Suppose that the state of A is initially $\omega(X)$ where X is a random variable which is stored in the system R. The state of RA is

$$\rho_{\mathsf{RA}} = \sum_{x \in \mathcal{A}_X} P_X(x) |x\rangle\langle x|_{\mathsf{R}} \otimes \omega(x)_{\mathsf{A}}$$

and $\rho_{\mathsf{A}} = \sum_{x \in \mathcal{A}_X} P_X(x)\omega(x)_{\mathsf{A}}$. If the system A now undergoes a time evolution represented by operation $\mathcal{N}^{\mathsf{A}\leftarrow\mathsf{A}}$, the state of RA becomes

$$\rho'_{\mathsf{RA}} = \sum_{x \in \mathcal{A}_X} P_X(x) |x\rangle\langle x|_{\mathsf{R}} \otimes \sigma(x)_{\mathsf{A}}, \text{ where } \sigma(x)_{\mathsf{A}} := \mathcal{N}^{\mathsf{A}\leftarrow\mathsf{A}}\omega(x)_{\mathsf{A}}.$$

In this case,

$$F(\rho_{\mathsf{RA}}, \rho'_{\mathsf{RA}}) = \mathrm{Tr}\sqrt{\sum_x P_X(x)^2 |x\rangle\langle x|_{\mathsf{R}} \otimes \omega(x)^{1/2}\sigma(x)\omega(x)^{1/2}} \tag{12.1}$$

$$= \mathrm{Tr}\sum_x P_X(x)|x\rangle\langle x|_{\mathsf{R}} \otimes \sqrt{\omega(x)^{1/2}\sigma(x)\omega(x)^{1/2}} \tag{12.2}$$

$$= \sum_x P_X(x) F(\omega(x)_{\mathsf{A}}, \sigma(x)_{\mathsf{A}}) = \mathbb{E}F(\omega(X)_{\mathsf{A}}, \sigma(X)_{\mathsf{A}}), \tag{12.3}$$

i.e. the expectation of the fidelity between the state of A before the operation and the state of A after the operation. The quantity $F_{op}(\mathcal{N}^{\mathsf{A}\leftarrow\mathsf{A}}, \rho_{\mathsf{A}})$ is a lower bound on this expectation.

Now, suppose that $\omega(x) = |x\rangle\langle x|$, so that prior to the operation system A also stores the random variable X

$$\rho_{\mathsf{RA}} = \sum_{x \in \mathcal{A}_X} P_X(x)|x\rangle\langle x|_{\mathsf{R}} \otimes |x\rangle\langle x|_{\mathsf{A}} \text{ and } \rho_{\mathsf{A}} = \sum_{x \in \mathcal{A}_X} P_X(x)|x\rangle\langle x|_{\mathsf{A}}.$$

Furthermore, let us suppose that the operation applied to A consists of measuring A in the computational basis, obtaining the result X, and then storing a random variable X' in A. The dependence of X' on X is described by the conditional probability distribution $P_{X'|X}$. Explicitly, this operation is

$$\mathcal{N}^{\mathsf{A}\leftarrow\mathsf{A}} : \rho_{\mathsf{A}} \mapsto \sum_{x',x} P_{X'|X}(x'|x)|x'\rangle\langle x|_{\mathsf{A}}\rho_{\mathsf{A}}|x\rangle\langle x'|_{\mathsf{A}}.$$

Note that this is a Kraus representation with Kraus operators

$$\{P_{X'|X}(x'|x)^{1/2}|x'\rangle\langle x| : x', x \in \mathcal{A}_X\}.$$

Now, $F(\omega(x)_{\mathsf{A}}, \sigma(x)_{\mathsf{A}}) = P_{X'|X}(x|x)^{1/2}$ and

$$F(\rho_{\mathsf{RA}}, \rho'_{\mathsf{RA}}) = \sum_x P_X(x)\sqrt{P_{X'|X}(x|x)} \le \sqrt{\sum_x P_X(x)P_{X'|X}(x|x)} = \sqrt{\Pr(X' = X)},$$

where the inequality is by Jensen's inequality. So, in this situation, when X is distributed according to P_X, $\Pr(X' = X)$ is bounded below by $F_{op}(\mathcal{N}^{\mathsf{A}\leftarrow\mathsf{A}}, \rho_{\mathsf{A}})^2$.

13 Data compression

Definition 1 (Classical data compression)**.** Suppose we have a probability distribution p on $\mathcal{A}$ Let Z be a random variable with $\mathcal{A}_Z = \mathcal{A}$ and distribution $P_Z = p$. We denote by $s_\epsilon(p)$ the smallest number k such that there exists an

- encoding function $c : \mathcal{A} \to \{1, \dots, k\}$ and
- decoding function $d : \{1, \dots k\} \to \mathcal{A}$

such that, if $\hat{Z} = d(c(Z))$, then $\Pr(\hat{Z} = Z) \geq 1 - \epsilon$.

Clearly, for any set $S \subseteq \mathcal{A}$ there is a d with image S if and only if $k \geq |S|$. If d has image S, then any c has $\Pr(\hat{Z} = Z) \leq \Pr(Z \in S)$ and this bound can be attained by picking c so that $c(z)$ is an element of $d^{-1}(z)$ for all $z \in S$. These observations show that $s_\epsilon(p)$ is simply the size of the smallest possible ϵ-*sufficient set* for p, i.e.

$$s_\epsilon(p) = \min \left\{ |S| : S \subseteq \mathcal{A}, \sum_{z \in S} p(z) \geq 1 - \epsilon \right\}.$$

Remark 2 (Form of minimal ϵ-sufficient sets)**.** It isn't hard to see that we can obtain a minimal ϵ-sufficient set by taking elements from $\mathcal{A}$ in order of decreasing probability until those accumulated have total probability $\geq 1 - \epsilon$.

Definition 3 (Quantum data compression)**.** Suppose we have a density operator ρ on $\mathcal{H}$. Let Q be a quantum system, with $\mathcal{H}_\mathsf{Q} = \mathcal{H}$, which has state ρ. We denote by $s_\epsilon(\rho)$ the smallest number k such that if K is a system with $d_\mathsf{K} = k$, then there exists an

- encoding operation $\mathcal{C}^{\mathsf{K} \leftarrow \mathsf{Q}}$ and
- decoding operation $\mathcal{D}^{\mathsf{Q} \leftarrow \mathsf{K}}$

such that, $F_{op}^2(\mathcal{D}^{\mathsf{Q} \leftarrow \mathsf{K}} \mathcal{C}^{\mathsf{K} \leftarrow \mathsf{Q}}, \rho_\mathsf{Q}) \geq 1 - \epsilon$. **Note that** we are using the **square** of F_{op} here.

13.1 Relating the quantum and classical cases

Theorem 4. Given a density operator ρ on $\mathcal{H}$ with $\dim(\mathcal{H}) = d$, let $\rho = \sum_{k=1}^{d} p(k) |\alpha_k\rangle\langle\alpha_k|$ be an eigendecomposition of ρ (so p is a distribution on $\{1, \dots, d\}$). Then, for any $\epsilon \in [0, 1]$,

$$s_\epsilon(p) \leq s_\epsilon(\rho) \leq s_{\epsilon/2}(p). \tag{13.1}$$

13.1.1 Proof of the upper bound in Theorem 4

Proposition 5. Let $\rho_{\mathsf{Q}} = \sum_k p(k)|\alpha_k\rangle\langle\alpha_k|_{\mathsf{Q}}$ be an eigendecomposition for ρ_{Q}, so p is a distribution on $\mathcal{A} = \{1,\ldots,d_{\mathsf{Q}}\}$. Given any $S \subseteq \mathcal{A}$, let $d_{\mathsf{K}} = |S|$ and let $V_{\mathsf{Q}\leftarrow\mathsf{K}}$ be an isometry which maps $\mathcal{H}_{\mathsf{K}}$ to $\mathrm{span}\{|\alpha_k\rangle_{\mathsf{Q}} : k \in S\}$ and let $\Pi_{\mathsf{Q}} := VV^\dagger = \sum_{k\in S}|\alpha_k\rangle\langle\alpha_k|_{\mathsf{Q}}$. Then

$$\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}} : \sigma_{\mathsf{K}} \mapsto V\sigma_{\mathsf{K}}V^\dagger, \tag{13.2}$$
$$\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}} : \rho_{\mathsf{Q}} \mapsto V^\dagger\rho_{\mathsf{Q}}V + |0\rangle\langle 0|_{\mathsf{K}}\mathrm{Tr}(\mathbb{1}_{\mathsf{Q}} - \Pi_{\mathsf{Q}})\rho_{\mathsf{Q}}. \tag{13.3}$$

are operations which achieve

$$F_{op}^2(\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}, \rho_{\mathsf{Q}}) \geq \left(\sum_{k\in S} p(k)\right)^2,$$

and it follows that $s_\epsilon(\rho) \leq s_{\epsilon/2}(p)$.

Proof. The isometric evolution $\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}$ is clearly an operation. The map $\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}$ is the sum of two completely positive maps and, as such, is completely positive. It is also trace preserving:

$$\mathrm{Tr}_{\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}\rho_{\mathsf{Q}} = \mathrm{Tr}_{\mathsf{K}}V^\dagger\rho_{\mathsf{Q}}V + \mathrm{Tr}(\mathbb{1}_{\mathsf{Q}} - \Pi_{\mathsf{Q}})\rho_{\mathsf{Q}} = \mathrm{Tr}_{\mathsf{Q}}VV^\dagger\rho_{\mathsf{Q}} + \mathrm{Tr}(\mathbb{1}_{\mathsf{Q}} - \Pi_{\mathsf{Q}})\rho_{\mathsf{Q}} = \mathrm{Tr}\rho_{\mathsf{Q}}$$

So $\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}$ is an operation, and there is a Kraus represention for $\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}$ of the form

$$\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}\rho_{\mathsf{Q}} = VV^\dagger\rho_{\mathsf{Q}}VV^\dagger + \ldots = \Pi_{\mathsf{Q}}\rho_{\mathsf{Q}}\Pi_{\mathsf{Q}} + \ldots.$$

From the expression for F_{op} in terms of Kraus operators (Proposition 15 in the previous handout), $F_{op}(\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}, \rho_{\mathsf{Q}})^2 \geq |\mathrm{Tr}\Pi_{\mathsf{Q}}\rho_{\mathsf{Q}}|^2 = \left(\sum_{k\in S} p(k)\right)^2$, and we know that there is a set S of size $s_{\epsilon/2}(p)$ such that $\left(\sum_{k\in S} p(k)\right)^2 \geq (1-\epsilon/2)^2 \geq 1-\epsilon$, so we achieve this fidelity with $d_{\mathsf{K}} = s_{\epsilon/2}(p)$. □

13.1.2 Proof of the lower bound in Theorem 4

We start with an almost obvious fact:

Proposition 6. Given d real numbers λ_j for $j \in \{0,\ldots,d-1\}$, such that $\lambda_j \geq \lambda_{j+1}$ and $r \in \mathbb{N}$,

$$\max\left\{\sum_{0\leq j<d}\lambda_j t_j : 0 \leq t_j \leq 1, \sum_{0\leq j<d} t_j = r\right\} = \sum_{0\leq j<r}\lambda_j.$$

Proof. Letting $l_j := 1 - t_j$, $\sum_{0\leq j<d} t_j = r$ is equivalent to $\sum_{0\leq j<r} l_j = \sum_{r\leq j<d} t_j =: R$, and

$$\sum_{0\leq j<d}\lambda_j t_j = \sum_{0\leq j<r}\lambda_j(1-l_j) + \sum_{r\leq j<d}\lambda_j t_j = \sum_{0\leq j<r}\lambda_j + \sum_{r\leq j<d}\lambda_j t_j - \sum_{0\leq j<r}\lambda_j l_j \tag{13.4}$$
$$\leq \sum_{0\leq j<r}\lambda_j + \lambda_r R - \lambda_{r-1}R \leq \sum_{0\leq j<r}\lambda_j, \tag{13.5}$$

where the inequalities are by the ordering of the eigenvalues and the positivity of the t_j, l_j. □

Proposition 7. For any $M \in \mathrm{Herm}(\mathcal{H})$ with eigendecomposition $M = \sum_{0\le j<d} \lambda_j |\alpha_j\rangle\langle\alpha_j|$, where $\lambda_j \ge \lambda_{j+1}$ and $d = \dim(\mathcal{H})$,

$$\max\{\mathrm{Tr} M\Pi : \Pi^\dagger\Pi = \Pi,\ \mathrm{rank}(\Pi) = r\} = \sum_{0\le j<r} \lambda_j \tag{13.6}$$

and this maximum is achieved by the projector $\Pi = \sum_{0\le j<r} |\alpha_j\rangle\langle\alpha_j|$.

Proof. It is clear that the projector $\sum_{0\le j<r} |\alpha_j\rangle\langle\alpha_j|$ achieves the stated bound, so it remains to establish that the bound holds for *any* rank-r projector Π.

$$\mathrm{Tr}\Pi M = \sum_{0\le j<d} \lambda_j t_j \text{ where } t_j := \langle\alpha_j|\Pi|\alpha_j\rangle.$$

Since the $|\alpha_j\rangle$ form an orthonormal basis, $\sum_{0\le j<d} t_j = \mathrm{Tr}\Pi = r$, and since $0 \le \Pi \le \mathbb{1}$ we have $0 \le t_j \le 1$. Therefore, the preceding proposition tells us that

$$\mathrm{Tr}\Pi M \le \sum_{0\le j<r} \lambda_j.$$

□

You may recognise the preceding proposition as a generalisation of the Rayleigh-Ritz theorem (which corresponds to the $r = 1$ case). We shall now use it prove that $s_\epsilon(p) \le s_\epsilon(\rho)$:

Proof. Without loss of generality we may assume that p satisfies $p(k) \ge p(k+1)$. Suppose that there exist encoding and decoding operations, $\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}$ and $\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}$, such that $1-\epsilon \le F_{op}^2(\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}, \rho_{\mathsf{Q}})$. We want to show that $s_\epsilon(p) \le d_{\mathsf{K}}$. Let

$$\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}} : \rho_{\mathsf{Q}} \mapsto \sum_i K_i \rho_{\mathsf{Q}} K_i^\dagger, \text{ where } K_i \in \mathcal{L}(\mathcal{H}_{\mathsf{Q}}, \mathcal{H}_{\mathsf{K}}), \text{ and} \tag{13.7}$$

$$\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}} : \sigma_{\mathsf{K}} \mapsto \sum_j L_j \sigma_{\mathsf{K}} L_j^\dagger, \text{ where } L_j \in \mathcal{L}(\mathcal{H}_{\mathsf{K}}, \mathcal{H}_{\mathsf{Q}}), \tag{13.8}$$

be Kraus representations for these operations. The operation $\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}$ has a Kraus representation

$$\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}} : \rho_{\mathsf{Q}} \mapsto \sum_{i,j} L_j K_i \rho_{\mathsf{Q}} K_i^\dagger L_j^\dagger, \text{ so } \sum_{i,j} K_i^\dagger L_j^\dagger L_j K_i = \mathbb{1}_{\mathsf{Q}}. \tag{13.9}$$

Let Π_j denote the projector onto the image of L_j, so $L_j = \Pi_j L_j$. The image of each L_j has dimension no larger than d_{K}, so $\mathrm{rank}(\Pi_j) \le d_{\mathsf{K}}$.

$$\begin{aligned}
1-\epsilon \le& F_{op}^2(\mathcal{D}^{\mathsf{Q}\leftarrow\mathsf{K}}\mathcal{E}^{\mathsf{K}\leftarrow\mathsf{Q}}, \rho_{\mathsf{Q}}) \overset{(a)}{=} \sum_{j,i} |\mathrm{Tr}\Pi_j L_j K_i \rho_{\mathsf{Q}}|^2 \overset{(b)}{=} \sum_{j,i} \left|\langle \Pi_j \rho_{\mathsf{Q}}^{1/2}, L_j K_i \rho_{\mathsf{Q}}^{1/2}\rangle\right|^2 \\
&\overset{(c)}{\le} \sum_{j,i} \langle \Pi_j \rho_{\mathsf{Q}}^{1/2}, \Pi_j \rho_{\mathsf{Q}}^{1/2}\rangle \langle L_j K_i \rho_{\mathsf{Q}}^{1/2}, L_j K_i \rho_{\mathsf{Q}}^{1/2}\rangle \le \sum_{j,i} (\mathrm{Tr}\Pi_j\rho_{\mathsf{Q}})(\mathrm{Tr} K_i^\dagger L_j^\dagger L_j K_i \rho_{\mathsf{Q}}) \\
&\overset{(d)}{\le} \left(\sum_{k=1}^{d_{\mathsf{K}}} p(k)\right) \sum_{j,i} \mathrm{Tr} K_i^\dagger L_j^\dagger L_j K_i \rho_{\mathsf{Q}} \overset{(e)}{=} \sum_{k=1}^{d_{\mathsf{K}}} p(k).
\end{aligned}$$

Equality (a) is by the expression for F_{op} in terms of Kraus operators (Proposition 15 in the previous handout) and $L_j = \Pi_j L_j$; (b) is by definition of the Hilbert-Schmidt inner product; (c) is Cauchy-Schwarz; (d) is by Proposition 7; (e) is by (13.9). Therefore, $s_\epsilon(p) \le d_{\mathsf{K}}$.

□