APPENDIX A

Mathematical supplement

A.1. Proof of the Hölder inequality for traces

We start with the proof of the Hölder inequality for matrices, Proposition 2.3. There are no short proofs in the case of matrices. The proof here is due to Fröhlich [1978] and it uses chessboard estimates. The proof of Proposition 2.3 can be found after that of Lemma A.3.

Lemma A.1 (Chessboard estimate). *For any $n \in \mathbb{N}$ and any matrices $a_1, \dots, a_{2n}$, we have*

$$\left|\mathrm{Tr}\, a_1 \dots a_{2n}\right| \leq \prod_{i=1}^{2n} \Big(\mathrm{Tr}\,(a_i a_i^*)^n\Big)^{1/2n}.$$

Proof. Since $(a, b) \mapsto \mathrm{Tr}\, a^* b$ is an inner product, we have the Cauchy–Schwarz inequality: $||trab|^2 \leq \mathrm{Tr}\, a^* a \mathrm{Tr}\, b^* b$. The following inequality follows:

$$\left|\mathrm{Tr}\, a_1 \dots a_{2n}\right|^2 \leq \mathrm{Tr}\,\big(a_1 \dots a_n a_n^* \dots a_1^*\big)\ \mathrm{Tr}\,\big(a_{2n}^* \dots a_{n+1}^* a_{n+1} \dots a_{2n}\big). \tag{A.1}$$

This allows to use a reflection positivity argument. By replacing a_i with $a_i/\sqrt{\mathrm{Tr}\,(a_i a_i^*)^n}$ it is enough to prove the inequality for matrices that satisfy $\mathrm{Tr}\,(a_i a_i^*)^n = 1$; the general result follows from scaling. Note that the set of such matrices is compact.

Let $a_1, \dots, a_{2n}$ be matrices that maximise $|\mathrm{Tr}\, a_1 \dots a_{2n}|$, with maximum number of matching neighbours $a_{i+1} = a_i^*$. Suppose there exists an index j such that $a_{j+1} \neq a_j^*$. Using cyclicity, we can assume that $j = n$. By the inequality (A.1), $a_1, \dots, a_n, a_n^*, \dots, a_1^*$ and $a_{2n}^*, \dots, a_{n+1}^*, a_{n+1}, \dots, a_{2n}$ are also maximisers. At least one has strictly more matching neighbours, hence a contradiction. The maximum is then $\mathrm{Tr}\,(aa^*)^n$ for some matrix $a \in \{a_1, \dots, a_n\}$, which is equal to 1. □

Chessboard estimates allow to prove what is essentially the case $r = 1$ of Hölder's inequality.

Corollary A.2. *We have*

$$|\mathrm{Tr}\, a_1 \dots a_n| \leq \prod_{i=1}^{n} \|a_i\|_{p_i}$$

for all n and all $p_i \geq 1$ such that $\sum_{i=1}^{n} \frac{1}{p_i} = 1$.

Proof. It suffices to consider rational p_i, by continuity. Let ℓ be a positive integer such that $2\ell/p_i$ is integer for all i. Let $a_i = U_i|a_i|$ be the polar decomposition of a_i, and let

$$b_i = |a_i|^{p_i/2\ell}, \qquad \hat{b}_i = U_i|a_i|^{p_i/2\ell}. \tag{A.2}$$

Then $a_i = \hat{b}_i b_i^{(2\ell/p_i)-1}$, and we have

$$\begin{aligned} \big|\mathrm{Tr}\, a_1 \dots a_n\big| &= \big|\mathrm{Tr}\, \hat{b}_1 \underbrace{b_1 \dots b_1}_{(2\ell/p_1)-1} \dots \hat{b}_n \underbrace{b_n \dots b_n}_{(2\ell/p_n)-1}\big| \\ &\leq \prod_{i=1}^{n} (\mathrm{Tr}\, |a_i|^{p_i})^{1/p_i} \qquad \text{(A.3)} \\ &= \prod_{i=1}^{n} \|a_i\|_{p_i}. \end{aligned}$$

The inequality follows from Lemma A.1 and from the identities

$$\mathrm{Tr}\, (b_i b_i^*)^{\ell} = \mathrm{Tr}\, (\hat{b}_i \hat{b}_i^*)^{\ell} = \mathrm{Tr}\, |a_i|^{p_i}. \tag{A.4}$$

□

Lemma A.3. *Let $r, r' \in [1, \infty]$ such that $\frac{1}{r} + \frac{1}{r'} = 1$. Then for any square matrix a, we have*

$$\|a\|_r = \max_{\|c\|_{r'}=1} \mathrm{Tr}\, c^* a.$$

Proof. The right side is smaller by Corollary A.2:

$$\big|\mathrm{Tr}\, c^* a\big| \leq \|c\|_{r'} \|a\|_r = \|a\|_r. \tag{A.5}$$

In order to check that this inequality is saturated, let $a = U|a|$ be the polar decomposition of a, and choose $c = \|a\|_r^{1-r} U|a|^{r-1}$. Then $\|c\|_{r'} = 1$ and $\mathrm{Tr}\, c^* a = \|a\|_r$. □

Proof of Proposition 2.3. Starting with Lemma A.3 and then using Corollary A.2 with $a_1 = c^*$, $a_2 = a$, $a_3 = b$ and $p_1 = r$, $p_2 = p$, $p_3 = q$, we have

$$\begin{aligned} \|ab\|_r &= \sup_{\|c\|_{r'}=1} \mathrm{Tr}\, c^* ab \\ &\leq \sup_{\|c\|_{r'}=1} \|c\|_{r'} \|a\|_p \|b\|_q. \end{aligned} \tag{A.6}$$

□

A.2. Trotter and Duhamel

We now review two useful expansions for the exponential of a sum of two non-commuting operators, namely the Trotter and Duhamel formulæ.

PROPOSITION A.4 (Trotter formula). *Let a, b be $n \times n$ matrices. Then*

$$e^{a+b} = \lim_{n\to\infty} \Big(e^{\frac{1}{n}a} \, e^{\frac{1}{n}b} \Big)^n = \lim_{n\to\infty} \Big[\big(1 + \tfrac{1}{n}a\big) \, e^{\frac{1}{n}b} \Big]^n.$$

PROOF. We prove the second formula — the mild changes for the first formula are straightforward. Let K_n be the matrix such that

$$\big(1 + \tfrac{1}{n}a\big) \, e^{\frac{1}{n}b} = 1 + \tfrac{1}{n}(a+b) + K_n. \tag{A.7}$$

It is clear that $\|K_n\| = O(\frac{1}{n^2})$. We have

$$\Big[\big(1 + \tfrac{1}{n}a\big) \, e^{\frac{1}{n}b} \Big]^n = \Big(1 + \tfrac{1}{n}(a+b) \Big)^n + R_n, \tag{A.8}$$

where R_n is a matrix whose norm satisfies

$$\|R_n\| \leq \sum_{k=0}^{n-1} \tbinom{n}{k} \, \|1 + \tfrac{1}{n}(a+b)\|^k \, \|K_n\|^{n-k} = O(\tfrac{1}{n}). \tag{A.9}$$

The first term in the right side of (A.8) converges to e^{a+b}. □

PROPOSITION A.5 (Duhamel formula). *Let a, b be $n \times n$ matrices. Then*

$$\begin{aligned} e^{a+b} &= e^{a} + \int_0^1 e^{ta} \, b \, e^{(1-t)(a+b)} \, dt \\ &= \sum_{k\geq 0} \int_{0<t_1<\dots<t_k<1} dt_1 \dots dt_k \, e^{t_1 a} \, b \, e^{(t_2-t_1)a} \, b \dots b \, e^{(1-t_k)a} . \end{aligned}$$

PROOF. Let $F(s)$ be the matrix-valued function

$$F(s) = e^{sa} + \int_0^s e^{ta} \, b \, e^{(s-t)(a+b)} \, dt. \tag{A.10}$$

We show that, for all s,

$$e^{s(a+b)} = F(s). \tag{A.11}$$

The derivative of $F(s)$ is

$$F'(s) = e^{sa} \, a + e^{sa} \, b + \int_0^s e^{ta} \, b \, e^{(s-t)(a+b)} \, (a+b) dt = F(s)(a+b). \tag{A.12}$$

On the other hand, the derivative of $\mathrm{e}^{s(a+b)}$ is $\mathrm{e}^{s(a+b)}\,(a+b)$. The identity (A.11) clearly holds for $s=0$ and, since both sides satisfy the same differential equation, they must be equal for all s.

We can iterate Duhamel's formula N times so as to get

$$\begin{aligned}\mathrm{e}^{a+b} &= \sum_{k=0}^{N} \int_{0<t_1<\cdots<t_k<1} \mathrm{d}t_1 \ldots \mathrm{d}t_k\, \mathrm{e}^{t_1 a}\, b\, \mathrm{e}^{(t_2-t_1)a}\, b \ldots b\, \mathrm{e}^{(1-t_k)a} \\ &+ \int_{0<t_1<\cdots<t_N<1} \mathrm{d}t_1 \ldots \mathrm{d}t_k\, \mathrm{e}^{t_1 a}\, b\, \mathrm{e}^{(t_2-t_1)a}\, b \ldots b \Big[\mathrm{e}^{(1-t_N)(a+b)} - \mathrm{e}^{(1-t_N)a} \Big].\end{aligned} \tag{A.13}$$

Using $\|\,\mathrm{e}^{ta}\,\| \leq \mathrm{e}^{t\|a\|}$, the last line is less than $2\,\mathrm{e}^{\|a\|+\|b\|}\,\frac{\|b\|^N}{N!}$ and so it vanishes in the limit $N\to\infty$. The summand is less than $\mathrm{e}^{\|a\|}\,\frac{\|b\|^k}{k!}$, so that the sum is absolutely convergent. ☐

A.3. About convex functions

Here is a simple result about convex functions on $\mathbb{R}$. Recall that the right and left derivatives of a convex function $f(s)$ at $s=0$ are given respectively by $\partial_+ f(0) = \inf_{s>0} \frac{f(s)-f(0)}{s}$ and $\partial_- f(0) = \sup_{s<0} \frac{f(s)-f(0)}{s}$.

Proposition A.6. *Let f_n be a sequence of continuously differentiable, convex functions on $\mathbb{R}$ such that $f_n \to f$ pointwise. For each $m \in [\partial_- f(0), \partial_+ f(0)]$ there is a sequence $s_n \to 0$ such that $m = \lim_{n\to\infty} f_n'(s_n)$.*

Proof. We claim the following: for any $\varepsilon, \delta > 0$ there is $N = N(\varepsilon,\delta)$ such that for any $n > N$ we have

$$f_n'(\delta) > \partial_+ f(0) - \varepsilon, \qquad f_n'(-\delta) < \partial_- f(0) + \varepsilon. \tag{A.14}$$

The result then follows using the mean value theorem for the continuous function $f_n'(s)$: if it is the case that $m \in [\partial_- f(0)+\varepsilon, \partial_+ f(0) - \varepsilon]$ then there is some $s_n \in [-\delta,\delta]$ satisfying $f_n'(s_n) = m$, otherwise we may take $s_n = \delta$ or $s_n = -\delta$.

We prove the claim for $\partial_+ f(0)$. We have

$$f_n'(\delta) \geq \frac{f_n(\delta) - f_n(\delta/2)}{\delta/2}, \qquad \frac{f(\delta/2) - f(0)}{\delta/2} \geq \partial_+ f(0). \tag{A.15}$$

Moreover,

$$\lim_{n\to\infty} \frac{f_n(\delta) - f_n(\delta/2)}{\delta/2} - \frac{f(\delta/2)-f(0)}{\delta/2} = \frac{f(\delta) - f(\delta/2)}{\delta/2} - \frac{f(\delta/2)-f(0)}{\delta/2} \geq 0. \tag{A.16}$$

So for n large enough we have

$$f_n'(\delta) \geq \frac{f_n(\delta) - f_n(\delta/2)}{\delta/2} \geq \frac{f(\delta/2)-f(0)}{\delta/2} - \varepsilon \geq \partial_+ f(0) - \varepsilon, \tag{A.17}$$

as claimed. □

A.4. Proof of the quantum Pinsker inequality

This section is devoted to the proof of Lemma 10.3 (quantum Pinsker's inequality). It relies on the convexity of the relative entropy, which can be proved using Lieb's concavity theorem, and on a clever path that allows to use the classical Pinsker inequality for Bernoulli variables.

LEMMA A.7. *Let $a, b, h \in \mathcal{M}_n$ such that $a, b \geq 0$, $[a, b] = 0$, and $h = h^*$. Assume that $\begin{pmatrix} a & h \\ h & b \end{pmatrix} \geq 0$. Then $h \leq a^{1/2} b^{1/2}$.*

PROOF. Let $u, v \in \mathbb{C}^n$ such that $\|u\| = \|v\| = 1$ and define $x = a^{-1/2} u$, $y = b^{-1/2} v$. Then, with $\mathbf{0}$ the zero vector in $\mathbb{C}^n$, we have

$$0 \leq \begin{pmatrix} x^* & \mathbf{0}^* \\ \mathbf{0}^* & y^* \end{pmatrix} \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} x & \mathbf{0} \\ \mathbf{0} & y \end{pmatrix} = \begin{pmatrix} x^* a x & x^* h y \\ y^* h x & y^* h y \end{pmatrix} = \begin{pmatrix} 1 & u^* a^{-1/2} h b^{-1/2} v \\ v^* b^{-1/2} h a^{-1/2} u & 1 \end{pmatrix}. \tag{A.18}$$

The latter is a 2×2 matrix; its determinant is nonnegative so that $|u^* a^{-1/2} h b^{-1/2} v| \leq 1$. This implies that $\|a^{-1/2} h b^{-1/2}\| \leq 1$. Next one can check that $a^{-1/4} b^{-1/4} h a^{-1/4} b^{-1/4}$ has the same eigenvalues as $a^{-1/2} h b^{-1/2}$ (if v is eigenvector of the first matrix, then $a^{-1/4} b^{1/4} v$ is eigenvector of the second matrix with the same eigenvalue). Then $\|a^{-1/4} b^{-1/4} h a^{-1/4} b^{-1/4}\| \leq 1$ and since this matrix is hermitian, we have

$$a^{-1/4} b^{-1/4} h a^{-1/4} b^{-1/4} \leq \mathbb{1} \qquad \Longleftrightarrow \qquad h \leq a^{-1/2} b^{-1/2}. \tag{A.19}$$

□

THEOREM A.8 (Lieb's concavity). *Let $a_1, a_2, b_1, b_2 \geq 0$ be $n \times n$ complex matrices and let $\alpha \in [0, 1]$. Then*

$$(a_1 + a_2)^\alpha \otimes (b_1 + b_2)^{1-\alpha} \geq a_1^\alpha \otimes b_1^{1-\alpha} + a_2^\alpha \otimes b_2^{1-\alpha}.$$

PROOF. Let $x(\alpha) = a_1^\alpha \otimes b_1^{1-\alpha}$, $y(\alpha) = a_2^\alpha \otimes b_2^{1-\alpha}$, $z(\alpha) = (a_1 + a_2)^\alpha \otimes (b_1 + b_2)^{1-\alpha}$. We need to show that $z(\alpha) \geq x(\alpha) + y(\alpha)$ for all $\alpha \in [0, 1]$. It is actually enough to show this in a dense subset since all expressions are continuous in α. This clearly holds for $\alpha \in \{0, 1\}$. We now show that if it holds for α and β, then it also holds for $\frac{\alpha+\beta}{2}$.

We have $x(\frac{\alpha+\beta}{2}) = x(\alpha)^{1/2} x(\beta)^{1/2}$, and the same relations for y and z. Then

$$\begin{pmatrix} x(\alpha) & x(\frac{\alpha+\beta}{2}) \\ x(\frac{\alpha+\beta}{2}) & x(\beta) \end{pmatrix} = \begin{pmatrix} x(\alpha)^{1/2} \\ x(\beta)^{1/2} \end{pmatrix} \begin{pmatrix} x(\alpha)^{1/2} & x(\beta)^{1/2} \end{pmatrix} \geq 0. \tag{A.20}$$

The latter inequality holds quite generally, only using that $x(\alpha)$ is hermitian. We have a similar inequality for y. Then

$$0 \leq \begin{pmatrix} x(\alpha) & x(\frac{\alpha+\beta}{2}) \\ x(\frac{\alpha+\beta}{2}) & x(\beta) \end{pmatrix} + \begin{pmatrix} y(\alpha) & y(\frac{\alpha+\beta}{2}) \\ y(\frac{\alpha+\beta}{2}) & y(\beta) \end{pmatrix} \leq \begin{pmatrix} z(\alpha) & x(\frac{\alpha+\beta}{2}) + y(\frac{\alpha+\beta}{2}) \\ x(\frac{\alpha+\beta}{2}) + y(\frac{\alpha+\beta}{2}) & z(\beta) \end{pmatrix}. \tag{A.21}$$

The second holds because the difference is equal to $\left(\begin{smallmatrix} z(\alpha)-x(\alpha)-y(\alpha) & 0 \\ 0 & z(\beta)-x(\beta)-y(\beta) \end{smallmatrix}\right)$, which is nonnegative by assumption. We now use Lemma A.7 and we get

$$x\big(\tfrac{\alpha+\beta}{2}\big) + y\big(\tfrac{\alpha+\beta}{2}\big) \leq z(\alpha)^{1/2} z(\beta)^{1/2} = z\big(\tfrac{\alpha+\beta}{2}\big). \tag{A.22}$$

We can start with $\alpha = 0$ and $\beta = 1$ and iterate the inequality, so it applies to all multiples of 2^{-k} for arbitrary k; this set is dense in $[0, 1]$. □

COROLLARY A.9. *Let $a_1, a_2, b_1, b_2 \geq 0$ be $n \times n$ complex matrices and let $\alpha \in [0, 1]$. Then*

$$\operatorname{Tr}\big((a_1 + a_2)^\alpha (b_1 + b_2)^{1-\alpha}\big) \geq \operatorname{Tr}(a_1^\alpha b_1^{1-\alpha}) + \operatorname{Tr}(a_2^\alpha b_2^{1-\alpha}).$$

PROOF. We use the following correspondence between $\mathcal{M}_n$ and $\mathbb{C}^n \otimes \mathbb{C}^n$:

$$\operatorname{Tr} a^{\mathrm{T}} b = \sum_{i,j=1}^n a_{i,j} b_{i,j} = \sum_{i,j=1}^n \langle i| \otimes \langle j| \, (a \otimes b) \, |i\rangle \otimes |j\rangle. \tag{A.23}$$

This allows to use Theorem A.8. □

We now introduce the relative entropy. It is usually applied to density matrices but it helps to define it more generally on all positive-definite matrices. One can check that it is always nonnegative; we allow the value $+\infty$.

DEFINITION A.10. *The* **relative entropy** $S(\cdot\|\cdot)$ *is the following function of two positive-definite matrices $a, b \in \mathcal{M}_n$:*

$$S(a\|b) = \operatorname{Tr} a\big(\log a - \log b\big).$$

LEMMA A.11. *We have for all $a, b \geq 0$ that*

$$S(a\|b) = \lim_{\varepsilon \to 0+} \tfrac{1}{\varepsilon}\big(\operatorname{Tr} a - \operatorname{Tr} a^{1-\varepsilon} b^\varepsilon\big).$$

PROOF. Let $f(\varepsilon) = \operatorname{Tr} a^{1-\varepsilon} b^\varepsilon$. The right side is the derivative $f'(0)$ which is equal to $S(a\|b)$. □

THEOREM A.12 (Joint convexity of the relative entropy). *If $a_1, a_2, b_1, b_2 \geq 0$ are complex matrices in $\mathcal{M}_n$, then*

$$S\big(a_1 + a_2 \| b_1 + b_2\big) \leq S(a_1\|b_1) + S(a_2\|b_2).$$

Since $S(\lambda a\|\lambda b) = \lambda S(a\|b)$, the joint convexity of the relative entropy follows immediately. And since the entropy is equal to $S(a) = S(a\|\mathbb{1})$, it is convex too.

PROOF. Starting with Lemma A.11, and using Corollary A.9, we have

$$\begin{aligned} S(a_1 + a_2\|b_1 + b_2) &= \lim_{\varepsilon\to 0+} \tfrac{1}{\varepsilon}\Big(\operatorname{Tr}(a_1 + a_2) - \operatorname{Tr}(a_1 + a_2)^{1-\varepsilon}(b_1 + b_2)^{\varepsilon}\Big) \\ &\le \lim_{\varepsilon\to 0+} \tfrac{1}{\varepsilon}\Big(\operatorname{Tr} a_1 - \operatorname{Tr} a_1^{1-\varepsilon} b_1^{\varepsilon} + \operatorname{Tr} a_2 - \operatorname{Tr} a_2^{1-\varepsilon} b_2^{\varepsilon}\Big) \\ &= S(a_1\|b_1) + S(a_2\|b_2). \end{aligned} \tag{A.24}$$

□

LEMMA A.13. *Let* $U = \operatorname{diag}\big(1, e^{\frac{2\pi i}{n}}, \dots, e^{\frac{2\pi i}{n}(n-1)}\big)$. *Then for any matrix* $a \in \mathcal{M}_n$ *we have*

$$\operatorname{diag} a = \frac{1}{n}\sum_{k=0}^{n-1} U^k a\, U^{-k}.$$

PROOF. The element (ℓ, m) of the left side is equal to

$$\frac{1}{n}\sum_{k=0}^{n-1} e^{\frac{2\pi i}{n}k\ell}\, a_{\ell,m}\, e^{-\frac{2\pi i}{n}km} = \frac{1}{n} a_{\ell,m}\sum_{k=0}^{n-1} e^{\frac{2\pi i}{n}k(\ell-m)} = \begin{cases} a_{\ell,\ell} & \text{if } \ell = m, \\ 0 & \text{if } \ell \neq m. \end{cases} \tag{A.25}$$

□

Let $P = \operatorname{diag}(1, \dots, 1, 0, \dots, 0)$ with k elements equal to 1, where $k \in \{1, \dots, n-1\}$. We consider the map $\Phi : \mathcal{M}_n \to \mathcal{M}_n$ defined by

$$\Phi(a) = \operatorname{diag}\big(\underbrace{\tfrac{1}{k}\operatorname{Tr} Pa \dots \tfrac{1}{k}\operatorname{Tr} Pa}_{k \text{ elements}}\ \underbrace{\tfrac{1}{n-k}\operatorname{Tr}(1-P)a \dots \tfrac{1}{n-1}\operatorname{Tr}(1-P)a}_{n-k \text{ elements}}\big). \tag{A.26}$$

The image $\Phi(a)$ is a simple matrix with just two values.

LEMMA A.14. *There exists a finite number* L *and unitary matrices* $U_1, \dots, U_L$ *such that*

$$\Phi(a) = \frac{1}{L}\sum_{\ell=1}^{L} U_\ell a U_\ell^{-1}.$$

PROOF. If a is diagonal we can consider the permutation $(1, \dots, k)(k+1, \dots, n)$ and its permutation matrix V. Then $\Phi(a) = \frac{1}{k(n-k)}\sum_{\ell=1}^{k(n-k)} V^\ell a\, V^{-\ell}$. Indeed, this amounts to average over diagonal matrices where the first k elements of a have been rotated, as well as the last $n - k$ elements.

For the general case we combine this with Lemma A.13 to get

$$\Phi(a) = \frac{1}{nk(n-k)} \sum_{\ell=1}^{k(n-k)} \sum_{m=0}^{n-1} V^{\ell} U^{m} a\, U^{-m} V^{-\ell}. \tag{A.27}$$

□

We can now prove the quantum Pinsker inequality.

Proof of Lemma 10.3. Let P be the projector onto the subspace of the eigenvectors of $\rho - \sigma$ with nonnegative eigenvalues and let Φ be the map defined in Eq. (A.26). By Lemma A.14 and the convexity of the relative entropy (Theorem A.12), we obtain that

$$S(\rho\|\sigma) \geq S\big(\Phi(\rho)\|\Phi(\sigma)\big). \tag{A.28}$$

The latter is equal to the classical relative entropy of two Bernoulli random variables with parameters $\operatorname{Tr} P\rho$ and $\operatorname{Tr} P\sigma$. Using the classical Pinsker inequality (see Exercise A.1), we get that it is greater than

$$\tfrac{1}{2}\big(|\operatorname{Tr} P\rho - \operatorname{Tr} P\sigma| + |\operatorname{Tr}(1-P)\rho - \operatorname{Tr}(1-P)\sigma|\big)^2 = \tfrac{1}{2}\|\rho - \sigma\|_1^2. \tag{A.29}$$

The last identity uses the fact that P is the projector onto the suitable eigensubspace of $\rho - \sigma$. □

We consider one more proposition, that gives the partial trace as a convex combination of unitary operations. It can be combined with Theorem A.12 to give inequalities involving the (relative) entropy in different domains.

Proposition A.15. *Let* Tr_1 *denote the partial trace with respect to the first space of* $\mathcal{H}_1 \otimes \mathcal{H}_2$*; on operators of the form* $a \otimes b$*, we have that* $\operatorname{Tr}_1 a \otimes b = (\operatorname{Tr} a) b$ *is an operator on* $\mathcal{H}_2$*. Then there exist unitary matrices* $U_1, \ldots, U_L$ *on* $\mathcal{H}_1 \otimes \mathcal{H}_2$ *such that*

$$\operatorname{Tr}_1 a = \frac{1}{L} \sum_{\ell=1}^{L} U_\ell a\, U_\ell^{-1}.$$

This can be proved as in Lemma A.14 with unitary matrices on $\mathcal{H}_1$ tensored with the identity.

Exercise A.1. *Prove the classical Pinsker inequality in the case of Bernoulli random variables.*

Bibliographical references

For the proof of the quantum Pinsker's inequality we mainly rely on Watrous [2018] and on Carlen [2010].