq \frac{e_1^\nu Y_1^\nu + 2D_{\mu\nu}}{Y_1^\mu} \leq \frac{e^\nu Q_\nu E_\nu - e_0 Y_0^L + 2\nu D_{\mu\nu}}{\nu Y_1^\mu} \equiv K^\nu, \tag{5.30b}$$

$$e_1^\mu \leq \frac{e_1^{\nu_1} Y_1^{\nu_1} + 2D_{\mu\nu_1}}{Y_1^\mu} \leq \frac{e^{\nu_1} Q_{\nu_1} E_{\nu_1} - e_0 Y_0^L + 2\nu_1 D_{\mu\nu_1}}{\nu_1 Y_1^\mu} \equiv K^{\nu_1}. \tag{5.30c}$$

Meanwhile, we have

$$\begin{aligned}
&Q_\mu E_\mu e^\mu - Q_\nu E_\nu e^\nu \\
&= \sum_{n=0}^{\infty}[\frac{\mu^n}{n!}Y_n^\mu e_n^m u - \frac{\mu^n}{n!}Y_n^\mu e_n^m u] \\
&= uY_1^\mu e_1^\mu - \nu Y_1^\nu e_1^\nu + \sum_{n=2}^{\infty}\frac{\mu^n Y_n^\mu e_n^\mu - \nu Y_n^\nu e_n^\nu}{n!} \\
&\geq (\mu-\nu)Y_1^\mu e_1^\mu - 2D_{\mu\nu}(\nu + \sum_{n=2}^{\infty}\frac{\nu^n}{n!}) + \sum_{n=2}^{\infty}\frac{\mu^n-\nu^n}{n!}Y_n^\mu e_n^\mu \\
&= (\mu-\nu)Y_1^\mu e_1^\mu - 2D_{\mu\nu}(e^\nu - 1).
\end{aligned} \tag{5.31}$$

Here we reasonably consider that $Y_0^\mu = Y_0^\nu$ and $e_0^\mu = e_0^\nu$, due to no difference for the vacuum pulse. We then obtain

$$e_1^\mu \leq \frac{e^\mu Q_\mu E_\mu - e^\nu Q_\nu E_\nu + 2D_{\mu\nu}(e^\nu - 1)}{(\mu - \nu)Y_1^\mu} \equiv K^{\mu\nu}. \tag{5.32}$$

Similarly, we have

$$e_1^\mu \leq \frac{e^\mu Q_\mu E_\mu - e^{\nu_1} Q_{\nu_1} E_{\nu_1} + 2D_{\mu\nu_1}(e^{\nu_1} - 1)}{(\mu - \nu_1)Y_1^\mu} \equiv K^{\mu\nu_1}. \tag{5.33}$$

Thus, we can estimate the upper bound of e_1^μ by

$$e_1^\mu \leq \min\{K^\mu, K^\nu, K^{\nu_1}, K^{\mu\nu}, K^{\mu\nu_1}\}. \tag{5.34}$$

For weak + vacuum decoy state protocol, $\nu_1 = 0$ and $Y_0 = Q_{\text{vac}}$. Then $Q_{\nu_1} E_{\nu_1} = e_0 Y_0$. Inequality (5.30c) becomes

$$e_1^\mu \leq \frac{1}{Y_1^\mu}, \tag{5.35}$$

which is hold by the definition of e_1. Also, $K^{\mu\nu_1} = K^\mu$. Thus, the upper bound of e_1^μ can be rewritten as

$$e_1^\mu \leq \min\{K^\mu, K^\nu, K^{\mu\nu}\}. \tag{5.36}$$

5.5.4 Numerical simulation

We simulate the weak + vacuum decoy state protocol with an imperfect source as shown in Fig. 5.5. If there is no Eve, the total gain and the error rate of the signal state and the decoy state are given by Eqs. (5.12) and (5.14). The rest of parameters follow the definition in standard decoy state protocol [120]. Submitting Eqs. (5.12) and (5.14) into Eqs. (5.28) and (5.36), the lower bound of yield and the upper bound of the error rate for the single photon pulse are shown in Fig. 5.5(a) and (b). The estimated key rate is then illustrated in Fig. 5.5(c). The detection parameters used in the simulations are the same as those in Fig. 5.3. The intensities of the signal state and the decoy state are optimized with step 0.01 from $\mu \in [0.01, 0.5]$, $\nu \in [0.01, 0.2]$. We only show the estimated Y_1^μ and e_1^μ where the final key rate is positive. No secure key can be generated for $D_{\mu\nu} = 10^{-1}$ and 10^{-2}. The results clearly show that the imperfection of source decreases the key rate between Alice and Bob rapidly. For example, when the source is perfect, the maximum distance can reach about 141 km. However, the maximum distances have decreased to $124, 92, 48$ km for $D_{\mu\nu} = 10^{-5}, 10^{-4}, 10^{-3}$. No positive key rate at any distance for $D_{\mu\nu} = 10^{-2}, 10^{-1}$.

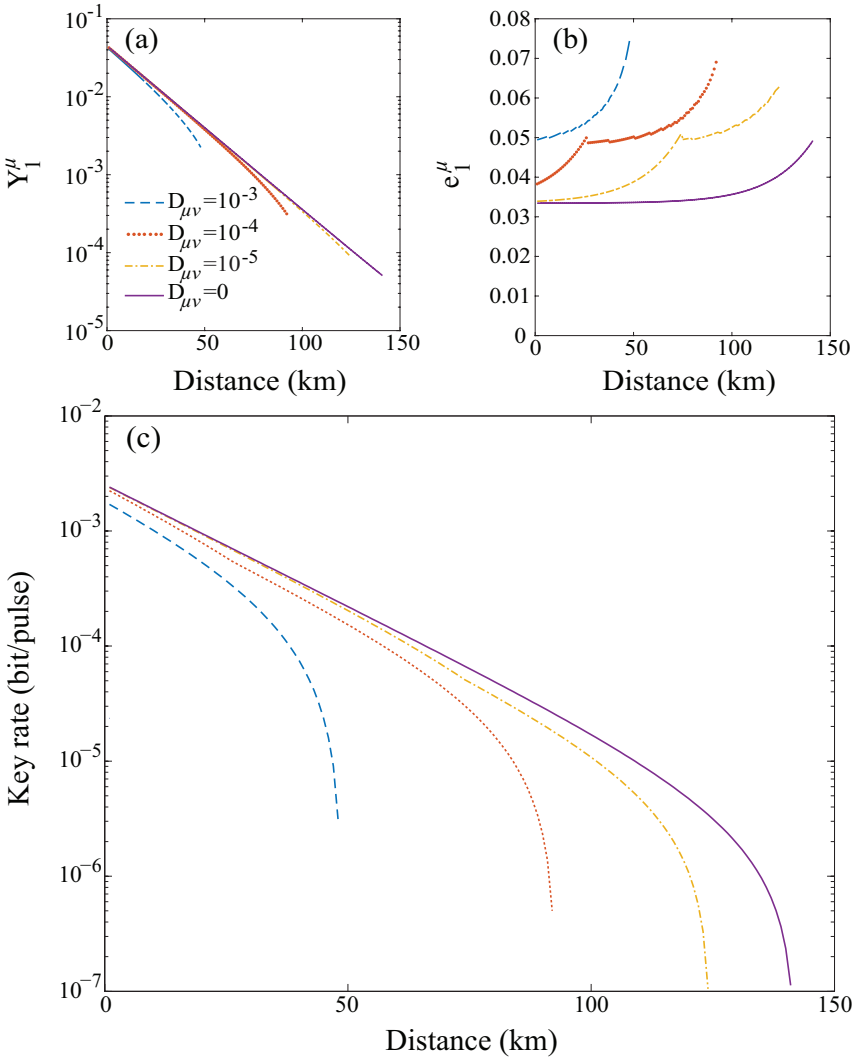

Figure 5.5: Estimated system parameters with imperfect source. (a) the yield and (b) the error rate of the signal state for the single photon pulse, and (c) the key rate are shown for different amounts of imperfection $D_{\mu\nu}$.

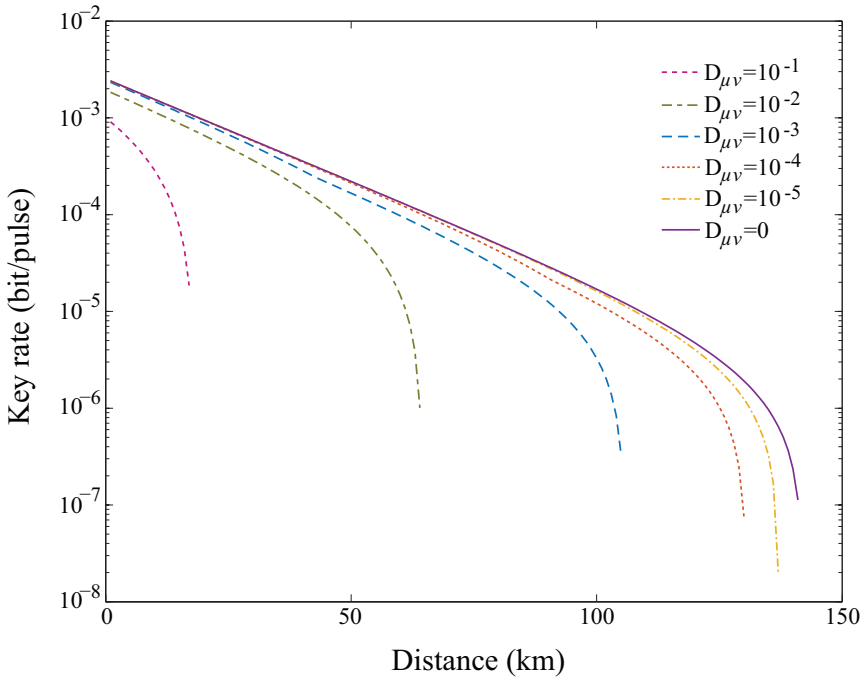

Figure 5.6: The estimated key rate assuming calibrated transmittance in Bob's optical devices. The detection parameters used here are the same as those in Fig. 5.3.

5.5.5 Theory improvement

From the modified security proof in the last section, it is shown that even for the relatively small imperfection $D_{\mu\nu} = 10^{-3}$, the maximum secure distance under the tightened key rate drops quickly from 141 km to 48 km. Here, we propose an advanced security proof to improve the final key rate with the imperfect source by setting a reasonable assumption. We could then loosen the security constraint when estimating Y_1^{μ} and e_1^{μ}, theoretically improving the secure key rate and the maximum secure distance.

In practical QKD systems based on a prepare-and-measurement protocol, Bob's devices are located within his protected zone. Thus, it is possible for Bob to calibrate the optical transmittance of his optical devices. Note that here we do not mean that Eve could not change the parameters of Bob's system (for example, change the SPD from Geiger mode to linear mode by performing the blinding attack), but mean that Bob could actively calibrate the transmittance of his partial devices or all devices. In fact, this assumption has been used to secure the single photon detector of Bob [124,125]. Thus, we think this assumption

is reasonable and practical. We can then have

$$Y_n(\lambda) \leq 1 - (1 - \eta_{\text{Bob}}^{\text{cal}})^n. \tag{5.37}$$

Please note that $\eta_{\text{Bob}}^{\text{cal}}$ is the calibrated transmittance in Bob, which should be equal to or lower than the total transmittance of Bob η_{Bob}. In the simulation, we could easily assume that Bob can calibrate the whole transmittance in his apparatus. We could then have $\eta_{\text{Bob}}^{\text{cal}} = \eta_{\text{Bob}}$. Thus, Eq. (5.22) becomes

$$\begin{aligned} |Y_n^{\omega} - Y_n^{\omega'}| &\leq 2D_{\omega\omega'}[1 - (1 - \eta_{\text{Bob}})^n], \\ |Y_n^{\omega} e_n^{\omega} - Y_n^{\omega'} e_n^{\omega'}| &\leq 2D_{\omega\omega'}[1 - (1 - \eta_{\text{Bob}})^n]. \end{aligned} \tag{5.38}$$

In the weak + vacuum decoy state protocol, the lower bound of Y_1^{μ} [Eq. (5.28)] and the upper bound of e_1^{μ} [Eq. (5.36)] can be obtained

$$\begin{aligned} Y_1^{\mu} \geq\ & \frac{\mu}{\mu\nu - \nu^2}[e^{\nu} Q_{\nu} - \frac{\nu^2}{\mu^2} e^{\mu} Q_{\mu} - \frac{\mu^2 - \nu^2}{\mu^2} Y_0 \\ & - 2D_{\mu\nu}(e^{\nu} - e^{\nu(1-\eta_{\text{Bob}})})], \\ e_1^{\mu} \leq\ & \min\{K^{\mu}, K^{\nu}, K^{\mu\nu}\}, \text{where} \\ K^{\mu} =\ & \frac{e^{\mu} Q_{\mu} E_{\mu} - e_0 Y_0}{\mu Y_1^{\mu}}, \\ K^{\nu} =\ & \frac{e^{\nu} Q_{\nu} E_{\nu} - e_0 Y_0 + 2\nu D_{\mu\nu}\eta_{\text{Bob}}}{\nu Y_1^{\mu}}, \\ K^{\mu\nu} =\ & \frac{e^{\mu} Q_{\mu} E_{\mu} - e^{\nu} Q_{\nu} E_{\nu} + 2D_{\mu\nu}(e^{\nu} - e^{\nu(1-\eta_{\text{Bob}})})}{(\mu - \nu) Y_1^{\mu}}. \end{aligned} \tag{5.39}$$

We can then estimate the final key rate following the same method provided above. The estimation result in Fig. 5.6 shows that when the transmittance of Bob's optical devices ($\eta_{\text{Bob}} = 4.5\%$) is calibrated, the final key rate and the maximum distance are improved. For instance, if $D_{\mu\nu} = 10^{-3}$, the maximum distance increases from 48 km to 105 km. We may also note that for $D_{\mu\nu} = 10^{-2}$ and 10^{-1}, the improved proof can generate the positive key rate up to 64 km and 18 km. We would remark here that the assumption of calibrated transmission loss for Bob's devices is not applicable to measurement-device-independent QKD (MDI QKD), in which the detection part is not in the protected zone and can be fully controlled by Eve.

5.6 Discussion and application examples

The security proof in Sec. 5.5 considers a type of imperfect source that could partially distinguish signal and decoy states in any degrees of freedom. Once these imperfections are experimentally measured, the security proof proposed in this work may provide a standard method to calculate the final key rate under such imperfections. Our security proof focuses on the imperfect modulation of signal and decoy states, but does not handle the distinguishability among different BB84 states. We currently assume the identical mismatch of signal and decoy states for each BB84 state. Removing this theoretical limitation could be future work.

Another limitation lies in our experiment. We have measured the distinguishability between signal and decoy states only in the time domain. However, the two modulation methods that we have tested might also introduce time-dependent spectral mismatch, which we have not measured. For the gain-switched semiconductor laser, a short pulse usually has a so-called chirp, a fast-changing wavelength modulation [91, 104]. The spectral and intensity modulation contribute simultaneously to the distinguishability, resulting in a joint distribution of $D_{\mu\nu}$ as explained later in this section. The external intensity modulator may also affect the spectrum of pulses [88, 93]. However, the requisite time-resolved spectroscopy is a more complex measurement [104, 133, 151], which could be investigated in the future. For the two devices tested, we henceforth assume distinguishability in the time domain only.

We now apply our security proof to the measurement results of the two sources tested in Sec. 5.3, and to one more published source measurement in Ref. 132. Both the initial method in Sec. 5.5.4 and the improved method in Sec. 5.5.5 are applied in each case. The purpose of this application is quantifying the imperfection of signal state and decoy state preparation, and showing its effect on the secure key rate. To compare the three source implementations, we arbitrarily assume that these different types of the sources are used in the same fiber-based QKD system with GYS parameters at the detection side. The resulting secure key rates are shown in Fig. 5.7. The key rate is estimated by the initial proof with Eq. (5.28) and Eq. (5.36) and the improved proof with Eq. (5.39). Note that in the case of $D_{\mu\nu} = 0.1400$, no secure key can be generated with the initial proof, but a positive key rate is possible with our improved proof. $D_{\mu\nu} = 0.4005$ cannot generate the positive key rate in either proof.

For the first case shown in Fig. 5.1(a), the corresponding value of $D_{\mu\nu}$ given by Eq. (5.19) is 0.4005. Both our security proofs are then applied. For the improved proof, we assume $\eta_{\mathrm{Bob}} = 4.5\%$. The secure key rate is *zero* under either key rate estimation. This verifies again that the modulation imperfection makes the system insecure.

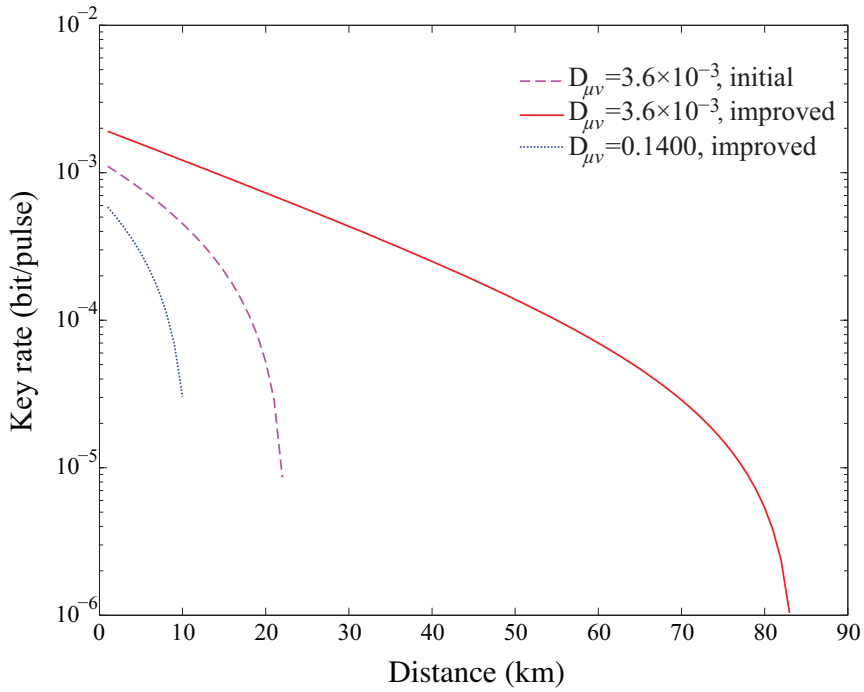

Figure 5.7: The estimated key rate for different experimental distinguishability of signal and decoy states. The detection parameters used here are the same as in Fig. 5.3.

On the contrary, the value of $D_{\mu\nu}$ for the case in Fig. 5.2 is only 3.6×10^{-3}. This non-zero value probably stems from the noise in our characterization apparatus. Nevertheless, we should conservatively treat all mismatch as belonging to the source under test. This non-zero value of $D_{\mu\nu}$ still indicates the certain degree of mismatch. As shown in Fig. 5.7, under the initial proof, the maximum distance drops to 22 km, while under the advanced proof with $\eta_{\text{Bob}} = 4.5\%$, the key rate is improved to 83 km. That is, owing to a much lower mismatch in this case, the positive key rate could be generated.

Another case of imperfect preparation for signal and decoy states is published in Ref. 132. In that study, the signal and decoy states are generated by individual laser diodes, which is a common technique [99, 106, 135, 187]. It shows that mismatches between signal and decoy states are both in the time domain and frequency domain for each individual BB84 state [132]. Being that our proof cannot handle the BB84 states individually, we have chosen a typical mismatch between the signal and decoy states in vertical polarization as reprinted in Fig. 5.8, and assumed arbitrarily that the other three BB84 polarization states have the mismatch identical to that. Even though Ref. 132 studies an imperfect source in a free-space QKD system, we remark that it is reasonable to expect mismatch for any QKD implementations that generate signal and decoy states by individual laser diodes [99, 106, 135, 187].

The security proof in Sec. 5.5 is able to handle mismatch in arbitrary degrees of freedom, being that we do not specify the dimensions of the probability $f_\omega(\lambda)$. $f_\omega(\lambda)$ can be a joint probability. For example, the joint probability distribution of ω state in the time and frequency domains can be $f_\omega(t, f)$, where t represents the time domain and f represents the frequency domain. Thus, $D_{\mu\nu}$ can be defined as

$$D_{\mu\nu} = \frac{1}{2} \sum_t \sum_f |f_\mu(t, f) - f_\nu(t, f)|. \tag{5.40}$$

Similarly, the calculation of $D_{\mu\nu}$ can be expanded to more than two dimensions. In the specific case showed in Fig. 5.8, time-resolved spectroscopy necessary to measure the joint probability was not performed. It has been arbitrarily assumed instead that the probability distributions in the time and frequency domains are independent, with a remark that this will need to be verified experimentally [132]. Then, $f_\omega(t, f) = f_\omega(t) f_\omega(f)$ can be calculated from the available experimental data. The corresponding $D_{\mu\nu}$ is 0.1400. With such value of $D_{\mu\nu}$, the initial security proof cannot generate the positive key rate for this case, while the improved proof with $\eta_{\text{Bob}} = 4.5\%$ could generate the secure key up to only 10 km, as shown in Fig. 5.7.

According to the above analysis and comparison of the three cases, the external intensity modulator shows the smallest mismatch between signal and decoy states, resulting in the

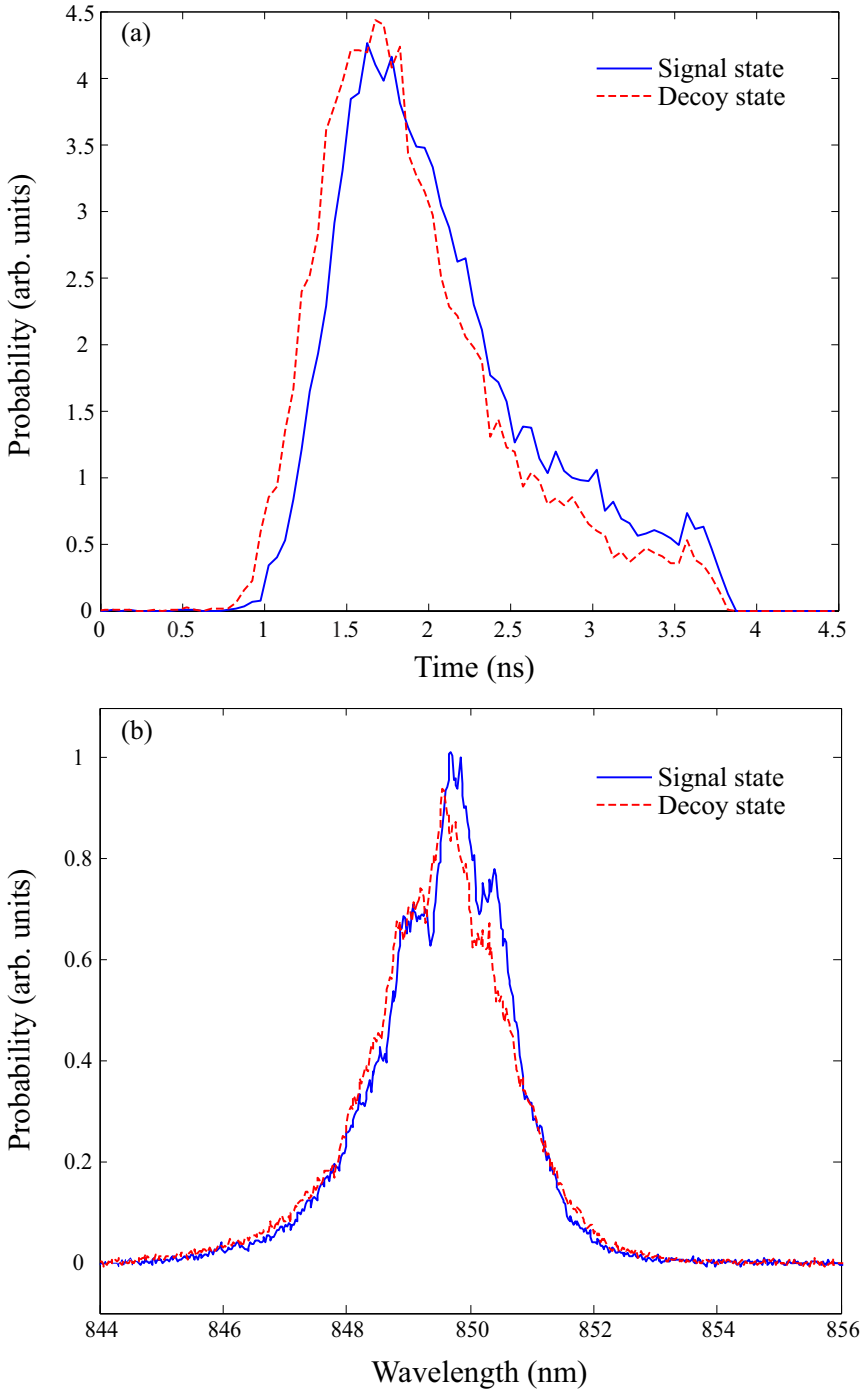

Figure 5.8: Mismatch of signal and decoy states for the vertical-polarization pair of laser diodes in (a) time domain and (b) frequency domain. Data reprinted from Ref. 132.

highest key rate and longest transmission distance. This indicates that modulating signal and decoy states by the intensity modulator could be a proper method to realize the decoy state protocol. However, Trojan-horse attack could read out the modulation information from the intensity modulator [165,172]. Thus, countermeasures against Trojan-horse attack are also necessary [113, 148]. The intensity modulator may also result in the non-zero vacuum state, which can be handled by our security proof. The secure key rate can then be estimated by applying Eqs. (5.27) and (5.34).

All the measurement results shown above contain a measurement error. The error may come from the thermal noise of electronic devices, the nonlinearity of optical-to-electrical converters and digital-to-analog converters in the oscilloscope. We have simply treated the measured $D_{\mu\nu}$ as the real mismatch. Thus, the key rates showed in Fig. 5.7 are conservative estimates. It is an open question of how to extract the real parameter from the noisy test results.

5.7 Conclusion

In this chapter, we have investigated the imperfect sources in QKD systems that implement the decoy state protocol. By testing two intensity modulation methods, we have found that the basic assumption about the indistinguishability of signal and decoy states does not hold in practice, especially in the case of laser diode pump current modulation. This pump-current modulation shows timing mismatch between the signal and decoy states. We have modeled a PNS attack based on the timing mismatch that breaks the security of the QKD system. To make the system robust against this loophole, we have modified the method of generating the secure key rate in the decoy state protocol to consider an imperfect source, in which a signal state and a decoy state are distinguishable in any degrees of freedom. Our proof proves that the distinguishability would reduce the secure key rate. We have applied our proof to three implementations of the decoy state protocol to estimate the secure key rate that, in some cases, has become reduced (also limiting the transmission distance) and in some cases, is just zero.

The evaluation of the three types of intensity modulation indicates that implementing the decoy state by an external intensity modulator is superior to the other methods. It leaves less distinguishability between signal and decoy states, consequently maintaining the higher key rate. We do not recommend the other two methods: pump-current modulation and individual-laser-diode generation, because they show a significant mismatch. Time-resolved spectroscopy should be performed in the future to check whether it enlarges the imperfection $D_{\mu\nu}$.

Our security proof with an imperfect source provides a general method to guarantee the security of practical quantum cryptographic systems. It may be employed as a standard tool to estimate the secure key rate, once the source imperfection in the decoy state protocol is quantified in all degrees of freedom. A conceptually similar evaluation method has been proposed for the Trojan-horse attack [113].

Chapter 6

Laser seeding attack on the source

6.1 Motivation

The security of the source is vital to any QKD system. Unfortunately, this important requirement is not always satisfied in practice. For example, the widely-used weak coherent source sometimes generates multiple photons, which enables the PNS attack (see Sec. 2.3.1). This problem can be practically solved by the decoy-state protocol (see Sec. 5.2). Nevertheless, other side channels created by Eve might also compromise the security of QKD systems.

In this work, we study Eve's capability to manipulate a laser diode in the apparatus of state preparation. We show that Eve can increase the output power of the laser diode, which increases the mean photon number of the optical pulses. In this attack, Eve injects photons into a laser diode to trigger a stronger stimulated emission than usual, which results in an increased intensity. We test two types of laser diodes and analyze the effect on the security of the standard QKD protocol.

6.2 Experimental scheme and principle

To investigate Eve's ability to control a semiconductor laser diode, we conduct our testing according to the scheme shown in Fig. 6.1. At Alice's side, the laser diode generates optical pulses as a testing target. An eavesdropper Eve employs a tunable laser (Agilent 8164B) to send continuous-wave (c.w.) bright light to Alice via a fiber. The tunable laser allows

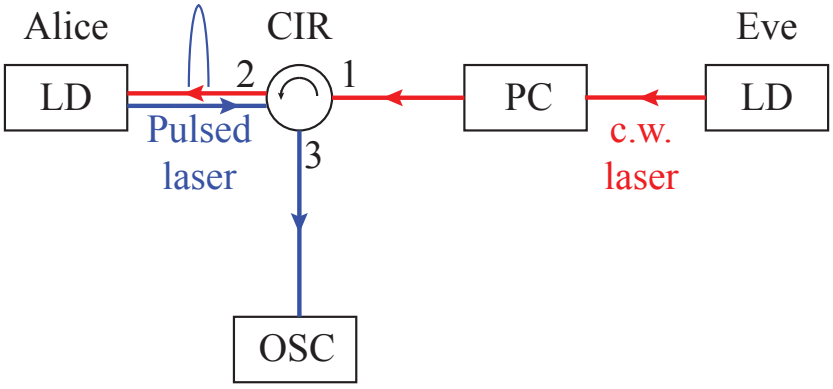

Figure 6.1: Experimental scheme of laser seeding test. The red path represents Eve's injection, and the blue path shows the direction of photon emission from Alice's laser. LD: laser diode; CIR: circulator; PC: polarization controller; OSC: oscilloscope.

the adjustment of its wavelength and the output power. Thus, Eve injects photons with a particular wavelength into Alice's laser. With this wavelength, the energy of injected photons matches the energy difference between the excited state and the ground state of electrons in Alice's laser, which satisfies the condition of stimulated emission in the laser diode.

A polarization controller is applied to adjust Eve's laser, matching the polarization of Alice's laser. This adjustment maximizes the injection efficiency. We employ a circulator to isolate the injected light From Eve and the emitting light from Alice's laser. Eve sends the tampering light from port 1 to port 2, while the output of Alice's laser goes through port 2 to port 3. After that, Alice's output energy is measured by an optical-to-electronic converter (with 1 GHz bandwidth) that connects to an oscilloscope (Agilent DSOX93303Q) with 33 GHz bandwidth. Hence, we can observe the energy of Alice's laser pulse with and without Eve's tampering. We test two ID300 short-pulse lasers from ID Quantique and one Thorlabs-LP1550-SAD2 laser diode as Alice's laser. All the laser diodes under test generate optical pulses with a repetition frequency of 1 MHz.

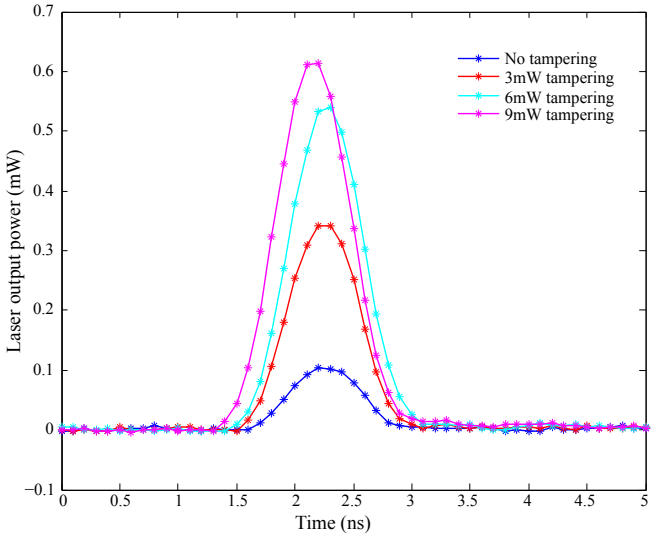

Figure 6.2: Waveforms of ID300's laser pulses with and without Eve's tampering.

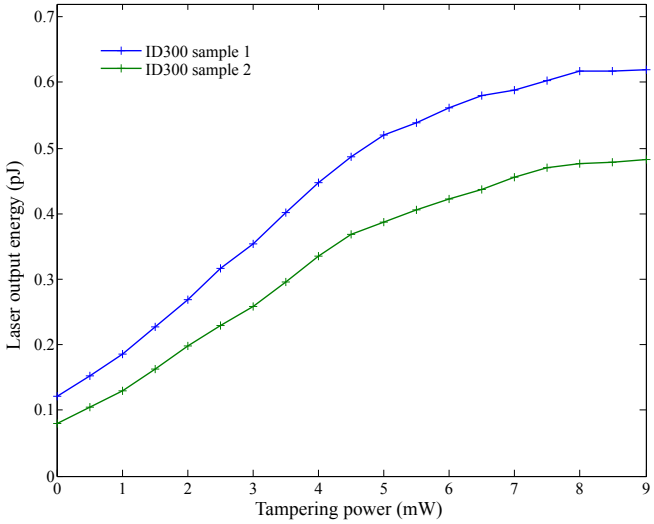

Figure 6.3: Alice's output energy versus Eve's tampering power from two samples of ID300 from ID Quantique.

6.3 Experimental results

Two samples of the ID300 laser have been tested, and the testing results from these two lasers show Eve's ability to control the output power. By following the testing scheme in Fig. 6.1, we first adjust the wavelength of Eve's laser to 1556.90 nm and 1557.18 nm for sample 1 and sample 2, respectively. Once finding the correct wavelength, we gradually increase the power of the c.w. laser. The energy of the optical pulses emitted from Alice increases as c.w. power increases. Figure 6.2 shows the waveforms of the pulses under Eve's attack. Compared to the original laser pulse, the amplitude of the pulse during the attack becomes much higher than usual, and increasing injection power results in an increasing output power. Secondly, the injected light also broadens the width of Alice's optical pulse. This is because a stimulated emission triggered by the injected light takes less time than the spontaneous emission which usually takes place.

We measure the pulse energy under various levels of tampering power as shown in Fig. 6.3. For each sample of the ID300 laser, the initial energy without light injection is 0.12 pJ for sample 1 and 0.08 pJ for sample 2. The injected c.w. laser then increases gradually to 9 mW, which triggers the increase of Alice's pulse energy. Finally, the output energy of Alice's laser rises to 0.62 pJ for sample 1 and 0.48 pJ for sample 2. Under 9 mW light injection, the pulse energy increases 5.2 times for sample 1 and 6.0 times for sample 2.

The increased intensity is verified again for the Thorlabs-LP1550-SAD2 laser, albeit less spectacularly. The c.w. light with a wavelength of 1551.32 nm is injected into Alice's laser and triggers a higher output power than usual. Figure 6.4 shows the waveforms of four specific cases: no tampering power, 3 mW tampering power, 6 mW tampering power, and 9 mW tampering power. It is obvious that, under the attack, the amplitude becomes higher along with the increase of the tampering power. The quantified increase of pulse energies under different amounts of tampering power is shown in Fig. 6.5. Initially, the energy of Alice's laser pulse is 2.217 pJ. The pulse energy then gradually rises under the attack, and reaches 2.388 pJ with 9 mW tampering power. The energy of the tampered pulse is 1.077 times higher than the original one.

6.4 Security analysis under the attack

This laser seeding attack exploits the vulnerability of the laser in Alice, so that Eve can manipulate and increase the output power of Alice's laser. The increased power results in a larger mean photon number than the value that Alice and Bob expected, which

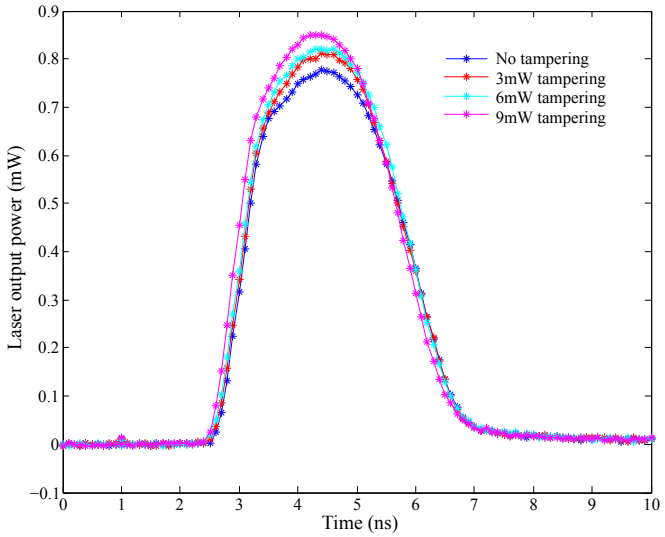

Figure 6.4: Waveforms of laser pulses emitted from Thorlabs-LP1550-SAD2 with and without Eve's tampering.

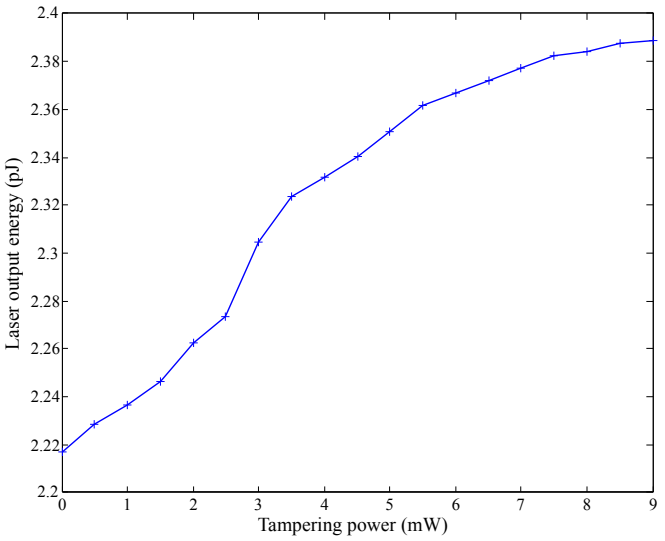

Figure 6.5: Alice's output energy versus Eve's tampering power from Thorlabs-LP1550-SAD2.

compromises the security of a QKD system. Here, we take the case of doubled intensities under the attack as an example to show the security threat to a QKD system. In this security analysis, we follow the security model of the BB84 QKD protocol with decoy states. As we have shown in Chapter 5, the key rate of a QKD system with a weak coherent laser source is given by the Gottesman-Lo-Lütkenhaus-Preskill (GLLP) model [70] as follows.

$$R \geq q\{-Q_\mu H_2(E_\mu) f(E_\mu) + P_1^\mu Y_1^\mu [1 - H_2(e_1^\mu)]\}, \tag{6.1}$$

where all the parameters also follow the same definitions described previously in Chapter 5. In particular, μ represents the average intensity of the signal state; Q_μ and E_μ are the total gain and error rate of the signal state; Y_1^μ and e_1^μ are the yield and the error rate of single-photon pulses; P_1^μ is the probability of single-photon pulses.

In the GLLP model, μ is a preset parameter that is shared between Alice and Bob. Once μ is decided and fixed, P_1^μ is automatically known according to the Poisson distribution of the weak coherent source. On the other hand, Q_μ and E_μ are experimental parameters that Bob obtains during the raw key exchange. The decoy states then help Alice and Bob tightly estimate the lower bound of Y_1^μ and the upper bound of e_1^μ [120]. As we have mentioned previously in Chapter 5, the commonly used decoy protocol applies a weak decoy state with average intensity ν and a vacuum state with intensity 0 [120]. If the intensities of the signal and decoy states are the same as the preset μ and ν, the GLLP model with the decoy state protocol works well. This model can produce the tight lower bound of the secret key rate.

However, it is evident that the laser seeding attack changes the values of μ and ν, and therefore the key rate estimated by Alice and Bob also changes. For example, we assume that the average intensities μ and ν become 2μ and 2ν under the attack, while Alice and Bob are not aware of the change. They still believe the average intensities are μ and ν. Hence, the probability of single-photon pulses is still supposed to be P_1^μ. In contrast, the total detection gain Q and error rate E are actual parameters that Bob receives from the quantum channel, so that $Q_{2\mu}$ and $E_{2\mu}$ are applied to estimate the key rate by Alice and Bob.

It is notable that, in principle, Alice and Bob could monitor the detection gain. However, in practice, they do not. There are several reasons for this. First, monitoring the detection gain would require characterizing the channel beforehand, which is tricky because Eve could also tamper with such characterization. Second, the typical security proofs of QKD do not consider any priori knowledge about the behavior of the quantum channel. They rely on the knowledge of the quantum states sent by Alice, the knowledge of the quantum measurements performed by Bob, and the observed experimental data. Additionally, a practical QKD system just requires a minimum detection rate below which the

protocol aborts. Therefore, it is reasonable that Alice and Bob estimate the key rate with preset μ, ν and P_1^μ with measured $Q_{2\mu}$ and $E_{2\mu}$. Apparently, this estimation is not precise. In contrast, the correct key rate with doubled intensities should be estimated by the parameters 2μ, 2ν and $P_1^{2\mu}$ with the measured $Q_{2\mu}$ and $E_{2\mu}$.

I simulate both the wrong and correct key rates as shown in Fig. 6.6. In the simulation, we assume the efficient BB84 protocol [109] with $q = 1$. For the source, we assume that the preset intensities are $\mu = 0.6$ and $\nu = 0.2$. For the detection, we use the detection parameters from Gobby-Yuan-Shields (GYS) experiment [66]: the dark count rate $Y_0 = 1.7 \times 10^{-6}$, the transmission in Bob's apparatus $\eta_{\text{Bob}} = 4.5\%$, the misalignment error rate $e_{\text{detector}} = 3.3\%$ and the error correction efficiency $f(E_\mu) = 1.22$. For comparison, I first plot the key rate with intensities $\mu = 0.6$ and $\nu = 0.2$, which shows the key rate during normal operation without the attack as the red dashed curve shown in Fig. 6.6. The blue curve in Fig. 6.6 illustrates the key rate that Alice and Bob estimate during the attack. Since the detection gain under doubled intensities is larger than usual, Alice and Bob generate a higher key rate than that in the normal case. It hints the incorrect estimation. Moreover, the correct key rate under doubled intensities is showed as the green curve in Fig. 6.6, which is much lower than the key rate that Alice and Bob estimate. It shows that with the current security analysis, Alice and Bob would distill a key that is insecure to a significant degree. That is, they overestimate the key rate.

It is remarkable to consider that lower bounds might improve. One might develop a better security analysis that provides higher rates for a certain distance. One day, the correct lower bound might reach the blue curve. Therefore, we cannot completely discard the reality that an improved future security analysis could make the blue line secure. However, the secure key rate is limited by the upper bound, which indicates that the highest key rate can be achieved by any current and future security analysis. That is, it guarantees that the key rates beyond the upper bound are certainly insecure. Hence, to determine the key-rate limitation in our case, I also calculate the upper bound of the key rate in the BB84 protocol with doubled intensities by

$$R^u = Y_1^{2\mu} 2\mu e^{-2\mu}. \tag{6.2}$$

The simulation result is drawn as the pink curve in Fig. 6.6. The upper bound is lower than the overestimated lower bound up to a distance of about 140 km. It shows that the key rate overestimated by Alice and Bob is definitely insecure, at least when the communication distance is shorter than 140 km.

Further security analysis is in progress as my collaborators are currently analyzing the threat of increased intensities for an MDI QKD system.

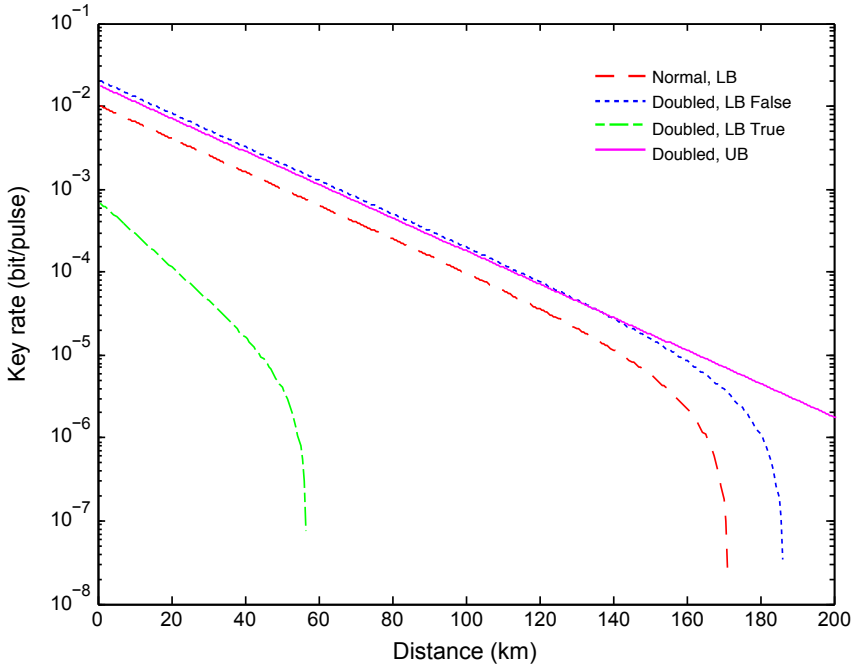

Figure 6.6: Key rates under the laser seeding attack with doubled intensities. LB: lower bound, UB: upper bound.

6.5 Conclusion

In this study, I show experimentally that Eve can tamper with the emitted power of a laser source by injecting light into the laser diode. This results in the increased intensities of signals transmitted to Bob. The security analysis based on the decoy BB84 QKD protocol illustrates that the increased intensities of the signal pulses can compromise the security of the QKD system.

Chapter 7

Laser damage attack on optical attenuators in QKD

7.1 Motivation

In practice, QKD systems widely employ weak coherent lasers as sources, with mean photon numbers attenuated to single-photon levels. This ensures that the most quantum states are sent using single-photon pulses, so an eavesdropper cannot split and measure the states independently. Additionally, the side effect of the minor multi-photon states is removed by applying decoy state protocol [112,120,175]. However, if an optical attenuation component can be altered and its attenuation decreased, either permanently or temporarily, it is possible again for an eavesdropper to obtain parts of the secret key. Please note that the optical attenuator is usually the last component in Alice's apparatus right before the states are sent to the quantum channel. For Eve, an attenuator is the first component her laser reaches when she injects laser power to Alice. In our experiment, we attempt to reduce the attenuation in three types of optical attenuators through optical damage. This experiment assesses the possibility of such an attack, which, if successful, would break the fundamental assumption in state preparation in QKD systems.

7.2 Optical power handling capacity of single-mode fibers

In our experiment, a high-power laser is the tool for Eve's attack. At the first step, how much optical power is allowed to be transmitted through a single-mode fiber needs to be clarified. Here, we set the restriction that Eve can only use a standard single-mode fiber, instead of a multi-mode fiber, in her attack, because this is the case for the quantum channel in the real scenario of QKD. The amount of optical power is limited by the inherent handling capability of the single-mode fiber and the safety requirement of the laser source.

The laser-induced damage threshold (LIDT) of the standard single-mode silica fiber has been discussed in Ref. 113. The theoretical LIDT relies on the softening point of silica [178], which is 5.5×10^4 W over 50 μm^2 fiber core area [113]. However, in reality, thermal damage likely happens at the fiber connection points or the interface between the fiber core and the cladding. The fiber fuse is triggered by high temperature at a fiber end facet. This fiber fuse can be realized by taking the fiber end against an absorptive material such as metal to accumulate heat, or also by using a flame (~ 2700 °C) to provide a high temperature. It has been experimentally demonstrated that 2–5 W continuous-wave (c.w.) laser can initiate a fiber fuse [87]. For our experiment, we tested a 20 m single-mode fiber ending without any termination. When no deliberate method is applied attempting to trigger a fiber fuse, the fiber can tolerate 9 W c.w. laser. Thus in practice, the fiber fuse threshold can deviate from the precise LIDT. A reasonable value of LIDT given in Ref. 113 is 12.8 W for c.w. laser.

The source of the high-power laser also needs to be protected from damage. The major threat comes from backward scatterings in the optical fiber, especially the backward stimulated Raman scattering (SRS) and the backward stimulated Brillouin scattering (SBS) [157]. Generally, during the light transmission, a fraction of incident light can be transferred from one optical field to another field with a frequency shift, due to the vibration of the transmission medium. The frequency-shifted light is called Stokes wave. The Stokes wave may rapidly increase over the transmission, which causes SRS and the SBS. The SRS and SBS can travel backward to the laser source such that the source may be destroyed. To keep the high-power laser source being safe, the thresholds of SRS and SBS are investigated for our case. The threshold is defined as the incident pump power P_{th}, at which point the backward Stokes power P_{s} becomes equal to the input power at the fiber output [157]

$$P_{\mathrm{th}} \exp(-\alpha L) = P_{\mathrm{s}}(L), \tag{7.1}$$

where α is the fiber loss (typically 0.2 dB/km at 1550 nm), and L is the transmission

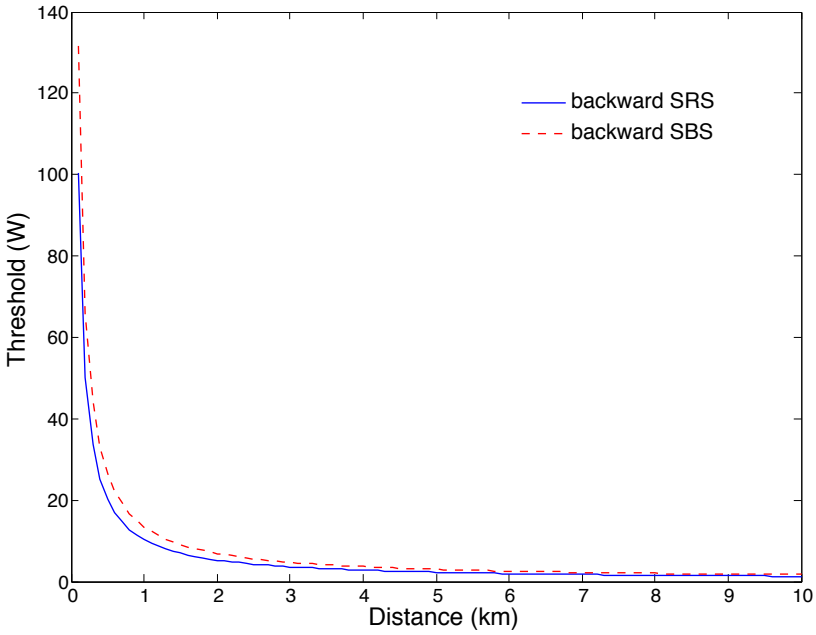

Figure 7.1: Simulated backward SRS and SBS thresholds.

distance. It indicates that the threshold is dependent on the fiber length.

For the backward SRS, threshold is given by [157]

$$P_{\text{th}}^{\text{SRS}} = \frac{20A_{\text{eff}}}{g_{\text{R}} \times L_{\text{eff}}}. \tag{7.2}$$

Here A_{eff} is the effective core area. For the standard single-mode fiber with the core diameter of 8 μm, $A_{\text{eff}} = 50\ \mu\text{m}^2$. g_{R} is the Raman-gain coefficient, which is 6.67×10^{-14} m/W at 1550 nm wavelength [32]. L_{eff} is the effective interaction length defined as

$$L_{\text{eff}} = [1 - \exp(-\alpha L)]/\alpha. \tag{7.3}$$

Thus, the threshold value is dependent on the transmission distance. The simulation result about $P_{\text{th}}^{\text{SRS}}$ versus transmission distance is given in Fig. 7.1. The result shows that the SRS threshold drops dramatically when the fibre length extends. However, the short-distance transmission until 1 km can handle more than 10 W optical power.

On the other hand, the backward SBS plays a key role in limiting the transmission power. SBS can occur at a much lower incident power level than that needed for SRS. The

input power threshold is quantified by [157]

$$P_{\text{th}}^{\text{SBS}} = \frac{21 A_{\text{eff}}}{g_{\text{B}} \times L_{\text{eff}}}, \tag{7.4}$$

where the Brillouin-gain coefficient $g_{\text{B}} = 5 \times 10^{-11}$ m/W [32]. This threshold allows only 0.2 W as maximum input power. Fortunately, the SBS threshold can increase considerably if the spectral width $\Delta\nu_p$ of pump laser is much larger than the full width at half maximum (FWHM) of the Brillouin-gain spectrum $\Delta\nu_B$ in the SBS. More specifically, the SBS threshold in Eq. (7.4) increases by a factor of $1 + \Delta\nu_p/\Delta\nu_B$. In our situation, we employ a laser diode with $\Delta\nu_p = 10$ GHz in our experiment. $\Delta\nu_B$ in the single-mode fiber at 1550 nm is 16 MHz. The SBS threshold in this case is then shown in Fig. 7.1. It is obvious that the SBS threshold follows the similar trend as the SRS threshold. Both thresholds are at the similar power levels at the same distance. For example, at the distance of 1 km, the SBS threshold is still 13.5 W.

Overall, a 10 W c.w. laser is allowed to safely transmit through 1 km single mode fiber. The shorter the distance, the larger handling capability of c.w. power. In our experiment, a laser amplifier provides up to 9 W c.w. power transmitting through a 20-m fiber. Its feasibility has been theoretically verified by the models above. The details about the experimental setup are elaborated in the next section.

7.3 Experimental setup

The high-power test of optical attenuators has been conducted using the setup shown in Fig. 7.2. The experimental scheme simulates a hacking scenario for a running QKD system. The test laser is a low power single-mode 1550nm fiber-pigtail laser diode (Gooch & Housego AA1406) acting as Alice's laser. This laser also provides the initial source of light for measuring the attenuation. Power meter A (Joinwit JW3208) monitors the power of the test laser. The input of an optical attenuator under test is connected to the test laser through a 50:50 beamsplitter (BS). The output of the optical attenuator under test is connected to a laser amplifier through a 99:1 BS. Power meter B (Thorlabs PM200 with S146C) is used as a monitor for the high power. Power meter C (Thorlabs PM200 with S154C) is applied to check the attenuation of the optical attenuator before and after optical damage. This scheme represents a real scenario in which Eve injects light to a source station in a QKD system via a quantum channel from the reversed direction of Alice's light.

The attenuation is determined by comparing the measurement from power meter A and power meter C, taking into account the additional 20 dB attenuation of the 99 : 1 BS. This measurement is first performed for each attenuator with the optical amplifier and the seed laser turned off, so the initial attenuation is calibrated before any attempted tampering. The same measurement is repeated after laser damage testing. After applying high-power laser, any attenuation change can then be attributed to optical damage within the setup. Please note that the beam splitter can tolerate high-power laser during our testing. Thus, it is reasonable to believe that all the attenuation change is due to the optical attenuator. In the case where the connection points of fibers are burnt, we treat it as an outcome of Eve's attack, which causes denial-of-service in an actual QKD system.

A custom 1550nm Erbium-Ytterbium high-power amplifier (EDFA) is employed in the experiment. The amplifier is designed using core pumping as the first stage and then double clad pumping as the second stage. The EDFA allows a high gain and uses an input seed power as low as 0.4 mW (−4.0 dBm) which is amplified to a maximum of about 9 W. The EDFA overcomes amplified spontaneous emission (ASE) [56] for high power double cladding erbium-ytterbium fiber amplifier at 1.0 mm. The presence of this ASE could lead to spurious lasing at 1.0 μm and limitation of the energy transfer process from the ytterbium ions to the erbium, which limits the output power level. As a result, the amplifier produces high output power up to $\sim$ 9 W (39.5 dBm) through standard single-mode fiber (SMF-28) with a high slope efficiency of 28%.

In our experiment, a single mode fiber-pigtailed laser (Qphotonics QFBGLD-1550-100) set to 20 mW continuous output power is used as a seed source for the EDFA. The exact power applied to the optical attenuator can be calculated from the measurement at power meter B, as shown in Fig. 7.2 multiplying by a factor of 99. In case that a fiber fuse occurs while applying high power during the experiment, a protection system is implemented. Two photodiode monitors are used to detect light from the fiber fuse around the fiber jacket. The monitors are placed at the output of the EDFA and the input of the attenuator. If a fiber fuse happens, then the monitor triggers a TTL voltage connected to the EDFA interlock, thus shutting off the laser and stopping the fiber fuse.

Three types of attenuators are tested. The first type is a variable optical attenuator (VOA) controlled by microelectromechanical systems (MEMS). The MEMS VOA is electronically adjustable from a maximum of 31–34 dB at a default control voltage of 0 V to a minimum of approximately 1 dB at 15 V control voltage. Physical disassembly shows two closely placed input and output fibers facing an electronically controlled reflector-lens assembly. The voltage setting controls the tilt of the reflector and changes the amount of optical power being reflected from the input fiber to the output.

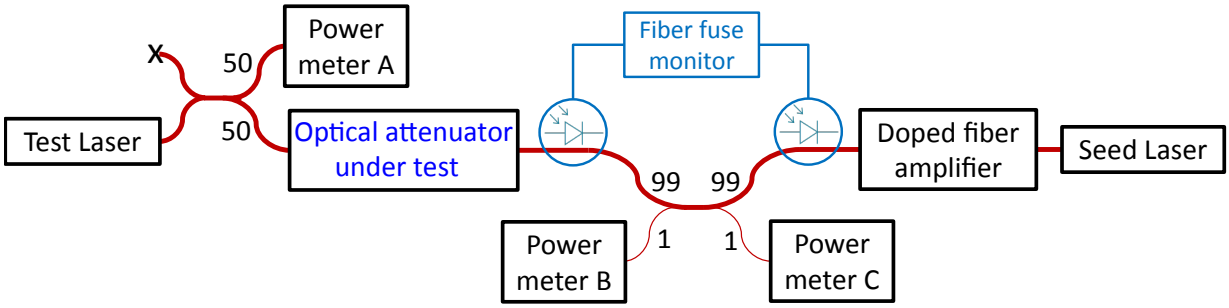

Figure 7.2: Experimental setup. The optical attenuator, as the testing target, is replacable. The test laser and the high-power laser are applied to the optical attenuator from different directions.

The second one is a fixed attenuator with about 25dB attenuation. Physical disassembly shows a dark material between the input and output fibers. The third attenuator is a variable attenuator adjustable between 1.5 dB to 80 dB. A screw tipped with a dark material placed between the input and output fibers, which can be inserted between the input and output. The adjustable position of the screw then changes the amount of light transmitted from the input to the output.

The testing procedure for each attenuator is as follows. The test laser is always on as a working QKD system. Eve shines the high-power laser starting from 300 mW and holds for at least 1 minute. After that, Eve turns off the high-power laser, and the attenuation is measured. If no attenuation change occurs, then the laser power is gradually increased, and we repeat the steps above. Until the measurable attenuation change happens, we will stop the testing. Since there might be a cooling down period in which the attenuation fluctuates, we record the measurement values during this period. After this period, we also record the reasonably constant attenuation.

7.4 Experimental results

7.4.1 Testing of individual attenuators

Over these three types of optical attenuators, the MEMS VOA shows a permanent change of attenuation after being shined upon by the high-power laser. The fixed attenuator shows a temporary decrease in attenuation, while the screw-tipped VOA appears to be minimally affected by the high-power laser.

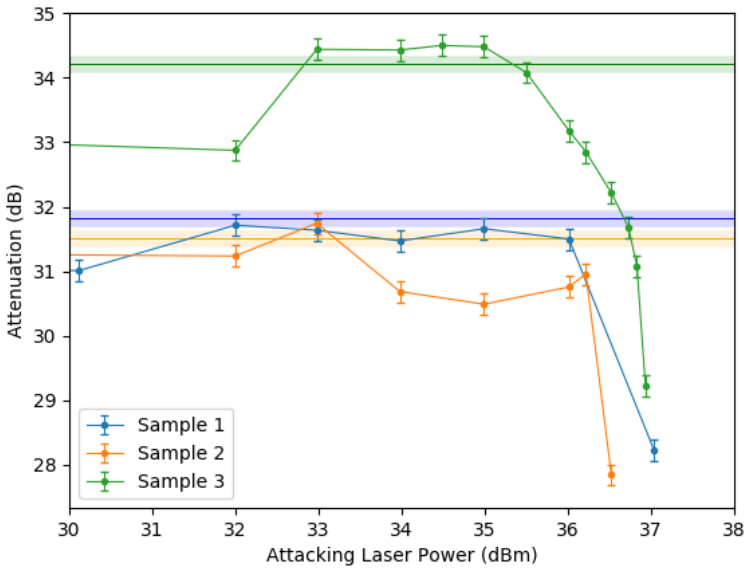

Figure 7.3: Selected samples of the MEMS VOA with a permanent decrease in attenuation after laser damage. The horizontal lines indicate the initial measured attenuation before experimentation.

From 8 tested samples of MEMS VOAs, we were successful in permanently decreasing the attenuation for three samples with an average decline of 3 dB after our testing as shown in Fig. 7.3. The permanent decrease in attenuation occurs when a laser power is about 4 W. If this power level is exceeded, then testing often results in catastrophic damage and has its attenuation permanently increased to a higher level of about 70 dB. The permanence of the damage in the MEMS VOA is confirmed by measuring the attenuation after a few hours and a few days. Moreover, as Fig. 7.4 shows, by measuring the attenuation across the entire adjustable range, the most permanent decrease in attenuation is 5 dB where the initial maximum attenuation is 34 dB when no voltage is applied. The difference in attenuation before and after the damage is gradually reduced for lower attenuations under higher voltages.

From Fig. 7.4, the noticeable decrease is in the range of the original attenuation from 35 dB to 20 dB, which could be the working range in a QKD system. It is then reasonable to assume that decreased attenuation can result in higher photon number in the emitted pulses than Alice expected. This gives Eve the chance to obtain additional photons from

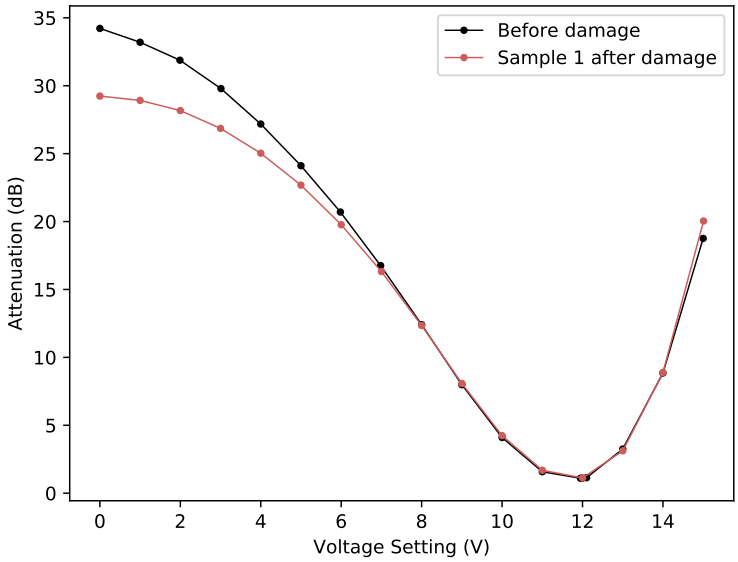

Figure 7.4: A sample of the MEMS VOA at various voltage settings, before and after laser damage.

Alice's pulses which also carry the secret key.

The fixed attenuator temporarily decreases its attenuation under 5 W hacking laser. The maximum decrease is about 2 dB, as shown in Fig. 7.5. This decrease in attenuation with oscillation then goes back to the initial state, which takes several minutes. Eve can exploit such a time slot to split photons from the output pulses. Thus, part of the secret information becomes accessible to Eve, but permanent damage also occurs at higher power.

As for the screw-tipped VOA, our testing shows almost no permanent change in the attenuation even at the highest available laser power of 9 W that is continuously applied for 20 minutes, although the attenuator reaches 234 $^\circ$ C at this power, as measured by an infrared camera. However, visual inspection of the optical blocking material at the adjustable screw tip shows a concave structure which fits the input fiber position where the laser power is delivered. This suggests that higher optical power is likely to further damage the screw, while a change in attenuation might be possible.

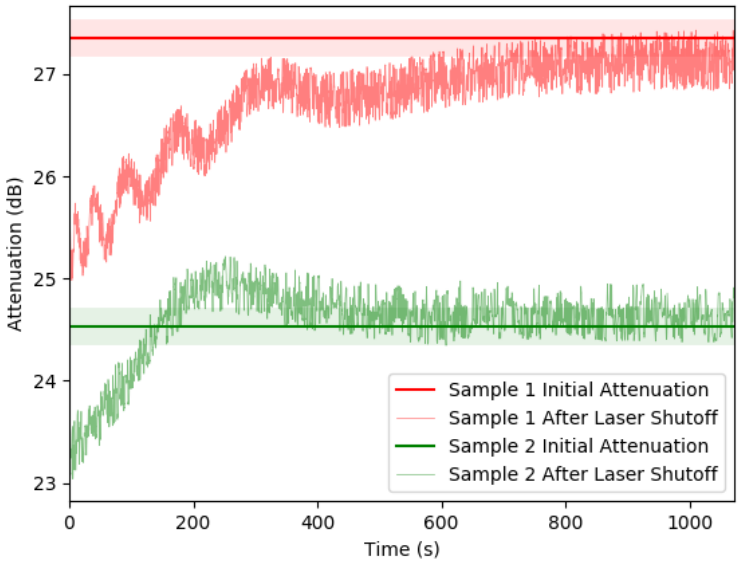

Figure 7.5: Samples of the fixed attenuator after being subjected to a damaging laser power where a temporary decrease in attenuation is observed.

Table 7.1: Results after optical damage for MEMS VOA samples. Testing voltage refers to the parameter used in step 2 of the experimental procedure. The Δ column refers to the change in attenuation observed at the testing voltage.

#	Type	Testing Voltage (V)	Att. Before (dB)	Att. After (dB)	Δ (dB)	Damage Threshold (W)
1	A	12.0	33.05	35.44	+2.39	4.5
2	A	12.0	33.88	32.95	−0.93	5.0
3	A	11.5	32.81	64.28	+31.47	5.6
4	B	14.0	38.79	32.32	−6.47	5.5
5	B	14.5	≈ 68	58.82	≈ -9.2	4.0
6	B	13.5	31.21	22.29	−8.92	2.8

7.4.2 Testing of attenuator assembly

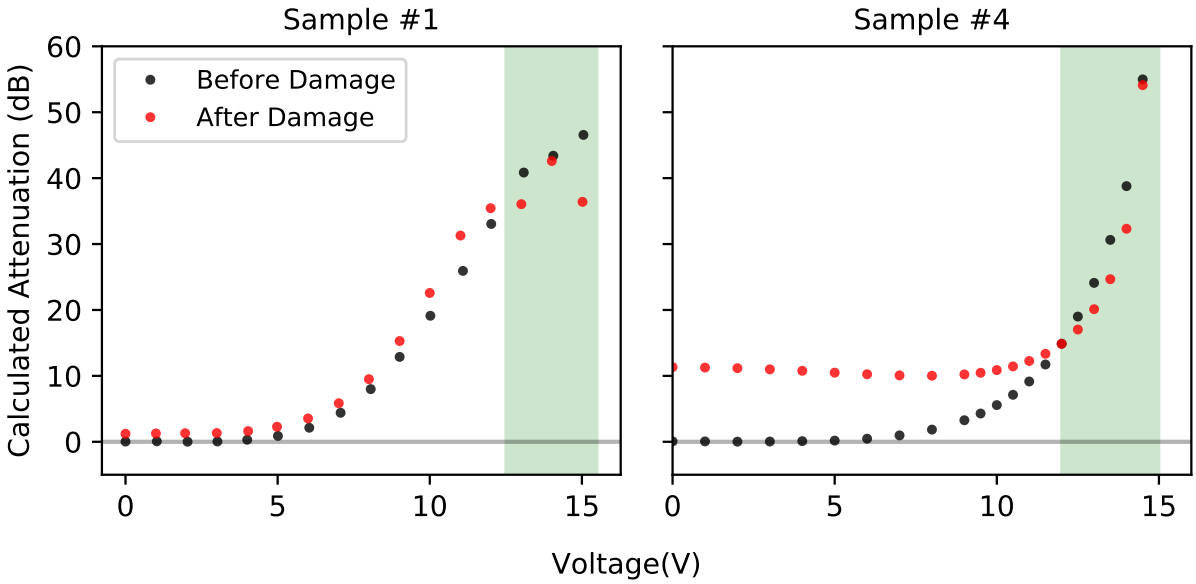

Figure 7.6: Typical VOA voltage-attenuation curves before and after optical damage. Permanent attenuation drop is observed within the green area.

From the testing above, we confirm that MEMS VOAs are especially vulnerable to the laser damage attack. The tested samples show a permanent decrease in attenuation after our testing. To investigate further the possible effects in a real QKD system, we subsequently test complete PCB-mounted attenuator assemblies for an actual QKD system provided by the third party. These MEMS VOAs are from two different manufacturers, labeled type A or B. Each pair of MEMS VOAs is attached using soft silicone glue to the PCB assembly, providing a thermally accurate representation of the actual QKD system.

We apply the same experimental method as before to test 6 attenuators assembled on 3 PCB boards. At the testing voltage, four out of the six attenuators show a permanent decrease of attenuation. One of the attenuators (3A) exhibited a catastrophic damage, where the attenuation after the optical damage is dramatically increased from its normal value. Another attenuator (1A in Table 7.1) did not show an attenuation drop at the testing voltage. However, it still has a range where the laser damage can decrease the attenuation as shown in Fig. 7.6(a). Therefore, for 5 out of 6 attenuators, there exists a range where the attenuation is decreased after laser damage. Typical results are illustrated in Fig. 7.6 with green areas showing decreased attenuation. Thus, we can reasonably believe that the

laser damage attack has been successful in these ranges of attenuation for the assembled samples.

Now we discuss two special samples in details. Attenuator #3 exhibited a near-total failure, where the attenuation after laser damage is dramatically increased from its normal value. The high attenuation almost blocks the incoming light, which would result in denial-of-service in a QKD system. The situation of attenuator #5 is unusual. Before our high-power testing, we followed the testing procedure to calibrate the attenuation-voltage curve first. However, the attenuation value becomes latched after 14.5 V which is still in the working range (0 - 15 V) according to its datasheet. Subsequent voltage adjustments down to even 0 V did not change this measured attenuation. Since the applied voltage is close to the maximum voltage, it is likely that the observed latching is due to the inherent variability between components. Despite this unexpected latching, a permanent decrease in attenuation still can be observed for this sample.

7.5 Possible MEMS VOA damage mechanisms

During the testing, we have observed temperature rise and physical change of the MEMS VOA under high optical power. To explain the details, we provide a brief schematic of the MEMS VOA in Fig. 7.7. Through observation by a thermal camera, we have found that the front end of the VOA casing with input and output fibers has a higher temperature, as shown in Fig. 7.8. The high-power laser was set to 2.8 W and turned on from $t = 0$ s to $t = 250$ s in Fig. 7.8. The outer casing of the VOA reached 120° C at this damage threshold power. The cap holding the input and output fibers start to extend outwards near the threshold of damage, which possibly makes the fibers inside the VOA out of alignment with the collimation lens. In a catastrophic damage scenario at 6.3 W optical power, the cap detaches itself from the attenuator casing with smoke emitted. Since the process of coupling a beam of light into a single-mode fiber is highly dependent on the relative positions of the involved optical elements [126], we hypothesize that the structural deformation under high temperatures is one of the possible major causes for the observed change in attenuation.

7.6 Conclusion

In this work, we test three types of attenuators by c.w. high-power laser. The testing power is up to 9 W through the single-mode fiber. The MEMS VOA shows a permanent decrease

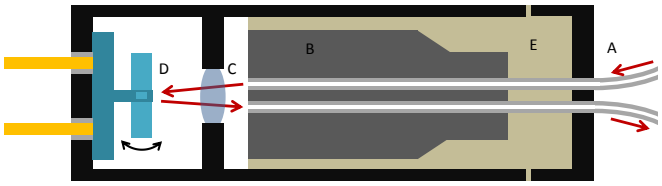

Figure 7.7: Simplified schematic of the MEMS variable optical attenuator with labeled parts (not to scale). (A) Incoming and outgoing single-mode fibers; (B) Glass sheath; (C) Collimation lens; (D) Voltage adjustable MEMS mirror on torsion mount; (E) Adhesive filler.

in attenuation. The fixed attenuator temporary declines its attenuation under 5 W laser. The screw-tipped VOA has minimal change during the testing. The decreased attenuation results in increased intensity of the transmitted states in a QKD system, which can be exploited by Eve. It helps Eve to discover the secret information in a QKD system. This type of attack could break the basic assumption about the mean photon number in a QKD protocol.

It has been shown that high-power c.w. laser can modify the characteristic of optical components in QKD systems. To protect a QKD system from laser damage, a watchdog monitor might not be enough, because the monitor could also be destroyed by the high-power laser [122]. We propose to add a special component, an optical fuse [171], at Alice's output. The optical fuse would only tolerate a certain amount of laser power but disconnects itself once the power crosses a threshold. In this way, the optical fuse physically blocks the injected high power, which effectively protects the system from the laser damage attack.

This study shows Eve's capability to change the characteristics of state preparation in a QKD system. The modification allows Eve to perform side-channel attacks on a source station, thus breaking the fundamental assumption in security proofs of QKD, even for those of MDI QKD protocols. The detailed analysis of the effect on BB84 QKD protocols has been presented in the previous Chapter 6. This would be a trigger to further investigate the implementation of nonleaky state preparation in MDI QKD systems.

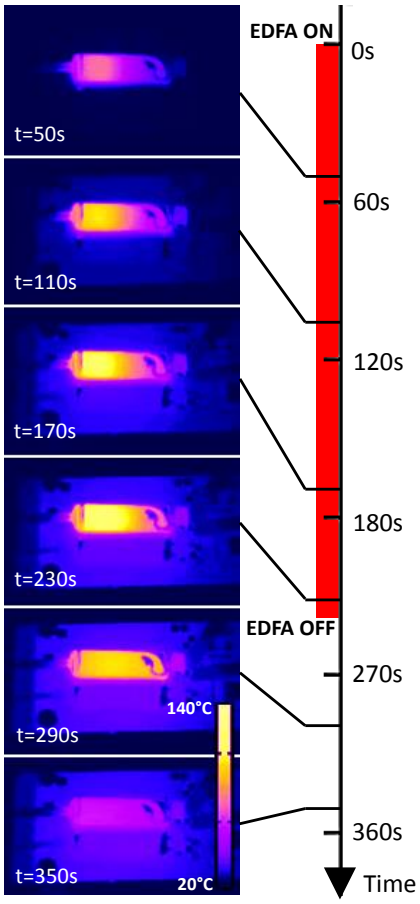

Figure 7.8: Temperature profile of the VOA from manufacturer B near the threshold of damage.

Chapter 8

Other projects

In this chapter, I will summarise other projects that I participated in. Please see reprints of research papers in the appendices for details.

8.1 Short pulse attack on continuous-variable quantum key distribution system

Homodyne detectors are widely used in a continuous-variable (CV) QKD system for measuring the transmitted photons. However, the imperfection of the homodyne detectors can be exploited by Eve to break the security of CV QKD systems [141]. In this work, we further investigate the vulnerability of a homodyne detector. By exploiting the finite bandwidth of electronic amplifier and limited response time of electronics in the homodyne detector, we propose and verify a new detector-side-channel attack, short pulse attack, in CV QKD. This attack is based on the fact that the efficiency of a homodyne detector is dependent on the input pulse temporal mode [94]. Thus, Bob's homodyne detector can respond nonlinearly when Eve manipulates input pulse widths. This behavior breaks the important linearity assumption in security proof of CV QKD. Hence, we have shown experimentally the nonlinear response of the homodyne detector. We have also simulated an intercept-resend attack with the nonlinear response of the homodyne detector to show its feasibility. The corresponding Qcrypt2017 abstract is included in Appendix A.

8.2 Effect of atmospheric turbulence on spatial-mode detector-efficiency mismatch

Previous research has shown that it is possible to hack a free-space QKD receiver by changing the spatial mode of the incoming beam to the receiver [146]. In this work, we considered an essential factor, atmospheric turbulence in the channel, into the attack, and then studied Eve's ability to perform spatial-mode detector-efficiency mismatch attack. The atmospheric turbulence is emulated in the lab by using a phase-only spatial light modulator. After that, we conducted the attack through the turbulence channel. The experimental results have shown that the stronger turbulence resulted in the weaker mismatch ratio between different detectors. Therefore, Eve introduced higher QBER under stronger turbulence during this attack. This result indicates that atmospheric turbulence disturbs Eve's attack, giving her a weaker ability for hacking. The study has thus shown that Eve can attack a free-space non-decoy state BB84 system only up to 250 m away from Bob with the turbulence at sea level. The corresponding Qcrypt2017 abstract is included in Appendix B.

Chapter 9

Conclusion and outlook

9.1 Conclusion

The work presented in this thesis was primarily motivated by the need to further investigate the practical security of quantum cryptography after the proposal of MDI QC. My Ph.D. research focuses on two main questions: 1) Since MDI QKD has some practical limitations, are there alternative methods to eliminate the insecurity of detection? 2) For a source that is assumed to be safe, can Eve exploit new loopholes to break its security even in MDI QKD? To address these two questions, I have studied the alternative countermeasures against imperfect detection and the loopholes in the source.

Alternative countermeasures against imperfect detection. Although MDI QKD removes all detector side-channel attacks, its deployment and application remain limited due to its incompatibility with current QKD systems and relatively low key rate. As a commercial alternative, ID Quantique proposed and implemented a random-detection-efficiency countermeasure as a patch in its existing Clavis2 system. Our experimental results showed that the countermeasure is not robust against the detector blinding attack. As another alternative, DDI QKD was proposed to relieve the implementation complexity of MDI QKD. However, we proved that DDI QKD is not as secure as is claimed. It is, in fact, insecure against detector side-channel attacks. Based on the imperfection of a beam splitter, we showed that DDI QKD is also insecure against side-channel attacks on Bob's linear optics network.

Loopholes in the source. As an essential assumption in MDI QKD, the source is required to be trusted. This assumption, however, cannot always be satisfied in practice.

We investigated the implementations of a decoy-state protocol widely used in a QKD system with a weak coherent source. The experimental results revealed that the intensity modulation, especially the pump-current modulation, causes a timing mismatch between signal and decoy states. A PNS attack was modeled to compromise the security of a QKD system. To act as an active Eve, we also experimentally demonstrated a laser seeding attack on the source. The results showed that Eve could increase the emission power of a laser diode. Furthermore, to create a stronger Eve, we demonstrated a laser damage attack on various attenuators. The experimental results showed that Eve can decrease the value of attenuation. The laser seeding attack and the laser damage attack both break the basic assumption about the mean photon number in QKD.

This work underscores the importance of further scrutinizing the practical security of quantum cryptography even in the age of MDI QC. Although the MDI concept overcomes all detector side channels, there is a clear need for additional practical security analysis. For example, other countermeasures against detector side channels for existing QKD systems must be verified. Moreover, new attack-resistant protocols must be thoroughly proved before security claims can be made. For MDI QKD itself, the practical security of the source should be deeply verified. This research revealed new vulnerabilities and loopholes in the source, which may become the new "Achilles' heel" in QKD.

In summary, this doctoral research focused on investigating the practical security of quantum cryptography, particularly as it relates to MDI QC. MDI QC represents significant progress in the practical security of QC and will continue to be the primary driver for practical analyses in the field going forward. This provided the primary motivation for the thesis work, where quantum hacking offered a promising avenue for evaluating and verifying the practical security of QKD systems. The more quantum hacking conducted, the more knowledge about the performance of practical QC gained. In this way, various devices were tested, and several new vulnerabilities were revealed. Thus, I believe that this thesis research will contribute to enhancing the security of quantum cryptography.

9.2 Outlook

Although the subject of this thesis is QC security, and in particular uncovering vulnerabilities to that security through quantum hacking, it is important to relieve a growing pessimism about the efficacy of QC to defend against future attacks in the quantum-computing era. Throughout this research, the practical security issues of QC were discussed with various stakeholders from industry, academia, and the general public. The fact that a technology proven to be unconditionally secure in theory can still be hacked

in practice understandably raises doubts around the meaningfulness and usefulness of QC. However, attempts to hack the QC system and expose vulnerabilities are not attempts to deny the significance and necessity of QC. QC takes unique advantage of quantum mechanics to ensure its information-theoretic security. Quantum hacking does not challenge the fundamental superiority of QC theory. Instead, quantum hacking focuses on imperfections in practical implementations of QC, which are general issues for all types of cryptographic systems. The purpose of researching quantum hacking is to provide more public information about the practical performance of QC systems, which is the first step to improve their practical security in the future.

On the other hand, while QC is indeed promising, it is not a "silver bulle" solution, nor is it universal. It cannot replace all other cryptography, nor should it. Currently, very widely used cryptographic systems still have advantages: they are compact, economical, mature, etc. Furthermore, although certain public-key cryptographic algorithms will be vulnerable to quantum computers, it has not been shown that a quantum computer can break all conventional cryptographic systems. Whether to use conventional or quantum cryptography can be decided on a case-by-case basis, and should depend on specific security requirements. For scenarios where super security is required, QC, of course, will be a good choice. For other scenarios where a certain amount of security risk is acceptable, conventional cryptography may provide a more convenient solution. Future cryptographic infrastructure designed to counter threats from quantum computers will benefit from both post-quantum cryptography and quantum cryptography.

Regarding QC, QKD is the most mature one and has already been commercialized. As QKD matures, standardization becomes a priority. The purpose of standardization is to ensure that QKD systems are appropriate for their intended use. Most importantly, standards increase the security level of QKD products. This can be done by testing the products and certifying their security performance. ETSI is working on proposing universally accepted QKD standards and certification [2]. These efforts promote the international standardization of QKD. Regarding certification, a QKD system should be tested against different types of attacks. The response of the tested object will show its security strength. To support this security certification, the more knowledge we gain about practical attacks, the better security verification we can provide. During my Ph.D., I participated in industry consulting for three QKD companies. These activities provide valuable experience and information to file the standard for QKD certification.

It is necessary to conduct practical security analysis for QC systems beyond QKD. Since they use similar optical components as QKD systems, there is a high chance that the potential threat from imperfect devices will occur in the wider area of QC as well. Therefore, rigorous investigation of practical security must be a high priority once the

implementations are available. This research provides a strong reference point for future security analysis of other QC implementations. However, considering practical imperfections in the theory of QC is another challenge. Thus, more efforts should be made in the future.

References

[1] Information Systems Audit and Control Association: ISACA Glossary of Terms, http://www.isaca.org/Knowledge-Center/Documents/Glossary/glossary.pdf, visited 1 November 2017.

[2] ETSI white paper no. 8: Quantum safe cryptography and security (2015), http://www.etsi.org/images/files/ETSIWhitePapers/QuantumSafeWhitepaper.pdf.

[3] Commercial QKD systems are available from at least three companies: ID Quantique (Switzerland), http://www.idquantique.com; QuantumCTek (China), http://www.quantum-info.com; Qasky (China), http://www.qasky.com/en/.

[4] ID Quantique user case: Geneva Government secures data transfer for elections, https://www.idquantique.com/wordpress/wp-content/uploads/user-case-gva-gov.pdf;.

[5] Quantis specification sheet, http://marketing.idquantique.com/acton/attachment/11868/f-0174/1/-/-/-/-/2016%2008%2024%20-%20QRNG%20paper.pdf; visited 19 January 2018.

[6] ID Quantique Cerberis datasheet, https://www.idquantique.com/wordpress/wp-content/uploads/Cerberis-Datasheet.pdf; visited 19 January 2018.

[7] Centauris specifiation sheet, http://marketing.idquantique.com/acton/attachment/11868/f-0100/1/-/-/-/-/Centauris-Datasheet.pdf; visited 19 January 2018.

[8] ID Quantique user case: private bank, https://marketing.idquantique.com/acton/attachment/11868/f-0210/1/-/-/-/-/Private%20Bank_%2010G%20DRC%20Use%20Case.pdf;.

[9] Technological Advantages, http://www.qasky.com/en/info.asp?base_id=1&second_id=1004; visited 19 January 2018.

[10] 20 MHz and 50 MHz Quantum key distribution terminal, http://www.qasky.com/en/display.asp?id=780; visited 19 January 2018.

[11] 1 GHz quantum key distribution terminal, http://www.qasky.com/en/display.asp?id=781; visited 19 January 2018.

[12] Application scheme for P2P network, http://www.qasky.com/en/display.asp?id=792; visited 19 January 2018.

[13] Application scheme for quantum network basic on optical switch, http://www.qasky.com/en/display.asp?id=791; visited 19 January 2018.

[14] Application scheme for full-time all-pass network, http://www.qasky.com/en/display.asp?id=790; visited 19 January 2018.

[15] QuantumCTek secure communication solutions, http://www.quantum-info.com/English/solution/overview/; visited 19 January 2018.

[16] QuantumCTek user cases, http://www.quantum-info.com/English/case/; visited 19 January 2018.

[17] Quantum secure communication "Beijing-Shanghai backbone" project, http://www.quantum-info.com/English/case/2017/0901/339.html; visited 19 January 2018.

[18] QuantumCTek profile, http://www.quantum-info.com/English/#cases; visited 19 January 2018.

[19] UK quantum technology hub for quantum communications technologies, https://www.quantumcommshub.net/; visited 19 January 2018.

[20] Main tasks of UK quantum communications hub, http://uknqt.epsrc.ac.uk/files/flyer-quantum-comms-hub/; visited 19 January 2018.

[21] Aims of UK quantum communications hub, https://www.quantumcommshub.net/about-us/; visited 19 January 2018.

[22] Scope of flagship on quantum technologies, http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/fetflag-03-2018.html; visited 19 January 2018.

[23] A 112 km QKD network based on trusted repeaters, http://www.sktelecom.com/en/press/press_detail.do?idx=1217; visited 19 January 2018.

[24] A 460 km QKD network based on trusted repeaters, https://www.fiercewireless.com/wireless/sk-telecom-develops-advanced-quantum-repeater; visited 19 January 2018.

[25] The Tokyo QKD network, http://www.uqcc.org/QKDnetwork/; visited 19 January 2018.

[26] BBC News: China launches quantum-enabled satellite Micius, http://www.bbc.com/news/world-asia-china-37091833; visited 19 January 2018.

[27] Funding announcement about the Canadian satellite from the Government of Canada, https://www.canada.ca/en/innovation-science-economic-development/news/2017/04/ministers_bains_andgarneaucelebrate809millionforthecanadianspace.html; visited 19 January 2018.

[28] Industry Specification Group for quantum key distribution belongs to Telecommunications Standards Institute, http://www.etsi.org/technologies-clusters/technologies/quantum-key-distribution; visited 19 January 2018.

[29] Clavis2 specification sheet, http://www.idquantique.com/images/stories/PDF/clavis2-quantum-key-distribution/clavis2-specs.pdf, visited 8 July 2016.

[30] M. Legre and G. Ribordy, "Apparatus and method for the detection of attacks taking control of the single photon detectors of a quantum cryptography apparatus by randomly changing their efficiency", international patent appl. WO 2012/046135 A2 (filed 2010-10-10, published 2012-04-12).

[31] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 98:230501, 2007.

[32] G. P. Agrawal. *Nonlinear fiber optics*. Academic press, 2007.

[33] G. P. Agrawal and N. K. Dutta. *Semiconductor Lasers*. Springer, 1993.

[34] C. H. Bennett, F. Bessette, L. Salvail, G. Brassard, and J. Smolin. Experimental quantum cryptography. *J. Cryptology*, 5:3–28, 1992.

[35] C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proc. International Conference on Computers, Systems, and Signal Processing (Bangalore, India)*, pages 175–179, 1984.

[36] C. H. Bennett and G. Brassard. Experimental quantum cryptography: the dawn of a new era for quantum cryptography: the experimental prototype is working. *ACM Sigact News*, 20:78–80, 1989.

[37] C. H. Bennett, G. Brassard, C. Crépeau, and M.-H. Skubiszewska. Practical quantum oblivious transfer. In *Advances in Cryptology: proceeding of Crypto'91*, pages 351–366, 1991.

[38] E. Biham, B. Huttner, and T. Mor. Quantum cryptographic network based on quantum memories. *Phys. Rev. A*, 54:2651–2658, 1996.

[39] E. Biham and T. Mor. Bounds on information and the security of quantum cryptography. *Phys. Rev. Lett.*, 79:4034, 1997.

[40] E. Biham and T. Mor. Security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 78:2256, 1997.

[41] M. Bloch, S. W. McLaughlin, J.-M. Merolla, and F. Patois. Frequency-coded quantum key distribution. *Opt. Lett.*, 32:301–303, 2007.

[42] J. Bogdanski, N. Rafiei, and M. Bourennane. Experimental quantum secret sharing using telecommunication fiber. *Phys. Rev. A*, 78:062307, 2008.

[43] G. Brassard, N. Lütkenhaus, T. Mor, and B. C. Sanders. Limitations on practical quantum cryptography. *Phys. Rev. Lett.*, 85(6):1330–1333, 2000.

[44] A. Broadbent, J. Fitzsimons, and E. Kashefi. Universal blind quantum computation. In *Proc. 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 517–526. IEEE Computer Society, Los Alamitos, CA, 2009.

[45] A. N. Bugge, S. Sauge, A. M. M. Ghazali, J. Skaar, L. Lydersen, and V. Makarov. Laser damage helps the eavesdropper in quantum cryptography. *Phys. Rev. Lett.*, 112:070503, Feb 2014.

[46] W. T. Buttler, R. J. Hughes, P. G. Kwiat, S. K. Lamoreaux, G. G. Luther, G. L. Morgan, J. E. Nordholt, C. G. Peterson, and C. M. Simmons. Practical free-space quantum key distribution over 1 km. *Phys. Rev. Lett.*, 81:3283, 1998.

[47] W.-F. Cao, Y.-Z. Zhen, Y.-L. Zheng, Z.-B. Chen, N.-L. Liu, K. Chen, and J.-W. Pan. Highly efficient quantum key distribution immune to all detector attacks.

[48] R. Chen, W. Bao, C. Zhou, H. Li, Y. Wang, and H. Bao. Biased decoy-state measurement-device-independent quantum cryptographic conferencing with finite resources. *Opt. Express*, 24:6594–6605, 2016.

[49] R. Cleve, D. Gottesman, and H.-K. Lo. How to share a quantum secret. *Phys. Rev. Lett.*, 83:648, 1999.

[50] S. Cova, M. Ghioni, A. Lotito, I. Rech, and F. Zappa. Evolution and prospects for single-photon avalanche diodes and quenching circuits. *J. Mod. Opt.*, 51(9):1267–1288, 2004.

[51] M. Curty, O. Gühne, M. Lewenstein, and N. Lütkenhaus. Detecting two-party quantum correlations in quantum-key-distribution protocols. *Phys. Rev. A*, 71:022306, 2005.

[52] M. Curty, M. Lewenstein, and N. Lütkenhaus. Entanglement as a precondition for secure quantum key distribution. *Phys. Rev. Lett.*, 92:217903, 2004.

[53] M. Curty, F. Xu, W. Cui, C. C. W. Lim, K. Tamaki, and H.-K. Lo. Finite-key analysis for measurement-device-independent quantum key distribution. *Nat. Commun.*, 5:3732, 2014.

[54] T. F. da Silva, D. Vitoreti, G. B. Xavier, G. C. do Amaral, G. P. Temporão, and J. P. von der Weid. Proof-of-principle demonstration of measurement-device-independent quantum key distribution using polarization qubits. *Phys. Rev. A*, 88:052303, 2013.

[55] F.-G. Deng, G.-L. Long, and X.-S. Liu. Two-step quantum direct communication protocol using the einstein-podolsky-rosen pair block. *Phys. Rev. A*, 68:042317, 2003.

[56] W. Dickson, L. Liebscher, et al. Amplified spontaneous emission of surface plasmon polaritons and limitations on the increase of their propagation length. *Opt. Lett.*, 35:1197–1199, 2010.

[57] A. R. Dixon, Z. L. Yuan, J. F. Dynes, A. W. Sharpe, and A. J. Shields. Gigahertz decoy quantum key distribution with 1 Mbit/s secure key rate. *Opt. Express*, 16(23):18790–18979, 2008.

[58] A. K. Ekert. Quantum Cryptography Based on Bell's Theorem. *Phys. Rev. Lett.*, 67(6):661–663, 1991.

[59] Y. Fu, H.-L. Yin, T.-Y. Chen, and Z.-B. Chen. Long-distance measurement-device-independent multiparty quantum communication. *Phys. Rev. Lett.*, 114:090501, 2015.

[60] C. A. Fuchs, N. Gisin, R. B. Griffiths, C. S. Niu, and A. Peres. Optimal eavesdropping in quantum cryptography. 1. Information bound and optimal strategy. *Phys. Rev. A*, 56(2):1163–1172, 1997.

[61] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov. Full-field implementation of a perfect eavesdropper on a quantum cryptography system. *Nat. Commun.*, 2:349, 2011.

[62] E. Gibney. Billion-euro quantum project takes shape, 2017.

[63] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy. Trojan-horse attacks on quantum-key-distribution systems. *Phys. Rev. A*, 73(2):022320, 2006.

[64] N. Gisin, S. Pironio, and N. Sangouard. Proposal for implementing device-independent quantum key distribution based on a heralded qubit amplifier. *Phys. Rev. Lett.*, 105(7):070501, 2010.

[65] M. Giustina, M. A. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán, et al. Significant-loophole-free test of bell's theorem with entangled photons. *Phys. Rev. Lett.*, 115:250401, 2015.

[66] C. Gobby, Z. L. Yuan, and A. J. Shields. Quantum key distribution over 122 km of standard telecom fiber. *Appl. Phys. Lett.*, 84(19):3762–3764, 2004.

[67] P. González, L. Rebón, T. Ferreira da Silva, M. Figueroa, C. Saavedra, M. Curty, G. Lima, G. B. Xavier, and W. A. T. Nogueira. Quantum key distribution with untrusted detectors. *Phys. Rev. A*, 92:022337, 2015.

[68] K. J. Gordon, V. Fernandez, P. D. Townsend, and G. S. Buller. A short wavelength gigahertz clocked fiber-optic quantum key distribution system. *IEEE J. Quantum Electron.*, 40(7):900–908, 2004.

[69] D. Gottesman and I. Chuang. Quantum digital signatures.

[70] D. Gottesman, H.-K. Lo, N. Lütkenhaus, and J. Preskill. Security of quantum key distribution with imperfect devices. *Quantum Inf. Comput.*, 4(5):325–360, 2004.

[71] F. Grosshans, G. Van Assche, J. Wenger, R. Brouri, N. J. Cerf, and P. Grangier. Quantum key distribution using gaussian-modulated coherent states. *Nature*, 421:238–241, 2003.

[72] B. Hensen, H. Bernien, A. E. Dreau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526:682–686, 2015.

[73] A. Huang, S. Sajeed, P. Chaiwongkhot, M. Soucarros, M. Legré, and V. Makarov. Testing random-detector-efficiency countermeasure in a commercial system reveals a breakable unrealistic assumption. *IEEE J. Quantum Electron.*, 52(11):8000211, 2016.

[74] R. J. Hughes, J. E. Nordholt, D. Derkacs, and C. G. Peterson. Practical free-space quantum key distribution over 10 km in daylight and at night. *New J. Phys.*, 4:43, 2002.

[75] B. Huttner, N. Imoto, N. Gisin, and T. Mor. Quantum cryptography with coherent states. *Phys. Rev. A*, 51:1863, 1995.

[76] W.-Y. Hwang. Quantum key distribution with high loss: Toward global secure communication. *Phys. Rev. Lett.*, 91(5):057901, 2003.

[77] H. Inamori. Security of practical time-reversed epr quantum key distribution. *Algorithmica*, 34:340–365, 2002.

[78] K. Inoue, E. Waks, and Y. Yamamoto. Differential phase shift quantum key distribution. *Phys. Rev. Lett.*, 89(3):037902, 2002.

[79] N. Jain, E. Anisimova, I. Khan, V. Makarov, C. Marquardt, and G. Leuchs. Trojan-horse attacks threaten the security of practical quantum cryptography. *New J. Phys.*, 16:123030, 2014.

[80] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs. Device calibration impacts security of quantum key distribution. *Phys. Rev. Lett.*, 107:110501, 2011.

[81] T. Jennewein, U. Achleitner, G. Weihs, H. Weinfurter, and A. Zeilinger. A fast and compact quantum random number generator. *Rev. Sci. Instrum.*, 71:1675–1680, 2000.

[82] M.-S. Jiang, S.-H. Sun, C.-Y. Li, and L.-M. Liang. Wavelength-selected photon-number-splitting attack against plug-and-play quantum key distribution systems with decoy states. *Phys. Rev. A*, 86:032310, 2012.

[83] J. Jogenfors, A. M. Elhassan, J. Ahrens, M. Bourennane, and J.-Å. Larsson. Hacking the Bell test using classical light in energy-time entanglementbased quantum key distribution. *Sci. Adv.*, 1:e1500793, 2015.

[84] P. Jouguet, S. Kunz-Jacques, and E. Diamanti. Preventing calibration attacks on the local oscillator in continuous-variable quantum key distribution. *Phys. Rev. A*, 87:062313, 2013.

[85] P. Jouguet, S. Kunz-Jacques, and E. Diamanti. Preventing calibration attacks on the local oscillator in continuous-variable quantum key distribution. *Phys. Rev. A*, 87:062313, 2013.

[86] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti. Experimental demonstration of long-distance continuous-variable quantum key distribution. *Nat. Photonics*, 7:378, 2013.

[87] R. Kashyap and K. Blow. Observation of catastrophic self-propelled self-focusing in optical fibres. *Electron. Lett.*, 24:47–49, 1988.

[88] T. Kawanishi, K. Kogo, S. Oikawa, and M. Izutsu. Direct measurement of chirp parameters of high-speed mach–zehnder-type optical modulators. *Opt. Commun.*, 195:399–404, 2001.

[89] A. Kerckhoffs. La cryptographie militaire. *J. des Sciences Militaires*, IX:5–38, January 1883.

[90] Y.-H. Kim. Single-photon two-qubit entangled states: Preparation and measurement. *Phys. Rev. A*, 67:040301, 2003.

[91] T. L. Koch and J. E. Bowers. Nature of wavelength chirping in directly modulated semiconductor lasers. *Electron. Lett.*, 20:1038–1040, 1984.

[92] B. Korzh, N. Walenta, T. Lunghi, N. Gisin, and H. Zbinden. Free-running ingaas single photon detector with 1 dark count per second at 10% efficiency. *Appl. Phys. Lett.*, 104:081108, 2014.

[93] F. Koyama and K. Iga. Frequency chirping in external modulators. *J. Lightwave Technol.*, 6:87–93, 1988.

[94] R. Kumar, E. Barrios, A. MacRae, E. Cairns, E. Huntington, and A. Lvovsky. Versatile wideband balanced detector for quantum optical homodyne tomography. *Opt. Commun.*, 285:5259–5267, 2012.

[95] A. Lamas-Linares and C. Kurtsiefer. Breaking a quantum key distribution system through a timing side channel. *Opt. Express*, 15:9388–9393, 2007.

[96] H.-W. Li, S. Wang, J.-Z. Huang, W. Chen, Z.-Q. Yin, F.-Y. Li, Z. Zhou, D. Liu, Y. Zhang, G.-C. Guo, W.-S. Bao, and Z.-F. Han. Attacking a practical quantum-key-distribution system with wavelength-dependent beam-splitter and multiwavelength sources. *Phys. Rev. A*, 84:062308, 2011.

[97] W.-Y. Liang, M. Li, Z.-Q. Yin, W. Chen, S. Wang, X.-B. An, G.-C. Guo, and Z.-F. Han. Simple implementation of quantum key distribution based on single-photon bell-state measurement. *Phys. Rev. A*, 92:012319, 2015.

[98] S.-K. Liao, W.-Q. Cai, J. Handsteiner, B. Liu, J. Yin, L. Zhang, D. Rauch, M. Fink, J.-G. Ren, W.-Y. Liu, et al. Satellite-relayed intercontinental quantum network.

[99] S.-K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, Z.-P. Li, F.-Z. Li, X.-W. Chen, L.-H. Sun, J.-J. Jia, J.-C. Wu, X.-J. Jiang, J.-F. Wang, Y.-M. Huang, Q. Wang, Y.-L. Zhou, L. Deng, T. Xi, L. Ma, T. Hu, Q. Zhang, Y.-A. Chen, N.-L. Liu, X.-B. Wang, Z.-C. Zhu, C.-Y. Lu, R. Shu, C.-Z. Peng, J.-Y. Wang, and J.-W. Pan. Satellite-to-ground quantum key distribution. *Nature*, 549:43, 2017.

[100] S.-K. Liao, H.-L. Yong, C. Liu, G.-L. Shentu, D.-D. Li, J. Lin, H. Dai, S.-Q. Zhao, B. Li, J.-Y. Guan, W. Chen, Y.-H. Gong, Y. Li, Z.-H. Lin, G.-S. Pan, J. S. Pelc, M. M. Fejer, W.-Z. Zhang, W.-Y. Liu, J. Yin, J.-G. Ren, X.-B. Wang, Q. Zhang, C.-Z. Peng, and J.-W. Pan. Long-distance free-space quantum key distribution in daylight towards inter-satellite communication. *Nat. Photonics*, 11:509–513, 2017.

[101] C. C. W. Lim, M. Curty, N. Walenta, F. Xu, and H. Zbinden. Concise security bounds for practical decoy-state quantum key distribution. *Phys. Rev. A*, 89:022307, 2014.

[102] C. C. W. Lim, B. Korzh, A. Martin, F. Bussières, R. Thew, and H. Zbinden. Detector-device-independent quantum key distribution. *Appl. Phys. Lett.*, 105:221112, 2014.

[103] C. C. W. Lim, N. Walenta, M. Legré, N. Gisin, and H. Zbinden. Random variation of detector efficiency: A countermeasure against detector blinding attacks for quantum key distribution. *IEEE J. Sel. Top. Quantum Electron.*, 21:6601305, 2015.

[104] R. Linke. Modulation induced transient chirping in single frequency lasers. *IEEE J. Quantum Electron.*, 21:593–597, 1985.

[105] W.-T. Liu, S.-H. Sun, L.-M. Liang, and J.-M. Yuan. Proof-of-principle experiment of a modified photon-number-splitting attack against quantum key distribution. *Phys. Rev. A*, 83:042326, 2011.

[106] Y. Liu, T.-Y. Chen, J. Wang, W.-Q. Cai, X. Wan, L.-K. Chen, J.-H. Wang, S.-B. Liu, H. Liang, L. Yang, C.-Z. Peng, K. Chen, Z.-B. Chen, and J.-W. Pan. Decoy-state quantum key distribution with polarized photons over 200 km. *Opt. Express*, 18(8):8587–8594, 2010.

[107] Y. Liu, T.-Y. Chen, L.-J. Wang, H. Liang, G.-L. Shentu, J. Wang, K. Cui, H.-L. Yin, N.-L. Liu, L. Li, X. Ma, J. S. Pelc, M. M. Fejer, C.-Z. Peng, Q. Zhang, and J.-W. Pan. Experimental measurement-device-independent quantum key distribution. *Phys. Rev. Lett.*, 111:130502, 2013.

[108] H.-K. Lo and H. F. Chau. Unconditional security of quantum key distribution over arbitrarily long distances. *Science*, 283(5410):2050–2056, 1999.

[109] H.-K. Lo, H. F. Chau, and M. Ardehali. Efficient quantum key distribution scheme and a proof of its unconditional security. *J. Cryptology*, 18:133–165, 2005.

[110] H.-K. Lo, M. Curty, and B. Qi. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.*, 108:130503, 2012.

[111] H.-K. Lo, M. Curty, and K. Tamaki. Secure quantum key distribution. *Nat. Photonics*, 8:595–604, 2014.

[112] H.-K. Lo, X. Ma, and K. Chen. Decoy state quantum key distribution. *Phys. Rev. Lett.*, 94(23):230504, 2005.

[113] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. Yuan, and A. J. Shields. Practical security bounds against the trojan-horse attack in quantum key distribution. *Phys. Rev. X*, 5:031030, 2015.

[114] N. Lütkenhaus and M. Jahma. Quantum key distribution with realistic states: photon-number statistics in the photon-number splitting attack. *New J. Phys.*, 4(5):44, 2002.

[115] L. Lydersen, M. K. Akhlaghi, A. H. Majedi, J. Skaar, and V. Makarov. Controlling a superconducting nanowire single-photon detector using tailored bright illumination. *New J. Phys.*, 13:113042, 2011.

[116] L. Lydersen, N. Jain, C. Wittmann, Ø. Marøy, J. Skaar, C. Marquardt, V. Makarov, and G. Leuchs. Superlinear threshold detectors in quantum cryptography. *Phys. Rev. A*, 84:032320, 2011.

[117] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov. Hacking commercial quantum cryptography systems by tailored bright illumination. *Nat. Photonics*, 4:686–689, 2010.

[118] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov. Thermal blinding of gated detectors in quantum cryptography. *Opt. Express*, 18:27938–27954, 2010.

[119] X. Ma, C.-H. F. Fung, and H.-K. Lo. Quantum key distribution with entangled photon sources. *Phys. Rev. A*, 76:012307, 2007.

[120] X. Ma, B. Qi, Y. Zhao, and H.-K. Lo. Practical decoy state for quantum key distribution. *Phys. Rev. A*, 72:012326, 2005.

[121] V. Makarov, A. Anisimov, and J. Skaar. Effects of detector efficiency mismatch on security of quantum cryptosystems. *Phys. Rev. A*, 74(2):022313, 2006. erratum ibid. **78**, 019905 (2008).

[122] V. Makarov, J.-P. Bourgoin, P. Chaiwongkhot, M. Gagné, T. Jennewein, S. Kaiser, R. Kashyap, M. Legré, C. Minshull, and S. Sajeed. Creation of backdoors in quantum communications via laser damage. *Phys. Rev. A*, 94:030302, 2016.

[123] V. Makarov and D. R. Hjelme. Faked states attack on quantum cryptosystems. *J. Mod. Opt.*, 52:691–705, 2005.

[124] Ø. Marøy, L. Lydersen, and J. Skaar. Security of quantum key distribution with arbitrary individual imperfections. *Phys. Rev. A*, 82:032337, 2010.

[125] Ø. Marøy, V. Makarov, and J. Skaar. Secure detection in quantum key distribution by real-time calibration of receiver. *Quantum Sci. Technol.*, 2:044013, 2017.

[126] J. Martin. *Coupling Efficiency and Alignment Sensitivity of Single Mode Optical Fibers*. University of Central Florida, 1979.

[127] D. Mayers and A. Yao. Quantum cryptography with imperfect apparatus. In *Proc. 39th Annual Symposium on Foundations of Computer Science*, pages 503–509. IEEE, 1998.

[128] T. Moroder, M. Curty, and N. Lütkenhaus. Detector decoy quantum key distribution. *New J. Phys.*, 11:045008, 2009.

[129] M. Mosca. Cybersecurity in an era with quantum computers: will we be ready? *IACR Cryptology ePrint Archive*, 2015:1075, 2015.

[130] J. Mulholland, M. Mosca, and J. Braun. The day the cryptography dies. *IEEE Security Privacy*, 15:14–21, 2017.

[131] A. Muller, T. Herzog, B. Huttner, W. Tittel, H. Zbinden, and N. Gisin. "Plug and play" systems for quantum cryptography. *Appl. Phys. Lett.*, 70(7):793–795, 1997.

[132] S. Nauerth, M. Fürst, T. Schmitt-Manderbach, H. Weier, and H. Weinfurter. Information leakage via side channels in freespace BB84 quantum cryptography. *New J. Phys.*, 11(6):065001, 2009.

[133] T. Niemi, S. Tammela, and H. Ludvigsen. Device for frequency chirp measurements of optical transmitters in real time. *Rev. Sci. Instrum.*, 73:1103–1107, 2002.

[134] M. Peev, C. Pacher, R. Alléaume, C. Barreiro, J. Bouda, W. Boxleitner, T. Debuisschert, E. Diamanti, M. Dianati, J. F. Dynes, S. Fasel, S. Fossier, M. Fürst, J.-D. Gautier, O. Gay, N. Gisin, P. Grangier, A. Happe, Y. Hasani, M. Hentschel, H. Hübel, G. Humer, T. Länger, M. Legré, R. Lieger, J. Lodewyck, T. Lorünser, N. Lütkenhaus, A. Marhold, T. Matyus, O. Maurhart, L. Monat, S. Nauerth, J.-B. Page, A. Poppe, E. Querasser, G. Ribordy, S. Robyr, L. Salvail, A. W. Sharpe, A. J. Shields, D. Stucki, M. Suda, C. Tamas, T. Themel, R. T. Thew, Y. Thoma, A. Treiber, P. Trinkler, R. Tualle-Brouri, F. Vannel, N. Walenta, H. Weier, H. Weinfurter, I. Wimberger, Z. L. Yuan, H. Zbinden, and A. Zeilinger. The SECOQC quantum key distribution network in Vienna. *New J. Phys.*, 11(7):075001, 2009.

[135] C.-Z. Peng, J. Zhang, D. Yang, W.-B. Gao, H.-X. Ma, H. Yin, H.-P. Zeng, T. Yang, X.-B. Wang, , and J.-W. Pan. Experimental long-distance decoy-state quantum key distribution based on polarization encoding. *Phys. Rev. Lett.*, 98:010505, 2007.

[136] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen, and U. L. Andersen. High-rate measurement-device-independent quantum cryptography. *Nat. Photonics*, 9:397–402, 2015.

[137] 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. Practical quantum key distribution with polarization entangled photons. *Opt. Express*, 12:3865, 2004.

[138] B. Qi. Trustworthiness of detectors in quantum key distribution with untrusted detectors. *Phys. Rev. A*, 91:020303, 2015.

[139] B. Qi, C.-H. F. Fung, H.-K. Lo, and X. Ma. Time-shift attack in practical quantum cryptosystems. *Quantum Inf. Comput.*, 7(1-2):73–82, 2007.

[140] B. Qi, L.-L. Huang, L. Qian, and H.-K. Lo. Experimental study on the gaussian-modulated coherent-state quantum key distribution over standard telecommunication fibers. *Phys. Rev. A*, 76:052323, 2007.

[141] H. Qin, R. Kumar, and R. Alléaume. Quantum hacking: Saturation attack on practical continuous-variable quantum key distribution. *Phys. Rev. A*, 94:012325, 2016.

[142] J.-G. Ren, P. Xu, H.-L. Yong, L. Zhang, S.-K. Liao, J. Yin, W.-Y. Liu, W.-Q. Cai, M. Yang, L. Li, et al. Ground-to-satellite quantum teleportation. *Nature*, 549:70, 2017.

[143] M. F. Riedel, D. Binosi, R. Thew, and T. Calarco. The european quantum technologies flagship programme. *Quantum Sci. Technol.*, 2:030501, 2017.

[144] D. Rosenberg, J. W. Harrington, P. R. Rice, P. A. Hiskett, C. G. Peterson, R. J. Hughes, A. E. Lita, S. W. Nam, and J. E. Nordholt. Long-distance decoy-state quantum key distribution in optical fiber. *Phys. Rev. Lett.*, 98(1):010503, 2007.

[145] A. Rubenok, J. A. Slater, P. Chan, I. Lucio-Martinez, and W. Tittel. Real-world two-photon interference and proof-of-principle quantum key distribution immune to detector attacks. *Phys. Rev. Lett.*, 111:130501, 2013.

[146] S. Sajeed, P. Chaiwongkhot, J.-P. Bourgoin, T. Jennewein, N. Lütkenhaus, and V. Makarov. Security loophole in free-space quantum key distribution due to spatial-mode detector-efficiency mismatch. *Phys. Rev. A*, 91:062301, 2015.

[147] S. Sajeed, A. Huang, S. Sun, F. Xu, V. Makarov, and M. Curty. Insecurity of detector-device-independent quantum key distribution. *Phys. Rev. Lett.*, 117:250505, 2016.

[148] S. Sajeed, I. Radchenko, S. Kaiser, J.-P. Bourgoin, A. Pappa, L. Monat, M. Legré, and V. Makarov. Attacks exploiting deviation of mean photon number in quantum key distribution and coin tossing. *Phys. Rev. A*, 91:032326, 2015.

[149] M. Sasaki, M. Fujiwara, H. Ishizuka, W. Klaus, K. Wakui, M. Takeoka, S. Miki, T. Yamashita, Z. Wang, A. Tanaka, K. Yoshino, Y. Nambu, S. Takahashi, A. Tajima, A. Tomita, T. Domeki, T. Hasegawa, Y. Sakai, H. Kobayashi, T. Asai, K. Shimizu, T. Tokura, T. Tsurumaru, M. Matsui, T. Honjo, K. Tamaki, H. Takesue, Y. Tokura, J. F. Dynes, A. R. Dixon, A. W. Sharpe, Z. L. Yuan, A. J. Shields, S. Uchikoga, M. Legré, S. Robyr, P. Trinkler, L. Monat, J.-B. Page, G. Ribordy, A. Poppe, A. Allacher, O. Maurhart, T. Länger, M. Peev, and A. Zeilinger. Field test of quantum key distribution in the Tokyo QKD Network. *Opt. Express*, 19(11):10387–10409, 2011.

[150] S. Sauge, L. Lydersen, A. Anisimov, J. Skaar, and V. Makarov. Controlling an actively-quenched single photon detector with bright light. *Opt. Express*, 19:23590–23600, 2011.

[151] R. Saunders, J. King, and I. Hardcastle. Wideband chirp measurement technique for high bit rate sources. *Electron. Lett.*, 30:1336–1338, 1994.

[152] 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. Experimental demonstration of free-space decoy-state quantum key distribution over 144 km. *Phys. Rev. Lett.*, 98(1):010504, 2007.

[153] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, et al. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, 115:250402, 2015.

[154] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26:1484–1509, 1997.

[155] P. W. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85(2):441–444, 2000.

[156] S. Singh. *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography*. Fourth Estate, London, 1999.

[157] R. G. Smith. Optical power handling capacity of low loss optical fibers as determined by stimulated raman and brillouin scattering. *Appl. Opt.*, 11:2489–2494, 1972.

[158] D. Stucki, C. Barreiro, S. Fasel, J.-D. Gautier, O. Gay, N. Gisin, R. Thew, Y. Thoma, P. Trinkler, F. Vannel, and H. Zbinden. Continuous high speed coherent one-way quantum key distribution. *Opt. Express*, 17(16):13326–13334, 2009.

[159] D. Stucki, N. Gisin, O. Guinnard, G. Ribordy, and H. Zbinden. Quantum key distribution over 67 km with a plug&play system. *New J. Phys.*, 4:41, 2002.

[160] D. Stucki, N. Walenta, F. Vannel, R. T. Thew, N. Gisin, H. Zbinden, S. Gray, C. R. Towery, and S. Ten. High rate, long-distance quantum key distribution over 250 km of ultra low loss fibres. *New J. Phys.*, 11(7):075003, 2009.

[161] S.-H. Sun, M.-S. Jiang, and L.-M. Liang. Passive Faraday-mirror attack in a practical two-way quantum-key-distribution system. *Phys. Rev. A*, 83(6):062331, 2011.

[162] S.-H. Sun, F. Xu, M.-S. Jiang, X.-C. Ma, H.-K. Lo, and L.-M. Liang. Effect of source tampering in the security of quantum cryptography. *Phys. Rev. A*, 92:022304, 2015.

[163] H. Takesue, S. W. Nam, Q. Zhang, R. H. Hadfield, T. Honjo, K. Tamaki, and Y. Yamamoto. Quantum key distribution over a 40-dB channel loss using superconducting single-photon detectors. *Nat. Photonics*, 1(6):343–348, 2007.

[164] K. Tamaki, M. Curty, G. Kato, H.-K. Lo, and K. Azuma. Loss-tolerant quantum cryptography with imperfect sources. *Phys. Rev. A*, 90:052314, 2014.

[165] K. Tamaki, M. Curty, and M. Lucamarini. Decoy-state quantum key distribution with a leaky source. *New J. Phys.*, 18:065008, 2016.

[166] Y.-L. Tang, H.-L. Yin, S.-J. Chen, Y. Liu, W.-J. Zhang, X. Jiang, L. Zhang, J. Wang, L.-X. You, J.-Y. Guan, D.-X. Yang, Z. Wang, H. Liang, Z. Zhang, N. Zhou, X. Ma, T.-Y. Chen, Q. Zhang, , and J.-W. Pan. Measurement-device-independent quantum key distribution over 200 km. *Phys. Rev. Lett.*, 113:190501, 2014.

[167] Y.-L. Tang, H.-L. Yin, S.-J. Chen, Y. Liu, W.-J. Zhang, X. Jiang, L. Zhang, J. Wang, L.-X. You, J.-Y. Guan, D.-X. Yang, Z. Wang, H. Liang, Z. Zhang, N. Zhou, X. Ma, T.-Y. Chen, Q. Zhang, and J.-W. Pan. Field test of measurement-device-independent quantum key distribution. *IEEE J. Sel. Top. Quantum Electron.*, 21:116–122, 2015.

[168] Y.-L. Tang, H.-L. Yin, X. Ma, C.-H. F. Fung, Y. Liu, H.-L. Yong, T.-Y. Chen, C.-Z. Peng, Z.-B. Chen, and J.-W. Pan. Source attack of decoy-state quantum key distribution using phase information. *Phys. Rev. A*, 88:022308, 2013.

[169] Y.-L. Tang, H.-L. Yin, Q. Zhao, H. Liu, X.-X. Sun, M.-Q. Huang, W.-J. Zhang, S.-J. Chen, L. Zhang, L.-X. You, Z. Wang, Y. Liu, C.-Y. Lu, X. Jiang, X. Ma, Q. Zhang, T.-Y. Chen, and J.-W. Pan. Measurement-device-independent quantum key distribution over untrustful metropolitan network. *Phys. Rev. X*, 6:011024, 2016.

[170] Z. Tang, Z. Liao, F. Xu, B. Qi, L. Qian, and H.-K. Lo. Experimental demonstration of polarization encoding measurement-device-independent quantum key distribution. *Phys. Rev. Lett.*, 112:190503, 2014.

[171] S.-i. Todoroki and S. Inoue. Observation of blowing out in low loss passive optical fuse formed in silica glass optical fiber circuit. *Jpn. J. Appl. Phys.*, 43:L728, 2004.

[172] A. Vakhitov, V. Makarov, and D. R. Hjelme. Large pulse attack as a method of conventional optical eavesdropping in quantum cryptography. *J. Mod. Opt.*, 48(13):2023–2038, 2001.

[173] U. Vazirani and T. Vidick. Fully device-independent quantum key distribution. *Phys. Rev. Lett.*, 113:140501, 2014.

[174] P. Wai and C. Menyak. Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence. *J. Lightwave Technol.*, 14:148–157, 1996.

[175] X.-B. Wang. Beating the photon-number-splitting attack in practical quantum cryptography. *Phys. Rev. Lett.*, 94(23):230503, 2005.

[176] H. Weier, H. Krauss, M. Rau, M. Fürst, S. Nauerth, and H. Weinfurter. Quantum eavesdropping without interception: an attack exploiting the dead time of single-photon detectors. *New J. Phys.*, 13:073024, 2011.

[177] C. Wiechers, L. Lydersen, C. Wittmann, D. Elser, J. Skaar, C. Marquardt, V. Makarov, and G. Leuchs. After-gate attack on a quantum cryptosystem. *New J. Phys.*, 13:013043, 2011.

[178] R. M. Wood. *Laser-Induced Damage of Optical Materials*. CRC Press, 2003.

[179] F. Xu, B. Qi, and H.-K. Lo. Experimental demonstration of phase-remapping attack in a practical quantum key distribution system. *New J. Phys.*, 12:113026, 2010.

[180] F. Xu, S. Sajeed, S. Kaiser, Z. Tang, L. Qian, V. Makarov, and H.-K. Lo. Experimental quantum key distribution with source flaws and tight finite-key analysis. arXiv:1408.3667.

[181] F. Xu, K. Wei, S. Sajeed, S. Kaiser, S. Sun, Z. Tang, L. Qian, V. Makarov, and H.-K. Lo. Experimental quantum key distribution with source flaws. *Phys. Rev. A*, 92:032305, 2015.

[182] H.-L. Yin, T.-Y. Chen, Z.-W. Yu, H. Liu, L.-X. You, Y.-H. Zhou, S.-J. Chen, Y. Mao, M.-Q. Huang, W.-J. Zhang, H. Chen, M. J. Li, D. Nolan, F. Zhou, X. Jiang, Z. Wang, Q. Zhang, X.-B. Wang, and J.-W. Pan. Measurement-device-independent quantum key distribution over a 404 km optical fiber. *Phys. Rev. Lett.*, 117:190501, 2016.

[183] H.-L. Yin, Y. Fu, H. Liu, Q.-J. Tang, J. Wang, L.-X. You, W.-J. Zhang, S.-J. Chen, Z. Wang, Q. Zhang, T.-Y. Chen, Z.-B. Chen, and J.-W. Pan. Experimental quantum digital signature over 102 km. *Phys. Rev. A*, 95:032334, 2017.

[184] H.-L. Yin, Y. Fu, H. Liu, Q.-J. Tang, J. Wang, L.-X. You, W.-J. Zhang, S.-J. Chen, Z. Wang, Q. Zhang, T.-Y. Chen, Z.-B. Chen, and J.-W. Pan. Experimental quantum digital signature over 102 km. *Phys. Rev. A*, 95:032334, 2017.

[185] H.-L. Yin, W.-L. Wang, Y.-L. Tang, Q. Zhao, H. Liu, X.-X. Sun, W.-J. Zhang, H. Li, I. V. Puthoor, L.-X. You, E. Andersson, Z. Wang, Y. Liu, X. Jiang, X. Ma, Q. Zhang, M. Curty, T.-Y. Chen, and J.-W. Pan. Experimental measurement-device-independent quantum digital signatures over a metropolitan network. *Phys. Rev. A*, 95:042338, 2017.

[186] J. Yin, Y. Cao, Y.-H. Li, S.-K. Liao, L. Zhang, J.-G. Ren, W.-Q. Cai, W.-Y. Liu, B. Li, H. Dai, et al. Satellite-based entanglement distribution over 1200 kilometers. *Science*, 356:1140–1144, 2017.

[187] Z.-Q. Yin, Z.-F. Han, W. Chen, F.-X. Xu, Q.-L. Wu, and G.-C. Guo. Experimental decoy state quantum key distribution over 120 km fibre. *Chin. Phys. Lett.*, 25:3547–3550, 2008.

[188] Z. Yuan, A. Sharpe, and A. Shields. Unconditionally secure one-way quantum key distribution using decoy pulses. *Appl. Phys. Lett.*, 90:011118, 2007.

[189] Y. Zhao, C.-H. F. Fung, B. Qi, C. Chen, and H.-K. Lo. Quantum hacking: Experimental demonstration of time-shift attack against practical quantum-key-distribution systems. *Phys. Rev. A*, 78(4):042333, 2008.

[190] Y.-H. Zhou, Z.-W. Yu, and X.-B. Wang. Making the decoy-state measurement-device-independent quantum key distribution practically useful. *Phys. Rev. A*, 93:042324, 2016.

Appendix A

Short pulse attack on continuous-variable quantum key distribution system (Qcrypt2017 Abstract)

Short pulse attack on continuous-variable quantum key distribution system

Hao Qin,[1,2,*] Anqi Huang,[1,3] and Vadim Makarov[2,1,3]

[1]*Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[2]*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[3]*Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*

As a coherent detection technique, homodyne detector (HD) is used in continuous-variable (CV) quantum key distribution (QKD) system for measurements, which is one of the main advantages over discrete-variable (DV) QKD using single photon detectors (SPDs) [1–3]. By using HD, CV QKD can be fully implemented with off-the-shelf telecom components [4–6]. The using of Local Oscillator (LO) in HD acts as single-mode filters, which enables CV QKD signals to be wavelength-multiplexed with intense classical channels over optical networks [7]. Moreover, unlike SPDs as vulnerable targets open for side channel attacks in DV QKD [8–10], CV QKD used to be believed robust against detector-based attack at early time. However, recently Ref [11, 12] has shown that an eavesdropper, Eve, can fully break the security of CV QKD taking advantage of saturation on HD's amplifier electronics. Although the concept of measurement device independent (MDI) is already introduced into CV QKD [13], there is still a large gap between practical implementation and theoretical proposal, there are even debates on whether MDI CV QKD can become practical regarding to its theoretical performances and current available technologies [14, 15]. Thus, it is worth studying detector based attacks in CV QKD to motivate the development of practical MDI CV QKD.

Here, we propose a new side channel attack on CV QKD implementing GG02 (Grosshans and Grangier, 2002) protocol [16] by exploiting HD's imperfections, such as the finite bandwidth of HD amplifiers and limited response time of HD electronics. In particular, we take advantage of the fact that HD's efficiency is dependent on the input pulse temporal mode [17] where Bob's HD can behave nonlinearly when Eve manipulates input pulse widths. In GG02 protocol, Alice modulates quadratures X and P of coherent states with a centered bivariate Gaussian modulation and sends them to Bob. Bob performs homodyne measurement on these coherent states and decodes them into continuous values as raw keys. In practice, in order to make sure Bob can correctly decode information and measure the shot noise (N_0), there is a trade off between electrical noise and bandwidth of HD. For this reason, most of CV QKD experiments [4–6] consist HD with only few MHz bandwidth to limit electrical noise, since Bob's HD must be shot noise limited. Meanwhile Alice needs to increase the pulse duration (typically 100 ns) and reduce repetition rate (1 MHz) to meet Bob's HD bandwidth requirement [5]. However, HD bandwidth will reduce the HD output efficiency significantly if it is smaller than the inverse temporal width of the signal temporal mode [17]. Such effects can be obvious when HD bandwidth is relatively small, which gives more space to a potential Eve to manipulate HD efficiency. The response time of the electronics is typically not faster than a few ns which means if the input pulse width is less than few ns, the HD efficiency also becomes very poor. In order to illustrate such effects, we perform simple experiments in which we vary width of optical pulses from 1 ns to 100 ns at 1550nm and send them to a classical optical photodiode (PD) detector with 3.5 GHz bandwidth and one of the port (PD1 or PD2) of our HD with 100 MHz bandwidth. For each measurement, we record the maximum amplitude value of PD output during the pulse duration as our measurement results (which is similar to sampling stage of CV QKD [5, 18]). As shown in Fig.1 when the pulse duration is longer than 4 ns there is no obvious degrading effect on the output, however when pulse duration becomes shorter, PD's outputs decrease in both case with 100 MHz at about 3 ns and 3.5 GHz bandwidth at about 1 ns. In order to compare the two cases, we normalize all the values by the amplitudes measurement with pulse width of 8 ns. Such observations confirm the predictions on the relation between HD efficiency and pulse width.

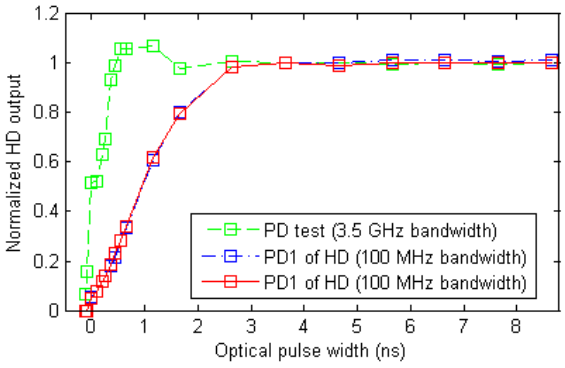

FIG. 1: Short pulse demonstration on classical PD with 3.5 GHz bandwidth and HD with 100 MHz. Each square corresponds to a measurement on the maximum amplitude value of PD (green), PD1 (blue) and PD2 (red) outputs.

By using such effects, Eve can thus manipulate Bob's HD efficiency by changing input pulse widths. If Bob's HD efficiency for certain parts of signal pulses becomes lower, then the linearity between all input quadratures

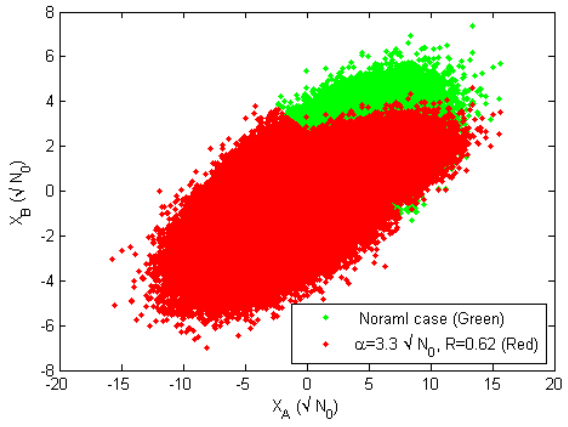

FIG. 2: Alice and Bob data distribution under the short pulse attack. Alice variance $10N_0$, Alice and Bob distance 40 km, fiber attenuation coefficient 0.2 dB/km, reconciliation efficiency 94%, Bob's HD efficiency 60%, electrical noise $0.01N_0$. Alice-Bob excess noise estimation based on green data $2.1N_0$, on red data $0.0084N_0$.

and Bob's HD outputs will not hold, which breaks the important linearity assumption in security proof of CV QKD. Assuming Alice and Bob implement GG02 CV QKD as in [5], we propose Eve's attack strategy as following: (1) Eve fully characterize Bob's HD, particularly, Eve builds the relationship between input pulse width and HD's output efficiency as a reference. As illustrated in Fig. 1, such relationship is determined by the HD's amplifier bandwidth and its electronics. (2) Eve cuts the quantum channel and measures the quadratures X and P sent by Alice with the help of a heterodyne detection [19, 20]. (3) According to her measurement results, Eve prepares corresponding signal pulses as in the intercept-resend (IR) attack [19, 20] with pulse width of 100 ns. Such"entanglement breaking" action will normally rise Alice and Bob's excess noise estimation to at least two units of shot noise ($2N_0$), which includes the vacuum noise of heterodyne detection and the vacuum noise in the new coherent states preparation. (4) Eve adjusts signal pulse widths according to following rules: 4.a) Eve sets a manipulating level ($\alpha > 0$) on her measurements. 4.b) For any Eve's measurement larger than α, she reduces the pulse width of corresponding re-prepared signal such that HD output efficiency reduces to a certain level, which relates to a efficiency reduction ratio R. The relation between HD efficiency and pulse width is determined in step (1). 4.c) For all the rest of resent signal pulses, Eve maintains their widths as Alice's pulse width (100 ns). 4.d) Eve sends all of these re-prepared signal pulses to Bob. (5) Bob performs HD measurements on Eve's resent pulses; Due to different pulse widths adjusted by Eve, Bob's HD output efficiency is not identical respect to different pulses widths (red dots in Fig. 2). (6) Alice and Bob then estimate excess noise on corrupted data which under certain conditions can lead them to underestimate the excess noise due to IR attack. Under such strategy, if the excess noise estimation can be biased below the null key threshold (collective attack [21]), then Eve's IR action won't be spotted by Alice and Bob, which fully breaks the security. We have confirmed this security break in our simulation as shown in Fig. 2, where red data corresponds to the mentioned strategy, Alice and Bob estimate excess noise as $0.0084N_0$ which is still under null key threshold ($0.091N_0$ with simulation parameters shown in Fig. 2). Overall, in our strategy Eve first performs a modified IR attack. By manipulating certain parts of resent signal pulse widths, Eve can force Bob's HD response to be non-linear, which violates the basic assumption of linear detection. Furthermore, Eve can set two target levels $\alpha_1 > 0$ and $\alpha_2 < 0$ to have more freedoms to influence Alice and Bob's data to achieve more powerful attack.

Regarding countermeasures, such attack can be prevented by MDI CV QKD [13, 22, 23]. However previous countermeasures against saturation attack may not be effective [11, 24, 25]. Since in this short pulse attack, Eve actually exploits nonlinear response of Bob's HD in the linear region that is characterized by Alice and Bob, the countermeasures of saturation attack only detect any actions that are happened beyond detection limits, which will not be enough to detect Eve's action in this new attack. On the other hand, the progress of CV QKD security proof includes additional steps such as symmetric test on Alice and Bob data [26], which may eventually cover such kind of attack. Above all, we propose a practical side channel attack targeting HD finite bandwidth and limited speed of electronics. We further propose our attack strategy and demonstrate in simulations that our attack can break the security of current GG02 CV QKD implementations.

* Electronic address: hao.qin@uwaterloo.ca

[1] Weedbrook, C., Pirandola, S., García-Patrón, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H., and Lloyd, S. *Rev. Mod. Phys.* **84**, 621–669 May (2012).
[2] Scarani, V., Bechmann-Pasquinucci, H., Cerf, N. J., Dušek, M., Lütkenhaus, N., and Peev, M. *Rev. Mod. Phys.* **81**, 1301–1350 Sep (2009).
[3] Diamanti, E. and Leverrier, A. *Entropy* **17**(9), 6072–6092 Auguest (2015).
[4] Lodewyck, J., Bloch, M., García-Patrón, R., Fossier, S., Karpov, E., Diamanti, E., Debuisschert, T., Cerf, N. J., Tualle-Brouri, R., McLaughlin, S. W., and Grangier, P. *Phys. Rev. A* **76**, 042305 Oct (2007).
[5] Jouguet, P., Kunz-Jacques, S., Leverrier, A., Grangier, P., and Diamanti, E. *Nat Photon* **7**(5), 378–381 May (2013).
[6] Fossier, S., Diamanti, E., Debuisschert, T., Villing, A.,

Tualle-Brouri, R., and Grangier, P. *New Journal of Physics* **11**(4), 045023 (2009).
[7] Kumar, R., Qin, H., and Allaume, R. *New Journal of Physics* **17**(4), 043027– (2015).
[8] Lydersen, L., Wiechers, C., Wittmann, C., Elser, D., Skaar, J., and Makarov, V. *Nat Photon* **4**(10), 686–689 October (2010).
[9] Gerhardt, I., Liu, Q., Lamas-Linares, A., Skaar, J., Kurtsiefer, C., and Makarov, V. *Nat Commun* **2**, 349– June (2011).
[10] Wiechers, C., Lydersen, L., Wittmann, C., Elser, D., Skaar, J., Marquardt, C., Makarov, V., and Leuchs, G. *New Journal of Physics* **13**(1), 013043– (2011).
[11] Qin, H., Kumar, R., and Alléaume, R. *Phys. Rev. A* **94**, 012325 Jul (2016).
[12] Qin, H., Kumar, R., and Alleaume, R. In *Proc. SPIE 9648, Electro-Optical and Infrared Systems: Technology and Applications XII; and Quantum Information Science and Technology*, volume 9648, 9648V–11, (2015).
[13] Pirandola, S., Ottaviani, C., Spedalieri, G., Weedbrook, C., Braunstein, S. L., Lloyd, S., Gehring, T., Jacobsen, C. S., and Andersen, U. L. *Nat Photon* **9**(6), 397–402 June (2015).
[14] Xu, F., Curty, M., Qi, B., Qian, L., and Lo, H.-K. *Nat Photon* **9**(12), 772–773 December (2015).
[15] Pirandola, S., Ottaviani, C., Spedalieri, G., Weedbrook, C., Braunstein, S. L., Lloyd, S., Gehring, T., Jacobsen, C. S., and Andersen, U. L. *Nat Photon* **9**(12), 773–775 December (2015).
[16] Grosshans, F., Van Assche, G., Wenger, J., Brouri, R., Cerf, N. J., and Grangier, P. *Nature* **421**(6920), 238–241 January (2003).
[17] Kumar, R., Barrios, E., MacRae, A., Cairns, E., Huntington, E., and Lvovsky, A. *Optics Communications* **285**(24), 5259–5267 November (2012).
[18] Li, H., Wang, C., Huang, P., Huang, D., Wang, T., and Zeng, G. *Opt. Express* **24**(18), 20481–20493 September (2016).
[19] Lodewyck, J., Debuisschert, T., García-Patrón, R., Tualle-Brouri, R., Cerf, N. J., and Grangier, P. *Phys. Rev. Lett.* **98**, 030503 Jan (2007).
[20] Cerf, N. J. and Grangier, P. *J. Opt. Soc. Am. B* **24**(2), 324–334 (2007).
[21] García-Patrón, R. and Cerf, N. J. *Phys. Rev. Lett.* **97**, 190503 Nov (2006).
[22] Li, Z., Zhang, Y.-C., Xu, F., Peng, X., and Guo, H. *Phys. Rev. A* **89**, 052301 May (2014).
[23] Ma, X.-C., Sun, S.-H., Jiang, M.-S., Gui, M., and Liang, L.-M. *Phys. Rev. A* **89**, 042335 Apr (2014).
[24] Zhengyu, L., Yichen, Z., Christian, W., and Hong, G. August (2016). Poster at QCrypt 2016.
[25] Huang, P., Huang, J., Wang, T., Li, H., Huang, D., and Zeng, G. *Phys. Rev. A* **95**(5), 052302 May (2017).
[26] Leverrier, A. *Phys. Rev. Lett.* **118**, 200501 May (2017).

Appendix B

Effect of atmospheric turbulence on spatial-mode detector-efficiency mismatch (Qcrypt2017 Abstract)

Effect of atmospheric turbulence on spatial-mode detector-efficiency mismatch

Poompong Chaiwongkhot,[1,2,*] Katanya B. Kuntz,[1,2] Anqi Huang,[1,3] Jean-Philippe Bourgoin,[1,2] Shihan Sajeed,[1,3] Norbert Lütkenhaus,[1,2] Thomas Jennewein,[1,2,4] and Vadim Makarov[2,3]

[1] *Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[2] *Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1 Canada*
[3] *Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1 Canada*
[4] *Quantum Information Science Program, Canadian Institute for Advanced Research, Toronto, ON, M5G 1Z8 Canada*

(Dated: April 26, 2017)

Introduction. In theory, quantum key distribution (QKD) is unconditionally secure; however in practice, a real system is never perfect. Therefore, it is important to study the flaws and vulnerabilities of a system, and find a solution or countermeasure to successful attacks. Recent studies have shown that it is possible to hack a QKD receiver by changing the spatial mode of the incoming beam to the receiver [1]. This attack depends on the ability of the eavesdropper, Eve, to precisely maintain a certain input angle to the receiver. It is well known that turbulence in the transmission channel can, in practice, hinder the performance of both legitimate parties' communication and the adversary's attack. While the assumption of a physical limitation of an eavesdropper (Eve) is not usually part of the security analysis of a QKD system, it is common in practice to have a secure surrounding where Eve could not present, such as in military operation. Therefore, the effect of turbulence on free-space QKD needs to be studied.

We experimentally emulated atmospheric turbulence in the lab using a phase-only spatial light modulator (SLM) to test whether such an attack would still succeed in a turbulent channel. We first verified the accuracy and reproducibility of the atmospheric turbulence emulated by our SLM setup. Then we performed a spatial mode attack for various strengths of the turbulence following a similar procedure as presented in Sajeed *et al.* [1]. From the result, we determined an upper bound on the level of turbulence and distance from adversary where such a spatial mode attack can still succeed on this specific receiver, assuming the adversary only has practical devices with today's technology. Therefore we can determine what atmospheric conditions makes our system safe from this type of attack.

Turbulence emulator. We use a phase-only SLM to emulate atmospheric turbulence in the lab. The advantage of using an SLM as opposed to performing the experiment outside is the ability to generate reproducible turbulence of various strengths without being affected by an unpredictable environment. We chose to generate the phase holograms that represent turbulence based on the Kolmogorov model [2] using a superposition of Zernike polynomials [3]. Zernike polynomials make a convenient basis choice as they directly relate to known optical aberrations, such as tip/tilt, defocus, astigmatism, etc.

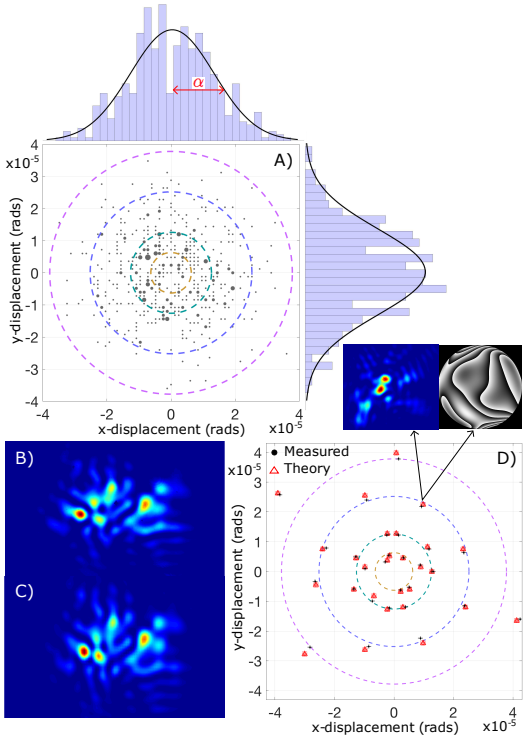

FIG. 1. Turbulence emulator characterization for $r_0 = 1$ cm and $D = 20$ cm. A) Simulated centroid displacements corresponding to 500 phase holograms (α = standard deviation). The size of the data point corresponds to the count frequency. B) Measured and C) Simulated far-field intensity distributions. D) Comparison between measured and simulated centroid displacements for hologram subset.

Another important advantage to using Zernike modes as the basis-set is that their weightings can be analytically calculated based on the strength of turbulence [4]. The radial phase function, $\phi(\rho, \theta)$, that describes each phase hologram is given by a weighted sum of several Zernike polynomials as $\phi(\rho, \theta) = \sum_i c_i Z_i$, where Z_i and c_i are the Zernike polynomial and corresponding coeffi-

* poompong.ch@gmail.com

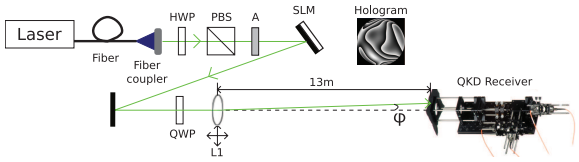

FIG. 2. Experimental setup of our spatial mode attack in a turbulent channel. HWP: half wave-plate, QWP: quarter wave-plate, PBS: polarization beam splitter, A: attenuator, SLM: spatial light modulator, L1: scanning lens, $\phi =$ scanning angle.

cient for the ith mode, respectively, following the Noll labelling convention [3].

We can use simple equations and devices, such as a CCD camera or wavefront sensor, to independently verify and characterize our turbulence emulator. This step is crucial before we can proceed with scanning a QKD receiver in a turbulent channel. It is vital to know whether the emulated turbulence generated by the SLM setup agrees with the predicted strength from theory and simulation results. Therefore, we calculated the theoretical far-field intensity distribution and centroid displacement for comparison with experimental results.

Figure 1 shows both the theoretical and experimental far-field intensity distributions and centroid displacements that emulates strong atmosphere turbulence corresponding to low-altitude sea level ($C_n^2 = 3.67 \times 10^{-14}\,\mathrm{m}^{-2/3}$). Each data point in Fig. 1A and 1D corresponds to a unique phase hologram and far-field distribution. This data illustrates we have excellent agreement between theory and experiment for turbulence emulated using our SLM setup. Therefore, we are confident our setup can accurately emulate reproducible turbulence of various strengths, and we can now attempt a spatial mode attack in a turbulent channel.

Spatial mode attack in a turbulent channel. We use our turbulence emulator to study the effect of turbulence on free-space detection efficiency mismatch. The experimental setup consists of two parts: the turbulence emulator (SLM) and the beam scanning (steering lens, L1), as shown in Fig. 2. Our source is a 532 nm continuous-wave laser that is first sent through polarization optics to generate horizontally-polarized light to ensure phase-only modulation from the SLM. The light after the SLM has a phase wavefront that represents a beam that has travelled through atmospheric turbulence. We use a quarter wave-plate to then rotate the polarization to circularly polarized so there will be a signal on all four detector channels in the QKD receiver. The scanning lens, L1, is mounted on a two-axis motorized translation stage to scan the angle of the outgoing beam. Finally, we place the receiver 13 m away from L1. The QKD receiver under test is a prototype for a quantum communication satellite [6] that has a passive basis choice to detect polarization-encoded light operating at 532 nm on four channels: horizontally **H**, vertically **V**, diagonally at $+45°$ **D** or anti-diagonally at $-45°$ **A**.

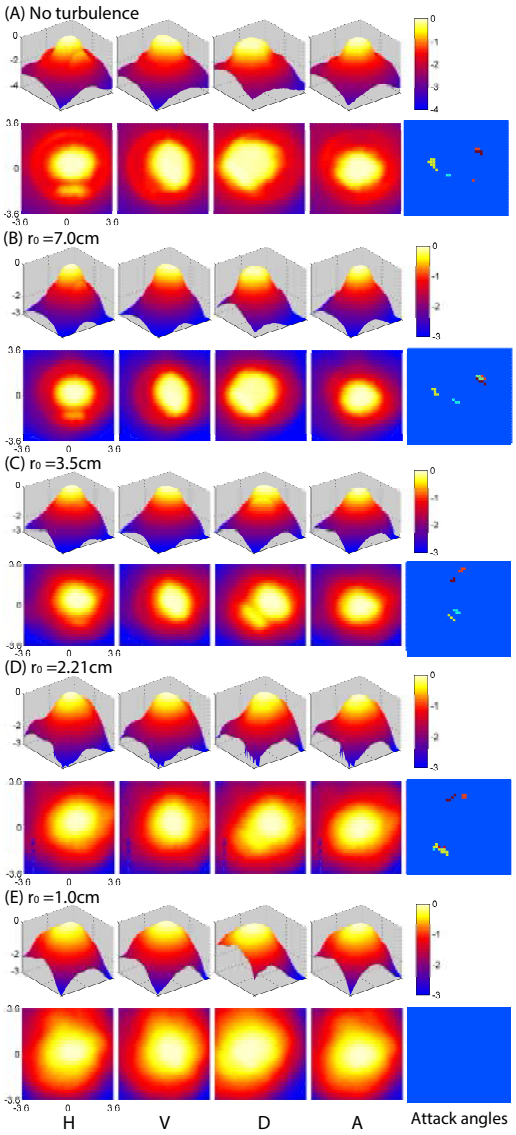

FIG. 3. Normalized count rates for each detector at different incoming beam angles, and the corresponding attack angles for different turbulence strengths. The color of the attack angles denote which detector: dark red: H-detector, red: V, yellow: D, light-blue: A, green: overlap between H and V detectors.

During the receiver alignment procedure, we first send a beam through the center of the lens, L1, to optimize and equalized the detection rates of all four detectors (along dashed line shown in Fig. 2). This initial alignment represents normal operation between Alice (sender) and Bob (receiver). We then adjust the position of lens L1,

TABLE I. Efficiency mismatch parameters for hacking data shown in Fig. 3 and Fig. 4. δ_k = minimum detection efficiency ratio, τ_k = detection efficiency lower bound.

Turbulence r_0	δ_k				τ_k			
	H	V	D	A	H	V	D	A
None	4	4	35	7	0.4	0.08	0.8	0.1
7.0 cm	2	3	30	3	0.4	0.3	0.5	0.5
3.5 cm	1.5	1.5	5	3	0.2	0.1	0.4	0.6
2.21 cm	1.5	1.4	3.5	3	0.2	0.3	0.1	0.5
1.0 cm	1.3	2	1.5	1.5	0.5	0.05	0.4	0.5

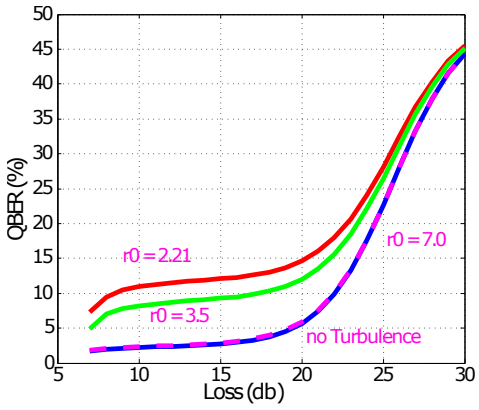

FIG. 4. Optimized quantum bit error rate (QBER) as a function of transmission loss for different turbulence strengths.

and record the detection efficiencies for different angles. The scan is performed in 90 µrad steps covering a range of ±3.6 mrad, which corresponds to a lateral displacement of ±48 mm at the front lens of the QKD receiver.

Our procedure follows the same method as described in [1] to find the potential attack angles where one channel has more probability to click than the other. In our attack model, Eve is restricted to today's technology and using a weak coherent state as her source. The attack angles shows in the right most plot of each sub figure Fig. 3, for example, the scan results without turbulence, Fig.3(A), for the **H** detector, the attack angles are where the **H** detector has a probability of clicking at least 4 times higher than **D** or **A** detectors ($\delta_H = 4$), and the normalized detection probability is greater than 0.4 ($\tau_H = 0.4$). The ratios for the other channels are shown in Tab. I. These parameters were then used in an optimization program to find a set of mean photon numbers that Eve could use for her resent signal to match Bob's expected detection probability while minimizing the quantum bit error rate (QBER).

To simulate the attack under turbulence, we sequentially cycle through a subset of phase holograms on the SLM for each attack angle. We assumed that Eve can measure and correct the tip/tilt component of turbulence using adaptive optics. The final normalized detection efficiency of each detector, η_k, is the weighted sum of the detection rates that resulted from each hologram. We then repeated this process for different turbulence strengths from very weak to strong turbulence corresponding to low-altitude sea level. The attack angles and respective parameters are shown in Fig. 3(B)-(E) and Tab. I.

It can be seen that the stronger turbulence is, the weaker the mismatch ratio (δ_k) and the normalized detection rate at each angle becomes. As a result, the optimized QBER for an attack under strong turbulence is higher overall. The minimized QBER under attack as a function of transmission loss between Alice and Bob is shown in Fig. 4. If we assume that the QBER threshold is 8 %, then the attack without turbulence is successful when the transmission loss between Alice and Bob is less than 22 dB. The weakest turbulence, $r_0 = 7.0$ cm, only slightly affects this result, and looks very similar to the no turbulence case. The strongest turbulence Eve can successfully attack is $r_0 = 3.5$ cm when the transmission loss is lower than 10 dB. This turbulence strength is equivalent to Eve having her resent setup 250 m away from Bob's receiver at sea level. Further more, the result for $r_0 = 2.21$ cm shows that there is no case where the transmission loss between Alice and Bob is low enough where Eve can attack without inducing a QBER that exceeds the threshold. Lastly, for $r_0 = 1.0$ cm, the mismatch ratio is too small ($\delta \leq 2$ for all channels). Therefore, the optimization program could not find a solution for an optimal QBER for any transmission loss.

Conclusion. In this study, we successfully emulated atmospheric turbulence in a lab environment using a phase-only spatial light modulator, and demonstrated a spatial mode detection efficiency mismatch attack in a turbulent channel. We showed the overall trend for the effectiveness of an attack under different turbulence strengths. We found that Eve can attack a free-space non-decoy state BB84 system from up to 250 m away at sea level. Our result implies that if Alice and Bob can establish a secure zone of approximately 250 m around this particular receiver system, then both parties can still exchange a key that is secure from this type of attack.

[1] S. Sajeed *et al.,* Phys. Rev. A **91**, 062301 (2015).
[2] L.C. Andrews and R.L. Phillips, *Laser beam propagation through random media* (SPIE Optical Engineering Press, 1998).
[3] R.J. Noll, J. Opt. Soc. Am., **66**, 207–211 (1976).
[4] L. Burger *et al.,* S. Afr. J. Sci. **104**, 129–134 (2008).
[5] R. Tyson, *Principles of adaptive optics* (CRC Press, 2010).
[6] J.-P. Bourgoin *et al.,* Phys. Rev. A **92**, 052339 (2015).