Photon Number Splitting Attack – Proposal and Analysis of an Experimental Scheme

*Ariel Ashkenazy, Yuval Idan, Dor Korn, Dror Fixler, Barak Dayan, and Eliahu Cohen**

Photon-number-splitting (PNS) is a well-known theoretical attack on quantum key distribution (QKD) protocols that employ weak coherent states produced by attenuated laser pulses. However, beyond the fact that it has not yet been demonstrated experimentally, its plausibility and effect on quantum bit error rate are questioned. In this work, an experimental scheme is presented for PNS attack employing demonstrated technological capabilities, specifically a single-photon Raman interaction (SPRINT) in a cavity-enhanced three-level atomic system. Several aspects of the proposed implementation are addressed, analytically and simulatively, and the eavesdropper's information gain by the attack is calculated. Furthermore, it is analytically shown that the scheme results in a small (yet non-zero) quantum bit error rate, and a comparison to purely theoretical analyses in the literature is presented. It is believed that the inherent nonlinearity of the PNS attack unavoidably affects the optical modes sent to the receiver, and accordingly will always result in some error rate.

1. Introduction

Quantum key distribution (QKD), offering a provably secure communication,[1] is one of the most developed quantum technologies to date. Many QKD protocols utilize single photons as information carriers, with BB84[2] being the first and probably most famous among them. In BB84, also known as the four-state protocol,[3] the information is encoded using any degree of freedom with two mutually unbiased bases (where measuring one property completely randomizes the other[4]). For each bit, the sender, usually referred to as Alice, randomly chooses the encoding basis. Then, the receiver, Bob, also randomly chooses the basis in which to measure. Afterwards, Bob publishes (publicly) which bases he had used, and Alice and Bob discard any measurement where their bases were different (half of the measurements, on average), producing the sifted key. Then, Alice and Bob select a random string of bits from the sifted key and compare their values. This process allows them to detect any eavesdropping attempt, since it inevitably introduces some quantum bit error rate (QBER).

With current technology, however, the use of single photons is impractical. Instead, QKD experiments use the same setup but with weak coherent states (WCS), i.e., laser pulses attenuated such that the average photon number, μ, is low.[4] The use of WCS for QKD opens the gate to the photon-number-splitting (PNS) attack,[3] where Eve (the eavesdropper) performs quantum non-demolition (QND) measurements of the number of photons contained in each pulse, and from any multi-photon pulse she extracts a single photon (or more) while forwarding the rest to Bob. She keeps her photons (in a quantum memory) until Alice and Bob publicly announce their bases, and then measures her photons in the correct basis.

In theory, it has been suggested that this attack will allow Eve to extract full information from all multi-photon pulses without introducing any QBER that can expose her attack to Alice and Bob.[5] In fact, even if we assume that Eve is not in possession of a quantum memory, she still benefits from a PNS attack since it allows her to gain information about the key whenever she guesses the correct basis. More advanced QKD protocols were invented to protect against PNS attack, notably the SARG04,[6,7] the decoy-state,[8,9] and the coherent one-way (COW)[10] protocols.

Even though the theoretical aspects of the attack were researched extensively,[5,11–13] an experimental demonstration of it has proved to be highly challenging and is still missing. To date, only a modified version of the PNS attack had been realized (in part) in a proof-of-principle experiment,[14] while some researchers deem the full-scale implementation of the PNS attack to be "unrealistic or even unphysical".[15] Undoubtedly, even detailed theoretical analyses of realistic physical implementations of a PNS attack are important since they can reveal

A. Ashkenazy, Y. Idan, D. Fixler, E. Cohen
Faculty of Engineering and the Institute of Nanotechnology and Advanced Materials
Bar-Ilan University
Ramat Gan 5290002, Israel
E-mail: eliahu.cohen@biu.ac.il

D. Korn, B. Dayan
AMOS and Department of Chemical and Biological Physics
Weizmann Institute of Science
Rehovot 76100, Israel

B. Dayan
Quantum Source Labs
Rehovot 7670402, Israel

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/qute.202300437

DOI: 10.1002/qute.202300437

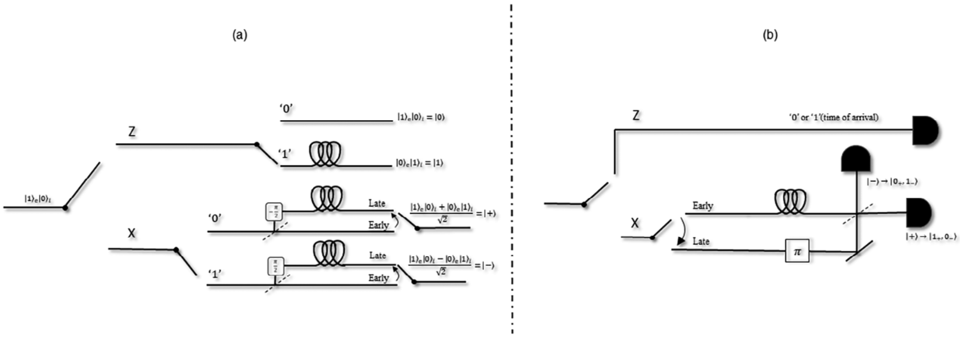
(a)
$|1\rangle_e|0\rangle_l$
Z
'0'
$|1\rangle_e|0\rangle_l = |0\rangle$
'1'
$|0\rangle_e|1\rangle_l = |1\rangle$
X
'0'
$-\frac{\pi}{2}$
Late
Early
$\frac{|1\rangle_e|0\rangle_l + |0\rangle_e|1\rangle_l}{\sqrt{2}} = |+\rangle$
'1'
$\frac{\pi}{2}$
Late
Early
$\frac{|1\rangle_e|0\rangle_l - |0\rangle_e|1\rangle_l}{\sqrt{2}} = |-\rangle$
(b)
Z
'0' or '1'(time of arrival)
$|-\rangle \rightarrow |0_+, 1_-\rangle$
Early
X
Late
π
$|+\rangle \rightarrow |1_+, 0_-\rangle$

Figure 1. a) A schematic diagram of the device employed for preparation of the states used in time-bin encoded BB84. The x-basis states are prepared using optical switches that combine the early and late modes onto a single spatial mode. b) Detection in the z-basis is done simply by measuring time-of-arrival, while for detection in the x-basis a reversed version of the preparation device is used.

inherent limitations on its feasibility and price in QBER to Bob's channel.

In this work, we present an implementation of the PNS attack using single-photon Raman interaction (SPRINT) with a single, cavity-enhanced atom. The SPRINT mechanism is well-established both theoretically and experimentally,[16–18] capable of extracting a single photon from an optical pulse.[19]

We give a full description of the proposed attack and present an analytical analysis of the detection statistics at Bob's and Eve's detectors. From the analysis, we extract Eve's information gain by the attack, as well as the QBER it introduces, as a function of the mean photon number in the pulse. Eve's information gain is shown to be close (and in some cases equal) to the theoretical limit, and thus the security of basic QKD protocols implemented with WCS is indeed vulnerable to this realization of the PNS attack. However, in contrary to the theoretical assumptions, our results show that some QBER is introduced by the attack. We believe that this is a consequence of the inherent nonlinearity of any PNS attack, which unavoidably changes the (in our case temporal) properties of the mode sent to Bob. These results demonstrate the significance of such realistic and detailed theoretical analyses of physical systems that can be harnessed for PNS attack, and can also serve to advance the debate regarding the general plausibility of this attack.

The rest of the paper is structured as follows. In the next section, we discuss some necessary preliminaries regarding time-bin encoded BB84 and SPRINT. In Section 3, we present our proposal for SPRINT-based PNS attack, including a full mathematical description of the attack, as well as the assumptions that we make in this initial analysis. Then, in Section 4, we describe our analysis results of the detection statistics following such an attack, as well as the QBER it introduces and Eve's information gain. A comprehensive discussion of the results and a comparison with the theoretical works on PNS attack is then presented in Section 5, followed by some concluding remarks.

2. Theoretical Background

2.1. Time-Bin Encoded BB84

In this work, we employ the time-bin encoded BB84 protocol, but we believe the results are quite general. The implementation of such a protocol with a single-photon source will serve as the benchmark for all of our analyses. In this implementation, the qubit is made up of two time-bins with the z-basis states being $|0\rangle \equiv |1\rangle_e |0\rangle_l$ for a photon in the early time-bin and $|1\rangle \equiv |0\rangle_e |1\rangle_l$ for being in the late time-bin, and the x-basis is comprised of the superposition states $|\pm\rangle \equiv \frac{1}{\sqrt{2}} \left(|1\rangle_e |0\rangle_l \pm |0\rangle_e |1\rangle_l\right)$, as described, for instance, in ref. [20]. Using this encoding scheme, the z-basis is usually referred to as the "time basis" while the x-basis is the "phase basis".[21] The states $\{|0\rangle, |+\rangle\}$ correspond to a bit value of '0', while a bit value of '1' is assigned to the $\{|1\rangle, |-\rangle\}$ states. As shown in **Figure 1a**, preparation of the z-basis states is done using a delay line, while the x-basis states are prepared using optical switches that first connect to the early mode and then to the late mode,[22] thus combining the two temporal modes onto a single spatial mode. Detection in the z-basis is done simply by measuring the time-of-arrival of the photon, while for measurement in the x-basis, a reversed version (up to an irrelevant global phase) of the preparation device is used, as depicted in Figure 1b. Note that, in realistic scenarios, Bob may get clicks in both detectors, due to dark current. These bits cannot be discarded, as this will allow Eve to perform an attack (in which she sends pulses with large photon number after her eavesdropping). Instead, we assume Bob assigns random bit values to these cases of double detection, see ref. [11].

Without the presence of an eavesdropper, it is simple to verify that Bob's probability to detect a state x, given that Alice has prepared a state y, $p(x|y)$, is:

$$\begin{aligned} &p(0|0) = p(1|1) = p(+|+) = p(-|-) = 1 \\ &p(0|1) = p(1|0) = p(+|-) = p(-|+) = 0 \end{aligned} \tag{1}$$

And it is also simple to verify that if they use different bases, then:

$$p(0|\pm) = p(1|\pm) = p(\pm|0) = p(\pm|1) = 0.5 \tag{2}$$

But of course, since after their public communication Alice and Bob completely discard all instances where they have used different bases, the sifted key consists of only the cases for which their bases agree. Therefore, the QBER, defined as the probability that Bob gets an incorrect bit value, is zero.

Now, if we assume Eve is present, she can perform an intercept-resend (I-R) attack, i.e., she detects either in the z-basis or the x-basis (using the scheme described in Figure 1b), and resends to Bob a state that corresponds to her detection result. With probability half, Eve's detection is in the wrong basis, and so Bob gets a state in a basis that is different from Alice's, resulting in a QBER of 0.25 (i.e., probability half that Eve's detection is in wrong basis times probability half of getting the wrong bit value when measuring in the wrong basis, Equation (2)).

Nowadays, the protocol is often implemented with WCS instead of a single-photon source. A coherent state is defined as:

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \tag{3}$$

where $\alpha \in \mathbb{C}$. A coherent state is a superposition of number states, and the probability of detecting n photons is:

$$p_\mu(n) = e^{-\mu} \frac{\mu^n}{n!} \tag{4}$$

where $\mu \equiv |\alpha|^2$ is the average photon number.[23]

Alice's preparation (using the same setup as in Figure 1a, but with WCS input) results in the following z-basis states:

$$\begin{aligned} |0\rangle &= |\alpha\rangle_e |0\rangle_l \\ |1\rangle &= |0\rangle_e |\alpha\rangle_l \end{aligned} \tag{5}$$

and x-basis states:

$$|\pm\rangle = |\frac{\alpha}{\sqrt{2}}\rangle_e |\frac{\pm\alpha}{\sqrt{2}}\rangle_l \tag{6}$$

A simple analysis (see Appendix A) shows the statistics at Bob's side to be:

$$\begin{aligned} &p(0|0) = p(1|1) = p(+|+) = p(-|-) = p_\mu(n \geq 1) \\ &p(0|1) = p(1|0) = p(+|-) = p(-|+) = 0 \end{aligned} \tag{7}$$

while detection in the wrong basis yields:

$$p(0|\pm) = p(1|\pm) = p(\pm|0) = p(\pm|1) = \frac{1}{2} p_\mu(n \geq 1) \tag{8}$$

So, for the sifted key, i.e., only the cases where Alice and Bob use the same basis, the QBER is zero.

If now Eve performs an I-R attack on the communication, she will introduce some QBER. The exact expression will depend on her strategy for cases of no-detection or double detection. For the cases of a single detection, Eve will resend a WCS in a state corresponding to her measurement result. There is a 50% chance that her detection is in the correct basis, in which case her measurement result is the correct bit value and so the state she sends out is the same as the original one and no QBER is introduced by her. In the other half of the cases, she sends out a WCS state in the wrong basis, and so Bob has a probability of error that is equal to $\frac{1}{2} p_\mu(n \geq 1)$, according to the above analysis. Thus, normalized by the detection probability at Bob's side ($p_\mu(n \geq 1)$), a single detection event at Eve's side, results in a QBER of 0.25. In principle, Eve may utilize some more advanced techniques as explored, for instance, in ref. [24, 25] or even attacks that affect directly Alice and Bob's setups, such as ref. [26, 27]. In addition, there is the risk of Trojan-horse attacks.[28] In a recent work,[29] a new variant of the beamsplitter attack[30,31] was proposed in the context of phase-matching QKD protocols.[32] In contrast to these beamsplitter attacks, the method analyzed herewith relies on active, nonlinear operations at Eve's side, and thus has fundamentally different properties and unique ramifications for the transmitted quantum states of light.

2.2. SPRINT

First proposed in ref. [33], SPRINT-based systems were investigated theoretically and experimentally for single-photon routing,[16] photon-atom qubit swapping,[17,18] and for deterministic extraction of a single photon from an optical pulse,[19] the latter being the basis for the SPRINT-based PNS attack proposed herein. The SPRINT mechanism occurs in a three-level system in a Λ configuration, where each transition is coupled to a different photonic mode. As depicted in **Figure 2**, the transition $|R\rangle \rightarrow |e\rangle$ is coupled to one mode, in our case the one propagating from left to right in a waveguide, while the transition $|T\rangle \rightarrow |e\rangle$ is coupled to another mode - here the opposite direction in the waveguide. Starting in the $|R\rangle$ state, a photon incident from the left will be reflected by the system due to destructive interference in the transmitted direction.[16] This reflection results in the system transitioning to the $|T\rangle$ state. Any additional photons incident from the left will be transmitted through the system without any disturbance, since the $|T\rangle$ state is not coupled to their propagation direction. Experimental realizations of SPRINT use a ^{87}Rb atom coupled to a microresonator to which light is coupled via a tapered nanofiber, and an extraction of a single photon out of coherent states of different intensities (with average photon number of 0.2 – 11) was demonstrated. For further details of the theoretical and experimental aspects of SPRINT the reader is referred to ref. [17, 19] and references therein. In this paper, we focus on theoretical, rather than experimental, aspects underlying a possible realization of PNS using SPRINT and the ensuing implications for QKD.

3. SPRINT-Based PNS Attack

3.1. Description and Assumptions

With the theoretical background presented above, the proposal of SPRINT-based PNS attack is straightforward – Eve places a SPRINT-based system on the quantum communication channel, with the reflection output arm of the system directed to Bob, and the transmission output arm connected to Eve's detectors (or quantum memory, if exists). For each qubit (which is comprised of two time-bins, early and late), Eve initializes the system in the $|R\rangle$ state. Thus, for single-photon pulses the photon is reflected to Bob, while for multi-photon pulses a single photon is reflected to Bob and the rest are transmitted to Eve. Note that in this implementation of the PNS attack, from all multi-photon pulses Eve steals all but one of the photons.

For the sake of this initial analysis of the attack, we make some conservative assumptions. First, we assume that the coupling to the SPRINT-system does not introduce any losses or phase-shifts, and it does not shift a photon from one time-bin (temporal mode) to another. It only separates a single photon from multi-photon pulses. We also assume perfect coupling, i.e., whenever

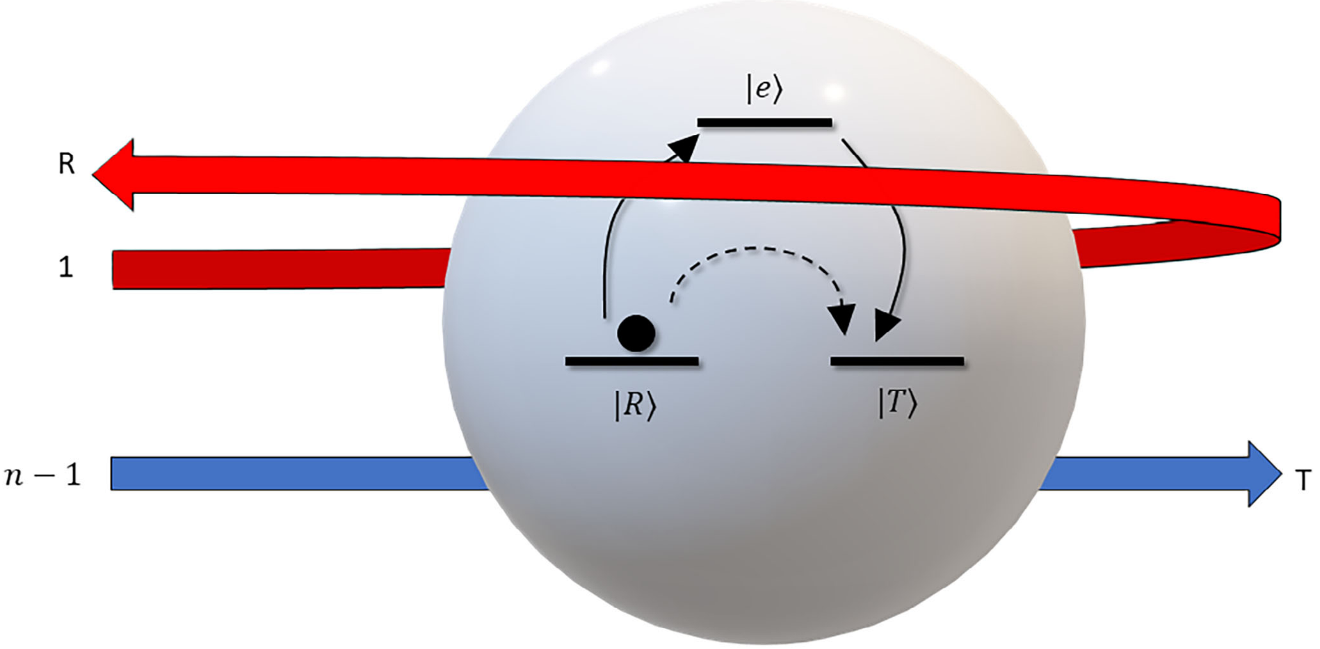

Figure 2. A schematic description of the SPRINT effect on a pulse containing n photons. The system is initially in the $|R\rangle$ state, and a destructive interference effect in the transmitted direction causes one photon (red) from the pulse to be reflected while the system transitions to the $|T\rangle$ state. The $|T\rangle$ state is a dark state in the sense that it is uncoupled to the incoming pulse, and accordingly the remaining $n-1$ photons (blue) are transmitted without interacting with the system.

the SPRINT-system starts in the $|R\rangle$ state, the first incident photon is reflected and the rest (if any) are transmitted, with certainty. In addition, we assume that Eve has full knowledge of the protocol that Alice and Bob are using and is able to optimize her attack accordingly. For simplicity, we assume that the losses in the quantum communication channel are negligible, and that Bob's and Eve's detectors are perfect, i.e., 100% quantum yield, no dark counts and no dead-time. Finally, we use the common assumption that Bob does not have a photon number resolving detector.[5] As for Eve's quantum memory, we do not assume anything but rather analyze the attack both with and without quantum memory.

We note that a different configuration may be proposed for the implementation of PNS attack with SPRINT-system, namely one in which Eve's detectors are placed at the reflection output arm, allowing her to block all incident single photon pulses, while stealing a single photon out of every multi-photon pulse. However, this scheme adds a significant loss into the channel, since in this configuration Bob will get a photon only when the original pulse contains two or more photons, $p_\mu(n \geq 2)$. Therefore, this configuration is less suitable for the present analysis in which the losses in the quantum communication channel are assumed to be negligible, and such a major increase of the loss will expose Eve's eavesdropping attempt.

As our analysis results show below, the QBER introduced by SPRINT-based PNS attack is not zero. It is reasonable to assume that Eve might want to limit the QBER she introduces into the system below a certain threshold. Therefore, we assume that Eve couples the quantum communication channel to the SPRINT-system only for a part $0 \leq \zeta \leq 1$ of the communicated qubits, and that Eve is non-aggressive, i.e., Eve does not block or perform an I-R attack on the single-photon pulses or on the $(1-\zeta)$ part of qubits that are not coupled to the system, so the QBER for these cases is zero.

3.2. Mathematical Description

For each optical spatial mode in the system, we have two temporal modes, early (e) and late (l). The whole system is described by the combined state of the early and late modes of the incident/transmitted spatial mode (T), the SPRINT-system (S), and the early and late modes of the reflected spatial mode (R). The SPRINT-system is initially in the reflecting state $|R\rangle$, and upon reflecting a single photon it transitions to the transmitting state, $|T\rangle$, which keeps the following incident photons in the same spatial mode. For the description of the state of the whole system, we will use the following notation:

$$|\psi\rangle = |\,\rangle_{T,e} \otimes |\,\rangle_{R,e} \otimes |\,\rangle_{S} \otimes |\,\rangle_{T,l} \otimes |\,\rangle_{R,l} \tag{9}$$

This notation and the assumptions above correspond to the description of the coupling of the incident photons to the SPRINT-system, when acting on the subspace of the early temporal mode (i.e., acting on $|\,\rangle_{T,e} \otimes |\,\rangle_{R,e} \otimes |\,\rangle_{S}$ subspace only), as the operator:

$$\hat{s}_e = I_{T,e} \otimes I_{R,e} \otimes |T\rangle\langle T| + |0,0,R\rangle\langle 0,0,R| \\ + \sum_{n=1}^{\infty} |n-1,1,T\rangle\langle n,0,R| \tag{10}$$

and, when acting on the late temporal mode (i.e., acting on $|\,\rangle_S \otimes |\,\rangle_{T,l} \otimes |\,\rangle_{R,l}$ subspace only):

$$\hat{s}_l = |T\rangle\langle T| \otimes I_{T,l} \otimes I_{R,l} + |R,0,0\rangle\langle R,0,0| + \sum_{n=1}^{\infty} |T,n-1,1\rangle\langle R,n,0| \tag{11}$$

extending the description given in ref. [19]. In these expressions for the operators, we do not describe the effect of the coupling to the SPRINT-system on a general $|n\rangle$ state at the reflected mode, since such an operation is never realized, as the reflected mode (for each temporal mode) is always assumed to initially be in the vacuum state, i.e. $I_R = |0\rangle\langle 0|$. This argument verifies that the operators are unitary, as $\hat{s}_e^{\dagger}\hat{s}_e = I_{T,e} \otimes I_{R,e} \otimes |T\rangle\langle T| + |0,0,R\rangle\langle 0,0,R| + \sum_{n=1}^{\infty} |n,0,R\rangle\langle n,0,R|$ which is indeed the identity operator for all the possible initial states of the system for the subspace of the early modes (the same can be easily verified for $\hat{s}_l$).

Thus, we can write the transformation for a coherent state (Equation (3)) in the early temporal mode as:

$$\hat{s}_e |\alpha,0,R\rangle = e^{-\frac{|\alpha|^2}{2}} \left(|0,0,R\rangle + \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n-1,1,T\rangle \right) \tag{12}$$

and for a coherent state in the late temporal mode:

$$\hat{s}_l |R,\alpha,0\rangle = e^{-\frac{|\alpha|^2}{2}} \left(|R,0,0\rangle + \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |T,n-1,1\rangle \right) \tag{13}$$

As for the mathematical description of other components of the system – A φ phase shift is described by the unitary transformation $\hat{U} = \exp[i\varphi\hat{n}]$, and its effect for a coherent state is $|\alpha\rangle \rightarrow |\alpha e^{i\varphi}\rangle$. A mirror introduces a phase shift of $\frac{\pi}{2}$.[23]

A beamsplitter, with reflectance r and transmittance t, transforms input annihilation operators to output operators according to $\hat{a}_{out} = t\hat{a}_{in} + ir\hat{b}_{in}$ and $\hat{b}_{out} = ir\hat{a}_{in} + t\hat{b}_{in}$, or equivalently $\hat{a}_{in} = t\hat{a}_{out} - ir\hat{b}_{out}$ and $\hat{b}_{in} = -ir\hat{a}_{out} + t\hat{b}_{out}$.[34]

Remembering that a general photon number state can be written as $|n\rangle = \frac{(\hat{a}^{\dagger})^n}{\sqrt{n!}}|0\rangle$, the effect of a beamsplitter on a general input state $|n,m\rangle$ is:

$$|n,m\rangle \rightarrow \frac{(t\hat{a}_{out}^{\dagger} + ir\hat{b}_{out}^{\dagger})^n (ir\hat{a}_{out}^{\dagger} + t\hat{b}_{out}^{\dagger})^m}{\sqrt{n!m!}} |0,0\rangle \tag{14}$$

which, for a balanced beamsplitter ($r = t = \frac{1}{\sqrt{2}}$) can be written as:

$$(2^{n+m}n!m!)^{-\frac{1}{2}} \sum_{k_a=0}^{n} \sum_{k_b=0}^{m} \binom{n}{k_a}\binom{m}{k_b} i^{k_a+k_b} \cdot \sqrt{(n-k_a+k_b)!(m-k_b+k_a)!}\, |n-k_a+k_b, m-k_b+k_a\rangle \tag{15}$$

where k_a (k_b) is the number of photons, out of the initial n (m) photons, reflected from the first (second) input mode.[35]

As described in Figure 1, early and late modes propagating in two different spatial modes can be combined into a single spatial mode using a simple optical switch that first connects the output fiber to the short arm (early bin) and later connects it to the long arm (late bin), thus getting both temporal modes on the same spatial mode without any loss. The opposite operation can also be implemented, separating two temporal modes propagating in the same spatial mode into two different spatial modes, by reversing the process.[22] Shifting between time-bins in the same spatial mode, i.e., transforming the quantum state of an early mode into a late mode, is just straightforward delay-line with integer multiples of the time-bin separation, it introduces no relative phase.[36]

When analyzing the detection statistics in the transmitted (reflected) mode, we trace over the SPRINT-system and the reflected (transmitted) modes to get the reduced density matrix, ρ_T (ρ_R), describing the transmitted (reflected) mode only. For that, the relation $tr(|B\rangle\langle A|) = \langle A|B\rangle$ is very useful.

Finally, a detector in some mode x is represented as a projection onto occupied states, $P_x = I_x - |0\rangle_x\langle 0|_x$, in the appropriate mode. The expectation value of such a projector thus gives the probability for a click in this mode's detector. Double clicks are assigned random bit values, as discussed above.

4. Analysis Results

4.1. Detection Statistics Following a SPRINT-Based PNS Attack

Using the above description of the SPRINT-based PNS attack, we analyzed the detection statistics in Bob's and Eve's sides.

The detection probabilities obtained in the reflected arm (Bob's side) are (detailed calculations are presented in Appendix B):

$$\begin{aligned} &p_R(0|0) = p_R(1|1) = p_\mu(n \geq 1) \\ &p_R(1|0) = p_R(0|1) = 0 \end{aligned} \tag{16}$$

$$p_R(+|0) = p_R(-|0) = p_R(+|1) = p_R(-|1) = \frac{1}{2}p_\mu(n \geq 1) \tag{17}$$

for z-basis states, and for x-basis states:

$$\begin{aligned} &p_R(0|+) = p_R(0|-) = p_{\frac{\mu}{2}}(n \geq 1) \\ &p_R(1|+) = p_R(1|-) = p_{\frac{\mu}{2}}(n = 0) \cdot p_{\frac{\mu}{2}}(n \geq 1) \end{aligned} \tag{18}$$

$$p_R(+|+) = p_R(-|-) = \frac{1}{2}p_\mu(n \geq 1) + e^{-\mu} \sum_{n=1}^{\infty} \frac{\mu^n \sqrt{n}}{2^n n!} \tag{19}$$

$$p_R(-|+) = p_R(+|-) = \frac{1}{2}p_\mu(n \geq 1) - e^{-\mu} \sum_{n=1}^{\infty} \frac{\mu^n \sqrt{n}}{2^n n!} \tag{20}$$

For the transmitted arm (Eve's side), we obtain the following detection statistics (detailed calculations are presented in Appendix C):

$$\begin{aligned} &p_T(0|0) = p_T(1|1) = p_\mu(n \geq 2) \\ &p_T(1|0) = p_T(0|1) = 0 \end{aligned} \tag{21}$$

$$p_T(+|0) = p_T(-|0) = p_T(+|1) = p_T(-|1) = \frac{1}{2}p_\mu(n \geq 2) \tag{22}$$

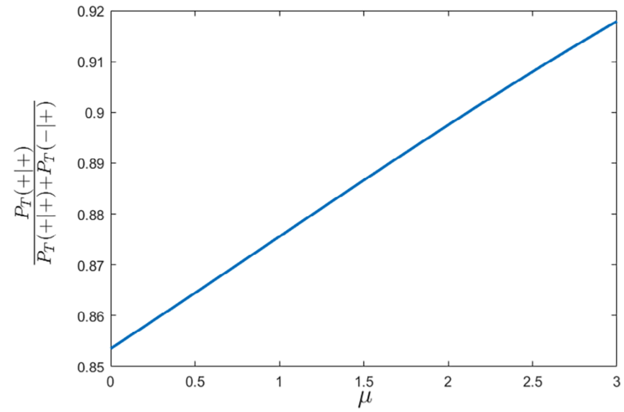

Figure 3. Ratio of correct bit detections at Eve's side, for x-basis states, following a SPRINT-based PNS attack.

for z-basis states, and for x-basis states:

$$p_T(0|+) = p_T(0|-) = \left(1 + e^{\frac{\mu}{2}}\right)A$$
$$p_T(1|+) = p_T(1|-) = \left(1 + e^{\frac{\mu}{2}}\right)A + p_{\frac{\mu}{2}}(n=1) - \tfrac{1}{2}p_\mu(n=1) \tag{23}$$

$$p_T(+|+) = p_T(-|-) = A + B_{m-1-K_A+K_B} + C \tag{24}$$

$$p_T(-|+) = p_T(+|-) = A + B_{n+j-m-K_B+K_A} + \tilde{C} \tag{25}$$

where the definitions of A, B_X, C, and $\tilde{C}$ are given in Appendix C.

The expressions for the detection probabilities of the x-basis states are rather complex. To gain some insight into these expressions, we have simulated $\frac{p_T(+|+)}{p_T(+|+)+p_T(-|+)}$, i.e., the ratio of correct bit detections to the total detection events at Eve's side, by truncating the sums (appearing in the definitions of A, B_X, C and $\tilde{C}$ to $n,j \leq 10$. This simulation result is given in **Figure 3**, showing that at least 85% of Eve's detections in the x-basis are correct, and the percentage increases with increasing the average photon number, μ.

4.2. QBER and Eve's Information Gain

The analytic results presented in the previous section can be used to calculate various interesting quantities. First, we look at the QBER introduced by the attack. Comparing Equation (7) with Equation (16), it is clear that the SPRINT-based attack has no effect on Bob's detection statistics for z-basis states, and the QBER for these states is zero. However, Equation (20) shows that for x-basis states there is a non-zero probability that Bob will get a wrong detection. Assuming Alice chooses between z-basis and x-basis with equal probability of 0.5, the QBER introduced by the SPRINT-based PNS attack is thus:

$$QBER = \frac{1}{2}\zeta\frac{p_R(-|+) + p_R(+|-)}{p_R(+|+) + p_R(-|+) + p_R(+|-) + p_R(-|-)} \tag{26}$$

where we have taken into account that the QBER is introduced only for the ratio ζ of bits for which Eve couples the SPRINT-system to the quantum communication channel, as discussed in Section 3.1. A graph of the QBER as a function of μ is presented in **Figure 4**. We note that the linear-like trend of Figures 3 and 4 is a bit deceiving – the behavior is only approximately linear for small values of μ because of the involved exponents. It can be readily verified that the linear approximation ceases to be valid for larger μ values, eventually taking the plot in Figure 3 to 1, and the $\zeta = 1$ plot in Figure 4 to 0.25 as expected.

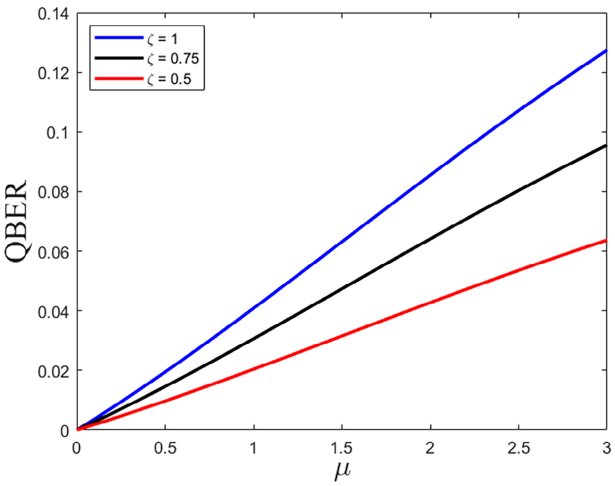

Figure 4. QBER introduced by SPRINT-based PNS attack, as a function of the average photon number, μ, for different values of ζ (the ratio of bits for which Eve couples the SPRINT-system to the quantum communication channel). In theoretical studies of the PNS attack, the QBER is usually assumed to be zero for all μ.

We now move to the detection statistics at Eve's side. Here also, we see that for z-basis states (Equation (21)) Eve gets full information whenever there is a multi-photon pulse, as was previously assumed in theoretical studies of PNS,[5] but for x-basis states she may sometime get a wrong detection (as shown in Figure 3). We are thus interested to calculate the probability for Eve to obtain the correct bit value given that Bob got a detection at his side for that bit. We call this quantity Eve's information gain, denote it with G, and it is clear that $0.5 \leq G \leq 1$, with the information gain obtained by simple I-R attack being 0.75. Note that in our implementation, Bob gets a detection whenever the pulse is not empty, $p_\mu(n \geq 1)$, and out of these cases, Eve will get a detection only if there are two or more photons in the pulse, $p_\mu(n \geq 2)$, and the SPRINT-system is coupled to the quantum communication channel. As mentioned above, we assume that Eve is non-aggressive, and so for the cases of single-photon pulses or when the SPRINT-system is not coupled to the channel, she just guesses the bit value.

The calculation of Eve's information gain depends on whether Eve has a quantum memory or not. Assuming she has a quantum memory, then for every pulse she manages to split she keeps her photons until Alice and Bob announce their bases and then she measures her photons in the correct basis. The expression for Eve's information gain in that case is:

$$G_M = \frac{\begin{array}{c}\frac{1}{2}p_\mu(n=1) + \frac{1}{2}(1-\zeta)p_\mu(n \geq 2) \\ +\frac{1}{4}\zeta\left[p_T(+|+) + p_T(-|-) + p_T(0|0) + p_T(1|1)\right]\end{array}}{p_\mu(n \geq 1)} \tag{27}$$

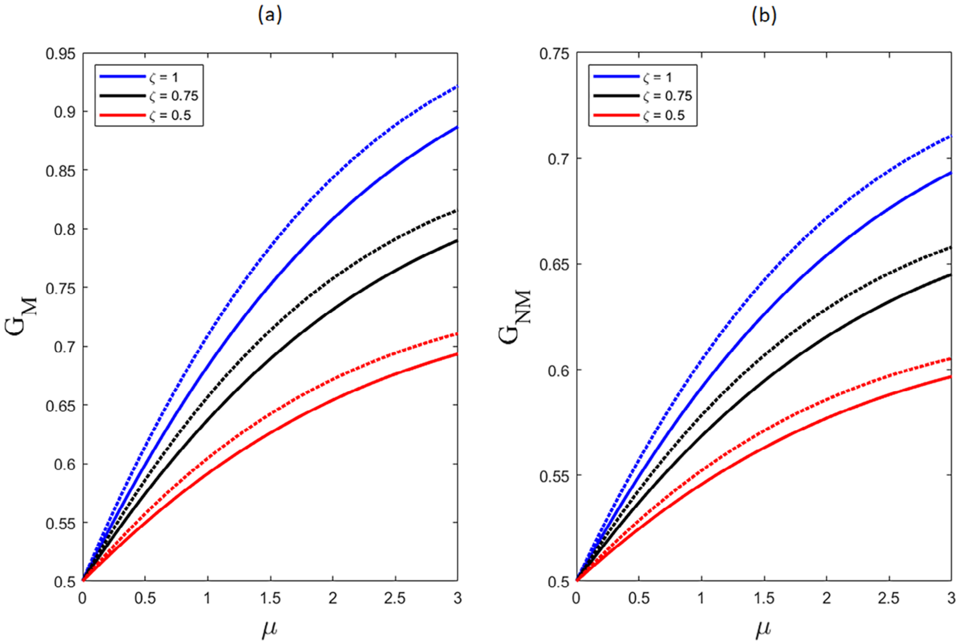

Figure 5. Solid line - Eve's information gain, (a) with and (b) without quantum memory, as a function of the average photon number μ, for different values of coupling ratio ζ. Dashed line - The theoretical limit of the information gain. In (b) we assume Eve randomly chooses the basis for her measurements.

where we assume that Alice prepares the states with equal probability of 0.25.

If Eve does not have a quantum memory, she randomly chooses the basis for her measurements. As can be seen from Equations (22) and (23), whenever Eve's detection is in the wrong basis she gets no information about the original state. Thus we assume that after the announcement of the bases by Alice, Eve chooses a random bit value for the bits she measured in the wrong basis (half of the cases, on average). Therefore, the expression for her information gain is:

$$G_{NM} = \frac{\begin{array}{c}\frac{1}{2}p_\mu(n=1) + \frac{1}{2}(1-\zeta)p_\mu(n\geq 2) + \frac{1}{4}\zeta p_\mu(n\geq 2)\\ +\frac{1}{8}\zeta\left[p_T(+|+) + p_T(-|-) + p_T(0|0) + p_T(1|1)\right]\end{array}}{p_\mu(n\geq 1)} \tag{28}$$

Figure 5a,b shows Eve's information gain with and without quantum memory, respectively, as a function of μ. In theoretical studies of the PNS attack, it is assumed that Eve gets full information whenever she is successful in extracting a photon ($\zeta p_\mu(n\geq 2)$). Thus, the maximal theoretical information gain is given by Equations (27) and (29) by inserting $p_T(+|+) = p_T(-|-) = p_\mu(n\geq 2)$. We show the theoretical limit, for both G_M and G_{NM}, as a dashed line in Figure 5.

Further investigation of Eve's detection statistics reveals that, if Eve is not in possession of a quantum memory, she can adopt a different strategy – Instead of randomly choosing the basis for her measurements, she measures all of the qubits in the z-basis. For half of the qubits, the z-basis is the correct basis, and so for half of the qubits Eve will get full information (Equation (21)), making it an *optimal strategy*, reaching the theoretical limit of Eve's information gain without quantum memory. The expression for Eve's information gain with that strategy is:

$$G_{NM}^{z-basis} = \frac{\begin{array}{c}\frac{1}{2}p_\mu(n=1) + \frac{1}{2}(1-\zeta)p_\mu(n\geq 2)\\ +\frac{1}{4}\zeta p_\mu(n\geq 2) + \frac{1}{4}\zeta\left[p_T(0|0) + p_T(1|1)\right]\end{array}}{p_\mu(n\geq 1)} \tag{29}$$

and a graph of it, as a function of μ, is presented in **Figure 6**.

5. Discussion

The detection statistics at Bob's side following a SPRINT-based PNS attack, Equations (16)–(20), show that z-basis states are un-

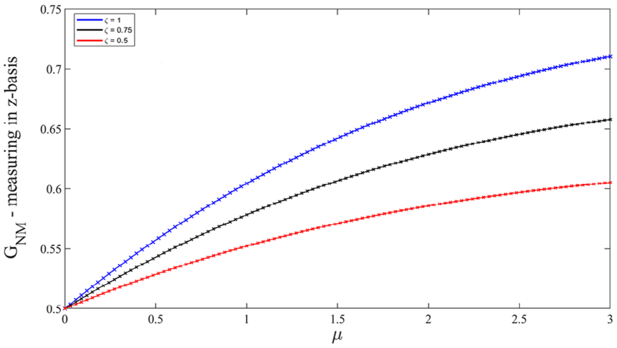

Figure 6. Eve's information gain when measuring in the z-basis only, as a function of the average photon number μ, for different values of coupling ratio ζ. Evidently, this is the optimal strategy for Eve if she does not possess a quantum memory, with her information gain reaching the theoretical limit (plotted as × marks).

affected by the attack, but for x-basis states some QBER is introduced. As presented in Figure 4, the resulting QBER is rather small (about 4% for $\mu \approx 1$), compared to the 25% QBER introduced by a simple I-R attack. The fact that the QBER increases with μ is expected, since as μ increases, so does the probability of Eve successfully splitting the pulse ($p_\mu(n \geq 2)$) and affecting the state reaching Bob's side.

In theoretical studies of PNS attack it was usually assumed that the attack introduces no QBER at all.[3,5,11] However, this assumption was challenged by some researchers,[37] claiming that, according to the laws of quantum mechanics, any unitary operator capable of (deterministically) extracting a photon out of a z-basis state without altering the state of the remaining photons, will necessarily be unable to do so for x-basis states. While, obviously, our results cannot be seen as a proof for this claim, they are surely supporting it, since they show that our implementation indeed fails to split x-basis states without affecting them, even though it does it perfectly for z-basis states.

The detection statistics at Bob's side reveal another interesting point, namely the breaking of symmetry between $p_R(0|\pm)$ and $p_R(1|\pm)$ (Equation (18)). This asymmetry is expected since an occupied early pulse changes the state of the SPRINT-system so that photons from the late pulse will not be reflected, and so the probability of Bob getting a detection is higher for the early time-bin. We note that this feature can be exploited by Alice and Bob to check the security of their communication against SPRINT-based PNS attack – Instead of immediately discarding all of the measurements where their bases were different, they keep all of the cases where an x-basis state was sent and a z-basis measurement was done. For these cases, they calculate the ratio of '0' detections to '1' detections. This ratio should be close to 1, as implied by Equation (8), and if it is higher than 1 it might be an indication that a PNS attack was carried out, causing the asymmetry in detections. Mathematically, the expected ratio following a SPRINT-based PNS attack is:

$$R = \frac{\#\text{`0'}}{\#\text{`1'}} = (1-\zeta) + \zeta \frac{p_R(0|\pm)}{p_R(1|\pm)} \tag{30}$$

This ratio is shown in **Figure 7**, as a function of μ.

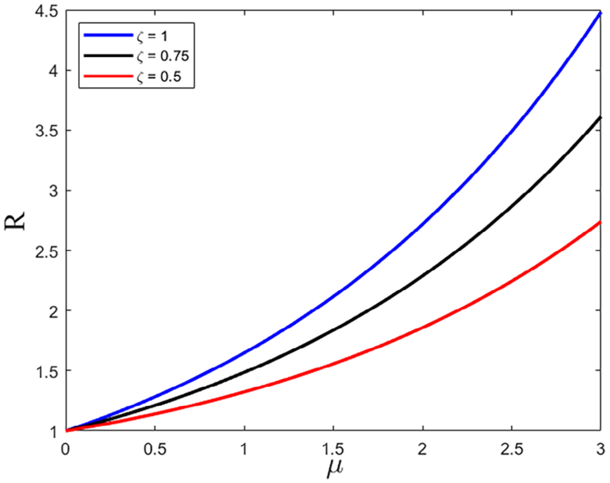

Figure 7. R, the ratio of '0' and '1' detections at Bob's side, following a SPRINT-based PNS attack, for all the cases where an x-basis state was sent and a z-basis measurement was done. The ratio is shown as a function of the average photon number μ, for different values of coupling ratio ζ.

To avoid this breaking of symmetry, Eve can try to reset the SPRINT-system to the reflecting state $|R\rangle$ by sending a strong pulse propagating from right to left, in between the early and late time-bins. That way the symmetry is restored since regardless of whether the early time-bin was occupied or not, the SPRINT-system gets back to the $|R\rangle$ state before the late time-bin arrives. However, in Appendix D we analyze this possibility and show that resetting the SPRINT-system in between the time-bins completely destroys the coherence of the x-basis states, resulting in $p_R(+|\pm) = p_R(-|\pm) = \frac{1}{2}p_\mu(n \geq 1)$, yielding a QBER of 25% (like a simple I-R attack) which will clearly be detected by Alice and Bob.

This result also implies the robustness of the COW protocol[10] against our suggested implementation of the PNS attack. The key point in COW is that the coherence is checked not only between the early and late time-bins, but also between the late time-bin of the previous qubit and the early time-bin of the current one. Since for the SPRINT-based PNS attack Eve must reset the system before each qubit, she inevitably destroys the coherence between the late time-bin of the previous qubit and the early time-bin of the current one, and that will be detected by Alice and Bob.

Looking at Eve's detection statistics, Equations (21)–(25), we again see that the attack works perfectly in the z-basis, with Eve getting full information whenever there is a multi-photon pulse ($p_\mu(n \geq 2)$), while for x-basis state Eve has some non-zero probability of getting the wrong result, supporting the claim of ref. [37] discussed earlier. Nevertheless, the information gain, Figure 5, is not too far from the theoretical limit, and for the case where Eve has no quantum memory, she can in fact obtain the optimal theoretical limit (Figure 6) by measuring only in the z-basis, as explained in Section 4.2. For $\mu \approx 1$ we have $G_M \approx 0.7$ and $G_{NM}^{z-basis} \approx 0.6$, and as μ increases Eve's information gain also increases, since the probability for a multi-photon pulse that can be split is higher. This should be compared to the information gain obtained in a simple I-R attack, 0.75, but is accompanied by a QBER of 25%. We note that Eve's information gain can be used for calculating the mutual information between her and Alice, see Appendix E.

To complete our discussion of the discrepancies between our suggested implementation and previous theoretical works, we point out two more differences – First, our implementation of the PNS attack is done in a single step, unlike the common theoretical description of the attack in which two steps are needed, i.e., a QND measurement of the number of photons followed by splitting of the pulse.[5] Second, in our proposed implementation, from all multi-photon pulses Eve steals all but one of the photons, as suggested by some theoretical studies of PNS,[38] while others assumed that Eve steals only a single photon.[5] We note that extracting more than a single photon can be advantageous for Eve, especially if she has no quantum memory – She can use another SPRINT-system to split all multi-photon pulses in her side and measure in both bases, such that after the classical communication between Alice and Bob she knows which basis was the correct one and obtain the information.

6. Conclusion

In this paper, we addressed a time-bin encoded QKD protocol and proposed an implementation of a PNS attack using the already-demonstrated SPRINT-system, under the assumptions that it is loss-free and perfectly coupled to the quantum communication channel. While, to the best of our knowledge, current experimental SPRINT-systems are yet to achieve these high standards, their imperfections are not fundamental, and it is reasonable to expect nearly perfect SPRINT-systems with the advancement of cavity quantum electrodynamics technologies.[19] Therefore, an understanding of the vulnerability of QKD to SPRINT-based PNS attack is highly desirable.

To that end, we presented a comprehensive analysis of the attack. We calculated Eve's information gain, proving that the SPRINT-system effectively separates each multi-photon pulse, granting Eve access to the information contained in it. Unlike previous theoretical studies of the PNS attack, our results show that for x-basis states Eve's information gain is not complete, and the QBER introduced by the attack is small but not zero. More generally, we postulate that the presence of non-zero QBER stems from the nonlinearity of the PNS attack, and is thus expected to be inherent to the attack, regardless of the specific implementation. These results raise new questions regarding the relationship between the separation of photons in an unknown quantum state and the information they contain. In future research, it could also be instructive to examine possible connections of PNS with quantum erasure[39–41] and quantum oblivion.[42,43] In all these effects, one refrains from recording all the available quantum information in order to allow for interference to occur.

It should be emphasized that in this initial analysis of SPRINT-based PNS attack, we assumed Alice and Bob's implementation of the QKD protocol to be perfect, i.e., using a nearly lossless quantum communication channel and perfect detectors. For real-world security analysis these elements should be taken into consideration. Specifically, if the original quantum communication channel is lossy, Eve can replace it with a lossless one, and then use the accessible loss[12] to block some of the single-photon pulses, increasing her overall information gain beyond what we have described herein.

Our results imply that an almost ideal PNS attack is in fact plausible with contemporary technologies. However, there are several known ways to defend against PNS attack, either by using more advanced QKD protocols,[6–10] or by using advanced technologies such as photon-number-resolving detectors. Rather, our results underscore the importance of adopting these countermeasures against PNS attacks to uphold the security of QKD systems.

Appendix A: Time-Bin Encoded BB84 with WCS

Here, we present the calculations of the detection statistics of time-bin encoded BB84 implemented with WCS, without the presence of an eavesdropper.

Replacing the single-photon source in Figure 1 with WCS, the state entering the preparation device is thus $|\alpha\rangle_e\,|0\rangle_l$. It is straightforward to verify that Alice's preparation results in the following z-basis states:

$$\begin{aligned} |0\rangle &= |\alpha\rangle_e\,|0\rangle_l \\ |1\rangle &= |0\rangle_e\,|\alpha\rangle_l \end{aligned} \tag{A1}$$

and x-basis states:

$$|\pm\rangle = |\frac{\alpha}{\sqrt{2}}\rangle_e\,|\frac{\pm\alpha}{\sqrt{2}}\rangle_l \tag{A2}$$

Then, if Bob chooses to detect in the z-basis, we get the following probabilities:

$$\begin{aligned} p(0|0) &= \langle\alpha|_e\,\langle 0|_l\left(P_e\otimes|0\rangle_l\,\langle 0|_l+\tfrac{1}{2}P_e\otimes P_l\right)|\alpha\rangle_e\,|0\rangle_l = 1-e^{-\mu} = p(1|1) \\ p(1|0) &= \langle\alpha|_e\,\langle 0|_l\left(|0\rangle_e\,\langle 0|_e\otimes P_l+\tfrac{1}{2}P_e\otimes P_l\right)|\alpha\rangle_e\,|0\rangle_l = 0 = p(0|1) \end{aligned} \tag{A3}$$

and:

$$\begin{aligned} p(0|\pm) &= \langle\tfrac{\alpha}{\sqrt{2}}|_e\langle\tfrac{\pm\alpha}{\sqrt{2}}|_l\Big(P_e\otimes|0\rangle_l\,\langle 0|_l+\tfrac{1}{2}P_e\otimes P_l\Big)|\tfrac{\alpha}{\sqrt{2}}\rangle_e\,|\tfrac{\pm\alpha}{\sqrt{2}}\rangle_l \\ &= \left(1-e^{-\frac{\mu}{2}}\right)e^{-\frac{\mu}{2}}+\tfrac{1}{2}\left(1-e^{-\frac{\mu}{2}}\right)^2 \\ &= \tfrac{1}{2}\left(1-e^{-\frac{\mu}{2}}\right)\left(1+e^{-\frac{\mu}{2}}\right) = \tfrac{1}{2}(1-e^{-\mu}) = p(1|\pm) \end{aligned} \tag{A4}$$

where $P_x = (I_x - |0\rangle_x\,\langle 0|_x)$ is the description of the detector as a projection on the occupied states of the mode x, and where double clicks are assigned random bit values (as discussed in the main text).

If Bob detects in the x-basis, the evolution of the states prepared by Alice through the measuring device is:

$$\begin{aligned} |0\rangle &\rightarrow |\tfrac{\alpha}{\sqrt{2}}\rangle_+\,|\tfrac{i\alpha}{\sqrt{2}}\rangle_- \\ |1\rangle &\rightarrow |\tfrac{\alpha}{\sqrt{2}}\rangle_+\,|\tfrac{-i\alpha}{\sqrt{2}}\rangle_- \\ |+\rangle &\rightarrow |\alpha\rangle_+\,|0\rangle_- \\ |-\rangle &\rightarrow |0\rangle_+\,|i\alpha\rangle_- \end{aligned} \tag{A5}$$

and so we get:

$$\begin{aligned} p(+|+) &= \langle\alpha|_+\,\langle 0|_-\left(P_+\otimes|0\rangle_-\,\langle 0|_-+\tfrac{1}{2}P_+\otimes P_-\right)|\alpha\rangle_+\,|0\rangle_- \\ &= 1-e^{-\mu} = p(-|-) \\ p(-|+) &= \langle\alpha|_+\,\langle 0|_-\left(|0\rangle_+\,\langle 0|_+\otimes P_-+\tfrac{1}{2}P_+\otimes P_-\right)|\alpha\rangle_+\,|0\rangle_- \\ &= 0 = p(+|-) \end{aligned} \tag{A6}$$

and:

$$\begin{aligned} p(+|0) &= \langle\tfrac{\alpha}{\sqrt{2}}|_+\langle\tfrac{i\alpha}{\sqrt{2}}|_-\Big(P_+\otimes|0\rangle_-\,\langle 0|_-+\tfrac{1}{2}P_+\otimes P_-\Big)|\tfrac{\alpha}{\sqrt{2}}\rangle_+\,|\tfrac{i\alpha}{\sqrt{2}}\rangle_- \\ &= \tfrac{1}{2}(1-e^{-\mu}) = p(+|1) \\ p(-|0) &= \langle\tfrac{\alpha}{\sqrt{2}}|_+\langle\tfrac{i\alpha}{\sqrt{2}}|_-\Big(|0\rangle_+\,\langle 0|_+\otimes P_-+\tfrac{1}{2}P_+\otimes P_-\Big)|\tfrac{\alpha}{\sqrt{2}}\rangle_+\,|\tfrac{i\alpha}{\sqrt{2}}\rangle_- \\ &= \tfrac{1}{2}(1-e^{-\mu}) = p(-|1) \end{aligned} \tag{A7}$$

Appendix B: Detection Statistics in the Reflected Arm

Here, we calculate the detection statistics in the reflected arm (Bob's side) for the case of time-bin encoded BB84 under SPRINT-based PNS attack. The analysis steps are – i) Alice's preparation of the state, ii) the SPRINT-system operates on the early time-bin, $\hat{s}_e$, iii) the SPRINT-system operates on the late time-bin, $\hat{s}_l$, iv) the SPRINT-system and the transmission output arm are traced over to obtain ρ_R, and finally v) detection in z-basis or x-basis.

Let us first look at the case where Alice sends out $|0\rangle$. We get:

$$(i)\ |0\rangle \Rightarrow |\psi\rangle = |\alpha, 0, R, 0, 0\rangle \tag{B1}$$

$$(ii), (iii)\ |\psi\rangle = e^{-\frac{\mu}{2}}\left(|0, 0, R, 0, 0\rangle + \sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n-1, 1, T, 0, 0\rangle\right) \tag{B2}$$

$$\begin{aligned}(iv)\rho = |\psi\rangle\langle\psi| = e^{-\mu}[&|0, 0, R, 0, 0\rangle\langle 0, 0, R, 0, 0|\\ &+\textstyle\sum_{n=1}^{\infty}\left(\frac{\alpha^n}{\sqrt{n!}}|n-1, 1, T, 0, 0\rangle\langle 0, 0, R, 0, 0|\right.\\ &\left.+\frac{(\alpha^*)^n}{\sqrt{n!}}|0, 0, R, 0, 0\rangle\langle n-1, 1, T, 0, 0|\right)\\ &+\textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{n!m!}}|n-1, 1, T, 0, 0\rangle\langle m-1, 1, T, 0, 0|\Big]\end{aligned} \tag{B3}$$

$$\begin{aligned}\rho_R = tr_{T_e,S,T_l}(\rho) = e^{-\mu}[&|0, 0\rangle\langle 0, 0|\langle 0, R, 0\,|\,0, R, 0\rangle\\ &+\textstyle\sum_{n=1}^{\infty}\left(\frac{\alpha^n}{\sqrt{n!}}|1, 0\rangle\langle 0, 0|\langle 0, R, 0|n-1, T, 0\rangle\right.\\ &\left.+\frac{(\alpha^*)^n}{\sqrt{n!}}|0, 0\rangle\langle 1, 0|\langle n-1, T, 0|0, R, 0\rangle\right)\\ &+\textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{n!m!}}|1, 0\rangle\langle 1, 0|\langle m-1, T, 0|n-1, T, 0\rangle\Big]\end{aligned} \tag{B4}$$

$$\begin{aligned}\Rightarrow \rho_R &= e^{-\mu}\left(|0, 0\rangle\langle 0, 0| + \textstyle\sum_{n=1}^{\infty}\frac{\mu^n}{n!}|1, 0\rangle\langle 1, 0|\right)\\ &= e^{-\mu}|0, 0\rangle\langle 0, 0| + (1-e^{-\mu})|1, 0\rangle\langle 1, 0|\end{aligned}$$

i.e., in the reflected arm we get a statistical mixture of vacuum (with probability of $e^{-\mu}$) and a single-photon state (with probability of $1-e^{-\mu}$). As we could expect.

It follows immediately that:

$$\begin{aligned}p_R(0|0) &= tr\left(\rho_R(P_e\otimes|0\rangle_l\langle 0|_l + \tfrac{1}{2}P_e\otimes P_l)\right) = 1-e^{-\mu}\\ p_R(1|0) &= tr\left(\rho_R(|0\rangle_e\langle 0|_e\otimes P_l + \tfrac{1}{2}P_e\otimes P_l)\right) = 0\end{aligned} \tag{B5}$$

When measuring in the x-basis, the evolution through the measuring device is $|1, 0\rangle \rightarrow \frac{|1,0\rangle+i|0,1\rangle}{\sqrt{2}}$ and $|0, 1\rangle \rightarrow \frac{|1,0\rangle-i|0,1\rangle}{\sqrt{2}}$. Thus, if the reflected arm is measured in the x-basis, the density matrix evolves in the measuring device to:

$$\rho_R = e^{-\mu}|0, 0\rangle\langle 0, 0| + (1-e^{-\mu})\frac{(|1, 0\rangle + i\,|0, 1\rangle)(\langle 1, 0| - i\,\langle 0, 1|)}{2} \tag{B6}$$

And so:

$$\begin{aligned}p_R(+|0) &= tr\left(\rho_R(P_+\otimes|0\rangle_-\langle 0|_- + \tfrac{1}{2}P_+\otimes P_-)\right) = \tfrac{1}{2}(1-e^{-\mu})\\ p_R(-|0) &= tr\left(\rho_R(|0\rangle_+\langle 0|_+\otimes P_- + \tfrac{1}{2}P_+\otimes P_-)\right) = \tfrac{1}{2}(1-e^{-\mu})\end{aligned} \tag{B7}$$

It is straightforward to generalize the above analysis to the case in which Alice sends out a $|1\rangle$ state, yielding:

$$\begin{aligned}|1\rangle \Rightarrow |\psi\rangle &= |0, 0, R, \alpha, o\rangle \Rightarrow \rho_R x\\ &= e^{-\mu}|0, 0\rangle\langle 0, 0| + (1-e^{-\mu})|0, 1\rangle\langle 0, 1|\\ p_R(1|1) &= 1-e^{-\mu}\\ p_R(0|1) &= 0\\ p_R(+|1) &= p_R(-|1) = \tfrac{1}{2}(1-e^{-\mu})\end{aligned} \tag{B8}$$

We continue by analysing the case in which Alice prepares the state $|+\rangle$. The evolution of the state is:

$$(i)\ |+\rangle \Rightarrow |\psi\rangle = |\frac{\alpha}{\sqrt{2}}, 0, R, \frac{\alpha}{\sqrt{2}}, 0\rangle \tag{B9}$$

$$(ii)\ |\psi\rangle = e^{-\frac{\mu}{4}}\left(|0, 0, R, \frac{\alpha}{\sqrt{2}}, 0\rangle + \sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^n n!}}|n-1, 1, T, \frac{\alpha}{\sqrt{2}}, 0\rangle\right) \tag{B10}$$

$$\begin{aligned}(iii)\ |\psi\rangle = &e^{-\frac{\mu}{2}}\left(|0, 0, R, 0, 0\rangle + \sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^n n!}}|0, 0, T, n-1, 1\rangle\right)\\ &+e^{-\frac{\mu}{4}}\sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^n n!}}|n-1, 1, T, \frac{\alpha}{\sqrt{2}}, 0\rangle\end{aligned} \tag{B11}$$

$$\begin{aligned}(iv)\rho_R = tr_{T_e,S,T_l}(|\psi\rangle\langle\psi|) = &e^{-\mu}\left(|0, 0\rangle\langle 0, 0| + \textstyle\sum_{n=1}^{\infty}\frac{\mu^n}{2^n n!}|0, 1\rangle\langle 0, 1|\right)\\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n\alpha^{*m}}{\sqrt{2^{n+m}n!m!}}\langle m-1, T, \frac{\alpha}{\sqrt{2}}|0, T, n-1\rangle|0, 1\rangle\langle 1, 0|\\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^{*n}\alpha^m}{\sqrt{2^{n+m}n!m!}}\langle 0, T, n-1|m-1, T, \frac{\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 0, 1|\\ &+e^{-\frac{\mu}{2}}\textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n\alpha^{*m}}{\sqrt{2^{n+m}n!m!}}\langle m-1, T, \frac{\alpha}{\sqrt{2}}|n-1, T, \frac{\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 1, 0|\end{aligned} \tag{B12}$$

$$\begin{aligned}\Rightarrow \rho_R = &e^{-\mu}\left(|0, 0\rangle\langle 0, 0| + (e^{\frac{\mu}{2}}-1)|0, 1\rangle\langle 0, 1|\right)\\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n=1}^{\infty}\frac{\alpha^n\alpha^*}{\sqrt{2^{n+1}n!}}\langle\frac{\alpha}{\sqrt{2}}|n-1\rangle|0, 1\rangle\langle 1, 0|\\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n=1}^{\infty}\frac{\alpha^{*n}\alpha}{\sqrt{2^{n+1}n!}}\langle n-1|\frac{\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 0, 1|\\ &+e^{-\frac{\mu}{2}}\textstyle\sum_{n=1}^{\infty}\frac{\mu^n}{2^n n!}\langle\frac{\alpha}{\sqrt{2}}|\frac{\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 1, 0|\end{aligned} \tag{B13}$$

$$\begin{aligned}\Rightarrow \rho_R = &e^{-\frac{\mu}{2}}\left(e^{-\frac{\mu}{2}}|0, 0\rangle\langle 0, 0| + (1-e^{-\frac{\mu}{2}})|0, 1\rangle\langle 0, 1|\right)\\ &+e^{-\mu}\textstyle\sum_{n=1}^{\infty}\frac{\mu^n\sqrt{n}}{2^n n!}(|0, 1\rangle\langle 1, 0| + |1, 0\rangle\langle 0, 1|)\\ &+(1-e^{-\frac{\mu}{2}})|1, 0\rangle\langle 1, 0|\end{aligned} \tag{B14}$$

We can now make sense of this result – the diagonal terms are the cases where a) both pulses are vacuum (we get $|0, 0\rangle\langle 0, 0|$), b) first is vacuum and second is occupied ($|0, 1\rangle\langle 0, 1|$), and c) first is occupied ($|1, 0\rangle\langle 1, 0|$), each case with its respective probability. We also get non-diagonal terms representing coherence between the early and late modes of the reflected arm. It is simple to see that ρ_R is indeed Hermitian and has a trace of 1, as needed.

It follows that:

$$\begin{aligned}p_R(0|+) &= tr\left(\rho_R(P_e\otimes|0\rangle_l\langle 0|_l + \tfrac{1}{2}P_e\otimes P_l)\right) = 1-e^{-\frac{\mu}{2}}\\ p_R(1|+) &= tr\left(\rho_R(|0\rangle_e\langle 0|_e\otimes P_l + \tfrac{1}{2}P_e\otimes P_l)\right) = e^{-\frac{\mu}{2}}(1-e^{-\frac{\mu}{2}})\end{aligned} \tag{B15}$$

If the reflected arm is measured in the x-basis, then the density matrix evolves in the measuring device to:

$$\begin{aligned}\rho_R = &e^{-\frac{\mu}{2}}\left(e^{-\frac{\mu}{2}}|0, 0\rangle\langle 0, 0| + (1-e^{-\frac{\mu}{2}})\frac{(|1,0\rangle-i|0,1\rangle)(\langle 1,0|+i\langle 0,1|)}{2}\right)\\ &+e^{-\mu}\textstyle\sum_{n=1}^{\infty}\frac{\mu^n\sqrt{n}}{2^n n!}\left(\frac{(|1,0\rangle-i|0,1\rangle)(\langle 1,0|-i\langle 0,1|)+(|1,0\rangle+i|0,1\rangle)(\langle 1,0|+i\langle 0,1|)}{2}\right)\\ &+(1-e^{-\frac{\mu}{2}})\frac{(|1,0\rangle+i|0,1\rangle)(\langle 1,0|-i\langle 0,1|)}{2}\end{aligned} \tag{B16}$$

We note that ρ_R is indeed Hermitian and has a trace of 1, as needed.

And so:

$$\begin{aligned} p_R(+|+) &= tr\Big(\rho_R(P_+ \otimes |0\rangle_- \langle 0|_- + \tfrac{1}{2}P_+ \otimes P_-)\Big) \\ &= \tfrac{1}{2}e^{-\frac{\mu}{2}}(1-e^{-\frac{\mu}{2}}) + e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} + \tfrac{1}{2}(1-e^{-\frac{\mu}{2}}) \\ &= \tfrac{1}{2}(1-e^{-\mu}) + e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} \end{aligned} \tag{B17}$$

It is obvious to see that $p(+|+) \geq 0$, but let us also verify that this probability is indeed smaller than 1 for all μ. Defining $x \equiv \frac{\mu}{2}$, we have:

$$\begin{aligned} &\tfrac{1}{2}(1-e^{-2x}) + e^{-2x}\textstyle\sum_{n=1}^{\infty} \frac{x^n\sqrt{n}}{n!} \leq \tfrac{1}{2}(1-e^{-2x}) \\ &+e^{-2x}\textstyle\sum_{n=1}^{\infty} \frac{x^n n}{n!} = \tfrac{1}{2}(1-e^{-2x}) + e^{-2x}xe^{x} \\ &= \tfrac{1}{2}(1-e^{-2x}) + xe^{-x} \leq \tfrac{1}{2}(1-e^{-2x}) + e^{-1} \leq \tfrac{1}{2} + e^{-1} < 1 \end{aligned} \tag{B18}$$

We also get:

$$\begin{aligned} p_R(-|+) &= tr\Big(\rho_R(|0\rangle_+ \langle 0|_+ \otimes P_- + \tfrac{1}{2}P_+ \otimes P_-)\Big) \\ &= \tfrac{1}{2}e^{-\frac{\mu}{2}}(1-e^{-\frac{\mu}{2}}) - e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} + \tfrac{1}{2}(1-e^{-\frac{\mu}{2}}) \\ &= \tfrac{1}{2}(1-e^{-\mu}) - e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} \end{aligned} \tag{B19}$$

Here also, it is simple to see that $p(-|+) \leq p(+|+) \leq 1$, but let us verify also that it is non-negative for all μ. Again we define $x \equiv \frac{\mu}{2}$, thus:

$$\begin{aligned} &\tfrac{1}{2}(1-e^{-2x}) - e^{-2x}\textstyle\sum_{n=1}^{\infty} \frac{x^n\sqrt{n}}{n!} \geq \tfrac{1}{2}(1-e^{-2x}) \\ &-e^{-2x}\textstyle\sum_{n=1}^{\infty} \frac{x^n n}{n!} = \tfrac{1}{2}(1-e^{-2x}) - xe^{-x} \geq 0 \end{aligned} \tag{B20}$$

Finally, we can verify that the probabilities sum up to 1, i.e.:

$$\begin{aligned} &p_R(+|+) + p_R(-|+) + p_R(00|+) = \tfrac{1}{2}(1-e^{-\mu}) + e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} \\ &+\tfrac{1}{2}(1-e^{-\mu}) - e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} + e^{-\mu} = 1 \end{aligned} \tag{B21}$$

where $p_R(00)$ is the probability of getting no detection in both detectors.

Generalizing to the case of $|-\rangle = |\frac{\alpha}{\sqrt{2}}, \frac{-\alpha}{\sqrt{2}}\rangle$:

$$(i)\ |-\rangle \Rightarrow |\psi\rangle = |\frac{\alpha}{\sqrt{2}}, 0, R, \frac{-\alpha}{\sqrt{2}}, 0\rangle \tag{B22}$$

$$(ii)\ |\psi\rangle = e^{-\frac{\mu}{4}}\left(|0, 0, R, \frac{-\alpha}{\sqrt{2}}, 0\rangle + \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{2^n n!}}|n-1, 1, T, \frac{-\alpha}{\sqrt{2}}, 0\rangle\right) \tag{B23}$$

$$\begin{aligned} (iii)\ |\psi\rangle = e^{-\frac{\mu}{2}}&\left(|0, 0, R, 0, 0\rangle + \sum_{n=1}^{\infty} \frac{(-\alpha)^n}{\sqrt{2^n n!}}|0, 0, T, n-1, 1\rangle\right) \\ &+ e^{-\frac{\mu}{4}}\sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{2^n n!}}|n-1, 1, T, \frac{-\alpha}{\sqrt{2}}, 0\rangle \end{aligned} \tag{B24}$$

$$\begin{aligned} (iv)\rho_R = tr_{T_e,S,T_l}(|\psi\rangle\langle\psi|) &= e^{-\mu}\Big(|0, 0\rangle\langle 0, 0| + \textstyle\sum_{n=1}^{\infty} \frac{\mu^n}{2^n n!}|0, 1\rangle\langle 0, 1|\Big) \\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n,m=1}^{\infty} \frac{(-\alpha)^n\alpha^{*m}}{\sqrt{2^{n+m}n!m!}}\langle m-1, T, \frac{-\alpha}{\sqrt{2}}|0, T, n-1\rangle|0, 1\rangle\langle 1, 0| \\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n,m=1}^{\infty} \frac{(-\alpha^*)^n\alpha^{m}}{\sqrt{2^{n+m}n!m!}}\langle 0, T, n-1|m-1, T, \frac{-\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 0, 1| \\ &+e^{-\frac{\mu}{2}}\textstyle\sum_{n,m=1}^{\infty} \frac{\alpha^n\alpha^{*m}}{\sqrt{2^{n+m}n!m!}}\langle m-1, T, \frac{-\alpha}{\sqrt{2}}|n-1, T, \frac{-\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 1, 0| \end{aligned} \tag{B25}$$

$$\begin{aligned} \Rightarrow \rho_R = e^{-\mu}&\Big(|0, 0\rangle\langle 0, 0| + (e^{\frac{\mu}{2}}-1)|0, 1\rangle\langle 0, 1|\Big) \\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n=1}^{\infty} \frac{(-\alpha)^n\alpha^{*}}{\sqrt{2^{n+1}n!}}\langle\frac{-\alpha}{\sqrt{2}}|n-1\rangle|0, 1\rangle\langle 1, 0| \\ &+e^{-\frac{3\mu}{4}}\textstyle\sum_{n=1}^{\infty} \frac{(-\alpha^*)^n\alpha}{\sqrt{2^{n+1}n!}}\langle n-1|\frac{-\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 0, 1| \\ &+e^{-\frac{\mu}{2}}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n}{2^n n!}\langle\frac{-\alpha}{\sqrt{2}}|\frac{-\alpha}{\sqrt{2}}\rangle|1, 0\rangle\langle 1, 0| \end{aligned} \tag{B26}$$

$$\begin{aligned} &\Rightarrow \rho_R = e^{-\frac{\mu}{2}}\Big(e^{-\frac{\mu}{2}}|0, 0\rangle\langle 0, 0| + (1-e^{-\frac{\mu}{2}})|0, 1\rangle\langle 0, 1|\Big) \\ &\qquad -e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!}(|0, 1\rangle\langle 1, 0| + |1, 0\rangle\langle 0, 1|) \\ &+(1-e^{-\frac{\mu}{2}})|1, 0\rangle\langle 1, 0| \end{aligned} \tag{B27}$$

Such that:

$$\begin{aligned} p_R(0|-) &= tr\Big(\rho_R(P_e \otimes |0\rangle_l\langle 0|_l + \tfrac{1}{2}P_e \otimes P_l)\Big) = 1 - e^{-\frac{\mu}{2}} \\ p_R(1|-) &= tr\Big(\rho_R(|0\rangle_e\langle 0|_e \otimes P_l + \tfrac{1}{2}P_e \otimes P_l)\Big) = e^{-\frac{\mu}{2}}(1-e^{-\frac{\mu}{2}}) \end{aligned} \tag{B28}$$

If the reflected mode is measured in the x-basis, the density matrix evolves to:

$$\begin{aligned} \rho_R = e^{-\frac{\mu}{2}}&\left(e^{-\frac{\mu}{2}}|0, 0\rangle\langle 0, 0| + (1-e^{-\frac{\mu}{2}})\tfrac{(|1,0\rangle - i|0,1\rangle)(\langle 1,0| + i\langle 0,1|)}{2}\right) \\ &-e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!}\left(\tfrac{(|1,0\rangle - i|0,1\rangle)(\langle 1,0| - i\langle 0,1|) + (|1,0\rangle + i|0,1\rangle)(\langle 1,0| + i\langle 0,1|)}{2}\right) \\ &+(1-e^{-\frac{\mu}{2}})\tfrac{(|1,0\rangle + i|0,1\rangle)(\langle 1,0| - i\langle 0,1|)}{2} \end{aligned} \tag{B29}$$

And so:

$$\begin{aligned} p_R(+|-) &= tr\Big(\rho_R(P_+ \otimes |0\rangle_- \langle 0|_- + \tfrac{1}{2}P_+ \otimes P_-)\Big) \\ &= \tfrac{1}{2}e^{-\frac{\mu}{2}}(1-e^{-\frac{\mu}{2}}) - e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} + \tfrac{1}{2}(1-e^{-\frac{\mu}{2}}) \\ &= \tfrac{1}{2}(1-e^{-\mu}) - e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} = p_R(-|+) \end{aligned} \tag{B30}$$

and:

$$\begin{aligned} p_R(-|-) &= tr\Big(\rho_R(|0\rangle_+ \langle 0|_+ \otimes P_- + \tfrac{1}{2}P_+ \otimes P_-)\Big) \\ &= \tfrac{1}{2}e^{-\frac{\mu}{2}}(1-e^{-\frac{\mu}{2}}) + e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} + \tfrac{1}{2}(1-e^{-\frac{\mu}{2}}) \\ &= \tfrac{1}{2}(1-e^{-\mu}) + e^{-\mu}\textstyle\sum_{n=1}^{\infty} \frac{\mu^n\sqrt{n}}{2^n n!} = p_R(+|+) \end{aligned} \tag{B31}$$

Appendix C: Detection Statistics in the Transmitted Arm

Here, we calculate the detection statistics in the transmitted arm (Eve's side) for the case of time-bin encoded BB84 under SPRINT-based PNS attack. The analysis steps are – i) Alice's preparation of the state, ii) the SPRINT-system operates on the early time-bin, $\hat{s}_e$, iii) the SPRINT-system operates on the late time-bin, $\hat{s}_l$, iv) the SPRINT-system and the reflection output arm are traced over to obtain ρ_T, and finally v) detection in z-basis or x-basis.

Let us first look at the case where Alice sends out $|0\rangle$. Using the analysis from Appendix B (Equation (B3)), we get:

$$\begin{aligned}&(i), (ii), (iii)\ |0\rangle \Rightarrow \rho = |\psi\rangle\langle\psi| = e^{-\mu}\Big[|0,0,R,0,0\rangle\langle 0,0,R,0,0| \\ &\quad + \textstyle\sum_{n,m=1}^{\infty} \frac{\alpha^n(\alpha^*)^m}{\sqrt{n!m!}}\ |n-1,1,T,0,0\rangle\langle m-1,1,T,0,0|\Big]\end{aligned} \tag{C1}$$

The evolution of the state, as seen in the transmitted arm, is thus:

$$\begin{aligned}&(iv)\rho_T = tr_{R_e,S,R_l}(\rho) = e^{-\mu}[|0,0\rangle\langle 0,0|\langle 0,R,0\,|\,0,R,0\rangle \\ &\quad + \textstyle\sum_{n=1}^{\infty}\Big(\frac{\alpha^n}{\sqrt{n!}}\ |n-1,0\rangle\langle 0,0|\langle 0,R,0|1,T,0\rangle \\ &\quad + \frac{(\alpha^*)^n}{\sqrt{n!}}\ |0,0\rangle\langle n-1,0|\langle 1,T,0|0,R,0\rangle\Big) \\ &\quad + \textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{n!m!}}\ |n-1,0\rangle\langle m-1,0|\langle 1,T,0|1,T,0\rangle\Big] \\ &\quad \Rightarrow \rho_T = e^{-\mu}\Big(|0,0\rangle\langle 0,0| + \textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{n!m!}}\ |n-1,0\rangle\langle m-1,0|\Big)\end{aligned} \tag{C2}$$

Thus:

$$\begin{aligned}p_T(0|0) &= tr\Big(\rho_T(P_e\otimes|0\rangle_l\langle 0|_l + \tfrac{1}{2}P_e\otimes P_l)\Big) \\ &= tr\Big(e^{-\mu}\textstyle\sum_{n=1}^{\infty}\sum_{m=2}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{n!m!}}\ |n-1,0\rangle\langle m-1,0|\Big) \\ &= e^{-\mu}\textstyle\sum_{n=2}^{\infty}\frac{\mu^n}{n!} \\ p_T(1|0) &= tr\Big(\rho_T(|0\rangle_e\langle 0|_e\otimes P_l + \tfrac{1}{2}P_e\otimes P_l)\Big) = 0\end{aligned} \tag{C3}$$

So, as expected, we get a detection in the transmitted arm only if the pulse contains two or more photons.

If the detection is in the x-basis, the evolution through the measuring device is:

$$\begin{aligned}\rho_T &= e^{-\mu}\ |0,0\rangle\langle 0,0| \\ &\quad + e^{-\mu}\textstyle\sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^{n-1}n}}\sum_{k_a=0}^{n-1}\frac{i^{k_a}}{\sqrt{(n-1-k_a)!(k_a)!}}\ |n-1-k_a,k_a\rangle \\ &\quad \cdot\textstyle\sum_{m=1}^{\infty}\frac{(\alpha^*)^m}{\sqrt{2^{m-1}m}}\sum_{K_A=0}^{m-1}\frac{(-i)^{K_A}}{\sqrt{(m-1-K_A)!(K_A)!}}\ \langle m-1-K_A,K_A|\end{aligned} \tag{C4}$$

where k_a, K_A represent the number of photons reflected from the original mode to the other mode. Note that ρ_T is indeed Hermitian and has a trace of 1.

And so:

$$\begin{aligned}p_T(+|0) &= tr\Big(\rho_T(P_+\otimes|0\rangle_-\langle 0|_- + \tfrac{1}{2}P_+\otimes P_-)\Big) \\ &= e^{-\mu}\textstyle\sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^{n-1}n}}\sum_{k_a=0}^{n-1}\frac{i^{k_a}}{\sqrt{(n-1-k_a)!(k_a)!}} \\ &\quad \cdot\textstyle\sum_{m=2}^{\infty}\frac{(\alpha^*)^m}{\sqrt{2^{m-1}m}}\frac{1}{\sqrt{(m-1)!}}\ \langle m-1,0|n-1-k_a,k_a\rangle \\ &\quad + \tfrac{1}{2}e^{-\mu}\textstyle\sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^{n-1}n}}\sum_{k_a=0}^{n-1}\frac{i^{k_a}}{\sqrt{(n-1-k_a)!(k_a)!}} \\ &\quad \cdot\textstyle\sum_{m=3}^{\infty}\frac{(\alpha^*)^m}{\sqrt{2^{m-1}m}} \\ &\quad \cdot\textstyle\sum_{K_A=1}^{m-2}\frac{(-i)^{K_A}}{\sqrt{(m-1-K_A)!(K_A)!}}\ \langle m-1-K_A,K_A|n-1-k_a,k_a\rangle\end{aligned} \tag{C5}$$

$$\begin{aligned}\rightarrow p_T(+|0) &= e^{-\mu}\Big[\textstyle\sum_{n=2}^{\infty}\frac{\mu^n}{2^{n-1}n!} + \frac{1}{2}\sum_{n=3}^{\infty}\frac{\mu^n}{2^{n-1}n}\sum_{k_a=1}^{n-2}\frac{1}{(n-1-k_a)!(k_a)!}\Big] \\ &= e^{-\mu}\Big[\textstyle\sum_{n=2}^{\infty}\frac{\mu^n}{2^{n-1}n!} + \frac{1}{2}\sum_{n=3}^{\infty}\frac{\mu^n}{n!}\sum_{k_a=1}^{n-2}\frac{(n-1)!}{2^{n-1}(n-1-k_a)!(k_a)!}\Big] \\ &= e^{-\mu}\Big[\textstyle\sum_{n=2}^{\infty}\frac{\mu^n}{2^{n-1}n!} + \frac{1}{2}\sum_{n=3}^{\infty}\frac{\mu^n}{n!}\Big(1-\frac{2}{2^{n-1}}\Big)\Big]\end{aligned} \tag{C6}$$

$$\rightarrow p_T(+|0) = e^{-\mu}\Big[\frac{\mu^2}{4} + \frac{1}{2}\textstyle\sum_{n=3}^{\infty}\frac{\mu^n}{n!}\Big] = \frac{1}{2}e^{-\mu}\textstyle\sum_{n=2}^{\infty}\frac{\mu^n}{n!} \tag{C7}$$

i.e., whenever the pulse contains two or more photons, there is a probability half of getting a + result. As expected. Also, it is obvious to see we get $p_T(-|0) = tr\left(\rho_T(|0\rangle_+\langle 0|_+\otimes P_- + \frac{1}{2}P_+\otimes P_-)\right) = p_T(+|0)$.

Generalizing the analysis to the case in which Alice sends out $|1\rangle$, we immediately get:

$$\begin{aligned}&p_T(0|1) = 0 \\ &p_T(1|1) = e^{-\mu}\textstyle\sum_{n=2}^{\infty}\frac{\mu^n}{n!} \\ &p_T(+|1) = p(-|1) = \tfrac{1}{2}e^{-\mu}\textstyle\sum_{n=2}^{\infty}\frac{\mu^n}{n!}\end{aligned} \tag{C8}$$

Now, for the case in which a $|+\rangle$ state is prepared by Alice, we get (using the analysis in Appendix B, Equation (B11)):

$$\begin{aligned}&(iii)\ |\psi\rangle = e^{-\frac{\mu}{2}}\left(|0,0,R,0,0\rangle + \sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^n n!}}\ |0,0,T,n-1,1\rangle\right) \\ &\qquad + e^{-\frac{\mu}{4}}\sum_{n=1}^{\infty}\frac{\alpha^n}{\sqrt{2^n n!}}\ |n-1,1,T,\frac{\alpha}{\sqrt{2}},0\rangle\end{aligned} \tag{C9}$$

$$\begin{aligned}&(iv)\rho_T = tr_{R_e,S,R_l}(|\psi\rangle\langle\psi|) = e^{-\mu}\ |0,0\rangle\langle 0,0| \\ &\quad + e^{-\mu}\textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{2^{n+m}n!m!}}\ |0,n-1\rangle\langle 0,m-1| \\ &\quad + e^{-\frac{\mu}{2}}\textstyle\sum_{n,m=1}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{2^{n+m}n!m!}}\ |n-1,\frac{\alpha}{\sqrt{2}}\rangle\langle m-1,\frac{\alpha}{\sqrt{2}}|\end{aligned} \tag{C10}$$

and so:

$$\begin{aligned}p_T(0|+) &= tr\Big(\rho_T(P_e\otimes|0\rangle_l\langle 0|_l + \tfrac{1}{2}P_e\otimes P_l)\Big) \\ &= e^{-\frac{3\mu}{4}}\textstyle\sum_{n=1}^{\infty}\sum_{m=2}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{2^{n+m}n!m!}}\ \langle m-1,0|n-1,\frac{\alpha}{\sqrt{2}}\rangle \\ &\quad + \tfrac{1}{2}e^{-\frac{\mu}{2}}\textstyle\sum_{n=1}^{\infty}\sum_{m=2}^{\infty}\frac{\alpha^n(\alpha^*)^m}{\sqrt{2^{n+m}n!m!}} \\ &\quad \cdot\Big(\langle m-1,\tfrac{\alpha}{\sqrt{2}}|n-1,\tfrac{\alpha}{\sqrt{2}}\rangle - e^{-\frac{\mu}{4}}\langle m-1,0|n-1,\tfrac{\alpha}{\sqrt{2}}\rangle\Big)\end{aligned} \tag{C11}$$

$$\Rightarrow p_T(0|+) = \left(e^{-\frac{\mu}{2}} \textstyle\sum_{n=2}^{\infty} \frac{\mu^n}{2^n n!}\right) e^{-\frac{\mu}{2}} + \frac{1}{2} e^{-\frac{\mu}{2}} \textstyle\sum_{n=2}^{\infty} \frac{\mu^n}{2^n n!} \left(1 - e^{-\frac{\mu}{2}}\right) \tag{C12}$$

i.e., we get the result of a '0' detection if (first term) the early pulse contains two or more photons and the late pulse is empty, or (second term) if the early pulse contains two or more photons and the late pulse is non-empty and the randomly assigned bit value is '0' (probability $\frac{1}{2}$).

The probability $p_T(1|+)$ is:

$$\begin{aligned} p_T(1|+) &= tr\left(\rho_T (|0\rangle_e \langle 0|_e \otimes P_l + \tfrac{1}{2} P_e \otimes P_l)\right) \\ &= e^{-\mu} \textstyle\sum_{n=1}^{\infty} \sum_{m=2}^{\infty} \frac{\alpha^n (\alpha^*)^m}{\sqrt{2^{n+m} n! m!}} \langle 0, m-1 | 0, n-1 \rangle \\ &\quad + e^{-\frac{\mu}{2}} \textstyle\sum_{n=1}^{\infty} \frac{\alpha^n \alpha^*}{\sqrt{2^{n+1} n!}} \left(\langle 0, \frac{\alpha}{\sqrt{2}} | n-1, \frac{\alpha}{\sqrt{2}} \rangle - e^{-\frac{\mu}{4}} \langle 0, 0 | n-1, \frac{\alpha}{\sqrt{2}} \rangle \right) \\ &\quad + \tfrac{1}{2} e^{-\frac{\mu}{2}} \textstyle\sum_{n=2}^{\infty} \frac{\mu^n}{2^n n!} \left(1 - e^{-\frac{\mu}{2}}\right) \end{aligned} \tag{C13}$$

$$\begin{aligned} \Rightarrow p_T(1|+) &= e^{-\frac{\mu}{2}} \left(e^{-\frac{\mu}{2}} \textstyle\sum_{n=2}^{\infty} \frac{\mu^n}{2^n n!} \right) + \frac{e^{-\frac{\mu}{2}} \mu}{2} \left(1 - e^{-\frac{\mu}{2}}\right) \\ &\quad + \tfrac{1}{2} e^{-\frac{\mu}{2}} \textstyle\sum_{n=2}^{\infty} \frac{\mu^n}{2^n n!} \left(1 - e^{-\frac{\mu}{2}}\right) \end{aligned} \tag{C14}$$

i.e., a result of '1' is obtained if (first term) the early pulse is empty and the late pulse contains two or more photons, or if (second term) the early pulse contains a single photon (changing the state of the SPRINT-system) and the late pulse is non-empty, or if (third term) the early pulse contains two or more photons and the late pulse is non-empty and the randomly assigned bit value is '1' (probability $\frac{1}{2}$). One can also verify that indeed $p_T(0|+) + p_T(1|+) + p_T(00|+) = 1$.

The above can be straightforwardly adapted to the case in which Alice sends out a $|-\rangle$ state, yielding:

$$\begin{aligned} p_T(0|-) &= p_T(0|+) \\ p_T(1|-) &= p_T(1|+) \end{aligned} \tag{C15}$$

If the transmitted arm is measured in the x-basis, then, for the case of a sent $|+\rangle$ state, the density matrix at the input of the beamsplitter of the measuring device is:

$$\begin{aligned} \rho_T = {}& ^{(\mathrm{I})} e^{-\mu} |0, 0\rangle \langle 0, 0| + ^{(\mathrm{II})} e^{-\mu} \\ & \cdot \textstyle\sum_{n,m=1}^{\infty} \frac{(-i)^{n-1} \alpha^n i^{m-1} (\alpha^*)^m}{\sqrt{2^{n+m} n! m!}} |0, n-1\rangle \langle 0, m-1| \\ & + ^{(\mathrm{III})} e^{-\frac{\mu}{2}} \textstyle\sum_{n,m=1}^{\infty} \frac{\alpha^n (\alpha^*)^m}{\sqrt{2^{n+m} n! m!}} |n-1, \frac{-i\alpha}{\sqrt{2}}\rangle \langle m-1, \frac{-i\alpha}{\sqrt{2}}| \end{aligned} \tag{C16}$$

At the output of the beamsplitter we will treat each term from the density matrix separately. We get for the first term:

$$\begin{aligned} & \rho_T^{(\mathrm{I})} \rightarrow e^{-\mu} |0, 0\rangle \langle 0, 0| \\ & \Rightarrow p_T^{(\mathrm{I})}(+|+) = tr\left(\rho_T^{(\mathrm{I})} (P_+ \otimes |0\rangle_- \langle 0|_- + \tfrac{1}{2} P_+ \otimes P_-)\right) = 0 = p_T^{(\mathrm{I})}(-|+) \end{aligned} \tag{C17}$$

For the second term:

$$\begin{aligned} \rho_T^{(\mathrm{II})} & \rightarrow e^{-\mu} \textstyle\sum_{n,m=1}^{\infty} \frac{2 (-i)^{n-1} \alpha^n i^{m-1} (\alpha^*)^m}{2^{n+m} \sqrt{n} \sqrt{m}} \\ & \cdot \textstyle\sum_{k_b=0}^{n-1} \frac{i^{k_b}}{\sqrt{(n-1-k_b)! k_b!}} |k_b, n-1-k_b\rangle \\ & \cdot \textstyle\sum_{K_B=0}^{m-1} \frac{(-i)^{K_B}}{\sqrt{(m-1-K_B)! K_B!}} \langle K_B, m-1-K_B| \end{aligned} \tag{C18}$$

$$\begin{aligned} \Rightarrow p_T^{(\mathrm{II})}(+|+) &= e^{-\mu} \textstyle\sum_{n=2}^{\infty} \frac{2\mu^n}{4^n n!} + \frac{1}{2} e^{-\mu} \sum_{n=3}^{\infty} \frac{\mu^n}{2^n n!} \left(1 - \frac{2}{2^{n-1}}\right) \\ &= e^{-\frac{\mu}{2}} \left(\tfrac{1}{2} e^{-\frac{\mu}{2}} \textstyle\sum_{n=2}^{\infty} \frac{\mu^n}{2^n n!} \right) \equiv A \end{aligned} \tag{C19}$$

which is half the probability that the early pulse is a vacuum and the late pulse contains two or more photons. As expected. It also straightforward to see that $p_T^{(\mathrm{II})}(-|+) = p_T^{(\mathrm{II})}(+|+)$

As for the third term:

$$\begin{aligned} \rho_T^{(\mathrm{III})} = {}& e^{-\mu} \textstyle\sum_{n=1}^{\infty} \frac{\sqrt{2n} \alpha^n}{2^n n!} \sum_{j=0}^{\infty} \frac{(-i)^j \alpha^j}{2^j j!} \sum_{k_a=0}^{n-1} \sum_{k_b=0}^{j} \binom{n-1}{k_a} \binom{j}{k_b} i^{k_a+k_b} \\ & \cdot \sqrt{(n-1-k_a+k_b)!(j-k_b+k_a)!} \, |n-1-k_a+k_b, j-k_b+k_a\rangle \\ & \cdot \textstyle\sum_{m=1}^{\infty} \frac{\sqrt{2m} (\alpha^*)^m}{2^m m!} \sum_{J=0}^{\infty} \frac{i^J (\alpha^*)^J}{2^J J!} \sum_{K_A=0}^{m-1} \sum_{K_B=0}^{J} \binom{m-1}{K_A} \binom{J}{K_B} (-i)^{K_A+K_B} \\ & \cdot \sqrt{(m-1-K_A+K_B)!(J-K_B+K_A)!} \, \langle m-1-K_A+K_B, J-K_B+K_A| \end{aligned} \tag{C20}$$

Note that $P_+ \otimes |0\rangle_- \langle 0|_- + \frac{1}{2} P_+ \otimes P_- = P_+ \otimes |0\rangle_- \langle 0|_- + \frac{1}{2} P_+ \otimes (I_- - |0\rangle_- \langle 0|_-)) = \frac{1}{2} P_+ \otimes (I_- + |0\rangle_- \langle 0|_-))$. And so we get, after some tedious work:

$$p_T^{(\mathrm{III})}(+|+) = B_{m-1-K_A+K_B} + C \tag{C21}$$

and:

$$p_T^{(\mathrm{III})}(-|+) = B_{J-K_B+K_A} + \tilde{C} \tag{C22}$$

where:

$$\begin{aligned} B_X \equiv {}& e^{-\mu} \textstyle\sum_{n=1}^{\infty} \frac{\sqrt{n} \alpha^n}{2^n n!} \sum_{j=0}^{\infty} \frac{(-i)^j \alpha^j}{2^j j!} \sum_{k_a=0}^{n-1} \sum_{k_b=0}^{j} \binom{n-1}{k_a} \binom{j}{k_b} i^{k_a+k_b} \\ & \cdot (n-1-k_a+k_b)!(j-k_b+k_a)! \textstyle\sum_{m=1}^{\infty} \frac{\sqrt{m} (\alpha^*)^m}{2^m m!} \sum_{J=0}^{\infty} \frac{i^J (\alpha^*)^J}{2^J J!} \\ & \cdot \textstyle\sum_{K_A=0}^{m-1} \sum_{K_B=0}^{J} \binom{m-1}{K_A} \binom{J}{K_B} (-i)^{K_A+K_B} \\ & \cdot \delta_{n-k_a+k_b, m-K_A+K_B} \delta_{j-k_b+k_a, J-K_B+K_A} (1 - \delta_{X,0}) \end{aligned} \tag{C23}$$

$$C \equiv e^{-\mu} \textstyle\sum_{n=1}^{\infty} \frac{\sqrt{n}}{n!} \sum_{j=0}^{\infty} \frac{\mu^{n+j}}{4^{n+j} j!} \sum_{m=1}^{n+j} \frac{\mathbf{1}}{\sqrt{m}} \binom{n-1+j}{m-1} (1 - \delta_{n-1+j,0}) \tag{C24}$$

$$\tilde{C} \equiv e^{-\mu} \textstyle\sum_{n=1}^{\infty} \frac{\sqrt{n}}{n!} \sum_{j=0}^{\infty} \frac{\mu^{n+j}}{4^{n+j} j!} \sum_{m=1}^{n+j} \frac{(-\mathbf{1})^{\mathbf{n+m}}}{\sqrt{m}} \binom{n-1+j}{m-1} (1 - \delta_{n-1+j,0}) \tag{C25}$$

where $\delta_{X,Y}$ is 1 only if $X = Y$, and is zero otherwise, such that in B_X the summation is only over terms for which $X \neq 0$ and $n - k_a + k_b = m - K_A + K_B$ and $j - k_b + k_a = J - K_B + K_A$, and for C and $\tilde{C}$ the summation is only over terms for which $n - 1 + j \neq 0$. The difference between C and $\tilde{C}$ is typed in boldface.

So over all we get:

$$p_T(+|+) = A + B_{m-1-K_A+K_B} + C \tag{C26}$$

and:

$$p_T(-|+) = A + B_{J-K_B+K_A} + \tilde{C} \tag{C27}$$

Finally, for the case where a state $|-\rangle$ was prepared by Alice, we can generalize the last analysis to get:

$$p_T(+|-) = A + \tilde{B}_{m-1-K_A+K_B} + \tilde{C} \tag{C28}$$

and:

$$p_T(-|-) = A + \tilde{B}_{J-K_B+K_A} + C \tag{C29}$$

where:

$$\begin{aligned}\tilde{B}_X \equiv\, & e^{-\mu} \sum_{n=1}^{\infty} \frac{\sqrt{n}\alpha^n}{2^n n!} \sum_{j=0}^{\infty} \frac{(i)^j \alpha^j}{2^j j!} \sum_{k_a=0}^{n-1} \sum_{k_b=0}^{j} \binom{n-1}{k_a}\binom{j}{k_b} i^{k_a+k_b} \\ & \cdot (n-1-k_a+k_b)!(j-k_b+k_a)! \sum_{m=1}^{\infty} \frac{\sqrt{m}(\alpha^*)^m}{2^m m!} \sum_{J=0}^{\infty} \frac{(-i)^J (\alpha^*)^J}{2^J J!} \\ & \cdot \sum_{K_A=0}^{m-1} \sum_{K_B=0}^{J} \binom{m-1}{K_A}\binom{J}{K_B} (-i)^{K_A+K_B} \\ & \cdot \delta_{n-k_a+k_b, m-K_A+K_B} \delta_{j-k_b+k_a, J-K_B+K_A} (1-\delta_{X,0})\end{aligned} \tag{C30}$$

The expressions can be simplified somewhat. Note that the Kronecker deltas in B and $\tilde{B}$ imply that $n+j = m+J \rightarrow J = n+j-m$, thus we can write:

$$\begin{aligned}B_X =\, & e^{-\mu} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \sum_{j=0}^{\infty} \frac{\mu^{n+j}}{4^{n+j}} \sum_{k_a=0}^{n-1} \sum_{k_b=0}^{j} \frac{(n-1-k_a+k_b)!(j-k_b+k_a)!}{(n-1-k_a)!(k_a)!(j-k_b)!(k_b)!} \\ & \cdot \sum_{m=1}^{n+j} \frac{1}{\sqrt{m}} \sum_{K_A=0}^{m-1} \sum_{K_B=0}^{n+j-m} \\ & \cdot [(m-1-K_A)!(K_A)!(n+j-m-K_B)!(K_B)!]^{-1} (\mathbf{-1})^{\mathbf{k_a-K_A}} \\ & \cdot \delta_{n-k_a+k_b, m-K_A+K_B} (1-\delta_{X,0})\end{aligned} \tag{C31}$$

and:

$$\begin{aligned}\tilde{B}_X =\, & e^{-\mu} \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} \sum_{j=0}^{\infty} \frac{\mu^{n+j}}{4^{n+j}} \sum_{k_a=0}^{n-1} \sum_{k_b=0}^{j} \frac{(n-1-k_a+k_b)!(j-k_b+k_a)!}{(n-1-k_a)!(k_a)!(j-k_b)!(k_b)!} \\ & \cdot \sum_{m=1}^{n+j} \frac{1}{\sqrt{m}} \sum_{K_A=0}^{m-1} \sum_{K_B=0}^{n+j-m} \\ & \cdot [(m-1-K_A)!(K_A)!(n+j-m-K_B)!(K_B)!]^{-1} (\mathbf{-1})^{\mathbf{k_b-K_B}} \\ & \cdot \delta_{n-k_a+k_b, m-K_A+K_B} (1-\delta_{X,0})\end{aligned} \tag{C32}$$

where the difference between the two expressions is typed in boldface. Moreover, it can be shown numerically (and possibly, with non-negligible effort, analytically) that $B_{m-1-K_A+K_B} = \tilde{B}_{n+j-m-K_B+K_A}$ and $B_{n+j-m-K_B+K_A} = \tilde{B}_{m-1-K_A+K_B}$, resulting in the expected symmetry:

$$\begin{aligned}p_T(+|+) = p_T(-|-) = A + B_{m-1-K_A+K_B} + C \\ p_T(-|+) = p_T(+|-) = A + B_{n+j-m-K_B+K_A} + \tilde{C}\end{aligned} \tag{C33}$$

To verify the validity of these rather complicated expressions, we ran a numerical simulation of the total probability, i.e., $p_T(+|+) + p_T(-|+) + p_T(00|+)$, with $p_T(00|+) = e^{-\mu}(1+\mu)$, as can be verified easily. In our simulation, we truncated the sums to $n, j \leq 10$. The result of this simulation is shown in **Figure C1**, proving that indeed the total probability sums to 1, up to a negligible difference attributed to the truncation of the analytical expressions.

Appendix D: Detection Statistics in the Reflected Arm, with Reset

The process of resetting the SPRINT-system (which can be thought of as projecting the SPRINT-system, regardless of its current state, onto the reflecting state, $|R\rangle$ ($\langle R| + \langle T|$)) is described by tracing over the SPRINT-system since all the information contained in its state is lost.

Here, we want to calculate the detection statistics in the reflected arm (Bob's side) for the case of time-bin encoded BB84 under SPRINT-based PNS attack, assuming that Eve resets the SPRINT-system in between the early and late time-bins. We will describe it "mathematically" as if there are two SPRINT-systems, S_e and S_l, with S_e (S_l) coupled only to the early (late) time-bin, using an optical switch. The notation we will use is:

$$|\psi\rangle = |\,\rangle_{T,e} \otimes |\,\rangle_{R,e} \otimes |\,\rangle_{S_e} \otimes |\,\rangle_{T,l} \otimes |\,\rangle_{R,l} \otimes |\,\rangle_{S_l} \tag{D1}$$

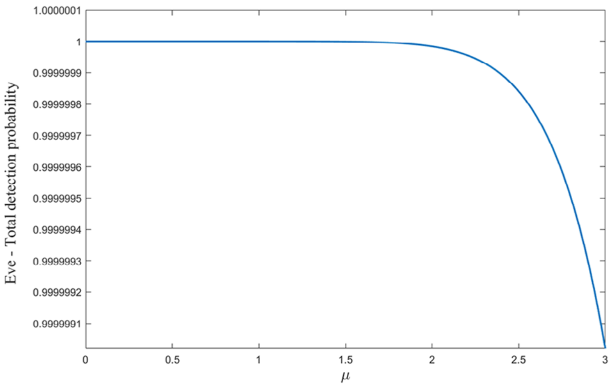

Figure C1. Simulation of the total detection probability in the transmitted arm (Eve's side), for x-basis states, as a function of the average photon number, μ. The total probability sums up to 1, as expected.

The analysis steps are – i) Alice's preparation of the state, ii) the early SPRINT-system operates on the early time-bin with $\hat{s}_e$, iii) the late SPRINT-system operates on the late time-bin with, $\hat{s}_l$, iv) the SPRINT-systems and the transmission output arm are traced over to obtain ρ_R, and finally v) detection in z-basis or x-basis.

It is obvious that the detection statistics might be affected by the reset only if the state sent by Alice is in the x-basis. Let us first look at the case where Alice sends out $|+\rangle$. We get:

$$(i)\ |+\rangle \Rightarrow |\psi\rangle = |\frac{\alpha}{\sqrt{2}}, 0, R, \frac{\alpha}{\sqrt{2}}, 0, R\rangle \tag{D2}$$

$$(ii)\ |\psi\rangle = e^{-\frac{\mu}{4}}\left(|0, 0, R, \frac{\alpha}{\sqrt{2}}, 0, R\rangle + \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{2^n n!}} |n-1, 1, T, \frac{\alpha}{\sqrt{2}}, 0, R\rangle\right) \tag{D3}$$

$$\begin{aligned}(iii)\ |\psi\rangle =\, & e^{-\frac{\mu}{2}}\left(|0, 0, R, 0, 0, R\rangle + \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{2^n n!}} |0, 0, R, n-1, 1, T\rangle\right) \\ & + e^{-\frac{\mu}{2}}\left(\sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{2^n n!}} |n-1, 1, T, 0, 0, R\rangle + \sum_{n=1}^{\infty} \frac{\alpha^n}{\sqrt{2^n n!}}\right. \\ & \left.\cdot \sum_{m=1}^{\infty} \frac{\alpha^m}{\sqrt{2^m m!}} |n-1, 1, T, m-1, 1, T\rangle\right)\end{aligned} \tag{D4}$$

$$\begin{aligned}(iv)\rho_R =\, & tr_{T_e, S_e, T_l, S_l}(|\psi\rangle\langle\psi|) \\ =\, & e^{-\mu}\left[|0,0\rangle\langle 0,0| + \sum_{n=1}^{\infty} \frac{\mu^n}{2^n n!}(|0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0|)\right] \\ & + e^{-\mu}\left(\sum_{n=1}^{\infty} \frac{\mu^n}{2^n n!} \sum_{m=1}^{\infty} \frac{\mu^m}{2^m m!} |1,1\rangle\langle 1,1|\right)\end{aligned} \tag{D5}$$

$$\begin{aligned}\Rightarrow \rho_R =\, & e^{-\frac{\mu}{2}} e^{-\frac{\mu}{2}} |0,0\rangle\langle 0,0| + e^{-\frac{\mu}{2}}\left(1 - e^{-\frac{\mu}{2}}\right) \\ & \cdot (|0,1\rangle\langle 0,1| + |1,0\rangle\langle 1,0|) \\ & + \left(1 - e^{-\frac{\mu}{2}}\right)^2 |1,1\rangle\langle 1,1|\end{aligned} \tag{D6}$$

which is a statistical mixture of the cases where both pulses are vacuum (first term), only one pulse is occupied (second and third terms), and both pulses are occupied (last term).

It follows immediately that:

$$\begin{aligned} p_R(0|+) &= tr\left(\rho_R(P_e \otimes |0\rangle_l \langle 0|_l + \tfrac{1}{2}P_e \otimes P_l)\right) \\ &= e^{-\frac{\mu}{2}}\left(1-e^{-\frac{\mu}{2}}\right)+\tfrac{1}{2}\left(1-e^{-\frac{\mu}{2}}\right)^2 = \tfrac{1}{2}(1-e^{-\mu}) \end{aligned} \tag{D7}$$

and that $p_R(0|+) = p_R(1|+)$.

Note that an input state $|1,1\rangle = a^\dagger b^\dagger |0,0\rangle$ to a beamsplitter results in $\frac{1}{2}(a^\dagger + ib^\dagger)(ia^\dagger + b^\dagger)|0,0\rangle = \frac{i}{\sqrt{2}}(|2,0\rangle + |0,2\rangle)$. Thus, if the detection is in the x-basis, the evolution through the detection apparatus is:

$$\begin{aligned} \rho_R &= e^{-\frac{\mu}{2}}e^{-\frac{\mu}{2}}|0,0\rangle\langle 0,0| \\ &+\tfrac{1}{2}e^{-\frac{\mu}{2}}\left(1-e^{-\frac{\mu}{2}}\right) \\ &\cdot[(|1,0\rangle - i|0,1\rangle)(\langle 1,0| + i\langle 0,1|) + (|1,0\rangle + i|0,1\rangle)(\langle 1,0| - i\langle 0,1|)] \\ &+\tfrac{1}{2}\left(1-e^{-\frac{\mu}{2}}\right)^2(|2,0\rangle + |0,2\rangle)(\langle 2,0| + \langle 0,2|) \end{aligned} \tag{D8}$$

And so:

$$\begin{aligned} p_R(+|+) &= tr\left(\rho_R(P_+ \otimes |0\rangle_- \langle 0|_- + \tfrac{1}{2}P_+ \otimes P_-)\right) \\ &= e^{-\frac{\mu}{2}}\left(1-e^{-\frac{\mu}{2}}\right)+\tfrac{1}{2}\left(1-e^{-\frac{\mu}{2}}\right)^2 = \tfrac{1}{2}(1-e^{-\mu}) \end{aligned} \tag{D9}$$

and $p_R(+|+) = p_R(-|+)$. It is also straightforward to generalize the analysis to the case in which Alice sends out a $|-\rangle$ state, obtaining:

$$p_R(0|-) = p_R(1|-) = p_R(+|-) = p_R(-|-) = \tfrac{1}{2}(1-e^{-\mu}) \tag{D10}$$

Appendix E: Mutual Information

Here, we show how Eve's information gain can be used to calculate the mutual information between Alice and her. The mutual information that Alice shares with Eve is defined as:

$$\begin{aligned} I(A;E) &= \sum_{a\in\{0,1\}}\sum_{e\in\{0,1\}} p(a,e)\log\left(\frac{p(a,e)}{p(a)p(e)}\right) \\ &= \sum_{e\in\{0,1\}}\left(p(0)p(e|0)\log\left(\frac{p(e|0)}{p(e)}\right) + p(1)p(e|1)\log\left(\frac{p(e|1)}{p(e)}\right)\right) \end{aligned} \tag{E1}$$

where the logarithm is taken to base 2.

Because Alice generates a random string, we know that for any a the marginal probability is $p(a) = 1/2$. That is also true for Eve, i.e., for any e the marginal probability is $p(e) = 1/2$. Thus:

$$\begin{aligned} I(A;E) &= \tfrac{1}{2}\textstyle\sum_{e\in\{0,1\}}(p(e|0)\log(2p(e|0)) + p(e|1)\log(2p(e|1))) \\ &= \tfrac{1}{2}[p(0|0)\log(2p(0|0)) + p(1|0)\log(2p(1|0)) \\ &\quad + p(0|1)\log(2p(0|1)) + p(1|1)\log(2p(1|1))] \end{aligned} \tag{E2}$$

Due to symmetry, we have $p(0|0) = p(1|1)$ and $p(0|1) = p(1|0)$. Also, we note that $p(0|0)$ is the probability that Eve gets the correct bit value, which is given by the information gain G_M calculated in Section 4.2. The complement probability, $p(1|0)$, is therefore given by $1 - G_M$. And so we get:

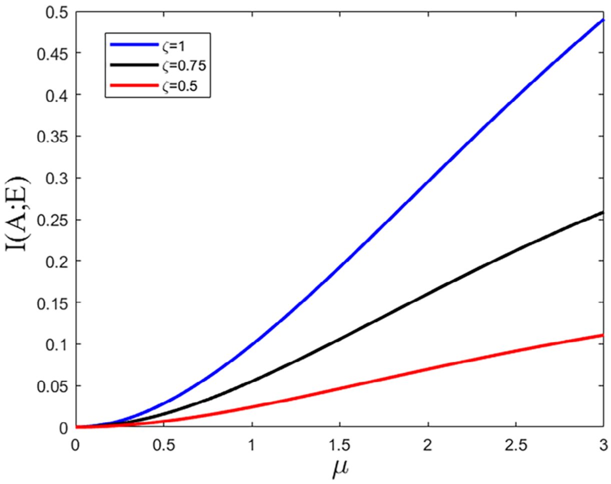

Figure E1. Alice and Eve's mutual information, Equation (E3), as a function of μ with different ζ values.

$$\begin{aligned} I(A;E) &= [p(0|0)\log(2p(0|0)) + p(1|0)\log(2p(1|0))] \\ &= G_M\log(2G_M) + (1-G_M)\log(2-2G_M) \end{aligned} \tag{E3}$$

Now, we can calculate the mutual information between Alice and Eve as a function of μ as plotted in **Figure E1**. These expressions for the mutual information can be helpful for various prospective security analyses.

Acknowledgements

A.A. and Y.I. contributed equally to this work. A.A. acknowledged scholarships from Israel's Council of Higher Education and from BIU's President. B.D. acknowledged support from the Israel Science Foundation and the Binational Science Foundation. E.C. was supported by the Israeli Innovation Authority under Project No. 73795, by the Pazy Foundation, by the Israeli Ministry of Science and Technology, and by the Quantum Science and Technology Program of the Israeli Council of Higher Education.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

light-matter interaction, photon number splitting, quantum key distribution, quantum optics, single-photon Raman interaction

Received: December 7, 2023
Revised: April 2, 2024
Published online: April 24, 2024

[1] P. W. Shor, J. Preskill, *Phys. Rev. Lett.* **2000**, *85*, 441.
[2] C. H. Bennett, G. Brassard, in Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, IEEE, New York, pp. 175–179.
[3] B. Huttner, N. Imoto, N. Gisin, T. Mor, *Phys. Rev. A* **1995**, *51*, 1863.
[4] C. H. Bennett, *Phys. Rev. Lett.* **1992**, *68*, 3121.
[5] G. Brassard, N. Lütkenhaus, T. Mor, B. C. Sanders, *Phys. Rev. Lett.* **2000**, *85*, 1330.
[6] V. Scarani, A. Acín, G. Ribordy, N. Gisin, *Phys. Rev. Lett.* **2004**, *92*, 057901.
[7] C. Branciard, N. Gisin, B. Kraus, V. Scarani, *Phys. Rev. A* **2005**, *72*, 032301.
[8] W.-Y. Hwang, *Phys. Rev. Lett.* **2003**, *91*, 057901.
[9] H.-K. Lo, X. Ma, K. Chen, *Phys. Rev. Lett.* **2005**, *94*, 230504.
[10] D. Stucki, N. Brunner, N. Gisin, V. Scarani, H. Zbinden, *Appl. Phys. Lett.* **2005**, *87*, 194108.
[11] N. Lütkenhaus, *Phys. Rev. A* **2000**, *61*, 052304.
[12] N. Lütkenhaus, M. Jahma, *New J. Phys.* **2002**, *4*, 44.
[13] H. Inamori, N. Lütkenhaus, D. Mayers, *Eur. Phys. J. D* **2007**, *41*, 599.
[14] W.-T. Liu, S.-H. Sun, L.-M. Liang, J.-M. Yuan, *Phys. Rev. A* **2011**, *83*, 042326.
[15] N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, *Rev. Mod. Phys.* **2002**, *74*, 145.
[16] I. Shomroni, S. Rosenblum, Y. Lovsky, O. Bechler, G. Guendelman, B. Dayan, *Science* **2014**, *345*, 903.
[17] S. Rosenblum, A. Borne, B. Dayan, *Phys. Rev. A* **2017**, *95*, 033814.
[18] O. Bechler, A. Borne, S. Rosenblum, G. Guendelman, O. E. Mor, M. Netser, T. Ohana, Z. Aqua, N. Drucker, R. Finkelstein, Y. Lovsky, R. Bruch, D. Gurovich, E. Shafir, B. Dayan, *Nat. Phys.* **2018**, *14*, 996.
[19] S. Rosenblum, O. Bechler, I. Shomroni, Y. Lovsky, G. Guendelman, B. Dayan, *Nat. Photonics* **2016**, *10*, 19.
[20] D. Rusca, A. Boaron, M. Curty, A. Martin, H. Zbinden, *Phys. Rev. A* **2018**, *98*, 052336.
[21] F. Bouchard, D. England, P. J. Bustard, K. Heshami, B. Sussman, *PRX Quantum* **2022**, *3*, 010332.
[22] J. Brendel, N. Gisin, W. Tittel, H. Zbinden, *Phys. Rev. Lett.* **1999**, *82*, 2594.
[23] C. C. Gerry, P. L. Knight, *Introductory Quantum Optics*, 1st edition, Cambridge University Press, Cambridge **2005**.
[24] H. Ko, B.-S. Choi, J.-S. Choe, C. J. Youn, *Quantum Inf. Process.* **2018**, *17*, 1.
[25] N. Jain, B. Stiller, I. Khan, D. Elser, C. Marquardt, G. Leuchs, *Contemp. Phys.* **2016**, *57*, 366.
[26] V. Makarov, D. R. Hjelme, *J. Mod. Opt.* **2005**, *52*, 691.
[27] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, V. Makarov, *Nat. Photonics* **2010**, *4*, 686.
[28] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, G. Ribordy, *Phys. Rev. A* **2006**, *73*, 022320.
[29] Y. Huang, Z. Du, X. Ma, *Adv. Quantum Technol.* **2024**, *7*, 2300275.
[30] M. Dušek, O. Haderka, M. Hendrych, *Opt. Commun.* **1999**, *169*, 103.
[31] H.-W. Li, S. Wang, J.-Z. Huang, W. Chen, Z.-Q. Yin, F.-Y. Li, Z. Zhou, D. Liu, Y. Zhang, G.-C. Guo, W.-S. Bao, Z.-F. Han, *Phys. Rev. A* **2011**, *84*, 062308.
[32] X. Ma, P. Zeng, H. Zhou, *Phys. Rev. X* **2018**, *8*, 031043.
[33] D. Pinotsi, A. Imamoglu, *Phys. Rev. Lett.* **2008**, *100*, 093603.
[34] M. O. Scully, M. S. Zubairy, *Quantum Optics*, 1st edition, Cambridge University Press, Cambridge **1997**.
[35] U. Leonhardt, *Measuring the quantum state of light*, 1st edition, Cambridge studies in modern optics, Cambridge University Press, Cambridge **1997**.
[36] P. C. Humphreys, B. J. Metcalf, J. B. Spring, M. Moore, X.-M. Jin, M. Barbieri, W. S. Kolthammer, I. A. Walmsley, *Phys. Rev. Lett.* **2013**, *111*, 150501.
[37] H. P. Yuen, *Quantum Semiclassical Opt.: J. Eur. Opt. Soc. Part B* **1996**, *8*, 939.
[38] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, J. L. Pereira, M. Razavi, J. Shamsul Shaari, M. Tomamichel, V. C. Usenko, G. Vallone, P. Villoresi, P. Wallden, *Adv. Opt. Photonics* **2020**, *12*, 1012.
[39] M. O. Scully, K. Drühl, *Phys. Rev. A* **1982**, *25*, 2208.
[40] M. O. Scully, B.-G. Englert, H. Walther, *Nature* **1991**, *351*, 111.
[41] Y.-H. Kim, R. Yu, S. P. Kulik, Y. Shih, M. O. Scully, *Phys. Rev. Lett.* **2000**, *84*, 1.
[42] A. C. Elitzur, E. Cohen, *Int. J. Quantum Inf.* **2015**, *12*, 1560024.
[43] A. Elitzur, E. Cohen, *Philos. Trans. R. Soc., A* **2016**, *374*, 20150242.