Ph219a/CS219a

Solutions
Nov 24, 2008

Problem 2.1

(a) The probability that Bob guesses wrong is

$$\begin{aligned} p_{error} &= \mathrm{Tr}[\ p_1\rho_1 E_2 + p_2\rho_2 E_1\ ] \\ &= \mathrm{Tr}[\ p_1\rho_1(\mathcal{I} - E_1) + p_2\rho_2 E_1\ ] \\ &= p_1 + \mathrm{Tr}[\ (p_2\rho_2 - p_1\rho_1)E_1\ ] \\ &= p_1 + \mathrm{Tr}\left[\sum_i (\lambda_i \left|i\right\rangle \left\langle i\right|) E_1\right] \\ &= p_1 + \sum_i \lambda_i \left\langle i\right| E_1 \left|i\right\rangle \end{aligned} \tag{1}$$

where $|i\rangle$ is the eigenstate of $(p_2\rho_2 - p_1\rho_1)$ corresponding to eigenvalue λ_i, as defined.

(b) Since E_1, E_2 are nonnegative and sum to identity ($E_1 + E_2 = \mathcal{I}$) their expectation values satisfy $0 \leq \langle i| E_1 |i\rangle \leq 1$. In order to minimize p_{error}, the coeffecient $\langle i| E_1 |i\rangle$ should be set to its minimum value($= 0$) when $|i\rangle$ corresponds to a positive eigenvalue and maximum value($= 1$) when $|i\rangle$ corresponds to a negative eigenvalue. In other words, Bob must choose E_1 to be the projector onto the subspace spanned by all eigenvectors of $(p_2\rho_2 - p_1\rho_1)$ corresponding to negative eigenvalues to obtain the optimal error probability

$$(p_{\mathrm{error}})_{\mathrm{optimal}} = p_1 + \sum_{\mathrm{neg}} \lambda_i \tag{2}$$

(c)From the definition of the L^1 norm, we have,

$$\|\ p_2\rho_2 - p_1\rho_1\ \|_1 = \sum_{i:\lambda_i \geq 0} \lambda_i - \sum_{i:\lambda_i < 0} \lambda_i \tag{3}$$

whereas

$$p_2 - p_1 = \mathrm{Tr}[\ p_2\rho_2 - p_1\rho_1\ ] = \sum_{i:\lambda_i \geq 0} \lambda_i + \sum_{i:\lambda_i < 0} \lambda_i \tag{4}$$

Combining equations (3) and (4), we have

$$\begin{aligned}\sum_{i:\lambda_i<0} \lambda_i &= \frac{1}{2}(p_1 - p_2) - \frac{1}{2} \parallel p_2\rho_2 - p_1\rho_1 \parallel_1 \\ &= \frac{1}{2} - p_1 - \frac{1}{2} \parallel p_2\rho_2 - p_1\rho_1 \parallel_1 \end{aligned} \tag{5}$$

Therefore

$$(p_{\text{error}})_{\text{optimal}} = p_1 + \sum_{i:\lambda_i<0} \lambda_i = \frac{1}{2} - \frac{1}{2} \parallel p_2\rho_2 - p_1\rho_1 \parallel_1 \tag{6}$$

When $\rho_1 = \rho_2$, equation(6) gives an optimal error probability of $(1 - |p_1 - p_2|)/2$. This makes sense, because when Bob is given identical states in both cases, his optimal strategy is to choose based on the *a priori* probabilities the state with the higher probability. Thus if $p_1 > p_2$, his probability of making an error is equal to the smaller probability p_2 and indeed, $(1 - |p_1 - p_2|)/2 = p_2$!

When ρ_1 and ρ_2 have support on orthogonal subspaces, then we expect that Bob should always be able to win by performing a POVM that distinguishes perfectly between the two states that Alice is sending. In terms of equation (6), this means that we can now diagonalize $(p_2\rho_2 - p_1\rho_1)$ by diagonalizing ρ_1 and ρ_2 separately so that $\parallel (p_2\rho_2 - p_1\rho_1) \parallel_1 = p_1 + p_2$. Therefore the optimal error probability is $\frac{1}{2}[1 - (p_1 + p_2)] = 0$ as expected!

(d) In this case,

$$p_2\rho_2 - p_1\rho_1 = \frac{1}{2}\begin{pmatrix} -\cos(2\alpha) & 0 \\ 0 & \cos(2\alpha) \end{pmatrix} \tag{7}$$

which has eigenvalues $\lambda_\pm = \pm\frac{1}{2}\cos(2\alpha)$. The eigenvector corresponding to the negative eigenvalue is $|0\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ so that

$$E_1 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{8}$$

which gives

$$(p_{\text{error}})_{\text{optimal}} = \frac{1}{2}(1 - \cos(2\alpha)) = \sin^2\alpha \tag{9}$$

(e) Bob's error probability is now given by

$$\begin{aligned} p_{\text{error}} &= \sum_i p_i\, p_{\text{error}}(i) \\ &= \sum_i p_i \left(\frac{1}{2} - \frac{1}{2}|p(2|i) - p(1|i)|\right) \\ &= \frac{1}{2} - \frac{1}{2}\sum_i |\, p_i p(2|i) - p_i p(1|i)\,| \\ &= \frac{1}{2} - \frac{1}{2}\sum_i |\,\mathrm{Tr}[\,(p_2\rho_2 - p_1\rho_1)E_i\,]\,| \end{aligned} \tag{10}$$

using the equalities $p_i p(2|i) = p_2 p(i|2) = p_2 \mathrm{Tr}[\rho_2 E_i]$ etc.

(f) Expanding $(p_2\rho_2 - p_1\rho_1)$ in its eigenbasis as $(p_2\rho_2 - p_1\rho_1) = \sum_j \lambda_j |j\rangle\langle j|$, we have,

$$\begin{aligned} p_{\text{error}} &= \frac{1}{2} - \frac{1}{2}\sum_i |\sum_j \lambda_j \langle j| E_i |j\rangle | \\ &\geq \frac{1}{2} - \frac{1}{2}\sum_i \sum_j |\lambda_j| \langle j| E_i |j\rangle \quad \text{(Triangle inequality!)} \\ &= \frac{1}{2} - \frac{1}{2}\sum_j |\lambda_j| \quad (\text{since } \sum_i E_i = \mathcal{I}) \\ &= \frac{1}{2} - \frac{1}{2} \parallel p_2\rho_2 - p_1\rho_1 \parallel_1 \end{aligned} \tag{11}$$

using the definition of the L^1 norm.

Problem 2.2

(a) Let $AB = X$. Then $ABBA = XX^\dagger$ and $BAAB = X^\dagger X$ when A and B are Hermitian. Now we show that for any matrix X, $X^\dagger X$ and $XX^\dagger$ have the same eigenvalues -
Say $|v\rangle$ is an eigenvector of $X^\dagger X$ corresponding to an eigenvalue λ_v, then,

$$XX^\dagger(X|v\rangle) = X(X^\dagger X)|v\rangle = X\lambda_v|v\rangle = \lambda_v(X|v\rangle) \tag{12}$$

we see that λ_v is also an eigenvalue of $XX^\dagger$ corresponding to eigenvector $X|v\rangle$. By a similar argument we can show that every (nonzero) eigenvalue of $XX^\dagger$ is also an eigenvalue of $X^\dagger X$. Using our definition $AB = X$, we see that $ABBA$ and $BAAB$ have the same eigenvalues. Therefore

$$\begin{aligned} \parallel AB \parallel_1 &= \text{Tr}[\sqrt{(AB)^\dagger AB}] = \text{Tr}[\sqrt{BAAB}] \quad \text{(for Hermitian } A, B) \\ &= \text{Tr}[\sqrt{ABBA}] \quad \text{(using the result above)} \\ &= \text{Tr}[\sqrt{(BA)^\dagger BA}] \\ &= \parallel BA \parallel_1 \end{aligned} \tag{13}$$

(b) Let $A_i = \rho^{1/2}E_i^{1/2}$ and $B_i = U\tilde{\rho}^{1/2}E_i^{1/2}$. Then the Schwarz inequality gives

$$\begin{aligned} |\text{Tr}[A_i^\dagger B_i]| &\leq \left(\text{Tr}[A_i^\dagger A_i]\right)^{1/2}\left(\text{Tr}[B_i^\dagger B_i]\right)^{1/2} \\ \Rightarrow |\text{Tr}[E_i\rho^{1/2}U\tilde{\rho}^{1/2}]| &\leq \sqrt{\text{Tr}[E_i\rho]}\sqrt{\text{Tr}[E_i\tilde{\rho}^{1/2}]} \end{aligned} \tag{14}$$

Using the definition of the overlap function,

$$\begin{aligned}
\text{Overlap}(\rho, \tilde{\rho}, E_i) &= \sum_i \sqrt{\text{Tr}[\rho E_i]}\sqrt{\text{Tr}[\tilde{\rho} E_i]} \\
&\geq \sum_i |\,\text{Tr}[\,E_i \rho^{1/2} U \tilde{\rho}^{1/2}\,]\,| \;\; \text{(using equation (14))} \\
&\geq |\sum_i \text{Tr}[\,E_i \rho^{1/2} U \tilde{\rho}^{1/2}\,]\,| \;\; \text{(triangle inequality!)} \\
&= |\,\text{Tr}[\rho^{1/2} U \tilde{\rho}^{1/2}]\,| \;\; (\text{since} \sum_i E_i = \mathcal{I})
\end{aligned} \tag{15}$$

as desired.

(c) Using the polar decomposition to write $\tilde{\rho}^{1/2}\rho^{1/2} = V\sqrt{\rho^{1/2}\tilde{\rho}\rho^{1/2}}$ and choosing the unitary U in part(b) to be $U \equiv V^{-1} = V^{\dagger}$, the LHS of equation() becomes,

$$\begin{aligned}
|\,\text{Tr}[\rho^{1/2} V^{\dagger} \tilde{\rho}^{1/2}]\,| &= |\text{Tr}[\tilde{\rho}^{1/2}\rho^{1/2}V^{\dagger}]| \\
&= |\,\text{Tr}[V^{\dagger}\tilde{\rho}^{1/2}\rho^{1/2}]\,| \\
&= |\,\text{Tr}[V^{\dagger}V\sqrt{\rho^{1/2}\tilde{\rho}\rho^{1/2}}]\,| \;\; \text{(using polar decomposition)} \\
&= \text{Tr}[\sqrt{\rho^{1/2}\tilde{\rho}\rho^{1/2}}]
\end{aligned} \tag{16}$$

Combining this with our result of part(b), we have,

$$\text{Overlap}(\rho, \tilde{\rho}, E_i) \geq \text{Tr}[\sqrt{\rho^{1/2}\tilde{\rho}\rho^{1/2}}] = \sqrt{F(\rho, \tilde{\rho})} \tag{17}$$

(d) Using the Bloch parametrization, the density operators ρ and $\tilde{\rho}$ can be written as $\rho = \frac{1}{2}(\mathcal{I} + \vec{\sigma}.\vec{P})$ and $\tilde{\rho} = \frac{1}{2}(\mathcal{I} + \vec{\sigma}.\vec{Q})$. Now lets define $M = \rho^{1/2}\tilde{\rho}\rho^{1/2}$ and calculate its trace and determinant in terms of the polarization vectors $\vec{P}$ and $\vec{Q}$:-

$$\begin{aligned}
\text{Tr}[M] &= \text{Tr}[\rho^{1/2}\tilde{\rho}\rho^{1/2}] = \frac{1}{4}\text{Tr}[(\mathcal{I} + \vec{\sigma}.\vec{Q})(\mathcal{I} + \vec{\sigma}.\vec{P})] \\
&= \frac{1}{4}\text{Tr}[\mathcal{I} + \vec{\sigma}.(\vec{Q} + \vec{P}) + (\vec{\sigma}.\vec{Q})(\vec{\sigma}.\vec{P})]
\end{aligned} \tag{18}$$

Recall that $\text{Tr}[\sigma] = 0$ and that $\text{Tr}[(\vec{\sigma}.\vec{Q})(\vec{\sigma}.\vec{P})] = \text{Tr}[\vec{P}.\vec{Q} + i\vec{\sigma}.(\vec{P} \times \vec{Q})] = 2\vec{P}.\vec{Q}$ (since $\text{Tr}[\,\mathcal{I}\,] = 2$ for single qubit operators). Thus

$$\text{Tr}[M] = \frac{1}{2}(1 + \vec{P}.\vec{Q}) \tag{19}$$

Also,

$$\det(M) = \det(\rho^{1/2}\tilde{\rho}\rho^{1/2}) = \det(\rho\tilde{\rho}) = \det(\tilde{\rho})\det(\rho) \tag{20}$$

Writing out the 2×2 matrices ρ and $\tilde{\rho}$ explicitly, we find that $\det(\rho) = \frac{1}{4}(1 - \vec{P}^2)$ so that

$$\det(M) = \frac{1}{16}(1 - \vec{P}^2)(1 - \vec{Q}^2) \tag{21}$$

Now recall, the trace and determinant of a matrix can be written in terms of its eigenvalues as

$$\prod_i \lambda_i = \det(M) \quad \sum_i \lambda_i = \text{Tr}[M] \tag{22}$$

The fidelity $F(\rho,\tilde{\rho}) \equiv F(\vec{P},\vec{Q})$ is given by

$$\begin{aligned} F(\vec{P},\vec{Q}) &= \left(\mathrm{Tr}\sqrt{(\rho^{1/2}\tilde{\rho}\rho^{1/2})}\right) \\ &= (\sqrt{\lambda_1}+\sqrt{\lambda_2})^2 \\ &= \lambda_1+\lambda_2+2\sqrt{\lambda_1\lambda_2} \\ &= \mathrm{Tr}[M]+2\sqrt{\det(M)} \\ &= \frac{1}{2}\left(1+\vec{P}.\vec{Q}+\sqrt{(1-\vec{P}^2)(1-\vec{Q}^2)}\right) \end{aligned} \tag{23}$$

where we have used equations (19) and (21).

Problem 2.3

(a) Writing out the operator E_{DK} in the standard basis, we have,

$$E_{DK} = \begin{pmatrix} 1-A & A\sin(2\alpha) \\ A\sin(2\alpha) & 1-A \end{pmatrix} \tag{24}$$

which has eigenvalues $\lambda_\pm = (1-A) \pm A\sin(2\alpha)$. For E_{DK} to be positive, the smaller of the two eigenvalues must be positive, ie.

$$\lambda_- = 1-A-A\sin(2\alpha) \geq 0 \quad \Rightarrow A \leq \frac{1}{1+\sin(2\alpha)} \tag{25}$$

Thus the largest possible value of A that Bob can choose (so that E_{DK} is still positive) is $A = \frac{1}{1+\sin(2\alpha)}$.

The probability of outcome E_{DK} when Alice sends $|u\rangle$ and $|v\rangle$ are respectively,

$$\begin{aligned} p(E_{DK}|u) &= \langle u|\, E_{DK}\, |u\rangle = (1-2A)+A+A|<u|v>|^2 = 1-A(1-\sin^2(2\alpha)) \\ p(E_{DK}|v) &= \langle v|\, E_{DK}\, |v\rangle = (1-2A)+A(\sin^2(2\alpha)+A = 1-A(1-\sin^2(2\alpha)) \end{aligned} \tag{26}$$

Assuming Alice chooses between these two states equiprobably, the probability of outcome E_{DK} is $p(E_{DK}) = \frac{1}{2}(p(E_{DK}|u)+p(E_{DK}|u)) = 1-A(1-\sin^2(2\alpha))$ which takes on the minimum value $(p(E_{DK}))_{\min} = \sin(2\alpha)$ when A takes on its maximum value $A = \frac{1}{1+\sin(2\alpha)}$.

(b) Lets denote the basis in which Eve is measuring as $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The probabilities of Eve's outcomes, conditioned on the state that Alice is sending are:

$$\begin{aligned} p(0|u) &= |<0|u>|^2 = \cos^2\alpha\ ,\ p(0|v) = |<0|v>|^2 = \sin^2\alpha \\ p(1|u) &= |<1|u>|^2 = \sin^2\alpha\ ,\ p(1|v) = |<1|v>|^2 = \cos^2\alpha \end{aligned} \tag{27}$$

(For example, $p(0|u)$ denotes the probability of Eve obtaining outcome $|0\rangle$ when Alice sends $|u\rangle$.)

Similarly, the probabilities of Bob's outcomes conditioned on the state he receives are:

$$\begin{aligned} p(E_{\neg v}|v) &= 0\ ,\ p(E_{\neg u}|v) = A(1-|<u|v>|^2) = A(1-\sin^2(2\alpha)) \\ p(E_{\neg u}|u) &= 0\ ,\ p(E_{\neg v}|u) = A(1-\sin^2(2\alpha)) \end{aligned} \tag{28}$$

(We have already computed $p(E_{DK}|u)$ and $p(E_{DK}|v)$ in part(a).)
There are ways in which Bob could obtain a wrong, but conclusive outcome. Lets compute their probabilities, assuming that Alice is making her choice equiprobably -
(i) Alice sends $|u\rangle$, Eve obtains outcome $|1\rangle$ and sends $|v\rangle$ to Bob, who obtains the outcome $\neg u$. This occurs with probability

$$p(u \to 1 \to \neg u) = \frac{1}{2}p(1|u)p(E_{\neg u}|v) = \frac{1}{2}(\sin^2\alpha)(A\cos^2(2\alpha)) = \frac{A\sin^2\alpha\cos^2(2\alpha)}{2} \tag{29}$$

(ii) Alice sends $|v\rangle$, Eve obtains outcome $|0\rangle$ and sends $|u\rangle$ to Bob, who obtains the outcome $\neg v$. This occurs with probability

$$p(v \to 0 \to \neg v) = \frac{1}{2}p(0|v)p(E_{\neg v}|u) = \frac{A\sin^2\alpha\cos^2(2\alpha)}{2} \tag{30}$$

Thus the probability of Bob obtaining an error conclusively is given by,

$$p(\text{conclusive, error}) = A\sin^2\alpha\cos^2(2\alpha) = (\sin^2\alpha)(1-\sin(2\alpha)) \tag{31}$$

where we have used the largest possible value of A to minimize the probability.
This tells us what fraction of Alice's messages went wrong due to Eve's interference. Now lets look at it from the point of view of being able to detect Eve's presence. Given either $|u\rangle$ or $|v\rangle$, the probability that Bob obtains a conclusive outcome is simply

$$\begin{aligned} p(\text{conclusive}) &= 1 - p(E_{DK}) = 1 - \frac{1}{2}(p(E_{DK}|u) + p(E_{DK}|v)) \\ &= 1-(1-A\cos^2(2\alpha)) = 1-\sin(2\alpha) \ \ \text{(using the max value of } A) \end{aligned} \tag{32}$$

Therefore the probability that Bob obtains an error, given that his outcome is conclusive is given by

$$\begin{aligned} p(\text{error}|\text{conclusive}) &= \frac{p(\text{conclusive, error})}{p(\text{conclusive})} \\ &= \frac{(\sin^2\alpha)(1-\sin(2\alpha))}{1-\sin(2\alpha)} = \sin^2\alpha \end{aligned} \tag{33}$$

which is also the probability of being able to detect Eve's presence.

Problem 2.4

(a) Consider how the protocol works for the special case $n = 1$, ie. when Alice encrypts a single qubit state ρ. The effect of Alice's encryption protocol can then be viewed as a quantum channel which applies one of the Pauli operators or the identity at random, each with probability $1/4$.

Notice that this channel (which we call $\mathcal{E}_1$) is simply the single qubit depolarizing channel[1] with $p = 3/4$. Recall that the effect of this channel is to map any state into the maximally mixed state -

$$\begin{aligned}\mathcal{E}_1(\rho) &= \frac{1}{4}\sum_{x=0}^{3}\sigma(x)\rho\sigma(x) \\ &= \frac{1}{4}(\rho + \sigma_1\rho\sigma_1 + \sigma_2\rho\sigma_2 + \sigma_3\rho\sigma_3) \\ &= \frac{1}{2}\mathcal{I} \end{aligned} \tag{34}$$

Now, in the case of a general n-qubit encryption, the channel $\mathcal{E}$ is nothing but the channel $\mathcal{E}_1$ acting separately on each qubit, ie. $\mathcal{E} = \mathcal{E}_1^{\otimes n}$. As we know, each qubit gets mapped to the maximally mixed state, in the process destroying all correlations between qubits. Therefore

$$\mathcal{E}(\rho) = (\frac{1}{2}\mathcal{I})^{\otimes n} = \frac{1}{2^n}\mathcal{I}^{\otimes n} \tag{35}$$

(b) If Alice and Bob are to achieve the same result with this new encryption scheme, then the channels $\mathcal{E}'$ and $\mathcal{E}$ must be the same. It follows from the Kraus representation theorem that two different Kraus representations ($\{\frac{1}{2^n}\sigma(x)\}$ and $\sqrt{p_a}U_a$) of this channel ($\mathcal{E}$) must be related by a unitary transformation. Lets call this transformation V, where

$$\sqrt{p_a}U_a = \sum_x V_{ax}\left(\frac{1}{2^n}\sigma(x)\right) \tag{36}$$

Since $\{x\}$ are $2n$-bit strings, there are 2^{2n} Kraus operators in the first representation and N in the second representation (since a takes values from 1 to N). The dimension of V is therefore $\max(N, 2^{2n})$ where we add zeros to the shorter set of Kraus operators so that both sets have the same cardinality, $\max(N, 2^{2n})$.
Multiplying both sides of equation(36) by their adjoints and taking the trace, we have,

$$\begin{aligned}\mathrm{Tr}[p_a U_a^\dagger U_a] &= \mathrm{Tr}[\frac{1}{2^{2n}}\sum_{x,y} V_{ax}V_{ay}^*\sigma(x)\sigma(y)] \\ \Rightarrow 2^n p_a &= \frac{1}{2^{2n}}\sum_x 2^n |V_{ax}|^2 \end{aligned} \tag{37}$$

We have simplified the left hand side using the fact that $\{U_a\}$ are unitaries acting on a n-qubit space, so $\mathrm{Tr}[U_a^\dagger U_a] = \mathrm{Tr}[\mathcal{I}] = 2^n$. To see how the right hand side simplifies, recall that single qubit Pauli operators have the property $\sigma_i\sigma_j = 2\delta_{ij}$ ie. they are orthogonal and square to identity. Generalizing this to n-qubits, we have $\sigma(x)\sigma(y) = \mathcal{I}^{\otimes n}\delta_{xy}$ which gives the RHS of equation (37). Thus we have,

$$p_a = \frac{1}{2^{2n}}\sum_x |V_{ax}|^2 \leq \frac{1}{2^{2n}} \tag{38}$$

since we are summing over the elements ath row the unitary matrix of dimension $\max(N, 2^{2n})$, with x ranging from 0 to 2^{2n}.

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

Problem 2.5

Given, a density operator ρ and a unital map $\mathcal{E}$ with operator elements $\{M_\mu\}$ -

$$\mathcal{E}(\rho) = \sum_\mu M_\mu \rho M_\mu^\dagger \tag{39}$$

Let U be the unitary that diagonalizes ρ into Δ and V diagonalize $\mathcal{E}(\rho)$ into Δ' so that the above equation becomes

$$\begin{aligned} V\Delta' V^\dagger &= \sum_\mu M_\mu U \Delta U^\dagger M_\mu^\dagger \\ \Rightarrow \Delta' &= \sum_\mu N_\mu \Delta N_\mu^\dagger \end{aligned} \tag{40}$$

where $\{N_\mu = V^\dagger M_\mu U\}$ form a set of operator elements for a TPCP, unital map ie.

$$\begin{aligned} \sum_\mu N_\mu^\dagger N_\mu &= U^\dagger \left[\sum_\mu M_\mu^\dagger (V^\dagger V) M_\mu \right] U \\ &= \mathcal{I} \ \text{(since } U, V \text{ are unitary, } \mathcal{E} \text{ is TP)} \\ \sum_\mu N_\mu N_\mu^\dagger &= V^\dagger \left[\sum_\mu M_\mu (U U^\dagger) M_\mu^\dagger \right] V \\ &= \mathcal{I} \ \text{(since } \mathcal{E} \text{ is unital)} \end{aligned} \tag{41}$$

The eigenvalues of ρ are simply the elements of Δ, ie. $[\lambda(\rho)]_i = \Delta_i$ and $[\lambda(\mathcal{E}(\rho))]_i = \Delta'_i$. Then equation(40) implies

$$\begin{aligned} \Delta'_i &= \sum_\mu \sum_j [N_\mu]_{ij} \Delta_j [N_\mu^\dagger]_{ji} \\ &= \sum_j \left(\sum_\mu [N_\mu]_{ij} [N_\mu^\dagger]_{ji} \right) \Delta_j \\ &= \sum_j D_{ij} \Delta_j \end{aligned} \tag{42}$$

where $D_{ij} = \sum_\mu [N_\mu]_{ij} [N_\mu^\dagger]_{ji}$ is a double stochastic matrix -

$$\begin{aligned} \sum_i D_{ij} &= \sum_i \sum_\mu [N_\mu^\dagger]_{ji} [N_\mu]_{ij} = [\mathcal{I}]_j = 1 \ \text{(since } \{N_\mu\} \text{ is a TPCP map)} \\ \sum_j D_{ij} &= \sum_j \sum_\mu [N_\mu]_{ij} [N_\mu^\dagger]_{ji} = [\mathcal{I}]_i = 1 \ \text{(since } \{N_\mu\} \text{ is a unital map)} \end{aligned} \tag{43}$$

Thus we see that $\Delta' \prec \Delta$ which proves that $\mathcal{E}(\rho) \prec \rho$.