# A Semi-Quantum Extended B92 Protocol and its Analysis

Walter O. Krawec[a] and Sam A. Markelon[a]

[a]University of Connecticut, Storrs, CT, USA

## ABSTRACT

Unlike purely classical communication, unconditionally secure key distribution is possible if Alice and Bob are both equipped with quantum hardware. The degree to which a protocol needs to be quantum is not only an interesting theoretical question, but also important for practical implementations. Indeed, one may wish to construct cheaper devices, or compensate for device malfunction. In this sense, studying limited resource QKD protocols is an important problem.

One direction to studying this is the semi-quantum model introduced by Boyer et al. in 2007 (PRL 99 140501). Several provably secure semi-quantum protocols were put forth. However, most of these protocols were proven secure in the perfect qubit scenario and not necessarily against practical attacks. Only recently, starting with seminal work of Boyer, Katz, Liss, and Mor in (PRA 96 062335) has research in the field of semi-quantum cryptography considered practical devices and imperfections, such as multi photon sources and imperfect detectors. In this work, we present a new SQKD protocol based on an Extended B92 protocol which is able to counter certain practical attacks. Furthermore, the techniques we use may see broad application to other limited-resource (S)QKD protocols.

**Keywords:** Quantum Key Distribution, Cryptography

## 1. INTRODUCTION

A quantum key distribution (QKD) protocol allows two parties, Alice ($A$) and Bob ($B$), to establish a shared secret key, secure against even all-powerful adversaries (referred to throughout as Eve ($E$)); see[1,2] for a general survey. This task is impossible using only classical communication. Thus, a natural question is "how quantum" must a protocol be to gain this advantage over classical communication? To study this, Boyer et al., introduced the notion of semi-quantum key distribution (SQKD) in.[3,4] In this model, one party, typically $A$, is "fully-quantum" in that she can perform any operation on qubits necessary. The other party, $B$ is "classical" in that he can only interact with the quantum channel in a limited, classical manner.

In more detail, such protocols utilize a two-way quantum channel allowing quantum information to travel from $A$ to $B$, then back to $A$. The "classical" user $B$ has two options when he receives a quantum state from $A$. These are:

1. `Measure and Resend`: He subjects the incoming state to a computational $Z$ basis measurement ($|0\rangle$, and $|1\rangle$), resending a computational basis state back to $A$.
2. `Reflect`: He reflects the state back to $A$ undisturbed.

Notice that, essentially, classical $B$ is restricted to either measuring and sending in a single, publicly known, basis, or disconnecting from the quantum channel. If both parties were restricted in this manner, the resulting protocol would be mathematically equivalent to a classical communication protocol. Rather interestingly, security of these protocols is possible in the theoretical, perfect qubit, setting.[5] Recently, results have been extended to practical scenarios, in particular with the seminal "Mirror protocol".[6] For a general survey on semi-quantum cryptography, the reader is referred to.[7]

Further author information: (Send correspondence to W.O.K.)
W.O.K.: E-mail: walter.krawec@uconn.edu

In this paper, we revisit a protocol we introduced in.[8] That SQKD protocol was designed to counter certain practical multi-photon attacks. However, as it was a B92-inspired protocol,[9] it was also susceptible to the Unambiguous State Discrimination (USD) attack.[10,11] Here, we extend this protocol to create a semi-quantum version of the Extended-B92 protocol[12] with an emphasis on practical device security. We show that the USD attack on this protocol is less effective. We also perform a security analysis against a certain subset of collective attacks, hinting at the protocol's overall security. Though we leave a general proof of security (even against arbitrary collective attacks) as future work.

## 2. THE PROTOCOL

The protocol we consider is an extension of one we introduced in.[8] Both protocols utilize certain "boxes," denoted $\mathcal{B}_b$ for $b \in \{0,1\}$ which, abstractly have a quantum input and output along with a classical input and output. This box, on receiving a quantum input $\rho_{in}$ and a classical input $c_{in} \in \{0,1\}$, will behave as follows:

- If $c_{in} = 0$, then the box will `Reflect` the input state, namely $\rho_{out} = \rho_{in}$. The classical output is simply $c_{out} = 0$.
- If $c_{in} = 1$, then the box performs a `Measure and Resend` type operation. With probability $P_{NC} = P_{NC}(\rho_{in})$, the box sets $c_{out} = 0$ and outputs:

$$\rho_{out} = \frac{1}{p_{NC}} \sum_{n \geq 0} q_n^{(b)}(\rho_{in}) \left|b\right\rangle \left\langle b\right|^{\otimes n}.$$

  Otherwise, with probability $1 - P_{NC}$ the box sets $c_{out} = 1$ and outputs:

$$\rho_{out} = \frac{1}{1 - p_{NC}} \sum_{n \geq 0} p_n^{(b)}(\rho_{in}) \left|b\right\rangle \left\langle b\right|^{\otimes n}.$$

The probability values $q_n^{(b)}(\rho_{in})$ and $p_n^{(b)}(\rho_{in})$ depend on the input state and also the box's construction. Our security proof will be performed for any values of these probability values, however to actually evaluate our resulting key-rate bound, we assume $A$ and $B$ are able to compute these values based on a given input state. That is, we are not considering a device-independent box - these boxes must be fully characterized. Note also that, even though our security analysis applies for any $q_n$ and $p_n$ that does not mean that any such values give a secure protocol - indeed, for some boxes, it may be that the resulting key-rate bound is always 0. Note, when clear, we will forgo writing the input state and simply write $q_n^{(b)}$ instead of $q_n^{(b)}(\rho_{in})$. We may also forgo writing the superscript. Note that these boxes may be placed into the context of "mirror-like" devices used in SQKD protocols.[6]

In our prior work,[8] we showed how such a box may be experimentally implemented using polarization encoding. Later, when evaluating, we will simulate such an implementation. Here, the $c_{out}$ value will be 1 if the detector "clicks" (thus $P_{NC}$ is the probability of a No Click). This implies that, if we have a pure state $\left|\psi\right\rangle = \sum_{i \in \{0,1\}^n} \alpha_i \left|i\right\rangle \otimes \left|e_i\right\rangle$, that is an adversarially prepared state consisting of $n$ qubits entangled with Eve's ancilla, then, if $c_{in} = 1$, the box transforms the state to:

$$\rho_{out} = \frac{1}{p_{NC}} \sum_{k \geq 0} q(k) \left|0\right\rangle^{\otimes N-k} \otimes P\left( \sum_{i:w(i)=k} \alpha_i \left|e_i\right\rangle \right),$$

assuming the detector did not click where, above, we use $q(j)$ to denote the probability of a detector clicking if $j$ photons hit it; $P(z) = zz^*$; and $w(i)$ is the Hamming weight of $i$, namely the number of non-zero bits of the string $i$. The above will be important for our simulations later, for more information see.[8]

In prior work, we utilized only a single box $\mathcal{B}_0$ creating a B92 style protocol. Here, we propose adding an additional box $\mathcal{B}_1$, implementing a semi-quantum version of the Extended B92 protocol.[12] The protocol operates as follows:

1. $A$ emits a quantum state $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

2. $B$ chooses a random $b \in \{0, 1\}$ specifying which box he is to use. Next, he chooses a random $k_B$ to be his candidate key bit. He sets $c_{in} = k_B$ for the corresponding box and observes the outcome $c_{out}$.

3. $A$ chooses a random basis $Z$ (spanned by $|0\rangle$ and $|1\rangle$) or $X$ (spanned by $|+\rangle$ or $|-\rangle$) and measures the incoming signal in that basis.

4. $B$ discloses which box he used ($b$) and the value of $c_{out}$. If $c_{out} = 1$, both users discard this iteration.

5. If $A$ observes $|-\rangle$, she sets her key-bit to be 1. Otherwise, if she observes $|1 - b\rangle$, she sets her key-bit to be 0. For any other observation, she informs $B$ to discard the iteration. Bob's key-bit is $k_B$.

That the protocol is correct is easy to see. Whenever $B$ sets $k_B = 0$, the corresponding box will `Reflect` and, so, $A$ will never observe $|-\rangle$ (assuming no noise of course). If $B$ chooses $k_B = 1$ and if $c_{out} = 0$, then the only state leaving the box is of the form $|b\rangle\langle b|$ (possibly multiple copies) and, so, $A$ can never observe $|1 - b\rangle$ (again, assuming no noise). In the next section, we show security of the protocol against a certain class of collective attacks (not all collective attacks, leaving that as future work) and analyze its performance against the USD attack.

## 3. SECURITY ANALYSIS

In this section, we perform an information theoretic security analysis in the asymptotic scenario against certain classes of collective attacks. To do so, we will utilize results in[13,14] which showed that the key-rate $r$ is:

$$r = \lim_{K\to\infty} \frac{\ell(K)}{K} \geq \inf_{\sigma_{AE}} S(B|E)_\sigma - H(B|A), \tag{1}$$

where the infimum is over all density operators $\sigma_{AE}$, resulting from a collective attack that induce the observed noise statistics. Above, we use $S(B|E)$ to denote the conditional von Neumann entropy and $H(B|A)$ the conditional Shannon entropy. Finally, $K$ is the size of the raw key before error correction and privacy amplification, while $\ell(K)$ is the size of the resulting secret key after these two processes are run.

Our first goal, therefore, is to derive a bound on $S(B|E)$ given a particular $\sigma_{BE}$. Initially, $A$ is required to send out a single state $|+\rangle$, however, due to device imperfections (e.g., her use of a weak coherent source), the state she actually sends may be some arbitrary mixed state $\rho_A$. However, since all parts of the protocol are public knowledge, in particular, the exact details of her source preparation devices, $E$ is, in the worst case, completely aware of the state $\rho_A$. Thus, we may assume, in the worst case, that $E$ is the one who actually prepares the signal sent to $B$. Furthermore, as in[8] we assume $E$ sends an $N$-photon state, entangled with her ancilla (i.e., we do not consider general collective attacks where the input state could be a mixture of photon numbers; we leave that as future work and only consider a particular multi-photon input state). It is to $E$'s advantage that this is a pure state. Thus, the state she prepares is:

$$|E\rangle = \sum_{i\in\{0,1\}^N} \alpha_i |i\rangle_T \otimes |\hat{e}_i\rangle_E , \tag{2}$$

where the $|\hat{e}_i\rangle$ are arbitrary, normalized, states in $E$'s ancilla. Note, later, we will denote by $\alpha_0$ to mean $\alpha_{0\cdots0}$ and $\alpha_1$ to mean $\alpha_{1\cdots1}$. The $N$-qubit $T$ register is sent to $B$'s lab. On return, $E$ is allowed to probe the returning signal. As in[8] we assume collective attacks and that $E$ sends only a single, or no, photons to $A$. The action of this unitary probe on certain, to be shown important, states we define as follows:

$$\begin{aligned} U|E\rangle &= |+, f_0\rangle + |-, f_1\rangle + |v, f_v\rangle \\ U|1^N, \hat{e}_{1^N}\rangle &= |0, e_0\rangle + |1, e_1\rangle + |v, e_v\rangle \\ U|0^N, \hat{e}_{0^N}\rangle &= |0, g_0\rangle + |1, g_1\rangle + |v, g_v\rangle , \end{aligned} \tag{3}$$

where $|v\rangle$ represents the vacuum state. The qubit register is sent to $A$ while the remaining portion (the $|f\rangle$, $|e\rangle$, and $|g\rangle$ states) are kept by $E$. Unitarity imposes certain restrictions on these states that will be important momentarily. Note that $U$'s action on other states that may be emitted by $B$'s lab will not be important for our analysis (though, analyzing them in more detail may lead to more optimistic key-rate bounds - a subject of interest for future work).

Since users are using either $\mathcal{B}_0$ or $\mathcal{B}_1$ randomly and disclosing the result over a public channel (thus leaking the information to $E$), we may take advantage of the concavity of von Neumann entropy to show:

$$S(B|E)_\sigma \geq \frac{1}{2}S(B|E)_{\sigma_0} + \frac{1}{2}S(B|E)_{\sigma_1},$$

where $\sigma_i$ is the density operator resulting from the use of $\mathcal{B}_i$.

In our prior work,[8] we showed that for the above form of collective attack, given the resulting density operator $\sigma_0$, it holds that:

$$S(B|E)_{\sigma_0} \geq \left(\frac{p^0_{0,0} + q_N\tilde{p}^0_{1,1}}{M}\right)\left[h\left(\frac{p^0_{0,0}}{p^0_{0,0} + q_N\hat{p}^0_{1,1}}\right) - h(\lambda_0)\right] \tag{4}$$

where:

$$\lambda_0 = \frac{1}{2}\left(1 + \frac{\sqrt{(p^0_{0,0} - q_N\hat{p}^0_{1,1})^2 + 4q_N Re^2\,\langle G|F_0\rangle}}{p^0_{0,0} + q_N\hat{p}^0_{1,1}}\right) \tag{5}$$

$$\tag{6}$$

and:

$$p^0_{0,0} = Pr\left(A \text{ observes } |1\rangle \mid A = Z \wedge c_{in} = 0\right) \tag{7}$$

$$\hat{p}^0_{1,1} = Pr\left(A \text{ observes } |-\rangle \mid A = X \wedge c_{in} = 1 \wedge c_{out} = 0 \wedge N \text{ photons leave his box}\right) \tag{8}$$

We further define the following probabilities:

$$p^0_{0,1} = Pr\left(A\text{observes } |-\rangle \mid A = X \wedge c_{in} = 0\right) \tag{9}$$

$$p^0_{1,0} = Pr\left(A \text{ observes } |1\rangle \mid A = Z \wedge c_{in} = 1 \wedge c_{out} = 0\right) \tag{10}$$

$$p^0_{1,1} = Pr\left(A \text{ observes } |-\rangle \mid A = X \wedge c_{in} = 1 \wedge c_{out} = 0\right) \tag{11}$$

$$\tag{12}$$

allowing us to define the normalization term $M$ as:

$$M = \sum_{i,j} p^0_{i,j}. \tag{13}$$

The above probabilities are all conditioning also on $B$ choosing $\mathcal{B}_0$ of course. Note that $p^0_{0,0}$ is an observable quantity. However $\hat{p}^0_{1,1}$ is not since $B$ can never be sure when $N$ photons leave his box. None the less, it may be bounded by $q_N\hat{p}_{1,1} \leq p_{1,1}$, where $p_{1,1}$ is the actual observed value on average over all number of photons leaving $B$'s lab. Similar bounds may be found, and will be used, for other unobservable quantities of this form.

The above $\lambda_0$ expression depends on a quantity $\langle G|F_0\rangle$ where, based on the analysis in,[8] these are: $|G\rangle = \frac{1}{\sqrt{2}}(|g_0\rangle - |g_1\rangle)$ and $|F_0\rangle = \frac{1}{\sqrt{2}}(|f_0\rangle - |f_1\rangle)$. Expanding the inner product yields:

$$\langle G|F_0\rangle = \frac{1}{2}(\langle g_0|f_0\rangle - \langle g_0|f_1\rangle - \langle g_1|f_0\rangle + \langle g_1|f_1\rangle).$$

Note that, due to unitarity of $U$ (see Equation 3), it holds that:

$$\alpha_0 = \frac{1}{\sqrt{2}}(\langle g_0|f_0\rangle + \langle g_0|f_1\rangle + \langle g_1|f_0\rangle - \langle g_1|f_1\rangle) + \langle e_v|f_v\rangle\,. \tag{14}$$

Solving for $\langle g_0|f_0\rangle$ in the above yields:

$$\langle g_0|f_0\rangle = \sqrt{2}(\alpha_0 - \langle g_v|f_v\rangle) - \langle g_0|f_1\rangle - \langle g_1|f_0\rangle + \langle g_1|f_1\rangle\,.$$

Finally, substituting into the equation for $\langle G|F_0\rangle$ yields the following:

$$\langle G|F_0\rangle = \frac{1}{\sqrt{2}}(\alpha_0 - \langle g_v|f_v\rangle) - \langle g_0|f_1\rangle - \langle g_1|f_0\rangle + \langle g_1|f_1\rangle\,. \tag{15}$$

Now, by the Cauchy-Schwarz inequality, we have: $|\langle g_v|f_v\rangle| \leq \sqrt{\langle g_v|g_v\rangle\langle f_v|f_v\rangle}$. Clearly $\langle f_v|f_v\rangle = Pr(A = vac \mid c_{in} = 0) = 1 - T$, the probability of observing a vacuum in the event $B$ choose to `Reflect`. Also, $\langle g_v|g_v\rangle$ is the probability of observing a vacuum in the event $N$ photons leave $B$'s lab and he choose `Measure and Resend` (this is regardless of the value of $c_{out}$). Consider:

$$Pr(A = vac \mid c_{in} = 1 \wedge c_{out} = 0) = \sum_n q_n Pr(A = vac \mid c_{in} = 1 \wedge c_{out} = 0 \wedge n \text{ photons leave } B\text{'s lab}).$$

The above, therefore, implies we may bound $\langle g_v|g_v\rangle \leq Pr(A = vac \mid c_{in} = 1 \wedge c_{out} = 0)/q_N = (1 - T^2)/q_N$, where we use $T$ to mean the probability of photon transmittance in one direction. Similarly, we may bound $\langle e_v|e_v\rangle$ needed later (except there we condition on the other box being used).

Let $Q = p^b_{1,0}$ and $Q_X = p^b_{0,1}$ (we assume a symmetry in $E$'s attack for $b = 0, 1$ - if the attack is asymmetric, users may abort, a common assumption in QKD security proofs). Note that $p_{0,1}$ is measuring the $X$ basis noise, thus we use $Q_X$ to denote this probability. Then, by Cauchy-Schwarz, and using the same observation we used to bound $\hat{p}_{1,1}$, we have $|\langle g_0|f_1\rangle| \leq \sqrt{q_N(1-Q)Q_X}$, $|\langle g_1|f_0\rangle| \leq \sqrt{q_N Q(1-Q_X)}$, and $|\langle g_1|f_1\rangle| \leq \sqrt{q_N Q Q_X}$. Therefore, assuming $\alpha_0$ is large enough (which may be enforced as we show later), we have:

$$Re\,\langle G|F_0\rangle \geq \frac{1}{\sqrt{2}}(\alpha_0 - [1-T^2]/\sqrt{q_N}) - \sqrt{q_N(1-Q)Q_X} - \sqrt{q_N Q(1-Q_X)} - \sqrt{q_N Q Q_X}. \tag{16}$$

Due to symmetry of the system, we may use the same analysis to derive an equivalent lower bound on $S(B|E)_{\sigma_1}$ using, instead, probabilities $p^1_{0,0}$ and $\hat{p}^1_{1,1}$, defined similarly but now conditioning on $B$ using box $\mathcal{B}_1$.

$$S(B|E)_{\sigma_1} \geq \left(\frac{p^1_{0,0} + q_N\tilde{p}^1_{1,1}}{M}\right)\left[h\left(\frac{p^1_{0,0}}{p^1_{0,0} + q_N\hat{p}^1_{1,1}}\right) - h(\lambda_1)\right] \tag{17}$$

where:

$$\lambda_1 = \frac{1}{2}\left(1 + \frac{\sqrt{(p^1_{0,0} - q_N\hat{p}^1_{1,1})^2 + 4q_N Re^2\,\langle E|F_1\rangle}}{p^1_{0,0} + q_N\hat{p}^1_{1,1}}\right) \tag{18}$$

$$\tag{19}$$

The resulting $\lambda_1$ expression involves $Re^2\,\langle E|F_1\rangle$ where $|E\rangle = \frac{1}{\sqrt{2}}(|e_0\rangle - |e_1\rangle)$ and $|F_1\rangle = \frac{1}{\sqrt{2}}(|f_0\rangle + |f_1\rangle)$. Similar to above, we find:

$$\langle E|F_1\rangle = -\frac{1}{\sqrt{2}}(\alpha_1 - \langle e_v|f_v\rangle) + \langle e_0|f_0\rangle + \langle e_0|f_1\rangle - \langle e_1|f_1\rangle\,. \tag{20}$$

Assuming $\alpha_1$ is large enough (and this can be bounded as we show later), we have:

$$|Re\,\langle E|F_1\rangle| \geq \frac{1}{\sqrt{2}}(\alpha_1 - [1-T^2]/\sqrt{q_N}) - \sqrt{q_N(1-Q)Q_X} - \sqrt{q_N Q(1-Q_X)} - \sqrt{q_N Q Q_X}. \tag{21}$$

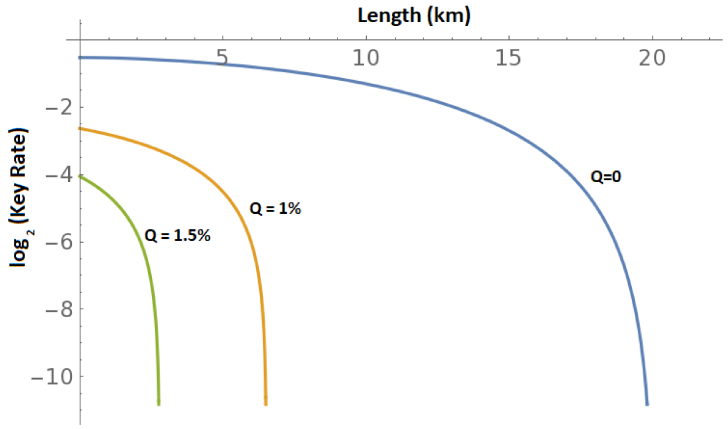


Figure 1. Key-rate of our protocol assuming ideal devices (namely $p_{dc} = 0$ and $\eta = 1$) for various noise levels. We assume $Q_X = Q$ in this setting, though our key-rate bound allows for alternative scenarios. The length here is from $A$ to $B$.

### 3.1 Evaluation

We assume a box implementation as in.[8] Let $p_{dc}$ be the dark count rate of the detector and $\eta$ be its efficiency. In this case, it was shown that:

$$\frac{P_{NC}}{\eta(1-p_{dc})} - \frac{1-\eta}{\eta} \leq |\alpha_i|^2 \leq \frac{P_{NC}}{1-p_{dc}} \tag{22}$$

for $i = 0, 1$ (again, taking $\alpha_0$ to mean $\alpha_{0^N}$ and $\alpha_1$ to mean $\alpha_{1^N}$ from Equation 2. (Of course, only one box was shown in[8] thereby only bounding $\alpha_0$; however the bound is symmetric and can be applied to $\alpha_1$ also.)

We simulate the expected statistics of the protocol over fiber assuming $E$'s attack is symmetric (an enforceable assumption). We use:

$$T = 10^{-.15\ell/10}$$

where $\ell$ is the distance between $A$ and $B$ (i.e., we assume a fiber channel). In this case, we have:

$$p^b_{0,0} = \frac{1}{2}T^2(1-Q) \tag{23}$$

$$p^b_{1,1} = \frac{1}{2}T^2(1-Q) \tag{24}$$

$$p^b_{0,1} = T^2Q \tag{25}$$

$$p^b_{1,0} = T^2Q \tag{26}$$

From,[8] we have $q^b_N = \alpha_b(1-p_{dc})$. Of course, only $b = 0$ was considered in our original work,[8] but the boxes are symmetric and so this bound also applies to box $b = 1$. We assume $Q_X = Q$ and $p_{NC} = .5$. This allows us to evaluate Equation 1 for various distances $\ell$ and noise scenarios $Q$. For ideal devices, namely $\eta = 1$ and $p_{dc} = 0$, the resulting key-rate bound is shown in Figure 1. Non-ideal devices are shown in Figure 2. Note we cannot compare maximal distance to our prior conference work as our original security analysis only considered noiseless attacks.

## 4. UNAMBIGUOUS STATE DISCRIMINATION ATTACK

Our original work in[8] used only a single box and, similar to B92, was therefore susceptible to the unambiguous state discrimination attack.[10,11] Such an attack induces no noise, yet, depending on the loss in the channel, gives Eve full information. We show, here, that our two-box protocol is better able to counter this attack (similar to the fully-quantum Extended B92[12]). This is where the second box becomes critical actually allowing users to gain vital noise statistics that are unavailable with only one box.

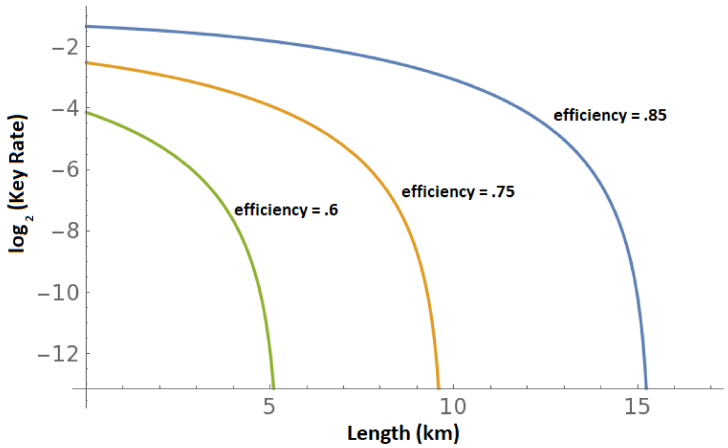


Figure 2. Key-rate of our protocol assuming non-ideal devices (namely $p_{dc} = 10^{-5}$ and $\eta$ varying as shown) for noiseless channels ($Q = 0$).

Let's decompose the initial state $E$ sends (Equation 2) as:

$$|E\rangle = \alpha_0 |0^N, \hat{e}_0\rangle + \alpha_1 |1^N, \hat{e}_1\rangle + \beta |\psi\rangle ,$$

where $\langle\psi|\psi\rangle = 1$ and, so, $|\alpha_0|^2 + |\alpha_1|^2 + |\beta|^2 = 1$. The USD attack induces no noise, therefore, Eve's second attack will take the form:

$$\begin{aligned} U |0^N, \hat{e}_0\rangle &= |0, g_0\rangle + |v, g_v\rangle \\ U |1^N, \hat{e}_1\rangle &= |1, e_1\rangle + |v, e_v\rangle \\ U |\psi\rangle &= |0, x_0\rangle + |1, x_1\rangle + |v, x_v\rangle . \end{aligned}$$

Furthermore, with the USD attack, we may assume $\langle g_0|g_0\rangle = \langle e_1|e_1\rangle = T$, the probability of photon transmittance in one direction. Note that, if only one box is used, say $\mathcal{B}_0$, then statistics on $|e_0\rangle$ could not be gathered and, so, Eve may set it to be a non-zero vector without inducing observable noise. This second box allows us to better bound $E$'s attack and, as we show, rule out the USD attack at least for ideal devices.

By linearity, we have:

$$\begin{aligned} U |E\rangle &= |0\rangle (\alpha_0 |g_0\rangle + \beta |x_0\rangle) + |1\rangle (\alpha_1 |e_1\rangle + \beta |x_1\rangle) + |v\rangle |\chi_v\rangle \\ &= \frac{1}{\sqrt{2}} |+\rangle (\alpha_0 |g_0\rangle + \alpha_1 |e_1\rangle + \beta |x_0\rangle + \beta |x_1\rangle) + \frac{1}{\sqrt{2}} |-\rangle (\alpha |g_0\rangle - \alpha_1 |e_1\rangle + \beta(|x_0\rangle - |x_1\rangle)) + |v\rangle |\chi_v\rangle \end{aligned}$$

where $|\chi_v\rangle$ is some state in $E$'s ancilla that is not relevant to our discussion.

Let $|y\rangle = |x_0\rangle - |x_1\rangle$ and $|v\rangle = \alpha_0 |g_0\rangle - \alpha_1 |g_1\rangle$. To induce no noise, from the above equation, it must be that:

$$0 = \frac{1}{2} || \alpha |g_0\rangle - \alpha_1 |e_1\rangle + \beta(|x_0\rangle - |x_1\rangle) ||^2 = \frac{1}{2} || |v\rangle + \beta |y\rangle ||^2 .$$

Expanding the above and solving for $Re \langle e_1|g_0\rangle$ yields:

$$Re \langle e_1|g_0\rangle = \frac{T(1-\beta^2) + \beta^2 \langle y|y\rangle + 2\beta Re \langle v|y\rangle}{2\alpha_0\alpha_1} \geq \frac{T(1-\beta^2) - 4\beta}{2\alpha_0\alpha_1}, \tag{27}$$

where, above, we used Cauchy-Schwarz along with the trivial bound that $\langle v|v\rangle , \langle y|y\rangle \leq 2$. The above lower-bound will be very important momentarily.

Now, let's consider a single iteration of the protocol assuming a key-bit was distilled. We are interested in the maximal information Eve can have on the key-bit given this attack. Note that, for our original protocol,[8] assuming $T$ was small enough, the above attack leaked *full information* even with perfect devices.

In the event $B$ distills a key-bit of 0 (i.e., he reflected), the state of $E$'s ancilla will be:

$$\rho_0 = \frac{1}{2} |0\rangle \langle 0|_P \otimes P(\alpha_0 |g_0\rangle + \alpha_1 |e_1\rangle + \beta |z\rangle)/2T + \frac{1}{2} |1\rangle \langle 1|_P \otimes P(\alpha_0 |g_0\rangle + \alpha_1 |e_1\rangle + \beta |z\rangle)/2T \tag{28}$$

where $|z\rangle = |x_0\rangle + |x_1\rangle$ and we use the $P$ register to denote the public information transmitted, namely which box $B$ used. On the other hand, if $B$ distills a key-bit of 1, the state of $E$'s ancilla will be:

$$\rho_1 = \frac{1}{2} |0\rangle \langle 0|_P \otimes |g_0\rangle \langle g_0| /T + \frac{1}{2} |1\rangle \langle 1|_P \otimes |e_1\rangle \langle e_1| /T. \tag{29}$$

Using a bound on the quantum Jensen-Shannon divergence,[15] we know that the quantum mutual information between $B$ and $E$, denoted $I(B:E)$ is upper-bounded by:

$$I(B:E) \leq \frac{1}{2}||\rho_0 - \rho_1||.$$

Let $|\phi\rangle = \frac{1}{\sqrt{2}}(\alpha_0 |g_0\rangle + \alpha_1 |e_1\rangle + \beta |z\rangle)$ then, taking advantage of basic properties of trace distance, we have:

$$I(B:E) \leq \frac{1}{4}(|| |\phi\rangle /\sqrt{T} - |g_0\rangle /\sqrt{T}|| + || |\phi\rangle /\sqrt{T} - |e_1\rangle /\sqrt{T}||) \leq \frac{1}{2}(\sqrt{1 - |\langle \phi|g_0\rangle|^2/T^2} + \sqrt{1 - |\langle \phi|e_1\rangle|^2/T^2}). \tag{30}$$

Expanding $\langle \phi|g_0\rangle$ and using Equation 27 yields:

$$\begin{aligned} \langle \phi|g_0\rangle &= \frac{1}{\sqrt{2}}(\alpha_0 T + \alpha_1 \langle g_0|e_1\rangle + \beta \langle z|g_0\rangle) \\ &\geq \frac{1}{\sqrt{2}}(T\alpha_0 - \beta\sqrt{T} + \alpha_1[T(1-\beta^2) - 4\beta]). \end{aligned}$$

Similarly, we find:

$$\langle \phi|e_1\rangle \geq \frac{1}{\sqrt{2}}(\alpha_1 T - \beta\sqrt{T} + \alpha_0[T(1-\beta^2) - 4\beta]). \tag{31}$$

As before, we may bound $\alpha_i$ using:

$$\frac{P_{NC}}{\eta(1-p_{dc})} - \frac{1-\eta}{\eta} \leq |\alpha_i|^2 \leq \frac{P_{NC}}{1-p_{dc}}, \tag{32}$$

which, with ideal devices, will be $P_{NC}$ (and, again with ideal devices, $P_{NC} = 1/2$). $\beta$ may be determined since $|\alpha_0|^2 + |\alpha_1|^2 + |\beta|^2 = 1$. Note, with only one box, only one of the $\alpha_i$ values may be determined, even with ideal settings. Combining with Equation 30 allows us to determine $E$'s maximal mutual information on $B$'s key bit based on $T$ and $\beta$.

The first important observation is that, with ideal devices, the USD attack gives Eve no information regardless of $T$. This is in stark contrast to our original protocol, using only one box, where Eve was able to get full information once $T$ was lower than a certain threshold even if ideal devices were used. For our two-box protocol, the USD attack is ineffective with ideal devices. For non-ideal devices, however, Eve is able to get partial information for various $T$ causing the key-rate to drop as shown in Figure 3.

## 5. CLOSING REMARKS

We introduced a new semi-quantum key distribution protocol inspired by the fully-quantum Extended B92 protocol.[12] This work extends our B92 style semi-quantum protocol in[8] to incorporate a second encoding scheme allowing us to better counter the USD attack. Furthermore, we refined our security analysis method to incorporate a larger class of attacks compared to our prior work in.[8] Though, we still leave an analysis of all collective attacks as future work. Furthermore, finding a better security proof technique may lead to more optimistic bounds on the performance of this protocol.

**Acknowledgments:** WK is partially supported by the NSF under grant number 1812070.

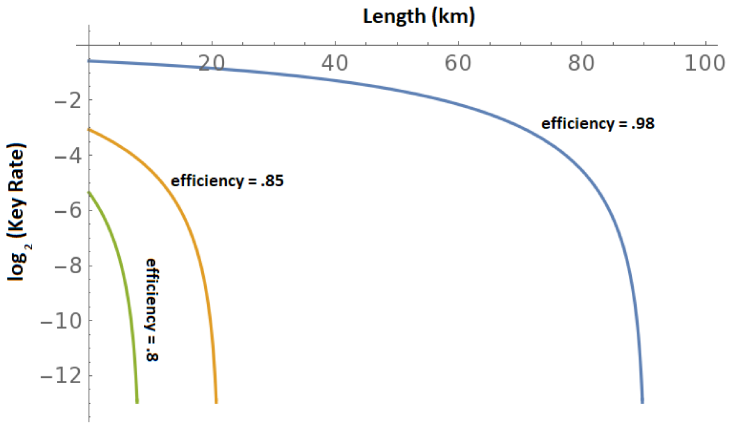


Figure 3. Key-rate of our protocol (using our bound on $I(B:E)$ along with the Devetak-Winter key-rate expression $I(B:A) - I(B:E)$ from[14] where, as this attack is noiseless, we have $I(B:A) = 1$ for various detector efficiencies (we set $p_{dc} = 0$ and $p_{NC} = .5$ for this graph).

## REFERENCES


[1] Scarani, V., Bechmann-Pasquinucci, H., Cerf, N. J., Dušek, M., Lütkenhaus, N., and Peev, M., "The security of practical quantum key distribution," *Rev. Mod. Phys.* **81**, 1301–1350 (Sep 2009).

[2] Pirandola, S., Andersen, U., Banchi, L., Berta, M., Bunandar, D., Colbeck, R., Englund, D., Gehring, T., Lupo, C., Ottaviani, C., et al., "Advances in quantum cryptography," *arXiv preprint arXiv:1906.01645* (2019).

[3] Boyer, M., Kenigsberg, D., and Mor, T., "Quantum key distribution with classical bob," *Phys. Rev. Lett.* **99**, 140501 (Oct 2007).

[4] Boyer, M., Gelles, R., Kenigsberg, D., and Mor, T., "Semiquantum key distribution," *Phys. Rev. A* **79**, 032341 (Mar 2009).

[5] Krawec, W. O., "Security proof of a semi-quantum key distribution protocol," in [*Information Theory (ISIT), 2015 IEEE International Symposium on*], 686–690, IEEE (2015).

[6] Boyer, M., Katz, M., Liss, R., and Mor, T., "Experimentally feasible protocol for semiquantum key distribution," *Physical Review A* **96**(6), 062335 (2017).

[7] Iqbal, H. and Krawec, W. O., "Semi-quantum cryptography," *Quantum Information Processing* **19**(3), 97 (2020).

[8] Krawec, W. O., "Practical security of semi-quantum key distribution," in [*Quantum Information Science, Sensing, and Computation X*], **10660**, 1066009, International Society for Optics and Photonics (2018).

[9] Bennett, C. H., "Quantum cryptography using any two nonorthogonal states," *Phys. Rev. Lett.* **68**, 3121–3124 (May 1992).

[10] Tamaki, K., Koashi, M., and Imoto, N., "Security of the bennett 1992 quantum-key distribution protocol against individual attack over a realistic channel," *Physical Review A* **67**(3), 032310 (2003).

[11] Ko, H., Choi, B.-S., Choe, J.-S., and Youn, C. J., "Advanced unambiguous state discrimination attack and countermeasure strategy in a practical b92 qkd system," *Quantum Information Processing* **17**(1), 17 (2018).

[12] Lucamarini, M., Di Giuseppe, G., and Tamaki, K., "Robust unconditionally secure quantum key distribution with two nonorthogonal and uninformative states," *Physical Review A* **80**(3), 032327 (2009).

[13] Renner, R., Gisin, N., and Kraus, B., "Information-theoretic security proof for quantum-key-distribution protocols," *Phys. Rev. A* **72**, 012332 (Jul 2005).

[14] Devetak, I. and Winter, A., "Distillation of secret key and entanglement from quantum states," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* **461**(2053), 207–235 (2005).

[15] Briët, J. and Harremoës, P., "Properties of classical and quantum jensen-shannon divergence," *Physical review A* **79**(5), 052311 (2009).