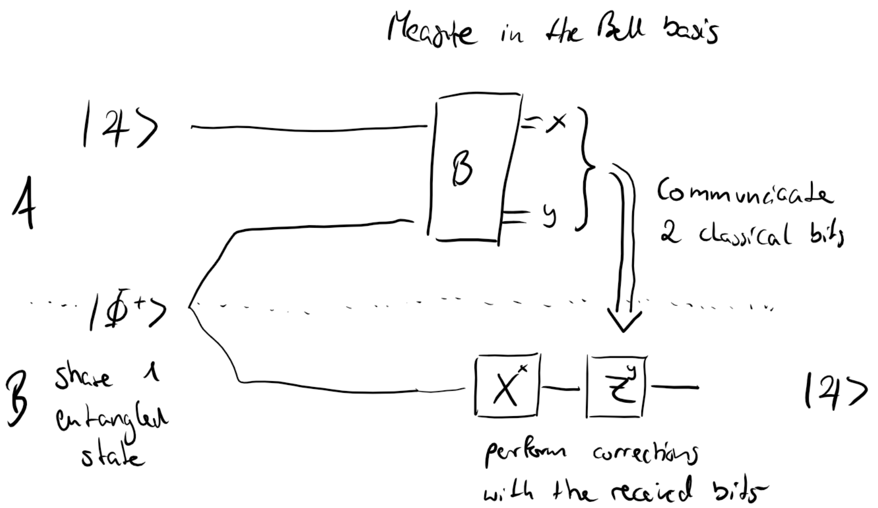
Exercise Sheet 3: Teleportation and Co.

Teleportation

In the lecture, you have seen the state teleportation protocol that allowed Alice to send an arbitrary quantum state to Bob, requiring that Alice and Bob share a maximally entangled state and Alice sends Bob two classical bits.

1 P. **Exercise 1.** Make a schematic drawing of the quantum state teleportation protocol.

Solution



We wish to use this sheet to construct some further variants of the teleportation protocol. For this, we first have to improve our understanding of the Bell states.

2 P. **Exercise 2.** Recall the Bell states

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \qquad |\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \tag{1}$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \qquad |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{2}$$

Show that you can interconvert the Bell states through local operations only, i.e. through operations that only act on one of the two systems at a time.

Solution

It is sufficient to show that $|\Phi^+\rangle$ can be converted into any other state by local operations, as we can then write any other state as the $|\Phi^+\rangle$ plus some local operations and then just add and remove local operations to obtain any other state. We have

$$\begin{aligned}
|\Phi^-\rangle &= (\mathbb{I} \otimes Z)|\Phi^+\rangle = (Z \otimes \mathbb{I})|\Phi^+\rangle \\
|\Psi^+\rangle &= (\mathbb{I} \otimes X)|\Phi^+\rangle = (X \otimes \mathbb{I})|\Phi^+\rangle \\
|\Psi^-\rangle &= (\mathbb{I} \otimes Y)|\Phi^+\rangle = (Y \otimes \mathbb{I})|\Phi^+\rangle,
\end{aligned}$$

where Y is proportional to ZX.

Let us now come to our first variation, where we restrict the classes of states that Alice wants to send.

4 P. **Exercise 3.** We now look at a restricted scenario and assume that both Alice and Bob know that the state that Alice wants to transmit is on the equator of the Bloch sphere, i.e. one of the states $|\alpha\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{-i\alpha}|1\rangle)$ for $0 \leq \alpha \leq 2\pi$. Moreover, Alice knows the value α, i.e. she prepares the state herself. In this case, we can actually reduce the amount of classical communication between Alice and Bob.

2 P. (a) Let us define $|\alpha^\perp\rangle = |\alpha + \pi\rangle$. Show that $|\alpha^\perp\rangle$ is orthogonal to $|\alpha\rangle$ and express the computational basis states $|0\rangle$ and $|1\rangle$ in terms of $|\alpha\rangle$ and $|\alpha^\perp\rangle$.

Solution

To establish orthogonality, we calculate the inner product

$$\begin{aligned}
\langle\alpha|\alpha^\perp\rangle &= \langle\alpha|\alpha+\pi\rangle \\
&= \frac{1}{\sqrt{2}}\left(\langle 0| + e^{i\alpha}\langle 1|\right)\frac{1}{\sqrt{2}}\left(|0\rangle + e^{-i\alpha-i\pi}|1\rangle\right) \\
&= \frac{1}{2}\left(\langle 0| + e^{i\alpha}\langle 1|\right)\left(|0\rangle - e^{-i\alpha}|1\rangle\right) \\
&= \frac{1}{2}\left(\langle 0|0\rangle + e^{i\alpha}\langle 1|0\rangle - e^{-i\alpha}\langle 0|1\rangle - \langle 1|1\rangle\right) \\
&= 0.
\end{aligned}$$

To re-express the computational basis states, we compute their overlap as well to obtain

$$\begin{aligned}
\langle\alpha|0\rangle &= \frac{1}{\sqrt{2}} & \langle\alpha^\perp|0\rangle &= \frac{1}{\sqrt{2}} \\
\langle\alpha|1\rangle &= \frac{e^{i\alpha}}{\sqrt{2}} & \langle\alpha^\perp|1\rangle &= -\frac{e^{i\alpha}}{\sqrt{2}}.
\end{aligned}$$

Combining this yields

$$\begin{aligned}
|0\rangle &= \frac{1}{\sqrt{2}}(|\alpha\rangle + |\alpha^\perp\rangle) \\
|1\rangle &= \frac{e^{i\alpha}}{\sqrt{2}}(|\alpha\rangle - |\alpha^\perp\rangle) \simeq \frac{1}{\sqrt{2}}(|\alpha\rangle - |\alpha^\perp\rangle),
\end{aligned}$$

where in the last step we removed the irrelevant global phase.

2 P. (b) Alice and Bob now devise the following protocol: As in the regular teleportation protocol, they share a maximally entangled state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Alice measures her half of this previously shared state in the basis $\{|\alpha\rangle, |\alpha^\perp\rangle\}$. Show that communicating the outcome of this measurement (one bit) is sufficient for Bob to faithfully obtain the state $|\alpha\rangle$.

Solution

Just as in the teleportation protocol, we can look at the possible post-measurement states that can be obtained depending on the outcome from Alice's measurement. Let us look at the case of $|\alpha\rangle$ first, thus projecting $|\Phi^+\rangle$ on $|\alpha\rangle \otimes \mathbb{I}$.

$$(\langle\alpha| \otimes \mathbb{I})|\Phi^+\rangle = \frac{1}{2}(((\langle 0| + e^{i\alpha}\langle 1|) \otimes \mathbb{I})(|00\rangle + |11\rangle) \tag{3}$$

$$= \frac{1}{2}(|0\rangle + e^{i\alpha}|1\rangle) = \frac{1}{\sqrt{2}}|-\alpha\rangle. \tag{4}$$

Renormalizing by the probability of obtaining this outcome, $p(\alpha) = 0.5$, the state that Bob has after Alice's measurement is $|-\alpha\rangle$. If Bob then performs a local X operation, he then gets

$$X|-\alpha\rangle = \frac{1}{\sqrt{2}}(|1\rangle + e^{i\alpha}|0\rangle) = \frac{e^{i\alpha}}{\sqrt{2}}(|0\rangle + e^{-i\alpha}|1\rangle) = e^{i\alpha}|\alpha\rangle. \tag{5}$$

Due to the fact that global phases are not physically relevant, this is equivalent to having obtained $|\alpha\rangle$, which is the target state.
Now let us look at the case of $|\alpha^\perp\rangle$.

$$(\langle\alpha^\perp| \otimes \mathbb{I})|\Phi^+\rangle = \frac{1}{2}(((\langle 0| - e^{i\alpha}\langle 1|) \otimes \mathbb{I})(|00\rangle + |11\rangle) \tag{6}$$

$$= \frac{1}{2}(|0\rangle - e^{i\alpha}|1\rangle) \tag{7}$$

If Bob performs a Z operation, he obtains the state (including renormalization)

$$Z\frac{1}{\sqrt{2}}(|0\rangle - e^{i\alpha}|1\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\alpha}|1\rangle) = |-\alpha\rangle. \tag{8}$$

This is the same state as after the $|\alpha\rangle$ measurement, so we know that an additional X operations leads to the target state $|\alpha\rangle$. In the end, if Alice sends the bit corresponding to measuring $|\alpha\rangle$, Bob performs an X, while if he receives the bit corresponding to $|\alpha^\perp\rangle$ he does XZ. In both cases, at the end, Bob faithfully obtained the state $|\alpha\rangle$.

1 P. **Exercise 4.** There is also another way Alice and Bob can reduce the amount of classical communication required in the original state teleportation protocol to a single bit when they accept that the protocol does not always succeed. How can they do it while still knowing if the protocol succeeded?

Solution

Alice and Bob can just agree on one specific configuration for the correction bits x and y, for example $x = y = 0$. If Alice measures these pre-agreed bits, she sends Bob the bit 1 for success, in which case Bob holds the correct teleported state. Otherwise, she sends the bit 0 indicating failure, in which case Bob needs to throw away his state and they restart the protocol. This particular protocol succeeds with probability 1/4.

1 P. **Exercise 5.** Use the state teleportation protocol to construct a new protocol that allows the following. At the beginning of the experiment, Bob holds a state $|\psi\rangle$ and at the end holds $U|\psi\rangle$ where U was applied by Alice. The protocol should require two maximally entangled states and four bits of classical communication.

Solution

Bob teleports the state to Alice, Alice applies the unitary and then teleports it back. This involves two teleportation steps and therefore requires two maximally entangled states and four bits of classical communication.

3 P. **Bonus Exercise 1.** Show that in the protocol of the previous exercise, if both Alice and Bob know that the operation U is one of $R_Z(\alpha) = e^{-i\alpha Z/2}$ (a Pauli-Z rotation) for $0 \leq \alpha \leq 2\pi$, then only three classical bits are required. (*Hint:* Establish the fact that $[Z, R_Z(\alpha)] = 0$ for all $0 \leq \alpha \leq 2\pi$. Use this to reason about the Z corrections involved in the previous exercise)

Solution

Let us look at the whole teleportation scheme. First Bob teleports his state $|\psi\rangle$ to Alice who then holds the state $Z^y X^x|\psi\rangle$ where x and y are the both bits Bob would normally send to Alice to complete the teleportation. Now, if Alice would right away apply her rotation we would obtain $R_Z(\alpha)Z^y X^x|\psi\rangle$. If she would teleport this state directly, Bob would hold the state $Z^{y'} X^{x'} R_Z(\alpha)Z^y X^x|\psi\rangle$, where x' and y' are now Alice' correction bits. As R_Z and Z commute, this is equal to $Z^{y'} X^{x'} Z^y R_Z(\alpha)X^x|\psi\rangle$. We see that it is actually sufficient if Bob does the Z correction at the end, but we still have to do the X^x correction on Alice's system.
We can deduce the following protocol: Bob teleports his state $|\psi\rangle$ to Alice and sends her one bit of information, namely the bit x which she uses to correct the state before she applies her rotation to obtain $Z^y R_Z(\alpha)|\psi\rangle$. Then, she teleports her state back to Bob, sending him both bits of information. Bob then holds the state $Z^{y'} X^{x'} Z^y R_Z(\alpha)|\psi\rangle$, which he can transform into $R_Z(\alpha)|\psi\rangle$ by applying the operation $Z^y X^{x'} Z^{y'}$, ending the protocol.

Let us now come to the last variant of teleportation we want to consider.

3 P. **Exercise 6.** Let us now consider the following situation: Both Alice and Bob each share a maximally entangled state with a third party, Charlie, meaning the system state is $|\Phi^+\rangle_{AC_1}|\Phi^+\rangle_{BC_2}$. Then, Charlie takes his two qubits and measures them in the Bell basis (compare Exercise 2). Show that it is sufficient for Charlie to communicate two bits so that at the end Alice and Bob share the state $|\Phi^+\rangle_{AB}$ (*Hint:* Use the result of Exercise 2). What can this protocol be used for?

Solution

We can write the state as

$$|\Phi^+\rangle_{AC_1}|\Phi^+\rangle_{BC_2} = \frac{1}{2}(|\Phi^+\rangle_{AB}|\Phi^+\rangle_{C_1C_2} + |\Phi^-\rangle_{AB}|\Phi^-\rangle_{C_1C_2} + |\Psi^+\rangle_{AB}|\Psi^+\rangle_{C_1C_2} + |\Psi^-\rangle_{AB}|\Psi^-\rangle_{C_1C_2}).$$

Hence, when Charlie performs a Bell measurement on his two qubits and obtains one of the Bell states as outcomes, Alice and Bob share the same state. It is then sufficient to communicate the outcome of the measurement (two bits) to either of the parties, as we know from Exercise 2 that it is sufficient that one of the parties applies a local operation to transform the state they share into $|\Phi^+\rangle$.
This protocol can be used to create entanglement between parties whose qubits have never interacted, and it allows us to create long distance entanglement from a collection of intermediate nodes.

Schatten norms

The next part of the sheet is dedicated to Schatten p-norms, a very important mathematical concept that we use a lot in this course. One way to define the Schatten p-norm with $p \in [1, \infty)$ for a matrix $A \in \mathbb{C}^{n\times n}$ is

$$\|A\|_p := \left(\mathrm{Tr}\left[|A|^p\right]\right)^{\frac{1}{p}}, \tag{9}$$

where $|A| := \sqrt{A^\dagger A}$ is the matrix absolute value. Note that here, $A^\dagger A \geq 0$ by construction, which means that the unique positive square root exists and $\sqrt{A^\dagger A}$ is well defined and positive semi-definite. Furthermore, the case $p = \infty$ is defined as the limit

$$\|A\|_\infty := \lim_{p\to\infty} \|A\|_p. \tag{10}$$

For the next exercise, we recall the definition of the p-norm for vectors $\boldsymbol{v} = (v_1, v_2, \ldots, v_d)$

$$\|\boldsymbol{v}\|_p = \left(\sum_{i=1}^{d} |v_i|^p\right)^{\frac{1}{p}}. \tag{11}$$

3 P. **Exercise 7.** In this exercise, we will relate the Schatten p norm of a matrix to the vector p norm of its vector of singular values. To do so, we first need some prerequisites.

1 P. (a) Let $B = UDU^\dagger$ a diagonalizable matrix, and let $f\colon \mathbb{C} \to \mathbb{C}$ be a complex-valued function with series expansion $f(x) = \sum_{k=0}^{\infty} f_k x^k$, with $f_k \in \mathbb{C}$ for each $k \in \mathbb{N}$. Then, we can define an analogous function f acting on matrices as

$$f(X) = \sum_{k=0}^{\infty} f_k X^k, \tag{12}$$

defined as an analogous series with the same coefficients. Show that

$$f(B) = f(UDU^\dagger) = Uf(D)U^\dagger, \tag{13}$$

i.e. applying a function to a diagonalizable matrix is like applying the function to the eigenvalues of said matrix.

Solution

The important observation is that the observation

$$B^2 = BB = UDU^\dagger UDU^\dagger = UD^2U^\dagger$$

extends to all powers of B,

$$B^k = UD^kU^\dagger.$$

Putting this into the expansion of f directly yields the desired formula.

2 P. (b) Let A be a Hermitian matrix and let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ be the vector of its singular values. Show that

$$\|A\|_p = \|\boldsymbol{\lambda}\|_p \tag{14}$$

for all p. (*Hint*: you can apply the result of the previous exercise to $f(X) = \sqrt{X}$ and $f(X) = X^p$.)

Solution

Let $A = U\Lambda V^\dagger$ be the singular value decomposition of A. Then,

$$\sqrt{A^\dagger A} = \sqrt{U\Lambda V^\dagger V \Lambda U^\dagger} = \sqrt{U\Lambda^2 U^\dagger} = U\sqrt{\Lambda^2}U^\dagger = U\Lambda U^\dagger.$$

Next, note that

$$(\sqrt{A^\dagger A})^p = (U\Lambda U^\dagger)^p = U\Lambda^p U^\dagger.$$

Therefore,

$$\|A\|_p^p = \mathrm{Tr}[(\sqrt{A^\dagger A})^p] = \mathrm{Tr}[U\Lambda^p U^\dagger] = \mathrm{Tr}[\Lambda^p] = \sum_{i=1}^d \lambda_i^p = \|\boldsymbol{\lambda}\|_p^p.$$

2 P. **Exercise 8.** Show that for a given vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_d)$

$$\|\boldsymbol{v}\|_\infty := \lim_{p\to\infty} \|\boldsymbol{v}\|_p = \max_{1\le i\le d} |v_i|.$$

What can we conclude from this about $\|A\|_\infty$?

Solution

Let us denote $v_{\max} = \max_{1\le i\le d} |v_i|$. We can pull it out of the norm to get

$$\lim_{p\to\infty} \|\boldsymbol{v}\|_p = \lim_{p\to\infty} \left(\sum_{i=1}^d |v_i|^p\right)^{\frac{1}{p}}$$
$$= v_{\max} \lim_{p\to\infty} \left(\sum_{i=1}^d \left(\frac{|v_i|}{v_{\max}}\right)^p\right)^{\frac{1}{p}}$$

The ratios $\frac{|v_i|}{v_{\max}}$ are at most 1 by construction, so we get an upper bound

$$\lim_{p\to\infty} \|\boldsymbol{v}\|_p \le v_{\max} \lim_{p\to\infty} (d)^{\frac{1}{p}} = v_{\max}.$$

Also, there is at least one term that has the value 1, again by construction, which leads to the lower bound

$$\lim_{p\to\infty} \|\boldsymbol{v}\|_p \ge v_{\max} \lim_{p\to\infty} (1)^{\frac{1}{p}} = v_{\max}.$$

Combining both results yields the desired statement. We can conclude that $\|A\|_\infty$ is the largest singular value of A.

Next, we wish to establish some properties of the Schatten p norms.

6 P. **Exercise 9.**

1 P. (a) Show that for all unitaries U and V,

$$\|UAV^\dagger\|_p = \|A\|_p. \tag{15}$$

Solution

We can absorb U and V into the unitaries in the singular value decomposition of A, hence $UAV^\dagger$ has the same singular values as A for all unitary U and V, and hence has the same Schatten p norm as per Exercise 7.

1 P. (b) Show that the Schatten p norms tensorize, i.e. that

$$\|A \otimes B\|_p = \|A\|_p \|B\|_p. \tag{16}$$

Solution

From the singular value decompositions of A and B it is obvious that the singular values of $A \otimes B$ are all the products of the singular values of A and B. If we denote the singular values of A with λ_i and of B with μ_j, it is easy to verify that

$$\begin{aligned}\|A \otimes B\|_p &= \left(\sum_{i=1}^{d} \sum_{j=1}^{d} (\lambda_i \mu_j)^p \right)^{\frac{1}{p}} \\ &= \left(\sum_{i=1}^{d} \sum_{j=1}^{d} \lambda_i^p \mu_j^p \right)^{\frac{1}{p}} \\ &= \left(\sum_{i=1}^{d} \lambda_i^p \right)^{\frac{1}{p}} \left(\sum_{j=1}^{d} \mu_j^p \right)^{\frac{1}{p}} \\ &= \|A\|_p \|B\|_p.\end{aligned}$$

2 P. (c) Prove the following variational formulation for the Schatten p-norm

$$\|A\|_p = \sup_{\|X\|_q = 1} \mathrm{Tr}[AX] \tag{17}$$

for $1/p + 1/q = 1$. You can make use of the variational formulation of the vector p norm

$$\|\boldsymbol{v}\|_p = \sup_{\|\boldsymbol{z}\|_q = 1} \langle \boldsymbol{v}, \boldsymbol{z} \rangle. \tag{18}$$

Solution

We again make use of the singular value decomposition of $A = U \Lambda V^\dagger$. The optimization then reads

$$\sup_{\|X\|_q = 1} \mathrm{Tr}[AX] = \sup_{\|X\|_q = 1} \mathrm{Tr}[\Lambda V^\dagger X U].$$

Now, we note that $\|V^\dagger X U\|_q = \|X\|_q$ because of the unitary invariance of the Schatten q norm. We can hence ignore the unitaries in our optimization and instead optimize over $Z = V^\dagger X U$ to obtain

$$\sup_{\|X\|_q = 1} \mathrm{Tr}[AX] = \sup_{\|Z\|_q = 1} \mathrm{Tr}[\Lambda Z].$$

As Λ is now a diagonal matrix, it is sufficient to optimize over diagonal Z. In this case, $\mathrm{Tr}[\Lambda Z]$ reduces to the inner product $\langle \boldsymbol{\lambda}, \boldsymbol{z} \rangle$ between the vectors $\boldsymbol{\lambda}$ and $\boldsymbol{z}$ containing the diagonal entries of Λ and Z. For diagonal matrices we immediately see that $\|Z\|_q = \|\boldsymbol{z}\|_q$, which yields an expression where we can just read of the variational definition of the vector p norm

$$\sup_{\|X\|_q = 1} \mathrm{Tr}[AX] = \sup_{\|\boldsymbol{z}\|_q = 1} \langle \boldsymbol{\lambda}, \boldsymbol{z} \rangle = \|\boldsymbol{\lambda}\|_p = \|A\|_p.$$

2 P. (d) Show that the Schatten norms are ordered, i.e. if $p \leq q$ then

$$\|A\|_p \geq \|A\|_q. \tag{19}$$

Solution

Let λ_i be the singular values of A and $\lambda_{\max}$ to be the maximal singular value. Then, we have that

$$\|A\|_p = \left(\sum_{i=1}^{d} \lambda_i^p\right)^{\frac{1}{p}}$$
$$= \lambda_{\max}\left(\sum_{i=1}^{d}\left(\frac{\lambda_i}{\lambda_{\max}}\right)^p\right)^{\frac{1}{p}}.$$

Here we observe that the ratios $\lambda_i/\lambda_{\max} \leq 1$ by definition. In this parameter range $p \leq q$ implies that

$$\left(\frac{\lambda_i}{\lambda_{\max}}\right)^p \geq \left(\frac{\lambda_i}{\lambda_{\max}}\right)^q.$$

Similarly, by definition we have that

$$\sum_{i=1}^{d}\left(\frac{\lambda_i}{\lambda_{\max}}\right)^p \geq 1$$

because one of the λ_i is identical to $\lambda_{\max}$. In this parameter range we have that

$$\left(\sum_{i=1}^{d}\left(\frac{\lambda_i}{\lambda_{\max}}\right)^p\right)^{\frac{1}{p}} \geq \left(\sum_{i=1}^{d}\left(\frac{\lambda_i}{\lambda_{\max}}\right)^q\right)^{\frac{1}{q}}.$$

Combining both facts yields the desired statement.

Last but not least, we want to observe two important applications for quantum information.

2 P. Exercise 10.

1 P. (a) Show that $\|\rho\|_1 = 1$ for any quantum state ρ.

Solution

As ρ is by definition positive semi-definite and Hermitian, the singular values coincide with the eigenvalues and $\mathrm{Tr}[\rho] = \|\rho\|_1 = 1$.

1 P. (b) Show that $\|\rho\|_2^2$ is the purity of the state.

Solution

Just by definition,

$$\|\rho\|_2^2 = \mathrm{Tr}[\rho^\dagger\rho] = \mathrm{Tr}[\rho^2].$$

There are some rather important inequalities that we will need later in the course which we establish as a bonus exercise. The next exercise gives a matrix generalization of Hölder's inequality for vectors which posits

$$\langle \boldsymbol{v}, \boldsymbol{w}\rangle \leq \langle |\boldsymbol{v}|, |\boldsymbol{w}|\rangle \leq \|\boldsymbol{v}\|_p \|\boldsymbol{w}\|_q \tag{20}$$

for $1/p + 1/q = 1$ and $|\boldsymbol{v}|$ being the vector $\boldsymbol{v}$ where all entries are replaced by their absolute values.

2 P. **Bonus Exercise 2.** Hölder's inequality for Schatten norms takes the form

$$\mathrm{Tr}[A^\dagger B] \leq \|A^\dagger B\|_1 \leq \|A\|_p \|B\|_q \tag{21}$$

for any $1 \leq p, q \leq \infty$ such that $1/p + 1/q = 1$. Prove this inequality. If we use the notation $\boldsymbol{\sigma}^\downarrow(A)$ to denote the vector of singular values of A in descending order, then you can use the following inequality:

$$\sum_{i=1}^{d} \sigma_i^\downarrow(AB) \leq \sum_{i=1}^{d} \sigma_i^\downarrow(A)\sigma_i^\downarrow(B). \tag{22}$$

Solution

Using the given inequality allows us to reduce to the vector case of Hölder's inequality:

$$\|A^\dagger B\|_1 = \|\boldsymbol{\sigma}^\downarrow(A^\dagger B)\|_1 \leq \|\boldsymbol{\sigma}^\downarrow(A)\boldsymbol{\sigma}^\downarrow(B)\|_1 \leq \|\boldsymbol{\sigma}^\downarrow(A)\|_p \|\boldsymbol{\sigma}^\downarrow(B)\|_q = \|A\|_p \|B\|_q.$$

2 P. **Bonus Exercise 3.** A different class of matrix norms is given by the *induced norms* which are defined via the vector p norms. We define

$$\|A\|_{p\to q} := \sup_{\|\boldsymbol{v}\|_p = 1} \|A\boldsymbol{v}\|_q. \tag{23}$$

These norms are useful, because they allow us to establish inequalities of the form

$$\|A\boldsymbol{v}\|_q \leq \|A\|_{p\to q} \|\boldsymbol{v}\|_p. \tag{24}$$

Show that

$$\|A\|_{2\to 2} = \|A\|_\infty. \tag{25}$$

Solution

$$\begin{aligned}
\|A\|_{2\to 2} &= \sup_{\|\boldsymbol{v}\|_p = 2} \|A\boldsymbol{v}\|_2 \\
&= \sup_{\|\boldsymbol{v}\|_p = 2} \sqrt{\langle \boldsymbol{v}, A^\dagger A \boldsymbol{v}\rangle} \\
&= \sup_{\|\boldsymbol{v}\|_p = 2} \sqrt{\lambda_{\max}(A^\dagger A)} \\
&= \sup_{\|\boldsymbol{v}\|_p = 2} \sigma_{\max}(A) \\
&= \|A\|_\infty.
\end{aligned}$$

Here we denoted with $\lambda_{\max}$ and $\sigma_{\max}$ the largest eigenvalue and singular value of a matrix.

2 P. **Bonus Exercise 4.** Another important property of norms is submultiplicativity. First, establish that

$$\|AB\|_p \leq \|A\|_\infty \|B\|_p \tag{26}$$

and then conclude that

$$\|AB\|_p \leq \|A\|_p \|B\|_p. \tag{27}$$

You can make use of the min-max-principle for singular values which posits that

$$\sigma_i^{\uparrow}(A) = \min_{S \subseteq \mathbb{C}^d, \dim(S)=i} \max_{\boldsymbol{v} \in S, \|\boldsymbol{v}\|_2=1} \|A\boldsymbol{v}\|_2, \tag{28}$$

where $\boldsymbol{\sigma}^{\uparrow}(A)$ is the vector of singular values in increasing order and the optimization is over subspaces of bounded dimension.

Solution

We have that

$$\begin{aligned}
\sigma_i^{\uparrow}(AB) &= \min_{S \subseteq \mathbb{C}^d, \dim(S)=i} \max_{\boldsymbol{v} \in S, \|\boldsymbol{v}\|_2=1} \|AB\boldsymbol{v}\|_2 \\
&\leq \min_{S \subseteq \mathbb{C}^d, \dim(S)=i} \max_{\boldsymbol{v} \in S, \|\boldsymbol{v}\|_2=1} \|A\|_\infty \|B\boldsymbol{v}\|_2 \\
&= \|A\|_\infty \sigma_i^{\uparrow}(B),
\end{aligned}$$

where we used that $\|A\|_\infty = \|A\|_{2\to 2}$. Now we just use the expression of the Schatten norm via the norm of the vector of singular values to conclude

$$\begin{aligned}
\|AB\|_p &= \left(\sum_{i=1}^d \sigma_i^{\uparrow}(AB)^p \right)^{\frac{1}{p}} \\
&\leq \|A\|_\infty \left(\|A\|_\infty^p \sum_{i=1}^d \sigma_i^{\uparrow}(B)^p \right)^{\frac{1}{p}} \\
&= \|A\|_\infty \|B\|_p.
\end{aligned}$$

We can conclude the submultiplicativity because $\|A\|_\infty \leq \|A\|_p$ for all $1 \leq p \leq \infty$.

Total Points: 25 (+9)