CS 593/MA 595 - Introduction to Quantum Computing

Ralph Razzouk

Spring 2024

rlphrazz@gmail.com

Contents

Homework 1

Problem 2.2
(Matrix representations: example) Suppose V is a vector space with basis vectors $|0\rangle$ and $|1\rangle$, and A is a linear operator from V to V such that $A|0\rangle = |1\rangle$ and $A|1\rangle = |0\rangle$. Give a matrix representation for A, with respect to the input basis $|0\rangle$, $|1\rangle$, and the output basis $|0\rangle$, $|1\rangle$. Find input and output bases which give rise to a different matrix representation of A.

Proof. We recall that

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

By Equation 2.12, a linear operator $A : V \to W$ has a matrix representation given by

$$A|v_j\rangle = \sum_i A_{ij}|w_i\rangle$$

Since we are working only in two dimensions, A will be of the form

$$A = \begin{pmatrix}A_{00} & A_{01}\\ A_{10} & A_{11}\end{pmatrix}.$$

Here, we are given a linear operator $A : V \to V$ ($W = V$) such that

$$A|0\rangle = 0|0\rangle + 1|1\rangle = |1\rangle \implies \begin{pmatrix}A_{00} & A_{01}\\ A_{10} & A_{11}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix} \implies A_{00} = 0, A_{10} = 1$$
$$A|1\rangle = 1|0\rangle + 0|1\rangle = |0\rangle \implies \begin{pmatrix}A_{00} & A_{01}\\ A_{10} & A_{11}\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} \implies A_{01} = 1, A_{11} = 0.$$

Thus, a matrix representation of A in the input and output bases of $\{|0\rangle, |1\rangle\}$ is

$$A = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

Consider the input and output bases of

$$\left\{|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \ |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}\right\}.$$

Applying A to our basis element, and by the linearity of A, we get

$$A|+\rangle = \frac{1}{\sqrt{2}}A(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}}(A|0\rangle + A|1\rangle) = \frac{1}{\sqrt{2}}(|1\rangle + |0\rangle) = |+\rangle$$
$$A|-\rangle = \frac{1}{\sqrt{2}}A(|0\rangle - |1\rangle) = \frac{1}{\sqrt{2}}(A|0\rangle - A|1\rangle) = \frac{1}{\sqrt{2}}(|1\rangle - |0\rangle) = -|-\rangle.$$

Thus

$$A|+\rangle = 1|+\rangle + 0|-\rangle = |+\rangle \implies \begin{pmatrix}A_{00} & A_{01}\\ A_{10} & A_{11}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} \implies A_{00} = 1, A_{10} = 0$$
$$A|-\rangle = 0|+\rangle - 1|-\rangle = -|-\rangle \implies \begin{pmatrix}A_{00} & A_{01}\\ A_{10} & A_{11}\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\-1\end{pmatrix} \implies A_{01} = 0, A_{11} = -1.$$

Thus, a matrix representation of A in the input and output bases of $\{|+\rangle, |-\rangle\}$ is

$$A = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

■

Problem 2.7
Verify that $|w\rangle \equiv (1,1)$ and $|v\rangle \equiv (1,-1)$ are orthogonal. What are the normalized forms of these vectors?

Proof. To show that two vectors are orthogonal, the inner product of these two vectors should be zero. Then

$$\langle v|w\rangle = 1 \cdot 1 + 1 \cdot (-1) = 0.$$

Thus, $|v\rangle$ and $|w\rangle$ are orthogonal.
The normalized form of a vector $|u\rangle$ is given by the following

$$\frac{|u\rangle}{\sqrt{\langle u|u\rangle}}$$

where $\sqrt{\langle u|u\rangle} = ||\,|u\rangle\,||$ is the norm or length of the vector.
The norms of $|v\rangle$ and $|w\rangle$ are

$$\sqrt{\langle v|v\rangle} = \sqrt{2},$$
$$\sqrt{\langle w|w\rangle} = \sqrt{2}.$$

Thus, the normalized forms of the vectors are

$$\frac{|v\rangle}{\sqrt{\langle v|v\rangle}} = \frac{1}{\sqrt{2}}(1,-1) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix},$$
$$\frac{|w\rangle}{\sqrt{\langle w|w\rangle}} = \frac{1}{\sqrt{2}}(1,1) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}.$$

■

Problem 2.10
Suppose $|v_i\rangle$ is an orthonormal basis for an inner product space V. What is the matrix representation for the operator $|v_j\rangle\langle v_k|$, with respect to the $|v_i\rangle$ basis?

Proof. The matrix representation of $|v_j\rangle\langle v_k|$ is an n-dimensional matrix ($\dim(V) = n$) with an entry of 1 in the jth row and kth column.

$$\begin{aligned}
|v_j\rangle\langle v_k| &= \mathbb{I}_V\,|v_j\rangle\langle v_k|\,\mathbb{I}_V \\
&= \left(\sum_p |v_p\rangle\langle v_p|\right)|v_j\rangle\langle v_k|\left(\sum_q |v_q\rangle\langle v_q|\right) \\
&= \sum_{p,q} |v_p\rangle\langle v_p|v_j\rangle\langle v_k|v_q\rangle\langle v_q| \\
&= \sum_{p,q} \delta_{pj}\delta_{kq}\,|v_p\rangle\langle v_q|
\end{aligned}$$

Thus

$$\left(|v_j\rangle\langle v_k|\right)_{pq} = \delta_{pj}\delta_{kq}$$

■

Problem 2.20
(Basis changes) Suppose A' and A'' are matrix representations of an operator A on a vector space V with respect to two different orthonormal bases, $|v_i\rangle$ and $|w_i\rangle$. Then the elements of A' and A'' are $A'_{ij} = \langle v_i|A|v_j\rangle$ and $A''_{ij} = \langle w_i|A|w_j\rangle$. Characterize the relationship between A' and A''.

Proof.

$$\begin{aligned}
A'_{ij} &= \langle v_i|A|v_j\rangle \\
&= \langle v_i|\mathbb{I}A\mathbb{I}|v_j\rangle \\
&= \left\langle v_i \middle| \left(\sum_{i'} |w_{i'}\rangle \langle w_{i'}|\right) A \left(\sum_{j'} |w_{j'}\rangle \langle w_{j'}|\right) \middle| v_j \right\rangle \\
&= \sum_{i',j'} \langle v_i|w_{i'}\rangle \langle w_{i'}|A|w_{j'}\rangle \langle w_{j'}|v_j\rangle \\
&= \sum_{i',j'} \langle v_i|w_{i'}\rangle A''_{ij} \langle w_{j'}|v_j\rangle .
\end{aligned}$$

■

Problem 2.26
Let $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Write out $|\psi\rangle^{\otimes 2}$ and $|\psi\rangle^{\otimes 3}$ explicitly, both in terms of tensor products like $|0\rangle|1\rangle$, and using the Kronecker product.

Proof. Let $|\psi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Then

$$\begin{aligned}
|\psi\rangle^{\otimes 2} &= |\psi\rangle \otimes |\psi\rangle \\
&= \frac{1}{2}((|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)) \\
&= \frac{1}{2}(|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle) \\
&= \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
|\psi\rangle^{\otimes 3} &= |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle \\
&= \frac{1}{2\sqrt{2}}((|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)) \\
&= \frac{1}{2\sqrt{2}}(|0\rangle|0\rangle|0\rangle + |0\rangle|0\rangle|1\rangle + |0\rangle|1\rangle|0\rangle + |0\rangle|1\rangle|1\rangle + |1\rangle|0\rangle|0\rangle + |1\rangle|0\rangle|1\rangle + |1\rangle|1\rangle|0\rangle + |1\rangle|1\rangle|1\rangle) \\
&= \frac{1}{2\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}
\end{aligned}$$

■

Problem 2

(a) Let $\mathcal{H}$ be a finite dimensional Hilbert space. We define the *dual* Hilbert space $\mathcal{H}^*$ to be the set of all linear transformations $\rho : \mathcal{H} \to \mathbb{C}$, that is:

$$\mathcal{H}^* := \{\rho : \mathcal{H} \to \mathbb{C} \mid \rho \text{ is a linear function } \}.$$

Recall that if $|\psi\rangle \in \mathcal{H}$, then $\langle\psi|$ is the linear transformation

$$\begin{aligned} \langle\psi| : \mathcal{H} &\to \mathbb{C} \\ |\psi\rangle &\mapsto \langle\psi|\phi\rangle \end{aligned}$$

where $\langle\psi|\phi\rangle$ is the inner product.

(i) Show that $\mathcal{H}^*$ is a vector space with the same dimension as $\mathcal{H}$.

(ii) Show that the function

$$\begin{aligned} F : \mathcal{H} &\to \mathcal{H}^* \\ |\psi\rangle &\to \langle\psi| \end{aligned}$$

is a bijection. Warning: it is NOT linear (it is *anti-linear* or *conjugate-linear*), so you more-or-less need to show injectivity and surjectivity directly. For surjectivity, use the previous problem part (in particular, the fact that $\mathcal{H}$ is finite dimensional).

In fact, we won't do this, but it's even possible to define an inner product structure on $\mathcal{H}^*$, and the map $|\psi\rangle \mapsto \langle\psi|$ becomes an anti-linear isometry. The moral of the story is that Hilbert spaces are ALMOST isometrically isomorphic to their dual spaces - the only finnicky thing is that the "isomorphism" is not linear, it's anti-linear! This is called the Riesz representation theorem. It's true more generally, *i.e.* for infinite dimensional Hilbert spaces too.

(b) Let $\mathcal{B}(\mathcal{H})$ be the set of all linear transformations $A : \mathcal{H} \to \mathcal{H}$.

(i) Show that $\mathcal{B}(\mathcal{H})$ is a vector space. What is its dimension?

(ii) Show that the map

$$\begin{aligned} \mathcal{H} \otimes \mathcal{H}^* &\to \mathcal{B}(H) \\ |\phi\rangle \otimes \langle\psi| &\mapsto |\phi\rangle\langle\psi| \end{aligned}$$

is a vector space isomorphism.

Proof. (a) (i) Let $\mathcal{H}$ and $\mathbb{C}$ be any two vector spaces over the same field $\mathbf{F}$. Let $\mathcal{H}^* = \mathcal{L}(\mathcal{H}, \mathbb{C})$ be the set of linear transformations $\rho : \mathcal{H} \to \mathbb{C}$.

To show that $\mathcal{H}^*$ is a vector space, we must first define an "*addition of linear transformations*" and a "*scalar multiplication of elements of* $\mathbf{F}$ *by linear transformations*". In other words, our "vectors" will be linear transformations from $\mathcal{H}$ to $\mathbb{C}$. A vector space is just a set with two binary operations, vector addition and scalar multiplication, that satisfy certain properties. We call the elements of a vector space as vectors.

Given two linear transformations $\rho, \sigma : \mathcal{H} \to \mathbb{C}$, we need to define a new linear transformation that is called the "sum of ρ and σ". Denote it by $\rho \oplus \sigma$, to distinguish the "sum of linear transformations" from the sum of vectors. Since we want $\rho \oplus \sigma$ to be a linear transformation (which is a special kind of function) from $\mathcal{H}$ to $\mathbb{C}$, in order to specify it we need to say what the value of $\rho \oplus \sigma$ is at every $\mathbf{h} \in \mathcal{H}$. Define

$$(\rho \oplus \sigma)(\mathbf{h}) = \rho(\mathbf{h}) + \sigma(\mathbf{h}),$$

where the sum on the right is taking place in $\mathbb{C}$. This makes sense, because ρ and σ are already functions from $\mathcal{H}$ to $\mathbb{C}$, so $\rho(\mathbf{h})$ and $\sigma(\mathbf{h})$ are vectors in $\mathbb{C}$, which we can add.

Is $\rho \oplus \sigma$ a linear transformation from $\mathcal{H}$ to $\mathbb{C}$? First, it *is* a function from $\mathcal{H}$ to $\mathbb{C}$. Now, to check that it is a linear transformation, we need to check that $\forall \mathbf{h}_1, \mathbf{h}_2 \in \mathcal{H},\ \forall \alpha \in \mathbf{F}$, we have

$$(\rho \oplus \sigma)(\mathbf{h}_1 + \mathbf{h}_2) = (\rho \oplus \sigma)(\mathbf{h}_1) + (\rho \oplus \sigma)(\mathbf{h}_2) \quad \text{and} \quad (\rho \oplus \sigma)(\alpha \mathbf{h}_1) = \alpha((\rho \oplus \sigma)(\mathbf{h}_1)).$$

Indeed, since ρ and σ are themselves linear transformations, we have

$$\begin{aligned}
(\rho \oplus \sigma)(\mathbf{h}_1 + \mathbf{h}_2) &= \rho(\mathbf{h}_1 + \mathbf{h}_2) + \sigma(\mathbf{h}_1 + \mathbf{h}_2) && \text{(by definition of } \rho \oplus \sigma) \\
&= \rho(\mathbf{h}_1) + \rho(\mathbf{h}_2) + \sigma(\mathbf{h}_1) + \sigma(\mathbf{h}_2) && \text{(by linearity of } \rho \text{ and } \sigma) \\
&= \rho(\mathbf{h}_1) + \sigma(\mathbf{h}_1) + \rho(\mathbf{h}_2) + \sigma(\mathbf{h}_2) \\
&= (\rho \oplus \sigma)(\mathbf{h}_1) + (\rho \oplus \sigma)(\mathbf{h}_2) && \text{(by definition of } \rho \oplus \sigma) \\
(\rho \oplus \sigma)(\alpha \mathbf{h}_1) &= \rho(\alpha \mathbf{h}_1) + \sigma(\alpha \mathbf{h}_1) && \text{(by definition of } \rho \oplus \sigma) \\
&= \alpha\rho(\mathbf{h}_1) + \alpha\sigma(\mathbf{h}_1) && \text{(by linearity of } \rho \text{ and } \sigma) \\
&= \alpha(\rho(\mathbf{h}_1) + \sigma(\mathbf{h}_1)) \\
&= \alpha((\rho \oplus \sigma)(\mathbf{h}_1)) && \text{(by definition of } \rho \oplus \sigma)
\end{aligned}$$

Thus, $\rho \oplus \sigma$ is indeed an element of $\mathcal{H}^* = \mathcal{L}(\mathcal{H}, \mathbb{C})$.

Now, we define a scalar multiplication, which I will denote by $\odot$ (again, to avoid confusion with the scalar multiplication from $\mathcal{H}$ and $\mathbb{C}$. Given $\rho : \mathcal{H} \to \mathbb{C}$ and $\alpha \in \mathbf{F}$, define $(\alpha \odot \rho)$ to be the function

$$(\alpha \odot \rho)(\mathbf{h}) = \alpha\rho(\mathbf{h}).$$

To have a vector space, the eight following axioms must be satisfied $\forall \rho, \sigma, \tau \in \mathcal{H}^*$ and $\forall \alpha, \beta \in \mathbf{F}$, we have

- **Associativity of transformation addition:** $(\rho \oplus \sigma) \oplus \tau = \rho \oplus (\sigma \oplus \tau)$.
- **Commutativity of transformation addition:** $\rho \oplus \sigma = \sigma \oplus \rho$.
- **Identity element of transformation addition:** There exists an element $\mathbf{0} \in \mathcal{H}^*$, called the zero transformation, such that $\rho + 0 = \rho$, $\forall \rho \in \mathcal{H}^*$.
- **Inverse elements of transformation addition** For every $\rho \in \mathcal{H}^*$, there exists an element $(-\rho) \in \mathcal{H}^*$, called the additive inverse of ρ, such that $\rho + (-\rho) = \mathbf{0}$.
- **Compatibility of scalar multiplication with field multiplication:** $\alpha(\beta\rho) = (\alpha\beta)\rho$.
- **Identity element of scalar multiplication:** $\mathbf{1}\rho = \rho$, where $\mathbf{1}$ denotes the multiplicative identity in $\mathbf{F}$.
- **Distributivity of scalar multiplication with respect to transformation addition:** $\alpha(\rho + \sigma) = \alpha\rho + \alpha\sigma$.
- **Distributivity of scalar multiplication with respect to field addition:** $(\alpha + \beta)\rho = \alpha\rho + \beta\rho$.

To prove that $\mathcal{H}^*$ has the same dimension as $\mathcal{H}$, let $\rho \in \mathcal{H}^*$ and let $\{e_1, \ldots, e_n\}$ be a basis for $\mathcal{H}$. Define $e^i \in \mathcal{H}^*$ by $e^i(e_j) = \delta_{ij}$. We want to show that $\{e^1, \ldots, e^n\}$ spans $\mathcal{H}^*$.

$$\rho(\mathbf{h}) = \rho(h_1 e_1 + \cdots + h_n e_n) = h_1\rho(e_1) + \cdots + h_n\rho(e_n).$$

If $h_1\rho(e_1) = \lambda_1, \ldots, h_n\rho(e_n) = \lambda_n$, then $\rho(\mathbf{h}) = h_1\lambda_1 e^1(e_1) + \cdots + h_n\lambda_n e^n(e_n) = \lambda_1 e^1(\mathbf{h}) + \cdots + \lambda_n e^n(\mathbf{h})$.

To show that the set $\{e^1, \ldots, e^n\}$ is linearly independent, suppose that $\mathbf{0} = c_1 e^1 + \ldots + c_n e^n$ is the zero mapping. Consider the image of $e_1 : 0(e_1) = c_1 * 1 + \ldots + c_n * 0 = c_1 \implies c_1 = 0$. By repeating the procedure for all $e_j, 2 \leq j \leq n$, we see that $c_1 = \cdots = c_n = 0$.

(ii) A function $F : A \to B$ is a bijection if and only if F is:

- **Injective:** $F(x) = F(y) \implies x = y$
- **Surjective:** $\forall b \in B, \exists a \in A \ / \ F(a) = b$

■

Problem 3
In this set of exercises, we dig into the spectral theorem/diagonalization in more detail.

(a) Prove the corollaries of the spectral theorem (Theorem 2.1 in Nielsen and Chuang) that I stated on Tuesday by doing exercises 2.17 and 2.18 . [Hint: being "corollaries" of course means that you should use the spectral theorem in your proof.]

(b) Differing slightly from the book (see page 70), in this problem let us define an orthogonal projector to be a linear operator $P : \mathcal{H} \to \mathcal{H}$ such that $P^2 = P$ and $P^* = P$. Show that P is an orthogonal projector if and only if P is unitarily diagonalizable, with all eigenvalues equal to either 0 or 1.

(c) Do exercises 2.29-2.32. [Hint: you can use the previous parts of this exercise, or, closely related, Exercise 2.28 (which you need not prove for this problem).]

Proof. (a) **Exercise 2.17:** Show that a normal matrix is Hermitian if and only if it has real eigenvalues.
Proof:

$\boxed{\implies}$ Let A be a Normal and Hermitian matrix. Then, by the spectral decomposition theorem, A has a diagonal representation given by $\sum_i \lambda_i |i\rangle \langle i|$, where the set of $|i\rangle$ form an orthonormal basis for V and each $|i\rangle$ is an eigenvector with eigenvalue λ_i. Since A is Hermitian, then $A^\dagger = A$. Then, we have

$$\begin{aligned} A^\dagger &= \left(\sum_i \lambda_i |i\rangle \langle i| \right)^\dagger \\ &= \sum_i \lambda_i^* |i\rangle \langle i| \\ &= A \\ &= \sum_i \lambda_i |i\rangle \langle i| \end{aligned}$$

Thus, $\lambda_i^* = \lambda_i \implies \lambda_i \in \mathbb{R}$ and A has real eigenvalues.

$\boxed{\impliedby}$ Let A be a normal matrix with real eigenvalues. Then, by the spectral decomposition theorem, A has a diagonal representation given by $\sum_i \lambda_i |i\rangle \langle i|$, where the set of $|i\rangle$ form an orthonormal basis for V and each $|i\rangle$ is an eigenvector with real eigenvalue λ_i. Then, we have

$$\begin{aligned} A^\dagger &= \left(\sum_i \lambda_i |i\rangle \langle i| \right)^\dagger \\ &= \sum_i \lambda_i^* |i\rangle \langle i| \\ &= \sum_i \lambda_i |i\rangle \langle i| \\ &= A \end{aligned}$$

Thus, A is Hermitian.

Exercise 2.18: Show that all eigenvalues of a unitary matrix have modulus 1, that is, can be written in the form $e^{i\theta}$ for some real θ.

Proof: Let U be a unitary matrix. It is then normal as $U^\dagger U = UU^\dagger = \mathbb{I}$. Then, by the spectral decomposition theorem, U has a representation given by $\sum_i \lambda_i |i\rangle \langle i|$, where the set of $|i\rangle$ form an

orthonormal basis for V and each $|i\rangle$ is an eigenvector with real eigenvalue λ_i. Then, we have

$$\begin{aligned} UU^\dagger = \mathbb{I} &= \left(\sum_i \lambda_i |i\rangle \langle i|\right)\left(\sum_{i'} \lambda_{i'} |i'\rangle \langle i'|\right)^\dagger \\ &= \sum_{i,i'} \lambda_i \lambda_{i'}^* |i\rangle \langle i|i'\rangle \langle i'| \\ &= \sum_{i,i'} \lambda_i \lambda_{i'}^* |i\rangle \delta_{i,i'} \langle i'| \\ &= \sum_i \lambda_i \lambda_i^* |i\rangle \langle i| \\ &= \sum_i |\lambda_i|^2 |i\rangle \langle i| \end{aligned}$$

Thus, $|\lambda_i|^2 = 1 \implies |\lambda_i| = 1$.

(b) Define an orthogonal projector to be a linear operator $P : \mathcal{H} \to \mathcal{H}$ such that $P^2 = P$ and $P^* = P$. Consider the eigenvalue equation $P|\psi\rangle = \lambda|\psi\rangle$. Applying P again, we get

$$P^2|\psi\rangle = P\lambda|\psi\rangle = \lambda^2|\psi\rangle$$

but also

$$P^2|\psi\rangle = P|\psi\rangle = \lambda|\psi\rangle$$

Hence, $\lambda^2 = \lambda \implies \lambda = 0, 1$.

If P is a projector, that means there's a subspace $\mathcal{H}$ onto which it projects. It maps every vector in $\mathcal{H}$ to itself. Therefore every vector in $\mathcal{H}$ is an eigenvector with eigenvalue 1. Every vector not in $\mathcal{H}$ is mapped to a vector in $\mathcal{H}$. Take any vector $|\psi\rangle$ and write

$$|\psi\rangle = P|\psi\rangle + (|\psi\rangle - P|\psi\rangle)$$

so the first term $P|\psi\rangle \in \mathcal{H}$. It is easy to see that the second term, $|\psi\rangle - P|\psi\rangle$, is in the kernel of P: the first term is mapped to $P|\psi\rangle$, and the second is mapped to $P|\psi\rangle - P^2|\psi\rangle$. But since $P|\psi\rangle$ is in $\mathcal{H}$, it must be fixed by P, so $P^2|\psi\rangle = P|\psi\rangle$; thus $P(|\psi\rangle - P|\psi\rangle) = 0$. In this way, every vector $|\psi\rangle$ is written as the sum of a vector in $\mathcal{H}$, which is an eigenvector with eigenvalue 1, and a vector in the kernel of P, which is an eigenvector with eigenvalue 0. So, forming a basis of the whole space by taking the union of a basis of $\mathcal{H}$ and a basis of the kernel of P, and the matrix of P with respect to that basis is

$$\begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & 0 & \\ & & & & & & \ddots \\ & & & & & & & 0 \end{bmatrix}$$

(and all off-diagonal entries are 0) where the number of 1's is the dimension of $\mathcal{H}$ and the number of 0's is the dimension of the kernel of P.

Thus, P is unitarily diagonalizable with all eigenvalues equal to either 0 or 1.

(c) **Exercise 2.29:** Show that the tensor product of two unitary operators is unitary.

Proof: An operator U is said to be unitary if $U^\dagger U = \mathbb{I}$. Suppose A and B are two unitary operators. We need to show that $A \otimes B$ is also unitary. We have

$$(A \otimes B)^\dagger (A \otimes B) = (A^\dagger \otimes B^\dagger)(A \otimes B) = (A^\dagger A) \otimes (B^\dagger B) = \mathbb{I} \otimes \mathbb{I}.$$

Sometimes, the definition of a unitary operator is given as $UU^\dagger = \mathbb{I}$. In that case

$$(A \otimes B)(A \otimes B)^\dagger = (A \otimes B)(A^\dagger \otimes B^\dagger) = (AA^\dagger) \otimes (BB^\dagger) = \mathbb{I} \otimes \mathbb{I}.$$

Thus, the tensor product of two unitary operators is unitary.

Exercise 2.30: Show that the tensor product of two Hermitian operators is Hermitian.

Proof: An operator H is said to be Hermitian if $H = H^\dagger$. Suppose A and B are two Hermitian operators. We have

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger = A \otimes B.$$

Thus, the tensor product of two Hermitian operators is Hermitian.

Exercise 2.31: Show that the tensor product of two positive operators is positive.

Proof: An operator P is said to be positive if $\langle\psi|P|\psi\rangle \geq 0$, for all $|\psi\rangle$. Suppose A and B are two positive operators. For any $|v\rangle \otimes |w\rangle$, we have

$$\langle v| \otimes \langle w| (A \otimes B) |v\rangle \otimes |w\rangle = \langle v|A|v\rangle \langle w|B|w\rangle \geq 0$$

Thus, the tensor product of two positive operators is positive.

Exercise 2.32: Show that the tensor product of two projectors is a projector.

Proof: An operator P is said to be a projector if $P^2 = P$. Suppose A and B are two projectors. We have

$$(A \otimes B)^2 = (A \otimes B)(A \otimes B) = A^2 \otimes B^2 = A \otimes B.$$

Thus, the tensor product of two projectors is a projector.

■

Problem 2.42
Verify that

$$AB = \frac{[A, B] + \{A, B\}}{2}$$

Proof.

$$\frac{[A, B] + \{A, B\}}{2} = \frac{(AB - BA) + (AB + BA)}{2} = \frac{(AB + AB) + (BA - BA)}{2} = AB$$

■

Problem 2.44
Suppose $[A, B] = 0$, $\{A, B\} = 0$, and A is invertible. Show that B must be 0.

Proof. If $[A, B] = 0$, then $AB = BA$. If $\{A, B\} = 0$, then $AB = -BA$. This implies that $AB = 0$. It must be that $A = 0$ or $B = 0$. Since A is invertible, this means $A \neq 0$. Thus, $B = 0$. ■

Problem 2.45
Show that $[A, B]^\dagger = [B^\dagger, A^\dagger]$.

Proof.

$$[A, B]^\dagger = (AB - BA)^\dagger = (AB)^\dagger - (BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = [B^\dagger, A^\dagger].$$

■

Problem 2.46
Show that $[A, B] = -[B, A]$.

Proof.

$$[A, B] = AB - BA = -(BA - AB) = -[B, A].$$

■

Problem 2.47
Suppose A and B are Hermitian. Show that $i[A, B]$ is Hermitian.

Proof. To show that an operator is Hermitian, we take the Hermitian of that operator and show it is equal to the operator itself.

$$(i[A, B])^\dagger = -i[A, B]^\dagger = -i[B^\dagger, A^\dagger] = -i[B, A] = i[A, B].$$

Thus, $i[A, B]$ is Hermitian. ■

Homework 2

Problem 2.57
(Cascaded measurements are single measurements) Suppose $\{L_l\}$ and $\{M_m\}$ are two sets of measurement operators. Show that a measurement defined by the measurement operators $\{L_l\}$ followed by a measurement defined by the measurement operators $\{M_m\}$ is physically equivalent to a single measurement defined by measurement operators $\{N_{lm}\}$ with the representation $N_{lm} \equiv M_m L_l$.

Proof. Consider a normalized initial quantum state $|\psi_0\rangle$. The state $|\psi_0\rangle$, after of measurement of L_l is given, by definition, to be

$$|\psi_0\rangle \mapsto |\psi_1\rangle = \frac{L_l |\psi_0\rangle}{\sqrt{\left\langle \psi_0 \middle| L_l^\dagger L_l \middle| \psi_0 \right\rangle}}.$$

The state $|\psi_1\rangle$, after of measurement of M_m, is given to be

$$\begin{aligned}
|\psi_1\rangle \mapsto |\psi_2\rangle &= \frac{M_m |\psi_1\rangle}{\sqrt{\left\langle \psi_1 \middle| M_m^\dagger M_m \middle| \psi_1 \right\rangle}} \\
&= \frac{M_m \left(\frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}} \right)}{\sqrt{\left\langle \frac{L_l^\dagger\langle\psi_0|}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}} \middle| M_m^\dagger M_m \middle| \frac{L_l|\psi_0\rangle}{\sqrt{\langle\psi_0|L_l^\dagger L_l|\psi_0\rangle}} \right\rangle}} \\
&= \frac{M_m L_l |\psi_0\rangle}{\sqrt{\left\langle \psi_0 \middle| L_l^\dagger L_l \middle| \psi_0 \right\rangle}} \frac{\sqrt{\left\langle \psi_0 \middle| L_l^\dagger L_l \middle| \psi_0 \right\rangle}}{\sqrt{\left\langle \psi_0 \middle| L_l^\dagger M_m^\dagger M_m L_l \middle| \psi_0 \right\rangle}} \\
&= \frac{M_m L_l |\psi_0\rangle}{\sqrt{\left\langle \psi_0 \middle| L_l^\dagger M_m^\dagger M_m L_l \middle| \psi_0 \right\rangle}}.
\end{aligned}$$

If we define $N_{lm} \equiv M_m L_l$, then the state $|\psi_0\rangle$, after measurement of N_{lm}, is equivalent to the state $|\psi_2\rangle$. In other words, the state $|\psi_2\rangle$, after the measurement of N_{lm}, yields the same state. ■

Problem 2.58
Suppose we prepare a quantum system in an eigenstate $|\psi\rangle$ of some observable M, with corresponding eigenvalue m. What is the average observed value of M, and the standard deviation?

Proof. By the definition of the expectation value, we have

$$\begin{aligned}
\langle M \rangle_{|\psi\rangle} &= \langle\psi|M|\psi\rangle \\
&= \langle\psi|m|\psi\rangle \quad \text{(definition of eigenvalue)} \\
&= m \langle\psi|\psi\rangle \\
&= m \qquad \text{(normalized quantum state).}
\end{aligned}$$

Calculating the expectation value of the square, we have

$$\begin{aligned}
\left\langle M^2 \right\rangle_{|\psi\rangle} &= \left\langle \psi \middle| M^2 \middle| \psi \right\rangle \\
&= \langle\psi|MM|\psi\rangle \\
&= m \langle\psi|M|\psi\rangle \\
&= m^2 \langle\psi|\psi\rangle \\
&= m^2.
\end{aligned}$$

Calculating the standard deviation, we have

$$\Delta(M) = \sqrt{\langle M^2 \rangle - \langle M \rangle^2} = \sqrt{m^2 - m^2} = 0.$$

■

Problem 2.59
Suppose we have qubit in the state $|0\rangle$, and we measure the observable X. What is the average value of X? What is the standard deviation of X?

Proof. The expectation value of X is

$$\begin{aligned} \langle X \rangle_{|0\rangle} &= \langle 0|X|0\rangle \\ &= \langle 0|1\rangle \quad \text{(definition of } X\text{)} \\ &= 0. \end{aligned}$$

Calculating the expectation value of the square, we have

$$\begin{aligned} \langle X^2 \rangle_{|0\rangle} &= \langle 0|X^2|0\rangle \\ &= \langle 0|XX|0\rangle \\ &= \langle 0|X|1\rangle \\ &= \langle 0|0\rangle \\ &= 1. \end{aligned}$$

Calculating the standard deviation, we have

$$\Delta(X) = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} = \sqrt{1 - 0} = 1.$$

■

Problem 2.60
Show that $\vec{v} \cdot \vec{\sigma}$ has eigenvalues ± 1, and that the projectors onto the corresponding eigenspaces are given by $P_\pm = (I \pm \vec{v} \cdot \vec{\sigma})/2$.

Proof. Let $|v\rangle$ be a unit vector and $|\sigma\rangle = (\sigma_1, \sigma_2, \sigma_3)$, where σ_i are the Pauli's sigma matrices. First, we compute $\vec{v} \cdot \vec{\sigma}$

$$\vec{v} \cdot \vec{\sigma} = v_1\sigma_1 + v_2\sigma_2 + v_3\sigma_3 = v_1 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + v_2 \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + v_3 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} v_3 & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 \end{bmatrix}.$$

Using the characteristic equation to find the eigenvalues, we have

$$\begin{aligned} \det(\vec{v} \cdot \vec{\sigma} - \lambda \mathbb{I}) &= 0 \\ \begin{vmatrix} v_3 - \lambda & v_1 - iv_2 \\ v_1 + iv_2 & -v_3 - \lambda \end{vmatrix} &= 0 \\ -(v_3 + \lambda)(v_3 - \lambda) - (v_1 + iv_2)(v_1 - iv_2) &= 0 \\ \lambda^2 - (v_1^2 + v_2^2 + v_3^2) &= 0 \\ \lambda^2 - 1 &= 0 \quad \text{(since } |v\rangle \text{ is a unit vector)} \\ \lambda_\pm &= \pm 1. \end{aligned}$$

Finding the eigenvectors, we have

- For $\lambda_- = -1$:

$$(\vec{v}\cdot\vec{\sigma} - \lambda_-\mathbb{I})\,|\lambda_-\rangle = 0 \begin{pmatrix} v_3+1 & v_1-iv_2 \\ v_1+iv_2 & -v_3+1 \end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = 0$$

$$\implies \begin{cases} (v_3+1)\alpha + (v_1-iv_2)\beta = 0 \\ (v_1+iv_2)\alpha + (-v_3+1)\beta = 0 \end{cases}$$

$$\implies \begin{cases} \beta = \frac{v_1+iv_2}{v_3-1}\alpha \\ \alpha = v_3 - 1 \text{ (arbitrary)} \end{cases}$$

$$\begin{aligned} |\lambda_-\rangle &= \frac{1}{\sqrt{|v_3-1|^2 + |v_1+iv_2|^2}}\begin{pmatrix} v_3-1 \\ v_1+iv_2\end{pmatrix} \\ &= \frac{1}{\sqrt{v_3^2 - 2v_3 + 1 + v_1^2 + v_2^2}}\begin{pmatrix} v_3-1 \\ v_1+iv_2\end{pmatrix} \\ &= \frac{1}{\sqrt{2(1-v_3)}}\begin{pmatrix} v_3-1 \\ v_1+iv_2\end{pmatrix} \end{aligned}$$

Let P_- be the projector of $|\lambda_-\rangle$, then

$$\begin{aligned} P_- &= |\lambda_-\rangle\langle\lambda_-| \\ &= \frac{1}{2(1-v_3)}\begin{pmatrix} v_3-1 \\ v_1+iv_2\end{pmatrix}\begin{pmatrix} v_3-1 & v_1-iv_2\end{pmatrix} \\ &= \frac{1}{2(1-v_3)}\begin{pmatrix} (v_3-1)^2 & (v_3-1)(v_1-iv_2) \\ (v_1+iv_2)(v_3-1) & (v_1+iv_2)(v_1-iv_2)\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1-v_3 & -v_1+iv_2 \\ -v_1-iv_2 & \frac{v_1^2+v_2^2}{1-v_3}\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1-v_3 & -v_1+iv_2 \\ -v_1-iv_2 & \frac{1-v_3^2}{1-v_3}\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1-v_3 & -v_1+iv_2 \\ -v_1-iv_2 & 1+v_3\end{pmatrix} \\ &= \frac{1}{2}(\mathbb{I} - \vec{v}\cdot\vec{\sigma}) \end{aligned}$$

- For $\lambda_+ = 1$:

$$(\vec{v}\cdot\vec{\sigma} - \lambda_+\mathbb{I})\,|\lambda_+\rangle = 0 \begin{pmatrix} v_3-1 & v_1-iv_2 \\ v_1+iv_2 & -v_3-1 \end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = 0$$

$$\implies \begin{cases} (v_3-1)\alpha + (v_1-iv_2)\beta = 0 \\ (v_1+iv_2)\alpha + (-v_3-1)\beta = 0 \end{cases}$$

$$\implies \begin{cases} \beta = \frac{v_1+iv_2}{v_3+1}\alpha \\ \alpha = v_3 + 1 \text{ (arbitrary)} \end{cases}$$

$$\begin{aligned}|\lambda_+\rangle &= \frac{1}{\sqrt{|v_3+1|^2+|v_1+iv_2|^2}}\begin{pmatrix} v_3+1 \\ v_1+iv_2\end{pmatrix} \\ &= \frac{1}{\sqrt{v_3^2+2v_3+1+v_1^2+v_2^2}}\begin{pmatrix} v_3+1 \\ v_1+iv_2\end{pmatrix} \\ &= \frac{1}{\sqrt{2(1+v_3)}}\begin{pmatrix} v_3+1 \\ v_1+iv_2\end{pmatrix}\end{aligned}$$

Let P_+ be the projector of $|\lambda_+\rangle$, then

$$\begin{aligned}P_+ &= |\lambda_+\rangle\langle\lambda_+| \\ &= \frac{1}{2(1+v_3)}\begin{pmatrix} v_3+1 \\ v_1+iv_2\end{pmatrix}\begin{pmatrix} v_3+1 & v_1-iv_2\end{pmatrix} \\ &= \frac{1}{2(1+v_3)}\begin{pmatrix} (v_3+1)^2 & (v_3+1)(v_1-iv_2) \\ (v_1+iv_2)(v_3+1) & (v_1+iv_2)(v_1-iv_2)\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1+v_3 & v_1-iv_2 \\ v_1+iv_2 & \frac{v_1^2+v_2^2}{1+v_3}\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1+v_3 & v_1-iv_2 \\ v_1+iv_2 & \frac{1-v_3^2}{1+v_3}\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1+v_3 & v_1-iv_2 \\ v_1+iv_2 & 1-v_3\end{pmatrix} \\ &= \frac{1}{2}(\mathbb{I}+\vec{v}\cdot\vec{\sigma})\end{aligned}$$

Thus,

$$P_\pm = \frac{1}{2}(\mathbb{I}\pm\vec{v}\cdot\vec{\sigma})$$

as claimed. ■

Problem 2.61
Calculate the probability of obtaining the result $+1$ for a measurement of $\vec{v}\cdot\vec{\sigma}$, given that the state prior to measurement is $|0\rangle$. What is the state of the system after the measurement if $+1$ is obtained?

Proof. From Problem 2.60, we have

$$P_\pm = \frac{1}{2}(\mathbb{I}\pm\vec{v}\cdot\vec{\sigma})$$

where

$$\vec{v}\cdot\vec{\sigma} = \begin{bmatrix} v_3 & v_1-iv_2 \\ v_1+iv_2 & -v_3\end{bmatrix} = v_3|0\rangle\langle 0| + (v_1-iv_2)|0\rangle\langle 1| + (v_1+iv_3)|1\rangle\langle 0| - v_3|1\rangle\langle 1|$$

The probability of obtaining the result $+1$ given that the state prior to measurement is $|0\rangle$ is

$$\begin{aligned}p(+1) &= \langle 0|P_+|0\rangle \\ &= \frac{1}{2}\langle 0|\mathbb{I}+\vec{v}\cdot\vec{\sigma}|0\rangle \\ &= \frac{1}{2}\left[\langle 0|\mathbb{I}|0\rangle + \langle 0|\vec{v}\cdot\vec{\sigma}|0\rangle\right] \\ &= \frac{1}{2}\left[1+\langle 0|v_3|0\rangle + \langle 0|(v_1+iv_3)|1\rangle\right] \\ &= \frac{1+v_3}{2}.\end{aligned}$$

The state of the system after the measurement if +1 is obtained is

$$|0\rangle \mapsto \frac{P_+ |0\rangle}{\sqrt{p(+1)}} = \frac{\frac{1+v_3}{2}|0\rangle + \frac{v_1+iv_2}{2}|1\rangle}{\sqrt{\frac{1+v_3}{2}}} = \frac{(1+v_3)|0\rangle + (v_1+iv_2)|1\rangle}{\sqrt{2(1+v_3)}}.$$

■

Problem 2.66
Show that the average value of the observable $X_1 Z_2$ for a two qubit system measured in the state $(|00\rangle + |11\rangle)/\sqrt{2}$ is zero.

Proof. The expectation value of the observable $X_1 Z_2$ when measured in the state $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$ is

$$\begin{aligned}
\langle X_1 Z_2 \rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}}\right) X_1 Z_2 \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) \\
&= \frac{1}{2}\left(\langle 00| + \langle 11|\right)\left(X_1 Z_2 |00\rangle + X_1 Z_2 |11\rangle\right) \\
&= \frac{1}{2}\left(\langle 00| + \langle 11|\right)\left(X_1 |00\rangle - X_1 |11\rangle\right) \\
&= \frac{1}{2}\left(\langle 00| + \langle 11|\right)\left(|10\rangle - |01\rangle\right) \\
&= \frac{1}{2}\left(\langle 00|10\rangle - \langle 00|01\rangle + \langle 11|10\rangle - \langle 11|01\rangle\right) \\
&= 0.
\end{aligned}$$

■

Problem 2

(a) A vector $|\psi\rangle$ in a tensor product Hilbert space $V \otimes W$ is called *separable* (or *unentangled*) if there exist vectors $|v\rangle \in V$ and $|w\rangle \in W$ such that $|\psi\rangle = |v\rangle \otimes |w\rangle$. Give an example of a state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes 2}$ on two qubits that is not separable (in other words, it is entangled). Justify your answer.

(b) Show that $V \otimes W$ has no entangled states if and only if V or W is 0 or 1 dimensional.

Proof. (a) The most common example of an entangled state is the Bell state $|\psi\rangle = \frac{1}{2}(|00\rangle + |11\rangle)$.

To determine if the pure state is entangled, one could try a brute force method of attempting to find satisfying states $|v\rangle \in V$ and $|w\rangle \in W$ such that $|\psi\rangle = |v\rangle \otimes |w\rangle$. This is inelegant, and hard work in the general case. A more straightforward way to prove whether this pure state is entangled is the calculate the reduced density matrix ρ for one of the qubits, i.e. by tracing out the other. The state is separable if and only if ρ has rank 1. Otherwise it is entangled. Mathematically, we can test the rank condition simply by evaluating $\mathrm{Tr}(\rho^2)$. The original state is separable if and only if this value is 1. Otherwise the state is entangled.

Suppose we have a pure separable state $|\phi\rangle = |v\rangle \otimes |w\rangle$. The reduced density matrix is

$$\rho = \mathrm{Tr}(|\phi\rangle\langle\phi|) = |v\rangle\langle v|,$$

and

$$\mathrm{Tr}(\rho^2) = \mathrm{Tr}(|v\rangle\langle v| \cdot |v\rangle\langle v|) = \mathrm{Tr}(|v\rangle\langle v|) = 1.$$

Thus, we have a separable state.

Considering our Bell state $|\psi\rangle$, then

$$\rho = \mathrm{Tr}(|\psi\rangle\langle\psi|) = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11|) = \frac{1}{2}\mathbb{I},$$

and

$$\mathrm{Tr}(\rho^2) = \frac{1}{4}\mathrm{Tr}(\mathbb{I}\cdot\mathbb{I}) = \frac{1}{2} \neq 1$$

Thus, the Bell state is not a separable state and is, hence, entangled.

(b) Suppose that $\dim(V) = m$ and $\dim(W) = n$.

$\boxed{\Longrightarrow}$ Suppose the product $V \otimes W$ has no entangled states. Then, for every element $|\psi\rangle \in V \otimes W$, there exists vectors $|v\rangle \in V$ and $|w\rangle \in W$ such that $|\psi\rangle = |v\rangle \otimes |w\rangle$.

Let us assume that $\dim(V) = m \geq 2$ and $\dim(W) = n \geq 2$ and suppose that $\{|v_1\rangle, |v_2\rangle\}$ are linearly independent in V and that $\{|w_1\rangle, |w_2\rangle\}$ are linearly independent in W. Seeking a contradiction, suppose that

$$|v_1\rangle \otimes |w_1\rangle + |v_2\rangle \otimes |w_2\rangle = |v\rangle \otimes |w\rangle . \tag{1}$$

Extending $\{|v_1\rangle, |v_2\rangle\}$ to a basis $\{|v_1\rangle, |v_2\rangle, \ldots, |v_m\rangle\}$ of V and $\{|w_1\rangle, |w_2\rangle\}$ to a basis $\{|w_1\rangle, |w_2\rangle, \ldots, |w_n\rangle\}$ of W, we have

$$\begin{aligned} |v\rangle &= \alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle + \cdots + \alpha_m |v_m\rangle \\ |w\rangle &= \beta_1 |w_1\rangle + \beta_2 |w_2\rangle + \cdots + \beta_n |w_n\rangle \end{aligned}$$

$$\implies |v\rangle \otimes |w\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i\beta_j |v_i\rangle \otimes |w_j\rangle . \tag{2}$$

Since the tensor product $V \otimes W$ has no entangled states, then a valid basis for the tensor product space is

$$\mathcal{B} = \{|v_i\rangle \otimes |v_j\rangle \,|\, 1 \leq i \leq m, 1 \leq j \leq n\}.$$

From the equations (1) and (2), we have that

$$\alpha_i\beta_j = \begin{cases} 1, & \text{for } i,j = 1 \\ 1, & \text{for } i,j = 2 \\ 0, & \text{otherwise} \end{cases} \implies \alpha_i = \begin{cases} 1, & \text{for } i = 1,2 \\ 0, & \text{for } i \neq 1,2 \end{cases}, \quad \beta_j = \begin{cases} 1, & \text{for } j = 1,2 \\ 0, & \text{for } j \neq 1,2 \end{cases}.$$

We may now write equation (1) as

$$\begin{aligned} |v_1\rangle \otimes |w_1\rangle + |v_2\rangle \otimes |w_2\rangle &= |v\rangle \otimes |w\rangle \\ &= (|v_1\rangle + |v_2\rangle) \otimes (|w_1\rangle + |w_2\rangle) \\ &= |v_1\rangle \otimes |w_1\rangle + |v_1\rangle \otimes |w_2\rangle + |v_2\rangle \otimes |w_1\rangle + |v_2\rangle \otimes |w_2\rangle \end{aligned}$$

which contradicts that $\mathcal{B}$ is a basis for $V \otimes W$.

Therefore, if the tensor product space $V \otimes W$ has no entangled states, then $\dim(V) = m \leq 1$ or $\dim(W) = n \leq 1$.

$\boxed{\Longleftarrow}$

- Without loss of generality, suppose that V is zero-dimensional. Then it must be that $|v\rangle = \mathbf{0}$ is the only element in V. Then any $|\psi\rangle \in V \otimes W$ is of the form

$$|\psi\rangle = \mathbf{0} \otimes |w\rangle = \mathbf{0}.$$

In addition, $\mathbf{0}$ is always separable. In fact

$$\mathbf{0} \otimes \mathbf{0} = \mathbf{0}.$$

Thus, the tensor product space $V \otimes W$, where V or W is zero-dimensional, has no entangled states.

- Without loss of generality, suppose that V is one-dimensional. Then it must be that $|v\rangle = \alpha |0\rangle$. Then any $|\psi\rangle \in V \otimes W$ is of the form

$$|\psi\rangle = |v\rangle \otimes |w\rangle = \alpha |0\rangle \otimes \sum_{j=1}^{n} \beta_j |w_j\rangle ,$$

which is always separable.

Thus, the tensor product space $V \otimes W$, where V or W is one-dimensional, has no entangled states.

Therefore, if V or W is zero or one dimensional, then the tensor product space $V \otimes W$ has no entangled states.

■

Problem 3

Let's work through the details of quantum state tomography via repeated measurements in the computational basis.

Let

$$|\psi\rangle = \sum_{b=0}^{2^n-1} z_b |b\rangle \in (\mathbb{C}^2)^{\otimes 2}$$

be some unknown state on n qubits, which we will assume is normalized. The goal of quantum state tomography is to determine what the amplitudes z_b are—up to a given error, with high confidence. We don't yet have the tools to do things at this level of precision quite yet, but we can at least ask about trying to determine, say, $|z_0|^2$ up to some given accuracy.

Since measurement collapses the state, we will assume that we are able to prepare copies of this state for free. On each copy, we will perform projective measurement in the computational basis. The outcomes will be independent and identically distributed. If we do this k times, we get a sequence of outcomes $(i_1, \ldots, i_k)$ where each $i_j \in \{0, \ldots, 2^n - 1\}$. From this, we may compute an empirical probability distribution $\tilde{p}_k$ on the set $\{0, \ldots, 2^n - 1\}$ simply by counting the different outcomes and dividing by k

$$\tilde{p}_k(i) := \frac{\#\{j | i_j = i\}}{k}. \tag{3}$$

Of course, the *true* distribution of outcomes is given by the Born rule:

$$p(i) = p(i |\psi\rangle) = |z_i|^2 = z_i z_i^*.$$

Let $\epsilon > 0$. We would like to know how many rounds of our experiment we need to perform—that is, how large k needs to be—in order for us to be able to *confidently* say that our empirical estimate $\tilde{p}_k(0)$ is within ϵ of the true value $p(0)$. This requires a little bit of explaining, basically having to do with the fact that $\tilde{p}_k(0)$ is itself a random variable (on the set $\{0, 1/k, 2/k, \ldots, k/k = 1\}$, but don't think too hard about this).

Let us say that we are $\delta - confident$ that our observed $\tilde{p}_k(0)$ is within ϵ if we pick k large enough so that

$$\text{Prob}(|\tilde{p}_k(0) - p(0)| \geq \epsilon) \leq \delta$$

Our goal is to find a lower bound on k (as a function of ϵ, but independent of everything else) that makes this inequality true.

To do so, we can use Chebyshev's inequality (see Appendix 1 in Nielsen and Chuang). This problem will walk you through this. The idea is exactly the same as trying to get a good estimate of the bias of an unfair coin with high confidence

(a) Let Y be the random variable on the set $\{0,1\}$ with $p(0) = 1 - |z_0|^2$ and $p(1) = |z_0|^2$. Show that $\mathbb{E}(Y) = \mathbb{E}(Y^2) = |z_0|^2$. Use this to show the variance $\text{var}(Y) = |z_0|^2 - |z_0|^4 = |z_0|^2(1 - |z_0|^2)$.

(b) Show that $\max_{0\leq p\leq 1} p(1-p) = 1/4$. Conclude that $var(Y) \leq 1/4$.

(c) Now let $Y_1, \ldots, Y_k$ be k i.i.d variables all having the same distribution as Y. Let X_k be the sample mean

$$\frac{1}{k}\sum_{i=1}^{k} Y_i.$$

Show that X_k is exactly the same thing as $\tilde{p}_k(0)$. (This should be very easy.)

(d) Use the fact that expectation values are linear to show $\mathbb{E}(X_k) = \mathbb{E}(\tilde{p}_k(0)) = p(0)$. (In the language of probability theory, this shows that $\tilde{p}_k(0)$ is an "unbiased estimator" of the true probability $p(0)$.)

(e) Since the Y_i are independent, the variance of their sum is the sum of their variances. Use this to show $\text{var}(X_k) = \frac{1}{k}\text{var}(Y)$.

(f) Now use Chebyshev's inequality to argue that we should take $k \geq \frac{1}{4\epsilon^2\delta}$.

(g) How big should k be if we want to be 95% confident that our estimate of $|z_0|^2$ is correct up to b bits?

Let me conclude by noting that there are better ways to do quantum state tomography!

Proof. (a) The expectation values of the random variables Y and Y^2 are, respectively,

$$\mathbb{E}(Y) = \sum_{i=1}^{2} ip(i) = (0)\cdot p(0) + (1)\cdot p(1) = |z_0|^2$$
$$\mathbb{E}\left(Y^2\right) = \sum_{i=1}^{2} i^2 p(i) = (0)^2\cdot p(0) + (1)^2\cdot p(1) = |z_0|^2.$$

The variance of a random variable is given by

$$\text{Var}(Y) = \mathbb{E}\left(Y^2\right) - \mathbb{E}(Y)^2 = |z_0|^2 - |z_0|^4 = |z_0|^2(1-|z_0|^2).$$

(b) To find the maximum value for $p(1-p)$, where $p \in [0,1]$, we can search for a maximum on that domain. Let $f(p) = p(1-p)$, then $f'(p) = 1-2p$. Setting $f'(p) = 0$, we get that the $p = \frac{1}{2}$. Plugging this back, we get $f\left(\frac{1}{2}\right) = \frac{1}{4}$. Thus, the maximum value for $f(p)$ is $\frac{1}{4}$ when $p = \frac{1}{2}$.

Let $p = |z_0|^2$, then $\text{Var}(Y) = p(1-p) = |z_0|^2(1-|z_0|^2) \leq \frac{1}{4}$.

(c) We have that

$$\tilde{p}_k(0) = \frac{\#\{j|i_j = 0\}}{k}.$$

The variables Y_j and i_j are equivalent, where $i_j = 0 \iff Y_j = 1$ and $i_j \neq 1 \iff Y_j = 0$.

Then,

$$X_k = \frac{1}{k}\sum_{i=1}^{k} Y_i = \frac{\#\{i|Y_i = 1\}}{k} = \frac{\#\{j|i_j = 0\}}{k} = \tilde{p}_k(0).$$

(d) We have

$$\begin{aligned}\mathbb{E}(X_k) &= \mathbb{E}(\tilde{p}_k(0)) \\ &= \frac{1}{k}\sum_{i=1}^{k}\mathbb{E}(Y_i) \\ &= \frac{1}{k}\sum_{i=1}^{k}|z_0|^2 \\ &= |z_0|^2 \\ &= p(0).\end{aligned}$$

(e) We have

$$\begin{aligned}\mathrm{Var}(X_k) &= \mathrm{Var}\left(\frac{1}{k}\sum_{i=1}^{k}Y_i\right) \\ &= \frac{1}{k^2}\mathrm{Var}\left(\sum_{i=1}^{k}Y_i\right) \\ &= \frac{1}{k^2}\sum_{i=1}^{k}\mathrm{Var}(Y_i) \\ &= \frac{1}{k^2}k\mathrm{Var}(Y) \\ &= \frac{1}{k}\mathrm{Var}(Y).\end{aligned}$$

(f) Chebyshev's inequality states that the probability that a random variable X deviates from its mean μ by more than $k\sigma$ is at most $\frac{1}{k^2}$, where k is any positive constant and σ is the standard deviation. It is expressed as

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

We want to reach

$$\mathrm{Prob}(|\tilde{p}_k(0) - p(0)| \geq \epsilon) \leq \delta.$$

In our case, we have

$$\begin{aligned}P(|X_k - \mathbb{E}(X_k)| \geq k\sigma) &\leq \frac{1}{k^2}, \\ P(|\tilde{p}_k(0) - p(0)| \geq \epsilon) &\leq \left(\frac{\sigma}{\epsilon}\right)^2 \qquad \left(k = \frac{\epsilon}{\sigma}\right).\end{aligned}$$

From part (b), we have that

$$\sigma^2 = \mathrm{Var}(X_k) = \frac{1}{k}\mathrm{Var}(Y) \leq \frac{1}{4k}$$

Thus, we need

$$\begin{aligned}\left(\frac{\sigma}{\epsilon}\right)^2 &\leq \frac{1}{4k\epsilon^2} \leq \delta \\ \implies k &\geq \frac{1}{4\delta\epsilon^2}.\end{aligned}$$

(g) For a 95% confidence, we need $\delta = 0.95$, and up to b bits means we need $\epsilon = \frac{1}{2^b}$. Replacing, we get

$$k \geq \frac{1}{4(0.95)2^{-2b}} = \frac{5}{19}\cdot 4^b.$$

■

Homework 3

Do the following exercises from Nielsen and Chuang: 4.6, 4.11, 4.12, 4.17, 4.34, 4.35, 4.38, 4.39. For 4.17 and 4.39, just draw your answer, you do not need to justify it.

Problem 4.6
(Bloch sphere interpretation of rotations) One reason why the $R_{\hat{n}}(\theta)$ operators are referred to as rotation operators is the following fact, which you are to prove. Suppose a single qubit has a state represented by the Bloch vector $\vec{\lambda}$. Then the effect of the rotation $R_{\hat{n}}(\theta)$ on the state is to rotate it by an angle θ about the $\hat{n}$ axis of the Bloch sphere. This fact explains the rather mysterious looking factor of two in the definition of the rotation matrices.

Proof. Suppose a single qubit has a state represented by an arbitrary Bloch vector $\vec{\lambda}$. Without loss of generality, we can express $\vec{\lambda}$ in a coordinate system such that $\hat{n}$ is aligned with the $\hat{z}$ axis, so it suffices to consider how the state behaves under application $R_{\hat{z}}(\theta)$. Let $\vec{\lambda} = (\lambda_x, \lambda_y, \lambda_z)$ be the vector expressed in this coordinate system. By Exercise 2.72, the density operator ρ corresponding to this Bloch vector is given by:

$$\rho = \frac{\mathbb{I} + \vec{\lambda} \cdot \vec{\sigma}}{2}.$$

Observing how ρ transforms under conjugation by $R_{\hat{z}}(\theta)$, we have

$$\begin{aligned} R_{\hat{z}}(\theta)\rho R_{\hat{z}}(\theta)^\dagger &= R_{\hat{z}}(\theta)\rho R_{\hat{z}}(-\theta) \\ &= R_{\hat{z}}(\theta)\left(\frac{\mathbb{I} + \vec{\lambda} \cdot \vec{\sigma}}{2}\right) R_{\hat{z}}(-\theta) \\ &= R_{\hat{z}}(\theta)\left(\frac{\mathbb{I} + \lambda_x\sigma_x + \lambda_y\sigma_y + \lambda_z\sigma_z}{2}\right) R_{\hat{z}}(-\theta). \end{aligned}$$

Using $\sigma_j\sigma_k = \delta_{jk} + i\sum_l \epsilon_{jkl}\sigma_l$ and $\sigma_j\sigma_k = -\sigma_k\sigma_j$, we have

$$\begin{aligned} R_{\hat{z}}(\theta)\sigma_x &= \left(\cos\left(\frac{\theta}{2}\right)\mathbb{I} - i\sin\left(\frac{\theta}{2}\right)\sigma_z\right)\sigma_x \\ &= \cos\left(\frac{\theta}{2}\right)\mathbb{I}\sigma_x - i\sin\left(\frac{\theta}{2}\right)\sigma_z\sigma_x \\ &= \cos\left(\frac{\theta}{2}\right)\sigma_x\mathbb{I} + i\sin\left(\frac{\theta}{2}\right)\sigma_x\sigma_z \\ &= \sigma_x\left(\cos\left(\frac{\theta}{2}\right)\mathbb{I} + i\sin\left(\frac{\theta}{2}\right)\sigma_z\right) \\ &= \sigma_x\left(\cos\left(-\frac{\theta}{2}\right)\mathbb{I} - i\sin\left(-\frac{\theta}{2}\right)\sigma_z\right) \\ &= \sigma_x R_{\hat{z}}(-\theta). \end{aligned}$$

Similarly, $R_{\hat{z}}(\theta)\sigma_y = \sigma_y R_{\hat{z}}(-\theta)$ and $R_{\hat{z}}(\theta)\sigma_z = \sigma_z R_{\hat{z}}(\theta)$. Then, we have

$$\begin{aligned} R_{\hat{z}}(\theta)\rho R_{\hat{z}}(\theta)^\dagger &= R_{\hat{z}}(\theta)\left(\frac{\mathbb{I} + \lambda_x\sigma_x + \lambda_y\sigma_y + \lambda_z\sigma_z}{2}\right) R_{\hat{z}}(-\theta) \\ &= \left(\frac{\mathbb{I}R_{\hat{z}}(\theta) + \lambda_x\sigma_x R_{\hat{z}}(-\theta) + \lambda_y\sigma_y R_{\hat{z}}(-\theta) + \lambda_z\sigma_z R_{\hat{z}}(\theta)}{2}\right) R_{\hat{z}}(-\theta) \\ &= \frac{\mathbb{I} + \lambda_x\sigma_x R_{\hat{z}}(-2\theta) + \lambda_y\sigma_y R_{\hat{z}}(-2\theta) + \lambda_z\sigma_z}{2}. \end{aligned}$$

By term-by-term calculation, we have

$$\begin{aligned}\sigma_x R_{\hat{z}}(-2\theta) &= \sigma_x \left(\cos\left(-\frac{2\theta}{2}\right) - i\sin\left(-\frac{2\theta}{2}\right)\sigma_z\right)\\ &= \sigma_x\left(\cos(\theta) + i\sin(\theta)\sigma_z\right)\\ &= \cos(\theta)\sigma_x + i\sin(\theta)\sigma_x\sigma_z\\ &= \cos(\theta)\sigma_x + i\sin(\theta)(-i\sigma_y)\\ &= \cos(\theta)\sigma_x + \sin(\theta)\sigma_y,\end{aligned}$$

$$\begin{aligned}\sigma_y R_{\hat{z}}(-2\theta) &= \sigma_y \left(\cos\left(-\frac{2\theta}{2}\right) - i\sin\left(-\frac{2\theta}{2}\right)\sigma_z\right)\\ &= \sigma_y\left(\cos(\theta) + i\sin(\theta)\sigma_z\right)\\ &= \cos(\theta)\sigma_y + i\sin(\theta)\sigma_y\sigma_z\\ &= \cos(\theta)\sigma_y + i\sin(\theta)(i\sigma_x)\\ &= \cos(\theta)\sigma_y - \sin(\theta)\sigma_x,\end{aligned}$$

and substituting in the inital expression, we get

$$\begin{aligned}R_{\hat{z}}(\theta)\rho R_{\hat{z}}(\theta)^\dagger &= \frac{\mathbb{I} + \lambda_x\sigma_x R_{\hat{z}}(-2\theta) + \lambda_y\sigma_y R_{\hat{z}}(-2\theta) + \lambda_z\sigma_z}{2}\\ &= \frac{\mathbb{I} + \lambda_x\left(\cos(\theta)\sigma_x + \sin(\theta)\sigma_y\right) + \lambda_y\left(\cos(\theta)\sigma_y - \sin(\theta)\sigma_x\right) + \lambda_z\sigma_z}{2}\\ &= \frac{\mathbb{I} + \left(\lambda_x\cos(\theta) - \lambda_y\sin(\theta)\right)\sigma_x + \left(\lambda_x\sin(\theta) + \lambda_y\cos(\theta)\right)\sigma_y + \lambda_z\sigma_z}{2}.\end{aligned}$$

From this, the new Bloch vector $\vec{\lambda}'$, after conjugation by $R_{\hat{z}}(\theta)$ is expressed as

$$\vec{\lambda}' = (\lambda_x\cos(\theta) - \lambda_y\sin(\theta), \lambda_x\sin(\theta) + \lambda_y\cos(\theta), \lambda_z).$$

Notice that

$$\vec{\lambda}' = \begin{pmatrix}\cos(\theta) & -\sin(\theta) & 0\\ \sin(\theta) & \cos(\theta) & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\lambda_x\\ \lambda_y\\ \lambda_z\end{pmatrix},$$

where the matrix is the 3-dimensional rotation matrix about $\hat{z}$ by an angle of θ.
Thus, the conjugation of ρ under $R_{\hat{z}}(\theta)$ has the equivalent effect to rotating the Bloch vector by θ about the z-axis, and hence, the effect of $R_{\hat{n}}(\theta)$ on a one qubit state is to rotate it by an angle θ about $\hat{n}$. ■

Problem 4.12
Give A, B, C, and α for the Hadamard gate

Proof. Since the Hadamard gate H is a unitary gate on a single qubit, then there exist unitary operators A, B, C on a single qubit such that $ABC = \mathbb{I}$ and $U = e^{i\alpha}AXBXC$, where α is some overall phase factor.

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = e^{i\frac{\pi}{2}}R_z(\pi)R_y\left(-\frac{\pi}{2}\right)R_z(0).$$

Thus,

$$\begin{aligned}A &= R_z(\pi)R_y\left(-\frac{\pi}{4}\right)\\ B &= R_y\left(\frac{\pi}{4}\right)R_z\left(-\frac{\pi}{2}\right)\\ C &= R_z\left(-\frac{\pi}{2}\right)\\ \alpha &= \frac{\pi}{2}.\end{aligned}$$

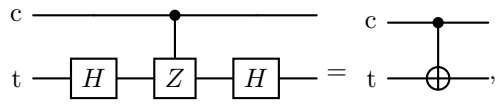
■

Problem 4.17
(Building CNOT from controlled-Z gates) Construct a CNOT gate from one controlled-Z gate, that is, the gate whose action in the computational basis is specified by the unitary matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

and two Hadamard gates, specifying the control and target qubits.

Proof. From Exercise 4.13, we have that $HZH = X$. To obtain a CNOT gate from a single controlled-Z gate, we can conjugate the target qubit with Hadamard gates

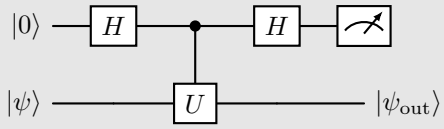


which is

$$\begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & Z \end{pmatrix} \begin{pmatrix} H & 0 \\ 0 & H \end{pmatrix} = \begin{pmatrix} HH & 0 \\ 0 & HZH \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & X \end{pmatrix} = CNOT.$$

■

Problem 4.34
(Measuring an operator) Suppose we have a single qubit operator U with eigenvalues ± 1, so that U is both Hermitian and unitary, so it can be regarded both as an observable and a quantum gate. Suppose we wish to measure the observable U. That is, we desire to obtain a measurement result indicating one of the two eigenvalues, and leaving a post-measurement state which is the corresponding eigenvector. How can this be implemented by a quantum circuit? Show that the following circuit implements a measurement of U:



Proof. We can obtain a measurement result indicating one of the two eigenvalues, while leaving a post-measurement state which is the corresponding eigenvector by using a controlled gate to entangle the system to a qubit whose measurement will collapse the state into +1 or -1, while also giving us the state of the original qubit. Additionally, since U is both Hermitian and unitary, then it is also an involutory matrix, i.e. $1 = U^\dagger U = UU = U^2$. The circuit will execute as follows

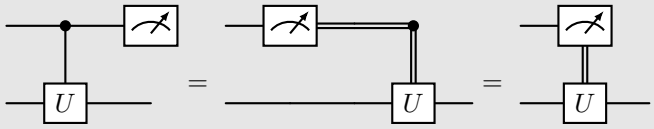
$$\begin{aligned} |0\rangle |\psi\rangle &\to^H \frac{1}{\sqrt{2}}(|0\rangle |\psi\rangle + |1\rangle |\psi\rangle) \\ &\to^{CU} \frac{1}{\sqrt{2}}(|0\rangle |\psi\rangle + |1\rangle U |\psi\rangle) \\ &\to^H \frac{1}{2}\left[|0\rangle |\psi\rangle + |1\rangle |\psi\rangle + |0\rangle U |\psi\rangle + |1\rangle U |\psi\rangle\right] \\ &= \frac{1}{2}\left[|0\rangle (\mathbb{I} + U) |\psi\rangle + |1\rangle (\mathbb{I} - U) |\psi\rangle\right]. \end{aligned}$$

- If the measurement value is $|0\rangle$, then the state is in $|\psi_{out}\rangle = (\mathbb{I}+U)|\psi\rangle$, where $U|\psi_{out}\rangle = U(\mathbb{I}+U)|\psi\rangle = (\mathbb{I}+U)|\psi\rangle$, and thus has an eigenvalue of +1.

- If the measurement value is $|1\rangle$, then the state is in $|\psi_{out}\rangle = (\mathbb{I}-U)|\psi\rangle$, where $U|\psi_{out}\rangle = U(\mathbb{I}-U)|\psi\rangle = -(\mathbb{I}-U)|\psi\rangle$, and thus has an eigenvalue of -1.

■

Problem 4.35
(Measurement commutes with controls) A consequence of the principle of deferred measurement is that measurements commute with quantum gates when the qubit being measured is a control qubit, that is:



(Recall that the double lines represent classical bits in this diagram.) Prove the first equality. The rightmost circuit is simply a convenient notation to depict the use of a measurement result to classically control a quantum gate.

Proof. Let the system be in the state $a|0\rangle|\psi\rangle + b|1\rangle|\psi\rangle$. Then the effect of the circuits are

- **Circuit 1:**

$$\begin{aligned} a|0\rangle|\psi\rangle + b|1\rangle|\psi\rangle &\to^{CU} a|0\rangle|\psi\rangle + b|1\rangle U|\psi\rangle \\ &\to^{M} \begin{cases} |0\rangle, & \text{with } p = |a|^2 \text{ and state } |\psi\rangle, \\ |1\rangle, & \text{with } p = |b|^2 \text{ and state } U|\psi\rangle \end{cases} \end{aligned}$$

- **Circuit 2:**

$$\begin{aligned} a|0\rangle|\psi\rangle + b|1\rangle|\psi\rangle &\to^{M} \begin{cases} |0\rangle, & \text{with } p = |a|^2 \text{ and state } |\psi\rangle, \\ |1\rangle, & \text{with } p = |b|^2 \text{ and state } |\psi\rangle \end{cases} \\ &\to^{CU} \begin{cases} |0\rangle, & \text{with } p = |a|^2 \text{ and state } |\psi\rangle, \\ |1\rangle, & \text{with } p = |b|^2 \text{ and state } U|\psi\rangle \end{cases} \end{aligned}$$

■

Problem 4.38
Prove that there exists a $d \times d$ unitary matrix U which cannot be decomposed as a product of fewer than $d-1$ two-level unitary matrices.

Proof. Suppose U is a $d \times d$ unitary matrix which can be decomposed using less than $d-1$ two-level unitaries. We can think of each two-level unitary as an "edge" linking some pair of nodes $|i\rangle$ and $|j\rangle$, interpreting each node as a vertex. Let $U = U_{d-1}U_{d-2}\cdots U_2U_1$, where U_k is a two-level unitary. The graph corresponding to U has at most $d-1$ edges corresponding to the U_k operators, but we have d vertices, then there must be two subsets of nodes on which U acts independently. Hence, U must be block diagonal in some rearrangement of the initial basis. A one-component graph would require more than $d-1$ edges. A non-block diagonal operator cannot be written with less than d operators U_k.
Clearly not every U has this form. To name one example, the Quantum Fourier Transform matrix doesn't. Thus, by contradiction, there exists a $d \times d$ matrix U which cannot be decomposed as a product of fewer than $d-1$ two-level unitary matrices.

■

Problem 4.39
Find a quantum circuit using single qubit operations and CNOTs to implement the transformation

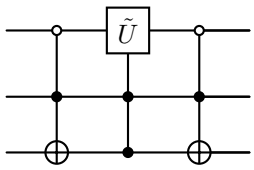
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & c \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & d \end{bmatrix}$$

where $\tilde{U} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ is an arbitrary 2×2 unitary matrix.

Proof. From the entries of the matrix, we can see that it acts non-trivially on the states $|010\rangle$ and $|111\rangle$. We write a Gray code connecting 010 and 111:

$$\begin{array}{c} ABC \\ 0\,1\,0 \\ 0\,1\,1 \\ 1\,1\,1 \end{array}$$

From this we read off the required circuit to be



■

Problem 2
Give an example of a unitary 2-qubit gate $U : \left(\mathbb{C}^2\right)^{\otimes 2} \to \left(\mathbb{C}^2\right)^{\otimes 2}$ that is "entangling," that is, can not be expressed as a tensor product $U_1 \otimes U_2$ where U_1 and U_2 are two 1-qubit gates $U_1, U_2 : \mathbb{C}^2 \to \mathbb{C}^2$. Justify your example.

Proof. An example of a unitary 2-qubit gate that is "entangling" is the CNOT gate, given by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & X \end{pmatrix}.$$

This gate is not separable (entangling) as it cannot be written as the tensor product of two matrices. In fact, if it separable, then

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} \implies \begin{cases} \mathbb{I}_2 = a_{11}B \\ X = a_{22}B \end{cases} \implies \mathbb{I}_2 = bX,$$

where $b = \frac{a_{11}}{a_{22}}$ is some scalar, which is a contradiction.
Therefore, CNOT is an entangling gate. ■

Homework 4

Only problems 1 and 7 will be graded for correctness. The other problems will be graded for completeness (that is, you have to make an "honest attempt" to solve them).

Problem 1
For this problem, you will need to use the definition of $BQP(\mathcal{G}, \delta)$ I gave in the lecture notes for lecture 4.2. Other definitions are equivalent to mine, but the problems I have written here are closely tied to my specific definition.

(a) In the lecture notes for lecture 4.2, I define $BQ(\mathcal{G}, \delta)$ to be like BQP, except we don't require there be any algorithm at all (much less a polynomial time one) to identify the quantum circuit C_x. Show that $BQ(\mathcal{G}, 0) = ALL$ if $\mathcal{G}$ is universal. Deduce that $BQ(\mathcal{G}, \delta) = ALL$ for any $0 < \delta < 1$ if $\mathcal{G}$ is universal. [Hint: your answer shouldn't need to be longer than one paragraph.]

(b) Show that if $\mathcal{G}$ is universal and $1/2 \leq \delta \leq 1$, then $BQP(\mathcal{G}, \delta) = ALL$. [Hint: your answer should only need to be two or three sentences.]

(c) Suppose, as I suggest but didn't say precisely in the notes, that we modify the definition of $BQP(\mathcal{G}, \delta)$ to the following:

> $L \in BQP(\mathcal{G}, \delta)$ if there exists a classical (deterministic) polynomial-time algorithm which, for each integer $n \geq 1$, outputs a description of a quantum circuit C_n over $\mathcal{G}$ such that, for all bit strings $x \in \{0,1\}^n$, measuring the first qubit $C_n |x0\cdots0\rangle$ in the computational basis satisfies
>
> $$prob\left(\text{Output}(C_n) = L(x) \mid |x0\cdots0\rangle\right) \geq 1 - \delta.$$

Show that these two definitions give equal complexity classes. [Hint: as suggested in class, the distinction is simply about whether or not we "hard-code" the value of x into the circuit. Your answer shouldn't need to be more than a few sentences and a couple pictures.]

(d) (**Extra credit**) Suppose we let $\delta = 1/2$ and modify the definition of $BQP(\mathcal{G}, \delta)$ so that the greater-than-or-equal to sign "$\geq$" is now a strict inequality "$>$". Is this a "reasonable" complexity class? Do there exist uncomputable problems in it?

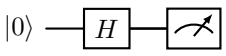
Proof. (a) Considering $\text{BQ}(\mathcal{G}, \delta)$ is equivalent to $\text{BQP}(\mathcal{G}, \delta)$, except that we remove the condition for our quantum circuit C_x over $\mathcal{G}$ to be prepared algorithmically. Now consider $\text{BQ}(\mathcal{G}, 0)$, then

$$\mathbb{P}(\text{OUTPUT}(C_x) = L(x)) \geq 1 = 1.$$

Thus, $\text{BQ}(\mathcal{G}, 0) = \text{ALL}$. Of course, here we are assuming that $\mathcal{G}$ is universal to be able to output a description of C_x.

Since, by definition, $\text{BQ}(\mathcal{G}, 0) \subseteq \text{BQ}(\mathcal{G}, \delta)$, then $\text{BQ}(\mathcal{G}, \delta) = \text{ALL}$.

(b) For any decision $L(x)$, let the algorithmically produced quantum circuit be the simple Hadamard gate,



Then, we have

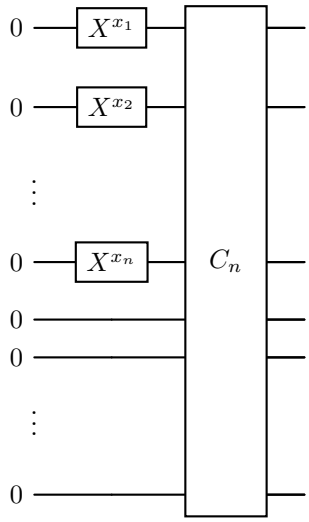
$$\mathbb{P}(\text{OUTPUT}(C_x) = L(x)) = \frac{1}{2} \geq 1 - \delta.$$

If $\frac{1}{2} \leq \delta \leq 1$, then $L \in \text{BQP}(\mathcal{G}, \delta)$. Since this works of all L, then $\text{BQP}(\mathcal{G}, \delta) = \text{ALL}$.

(c) Call the definition from the lecture notes as Def_1 and the definition in the problem as Def_2.

$\boxed{\text{Def}_2 \implies \text{Def}_1}$

Suppose we have a circuit C_n for all bit-strings x of length n with input $|x0\cdots0\rangle$. Our goal is to find C_x. Let $x = x_1x_2\cdots x_n$, then C_x is given by



■

Problem 2
Let $x \in \{0,1\}^n$ be a bit string of length n. Show that

$$H^{\otimes n}|x\rangle = \frac{1}{2^{\frac{n}{2}}} \sum_{y\in\{0,1\}^n} (-1)^{x\cdot y}|y\rangle .$$

Proof. We know that

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \qquad \text{and} \qquad H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

We have

$$\begin{aligned}
H^{\otimes n}|x\rangle &= H^{\otimes n}|x_0x_1\cdots x_n\rangle \\
&= H|x_0\rangle \otimes H|x_1\rangle \otimes \cdots \otimes H|x_{n-1}\rangle \\
&= \bigotimes_{i=0}^{2^n-1} H|x_i\rangle \\
&= \bigotimes_{i=0}^{2^n-1} \frac{|0\rangle + (-1)^{x_i}|1\rangle}{\sqrt{2}} \\
&= \frac{1}{\sqrt{2^n}} \bigotimes_{i=0}^{2^n-1} |0\rangle + (-1)^{x_i}|1\rangle \\
&= \frac{1}{\sqrt{2^n}} \left[(-1)^{x_0}|00\cdots0\rangle + (-1)^{x_{n-1}}|00\cdots1\rangle + \cdots + (-1)^{x_0\cdot x_1\cdots x_{n-1}}|11\cdots1\rangle\right] \\
&= \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} (-1)^{x\cdot y}|y\rangle ,
\end{aligned}$$

where $x\cdot y = x_0y_0 \oplus x_1y_1 \oplus \cdots \oplus x_{n-1}y_{n-1}$. ■

Problem 1.1
(Probabilistic classical algorithm) Suppose that the problem is not to distinguish between the constant and balanced functions with certainty, but rather, with some probability of error $< 1/2$. What is the performance of the best classical algorithm for this problem?

Proof. A Boolean function $f : \{0,1\}^n \to \{0,1\}$ is said to be balanced if $f(x) = 1$ for exactly half of all possible 2^n values of x.
Evaluating only once does not give us any information about whether or not f is constant or balanced, so our success rate after a single evaluation is $\epsilon = \frac{1}{2}$, since it is a random guess.
Consider now the case where we do two evaluations:

- If we obtain two different results, then we immediately conclude that f is balanced.
- If we obtain two results that are the same, then if f is balanced, the probability that the first evaluation returned the given result is $\frac{1}{2}$, and the probability that the second evaluation returned the same result as the first is $\frac{2^{n-1}-1}{2^n-1}$ (as there are 2^{n-1} of each result of 0 and 1, 2^n total results, $2^{n-1}-1$ of the given result left after the first evaluation, and $2^n - 1$ total uninvestigated cases after the first evaluation).

Therefore, if f is balanced, this occurs with probability $\frac{1}{2}\frac{2^{n-1}-1}{2^n-1}$, which we can see is

$$2^{n-1} < 2^n \implies \frac{2^{n-1}-1}{2^n-1} < 1 \implies \frac{1}{2}\left(\frac{2^{n-1}-1}{2^n-1}\right) < \frac{1}{2}.$$

Thus, if we get the same result in two evaluations, we can conclude that f is constant with error $\epsilon < \frac{1}{2}$.
Therefore, only 2 evaluations are required for this algorithm. ■

Problem 4
Suppose Alice has a state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$. Show that if she and Bob share n Bell pairs $|\beta_{00}\rangle$, then to teleport her state $|\psi\rangle$ to Bob, it suffices to simply teleport each qubit in her state to Bob one at a time using the single qubit teleportation protocol. [Hint: induction. $n = 2$ is the most interesting case.]

Proof. Denote Alice's state as

$$|\psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle ,$$

where α_i are complex amplitudes and $|i\rangle$ represents the computational basis states of n qubits. The single qubit teleportation protocol works as follows:

1. Alice performs a Bell measurement on the qubit she wants to teleport and her corresponding qubit from the shared Bell pair.
2. Alice sends the classical result of the measurement to Bob.
3. Depending on the measurement outcome, Bob applies a set of unitary operations (identity, X, Z, or XZ) to his remaining qubit to recover the state $|\psi\rangle$.

Now, let's proceed with the proof by induction.

- **Base case:** When $n = 1$, Alice has a single qubit state $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$. Teleporting this state using the single qubit teleportation protocol is straightforward and well-known.
- **Inductive hypothesis:** Assume that for some $k \geq 1$, teleporting an k-qubit state using k shared Bell pairs via the single qubit teleportation protocol is successful.
- **Inductive step:** Consider the case when $n = k + 1$. We want to show that if Alice and Bob share $k + 1$ Bell pairs, Alice can teleport her $(k + 1)$-qubit state to Bob using the single qubit teleportation protocol applied sequentially.

- **Proof:** Alice can isolate the first qubit of her state $|\psi\rangle$ and the first qubit from the first Bell pair, then perform a Bell measurement and send the result to Bob. Bob receives this classical information and performs the necessary unitary operations on his remaining qubits to obtain the state of Alice's first qubit. Now, Alice and Bob are left with k qubits each. By the inductive hypothesis, Alice can use the remaining k Bell pairs to teleport her remaining k-qubit state to Bob's k qubits one at a time. Thus, the claim holds for all n.

Therefore, to teleport Alice's n-qubit state $|\psi\rangle$ to Bob, it suffices to teleport each qubit in her state to Bob one at a time using the single qubit teleportation protocol. ■

Problem 6
I said something slightly misleading/incomplete in class during lecture 5.2 when I sketched the idea of the proof that BQP is in $PSPACE$. Recall that I proved a lemma: for any $\epsilon > 0$, any bit strings $x, y \in \{0,1\}^n$ and any quantum circuit C on n qubits (over some fixed gate set $\mathcal{G}$), we may compute an approximation to the amplitude $\langle y|C|x\rangle$ in $PSPACE$ in the size of the circuit C.
After that, I very hastily explained that from here, we could, in $PSPACE$, decide whether or not the probability that the first qubit of the output of $C|0\cdots 0\rangle$ returns 0 is $< 1/3$ or $\geq 2/3$, and thus, decide whether the circuit is answering YES or NO.
I had insinuated that we only needed the case $\epsilon = 1/3$ in the lemma, but this is not true strictly speaking as I explained it (but see problem 7(b) below, in which case it is!). We need to be able to compute these amplitudes to precision $\frac{1}{2^n}$ for the argument that $BQP \subseteq PSPACE$ to work. Note: The amplitudes in the computational basis of course in principal determine the *marginal* probability that the first bit returns, say, 1, but we need to know the amplitudes to exponential precision if we want to know this marginal probability to $O(1)$ precision.
With this in mind, prove the following: for any quantum circuit C on n qubits (over some fixed gate set $\mathcal{G}$) and any two bit strings $x, y \in \{0,1\}^n$, we may compute a complex number z such that

$$|z - \langle y|C|x\rangle| < \frac{1}{2^n}$$

in $PSPACE$ (as a function of the size of C).

Proof. The quantum circuit C is related to a decision problem in BQP, then the size of C is a polynomial. Suppose $C = U_1 U_2 \cdots U_{p(n)}$, then

$$\langle y|C|x\rangle = \sum_{x_1,\ldots,x_{p(n)-1}} \langle y|U_{p(n)}|x_{p(n)-1}\rangle \langle x_{p(n)-1}|U_{p(n)-2}\rangle \cdots \langle U_2|x_1\rangle \langle x_1|U_1|x\rangle .$$

We want to compute an approximation of z within an error of $\frac{1}{2^n}$. Suppose every matrix element is represented with an accuracy of $\frac{1}{2^m}$. Moreover, suppose we have an error in the multiplication of two terms (rounding error) which is $O\left(\frac{1}{2^m}\right)$. We have the multiplication of $p(n)$ terms, resulting in a total error of $\frac{p(n)}{2^m}$. Computing the sum means adding these numbers, resulting in $(2^n)^{p(n)}$ total addition operations. Suppose, again, that each addition operation results in an error $\frac{1}{2^m}$.

- If addition is perfect and multiplication is error-prone, then the total error is $\frac{p(n)}{2^m}$. Then, we need to have

$$\begin{aligned} \frac{p(n)}{2^m} &< \frac{1}{2^n} \\ 2^m &> p(n)2^n \\ m &> n + \log(p(n)) \\ \implies m &\in \Theta(n). \end{aligned}$$

- If multiplication is perfect and addition is error-prone, then the total error is $\frac{(2^n)^{p(n)}}{2^m}$. Then, we need to have

$$\begin{aligned}\frac{(2^n)^{p(n)}}{2^m} &< \frac{1}{2^n}\\ 2^m &> (2^n)^{p(n)+1}\\ m &> n(p(n)+1)\\ \implies m &\in \Theta(q(n)),\end{aligned}$$

  where $q(n)$ is a polynomial in n.

Since m represents the number of bits needed to store z, then we need a polynomial space in n to approximate $\langle y|C|x\rangle$ to n bits.
Therefore, we may compute a complex number z such that

$$|z - \langle y|C|x\rangle| < \frac{1}{2^n}$$

in $PSPACE$. ■

Problem 7
In this problem we will explore some examples of "BQP-universal" problems.

(a) Suppose you had the power to decide the following problem: given a description of a quantum circuit C on n qubits, decide if the probability that C outputs 1 in its first qubit when input the basis state $|0\cdots 0\rangle$ is greater than or equal to 2/3.

Show that you could use your power (together with classical polynomial time effort) to solve every problem in BQP. (This should only take one short paragraph to explain. It should follow essentially from the definition of BQP).

Proof. Assume $L \in BQP(\mathcal{G}, \delta)$, then we can algorithmically generate a quantum circuit C_x such that

$$\mathbb{P}(\text{OUTPUT}(C_x) = L(x) \mid |00\cdots 0\rangle) \geq 1 - \delta,$$

where $0 < \delta < \frac{1}{2}$.
Having the freedom to choose δ, take $\delta = \frac{1}{3}$, and now we have $L(x) = 0$ or $L(x) = 1$. It must be that if

$$\mathbb{P}(\text{OUTPUT}(C_x) = 1) \geq \frac{2}{3},$$

then $L(x) = 1$. Otherwise, $L(x) = 0$.
In fact, we have

- If $\mathbb{P}(\text{OUTPUT}(C_x) = 1) \geq \frac{2}{3}$ and $L(x) = 0$, then it means that

$$\mathbb{P}(\text{OUTPUT}(C_x) = L(x)) < \frac{1}{3},$$

  which is a bad circuit to begin with.

- If $\mathbb{P}(\text{OUTPUT}(C_x) = 1) < \frac{2}{3}$ and $L(x) = 1$, then it means that

$$\mathbb{P}(\text{OUTPUT}(C_x) = L(x)) < \frac{2}{3},$$

  which is also a contradiction.

■

Homework 5

Problem 4.50
Let $H = \sum_k^L H_k$, and define

$$U_{\Delta t} = \left[e^{-iH_1\Delta t}e^{-iH_2\Delta t}\cdots e^{-iH_L\Delta t}\right]\left[e^{-iH_L\Delta t}e^{-iH_{L-1}\Delta t}\cdots e^{-iH_1\Delta t}\right].$$

(a) Prove that $U_{\Delta t} = e^{-2iH\Delta t} + O(\Delta t^3)$.

(b) Use the results in Box 4.1 to prove that, for a positive integer m,

$$E\left(U_{\Delta t}^m, e^{-2miH\Delta t}\right) \leq m\alpha\Delta t^3,$$

for some constant α.

Proof. (a) We have $H = \sum_k^L H_k$. The Baker-Campbell-Hausdorf formula states that

$$\mathrm{e}^{(A+B)\Delta t} = \mathrm{e}^{A\Delta t}\mathrm{e}^{B\Delta t}\mathrm{e}^{-\frac{1}{2}[A,B]\Delta t^2} + O(\Delta t^3).$$

By repeated application of the Baker-Campbell-Hausdorf formula, knowing that $[H_i, H_i] = 0$, we have

$$\begin{aligned}
U_{\Delta t} &= \left[e^{-iH_1\Delta t}\cdots e^{-iH_L\Delta t}\right]\left[e^{-iH_L\Delta t}\cdots e^{-iH_1\Delta t}\right] \\
&= \left[e^{-iH_1\Delta t}\cdots e^{-iH_{L-1}\Delta t}\right]e^{-2iH_L\Delta t}\left[e^{-iH_{L-1}\Delta t}\cdots e^{-iH_1\Delta t}\right] + O(\Delta t^3) \\
&= \left[e^{-iH_1\Delta t}\cdots e^{-iH_{L-2}\Delta t}\right]e^{-2i(H_L+H_{L-1})\Delta t}\left[e^{-iH_{L-2}\Delta t}\cdots e^{-iH_1\Delta t}\right] + O(\Delta t^3) \\
&= e^{-2iH\Delta t} + O(\Delta t^3).
\end{aligned}$$

(b) We have

$$\begin{aligned}
E\left(U_{\Delta t}^m, e^{-2miH\Delta t}\right) &\leq mE\left(U_{\Delta t}, e^{-2iH\Delta t}\right) \\
&= m||\left(U_{\Delta t} - e^{-2iH\Delta t}\right)|\psi\rangle\,|| \\
&= m||O(\Delta t^3)\,|\psi\rangle\,|| \\
&= m\alpha\Delta t^3.
\end{aligned}$$

■

Problem 4.1
(Computable phase shifts) Let m and n be positive integers. Suppose $f : \{0,\ldots,2m-1\} \to \{0,\ldots,2n-1\}$ is a classical function from m to n bits which may be computed reversibly using T Toffoli gates, as described in Section 3.2.5. That is, the function $(x, y) \to (x, y \oplus f(x))$ may be implemented using T Toffoli gates. Give a quantum circuit using $2T + n$ (or fewer) one, two, and three qubit gates to implement the unitary operation defined by

$$|x\rangle \to \exp\left(\frac{-2\mathrm{i}\pi f(x)}{2^n}\right)|x\rangle.$$

Proof. We first need to compute $\hat{P}\,|x, y\rangle \equiv |x, y \oplus f(x)\rangle$ using T Toffoli gates. Then we expand $q = \sum_{j=0}^{n-1} q_j 2^j$ with q_j being the digits of a binary representation of $0 \leq q \leq 2^{n-1}$. Observe that

$$\begin{aligned}
\hat{O}\,|x, q\rangle &\equiv \mathrm{e}^{\frac{2\pi\mathrm{i}yn}{2^n}}\prod_{j=0}^{n-1}\mathrm{e}^{-\pi\mathrm{i}q_j 2^{j-n+1}}\,|x, q\rangle \\
&= \mathrm{e}^{-\frac{2\pi\mathrm{i}}{2^n}\left(\sum_{j=0}^{n-1} q_j 2^j - y\right)}\,|x, q\rangle \\
&= \mathrm{e}^{-\frac{2\pi\mathrm{i}}{2^n}(q-y)}\,|x, q\rangle.
\end{aligned}$$

For $q = y \oplus f(x)$, we obtain

$$\hat{O}\,|x, y \oplus f(x)\rangle = \mathrm{e}^{-\frac{2\pi \mathrm{i}}{2^n} f(x)}\,|x, y \oplus f(x)\rangle\,.$$

The operation $\prod_{j=0}^{n-1} \mathrm{e}^{-\pi \mathrm{i} q_j 2^{j-n+1}}$ can be implemented using n single-qubit phase gates $\hat{U}_{n-1} \cdots \hat{U}_0$, each acting as

$$\hat{U}_j = \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{-\pi \mathrm{i} 2^{j-n+1}} \end{pmatrix},$$

while the y-dependent phase can be neglected.
Finally, we uncompute $P^\dagger\,|x, y \oplus f(x)\rangle = |x, y\rangle$ using T Toffoli gates.
The total cost of the operation $\hat{P}^\dagger \hat{O} \hat{P}$, and therefore requires $2T + n$ gates, as required. ■

Problem 3
Show that if A is a finite abelian group, then the dual group $\widehat{A} \cong A$. [Hint: do it in two steps. First, use the fact that every finite abelian group A is a direct sum of cyclic groups to reduce to the case that $A = \mathbb{Z}/N\mathbb{Z}$. Then argue however you want that $\widehat{\mathbb{Z}/N\mathbb{Z}}$ is a cyclic group of order N.]

Before I start my proof of this, I will state the definition of a **character** (since it wasn't done in class).

Definition 1. Let V be a finite-dimensional vector space over a field F and let $\rho : G \to \mathrm{GL}(V)$ be a representation of a group G on V. The **character** of ρ is the function $\chi_\rho : G \to F$ given by

$$\chi_\rho(g) = \mathrm{Tr}(\rho(g)),$$

where Tr is the trace.

Proof. This proposition is known as the self-duality of finite Abelian groups. The weak fundamental theorem of group theory states that, if G is a finite Abelian group, then either G is cyclic or $G \cong H \times K$ for non-trivial groups H and K. These two possibilities are not mutually exclusive, since, for example, $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$. Since G is a finite Abelian group, then $|G| < \infty$, on which we perform induction.

- For $|G| = 1$, this is trivially true.
- For $|G| > 1$, since G is a finite Abelian group, then either G is cyclic or $G \cong H \times K$ for non-trivial groups H and K
  - If G is a cyclic group, we need to prove that $G \cong \widehat{G}$. We can assume without loss of generality that $G = \mathbb{Z}/N\mathbb{Z}$. A character $\chi \in \widehat{G}$ is determined by what it does to the generator $1 \in \mathbb{Z}/N\mathbb{Z}$. Furthermore, since $\chi(N) = \chi(1)^N$ and $N = 0$ in $\mathbb{Z}/N\mathbb{Z}$, then $\chi(1)^N = 1$. In fact, choose a Nth root of unity and assign it to $\chi(1)$. Define $\mathrm{e}(t) = \mathrm{e}^{2\pi i t}$, then the Nth roots of unity are of the form $\mathrm{e}\left(\frac{a}{N}\right)$, with $a \in \mathbb{Z}/N\mathbb{Z}$. Define χ_a by

    $$\chi_a(x) = \mathrm{e}\left(\frac{ax}{N}\right).$$

    We can verify that χ_a is a character on $\mathbb{Z}/N\mathbb{Z}$, and that $\chi_a(1) = \mathrm{e}\left(\frac{a}{N}\right)$. Hence, the set $\{\chi_a : a \in \mathbb{Z}/N\mathbb{Z}\}$ is all of $\widehat{G}$. Finally, we can also verify that the map $a \mapsto \chi_a$ is a bijective homomorphism. Therefore, $G \cong \widehat{G}$
  - If $G = H \times K$, then let $\chi \in \widehat{G}$ and define $\nu \in H$ and $\eta \in K$ by $\nu(h) = \chi(h, 1)$ and $\eta(k) = \chi(1, k)$. Let $\nu \in H$ and $\eta \in K$. Define $\chi \in G$ by $\chi(h, k) = \nu(h)\eta(k)$. Then the maps $\chi \mapsto (\nu, \eta)$ and $(\nu, \eta) \mapsto \chi$ described are inverses of each other and are both homomorphisms. Thus, $G \cong H \times K \implies \widehat{G} \cong \widehat{H} \times \widehat{K}$. Since G is finite, then we have that $1 < |H|, |K| < |G|$, and by strong induction, we have $H \cong \widehat{H}$ and $K \cong \widehat{K}$. Thus $H \times K \cong \widehat{H} \times \widehat{K}$. Therefore, $G \cong \widehat{G}$.

■

Problem 5.6
(Approximate quantum Fourier transform) The quantum circuit construction of the quantum Fourier transform apparently requires gates of exponential precision in the number of qubits used. However, such precision is never required in any quantum circuit of polynomial size. For example, let U be the ideal quantum Fourier transform on n qubits, and V be the transform which results if the controlled-R_k gates are performed to a precision $\Delta = 1/p(n)$ for some polynomial $p(n)$. Show that the error $E(U,V) \equiv \max_{|\psi\rangle} ||(U-V)|\psi\rangle||$ scales as $\Theta(n^2/p(n))$, and thus polynomial precision in each gate is sufficient to guarantee polynomial accuracy in the output state.

Proof. In the quantum Fourier transform circuit of n qubits, we have $m = \frac{n(n+1)}{2} = \Theta(n^2)$ controlled-R_k gates. We are given that each approximation of the ideal transform is performed to a precision of $\Delta = \frac{1}{p(n)}$. Thus, $E(U,V) \leq m\Delta = \Theta\left(\frac{n^2}{p(n)}\right)$. ■

Problem 5.8
Suppose the phase estimation algorithm takes the state $|0\rangle|u\rangle$ to the state $|\widetilde{\varphi_u}\rangle|u\rangle$, so that given the input $|0\rangle\left(\sum_u c_u|u\rangle\right)$, the algorithm outputs $\sum_u c_u|\widetilde{\varphi_u}\rangle|u\rangle$. Show that if t is chosen according to (5.35), then the probability for measuring φ_u accurate to n bits at the conclusion of the phase estimation algorithm is at least $|c_u|^2(1-\epsilon)$.

Proof. We measure φ of the state $|u\rangle$ with a probability of $|c_u|^2$. If t is of the form of (5.35), then each $\widehat{\varphi}$ is accurate to φ up to n bits with probability $1-\epsilon$. Thus, the total probability of measuring φ up to n bits is $|c_u|^2(1-\epsilon)$. ■

ONLINE QUIZ

Homework 7

Problem 1
In this problem, you'll prove the two mathematical facts we needed to know in order for Simon's algorithm to work.

(a) Let A be a finite abelian group. Prove that if $g_1, \ldots, g_l$ are l independently and uniformly randomly chosen elements of A, then the probability that $\langle g_1, g_2, \ldots, g_l \rangle = A$ is at least $1 - \frac{|A|}{2^l}$. [Hint: as an intermediate step, you might use Lagrange's theorem to argue that the probability that $g_{i+1} \notin \langle g_1, \ldots, g_i \rangle$ is at least $1/2$ whenever $\langle g_1, \ldots, g_i \rangle \neq A$.]

(b) Let $A = (\mathbb{Z}/2\mathbb{Z})^n$ be an n dimensional vector space over $\mathbb{Z}/2\mathbb{Z}$. Let $s \in A$ be a non-zero element and suppose $g_1, \ldots, g_l \in A$ generate what I called $\langle s \rangle^\perp$, which is defined as

$$\langle s \rangle^\perp = \{a \in A \mid a \cdot s = 0 \bmod 2\},$$

where $a \cdot s$ is the modulo 2 dot product of $a = (a_1, \ldots, a_n)$ and $s = (s_1, \ldots, s_n)$. Prove that s is the unique non-zero solution to the system of equations

$$\begin{aligned} g_1 \cdot x &= 0 \mod 2 \\ g_2 \cdot x &= 0 \mod 2 \\ &\vdots \\ g_l \cdot x &= 0 \mod 2. \end{aligned}$$

Proof. (a) Let A be a finite abelian group and take $g_1, \ldots, g_l$ to be l independently and uniformly chosen elements of A. We will prove this by mathematical induction.

- **For $l = 1$:** We have only one element from A, namely g_1. Then $\mathbb{P}(\langle g_1 \rangle = A) = 1$ since we can construct the group from non-zero elements of the group.
- **For $l = k$:** We now make the inductive step and assume that, for $l = k \geq 1$,

$$\mathbb{P}(\langle g_1, g_2, \ldots, g_k \rangle = A) \geq 1 - \frac{|A|}{2^k}.$$

- **For $l = k + 1$:** We now need to show that

$$\mathbb{P}(\langle g_1, g_2, \ldots, g_k, g_{k+1} \rangle = A) \geq 1 - \frac{|A|}{2^{k+1}}.$$

From the inductive step, we know that $\mathbb{P}(\langle g_1, g_2, \ldots, g_k \rangle = A) \geq 1 - \frac{|A|}{2^k}$. If $\langle g_1, g_2, \ldots, g_k \rangle \neq A$, then $g_{k+1} \notin \langle g_1, g_2, \ldots, g_k \rangle$. Then $\mathbb{P}(g_{k+1} \notin \langle g_1, g_2, \ldots, g_k \rangle) = 1 - \frac{|\langle g_1, g_2, \ldots, g_k \rangle|}{|A|}$ (by Lagrange's Theorem). Thus

$$\begin{aligned} \mathbb{P}(\langle g_1, g_2, \ldots, g_k \rangle = A) &\geq \mathbb{P}(\langle g_1, g_2, \ldots, g_k \rangle = A) \times \mathbb{P}(g_{k+1} \notin \langle g_1, g_2, \ldots, g_k \rangle) \\ &\geq \left(1 - \frac{|A|}{2^k}\right)\left(1 - \frac{|\langle g_1, g_2, \ldots, g_k \rangle|}{|A|}\right) \\ &\geq \left(1 - \frac{|A|}{2^k}\right)\left(\frac{1}{2}\right) \\ &\geq 1 - \frac{|A|}{2^{k+1}}. \end{aligned}$$

Thus, with proof by induction, we have

$$\mathbb{P}(\langle g_1, g_2, \ldots, g_k \rangle = A) \geq 1 - \frac{|A|}{2^k}.$$

(b) For $1 \leq i \leq l$, s satisfies the system of equations given by $\langle s \rangle^{\perp} = \{g_i \in A \mid g_i \cdot s \equiv 0 \bmod 2\}$, where $g_i \in A$. Let x be a non-zero solution, then $x \cdot s \equiv 0 \bmod 2$ and $x \in \langle s \rangle^{\perp}$ is orthogonal to s. Any element in x can be expressed as a linear combination of the elements $g_1, g_2, \ldots, g_l$ as

$$x = \sum_{i=1}^{l} c_i g_i,$$

where $c_i \in \mathbb{Z}/2\mathbb{Z}$. We then have

$$\begin{aligned} x \cdot s &= \left(\sum_{i=1}^{l} c_i g_i \right) \cdot s \\ &= \sum_{i=1}^{l} c_i \left(g_i \cdot s \right) \\ &\equiv 0 \bmod 2. \end{aligned}$$

Thus, x satisfies the equation if and only if s does, but since s must be a solution, then any other solution must be orthogonal to it, which is a contradiction.

Therefore, s is the unique non-zero solution to system of equations defined.

■

Problem 2
List all of the numbers $1 \leq x \leq 100$ such that Shor's factoring algorithm actually needs to use a quantum computer in order to find a factor.

Proof. Needing a quantum computer to find a factor of a number N is essentially when we do not have an efficient classical algorithm to find a factor of N. The two classically efficient algorithms for finding a factor of a number N are if N is even or if the number N is some power of a unique prime, *i.e.* $N = p^r$, like 2^6, 7^3, etc. Thus, removing all even numbers, primes, and powers of primes (and 1 trivially) from the list, we are left with

$$\{15, 21, 33, 35, 39, 45, 51, 55, 57, 63, 65, 69, 75, 77, 85, 87, 91, 93, 95, 99\}.$$

■

Problem A4.17
(Reduction of order-finding to factoring) We have seen that an efficient order-finding algorithm allows us to factor efficiently. Show that an efficient factoring algorithm would allow us to efficiently find the order modulo N of any x co-prime to N.

Proof. Suppose x and N are coprime with $N = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}$. By the Chinese Remainder Theorem, we can identify $\mathbb{Z}/N\mathbb{Z}$ with a sum of cyclic groups of prime power order. Our goal is to find the smallest r such that

$$x^r \equiv 1 \pmod{N}.$$

Suppose we have an efficient factoring algorithm and let p_1, p_2, …, p_n be the prime factors of N as above. Then, by Euler's theorem

$$x^{\phi(N)} \equiv 1 \pmod{N},$$

where $\phi(N)$ is the Euler totient function which returns the number of positive integers up to N that are relatively prime with N, and is given by

$$\phi(N) = N \prod_{p|N} \left(1 - \frac{1}{p}\right),$$

where the product is over the distinct prime numbers dividing N. Notice that if N is prime, then every number less than N is clearly relatively prime with N, and thus $\phi(N) = N - 1$. Additionally, the Euler totient function is multiplicative, so if $gcd(m, n) = 1$, then $\phi(mn) = \phi(m)\phi(n)$.
Then,

$$\begin{aligned}\phi(N) &= \phi\left(p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}\right) \\ &= \phi\left(p_1^{a_1}\right)\phi\left(p_2^{a_2}\right)\cdots\phi\left(p_n^{a_n}\right) \\ &= p_1^{a_1-1}(p_1-1)p_2^{a_2-1}(p_2-1)\cdots p_n^{a_n-1}(p_n-1) \\ &= \prod_i p_i^{a_i-1}(p_i-1) \\ &= \prod_i \phi\left(p_i^{a_i}\right).\end{aligned}$$

In particular, if the order of x is r, then r must divide $\phi\left(p_i^{a_i}\right)$. Since we have an efficient factoring algorithm, then all we have to do is find a factorization of $p_i - 1$.
Suppose $p_i - 1$ is a product of prime powers $q_j^{b_j}$, then

$$\phi\left(p_i^{a_i}\right) = p_i^{a_i-1}(p_i-1) = p_i^{a_i-1}\prod_j q_j^{b_j}.$$

Iterating through all the divisors of $\phi\left(p_i^{a_i}\right)$, we find the smallest r such that $x^r \equiv 1 \pmod{p_i^{a_i}}$. This last part can be done efficiently since the powers a_i and b_j are relatively smaller than both x and N.

■

Problem 5.13
Prove (5.44). (*Hint:* $\sum_{s=0}^{r-1}\exp(-2\pi isk/r) = r\delta_{k0}$.) In fact, prove that

$$\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1} \mathrm{e}^{2\pi isk/r}\,|u_s\rangle = \left|x^k \bmod N\right\rangle.$$

Proof. Starting with the left hand side of Equation (5.44), we have

$$\begin{aligned}\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|u_s\rangle &= \frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}\left[\frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}\mathrm{e}^{\frac{2\pi isk}{r}}\left|x^k \bmod N\right\rangle\right] \\ &= \frac{1}{r}\sum_{s=0}^{r-1}\sum_{k=0}^{r-1}\mathrm{e}^{\frac{-2\pi isk}{r}}\left|x^k \bmod N\right\rangle \\ &= \frac{1}{r}\sum_{k=0}^{r-1}\sum_{s=0}^{r-1}\mathrm{e}^{\frac{-2\pi isk}{r}}\left|x^k \bmod N\right\rangle \\ &= \frac{1}{r}\sum_{k=0}^{r-1} r\delta_{k0}\left|x^k \bmod N\right\rangle \\ &= \sum_{k=0}^{r-1}\delta_{k0}\left|x^k \bmod N\right\rangle \\ &= |1\rangle.\end{aligned}$$

Additionally, we have

$$\begin{aligned}
\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1} \mathrm{e}^{\frac{2\pi i s k}{r}} |u_s\rangle &= \frac{1}{\sqrt{r}}\sum_{s=0}^{r-1} \mathrm{e}^{\frac{2\pi i s k}{r}} \left[\frac{1}{\sqrt{r}}\sum_{k=0}^{r-1} \mathrm{e}^{\frac{2\pi i s k}{r}} \left|x^k \bmod N\right\rangle\right] \\
&= \frac{1}{r}\sum_{s=0}^{r-1}\sum_{k'=0}^{r-1} \mathrm{e}^{\frac{2\pi i s (k-k')}{r}} \left|x^{k'} \bmod N\right\rangle \\
&= \frac{1}{r}\sum_{k'=0}^{r-1}\sum_{s=0}^{r-1} \mathrm{e}^{\frac{2\pi i s (k-k')}{r}} \left|x^{k'} \bmod N\right\rangle \\
&= \frac{1}{r}\sum_{k'=0}^{r-1} r\delta_{kk'} \left|x^{k'} \bmod N\right\rangle \\
&= \sum_{k'=0}^{r-1} \delta_{kk'} \left|x^{k'} \bmod N\right\rangle \\
&= \left|x^k \bmod N\right\rangle .
\end{aligned}$$

■

Problem 5.16
For all $x \geq 2$ prove that $\int_x^{x+1} 1/y^2 \,\mathrm{d}y \geq 2/3x^2$. Show that

$$\sum_q \frac{1}{q^2} \leq \frac{3}{2}\int_2^\infty \frac{1}{y^2}\,\mathrm{d}y = \frac{3}{4}$$

and thus that (5.58) holds.

Proof. We have

$$\begin{aligned}
\int_x^{x+1} \frac{1}{y^2}\,\mathrm{d}y &= -\frac{1}{y}\Big|_x^{x+1} \\
&= -\frac{1}{x+1} + \frac{1}{x} \\
&= \frac{1}{x(x+1)}.
\end{aligned}$$

Consider $\frac{1}{x(x+1)} - \frac{2}{3x^2} = \frac{x-1}{3x^2(x+1)}$. For values of $x \geq 2$, the right hand side is always positive, which means that

$$\frac{1}{x(x+1)} = \int_x^{x+1} \frac{1}{y^2}\,\mathrm{d}y \geq \frac{2}{3x^2}.$$

Now, we have

$$\begin{aligned}
\frac{3}{2}\int_2^\infty \frac{1}{y^2}\,\mathrm{d}y &= \frac{3}{2}\sum_{q=2}^{\infty}\int_q^{q+1} \frac{1}{y^2}\,\mathrm{d}y \\
&\geq \frac{3}{2}\sum_{q=2}^{\infty} \frac{2}{3q^2} \\
&= \sum_{q=2}^{\infty} \frac{1}{q^2}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}\frac{3}{2}\int_2^\infty \frac{1}{y^2}\,\mathrm{d}y &= \frac{3}{2}\left.-\frac{1}{y}\right|_2^\infty \\ &= \frac{3}{2}\left(-\frac{1}{\infty}+\frac{1}{2}\right) \\ &= \frac{3}{4}.\end{aligned}$$

Thus,

$$\sum_q \frac{1}{q^2} \le \frac{3}{2}\int_2^\infty \frac{1}{y^2}\,\mathrm{d}y = \frac{3}{4}.$$

Finally, from Equation (5.58), we have

$$1-\sum_{q=2}^{\infty}\frac{1}{q^2} \ge 1-\frac{3}{4}=\frac{1}{4}.$$

Thus, Equation (5.58) holds. ■

Problem 5.17
Suppose N is L bits long. The aim of this exercise is to find an efficient classical algorithm to determine whether $N = a^b$ for some integers $a \ge 1$ and $b \ge 2$. This may be done as follows:

(a) Show that b, if it exists, satisfies $b \le L$.

(b) Show that it takes at most $O\left(L^2\right)$ operations to compute $\log_2(N)$, $x = y/b$ for $b \le L$, and the two integers u_1 and u_2 nearest to 2^x.

(c) Show that it takes at most $O\left(L^2\right)$ operations to compute u_1^b and u_2^b (use repeated squaring) and check to see if either is equal to N.

(d) Combine the previous results to give an $O\left(L^3\right)$ operation algorithm to determine whether $N = a^b$ for integers a and b.

Proof. (a) We have $N = a^b$. Taking the logarithm on both sides, we get $L = b\log(a)$.

- If $a = 1$, then $L = 1$ and $b = 0$.
- If $a \ge 2$, then $\log(a) \ge 1$ and b is a positive integer with $b \le L$.

(b) To calculate two estimates of $x = \log N/b$, we need $O(1)$ operations to find y, $O(L^2)$ operations to compute $x = y/b$ for a specific b, and $O(1)$ operations to calculate 2^x and find the nearest integers u_1 and u_2.

(c) When taking the square of a number, we roughly multiply the number of digits by two. Considering the $\log_2(b)$ loops, $a \times a$ takes

- $O\left(L^2\right)$ in the first iteration.
- $O\left((2L)^2\right)$ in the second iteration.
- $O\left((4L)^2\right)$ in the third iteration.
- $O\left((2^{k-1}L)^2\right)$ in the kth iteration.

Assuming the number of iterations k is relatively small compared to the number of digits L of N, then we need $O(L^3)$ operations to compute u_1^b and u_2^b using repeated squaring and to check to see if either is equal to N.

(d) We need to do parts (b) and (c) L times, requiring a total of $O(L^3)$ operations.

■

Homework 8

This is the last homework of the semester. It is "quasi-optional." This means that you don't have to do it, but there are some reasons you might want to:

- You can use this assignment to replace a previous assignment. For example, if you didn't submit an old assignment, or you did poorly on one, you can use your score on this one to replace that old score. **If this is what you would like to do, please let me know by sending me an email.**
- Acing this assignment is a requirement if you want an A+ in the course. Some of you have been doing very well on all of the requirements in the class, but if you want an A+, then I am requiring you to do this assignment (and ace it). If you don't do this assignment, then the highest grade you can expect to earn is an A, although not doing it won't otherwise affect your grade.

And to reiterate: you are still allowed to submit your solutions in teams of two.
The assignment consists of just one long problem, which will have you work through the structure of *CSS (Calderbank-Shor-Steane) codes*, a special subclass of Pauli stabilizer codes. CSS codes are the most popular codes that people like to (try to!) implement on current quantum computers. The toric code is a special case, and the results of this problem can be easily used to rederive the calculations I did in class for toric code.

Problem Final

A *CSS code* is a Pauli stabilizer code generated by a set of stabilizers with the property that each is either "pure X" or "pure Z". To this end, let $A = \{(\vec{x}_1, 0), \ldots, (\vec{x}_l, 0)\} \subset \mathbb{F}_2^n \oplus \mathbb{F}_2^n$ and $B = \{(0, \vec{z}_1), \ldots, (0, \vec{z}_k)\} \subset \mathbb{F}_2^n \oplus \mathbb{F}_2^n$, and define

$$S = S(A, B) := \{X(\vec{x}_i), Z(\vec{z}_j) \mid (\vec{x}_i, 0) \in A, (0, \vec{z}_j) \in B\}.$$

In this exercise, we will establish fairly explicit formulas for the distance and number of logical qubits for the stabilizer code

$$\mathcal{C}_S = \{|\psi\rangle \in \left(\mathbb{C}^2\right)^{\otimes n} \mid s\,|\psi\rangle = |\psi\rangle\,, s \in S\}.$$

(a) By a minor abuse of notation, let A denote the $n \times l$ matrix (with entries in $\mathbb{F}_2 = \{0, 1\}$) whose i^{th} column is $\vec{x}_i$, where $\left(\vec{x}_i, \vec{0}\right) \in A$. Similarly, let B denote the $n \times k$ matrix whose j^{th} column is $\vec{z}_j$, where $\left(\vec{0}, \vec{z}_j\right) \in B$.

Show that $\mathcal{C}_S \neq \{\vec{0}\}$ is a non-trivial subspace of $\left(\mathbb{C}^2\right)^{\otimes n}$ if and only if $B^T A = 0$ if and only if $A^T B = 0$.

[Hint: remember, I explained in class that $\mathcal{C}_S \neq \{\overrightarrow{0}\}$ if and only if $-1 \notin \langle S \rangle \leq G_n$, where G_n is the Pauli group on n qubits. You need to think through what our choice of stabilizer generators has to do with the symplectic inner product.]

For the rest of the problem, let us assume that $B^T A = 0$.

(b) Let $X = \text{span}(A) \leq \mathbb{F}_2^n \oplus \mathbb{F}_2^n$, $Z = \text{span}(B) \leq \mathbb{F}_2^n \oplus \mathbb{F}_2^n$, and $W = \text{span}(A \cup B) = X \oplus Z \leq \mathbb{F}_2^n \oplus \mathbb{F}_2^n$. Show that $W^\perp = X^\perp \cap Z^\perp$.

Note: as should be expected, for a subspace $V \leq \mathbb{F}_2^n \oplus \mathbb{F}_2^n$, the notation $V^\perp$ means the symplectic complement. That is

$$V^\perp = \left\{(\vec{a}, \vec{b}) \in \mathbb{F}_2^n \oplus \mathbb{F}_2^n \mid \omega\left((\vec{a}, \vec{b}), (\vec{v}_1, \vec{v}_2)\right) = 0, \quad (\vec{v}_1, \vec{v}_2) \in V\right\}$$

where

$$\omega\left((\vec{a}, \vec{b}), (\vec{v}_1, \vec{v}_2)\right) = \vec{a} \cdot \vec{v}_2 + \vec{b} \cdot \vec{v}_1 \pmod 2.$$

(c) Recall that the *symplectic weight* $wt((\vec{x}, \vec{z}))$ of a vector $(\vec{x}, \vec{z}) = ((x_1, \ldots, x_n), (z_1, \ldots, z_n)) \in \mathbb{F}_2^n \oplus \mathbb{F}_2^n$ is the number of nonzero columns of the matrix

$$\begin{pmatrix} x_1 & \cdots & x_n \\ z_1 & \cdots & z_n \end{pmatrix}.$$

Show

$$\begin{aligned} \min_{(\vec{x}, \vec{z}) \in W^\perp - W} wt((\vec{x}, \vec{z})) &= \min\left\{ \min_{(\vec{x}, \overrightarrow{0}) \in W^\perp - W} |\vec{x}|, \min_{(\overrightarrow{0}, \vec{z}) \in W^\perp - W} |\vec{z}| \right\} \\ &= \min\left\{ \min_{\vec{x} \in \ker(B^T) - \text{im}(A)} |\vec{x}|, \min_{\vec{z} \in \ker(A^T) - \text{im}(B)} |\vec{z}| \right\} \end{aligned}$$

where $|\vec{x}|$ is the usual Hamming weight of a vector in $\mathbb{F}_2^n$.

(d) Show

$$W^\perp = \left(\ker(B^T) \oplus \{\vec{0}\}\right) + \left(\{\vec{0}\} \oplus \ker(A^T)\right)$$

and use this to show

$$W^\perp/W = \frac{X^\perp \cap Z^\perp}{X+Z} = \frac{\Big(\ker(B^T)\oplus\{\vec{0}\}\Big)+\Big(\{\vec{0}\}\oplus\ker(A^T)\Big)}{\Big(\mathrm{im}(A)\oplus\{\vec{0}\}\Big)+\Big(\{\vec{0}\}\oplus\mathrm{im}(B)\Big)} \cong \frac{\ker(B^T)}{\mathrm{im}(A)}\oplus\frac{\ker(A^T)}{\mathrm{im}(B)}.$$

(e) Show

$$\dim\left(\frac{\ker(B^T)}{\mathrm{im}(A)}\right) = \dim\left(\frac{\ker(A^T)}{\mathrm{im}(B)}\right).$$

(f) Conclude that the number of logical qubits in $\mathcal{C}_S$ is

$$\frac{1}{2}\dim W^\perp/W = \dim\left(\frac{\ker(B^T)}{\mathrm{im}(A)}\right).$$

Proof. ■

That concludes the problem. What follows is for your enjoyment.
We can summarize/interpret the results of this problem as follows: part (a) shows that a CSS code on n qubits (together with a choice of generators for it) is essentially the same thing as matrices A and B such that the composition

$$\mathbb{F}_2^l \xrightarrow{A} \mathbb{F}_2^n \xrightarrow{B^T} \mathbb{F}_2^k$$

is $B^T A = 0$. In algebraic topology, this would be called a "chain complex of based vector spaces over $\mathbb{F}_2$." Such things can be found in many places, thanks to the following: any choice of a positive integer k and "CW complex" Δ gives rise to one, by looking at the cellular chain complex (with $\mathbb{F}_2$ coefficients) in dimension k:

$$C_{k+1}^{\text{cellular}}(\Delta;\mathbb{F}_2) \xrightarrow{\partial} C_k^{\text{cellular}}(\Delta;\mathbb{F}_2) \xrightarrow{\partial} C_{k-1}^{\text{cellular}}(\Delta;\mathbb{F}_2)$$

Part (f) shows that the number of logical qubits is exactly the rank of the homology of the chain complex. Part (c), roughly, shows that the distance of the code is determined by identifying the smallest Hamming weight representative of any non-trivial homology class. More precisely, we have to find the minimum weight we see among the non-trivial homology classes in the above chain complex, as well as in the dual cochain complex

$$C^{k+1}_{\text{cellular}}(\Delta;\mathbb{F}_2) \xleftarrow{\delta} C^{k}_{\text{cellular}}(\Delta;\mathbb{F}_2) \xleftarrow{\delta} C^{k-1}_{\text{cellular}}(\Delta;\mathbb{F}_2)$$

If this sounds interesting to you, then you should take MA 572 - Algebraic Topology!
As far as current state of the art: as of roughly 2020, "good quantum LDPC codes" are now known to exist. LD means "low-density" and PC means "parity check" (a synonym for CSS). More precisely, this means there exists an infinite family of CSS codes with generators $(A_n, B_n)_{n\in\mathbb{N}}$ with the following properties:

- the n^{th} code uses $O(n)$ many physical qubits.
- the columns of the matrices A_n and B_n have $O(1)$ many non-zero entries ("low-density").
- there are $\theta(n)$ many logical qubits.
- the distance is $\theta(n)$.

The last two properties are why the codes are called "good."