Ph219/CS219: Quantum Computation

Fall 2005

Solutions to Problem Set 2

Problem 2.1

Let U (resp. V) be the unitary that diagonalizes ρ (resp. $\rho' \equiv \mathcal{E}(\rho)$) and let Δ (resp. Δ') be the corresponding diagonal matrix. We can write

$$\begin{aligned} \mathcal{E}(\rho) = \sum_\mu M_\mu \, \rho \, M_\mu^\dagger \quad &\Rightarrow \quad V \Delta' V^\dagger = \sum_\mu M_\mu \, U \Delta U^\dagger \, M_\mu^\dagger \\ &\Rightarrow \quad \Delta' = \sum_\mu N_\mu \, \Delta \, N_\mu^\dagger \, , \end{aligned} \tag{1}$$

where $N_\mu \equiv V^\dagger M_\mu U$. We note that $\sum_\mu N_\mu^\dagger N_\mu = \sum_\mu N_\mu N_\mu^\dagger = I$ and therefore the operation defined by the Kraus operators $\{N_\mu\}$ is also unital.

We want to show that $\rho' \prec \rho \Leftrightarrow \lambda(\rho') \prec \lambda(\rho)$. But $[\lambda(\rho)]_i = \Delta_i$ and $[\lambda(\rho')]_i = \Delta'_i$, where $[\lambda]_i$ denotes the i-th component of the vector λ and Δ_i denotes the i-th diagonal entry of the diagonal matrix Δ. Using Eq. (1) we have

$$\begin{aligned} \Delta'_i &= \sum_{\mu,j} [N_\mu]_{ij} \, \Delta_j \, [N_\mu^\dagger]_{ji} \\ &= \sum_j \left(\sum_\mu [N_\mu]_{ij} \, [N_\mu^\dagger]_{ji} \right) \Delta_j \\ &= \sum_j D_{ij} \, \Delta_j \, , \end{aligned} \tag{2}$$

where $D_{ij} \equiv \sum_\mu [N_\mu]_{ij} \, [N_\mu^\dagger]_{ji} = \sum_\mu |[N_\mu]_{ij}|^2$. We can verify D is doubly stochastic by calculating $\sum_j D_{ij} = [\sum_\mu N_\mu N_\mu^\dagger]_i = [I]_i = 1$ and $\sum_i D_{ij} = [\sum_\mu N_\mu^\dagger N_\mu]_j = [I]_j = 1$. Thus we have showed $\lambda(\rho') \prec \lambda(\rho)$. $\square$

Problem 2.2

(a) Consider how the channel would act in the special case when we have a single qubit ($n = 1$): It applies one of the three Pauli operators or the identity at random, all with equal probability 1/4. Let's call this channel η. Then we can easily verify that[1] $\eta(\rho) = \frac{1}{4} \sum_{x=0}^{3} \sigma(x) \rho \sigma(x) = \frac{1}{2} I$. In other words, the channel takes *any* input

[1] Or, see §3.4.1 in http://www.theory.caltech.edu/people/preskill/ph229/notes/chap3.ps

density matrix ρ and replaces it with the completely mixed state $\frac{1}{2}I$. In fact, for any matrix A input to the channel η the output is proportional to the completely mixed state $\frac{1}{2}I$ (with proportionality constant $\mathrm{tr}A$).

The channel $\mathcal{E}$ acting on n qubits is nothing but the channel η applied separately on each qubit, i.e. $\mathcal{E} = \eta^{\otimes n}$. But the result of η acting on some qubit is to produce a state proportional to the completely mixed state—all correlations among qubits are destroyed. Therefore

$$\mathcal{E}(\rho) = \left(\frac{1}{2}I\right)^{\otimes n} = \frac{1}{2^n}I^{\otimes n} . \tag{3}$$

(b) It follows from the Kraus representation theorem that the two different Kraus representations $\{\frac{1}{2^n}\sigma(x)\}$ and $\{\sqrt{p_a}U_a\}$ of the channel $\mathcal{E}$ must be related by a unitary transformation V. We observe that V must have dimension $\max(N, 2^{2n})$, since 2^{2n} and N is the number of Kraus operators in the two representations and a unitary taking operators from one set to the other can only be defined if the two sets have the same cardinality (we just add zero elements in the shorter set until they both have $\max(N, 2^{2n})$ elements). We can write

$$\begin{aligned}
\sqrt{p_a}U_a &= \sum_x V_{ax}\tfrac{1}{2^n}\sigma(x) \\
\Rightarrow \mathrm{tr}\left(\sqrt{p_a}U_a U_a^\dagger \sqrt{p_a}\right) &= \mathrm{tr}\left(\sum_{x,y} V_{ax}\tfrac{1}{2^n}\sigma(x)\sigma(y)\tfrac{1}{2^n}V_{ay}^*\right) \\
\Rightarrow 2^n p_a &= \tfrac{1}{2^{2n}}\sum_{xy} V_{ax}V_{ay}^*\,\mathrm{tr}\left(\sigma(x)\sigma(y)\right) ,
\end{aligned} \tag{4}$$

where we have used that $U_a U_a^\dagger = I$ since U_a is unitary (acting on a 2^n-dimensional space), and have interchanged the order of the trace and the summation since trace is linear.

Now, for qubits we know that $\mathrm{tr}\left(\sigma_i\sigma_j\right) = 2\delta_{ij}$, since the product of any two different Pauli matrices is a new Pauli matrix (which is traceless) and Pauli matrices square to the identity. We can generalize this to tensor products of Pauli matrices $\sigma(x)$ as $\mathrm{tr}\left(\sigma(x)\sigma(y)\right) = 2^n\delta_{xy}$, since if the $2n$-bit strings x and y differ in at least one place then a (non-identity) Pauli matrix will appear at the corresponding location making the trace zero. Substituting in Eq. (4) we get

$$\begin{aligned}
2^n p_a &= \tfrac{1}{2^{2n}}2^n\sum_x V_{ax}V_{ax}^* \\
\Rightarrow p_a &= \tfrac{1}{2^{2n}}\sum_x |V_{ax}|^2 \le \tfrac{1}{2^{2n}} ,
\end{aligned} \tag{5}$$

since $\sum_x |V_{ax}|^2 \le 1$ as x ranges over 2^{2n} of the $\max(N, 2^{2n})$ elements in the a-th row of V. $\square$

Problem 2.3

(a) Bob's probability of error is

$$\begin{aligned} p_{\text{error}} &= \operatorname{tr}\left(p_1\rho_1 E_1 + p_2\rho_2 E_1\right) \\ &= \operatorname{tr}\left(p_1\rho_1(I - E_2) + p_2\rho_2 E_1\right) \\ &= p_1 + \operatorname{tr}\left((p_2\rho_2 - p_1\rho_1)E_1\right) \ . \end{aligned} \tag{6}$$

Writing $p_2\rho_2 - p_1\rho_1$ in diagonal form as $p_2\rho_2 - p_1\rho_1 = \sum_i \lambda_i |i\rangle\langle i|$ and substituting back in Eq. (6) we obtain

$$p_{\text{error}} = p_1 + \sum_i \lambda_i \langle i|E_1|i\rangle \ . \tag{7}$$

(b) We note that $\langle i|E_1|i\rangle, \langle i|E_2|i\rangle \geq 0$ since E_1, E_2 are non-negative, and also that $\langle i|i\rangle = \langle i|E_1|i\rangle + \langle i|E_2|i\rangle = 1 \Rightarrow 0 \leq \langle i|E_1|i\rangle \leq 1$. Therefore to minimize the probability of error given by Eq. (7) we must set all coefficients $\langle i|E_1|i\rangle$ multiplying positive eigenvalues to the minimum value zero and all such coefficients multiplying negative eigenvalues to the maximum value one. In other words, we must choose E_1 to be the projection onto the subspace spanned by all eigenvectors of $p_2\rho_2 - p_1\rho_1$ with negative eigenvalues, and the optimal probability of error becomes

$$(p_{\text{error}})_{\text{opt}} = p_1 + \sum_{i:\lambda_i<0} \lambda_i \ . \tag{8}$$

(c) By the definition of the trace norm

$$||p_2\rho_2 - p_1\rho_1||_{\text{tr}} = \sum_{i:\lambda_i\geq 0} \lambda_i \ - \sum_{i:\lambda_i<0} \lambda_i \ , \tag{9}$$

whereas

$$p_2 - p_1 = \operatorname{tr}\left(p_2\rho_2 - p_1\rho_1\right) = \sum_{i:\lambda_i\geq 0} \lambda_i \ + \sum_{i:\lambda_i<0} \lambda_i \ . \tag{10}$$

Combining Eqs. (9) and (10) we get

$$\sum_{i:\lambda_i<0} \lambda_i = \tfrac{1}{2}(p_2 - p_1) - \tfrac{1}{2}||p_2\rho_2 - p_1\rho_1||_{\text{tr}} \ , \tag{11}$$

and so

$$(p_{\text{error}})_{\text{opt}} = \tfrac{1}{2} - \tfrac{1}{2}||p_2\rho_2 - p_1\rho_1||_{\text{tr}} \ . \tag{12}$$

When $\rho_1 = \rho_2$, Eq. (12) gives an optimal error probability of $(1 - |p_2 - p_1|)/2$. But this makes sense, since Bob is given identical states in both cases and so his optimal strategy is to choose based on the a priori probabilities the possibility with the largest probability. Say $p_2 > p_1$. Then his probability of making an error is equal to the smaller probability p_1 and indeed $(1-(p_2-p_1))/2 = p_1$!

If instead ρ_1 and ρ_2 have support on orthogonal subspaces, then we expect that Bob should be able to always win by performing a POVM that projects on either one of the two subspaces with certainty for any of the two states Alice may give him. In fact, now $p_2\rho_2 - p_1\rho_1$ is trivially diagonalized if we diagonalize ρ_1 and ρ_2 separately. Moreover, the sum of the positive eigenvalues equals p_2 whereas the sum of the negative eigenvalues equals $-p_1$. Therefore, from Eq. (8) the optimal error probability is $p_1 - p_1 = 0$ as expected.

(d) In this case

$$p_2\rho_2 - p_1\rho_1 = \frac{1}{2}\begin{pmatrix} -cos(2\alpha) & 0 \\ 0 & cos(2\alpha) \end{pmatrix} , \tag{13}$$

and so the eigenvalues are $\lambda_{1,2} = \pm\frac{1}{2}cos(2\alpha)$. The eigenvector corresponding to the negative eigenvalue is $|0\rangle \equiv \binom{1}{0}$ and so

$$E_1 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} . \tag{14}$$

The optimal error probability is then

$$(p_{\text{error}})_{\text{opt}} = \frac{1}{2}(1 - cos(2\alpha)) = sin^2(\alpha) . \tag{15}$$

(e) We have

$$\begin{aligned} p_{\text{error}} &= \sum_i p_i \; p_{\text{error}}(i) \\ &= \sum_i p_i \left(\tfrac{1}{2} - \tfrac{1}{2}|p(2|i) - p(1|i)|\right) \\ &= \tfrac{1}{2} - \tfrac{1}{2}\sum_i |p_i\, p(2|i) - p_i\, p(1|i)| \\ &= \tfrac{1}{2} - \tfrac{1}{2}\sum_i |\text{tr}\left((p_2\rho_2 - p_1\rho_1)\, E_i\right)| . \end{aligned} \tag{16}$$

(f) We substitute $p_2\rho_2 - p_1\rho_1 = \sum_j \lambda_j |j\rangle\langle j|$ into Eq. (16) to get

$$\begin{aligned} p_{\text{error}} &= \tfrac{1}{2} - \tfrac{1}{2}\sum_i |\text{tr}\left(\left(\sum_j \lambda_j |j\rangle\langle j|\right) E_i\right)| \\ &= \tfrac{1}{2} - \tfrac{1}{2}\sum_i |\sum_j \lambda_j \langle j|E_i|j\rangle| \\ &\geq \tfrac{1}{2} - \tfrac{1}{2}\sum_i \sum_j |\lambda_j| \langle j|E_i|j\rangle \end{aligned} \tag{17}$$

where we have used the triangle inequality to upper bound the norm of the sum by the sum of the norms. But now we note that $\sum_i E_i = I$ and so

$$p_{\text{error}} \geq \frac{1}{2} - \frac{1}{2}\sum_j |\lambda_j| = \frac{1}{2} - \frac{1}{2}||p_2\rho_2 - p_1\rho_1||_{\text{tr}} , \tag{18}$$

using the definition of the trace norm. □

Problem 2.4

(a) We know that any TPCP map $\mathcal{E}$ with Kraus operators $\{M_\mu\}$ has a unitary representation. That is, there exists a unitary U_{AC} between the system space $\mathcal{H}_A$ and an ancillary space $\mathcal{H}_C$ such that each M_μ takes the form $M_\mu = {}_C\langle\mu|U_{AC}|0\rangle_C$, where $|0\rangle_C$ is a fixed reference state and $\{|\mu\rangle_C\}$ is an orthonormal basis in $\mathcal{H}_C$ [2].

By definition, the dual operation $\mathcal{E}^*$ acts on an operator a_A as

$$\begin{aligned} \mathcal{E}^*(a) &= \sum_\mu M_\mu^\dagger \, a_A \, M_\mu \\ &= \sum_\mu {}_C\langle 0|U_{AC}^\dagger|\mu\rangle_C \; a_A \;\; {}_C\langle\mu|U_{AC}|0\rangle_C \\ &= \sum_\mu {}_C\langle 0|U_{AC}^\dagger \, (a_A \otimes I_C) \, U_{AC}|0\rangle_C \; , \end{aligned} \tag{19}$$

since $\sum_\mu |\mu\rangle_C \, {}_C\langle\mu| = I_C$.

(b) By definition, an operation $\mathcal{E}$ on $\mathcal{H}_{AB}$ is semicausal if there exists an operation $\tilde{\mathcal{E}}$ on $\mathcal{H}_A$ such that for all operators a_A and for all density matrices ρ_{AB},

$$\mathrm{tr}_B\left((a_A \otimes I_B)\, \mathcal{E}(\rho_{AB})\right) = a_A \tilde{\mathcal{E}}(\rho_A) \; , \tag{20}$$

where $\rho_A \equiv \mathrm{tr}_B\left(\rho_{AB}\right)$. In other words, any change in Bob's half of the state ρ_{AB} before $\mathcal{E}$ acts cannot make any difference on Alice's side—the evolution of her subsystem A due to $\mathcal{E}$ only depends on the state in her subsystem ρ_A and therefore can be described by an effective operation $\tilde{\mathcal{E}}$ acting on $\mathcal{H}_A$ alone.

Going to the Heisenberg picture, the dual description of a semicausal $\mathcal{E}^*$ will be

$$\mathrm{tr}_B\left(\mathcal{E}^*(a_A \otimes I_B)\, \rho_{AB}\right) = \tilde{\mathcal{E}}^*(a_A)\, \rho_A \; , \tag{21}$$

again for all operators a_A and for all density matrices ρ_{AB}. Using $\rho_A \equiv \mathrm{tr}_B\left(\rho_{AB}\right)$ we can rewrite Eq. (21) as

$$\mathrm{tr}_B\left(\mathcal{E}^*(a_A \otimes I_B)\, \rho_{AB}\right) = \mathrm{tr}_B\left(\left(\tilde{\mathcal{E}}^*(a_A) \otimes I_B\right)\; \rho_{AB}\right) \; . \tag{22}$$

Since Eq. (23) is true for all ρ_{AB}, we conclude that

$$\mathcal{E}^*(a_A \otimes I_B) = \tilde{\mathcal{E}}^*(a_A) \otimes I_B \; . \tag{23}$$

[2] For a TPCP map the outcome of the fictitious measurement on $\mathcal{H}_C$ is unknown—after the unitary U_{AC} is applied we trace over system C using the $\{|\mu\rangle_C\}$ basis.

(c) By definition, an operation $\mathcal{E}$ on $\mathcal{H}_{AB}$ is semilocal if there exists a CP map $\mathcal{G}_{AC}$ on $\mathcal{H}_{AC}$ and a TPCP map $\mathcal{F}_{BC}$ on $\mathcal{H}_{BC}$ such that

$$\mathcal{E} = \mathrm{tr}_C\left((\mathcal{F}_{BC} \otimes I_A) \circ (\mathcal{G}_{AC} \otimes I_B) \right) \ . \tag{24}$$

$\mathcal{G}_{AC}$ is an operation from Alice's system A to the message system C. It is not necessarily trace-preserving, since Alice may use the one-way quantum channel to tell Bob what the result of a measurement she potentially did was. On the other hand, $\mathcal{F}_{BC}$ is an operation from Bob's system B to the message system C and is necessarily trace-preserving as Bob cannot communicate back to Alice. Actually, Bob eventually discards the message system C. So we may think of tr_C as being part of the operation $\mathcal{F}_{BC}$ so that

$$\mathcal{E} = (\mathcal{F}_{BC} \otimes I_A) \circ (\mathcal{G}_{AC} \otimes I_B) \ . \tag{25}$$

Taking the dual of $\mathcal{E}$ amounts to taking the duals of $\mathcal{G}_{AC}$ and $\mathcal{F}_{BC}$ separately and then reversing their order [3]. Thus

$$\mathcal{E}^* = (\mathcal{G}^*_{AC} \otimes I_B) \circ (\mathcal{F}^*_{BC} \otimes I_A) \ . \tag{26}$$

(d) We want to calculate

$$\mathcal{E}^*\left(a_A \otimes I_B\right) = \left[(\mathcal{G}^*_{AC} \otimes I_B) \circ (\mathcal{F}^*_{BC} \otimes I_A) \right] \left(a_A \otimes I_B\right) \tag{27}$$

and show that it equals $\tilde{\mathcal{E}}^*(a_A) \otimes I_B$ for some $\tilde{\mathcal{E}}^*$. First, we have $\mathcal{F}^*_{BC}(I_B) = I_{BC}$ since $\mathcal{F}_{BC}$ is trace-preserving. Therefore

$$\left[(\mathcal{F}^*_{BC} \otimes I_A) \right] \left(a_A \otimes I_B\right) = a_A \otimes I_B \otimes I_C \ . \tag{28}$$

Substituting in Eq. (27) we get

$$\begin{aligned} \mathcal{E}^*\left(a_A \otimes I_B\right) &= \left[(\mathcal{G}^*_{AC} \otimes I_B) \right] \left(a_A \otimes I_B \otimes I_C\right) \\ &= \tilde{\mathcal{E}}^*(a_A) \otimes I_B \ , \end{aligned} \tag{29}$$

where $\tilde{\mathcal{E}}^*(a_A) \equiv \mathcal{G}^*_{AC}(a_A \otimes I_C)$. $\square$

[3] This is intuitive as the order in which operators are applied in the Heisenberg picture is the reverse of the order they appear in the Schroedinger picture. You can explicitly verify the fact by considering the Kraus decomposition of $\mathcal{E}$ (say, $\{E_\mu\}$) in terms of the Kraus decompositions of $\mathcal{G}_{AC}$ (say, $\{[G_{AC}]_\nu\}$) and $\mathcal{F}_{BC}$ (say, $\{[F_{BC}]_\lambda\}$) and then taking the Hermitian adjoint of each E_μ to find the Kraus decomposition of $\mathcal{E}^*$ in terms of $\mathcal{G}^*_{AC}$ and $\mathcal{F}^*_{BC}$.