

Problem 1: Schur convexity & concavity

For a probability vector $\vec{p} = (p_1, \dots, p_D)$, $p_k \geq 0$, $\sum p_k = 1$, define $\vec{p}^\downarrow = (p_1^\downarrow, \dots, p_D^\downarrow)$ as the vector obtained by ordering the entries of \vec{p} in descending order.

We say that \vec{p} majorizes \vec{q} – denoted by $\vec{p} \succeq \vec{q}$ (or $\vec{q} \preceq \vec{p}$) – if and only if

$$\forall d = 1, \dots, D : \sum_{i=1}^d p_i^\downarrow \geq \sum_{i=1}^d q_i^\downarrow \quad (1)$$

with equality for $d = D$. (Intuitively, this says that the distribution \vec{q} is more flat (“more random”) than \vec{p} – try to convince yourself of this intuition by looking at some examples.)

An important property is that majorization introduces a natural ordering on probability distributions: It can be proven that $\vec{p} \succeq \vec{q}$ if and only if a random source with distribution \vec{q} can be obtained by randomizing a source with distribution \vec{p} , that is, there exists a *doubly stochastic matrix* S_{ij} (i.e., $\sum_i S_{ij} = \sum_j S_{ij} = 1$) such that $\vec{q} = D\vec{p}$. *Birkhoff’s theorem* states that any such S can in turn be written as $S = \sum r_i \Pi_i$, with probabilities $r_i \geq 0$, $\sum r_i = 1$, and permutations Π_i (i.e., S can be implemented by applying the permutation Π_i with probability r_i ; again, this is an if and only if statement – the converse should be obvious).

We thus arrive at the following characterization of majorization:

$$\vec{p} \succeq \vec{q} \iff \exists r_i, \Pi_i : \vec{q} = \sum_i r_i \Pi_i \vec{p}. \quad (2)$$

This does not need to be proven, and can be used for the problem.

1. Let $F(\vec{x}) = \sum_i f(x_i)$, where $f(x)$ is a convex function. Prove that F is *Schur convex*, that is,

$$\vec{q} \preceq \vec{p} \implies F(\vec{q}) \leq F(\vec{p}). \quad (3)$$

2. Prove that the Shannon entropy $H(\vec{p}) = -\sum p_i \log p_i$ and the α Rényi entropies

$$H_\alpha(\vec{p}) = \frac{\log \sum_i p_i^\alpha}{1 - \alpha},$$

$\alpha \neq 1$, are *Schur concave functions*, i.e.,

$$\vec{q} \preceq \vec{p} \implies F(\vec{q}) \geq F(\vec{p}). \quad (4)$$

Problem 2: Truncation error vs. Rényi entropy

In this problem, we determine the error in approximating a bipartite pure state with a given Rényi entanglement entropy by a state with a lower Schmidt rank. This step is central in obtaining an parameter-efficient approximation to quantum many-body states which satisfy an entanglement area law for a suitable α -Rényi entropy. This follows the derivation in <https://arxiv.org/abs/cond-mat/0505140>.

Throughout this problem, we consider some fixed α with $0 \leq \alpha < 1$.

1. Show that majorization introduced a *partial order* on the space of probability distributions (in particular, $\vec{p} \preceq \vec{q}$ and $\vec{q} \preceq \vec{r}$ implies $\vec{p} \preceq \vec{r}$), but not a total order (i.e., there are \vec{p} and \vec{q} which are not related by majorization).
2. Fix some $\chi \geq 1$. Consider all probability distributions \vec{p} with a fixed value of

$$\epsilon(\chi) := \sum_{i > \chi+1} p_i^\downarrow. \quad (5)$$

We will not determine the distribution \vec{p} satisfying (5) which minimized $H_\alpha(\vec{p})$.

- (a) Show that there is a one-parameter family of distributions \vec{p} , parametrized by the value $p_{\chi+1}^\downarrow =: h$, which majorizes all other distributions with this property (this is non-trivial because of part 1), and explicitly derive these extremal distributions.
- (b) Compute $H_\alpha(\vec{p})$ for these distributions. Find a suitable lower bound to this quantity and minimize it as a function of the parameter h .
- (c) What is the interpretation of this entropy for a given $\epsilon(\chi)$ (in the light of the results of problem 1)?
3. Use this to derive the maximum possible $\epsilon(\chi)$ for all distributions \vec{p} with a given value of $S_\alpha(\vec{p})$ (for one given α).
4. Let $|\psi\rangle \in \mathbb{C}^D \otimes \mathbb{C}^D$ be a bipartite state with Schmidt decomposition

$$|\psi\rangle = \sum_{i=1}^D s_i |i\rangle |i\rangle, \quad (6)$$

where $s_1 \geq s_2 \dots$, and α -Rényi entanglement entropy $S_\alpha(\text{tr}_B |\psi\rangle\langle\psi|) = E$. What is the error

$$\epsilon = \| |\psi\rangle - |\psi(\chi)\rangle \|_2 \quad (7)$$

made when approximating $|\psi\rangle$ by

$$|\psi(\chi)\rangle = \sum_{i=1}^{\chi} s_i |i\rangle |i\rangle ? \quad (8)$$

5. What is the error for the *normalized* approximation, $\| |\psi\rangle - |\hat{\psi}(\chi)\rangle \|_2$, with $|\hat{\psi}(\chi)\rangle = \frac{|\psi(\chi)\rangle}{\| |\psi(\chi)\rangle \|_2}$?
6. If we want to obtain an approximation with a given accuracy ϵ_0 in (7), how do we have to scale χ ?