Work #9:  Nov. 27 — Dec. 6

12-hour biweek

Obligatory problems are marked with [**]
1. [Basic Adversary Method.]

(a) [**] Prove the Basic Adversary Method Theorem (generalizing the Super-Basic Adversary Method Theorem) stated towards the end of the Lecture 20 video. Of course, you should mimic the proof of the Super-Basic Adversary Method Theorem.

(b) Use that theorem to show a quantum query lower bound of $\geq \sqrt{N/k}$ for the following promise-decision problem (assuming $1 \leq k \leq N/2$): Output “yes” if the input string $w \in \{0, 1\}^N$ has at least $k$ 1’s; output “no” if it is the all-0’s string.
2. [Product probability spaces.]

(a) Let \( p \in \mathbb{R}^d \) be a probability distribution on \( [d] = \{1,2,\ldots,d\} \). Let \( q \in \mathbb{R}^e \) be a probability distribution on \( [e] = \{1,2,\ldots,e\} \). Prove that the Kronecker product \( p \otimes q \) (which is a vector naturally indexed by the set \([d] \times [e]\)) is the associated “product probability distribution” on \( [d] \times [e] = \{(i,j) : 1 \leq i \leq d, 1 \leq j \leq e\} \); i.e., it’s the distribution gotten by drawing \( i \) from \( p \) and \( j \) from \( q \) independently.

(b) \([**]\) Let \((p_1,|\psi_1\rangle),\ldots,(p_m,|\psi_m\rangle)\) be the mixed state of a \( d\)-dimensional particle (meaning we have probability \( p_i \) of pure state \(|\psi_i\rangle \in \mathbb{C}^d, i = 1\ldots m\)). Similarly, let \((q_1,|\phi_1\rangle),\ldots,(q_n,|\phi_n\rangle)\) be the mixed state of an \( e\)-dimensional particle. Write \( \rho \in \mathbb{C}^{d \times d} \) for the density matrix of the first mixed state and \( \sigma \in \mathbb{C}^{e \times e} \) for the density matrix of the second. Suppose the particles were created completely separately and independently, but we now decide to view them as a joint \( de\)-dimensional state. Recalling the rules of how to do this for pure states, show that the resulting \( de\)-dimensional mixed state has density matrix \( \rho \otimes \sigma \), the Kronecker product of \( \rho \) and \( \sigma \).
3. **[Positive semidefinite matrices.]** A Hermitian matrix \( M \in \mathbb{C}^{d\times d} \) is said to be *positive*, or *positive semidefinite* (denoted \( M \geq 0 \) or \( M \succeq 0 \)) if \( \langle u|M|u \rangle \geq 0 \) for all vectors \( |u\rangle \in \mathbb{C}^d \).

(a) Prove that \( M \geq 0 \) if and only if \( \langle u|M|u \rangle \geq 0 \) holds for all *unit* vectors \( |u\rangle \in \mathbb{C}^d \).

(b) Let \( M \in \mathbb{C}^{d\times d} \) be a diagonal matrix (meaning all off-diagonal entries are 0). Verify that \( M \) is Hermitian if and only if all its diagonal entries are real. In this case, prove that \( M \geq 0 \) if and only if each of its diagonal entries is nonnegative.

(c) Let \( A \in \mathbb{C}^{k\times d} \) be any matrix (possibly rectangular). First show that \( A^\dagger A \) is Hermitian; then show that \( A^\dagger A \geq 0 \).

(d) Let \( R,X \in \mathbb{C}^{d\times d} \) be positive semidefinite matrices. Prove that \( \langle R,X \rangle \geq 0 \). (See Equation (1) if you forget the definition of \( \langle R,X \rangle \).) You may use the fact that every Hermitian matrix \( M \) can be represented as \( M = \sum_{i=1}^{d} \lambda_i |\psi_i\rangle\langle \psi_i| \) for some real \( \lambda_1, \ldots, \lambda_d \) and some orthonormal basis \( |\psi_1\rangle, \ldots, |\psi_d\rangle \).
4. [The basics of quantum random variables.] Let \( \rho \in \mathbb{C}^{d\times d} \) be a density matrix. Recall that for an observable (i.e., Hermitian matrix) \( X \in \mathbb{C}^{d\times d} \), we define

\[
\mathbf{E}_\rho[X] = \langle \rho, X \rangle = \text{tr}\left(\rho^\dagger X\right) = \text{tr}(\rho X) = \sum_{i,j=1}^{d} \rho_{ij} X_{ij}.
\] (1)

In this problem, we will extend the above notation to allow for a non-Hermitian matrix \( X \). This is not “physically meaningful” (since there is no measurement instrument corresponding to a non-Hermitian matrix \( X \)), but it will be mathematically convenient to let us reason about observables.

(a) [**] Prove that \( \mathbf{E}_\rho[1] = 1 \), where \( 1 \) denotes the \( d \times d \) identity matrix.

(b) Prove that \( \mathbf{E}_\rho[X^\dagger] = \mathbf{E}_\rho[X]^* \).

(c) [**] Let \( X, Y \in \mathbb{C}^{d\times d} \) be Hermitian and let \( \alpha, \beta \in \mathbb{C} \). Prove “linearity of expectation”:

\[
\mathbf{E}_\rho[\alpha X + \beta Y] = \alpha \mathbf{E}_\rho[X] + \beta \mathbf{E}_\rho[Y].
\]

Also, show that \( \alpha X + \beta Y \) is Hermitian if \( \alpha, \beta \in \mathbb{R} \) (otherwise, we can’t be sure).

(d) [**] Prove that \( \mathbf{E}_\rho[A^\dagger A] \geq 0 \) for any matrix \( A \in \mathbb{C}^{k\times d} \). (You may use Problem 3.)

(e) [**] Let \( \sigma \in \mathbb{C}^{d\times d} \). Referring to Problem 2, prove that \( \mathbf{E}_{\rho \otimes \sigma}[X \otimes Y] = \mathbf{E}_\rho[X] \mathbf{E}_\sigma[Y] \).

(This generalizes the classical probability fact that if \( x \) and \( y \) are independent random variables then \( \mathbf{E}[xy] = \mathbf{E}[x] \mathbf{E}[y] \).

(f) [**] Let \( X, Y \in \mathbb{C}^{d\times d} \), not necessarily Hermitian. Define their covariance with respect to \( \rho \) to be

\[
\text{Cov}_\rho[X, Y] = \mathbf{E}_\rho[(X - \mu_X \mathbb{1})(Y - \mu_Y)^\dagger],
\]

where \( \mu_X = \mathbf{E}_\rho[X], \mu_Y = \mathbf{E}_\rho[Y] \). Prove that \( \text{Cov}_\rho[X, Y] = \mathbf{E}_\rho[X^\dagger Y] - \mu_X^* \mu_Y \).

(g) [**] Prove that covariance is “translation-invariant” in each argument, meaning \( \text{Cov}_\rho[X + \alpha \mathbb{1}, Y + \beta \mathbb{1}] = \text{Cov}_\rho[X, Y] \) for all \( \alpha, \beta \in \mathbb{C} \). Prove also that \( \text{Cov}[\alpha X, \beta Y] = \alpha^* \beta \text{Cov}[X, Y] \).

(h) [**] Let \( X \in \mathbb{C}^{d\times d} \), not necessarily Hermitian. Define the variance of \( X \) with respect to \( \rho \) to be

\[
\text{Var}_\rho[X] = \text{Cov}_\rho[X, X].
\]

Show that \( \text{Var}_\rho[X] \geq 0 \) always, that \( \text{Var}_\rho[X] \) is translation-invariant, and that \( \text{Var}_\rho[\alpha X] = |\alpha|^2 \text{Var}_\rho[X] \).

(i) We wish to prove the quantum Cauchy–Schwarz inequality: For \( X, Y \in \mathbb{C}^{d\times d} \),

\[
|\text{Cov}_\rho[X, Y]|^2 \leq \text{Var}_\rho[X] \text{Var}_\rho[Y].
\] (2)

It’s a little annoying to handle the cases when \( \text{Var}_\rho[X] = 0 \) or \( \text{Var}_\rho[Y] = 0 \), so let’s assume we don’t need to worry about these cases. Otherwise, show that in attempting to prove the above, we may assume without loss of generality that \( \text{Var}_\rho[X] = \text{Var}_\rho[Y] = 1 \) and that \( \text{Cov}_\rho[X, Y] \) is a nonnegative real. (Hint: consider multiplying \( X \) and \( Y \) by scalars.)

(j) Show that it also suffices to assume \( \mathbf{E}_\rho[X] = \mathbf{E}_\rho[Y] = 0 \). (Hint: consider subtracting scalar multiples of the identity.)

(k) Thus it remains to show \( \text{Cov}_\rho[X, Y] \leq 1 \) assuming \( \text{Var}_\rho[X] = \text{Var}_\rho[Y] = 1, \text{Cov}_\rho[X, Y] \in \mathbb{R}^{\geq 0} \), and \( \mathbf{E}_\rho[X] = \mathbf{E}_\rho[Y] = 0 \). Prove this.
5. **[The Uncertainty Principle.]** Let $X, Y \in \mathbb{C}^{d \times d}$ be observables; i.e., Hermitian matrices.

(a) [**] Prove that $X^2$ and $Y^2$ are Hermitian.

(b) [**] Prove that $XY$ is Hermitian if and only if $X$ and $Y$ commute (i.e., $XY = YX$).

(c) [**] Let $[X, Y]$ denote $XY + YX$ (this is nonstandard notation). Prove that $\frac{1}{2}[X, Y]$ is Hermitian. (This matrix is the “symmetrization” of $XY$, or perhaps “Hermitianization”.)

(d) [**] Let $[X, Y]$ denote the matrix $XY - YX$, called the “commutator” of $X$ and $Y$ because it’s 0 if and only if $X$ and $Y$ commute (this is standard notation). Prove that $\frac{1}{2i}[X, Y]$ is Hermitian.

(e) [**] Prove that $XY = \frac{1}{2}[X, Y] + i \cdot \frac{1}{2i}[X, Y]$.

(f) In 1927, Werner Heisenberg stated his famous Uncertainty Principle for two particular observables of a quantum particle, its “position” and “momentum”. In 1928, Earle Kennard properly mathematically proved Heisenberg’s Uncertainty Principle. In 1929, Bob Robertson generalized the Uncertainty Principle to a statement about any two observables. Specifically, he proved the following:

$$\sigma_{\mu}[X] \cdot \sigma_{\mu}[Y] \geq \left| \mathbb{E}_{\rho} \left[ \frac{1}{2i}[X, Y] \right] \right|^2,$$

where $\sigma_{\mu}[X] = \sqrt{\text{Var}_\rho[X]}$ is the standard deviation of the observable $X$ (and similarly for $\sigma_{\mu}[Y]$). Here $\text{Var}_\rho[X]$ is as defined in Problem 4h.

Show that if we want to establish (3), we can reduce to the case that $\mathbb{E}_{\rho}[X] = \mathbb{E}_{\rho}[Y] = 0$. (Hint: use Problem 4h.)

(g) [**] Having made this reduction, prove the Uncertainty Principle (3). (Hint: use the Cauchy–Schwarz inequality (2) and the decomposition from Problem (5e).)
6. **[The SWAP test.]** We’ve previously discussed the SWAP gate operating on two qubits, but it also makes sense as an operator on two qudits. In general, a two-qudit state looks like

\[ |\psi\rangle = \sum_{i,j=1}^{d} \alpha_{ij} |i\rangle \otimes |j\rangle \in \mathbb{C}^{d^2}. \] (4)

(Mathematicians would probably prefer to write \( \mathbb{C}^{d^2} \) as \( \mathbb{C}^{d} \otimes \mathbb{C}^{d} \) here.) The SWAP operator is the linear transformation defined by

\[ \text{SWAP} |\psi\rangle = \sum_{i,j=1}^{d} \alpha_{ij} |j\rangle \otimes |i\rangle \]

when \( |\psi\rangle \) is as in Equation (4).

(a) [**] Explicitly write the matrix for SWAP in the case of \( d = 3 \). Label the rows and columns using a natural order like \( |11\rangle, |12\rangle, |13\rangle, |21\rangle, \ldots, |33\rangle \).

(b) We’re used to SWAP being a quantum gate and thus unitary. Prove that SWAP is also a Hermitian matrix, hence a valid observable for density matrices \( \varrho \) on \( \mathbb{C}^{d^2} \) (or \( \mathbb{C}^{d} \otimes \mathbb{C}^{d} \), if you prefer).

(c) [**] Suppose \( |u_1\rangle, \ldots, |u_d\rangle \) is any orthonormal basis for \( \mathbb{C}^{d} \). This means that the set of all vectors \( |u_i\rangle \otimes |u_j\rangle \) (\( 1 \leq i,j \leq d \)) is an orthonormal basis for \( \mathbb{C}^{d^2} \). Show that SWAP is “basis-independent” in the sense that

\[ |\phi\rangle = \sum_{i,j=1}^{d} \beta_{ij} |u_i\rangle \otimes |u_j\rangle \implies \text{SWAP} |\phi\rangle = \sum_{i,j=1}^{d} \beta_{ij} |u_j\rangle \otimes |u_i\rangle. \]

(d) [**] Suppose you have some quantum apparatus that produces a \( d \)-dimensional particle in a mixed state with density matrix \( \rho \in \mathbb{C}^{d \times d} \). Write the eigenvalues of \( \rho \) as \( \lambda_1, \ldots, \lambda_d \), with associated eigenvectors \( |u_1\rangle, \ldots, |u_d\rangle \). Let \( \varrho = \rho \otimes \rho \), which is the \( d^2 \)-dimensional density matrix corresponding to the state you get if you run your quantum apparatus two times independently and then treat the two particles as a joint system. Prove that

\[ \mathbf{E}_\varrho[\text{SWAP}] = \sum_{i=1}^{d} \lambda_i^2. \]

(e) [**] The quantity \( \sum_{i=1}^{d} \lambda_i^2 \) is called the purity of the mixed state \( \rho \). Show that the maximum possible value of the purity is 1 and it occurs when \( \rho \) is a pure state. Show also that the minimum possible value of the purity is \( 1/d \), and it occurs when \( \rho \) is the maximally mixed state \( \frac{1}{d} \mathbb{1}_{d \times d} \).

(f) Let \( p \in \mathbb{R}^d \) be a probability distribution, and consider the following experiment: make two independent draws from \( i, j \) from \( p \), and let \( S \) be the random variable which is 1 if \( (i,j) = (j,i) \) and is 0 otherwise. Show that \( \mathbf{E}[S] = \sum_{i=1}^{d} p_i^2 \). Prove that this quantity has maximal value 1, occurring when \( p \) has all of its probability on a single outcome; and, prove that this quantity has minimal value \( 1/d \), occurring when \( p \) is the uniform distribution \( \frac{1}{d} \mathbb{1} = (1/d, \ldots, 1/d) \).
7. [Zero-error state discrimination.] Back in Lecture 4.5, we considered the following task. There were two fixed qubit states $|u\rangle, |v\rangle \in \mathbb{R}^2$ which we assumed had real amplitudes for simplicity. We were given access to an unknown qubit state $|\psi\rangle \in \mathbb{R}^2$ (with real amplitudes) and were promised that either $|\psi\rangle = |u\rangle$ or $|\psi\rangle = |v\rangle$. Our goal was to try to guess which is the case. In Lecture 4.5 we saw the optimal algorithm allowing for “two-sided error”, and the optimal algorithm allowing for “one-sided error”. We also saw a natural “zero-sided error” algorithm, but observed that it couldn’t be optimal. In this problem we will see the optimal zero-sided error algorithm (though we won’t prove its optimality). Assume henceforth that the angle between $|u\rangle$ and $|v\rangle$ is $0 < \theta < \pi/2$. Also, write $|u\rangle_\perp$ for a unit vector perpendicular to $|u\rangle$, and $|v\rangle_\perp$ for a unit vector perpendicular to $|v\rangle$.

(a) [**] Let $\Pi_1 = |u\rangle_\perp\langle u\rangle_\perp$, the linear operator on $\mathbb{R}^2$ that projects onto the $|u\rangle_\perp$ vector. Show that $\Pi_1 = 1 - |u\rangle\langle u|$ (where 1 denotes the $2 \times 2$ identity matrix) and that this is a positive operator. We’ll similarly let $\Pi_2 = |v\rangle_\perp\langle v\rangle_\perp$.

(b) [**] The idea of the algorithm is to define $E_1 = \frac{1}{c}\Pi_1$ and $E_2 = \frac{1}{c}\Pi_2$, where $c$ is a positive scalar that is just large enough such that $E_0 = 1 - E_1 - E_2$ is a positive operator. Having done this, $\{E_0, E_1, E_2\}$ becomes a valid POVM. Suppose we then measure the unknown state $\rho = |\psi\rangle\langle \psi|$ with this POVM. Show that when $|\psi\rangle = |u\rangle$, the probability of outcome 1 is 0, and similarly when $|\psi\rangle = |v\rangle$, the probability of outcome 2 is 0.

(c) [**] In light of the previous problem, we see that if we get outcome 1 we can safely guess $|\psi\rangle = |v\rangle$, and if we get outcome 2 we can safely guess $|\psi\rangle = |u\rangle$. If we get outcome 0, we will guess “don’t know”. Our goal, therefore, is to minimize the probability of getting outcome 0. Show that this probability is $1 - \frac{1 - \cos^2 \theta}{c}$.

(d) [**] In light of the previous problem, we clearly want $c$ to be as small as possible. As mentioned, we have the restriction that $E_0$ must be a positive operator. Show that if $|w\rangle \in \mathbb{R}^2$ is any unit vector, $\langle w|E_0|w\rangle = 1 - \frac{\sin^2 \theta_1 + \sin^2 \theta_2}{c}$, where $\theta_1$ is the angle from $|u\rangle$ to $|w\rangle$ and $\theta_2$ is the angle from $|w\rangle$ to $|v\rangle$. We have the restriction $\theta_1 + \theta_2 = \theta$. Hence the least possible $c$ for which $E_0$ is positive is the least $c$ such that $1 - \frac{\sin^2 \theta_1 + \sin^2 \theta_2}{c} \geq 0$ whenever $\theta_1 + \theta_2 = \theta$. Show that this least $c$ is $c = 1 + \cos \theta$.

(e) [**] Deduce that there is a zero-sided error qubit discrimination algorithm with failure probability $\cos \theta$, as claimed at the end of Lecture 4.5.
8. **Quantum information theory.** Learn more about it by watching these lectures of Reinhard Werner on Tobias Osborne’s YouTube channel.
9. [A primer on the statistics of longest increasing subsequences and quantum states.] Take a look at this survey paper describing some research on quantum learning/statistics.