Assignment 2

MATH 667 – Quantum Information Theory

Mark Girard

26 February 2016

Problem 1

Problem 1 (Naimark's Theorem)**.** Consider the the rank 1 matrices

$$F_a = \frac{2}{3} \left|\uparrow_{\hat{n}_a}\right\rangle \left\langle\uparrow_{\hat{n}_a}\right|, \quad a = 1, 2, 3,$$

where $\left|\uparrow_{\hat{n}_a}\right\rangle \in \mathbb{C}^2$ for every a and the unit vectors $\hat{n}_a \in \mathbb{R}^3$ satisfy $\hat{n}_1 + \hat{n}_2 + \hat{n}_3 = 0$.

(a) Show that the set $\{F_a\}$ is a POVM.

(b) Find an orthonormal basis $\{|v_a\rangle\}$ in $\mathbb{C}^3$ such that $\{F_a\}$ is realized by $\{E_a\}$ where $E_a = |v_a\rangle \langle v_a|$.

Solution. (a) We need to show that $\sum_a F_a = I$. First note that $\sum_a \operatorname{Tr} F_a = \frac{2}{3} \sum_a \operatorname{Tr} \left|\uparrow_{\hat{n}_a}\right\rangle \left\langle\uparrow_{\hat{n}_a}\right| = \frac{2}{3} \cdot 3 = 2$. Next, we show that $\sum_a \langle F_a, \sigma_j\rangle = 0$ for each $j = 1, 2, 3$, where the operators σ_j are the standard Pauli operators. Recall that, for every $j = 1, 2, 3$ and every unit vector $\hat{n} \in \mathbb{R}^3$, we have $\left\langle\uparrow_{\hat{n}}\right| \sigma_j \left|\uparrow_{\hat{n}}\right\rangle = \hat{n} \cdot \hat{e}_j$ where $\{\hat{e}_j\}$ are the standard unit vectors in $\mathbb{R}^3$. Thus

$$\begin{aligned}
\sum_a \langle F_a, \sigma_j\rangle &= \frac{2}{3} \sum_a \operatorname{Tr}\left[\left|\uparrow_{\hat{n}_a}\right\rangle \left\langle\uparrow_{\hat{n}_a}\right| \sigma_j\right] \\
&= \frac{2}{3} \sum_a \left\langle\uparrow_{\hat{n}_a}\right| \sigma_j \left|\uparrow_{\hat{n}_a}\right\rangle \\
&= \frac{2}{3} \sum_a \hat{n}_a \cdot \hat{e}_j = 0,
\end{aligned}$$

which follows from the assumption that $\sum_a \hat{n}_a = 0$.

(b) The set of vectors $\left\{\left|\uparrow_{\hat{n}_1}\right\rangle, \left|\uparrow_{\hat{n}_2}\right\rangle, \left|\uparrow_{\hat{n}_3}\right\rangle\right\}$ must be linearly dependent, since it is a set of three vectors in a space of dimension 2. This implies that there is a non-trivial linear combination of these vectors such that

$$\alpha_1 \left|\uparrow_{\hat{n}_1}\right\rangle + \alpha_2 \left|\uparrow_{\hat{n}_2}\right\rangle + \alpha_3 \left|\uparrow_{\hat{n}_3}\right\rangle = 0,$$

where $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{C}$ are not all zero. We may assume that $|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 = 1$ without loss of generality. Define the vectors

$$|v_a\rangle = \sqrt{\frac{2}{3}} \left\langle 0\right| \uparrow_{\hat{n}_a}\rangle \left|0\right\rangle + \sqrt{\frac{2}{3}} \left\langle 1\right| \uparrow_{\hat{n}_a}\rangle \left|1\right\rangle + \overline{\alpha}_a \left|2\right\rangle$$

for every $a = 1, 2, 3$ and define the operator

$$U = \sum_a |v_a\rangle \langle a| .$$

In block matrix form, this matrix is

$$U = \begin{pmatrix} \sqrt{\frac{2}{3}} \left|\uparrow_{n_1}\right\rangle & \sqrt{\frac{2}{3}} \left|\uparrow_{n_2}\right\rangle & \sqrt{\frac{2}{3}} \left|\uparrow_{n_3}\right\rangle \\ \overline{\alpha}_1 & \overline{\alpha}_2 & \overline{\alpha}_3 \end{pmatrix}.$$

We have that

$$\begin{aligned} UU^* &= \begin{pmatrix} \sqrt{\frac{2}{3}} \left|\uparrow_{n_1}\right\rangle & \sqrt{\frac{2}{3}} \left|\uparrow_{n_2}\right\rangle & \sqrt{\frac{2}{3}} \left|\uparrow_{n_3}\right\rangle \\ \overline{\alpha}_1 & \overline{\alpha}_2 & \overline{\alpha}_3 \end{pmatrix} \begin{pmatrix} \sqrt{\frac{2}{3}} \left\langle\uparrow_{n_1}\right| & \alpha_1 \\ \sqrt{\frac{2}{3}} \left\langle\uparrow_{n_2}\right| & \alpha_2 \\ \sqrt{\frac{2}{3}} \left\langle\uparrow_{n_3}\right| & \alpha_3 \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{3} \sum_a \left|\uparrow_{\hat{n}_a}\right\rangle \left\langle\uparrow_{\hat{n}_a}\right| & \sqrt{\frac{2}{3}} \sum_a \alpha_a \left|\uparrow_{\hat{n}_a}\right\rangle \\ \sqrt{\frac{2}{3}} \sum_a \overline{\alpha}_a \left\langle\uparrow_{\hat{n}_a}\right| & \sum_a |\alpha_a|^2 \end{pmatrix} \\ &= \begin{pmatrix} I & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

and thus U is a unitary operator. It follows that $\{|v_a\rangle\}$ must be an orthonormal basis of $\mathbb{C}^3$.

Problem 2

Problem 2 (The Density Operator of a Qutrit)**.** Consider a density operator for a qutrit; that is, $\rho : \mathbb{C}^3 \to \mathbb{C}^3$. Let $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_8)$ be a vector of matrices where λ_i $(i = 1, 2, \ldots, 8)$ are some Hermitian traceless 3×3 generalizations of the Pauli matrices (e.g. the Gell-Mann matrices) satisfying the condition $\mathrm{Tr}(\lambda_i \lambda_j) = 2\delta_{ij}$ (note that also the Pauli matrices satisfy this orthogonality condition).

(a) Show that ρ can be written as:

$$\rho = \frac{1}{3} I + \vec{P} \cdot \vec{\lambda},$$

where $\vec{P}$ is a vector in $\mathbb{R}^8$ and I is the 3×3 identity matrix.

(b) Show that $|\vec{P}| \leq \frac{1}{\sqrt{3}}$.

(c) Show that if ρ is a pure state then $|\vec{P}| = \frac{1}{\sqrt{3}}$.

(d) Is it true that every $\vec{P}$ with $|\vec{P}| \leq \frac{1}{\sqrt{3}}$ corresponds to a density matrix?

Solution. (a) The space of 3×3 hermitian matrices is 9-dimensional as a real vector space. Since each λ_k is traceless and $\mathrm{Tr}\, \lambda_i \lambda_j = \delta_{ij}$ holds for each pair (i, j), it follows that $\{I, \lambda_1, \ldots, \lambda_8\}$ is an orthogonal basis. Hence every qutrit density operator can be written uniquely as

$$\rho = p_0 I + p_1 \lambda_1 + \cdots + p_8 \lambda_8$$

for some real numbers $p_0, \ldots, p_8$. Note that $\mathrm{Tr}\, \rho = p_0 \,\mathrm{Tr}\, I = 3p_0$. Since it must hold that $\mathrm{Tr}\, \rho = 1$, it follows that $p_0 = \frac{1}{3}$. Defining $\vec{P} = (p_1, \ldots, p_8)$ yields the desired result.

(b) Recall that $\mathrm{Tr}\, \rho^2 \leq 1$ for all density operators ρ, and equality holds if and only if ρ is pure. We have

$$\rho^2 = \frac{1}{9} I + \frac{1}{3} \vec{P} \cdot \vec{\lambda} + \sum_{i,j} P_i P_j \lambda_i \lambda_j,$$

and thus

$$\mathrm{Tr}\, \rho^2 = \frac{1}{9} \cdot 3 + 0 + 2 \sum_i P_i^2 = \frac{1}{3} + 2|P|^2.$$

Noting that $\mathrm{Tr}\, \rho^2 \leq 1$ must hold, it follows that $|P|^2 \leq \frac{1}{3}$ and thus $|P| \leq \frac{1}{\sqrt{3}}$.

(c) Using the same observation as above, since $\mathrm{Tr}\, \rho^2 = 1$ holds for any pure state, it follows that $|P| = \frac{1}{\sqrt{3}}$.

(d) No. Consider the following 3×3 matrix, which is traceless and has the property that $\mathrm{Tr}\, \lambda_1^2 = 2$

$$\lambda_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

We may then construct the remaining desired set of traceless matrices $\lambda_2, \ldots, \lambda_8$ using this λ_1. Choose $\vec{P} = \frac{1}{\sqrt{3}}(1, 0, \ldots, 0)$ and consider the matrix

$$\frac{1}{3} I + \vec{P} \cdot \vec{\lambda} = \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

which clearly has a negative eigenvalue even though $|\vec{P}| = \frac{1}{\sqrt{3}}$. It follows that $|\vec{P}| \leq \frac{1}{\sqrt{3}}$ is not a sufficient criterion for $\frac{1}{3} I + \vec{P} \cdot \vec{\lambda}$ to be a density operator.

Problem 3

Problem 3 (Teleportation & Superdense Coding)**.** Consider the maximally entangled state in $\mathbb{C}^2 \otimes \mathbb{C}^d$ shared between Alice and Bob:

$$|\psi\rangle^{AB} = \frac{1}{\sqrt{d}} \left(|0\rangle^A |0\rangle^B + \cdots + |d-1\rangle^A |d-1\rangle^B \right).$$

(a) Find a protocol for faithful teleportation of a qudit from Alices lab to Bobs lab. What are the Kraus operators representing Alices joint measurement? What are the unitary operators performed by Bob? How many classical bits Alice transmits to Bob?

(b) Argue that the protocol you found in (a) can be used to teleport any mixed state ρ in d dimensions.

(c) Find a protocol for a super-dense coding using the above state. How many classical bits Alice can encode in her qudit?

Solution. (a) Let $|\varphi\rangle^{A'} = \sum_{m=0}^{d-1} \alpha_m |m\rangle^{A'}$ be an arbitrary state of a qudit (of an ancillary system labeled A') that Alice wants to teleport. Consider the unitary operators defined by

$$X = \sum_{j=0}^{d} |j+1\rangle \langle j| \qquad \text{and} \qquad Z = \sum_{j=0}^{d} \omega^j |j\rangle \langle j|$$

where $\omega = e^{i2\pi/d}$ is the principal d^{th} root of unity. For each $k, l \in \{0, \ldots d-1\}$, define unitary operators $U_{kl} = X^k Z^l$ and unit vectors $|\psi_{kl}\rangle = (I \otimes U_{kl}^*) |\psi\rangle$. Note that $ZX = \omega XZ$ and $U_{kl}^* = Z^{-l} X^{-k}$. This set of d^2 vectors form an orthonormal basis, since

$$\begin{aligned} \langle \psi_{k',l'} | \psi_{kl} \rangle &= \langle \psi | I \otimes (X^{k'} Z^{l'} Z^{-l} X^{-k}) |\psi\rangle \\ &= \omega^{k(l-l')} \langle \psi | I \otimes (X^{k'-k} Z^{l'-l}) |\psi\rangle \\ &= \delta_{k,k'} \delta_{l,l'}. \end{aligned}$$

Define the operators $M_{kl} = |\psi\rangle \langle\psi| U_{kl}$ for each k, l. Note that

$$\sum_{kl} M_{kl}^* M_{kl} = \sum_{kl} |\psi_{kl}\rangle \langle \psi_{kl}| = I$$

so these operators define a valid measurement. Suppose Alice performs the corresponding measurement on her systems. If Alice obtains outcome (k, l), Bob performs the unitary $\overline{U_{kl}} = (U_{kl}^T)^*$ such that the resulting state is

$$\begin{aligned} (M_{kl} \otimes \overline{U_{kl}})(|\varphi\rangle^{A'} \otimes |\psi\rangle^{AB}) &= (|\psi\rangle \langle\psi|^{A'A} \otimes I^B) \left(I^{A'} \otimes U_{kl}^A \otimes \overline{U_{kl}}^B \right) \left(|\varphi\rangle^{A'} \otimes |\psi\rangle^{AB} \right) \\ &= (|\psi\rangle \langle\psi|^{A'A} \otimes I) \left(|\varphi\rangle^{A'} \otimes |\psi\rangle^{AB} \right) \\ &= \frac{1}{d} \sum_{i,j,m} (\alpha_m \langle i|m\rangle \langle i|j\rangle) \left(|\psi\rangle^{A'A} \otimes |j\rangle^B \right) \\ &= \frac{1}{d} \sum_m \alpha_m |\psi\rangle^{A'A} \otimes |m\rangle^B = \frac{1}{d} |\psi\rangle^{A'A} \otimes |\varphi\rangle^B, \end{aligned}$$

where we use the fact that $U \otimes \overline{U} |\psi\rangle = |\psi\rangle$ for every unitary operator U acting on the maximally entangled state vector $|\psi\rangle$.

In this scheme, Alice tells Bob which of the d different outcomes she obtains, so she must transmit $\log_2 d$ bits of information.

(b) Define a channel Λ on the system $A' \otimes A \otimes B$ by $\Lambda(\sigma) = \sum_{kl}(M_{kl} \otimes \overline{U_{kl}})\sigma(M_{kl}^* \otimes U_{kl}^{\mathrm{T}})$. It is clear that

$$\Lambda(|\varphi\rangle\langle\varphi|^{A'} \otimes |\psi\rangle\langle\psi|^{AB}) = |\psi\rangle\langle\psi|^{A'A} \otimes |\varphi\rangle\langle\varphi|^{B}$$

holds for every pure state $|\varphi\rangle$. By linearity, it holds that $\Lambda(\rho^{A'} \otimes |\psi\rangle\langle\psi|^{AB}) = |\psi\rangle\langle\psi|^{A'A} \otimes \rho^{B}$ for every density operator ρ on $\mathbb{C}^d$.

(c) Alice and Bob start by sharing the maximally entangled state $|\psi\rangle^{AB}$ of two qudits. She can send one of the d^2 messages of the form $(k, l) \in \{0, \ldots, d-1\}^2$. Alice chooses her message to be (k, l) and performs the unitary U_{kl}^* on her half of the system and subsequently transmit her qudit to Bob. The resulting state that Bob obtains is

$$U_{kl}^* \otimes I\,|\psi\rangle\,.$$

Bob then performs the measurement corresponding to the measurement operators $\{M_{k'l'}\}$ defined by $M_{k'l'} = |\psi\rangle\langle\psi|\,(U_{k'l'} \otimes I)$ (the same measurement that Alice performed is the previous problem). Note that he obtains the result (k, l) with unit probability, since

$$M_{k'l'}(U_{kl}^* \otimes I)\,|\psi\rangle = \underbrace{\langle\psi|\,(U_{k'l'}U_{kl}^* \otimes I)\,|\psi\rangle}_{\delta_{kk'}\delta_{ll'}}\,|\psi\rangle\,,$$

and thus faithfully receives the intended message sent to him by Alice.

Problem 4

Problem 4. Consider the entangled state shared between Alice and Bob:

$$|\psi\rangle^{AB} = \sqrt{p_0}\,|0\rangle^A\,|0\rangle^B + \cdots + \sqrt{p_{n-1}}\,|n-1\rangle^A\,|n-1\rangle^B\,.$$

1. Prove the following theorem: Faithful teleportation of a qudit is possible if, and only if,

$$E_t(|\psi\rangle^{AB})\log_2 p_{\max} \geq \log_2 d,$$

where $p_{\max} = \max_k p_k$. That is, teleportation is possible if, and only if, none of the Schmidt coefficients are greater than $1/d$. This also implies that the Schmidt rank n is greater or equal to d.

2. Find a protocol for faithful teleportation of a qubit from Alices lab to Bobs lab assuming Alice and Bob sharing the partially entangled state

$$|\psi\rangle^{AB} = \frac{1}{\sqrt{2}}\,|00\rangle^{AB} + \frac{1}{2}\,|11\rangle^{AB} + \frac{1}{2}\,|22\rangle^{AB}\,.$$

In particular, determine the projective measurement performed by Alice and the unitary operators performed by Bob. What is the optimal classical communication cost? That is, how many classical bits Alice has to send to Bob?

Solution. We suppose that the teleportation scheme is performed by the following operations: (1) Alice and Bob first convert $|\psi\rangle$ to a maximally entangled state of two qudits (i.e. two d-dimensional systems), and (2) Alice and Bob use the maximally entangled state to teleport $|\varphi\rangle$ (an arbitrary pure state of a qudit).

(a) *Proof.* Suppose that $-\log_2 p_{\max} \geq \log_2 d$. It follows that $p_k \leq 1/d$ for every k and thus $n \geq d$. Without loss of generality, we may suppose that the p_k are in decreasing order, i.e. $p_1 \geq \cdots \geq p_{n-1} \geq 0$. For each $k = 1, \ldots, d-1$, it holds that

$$\sum_{i=0}^{k} p_i \leq \frac{k}{d} = \sum_{i=0}^{k} \frac{1}{d}$$

and for each $k = d, \ldots, n-1$ it holds that

$$\sum_{i=0}^{k} p_i \leq 1 = \sum_{i=0}^{d-1} \frac{1}{d}.$$

It follows that $(p_1, \ldots, p_{n-1}) \preceq \left(\frac{1}{d}, \ldots, \frac{1}{d}, 0, \ldots, 0\right)$ and thus $|\psi\rangle$ can be converted into the maximally entangled state of two qudits via LOCC.

On the other hand, suppose that $|\psi\rangle$ can be converted into the maximally entangled state of two qudits via LOCC. By the majorization condition, it follows that $p_1 \leq \frac{1}{d}$. Since $p_{n-1} \leq \cdots \leq p_1$, it holds that $p_{\max} = \max_k\{p_k\} \leq \frac{1}{d}$ and thus $-\log_2 p_{\max} \geq \log_2 d$. □

(b) Consider the following operators:

$$M_1 = \frac{1}{\sqrt{2}}\,|0\rangle\langle 0| + |1\rangle\langle 1| = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$M_2 = \frac{1}{\sqrt{2}}\,|0\rangle\langle 0| + |2\rangle\langle 2| = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Note that

$$M_1^* M_1 = \frac{1}{2} |0\rangle \langle 0| + |1\rangle \langle 1| = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$M_2^* M_2 = \frac{1}{2} |0\rangle \langle 0| + |2\rangle \langle 2| = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and thus $M_1^* M_1 + M_2^* M_2 = I$, so the operators $\{M_1, M_2\}$ define a valid measurement. Furthermore consider the unitary operators

$$U_1 = |0\rangle \langle 0| + |1\rangle \langle 1| + |2\rangle \langle 2|$$

and

$$U_2 = |0\rangle \langle 0| + |1\rangle \langle 2| + |2\rangle \langle 1|,$$

so that $U_1 = I$ is the identity matrix and U_2 is the permutation matrix that swaps 1 and 2. It holds that

$$M_1 \otimes U_1 |\psi\rangle = M_2 \otimes U_2 |\psi\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle).$$

The resulting state can be used by Alice and Bob to perform perfect teleportation using the standard means.

To first convert $|\psi\rangle$ into the Bell state, Alice must transmit one bit of information (either 1 or 2). To perform the teleportation protocol, Alice must transmit 2 bits of information to Bob. The composition of these two actions requires that Alice send a total of 3 bits to Bob. It's possible that Alice and Bob could transport the state $|\varphi\rangle$ directly without converting $|\psi\rangle$ into a Bell state first and use fewer bits, but that is beyond the scope of this assignment.

Problem 5

Problem 5 (Majorization)**.** .

(a) Find the extreme points of the convex set of all $n \times n$ column stochastic matrices.

(b) Show that a matrix A is doubly stochastic if and only if $A\vec{x}$ is majorized by $\vec{x}$ for all vectors $\vec{x}$ (do not use Birkhoff's theorem).

(c) Show that $\vec{x} \prec \vec{y}$ if and only if $\vec{x} = D\vec{y}$ for some doubly stochastic matrix D.

Solution. Note that in this solution we will be concerned with $n \times n$ matrices and n-dimensional vectors for some fixed positive integer n.

(a) We will show that the extreme points of the set of $n \times n$ column stochastic matrices is the set of $n \times n$ matrices that have exactly one 1 in each column and zeros elsewhere.

Proof. It is clear that these matrices are indeed column stochastic. To show that these matrices are also the extreme points of the set of column stochastic matrices, we will show that matrices not of this form are not extreme points.

Suppose A is not of this form. There is at least one pair of indices (i_1, j_1) such that $a_{i_1 j_1} \in (0, 1)$. Since A is column stochastic, there must be at least one other entry $a_{i_2 j_1}$ in the same column but different row, with $i_1 \neq i_2$, such that $a_{i_2 j_1} \in (0, 1)$. Let $\varepsilon = \min\{a_{i_1 j_1}, a_{i_2 j_1}\}$ and define a new $n \times n$ matrix B by

$$B = \varepsilon \left(E_{i_1 j_1} - E_{i_2 j_1}\right)$$

where E_{ij} is the $n \times n$ matrix with a 1 in the (i, j)-entry and zeros elsewhere for every i, j. The sum of the columns of B are all zero, so it follows that both $A + B$ and $A - B$ are column stochastic. Furthermore, we see that

$$A = \frac{1}{2}(A - B) + \frac{1}{2}(A + B)$$

and thus A is not an extreme point. □

(b) For each $i, j \in \{1, \ldots, n\}$, let $a_{ij} = \vec{e_i}^{\mathrm{T}} A \vec{e_j}$ denote the matrix elements of A, where the vectors $\{\vec{e_j}\}$ are the standard basis vectors of $\mathbb{R}^n$.

Proof. Suppose A is doubly stochastic. Let $\vec{x}$ be a vector and let $\vec{y} = A\vec{x}$. Without loss of generality, we may assume that the elements of both $\vec{x}$ and $\vec{y}$ are in decreasing order. Indeed, if $\vec{x}$ and $\vec{y}$ are not in decreasing order, let P_x and P_y be permutation matrices that order the entries of $\vec{x}$ and $\vec{y}$ respectively (i.e. such that $\vec{x}^{\downarrow} = P_x \vec{x}$ and $\vec{y}^{\downarrow} = P_y \vec{y}$). Note that $\vec{x} \prec \vec{y} \iff \vec{x}^{\downarrow} \prec \vec{y}^{\downarrow}$ and that

$$\vec{y}^{\downarrow} = A' \vec{x}^{\downarrow}$$

where $A' = P_y A P_x^{-1}$ is another doubly stochastic matrix.

From the above argument, it is clear that we only need to prove the case where the entries of both $\vec{x}$ and $\vec{y}$ are in decreasing order. For every $i, k \in \{1, \ldots, d\}$, define the values

$$t_{ik} = \sum_{j=1}^{k} a_{ij}.$$

It holds that $0 \leq t_{ik} \leq 1$ for every i and k. Furthermore, note that $k = \sum_{i=1}^{n} t_{ik}$ for every k, since this is the sum of the entries in the first k columns of A (and each column must sum to 1). Now, for every k we have that

$$\sum_{j=1}^{k} y_j - \sum_{j=1}^{k} x_j = \sum_{j=1}^{k} \sum_{i=1}^{n} a_{ij} x_i - \sum_{j=1}^{k} x_j$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} t_{ik} x_i - \sum_{j=1}^{k} x_j \\
&= \sum_{i=1}^{n} t_{ik} x_i - \sum_{j=1}^{k} x_j + \left(k - \sum_{i=1}^{n} t_{ik} \right) x_k \\
&= \sum_{i=1}^{k} \underbrace{(t_{ik} - 1)}_{\leq 0} \underbrace{(x_i - x_k)}_{\geq 0} + \sum_{i=k+1}^{n} \underbrace{t_{ik}}_{\geq 0} \underbrace{(x_i - x_k)}_{\leq 0} \leq 0.
\end{aligned}
$$

Thus $\sum_{j=1}^{k} y_j \leq \sum_{j=1}^{k} x_j$ holds for every k. Furthermore, it is clear that $\sum_{j=1}^{n} y_j = \sum_{j=1}^{n} x_j$ since A is doubly stochastic. It follows that $\vec{y} \prec \vec{x}$.

Now suppose that $A\vec{x} \prec \vec{x}$ for all vectors $\vec{x}$. We will show that A has nonnegative entries and that all of the rows and columns of A sum to unity. Consider the standard basis vectors $\{\vec{e}_j\}$. These are probability vectors, so it follows that $A\vec{e}_j$ is a probability vector for each j since it holds that $A\vec{e}_j \prec \vec{e}_j$. For each j, summing the entries of the j^{th} column of A yields

$$
\sum_{i=1}^{n} a_{ij} = \sum_{i=1}^{n} \vec{e}_i^{\mathrm{T}} A \vec{e}_j = 1
$$

since the sum of the entries of $A\vec{e}_j$ must sum to 1. Furthermore, this shows that all of the entries of A are nonnegative since each $A\vec{e}_j$ must have nonnegative entries. To show that the rows of A also sum to unity, consider the vector

$$
\vec{x} = \frac{1}{d} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.
$$

Recall that every d-dimensional probability vector majorizes x. Since it must hold that $A\vec{x} \prec \vec{x}$, it follows that $A\vec{x} = \vec{x}$. For every i, the sum of the i^{th} row of A is

$$
\sum_{j=1}^{n} a_{ij} = \vec{e}_i^{\mathrm{T}} A \left(\sum_{j=1}^{n} \vec{e}_j \right) = d\vec{e}_i^{\mathrm{T}} A x = d\vec{e}_i^{\mathrm{T}} x = d\frac{1}{d} = 1,
$$

as desired. □

(c) *Proof.* [1] If there exists a doubly stochastic matrix D such that $\vec{x} = D\vec{y}$, we see that $\vec{x} \prec \vec{y}$ is clear from part (b). We prove the converse by induction on n. If $n = 1$, the statement is trivial so we may assume that $n \geq 2$ and assume that the statement holds for all vectors in $\mathbb{R}^{n-1}$. Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ and suppose that $\vec{x} \prec \vec{y}$. Without loss of generality, we may suppose that the entries of $\vec{x}$ and $\vec{y}$ are in decreasing order. It is clear that $x_1 \leq y_1$ and that $y_n \leq x_k$. Hence $y_n \leq x_1 \leq y_1$ since $x_n \leq x_1$.

Let $s \in \{1, \ldots, n\}$ be the smallest integer satisfying $y_s \leq x_1 \leq y_1$. In the case that $s = 1$, it follows that

$$
\vec{w} = \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \prec \begin{pmatrix} y_2 \\ \vdots \\ y_n \end{pmatrix} = \vec{z}
$$

where $\vec{w}$ and $\vec{z}$ are vectors in $\mathbb{R}^{n-1}$. By the induction hypothesis, there exists an $(n-1) \times (n-1)$ doubly stochastic matrix B satisfying $\vec{w} = B\vec{z}$ and we can define an $n \times n$ doubly stochastic matrix D by

$$
D = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}
$$

[1] The idea for this proof came from the proof of Theorem 13.2 in *Theory of Quantum Information* by John Watrous

that satisfies $\vec{x} = D\vec{y}$.

In the case that $s > 1$, it holds that $y_s \leq x_1 < y_1$ and so there must exist a real number $p \in [0, 1)$ such that $x_1 = py_1 + (1-p)y_s$. Define the vectors $\vec{w} = (w_2, \ldots, w_n)$ and $\vec{z} = (z_2, \ldots, z_n)$ in $\mathbb{R}^{n-1}$ as

$$\begin{aligned}\vec{w} &= (x_2, \ldots, x_n)\\ \vec{z} &= (y_2, \ldots, y_{s-1}, (1-p)y_1 + py_s, y_{s+1}, \ldots, y_n).\end{aligned}$$

Note that $y_j > x_1$ holds for every $j = 1, \ldots, s-1$, since s is the smallest value for which $y_s \leq x_1$ holds. Hence, for every $k \in \{2, \ldots, s-1\}$ it holds that

$$\sum_{j=1}^{k} z_j = z_2 + \cdots + z_k = y_2 + \cdots + y_m > (k-1)x_1 \geq x_2 + \cdots x_m = w_2 + \cdots + w_k = \sum_{j=1}^{k} w_j.$$

Furthermore, for every $k \in \{s, \ldots, n\}$ it holds that

$$\begin{aligned}\sum_{j=1}^{k} z_j = z_2 + \cdots + z_m &= (1-p)y_1 + y_2 + \cdots + y_{s-1} + py_s + y_{s+1} + y_m\\ &= y_1 + \cdots + y_m - (py_1 + (1-p)y_s)\\ &= y_1 + \cdots + y_m - x_1\\ &\geq x_1 + \cdots x_m - x_1\\ &= x_2 + \cdots + x_m = w_2 + \cdots + w_m = \sum_{j=1}^{k} w_j.\end{aligned}$$

Since $\vec{x} \prec \vec{y}$, it must hold that $x_1 + \cdots + x_n = y_1 + \cdots + y_n$. We see that

$$\sum_{j=2}^{n} z_j = \sum_{j=1}^{n} y_j - (py_1 + (1-p)y_s) = \sum_{j=1}^{n} x_j - x_1 = \sum_{j=2}^{n} w_j.$$

This shows that $\vec{w} \prec \vec{z}$. By the induction hypothesis, there exists an $(n-1) \times (n-1)$ doubly stochastic matrix B such that $\vec{w} = B\vec{z}$. Define an $n \times n$ matrix A by

$$\begin{aligned}A\vec{e}_1 &= p\vec{e}_1 + (1-p)\vec{e}_s\\ A\vec{e}_s &= (1-p)\vec{e}_1 + p\vec{e}_s\end{aligned}$$

and $A\vec{e}_j = \vec{e}_j$ whenever $j \neq 1$ and $j \neq s$. This matrix is clearly doubly stochastic, so we can define another $n \times n$ doubly stochastic matrix D by

$$D = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} A,$$

which is doubly stochastic since it is the product of doubly stochastic matrices. It holds that

$$A\vec{y} = \begin{pmatrix} x_1 \\ \vec{z} \end{pmatrix}$$

and thus

$$D\vec{y} = \begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} x_1 \\ \vec{z} \end{pmatrix} = \begin{pmatrix} x_1 \\ \vec{w} \end{pmatrix} = \vec{x}.$$

□