Using Modeling and Simulation to Study Photon Number Splitting Attacks

Logan O. Mailloux  
*Air Force Institute of Technology*

Douglas D. Hodson  
*Air Force Institute of Technology*

Michael R. Grimaila  
*Air Force Institute of Technology*

Ryan D. Engle  
*Air Force Institute of Technology*

Colin V. McLaughlin  
*Naval Research Laboratory*

See next page for additional authors

Follow this and additional works at: [https://scholar.afit.edu/facpub](https://scholar.afit.edu/facpub)

**Recommended Citation**

This Article is brought to you for free and open access by AFIT Scholar. It has been accepted for inclusion in Faculty Publications by an authorized administrator of AFIT Scholar. For more information, please contact richard.mansfield@afit.edu.
Authors
Logan O. Mailloux, Douglas D. Hodson, Michael R. Grimaila, Ryan D. Engle, Colin V. Mclaughlin, and Gerald B. Baumgartner

This article is available at AFIT Scholar: https://scholar.afit.edu/facpub/25
Using Modeling and Simulation to Study Photon Number Splitting Attacks

LOGAN O. MAILLOUX¹, (Member, IEEE), DOUGLAS D. HODSON¹, MICHAEL R. GRIMAILA¹, (Senior Member, IEEE), RYAN D. ENGLE¹, (Member, IEEE), COLIN V. MCLAUGHLIN², AND GERALD B. BAUMGARTNER³

¹Air Force Institute of Technology, Wright-Patterson AFB, OH 45433-7765, USA
²Naval Research Laboratory, Washington, DC 20375, USA
³Laboratory for Telecommunications Sciences, College Park, MD 20740, USA

Corresponding author: L. O. Mailloux (logan.mailloux@afit.edu)

This work was supported in part by the Laboratory for Telecommunication Sciences under Grant 5743400-304-6448 and in part by the Department of Defense High Performance Computing Modernization Program within the Air Force Research Laboratory, Wright-Patterson AFB, OH, USA.

ABSTRACT Quantum key distribution (QKD) is an innovative technology, which exploits the laws of quantum mechanics to generate and distribute unconditionally secure shared cryptographic keying material between two geographically separated parties. The unique nature of QKD that ensures eavesdropping on the key distribution channel necessarily introduces detectable errors and shows promise for high-security environments, such as banking, government, and military. However, QKD systems are vulnerable to advanced theoretical and experimental attacks. In this paper, the photon number splitting (PNS) attack is studied in a specialized QKD modeling and simulation framework. First, a detailed treatment of the PNS attack is provided with emphasis on practical considerations, such as performance limitations and realistic sources of error. Second, ideal and non-ideal variations of the PNS attack are studied to measure the eavesdropper’s information gain on the QKD-generated secret key bits and examine the detectability of PNS attacks with respect to both quantum bit error rate and the decoy state protocol. Finally, this paper provides a repeatable methodology for efficiently studying advanced attacks, both realized and notional, against QKD systems and more generally quantum communication protocols.

INDEX TERMS Quantum key distribution, photon number splitting attack, decoy state protocol.

I. INTRODUCTION Quantum Key Distribution (QKD) is the most mature application of quantum information science and heralded as a revolutionary technology offering the means for two geographically separated parties to generate unconditionally secure shared cryptographic keying material. Unlike conventional key distribution techniques (i.e., RSA), the security of QKD rests on the laws of quantum mechanics and not computational complexity. In theory, these attributes make QKD well suited for high-security applications such as banking, government, and military environments. However, QKD is a nascent technology where implementation defects, practical engineering limitations, and poor design decisions can result in vulnerabilities [1]. These vulnerabilities are subject to a growing number of theoretical and realized attacks from sophisticated “quantum hacking” groups [2]. Moreover, understanding the security and performance impact of these attacks is critical for QKD system certification in strictly controlled high-security environments [3].

Arguably the most powerful QKD attack to date, the Photon Number Splitting (PNS) attack is a theoretical attack designed to gain full information on the QKD-generated shared secret key bits without introducing detectable errors [4], [5]. While the PNS attack cannot be fully realized with current technology, in his defining work on the topic Lütkenhaus states: “the PNS attack can be well approximated with linear optics, a rudimentary QND measurement and a short-time quantum memory” [5]. Thus, in this work we study both the ideal PNS attack and non-ideal versions through Modeling and Simulation (M&S). More specifically, the PNS attack is modeled with performance limitations and realistic sources of error while simulation is used to measure
the eavesdropper’s information gain on the QKD system’s secret key bits. Additionally, the detectability of the PNS attack is examined with respect to both Quantum Bit Error Rate (QBER) and the decoy state protocol.

This paper is organized as follows: Section II provides a brief history of QKD and discusses vulnerabilities which permit the PNS attack. In Section III, the PNS attack is decomposed and modeled in a parameterized fashion to account for technology induced performance limitations and realistic sources of error. Simulation results for the eavesdropper’s information gain and the attack’s detectability are presented in Section IV. Lastly, conclusions and future work are described in Section V. This work is related to the Author’s previous works [6]–[8] and extends the initial PNS attack assessment accomplished in [9]. For accessible engineering oriented introductions to QKD, please see [10] and [11]. For more comprehensive reviews see [12] and [13].

II. BACKGROUND

The genesis of QKD can be traced back to Stephen Wiesner, who proposed the idea of encoding information on polarized photons in the late 1960s [14]. In 1984, Bennett and Brassard extended this idea to create the first QKD protocol, known as “BB84,” to generate unconditionally secure shared cryptographic keying material between two geographically separated parties [15]. From its relatively unnoticed beginnings, QKD has gained global interest as an emerging cyber security technology.

![QKD System Context Diagram](image)

**FIGURE 1.** QKD System Context Diagram. The sender “Alice” and receiver “Bob” generate shared secret key, $K$, to encrypt sensitive information.

As illustrated in Table 1, BB84 is a prepare-and-measure protocol where Alice prepares photons in one of four polarization states (e.g., $\leftrightarrow$, $\uparrow$, $\downarrow$, or $\wedge$) according to a randomly selected bit value (0 or 1) and basis ($\oplus$ for the pair $\leftrightarrow$ or $\uparrow$ or $\ominus$ for the pair $\downarrow$ or $\wedge$). The encoded photons are sent over the quantum channel to Bob, where he measures each photon using a randomly selected measurement basis ($\oplus$ or $\ominus$). If Alice’s encoding and Bob’s decoding bases match, the photon’s bit value is read correctly with a high probability. Otherwise a random result occurs (i.e., equal likelihood of a 0 or 1). This is due to the inherent uncertainty in the measurement of a quantum system [12].

<table>
<thead>
<tr>
<th>Alice Prepares</th>
<th>Bob Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bit</strong></td>
<td><strong>Basis</strong></td>
</tr>
<tr>
<td>0</td>
<td>$\oplus$</td>
</tr>
<tr>
<td>1</td>
<td>$\ominus$</td>
</tr>
<tr>
<td>0</td>
<td>$\ominus$</td>
</tr>
<tr>
<td>1</td>
<td>$\oplus$</td>
</tr>
<tr>
<td>0</td>
<td>$\ominus$</td>
</tr>
<tr>
<td>1</td>
<td>$\oplus$</td>
</tr>
</tbody>
</table>

**TABLE 1.** Example BB84 protocol.

A. IMPLEMENTATION NON-IDEALITIES AND SYSTEM VULNERABILITIES

While the “unconditionally secure” nature of QKD depends on formal security proofs [12], real-world systems have implementation non-idealities which deviate from their theoretical underpinnings [1], [13]. These non-ideal implementations can lead to vulnerabilities which raise serious concerns regarding the security of QKD systems [3]. For example, state of the art commercially viable QKD systems do not employ on-demand single photon sources [18], [19]. These systems utilize reliable, low cost laser sources to generate classical optical pulses with millions of photons and attenuate them to weak coherent pulses with a Mean Photon Number (MPN) $< 1$. These low energy levels are represented by a Poisson distribution $P(n|\mu) = \mu^n e^{-\mu}/n!$ where $\mu = \text{MPN}$ and $n$ represents the number of photons in each weak coherent pulse [12]. This means, for example, when the MPN of $\mu = 0.5$, $\sim 61\%$ of the pulses have no photons, $\sim 30\%$ of the pulses have one photon, and $\sim 9\%$ of the pulses have two or more photons. Thus, $\sim 23\%$ of the non-empty pulses are insecure multi-photon pulses.

This implementation non-ideality represents a significant security vulnerability which exposes information about the “unconditionally secure” QKD-generated key to eavesdroppers. These insecure multi-photon pulses have been...
the subject of a number of experimental and theoretical attacks since the first QKD was built in 1989 [20]. More specifically, the PNS attack was designed to take advantage of the multi-photon vulnerability in order to gain complete information on Alice and Bob’s shared secret key bits [4], [5].

III. THE PNS ATTACK MODEL

In this section, the PNS attack is functionally decomposed and modeled to accurately account for technological performance limitations and realistic sources of error.

A. THE PNS ATTACK

Suggestions of a PNS-like attack were first made in the early 1990’s [20], [21]; however, it was not until 2000 that the attack was formally defined [4], [5] with additional insights offered shortly thereafter [22]–[26]. In accordance with QKD security proofs, the PNS attack is conducted by an all-power eavesdropper “Eve” who is only constrained by the laws of quantum mechanics. Unique to this attack, Eve does not introduce errors on the quantum channel (i.e., increase the QBER), which is how attacks are typically detected by QKD systems.

**FIGURE 2.** Provides a simplified depiction of the PNS attack conducted against a QKD system (i.e., the transmitter Alice and receiver Bob).

Fig. 2 provides a simplistic depiction of Eve conducting the PNS attack, where she is actively eavesdropping on the quantum channel. In this context, Eve is able to interfere with the quantum channel (i.e., block, manipulate, or fabricate quantum signals) and eavesdrop on the classical channel (i.e., listen but not introduce or spoof messages). According to QKD security proofs, the all-powerful Eve is able to employ any conceivable technologies as long as they do not violate the known laws of quantum mechanics [4], [5]. For example, it is valid (within the laws of quantum mechanics) for Eve to transmit photons to Bob without loss or error even though no technological means currently exists [12].

The PNS attack is typically described in three steps:

(i) Eve determines the number of photons, n, in each optical pulse generated by Alice. If \( n \geq 2 \), the multi-photon pulse is split with one photon stored internally. Otherwise, the pulse is blocked.

(ii) For each multi-photon pulse (i.e., \( n \geq 2 \)), Eve transmits the remaining \( n - 1 \) photons to Bob via a lossless channel.

(iii) Eve listens to the classical channel for Alice’s and Bob’s sifting information (i.e., the bases announcements). Once Eve knows the basis measurement information, she is able to correctly measure each stored photon to obtain a complete copy of the shared secret key bits.

B. CONCEPTUAL MODEL

Fig. 3 provides a conceptual model of the Eve’s physical architecture with the three PNS steps, denoted as (i), (ii), and (iii). While Eve’s construction may vary, this depiction provides a complete representation based on Eve’s architecture as most often described in the available literature [4], [5], [22]–[26]. First, note Eve now constitutes two parts, Eve and Eve′, connected by a classical communications channel, as well as, the third party “Charlie” in support of lossless quantum teleportation.

In order to conduct step (i) of the PNS attack, Eve determines the number of photons, n, in each pulse generated by Alice through a Quantum Non-Demolition (QND) measurement. This specialized measurement uses a projection of Alice’s weak coherent optical pulse to determine the number of photons in each pulse without disturbing the encoded state of each photon [4]. This allows Eve to avoid introducing detectable errors. When \( n = 1 \), the pulse is simply blocked by Eve; however, when \( n \geq 2 \) Eve is configured to split one (1) photon from each multi-photon pulse and store it in her quantum memory, while the remaining \( n - 1 \) photons are transmitted to Bob. Based on well-stated theory [5] and promising experimental result [27]–[29] there are no perceived performance limitations associated with conducting Eve’s QND measurements.

When considering quantum memory in the context of reading, writing, and storing data as single photons, there are various technologies in research and development [30]. While these technologies are currently not sufficient to support a complete PNS attack (i.e., storing 100,000’s of photons for ~60 seconds), the field is rapidly evolving [31]. Thus, it is not unreasonable for Eve’s quantum memory and supporting measurement system to provide sufficient storage, accurate read/write capability, immediate response times, and low error rates when considering the rapidly advancing field of quantum computing [32].

In step (ii), lossless transmission of photons from Eve to Eve′ is often described as the instantaneous, on-demand, error-free quantum teleportation of photons. While this idealized transmission is allowable within the laws of quantum mechanics, there are no known technological means for achieving this functionality. More realistically, Eve uses the third party, Charlie, to distribute entangled photon pairs, along with a complex Bell state measurement to achieve a performance limited version of quantum teleportation. With respect to Charlie’s entangled photon pairs, he must be configured to reliably generate and preposition well-synchronized photon pairs at both Eve and Eve′. This means the entangled photons must be intrinsically correlated such that measuring one photon at Eve simultaneously collapses the corresponding photon at Eve′ into a known
state regardless of the distance between them. While not trivial, current technologies can generate entangled photon pairs [12]. However, these photons must also overcome potentially significant losses and noise along their respective paths from Charlie to Eve and Eve’’. Despite seemingly problematic circumstances, well-synchronized entangled photon pairs at Eve and Eve’’ can be achieved through advanced techniques such as quantum purification [33].

Once the entangled photons are prepositioned, Eve performs a Bell state measurement with the \( n - 1 \) encoded photons to be teleported. This results in one of four measurement outcomes: \(|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, \) or \(|\Psi^-\rangle\), which must be transmitted to Eve’ via her classical channel [33]. However, only two of the four outcomes are definitive, where the non-definitive outcomes can increase the system’s QBER. Thus, performance of the described quantum teleportation scheme is generally limited to 50%.

In order to remedy this shortcoming, Grice proposed using ancillary entangled photon pairs to achieve higher success rates (this phenomenon is detailed further in Section IV) [34]. Additionally, because of Eve’s lossless quantum teleportation, Eve’ must also apply gain matching as to not exceed Bob’s expected detection rate. This can be achieved through a relatively simple variable optical attenuator configured to precisely control Eve’s quantum throughput. While step (ii) is seemingly complex – requiring multiple quantum processes and advanced compensations – reasonable performance can be achieved [5].

In step (iii), Eve listens to Alice’s and Bob’s classical channel for encoding basis information. Eve then uses this information to correctly measure each stored photon (i.e., in the matching prepare-and-measure basis). This measurement allows Eve to gain complete information on each of Bob’s detected qubits. In this way, the PNS attack allows Eve to gain information on the QKD system’s raw secret key bits before any post-processing activities such as entropy estimation and privacy amplification occur.

C. THE PNS MODEL’S DECOMPOSED FUNCTIONALITY

Fig. 4 illustrates the modeled PNS attack against a decoy state QKD system (for a detailed discussion of the decoy state QKD system model see [8]). The PNS attack model is designed in a parameterized fashion to account for the ideal, theoretical PNS attack, as well as, non-ideal versions that account for performance limitations and realistic sources of error. The model also accounts for expected timing delays in Eve such as propagation delays through each of Eve’s optical components and transmission over her classical channel. The model is decomposed into five functions derived from the three PNS attack steps described in III.A and the conceptual model presented in III.B.
Eve’s QND measurement is modeled with no significant performance limitations. This means that each weak coherent optical pulse with an associated Mean Photon Number (MPN), is transformed into a Fock state pulse representation with a discrete number of photons [35]. This allows the specific number of photons per pulse to be determined probabilistically and each photon to be treated independently such that its encoded state is retained during the QND measurement. If desired, performance limitations can be introduced during QND measurement to represent device non-idealities such as multi-photon splits which cause the encoded qubit state to erroneously collapse.

2) QUANTUMTELEPORTATION()

The modeled quantum teleportation function is designed in a parameterized fashion to transfer photons from Eve to Eve’. By default, the function is configured to instantaneously transfer Fock state pulses (i.e., independent photons) from Eve to Eve’ without loss or error. However, the function is configurable to represent practical performance limitations of the teleportation process, where the non-idealities of Eve’s Bell state measurement are captured by introducing erroneous results into Eve’s quantum teleportation process. In this way, Eve can simulate theoretical or experimental configurations of the PNS attack. Moreover, this functionality allows notional quantum communication means to be further explored.

3) MEASUREMENTOUTPUT()

Assuming Eve is able to store and measure photons without error, the output of Eve’s PNS attack is a copy of Alice’s and Bob’s raw secret key bits (i.e., a string of binary 0’s and 1’s). Measurement errors can be added as desired by the researcher.

4) GAINMATCHING()

Eve’ conducts gain matching by applying a specific attenuation to Eve’s teleported pulses in order to meet Bob’s expected detection rate. Eve’s attenuation is specific to the modeled QKD system’s architectural implementation (i.e., losses over the quantum channel and within Bob). The attenuation is applied universally to all pulses propagating through Eve’ to Bob regardless of their photon number.

5) INDUCEERRORS()

Due to engineering limitations, manufacturing defects, and dynamic operational environments, device non-idealities are expected to cause errors (e.g., fielded QKD systems typically have measured error rates of 3-5% QBER [12]). These errors may occur anywhere within the optical path, however, they are only realized where measurements occur. Thus, the model is configured such that errors can be introduced at any modeled optical component and realized at Bob’s detectors [7]. For example, the modeler may want to study the impact of errors at Eve’s QND measurement.

For additional details on the PNS attack model, please see the associated video posted at the IEEE Access website.

IV. PNS ATTACK RESULTS

In this Section, we present simulation results for both ideal and non-ideal versions of the PNS attack. In particular, Eve’s information gain and her detectability are studied for several PNS attack configurations using the model described in Section III. As an example of the model’s ability flexibility, Grice’s proposed method for increasing the success rate of Eve’s quantum teleportation is studied [34].

A. QUANTUM TELEPORTATION PERFORMANCE LIMITATIONS

While there are many technological challenges in realizing the ideal PNS attack, arguably the most significant is Eve’s ability to successfully perform quantum teleportation [9]. Teleportation of arbitrarily encoded photons is a non-trivial task, typically limited to a 50% success rate, due to the inability to discriminate between all four Bell state measurement outcomes (i.e., $|\Phi^+\rangle$, $|\Phi^-\rangle$, $|\Psi^+\rangle$, $|\Psi^-\rangle$) [34]. However, recently Grice proposed a method for increasing the success rate of Eve’s Bell state measurement using ancillary photons [34]. More specifically, the performance of Eve’s quantum teleportation (i.e., her Bell state measurement) can be significantly improved with the addition of ancillary photons.

$$\text{Ancillary Photons} = 2^N - 2$$

(1)

where $N = 1, 2, 3, \ldots, n$. These ancillary entangled photon pairs, known as ancilla, contribute to an arbitrarily high quantum teleportation success rate:

$$\text{Success Rate} = 1 - 1/2^N$$

(2)

Table 2 details the addition of entangled photon pairs as they contribute to Eve’s improved Bell state measurement and thus her quantum teleportation success rate. For example, the addition of one entangled photon pair ($N = 2$) increases
Eve’s quantum teleportation success rate from 50% to 75%. While the addition of several entangled photons pairs is not practical, configurations of up to 63 ancilla are considered to highlight and study the gap between theoretical attacks and what is currently feasible. Moreover, examining Grice’s proposed method is merely example of how the model can be used to study the PNS attack and its impact on QKD systems.

In each of these eight cases, 1,000 simulation runs were completed in order to conduct a thorough assessment of Eve’s PNS attack and provide statistically significant analysis. These simulation results are detailed in the following subsections.

### B. ASSESSMENT OF EVE’S INFORMATION GAIN

Fig. 5 presents simulation results for each of the eight PNS attack configurations examined in this study. In general, as the number of additional ancilla increases, Eve’s quantum teleportation success rate improves and her information gain approaches 100%. Eve’s information gain is primarily dependent upon the success rate of her Bell state measurement scheme \(N = 1, 2, 3, 4, 5, 6, 7\); however, it is also supplemented by the likelihood of randomly guessing the correct bit value (0 or 1) when quantum teleportation is unsuccessful. For example, Eve’s 0-Ancilla PNS configuration \(N = 1\) results in \(\sim 75\%\) information gain because the Bell state measurement success rate is 50% and a 25% contribution occurs due to inherent randomness in Eve’s choice of basis and bit value.

More specifically, Eve will successfully teleport the correct basis and bit value according to the Bell state measurement success rates of Table 2; however, in each case there remains a percentage of unsuccessful attempts. When unsuccessful, Eve must still apply a basis and bit value to each qubit to be teleported. This is accomplished randomly from the perspective of Bob’s measurement. Unsuccessful attempts with non-matching bases are sifted from Bob’s detections and have no impact on Eve’s information gain (or her induced QBER). Conversely, randomly assigned matching prepare-and-measure bases contribute to Eve’s information gain where the random assignment of a 0 or 1 results in Eve gaining additional information on the secret key bits. Because Eve’s only bit options are 0 or 1, she will successfully gain information on approximately 1/2 of the unsuccessful teleportation attempts. Additionally, note that errors during quantum exchange will slightly reduce Eve’s effective information gain; however, assuming Eve knows Alice’s and Bob’s error reconciliation technique, she can recover this information during post processing activities.

In the semi-plausible 1-Ancilla and 3-Ancilla PNS attack configurations of Fig. 5, Eve is able to gain approximately 88% and 94%, respectively, of Alice’s and Bob’s raw key information. In addition, several PNS configurations with increasingly large numbers of supplemental ancilla (i.e., 7, 15, 31, 63 entangled photon pairs) are shown to
FIGURE 6. Simulation results of Eve conducting PNS attacks against the modeled decoy state QKD system (i.e., Alice and Bob of Fig. 4). The performance of several PNS attack configurations with increasing numbers of ancilla (additional entangled photon pairs to improve performance of Eve’s PNS attack) are shown with respect to the QKD system’s measured QBER. Additionally, the QBER and noise thresholds are shown to further illustrate detectability of the PNS attacks.

demonstrate how Eve’s information gain can approach the theoretical attack’s ideal performance. In the ideal PNS attack, Eve is able to gain information on 100% of the QKD-generated raw key bits (minus occasional errors due to noise).

While the impact of the theoretical PNS attack is well-understood, these results demonstrate how much information PNS attacks can gain on Alice’s and Bob’s shared secret key bits with consideration for non-ideal implementations. Next, the detectability of Eve’s PNS attacks are explored with respect to both QBER and the decoy state protocol.

C. EVE’S DETECTABILITY DUE TO INDUCED QBER

The unconditionally secure nature of QKD is based on uncertainty in the qubit’s measurement result when preparing photons in two conjugate bases as with the BB84 protocol (i.e., $\oplus$ or $\otimes$) [36]. This is because Eve must randomly select her measurement basis while eavesdropping on the quantum channel (when not conducting the ideal PNS attack). Thus, her random choice necessarily introduces detectable errors and increases the system’s measured QBER above allowable thresholds. For example, during an intercept-resend attack, Eve intercepts Alice’s encoded photons, measures them, and retransmits the measured values to Bob. For each measured photon, Eve must pick a measurement basis ($\oplus$ or $\otimes$) and will inevitably be wrong 50% of the time. When Eve correctly guesses the matching basis, she can accurately measure and retransmit the correct encoded value with minimal error; however, when she guesses the non-matching basis, she is prone to error because Bob will only measure the correct bit value (0 or 1) 50% of the time. Thus, Eve will introduce a 25% QBER when attempting the intercept-resend attack, which is readily detectable compared to the established QKD QBER security threshold of 11% [13]. Since QBER is the primary detection mechanism for unconditionally secure key distribution, it is important to understand how theoretical and non-ideal PNS attacks impact the system’s QBER.

Fig. 6 depicts the measured QBERs for the eight PNS attack configurations studied. Additionally, a QBER security threshold of 11% [13] is shown along with an expected noise level of 5% [12]. While the security threshold can be further lowered (e.g., 8%), it is often difficult to well-characterize the system’s noise level due to uncontrollable physical disturbances over miles of optical fiber (e.g., vibrations from nearby road, train, or subway traffic). Thus, to avoid false positives (i.e., indications of an attack when none is occurring) security thresholds are expected to remain at 11%.

As expected, the performance limited 0-Ancilla PNS configuration has a relatively high error rate with a measured QBER of $\sim$25%. More specifically, Eve’s $\sim$25% induced QBER is the result of her correctly guessing Alice’s and Bob’s prepare-and-measure basis but randomly assigning the wrong bit value (0 or 1). This is because Eve will not induce errors when she successfully performs the Bell state measurement or when she randomly (from Bob’s perspective) selects the wrong basis measurement (because they will be sifted out). Thus, Eve only introduces errors to the system’s measured QBER when she randomly selects the matching basis and the wrong bit value, which occurs 25% of the time when $N = 1$.

In the 1-Ancilla configuration a QBER of $\sim$12.5% is observed, which is clearly above the security threshold;
however, the 3-Ancilla PNS results demonstrate a QBER of 6.25% which falls below the security threshold and only slightly above expected noise level. As additional ancilla (i.e., 7, 15, 31, and 63) are considered, the attack becomes increasingly undetectable approaching the ideal PNS attack. Recall, the ideal PNS attack does not introduce errors on the quantum channel because Eve makes informed measurement decisions after quantum exchange by listening for sifting basis information as described in step (iii) of the PNS attack. Because the PNS attack is not detectable using the measured QBER, an alternate detection method was introduced in 2003, namely the decoy state protocol [37].

**TABLE 3.** Target QKD system model configuration.

<table>
<thead>
<tr>
<th>State Name</th>
<th>State Description</th>
<th>Mean Photon Number (MPN) Value</th>
<th>Occurrence Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>Facilitates higher secret key distribution rates by using an increased MPN – 0.5 is greater than the value 0.1 typically used in non-decoy state enabled QKD systems</td>
<td>0.5</td>
<td>99%</td>
</tr>
<tr>
<td>Decoy State</td>
<td>Detects PNS attacks through differential analysis between the signal and decoy states</td>
<td>0.1</td>
<td>0.5%</td>
</tr>
<tr>
<td>Vacuum</td>
<td>Characterizes environmental noise on the quantum channel</td>
<td>0.0</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

**D. EVE’S DETECTABILITY WITH THE DECOY STATE PROTOCOL**

Shortly after the PNS attack was formalized, the decoy state protocol was introduced as a mitigation technique [37]. It was quickly improved upon in a series of works [38]–[42] and is now employed in a number of record-holding, high-performance QKD systems [18], [19]. Decoy state enabled QKD systems typically employ three states: Signal, Decoy, and Vacuum as described in Table 3. The decoy state protocol extends the BB84 protocol by configuring Alice to randomly transmit signal, decoy, and vacuum states according to their prescribed occurrence percentages and respective MPNs, where each state must be indistinguishable such that Eve cannot distinguish a decoy state pulse from a signal or vacuum state pulse.

The decoy state QKD system model used in this study (as illustrated in Fig. 4) is based on decoy state protocol configuration as presented in Table 3, which is primarily based on operational characteristics from [18] and [19] and consistent with a comprehensive survey of practically-oriented fielded decoy state QKD systems. Modeling parameters of note include: 10 dB channel loss; 3.5 dB receiver loss; 10% detector efficiency; 2.5E-6 dark count rate, and 0.008 after pulse rate. A detailed description of the decoy state enabled QKD system model is available in [7] and [8].

The decoy state protocol detects PNS attacks by comparing signal and decoy photon number dependent yields described by the security condition [38]

\[ Y_n = Y_n^{\text{signal}} = Y_n^{\text{decoy}} \]  

where \( n = 1, 2, 3, \ldots, N \). However, these photon number dependent yields cannot be measured directly without using costly photon number resolving detectors [43], thus the author’s efficiency-based security condition is used [44]

\[ \eta^{\text{signal}} = \eta^{\text{decoy}}. \]  

Using this security condition the signal and decoy state efficiencies can be directly calculated from standardized system measurements such the state’s MPN \( \mu \), gain \( Q_\mu \) and dark count rate \( Y_0 \) as described for the signal state efficiency:

\[ \eta^{\text{signal}} = \frac{-\ln |1 + Y_0 - Q_\mu|}{\mu} \]  

where the system’s measured dark count rate is

\[ Y_0 = \frac{\text{Number of vacuum state detections}}{\text{Number of vacuum state pulses sent}} \]  

And the measured signal state gain is

\[ Q_\mu = \frac{\text{Number of signal state detections}}{\text{Number of signal state pulses sent}}. \]

The decoy state’s efficiency \( \eta^{\text{decoy}} \) is likewise defined using the dark count \( Y_0 \), the operational decoy state gain \( Q_\nu \), and the decoy state MPN \( \nu \).

Fig. 7 presents simulation results for the decoy state enabled QKD system under normal operations and when subject to PNS attacks. In this case, only the ideal PNS configuration is shown in detail since all eight PNS attack configurations are very similar. This is because the signal and decoy state efficiencies depend on the number of detections (i.e., the state’s measured gain) and not Eve’s information gain or induced QBER.

On the left side of Fig. 7, signal and decoy efficiency results from 1,000 simulation runs are shown for normal operating conditions without PNS attacks. Overlap between the signal and decoy state efficiencies implies secure operation with variation primarily driven by the number of detections per state during each simulation run. The signal state demonstrates relatively little variation compared to the decoy state because of its higher occurrence percentage (99% compared 0.5%) and MPN (0.5 to 0.1). More specifically, the signal state has nearly 40,000 detections per round of quantum exchange, while the decoy state has <100 detections.

On the right side of Fig. 7, the impact of the PNS attack is demonstrated as the decoy state efficiency is significantly reduced outside normal operational parameters. This is because Eve’s PNS attack unavoidably blocks a disproportionally large percentage of decoy state pulses due to its lower MPN (0.1 compared to 0.5). Moreover, the negative impact of the PNS attack on the decoy state efficiency is accentuated when employing a very small occurrence percentage such as 0.5%. In some cases, the PNS
attack blocked nearly all of the decoy state pulses driving the state’s efficiency towards zero. In each of the eight configurations studied, the decoy state protocol can detect the PNS attack with high statistical confidence $P < 0.001$ over 1,000 trials.

V. CONCLUSIONS
In this article, we provide a comprehensive discussion of the PNS attack and present a flexible PNS model capable of studying the impact of implementation non-idealities such as practical limitations in quantum teleportation, QND measurements, or quantum memory. This work is useful for the study of QKD implementation security for system certification, as well as, understanding how advanced quantum attacks can impact theoretical security models. Lastly, this work presents an efficient and repeatable methodology for studying attacks against QKD systems, and more generally, quantum communication protocols. Future work includes using the PNS attack model to study decoy state QKD systems and configurations.

DISCLAIMER
The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the U.S. Government.

REFERENCES


LOGAN O. MAILLOUX (M’12) received the B.S. degree in 2002, the M.S. degree in 2008, and the Ph.D. degree in 2015. He is a Commissioned Officer in the United States Air Force and an Assistant Professor with the Air Force Institute of Technology, Wright-Patterson AFB, OH, USA. His research interests include system security engineering, complex information communication and technology implementations, and quantum key distribution systems. He is a member of Tau Beta Pi, Eta Kappa Nu, INCOSE, and ACM.

DOUGLAS D. HODSON received the B.S. degree in 1985, the M.S. degree in 1987, and the Ph.D. degree in 2009. He is an Assistant Professor of Software Engineering with the Air Force Institute of Technology, Wright-Patterson AFB, OH, USA. His research interests include computer engineering, software engineering, real-time distributed simulation, and quantum communications. He is also a DAGSI Scholar and a member of Tau Beta Pi.

MICHAEL R. GRIMAILA (SM’05) received the B.S. degree in 1993, the M.S. degree in 1995, and the Ph.D. degree in 1999. He is a Professor and the Head of the Systems Engineering Department and a member of the Center for Cyberspace Research with the Air Force Institute of Technology, Wright-Patterson AFB, OH, USA. His research interests include computer engineering, mission assurance, quantum communications and cryptography, data analytics, network management and security, and systems engineering. He is a member of Tau Beta Pi, Eta Kappa Nu, and ACM, and a fellow of the ISSA.

RYAN D. ENGLE (M’14) received the B.S. degree in 2007 and the M.S. degree in 2015. He is currently pursuing the Ph.D. degree with the Air Force Institute of Technology, Wright-Patterson AFB. He is a Commissioned Officer in the U.S. Air Force. His research interests include software engineering, computer engineering, and model-based quantum key distribution systems. He is a member of ACM.

COLIN V. MCLAUGHLIN received the B.A. degree in 2003 and the M.S. degree in 2010. He is a Research Physicist with the United States Naval Research Laboratory, Washington, DC, USA. He specializes in photonic communication devices and systems.

GERALD B. BAUMGARTNER received the B.S. degree in 1971, the M.S. degree in 1973, and the Ph.D. degree from the Illinois Institute of Technology, in 1980. He is a Research Physicist with the Laboratory for Telecommunications Sciences, College Park, MD, USA. His research interests include quantum optics, quantum communications, quantum information, communications security, communications system modeling, and simulation and statistical signal processing. He is a member of the American Physical Society, the Optical Society of America, and the Society for Industrial and Applied Mathematics.