Chapter 6

Quantum frequentist estimation

In order to better grasp the intuitions behind the theory that will be developed in this chapter let us be guided by the following simple example.

Example 6.1 Consider a single qubit system, and the family of states

$$|\psi_\theta\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\theta}|1\rangle\right), \tag{6.1}$$

parameterized by an angle $\theta \in [0, 2\pi]$. In the Bloch sphere picture these states correspond to states on the equator. Assume we are given N copies of the state so that

$$\rho_\theta = (|\psi_\theta\rangle\langle\psi_\theta|)^{\otimes N}. \tag{6.2}$$

We want to know how to choose the optimal measurement and estimator in order to estimate θ with the lowest uncertainty possible. Consider two exemplary measurements, corresponding to the following choices of basis, a) $\{|0\rangle, |1\rangle\}$, b) $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. In case of measurement a) we see that $p_\theta(0) = 1/2$, $p_\theta(1) = 1/2$, so that the measurements results do not carry any information on the parameter θ. On the other hand, measurement b) leads to probability distributions $p_\theta(\pm) = |\langle +|\psi_\theta\rangle|^2 = \frac{1}{2}(1 \pm \cos\theta)$. which seems much more reasonable. We can caluclate the correspondin FI for this measurement which yields:

$$F = \frac{1}{p_\theta(+)}\left(\frac{\mathrm{d}p_\theta(+)}{\mathrm{d}\theta}\right)^2 + \frac{1}{p_\theta(-)}\left(\frac{\mathrm{d}p_\theta(-)}{\mathrm{d}\theta}\right)^2 = 1. \tag{6.3}$$

As a result given N copies, $F^{(N)} = N$, and the CR bound implies that $\Delta^2\tilde{\theta} \geq 1/N$.

The question is whether this measurement is optimal, or maybe some other measurement could result in higher FI? Note that we have just considered measurements on a single qubit, while in principle given N copies one could consider also collective measurements on all N copies simultaneously. Might that be helpful? In the following sections we will develop tools that will allow us to answer these questions.

In this chapter we will pursue the frequentist approach and try to find a fundamental lower bound on achievable estimation uncertainty. The following theorem is a generalization of the classical Cramér-Rao bound and we first focus on the single-parameter case.

6.1 Quantum Cramér-Rao bound

Theorem 6.1 (Quantum Cramér-Rao bound)**.** Given a family of states ρ_θ, arbitrary measurements and locally unbiased estimators the estimation variance is lower bounded by:

$$\Delta^2\tilde{\theta} \geq \frac{1}{F_Q}, \tag{6.4}$$

where F_Q is the quantum Fisher information (QFI) that is defined as

$$F_Q = \mathrm{Tr}(\rho_\theta \Lambda_\theta^2), \tag{6.5}$$

where Λ_θ is the symmetric logarithmic derivative (SLD) operator defined implicitly via the following equation

$$\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta} = \frac{1}{2}\left(\Lambda_\theta\rho_\theta + \rho_\theta\Lambda_\theta\right). \tag{6.6}$$

Proof. First note, that the SLD operator defined in the theorem can be written explicitly if one consider it in the ρ_θ eigenbasis. Let $\rho_\theta = \sum_i \lambda_i |e_i\rangle\langle e_i|$ then according to defining equation for SLD we have:

$$\langle e_i|\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta}|e_j\rangle = \frac{1}{2}\langle e_i|\left(\rho_\theta\Lambda_\theta + \Lambda_\theta\rho_\theta\right)|e_j\rangle = \\ = \frac{1}{2}\left(\lambda_i\langle e_i|\Lambda_\theta|e_j\rangle + \langle e_i|\Lambda_\theta|e_j\rangle\lambda_j\right). \quad (6.7)$$

As a result we get an explicit formula for the SLD operator:

$$\langle e_i|\Lambda_\theta|e_j\rangle = \frac{2\langle e_i|\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta}|e_j\rangle}{\lambda_i+\lambda_j}. \quad (6.8)$$

Since ρ_θ is hermitian, then from the above formula it in particular follows that Λ_θ is also a hermitian operator.

Consider a measurement $\{\Pi_x\}$ and the corresponding probability distribution $p_\theta(x) = \mathrm{Tr}(\rho_\theta\Pi_x)$. We want to derive an upper bound on the resulting Fisher information valid for arbitrary measurements:

$$F = \int \mathrm{d}x\,\frac{\left[\mathrm{Tr}(\Pi_x\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta})\right]^2}{\mathrm{Tr}(\Pi_x\rho_\theta)} = \\ = \int \mathrm{d}x\,\frac{\left[\mathrm{Tr}\left[\frac{1}{2}\Pi_x(\Lambda_\theta\rho_\theta+\rho_\theta\Lambda_\theta)\right]\right]^2}{\mathrm{Tr}(\Pi_x\rho_\theta)}. \quad (6.9)$$

Let us just focus on the term in the enumerator. Since all ρ_θ, Π_x, Λ_θ are hermitian we may equivalently write

$$\left|\mathrm{Tr}\left[\frac{1}{2}\Pi_x(\Lambda_\theta\rho_\theta+\rho_\theta\Lambda_\theta)\right]\right| = \\ = \left|\mathrm{Tr}\left(\frac{1}{2}[\Pi_x\Lambda_\theta\rho_\theta + (\Pi_x\Lambda_\theta\rho_\theta)^\dagger]\right)\right| = \\ = \left|\mathrm{Re}\mathrm{Tr}\left(\Pi_x\Lambda_\theta\rho_\theta\right)\right| \leq \left|\mathrm{Tr}\left(\Pi_x\Lambda_\theta\rho_\theta\right)\right|. \quad (6.10)$$

We now make use of the Cauchy-Schwarz inequality with respect to the Hilbert-Schmidt matrix scalar product:

$$\left|\mathrm{Tr}(AB^\dagger)\right|^2 \leq \mathrm{Tr}\left(A^\dagger A\right)\mathrm{Tr}\left(B^\dagger B\right), \quad (6.11)$$

where we set $A = \sqrt{\Pi_x}\sqrt{\rho_\theta}$, $B = \sqrt{\Pi_x}\Lambda_\theta\sqrt{\rho_\theta}$ and obtain:

$$\left|\mathrm{Tr}\left(\Pi_x\Lambda_\theta\rho_\theta\right)\right|^2 \leq \mathrm{Tr}(\rho_\theta\Pi_x)\mathrm{Tr}(\sqrt{\rho_\theta}\Lambda_\theta\Pi_x\Lambda_\theta\sqrt{\rho_\theta}). \quad (6.12)$$

Substituting this inequality to (6.10) and (6.9) we finally arrive at:

$$F \leq \int \mathrm{d}x\,\mathrm{Tr}(\rho_\theta\Lambda_\theta\Pi_x\Lambda_\theta) = \mathrm{Tr}(\rho_\theta\Lambda_\theta^2) = F_Q, \quad (6.13)$$

where we have made use of the completeness property of measurement operators Π_x. This way we have proved that whatever measurement is chosen $F \leq F_Q$. Making use of the classical CR bound we therefore obtain

$$\Delta^2\tilde{\theta} \geq \frac{1}{F} \geq \frac{1}{F_Q}. \quad (6.14)$$

□

Remark. Note that in the classical case,

$$\frac{\mathrm{d}p_\theta(x)}{\mathrm{d}\theta} = \frac{\mathrm{d}\log p_\theta(x)}{\mathrm{d}\theta}p_\theta(x) \quad (6.15)$$

so the object that multiplies the probability distribution and yields its derivative is the logarithmic derivative. The Λ_θ is therefore the operator analog of the logarithmic derivative. Due to non-commutativity this choice is not unique and hence the name SLD indicates that we define it in a symmetric way. It is possible to define e.g. right logarithmic derivative (RLD) via $\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta} = \rho_\theta(x)\Lambda_\theta^R$, and the derivation of the CR bound will also be valid. Still in general the RLD does not exist (notice that it will only exist if the kernel of $\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta}$ is the same as the kernel od $\rho_\theta(x)$). Moreover, even if it exists, it may be shown that in the single parameter case the resulting bound is never tighter than the one based on the SLD. In multi-parameter case, however, it might happen that RLD provides a tighter bound, see Sec. 6.4 for more information.

Additivity of the QFI Similarly to the FI the QFI is additive. Consider a family of states of a bipartite system that are products of states of individual systems:

$$\rho_\theta^{(12)} = \rho_\theta^{(1)} \otimes \rho_\theta^{(2)}. \tag{6.16}$$

Let $\Lambda_\theta^{(1)}$, $\Lambda_\theta^{(2)}$ be SLD operators corresponding to $\rho_\theta^{(1)}$ and $\rho_\theta^{(2)}$ respectively. Then:

$$\frac{\mathrm{d}\rho_\theta^{(12)}}{\mathrm{d}\theta} = \frac{\mathrm{d}\rho_\theta^{(1)}}{\mathrm{d}\theta} \otimes \rho_\theta^{(2)} + \rho_\theta^{(1)} \otimes \frac{\mathrm{d}\rho_\theta^{(2)}}{\mathrm{d}\theta} = \frac{1}{2}\left(\Lambda_\theta^{(1)}\rho_\theta^{(1)} + \rho_\theta^{(1)}\Lambda_\theta^{(1)}\right) \otimes \rho_\theta^{(2)} + \rho_\theta^{(1)} \otimes \frac{1}{2}\left(\Lambda_\theta^{(2)}\rho_\theta^{(2)} + \rho_\theta^{(2)}\Lambda_\theta^{(2)}\right) = \frac{1}{2}\left(\Lambda_\theta^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes \Lambda_\theta^{(2)}\right)\rho_\theta^{(1)} \otimes \rho_\theta^{(2)} + \leftrightarrow, \tag{6.17}$$

from which we see that the SLD for the joined state equals

$$\Lambda_\theta^{(12)} = \Lambda_\theta^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes \Lambda_\theta^{(2)}. \tag{6.18}$$

Consequently the QFI reads:

$$F_Q^{(12)} = \mathrm{Tr}(\rho_\theta^{(12)}\Lambda_\theta^{(12)2}) = = \mathrm{Tr}(\rho_\theta^{(1)}\Lambda_\theta^{(1)2}) + \mathrm{Tr}(\rho_\theta^{(2)}\Lambda_\theta^{(2)2}) + + 2\mathrm{Tr}(\rho_\theta^{(1)}\Lambda_\theta^{(1)}) \cdot \mathrm{Tr}(\rho_\theta^{(2)}\Lambda_\theta^{(2)}) = F_Q^{(1)} + F_Q^{(2)}, \tag{6.19}$$

where we have used the fact that $\mathrm{Tr}(\rho_\theta\Lambda_\theta) = \mathrm{Tr}\frac{\mathrm{d}\rho_\theta}{\mathrm{d}\theta} = 0$.

In particular, when we consider N copies of a quantum state $\rho_\theta^{\otimes N}$, the resulting QFI reads $F_Q^{(N)} = NF_Q$, where F_Q is the QFI corresponding to the single state ρ_θ.

Pure state case Consider a special case where the states in which the parameter is encoded are pure, $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$. In this case the SLD operator may be written explicitly as

$$\Lambda_\theta = 2\left(|\dot{\psi}_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\dot{\psi}_\theta|\right), \tag{6.20}$$

where $|\dot{\psi}_\theta\rangle = \frac{\mathrm{d}|\psi_\theta\rangle}{\mathrm{d}\theta}$. Let us check this:

$$\frac{1}{2}\left(\Lambda_\theta|\psi_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\psi_\theta|\Lambda_\theta\right) = |\dot{\psi}_\theta\rangle\langle\psi_\theta| + |\psi_\theta\rangle\langle\dot{\psi}_\theta| + (\langle\dot{\psi}_\theta|\psi_\theta\rangle + \langle\psi_\theta|\dot{\psi}_\theta\rangle)|\psi_\theta\rangle\langle\psi_\theta| = = \frac{\mathrm{d}|\psi_\theta\rangle\langle\psi_\theta|}{\mathrm{d}\theta}, \tag{6.21}$$

where we have used the identity $0 = \frac{\mathrm{d}\langle\psi_\theta|\psi_\theta\rangle}{\mathrm{d}\theta} = \langle\dot{\psi}_\theta|\psi_\theta\rangle + \langle\dot{\psi}_\theta|\psi_\theta\rangle$.

The resulting QFI reads:

$$F_Q = \mathrm{Tr}\left(|\psi_\theta\rangle\langle\psi_\theta|\Lambda_\theta^2\right) = \langle\psi_\theta|\Lambda_\theta^2|\psi_\theta\rangle = = 4\left(\langle\psi_\theta|\dot{\psi}_\theta\rangle^2 + \langle\dot{\psi}_\theta|\psi_\theta\rangle^2 + \langle\dot{\psi}_\theta|\dot{\psi}_\theta\rangle + |\langle\psi_\theta|\dot{\psi}_\theta\rangle|^2\right). \tag{6.22}$$

Since $\langle\dot{\psi}_\theta|\psi_\theta\rangle + \langle\dot{\psi}_\theta|\psi_\theta\rangle = 0$, if we square it, we obtain the following identity:

$$\langle\psi_\theta|\dot{\psi}_\theta\rangle^2 + \langle\dot{\psi}_\theta|\psi_\theta\rangle^2 + |\langle\psi_\theta|\dot{\psi}_\theta\rangle|^2 = -|\langle\dot{\psi}_\theta|\psi_\theta\rangle|^2. \tag{6.23}$$

Substituting this to (6.22) we finally arrive at:

$$F_Q = 4\left(\langle\dot{\psi}_\theta|\dot{\psi}_\theta\rangle - |\langle\dot{\psi}_\theta|\psi_\theta\rangle|^2\right). \tag{6.24}$$

The above formula has a very intuitive interpretation. There is more information on the parameter accessible in the state the bigger is the derivative $|\dot{\psi}_\theta\rangle$. Still, since states are anyway normalized, the real change has to happen in the direction perpendicular to the state itself, and that is why we need to substitute the component representing the change in the direction of the state itself.

Note, that while in the derivation we have used an explicit formula for the SLD, the SLD for pure states is not a unique operator. In general only for full rank state the SLD is uniquely defined via (6.7), otherwise we may always add terms that are outside of the support of the ρ_θ and this will still lead to the correct formula for the derivative.

Remark (Time-Energy uncertainty relation)**.** Consider the problem of estimating the time of evolution of a quantum state evolving under a known Hamitonian. Formally consider the family of states:

$$|\psi_t\rangle = e^{-iHt/\hbar}|\psi\rangle, \tag{6.25}$$

where now it is time t which is the parameter to be estimated. As this is a pure state model we can easily calculate QFI using [?] and get

$$F_Q = \frac{4}{\hbar^2}\left(\langle\psi_t|H^2|\psi_t\rangle - |\langle\psi(t)|H|\psi(t)\rangle|^2\right), \tag{6.26}$$

which is proportional to the variance of the Hamiltonian on the state. The QCR bound now takes the form

$$\Delta^2\tilde{t}\Delta^2 H \geq \frac{\hbar^2}{4}, \tag{6.27}$$

which may be viewed as a formal statement of the time-energy uncertainty relation. The fact that there is no time operator in quantum mechanics does not cause any problem here, since $\Delta^2\tilde{t}$ is the variance of an estimator and not of an operator.

Example 6.1 (continued) Let us calculate the QFI information for the qubit model. First note that $|\dot{\psi}_\theta\rangle = ie^{i\theta}|1\rangle/\sqrt{2}$, hence

$$F_Q = 4\left(\frac{1}{2} - \frac{1}{4}\right) = 1. \tag{6.28}$$

Given N copies the QFI equals $F_Q^{(N)} = N$, and hence the QCR bound implies that $\Delta^2\tilde{\theta} = 1/N$. Note that this is the same value we have obtained, when we calculated FI for the measurement in $|\pm\rangle$ basis. It implies that this measurement is indeed optimal (one can check that actually any measurement in the basis where vectors lie in the equatorial plane of the Bloch sphere will be optimal, so this choice was not unique).

This observation has some far reaching consequences. As a by product we have also proven, that collective measurements are not necessary to achieve the optimal precision—note that QFI for N copies $\rho_\theta^{\otimes N}$ is just N times QFI for a single copy, and hence we can find a measurement on a single copy that makes the corresponding FI equal to QFI of the state, it implies that if we repeat the measurement on N copies we will get N times larger FI, and as result the same value as QFI for the N copy state—note that that the derivation of the quantum CR bound allowed for arbitrary measurements, so when considered for the $\rho^{\otimes N}$ we have taken into account the possibility of collective measurements.

Saturability of the quantum CR bound In the single qubit example from previous section we have seen that there was a simple measurement for which the corresponding FI was equal to the QFI. Inspecting the derivation of the QCR bound, we see that in order to saturate the Cauchy-Schwarz inequality we need to satisfy

$$\sqrt{\Pi_x}\sqrt{\rho_\theta} = \lambda_x\sqrt{\Pi_x}\Lambda_\theta\sqrt{\rho_\theta}, \tag{6.29}$$

where λ_x is some proportionality constant. Moreover, iff λ is real then inequalitiy (6.10) will also be saturated. This can be seen as follows:

$$|\mathrm{Re}\mathrm{Tr}(\Pi_x\Lambda_\theta\rho_\theta)| = |\mathrm{Re}\mathrm{Tr}(\sqrt{\rho_\theta}\sqrt{\Pi_x}\sqrt{\Pi_x}\Lambda_\theta\sqrt{\rho_\theta})| = |\mathrm{Re}\lambda_x\mathrm{Tr}(\sqrt{\rho_\theta}\Lambda_\theta\Pi_x\Lambda_\theta\sqrt{\rho_\theta})|. \tag{6.30}$$

Note that the operator under the trace is hermitian so the trace is real. Hence if and only if $\lambda \in \mathbb{R}$ we can remove Re without changing the value of the expression.

Let $\Lambda_\theta = \sum_x l_x|x\rangle\langle x|$ be the eigendecomposition of Λ_θ so that $|x\rangle$ form orthonormal eigenbasis. Now, let us consider a measurement which corresponds to a projection measurement in the eigenbasis of the SLD operator: $\Pi_x = |x\rangle\langle x|$. Note that since this is a projective measurement $\sqrt{\Pi_x} = \Pi_x$. We have:

$$|x\rangle\langle x|\Lambda_x\sqrt{\rho_\theta} = l_x|x\rangle\langle x|\sqrt{\rho_\theta}, \tag{6.31}$$

and hence indeed we satisfy all saturability conditions provided we set $\lambda_x = 1/l_x$.

Remark. Even though we have proven that there always exist a projective measurement for which FI equals to QFI, we need to keep in mind that we still need to satisfy the classical requirement of existence of the estimator that satisfies the classical CR bound in order to claim that actually the QCR bound is saturated. In particular even if there is a single copy measurement for which FI equals the QFI it does not mean we can saturate QCR bound using single copy measurements. We may still need in general to have many repetiions of the experiment (many copies of a quantum state) to really be sure that the estimator that asymptotically saturates the CR bound (e.g. max-likelihood estimator) exists.

6.2 Multi-parameter case

Let us now consider a multi-parameter estimation problem, where the family of states $\rho_{\boldsymbol{\theta}}$ is parametrized by K real parameters $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_K\}$. Similarly as in the classical case the following multiparameter generalization of the CR bound holds.

Theorem 6.2 (Multiparameter quantum Cramér-Rao bound)**.** Given a family of states $\rho_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_K\}$, the following matrix inequality holds:

$$\mathbb{C} \geq \mathbb{F}_Q^{-1}, \tag{6.32}$$

where $\mathbb{C}$ is the $K \times K$ covariance matrix corresponding to estimation involving any locally unbiased estimators and arbitrary measurements and $\mathbb{F}_Q$ is the QFI matrix defined as:

$$(\mathbb{F}_Q)_{ij} = \frac{1}{2}\mathrm{Tr}\left[\rho_{\boldsymbol{\theta}}(\Lambda_{\boldsymbol{\theta},i}\Lambda_{\boldsymbol{\theta},j} + \Lambda_{\boldsymbol{\theta},j}\Lambda_{\boldsymbol{\theta},i})\right], \tag{6.33}$$

where $\Lambda_{\boldsymbol{\theta},i}$ is the SLD corresponding to parameter θ_i:

$$\frac{\mathrm{d}\rho_{\boldsymbol{\theta}}}{\mathrm{d}\theta_i} = \frac{1}{2}\left(\rho_{\boldsymbol{\theta}}\Lambda_{\boldsymbol{\theta},i} + \Lambda_{\boldsymbol{\theta},i}\rho_{\boldsymbol{\theta}}\right). \tag{6.34}$$

Proof. The proof utilizes the same steps that could be found in earlier derivations of the single parameter quantum CR bound and multiparameter classical CR bound. We provide the proof below, without comments as we basically repeat the steps that were employed in the earlier proofs:

$$v^T\mathbb{F}v = \int \mathrm{d}x\, \frac{\left[\sum_i v_i \mathrm{Re}\mathrm{Tr}(\Pi_x\Lambda_{\boldsymbol{\theta},i}\rho_{\boldsymbol{\theta}})\right]^2}{\mathrm{Tr}\rho_{\boldsymbol{\theta}}\Pi_x} \leq$$
$$\int \mathrm{d}x\, \frac{\left|\mathrm{Tr}(\sum_i v_i\Pi_x\Lambda_{\boldsymbol{\theta},i}\rho_{\boldsymbol{\theta}})\right|^2}{\mathrm{Tr}\rho_{\boldsymbol{\theta}}\Pi_x} \leq$$
$$\int \mathrm{d}x \frac{\mathrm{Tr}(\rho_{\boldsymbol{\theta}}\Pi_x)\mathrm{Tr}\left(\sum_{ij} v_i v_j \sqrt{\rho_{\boldsymbol{\theta}}}\Lambda_{\boldsymbol{\theta},i}\sqrt{\Pi_x}\sqrt{\Pi_x}\Lambda_{\boldsymbol{\theta},j}\sqrt{\rho_{\boldsymbol{\theta}}}\right)}{\mathrm{Tr}(\rho_{\boldsymbol{\theta}}\Pi_x)} =$$
$$= \sum_{ij} v_i \mathrm{Tr}\left(\rho_{\boldsymbol{\theta}}\Lambda_{\boldsymbol{\theta},\boldsymbol{i}}\Lambda_{\boldsymbol{\theta},\boldsymbol{j}}\right) v_j =$$
$$= v^T\mathbb{F}_Q v. \tag{6.35}$$

□

Remark. In the multiparameter case the QCR is not in general saturable. This is due to the fact that different SLDs corresponding to different parameters might not commute, so it is not clear whether there exist a single measurement that provides the optimal FI for all the parameters simultaneously. In fact, tighter bounds exist that are more informative and take into account the necessary trade-offs due to incompatibility of measurements which are optimal for different parameters, see Sec. 6.4.

6.3 Natural metric in the space of quantum states

Bures metric Since the QFI is a measure of distinguishability of quantum states we may employ it as a natural measure of distance between quantum states. Let us define infinitesimal distance between states ρ and $\rho + d\rho$ as:

$$d_B(\rho, \rho + \mathrm{d}\rho)^2 = \frac{1}{4}\mathrm{Tr}(\rho\, \mathrm{d}\Lambda^2), \tag{6.36}$$

where $\mathrm{d}\Lambda$ is defined via:

$$\mathrm{d}\rho = \frac{1}{2}(\mathrm{d}\Lambda\, \rho + \rho\, \mathrm{d}\Lambda). \tag{6.37}$$

d_B is called the Bures distance and the resulting metric in the space of quantum states is called the Bures metric. When restricted to pure states, the Bures metric is refered to as the Fubini-Study metric.

Fidelity In quantum information theory a commonly used measure of similarity of two quantum states is the so called *fidelity*. Given two pure states $|\psi_1\rangle$, $|\psi_2\rangle$ the fidelity is defined as:

$$\mathcal{F}(|\psi_1\rangle, |\psi_2\rangle) = |\langle\psi_1|\psi_2\rangle|^2 \quad (6.38)$$

and may be interpreted as the probability of observing state $|\psi_1\rangle$ as state $|\psi_2\rangle$ or vice versa.

Fidelity is generalized to mixed states using the following formula:

$$\mathcal{F}(\rho_1, \rho_2) = \max_{|\psi_1\rangle, |\psi_2\rangle} |\langle\psi_1|\psi_2\rangle|^2, \quad (6.39)$$

where $|\psi_i\rangle \in \mathcal{H}\otimes\mathcal{H}_E$ are purifications of states ρ_i: $\rho_i = \mathrm{Tr}_E|\psi_i\rangle\langle\psi_i|$, where E represents ancillary Hilbert space used for purification.

Theorem 6.3 (Uhlman theorem)**.** An explicit formula for the fidelity between two mixed states reads:

$$\mathcal{F}(\rho_1, \rho_2) = (\mathrm{Tr}|\sqrt{\rho_1}\sqrt{\rho_2}|)^2, \quad (6.40)$$

where $|A| = \sqrt{A^\dagger A}$.

Proof. Let $|\psi_i\rangle \in \mathcal{H}_S \otimes \mathcal{H}_E$ be purifications of ρ_i. We can rewrite each $|\psi_i\rangle$ as a matrix A_i, such that $(A_i)^k_l = \psi_{i,kl}$, were k, l represent indices corresponding to spaces $\mathcal{H}$ and $\mathcal{H}_E$. Then $\rho_i = \mathrm{Tr}_E|\psi_i\rangle\langle\psi_i| = A_i A_i^\dagger$. Now,

$$\mathcal{F}(\rho_1, \rho_2) = \max_{A_1, A_2} |\mathrm{Tr}(AB^\dagger)|^2, \ \rho_i = A_i A_i^\dagger. \quad (6.41)$$

Note that changing a purification for a given ρ corresponds to replacing $A \to AU$, where U is a unitary. Let us now consider the polar decomposition of A_i, $A_i = \sqrt{\rho_i}U_i$, and observe that

$$|\mathrm{Tr}\left(A_1 A_2^\dagger\right)| = |\mathrm{Tr}\left(\sqrt{\rho_1}U_1 U_2^\dagger \sqrt{\rho_2}\right)| =$$
$$|\mathrm{Tr}\left(\sqrt{\rho_2}\sqrt{\rho_1}U_1 U_2^\dagger\right)| \leq \mathrm{Tr}|\sqrt{\rho_1}\sqrt{\rho_2}|. \quad (6.42)$$

Note that the above inequality can be saturated if we choose purifications such that $U_1 U_2^\dagger = \mathbb{1}$.

In the above derivation we have used the property that for any hermitian matrix A and unitary U, $|\mathrm{Tr}(AU)| \leq \mathrm{Tr}|A|$. This can be seen as follows. Let $A = \sum_i a_i |i\rangle\langle i|$ be eigendecomposition of A. Since trace is basis independent we can perform it using the basis $|i\rangle$: $|\mathrm{Tr}(AU)| = |\sum_i a_i \langle i|U|i\rangle|$. Absolute value of any matrix element of a unitray matrix is smaller or equal to 1. Hence $|\sum_i a_i\langle i|U|i\rangle| \leq |\sum_i a_i| \leq \sum_i |a_i| = \mathrm{Tr}|A|$. □

Relation between the Fidelity and the QFI We now prove a theorem that provides a link between the fidelity and the QFI, by showing that infinitesimal change in the fidelity when a quantum state is changed is proportional to the QFI.

Theorem 6.4. For two infinitesimally close states ρ_θ, $\rho_{\theta+\mathrm{d}\theta}$,

$$\mathcal{F}(\rho_\theta, \rho_{\theta+\mathrm{d}\theta}) = 1 - \frac{1}{4}F_Q(\rho_\theta)\mathrm{d}\theta^2 + O(\mathrm{d}\theta^3). \quad (6.43)$$

Proof. Consider:

$$\sqrt{\mathcal{F}(\rho_\theta, \rho_{\theta+\mathrm{d}\theta})} = \mathrm{Tr}\left[\sqrt{\rho_\theta}\left(\rho_\theta + \dot{\rho}_\theta \mathrm{d}\theta\right)\sqrt{\rho_\theta}\right]) =$$
$$= \mathrm{Tr}\left[\rho_\theta^2 + \sqrt{\rho_\theta}\dot{\rho}_\theta\sqrt{\rho_\theta}\mathrm{d}\theta\right], \quad (6.44)$$

where we have made a replacement $\rho_{\theta+\mathrm{d}\theta} = \rho_\theta + \dot{\rho}_\theta \mathrm{d}\theta$. We now want to expand the above quantity up to the second order in $\mathrm{d}\theta$. Since we deal with operators we have to be careful. We write:

$$\mathrm{Tr}\left[\rho_\theta^2 + \sqrt{\rho_\theta}\dot{\rho}_\theta\sqrt{\rho_\theta}\mathrm{d}\theta\right] = \rho_\theta + A\mathrm{d}\theta + B\mathrm{d}\theta^2 + O(\mathrm{d}\theta^3), \quad (6.45)$$

where A and B are operators we want to determine now. Let us take square of the both sides of the above equations and compare terms in the leading orders in $\mathrm{d}\theta$. As a result we obtain the following equations:

$$\rho_\theta A + A\rho_\theta = \sqrt{\rho_\theta}\dot{\rho}\sqrt{\rho_\theta} \qquad (6.46)$$

$$A^2 + \rho_\theta B + B\rho_\theta = 0. \qquad (6.47)$$

Solving the above equations in the eigenbasis of $\rho_\theta = \sum_i p_i |i\rangle\langle i|$ we get:

$$A_{ij} = \frac{\sqrt{p_i p_j}}{p_i + p_j}\dot{\rho}_{\theta,ij} \to \mathrm{Tr}(A) = 0, \qquad (6.48)$$

while

$$B_{ij} = -\frac{\sum_k A_{ik}A_{kj}}{p_i + p_j} = - \frac{1}{p_i + p_j}\sum_k \frac{\sqrt{p_i p_j}}{p_i + p_k}\frac{\sqrt{p_k p_j}}{p_k + p_j}\dot{\rho}_{\theta,ik}\dot{\rho}_{\theta,kj}. \qquad (6.49)$$

As a result

$$\mathrm{Tr}B = -\sum_{ik}\frac{p_k}{2(p_i + p_k)}|\dot{\rho}_{\theta,ik}|^2. \qquad (6.50)$$

Note that from the definition of the SLD, $\dot{\rho}_{\theta,ik} = \frac{1}{2}\Lambda_{\theta,ik}(p_i + p_k)$, and hence $\mathrm{Tr}B = -\frac{1}{2}\mathrm{Tr}\rho_\theta\Lambda_\theta^2 = -\frac{1}{8}F_Q$. So we get

$$\sqrt{\mathcal{F}(\rho_\theta, \rho_{\theta+\mathrm{d}\theta})} = 1 - \frac{1}{8}F_Q \mathrm{d}\theta^2 + O(\mathrm{d}\theta^3) \qquad (6.51)$$

which yields the desired theorem when squared. □