Q. Inf. Science 3 (8.S372 / 18.S996) — Fall 2020

Assignment 2

Due: **Friday**, *Sep 18, 2020 at* ***5pm***

Turning in your solutions: Upload a single pdf file (typed or neatly handwritten) to canvas.

Collaboration policy: You may work individually or together in small groups but should write up your solutions individually.

Strongly recommended collaboration approach: We recommend that you find 1-2 other people to work with. You can use psetpartners.mit.edu to find partners if you don't already know people in the class.

Then for each problem, attempt it first on your own, and work until you get stuck. When you meet with the group, discuss each problem *even if you've already solved it.* If the whole group is stuck then we can answer questions on Piazza, in office hours, and can also schedule meetings at other times.

1. **Quantum channels**

   (a) Show that any linear operator $\mathcal{N}$ from $L(\mathbb{C}^{d_1})$ to $L(\mathbb{C}^{d_2})$ can be written in the form $\mathcal{N}(X) = \sum_a A_a X B_a^\dagger$ for some matrices A_a, B_a. What dimension are these matrices?

   (b) **Non-uniqueness of Kraus operators.** When we write a channel in the Stinespring representation as $\mathcal{N}(\rho) = \operatorname{tr}_E V\rho V^\dagger$, the outcome is the same if we perform a further isometry on system E before tracing it out. What effect does this have on the Kraus operators?

   (c) **Adjoint**. Define the *Hilbert-Schmidt* inner product between two matrices to be

   $$\langle X, Y\rangle := \operatorname{tr}\big[X^\dagger Y\big]. \tag{1}$$

   The adjoint of a superoperator $T \in L(L(A), L(B))$ with respect to this inner product is defined by the expression

   $$\langle X, T(Y)\rangle = \langle T^\dagger(X), Y\rangle. \tag{2}$$

   This is also known as the Heisenberg picture for quantum operations.

   i. If $T(\rho) = \sum_{i\in[k]} A_i \rho A_i^\dagger$ then what are the Kraus operators of $T^\dagger$?

   ii. tr_C is a quantum channel from $B \otimes C$ to B. What is $\operatorname{tr}_C^\dagger$?

   iii. Write down a quantum operation T that is not unitary and that satisfies $T = T^\dagger$.

2. **Types.** Given a sequence $x^n = x_1, x_2, \ldots, x_n \in [d]^n$ and a symbol $a \in [d]$, let $N(a|x^n)$ be the number of occurrences of a in x^n. The *type* (or empirical probability distribution) of x^n is the distribution that results from choosing a random letter from x^n, i.e. $P_{x^n}(a) = N(a|x^n)/n$. Here we use P_{x^n} to denote the type of x^n. Let $\mathcal{P}_n$ denote the set of all possible types of sequences in $[d]^n$; equivalently $\mathcal{P}_n$ is the set of probability distributions on $[d]$ whose entries are integer multiples of $1/n$. Let $\mathcal{T}_p^n := \{x^n : P_{x^n} = p\}$. Note that

$$|\mathcal{T}_p^n| = \binom{n}{np} := \frac{n!}{np_1!np_2!\cdots np_d!}. \tag{3}$$

   (a) List the elements of $\mathcal{P}_3$ when $d = 3$.

   (b) Prove the upper bound
   $$|\mathcal{P}_n| \leq (n+1)^{d-1}. \tag{4}$$

   (c) Prove that for $x^n \in \mathcal{T}_p^n$,
   $$p^n(x^n) := p(x_1)\cdots p(x_n) = 2^{-nH(p)}, \tag{5}$$
   where $H(p) := \sum_x p(x)\log(1/p(x))$.

   (d) For types $p, q \in \mathcal{P}_n$, compute $p^n(\mathcal{T}_q^n)$ where we use the notation $p^n(S)$ to mean $\sum_{x^n\in S} p^n(x^n)$. Express your answer in terms of $H(q)$ and $D(q\|p) = \sum_x q(x)\log\frac{q(x)}{p(x)}$.

   (e) It turns out that $p^n(\mathcal{T}_q^n)$ takes on its maximum value (as a function of q) when $q = p$. You do not need to prove this. Use this fact, along with the previous parts, to prove that
   $$\frac{2^{nH(p)}}{(n+1)^{d-1}} \leq |\mathcal{T}_p^n| \leq 2^{nH(p)}. \tag{6}$$

   (f) Pinsker's inequality (which you can use without proof) states that
   $$D(q\|p) \geq \frac{1}{2\ln 2}\|p-q\|_1^2. \tag{7}$$
   Combine this with the last two parts to prove that
   $$p^n(\mathcal{T}_q^n) \leq e^{-n\frac{\|p-q\|_1^2}{2}}. \tag{8}$$

   (g) One consequence of (8) is a weak version of a Chernoff bound. Suppose that we have a coin with probability a of heads and probability $1-a$ of tails. If we flip it n times show that the probability of $\geq nb$ heads for $b > a$ decreases exponentially with n.

   (h) We can also use types to define a sharper version of typical sets. Define
   $$\mathcal{T}_{p,\delta}^n = \bigcup_{q:\|p-q\|_1\leq\delta} \mathcal{T}_q^n. \tag{9}$$
   Prove that $1 - p^n(\mathcal{T}_{p,\delta}^n)$ is exponentially small for fixed p and fixed $\delta > 0$.

3. **Gibbs distributions** In this problem we define entropy with log base-e, i.e. ln. Also

let $\exp(x) := e^x$.

(a) Consider a classical system whose state lies in the set Ω. For simplicity assume that Ω is finite. The energy is defined by function $E : \Omega \to \mathbb{R}$. The Gibbs distribution at temperature T is

$$g_T(x) := \frac{e^{-E(x)/T}}{\sum_{x'\in\Omega} e^{-E(x')/T}}. \tag{10}$$

For given E, T, define the free energy of a probability distribution p by

$$F(p) := \mathop{\mathbb{E}}_{x\sim p}[E(x)] - TH(p) = \sum_{x\in\Omega} p(x)[E(x) + T\ln(p(x))] \tag{11}$$

Prove that g_T is a local minimum of the free energy. There are a few different ways to do this; probably calculus is the most straightforward.

(b) Now repeat the above exercise quantumly. Let H be a finite-dimensional Hermitian matrix. Define the Gibbs state

$$\gamma_T := \frac{e^{-H/T}}{\mathrm{tr}[e^{-H/T}]} \tag{12}$$

and the free energy

$$F(\rho) := \mathrm{tr}[H\rho] - TS(\rho). \tag{13}$$

Prove that γ_T minimizes F.

(c) Is F concave, convex or neither? Does this tell us anything about whether g_T and γ_T are global minima of F?

(d) For any state ρ, interpret $F(\rho) - F(\gamma_T)$ as a relative entropy. Use this to derive a robust version of (c), showing that even approximate minimizers of F are close to γ_T. You may use without proof the quantum Pinsker inequality $D(\rho\|\sigma) \geq \frac{1}{2}\|\rho - \sigma\|_1^2$; note that this formulation uses entropies defined with the natural log ($D(\rho\|\sigma) = \mathrm{tr}\,\rho[\ln(\rho) - \ln(\sigma)]$), and that the usual relative entropy has an extra factor of $\frac{1}{\ln 2}$ on the RHS.