Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science

6.453 Quantum Optical Communication

Lecture Number 16
Fall 2016

Jeffrey H. Shapiro

Date: Thursday, November 3, 2016

Reading: For quantum key distribution:

- D. Bouwmeester, A. Ekert, and A. Zeilinger, eds. *The Physics of Quantum Information* (Springer, Berlin, 2000) Chap. 2.
- M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2003) Chap.2.

Introduction

In this lecture we will examine the first commercial application of quantum optical communication: quantum cryptography. We'll start with a purely classical case, and show that a one-time pad provides totally secure digital communication. The difficulty of creating a shared one-time pad between remote sites will then lead us into quantum cryptography, or, more properly, quantum key distribution. The two most prominent examples of quantum key distribution will be discussed: the Bennett-Brassard 1984 (BB84) protocol, and the Ekert protocol. The latter makes use of polarization-entangled photons, and hence it gives us the opportunity to discuss another feature of entanglement, namely its exceeding the classical bounds of "hidden-variable" theories.

Secure Communication with a One-Time Pad

Suppose that Alice has a digital message that she wants to send to Bob in a secure manner. Also suppose that Alice knows her transmission is being monitored by an eavesdropper (Eve), such that whatever Alice transmits will be received by *both* Eve and Bob. How can Alice communicate to Bob without Eve's getting the message too? If Alice and Bob have shared in advance a list of statistically independent, identically distributed, completely random bits—a one-time pad—then they can easily communicate in complete security. Let Alice's message be an N-bit string $\{p_n : 1 \leq n \leq N\}$, where each $p_n = 0$ or 1. In cryptography parlance this is the plaintext. We will use $\{k_n : 1 \leq N\}$ to denote the one-time pad that Alice employs to encrypt her

message. Here, $\{k_n\}$ is a set of statistically independent random variables that are each equally likely to be 0 or 1. Alice sends Bob the ciphertext formed by modulo-2 addition of the one-time pad, on a bit-by-bit basis, to the plaintext, viz.,

$$c_n = p_n \oplus k_n, \quad \text{for } 1 \leq n \leq N, \tag{1}$$

is the ciphertext.

Consider the statistics of $\{c_n\}$ where, for simplicity and without loss of generality, we assume that the plaintext is non-random. When $p_n = 0$, we have that $c_n = 0 \oplus k_n = k_n$ is equally likely to be 0 or 1. But, when $p_n = 1$, we have that $c_n = 1 \oplus k_n$, so that $c_n = 1$ if $k_n = 0$ and $c_n = 0$ if $k_n = 1$. In this case c_n is *still* equally likely to be 0 or 1. Moreover, because the $\{k_n\}$ are all statistically independent, so too are the $\{c_n\}$. As a result, when Eve listens in on the communication channel and obtains the ciphertext $\{c_n\}$, she gains *no* information about Alice's message. How then does Bob retrieve the message? It's easy. Bob has the *same* one-time pad that Alice used. Thus, he takes the ciphertext and adds (modulo-2) the one-time pad to decode the ciphertext. This gives him

$$c_n \oplus k_n = (p_n \oplus k_n) \oplus k_n = p_n \oplus (k_n \oplus k_n) = p_n \oplus 0 = p_n, \quad \text{for } 1 \leq n \leq N, \tag{2}$$

because modulo-2 addition is associative (second equality) and the modulo-2 sum of a bit with itself is 0 (third equality).

So, Bob recovers Alice's plaintext and Eve has learned nothing. Why then doesn't the one-time pad solve all secure communication problems? That too is easy to explain. You can only use a one-time pad once, if security is to be maintained. Reusing a one-time pad leads to a security breakdown, although we will not take the time to show that this is so. Alice and Bob thus need one bit of key for every bit they want to communicate. Moreover, Alice and Bob are presumed to be at different locations. Unless they have a secure way to acquire a common one-time pad—say by means of a trusted courier—they would have to rely on a reservoir of key bits that they shared at some earlier time when they were co-located. But that is problematic, as Alice and Bob might *never* have been at the same location. Quantum cryptography, in the guise of quantum key distribution, provides the means for Alice and Bob to accumulate their shared one-time pad in a secure manner even though they are at remote locations.

Bennett-Brassard 1984 Quantum Key Distribution

Consider transmitting an arbitrary single-photon polarization qubit of a single-mode quantum electromagnetic field from Alice to Bob. If Eve intercepts this photon and does polarization analysis on it, we know she cannot perfectly determine the state of that qubit. Likewise, because of the no-cloning theorem, Eve can't make perfect copies of the unknown polarization qubit. These considerations provide Alice and

Alice's Basis	Alice's Bit	Bob's Basis	Bob's Bit
H/V	H	$\pm 45^\circ$	$\pm 45^\circ$ equally likely
H/V	V	$\pm 45^\circ$	$\pm 45^\circ$ equally likely
H/V	H	H/V	H
H/V	V	H/V	V
$\pm 45^\circ$	$+45^\circ$	$\pm 45^\circ$	$+45^\circ$
$\pm 45^\circ$	-45°	$\pm 45^\circ$	-45°
$\pm 45^\circ$	$+45^\circ$	H/V	H/V equally likely
$\pm 45^\circ$	-45°	H/V	H/V equally likely

Table 1: Polarization-qubit measurement results for BB84 QKD with no eavesdropping.

Bob with the means to detect the presence of any eavesdropping as they attempt to create a shared random bit string for use as a one-time pad.

The basic Bennett-Brassard 1984 (BB84) quantum key distribution (QKD) protocol makes use of this sensitivity to eavesdropping as follows. For each bit interval, Alice sends a single-photon polarization qubit that is equally likely to be in any of the following four polarizations: horizontal (H), vertical (V), $+45^\circ$, or -45°. We'll ignore loss, and (for now) assume that Eve is not present. Thus Alice's photon will arrive undisturbed at Bob's location, where he does polarization analysis that he randomly chooses, in an equally likely manner, to be in *either* the H/V basis *or* the $\pm 45^\circ$ basis. What happens? If Alice and Bob chose the same basis, say H/V, and Alice sends a horizontally-polarized photon, then Bob's polarization analysis system will, with probability one, record a click for the detector that heralds the presence of a horizontally-polarized photon. With this same choice of basis for Alice and Bob, if Alice sends a vertically-polarized photon, then Bob will definitely detect that photon as being vertically polarized. Similarly, if Alice and Bob both used the $\pm 45^\circ$ basis, then Bob's measurement will yield the same polarization as the photon that Alice sent. On the other hand, if Alice and Bob have chosen different bases, then very different behavior occurs. For example, if Alice sends an H-polarized photon and Bob uses the $\pm 45^\circ$ basis, he gets a completely random outcome:

$$\Pr(\text{ Bob } = \pm 45^\circ \mid \text{ Alice } = H\,) = |\langle \pm 45^\circ | H \rangle|^2 = 1/2, \tag{3}$$

where $\{|H\rangle, |V\rangle, |+45^\circ\rangle, |-45^\circ\rangle\}$ denote the polarization qubit states and we have used

$$|\pm 45^\circ\rangle = \frac{|H\rangle \pm |V\rangle}{\sqrt{2}}. \tag{4}$$

From similar calculations we can flesh out the entire set of measurement probabilities shown in Table 1.

Let's continue with the BB84 protocol. After a long string of photons have been sent from Alice to Bob, Bob tells Alice which bases he used for each bit. Alice

tells Bob for which bits his basis coincided with hers. This conversation is carried out over a public, classical communication channel that Eve can safely and easily monitor but not modify. However, the basis reconciliation information that Alice and Bob have shared does *not* tell Eve which polarization Alice sent nor does it tell her which polarization Bob received. In other words, knowing that Alice and Bob both used the H/V basis on the tenth bit, doesn't tell Eve whether that bit was H- or V-polarized. She *could* have gotten that information had she intercepted the tenth bit, measured it in the H/V basis and, after her measurement, sent a photon to Bob that was H-polarized if her measurement result was H and sent a V-polarized photon to Bob if her measurement result was V. BUT, Eve cannot know, in advance, that Alice and Bob would use the H/V basis on the tenth photon. If she intercepts that photon and measures it in the H/V basis, but Alice and Bob were both using the $\pm 45^\circ$ basis, there will be a probability 1/2 that Bob's $\pm 45^\circ$-basis measurement on the H- or V-polarized photon that Eve sends him will *disagree* with what Alice sent. In other words, Eve's presence will create some errors on the Alice/Bob transmissions in which they used the same basis. This is the security afforded by quantum mechanics. If Alice and Bob use the public channel to check some of their bit values for cases in which they used the same basis, they will detect some errors if there has been any eavesdropping. If there was no eavesdropping, then Alice and Bob can construct their shared one-time pad by, e.g., making the assignments $H \to 0$, $V \to 1$, $+45^\circ \to 0$, and $-45^\circ \to 1$ for the photons in which they used the same basis.

BB84 QKD is obviously secure when Alice and Bob have perfect equipment, i.e., a single photon source, lossless propagation, and perfect photon counters. Real BB84 is done with attenuated laser sources—producing less than one photon, on average, per pulse—lossy propagation, and imperfect photodetectors. These systems must do all of the following steps after a string of photons have been sent from Alice to Bob.

Step 1: sifting Bob announces to Alice the bit intervals in which he made detections[1] and which basis he used in each of those bit intervals. Alice responds by telling Bob in which of those intervals he used the same basis that she did. Alice and Bob then discard all intervals other than those in which Bob detected a photon using the same basis that Alice did.

Step 2: error detection and correction Because of imperfections in their equipment, Alice and Bob may disagree on the values of some of their sifted bits, even when there is no eavesdropping. Because they must have identical bit strings to use as one-time pads, they use the public channel to detect and correct errors in their sifted bits. This is done by a process of checksum exchanges, i.e., the sifted bits themselves are *not* communicated, as that would make their values immediately available to Eve, who is free to listen on the public channel.

Step 3: privacy amplification The public communication that Alice and Bob have

[1]Because of the loss, Bob may not detect a photon in each bit interval.

carried out in Step 2 reveals *some* information to Eve about their shared key. So, before using the key, Alice and Bob perform a step of privacy amplification. This entails a distillation of the key—a reduction in the number of bits—by a procedure which leaves Alice and Bob with identical, but shorter, random bit strings, while reducing Eve's information about the distilled key to as low a value as is desired.[2]

The Clauser-Horne-Shimony-Holt Inequality

In preparation for our discussion of the Ekert protocol for quantum key distribution, we need to understand the Clauser-Horne-Shimony-Holt (CHSH) inequality, which distinguishes classical physics from quantum mechanics. In particular, the CHSH inequality (which is a particular form of a Bell's inequality) allows us to show that entanglement precludes the notion of a local reality for the polarizations of the photons in a singlet state. (Were there such a local reality, then we could avoid the action at a distance explanation that seems to be needed for the two-mode polarization analysis of the singlet state we treated in a previous lecture.)

Slide 6 shows the setup for the CHSH inequality. Charlie has prepared a pair of photons in the polarization singlet state

$$|\psi^-\rangle = \frac{|H\rangle|V\rangle - |V\rangle|H\rangle}{\sqrt{2}}, \tag{5}$$

and sends one to Alice and the other to Bob. Alice and Bob do ideal polarization analyses such that Alice's output is $a(\theta_A) = 1$ if she detects a photon emerging from a polarizer set for linear polarization at angle θ_A; if she does not get a click on that detector, her output is $a(\theta_A) = -1$.[3] Bob's situation is analogous. His output is $b(\theta_B) = 1$ if he detects a photon emerging from his angle-θ_B polarizer; if he does not get a detector click, then his output is $b(\theta_B) = -1$. This experiment is repeated many times—i.e., Charlie prepares a singlet and sends one photon to Alice and one to Bob, and Alice and Bob do polarization analysis on the photons that they receive—and the results are multiplied and averaged to obtain

$$C(\theta_A, \theta_B) = \langle a(\theta_A) b(\theta_B) \rangle. \tag{6}$$

Moreover, this whole procedure is repeated until $C(\theta_A, \theta_B)$ has been determined for $\theta_A = 0, -\pi/4$ and $\theta_B = 3\pi/8, \pi/8$.

[2] If Eve has eavesdropped on the photons that Alice sent, her presence may create a sufficiently large error rate (seen in Step 2) that Alice and Bob will choose to abort their QKD protocol, *or* they will discover that after privacy amplification they have no key left once they have reduced Eve's information to an acceptably low level.

[3] We are assuming that Alice and Bob have unity quantum efficiency detectors, and that there is no propagation loss from Charlie to them. Thus the absence of a click on Alice's detector means that the photon would have been detected in the orthogonal $\theta_A + \pi/2$ polarization.

Here is where we inject the notion of local reality, by means of a "hidden variable." Suppose that there is some classical parameter, μ, such that if we knew the value of μ then we would know the polarization of each photon in the singlet state. In other words, with knowledge of μ we will have that Alice and Bob's measurement outcomes can be expressed as $a(\theta_A, \mu)$ and $b(\theta_B, \mu)$, respectively, where these are *deterministic* functions of their two arguments. Of course *without* knowledge of μ we know that Alice and Bob's measurement outcomes, $a(\theta_A)$ and $b(\theta_B)$, are random. So, we will take μ to be a classical random variable with some probability density function $p(\mu)$. With this in mind, let us see what constraint local realism (hidden-variable theory) places on the following combination of $C(\theta_A, \theta_B)$ values:

$$S \equiv |C(0, 3\pi/8) + C(-\pi/4, 3\pi/8) + C(-\pi/4, \pi/8) - C(0, \pi/8)|. \tag{7}$$

Substituting in for the $C(\theta_A, \theta_B)$, using the hidden-variable formalism, and rearranging terms then gives us

$$S = \left| \int \mathrm{d}\mu \, \{[a(0,\mu) + a(-\pi/4,\mu)]b(3\pi/8,\mu) + [a(-\pi/4,\mu) - a(0,\mu)]b(\pi/8,\mu)\}p(\mu) \right|. \tag{8}$$

Because $a(\theta_A, \mu)$ must equal either $+1$ or -1, then $a(0,\mu) + a(-\pi/4,\mu)$ must either equal -2, 0, or 2. But, if this term equals -2 or 2, then we must have $a(0,\mu) = a(-\pi/4,\mu)$, whence $a(-\pi/4,\mu) - a(0,\mu) = 0$. Likewise if $a(0,\mu) + a(-\pi/4,\mu) = 0$, then we must have $a(0,\mu) = -a(-\pi/4,\mu)$ and so we see that $a(-\pi/4,\mu) - a(0,\mu)$ has to be -2 or 2. Thus, we can segregate the $\mathrm{d}\mu$ integral into two non-overlapping regions: $\mathcal{M}_1$ where the first bracketed term is non-zero, and $\mathcal{M}_2$ where the second bracketed term is non-zero. Once we do this is separation, the rest of the derivation is straightforward:

$$\begin{aligned}
S &= \left| \int_{\mathcal{M}_1} \mathrm{d}\mu \, [a(0,\mu) + a(-\pi/4,\mu)]b(3\pi/8,\mu)p(\mu) \right. \\
&+ \left. \int_{\mathcal{M}_2} \mathrm{d}\mu \, [a(-\pi/4,\mu) - a(0,\mu)]b(\pi/8,\mu)p(\mu) \right| \qquad (9) \\
&\leq \int_{\mathcal{M}_1} \mathrm{d}\mu \, |[a(0,\mu) + a(-\pi/4,\mu)]b(3\pi/8,\mu)|p(\mu) \\
&+ \int_{\mathcal{M}_2} \mathrm{d}\mu \, |[a(-\pi/4,\mu) - a(0,\mu)]b(\pi/8,\mu)|p(\mu) \qquad (10) \\
&= \int_{\mathcal{M}_1} \mathrm{d}\mu \, 2|b(3\pi/8,\mu)|p(\mu) + \int_{\mathcal{M}_2} \mathrm{d}\mu \, 2|b(\pi/8,\mu)|p(\mu) \qquad (11) \\
&= \int_{\mathcal{M}_1} \mathrm{d}\mu \, 2p(\mu) + \int_{\mathcal{M}_2} \mathrm{d}\mu \, 2p(\mu) = 2\int \mathrm{d}\mu \, p(\mu) = 2. \qquad (12)
\end{aligned}$$

Inequality (12) is called the Clauser-Horne-Shimony-Holt inequality. It states that $S \leq 2$ for any hidden variable (local reality) theory of the dual-polarization analysis experiment. Let's see how S behaves when we treat this experiment quantum mechanically. We will use

$$|\mathbf{i}_k\rangle_k \equiv \cos(\theta_k)|H\rangle_k + \sin(\theta_k)|V\rangle_k \quad \text{and} \quad |\mathbf{i}'_k\rangle_k \equiv \sin(\theta_k)|H\rangle_k - \cos(\theta_k)|V\rangle_k \tag{13}$$

to denote the single-photon, polarization-qubit bases at angles $\{\theta_k : k = A, B\}$ for Alice ($k = A$) and Bob ($k = B$). To get the quantum expression for S, we start with

$$\begin{aligned} C(\theta_A, \theta_B) &= \Pr(a(\theta_A) = 1, b(\theta_B) = 1) + \Pr(a(\theta_A) = -1, b(\theta_B) = -1) \\ &- \Pr(a(\theta_A) = -1, b(\theta_B) = 1) - \Pr(a(\theta_A) = 1, b(\theta_B) = -1), \end{aligned} \tag{14}$$

and use our Axiom 3 to calculate the probabilities on the right-hand side, obtaining

$$\begin{aligned} C(\theta_A, \theta_B) &= |\langle\psi^-|(|\mathbf{i}_A\rangle_A \otimes |\mathbf{i}_B\rangle_B)|^2 + |\langle\psi^-|(|\mathbf{i}'_A\rangle_A \otimes |\mathbf{i}'_B\rangle_B)|^2 \\ &- |\langle\psi^-|(|\mathbf{i}'_A\rangle_A \otimes |\mathbf{i}_B\rangle_B)|^2 - |\langle\psi^-|(|\mathbf{i}_A\rangle_A \otimes |\mathbf{i}'_B\rangle_B)|^2. \end{aligned} \tag{15}$$

Performing the inner product evaluations leaves us with

$$\begin{aligned} C(\theta_A, \theta_B) &= \left|\frac{\cos(\theta_A)\sin(\theta_B) - \sin(\theta_A)\cos(\theta_B)}{\sqrt{2}}\right|^2 \\ &+ \left|\frac{-\sin(\theta_A)\cos(\theta_B) + \cos(\theta_A)\sin(\theta_B)}{\sqrt{2}}\right|^2 \\ &- \left|\frac{\sin(\theta_A)\sin(\theta_B) + \cos(\theta_A)\cos(\theta_B)}{\sqrt{2}}\right|^2 \\ &- \left|\frac{-\cos(\theta_A)\cos(\theta_B) - \sin(\theta_A)\sin(\theta_B)}{\sqrt{2}}\right|^2 \end{aligned} \tag{16}$$

$$= \sin^2(\theta_A - \theta_B) - \cos^2(\theta_A - \theta_B) = -\cos[2(\theta_A - \theta_B)]. \tag{17}$$

From this point it is trivial to get the quantum mechanical result for the CHSH parameter S:

$$S \equiv |C(0, 3\pi/8) + C(-\pi/4, 3\pi/8) + C(-\pi/4, \pi/8) - C(0, \pi/8)| \tag{18}$$

$$= |\cos(3\pi/4) + \cos(5\pi/4) + \cos(3\pi/4) - \cos(\pi/4)| \tag{19}$$

$$= \left|-\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\right| = 2\sqrt{2} > 2. \tag{20}$$

Experiments with singlet states produced by spontaneous parametric downconversion *do* give S parameter values in excess of 2 and approaching $2\sqrt{2}$. However, the sources used in these experiments produce states

$$\frac{1}{N+1}\,|0\rangle|0\rangle + \sqrt{\frac{N}{(N+1)^3}}\,(|H\rangle|V\rangle - |V\rangle|H\rangle) + \text{higher order terms}, \quad \text{with } N \ll 1,$$

where $|0\rangle|0\rangle$ denotes the two-mode vacuum state. Also, these experiments are done with detectors that have dark counts and whose quantum efficiencies are less than 1. Thus, a slightly more complicated version of the CHSH inequality must be used, to account for the fact that the absence of a click for polarization θ_k does *not* automatically mean that there was a photon present in the orthogonal polarization. The result of these and other imperfections in the actual experiments is that it is only last year that a *loophole-free* Bell-inequality measurements, i.e., ones that fully refutes hidden-variable theory, were reported.[4]

Ekert Protocol Quantum Key Distribution

Armed with our understanding of the CHSH inequality, we are ready to look into Ekert protocol (entangled state) QKD. The setup is shown in Slide 10. A dual parametric amplifier source is used to create a stream of polarization-entangled photon pairs that are in singlet states. One photon from each pair is sent to Alice with the other being sent to Bob. To ensure that there has been no eavesdropping, Alice and Bob first measure the CHSH S parameter. If there has been any eavesdropping, the entanglement between Alice and Bob's photon will have been destroyed—e.g.,by Eve's using a measurement and resend strategy—and so Alice and Bob will *not* get an S value in excess of 2. Once Alice and Bob have determined that there has been no eavesdropping, they can revert to the setup in which they have detectors for H/V and $\pm 45^\circ$ polarizations. When a single photon arrives at Alice's station it will be detected—in a completely random fashion—at one of her four detectors. The same thing occurs at Bob's station. However, because their two photons come from a singlet, if Alice and Bob get their detections in the *same* basis, then each will know what polarization the other has received, e.g., if Alice's H detector clicked and Bob got a click in his H/V basis, then his V detector had to be the one that clicked, etc. As a result, Alice and Bob can use the rest of the protocol stack that BB84 employs. Bob tells Alice which basis gave him a detection in each bit interval. Alice tells Bob in which intervals she got a detection in the same basis. At this point they discard the bit intervals in which their bases did not agree, and then proceed to the error

[4]See B. Hensen, *et al.*, "Loophole-free Bell inequality violation using electron spins separated by 1.3 km," Nature **526,** 682–686 (2015); L. K. Shalm, *et al.*, "Strong loophole-free test of local realism," Phys. Rev. Lett. **115,** 250402 (2015); and M. Giustina, *et al.*, Phys. Rev. Lett. **115,** 250401 (2015).

detection, correction, and privacy amplification steps of the QKD protocol.[5]

The Road Ahead

We've reached the end of the road, insofar as simple one- or two-mode (harmonic oscillator) descriptions of quantum optical communication is concerned. Next lecture we will describe how to quantize the full multi-mode electromagnetic field.

[5]Strictly speaking, what we have just presented is a hybrid of the original Ekert protocol and BB84. In Ekert's original proposal, Alice and Bob randomly choose their polarization analysis angles from the sets $\theta_A \in \{0, \pi/8, \pi/4\}$ and $\theta_B \in \{\pi/8, \pi/4, 3\pi/8\}$, respectively. They communicate over the public channel to determine the photons for which they used the same polarization angles after which they segregate their data into two groups: one in which they used different angles and one in which they employed the same angle. Then, by disclosing the polarization angles *and* the measurement results for the first group they are able to evaluate the CHSH S parameter. Once assured that there has been no eavesdropping, they can develop a shared string of random bits from the second group, because those measurements—in the absence of eavesdropping and with perfect equipment—are always anti-correlated.

6.453 Quantum Optical Communication

Fall 2016