Quantum Cryptography Using Any Two Nonorthogonal States

Charles H. Bennett
IBM Research Division, T. J. Watson Research Center, Yorktown Heights, New York 10598
(Received 23 December 1991)

Quantum techniques for key distribution—the classically impossible task of distributing secret information over an insecure channel whose transmissions are subject to inspection by an eavesdropper, between parties who share no secret initially—have been proposed using (a) four nonorthogonally polarized single-photon states or low-intensity light pulses, and (b) polarization-entangled or spacetime-entangled two-photon states. Here we show that in principle any *two* nonorthogonal quantum states suffice, and describe a practical interferometric realization using low-intensity coherent light pulses.

PACS numbers: 03.65.Bz, 42.50.Wm, 89.70.+c

Key distribution is the term applied to techniques allowing two parties to acquire a random bit sequence (the "key") with a high level of confidence that no one else knows it or has significant partial information about it. One party (henceforth "Alice"), for example, might generate the key by a physically random process, make a copy of it, and hand deliver the copy to the other party (henceforth "Bob"). Such shared secret key bits, although random and meaningless in themselves, are a valuable resource because they allow the communicating parties to achieve, with provable security, two of the main goals of cryptography: encrypting a subsequent meaningful message to make it unintelligible to a third party [1], and certifying to the legitimate receiver that a message (plain or encrypted) has not been altered in transit [2].

If two parties share no secret information initially and communicate solely through classical messages monitored by an eavesdropper, it is impossible for them to arrive at a certifiably secret key [3]. However, it becomes possible to do so if they exchange both classical public messages (which can be monitored but not altered or suppressed by the eavesdropper) and quantum transmissions having the property that they can be suppressed or altered, but cannot in principle be monitored without disturbance [4]. Various types of quantum transmissions have been shown to suffice: a random sequence of spin-$\frac{1}{2}$ particles or single photons in four non-orthogonal polarization states (e.g., $\leftrightarrow$, $\updownarrow$, $\circlearrowright$, and $\circlearrowleft$); an analogous random sequence of low-intensity polarized coherent or incoherent light pulses [5]; a sequence of polarization-entangled Einstein-Podolsky-Rosen [6] (EPR) two-photon states [7]; and an analogous sequence of spacetime-entangled two-photon states produced, for example, by parametric down-conversion [8–11].

An earlier paper [12] has shown a general equivalence between EPR-based [7,12] key distribution and non-EPR schemes using on nonorthogonal states [4], arising from the fact that *measuring* one member of an EPR pair is equivalent to *preparing* the other member in a random state corresponding to the result of the measurement. Table I shows a typical key distribution scheme in both its EPR and non-EPR versions. The end result in either case is a sequence of random key bits, with evidence either that it is shared and secret, or else that it has been

TABLE I. EPR and non-EPR key distribution.

EPR	non-EPR														
1a	–	○	+	○	+	+	+	+	+	○	○	+	○	○	○	+
2a	2b	↺	|	↻	↔	|	|	↔	↔	↻	↺	|	↻	↺	↺	|
3	3	+	○	○	+	+	○	○	+	○	+	○	○	○	○	+
4	4	|		↻	↔	|	↺	↺	↔	↻	|	↻	↻		↺	|
5	5	+		○	+	+	○	○	+	○	+	○	○			○	+
6	6			√	√	√			√	√			√			√	√
7	7			↻	↔	|			↔	↻			↻			↺	|
8	8			1	0	1			0	1			1			0	1
9	9			1	0				0				1				1
10	10				0	1			0	1			1			0	1

[1a] In the EPR version, Alice chooses a random basis for measuring one member of each EPR pair of photons: rectilinear (+), or circular (○). The other photon of each EPR pair is measured by Bob in step 3.

[2a] Alice's measurement results in effect determine, through the EPR correlations, a random sequence of states for Bob's photon: horizontal ($\leftrightarrow$), vertical ($\updownarrow$), right-circular ($\circlearrowright$), and left-circular ($\circlearrowleft$).

[2b] In the non-EPR version, Alice *prepares* a random sequence of photons polarized $\leftrightarrow$, $\updownarrow$, $\circlearrowright$, and $\circlearrowleft$, and sends them to Bob.

[3] Bob measures his photon using a random sequence of bases.

[4] Results of Bob's measurements. Some photons are shown as not having been received owing to imperfect detector efficiency. (Realistic detectors would also generate occasional errors due to dark counts, which can be found and corrected as described in [5].)

[5] Bob tells Alice which basis he used for each photon he received.

[6] Alice tells him which bases were correct.

[7] Alice and Bob keep only the data from these correctly measured photons, discarding all the rest.

[8] This data is interpreted as a binary sequence according to the coding scheme $\leftrightarrow = \circlearrowleft = 0$ and $\updownarrow = \circlearrowright = 1$.

[9] Bob and Alice test their key by publicly choosing a random subset of bit positions and verifying that this subset has the same parity in Bob's and Alice's versions of the key (here the parity is odd). If their keys had differed in one or more bit positions, this test would have discovered that fact with probability $\frac{1}{2}$.

[10] Remaining secret key after Alice and Bob have discarded one bit from the chosen subset in step 9, to compensate for the information leaked by revealing its parity. Steps 9 and 10 are repeated k times, with k independent random subsets, to certify with probability $1-2^{-k}$ that Alice's and Bob's keys are the identical, at the cost of reducing the key length by k bits.

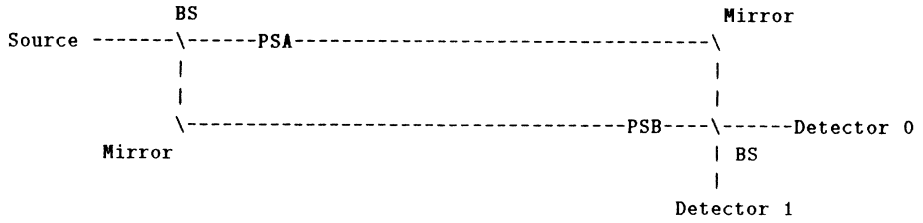

FIG. 1. Interferometric quantum key distribution using four nonorthogonal one-photon states. Alice's source at left supplies single photon states, which are split by a symmetric beam splitter BS into two arms of a Mach-Zehnder interferometer. Alice applies (PSA) a random 0-, 90-, 180-, or 270-deg phase shift in one arm; Bob (PSB) a random 0- or 90-deg phase shift in the other arm. After the quantum transmission, Alice and Bob agree publicly to keep only those instances in which their phase shifts differ by 0 or 180 deg, causing the photon to behave deterministically at the second beam splitter.

disturbed by eavesdropping, and should be discarded.

Although the most familiar example of the EPR effect involves two-particle states with nonclassical spin correlations, it has been known for some time [8,9,13] that other two-particle states can be prepared, for example, by parametric down-conversion, which exhibit entirely analogous correlations in *phase* that can be observed interferometrically. Recently Franson [10] and Ekert, Rarity, and Tapster [11] have pointed out that these correlations too can be used for key distribution. Here we note the existence of one-particle versions of these two-particle schemes. As in the spin-based key distribution of Table I, the one-particle version involves random preparations by one party and random measurements by the other, while the two-particle scheme uses random measurements by both parties. A one-particle interferometric key distribution scheme, involving preparation and measurement of four states nonorthogonal with respect to phase, is shown in Fig. 1.

In [12] the security of non-EPR key distribution schemes is derived from the fact that any measurement which fails to disturb each of two nonorthogonal states also fails to yield any information distinguishing them. This naturally suggests the possibility that key distribution might be performed using only *two* nonorthogonal states, instead of the four used in Table I and Fig. 1. Here we show that key distribution is possible in principle using any two nonorthogonal states of a quantum system.

Let $|u_0\rangle$ and $|u_1\rangle$ be two distinct, nonorthogonal states, and let $P_0 = 1 - |u_1\rangle\langle u_1|$ and $P_1 = 1 - |u_0\rangle\langle u_0|$ be (noncommuting) projection operators onto subspaces orthogonal to $|u_1\rangle$ and $|u_0\rangle$, respectively (note reversed order of indices). Thus P_0 annihilates $|u_1\rangle$, but yields a positive result with probability $1 - |\langle u_0|u_1\rangle|^2 > 0$ when applied to $|u_0\rangle$, and vice versa for P_1.

To begin the key distribution, Alice prepares and sends Bob a random binary sequence of quantum systems, using states $|u_0\rangle$ and $|u_1\rangle$ to represent the bits 0 and 1, respectively. Bob then decides, randomly and independently of Alice for each system, whether to subject it to a measurement of P_0 or P_1. Next Bob publicly tells Alice in which instances his measurement had a positive result (but not, of course, which measurement he made), and the two parties agree to discard all the other instances.

If there has been no eavesdropping, the remaining instances, a fraction approximately $(1 - |\langle u_0|u_1\rangle|^2)/2$ of the original trials, should be perfectly correlated, consisting entirely of instances in which Alice sent $|u_0\rangle$ and Bob measured P_0, or Alice sent $|u_1\rangle$ and Bob measured P_1. However, before Alice and Bob can trust this data as key, they must, as in other key distribution schemes, sacrifice some of it to verify that their versions of the key are indeed identical. This also certifies the absence of eavesdropping, which would necessarily have disturbed the states $|u_0\rangle$ or $|u_1\rangle$ in transit, causing them sometimes to yield positive results when later subjected to measurements P_1 or P_0, respectively.

Figure 2 shows a practical interferometric realization, in which the two nonorthogonal states $|u_0\rangle$ and $|u_1\rangle$ are dim coherent light pulses differing in phase relative to an accompanying bright reference pulse (bright coherent states, typically nearly orthogonal, become significantly nonorthogonal when attenuated to < 1 expected photon intensity, because all such dim states include a significant component of the zero photon number state). Beginning at the left side of the figure, Alice uses an arrangement of unsymmetric beam splitters (UBS) and mirrors to split an initial coherent pulse into two pulses separated in time: a dim signal pulse of intensity $\mu < 1$ expected photons followed by a bright reference pulse of $M > 1$ expected photons. The signal pulse is phase shifted (PSA) 0 or 180 deg to encode the bits 0 and 1, then launched into a single mode optical fiber. The brighter reference pulse is not phase shifted, but is delayed by a fixed time Δt then launched into the same fiber. At the receiving end of the apparatus, Bob uses a half-interferometer similar to Alice's to split the incoming beam again, in the same ratio as before, into a dim part and a bright part. As before the dim part is phase shifted (PSB) by 0 or 180 deg, randomly and independently of Alice's phase shifts, while the bright part is delayed by Δt. Finally the two parts are caused to interfere as they enter a detector.

The wave entering the detector consists of three pulses separated by times Δt. The first pulse, a very dim pulse

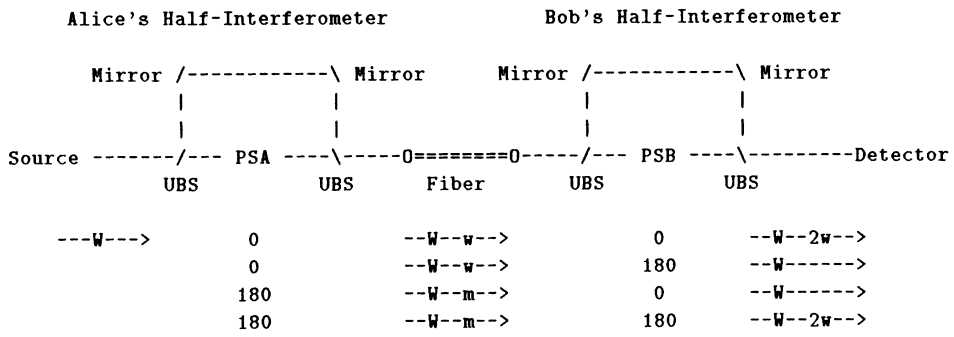

FIG. 2. Interferometric quantum key distribution using two nonorthogonal low-intensity coherent states. Source at left supplies coherent pulse (wave form -W-) of $M > 1$ expected photons intensity to Alice's half-interferometer, where unsymmetric beam splitters (UBS), mirrors, and phase shifter (PSA = 0 or 180 deg) produce a dim signal pulse (-w- or, phase shifted, -m-), followed by a bright reference pulse -W-. Sent to Bob through a single mode optical fiber, the pulses enter Bob's half-interferometer, where, depending on whether the sum of Alice's and Bob's phase shifts (PSA + PSB) is 0 or 180 deg, the signal pulse undergoes constructive (wave form -2w-) or destructive (wave form ---) interference with the attenuated reference pulse before entering the detector. Arriving before this interference pulse is a very dim pulse (not shown) attenuated by both Alice and Bob but delayed by neither. Arriving after the interference pulse is a bright twice-delayed reference pulse (wave form -W-) which Bob monitors to be sure the reference pulses are not being suppressed. Also not shown are two unused beams leaving the rightmost beam splitter of each half-interferometer in the downward direction.

which has been attenuated both by Bob and by Alice but delayed by neither, is not considered further.

The second pulse, containing the important key information, is a dim pulse consisting of the superposition of the beam delayed by Alice and attenuated by Bob, and the beam delayed by Bob and attenuated by Alice. If Alice's and Bob's phase shifts are equal, constructive interference will occur and the superposed pulse will generate a count with probability $\approx 4\mu Tq$ expected photons, where T is the transmission coefficient of the fiber and q the quantum efficiency of the detector. If Alice's and Bob's phase shifts differ, the superposed pulse will have much lower intensity, ideally zero in the limit of perfect interferometer alignment (the coherence time of the light source is not an issue here, since the two interfering pulses are exactly proportional, being attenuated versions of the same source pulse).

Finally, at a delay Δt after the superposed pulse, a bright pulse, which has been delayed by both Alice and Bob but attenuated by neither, will arrive at Bob's detector. Bob confirms its arrival, with approximately the expected intensity MT, which he can do reliably if $MTq > 1$. This third pulse contains no phase information, but serves to confirm that the reference pulse has actually arrived. It thus protects against an attack in which an eavesdropper ("Eve") would measure each signal-reference pulse pair by an apparatus similar to Bob's, resend a correctly fabricated pulse pair whenever she was successful, and suppress both the signal and reference pulses when she was unsuccessful, thereby eavesdropping on the channel without creating errors in Bob's subsequent measurement results. Eve cannot suppress the reference pulse without immediately being caught. But if she suppresses only the signal pulse, the uncancelled reference pulse will still produce a count in Bob's detector with probability, μTq, and half these counts will result in errors in Bob's key.

The encoding of each bit in the phase difference between a dim signal pulse and an accompanying bright reference pulse gives a practical way to implement operators analogous to P_0 and P_1, which yield a guaranteed null result only on the two legitimate signals $|u_1\rangle$ and $|u_0\rangle$, respectively, but not on fake signals (e.g., the vacuum state) that an eavesdropper might substitute. The separation of the signal and reference pulses in time also allows them to be transmitted through the same optical fiber [14], thereby automatically compensating for environmental phase drift in the fiber that would otherwise make such a large interferometer unmanageable.

Since any pair of coherent or incoherent optical signals become significantly nonorthogonal at low intensity, it would seem that almost any source of two kinds of dim light flash, for example, a very attenuated red versus green traffic light, could be used for key distribution without the complications of interferometry. Alice would randomly send red and green flashes of < 1 photon intensity, and Bob would publicly report which flashes he saw, but not their colors, which would constitute the secret key. Because of the low intensity Bob can be confident that a passive Eve standing beside him and watching the same signal source would not see the same subset of flashes, and so would be at least partially ignorant of the key he agrees on with Alice (this partial ignorance can later be amplified to near-total ignorance by hashing techniques similar to that used in steps 9 and 10 of Table I [5,15]) [16].

However, a more intrusive Eve, who stands between Alice and Bob, can thoroughly subvert the scheme by in-

tercepting all Alice's flashes, and resending a flash to Bob only when she succeeds in seeing Alice's flash herself, while simply stopping the others. To compensate for their reduced number, Eve's fabricated flashes must be proportionately brighter, so that Bob's probability of seeing will be the same as before (a cautious Eve would need to fabricate flashes with non-Poissonian photon number statistics, to simulate a Poisson distribution of lesser mean). In terms of the projection operator formalism discussed earlier, the red and green scheme fails because the two signals Alice sends here are not pure states, but statistical mixtures in which the phase of the electric field is random. Therefore any operator P_0 which annihilates all Alice's red flashes will also annihilate the vacuum state, since it may be viewed as a superposition of two red flashes of opposite phase; the same holds for P_1 and green flashes. Eve can thus safely substitute the vacuum state for any flash she fails to detect. By contrast, in the interferometric scheme of Fig. 2, there is no fake signal an eavesdropper can substitute to hide her failure to detect the original signal, and the scheme remains secure.

These considerations may be generalized to conclude that key distribution is possible not only using any two nonorthogonal pure states $|u_0\rangle$ and $|u_1\rangle$, but also any two nonorthogonal mixed states ρ_0 and ρ_1 which span disjoint subspaces of Hilbert space, therefore allowing Bob to find two operators P_0 and P_1 such that P_0 annihilates ρ_1 and P_1 annihilates ρ_0 but no state is annihilated by both operators. The requirement of spanning disjoint subspaces is not present in key distribution schemes using *more than two* mixed states, allowing such schemes (e.g., the scheme [5] which uses four nonorthogonal incoherent states) to be carried out with simple square-law detection of the optical signals, rather than interferometric homodyne detection as used in Fig. 2.

The author thanks Leonard Mandel, Joshua Rothenberg, and Gilles Brassard for helpful discussions. Part of this work was done while the author was visiting California Institute of Technology as a Sherman Fairchild Scholar.

[1] The classic "one-time pad" technique encrypts a message with absolute security as the bitwise exclusive-or (XOR) of the message with a random secret key of the same length. The resulting cryptogram conveys no information to an eavesdropper except for its length, but the original message can be recovered by again XORing the secret key. The method's security depends on never reusing the key. If the same key is used for a second message, an eavesdropper can learn the XOR of the messages by taking the XOR of the cryptograms. For redundant messages such as English, this is often enough to reveal both messages.

[2] M. N. Wegman and J. L. Carter, J. Comput. Syst. Sci. **22**, 265 (1981), show how to generate a key- and message-dependent "authentication tag," analogous to a check sum, whose value cannot be predicted by anyone ignorant of the key, and which thus certifies that the accompanying message has not been altered in transit. Like one-time-pad encryption, authentication uses up key bits and renders them unfit for reuse.

[3] So-called public-key cryptosystems allow users to agree on a secret key solely through public messages, and are widely used for that purpose, but they provide only conditional security: By sufficient computational effort the eavesdropper can always break the system and learn the key.

[4] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), p. 175.

[5] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, J. Cryptology **5**, 3 (1992).

[6] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935); D. Bohm, *Quantum Theory* (Prentice Hall, Englewood Cliffs, NJ, 1951).

[7] A. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[8] J. D. Franson, Phys. Rev. Lett. **62**, 2205 (1989).

[9] M. A. Horne, A. Shimony, and A. Zeilinger, Phys. Rev. Lett. **62**, 2209 (1989).

[10] J. D. Franson (private communication).

[11] A.K. Ekert, J. Rarity, and P. Tapster, in Proceedings of the European Science Foundation Conference, Davos, Switzerland, 1991 (unpublished).

[12] C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).

[13] R. Ghosh and L. Mandel, Phys. Rev. Lett. **59**, 1903 (1987).

[14] M. Shirasaki, A. Haus, and D. Liu Wong, in *Digest of Conference on Lasers and Electro-Optics* (Optical Society of America, Washington, DC, 1987), paper No. THO1, have described a similar interferometer, consisting of two half-interferometers joined by a fiber.

[15] C. H. Bennett, G. Brassard, and J.-M. Robert, SIAM J. Comput. **17**, 210–229 (1988).

[16] U. Maurer, in *Proceedings of the Twenty-Third ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 1991), p. 561, has shown more generally that if Alice, Eve, and Bob listen to the same random source through noisy channels that are even slightly independent, Alice and Bob can derive a secret key by subsequent public discussion.