1 Majorization and Random Permutations

a) The probability vector (p-vector) that is majorized by all p-vectors (including itself) is the uniform p-vector $u = (\frac{1}{n}, \ldots, \frac{1}{n})$.

We show that all other p-vectors majorize it by contradiction. Assume some p-vector $y^\downarrow = (y_1^\downarrow, \ldots, y_n^\downarrow)$ does not majorize u. Then for some $1 \le k \le n$, $\sum_{i=1}^{k} y_i^\downarrow < \sum_{i=1}^{k} \frac{1}{n}$, and there exists an element $y_j^\downarrow$ that is less than $\frac{1}{n}$. Conversely, we have for the same k $\sum_{i=k+1}^{n} y_i^\downarrow > \sum_{i=k+1}^{n} \frac{1}{n}$, and there exists an element $y_l^\downarrow$ that is greater than $\frac{1}{n}$. This contradicts our definition of $y^\downarrow$ being sorted in non-increasing order. Therefore all p-vector majorize u.

We also show u uniquely has this property by contradiction. Assume some vector has this property and is non-uniform. Then one of its elements must be greater than $1/n$. When sorted in non-increasing order, this vector is not majorized by u, which is a contradiction. Therefore u is uniquely majorized by all other vectors.

b) Suppose we are given some $n \times n$ doubly stochastic matrices $A_1, A_2, \ldots, A_m$, where the elements of a matrix A_k is denoted a_{ij}^k. We can calculate an arbitrary convex combination as $A = (a_{ij}) = \sum_{j+1}^{m} q_j A_j$ where $q_i \ge 0$ and $\sum_{j=1}^{m} q_j = 1$. Then we have the following:

$$a_{ij} = \sum_{k=1}^{m} q_k a_{ij}^k$$

$$\sum_{i=1}^{n} a_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{m} q_k a_{ij}^k = \sum_{k=1}^{m} q_k \sum_{i=1}^{n} a_{ij}^k = \sum_{k=1}^{m} q_k \cdot 1 = 1$$

$$\sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} \sum_{k=1}^{m} q_k a_{ij}^k = \sum_{k=1}^{m} q_k \sum_{j=1}^{n} a_{ij}^k = \sum_{k=1}^{m} q_k \cdot 1 = 1$$

Therefore, the convex combination of doubly stochastic matrices is also doubly stochastic.

c) If $Ax \prec x$ for all vectors x, then this must also hold for "deterministic" p-vector such as $D_k = (d_1, \ldots, d_n)$, in which the machine is always to be found in configuration k: $d_i = \delta_{ik}$. It must also hold for the uniform p-vector u.

For the deterministic states, the probability mass of unity can be concentrated in any of the n configurations. The resulting vector $E = AD_K = (e_1, \ldots, e_n)$ must be majorized by D_K and their elements must have the same sum. Therefore, every column k of A, $1 \le k \le n$, must sum to unity.

$$e_i = \sum_{j=1}^{n} A_{ij} d_j = A_{ik}$$

$$\sum_{i=1}^{n} e_i = 1 = \sum_{i=1}^{n} A_{ik}$$

For the uniform vector u, the resulting vector $V = Au = (v_1, \ldots, v_n)$ must be majorized by u and their elements must have the same sum. Therefore, every row of A must sum to unity.

$$v_i = \sum_{j=1}^{n} \frac{1}{n} A_{ij} = \frac{1}{n} \sum_{j=1}^{n} A_{ij}$$

$$\sum_{i=1}^{n} v_i = 1 = \sum_{i=1}^{n} \frac{1}{n} \sum_{j=1}^{n} A_{ij} = \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{n} A_{ij} = \frac{1}{n} \sum_{j=1}^{n} 1$$

In the last line we use the fact above that all columns must sum to unity.

Therefore $Ax \prec x$ for all x implies that A is doubly stochastic.

d) Assume A is doubly stochastic. To show that every vector x majorizes $y = Ax$, we must show that every maximal subset of $x^{\downarrow}$ majorizes the same maximal subset of $y^{\downarrow}$, and also that the sums of $y_i^{\downarrow}$ and $x_i^{\downarrow}$ are equal.

First we have by definition:

$$y_i^{\downarrow} = \sum_{j=1}^{n} A_{ij} x_j$$

Then we show that the same maximal subsets are majorized by induction, with our inductive hypothesis as:

$$\sum_{i'=1}^{k} y_{i'}^{\downarrow} \leq \sum_{i'=1}^{k} x_{i'}^{\downarrow}$$

We can show our base case $k = 1$ using the fact that all rows A_i sum to unity.

$$y_1^{\downarrow} = \sum_{j=1}^{n} A_{1j} x_j \leq \sum_{j'=1}^{n} A_{1j'} x_1^{\downarrow} = x_1 \sum_{j'=1}^{n} A_{1j'} = x_1$$

Then we assume our hypothesis is true for any $k-1$ and show it is true for k.

$$\sum_{i'=1}^{k} y_{i'}^{\downarrow} = \sum_{i'=1}^{k-1} y_{i'}^{\downarrow} + y_k^{\downarrow} \leq \sum_{i'=1}^{k-1} x_{i'}^{\downarrow} + \sum_{j'=1}^{n} A_{kj'} x_{j'}^{\downarrow} \leq \sum_{i'=1}^{k-1} x_{i'}^{\downarrow} + x_k^{\downarrow} = \sum_{i'=1}^{k} x_{i'}^{\downarrow}$$

This completes the inductive step, so it is true for all $1 \leq k \leq n$.

Since we know every column j in A sums to unity, we can show that the sum of elements in x and y are the same:

$$\sum_{i'=1}^{n} y_{i'}^{\downarrow} = \sum_{i'=1}^{n} \sum_{j'=1}^{n} A_{i'j'} x_{j'}^{\downarrow} = \sum_{j'=1}^{n} = x_{j'} \sum_{i'=1}^{n} A_{i'j'} = \sum_{j'=1}^{n} x_{j'}$$

e) For a permutation A to move the probability of one configuration p_i to another configuration p_j in a state vector p, there must be a 1 in entry A_{ij}. A random permutation contains at most n such moves; any probabilities which are not permuted have a '1' in some entry A_{kk}. Then each row and column of such a matrix A contains exactly one '1' entry and is zero everywhere else; therefore A is doubly stochastic.

A *random permutation* can be written as a convex combination of permutations, which is also doubly stochastic using the result of Part (b).

The state $s = \left[\frac{1}{12} \frac{1}{2} \frac{1}{12} \frac{1}{3}\right]^T$ cannot evolve into state $t = \left[\frac{1}{2} \frac{1}{6} \frac{1}{6} \frac{1}{6}\right]^T$, using any doubly stochastic transformation because s does not majorize t as it should from the result in Part (d).

2 Paulis, Cliffords, and Toffolis

a)

$$\begin{aligned} P^\dagger(a,b,k) &= (-i)^k((Z^{b_1})^\dagger(X^{a^1})^\dagger)\otimes\cdots\otimes((Z^{b_n})^\dagger(X^{a_n})^\dagger) \\ P(a,b,k)P\dagger(a,b,k) &= (-i)^k i^k(X^{a_1}Z^{b_1}(Z^{b_1})^\dagger(X^{a_1})^\dagger)\otimes\cdots\otimes(X^{a_n}Z^{b_n}(Z^{b_n})^\dagger(X^{a_n})^\dagger) \\ &= 1^k I\otimes\cdot\otimes I = I \end{aligned}$$

b)

$$P(a,b,k)P(c,d,l) = i^{k+l}(X^{a_1}Z^{b_1}X^{c_1}Z^{d_1})\otimes\ldots\otimes(X^{a_n}Z^{b_n}X^{c_n}Z^{d_n})$$

For each term in the tensor product, if the parity of b_ic_i and a_id_i are the same, then either $a_i = d_i = b_i = c_i$ or $a_i = d_i \neq b_i = c_i$. In either case, $X^2 = Z^2 = XZXZ = I$. If the parities are different, then either $((a_i = d_i) \wedge (b_i \neq c_i))$ or $((a_i \neq d_i) \wedge (b_i = c_i))$. In both cases there is a single X and a single Z. We use the fact that X and Z anti-commute to show that $(X^{a_i}Z^{b_i}X^{c_i}Z^{d_i}) = -(X^{c_i}Z^{d_i}X^{a_i}Z^{b_i})$.

If there are an even number of tensor product terms with different parities between b_ic_i and a_id_i, then the negative signs will cancel and $P(a,b,k)$ commutes with $P(c,d,l)$. Otherwise, there will be a negative sign left over and $P(a,b,k)$ anti-commutes with $P(c,d,l)$. This corresponds exactly to the factor $(-1)^m$ where $m = \left(\sum_{i=1}^n a_id_i + \sum_{i=1}^n b_ic_i\right) \mod 2$, as desired.

c) If $P(a,b,k)$ is Hermitian, then $P = P^\dagger$ and we know that $i^k = (-i)^k$ so that k must be even and we neglect this phase factor below. We then have the following:

$$\begin{aligned} R(P(a,b,k))R(P(a,b,k))\dagger &= \frac{1}{2}(I+iP(a,b,k))(I^\dagger - iP^\dagger) = \frac{1}{2}(I^2+iP-iP+P^2) = \frac{1}{2}(I+P^2) \\ P^2 &= (X^{a_1}Z^{b_1}X^{a_1}Z^{b_1})\otimes\ldots\otimes(X^{a_n}Z^{b_n}X^{a_n}Z^{b_n}) \end{aligned}$$

For each tensor product term in P^2, there are four cases which all lead to identity.

- If $a_i = b_i = 0$ then each term is I.
- If $a_i \neq b_i$, then each term is $X^2 = Z^2 = I$.
- If $a_i = b_i = 1$ then each term is $XZXZ = X^2 = I$.

Therefore, P^2 is I tensored with itself n times, and $RR^\dagger = \frac{1}{2}(I+I) = I$. In conclusion, if $P(a,b,k)$ is hermitian, then $R(P(a,b,k))$ is unitary.

d) $R(P(a,b,c))P(c,d,l)R(P(a,b,c))^\dagger$

$$\begin{aligned}
&= \frac{1}{2}(I + iP(a,b,k))P(c,d,l)(I - iP(a,b,k)) \\
&= \frac{1}{2}(P(c,d,l) + iP(a,b,k)P(c,d,l) - iP(c,d,l)P(a,b,k) + P(a,b,k)^2P(c,d,l)) \\
&= \frac{1}{2}(2P(c,d,l)) = P(c,d,l)
\end{aligned}$$

In the next to last equation we used the fact that $P(a,b,k)$ and $P(c,d,l)$ commute.

e) $R(P(a,b,c))P(c,d,l)R(P(a,b,c))^\dagger$

$$\begin{aligned}
&= \frac{1}{2}(I + iP(a,b,k))P(c,d,l)(I - iP(a,b,k)) \\
&= \frac{1}{2}(P(c,d,l) + iP(a,b,k)P(c,d,l) - iP(c,d,l)P(a,b,k) + P(a,b,k)^2P(c,d,l)) \\
&= \frac{1}{2}(2P(c,d,l) + iP(a,b,k)P(c,d,l)) \\
&= P(c,d,l) + iP(a,b,k)P(c,d,l)
\end{aligned}$$

In the next to last equation we used the fact that $P(a,b,k)$ and $P(c,d,l)$ anti-commute. I don't know how to get rid of the $P(c,d,l)$ term. Sorry.

3 Distinguishing Paulis

a) Since we are only allowed to measure once, we can only distinguish orthogonal states with perfect certainty. There are two such states for a single qubit, $|0\rangle$ and $|1\rangle$. However, we need to encode output for one of four Pauli operators, so we need at least 2 qubits.

In particular, given an input state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$,

$$I|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, X|\psi\rangle = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}, Y|\psi\rangle = \begin{bmatrix} -i\beta \\ i\alpha \end{bmatrix}, Z|\psi\rangle = \begin{bmatrix} \alpha \\ -\beta \end{bmatrix}$$

No single measurement can distinguish I from Z or X from Y.

b) The tensor product of each Pauli matrix with a 2×2 identity matrix applied to the entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ produces the four vectors below. Each pairwise inner product is the zero vector, hence the four vectors are orthogonal.

$$(I \otimes I)\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$(X \otimes I)\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$(Y \otimes I)\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ i \\ -i \\ 0 \end{bmatrix}$$

$$(Z \otimes I)\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

c) Using the result in the previous part, we can use the entangled state $|\psi_0 0\rangle \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ as input and the following 2-qubit gate V to distinguish between the Pauli operators as a black box U.

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$$

We can now verify that each Pauli operator as U on this input produces a different pure 2-qubit state in the computational basis. When measured, this 2-qubit returns 2 classical bits deterministically which can encode one of four choices of Pauli operator.

$$VI|\psi_{00}\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; VX|\psi_{00}\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}; VY|\psi_{00}\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}; VZ|\psi_{00}\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$