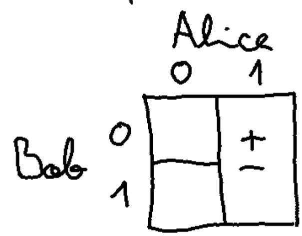
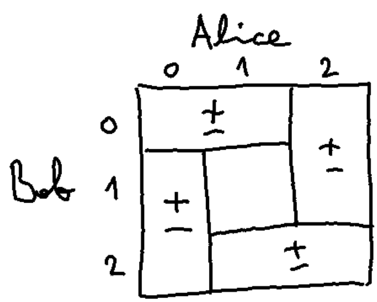
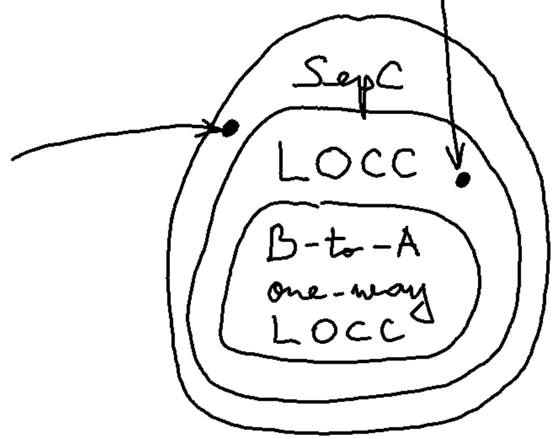
Lecture 11: Majorization and Nielsen's theorem

<u>Remark</u>: You saw in Problem 10.2 a set of orthogonal product states that can be perfectly discriminated by a separable measurement or a one-way LOCC from Alice to Bob, but not by one-way LOCC from Bob to Alice:



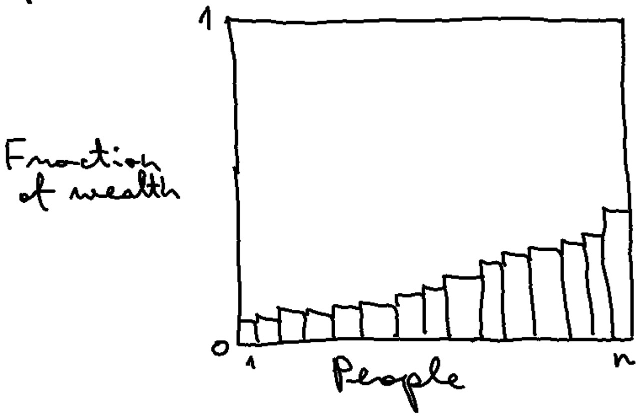
This idea can be extended to orthogonal product states that cannot be perfectly discriminated even by two-way LOCC:





How to measure wealth inequality?

We can model the distribution of wealth by a probability distribution:



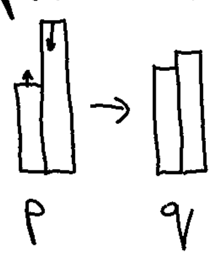
$p(i) \geq 0$

$$\sum_{i=1}^{n} p(i) = 1$$

Given two distributions p and q, how we tell which one is more equal?

Clearly, $p = (\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n})$ is the most equal and $q = (0, 0, \cdots, 0, 1)$ is the least equal. What about the rest?

Robin Hood transfer: if a richer person gives to a poorer one, the distribution becomes more equal:

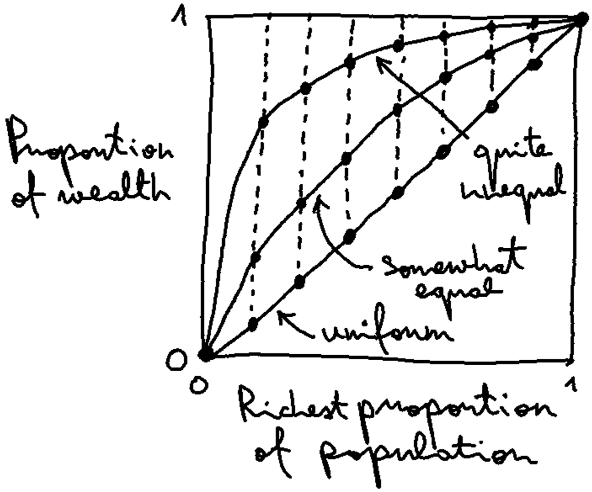


$$p = Mq \quad \text{where } M = cI + (1-c)X = \begin{pmatrix} c & 1-c \\ 1-c & c \end{pmatrix}$$

for some $0 < c < 1$.

Any sequence of such moves makes the distribution more equal. Note that the overall transformation is a convex combination of permutations.

Another way to compare wealth distributions is to look at what proportion of wealth is owned by the richest:



We are plotting the cumulative wealth of the richest fraction of population:

$$p(1) \geq p(2) \geq \cdots \geq p(n)$$

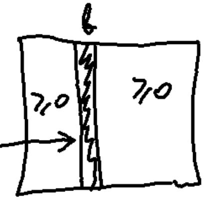
$$f(k) = \sum_{i=1}^{k} p(i)$$

If one curve is point-wise below another, the wealth distribution is more equal. Robin Hood pushes the curve downwards and makes the distribution more equal.

Stochastic and doubly stochastic operators

Def (p. 233) $A \in L(\mathbb{R}^\Sigma)$ is stochastic if

1. $A(a,b) \geq 0, \ \forall a,b \in \Sigma$
2. $\sum_{a\in\Sigma} A(a,b) = 1, \ \forall b \in \Sigma$ (columns sum to 1)



A is doubly stochastic if 1., 2., and

3. $\sum_{b\in\Sigma} A(a,b) = 1, \ \forall a \in \Sigma$ (rows sum to 1)

A is a permutation if 1., 2., 3., and $A(a,b) \in \{0,1\}, \forall a,b \in \Sigma$. So each row and column of a permutation operator contains exactly one entry equal to 1 and the rest are 0.

Note that any convex combination of permutations is doubly stochastic since each row and column is a convex combination of the standard basis vectors. Surprisingly, the converse is also true.

Thm 4.28 (Birkhoff – von Neumann):
$A \in L(\mathbb{R}^\Sigma)$ is doubly stochastic iff there exists a probability vector $p \in P(\mathrm{Sym}(\Sigma))$ such that

$$A = \sum_{\pi \in \mathrm{Sym}(\Sigma)} p(\pi) V_\pi$$

where $\mathrm{Sym}(\Sigma)$ is the set of all permutations acting on Σ and $V_\pi(a,b) = \delta_{a,\pi(b)}$.

Majorization for real vectors

Def 4.29: Let $u, v \in \mathbb{R}^\Sigma$. Then u *majorizes* v, $v \prec u$, if $v = Au$, for some doubly stochastic $A \in L(\mathbb{R}^\Sigma)$.

Let $r(u)$ denote the *reverse sorting* of u:

$$r_1(u) \geq r_2(u) \geq \dots \geq r_n(u)$$

and $\{r_1(u), r_2(u), \dots, r_n(u)\} = \{u(a) : a = 1, \dots, n\}$.

Thm 4.30: Let $u, v \in \mathbb{R}^\Sigma$. Then these are equivalent:

1. $v \prec u$
2. $\sum_{i=1}^m r_i(v) \leq \sum_{i=1}^m r_i(u)$, for all $m \in \{1, \dots, n-1\}$ and $\sum_{i=1}^n r_i(v) = \sum_{i=1}^n r_i(v)$.

Majorization for Hermitian operators

For probability distributions, permutations are precisely the reversible transformations. Similarly, for quantum states they correspond to unitary change of basis.

<u>Def 4.2</u>: $\Phi \in C(\mathcal{X})$ is a <u>mixed-unitary channel</u> if

$$\Phi(X) = \sum_{a \in \Sigma} p(a)\, U_a X U_a^*,$$

for some $p \in P(\Sigma)$ and $U_a \in U(\mathcal{X})$.

<u>Def 4.31</u>: Let $X, Y \in \mathrm{Herm}(\mathcal{X})$. Then X <u>majorizes</u> Y, $Y \prec X$, if $Y = \Phi(X)$, for some mixed-unitary $\Phi \in C(\mathcal{X})$.

<u>Thm 4.32</u> (Uhlmann): Let $X, Y \in \mathrm{Herm}(\mathcal{X})$. Then the following are equivalent:

1. $Y \prec X$
2. $\lambda(Y) \prec \lambda(X)$,

where $\lambda(X)$ denotes the spectrum of X.

Note that for diagonal matrices X, the spectrum $\lambda(X)$ is just the set of diagonal entries of X. Hence, majorization for diagonal matrices reduces to majorization for vectors.

If $\mathrm{vec}(X) = |u\rangle_{AB}$ and $\mathrm{vec}(Y) = |v\rangle_{AB}$, can we use the mixed-unitary channel Φ s.t. $Y = \Phi(X)$ to devise a one-way LOCC protocol for transforming $|u\rangle_{AB}$ to $|v\rangle_{AB}$? This is what Nielsen's thm is about!

Theorem 6.33 (Nielsen) Let $|u\rangle, |v\rangle \in S(\mathcal{X} \otimes \mathcal{Y})$. The following are equivalent:

1. $\operatorname{Tr}_{\mathcal{Y}}\big[|u\rangle\langle u|\big] \prec \operatorname{Tr}_{\mathcal{Y}}\big[|v\rangle\langle v|\big]$.
2. There exists a one-way LOCC protocol $\Xi \in \mathrm{LOCC}(\mathcal{X}:\mathcal{Y})$ from Bob to Alice such that $\Xi(|u\rangle\langle u|) = |v\rangle\langle v|$.
3. Same, but one-way LOCC from Alice to Bob.
4. Same, but $\Xi \in \mathrm{SepC}(\mathcal{X}:\mathcal{Y})$.

Proof: $1 \Rightarrow 2$ Recall that, for any $A, B \in L(\mathcal{Y}, \mathcal{X})$,

$$\operatorname{Tr}_{\mathcal{Y}}\big[\operatorname{vec}(A)\operatorname{vec}(B)^*\big] = AB^*. \qquad \text{(Exercise)}$$

In particular, if $X, Y \in L(\mathcal{Y}, \mathcal{X})$ are such that $\operatorname{vec}(X) = |u\rangle$ and $\operatorname{vec}(Y) = |v\rangle$ then 1. is equivalent to

$$XX^* \prec YY^*.$$

By Def. 4.31 of majorization for Hermitian operators,

$$XX^* = \sum_{a \in \Sigma} p(a) W_a YY^* W_a$$

where $p \in P(\Sigma)$ is a probability distribution and $W_a \in U(\mathcal{X})$. We need to convert this somehow into a one-way LOCC protocol. Note that we can write

$$XX^* = \underbrace{\Big(\sum_{a \in \Sigma} \sqrt{p(a)}\,(W_a Y) \otimes \langle a|\Big)}_{Z \in L(\mathcal{Y} \otimes \mathcal{Z}, \mathcal{X})} \cdot \underbrace{\Big(\sum_{a' \in \Sigma} \sqrt{p(a')}\,(Y^* W_{a'}^*) \otimes |a'\rangle\Big)}_{Z^*}.$$

Given that $XX^* = ZZ^*$, how are X and Z related? Let

$$X = \sum_{k=1}^{r} s_k |x_k\rangle\langle y_k|$$

be the singular value decomposition of X where $r = \operatorname{rank}(X)$, $s_k > 0$, and $|x_k\rangle \in S(\mathcal{X})$ and $|y_k\rangle \in S(\mathcal{Y})$ are orthonormal sets in $\mathcal{X}$ and $\mathcal{Y}$. Note that

$$XX^* = \sum_{j,k=1}^{r} s_j s_k |x_j\rangle\langle y_j|y_k\rangle\langle x_k| = \sum_{k=1}^{r} s_k^2 |x_k\rangle\langle x_k| = ZZ^*,$$

so XX^* and ZZ^* have eigenvalues s_k^2 with eigenvectors $|x_k\rangle$. Hence, Z has singular value decomposition

$$Z = \sum_{k=1}^{r} s_k |x_k\rangle\langle w_k|$$

for some orthonormal basis $\{|w_1\rangle, \dots, |w_r\rangle\}$ of $\mathcal{Y} \otimes \mathcal{Z}$. Let $V \in U(\mathcal{Y}, \mathcal{Y} \otimes \mathcal{Z})$ be an isometry such that $V|y_k\rangle = |w_k\rangle$, for all k. Then

$$XV^* = Z = \sum_{a \in \Sigma} \sqrt{p(a)} (W_a Y) \otimes \langle a|.$$

Let us use this observation to devise a one-way LOCC protocol from Bob to Alice. Let

$$\{B_a : a \in \Sigma\} \subset L(\mathcal{Y}) \text{ and } \{U_a : a \in \Sigma\} \subset U(\mathcal{X})$$

be Bob's measurement and Alice's basis change. Note that

$$\begin{aligned}(U_a \otimes B_a)|u\rangle &= (U_a \otimes B_a) \operatorname{vec}(X) \\ &= \operatorname{vec}(U_a X B_a^T).\end{aligned}$$

We would like this to be $\sqrt{p(a)} \operatorname{vec}(Y)$ since then

$$\begin{aligned}\sum_{a \in \Sigma} (U_a \otimes B_a)|u\rangle\langle u|(U_a \otimes B_a)^* &= \sum_{a \in \Sigma} p(a) \operatorname{vec}(Y) \operatorname{vec}(Y)^* \\ &= \operatorname{vec}(Y) \operatorname{vec}(Y)^* \\ &= |v\rangle\langle v|.\end{aligned}$$

So what we want is that

$$U_a X B_a^{\mathsf{T}} = \sqrt{p(a)}\, Y.$$

Recall that

$$X V^* = \sum_{a\in\Sigma} \sqrt{p(a)}\, (W_a Y) \otimes \langle a|,$$

hence

$$\underbrace{W_a^*}_{U_a} X \underbrace{V^* (I_Y \otimes |a\rangle_Z)}_{B_a^{\mathsf{T}}} = \sqrt{p(a)}\, Y.$$

So we take $U_a := W_a^*$ and $B_a := (I_Y \otimes \langle a|_Z) \overline{V}$.

Note that $\sum_{a\in\Sigma} B_a^* B_a = V^{\mathsf{T}} \sum_{a\in\Sigma} (I_Y \otimes |a\rangle\langle a|_Z) \overline{V} = V^{\mathsf{T}} \overline{V} = I_Y$,

so $\{B_a : a \in \Sigma\}$ is a valid measurement.

(2 ⟹ 3) Same, but exchange Alice and Bob.

(3 ⟹ 4) Every LOCC channel is separable.

(4 ⟹ 1) Let $\Xi \in \mathrm{SepC}(\mathcal{X}:\mathcal{Y})$ with Kraus operators $\{A_a : a \in \Sigma\} \subset L(\mathcal{X})$ and $\{B_a : a\in\Sigma\} \subset L(\mathcal{Y})$ such that

$$\Xi(|u\rangle\langle u|) = \sum_{a\in\Sigma} (A_a \otimes B_a) |u\rangle\langle u| (A_a \otimes B_a)^* = |v\rangle\langle v|.$$

Since $|v\rangle\langle v|$ is rank-one, each term must be of the form $p(a)\,|v\rangle\langle v|$, for some probability distr. $p \in P(\Sigma)$.
Equivalently,

$$\mathrm{vec}(A_a X B_a^{\mathsf{T}})\, \mathrm{vec}(A_a X B_a^{\mathsf{T}})^* = p(a)\, \mathrm{vec}(Y)\, \mathrm{vec}(Y)^*.$$

Taking partial trace over Y,

$$A_a X B_a^T \bar{B}_a X^* A_a^* = p(a) YY^*.$$

What we want to show is $XX^* \prec YY^*$, which by Theorem 4.32 is equivalent to

$$\lambda(XX^*) \prec \lambda(YY^*).$$

Note that $\sum_{k=1}^{n} \lambda_k(XX^*) = \mathrm{Tr}[XX^*] = 1 = \mathrm{Tr}[YY^*] = \sum_{k=1}^{n} \lambda_k(YY^*)$, since $|u\rangle$ and $|v\rangle$ are unit vectors. What remains to show is that, for all $m \in \{1, \dots, n\}$,

$$\sum_{k=m}^{n} \lambda_k(YY^*) \leq \sum_{k=1}^{m} \lambda_k(XX^*).$$

Note that $\lambda_k(cM) = c\lambda_k(M)$ for any $c > a$ and $M \in L(\mathcal{X})$, so

$$\sum_{k=m}^{n} \lambda_k(YY^*) = \sum_{k=m}^{n} \sum_{a\in\Sigma} \lambda_k(p(a) YY^*)$$
$$= \sum_{a\in\Sigma} \sum_{k=m}^{n} \lambda_k\Big(\underbrace{A_a X B_a^T \bar{B}_a X^* A_a^*}_{P_a \in \mathrm{Pos}(\mathcal{X})}\Big).$$

Since we are summing over the $n-m+1$ smallest eigenvalues of P_a, for any projector $\Pi_{a,m} \in \mathrm{Pos}(\mathcal{X})$ such that $\mathrm{rank}(\Pi_{a,m}) \geq n-m+1$,

$$\sum_{k=m}^{n} \lambda_k(P_a) \leq \mathrm{Tr}[\Pi_{a,m} \cdot P_a].$$

Let us choose $\Pi_{a,m}$ so that $\Pi_{a,m} A_a |x_i\rangle = 0$, for all $i \in \{1, \dots, m-1\}$, where $|x_i\rangle$ is the i-th left singular vector of X (or 0 if $i > r = \mathrm{rank}(X)$). Let us truncate the singular value decomposition of X and define

$$X_m := \sum_{k=m}^{r} s_k |x_k\rangle\langle y_k|.$$

By the definition of $\Pi_{a,m}$,

$$\mathrm{Tr}\left[\Pi_{a,m} \cdot P_a\right] = \mathrm{Tr}\left[\Pi_{a,m} \cdot P_{a,m}\right]$$

where $P_{a,m} := A_a X_m B_a^T \overline{B}_a X_m^* A_a^*$.

So far we have $\sum_{k=m}^{n} \lambda_k(YY^*) = \sum_{a\in\Sigma} \lambda_k(P_a)$, where

$$\lambda_k(P_a) \leq \mathrm{Tr}\left[\Pi_{a,m} \cdot P_{a,m}\right] \leq \mathrm{Tr}\left[P_{a,m}\right]$$

Note that

$$\sum_{a\in\Sigma} \mathrm{Tr}\left[P_{a,m}\right] = \sum_{a\in\Sigma} \mathrm{Tr}\left[A_a X_m B_a^T \overline{B}_a X_m^* A_a^*\right] \quad \text{Exercise}$$

$$= \mathrm{Tr}\left[\sum_{a\in\Sigma} (A_a \otimes B_a)\, \mathrm{vec}(X_m)\, \mathrm{vec}(X_m)^* (A_a \otimes B_a)\right]$$

$$= \mathrm{Tr}\left[\Xi\left(\mathrm{vec}(X_m)\, \mathrm{vec}(X_m)^*\right)\right] \quad \text{Since } \Xi \text{ is trace-preserving}$$

$$= \mathrm{Tr}\left[\mathrm{vec}(X_m)\, \mathrm{vec}(X_m)^*\right]$$

$$= \mathrm{Tr}\left[X_m X_m^*\right]$$

$$= \mathrm{Tr}\left[\sum_{j,k=m}^{r} s_j |x_j\rangle\langle y_j| \cdot s_k |y_k\rangle\langle x_k|\right] = \sum_{k=m}^{r} s_k^2$$

Note that $XX^* = \sum_{i,j=1}^{n} s_i |x_i\rangle\langle y_i| \cdot s_j |y_j\rangle\langle x_j|$

$$= \sum_{i=1}^{n} s_i^2 |x_i\rangle\langle x_i|,$$

Since rank$(X) = r$

so $s_k^2 = \lambda_k(XX^*)$ and $\sum_{k=m}^{r} s_k^2 = \sum_{k=m}^{n} s_k^2 = \sum_{k=m}^{n} \lambda_k(XX^*)$.

Hence $\sum_{k=m}^{n} \lambda_k(YY^*) \leq \sum_{k=m}^{n} \lambda_k(XX^*)$. $\square$