**18.435/2.111 Homework 3 Solutions**

**1a:**

$$\sum_{i=1}^{k} E_i = \sum_{i=1}^{k} T\Pi_i T^\dagger \quad (1)$$

$$= T(\sum_{i=1}^{k} \Pi_i)T^\dagger \quad (2)$$

$$= TT^\dagger \quad (3)$$

$$= I_n \quad (4)$$

**1b:** I have received feedback that this problem was confusing in terms of dimensions, so I will try to be precise. The confusion arises when we try to interpret a *ket* $|v\rangle$ in terms of the more familiar *vector* or *n-tuple* $[\ v_1 \quad v_2 \quad \cdots \quad v_n\ ]^T$ or when we interpret an *operator* as a familiar *matrix*. Both the matrix and the vector representations for operators and kets arise when we define a basis for the space of interest. If we define a basis $\{|i\rangle\}_{i=1\cdots d}$ we can say that $\mathcal{H}_d$ is the span of this basis. We can further define the space $\mathcal{H}_n$ as the span of the first $n$ kets of the basis. If we want to represent $|v\rangle$ as a vector, we can do it in this basis by defining $v_i = \langle i|v\rangle$.

Similarly, for any operator $A : \mathcal{H}_d \mapsto \mathcal{H}_d$, we can fully describe it by its action on the basis kets $\{|i\rangle\}_{i=1}^{d}$. We define the *matrix elements* $A_{ij} = \langle i|A|j\rangle$ and we can always write the operator as $A = \sum_{ij} A_{ij} |i\rangle \langle j|$. Let's interpret $T$ as an operator. We can see from the matrix elements that $T = \sum_{i=1}^{n} |i\rangle \langle i|$.

Now to the problem of interest. We have been told that $|v\rangle$ is in the subspace generated by the first $n$ basis vectors, *i.e.* $v_i = 0$ for $i > n$. Viewed in terms of its basis elements, it should be clear that $T|v\rangle = |v\rangle$. In that case, we insert $T$ on either side of $\Pi_i$ and see the desired equality:

$$\langle v| \Pi_i |v\rangle = \langle v| T\Pi_i T |v\rangle = \langle v| E_i |v\rangle \quad (5)$$

You might complain that I've left off the adjoint $^\dagger$. When we interpret $T$ in bra-ket notation, as above, we can see that $T = T^\dagger$.

I can see why this may be confusing. It may be easier to see if we were to rewrite the desired equation as

$${}_{\mathcal{H}_d}\langle v|\Pi_i |v\rangle_{\mathcal{H}_d} = {}_{\mathcal{H}_n}\langle v|E_i |v\rangle_{\mathcal{H}_n} \quad (6)$$

where I labelled with the bras and kets with subscripts indicating which space they live in. In that case, $T$ is defined as $T = \sum_{i=1}^{n} |i\rangle_{\mathcal{H}_n} {}_{\mathcal{H}_d}\langle i|$. We can see that $T \neq T^\dagger$. In this form, we should derive the desired result by noting that when $|v\rangle_{\mathcal{H}_d}$ is spanned by the first $n$ basis vectors, we have $|v\rangle_{\mathcal{H}_n} = T^\dagger |v\rangle_{\mathcal{H}_n} = T^\dagger T |v\rangle_{\mathcal{H}_d}$. Using this, we can get

$$ {}_{\mathcal{H}_d}\langle v|\Pi_i|v\rangle_{\mathcal{H}_d} = {}_{\mathcal{H}_d}\langle v|T^\dagger T \Pi_i T^\dagger T|v\rangle_{\mathcal{H}_d} = {}_{\mathcal{H}_n}\langle v|E_i|v\rangle_{\mathcal{H}_n} \tag{7} $$

**1c:** First, let's understand the probabilities associated with the $k+1$ outcomes of the consecutive projective measurement. The first measurement has two outcomes, namely $N$ and NOT $N$. We can state $p(N) = \langle\psi| T^\dagger T |\psi\rangle$ and $p(\text{NOT } N) = \langle\psi| (I - T^\dagger T) |\psi\rangle = 1 - p(N)$. Now, to derive the probabilities of the subsequent measurement (for if we observed outcome $N$), first let's write down the resultant state after outcome $N$: $(T^\dagger T)/\sqrt{p(N)}\,|\psi\rangle$. From here, we can write the conditional probabilities for the $k$ outcomes associated with $\Pi_i$:

$$ p(i|N) = \frac{\langle\psi| T^\dagger T \Pi_i T^\dagger T |\psi\rangle}{p(N)}. \tag{8} $$

We use the definition of conditional probability to conclude

$$ p(i) = p(i|N)p(N) = \langle\psi| T^\dagger T \Pi_i T^\dagger T |psi\rangle . \tag{9} $$

Combining these, we can define a POVM by the $k+1$ operators $\{I - T^\dagger T, T^\dagger T \Pi_i T^\dagger T\}$. These form a valid POVM which we can see by summing them to the identity:

$$ \begin{aligned} I - T^\dagger T + \sum_{i=1}^{k} T^\dagger T \Pi_i T^\dagger T &= I - T^\dagger T + T^\dagger T (\sum_{i=1}^{k} \Pi_i) T^\dagger T && (10)\\ &= I - T^\dagger T + T^\dagger T && (11)\\ &= I. && (12) \end{aligned} $$

**2:** Let's first define what we mean by $E^{1/2}$. We can see by its construction that $E$ is a positive matrix (since it is defined as $E = A^\dagger A$), so we know that it can be written as $E = V^\dagger D V$, where $V$ is unitary and $D$ is a diagonal matrix where the diagonal entries are non-negative real numbers. By convention, we write the diagonal elements of $D$ in descending order. This is the spectral decomposition of $E$. We define $E^{1/2} \equiv V^\dagger D^{1/2} V$. (This definition is unique if $E$ has distinct eigenvalues, but this will not be important for the existence proof we need here.)

Now let's write down the singular value decomposition of $A = USW$, where $U$ and $W$ are unitary and $S$ is a diagonal matrix with non-negative real diagonal entries. Again, we write the diagonal elements of $S$ in descending order. By writing $E = A^\dagger A = W^\dagger S U^\dagger U S W = W^\dagger S^2 W$, we see that $S = D^{1/2}$.

If we write $A = UD^{1/2}W = UWW^\dagger D^{1/2}W$, we can define $U' = W^\dagger U^\dagger$ and we see that $U'A = W^\dagger U^\dagger U W W^\dagger D^{1/2} W = W^\dagger D^{1/2} W$. This is what we wanted to show, as the left hand side is $E^{1/2}$. (Note, we haven't claimed $W = V$. This is only true the eigenvalues of $E$ are distinct and $E^{1/2}$ is uniquely defined. This is a technicality that is not too important to our purposes but is certainly worth understanding.)

**3:** Let's start by noting that $Q^2 = R^2 = S^2 = T^2 = I$. This follows immediately from the fact that the eigenvalues are $\pm 1$. With this, the first result is straight algebra:

$$\begin{aligned}
(Q &\otimes S + R\otimes S + R\otimes T - Q\otimes T)^2 \\
&= ((R+Q)\otimes S + (R-Q)\otimes T)^2 \\
&= (Q+R)^2\otimes I + (R-Q)^2 \otimes I \\
&\quad +(R+Q)(R-Q)\otimes ST + (R-Q)(R+Q)\otimes TS \\
&= 4I\otimes I + QR\otimes I + RQ\otimes I - QR\otimes I - RQ\otimes I \\
&\quad +(R^2 - Q^2 - RQ + QR)\otimes ST + (R^2 - Q^2 - QR + RQ)\otimes TS \\
&= 4I + QR\otimes ST - RQ\otimes ST - QR\otimes TS + RQ\otimes TS \\
&= 4I + [Q,R]\otimes ST - [Q,R]\otimes TS \\
&= 4I + [Q,R]\otimes [S,T].
\end{aligned}$$

The bound is easiest to see if we define

$$X = Q\otimes S + R\otimes S + R\otimes T - Q\otimes T.$$

We know that the variance of a random variable is non-negative, so $\langle X^2\rangle - \langle X\rangle^2 \geq 0$. We rearrange this and take the square root to get $\langle X\rangle \leq \langle X^2\rangle^{1/2}$. So to finish the proof of the bound, we need to show that $\langle X^2\rangle \leq 8$.

To continue, let's remind ourselves about the meaning of the notation $\langle A\rangle$. Above, I used it as the expected value. This is a shorthand referring to the expected value of an observable (i.e. Hermitian operator) $A$ when the quantum state is $|\psi\rangle$: $\langle A\rangle = \langle\psi| A |\psi\rangle$. It should be clear from linear algebra that $\langle A\rangle \leq \lambda_{max}(A)$ where $\lambda_{max}(A)$ is the largest eigenvalue of $A$.

Now let's look at $\lambda_{max}([Q,R])$. Writing out the commutator, we can see that $\lambda_{max}(QR - RQ) \leq \lambda_{max}(RQ) + \lambda_{min}(RQ) \leq 2$. A similar argument shows that $\lambda_{max}([S,T]) \leq 2$. Finally, since $\lambda_{max}(A\otimes B) = \lambda_{max}(A)\lambda_{max}(B)$, we have $\lambda_{max}([Q,R]\otimes[S,T]) \leq 4$.

Combining all of this together,

$$\begin{aligned}\langle X^2 \rangle &= \langle 4I + [Q,R] \otimes [S,T] \rangle \\ &\leq 4 + \lambda_{max}([Q,R] \otimes [S,T]) \\ &\leq 8.\end{aligned}$$

. This is the desired result.

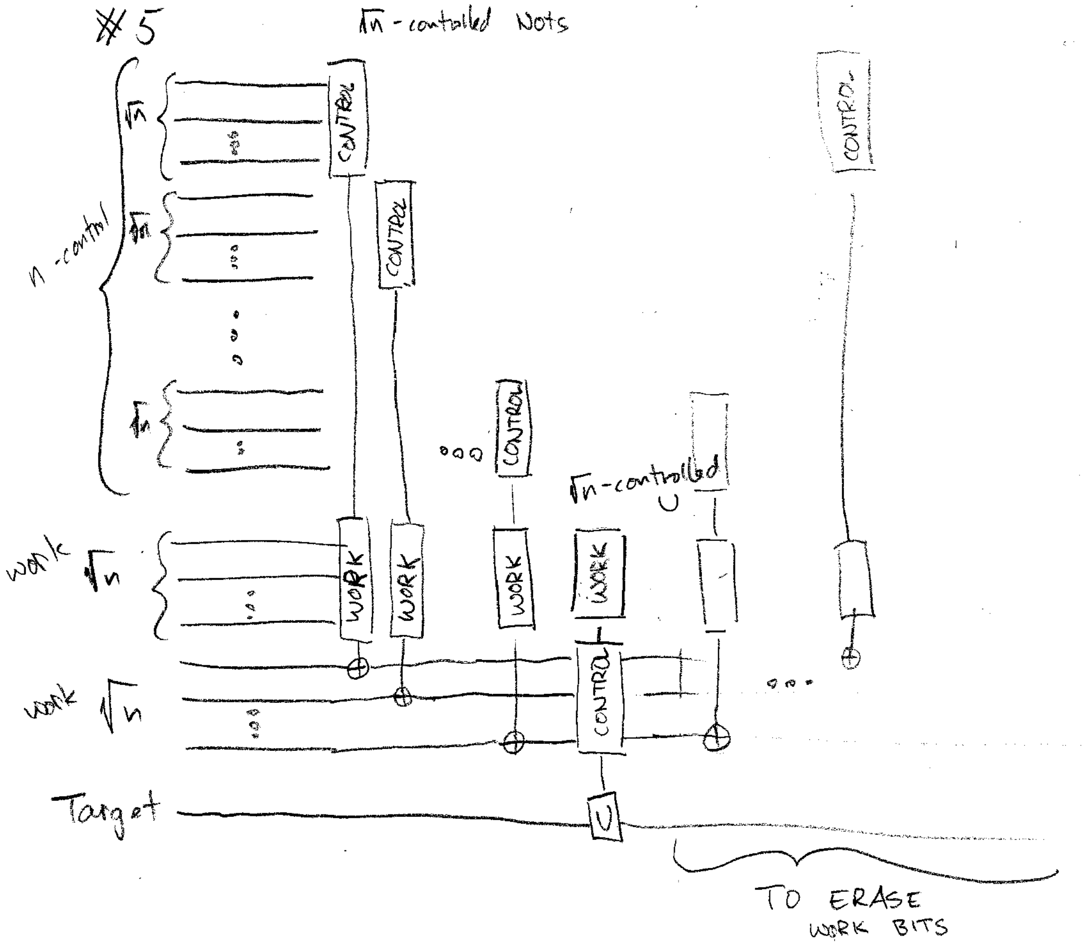
**4:** See attached diagram.

**5:** As mentioned in the hint, we can do this in $O(\sqrt{n})$ or $O(\log n)$ work bits, each with $O(n)$ gates. We'll describe the construction for $O(\sqrt{n})$.

We know from class that we can design a $\sqrt{n}$-controlled $U$ gate using $\sqrt{n}$ work bits and $O(\sqrt{n})$ gates. We'll need $2\sqrt{n}$ work bits. The first $\sqrt{n}$ bits will be used for each $\sqrt{n}$-controlled gate. The second $\sqrt{n}$ bits will be used as follows. Divide the $n$ control bits into $\sqrt{n}$ groups of $\sqrt{n}$ bits each. Using $\sqrt{n}$-controlled Nots and the first $\sqrt{n}$ work bits, set a work bit to one for each set of $\sqrt{n}$ control bits. When we have done this for each group, we will use the second set of work bits as control for a $\sqrt{n}$-controlled $U$ gate acting on the target bit. Finally, we need to reverse the operations to erase the work bits. Since the control-NOT is its own inverse, we simply repeat the task.

See attached diagram.

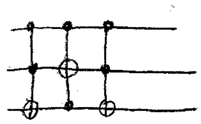
# #5



√n - controlled Nots
n-control
√n
√n
√n
CONTROL
CONTROL
CONTROL
CONTROL
work √n
WORK
WORK
WORK
WORK
√n-controlled U
work √n
CONTROL
Target
U
TO ERASE
WORK BITS


# #4

(a)



(b)