Supporting Document

Mandatory Technical Document

Evaluation Method for Protection Profile for Prepare and Measure Quantum Key Distribution Modules

May 2025

Version 1.0

Reference

QF-TD-QKD-2025-002_SDv1.0

Disclaimer

The present document has been produced and approved by the Quantum Key Distribution Technology Promotion Committee and represents the views of those members who participated in this committee.
It does not necessarily represent the views of the entire Quantum Forum membership.

Copyright Notification

Acknowledgement

This work was partly supported by the following national projects:
"Research and Development for Construction of a Global Quantum Cryptography Network (JPJ008957)" in "R&D of ICT Priority Technology (JPMI00316)" of Ministry of Internal Affairs and Communication (MIC), Japan.; and
Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), "Photonics and Quantum Technology for Society 5.0" (Funding agency: QST).

Sincere gratitude is extended to the members of European Telecommunications Standards Institute (ETSI), Industry Specification Group on Quantum Key Distribution (ISG-QKD) for their insightful discussions on this document.

Authors

Masato Koashi
University of Tokyo
Akihisa Tomita
Hokkaido University
Go Kato
National Institute of Information and Communications Technology
Mikio Fujiwara
National Institute of Information and Communications Technology
Masahide Sasaki
National Institute of Information and Communications Technology
Ken-ichiro Yoshino
NEC Corporation
Shinya Hirashita
NEC Corporation
Yoshimichi Tanizawa
Toshiba Corporation
Akira Murakami
Toshiba Digital Solutions Corporation
Kenji Yamaya
ECSEC Laboratory Inc.

Reviewers: *QKD Technical Review Committee*

Kiyoshi Tamaki, Chair
University of Toyama
Toyohiro Tsurumaru, Vice chair
Mitsubishi Electric Corporation
Toshimori Honjo
Nippon Telegraph and Telephone Corporation
Kaoru Kenyoshi
National Institute of Information and Communications Technology
Ryutaroh Matsumoto
Institute of Science Tokyo
Takao Saito
ECSEC Laboratory Inc.

Foreword

This is a Supporting Document (SD), intended to complement the Common Criteria (CC) 2022, Revision 1 and the associated Common Evaluation Methodology for Information Technology Security Evaluation (CEM).

Supporting documents may be "Guidance Documents", that highlight specific approaches and application of the standard to areas where no mutual recognition of its application is required, and as such, are not of normative nature, or "Mandatory Technical Documents", whose application is mandatory for evaluations whose scope is covered by that of the supporting document. The usage of the latter class is not only mandatory, but certificates issued as a result of their application are recognized under the Common Criteria Recognition Arrangement (CCRA). This SD shall be considered a Mandatory Technical Document.

Conventions

Citations from CC and CEM are indicated by square brackets.

Document titles and citations are shown in *italics*.

Terminology

Glossary

For definitions of standard CC terminology see [CC] part 1.

Term	Meaning
ADV	Assurance class: Development
AGD	Assurance class: Guidance Documents
ASE	Assurance class: Security Target
ATE	Assurance class: Test
AVA	Assurance class: Vulnerability Assessment

Acronyms

Acronym	Meaning
BB84	Bennett-Brassard 84 Protocol
CC	Common Criteria
CCRA	Common Criteria Recognition Arrangement
CEM	Common Evaluation Methodology
CV-QKD	Continuous-Variable Quantum Key Distribution

Acronym	Meaning
DSC	Dedicated Security Component
DV-QKD	Discrete Variable Quantum Key Distribution
EA	Evaluation Activity
EAL	Evaluation Assurance Level
EB-QKD	Entanglement-Based Quantum Key Distribution
HCD	Hard Copy Device
PP	Protection Profile
QKD	Quantum Key Distribution
SAR	Security Assurance Requirement
SFR	Security Functional Requirement
SD	Supporting Document
ST	Security Target
TOE	Target Of Evaluation
TSF	TOE Security Functionality
TSFI	TOE Security Functional Interface

References

[CC1] Common Criteria for Information Technology Security Evaluation, Part 1: Introduction and general model; November 2022, CC:2022, Revision 1, CCMB-2022-11-001

[CC2] Common Criteria for Information Technology Security Evaluation, Part 2: Security functional components; November 2022, CC:2022, Revision 1, CCMB-2022-11-002

[CC3] Common Criteria for Information Technology Security Evaluation, Part 3: Security assurance requirements; November 2022, CC:2022, Revision 1, CCMB-2022-11-003

[CC4] Common Criteria for Information Technology Security Evaluation, Part 4: Framework for the specification of evaluation methods and activities; November 2022, CC:2022, Revision 1, CCMB-2022-11-004

[CC5] Common Criteria for Information Technology Security Evaluation, Part 5: Pre-defined packages of security requirements; November 2022, CC:2022, Revision 1, CCMB-2022-11-005

[CCEI] Errata and Interpretation for CC:2022 (Release 1) and CEM:2022 (Release 1), 002, Version 1.1, February 1, 2024

[CEM] Common Methodology for Information Technology Security Evaluation, Evaluation methodology, November 2022, CEM:2022 Revision 1, CCMB-2022-11-006

[ISO/IEC 23837-2] ISO/IEC 23837-2:2023 Information security—Security requirements, test and evaluation methods for quantum key distribution—Part 2: Test and evaluation methods, Edition 1, 2023

[ETSISP] STABLE DRAFT Title: Quantum Key Distribution; Security Proofs, ETSI GS QKD 005 V1.3.2 (2021-03)

[AIS20/31] A proposal for: Functionality classes for random number generators, Version 2.0, 18 September 2011

[SP800-22] A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, Revision 1a, April 2010

[PP-EAL4] Common Criteria Protection - Profile Pair of Prepare and Measure Quantum Key Distribution Modules, ETSI GS QKD 016 V2.1.1 (2024-01)

[PP-EAL2] Prepare and Measure Quantum Key Distribution Modules Protection Profile, Version 1.00

[DSCSD] Supporting Document: Evaluation Activities for collaborative Protection Profile for Dedicated Security Component: Mandatory Technical Document, Version 2.0, October 28, 2024

[HCDSD] Supporting Document Mandatory Technical Document Evaluation Activities for collaborative Protection Profile for Hardcopy Devices Version 1.0e, 4 March 2024

[NDSD] Evaluation Activities for Network Device cPP Version: 3.0e Date: 06-December-2023

Table of Contents

1. Evaluation method introduction

1.1. Evaluation method identifier

Title: Supporting Document Mandatory Technical Document Evaluation Method for Protection Profile for Prepare and Measure Quantum Key Distribution Modules

Version: 1.0

Date: May 2025

1.2. Entity responsible for the evaluation method

Quantum Key Distribution Technology Promotion Committee, Quantum Forum

1.3. Technology area and scope of supporting document

This document defines the refinements of SARs and evaluation activities for Quantum Key Distribution (QKD) protocol implementation evaluation in accordance with Common Criteria. Currently, this document supports only the decoy-state BB84 protocol (which is one of the DV-QKD protocols). This document further focuses on a specific implementation called time-bin encoding, in which a pair of optical pulses are transmitted in each of the repeated rounds of communication. In some sections, however, other encoding schemes will be mentioned in the context of discussing general issues in QKD, such as vulnerability analysis. The QKD protocol is a security functional requirement of the PP/ST for QKD modules and is implemented in the QKD module. This document provides evaluation method for the QKD protocol implementation. Other security functions implemented in the QKD module shall be evaluated based on SARs in [CC] ,[CEM] and other supporting documents.

1.4. Evaluation method overview

1.4.1. PP Reference

This document refers to [PP-EAL4] and [PP-EAL2].

This document may be applied to the CC evaluation of TOEs claiming to comply with one of the above PPs. The developer and the evaluator shall select the content and presentation elements of the required developer evidence and work units that correspond to the assurance components in the assurance package of the compliant PP. Table 1-1 shows the corresponding content and presentation elements and work units to be selected for each PP. In other words, when evaluating the TOE that conforms to [PP-EAL2], refer to the left column of Table 1-1, and when evaluating the TOE conforms to [PP-EAL4], refer to the right column of Table 1-1.

The content and presentation elements of the required developer evidence are detailed in Section 3. Evaluation activities of work units are defined in Section 8.

Table 1-1 Correspondence between the PPs and content and the presentation elements and the work units

[PP-EAL2]	[PP-EAL4]
Content and presentation elements in Section 3: Required inputs	
ADV_FSP.2.1C	ADV_FSP.4.1C
ADV_FSP.2.3C	ADV_FSP.4.3C

[PP-EAL2]	[PP-EAL4]
ADV_ARC.1.2C	
ADV_ARC.1.4C	
ADV_ARC.1.5C	
AGD_OPE.1.3C	
AGD_OPE.1.5C	
ATE_COV.1.1C	ATE_COV.2.2C
ATE_FUN.1.1C	
Work units in Section 8 Evaluation activities	
ASE_REQ.2-11	
ADV_ARC.1-2	
ADV_ARC.1-4	
ADV_ARC.1-5	
ADV_FSP.2-4	ADV_FSP.4-5
ADV_FSP.2-5	ADV_FSP.4-6
ADV_FSP.2-9	ADV_FSP.4-11
ADV_FSP.2-10	ADV_FSP.4-12
ADV_TDS.1-7	ADV_TDS.3-15
ADV_TDS.1-8	ADV_TDS.3-16
AGD_OPE.1-3	
AGD_OPE.1-5	
ATE_COV.1-1	ATE_COV.2-4
ATE_FUN.1-1	
AVA_VAN.2-3	AVA_VAN.5-3

The two PPs define equivalent functional requirements, but some SFR identifications are different. The correspondence between their SFR identifications is detailed in Table 1-2. The functional tests for SFRs defined in [PP-EAL2] are described in Section 10.

Table 1-2 Correspondence of SFR identifications

[PP-EAL2]	[PP-EAL4]
FCS_QKD.1	
FPT_ITQ.1	FPT_ITC.1
FPT_EMS.1	
FPT_PHP.3	
FPT_FLS.1	FPT_FLS.1/Fail FPT_FLS.1/EoL
FCS_CKM.6	FCS_CKM.6/EXP FCS_CKM.6/QAK
FCS_COP.1	FCS_COP.1/CCI
FCS_RNG.1	

1.4.2. Concept of the evaluation for the QKD protocol implementation

The security of QKD protocols is mathematically proven as information-theoretical security, meaning that the keys exchanged are secure against attackers who have unbounded computing resources. Security proofs demonstrate that a QKD protocol remains secure under assumptions on the characteristics of the devices used in the QKD system and the conditions on processing in the QKD protocol. Security proofs should preferably take imperfections of the QKD system into account. Unfortunately, however, such security proofs rely on highly precise device characterization techniques, which still require further research and development. Therefore, it is often the case that most assumptions represent perfect devices and the ideal conditions on processing. In this document, such representations are referred to as assumptions of "ideal characteristics". On the other hand, assumptions that represent realistic devices and practical conditions on processing are referred to as assumptions of "realistic characteristics".

Each assumption of ideal characteristics is usually simpler in its description compared to a corresponding assumption of realistic characteristics. It is likely that security proofs taking into account realistic device characteristics and practical processing conditions would require more assumptions than those based on ideal characteristics. Therefore, it would be reasonable to recognize that each assumption of ideal characteristics defines an "assumption family", and each family may include one or more assumptions of realistic characteristics. For the QKD protocol implemented in the TOE, the assumptions in security proofs (whether ideal or realistic) are not always completely fulfilled, and there are deviations between the assumptions and the corresponding characteristics to them in the implementation of the TOE, referred to as "implementation characteristics". Such deviations may compromise the implemented QKD protocol and should be treated as potential vulnerabilities in the QKD protocol. Note that the requirements in the PPs and the assumptions in security proofs are different. The requirements in the PPs are unconditionally fulfilled for any TOE to pass the evaluation, but the assumptions in security proofs are not fulfilled in some cases.

To conduct vulnerability analysis and testing upon the TOE, it is often necessary to restate, modify, or relax the assumptions in security proofs to be testable and preferably quantitative in terms of physical parameters or characteristics rather than remaining in strict and abstract descriptions that current technology cannot implement. This document addresses the assumptions commonly used in security proofs of many QKD protocols, whose concrete descriptions are provided in Section 9 and considers their corresponding testable physical parameters or characteristics, hereafter referred to as "testable parameters/characteristics". In Section 3, commonly used assumption families are listed in Table 3-1, and the testable parameters/characteristics are mapped to each family. Functional tests are derived in Section 10. When designing the functional tests and determining pass/fail criteria, it is often considered that the achievable key generation rate should be practically relevant and not unnecessarily restricted by the criteria of the functional tests.

The testable parameters/characteristics mapped to the assumptions in the security proofs can be linked to appropriate functional test(s) in Section 10.

Regarding parameters/characteristics that are not tested, the developer shall create and provide the guidance to the TOE user to ensure that the performance of TOE components related to those parameters/characteristics are adequately maintained.

Therefore, in the evaluation activity for the QKD protocol implementation:

(1) In ASE class:

The evaluator checks that the QKD protocol linked to the correct security proof already verified is assigned to

the SFR (see Subsection 8.2).

(2) In ADV class:

a) The evaluator examines that the behaviour of the QKD protocol described in the security proof is completely and accurately instantiated in the functional specification and the TOE design (see Subsections 3.1 and Subsection 8.3).

b) The evaluator examines that the assumptions of the security proof are completely and accurately described in terms of the testable parameters/characteristics in the functional specification or the TOE design (see Subsections 3.1 and Subsection 8.3).

(3) In AGD class:

The evaluator examines that the operational user guidance to provides a routine inspection measure to ensure the performance of TOE components related to parameters/characteristics that are not tested (see Subsection 3.2 and Subsection 8.4).

(4) In ATE class:

a) The developer tests functional tests described in Section 10 as developer's tests (see also Subsection 3.3).

b) The evaluator examines that the developer's tests demonstrate the behaviour of the QKD protocol implementation described in the functional specification and the TOE design (see Subsection 3.3, Subsection 8.5 and Section 10).

c) The evaluator examines that the developer's tests demonstrate the testable parameters/characteristics described in the functional specification and the TOE design (see Subsection 3.3, Subsection 8.5 and Section 10).

(5) In AVA class:

a) The evaluator assesses vulnerabilities caused by the deviations between the assumptions in the security proof and the corresponding testable parameters/characteristics, and identifies possible potential vulnerabilities in the TOE. It is not necessary to determine how this affects the security parameters (see Subsection 8.6).

b) The evaluator conducts penetration testing for the identified potential vulnerabilities.

1.4.3. Approval of security proof

This document assumes that the developer or the sponsor has submitted the security proof associated with the QKD protocol to a responsible organization prior to evaluation process. Evaluation of the security proofs themselves is not part of the evaluation for QKD protocol implementation. The security proof shall be approved by the responsible organization. The responsible organization may take the opinion of experts, such as a standards developing organization, into account for approval of the security proof. The developer or the sponsor shall provide the evaluation body with the complete, correct, and comprehensible security proof and a detailed correspondence of the assumptions in the security proof to the implementation as evaluation evidence.

2. Evaluation method dependencies

This document does not depend on any other evaluation method.

3. Required inputs

The required inputs from the developer are shown in the SARs refinements below.

3.1. ADV: Development

3.1.1. ADV_FSP.2.1C, ADV_FSP.4.1C

ADV_FSP.2.1C	*The functional specification shall completely represent the TSF.*
ADV_FSP.4.1C	*The functional specification shall completely represent the TSF.*
Refinement:	**The functional specification shall completely identify the assumptions in the security proof. The identification of assumptions should be consistent with the identification of the assumption families in Table 3-1 and their detailed descriptions in Section 9 of this document.**

The refinements of ADV_FSP.2.1C and ADV_FSP.4.1C aim at providing the knowledge for conducting vulnerability analysis and testing upon the TOE, as described in the AVA and ATE classes, and require that the assumptions in the security proof are completely identified in the functional specification.

3.1.2. ADV_FSP.2.3C, ADV_FSP.4.3C

ADV_FSP.2.3C	*The functional specification shall identify and describe all parameters associated with each TSFI.*
ADV_FSP.4.3C	*The functional specification shall identify and describe all parameters associated with each TSFI.*
Refinement:	**The functional specification shall identify and describe the testable parameters/characteristics, which can be mapped to each of all the identified assumptions in the security proof.**

The refinements of ADV_FSP.2.3C and ADV_FSP.4.3C require the identification and the description of feasible, concrete, and preferably quantitative testing methods for the assumptions in the security proof. Therefore, testable physical parameters or characteristics shall be mapped to the assumptions in the security proof. For example, if the security proof assumes that "the phase of the pulses are completely random", a good concrete description would be "the phase of the pulses are indistinguishable from a random state using specified statistical methods". The description of the testable parameters/characteristics should reflect either design target values or estimated values based on existing knowledge of them.

In this document, the assumptions commonly used in relevant security proofs of the QKD protocol are addressed and their concrete definitions are provided in Section 9. Their assumption names and the corresponding testable parameters/characteristics are shown in Table 3-1. This mapping is provisional and may contain some differences between the meanings of the assumptions and the corresponding testable parameters/characteristics. However, identifying the corresponding testable parameters/characteristics in ADV activity and demonstrating them in ATE activity is useful for AVA activity.

Based on the mapping, each assumption in the security proof can be identified as either of two types:

(i) the assumption is described quantitatively and verifiable by functional tests,

(ii) otherwise.

If type (i) is the case, no vulnerability analysis is required. Otherwise, an assessment of vulnerabilities against the attacks identified for the assumption is necessary.

If the security proof requires a privacy amplification ratio based on the assumptions of realistic characteristics, the developer shall describe the testable parameters/characteristics corresponding to the implemented privacy amplification ratio.

Table 3-1: Assumption families commonly used in security proofs and testable parameters/characteristics mapped to each family.

Classification	Assumption family and testable parameters/characteristics mapped to it
QKD transmitter	**Phase randomization** This family involves assumptions of the phase distribution of the light source, which is ideally indistinguishable from a uniform random distribution. Detailed description is given in Subsubsection 9.1.1. The assumptions can be tested by observing interference between the light pulses. This measurement tests the phase characteristic of output from the light pulse source, rather than phase characteristic of output from the QKD transmitter. In other words, the phase characteristic before attenuating is measured. (Optional) If the phase characteristic is expressed using statistical characteristics, the statistical method should be identified. Functional tests are described in Subsubsection 10.3.2.1 (also refer to ISO/IEC23837(2) 7.7). The related penetration test is described in Subsubsection 11.1.1.
	Photon statistics and intensity This family involves assumptions of the photon number statistics, ideally such that the photon number in each encoded pulse emitted from the QKD transmitter follows a Poisson distribution with a given mean photon number μ. Detailed description is given in Subsubsection 9.1.2. For decoy method, the test is sufficient to measure the ratio of probabilities p(1)/p(2) for signal pulses and decoy pulses, where p(n) is the probability that a pulse contains n photons. Functional tests are described in Subsubsection 10.3.2.2 (also refer to ISO/IEC23837(2) 7.2).
	Degrees of freedom This family involves assumptions of the degrees of freedom of light used by the QKD transmitter to encode the information, ideally such that the characteristics of the intentionally unused degrees of freedom for encoding are independent of the encoded photon state. Detailed description is given in Subsubsection 9.1.3. These assumptions can be tested by measuring the characteristics of each encoded photon state of degrees of freedom other than those used to encode the information. For example, if the polarization of photon pulses is used to encode, the measurement includes the spectrum (wavelength), time waveform, and phase of the photon pulses. Functional tests are described in Subsubsection 10.3.2.3 (also refer to ISO/IEC23837(2) 7.6). The related penetration test is described in Subsubsection 11.1.2 (tentative).
	Security and cryptographic boundaries This family involves assumptions of the cryptographic boundaries of the QKD transmitter, ideally such that no reading of the internal settings of the QKD transmitter unit can be conducted from the outside, nor any modification of its internal components. Detailed description is given in Subsubsection 9.1.4. These assumptions correspond to the assumptions of the PP concerning physical protected environment. If the transmitter implements a countermeasure against optical injection attacks, its functionality must be verified.

Classification	Assumption family and testable parameters/characteristics mapped to it
	Functional tests are described in Subsubsection 10.3.2.4 (also refer to ISO/IEC23837(2) 7.8, 7.9, and 7.10). The related penetration test is described in Subsubsection 11.1.3.
	Accuracy of the encoding This family involves assumptions of the accuracy of the encoding, ideally such that the QKD transmitter modulates a characteristic of the photon state to the expected value. Detailed description is given in Subsubsection 9.1.5. This assumption is tested in terms of fidelity or distance between the ideal photon states and those under examination. Functional tests are described in Subsubsection 10.3.2.5 (also refer to ISO/IEC23837(2) 7.5). The related penetration test can be performed by the method described in Subsubsection 11.1.4.
	Independence of adjacent pulses This family involves assumptions of the correlation between adjacent pulses, ideally such that the intensity of emitted pulses is independent of the intensity modulation pattern. Detailed description is given in Subsubsection 9.1.6. These assumptions can be tested by measuring correlation of the pulse intensities to the adjacent pulse states. Functional tests are described in Subsubsection 10.3.2.6 (also refer to ISO/IEC23837(2) 7.4).
QKD receiver	**Detection efficiency** This family involves assumptions of the detection efficiency of the detectors, ideally such that it is independent of each basis or bit value. Detailed description is given in Subsubsection 9.2.1. These assumptions are tested by measuring the detection efficiencies of the photon detectors. If mechanisms are implemented to counteract differences in the detection efficiency, a function test should be performed to confirm the validity of the mechanisms (reference to specification of the mechanisms). Functional tests are described in Subsubsection 10.3.3.1 (also refer to ISO/IEC23837(2) 8.2).
	Degrees of freedom This family involves assumptions of the degrees of freedom of the detection unit used by the QKD receiver, ideally such that the detection unit reacts always in the same way irrespective of the degree of freedom into which the quantum signal is encoded. Detailed description is given in Subsubsection 9.2.2. These assumptions can be tested by measuring the detection efficiency of the photon detectors. In this measurement, photon characteristics are varied in the designed range for all the degrees of freedom of a photon. Functional tests are described in Subsubsection 10.3.3.2 (also refer to ISO/IEC23837(2) 8.2). The related penetration tests on {time, wavelength, polarization}-shift attacks are described in Subsubsection 11.2.1, which can be waived, if the receiver passes the function test described in Subsubsection 10.9.2.2.
	Security boundary on optical channel This family involves assumptions that no reading of the internal settings of the QKD receiver unit can be conducted from the outside, nor any modification of its internal components. Detailed description is given in Subsubsection 9.2.3. These assumptions correspond to the assumptions of the PP concerning physically protected environment. If the receiver implements a countermeasure against attacks, such as Trojan horse attack and back-flash attack, its functionality must be verified. Functional tests are described in Subsubsection 10.3.3.3 (also refer to ISO/IEC23837(2) 8.3, 8.4 and 8.5). The related penetration tests described in Subsubsection 11.2.2 can be waived, if the receiver passes the function test.

Classification	Assumption family and testable parameters/characteristics mapped to it
	Accuracy of the demodulation This family involves assumptions that an ideal receiver can perfectly distinguish the two optical modes used for encoding on the chosen basis. Detailed description is given in Subsubsection 9.2.4. The functional test is described in Subsubsection 10.3.3.4. The penetration test for the attack is described in Subsubsection 11.2.4.
	Single-photon sensitivity This family involves assumptions of the detection efficiency in the context of bright illumination attacks, ideally such that the single photon sensitivity of the QKD receiver is not controlled by injected bright light. Detailed description is given in Subsubsection 9.2.5. These assumptions can be tested by measuring the detection efficiency of the photon detector under the illumination of bright light. A set of the functional tests are given in Subsubsection 10.3.3.5 and 10.9.1.1 to evaluate the resistance against the bright illumination attack (also refer to ISO/IEC23837(2) 8.6). The penetration test for the attack is described in Subsubsection 11.2.3.
	Recovery or dead time This family involves assumptions of the dead time of the photon detector, ideally such that the photon detector in the QKD receiver always detects a single photon. In other words, the raw data excludes the detection events during the dead time of any photon detectors. Detailed description is given in Subsubsection 9.2.6. The test is similar to that for single-photon sensitivity, but attack should be done during the dead-time of the photon detectors. The tests should consider properly dead-time width, detection window width of the photon detectors, and gate pulse width for gate-mode detectors (if any). The functional tests are described in Subsubsection10.3.3.6. (also refer to ISO/IEC23837(2) 8.7). The penetration test for the attack is described in Subsubsection 11.2.5
Whole of the TOE	**Calibration** This family involves assumptions of calibration, ideally such that the optical signals exchanged in the Calibration phase and the data exchanged in the Post-Processing phase cannot be exploited by attacker to enhance her attack against the QKD system. Detailed description is given in Subsubsection 9.3.1. Functional tests are described in Subsubsection 10.3.4.1 (also refer to ISO/IEC23837(2) 9.1 and 9.2). The related penetration tests described in Subsubsection 11.3.1 can be waived, if the receiver passes the function test. This assumption corresponds to the specification of the calibration. The developer should refer the specification in the functional specification or the TOE design.
	Stability of the light source and the photon detector This family involves assumptions of the stabilities, ideally such that the QKD transmitter and the QKD receiver are typically assumed to remain stable, and the characteristics are the same as when they were characterised. Detailed description is given in Subsubsection 9.3.2. Functional tests are described in Subsubsection 10.3.4.2 (also refer to ISO/IEC23837(2) 7.3). This family corresponds to the stability of the light source in QKD transmitter and the photon detectors in QKD receiver. The developer should refer the user guidance statement which is required by the refinement of AGD_OPE.1.5C in Subsubsection 3.2.2.

Classification	Assumption family and testable parameters/characteristics mapped to it
	Robustness against provoked damage This family involves assumptions of robustness, ideally such that the light source in the QKD transmitter and the photon detectors in the QKD receiver works properly. Detailed description is given in Subsubsection 9.3.3. Functional tests are described in Subsubsection 10.3.4.3 (also refer to ISO/IEC23837(2) 8.9). This assumption corresponds to robustness of the light source of the QKD transmitter and the photon detectors in the QKD receiver. But no countermeasures are currently known to completely prevent damage to the light source or the photon detector. The developer should refer the user guidance statement which is required by the refinement of AGD_OPE.1.5C in Subsubsection 3.2.2.
	Authenticated classical channel This family involves assumptions of the authenticated classical channel, ideally such that the authenticated classical channel provides assured identification of the end point from which channel data was sent and protection of the channel data from modification. Functional tests described in Subsection 10.2.
	Random number generator This family involves assumptions of the random number generator, ideally such that the random number generator provides random bits that meets the defined quality metric. Functional tests described in Subsection 10.6.

3.1.3. ADV_ARC.1.2C

ADV_ARC.1.2C	*The security architecture description shall describe the security domains maintained by the TSF consistently with the SFRs.*

Refinement: **The developer shall describe how to isolate the environment used by untrusted users.**

Security domains refer to environments supplied by the TSF to separate domains for use by potentially-harmful entities; for example, a typical secure operating system supplies a set of resources (address space, per-process environment variables) for use by processes with limited access rights and security properties. Such domains depend on the SFR described in the ST. For example, in the ST which is compliant to [PP-EAL4], the Administrator and Maintainer are trusted due to assumption A,Maint. But Key Requester and Auditor may not be trusted. If the processes run by such untrusted users exist, it may be harmful. So the environments used by such process shall be security domains.

On the other hand, in the ST which is compliant to [PP-EAL2], the operator and all IT products are trusted due to A.OPERATOR and A.IT_PRODUCTS. Since any actions on behalf of users are not allowed before the user is authenticated, no processes run by untrusted users exist if the developer implements SFRs completely and accurately. Therefore, security domains are not necessary.

3.1.4. ADV_ARC.1.4C

ADV_ARC.1.4C	*The security architecture description shall demonstrate that the TSF protects itself from tampering.*

Refinement: **Active probing attacks via the QKD link are considered as attacks that tamper behaviour of**

	the TSF. The security architecture that resists such attacks is one of TSF's self-protection mechanisms. The security architecture description shall contain how the TSF resists active probing attacks and achieves self-protection.

The self-protection mechanism shown in the refinement is related to FPT_PHP.3 in the PP. Even if the specifications for implementing FTP_PHP.3 are shown in the functional specification and the TOE design, the developer shall comprehensively describe which specification resists what type of the attack and how resists the attack in the security architecture description.

3.1.5. ADV_ARC.1.5C

ADV_ARC.1.5C	*The security architecture description shall demonstrate that the TSF prevents bypass of the SFR-enforcing functionality.*
Refinement:	**Side channel attacks over the QKD link are considered as bypass of the SFR-enforcing functionality. The security architecture that prevents such side channel is one of TSF's bypass prevention mechanisms. The security architecture description shall contain how the TSF prevents side channel attacks.**

The bypass prevention mechanism shown in the refinement is related to FPT_EMS.1 in the PP. Even if the specifications for implementing FPT_EMS.1 are shown in the functional specification and the TOE design, the developer should comprehensively describe which specification prevents what type of the attack and how counters the attack in the security architecture description.

3.2. AGD: Guidance documents

3.2.1. AGD_OPE.1.3C

AGD_OPE.1.3C	*The operational user guidance shall describe, for each user role, the available functions and interfaces, in particular all security parameters under the control of the user, indicating secure values as appropriate.*
Refinements:	**The operational user guidance shall provide a procedure for each user role to limit the value of the key establishment attempt counter to secure range. If applicable, the guidance shall contain secure value of the attempt counter threshold. And any security implications related to the management of attempt counter limit shall be detailed.**

3.2.2. AGD_OPE.1.5C

AGD_OPE.1.5C	*The operational user guidance shall identify all possible modes of operation of the TOE (including operation following failure or operational error), their consequences and implications for maintaining secure operation.*
Refinements:	**The operational user guidance shall provide the TOE user with routine inspection measures to ensure that there is no performance degradation due to aging in and no damage to the light source in the QKD transmitter and the photon detector in the QKD receiver. The**

guidance shall contain necessary user actions to maintain secure operation if any performance degradation or damage is identified in either component.

3.3. ATE: Tests

3.3.1. ATE_COV.1.1C

ATE_COV.1.1C	*The evidence of the test coverage shall show the correspondence between the tests in the test documentation and the TSFIs in the functional specification.*

Refinements: The evidence of the test coverage shall contain the correspondence between the tests in the test documentation and the testable parameters/characteristics mapped to the assumptions in the security proof in the functional specification.

See Table 3-1 for the mapping between the assumptions in the security proof and the testable parameters/characteristics.

3.3.2. ATE_COV2.2C

ATE_COV.2.2C	*The analysis of the test coverage shall demonstrate that all TSFIs in the functional specification have been tested.*

Refinements: The analysis of the test coverage shall contain the correspondence between the tests in the test documentation and the testable parameters/characteristics mapped to the assumptions in security proof in the functional specification.

See Table 3-1for the mapping between the assumptions in the security proof and the testable parameters/characteristics.

3.3.3. ATE_FUN.1.1C

ATE_FUN.1.2C	*The test documentation shall consist of test plans, expected test results and actual test results.*

Refinements: The test plan shall include functional tests described in Section 10.

4. Required tool types

The functional tests and the penetration tests identified in this document require some optical tools. The tools are listed in Table 12-1. The developer and the evaluator may choose the required tools for each functional test or the penetration test.

5. Required evaluator competences

The evaluator for the QKD modules shall be able to link the penetration tests shown in Section 11 to the implementation of the TOE and shall be able to judge the validity of the results of the penetration tests. The following knowledge is required.

1. Basic knowledge of the QKD Protocol
2. Knowledge of CC to understand PP and SD
3. Knowledge of QKD module
    A. to understand vendor QKD protocol implementation and QKD module security architecture
    B. to select appropriate penetration tests that can test vulnerabilities of the TOE, understanding vendor's QKD protocol implementation and QKD module security architecture and to build the penetration test step by step

6. Requirements for reporting

The evaluation activities in this document start from SARs and are defined in conjunction with CEM work units. Therefore, the evaluator may include the report of the evaluation activity in the report of the CEM work unit.

7. Rationale for the evaluation method

A rationale is given at the level of the evaluation method below to show that the derivation of the evaluation activities in an evaluation method, from the original work units in the CEM, is appropriate.
This may be given either at the level of the evaluation method, or at the level of individual evaluation activities.

The following rationale shows that the evaluation activities in this evaluation method are appropriately derived from the original work units of the CEM at the level of the evaluation method.

The evaluation method shall include a rationale that the derivation of the evaluation activities from work units in the CEM.

That rationale may contain an explanation of why work units were modified for the scope and depth of an evaluation of a specific technology or TOE type.

The TOE to which this SD applies is QKD modules that implements Decoy-state BB84 with time-bin encoding, one of the QKD protocols. The deviations between the assumptions in the security proof and the actual TOE characteristics corresponding to the assumptions may compromise the security of the QKD protocol and should be treated as potential vulnerabilities in the QKD modules. Since there is no original work unit that handles such deviations, the evaluation activity was derived.

The rationale shall further state how the evaluation activities it contains address all aspects of the action elements in CC Part 3 to which they apply.

The developer action elements use the elements already defined in CC Part 3 without modification.
The content and presentation elements are defined in Section 3 by detailing the information required for the evaluation activity.
The evaluator action elements use the elements already defined in CC Part 3 without modification.

It shall also justify that the manner in which the action elements or work units are addressed is complete with respect to the evaluation context in which the evaluation method is intended to be applied.

The unique context of the evaluation for the QKD protocol implementation involves assessing the deviations between assumptions in security proofs and the actual implementation characteristics of the TOE, as well as evaluating the vulnerabilities arising from these deviations.
Section 1 addresses these aspects by mapping the testable parameters/characteristics to the assumptions and integrating them into functional tests, ensuring a robust evaluation framework tailored to QKD protocol implementation. The evaluation framework is structured using assurance classes: ASE, ADV, ATE, AGD and AVA. Each class provides specific procedures for evaluating the QKD protocol, including identification of a security proof, examining functional specifications, conducting functional tests, maintaining performance through the operational user guidance and assessing potential vulnerabilities.
In Section 3, the content and presentation elements are detailed, and the necessary evaluation evidence is identified.
In Section 8, the following evaluation activities are also defined:

- identification of security proofs and QKD protocols in the ASE class,

- instantiation in specifications and design documents in the ADV class,
- demonstration through testing in the ATE class,
- maintaining performance through the operational user guidance in AGD class, and
- vulnerability analysis in the AVA class.

The developer action element and evaluator action element remain unchanged.

By maintaining these elements, it ensures that evaluators can effectively obtain the necessary evaluation evidence, assess the deviations between security proofs and implementation, and thoroughly evaluate the behaviour of the QKD protocol, thereby ensuring that all elements and work units are completely addressed.

8. Evaluation activities

8.1. Objective

Evaluation Activities (EA) aims to support evaluation for the SFRs of QKD protocols associated with security proofs in the ADV class, evaluation of developer tests in the ATE class and analysing vulnerabilities in the AVA class. The evaluator performs the work units as presented in the CEM. In addition, the evaluator ensures evaluation evidence satisfies EAs specified in the following subsections.

8.2. ASE: Security Target Evaluation

8.2.1. ASE_REQ.2-11

ASE_REQ.2-11	*The evaluator shall examine the statement of security requirements to determine that all assignment operations are performed correctly.*
Evaluation Activity:	**For assignment of QKD protocol, the evaluator shall check the protocol is associated with to security proofs approved by a responsible organization.**

This evaluation activity is related to the assignment for the extended SFR FCS_QKD.1.1 defined in [PP-EAL4] or [PP-EAL2].

8.3. ADV: Development

8.3.1. ADV_ARC.1-2

ADV_ARC.1-2	*The evaluator shall examine the security architecture description to determine that it describes the security domains maintained by the TSF.*
Evaluation Activity:	**The evaluator shall examine the security architecture description to determine how the TSF isolates the environment used by untrusted users.**

8.3.2. ADV_ARC.1-4

ADV_ARC.1-4	*The evaluator shall examine the security architecture description to determine that it contains information sufficient to support a determination that the TSF is able to protect itself from tampering by untrusted active entities.*
Evaluation Activity:	**The evaluator shall examine the security architecture description to determine how the TSF resists active probing attacks and achieves self-protection.**

8.3.3. ADV_ARC.1-5

ADV_ARC.1-5	*The evaluator shall examine the security architecture description to determine that it presents an analysis that adequately describes how the SFR-enforcing mechanisms cannot be bypassed.*
Evaluation Activity:	**The evaluator shall examine the security architecture description to determine how the TSF prevents side channel attacks.**

8.3.4. ADV_FSP.2-1, ADV_FSP.4-1

ADV_FSP.2-1	*The evaluator shall examine the functional specification to determine that the TSF is fully represented.*
ADV_FSP.4-1	*The evaluator shall examine the functional specification to determine that the TSF is fully represented.*
Evaluation Activity:	**The evaluator shall examine the functional specification to determine that it completely identifies all assumptions in the security proof.**

8.3.5. ADV_FSP.2-4, ADV_FSP.4-5

ADV_FSP.2-4	*The evaluator shall examine the presentation of the TSFI to determine that it completely identifies all parameters associated with every TSFI.*
ADV_FSP.4-5	*The evaluator shall examine the presentation of the TSFI to determine that it completely identifies all parameters associated with every TSFI.*
Evaluation Activity:	**The evaluator shall examine the functional specification to determine that it completely identifies all the testable parameters/characteristics mapped to the assumptions in the security proof.**

See Table 3-1 for the correspondence between the assumptions and the testable parameters/characteristics.

8.3.6. ADV_FSP.2-5, ADV_FSP.4-6

ADV_FSP.2-5	*The evaluator shall examine the presentation of the TSFI to determine that it completely and accurately describes all parameters associated with every TSFI.*
ADV_FSP.4-6	*The evaluator shall examine the presentation of the TSFI to determine that it completely and accurately describes all parameters associated with every TSFI.*
Evaluation Activity:	**The evaluator shall examine the functional specification to determine that it completely and accurately describes all the testable parameters/characteristics mapped to the assumptions in the security proof.**

The security proof document contains the security proof and defines the assumptions of the security proof. Generally, the functional specification document which is created during developing the TOE and the security proof document are issued separately. For this reason, the developer needs to describe the assumptions in the security proof in the functional specification document without omission or excess.

The accuracy of the descriptions of the testable parameters/characteristics does not mean that the values of the ideal/realistic characteristics in the assumption of the security proof and the corresponding values of the testable parameters/characteristics in the functional specification match exactly.

In some cases, the assumption is described with the values of the ideal characteristics, but the corresponding values in the functional specification are design target values or estimated values based on existing knowledge of the testable parameters/characteristics. In this case, it is considered accurate if the functional specification describes the values of the testable parameters/characteristics in Table 3-1.

On the other hand, the assumption is described with the values of realistic characteristics (especially when

determining the privacy amplification ratio), the corresponding values in the functional specification shall be consistent. For example, in the case of phase randomization, if the privacy amplification ratio is determined by assuming that the deviation between the realistic phase and the ideal randomized phase is 10, the deviation between the realistic phase and the ideal random phase shall be described also in the functional specification, and the value shall be within 10.

8.3.7. ADV_FSP.2-9, ADV_FSP.4-11

ADV_FSP.2-9	*The evaluator shall examine the functional specification to determine that it is a complete instantiation of the SFRs.*
ADV_FSP.4-11	*The evaluator shall examine the functional specification to determine that it is a complete instantiation of the SFRs.*
Evaluation Activity:	The evaluator shall examine the functional specification to determine that it is a complete instantiation of external behaviour of the QKD protocol described in the security proof.

8.3.8. ADV_FSP.2-10, ADV_FSP.4-12

ADV_FSP.2-10	*The evaluator shall examine the functional specification to determine that it is an accurate instantiation of the SFRs.*
ADV_FSP.4-12	*The evaluator shall examine the functional specification to determine that it is an accurate instantiation of the SFRs.*
Evaluation Activity:	The evaluator shall examine the functional specification to determine that it is an accurate instantiation of external behaviour of the QKD protocol described in the security proof.

8.3.9. ADV_TDS.1-7, ADV_TDS.3-15

ADV_TDS.1-7	*The evaluator shall examine the TOE security functional requirements and the TOE design, to determine that all ST security functional requirements are covered by the TOE design.*
ADV_TDS.3-15	*The evaluator shall examine the TOE security functional requirements and the TOE design, to determine that all ST security functional requirements are covered by the TOE design.*
Evaluation Activity:	The evaluator shall examine the security proof and the TOE design, to determine that all behaviour(s) of the QKD protocol described in the security proof are covered by the TOE design.

8.3.10. ADV_TDS.1-8, ADV_TDS.3-16

ADV_TDS.1-8	*The evaluator shall examine the TOE design to determine that it is an accurate instantiation of all security functional requirements.*
ADV_TDS.3-16	*The evaluator shall examine the TOE design to determine that it is an accurate instantiation of all security functional requirements.*
Evaluation Activity:	The evaluator shall examine the TOE design to determine that it is an accurate instantiation of all behaviour of the QKD protocol described in the security proof.

8.4. AGD: Guidance documents

8.4.1. AGD_OPE.1-3

AGD_OPE.1-3	*The evaluator shall examine the operational user guidance to determine that it describes, for each user role, the available security functionality and interfaces, in particular all security parameters under the control of the user, indicating secure values as appropriate.*

Evaluation Activity: The evaluator shall examine the operational user guidance to determine that it describes for each user role, a procedure to limit the value of the key establishment attempt counter to secure range. If the TSF provides management function of the attempt counter threshold, the evaluator shall examine that the description contains secure value of the attempt counter threshold.

Evaluation Activity: The evaluator shall examine the operational user guidance to determine that it details security implications related to the management of the attempt counter limit.

8.4.2. AGD_OPE.1-5

AGD_OPE.1-5	*The evaluator shall examine the operational user guidance and other evaluation evidence to determine that the guidance identifies all possible modes of operation of the TOE (including, if applicable, operation following failure or operational error), their consequences and implications for maintaining secure operation.*

Evaluation Activity: The evaluator shall examine that the operational user guidance provides the TOE user with routine inspection measures to ensure that there is no performance degradation due to aging in and no damage to the light source in the QKD transmitter and the photon detector in the QKD receiver.

Evaluation Activity: The evaluator shall examine that the operational user guidance contains necessary user action to maintain secure operation if any performance degradation or damage is identified in either component.

8.5. ATE: Tests

8.5.1. ATE_COV.1-1

ATE_COV.1-1	*The evaluator shall examine the test coverage evidence to determine that the correspondence between the tests identified in the test documentation and the TSFIs described in the functional specification is accurate.*

Evaluation Activity: The evaluator shall examine the test coverage evidence to determine that the correspondence between the tests identified in the test documentation and the actual characteristics corresponding to the assumptions in the security proof described in the functional specification is accurate.

Evaluation Activity: The evaluator shall examine the test documentation to determine that functional tests described in Section 10 are performed by the developer.

8.5.2. ATE_COV.2-4

ATE_COV.2-4	*The evaluator shall examine the test coverage analysis to determine that the correspondence between the interfaces in the functional specification and the tests in the test documentation is complete.*

Evaluation Activity: The evaluator shall examine the test coverage evidence to determine that the correspondence between the tests identified in the test documentation and the testable parameters/characteristics mapped to the assumptions in the security proof described in the functional specification is complete.

Evaluation Activity: The evaluator shall examine the test documentation to determine that functional tests described in Section 10 are performed by the developer.

8.5.3. ATE_FUN.1-1

ATE_FUN.1-1	*The evaluator shall check that the test documentation includes test plans, expected test results and actual test results.*

Evaluation Activity: The evaluator shall check the test plan includes the functional tests described in Section 10.

8.6. AVA: Vulnerability Assessment

8.6.1. AVA_VAN.2-3, AVA_VAN.5-3

AVA_VAN.2-3	*The evaluator shall examine sources of information publicly available to identify potential vulnerabilities in the TOE.*
AVA_VAN.5-3	*The evaluator shall examine sources of information publicly available to identify potential vulnerabilities in the TOE.*

Evaluation Activity: The evaluator shall examine Section 9 in this document to identify potential vulnerabilities in the TOE.

9. Identifying potential vulnerabilities in the TOE

The assumptions in security proofs (whether of ideal characteristics or realistic characteristics) are not always fully met, and there are deviations between these assumptions and the corresponding implementation characteristics of the TOE. Such deviations may compromise the security of QKD protocol and should be treated as potential vulnerabilities in the TOE. This section addresses commonly used assumptions in security proofs of many QKD protocols, identifies known attacks (see Section 11 in detail), and provides vulnerability assessments. To conduct vulnerability analysis and testing upon the TOE, the testable parameters/characteristics are mapped to the assumptions as shown in Section 3.

If an assumption is described quantitatively and can be verified through functional tests, no vulnerability analysis is required. This is because these functional tests ensure that appropriate countermeasures are implemented completely and accurately, and by reflecting the corresponding testable parameters in the privacy amplification ratio, it can be proven through a security proof that the identified attacks do not compromise the security of the QKD protocol. Otherwise, an assessment of vulnerabilities against attacks identified for the assumption is necessary, based on the corresponding testable parameters/characteristics.

9.1. QKD transmitter

9.1.1. Phase randomization

Description of assumption family:

It is assumed that a pulse (or a pulse pair) emitted by an ideal QKD transmitter is phase-randomized. As a result, the quantum state of encoded pulses is invariant under optical phase shifts. In the case of polarization encoding, the quantum state of an emitted optical pulse is invariant under any amount of polarization-independent optical phase shift. In the case of time-bin encoding (i.e. phase encoding) on a pulse pair, the quantum state of an emitted pulse pair is invariant under any amount of common optical phase shift applied to both pulses. A security proof involves assessment of how much the attacker may learn about the information encoded on an optical pulse (e.g., bit, basis and nominal intensity of decoy-state) by measuring the pulse. If the security proof adopts the above ideal assumption, it amounts to assume that the attacker can gain no more information from attack strategies sensitive to the common optical phase.

Description of the attack method:

Source attacks with phase information: An attacker prepares a light source that emits pulses whose optical phases are correlated to those of the pulses emitted from the QKD transmitter. They send the target pulse from the QKD transmitter and another pulse from their light source to an interferometer to acquire a measurement outcome. If the QKD transmitter does not satisfy the ideal assumption, the outcome may depend on the optical phase difference between the two pulses. The attacker can then estimate the information encoded on the target pulse based on the measurement outcome. This knowledge may allow for a higher probability of QKD key estimation. For example, if the phase of the pulse encoding bit "1" is not random and has a unique phase, the attacker can estimate in which pulse bit "1" is encoded.

Assessment:

If the security proof specifies relevant quantitative assumptions and those assumptions can be verified by functional

testing, no further analysis is required.
If it is not the case, assessment of vulnerabilities against the attack described above is necessary, which is detailed in the following.

The known attacks assume that the attacker can prepare a light source that emits pulses whose optical phases are correlated to those of the pulses emitted from the QKD transmitter. If the light source used in the QKD transmitter is a gain-switched laser or other pulsed lasers in which the laser oscillation ceases after emission of each pulse, penetration tests are not necessary because preparation of such a light source is not possible with current technology. There has been no demonstration of injection locking with a single seed pulse that contains only one photon or less on average.
If the light source used in the QKD transmitter is a laser that keeps laser oscillation continually, such as a mode-locked laser or a CW laser followed by light intensity modulation, the source attacks with phase information described above using a phase-correlated light source may be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack is described in Subsubsection 11.1.1.

References:

H. -K. Lo and J. Preskill, "Phase randomization improves the security of quantum key distribution", arXiv:quant-ph/0504209.
H. -K. Lo and J. Preskill, "Security of quantum key distribution using weak coherent states with nonrandom phases", Quant. Inf. Comput. **8** 431-458 (2007).
Y. -L. Tang *et al.*, "S Source attack of decoy-state quantum key distribution using phase information", Phys. Rev. A **88**, 022308 (2013).

9.1.2. Photon statistics and intensity

Description of assumption family:

It is assumed that the photon number contained in each of the encoded pulse emitted from an ideal QKD transmitter follows a Poisson distribution with a given mean photon number μ. This is an assumption on infinite number of parameters $p(n)$, which are probabilities of the pulse containing $n = 0,1, \dots, \infty$ photons. Some security proofs adopt relaxed assumptions. Some assume that the mean photon number μ is unknown but satisfies $\mu_0 \leq \mu \leq \mu_1$, where μ_0 and μ_1 are known lower and upper bounds. Others directly assume a set of inequalities fulfilled by $p(n)$ instead of requiring an exact Poisson distribution.

Description of the attack method:

Photon-number-splitting (PNS) attack: Although the BB84 protocol was originally designed to encode information on a single photon, most of the current QKD transmitters use lasers which may emit multiple photons at the same time. The attacker can exploit such an occasion to extract one photon and store it in a quantum memory, while they let the remaining photons be received by the QKD receiver possibly with a better efficiency than the transmissivity of the actual quantum channel to enhance the effectiveness of the attack. Since the quantum state of the photons received by the receiver is not disturbed, this attack causes no increase in the bit error rates. After the basis used for each pulse is announced, the attacker can measure the stored photon to know the bit value encoded in the photon.

The decoy-state BB84 protocol counters this type of attacks by monitoring the detection rates for emitted pulses with different intensities. Since the amount of the trace that should be inevitably left by the PNS attack is estimated under the assumptions of Poisson distribution for the emitted photon number, unexpected deviation from Poisson distribution opens a risk of making the PNS attack effective.

Conditional beam-splitting attack: A weaker version of the PNS attacks implemented by linear optical devices (optical switches and beam splitters), photon detectors and feed-forward electronics. This attack cannot implement the heralded extraction and storing of a single photon in the PNS attacks, but the extracted photon must be immediately measured in a basis. Otherwise, the function provided by this attack differs from the ideal PNS attacks only quantitatively depending on the internal losses and efficiency of optical devices and the bandwidths of the detectors, the switches, and the electronics.

Assessment:

If the security proof specifies relevant quantitative assumptions and those assumptions can be verified by functional testing, no further analysis is required.

If it is not the case, assessment of vulnerabilities against the attack described above is necessary, which is detailed in the following.

The PNS attack described above is possible in principle, but this is not feasible with current technology. Therefore, penetration test(s) are not necessary.

The conditional beam-splitting attack described above can be performed with current technology. However, there is no known detailed strategy to make it work on the decoy-state BB84 protocol. Therefore, penetration test(s) are not necessary.

References:

M. Dušek, et al., “Generalized beam-splitting attack in quantum cryptography with dim coherent states”, Opt. Comm. 169, 103 (1999).

J. Calsamiglia, et al., “Removal of a single photon by adaptive absorption”, Phys. Rev. A 64, 043814 (2001).

J. Calsamiglia, et al., “Conditional beam-splitting attack on quantum key distribution”, Phys. Rev. A 65, 012312 (2001).

9.1.3. Degrees of freedom

Description of assumption family:

It is assumed that pulses from an ideal QKD transmitter leak no information on their encoding to any degrees of freedom of light other than the degree of freedom that the protocol uses for encoding (e.g. polarization or time-bin).

Description of the attack method:

Intercept-resend attack with side information: The attacker intercepts the encoded pulse(s) by detecting a photon that is in a specific mode of the proper degree of freedom and matches to a mode description in other degrees of freedom at the same time. When detection succeeds, the attacker resends another photon in the same mode of the proper degree of freedom. For example, in the case of time-bin encoding when the pulses from the transmitter nominally have V polarization, the attacker may detect a H-polarized photon in the time-bin mode corresponding

to the Z-basis and the bit value 0. When detection succeeds, the attacker resends a V-polarized photon in the time-bin mode corresponding to the Z-basis and the bit value 0. If the pulses from the transmitter have nonnegligible H polarization components only for the Z-basis state, this attack causes no bit errors in the X basis.

Side-channel filtering attack: The attacker places in the optical channel a linear optical transmission filter whose transmissivity does not have dependency in the proper degree of freedom but has dependency in other degrees of freedom. For example, in the case of time-bin encoding, the attacker may place a spectral filter. If the two Z-basis states with bit values 0 and 1 have different transmissivity, the bit recorded upon successful detection at the QKD receiver will be biased. If the same filter does not affect the X-basis states, it causes no bit errors in the X basis.

Quantum nondemolition measurement (QND) attack: There is a quantum process called QND measurement which, when applied to an input photon, produces a measurement outcome and leaves a photon with minimal backaction of the measurement. A variant of this type of measurement may provide a wider variety of input-output relation than the intercept-resend attack, which the attacker may exploit to acquire larger bit information with a smaller increase in the bit error rates.

Photon-number-splitting (PNS) attack and Conditional beam-splitting attack: See the description in Subsubsection 9.1.2. The decoy-state BB84 protocol counters these types of attacks by monitoring the detection rates for emitted pulses with different intensities. Any leak of the information related to the intensity of the pulses through other degrees of freedom opens a risk of making these attacks effective.

Assessment:

If the security proof specifies relevant quantitative assumptions and those assumptions can be verified by functional testing for a degree of freedom, no further analysis is required for the degree of freedom.
If it is not the case, assessment of vulnerabilities against the attack described above is necessary, which is detailed in the following.
The intercept-resend attack with side information described above can be performed with current technology. It is necessary to conduct and pass penetration tests.
The side-channel filtering attack described above can be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack is described in Subsubsection 11.1.2.
The QND attack described above is possible in principle, but this is not feasible with current technology. Therefore, penetration test(s) are not necessary.
The PNS attack described above is possible in principle, but this is not feasible with current technology. Therefore, penetration test(s) are not necessary.
The conditional beam-splitting attack described above can be performed with current technology. However, there is no known detailed strategy to make it work on the decoy-state BB84 protocol. Therefore, penetration test(s) are not necessary.

9.1.4. Security and cryptographic boundaries

Description of assumption family:

It is assumed that an ideal QKD transmitter allows no reading of its internal settings and no modification of its internal components.

Description of the attack method:

- An attacker reads/writes internal settings of the QKD transmitter.
- An attacker modifies internal components of the QKD transmitter.
- An attacker reads internal confidential data from internal components of the QKD transmitter.
- An attacker observes internal states of the QKD transmitter.

An attacker uses these adverse actions to disclose the QKD key or compromise the QKD transmitter.
One well-known attack method is the Trojan horse attack. An attacker injects light into the QKD transmitter via the QKD link, observes the reflected light, and estimates the status of modulation optics in the QKD transmitter. From this, the attacker guesses the basis, the bit value, and the intensity choices made by the QKD transmitter.

Assessment:

The security and cryptographic boundaries of the QKD transmitter are physically protected due to the assumption of each PP. i.e. A.SecureOp of [PP-EAL4] or A.PHYSICAL of [PP-EAL2]. So an attacker cannot access internal components of the QKD transmitter directly.
The internal settings of the QKD transmitter are protected by user identification and authentication functions and access control functions via user interface(s). So an attacker cannot access internal settings of the QKD transmitter via user interface(s).
If above assumptions are achieved and above functions are implemented completely and accurately, no potential vulnerabilities exist in above point of view.
However, the QKD link is not physically protected and not access controlled. An attacker may observe internal state (e.g. choice of encoding basis) of the QKD transmitter via injecting probing light through the QKD link (Trojan horse attack). An attacker may also attempt to modify the characteristics of internal components (e.g. laser source) via irradiation through the QKD link.

Trojan horse attack countermeasures are implemented in several steps.

1. A light injection monitor is implemented that monitors the light intensity injected into the QKD transmitter.
2. When the light injection monitor detects strong light, the TSF will automatically respond to prevent information leakage due to light reflection. e.g. the TSF performs "emergency stop of the QKD link" (FPT_PHP.3).
3. When light is injected below the detection limit, the maximum reflected light intensity is estimated based on the transmission and reflection characteristics of the QKD transmitter components.

If the security proof specifies quantitative assumptions on the reflected light intensity and those assumptions can be verified by functional testing, no further analysis is required on the Trojan horse attack.
If it is not the case, assessment of vulnerabilities against the Trojan horse attack is necessary, which is detailed in the following.

Among variants of the Trojan horse attacks, one that estimates the choice of pulse intensity must accompany the PNS attack or its variant described in Subsubsection 9.1.2. The PNS attack is possible in principle, but this is not feasible with current technology. The conditional beam-splitting attack can be performed with current technology, but there is no known detailed strategy to make it work on the decoy-state BB84 protocol. Therefore, penetration

tests for this variant of the Trojan horse attack is not necessary.

Variants of the Trojan horse attacks that estimate the choice of the basis and the bit value can be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack is described in Subsubsection 11.1.3.

On the other hand, assessment of the threat of light irradiation altering the characteristics of internal components is given as the following. If a light injection monitor is implemented, strong light injection will be detected. In other words, the intensity of light injection is limited by the detection threshold of the light injection monitor. It has not been reported that the characteristics of linear optical components are affected by light injection with the intensity below the detection threshold. On the other hand, a laser can be affected by injected light due to its nonlinear dynamics. The effect of nonlinearity is most significant when the frequency of the injected light is the same as that of the laser. This situation has been analyzed as the effect of feedback light on a laser. It is reported that the feedback effects are negligible when the light is reinjected into the laser as a fraction smaller than one 10^{-6} of the emitted light. Therefore, the effect of the injected light does not need to be considered if the estimated intensity is less than the above criterion. Otherwise, the TOE should be tested using the test specified in Subsubsection 10.3.2.4.

Reference

K.Stubkjaer and M. Small, "Noise properties of semiconductor lasers due to optical feedback", *IEEE Journal of Quantum Electronics*, vol. 20, no. 5, pp. 472-478, May 1984, doi: 10.1109/JQE.1984.1072428.

K. I. Kallimani and M. J. O'Mahony, "Relative intensity noise for laser diodes with arbitrary amounts of optical feedback", *IEEE Journal of Quantum Electronics*, vol. 34, no. 8, pp. 1438-1446, Aug. 1998, doi: 10.1109/3.704337.

9.1.5. Accuracy of the encoding

Description of assumption family:

Encoding of the QKD transmitter is performed by modulating the assumed degree of freedom of light. An ideal QKD transmitter carries out the modulation accurately as implied by the QKD protocol and by the chosen values of protocol parameters. In the case of the decoy-state BB84 protocol, a degree of freedom formed by a pair of optical modes, such as the polarization and the time bin, is used for the encoding of the four states of the BB84 protocol. A set of values for the pulse intensity are specified as protocol parameters of the decoy-state BB84 protocol.

Description of the attack method:

Intercept-resend attack on the monitoring basis: The attacker intercepts the encoded pulse(s) from the QKD transmitter and makes a photon detection to distinguish the two states for the basis used for monitoring, determining a bit value. When the detection was successful, the attacker prepares a stronger optical pulse with the proper modulation corresponding to the determined bit value and sends it to the QKD receiver. This attack introduces no additional errors, but the determined bit value may partially reveal the bit value chosen by the QKD transmitter on the other basis if the encoding is not accurate.

Photon-number-splitting (PNS) attack and Conditional beam-splitting attack: See the description in Subsubsection 9.1.2. The decoy-state BB84 protocol counters these types of attacks by monitoring the detection rates for emitted pulses with different intensities. Unexpected deviation of the modulation intensity opens a risk of

making these attacks effective.

Assessment:
If the security proof specifies relevant quantitative assumptions and those assumptions can be verified by functional testing, no further analysis is required.
If it is not the case, assessment of vulnerabilities against the attack described above is necessary, which is detailed in the following.
The PNS attack described above is possible in principle, but this is not feasible with current technology. Therefore, penetration test(s) are not necessary.
The conditional beam-splitting attack described above can be performed with current technology. However, there is no known detailed strategy to make it work on the decoy-state BB84 protocol. Therefore, penetration test(s) are not necessary.
The intercept-resend attack on the monitoring basis can be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack is described in Subsubsection 11.1.4.

9.1.6. Independence of adjacent pulses

Description of assumption family:
The internal states of light source and the modulation components of an ideal QKD transmitter in one communication round are statistically independent of those in the other rounds.

Description of the attack method:
If the internal states such as the choices of bases, bit values, and pulse intensities in one communication round are correlated to the optical pulses emitted in other rounds, the latter pulses serve as a side channel from which the attacker may extract information on the encoding.

Assessment:
If the security proof specifies relevant quantitative assumptions and those assumptions can be verified by functional testing, no further analysis is required.
If it is not the case, assessment of vulnerabilities against the attack described above is necessary, which is detailed in the following.
When there are correlations between the internal states of a round and the emitted pulses in other rounds, an attacker may measure the pulses and estimate the status of modulation optics in the QKD transmitter in the former round. From this, the attacker may guess the basis, the bit value, and the intensity choices made by the QKD transmitter. This threat is equivalent to that of the Trojan horse attack on the transmitter described in Subsection 9.2.3 except that the role of the reflection of the injected pulse is substituted by the emitted pulses in the other rounds. This attack can be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack can be designed as a modification of the one described in Subsubsection 11.1.3.

9.2. QKD receiver

9.2.1. Detection efficiency

Description of assumption family:

It is assumed in most security proofs that the detection efficiency of the detectors is independent of each basis or bit value.

Description of the attack method:

An attacker can eavesdrop on the bit value transmitted from the QKD transmitter using man-in-the-middle attack with a certain probability. This certain probability is taken into account in the privacy amplification and the eavesdropped bits are removed from the final QKD key, so this attack method is ineffective.

However, if the detection efficiency differs depending on the basis, an attacker optimizes the eavesdropping strategy for each basis, he may succeed in eavesdropping more than a certain probability.

Or, for example, if the detection efficiency of bit 0 is lower, the attacker can estimate with high probability that the raw key is bit 1 without even eavesdropping.

Passive attack based on sifted key inference:

If the probability for a sifted key bit to have one value, say, 0, is larger than that for the other value, say, 1, distribution of the sifted key is not uniform, and the attacker may exploit that information for guessing the value of the QKD key. This involves no active intervention on the quantum channel and hence leads to no increase in the observed bit error rate.

Assessment:

If the security proof specifies relevant quantitative assumptions and those assumptions can be verified by functional testing, no further analysis is required.

If it is not the case, assessment of vulnerabilities against the attack described above is necessary, which is detailed in the following.

If the two photon detectors for bit values 0 and 1 used for generation of sifted key bits have different detection efficiencies, the probability of a sifted key bit to have one value is larger than that of the other value, leading to potential vulnerability against the passive attack based on sifted key inference described above, can be performed with current technology. The TOE may adopt some mechanisms to cancel out the difference in detection efficiencies, such as inserting an optical attenuator before a detector or randomly changing assignment of the bit values to the two detectors. Since these mitigating mechanisms involve physical means, the cancellation should still be imperfect.

If the TOE passes the functional test described in Subsubsection 10.9.2, the probability that the passive attack based on sifted key inference will succeed is expected to be extremely low, based on the rationale provided in Subsubsection 13.2.1. Therefore, the penetration test for this attack can be waived.

9.2.2. Degrees of freedom

Description of assumption family:

An ideal detection unit reacts always in the same way irrespective of the degree of freedom into which the quantum signal is encoded. For polarization coding, for example, the detectors monitoring the various polarization modes are assumed to behave the same for all the pulses' degrees of freedom, such as timing, wavelength or spatial mode.

Description of the attack method:

The attacker modifies the degrees of freedom of the optical pulse on the quantum channel.

For example:

- delays the optical pulse;
- shifts wavelength phase of the optical pulse;
- shifts polarization of the optical pulse.

If detection efficiency of the photon detector changes depending on these degrees of freedom, in the extreme case the QKD receiver will be unable to receive bit 0, the attacker can presume that all raw key is bit 1. Even if it is not so extreme, if detection efficiency for bit 0 of the detector decreases, the attacker can estimate with high probability that the raw key is bit 1.

Assessment:

The degrees of freedom of light pulse include polarization, spatial mode, timing, and wavelength. However, since the spatial mode is defined by the input single-mode fibre and there is only one spatial mode incident on the photon detectors, it is impossible to attack using the difference in detection efficiency depending on the spatial mode.

The time-shift attack, wavelength-dependent attack, and polarization--dependent attack described above can be performed with current technology. The penetration tests are described in Subsubsection 11.2.1 for the attacks. However, if the TOE passes the functional test described in Subsubsection 10.9.2, the probabilities that the TOE will not pass the penetration test for the time-shift attack is expected to be extremely low, based on the rationale provided in Subsubsection 13.2.1. Therefore, the penetration test for these attacks can be waived.

References:

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

9.2.3. Security boundary on optical channel

Description of assumption family:

It is assumed that an ideal QKD transmitter allows no reading of its internal settings and no modification of its internal components.

Description of the attack method:

- An attacker reads/writes internal settings of the QKD receiver.
- An attacker modifies internal components of the QKD receiver.
- An attacker reads internal confidential data from internal components of the QKD receiver.
- An attacker observes internal states of the QKD receiver.

An attacker uses these adverse actions to disclose the QKD key or compromise the QKD receiver.

Assessment:

The security and cryptographic boundaries of the QKD receiver are physically protected due to the assumption of each PP. i.e. A.SecureOp of [PP-EAL4] or A.PHYSICAL of [PP-EAL2]. So an attacker cannot access internal components of the QKD receiver directly.

The internal settings of the QKD receiver are protected by user identification and authentication functions and

access control functions via user interface(s). So an attacker cannot access internal settings of the QKD receiver via user interface(s).
If above assumptions are achieved and above functions are implemented completely and accurately, no potential vulnerabilities exist in above point of view.
However, the QKD link is not physically protected and not access controlled. An attacker may observe or modify internal state of the QKD receiver via the QKD link, e.g. choice of encoding basis.
The attack method that exploits modification of the internal state is used in some attacks. An example is bright illumination attack described in Subsubsection 9.2.5.
One of attack method that exploits observation of the internal state is Back-flash attack. This attack is applicable when different detectors are implemented for each photon state. It is known that detectors emit weak light by themselves in response to detection. If the emission varies depending on detectors, an attacker can obtain information on detector detection events by observing the emission. The Back-flash attack can be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack is described in Subsubsection 11.2.2.

Another attack method that exploits observation of the internal state is Trojan horse attack. This attack is applicable when the modulator is used to select the basis and same detector is used for all basis. An attacker injects light into the QKD receiver via the QKD link, observes the reflected light, and estimates the basis state.

Trojan horse attack countermeasures are implemented in several steps.
1. A light injection monitor is implemented that monitors the light intensity injected into the QKD receiver.
2. When the light injection monitor detects strong light, the TSF will automatically respond to prevent information leakage due to light reflection. e.g. the TSF performs "emergency stop of the QKD link" (FPT_PHP.3).
3. When light is injected below the detection limit, the maximum reflected light intensity is estimated based on the transmission and reflection characteristics of the QKD receiver components.

If the security proof specifies quantitative assumptions on the reflected light intensity and those assumptions can be verified by functional testing, no further analysis is required on the Trojan horse attack.

If it is not the case, assessment of vulnerabilities against the Trojan horse attack is necessary, which is as follows.
The Trojan horse attack can be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack is described in Subsubsection 11.2.2.

9.2.4. Accuracy of the demodulation

Description of assumption family:

The decoy-state BB84 protocol dictates that the receiver chooses between the two measurement bases. A measurement basis is usually determined by a set of optical components in front of two photon detectors. An ideal receiver can perfectly distinguish the two optical modes used for encoding on the chosen basis.

Description of the attack method:

Intercept-resend attack on the monitoring basis: The attacker intercepts the encoded pulse(s) from the QKD

transmitter and makes a photon detection to distinguish the two states for the basis used for monitoring, determining a bit value. When the detection was successful, the attacker prepares a stronger optical pulse with the proper modulation corresponding to the determined bit value and sends it to the QKD receiver. This attack introduces no additional errors, but the determined bit value may partially reveal the bit value determined by the QKD receiver on the other basis if the measurement bases are not accurate.

Assessment:

If the security proof makes no assumption on the accuracy of measurement bases, or if it specifies relevant quantitative assumptions and those assumptions can be verified by functional testing, no further analysis is required. If it is not the case, assessment of vulnerabilities against the attack described above is necessary, which is detailed in the following.

The intercept-resend attack on the monitoring basis can be performed with current technology. It is necessary to conduct and pass penetration tests. The penetration test for the attack is described in Subsubsection 11.2.4.

9.2.5. Single-photon sensitivity

Description of assumption family:

It is assumed that the single photon sensitivity of the QKD receiver is not controlled by injected bright light.

Description of the attack method:

An attacker injects bright light to the QKD receiver at the timing when the light pulse of the QKD transmitter should be received. After that, the attacker injects trigger light that encodes his own bit values. If the detection efficiency of the single photon detector in the QKD receiver decreases due to the injected bright light, the QKD receiver will not be able to receive the photons transmitted by the QKD transmitter. In this situation, the attacker can force to receive intended bit value to the QKD receiver using own trigger light at a later timing.

There are two distinct types of bright illumination attacks. In the ideal case, those attacks are described as follows. (Here we assume a decoy state BB84 protocol in which the sifted key is generated from the Z basis only and the X basis is only used to monitor eavesdropping.)

1). Bright light puts all detectors in linear mode. When a strong control light with the same optical mode as a signal light used in the Z-basis is injected, all the light in the Z-basis is incident on the corresponding detector. The intensity of the control light is chosen so that the intensity at the incidence is slightly above the threshold of the linear mode detector. In this case, no detection occurs in the X-basis, as the control light is equally divided between the two detectors.

2). This attack is only valid for devices with passive base selection. The passive basis selection device uses a pair of photon detectors for the measurement of the X-basis and a pair of photon detectors for the measurement of the Z-basis. Bright light reduces only the quantum efficiency of the detectors in the X-basis to zero. When the control light with the same mode as a signal light used in the Z-basis is injected, all the light branched to the Z-basis detection is incident on the corresponding detector and only that detector causes detection. The X-basis detectors do not cause detection due to the bright light.

By using attack type 1) or 2) at the resending step of the intercept-resend attack, the attacker can learn the value of the sifted key in the Z basis without increasing the bit error in the X basis.

Assessment:

One of countermeasure against bright-illumination attack is to implement a light injection monitor in the TOE. The light injection monitor is a function implemented inside the QKD receiver, which detects and alarms when bright light is input. If no countermeasures are implemented, FPT_PHP.3 is not fulfilled, and ADV activity of the evaluation will be failed. Based on the rationale provided below, a set of the functional tests in Subsection 10.3.3.5 and Subsubsection 10.9.2.1, instead of a penetration test, can suffice to evaluate how well the TOE withstand bright-illumination attacks.

Feasibility of attack type 1) can be evaluated from the functional test in Subsubsection 10.9.2.1. When a detector goes to the linear mode under illumination of a bright pulse with an intensity μ, it is no longer sensitive to a small signal input. It follows that if the intensity of the bright pulse is increased from zero to μ, a significant decrease of sensitivity to a small signal input should be observed. On the other hand, the functional test requires that there be no loss of sensitivity in the intensity range at which the light injection monitor is not activated. Hence, if the TOE passed the functional test of Subsubsection 10.9.2.1, the attack type 1) should fail because it is impossible to make the detectors transition to the linear mode without triggering the light injection monitor.

The attack type 2) may still effective even if the detectors are not switched to the linear mode. This attack can be performed with current technology, and the penetration tests described in Subsubsection 11.2.2 can be identified for the attack. However, if the TOE passes the functional test described in Subsubsection 10.9.2.1, the probability that the TOE will not pass the penetration test described in Subsubsection 11.2.3 is expected to be extremely low, based on the rationale provided in Subsubsection 13.2.2.

References:

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).
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 (2010).
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).
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. A84, 032320 (2011).

9.2.6. Recovery or dead time

Description of assumption family:

It is assumed that detection efficiency of the photon detector is not affected by past detection events.
This assumption can be seen as a subset of the assumption on single photon sensitivity. After a detection event, a single photon detector takes some time to recover (referred to as dead-time). During this time, it loses its single photon sensitivity.

Description of the attack method:

An attacker injects blinding light to the QKD receiver outside the detection window of the QKD receiver and

aiming at the timing when the transmitted pulse arrives in the dead-time of the detector. The timing should be a little before the detection window.
If the blinding light is encoded a specific photon state used in the TOE, the light blinds a specific photon detector. For example, when the blind light is encoded bit “1” with Z-basis , the QKD receiver will not be able to receive bit “1” with Z-base. As a result, the bit "1" is lost with in Z-basis, and an attacker can predict that the bit in the sifted key is "0" with high probability.
Experiments show that this attack is successful even with blind light of 20 photons or less.

Assessment:
Some TOEs implement countermeasures to ensure that the photon detector is not blinded.
For Example,
(a) The QKD receiver monitors the terminal voltage of the detector. Since the terminal voltage is temporarily dropped due to photo detection, if a terminal voltage of one detector drops, the QKD receiver disables all detections until the terminal voltage recovers.
(b) If a gated mode photon detectors are implemented, the QKD receiver controls the gate signal to not detect photons before the detection window.

However, even if (a) is implemented, the terminal voltage drop width and the dead-time width may not be exactly same. The time of descent may be shorter than the dead time, and the detector may be activated during the dead time.
If (b) is implemented, an attacker may inject strong blind light to force detection and blind the detector. To counter such attacks, light monitors can also be implemented to detect the stronger light. However, the sensitivity of the light monitor may not be sufficient to detect the blind light and fail to prevent blinding.
Therefore, the penetration tests shall demonstrate the TOE counters such attacks. See Subsubsection 11.2.5.

References:
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).

9.3. Whole of the TOE

9.3.1. Calibration

Description of assumption family:
It is commonly assumed that the optical signals exchanged in the calibration phase cannot be exploited by the attacker to enhance her attack against the QKD system. However, a slack execution of these phases, lacking coordination between the users or leaking more information than strictly necessary to the eavesdropper can compromise the security of the whole QKD system.

Description of the attack method:
For example, if the QKD receiver uses two detectors, the detection timing is adjusted using two pulses, one to adjust the detection timing of detector-A and the other to adjust the detection timing of detector-B. An attacker imposes a time delay only on the training pulse for detector-A. Then the detector-A is adjusted to non-optimal detection timing. If this reduces the detection efficiency of the detector-A, an attacker may succeed in the detection

efficiency mismatch attack shown in Subsubsection 9.2.1.

Assessment:
The penetration tests shall demonstrate the TOE counters such attacks. See Subsubsection 11.3.1.

9.3.2. Stabilities of the light source and the photon detector

Description of assumption family:
The light source of the QKD transmitter and the photon detectors of the QKD receiver are typically assumed to be stable and the characteristics are the same as when they were characterised. However, in practice, the light source and the photon detector may deteriorate over time and the security of the TOE cannot be guaranteed with the deteriorated characteristics of the device.

Description of the attack method:
If the light source of the QKD transmitter deteriorates over time, in extreme cases, the optical phase will become skewed, making it easier for attackers to predict the transmitted optical phase.
If the photon detector in the QKD receiver deteriorates over time, in an extreme case, the QKD receiver will be unable to receive bit 0, the attacker can estimate that all raw key is bits 1. Even if it is not so extreme, if detection efficiency for bit 0 of the detector decreases, the attacker can estimate with high probability that the raw key is bit 1.

Assessment:
The developer shall provide a routine inspection measure to ensure that the light source of the QKD transmitter and the photon detectors in the QKD receiver is no performance degradation for the TOE user. If the performance of the light source and photon detector are maintained through the regular inspection, penetration tests are not necessary.

9.3.3. Robustness against provoked damage

Description of assumption:
It is assumed that the light source of the QKD transmitter and the photon detectors in the QKD receiver works properly.

Description of the attack method:
An attacker injects strong light to the QKD receiver or the QKD transmitter via the quantum channel.
If the photon detector for bit 0 in the QKD receiver is permanently damaged due to the attack, in an extreme case, the QKD receiver will be unable to receive bit 0, the attacker can estimate that all sifted key is bits 1. Even if it is not so extreme, if detection efficiency for bit 0 of the detector decreases, the attacker can estimate with high probability that the sifted key is bit 1. If the modulator in the QKD transmitter is permanently damaged due to the attack, in an extreme case, only the unmodulated state is sent from the QKD transmitter. The attacker predicts the bit values in the sifted key with a high probability.

Assessment:

No countermeasures are currently known to completely prevent damage to the optical devices.
Theoretically, for example after injecting very strong light into the QKD receiver, a penetration test can be considered that demonstrates that there is no significant difference in the detection efficiency of bit 0 and bit 1, but this test means a fracture test. Therefore, it is difficult to ensure that the TOE counters this attack method completely.
At the least, the developer shall provide a routine inspection measure to ensure that the optical devices in the QKD transmitter and QKD receiver is undamaged for the TOE user. If the performance of the light source and photon detector are maintained through the regular inspection, penetration tests are not necessary.

9.3.4. Authenticated classical channel

Description of assumption:

The authenticated classic channel is assumed to assure the identification of the endpoint that sent the channel data and to protect the integrity of the channel data.

Description of the attack method:

There are various methods of attack for authenticated classical channel. For example, an attacker could install a packet sniffer on the classical communication channel between the QKD transmitter and the QKD receiver, and then impersonate the QKD transmitter and QKD receiver to eavesdrop on or tamper with the communication content.

Assessment:

The penetration tests shall demonstrate the TOE counters such attacks. For [PP-EAL4] and [PP-EAL2], the protocol to be implemented in the authenticated classical channel is not specified, and the TOE developer decides the protocol. The evaluator shall search for vulnerabilities in the implemented protocols and conduct penetration tests in accordance with [CEM].

9.3.5. Random number generators

Description of assumption:

It assumed that the random number generator provides random bits that meets the defined quality metric.

Description of the attack method:

- An attacker reads raw random bits from internal components of the QKD receiver or the QKD transmitter.
- An attacker modifies raw random bits in internal components of the QKD receiver or the QKD transmitter.
- An attacker reads digitized random numbers from internal components of the QKD receiver or the QKD transmitter.
- An attacker modifies digitized random numbers in internal components of the QKD receiver or the QKD transmitter.

An attacker uses these adverse actions to compromise the QKD receiver or the QKD transmitter.

Assessment:

The QKD receiver and the QKD transmitter are physically protected due to the assumption of each PP. i.e.

A.SecureOp of [PP-EAL4] or A.PHYSICAL of [PP-EAL2]. So an attacker cannot access internal components of the QKD receiver or the QKD transmitter directly.
If above assumptions are achieved, no potential vulnerabilities exist in above point of view.

10. Functional Tests

This section describes the functional tests that developer shall conduct. The functional tests described here have been deemed necessary by experts in the context of the related SFRs for testing upon the TOE. Additionally, functional tests that are required to be conducted in the process of identifying vulnerabilities in Section 9 are outlined in Subsection 10.9. The evaluator examines that the developer conducted the functional tests in this section according to the evaluation activity in Subsection 8.5.

10.1. FCS_QKD.1

Test 1:

This test demonstrates the establishment of identical QKD keys according to the QKD protocol (FCS_QKD.1.1).

Step 1 The tester shall start QKD session.

Step 2 The tester shall continue the QKD session until 200,000 bits or more QKD keys are established.

Step 3 The tester shall retrieve the QKD keys from the QKD transmitter and the QKD receiver.

Step 4 The tester shall compare the QKD keys retrieved from the QKD transmitter and the QKD keys retrieved from the QKD receiver.

Step 5 If the QKD keys match, the test result is PASS, otherwise the test result is FAIL.

Test 2:

The developer shall demonstrate that post-processing consistent with the functional specification is correctly implemented based on each post-processing algorithm assigned to FCS_QKD.1.4 and each privacy amplification algorithm assigned to FCS_QKD.1.6. Actual functional tests depend on the assignment of the SFR and description of the functional specification. The functional tests might be as follows in a typical case where the raw key are sequentially converted to sifted key through a sifting scheme, to reconciled key through error correction scheme, and then to QKD keys through a privacy amplification scheme.

Step 1. Sifting scheme

The correctness of the implementation of this scheme shall be demonstrated by verifying that bits whose basis does not match are removed after sifting for raw key.

Step 2. Error correction scheme

The correctness of the implementation of this scheme shall be demonstrated by verifying that the erroneous bits are corrected after error correction under situation where the communication errors of basis-matched raw key occur. Note that too many errors may cause the QKD session to be re-executed based on FCS_QKD.1.2, making it impossible to observe error correction behaviour. The error correction scheme may be followed by a process of the consistency check of the pair of the reconciled key. In such a case, the correctness of the implementation of the consistency check shall be demonstrated by verifying that the same hash value is obtained when the same string is input into the implemented hash function and that different hash values are obtained when different strings are input.

Step 3. Privacy amplification scheme

The correctness of the implementation of this scheme shall be demonstrated by verifying that the reconciled key is shortened by privacy amplification according to the privacy amplification ratio that is deduced based on FCS_QKD.1.5.

Test 3: This test demonstrates the function of repeated executions of key establishment by the QKD protocol and the behaviour of the attempt counter for all attempts for key establishment (FCS_QKD.1.2).

Step 1 The tester shall query the key establishment attempt counter and record the value.

Step 2 The tester shall start QKD session and attempt the key establishment multiple times.

Step 3 The tester shall stop QKD session after the key is established.

Step 4 The tester ensure that value of the attempt counter is incremented by the count of attempts.

Step 5 If the value of the attempt counter is incorrect, the test result is FAIL, otherwise proceed to the next step.

Step 6 If the TSF supports the function of automatic repeated executions of key establishment if the QKD protocol is aborted or sufficient key length is not established, the tester shall also perform the following steps:

Step 7 The tester shall configure conditions in which repeated executions of key establishment are required. In order to support this step, the developer may provide a test function dedicated to fulfilling the condition. For example, the TOE forces to abort the first key establishment attempt.

Step 8 The tester shall start QKD session and attempt the key establishment.

Step 9 The tester shall stop QKD session after the key is established.

Step 10 The tester ensure that value of the attempt counter is added by automatically repetitions count.

Step 11 The tester shall iterate step 7 to 10 until all the conditions in which repeated executions of key establishment are required are covered.

Step 12 For all iterations, if the value of the attempt counter is correct, the test result is PASS, otherwise the test result is FAIL.

Test 4:

This test demonstrates the behaviour of the FCS_QKD.1.2 functionality when the threshold of the attempt counter is exceeded.

If the TOE does not support the management function that modifies threshold of the attempt counter, in order to support this test, the developer shall provide a test-dedicated interface to force any value to the threshold of the attempt counter.

Step 1 The tester shall modify the threshold of the attempt counter to lower value. If multiple QKD key establishments are performed in parallel within one QKD session, the threshold value shall be set such that the attempt counter reaches the threshold plus one when below all QKD key establishments fail.

Step 2 The tester shall start QKD session.

Step 3 The tester shall force to fail key establishment and repeat key establishment until the attempt counter reaches the threshold plus one.

Step 4 The tester shall ensure that the QKD protocol execution is no longer allowed.

Step 5 If the QKD protocol execution is denied, the test result is PASS, otherwise the test result is FAIL.

10.2. FPT_ITQ.1

Test 1: This test demonstrates the behaviour of the authenticated classical channel (FPT_ITQ.1).

Step 1 The tester shall start QKD session.

Step 2 The tester shall modify information to be protected on the classical channel during the QKD session.

Step 3 The tester shall ensure the TSF detects the modification and takes action after detection to be implemented.

Step 4 The tester shall iterate step 1 to 3 until all information to be protected, such as bases information and error correction information are covered. For example, exchanging bases and exchanging error correction information. If the information is transmitted in both directions, the integrity check of the QKD transmitter and the integrity check of the QKD receiver shall be tested respectively.

Step 5 For all iterations, if the action is consistent to the functional specification or the TOE design, the test result is PASS, otherwise the test result is FAIL.

10.3. FPT_EMS.1

10.3.1.Overview of functional tests of assumption families

If the assumption in the security proof is described with the values of realistic characteristics, the corresponding values of the testable parameters/characteristics shall be demonstrated by the functional tests. The functional tests corresponding to the assumption family are shown in Table 10-1. The developer may use one or more tests shown in Clauses 7, 8 and 9 in [ISO/IEC 23837-2] for the above purpose. The threshold values (expected values) of these tests in developer's test plan document shall be consistent with the functional specification or the TOE design and with values of realistic characteristics of the assumptions in the security proof.

Table 10-1 Correspondence of assumption families and functional tests

QKD transmitter	
Assumption family	**Phase randomization**
Functional tests	Subsubsection 10.3.2.1, [ISO/IEC 23837-2] 7.7
Testable parameter(s)	The difference between the probability distribution of the measured intensity after passing through an asymmetric Mach-Zehnder interferometer and the theoretical probability distribution d_{phase}
Assumption family	**Photon statistics and intensity**
Functional tests	Subsubsection 10.3.2.2, [ISO/IEC 23837-2] 7.2
Testable parameter(s)	Deviation between the measured value of k-th order correlation function and the theoretically expected value $\Delta^{(k)}$
Assumption family	**Degrees of freedom**
Functional tests	Subsubsection 10.3.2.3, [ISO/IEC 23837-2] 7.6
Testable parameter(s)	The maximum absolute value of the difference in time of arrival, spectrum, azimuthal angle and ellipticity of the polarization between two encoded states $\delta_{\max,t}, \delta_{\max,\lambda}, \delta_{\max,\theta}, \delta_{\max,\varepsilon}$
Assumption family	**Security and cryptographic boundaries**
Functional tests	Subsubsection 10.3.2.4, [ISO/IEC 23837-2] 7.8, 7.9, 7.10
Testable parameter(s)	The minimum value of isolation measured under different conditions (input power and wavelength) in the isolation component being tested p_{minIso} The maximum values of injection power for CW light and pulsed light indicating exceptional events $p_{\mathrm{maxCWcont}}, p_{\mathrm{maxPulse}}$ The maximum values of deviations in intensity, spectrum, and phase induced by laser injection

	$d_{\mathrm{maxInt}}, d_{\mathrm{maxSpec}}, d_{\mathrm{maxPhase}}$
Assumption family	**Accuracy of the encoding**
Functional tests	Subsubsection 10.3.2.5, [ISO/IEC 23837-2] 7.5
Testable parameter(s)	The minimum fidelity between the measured density matrix and the ideal density matrix assumed in the QKD protocol F
Assumption family	**Independence of adjacent pulses**
Functional tests	Subsubsection 10.3.2.6, [ISO/IEC 23837-2] 7.4
Testable parameter(s)	Deviation of the average intensity of light pulses prepared with the same intensity setting $\delta_{k,j,i}$
QKD receiver	
Assumption family	**Detection efficiency**
Functional tests	Subsubsection 10.3.3.1, [ISO/IEC 23837-2] 8.2
Testable parameter(s)	The maximum value of detection probability mismatch between two encoded states $\sigma_{\max}$
Assumption family	**Degrees of freedom**
Functional tests	Subsubsection 10.3.3.2, [ISO/IEC 23837-2] 8.2
Testable parameter(s)	The maximum value of detection probability mismatch between two encoded states $\sigma_{\max}$
Assumption family	**Security boundary on optical channel**
Functional tests	Subsubsection 10.3.3.3, [ISO/IEC 23837-2] 8.3, 8.4, 8.5
Testable parameter(s)	The maximum value of back-flash probability P_{maxBF} The minimum value of isolation measured under different conditions (input power and wavelength) in the isolation component being tested p_{minIso} The maximum values of injection power for CW light and pulsed light indicating exceptional events $p_{\mathrm{maxCWcont}}, p_{\mathrm{maxPulse}}$
Assumption family	**Accuracy of the demodulation**
Functional tests	Subsubsection 10.3.3.4
Testable parameter(s)	The maximum quantum bit error rate in each basis $QBER_{\max}$
Assumption family	**Single-photon sensitivity**
Functional tests	Subsubsection 10.3.3.5, [ISO/IEC 23837-2] 8.6
Testable parameter(s)	The ratio of photon detection efficiency with and without blind light κ
Assumption family	**Recovery or dead time**
Functional tests	Subsubsection10.3.3.6, [ISO/IEC 23837-2] 8.7
Testable parameter(s)	None. Verify that no detection signals are output during dead time.

Whole of the TOE	
Assumption family	**Calibration**
Functional tests	Subsubsection 10.3.4.1 , [ISO/IEC 23837-2] 9.2
Testable parameter(s)	The maximum value of basis bias and bit bias when a tampering device inserted into a quantum channel causes temporal shifts in the detection efficiency $b_{\max 0}, b_{\max 1}, B_{\max}$
Assumption family	**Stabilities of the light source and the photon detector**
Functional tests	**The light source:** Subsubsection 10.3.4.2 , [ISO/IEC 23837-2] 7.3
	The photon detector: None. The TOE user shall periodically inspect the photon detector for performance degradation.
Testable parameter(s)	**The light source:** Mean photon number at each intensity
	The photon detector: None
Assumption family	**Robustness against provoked damage**
Functional tests	Subsubsection 10.3.4.3 , [ISO/IEC 23837-2] 8.9
Testable parameter(s)	The maximum value of the mismatch in detection efficiency of each photon detector after injecting light into the receiver $\sigma_{\max \mathrm{Mis}}$
Assumption family	**Authenticated classical channel**
Functional tests	Subsection 10.2
Testable parameter(s)	None
Assumption family	**Random number generator**
Functional tests	Subsection 10.6
Testable parameter(s)	None

10.3.2.Assumption families of the QKD transmitter

10.3.2.1. Phase randomization

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 7.7 EA to test the uniform distribution of the global phase of optical pulses.

When conducting this test, unattenuated light may be used.

10.3.2.2. Photon statistics and intensity

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 7.2 EA to test the photon-number distribution of optical pulses.

10.3.2.3. Degrees of freedom

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 7.6 EA to test the indistinguishability of encoded states.

When conducting this test, unattenuated light may be used.

10.3.2.4. Security and cryptographic boundaries

The test of this assumption family may be conducted according to following tests.

- [ISO/IEC238737-2] 7.8 EA to test the degree of optical isolation of the TX module
  The developer only needs to measure the characteristics of the isolator.
- [ISO/IEC238737-2] 7.9 the sensitivity of the injected light monitor in the TX module
  The developer only needs to measure the characteristics of the light injection monitor.
- [ISO/IEC23837-2] 7.10 the robustness of the TX module against laser injection
  The developer measures that the characteristics of the transmitted light do not change even when light is injected into the QKD transmitter. When conducting this test, unattenuated light may be used.

10.3.2.5. Accuracy of the encoding

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 7.5 EA to test the accuracy of state encoding.
When conducting this test, unattenuated light may be used.

Note 1
It is necessary to estimate the density matrices of the photon states at the transmitter output to perform [ISO/IEC 23837-2] 7.5 EA. This application note provides a method for the density matrix estimation.

This method requires transmitting optical pulses with a fixed quantum state from the QKD transmitter. The developer shall provide a function dedicated for this transmission.
This method also requires a reference receiver that outputs correctly for inputs in the correct state. The receiver measures the states in X-, Y-, and Z- basis. The tester shall prepare such a receiver.

Method (state tomography):
In the following, the TOE is assumed to use X basis and Z basis to perform the BB84 protocol. The tester selects one of the four states Φ_i $(\Phi \in \{X, Z\}, i = \{0,1\})$ and outputs it from the transmitter.
The transmitted light is measured in the X,Y,Z basis using a reference receiver to obtain the detection rate $P(\Psi_j|\Phi_i), (\Psi \in \{X, Y, Z\}, j = \{0,1\}))$.
Then, the density matrix on the basis Φ can be calculated. For example, the density matrix on X basis is reconstructed as a linear expansion with Pauli matrices $\hat{\sigma_i}'$s as

$$\rho = \frac{1}{2}\sum_{i=0}^{3} \frac{S_i}{S_0}\hat{\sigma_i},$$

where

$$S_0 = 2n_0$$
$$S_1 = 2(n_1 - n_0)$$
$$S_2 = 2(n_2 - n_0)$$
$$S_3 = 2(n_3 - n_0)$$

$$n_0 = \frac{N}{2}(\langle 0|\rho_{\mathrm{X}}|0\rangle + \langle 1|\rho_{\mathrm{X}}|1\rangle) = P(Z_0|\mathrm{X}_0) + P(Z_0|\mathrm{X}_1) + P(Z_1|X_0) + P(Z_1|X_1)$$

$$n_1 = N\langle 0|\rho_{\mathrm{X}}|0\rangle = P(Z_0|X_0) + P(Z_0|X_1)$$
$$n_2 = N\langle \mathrm{X}_1|\rho_{\mathrm{X}}|X_1\rangle = P(X_1|X_0) + P(X_1|X_1)$$
$$n_3 = N\langle Y_1|\rho_X|Y_1\rangle = P(Y_1|X_0) + P(Y_1|X_1)$$

The accuracy of the encoding is characterized by the fidelity between the intended state and the emitted state.

References

Daniel F. V. James, *et al*, Physical review A, **64**, 052312 (2001)

Weiyang Zhang, Yu Kadosawa, Akihisa Tomita, Kazuhisa Ogawa, and Atsushi Okamoto, "State preparation robust to modulation signal degradation by use of a dual parallel modulator for high-speed BB84 quantum key distribution systems", Opt. Express **28**, 13965-13977 (2020).

10.3.2.6. Independence of adjacent pulses

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 7.4 EA to test the independence of the intensities of optical pulses.

When conducting this test, unattenuated light may be used.

10.3.3.Assumption families of the QKD receiver

10.3.3.1. Detection efficiency

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 8.2 EA to test the consistency of detection probability in the RX module.

When conducting this test, the developer also measures wavelength dependency and time dependency.

10.3.3.2. Degrees of freedom

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 8.2 EA to test the consistency of detection probability in the RX module. In addition, the developer conducts the functional test described in Subsubsection10.9.2.2. However, it is not necessary to carry out each test separately, and it is acceptable to standardize them as long as the same information can be obtained.

10.3.3.3. Security boundary on optical channel

The test of this assumption family may be conducted according to following tests.

- [ISO/IEC 23837-2] 8.3 EA to test information leakage of back-flashes from the RX module
- [ISO/IEC 23837-2] 8.4 EA to test the degree of optical isolation of the RX module
- [ISO/IEC 23837-2] 8.5 EA to test the sensitivity of the injected light monitor in the RX module

10.3.3.4. Accuracy of the demodulation

This test demonstrates the accuracy of the demodulation in the receiver.

Input the signal transmitted from a transmitter that has passed functional testing 10.3.2.5. (Accuracy of the encoding) or its replacement into the receiver under test. From the measurement results of the receiver, calculate the quantum bit error rates $QBER_X$ and $QBER_Z$ for each basis. The larger of the two is designated as $QBER_{max}$.

10.3.3.5. Single-photon sensitivity

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 8.6 EA to test the robustness of the RX module against bright light blinding. In addition, the developer conducts the functional test described in Subsubsection 10.9.2.1. However, it is not necessary to carry out each test separately, and it is acceptable to standardize them as long as the same information can be obtained.

10.3.3.6. Recovery or dead time

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 8.7 EA to test the appropriateness of dead time settings of SPDs.

10.3.4.Assumption families of the whole of the TOE

10.3.4.1. Calibration

The test of this assumption family may be conducted according to [ISO/IEC 23837-2] 9.2 EA to test the inducibility of detection probability mismatch.

10.3.4.2. Stabilities of the light source and the photon detector

The light source

The test of the light source of this assumption family may be conducted according to [ISO/IEC 23837-2] 7.3 EA to test the mean photon number and stability of optical pulses.
When conducting this test, unattenuated light may be used.

The photon detector

There are no tests for the photon detector for this assumption family. The TOE user shall periodically inspect the photon detector to ensure that there is no deterioration in their performance to guarantee their stability.

10.3.4.3. Robustness against provoked damage

This assumption family does not require functional tests. The developer shall provide the guidance document for consumers to regularly maintain and inspect the QKD receiver and the QKD transmitter. Contents of the guidance are evaluated for appropriateness by the work units shown in section 3.2.

10.3.5.Functional tests for assumptions other than assumption families

Characteristics that are not subject of tests of assumption families described above, but that corresponds to the assumptions in the security proof, shall be also demonstrated by the functional testes. The developer may use one or more tests shown in Clause 7, 8 and 9 in [ISO/IEC 23837-2] for the above purpose. The threshold value (expected value) of these tests in developer's test plan document shall be consistent with the functional specification or the TOE design.

10.4. FPT_PHP.3

This test demonstrates the behaviour of the light injection monitor (FPT_PHP.3).
If a light injection monitor or filter is implemented for these SFRs, its actual characteristics shall be demonstrated by the functional test. If the light injection monitor is implemented, the test demonstrates power of injected light detected by the light injection monitor. If the sensitivity of the monitor changes with parameters such as wavelength, the sensitivity shall be demonstrated by changing the parameters. And more, the test demonstrates that the TSF automatically responds to monitor detection consistent with its functional specification. If the filter is implemented, such as a wavelength filter, the test demonstrates attenuation characteristics of the filter. In this case, filtering out

itself is automatic responses of the TSF, so no additional testing that demonstrates automatic response is required. The developer may use one or more tests shown in Clause 7, 8 and 9 in [ISO/IEC 23837-2] for the above purpose. The threshold value (expected value) of these tests in developer's test plan document shall be consistent with the functional specification or the TOE design.

10.5. FPT_FLS.1

This SFR requires that a secure state is preserved when some types of failures occur. In [PP-EAL4] case, state control is also required, but in [PP-EAL2] case, no state control is required. The common test scenario may be as follows, but if the TOE claims [PP-EAL4] compliant, the test scenario should be more refined in the developer's test plan document in order to demonstrate the state control.

Test 1:

This test demonstrates the FPT_FLS.1 functionality of maintaining the secure state when the failures occur.

Step 1 The tester shall reproduce the situation in which each failure occurs.

Step 2 The tester shall verify that the defined secure state in each PP is preserved.

It is expected that the developer shall provide test tools (e.g. debugger) or dedicated test interfaces that can access TSF data in order to reproduce the failure situation such as authentication failure of the classical channel. Depending on assignment of self-test SFR, the failure situation cannot be reproduced even if test tools or test interfaces are provided. For example, it is so difficult to reproduce the failure of the physical random number generator. It is acceptable to exclude such failures from this test. Assurance for such self-test function and secure state preservation function are provided only by document examination.

10.6. FCS_RNG.1

The developer shall test the random number generator according to the random number generator standard associated with the SFR. For example, the standard may be AIS31 or SP800-90B.

10.7. FCS_COP.1 and FCS_CKM.6

Depending on certification scheme, specific algorithm verification program may be required for crypto algorithms specified in FCS component. The developer should contact each certification body for the required algorithm testing.

The test for the destruction of the cryptographic keys specified in FCS_CKM.6 can be tested with reference to the supporting documents [DSCSD] or [HCDSD].

10.8. Other SFR in the Functional Package

Functional tests of the identification and authentication specified in FIA_UIA_EXT.1 and the secure channel protocol specified in FTP_ITC.1 shall be tested with reference to the supporting document [NDSD].

The use of trusted channels specified in FDP_ETC_EXT.2.1 is included in the tests for FTP_ITC.1.

Test 1:

This test demonstrates the FDP_ETC_EXT.2 functionality that is exporting QKD keys and that the exported keys are not re-used.

Step 1 The tester shall start QKD session.

Step 2 The tester shall ensure that the TOE automatically exports the QKD key to the key manager during the QKD session. The export shall be done only once and it shall be ensured that the exported QKD keys are not re-used.

10.9. Functional tests related with vulnerability analysis

10.9.1.QKD transmitter

At the moment, functional tests related with vulnerability analysis on the QKD transmitter have not been identified yet.

10.9.2.QKD receiver

10.9.2.1. Single-photon sensitivity

Purpose of test

The test of this assumption family demonstrates the countermeasure against the bright illumination attack.

Target of test

A set of receivers used for the TOE.

Test equipment and configuration

- QKD transmitter
  TOE QKD transmitter or its replacement.
- Fiber spool
  Fiber spool for the distance envisaged, e.g. 50 km
- Optical coupler
  For 1550nm
- Laser for blinding
  Light source capable of outputting pulsed and CW light in the 600-2000 nm range.

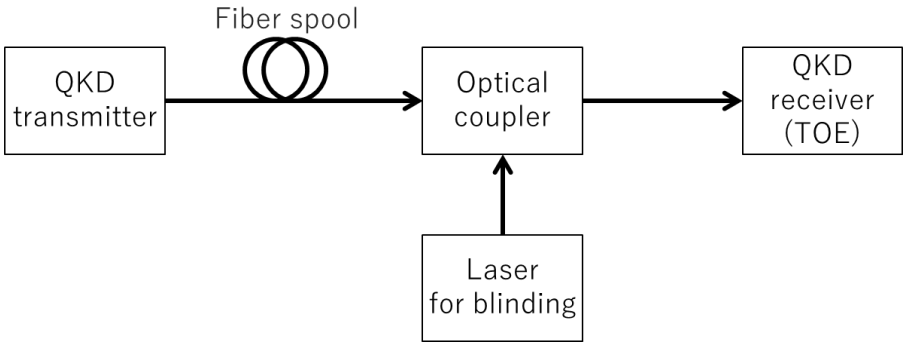

Figure 10-1 Test for bright illumination attack

Test method

Step 1 Establish a QKD link.

Step 2 Record the difference $R_i(0)$ between the photon detection rate when no Blind light is input to the receiver and the photon detection rate when the signal light from the transmitter is input to the receiver. i indicates the index of the multiple implemented photon detectors, typically $i = 0,1$ or $i = 0,1,2,3$. The photon detection rate refers to the number of photons detected per unit of time. The input signal light shall be of the intensity and state used in normal key generation.

Step 3 The Blind light source is set to CW (continuous light) mode. Set the Blind light intensity and wavelength to minimum values according to Table 10-2. The intensity is gradually increased and the difference R_i between the photon detection rate with signal light input and without signal light input at each intensity is recorded. The upper limit of the Blind light intensity shall be limited to the maximum value that is not detected by the light injection monitor implemented inside the QKD receiver.

Step 4 Gradually increase the wavelength and repeat step 3.

Step 5 The Blind light source is set to pulse oscillation mode. Set the Blind light intensity, wavelength, repetition frequency, pulse width and pulse incidence timing to minimum values according to Table 10-2. The intensity is gradually increased and the difference R_i between the photon detection rate with signal light input and without signal light input at each intensity is recorded.

Step 6 Gradually increase the wavelength and repeat step 5.

Step 7 Gradually increase the repetition frequency and repeat steps 5 and 6.

Step 8 Gradually increase the pulse width and repeat steps 5~7.

Step 9 Gradually increase the pulse injection timing and repeat steps 5-8.

Step 10 Repeat 1-9 in all modes permitted by the administrator.

The various parameters to be varied in steps 6-9 do not necessarily have to be varied in this order and the order may be interchanged.

Table 10-2 Parameter of blind light

Items	Description
Wavelength	
Minimum	600nm
Maximum	2000nm
Step	10nm
Notes	Evaluation of the transmission characteristics of the QKD receiver (e.g. filters) in advance, for wavelengths with losses <30 dB.
Pulse width	
Minimum	1/10th of the inverse of the APD bandwidth. Ex: 0.1ns
Maximum	Up to the reciprocal of the clock frequency Ex: 0.5ns
Step	At least 3 points per digit (e.g. 1, 2, 5 times).
Notes	"Up to the reciprocal of the clock frequency" is not a problem.
Repetition frequency	

Minimum	--
Maximum	Clock frequency (Ex. 1GHz)
Step	--
Notes	--
Intensity	
Minimum	Minimum pulse energy of received light assumed by the device (distance dependent) Ex: 0.5 Photon/pulse -10dB(Assume 50 km) =0.05 Photon/pulse = 6.4e-21[J]/pulse
Maximum	Up to the light injection monitor detection threshold (see also Notes).or Until they are failed.
Step	At least 3 points per digit (e.g. 1, 2, 5 times).
Notes	If the light injection monitor bandwidth is wide, the clock frequency is taken into account and evaluated up to the effective light injection monitor detection threshold.
Timing of incident	
Minimum	0ns（basis）
Maximum	clock cycle Ex: 1ns
Step	1/10th of a clock Ex: 0.1ns
Notes	--

Acceptance criteria

Indicator:

κ :The ratio of R_i to $R_i(0)$

$$\kappa = R_i \,/\, R_i(0)$$

For κ calculated from all R_i obtained in steps 3~10, $0.95 < \kappa < 1.05$

10.9.2.2. Degrees of freedom

Purpose of test

The test of this assumption family demonstrates the countermeasure against the time shift attack.

Target of test

A set of receivers used for TOE

Test equipment and configuration

- QKD transmitter
  TOE QKD transmitter or its replacement.
  The configuration for this test is shown in Figure 10-2.

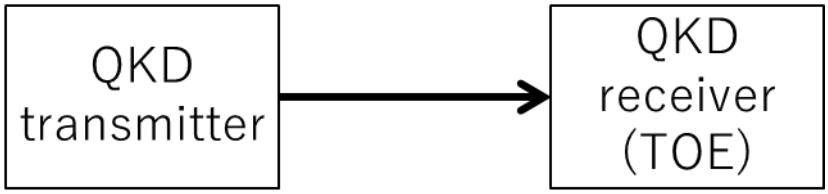

Figure 10-2 Test for time shift attack

Test method

Step 1 Establish a QKD link.

Step 2 Record the photon detection rates, denoted as $R_{\mathrm{X0}}(t), R_{\mathrm{X1}}(t), R_{\mathrm{Z0}}(t)$, and $R_{\mathrm{Z1}}(t)$, for each basis (X, Z) and each bit value (0, 1) while varying the timing t of the gate pulse in the receiver or the incidence timing of the signal light. Here, the photon detection rate refers to the probability of photon detection per round. Both dark counts and photon detections are included in the photon detection rates without distinction.

Step 3 Calculate the ratio $r_{\mathrm{X}}(t), r_{\mathrm{Z}}(t)$ of the two photon detection rates in the same basis at each timing t.

$$r_{\mathrm{X}}(t) = \frac{R_{\mathrm{X1}}(t)}{R_{\mathrm{X0}}(t)}$$

$$r_{\mathrm{Z}}(t) = \frac{R_{\mathrm{Z1}}(t)}{R_{\mathrm{Z0}}(t)}$$

Step 4 Let the maximum and minimum values obtained from the two bases and all timings t above be r_{max} and r_{min}, respectively

Table 10-3 Parameters of the time shift attack

Items	Description
Gate timing	
Minimum	0ps (basis)
Maximum	Clock frequency Ex. 800ps
Step	Below 1/10 of clock frequency Ex. 25ps
Notes	Evaluate the satellites (and possibly have them pick up the satellites).
Intensity	
Minimum	0
Maximum	Strength at normal operation or just before Bob side strength light injection monitor
Step	None (only the above two)
Notes	The evaluation of the minimum is based on the smallest difference in dark counts.

Acceptance criteria

Pass at $r_{max} < 1.2$ and $r_{min} > 0.8$.

11. Penetration Tests

This section outlines penetration tests to exploit vulnerabilities in the TOE for the assumption families. These penetration tests are derived from attacks which have been known in literature.

11.1. QKD transmitter

11.1.1.Exploitation of imperfect phase randomization

Test 1: Source attacks with phase information

This test requires an auxiliary laser source that emits pulses in the same mode as the pulses from the QKD transmitter.

This test composed of the following two phases:

Phase 1: Using a train of pulses emitted from the QKD transmitter, the tester adjusts the phase of the auxiliary laser source via injection locking or a feedback loop with a relative phase measurement.

Phase 2: Using the auxiliary laser source, the tester carries out attacks on the rest of the pulses from the QKD transmitter as described in the References.

After Phase 1, the tester should verify whether any correlations are made between the phases of the pulses from the QKD transmitter and those from the auxiliary laser source. If no correlations are observed, the test result is PASS with no need for proceeding to Phase 2.

References:

H. -K. Lo and J. Preskill, “Phase randomization improves the security of quantum key distribution”, arXiv:quant-ph/0504209.

H. -K. Lo and J. Preskill, “Security of quantum key distribution using weak coherent states with nonrandom phases”, Quant. Inf. Comput. **8** 431-458 (2007).

Y. -L. Tang *et al.*, “S Source attack of decoy-state quantum key distribution using phase information”, Phys. Rev. A **88**, 022308 (2013).

11.1.2.Exploitation of degrees of freedom not intentionally used

Test 1: Intercept-resend attack with side information

(*Tentative*)

The following procedure is used when the TOE uses time-bin encoding and the pulses from the QKD transmitter nominally have V polarization.

Step 1. The tester shall place a polarization filter that only passes H polarization followed by a wave plate to change the polarization to V.

Step 2. The tester shall distinguish the bit value in the Z-basis state through photon detection using the same apparatus as the QKD receiver.

Step 3. When the detection has been successful. the tester shall prepare a weak laser pulse in the Z-basis with the observed bit value and send it to the QKD receiver. The tester shall record the observed bit value. When the detection has failed, the tester shall send no light to the QKD receiver.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

Test 2: Side-channel filtering attack:

(*Tentative*)

The following procedure is used when the TOE uses time-bin encoding.

Step 1. The tester shall prepare a transmission filter (for polarization/temporal modes/spectral modes) that has different transmissivities for the two Z-basis states with bit values 0 and 1.

Step 2. The tester shall insert the filter in the quantum channel between the QKD transmitter and the QKD receiver.

Step 3. The tester shall record the bit value with the higher transmissivity.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.1.3.Exploitation of invalid security and cryptographic boundaries

Test 1: Trojan horse attack

The following procedure is used when the TOE decoy-state BB84 protocol generates sifted keys from the Z-basis only and the X-basis is only used to monitor eavesdropping.

Step 1. The tester prepares a light source (probe light source, henceforth) for probing the internal state of the QKD transmitter to learn its choices of the basis and the bit value.

Step 2. For each of the M rounds comprising a QKD session, the tester shall inject light from the probe light source to the QKD transmitter and make a measurement on the reflected light to obtain an outcome (probe outcome, henceforth). Then, depending on the outcome, the tester shall choose one from the following options. (Some options may not be available for a high-speed TOE).

i) Block the encoded pulse(s) the QKD transmitter sends out for the round and send a brighter encoded pulse(s) instead to the QKD receiver.

Note that this option is effective if, from the probe outcome, it is highly probable that the QKD transmitter chose the Z-basis and the encoded bit value can be guessed with high confidence.

ii) Measure the encoded pulse(s) the QKD transmitter sends out for the round on the Z basis. If the bit value was successfully determined, send the corresponding bright encoded pulse on the Z-basis to the QKD receiver. Otherwise, sends no light to the QKD receiver.

Note that this option is effective if, from the probe outcome, it is highly probable that the QKD transmitter chose the Z-basis.

iii) Let the encoded pulse(s) from the QKD transmitter pass through for the round.

Note that choosing this option over the next option is effective if the encoded bit value can be guessed with high confidence from the probe outcome.

iv) Block the encoded pulse(s) the QKD transmitter sends out for the round and sends no light to the QKD receiver.

Step 3. After following the procedure described in one of the above options, the tester shall determine a bit value

which is more likely and record it.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.1.4.Exploitation of inaccuracy in encoding

Test 1: Intercept-resend attack on the monitoring basis

(*Tentative*)

The following procedure is used when the TOE decoy state BB84 protocol generates sifted keys from the Z-basis only and the X-basis is only used to monitor eavesdropping.

The tester shall in advance determine the states of the optical pulse(s) emitted from the QKD transmitter for the bit values 0 and 1 in the X-basis. The tester shall then construct the averaged density operator of a single photon in the encoded degree of freedom to determine the two orthogonal modes that diagonalize the operator. The basis formed by the two modes is called X'-basis henceforth.

The tester shall choose the rate t at which the attack is conducted. During a QKD session consisting of M rounds, the tester shall select Mt rounds and carry out the following attacks for each round.

Step 1. The tester shall measure the pulse(s) from the QKD transmitter on the X'-basis to distinguish the two orthogonal modes.

Step 2. If the measurement at Step 1 has succeeded in the distinction, the tester shall prepare a bright pulse in the corresponding mode and send it to the QKD receiver. If a sifted key bit was produced in the round, the tester shall guess the bit value from the measurement outcome and record it.

Step 3. If the measurement at Step 1 has failed, the tester sends the vacuum to the QKD receiver.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.2. QKD receiver

11.2.1.Exploitation of detection efficiency mismatch for different degrees of freedom

Test: {time, polarization, wavelength} shift attack

Target of test

A set of the QKD transmitter and the QKD receiver that used in TOE

Test equipment and configuration

- QKD transmitter (TOE)
- QKD receiver (TOE)
-Shifting device; Timing controller for time shift attack, Polarization controller for polarization shift attack, Wavelength controller for wavelength shift attack

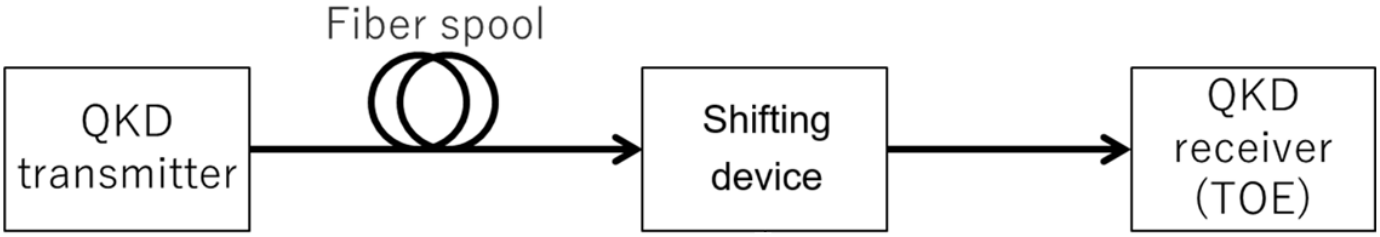

Figure 11-1 Test configuration for {time, polarization, wavelength} shift attack

Test method

The tester chooses a degree of freedom from time, polarization, and wavelength for attack, and inserts the shifting device for the degree of freedom on the optical channel between the transmitter and the receiver.

Step 1. The tester shifts the degree of freedom with an amount of shift.

Step 2. The tester sends bit strings from the transmitter and records the detection events.

Step 3. The tester set the bit value to a fixed value (0 or 1) for the detection events.

Repeat Step 2 and Step 3 for M rounds.

The tester changes the amount of shift and continues step 2 and step 3.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.2.2.Exploitation of invalid security boundary of optical channel

Test 1: Back-flash attack

Target of test

A set of the QKD transmitter and the QKD receiver that used in TOE

Test equipment and configuration

- QKD transmitter (TOE)
- QKD receiver (TOE)
- Optical circulator
- Wavelength division multiplexing (WDM) filter
- Single photon detector (SPD)

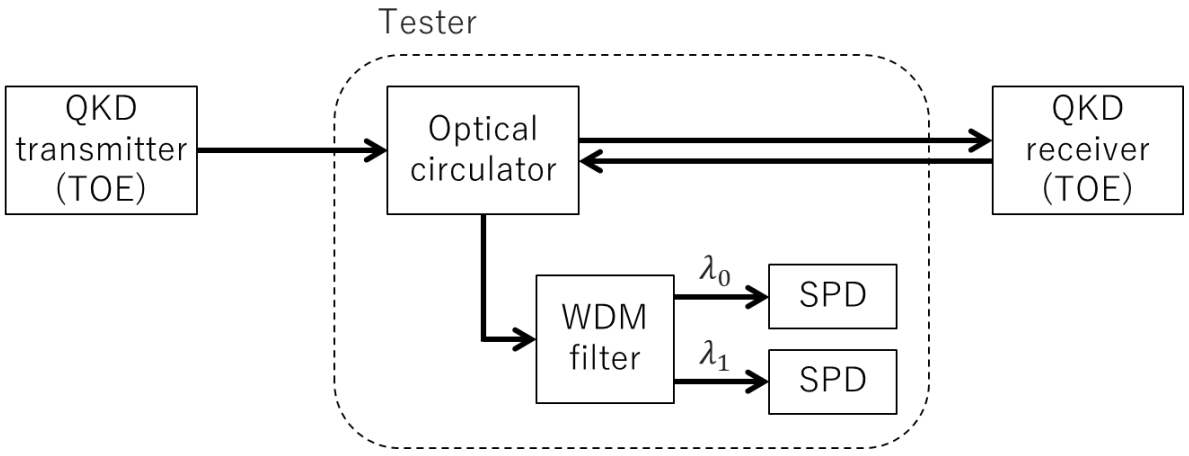

Figure 11-2 Test configuration for Back-flash Attack

Test method

The following is the procedure when the wavelength of backflash light differs for each photon detector. Let $\lambda_0(\lambda_1)$ be the wavelength of the backflash light from the photon detector corresponding to bit value $0(1)$ in the TOE. If degrees of freedom other than wavelength differ, conduct the test by replacing the WDM filter with an element that separates the corresponding degrees of freedom.

Step 1. Establish a QKD link.

Step 2. During an M-round QKD session, the tester uses a circulator, WDM filter, and photon detectors to detect return light from the QKD receiver and record the bit value. The WDM filter can separate λ_0 and λ_1, and when a photon is detected by the photon detector connected to the output port corresponding to $\lambda_0(\lambda_1)$ of the WDM filter, a bit value of $0(1)$ is assigned.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

Test 3: Trojan horse attack

The following procedure is used when the TOE decoy-state BB84 protocol generates sifted keys from the Z-basis only and the X-basis is only used to monitor eavesdropping.

Step 1. The tester prepares a light source (probe light source, henceforth) for probing the internal state of the QKD receiver to learn its choices of the basis.

Step 2. For each of the M rounds comprising a QKD session, the tester shall inject light from the probe light source to the QKD receiver and make a measurement on the reflected light to obtain an outcome (probe outcome, henceforth) . Then, depending on the outcome, the tester shall choose one from the following options.

i) Measure the encoded pulse(s) the QKD transmitter sends out for the round on the Z basis. If the bit value was successfully determined, send the corresponding bright encoded pulse on the Z-basis to the QKD receiver. Otherwise, sends no light to the QKD receiver.

Note that this option is effective if, from the probe outcome, it is highly probable that the QKD receiver chose the Z-basis.

ii) Block the encoded pulse(s) the QKD transmitter sends out for the round and sends no light to the QKD receiver.

Step 3. After following the procedure described in one of the above options, the tester shall determine a bit value which is more likely and record it.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.2.3.Exploitation of single photon sensitivity attack

Test 1: Bright illumination attack

Target of test

A set of the QKD transmitter and the QKD receiver that used in TOE

Test equipment and configuration

- QKD transmitter (TOE)
- QKD receiver (TOE)
- QKD receiver (Tester)
- Laser for blind
- Laser for control
- State modulator
- Optical coupler

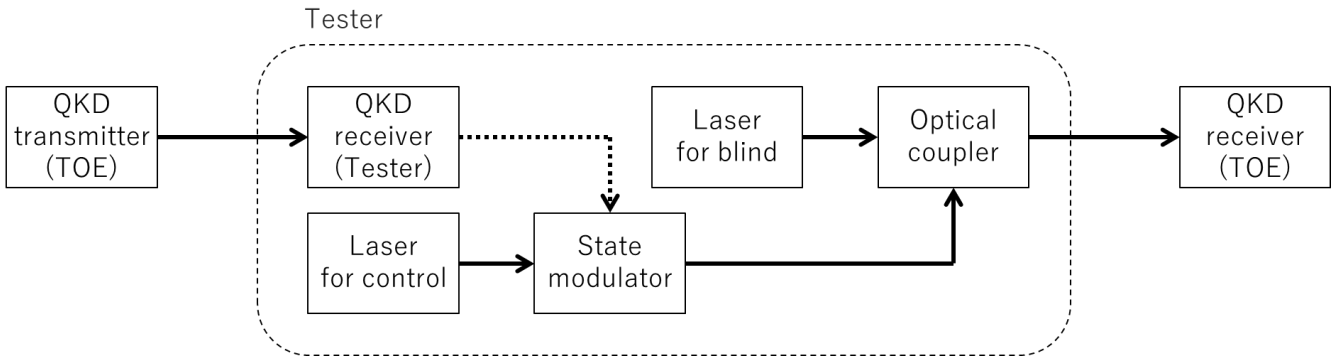

Figure 11-3 Test configuration for Bright Illumination Attack

Test method

The following procedure is used when the TOE decoy BB84decoy state BB84 protocol generates sifted keys from the Z-basis only and the X-basis is only used to monitor eavesdropping.

The tester chooses the rate t at which the bright light attack is conducted. During a QKD session consisting of M rounds, select Mt rounds and carry out the following attacks for each round

Step 1. The tester measures the optical pulses transmitted from the QKD transmitter (TOE) using the QKD receiver (Tester) in the Z-basis and records the bit values.

Step 2. The tester uses the laser for Control and the State modulator to generate optical pulse based on the bit values in Step 1.

Step 3. The tester injects a strong light into the QKD receiver (TOE) using the laser for Blind and the APD in the QKD receiver is changed to linear mode.

Step 4. The tester injects optical pulses for control generated in Step 2 into the APD, which has been changed to

linear mode in Step 3.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.2.4.Exploitation of inaccuracy in demodulation

Test 1: Intercept-resend attack on the monitoring basis

(*Tentative*)

The following procedure is used when the TOE decoy-state BB84 protocol generates sifted keys from the Z-basis only and the X-basis is only used to monitor eavesdropping.

The tester shall in advance determine the two orthogonal modes that are distinguished in the nominal X-basis measurement of the QKD receiver. The basis formed by the two modes is called X'-basis henceforth.

The tester shall choose the rate t at which the attack is carried out. During a QKD session consisting of M rounds, the tester shall select Mt rounds and carry out the following attacks for each round.

Step 1: The tester shall measure the pulse(s) from the QKD transmitter on the X'-basis to distinguish the two orthogonal modes.

Step 2: If the measurement at Step 1 has succeeded in the distinction, the tester shall prepare a bright pulse in the corresponding mode and send it to the QKD receiver. If a sifted key bit was produced in the round, the tester shall guess the bit value from the measurement outcome and record it.

Step 3: If the measurement at Step 1 has failed, the tester sends the vacuum to the QKD receiver.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.2.5.Exploitation of detector dead time

Target of test

A set of the QKD transmitter and the QKD receiver that are used in TOE

Test equipment and configuration

- QKD transmitter (TOE)
- QKD receiver (TOE)
- Laser for blind
- State modulator
-Timing controller
- Optical coupler

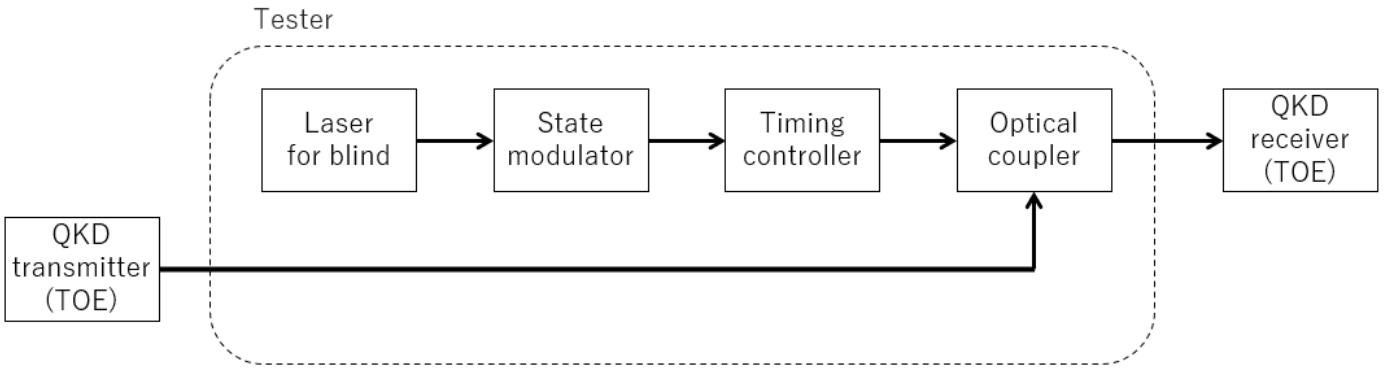

Figure 11-4 Test configuration for Detector Dead-time Attack

Test method

The following procedure is used when the TOE decoy-state BB84 protocol generates sifted keys from the Z-basis only and the X-basis is only used to monitor eavesdropping.

The tester chooses the rate t at which the bright light attack is conducted. During a QKD session consisting of M rounds, select Mt rounds and carry out the following attacks for each round

Step 1. The tester uses the laser for blind and the State modulator to generate a strong optical pulse of the state corresponding to the bit value 0 in Z-basis.

Step 2. The tester injects the optical pulse generated in Step 1 into the QKD receiver (TOE) at the time out of the detection window of the QKD receiver (TOE) to change the APD for bit value 0 in the QKD to linear mode.

Step 3. The tester records the detection events of the QKD receiver (TOE) and register the bit value of the key as 1.

Step 4. The tester uses the laser for blind and the State modulator to generate a strong optical pulse of the state corresponding to the bit value 1 in Z-basis.

Step 5. The tester injects the optical pulse generated in Step 3 into the QKD receiver (TOE) at the time out of the detection window of the QKD receiver (TOE) to change the APD for bit value 1 in the QKD to linear mode.

Step 6. The tester records the detection events of the QKD receiver (TOE) and register the bit value as 0.

Acceptance criteria

The tester shall use the criteria described in Subsection 11.4.

11.3. Whole of the TOE

11.3.1.Exploitation of invalid calibration

The object to be calibrated and the method of calibration vary from device to device. The tester should obtain information on TOE calibration and structure the test accordingly. As a simple example, this SD will treat the case where the polarization is loaded with information and the detection timing is adjusted for each polarization.

Target of test

A set of the QKD transmitter and the QKD receiver that are used in TOE

Test equipment and configuration

- QKD transmitter (TOE)
- QKD receiver (TOE)
- Polarization beam splitter
- Optical delay
- Polarization beam combiner

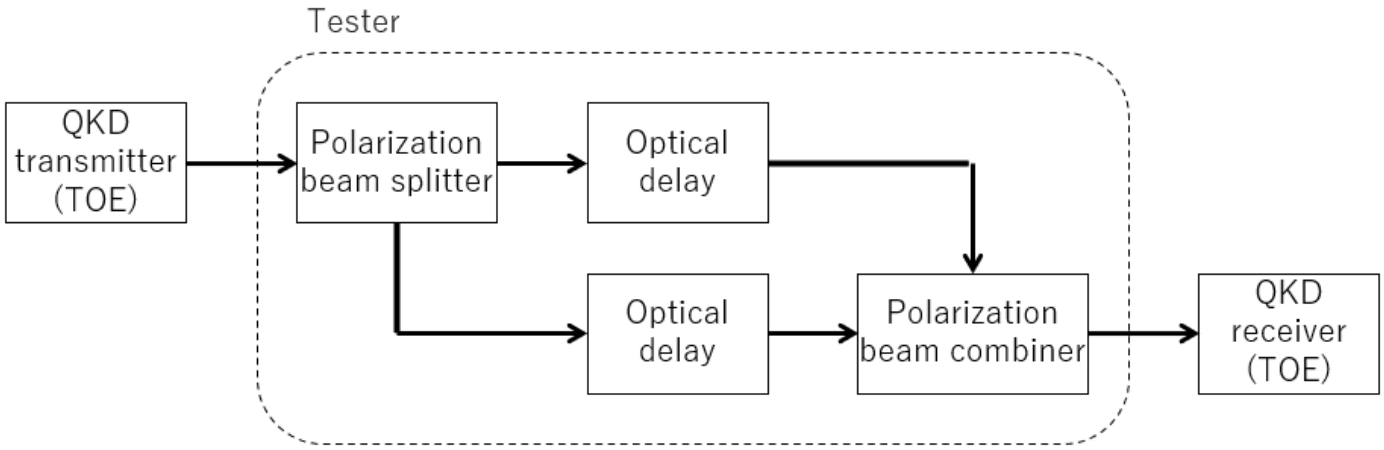

Figure 11-5 Test configuration for Invalid Calibration

Test method

The following procedure is used when the TOE decoy-state BB84 protocol generates sifted keys from the Z-basis only and the X-basis is only used to monitor eavesdropping. The bit value 0 and 1 are assigned to be horizontal and vertical polarization states, respectively. This assignment can be altered.

Step 1 In the calibration process, the tester sets a value for difference of optical delays, which should be determined to maximize the difference of the photon detectors.

Step 2 In the key generation process, the tester sets the optical delays to yield the same delay value. Then, the test proceeds Time-shift Attack which is described in Subsubsection 10.9.2.2 (see also reference below).

Reference

Yi Zhao, Chi-Hang Fred Fung, Bing Qi, Christine Chen, and Hoi-Kwong Lo, "Quantum hacking: Experimental demonstration of time-shift attack against practical quantum-key-distribution systems", Phys. Rev. **A 78,** 042333 (2008)

11.4. Acceptance criteria

Parameters, variables and functions are defined as follows:

α : Pre-defined significance level per single main test

ϵ : Security parameters set in the above QKD session

$\boldsymbol{\kappa}_{\mathrm{sif}}, N$: Sifted key and its lengths generated by the TOE in the above QKD sessions.

$\boldsymbol{\kappa}_{\mathrm{QKD}}, K$: QKD key generated by the TOE in the above QKD session and its length.

f_{PA} : The hush function used by TOE for privacy amplification in the above QKD session

N_1 : Number of rounds in which a bit value is recorded in the attack and a sifted key bit is generated by the TOE.

$\boldsymbol{\kappa}_{1,\mathrm{att}}$: The N_1-bit string consisting of the bit values recorded in the above N_1 rounds.

N_{EC} : Number of bits transmitted by the TOE for error correction in the above QKD session.

wt($\boldsymbol{b}$): Number of '1's in bitstring $\boldsymbol{b}$

$H(x) \coloneqq -x\ \log_2 x - (1-x)\log_2(1-x)$: binary entropy function

The tester chooses the largest possible non-negative integer N_{err} to satisfy the following condition.

$$N_1 H(N_{\mathrm{err}}/N_1) + (N - N_1) \leq K + N_{\mathrm{EC}} + \log_2(\alpha - \epsilon)$$

However, if the information transmitted by the TOE for error correction is encrypted and the formula for determining the QKD key length after privacy amplification does not include N_{EC}, the above inequality with $N_{\mathrm{EC}} = 0$ is used instead.

Note that if there is no non-negative integer N_{err} satisfying the condition, the attack fails.

The tester determines the success or failure of the attack in one of the following ways

(1) Take the sifted key $\boldsymbol{\kappa}_{\mathrm{sif}}$ from the TOE and extract the bits corresponding to the N_1 rounds in which the above attack was conducted to form the N_1 bit sequence $\boldsymbol{\kappa}_{1,\mathrm{sif}}$. The attack is successful if $\mathrm{wt}\left(\boldsymbol{\kappa}_{1,\mathrm{att}} - \boldsymbol{\kappa}_{1,\mathrm{sif}}\right) \leq N_{\mathrm{err}}$. If not satisfied, the attack fails.

(2) Take the QKD key $\boldsymbol{\kappa}_{\mathrm{QKD}}$ from the TOE. The sifted key $\boldsymbol{\kappa}_{\mathrm{sif}}$ actually generated by the TOE is not used. For an N-bit string $\boldsymbol{\kappa}$, let $f_1(\boldsymbol{\kappa})$ be the N_1-bit string formed by concatenating the bits corresponding to the N_1 rounds. The attack is successful if there exists an N-bit string $\boldsymbol{\kappa}$ that satisfies $f_{\mathrm{PA}}(\boldsymbol{\kappa}) = \boldsymbol{\kappa}_{\mathrm{QKD}}$ and $\mathrm{wt}\left(\boldsymbol{\kappa}_{1,\mathrm{att}} - f_1(\boldsymbol{\kappa})\right) \leq N_{\mathrm{err}}$. The attack fails if there are no such strings.

Notes

If the QKD key has ϵ-security, then the probability of the successful attack is lower than α under the above decision conditions.

In choosing the rate t at which to carry out the attack, note that as t is increased, N_1 increases, but K decreases due to an increase in bit errors in the X-basis.

12. Calculating attack potential

The evaluator shall calculate the attack potential according to [CEM] Appendix B.6. This section presents specific interpretations of attack potential calculations for the evaluation of QKD protocol implementation.

1. Elapsed Time

For FCS_QKD.1, the time required to overturn assumptions in security proofs. If the assumption(s) are overturned, QKD protocol cannot enforce its security proof, the security parameters in FCS_QKD.1.1 is not maintained, and FCS_QKD.1 is violated. Logically, this elapsed time is almost equal to the time it takes to fail the penetration test in Section 11. In many cases, violations of this SFR are accompanied by violations of FPT_EMS.1 and FPT_PHP.1. It is not required to consider the elapsed time until disclosure of the QKD key. The QKD key is of indefinite length, and hence if QKD protocol is run continuously, the elapsed time become infinite. Or the elapsed time for a length 2L QKD key is twice of the elapse time for a length L QKD key. That is, the elapsed time until disclosure of the QKD key only represent the QKD key length. If vulnerabilities are identified against other SFRs, the elapsed time is as defined in [CEM].

2. Specialist Expertise

The specialist expertise is as defined in [CEM]. Attack methods against QKD protocol implementations require optimization of attack methods based on knowledge of the QKD protocol and the implementation structures. At last, experts-level knowledge is required for the known vulnerabilities identified in Section 9.

3. Knowledge of the TOE

The knowledge of the TOE is as defined in [CEM].

4. Window of opportunity

The window of opportunity is as defined in [CEM]. At least for the known vulnerabilities identified in Section 4, attackers can attempt attacks by accessing the QKD link that are located in public area. Therefore, the attack opportunity is "unlimited".

5. IT hardware/software or other equipment

The equipment is as defined in [CEM]. For example, equipment which may be required for attack methods identified in Section 4 are classified as follows. This guide is based on price of each equipment.

Table 12-1 List of equipment

Classification	Equipment
Standard	Optical amplifier
	Photodetector
	Optical power meter
	Polarization analyser
	Beam splitter
	Polarizing beam splitter
	Circulator

Classification	Equipment
	Delay interferometer
	Optical delay line
	Polarization controller
	Phase modulator
	Intensity modulator
	Variable Optical Attenuator
Specialised	Tuneable laser
	Single photon detector
	High-end oscilloscope
	Optical spectrum analyser
	Spectrum analyser
	Time interval analyser

6. Example of ratings

In typical cases, expected rating for the known vulnerabilities identified in Section 4, the rating is shown below.

Table 12-2 Example of ratings

Elapsed Time	<= 2weeks	2	In situations where an attack is successful, the elapsed time is not so long. If we estimate longer, it is two weeks.
Specialist Expertise	experts	6	As in 2, the typical rating is experts.
Knowledge of the TOE	public	0	As in 3, the typical rating is public.
Window of opportunity	unlimited	0	As in 4, the typical rating is unlimited.
Equipment	specialised	4	As in 5, the typical rating is specialised.
Total		12	

13. Rationale for waiving penetration test

13.1. QKD transmitter

At the moment, no rationale for waving penetration test has been provided in the context of vulnerability analysis on the QKD transmitter.

13.2. QKD receiver

13.2.1.Detection efficiency

Rationale: Success conditions for penetration tests exploiting detector efficiency mismatch

This section deals with a penetration test with attacks exploiting difference in the detection efficiencies between the two photon detectors for bit values 0 and 1 used for generation of sifted key bits on the Z basis. Although this section assesses the threat caused by innate efficiency mismatch, this rationale is written such to also address cases where the mismatch arises from adversary intervention affecting the degrees of freedom of the optical pulses in the quantum channel, as detailed in Subsection 9.2.1.

In the following, the probability for the TOE to fail the penetration test is estimated using conditions obtained from functional tests and using a set of plausible assumptions. A concise sufficient condition for the failing probability to be negligibly small is given. Here we consider a decoy-state BB84 protocol in which the sifted key is generated from the Z basis only and the X basis is only used to monitor eavesdropping. The basis selection used in the QKD receiver may be an active or a passive one.

Suppose that a penetration test is conducted on a QKD session. Define parameters, variables, and functions as follows:

M : Number of communication rounds in the QKD session

$\bar{\mu}$: The mean photon number in a pulse (or a pair of pulses), averaged over the Z-basis signals.

$Q_{Z0(1)}$: Probability for a round to produce a sifted key bit with a bit value $0(1)$

N : Length of the sifted key produced in the QKD session

$\boldsymbol{b}$: N-bit sifted key produced in the QKD session

K : Length of the QKD key produced in the QKD session

N_{EC} : Length of the bit strings communicated for the error reconciliation that is accounted for in the privacy amplification. If the string is encrypted and is not accounted for in the privacy amplification, assume $N_{\mathrm{EC}} = 0$.

$H(x) \coloneqq -x \log_2 x - (1-x)\log_2(1-x)$: Binary entropy function

$D(x||y) \coloneqq x \log_2 \frac{x}{y} + (1-x)\log_2 \frac{1-x}{1-y}$: Kullback–Leibler divergence

$\mathrm{wt}(\boldsymbol{a})$: Number of '1's in bit string $\boldsymbol{a}$.

This rationale takes the following assumptions. These assumptions are expected to be true for penetration tests in Subsubsection. 11.2.1, but the evaluator should confirm the validity of them before applying this Rationale.

(A1) The criteria for the TOE to fail the penetration test is given as direct or indirect confirmation of the N-bit sifted key $\boldsymbol{b}$ belonging to a predicted set $\Omega \subset \{0,1\}^N$ satisfying $|\Omega| \leq 2^{K+N_{\mathrm{EC}}}$.

(A2) The privacy amplification ratio determined by the TOE correctly accounts for the fact that a sifted key bit may have leaked completely if the signal emitted from the transmitter included multiple photons.

(A3) The probability for the TOE to produce a sifted key bit when the transmitter emits two or more photons in an optical mode is no lower than that when it emits one or no photon in the same mode.

(A4) The photon number distribution in a signal emitted from the transmitter is well approximated by a Poisson distribution.
(A5) The mean photon number in every signal emitted from the transmitter does not exceeds unity.
In each round of a QKD session, the attacker may attack on the optical pulse(s) to modify the probabilities Q_{Z0} and Q_{Z1}. Suppose that a functional test assures that

$$\frac{\gamma}{1-\gamma} \leq \frac{Q_{Z1}}{Q_{Z0}} \leq \frac{1-\gamma}{\gamma}$$

holds for a positive constant $\gamma \leq 1/2$. Note that Q_{Z0} and Q_{Z1} may be different for different rounds.

Suppose that after the QKD session, the QKD receiver has produced a sifted key $\boldsymbol{b}$ of length N from a specific set of N rounds. On condition of those locations of the N rounds, we consider the conditional probability of the N-bit string $\boldsymbol{b} = b^{[1]} \cdots b^{[N]}$ over the 2^N values. Each bit is independent of the others, and $\mathrm{Prob}\{b^{[j]} = c\} = Q_{Zc}/(Q_{Z0} + Q_{Z1})$ where the values of Q_{Z0} and Q_{Z1} are for the round at which the jth sifted key bit $b^{[j]}$ was produced. Define a constant bit $c^{[j]}$ by $c^{[j]} = 0$ for $Q_{Z0} \geq Q_{Z1}$ and $c^{[j]} = 1$ for $Q_{Z0} < Q_{Z1}$. Then we have

$$p^{[j]} := \mathrm{Prob}\{b^{[j]} \neq c^{[j]}\} = 1 - \frac{Q_{Zc^{[j]}}}{(Q_{Z0} + Q_{Z1})} \geq \gamma,$$

and hence the expectation value of $\mathrm{wt}(\boldsymbol{b} - \boldsymbol{c})$ is no smaller than γN. From Hoeffding's inequality, we have, for all $\delta > 0$,

$$\mathrm{Prob}\{\mathrm{wt}(\boldsymbol{b} - \boldsymbol{c}) \leq (\gamma - \delta)N\} \leq \exp(-2\delta^2 N).$$

On the other hand, for any N-bit string $\boldsymbol{\Delta}$ with $\mathrm{wt}(\boldsymbol{\Delta}) \geq (\gamma - \delta)N$,

$$\mathrm{Prob}\{\boldsymbol{b} - \boldsymbol{c} = \boldsymbol{\Delta}\} \leq \gamma^{(\gamma-\delta)N}(1-\gamma)^{(1-\gamma+\delta)N} = 2^{-N(D(\gamma||\gamma-\delta)+H(\gamma-\delta))},$$

where $D(\gamma||\gamma - \delta) > 0$. Hence, according to (A1), the probability P_{fail} for the TOE to fail the penetration test satisfies

$$\begin{aligned} P_{\mathrm{fail}} &= \mathrm{Prob}\{\boldsymbol{b} \in \Omega\} \\ &\leq \mathrm{Prob}\{\boldsymbol{b} \in \Omega, \mathrm{wt}(\boldsymbol{b} - \boldsymbol{c}) \leq (\gamma - \delta)N\} + \mathrm{Prob}\{\boldsymbol{b} \in \Omega, \mathrm{wt}(\boldsymbol{b} - \boldsymbol{c}) \geq (\gamma - \delta)N\} \\ &\leq e^{-2\delta^2 N} + |\Omega| 2^{-N(D(\gamma||\gamma-\delta)+H(\gamma-\delta))} \leq e^{-2\delta^2 N} + 2^{-N(D(\gamma||\gamma-\delta)+H(\gamma-\delta))+K+H_{\mathrm{EC}}}. \end{aligned}$$

Hence, if

$$H(\gamma) > \frac{K + H_{\mathrm{EC}}}{N}$$

holds, we may choose $\delta > 0$ such that $NH(\gamma - \delta) = K + H_{\mathrm{EC}}$ holds, which shows that the probability P_{fail} is negligibly small.

We may further rewrite the condition by using assumptions (A2)-(A5). Let be the number of photons in the pulse(s) sent out by the transmitter in a round. Let "*tran_Z*" denote the event where the transmitter chooses the Z basis, and "*sif_suc*" denote the event where the TOE produces a sifted key bit. Then (A2) implies that

$$N - (K + H_{\mathrm{EC}}) > N\,\mathrm{Prob}\{n \geq 2 | sif\_suc\}$$

holds except a negligibly small probability. Since (A3) implies $\mathrm{Prob}\{sif\_suc | tran\_Z,\ n \geq 2\} \geq \mathrm{Prob}\{sif\_suc | tran\_Z,\ n \leq 1\}$, we have $\mathrm{Prob}\{sif\_suc | tran\_Z,\ n \geq 2\} \geq \mathrm{Prob}\{sif\_suc | tran\_Z\}$ and hence

$$\mathrm{Prob}\{n \geq 2 | sif\_suc\} = \mathrm{Prob}\{n \geq 2 | sif\_suc, tran\_Z\} \geq \mathrm{Prob}\{n \geq 2 |\ tran\_Z\}.$$

From (A4), we may write $\mathrm{Prob}\{n \geq 2 |\ tran\_Z\} = \sum_i p_i f(\mu_i)$ with $f(\mu) := 1 - e^{-\mu}(1 + \mu)$. Since $f''(\mu) \geq 0$ for $0 \leq \mu \leq 1$, (A5) implies that, for $\bar{\mu} := \sum_i p_i \mu_i$,

$$\mathrm{Prob}\{n \geq 2 |\ tran\_Z\} \geq f(\bar{\mu}).$$

Combining all the inequalities, we conclude that the probability P_{fail} is negligibly small if

$$H(\gamma) > e^{-\bar{\mu}}(1 + \bar{\mu}).$$

13.2.2.Single-photon sensitivity

Rationale: Success conditions for penetration test of bright illumination attacks.

This section deals with the bright illumination attack of the type 2).
In the following, the probability for the TOE to fail the penetration test is estimated using conditions obtained from functional tests and using a set of plausible assumptions. A concise sufficient condition for the failing probability to be negligibly small is given. Here we consider a decoy-state BB84 protocol in which the sifted key is generated from the Z basis only and the X basis is only used to monitor eavesdropping. The basis selection used in the QKD receiver may be an active or a passive one. In the case of passive basis selection, the beam splitter for selecting the Z- and X-basis is assumed to have a splitting ratio favourable for the Z-basis.

Suppose that a penetration test is conducted on a QKD session. Define parameters, variables, and functions as follows:
p_Z, p_X : Selection probabilities of the basis for the QKD transmitter
r_Z, r_X : Coupling efficiency of the beam splitter for the passive basis selection of the QKD receiver
η_Z, η_X : Quantum efficiency of the two photon detectors in each base of the QKD receiver (Assume that the two have the same quantum efficiency.)
M : Number of communication rounds in the QKD session
t : Parameter indicating the frequency of Intercept-resend attacks $(0 \leq t \leq 1)$
$Q_{Z(X)}$: Probability of a successful Z(X)-basis detection for a round with no intercept-resend attacks.
$\tilde{Q}_{Z(X)}$: Probability of a successful Z(X)-basis detection for a round with intercept-resend attacks.
N : Length of the sifted key produced in the QKD session
$\boldsymbol{b}$: N-bit sifted key produced in the QKD session
K : Length of the QKD key produced in the QKD session
H_{EC} : Length of the bit strings communicated for the error reconciliation that is accounted for in the privacy amplification. If the string is encrypted and is not accounted for in the privacy amplification, assume $H_{\mathrm{EC}} = 0$.
$H(x) \coloneqq -x \log_2 x - (1-x)\log_2(1-x)$: Binary entropy function
$v(t)$: Fraction of the bits in the sifted key that could be compromised by an attacker.
$e(t)$: Bit error rate in the X basis.

This rationale takes the following assumptions. These assumptions are expected to be true for penetration tests in Subsection 11.2.3, but the evaluator should confirm the validity of them before applying this Rationale.
(A1) The criteria for the TOE to fail the penetration test is given as direct or indirect confirmation of the N-bit sifted key $\boldsymbol{b}$ belonging to a predicted set $\Omega \subset \{0,1\}^N$ satisfying $|\Omega| \leq 2^{K+H_{\mathrm{EC}}}$.
(A2) When the observed bit error rate in the X basis is e, the privacy amplification ratio determined by the TOE satisfies $K + H_{\mathrm{EC}} \leq N\big(1 - H(e)\big)$.
(A3) The photon detectors may have their quantum efficiencies modified due to the bright illumination attack, they do not switch to linear mode and their response is well approximated by the standard model of an on-off detector: the detection probability when a laser pulse with average photon number μ is incident on an on-off detector with quantum efficiency η is given by $\eta\mu\xi(\eta\mu)$, where

$$\xi(x) \coloneqq \frac{1 - e^{-x}}{x}.$$

Here, $\xi(x)$ is a decreasing function and $x\xi(x)$ is an increasing function of x.
Suppose that among the M rounds in the QKD session, an attacker performs an intercept-resend attack for Mt rounds.

Round with no intercept-resend attack:

With probability $p_Z Q_Z$, the Z-basis communication succeeds and a sifted key bit is generated. The attacker has no knowledge of this bit value.

Round in which the intercept-resend attack took place:
With probability $p_Z \tilde{Q}_Z$ the communication in the Z-basis succeeds and a sifted key is generated.
With probability $p_X \tilde{Q}_X$ the communication in the X-basis succeeds and the occurrence of a bit errors is recorded. Due to the intercept-resend attack, a bit error occurs here with probability 1/2.

In this QKD session,
$N = Mtp_Z \tilde{Q}_Z + M(1-t)p_Z Q_Z$ bits of sifted keys are generated, of which at least $M(1-t)p_Z Q_Z$ bits are not compromised by the attacker at all. That is, the fraction of bits in the sifted key that may be compromised is at most $v(t) \coloneqq \frac{t\tilde{Q}_Z}{t\tilde{Q}_Z + (1-t)Q_Z}$. The probability of the sifted key $\boldsymbol{b}$ to take any specific N-bit string is no greater than $2^{-N(1-v(t))}$ and hence

$$\mathrm{Prob}\{\boldsymbol{b} \in \Omega\} \leq |\Omega| 2^{-N(1-v(t))}.$$

The bit error rate observed in the X basis is at least $e(t) \coloneqq \frac{1}{2}\frac{t\tilde{Q}_X}{t\tilde{Q}_X + (1-t)Q_X}$. Hence, according to (A1) and (A2), the probability P_{fail} for the TOE to fail the penetration test satisfies

$$P_{\mathrm{fail}} = \mathrm{Prob}\{\boldsymbol{b} \in \Omega\} \leq 2^{K + H_{\mathrm{EC}} - N(1-v(t))} \leq 2^{-N\left(H(e(t)) - v(t)\right)}$$

Hence, if

$$H\left(e(t)\right) > v(t)$$

holds, the probability P_{fail} is negligibly small. With $t' \coloneqq 2e(t)$ and $\gamma = \frac{Q_Z \tilde{Q}_X}{Q_X \tilde{Q}_Z}$, it holds that $v(t) = \frac{t'}{t' + (1-t')\gamma}$, which leads to a necessary condition for a successful attack,

$$H\left(\frac{t'}{2}\right) < \frac{t'}{t' + (1-t')\gamma}.$$

If $\gamma > 0.285$, there is no t' in $[0,1]$ that satisfies this inequality, so the attack will fail no matter how the attack frequency t is chosen. This means that the attack in a penetration test succeeds only if it holds that

$$\gamma = \frac{Q_Z \tilde{Q}_X}{Q_X \tilde{Q}_Z} \leq 0.285.$$

Relation to the functional test:
From the quantum efficiencies of the detectors without bright illumination attack,

$$\frac{Q_Z}{Q_X} = \frac{r_Z \eta_Z}{r_X \eta_X}$$

holds. At the resending step, light with an average photon number greater than unity can also be used. When the pulse intensity incident on the QKD receiver is μ, it holds that

$$\tilde{Q}_Z = r_Z \eta_Z \mu \xi(r_Z \eta_Z \mu)$$

according to (A3). The two photon detectors in the X-basis may have their quantum efficiencies modified due to the bright illumination attack. Denoting the modified quantum efficiencies by $\tilde{\eta}_{X0}$ and $\tilde{\eta}_{X1}$, the probability of successful detection in the X basis is given by

$$\tilde{Q}_X = \frac{r_X \tilde{\eta}_{X0} \mu}{2} \xi\left(\frac{r_X \tilde{\eta}_{X0} \mu}{2}\right) + \frac{r_X \tilde{\eta}_{X1} \mu}{2} \xi\left(\frac{r_X \tilde{\eta}_{X1} \mu}{2}\right).$$

The functional test of Sec. 10.9.2.1 guarantees that $\tilde{\eta}_{X0} \geq \kappa\, \eta_X$ and $\tilde{\eta}_{X1} \geq \kappa\, \eta_X$ using a parameter $\kappa (\leq 1)$, which takes a value close to unity. It follows that

$$\tilde{Q}_{\mathrm{X}} \geq r_X \kappa\, \eta_X \mu \xi\left(\frac{r_X \kappa\, \eta_X \mu}{2}\right),$$

which leads to

$$\gamma = \frac{Q_{\mathrm{Z}} \tilde{Q}_{\mathrm{X}}}{Q_{\mathrm{X}} \tilde{Q}_{\mathrm{Z}}} \geq\ \kappa \frac{\xi\left(\frac{r_X \kappa\, \eta_X \mu}{2}\right)}{\xi(r_Z \eta_Z \mu)}.$$

Since $\xi(x)$ is a decreasing function, $\gamma \geq \kappa$ is assured if the ratio of passive basis selection probabilities in normal operation satisfies

$$\frac{Q_{\mathrm{Z}}}{Q_{\mathrm{X}}} = \frac{r_Z \eta_Z}{r_X \eta_X} \geq \frac{1}{2}.$$

Therefore, it can be concluded that if the TOE has passed the functional test with $\kappa > 0.285$, the probability P_{fail} for the TOE to fail the penetration test is negligibly small.

Revision history

Version	Date	Description
1.0	13.05.2025	First issue.

Review history

Summary of editing and reviewing processes

Core parts of the document were drafted by the QKD CC/PP Study Group under the MIC project.
Editing and reviewing of the document were conducted by
- QKD Implementation Security Study Group
- QKD Technical Review Committee
under the Quantum Forum (General Incorporated Association).
The drafts of the document were presented and discussed in ETSI ISG-QKD meetings.

Activity record of QKD Technical Review Committee

- 1st meeting (Nov. 19, 2024, 15:00~16:00)
  Discussion on review policy and schedule
- 1st round review on SD v0.43 (Nov. 19 - 29, 2024)

- 2nd meeting (Dec. 12, 2024, 18:00~19:00, jointly with QKD CC/PP SG)
  Discussion on the revised edition
- 2nd round review on SD v0.44 (Dec. 12- 23, 2024)

- 3rd meeting (Jan. 23, 2025, 18:00~20:40, jointly with QKD CC/PP SG)
  Discussion on the revised edition
- 3rd round review on SD v0.53 (Feb. 10- 17, 2025)

- 4th meeting (Feb. 20, 2025, 17:00~19:30, jointly with QKD CC/PP SG)
  Discussion on the revised edition
- 4th round review on SD v0.59 (Mar. 7- 21, 2025)

- 5th meeting (Mar. 21, 2025, 16:00~19:00, jointly with QKD CC/PP SG)
  Discussion on the revised edition
- 5th round review on SD v0.63r2 (Apr. 4- 9, 2025)

- 6th meeting (Apr. 10, 2025, 16:00~18:00, jointly with QKD CC/PP SG)
  Discussion on the revised edition
- 6th round review on SD v0.65r4 (Apr. 16- 21, 2025)

- 7th meeting (Apr. 17, 2025, 16:00~18:00, jointly with QKD CC/PP SG)
  Discussion on the revised edition

- 8th meeting (Apr. 24, 2025, 16:00~18:00, jointly with QKD CC/PP SG)
  Discussion on the revised edition

Discussion record in ETSI

- ISG-QKD#36f, Nov. 5, 2024
  Introduction on QKD module certification activities in Japan
- ISG-QKD#37, Dec. 2-4, 2024
  Discussion on SD v0.43
- ISG-QKD#37b, Jan. 7, 2025
  Discussion on SD v0.47
- ISG-QKD#37c, Feb. 4, 2025
  Discussion on SD v0.51
- ISG-QKD#37d, Mar. 4, 2025
  Discussion on SD v0.58
- ISG-QKD#37e, Apr. 1, 2025
  Discussion on SD v0.62
- ISG-QKD#37e, May. 6, 2025
  Discussion on SD v0.67

End of document