Assignment
Quantum Cyber Security

Due: 12:00 Thursday 28 March, 2024

This assignment counts for **25% of the course** and you must answer **all three** questions. The weights of each question and sub-question are given (number of marks), but note that this is **not** indicative of how difficult the corresponding sub-question is. Note also that notation is set individually in each problem, and the same letters may have different meanings in each problem.

Important message:

Please remember the good scholarly practice requirements of the University regarding work for credit. You can find guidance at the School page https://web.inf.ed.ac.uk/infweb/admin/policies/academic-misconduct. This page also has links to the relevant University pages.

1. In your submission please include the steps that lead to your answers.

   (a) Evaluate the binary entropy $h(p)$ for Bernoulli processes with $p = 1/4$ and $p = 1/2$. [3 marks]

   Solution: The binary entropy function is defined as

   $$h(p) = -p \log_2 p - (1-p) \log_2 (1-p).$$

   Therefore, $h(1/4) = 2 - \frac{3}{4} \log_2 3 \approx 0.811$ and $h(1/2) = 1$.

   (b) Alice sends Bob a quantum state ρ_j with probability p_j, where $j \in \{1, \ldots, n\}$, so that Bob has the mixed state $\rho = \sum_{j=1}^{n} p_j \rho_j$. Use the Holevo bound to show that the maximum amount of information transmitted by N qubits is N bits. How much information can be transmitted if the states are instead composed of N qu*trits*, where the state space of a qutrit is defined as a three-dimensional complex Hilbert space? [3 marks]

   Solution: The Holevo bound states that the maximum information accessible to Bob is bounded by

   $$I_{\text{acc}} \leq S(\rho) - \sum_{j=1}^{n} p_j S(\rho_j).$$

   Since $S(\rho_j) \geq 0$ for all j, we have in particular that $I_{\text{acc}} \leq S(\rho)$. The dimension of the state space for the case of N qubits is by definition 2^N, and thus $S(\rho) \leq \log_2(2^N) = N$. Combining the inequalities, we finally obtain $I_{\text{acc}} \leq N$ bits of information.

In the case of N qutrits, the state space has dimension 3^N by definition, and thus we have the bound

$$I_{\text{acc}} \leq S(\rho) \leq \log_2(3^N) = N \log_2 3 \approx 1.585N$$

bits of information. Equivalently, N qutrits can transmit at most N trits of information.

(c) The phase-flip channel, which does nothing with probability p and flips the phase of $|1\rangle$ to $-|1\rangle$ with probability $1-p$, has Kraus operators

$$E_0 = \sqrt{p}I, \quad E_1 = \sqrt{1-p}Z,$$

where Z is the Pauli operator defined by $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$. Evaluate the action of the phase-flip channel with $p = 1/2$ on the state $\rho = |-\rangle\langle -|$, where $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$.

[2 marks]

Solution: The action of a channel with Kraus operators $\{E_j\}$ on a state ρ is

$$\rho \mapsto \sum_j E_j \rho E_j^\dagger.$$

In our case, $E_0^\dagger = E_0$ and $E_1^\dagger = E_1$, $p = 1/2$, and $\rho = |-\rangle\langle -|$. Thus, by noting $Z|-\rangle = |+\rangle$, applying the phase-flip channel gives

$$\begin{aligned}\rho \mapsto p\rho + (1-p)Z\rho Z &= \frac{1}{2}(|-\rangle\langle -| + Z|-\rangle\langle -|Z) \\ &= \frac{1}{2}(|-\rangle\langle -| + |+\rangle\langle +|) \\ &= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}I.\end{aligned}$$

That is, the channel takes $|-\rangle$ to the maximally mixed state.

(d) Charlie is given one of two possible states

$$\rho = |0\rangle\langle 0| \quad \text{or} \quad \sigma = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 1|).$$

Evaluate the fidelity $F(\rho, \sigma)$ of the two states. Using the fidelity, what can we say about the maximum probability with which Charlie can correctly identify the state?

[2 marks]

Solution: Let us first recognise that there are two alternative definitions of fidelity which may be used to obtain the correct solution, provided the choice of definition is consistent. The first of these (here denoted by F') is defined by

$$F'(\rho, \sigma) = \left(\operatorname{tr}\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}\right)^2,$$

and the second is simply $F(\rho, \sigma) = \sqrt{F'(\rho, \sigma)}$. Here, we will choose to adopt the latter definition, as is done in the lecture slides. Note: There was a typo contained in the lectures which stated $F(\psi, \varphi) = |\langle\psi|\varphi\rangle|^2$, while in fact $F(\psi, \varphi) = |\langle\psi|\varphi\rangle|$ is the correct simplified expression for pure states under this convention.

Since ρ is the density matrix for the pure state $|0\rangle$, the fidelity can be expressed in the simplified form

$$F(\rho, \sigma) = \sqrt{\langle 0|\sigma|0\rangle} = \frac{1}{\sqrt{2}}.$$

The maximum probability with which Charlie can identify the correct state is given by

$$p_{\text{guess}}^{\max} = \frac{1}{2}(1 + D(\rho, \sigma)),$$

where $D(\rho, \sigma)$ is the trace distance between ρ and σ. The trace distance is bounded above in terms of the fidelity as

$$D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2} = \sqrt{1 - \frac{1}{2}} = \frac{1}{\sqrt{2}},$$

and so $p_{\text{guess}}^{\max} \leq (2 + \sqrt{2})/4 \approx 0.854$. In fact, since $\sigma = |+\rangle\langle+|$ is also a pure state, the upper bound on the trace distance is in fact an equality, leading to $p_{\text{guess}}^{\max} = (2+\sqrt{2})/4$.

2. **Quantum Coin Flipping:**
Recall the quantum coin flipping protocol of Ambainis mentioned in the lecture notes where the following four qutrit states have been used:

$$|\phi_{a,x}\rangle = \begin{cases} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), & a = 0, x = 0 \\ \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), & a = 0, x = 1 \\ \frac{1}{\sqrt{2}}(|0\rangle + |2\rangle), & a = 1, x = 0 \\ \frac{1}{\sqrt{2}}(|0\rangle - |2\rangle), & a = 1, x = 1 \end{cases}$$

where $|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $|1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $|2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

(a) We want to first look at Bob's cheating strategy. Compute the mixed states ρ_0 and ρ_1 corresponding to the mixture of states Bob receives from Alice for the choice of random bit a being 0 and 1 respectively.

[3 marks]

Solution: If $a = 0$, Alice sends a mixed state that is equal to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ with probability 1/2 and $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ with probability 1/2. So the density matrix associated with this is: $\rho_0 = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$ if you write everything in the qutrit basis ($\{|0\rangle, |1\rangle, |2\rangle\}$ basis). If $b = 1$, she sends a mixed state that is equal to $\frac{1}{\sqrt{2}}(|0\rangle + |2\rangle)$ with probability 1/2 and $\frac{1}{\sqrt{2}}(|0\rangle - |2\rangle)$ with probability 1/2, which will result in $\rho_1 = \frac{1}{2}(|0\rangle\langle 0| + |2\rangle\langle 2|)$. The matrix forms of these two mixed states are:

$$\rho_0 = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \rho_1 = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

(b) Write down the matrix form of $\rho_0 - \rho_1$. Then calculate the trace norm of $\| \rho_0 - \rho_1 \|_{tr}$.
Note: The trace distance is related to the trace norm in this way:
$T(\rho_0, \rho_1) = \frac{1}{2}\|\rho_0 - \rho_1\|_{tr}$.

[2 marks]

Solution: Let's first write down the matrix $\rho_0 - \rho_1$, based on the last section:

$$\rho_0 - \rho_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

As we have seen in the lecture, one way to calculate the trace norm is to find the eigenvalues of the matrix $\rho_0 - \rho_1$ and then use the formula $\| A \|_{tr} = \sum_i |\lambda_i|$ to get the trace norm. Since the matrix $\rho_0 - \rho_1$ is diagonal, calculating the eigenvalues is easy ($\lambda_0 = 0, \lambda_1 = 1/2, \lambda_2 = -1/2$). So we have:

$$\| \rho_0 - \rho_1 \|_{tr} = 0 + 1/2 + 1/2 = 1$$

(c) Now use the Holevo-Helstrom bound for the maximal probability of distinguishing two mixed states, i.e. the following equation:

$$P_{opt}^{dist} = \frac{1}{2} + \frac{1}{4} \| \rho_0 - \rho_1 \|_{tr},$$

to obtain the maximum cheating probability of Bob for this protocol and hence determine the minimum bias for a dishonest Bob.

[1 mark]

Solution: We just substitute the trace norm from the previous section in the Holevo-Helstrom bound and get the maximal distinguishing probability to be $P_{opt}^{dist} = 3/4$. So the probability that bias on the protocol against dishonest Bob will be 0.25.

(d) In this part, we will look at a weak coin-flipping protocol. Let's assume that outcome 0 means that Bob wins, and outcome 1 is a win for Alice. The protocol is as follows:

- Step 1: Alice prepares a pair of systems in an entangled state $|\psi_{AB}\rangle \in \mathcal{H}^A \otimes \mathcal{H}^B$, being $|\psi_{AB}\rangle = \frac{\sqrt{3}}{2} |00\rangle + \frac{1}{2} |11\rangle$ and sends subsystem B to Bob.
- Step 2: Bob performs a 2-outcome POVM measurement $\{E_0, E_1\}$ on the qubit he received (System B), and sends a classical bit b that is the outcome bit to Alice. Let $E_0 = \frac{2}{3} |0\rangle \langle 0|$. (You can find what's E_1 using properties of POVMs.)
- Step 3: If $b = 0$ then Bob sends his system (B) back to Alice, if $b = 1$ then Alice sends her system (A) to Bob. The party that receives the system then performs the projective measurements $\{|\psi_b\rangle \langle \psi_b| , I - |\psi_b\rangle \langle \psi_b|\}$, where the $|\psi_b\rangle$ is defined as follows:

$$|\psi_b\rangle = \frac{I \otimes \sqrt{E_b} |\psi_{AB}\rangle}{\sqrt{\langle \psi_{AB}| I \otimes E_b |\psi_{AB}\rangle}}$$

I) Write down the reduced density matrix $\rho_B = Tr_A[|\psi_{AB}\rangle \langle \psi_{AB}|]$ that is being sent from Alice to Bob. Also, write down the state $|\psi_b\rangle$ for both cases where $b = 0$ and $b = 1$.

II) Following all the steps of the protocol, explain why this protocol is correct (achieves weak coin flipping) if both Alice and Bob are honest and they follow the protocol.

III) You may have noticed that in the above protocol, there is an extra measurement that allows Alice and Bob to catch each other cheating! Explain how they can detect each other's cheating by describing the four possible outcomes of the protocol. (You can explain your answer by trying to give an attack where either Alice or Bob are trying to cheat.)

[6 marks]

Note: Each of the above subquestions counts for 2 marks.

Solution: I) Let's begin with writing down ρ_{AB}:

$$\rho_{AB} = |\psi\rangle_{AB} \langle\psi|_{AB} = \frac{3}{4} |00\rangle \langle 00| + \frac{1}{4} |11\rangle \langle 11| + \frac{\sqrt{3}}{4}[|00\rangle \langle 11| + |11\rangle \langle 00|]$$

Now we take the partial trace of the first subsystem (A), which will give us the following mixed state:

$$\rho_B = Tr_A[|\psi_{AB}\rangle \langle\psi_{AB}|] = \frac{3}{4} |0\rangle \langle 0| + \frac{1}{4} |1\rangle \langle 1|$$

We can calculate the $|\psi_b\rangle$ directly from the given formula in Step 3. Let's first look at the case $b = 0$. We have:

$$|\psi_0\rangle = \frac{I \otimes \sqrt{E_0} |\psi_{AB}\rangle}{\sqrt{\langle\psi_{AB}| I \otimes E_0 |\psi_{AB}\rangle}}$$

Given that $E_0 = 2/3 |0\rangle \langle 0|$, then $\sqrt{(E_0)} = \sqrt{(2/3)} |0\rangle \langle 0|$, and we have:

$$|\psi_0\rangle = \frac{I \otimes \sqrt{2/3} |0\rangle \langle 0| (\frac{\sqrt{3}}{2} |00\rangle + \frac{1}{2} |11\rangle)}{\sqrt{(\frac{\sqrt{3}}{2} \langle 00| + \frac{1}{2} \langle 11|) I \otimes 2/3 |0\rangle \langle 0| (\frac{\sqrt{3}}{2} |00\rangle + \frac{1}{2} |11\rangle)}} = \frac{\frac{1}{\sqrt{2}} |00\rangle}{\frac{1}{\sqrt{2}}} = |00\rangle$$

Note that we act the operator $I \otimes \sqrt{2/3} |0\rangle \langle 0|$ from the right on the 2-qubit state and on the first qubit identity doesn't do anything, on the second qubit we project into the zero state.
Similarly, for $|\psi_1\rangle$ where the $E_1 = I - E_0 = \frac{1}{3} |0\rangle \langle 0| + |1\rangle \langle 1|$ is applied, in the enumerator we get the non-normalised state $\frac{1}{2}(|00\rangle + |11\rangle)$ and the denominator gives correct normalisation factor. Hence we get the state:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

There is another way to obtain these states, and that's by following the protocol and applying the measurements, which we discuss in the next section.

II) Let's see what will happen in this protocol when both Alice and Bob are honest. If Alice is honest, she is actually sending the subsystem B of the states that they both expect Alice to prepare, in our case the $|\psi_{AB}\rangle$ specified in Step 1, In Step 2, Bob will have the state ρ_B that we calculated in the first section of this question. Now if Bob is honest, he will perform the specified POVM, without trying to reveal Alice's side of the state through any other POVMs (which would be his general cheating strategy if he was dishonest). Let's say Bob obtains measurement result 0, then by announcing the correct measurement result and sending back his post-measurement state, Alice will have the 2-qubit state $|\psi_0\rangle$ and will perform the projective measurements $\{|\psi_b\rangle \langle\psi_b|, I - |\psi_b\rangle \langle\psi_b|\}$, and get the outcome 0 deterministically. Thus they will agree on the random bit 0, which is the outcome of Bob's measurement outcome (The same will happen for outcome 1). The point is that Alice will always get a deterministic outcome of the state she receives is the one that is expected by the protocol. Also note that $Tr[\rho_B E_0] = 1/2$, which means the measurement outcome is a random equal probability coin. So the protocol is correct.

III) This protocol has an interesting property and that's cheat sensitivity, which means that if one of the parties tries to cheat, they will be fought by the other party with some probability. Let's see all the four possible outcomes of the protocol:

1. $b = 0$, Alice finds $|\psi_0\rangle\langle\psi_0|$; Bob wins.
2. $b = 0$, Alice finds $I - |\psi_0\rangle\langle\psi_0|$; Alice catches Bob cheating.
3. $b = 1$, Bob finds $|\psi_1\rangle\langle\psi_1|$; Alice wins.
4. $b = 1$, Bob finds $I - |\psi_1\rangle\langle\psi_1|$; Bob catches Alice cheating.

So the second measurement acts as a check to ensure that the party who has the 2-qubit state at the end of the protocol, has the states as expected. And when the state deviates from the $|\psi_b\rangle$, the second measurement catches it, with some probability. Also, note that while this protocol is sufficient for weak coin-flipping, it is not a good strong coin-flipping protocol, because Bob can always choose to lose by simply announcing b = 1, and hence bias the result to one side. Let's say Bob wants to cheat when he receives outcome 1 (which means he will lose) and wants to announce 0 (claiming he won). But outcome 0 means that he has to send his state to Alice, which means Alice will perform the projection $|\psi_0\rangle\langle\psi_0|$ (and the complementary measurement) and since the state she will receive is not $|\psi_0\rangle$, she will detect the cheating with some probability. Bob can of course deviate from the state that he has left with in order to pass Alice's check, but he cannot perfectly retrieve the correct state if he wants to bias the outcome. You can give any example of such attacks or similar attacks by Alice when they cheat. As long as you can show that the catching probability is non-zero, that will work!
Also as you can guess, there is a tradeoff between the maximum winning probability and the maximum probability of not getting caught while cheating. Another interesting point about this protocol is that it is not symmetric, meaning that Alice and Bob have different maximum success probabilities. Alice's maximum success probability in this case is 3/8 and Bob's is 3/4. Although it's not super easy to prove these maximum bounds, so you don't need to do that in your solution.

3. Consider the following two functions:

$$A(x,i,j) = (-1)^{x\oplus i}(-1)^{x\cdot j} \text{ and } B(y,i,j) = (-1)^{y\oplus i}(-1)^{y\cdot j}\,.$$

All inputs x, y, i, j are single (classical) bits. Assume that these deterministic functions, correspond to the outcomes of Alice (the $A(x,i,j)$) and Bob (the $B(y,i,j)$) at some QKD protocol. The i, j occur with same probability, and are unknown to Alice and Bob.

(a) Show that the expectation values and the correlations that they obtain are the same as for the BBM92 QKD protocol. Recall that in that protocol, the expectation value of each observable A, B was zero (equal probability for each outcome), the expectation value of the correlator when measuring in the same basis was one (perfect correlation), while the expectation value of the correlator when measuring in a different basis is zero (completely uncorrelated).
Note: to compute the expectation value of an observable you need to sum over the hidden variables i, j, while to compute the correlator you need to multiply the two observables and then sum over the hidden variables i, j.

[2 marks]

Solution: Need to compute:

$$\sum_{i,j} A(x,i,j) = \sum_i (-1)^i \sum_j (-1)^x (-1)^{x\cdot j} = 0 \times \sum_j (-1)^x (-1)^{x\cdot j} = 0$$

$$\sum_{i,j} B(x,i,j) = \sum_i (-1)^i \sum_j (-1)^y (-1)^{y\cdot j} = 0$$

$$\sum_{i,j} A(x,i,j)B(y,i,j) = \sum_{i,j} (-1)^{x\oplus y}(-1)^{(x\oplus y)\cdot j}$$

If $x \oplus y = 0$ (same basis)

$$\sum_{i,j} 1 = 4$$

and given that each of the four (i,j) occurs with same probability this is just one.
If $x \oplus y = 1$ (different basis)

$$\sum_{i,j} (-1)(-1)^j = \sum_j (-1)^j \sum_i (-1) = 0$$

(b) Explain why the result of the previous question means that the BBM92 protocol is not device independent.

[1 mark]

Solution: The above functions are deterministic. Eve could toss two coins to obtain randomly a value for i, j and then send classical information to Alice and Bob. They would recover the same expectation values as if they used the BBM92 protocol, but in this case, Eve knows their exact outcomes and thus can cheat perfectly.

This of course does not make BB92 insecure, since in that protocol Alice and Bob know exactly what measurements they perform (and are not the ones that Eve wants them). This only demonstrates that the protocol would not be secure if Alice and Bob had black-box access to their quantum system.