Robustness of differential-phase-shift quantum key distribution against photon-number-splitting attack

Kyo Inoue and Toshimori Honjo
NTT Basic Research Laboratories, NTT Corporation, 3-1 Morinosato Wakamiya, Atsugi, 243-0198, Japan
(Received 18 January 2005; published 4 April 2005)

A photon-number-splitting (PNS) attack against differential-phase-shift (DPS) quantum key distribution (QKD) is described. In the conventional 1984 Bennett-Brassard protocol, using weak laser light, the PNS attacks, which involve installing a lossless transmission line and blocking pulses from which extra photons cannot be picked up, impose a limit on the transmission distance. In contrast, use of a coherent pulse train in DPS QKD prevents the PNS attack and removes the distance limitation imposed by it. We carried out a DPS QKD experiment that simulated the situation where some pulses are blocked. The result showed that extra bit errors are induced in an eavesdropped condition, indicating the robustness of DPS QKD against PNS attacks.

DOI: 10.1103/PhysRevA.71.042305 PACS number(s): 03.67.Dd, 42.50.Dv

I. INTRODUCTION

Quantum key distribution (QKD) provides an unconditionally secure secret key to two legitimate parties (Alice and Bob) for ciphering and deciphering messages [1]. Photons are usually used to carry key bit information. Although a single-photon source that emits just one photon in one pulse is desired for highly secure QKD systems [2,3], such a device is difficult to implement and a weak laser pulse is usually used instead in actual experiments. However, use of a weak laser pulse allows an eavesdropper (Eve) to conduct a photon-number-splitting (PNS) attack [4,5]. Even a highly attenuated pulse, e.g., 0.1 photon per pulse on average, has a finite probability that it contains more than two photons according to the photon statistics of laser light. Eve can extract a part of the key bit information from these extra photons without being detected by Alice and Bob. The PNS attack is especially effective for the well-known and widely employed QKD 1984 Bennet-Brassard (BB84) protocol [6], which operates over lossy transmission lines. The transmission distance in the BB84 scheme using weak laser light with practical system parameters is limited to around 50 km due to the possibility of a PNS attack [5]. Entanglement-based QKD schemes are robust against such attacks [1], but are hard to implement at the present. Recently, a protocol modified from the BB84 scheme has been proposed [7], which is also robust against a PNS attack. The unique sifting procedure of the modified BB84 protocol prevents Eve from obtaining full information by means of the PNS attack, and the transmission distance is enlarged as a result.

In this paper, we study the robustness of a QKD protocol called differential-phase-shift (DPS) QKD against PNS attacks. DPS QKD is a recently proposed QKD scheme that uses a weak coherent pulse train as a signal carrier [8,9]. It is shown that the use of a pulse train prevents Eve from blocking photons and removes the distance limitation imposed by the PNS attack.

II. PHOTON-NUMBER-SPLITTING ATTACK

A. BB84 protocol

First, we briefly describe how a PNS attack is carried out against the conventional BB84 scheme using strongly attenuated laser light. Alice sends out weak coherent states into the transmission line. Eve probes the states just after Alice's output using a photon-number quantum-nondemolition measurement and judges whether one state contains more than two photons or not. From states that contain more than two photons, she takes out one photon, stores it, and then lets the remaining photons go to Bob through a lossless transmission line. For states that contain one or no photons (i.e., states from which she does not take a photon), on the other hand, she blocks them as long as the blocking does not reduce the photon number received by Bob. After the photon transmission, Alice and Bob disclose the basis information. Eve listens in and then measures the stored photons according to that basis information.

This PNS attack restricts the transmission distance between Alice and Bob as follows [5]. In the normal condition, Bob's raw detection rate per state is $R_n = L\mu$, provided that Bob's detectors are perfect, where L is the transmission loss and μ the mean photon number sent from Alice. When Eve conducts a PNS attack, Eve's detection rate per state is $R_e = p_2$, where p_2 is the probability that one pulse contains more than two photons at Alice's output. When $R_n = R_e$, Eve can obtain all information about the sifted key without being detected by Alice and Bob. For $\mu = 0.1$ and a transmission loss of 0.25 dB/km, for example, the condition $R_n = R_e$ is satisfied for a transmission length of about 50 km, which is regarded as the maximum distance between Alice and Bob in BB84 QKD systems. When Bob's detection efficiency is not perfect (typically 10% in practice in the fiber communication wavelength), the possible distance is further reduced.

B. DPS QKD

The distance limitation in BB84 systems imposed by PNS attack arises because of Eve's strategy of blocking pulses that contain one or no photon. With this strategy, she can realize the condition that all photons received by Bob are identical to those stored by her, provided that the probability of more than two photons in a state is equal to the transmission loss. If Eve cannot employ this strategy, her PNS attack will not be so successful. This is the case for the differential-

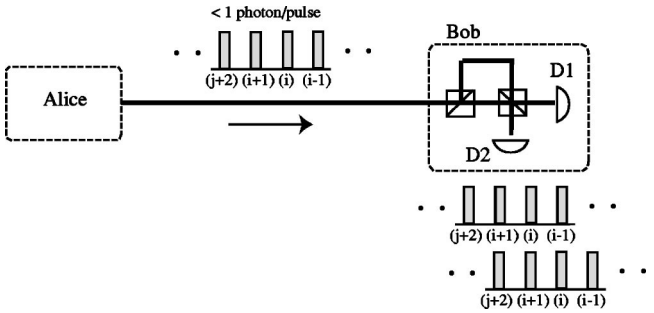

FIG. 1. The configuration of DPS QKD. D1 and D2 are photon detectors.

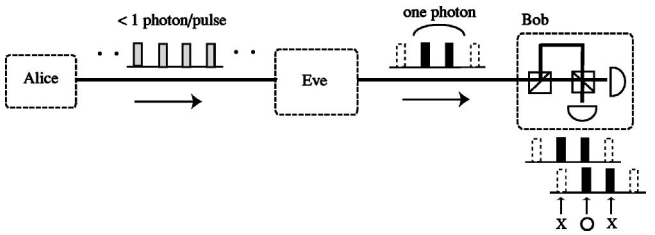

FIG. 2. State change through photon-number-splitting attack in DPS QKD.

phase-shift QKD protocol in fact. We discuss the PNS attack against DPS QKD in this section.

First, we briefly describe the setup and how DPS QKD works [8]. The basic configuration is shown in Fig. 1. Alice sends out a coherent pulse train in which each pulse has less than one photon on average and is phase modulated by 0 or π. The coherence time of the pulse train is much longer than the pulse interval. Bob receives the pulse train with an asymmetric Mach-Zehnder interferometer whose path-length difference is equal to the time interval of the incoming pulses. At the recombining coupler of the interferometer, neighboring pulses interfere with each other, and photons are detected by either detector D1 or D2 depending on the phase difference between neighboring pulses. A secret key bit is created from which a detector counts a photon, i.e., from whether the phase difference is 0 or π. Since the transmitted photon number is less than one per pulse on average, a photon is not detected at every time slot and at which time slot a photon is detected is probabilistic. This uncertainty originates the security of DPS QKD.

A PNS attack against the above DPS QKD can be considered as follows. Since bit information is embedded over two sequential pulses, a photon positioned over two neighboring pulses is needed for Eve to extract bit information. Thus, she probes the transmitted signal and judges whether more than two photons are positioned over two pulses or not (though we do not know how to do it). From pulses that contain more than two photons, she takes out one photon, stores it, and lets the two pulses go to Bob through a lossless transmission line. For other pulses, she blocks them. Then after the signal transmission, Eve measures the stored photons based on time information disclosed by Bob. That is a PNS attack analogous to one against conventional QKD schemes.

Unfortunately for Eve, however, this PNS attack is revealed as follows. When Eve conducts the above PNS attack, Bob receives a signal in which one photon is definitely positioned over two sequential pulses and no photon in other pulses, as illustrated in Fig. 2. In other words, the PNS attack changes the transmitted signal from a condition where every pulse has a finite probability of having a photon into a condition where some two sequential pulses definitely have one photon and other pulses have no probability of having photons. For such a signal, Bob's detectors click possibly at three time slots as illustrated in Fig. 2. A click at the middle time slot occurs according to interference between the two pulses, which gives Bob a correct answer. On the other hand, a click at the first or last time slots randomly occurs because a pulse having a photon probability interferes with a vacant pulse at these time slots. Thus, bit errors can be introduced from the detection events at the first and last time slots. The probability that the detectors click at the first or last time slot is 0.5, and then the error probability is 0.25. The eavesdropping is revealed from this error rate.

The above error rate 0.25 comes from a signal condition where two isolated pulses go to Bob. Eve may try to reduce this error rate by sending three (or more) isolated sequential pulses, i.e., one photon positioned over three pulses, to Bob. For such a signal, Bob counts a photon possibly at four time slots with a probability ratio of 1:2:2:1. The detection events at the first and fourth time slots randomly occur, which can induce an error with $1/6$ probability. Thus, the error rate is reduced from that for two isolated pulses. However, Eve cannot obtain full information in this eavesdropping. For sending three pulses in the framework of PNS attacks, she measures if two photons are positioned over three pulses, takes out one photon from three pulses that contain two photons, stores it, and lets the three pulses go to Bob. She keeps the photon positioned over three pulses until Bob counts the other photon and discloses its detection time. Note that she must keep her photon in a state where those three pulses have an equal probability of having a photon, because she cannot predict which phase difference Bob will measure. Then after listening in to the detection time, Eve measures the corresponding phase difference. This measurement is conducted such that she makes the first two pulses (or the last two pulses) interfere with each other when Bob measures the phase difference between the first two pulses (or between the last two pulses). In this measurement, one pulse out of the three is inevitably discarded, which means that Eve misses the photon with a probability of $1/3$ and cannot obtain full information from her stored photon. Since Alice and Bob can create a secure key through privacy amplification under such a condition, this eavesdropping is not successful.

In any event, the discussion above indicates that a PNS attack that blocks pulses cannot be carried out against DPS QKD. Eve would have to adopt a strategy that does not block pulses. A possible one is a passive photon-splitting attack as illustrated in Fig. 3, where Eve inserts a beam splitter on the transmission line just after Alice and partially splits the transmitted signal. At the beam-splitter output that goes to Bob, she installs a lossless transmission line in order to compensate signal loss due to beam splitting. The split pulses are passed through an asymmetric interferometer and then detected. After the signal transmission from Alice to Bob, Bob

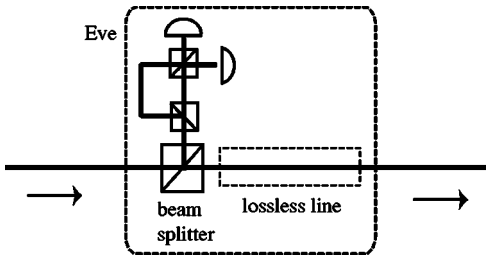

FIG. 3. Possible photon-splitting attack against DPS QKD.

discloses time slots at which he detected photons. Eve listens in and obtains bit information from her detection events corresponding to Bob's detection time.

In the above eavesdropping scenario, Eve obtains key bit information when she detects a photon at a time slot for which Bob also detects a photon. The probability of Bob detecting a photon at a particular time slot is $L\mu$, and that of Eve doing so is $\alpha\mu$, where α is the beam-splitting ratio in Eve. Thus, the probability that photons are detected by both Bob and Eve at a corresponding time slot is $L\mu^2\alpha$. Then the probability that Eve obtains key bit information relative to Bob is $L\mu^2\alpha/L\mu=\mu\alpha$. Here, $\mu,\alpha<1$; thus the amount of Eve's information is always less than Bob's regardless of the transmission loss. This indicates that the transmission distance in DPS QKD is not limited by the photon-splitting attack, unlike the case in the conventional BB84 QKD using weak laser light.

The discussion in this section concludes that DPS QKD is more robust than the conventional BB84 protocol against PNS attacks.

III. EXPERIMENT

The point of the robustness of DPS QKD discussed above is that Eve cannot block pulses in an attempt to obtain the same amount of key bit information as Bob's. If she blocks pulses, bit errors are induced in Bob's key bits and the eavesdropping is revealed. To confirm this, we carried out an experiment with the configuration shown in Fig. 4. cw light from an external cavity laser diode (wavelength =1.55 μm) was intensity modulated by a $LiNbO_3$ modulator driven by a pulse pattern generator. We generated a consecutive pulse train of 1 GHz repetition rate or a repetitive pulse pattern of (0001100000) at 1 gigabit/s, with a pulse width of 125 ps in both cases. The former pulse pattern simulated a situation

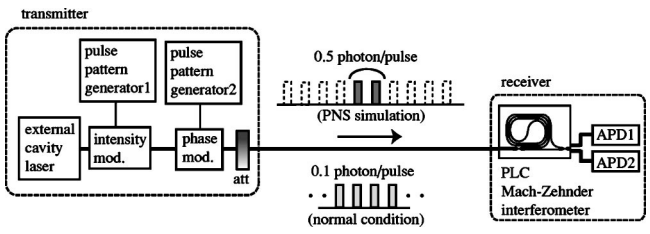

FIG. 4. The experimental setup simulating a photon-number-splitting attack that blocks pulses. PLC, planar lightwave circuit; att, attenuator.

where the system works in the normal condition, and the latter that where Eve blocks some pulses in a PNS attack. Each pulse was randomly phase modulated by 0 or π, attenuated, and then sent out toward the receiver. The output power was adjusted so that one pulse had 0.1 and 0.5 photons on average for the former and latter pulse patterns, respectively. With this power adjustment, the photon-counting rate in the receiver was the same for both situations. The pulses were transmitted through a 20 km fiber, and then input to the receiver, which consisted of a planar lightwave circuit (PLC) Mach-Zehnder interferometer module and gated avalanche photo diode (APD) photon detectors [9]. The interferometer had a path length difference of 20 cm, which corresponded to the time interval of the transmitted pulses. The APD gating rate was 5 MHz, at which the afterpulse effect was sufficiently small.

Using the above setup, we conducted signal transmission, created sifted key strings in the transmitter and receiver, and then evaluated the bit error rate by comparing those key strings. The obtained error rates were 3.9% for the normal situation and 28.3% for the pulse-blocked situation. The bit errors in the normal condition were due to imperfections in the experiment, such as imperfect interference in the interferometer, the timing jitter, and the APD dark count. The extra errors induced in the pulse-blocked condition resulted from there being no interfering detection events due to vacant pulses. The expected error rate induced by the pulse blocking is 25%, as described in the previous section. Taking the experimental error into account, the error rate in the pulse-blocked condition should be $25+3.9\times3/4=28.0\%$, which is nearly equal to the value obtained in our experiment. Thus, the above experiment confirmed that blocking pulses in DPS QKD induces bit errors and makes Alice and Bob aware of the PNS attack.

IV. SUMMARY AND DISCUSSION

Photon-number-splitting attacks against differential-phase-shift QKD were studied. The use of a coherent pulse train, in which every pulse has an equal probability of having a photon, prevents Eve from blocking pulses from which she cannot take out extra photons. As a result, the transmission distance in DPS QKD systems is not limited by the PNS, unlike in conventional BB84 systems using weak laser light.

The mechanism of this robustness of DPS QKD is basically the same as that of the Bennet 1992 (B92) QKD systems using a strong reference pulse [5,10]. In the original B92 protocol, a weak coherent pulse (signal pulse) is sent from Alice to Bob together with a strong reference pulse. When Eve blocks a signal pulse, a noninterfering detection event occurs from the reference pulse, which induces bit errors and makes Alice and Bob aware of the eavesdropping. However, the power ratio of the signal pulse to the reference pulse should be quite large in long-distance systems, e.g., $10^{-4}/4$ for an 80 km system [5]. That requirement is hard to satisfy in practice. DPS QKD provides the same robustness as the B92 protocol using a strong pulse without such difficulty.

[1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).
[2] E. Waks, K. Inoue, C. Santori, D. Fattal, J. Vuckovic, G. Solomon, and Y. Yamamoto, Nature (London) **420**, 762 (2002).
[3] A. Beveratos, R. Brouri, T. Gacoin, A. Villing, J.-P. Poizat, and P. Gransgier, Phys. Rev. Lett. **89**, 187901 (2002).
[4] G. Brassard, N. Lutkenkaus, T. Mor, and B. C. Sanders, Phys. Rev. Lett. **85**, 1330 (2000)
[5] A. Acin, N. Gisin, and V. Scarani, Phys. Rev. A **69**, 012309 (2004).
[6] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 175.
[7] V. Scarani, A Acin, G. Ribordy, and N. Gisin, Phys. Rev. Lett. **92**, 057901 (2004).
[8] K. Inoue, E. Waks, and Y. Yamamoto, Phys. Rev. A **68**, 022317 (2003).
[9] T. Honjo, K. Inoue, and H. Takahashi, Opt. Lett. **29**, 2797 (2004).
[10] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).