Unconditional security of coherent-state quantum key distribution with strong phase-reference pulse

Masato Koashi

CREST Research Team for Photonic Quantum Information, School of Advanced Sciences, The Graduate University for Advanced Studies (SOKENDAI), Hayama, Kanagawa, 240-0193, Japan

We prove the unconditional security of a quantum key distribution protocol in which bit values are encoded in the phase of a weak coherent-state pulse relative to a strong reference pulse. In contrast to implementations in which a weak pulse is used as a substitute for a single-photon source, the achievable key rate is found to decrease only linearly with the transmission of the channel.

PACS numbers: 03.67.Dd 03.67.-a

Quantum key distribution provides a way to distribute a secret key between two distant parties, Alice and Bob, even if the quantum channel between them suffers from small noises. As long as the law of quantum mechanics is valid, an eavesdropper, Eve, cannot force Alice and Bob to accept a key on which she has a nonnegligible amount of information. A proof of such unconditional security was first provided by Mayers [1], followed by other proofs [2, 3, 4, 5, 6, 7, 8]. While a perfect single-photon source is assumed in the earlier proofs, recent proofs [3, 4, 5, 6, 7, 8] cover the use of a weak laser pulse in a coherent state as a substitute for a single photon. This is good news in the practical point of view, but comes with a price: the multiphoton components of the weak pulse allow Eve a so-called photon-number splitting attack [3, 4, 5, 6]. In order to achieve the security under this attack, Alice must lower the amplitude of her weak pulse as the loss in the channel increases. As a result, there is a bound [10] on the achievable key rate which scales as $O(n^2)$ with channel transmission $n$.

In this paper, we prove an unconditional security of a scheme using a weak coherent pulse and achieving a rate that scales as $O(n)$. The scheme is essentially the one proposed by Bennett [11], in which a strong pulse is transmitted as a phase reference together with a weak pulse containing the bit information in the relative phase. We made a minor modification to introduce a second local oscillator (LO) for Bob. This makes the analysis simpler, and allows us to assume a realistic threshold detector that may be noisy, inefficient, sensitive to multimodes of light, and only discriminates the vacuum from one or more photons.

The scheme is depicted in Fig. 1(a). Suppose that Alice’s LO emits a strong pulse in a coherent state with complex amplitude $|\alpha_0|e^{i\phi_A}$. Using an asymmetric beamsplitter (BS1), Alice extracts a weak pulse with very small amplitude $\alpha = |\alpha|e^{i\phi_A}$, and encodes a randomly chosen bit value 0 or 1 by applying phase shift 0 or $\pi$, resulting in state $|\alpha\rangle$ or $|-\alpha\rangle$, respectively. Together with this signal, she sends the strong pulse from the other output of BS1 to Bob as a phase reference.

On the receiver’s side, Bob chooses randomly a bit value 0 or 1, and applies phase shift 0 or $\pi$ to the weak signal pulse, respectively. Instead of using the reference pulse from Alice directly, Bob uses another LO and tries to lock its phase to Alice’s one. Suppose that Bob’s LO produces a strong pulse with complex amplitude $|\beta_0|e^{i\phi_B}$. Combining a portion of this pulse and the reference pulse from Alice, he conducts a series of interference experiments (M) to infer the phase difference $\phi_A - \phi_B$. He then applies a phase shift equal to this estimated value $\phi^*$ to his LO, and mixes it with the weak signal from Alice at BS2. The mixed signal is measured by a threshold detector, which gives a “click” whenever it receives one or more photons. Bob reports the outcome of the detector to Alice over an authenticated public channel. The click implies a conclusive result, and both parties accept their bits. No click implies an inconclusive result, and they discard the bits.

The security analysis in this paper is valid even if LOs with phases $\phi_A$ and $\phi_B$ are available to Eve. Then, the reference pulse from Alice gives no information to Eve. The only effect of Eve’s attack on this pulse is to disturb the measurement outcome $\phi^*$ to be deviated from the desired value, as $\phi^* = \phi_A - \phi_B - \Delta\phi$. But exactly the

![FIG. 1: (a) A scheme with a strong reference pulse. (b) An equivalent scheme except that Eve’s region is extended.](image-url)
same effect can be obtained by just applying the phase shift $\Delta \phi$ to the weak signal from Alice (Eve may simulate $M$ by herself). Hence we can safely assume that Eve simply ignores the strong reference pulse. Similarly, any imperfection in the estimation process $M$, including the fundamental limitation arising from finiteness of the amplitudes of the two LOs, has the same effect as introducing a noise source applying a phase shift $\Delta \phi$ on the weak signal while assuming a perfect estimation, $\phi^* = \phi A - \phi B$.

The major imperfections in the detector can be treated as follows. Suppose that the quantum efficiency of the detector is $\eta_0$, the transmission coefficient of BS2 is $\eta_{BS2}$ and the amplitude of LO incident on BS2 is $(1 - \eta_{BS2})^{-1} \eta_1^{-1} \beta$. Then, the same measurement can be implemented by inserting a lossy medium (BS3) with transmission $\eta_{BS3} \eta_0^{-1}$, then mixing LO with amplitude $(1 - \eta_0)^{-1} \beta$ by a beamsplitter BS4 with transmission $\eta_0$, followed by a detector with unit efficiency. Here we take the limit of $\eta_0 \to 1$. The dark counting of the detector or the detection of stray photons can be simulated by a device (P) that inserts a photon in a mode that is orthogonal to the modes of the LOs. We thus finally arrive at a scheme with an ideal threshold detector and a locked pair of LOs, as in Fig. 1(b). In this figure, the region accessible by Eve is extended for the sake of simplicity. If a protocol is secure with this scheme, the same protocol accessible by Eve is extended for the sake of simplicity. If a protocol is secure with this scheme, the same protocol is also secure.

Bob’s decision process in the scheme in Fig. 1(b) can be regarded as a generalized measurement on the light entering his site with three outcomes, 0, 1, and 2, where the last one means “inconclusive”. Let $\mathcal{H}_B = \mathcal{H}_0 \otimes \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n \otimes \cdots$ be the Hilbert space for the light modes received by Bob that are sensible by the detector. The mode $\nu = 0$ represents the pulse mode of Bob’s LO, and the modes with $\nu \geq 1$ are orthogonal to it. Let us write the coherent state $|\beta\rangle_0 |0\rangle_1 |0\rangle_2 \cdots$ simply as $|\beta\rangle$. Then, the generalized measurement is described by the POVM \{\$F_0, F_1, F_2\}, where \[
F_0 = (1 - |\beta\rangle \langle \beta|)/2, \quad F_1 = (1 - |\beta\rangle \langle \beta|)/2,
\]
and $F_2 = 1 - F_0 - F_1$. If everything is ideal except for the transmission $\eta$ in the channel, Alice’s signal is received by Bob in coherent states $| \pm \sqrt{\eta} \alpha \rangle$, and they can agree on a key without errors by choosing $\beta = \sqrt{\eta} \alpha$.

Before describing the proof of unconditional security, we introduce several notations. We decompose $\mathcal{H}_B$ as $\mathcal{H}_B = \mathcal{K}_B \oplus \mathcal{H}_{ex}$, where $\mathcal{K}_B$ is the two-dimensional subspace spanned by $|\beta\rangle$ and $|-\beta\rangle$. We assume $\alpha$ and $\beta$ to be real and positive without loss of generality. Let $|\{\mu_i\}; k\}_{i=1,2,\ldots}$ be an arbitrary complete orthonormal basis for $\mathcal{H}_{ex}$. We identify $\mathcal{K}_B$ as a qubit, and define its $X$ basis as $|0,e\rangle_B \equiv (|\beta\rangle + |\beta\rangle) / (2\sqrt{2})$, $|1,e\rangle_B \equiv (|\beta\rangle - |\beta\rangle) / (2\sqrt{2})$, where $2\sqrt{2} - 1 = 1 - 2s^2 = -|\beta\rangle$ with $\beta = e^{-2\beta\eta^2}$. The $Z$-basis states are denoted as $|j,e\rangle_B \equiv (|0,e\rangle_B + (-1)^j |1,e\rangle_B) / \sqrt{2} (j = 0, 1)$. For Alice’s side, we denote by $\mathcal{H}_A$ the Hilbert space of the light modes emitted from her site. We also introduce an auxiliary qubit in Alice’s site, with Hilbert space $\mathcal{K}_A$. We denote the $X$- and the $Z$-basis states as $|j,e\rangle_A$ and $|j,e\rangle_A (j = 0, 1)$. We sometimes denote the projection $|\Phi\rangle\langle\Phi|$ as $P(|\Phi\rangle)$.

The key idea in the security proof is a trade-off in increasing completely positive map, which is specified by Kraus operators $A_j : \mathcal{H}_B \to \mathcal{K}_B (j = 0, 1, 2, \ldots)$ defined by $A_j = s_j |0,e\rangle_B |0,e\rangle_B + c_j |1,e\rangle_B |1,e\rangle_B$ for $j = 0$ and $A_j = |z,e\rangle_B |z,e\rangle_B$ otherwise. Since $\sum_j A_j^\dagger A_j \leq 1$, there exists a filter with the following property. It takes any state $\rho$ acting on $\mathcal{H}_B$ as an input, and it accepts with probability $p = \sum_j \text{Tr}(A_j^\dagger A_j \rho)$ while it rejects with probability $1 - p$. Whenever it accepts, it returns the output state $\sum_j A_j \rho A_j^\dagger / p$ acting on $\mathcal{K}_B$. This filter is related to the POVM \{\$F_0, F_1, F_2\} by \[
F_k = \sum_j A_j^\dagger |k,e\rangle_B \langle k,e| A_j \quad (1)
\]
for $k = 0, 1$, which is easily confirmed. This relation implies that we can implement the measurement \{\$F_0, F_1, F_2\} by applying the filter and conducting $Z$-basis measurement on the output state when it accepts (if it rejects, we assume that the outcome is “2”).

With the above decomposition of Bob’s measurement, we can prove the unconditional security by a method similar to the cases of qubit-based B92 protocols. We introduce a protocol based on entanglement distillation, which is later shown to be equivalent to the real protocol. In the new protocol, (1) Alice prepares state $|0_e\rangle_A |\alpha\rangle + |1_e\rangle_A |\alpha\rangle |\alpha\rangle / \sqrt{2}$ on $\mathcal{K}_A \otimes \mathcal{H}_A$. We assume that Alice produces $2N$ copies of this state. (2) Eve receives $2N$ pulses (corresponding to $\mathcal{H}_B^{2N}$) from Alice, and prepares a state on $\mathcal{H}_B^{2N}$, which may be entangled to Eve’s system. (3) After Bob has received $2N$ pulses (corresponding to $\mathcal{H}_B^{2N}$), Alice and Bob randomly permute the order of $2N$ pairs of systems by public discussion. (4) For the first $N$ pairs (check pairs), Alice measures each qubit $\mathcal{K}_A$ on $Z$ basis, and Bob performs the POVM \{\$F_0, F_1, F_2\} on each pulse ($\mathcal{H}_B$). They disclose all the results, and learn the number $n_{err}$ of error events where the combination of Alice’s and Bob’s outcomes are $(0, 1)$ or $(1, 0)$. (5) For the other $N$ pairs (data pairs), Bob applies the above filter to each pulse, and discloses each result (accept or reject). Let $n_{fil}$ be the number of events where the filter has accepted. (6) Alice and Bob now have $n_{fil}$ pairs of qubits ($\mathcal{K}_A \otimes \mathcal{K}_B$), from which they try to extract a number of pairs in the maximally entangled state $|0_e\rangle_A |0_e\rangle_B + |1_e\rangle_A |1_e\rangle_B / \sqrt{2}$. To do so, they estimate the number $n_{th}$ of pairs with a bit error (represented by the subspace spanned by $|0_e\rangle_A |1_e\rangle_B, |1_e\rangle_A |0_e\rangle_B$) and the number $n_{ph}$ of pairs with a phase error (the subspace spanned by $|0\rangle_A |1_e\rangle_B, |1\rangle_A |0_e\rangle_B$), from the knowledge of $n_{fil}$ and $n_{err}$. If neither number of errors is too high, they run an entanglement distillation protocol (EDP) and then measure on $Z$ basis to determine the
final key. As in the proof of BB84 [4], if the estimation of the upper bounds for \( n_{\text{hit}} \) and \( n_{\text{ph}} \) is correct except for a probability that becomes exponentially small as \( N \) increases, this protocol is essentially secure.

According to the argument by Shor and Preskill [4], if we choose an appropriate EDP scheme, Alice and Bob can conduct \( Z \)-basis measurement on the \( n_{\text{hit}} \) pairs immediately after step (5) and decide the final key by a public discussion without compromising the security. Then, Eq. (1) shows that Bob’s measurement on each data qubit is also the POVM \( \{F_0, F_1, F_2\} \). Alice’s measurement can be further brought forward to the end of step (1), then this step is equivalent to just preparing state \( |\alpha\rangle \) or \( |-\alpha\rangle \) randomly. The new protocol is thus equivalent to the prepare-measure protocol implemented as in Fig. (II b).

The remaining task for the security proof is to establish an exponentially good way of estimating \( n_{\text{hit}} \) and \( n_{\text{ph}} \). Since \( n_{\text{err}} \) and \( n_{\text{hit}} \) are the results of the same measurement applied to the (randomly assigned) check pairs and to the data pairs, we can apply a classical probability estimate to see that \( |n_{\text{hit}} - n_{\text{err}}| \leq N\epsilon \) holds except for a small probability which is asymptotically smaller than \( \sim\exp(-N\epsilon^2) \). The estimation of \( n_{\text{ph}} \) can be done by considering what could have happened if Alice and Bob measured their \( n_{\text{hit}} \) pairs of data qubits in \( X \) basis and determined \( n_{\text{ph}} \) by discussion, just after the step (5). In this scenario, they obtain three numbers \( (n_{\text{hit}}, n_{\text{ph}}, n_{\text{err}}) \).

The following argument shows that some combinations of \( (n_{\text{hit}}, n_{\text{ph}}, n_{\text{err}}) \) are exponentially rare for any attack by Eve, and hence gives an (exponentially reliable) upper bound \( n_{\text{ph}}(n_{\text{hit}}, n_{\text{err}}) \) for \( n_{\text{ph}} \) as a function of the other two.

We can regard \( n_{\text{ph}} \) as the number of events where a measurement on \( \mathcal{K}_A \otimes \mathcal{H}_B \) produced the outcome corresponding to the element of a POVM \( M_{\text{ph}} \equiv \sum A_j|\alpha\rangle\langle\alpha| + A_j|\beta\rangle\langle\beta| \otimes P(|\alpha\rangle\langle\alpha|) \otimes P(|\beta\rangle\langle\beta|) \) for \( j = 1, 2 \) and \( \epsilon > 0 \). Similarly, \( n_{\text{hit}} \) and \( n_{\text{err}} \) are also obtained by the projection measurement \( \{P_{01}, P_{10}, P_{01}, P(|\alpha\rangle\langle\alpha|)\otimes1_{\text{ex}}, P(|\beta\rangle\langle\beta|)\otimes1_{\text{ex}}\} \), where \( P_{ij} \equiv P(|\alpha\rangle\langle\alpha|) \) for \( i,j = 1, 2 \), followed by a classical procedure composed of Bernoulli trials. If we denote the results of the \( N \) projection measurements as \( (n_1, n_1, \ldots, n_N) \), these numbers should be related to \( n_{\text{ph}} \) and \( n_{\text{hit}} \) as

\[
|n_{\text{ph}} - m_1 - n_-| \leq 2\epsilon \\
|n_{\text{hit}} - m_0 - m_1 - n_-| \leq 2\epsilon \\
|n_{\text{err}} - m_0 - m_1 - n_-| \leq 2\epsilon
\]

with probability at least \( 1 - \exp(2N\epsilon^2) \). Since the marginal state \( \rho_A \) on \( \mathcal{K}_A \) cannot be altered by Eve, the \( X \)-basis measurement on \( \mathcal{K}_A \) is another Bernoulli trial.

For the check pairs, \( n_{\text{err}} \) corresponds to \( M_{\text{err}} \equiv \sum P_{ij} \equiv P(|\alpha\rangle\langle\alpha|) \otimes P(|\alpha\rangle\langle\alpha|) + P(|\beta\rangle\langle\beta|) \otimes P(|\beta\rangle\langle\beta|) \), where we have introduced a basis \( \{\Gamma_{ij}\}_{i,j = 0, 1} \) of \( \mathcal{K}_A \otimes \mathcal{K}_B \) by \( \Gamma_{ij} \equiv c_\delta|\alpha\rangle\langle\alpha| + (1-\delta)|\beta\rangle\langle\beta| \). It implies that \( n_{\text{err}} \) could also be obtained by the global projection measurement \( \{Q_{00}, Q_{11}, Q_{01}, Q_{10}\} \), followed by Bernoulli trials. If we write the results of \( N \) projection measurements as \( (n'_1, n'_2, n'_3, n'_4) \), we obtain

\[
|n_{\text{err}} - (n'_1, n'_2, n'_3, n'_4)/2| \leq 2\epsilon N
\]

If we compare the projection measurements on the data pairs and the check pairs, we further notice that \( n_+ \) and \( n'_+ \) are the results of an identical measurement, namely, projection onto the space \( \mathcal{H}_+ \) spanned by \( \{0_x\}_A|0_x\}_B, |1_x\}_A|1_x\}_B \}. We can thus apply the classical probability estimate. \( \delta_+ \) and \( \delta'_+ \) comes from projection to nonorthogonal states. For such a case, it was shown [6] that combination \( (\delta_+, \delta'_+) \) is exponentially rare unless there exists a state \( \rho \) on \( \mathcal{H}_+ \) satisfying \( \text{Tr}[\rho |\Gamma_{ij}\rangle\langle\Gamma_{ij}|] = \delta_+ \) and \( \text{Tr}[\rho |\Gamma_{ij}\rangle\langle\Gamma_{ij}|] = \delta'_+ \). Using these arguments, we obtain

\[
|n_+ - n'_+| \leq 2\epsilon
\]

\[
\delta_+ \geq c_\delta \delta_+ + \delta'_+ \mid 1-\delta_+ \mid - 2c_\delta s_{\beta} \sqrt{\delta_+ \mid 1-\delta_+ \mid} - \epsilon'.
\]

We are interested in the secret key gain in the limit \( N \to \infty \). Setting \( \epsilon \) and \( \epsilon' \) to be zero, we obtain \( 2n_{\text{err}} \geq n_{\text{hit}} - 2c_\beta s_{\beta} (\sqrt{\delta_+} (1-\delta_+) + n_- \sqrt{\delta_-} (1-\delta_-)) \). Then we can eliminate \( n_+ \) and \( \delta_+ \) to be left with two free parameters \( m_0 \) and \( m_1 \). From this point, in general, we may have to numerically minimize \( n_{\text{err}} \) over the two parameters. It turned out that in most of interesting cases \( m_0 = m_1 = 0 \) gives the minimum. Once we obtain the minimum of \( n_{\text{err}} \) as a function of \( (n_{\text{ph}}, n_{\text{hit}}) \), we can determine \( n_{\text{ph}}(n_{\text{hit}}, n_{\text{err}}) \). The length of the final key is given
by $n_{\text{key}}(n_{\text{fil}}; n_{\text{err}}) = n_{\text{fil}}[1-h(n_{\text{err}}/n_{\text{fil}})h(n_{\text{ph}}/n_{\text{fil}})]$ when this value is nonnegative and $2n_{\text{ph}} \leq n_{\text{fil}}$.

Figure 2 shows the parameter region $(n_{\text{fil}}, n_{\text{err}})$ where the key gain $G \equiv n_{\text{key}}/N$ is positive, for a few choices of $\alpha$ and $\beta = \sqrt{\pi} \alpha$. When Alice chooses $|\alpha|^2 = 0.5$, the tolerable error rate $n_{\text{err}}/n_{\text{fil}}$ is less than 1%. For a smaller amplitude $|\alpha|^2 = 0.001$, the tolerable rate increases to $\sim 7\%$. Choosing a smaller value for $|\alpha|^2$ than this example does not improve the tolerable rate significantly. For either case in Fig. 2, taking a smaller value of $\eta$ gives little change in the shape of region, except for the normalization factor in the abscissa $n_{\text{fil}} = N(1 - e^{4\eta |\alpha|^2})/\sqrt{2} \sim 2\eta |\alpha|^2 N$. This allows us to choose a fixed $|\alpha|^2$ in the limit of $\eta \to 0$ as long as $n_{\text{err}}/n_{\text{fil}}$ is fixed, leading to the key gain $G$ proportional to $\eta$.

In Fig. 2(b), we notice that the region extends far into the area with high $n_{\text{err}}/n_{\text{fil}}$ and $n_{\text{fil}}$, but ordinary sources of errors never achieve this region. For example, Errors in phase $[\Delta \phi$ in Fig. 1(b)] result in curve B. Errors by spurious countings [device P in Fig. 1(b)], which is modeled as $n_{\text{fil}} = N\lambda + (1 - \lambda) n_{\text{fil}}$ and $n_{\text{err}} = N\lambda/2$, follow curve A.

In order to achieve a high key gain, we can optimize over $|\alpha|^2$ for a given model of errors. Here we assume that all errors are spurious countings (curve A), and take $\lambda = \gamma + 1 - \exp[-|\alpha|^2\eta(1 - 2\zeta)]$. The first term is the contribution independent of $|\alpha|^2$, such as the dark counting rate of the detector. The rest represents “misalignment errors”, which are caused by a stray light proportional to the strength of LO. A mode mismatch between Alice’s and Bob’s LO is an example of this type of errors. We chose the parameter $\zeta$ such that $n_{\text{err}}/n_{\text{fil}} \to \zeta$ for $\eta \to 0$ when $\gamma = 0$. Assuming this model, we optimized $G$ over $|\alpha|^2$, which is shown in Fig. 3. For $\gamma = \zeta = 0$ [curve (a)], the key gain decreases as $G \sim O(\eta)$, which should be compared to the $O(\eta^2)$ decrease in the case [c] of a coherent-state source is simply substituted for a single-photon source in BB84 [curve (d)]. When $\eta$ is small, the optimal choice is $|\alpha|^2 \sim 0.23$, which gives $n_{\text{fil}}/N \sim 0.91(\eta/2)$. The raw key is shorter by factor $\sim 0.69$, leading to $G \sim 0.29(\eta/2)$. This value is smaller than the ideal BB84 $G = \eta/2$ [curve (c)] by a constant factor. If we include a small alignment error ($\zeta = 3\%$), the key rate drops by a constant factor but the $O(\eta)$ dependence remains [curve (b)]. This tendency continues up to $\zeta \sim 7\%$, at which the key gain is zero for any value of $|\alpha|^2$. Finally, if we include a contribution of dark counting $\gamma$, each curve drops to zero when the overall counting rate is comparable to $\gamma$.

In summary, we have shown that by encoding on the phase of a weak coherent pulse relative to a strong reference pulse, we can achieve a key rate of $O(\eta)$ with unconditional security, which is an advantage over the coherent-state BB84. There are several proposals [12] to improve the performance of the coherent-state BB84, and their unconditional security is an interesting problem.

The author thanks N. Imoto, H.-K. Lo, D. Mayers, J. Preskill, K. Tamaki, and especially N. Lütkenhaus for helpful discussions.

[1] D. Mayers, Lect. Notes Comput. Sci. 1109, 343-357 (1996).
[2] C. H. Bennett and G. Brassard, in Proceeding of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India (IEEE, New York, 1984), pp.175-179.
[3] E. Biham et al., quant-ph/9912053 (1999).
[4] W. -Y. Hwang, Phys. Rev. Lett, 85, 441 (2000).
[5] H. Inamori, N. Lütkenhaus, and D. Mayers, quant-ph/0107017.
[6] D. Gottesman, H. -K. Lo, N. Lütkenhaus, and J. Preskill, quant-ph/0212066.
[7] K. Tamaki, M. Koashi, and N. Imoto, Phys. Rev. Lett. 90, 167904 (2003).
[8] K. Tamaki and N. Lütkenhaus, quant-ph/0308048.
[9] G. Brassard et al., Phys. Rev. Lett. 85, 1330 (2000).
[10] N. Lütkenhaus, Phys. Rev. A 61, 052304 (2000).
[11] C. H. Bennett, Phys. Rev. Lett. 68, 3121 (1992).
[12] C. H. Bennett, D. P. Divincenzo, J. A. Smolin, and W. K. Wooters, Phys. Rev. A 54, 3824 (1996).
[13] W. -Y. Hwang, Phys. Rev. Lett. 91, 057901 (2003). V. Scarani et al., Phys. Rev. Lett. 92, 057901 (2004).